

# The structure of CCT–Hsc70<sub>NBD</sub> suggests a mechanism for Hsp70 delivery of substrates to the chaperonin

Jorge Cuéllar<sup>1</sup>, Jaime Martín-Benito<sup>1</sup>, Sjors H W Scheres<sup>1</sup>, Rui Sousa<sup>2</sup>, Fernando Moro<sup>3</sup>, Eduardo López-Viñas<sup>4,5</sup>, Paulino Gómez-Puertas<sup>4</sup>, Arturo Muga<sup>3</sup>, José L Carrascosa<sup>1</sup> & José M Valpuesta<sup>1</sup>

**Chaperones, a group of proteins that assist the folding of other proteins, seem to work in a coordinated manner. Two major chaperone families are heat-shock protein families Hsp60 and Hsp70. Here we show for the first time the formation of a stable complex between chaperonin-containing TCP-1 (CCT) and Hsc70, two eukaryotic representatives of these chaperone families. This interaction takes place between the apical domain of the CCT $\beta$  subunit and the nucleotide binding domain of Hsc70, and may serve to deliver the unfolded substrate from Hsc70 to the substrate binding region of CCT. We also show that a similar interaction does not occur between their prokaryotic counterparts GroEL and DnaK, suggesting that in eukarya the two types of chaperones have evolved to a concerted action that makes the folding task more efficient.**

Protein folding is usually a problem in the crowded environment of the cell, and nature has devised a group of proteins termed molecular chaperones that assist in the folding of other proteins<sup>1</sup>. Chaperones mostly function by protecting the aggregation-prone regions of their substrates or by providing them with a proper environment so that they can fold by themselves. It is increasingly evident that chaperones work in a coordinated manner in the protein-folding pathways<sup>1,2</sup>, and an example of these networks is the cooperation between chaperones belonging to the Hsp60 (heat-shock protein 60) and Hsp70 families<sup>3</sup>.

Hsp60 chaperones, also termed chaperonins, are large multimers (800–1,000 kDa) composed of ~60-kDa subunits usually arranged in two back-to-back rings, each one enclosing a cavity where folding takes place. Each monomer is in turn composed of three domains: the equatorial domain, responsible for most of the intra- and all the inter-ring interactions and for binding the nucleotide; the apical domain with the peptide binding recognition motif; and the flexible intermediate domain, which acts as a hinge and transmits the ATP-driven conformational changes in the equatorial domain to the apical domain. The chaperonins can assume two functionally and structurally distinct conformations: an open, substrate-receptive conformation, and a closed state, induced upon ATP binding and hydrolysis, wherein folding occurs. Chaperonins are found in all organisms and are classified in two different types. Type I chaperonins are found in eubacteria, mitochondria and chloroplasts, where they form homoheptameric structures, epitomized by the chaperonin GroEL from *Escherichia coli*<sup>4</sup>. The second type of chaperonin is found in archaea, where they can form octamers or nonamers composed of 1–3 different subunits. However, the most complex of these chaperonins is found in

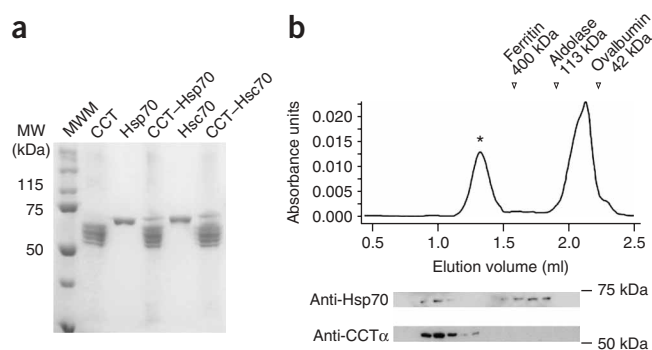
the eukaryotic cytosol and is termed CCT<sup>5</sup> (also known as TRiC). CCT is composed of eight different subunits, and, unlike the rest of the chaperonins, seems to have a more specialized role in the folding of proteins.

The Hsp70 family of chaperones is conserved in all organisms except archaea, and these proteins assist the folding and disaggregation of proteins. These chaperones bind to the unfolded protein using an ATP-driven mechanism that controls and is controlled by the allosteric interactions between their N-terminal ATP binding domain (NBD) and their C-terminal protein binding domain<sup>6</sup> (PBD). In the mammalian cytosol there are two types of Hsp70 chaperones: the constitutive isoforms, termed Hsc70 in mammals, and Hsp70, which is expressed under stress conditions (for the sake of simplicity and unless stated otherwise, the two eukaryotic cytosolic chaperones will be termed Hsc70 hereafter, whereas Hsp70 will be reserved to define all Hsp70 chaperones from prokaryotic and eukaryotic origin). Other paralogs of Hsc70 exist in the endoplasmic reticulum and mitochondria.

The existence of cooperation between Hsp60 and Hsp70 chaperones has been known for some time. In *E. coli* chaperonin GroEL collaborates with the Hsp70 chaperone DnaK<sup>7</sup>, and in eukaryotes CCT cooperates with Hsc70 (refs. 8–10). CCT and Hsc70 have been described to collaborate in the folding of a range of proteins such as the Von Hippel-Lindau protein<sup>11</sup> (VHL) or the assembly of polyglutamine (poly-Q) proteins into nontoxic oligomers<sup>12–14</sup>. Although it has been proposed that this cooperation may occur by the release of the Hsc70-bound substrate into the chaperonin cavity, nothing is known about the substrate-transfer mechanism, or even whether there

<sup>1</sup>Centro Nacional de Biotecnología, CSIC, Campus de la Universidad Autónoma de Madrid, Darwin, 3, 28049 Madrid, Spain. <sup>2</sup>Department of Biochemistry, University of Texas Health Science Center, 7703 Floyd Curl Drive, San Antonio, Texas 78229, USA. <sup>3</sup>Departamento de Bioquímica y Biología Molecular, Universidad del País Vasco, 48080 Bilbao, Spain. <sup>4</sup>Centro de Biología Molecular Severo Ochoa, CSIC-UAM, 28049 Madrid, Spain. <sup>5</sup>CIBER on Physiopathology of Obesity and Nutrition (CB06/03/0026), ISCIII Madrid, Spain. Correspondence should be addressed to J.M.V. (jmv@cnb.csic.es).

Received 12 November 2007; accepted 19 June 2008; published online 27 July 2008; doi:10.1038/nsmb.1464



**Figure 1** Binding of Hsc70 to CCT. **(a)** CCT was incubated with either Hsc70 or its heat-shock isoform Hsp70, and the solution was run on a native gel (data not shown and **Supplementary Fig. 1a**). The CCT band, which runs with a different mobility from that of the Hsp70 chaperones, was excised and loaded onto an SDS gel. The CCT-Hsp70 and CCT-Hsc70 lanes correspond to the experiment described above. The lanes corresponding to CCT, Hsp70 and Hsc70 are mobility controls. MW, molecular weight; MWM, molecular weight marker. **(b)** Gel-filtration experiment with a solution of CCT and Hsc70 $\Delta$ C. Some of the fractions collected were subjected to SDS gel electrophoresis and the bands probed with either an anti-CCT $\alpha$  or an anti-Hsp70/Hsc70 monoclonal antibody.

is a stable interaction between the two types of chaperones. The work presented here reveals the formation of a specific complex between CCT and Hsc70, the first to be observed among the two groups of molecular chaperones.

## RESULTS

### CCT interacts with Hsc70 or Hsp70

To determine whether the cooperation between eukaryotic Hsc70 and CCT would involve a stable intermolecular interaction, we incubated CCT with Hsc70 (or the stress-induced Hsp70) at a 1:10 molar ratio and then resolved the mixture on a native polyacrylamide gel. The CCT oligomer, ~960 kDa, which runs with a mobility distinct from that of Hsc70, 70 kDa (**Supplementary Fig. 1a** online), was excised and run on a denaturing acrylamide gel. This revealed—besides the bands corresponding to the eight CCT subunits—a band corresponding to the constitutive Hsc70 or its heat-shock isoform Hsp70 (**Fig. 1a** and **Supplementary Fig. 1**). In an attempt to corroborate this result, we used a different technique, gel filtration, which allows separation of the Hsc70 bound to CCT from the free Hsc70. Because full-length Hsc70 tends to aggregate through its C-terminal oligomerization domain<sup>15</sup> (**Supplementary Fig. 1**), we used a fully functional deletion mutant of Hsc70 that lacks the 10-kDa C-terminal oligomerization domain<sup>16</sup> (termed ‘Hsc70 $\Delta$ C’; **Supplementary Fig. 1**). Mixtures of Hsc70 $\Delta$ C and CCT were chromatographed on Superose 6. The chromatography profile revealed two peaks, one corresponding to free Hsc70 $\Delta$ C and another corresponding to CCT complexed to Hsc70 $\Delta$ C (**Fig. 1b**). These results confirm the existence of a stable complex between CCT and Hsc70.

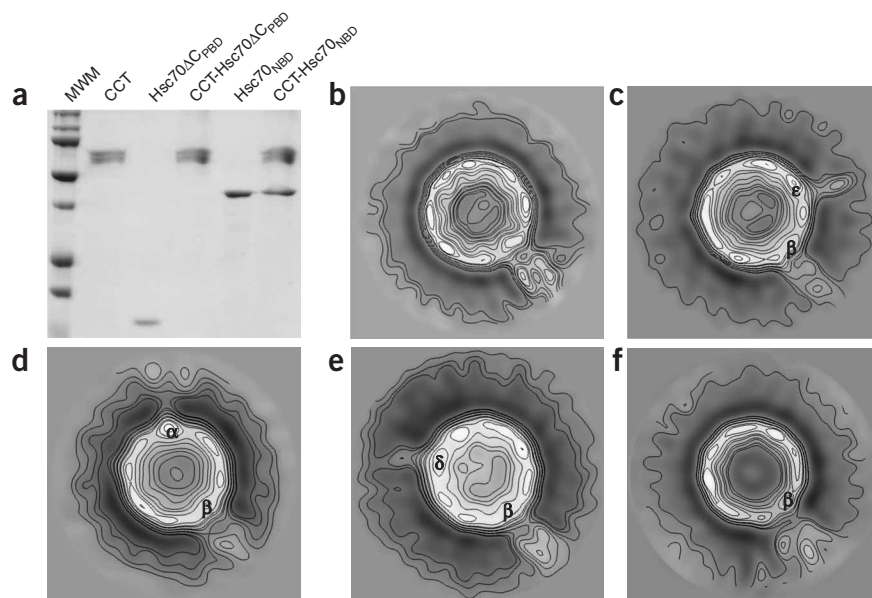
### The interaction between CCT and Hsc70 is specific

Having discovered the formation of a stable complex, we sought to characterize the nature of this interaction. As with all Hsp70 chaperones, Hsc70 is composed of an N-terminal NBD (Hsc70<sub>NBD</sub>; ~45 kDa) and a C-terminal PBD (Hsc70<sub>PBD</sub>; ~25 kDa). We asked whether one or both of these domains are involved in CCT binding. To answer this question, we treated Hsc70 $\Delta$ C with papain under controlled conditions and separated the two Hsc70 domains, Hsc70<sub>NBD</sub> and

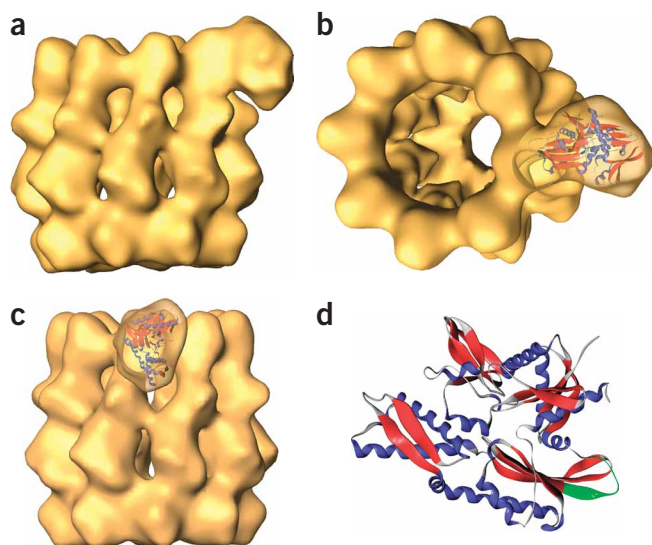
Hsc70 $\Delta$ C<sub>PBD</sub>, which we subsequently purified (**Supplementary Fig. 2** online). These two domains were then independently incubated with CCT, and we assayed their CCT binding activity as described above. The CCT binding assays revealed that, whereas Hsc70 $\Delta$ C<sub>PBD</sub> does not interact with CCT, Hsc70<sub>NBD</sub> does bind to the eukaryotic chaperonin (**Fig. 2a**).

As mentioned, CCT is composed of eight different, albeit homologous, subunits (CCT $\alpha$ , CCT $\beta$ , CCT $\gamma$ , CCT $\delta$ , CCT $\epsilon$ , CCT $\zeta$ , CCT $\eta$  and CCT $\theta$ ), whose arrangement within the ring is known<sup>17</sup>. This heteromeric composition could be important for the allosteric mechanism of CCT<sup>18</sup> and suggests a specific role for the CCT subunits, either in recognition of different substrates or in interaction with other proteins. Indeed, CCT interacts through specific subunits with unfolded actin<sup>19</sup> and tubulin<sup>20</sup>, or with the cochaperones prefoldin<sup>21</sup> or phosphoducin-like protein 1 (PhLP1; ref. 22).

We wanted to determine whether there are also specific subunit(s) of the CCT oligomer involved in the interaction with Hsc70<sub>NBD</sub>, and



**Figure 2** The specificity of the CCT-Hsc70 interaction. **(a)** A CCT binding assay as in **Figure 1a** with two purified Hsc70 domains, Hsc70 $\Delta$ C<sub>PBD</sub> and Hsc70<sub>NBD</sub>. CCT, Hsc70 $\Delta$ C<sub>PBD</sub> and Hsc70<sub>NBD</sub> were also run as mobility controls. MWM, molecular weight marker. **(b)** Two-dimensional average image of the end-on view of the CCT-Hsc70<sub>NBD</sub> complex (1,190 particles). **(c–f)** Two-dimensional average images of the immunocomplexes formed by the CCT-Hsc70<sub>NBD</sub> complex and the anti-CCT $\epsilon$  91A (2,895 particles; **c**), the anti-CCT $\delta$  8g (3,253 particles; **d**), the anti-CCT $\beta$  4E217 (2,523 particles; **e**) and the anti-CCT $\beta$  monoclonal antibody to CCT $\beta$  reveals a larger stain-excluding mass protruding from this CCT subunit that could accommodate both the Hsc70<sub>NBD</sub> and the antibody. The arrangement of CCT subunits corresponds to that previously described<sup>17</sup>.



**Figure 3** The three-dimensional structure of the CCT–Hsc70<sub>NBD</sub> complex. (a) Side view of the three-dimensional reconstruction of the CCT–Hsc70<sub>NBD</sub> complex. (b,c) Side and tilted views of the docking of the atomic structure of the Hsc70<sub>NBD</sub> domain (the NBD domain of Hsc70 was used; PDB 1YUW) into the corresponding mass of the reconstructed CCT–Hsc70<sub>NBD</sub> complex. (d) Atomic structure of the Hsc70<sub>NBD</sub> domain, as in b and c. The  $\alpha$ -helices are blue and the  $\beta$ -strands red. The putative CCT binding site is green.

part of this strand (**Fig. 3d**) as involved in the interaction between Hsc70<sub>NBD</sub> and the apical domain of CCT $\beta$ .

### The interaction between CCT and Hsc70 is unique

The sequence corresponding to the putative CCT binding site of Hsc70<sub>NBD</sub> (V<sub>189</sub>GAERNV<sub>195</sub> in the bovine Hsc70) is distinct in eukaryotes and prokaryotes. Furthermore, this sequence is well conserved in the Hsc70 family but varies in their prokaryotic counterparts or in other members of the eukaryotic family, such as Hsp110/Sse (**Supplementary Fig. 4** online), which is considered an Hsp70-like chaperone because its sequence and overall structure is similar to that of the Hsp70 family chaperones<sup>25</sup>. Hsp110 chaperones have been shown recently to function as a nucleotide-exchange factor of the Hsc70 chaperones through a complex with them<sup>26–28</sup>. A CCT binding assay carried out with the yeast Hsp110 homolog, SSE1, revealed that this protein does not interact with the cytosolic chaperonin (**Fig. 4a**), suggesting that CCT interacts specifically with Hsc70.

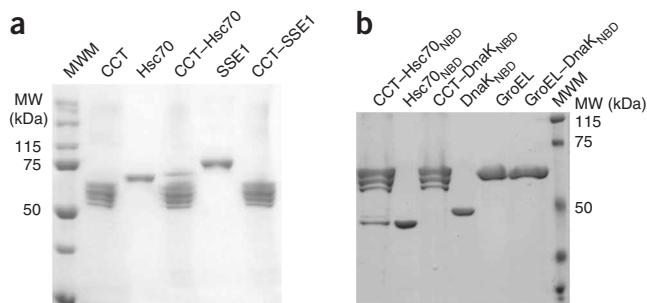
We also wanted to know whether GroEL and DnaK, the prokaryotic counterparts of CCT and Hsc70, respectively, interact with each other as in the case of the eukaryotic chaperones. Using a binding assay similar to the one used for CCT, we observed that neither full-length DnaK (**Supplementary Fig. 5** online) nor its ATPase domain (DnaK<sub>NBD</sub>; **Fig. 4b**) form a stable complex with GroEL or CCT, despite the fact that cooperation between GroEL and DnaK has been observed<sup>7</sup>. Moreover, the observation that Hsc70 also does not form a stable complex with GroEL (**Supplementary Fig. 5**) suggests that there is a unique interaction between Hsc70 and CCT that does not take place between other Hsp60 and Hsp70 chaperones.

To confirm the role of the putative CCT binding sequence (V<sub>189</sub>GAERNV<sub>195</sub>) we generated two mutants: Hsc70 $\Delta$ C<sub>DnaK189–192</sub>, in which the sequence V<sub>189</sub>GAE<sub>192</sub> of Hsc70 was changed to the corresponding sequence of DnaK from *E. coli* (T<sub>185</sub>GN<sub>187</sub>; **Fig. 5a**), and Hsc70 $\Delta$ C<sub>DnaK189–195</sub>, in which Hsc70 V<sub>189</sub>GAERNV<sub>195</sub> was

for that we used EM. A CCT–Hsc70<sub>NBD</sub> complex was generated and subjected to negative staining. We selected several hundred end-on views of the complex and generated a two-dimensional average image, revealing a stain-excluding mass of similar size to Hsc70<sub>NBD</sub> protruding from only one of the CCT subunits (**Fig. 2b**). We then used immunomicroscopy to determine which CCT subunit binds to Hsc70. Aliquots of the CCT–Hsc70<sub>NBD</sub> complex were independently incubated with four different monoclonal antibodies that react specifically against CCT $\epsilon$ <sup>20</sup> ( $\epsilon$ AD1), CCT $\alpha$ <sup>23</sup> (91A), CCT $\delta$ <sup>19</sup> (8g) and CCT $\beta$  (4E217), and the particles selected were used to generate the corresponding two-dimensional average images of the end-on views of the CCT–Hsc70<sub>NBD</sub> complex bound to the above-mentioned antibodies (**Fig. 2c–f**). Whereas, in the first three average images (**Fig. 2c–e**), two stain-excluding masses (those of Hsc70<sub>NBD</sub> and the specific monoclonal antibody) protrude from the CCT structure, in the fourth one (**Fig. 2f**), only a single but larger stain-excluding mass (which could accommodate both Hsc70<sub>NBD</sub> and the antibody) protrudes from the CCT oligomer. On the basis of the known intra-ring subunit arrangement<sup>17</sup>, the combination of the four average images identified CCT $\beta$  as the unique subunit involved in the interaction with Hsc70<sub>NBD</sub>. These results clearly indicate that a stable CCT–Hsc70 complex is formed through a specific interaction between Hsc70<sub>NBD</sub> and CCT $\beta$ .

### Three-dimensional structure of the CCT–Hsc70<sub>NBD</sub> complex

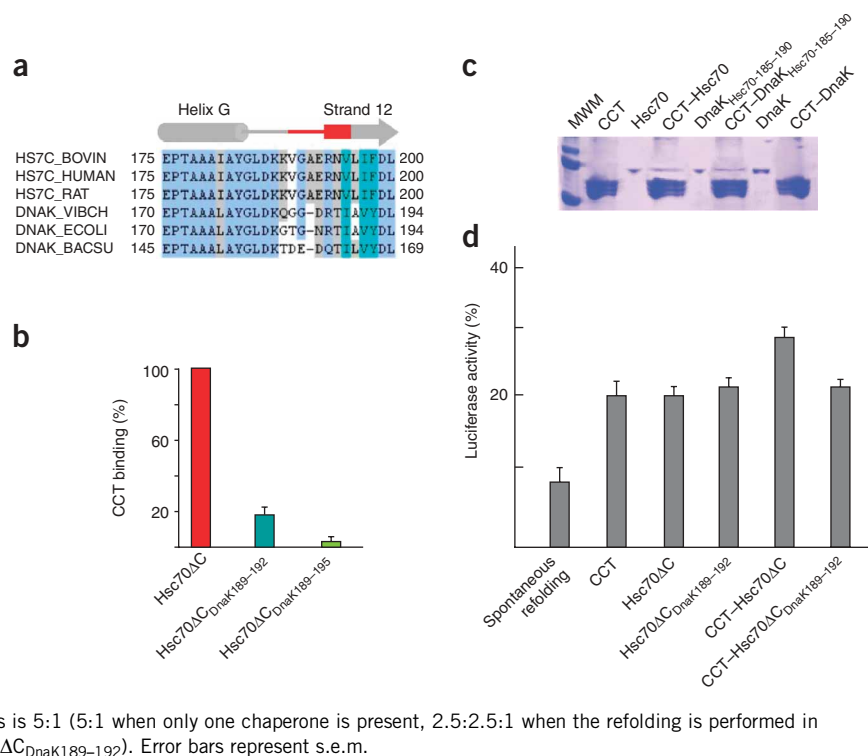
We sought to characterize in more detail the interaction between the two chaperones, and for that purpose aliquots of the CCT–Hsc70<sub>NBD</sub> complex were vitrified and subjected to cryo-EM. After selecting 4,097 particles, three-dimensional maximum-likelihood classification<sup>24</sup> (ML3D) served to separate CCT–Hsc70<sub>NBD</sub> complexes from uncomplexed CCT particles (**Supplementary Fig. 3** online), and a total of 1,218 particles were used to generate a three-dimensional reconstruction of this complex (**Fig. 3a**). This reconstruction revealed the typical double-ring structure of the cytosolic chaperonin, with an additional triangular-shaped mass protruding from the apical domain of one of its subunits (in this case, CCT $\beta$ ), not far from the entrance of the cavity. The similarity between the shape of the protruding mass and the atomic structure of the Hsc70<sub>NBD</sub> domain<sup>16</sup> prompted us to perform a docking of this structure into the corresponding mass of the CCT–Hsc70<sub>NBD</sub> complex (**Fig. 3b,c**). The best docking pointed to part of the loop between  $\alpha$ -helix G,  $\beta$ -strand 12 and the N-terminal



**Figure 4** The specific interaction between Hsc70 and CCT. (a) CCT binding assay of Hsp70-like chaperone SSE1. The CCT–Hsc70 and CCT–SSE1 lanes correspond to the experiments described in the text. The CCT, Hsc70 and SSE1 lanes correspond to mobility controls. (b) Binding experiment of the chaperonins CCT and its prokaryotic counterpart GroEL by Hsc70<sub>NBD</sub> and its prokaryotic homolog DnaK<sub>NBD</sub>. The lanes corresponding to GroEL, Hsc70<sub>NBD</sub> and DnaK<sub>NBD</sub> are mobility controls. The CCT–Hsc70<sub>NBD</sub>, CCT–DnaK<sub>NBD</sub> and GroEL–DnaK<sub>NBD</sub> lanes correspond to the experiments described in the text.



**Figure 5** The specific interaction between Hsc70 and CCT. **(a)** Alignment of several Hsc70s (from bovine, human and rat) and DnaK (*Vibrio cholerae*, *Escherichia coli* and *Bacillus subtilis*) sequences around the putative CCT binding site (red). **(b)** Percentage of CCT binding of the Hsc70 $\Delta$ C<sub>DnaK189–192</sub> and Hsc70 $\Delta$ C<sub>DnaK189–195</sub> mutants as compared to Hsc70 $\Delta$ C (100%). The values correspond to the average of three experiments (**Supplementary Fig. 5**). **(c)** CCT binding of a DnaK chimera containing the CCT binding domain of Hsc70 (DnaK<sub>Hsc70-185–190</sub>). Hsc70 (as a positive control), DnaK (as a negative control) and the DnaK<sub>Hsc70-185–190</sub> chimera were incubated with CCT, and the CCT binding assay described in **Figure 1a** was performed. The lanes corresponding to Hsc70, DnaK<sub>Hsc70-185–190</sub> and DnaK are mobility controls. The CCT–Hsc70, CCT–DnaK<sub>Hsc70-185–190</sub> and CCT–DnaK lanes correspond to the experiments described in the text. **(d)** Refolding activity of the chaperones CCT, Hsc70 $\Delta$ C and Hsc70 $\Delta$ C<sub>DnaK189–192</sub>, and the combined action of CCT with Hsc70 $\Delta$ C or Hsc70 $\Delta$ C<sub>DnaK189–192</sub>. The refolding activity is measured as the percentage of luciferase activity compared to the same amount of native luciferase. Note that the chaperone:luciferase molar ratio in all experiments is 5:1 (5:1 when only one chaperone is present, 2.5:2.5:1 when the refolding is performed in the presence of CCT and Hsc70 $\Delta$ C or CCT and Hsc70 $\Delta$ C<sub>DnaK189–192</sub>). Error bars represent s.e.m.



substituted with the corresponding region from DnaK (T<sub>185</sub>GNRTI<sub>191</sub>). CCT binding assays with these mutants revealed that Hsc70 $\Delta$ C<sub>DnaK189–192</sub> retains only 18% of the wild-type CCT binding ability, and Hsc70 $\Delta$ C<sub>DnaK189–195</sub> only 6% (**Fig. 5b** and **Supplementary Fig. 6** online). We also generated the complementary DnaK mutant (DnaK<sub>Hsc70-185–190</sub>), in which the DnaK T<sub>185</sub>GNRTI<sub>190</sub> sequence was exchanged with the corresponding sequence from Hsc70. The CCT binding assay revealed that this mutant binds CCT (**Fig. 5c**), albeit not with the same efficiency as Hsc70 (64% of the eukaryotic homolog). The observation that this DnaK mutant binds CCT confirms that the V<sub>189</sub>GARENV<sub>195</sub> sequence of Hsc70 is involved in the interaction with the chaperonin CCT, although its lower efficiency suggests that additional factors such as other CCT binding domains and/or a requirement for the structural framework provided by the Hsc70<sub>NBD</sub> domain have a role in the interaction between Hsc70 and CCT. These results confirm the existence of a CCT binding site in Hsc70 and strengthen the notion of a stable complex being formed between this chaperone and CCT.

### The CCT–Hsc70 complex is more efficient in protein folding

We reasoned that the formation of a stable complex between CCT and Hsc70 might have a role in promoting a higher efficiency in the folding of substrates. To test this hypothesis, we denatured luciferase, a protein that can be refolded by both Hsc70 and CCT<sup>9</sup>, and diluted the denaturing agent in the presence of CCT, Hsc70 or a combination of the two chaperones (the chaperone:luciferase molar ratio was kept constant at 5:1). The refolding ability of CCT, Hsc70 $\Delta$ C or the CCT binding–defective mutant Hsc70 $\Delta$ C<sub>DnaK189–192</sub> was similar (**Fig. 5d**). When both CCT and Hsc70 $\Delta$ C were incubated in the presence of unfolded luciferase, the combined refolding activity was higher than that of the two chaperones on their own (note that the total amount of chaperone is kept constant, meaning that in the experiment with both CCT and Hsc70 $\Delta$ C or Hsc70 $\Delta$ C<sub>DnaK189–192</sub> only

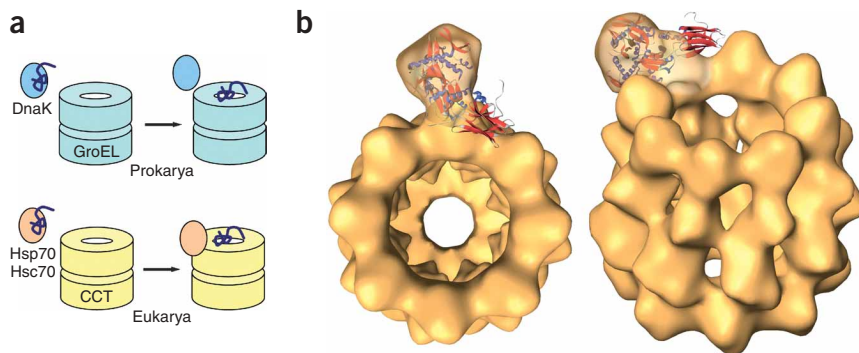
half the molar concentration of each of the two chaperones was used compared to the experiments using a single chaperone). This increase in efficiency does not take place in the presence of the CCT binding–defective mutant Hsc70 $\Delta$ C<sub>DnaK189–192</sub>, suggesting that the formation of a stable complex between CCT and Hsc70 increases the efficiency of the folding process.

Finally, because the two chaperones are ATPases, we reasoned that the formation of the CCT–Hsc70 complex might be modulated by ATP binding and hydrolysis. However, our results reveal that the interaction between the two chaperones is a nucleotide-independent process (**Supplementary Fig. 7** online).

### DISCUSSION

Many processes in the living cell are performed as a coordinated action between multiple proteins, often forming multicomponent complexes. The chaperone-assisted protein-folding pathway is a clear example of this, as many different molecular chaperones have been identified to form a network of interactions that is ultimately aimed at the folding of their common substrates<sup>1</sup>. These studies indicate that eukaryotes seem to have evolved to use more complex protein-folding pathways than prokaryotes. The two most abundant classes of molecular chaperones are the Hsp70s and the Hsp60s. Although a well-characterized set of cochaperones and cofactors has been described for prokaryotic Hsp70s such as DnaK, their eukaryotic homolog, Hsc70, has been found to interact with a much larger set of proteins, including chaperones Hsp90, Hsp100 and Hsp110, nucleotide-exchange factors such as Bag-1 and HspBP1, and other proteins such as Hip, Hop and CHIP<sup>29</sup>.

In the case of prokaryotic Hsp60s, type I chaperonins form an ATP-dependent complex with cochaperonins (or Hsp10s)<sup>30</sup>, and type II chaperonins interact with prefoldin, a small oligomer that transfers unfolded proteins to the chaperonin for its subsequent folding<sup>31,32</sup>. But again, the eukaryotic homolog, CCT, is involved in a much more



**Figure 6** A possible mechanism of substrate release by Hsc70 into the CCT cavity. **(a)** Whereas in prokaryotes DnaK cooperates with the chaperonin GroEL in the folding of denatured proteins without a stable interaction between the two chaperones, in eukaryotes the substrate-bound Hsc70 chaperone interacts specifically with CCT and liberates its substrate into the chaperonin cavity, thereby possibly making the folding process more efficient. **(b)** Two views of the docking of the atomic structure of Hsc70 (ref. 16; PDB 1YUW) into the three-dimensional reconstruction of the CCT–Hsc70<sub>NBD</sub> complex.

intricate set of interactions with other chaperones: several cofactors (A, B, C, D and E), which are needed for the correct incorporation of the folded tubulin into the protofilaments<sup>33</sup>; PhLPs, which assist in the folding of specific proteins such as G $\beta$ -transducin<sup>34</sup> and cytoskeletal proteins<sup>35</sup>; and Hop<sup>36</sup>, which in turn forms a complex with Hsp90 and Hsc70, and has a role in the formation of mature progesterone and glucocorticoid receptors<sup>37</sup>. The versatility of interactions with CCT is also reflected in the increased complexity of its molecular composition. Whereas CCT forms octameric rings composed of eight different subunits, which may interact specifically with both unfolded substrates<sup>19,20</sup> and cochaperones<sup>21,22</sup>, bacterial type I chaperonins form homoheptameric rings, and archaeal type II chaperonins form octamers or nonamers that may be composed of one to three different subunits.

However, despite the many interactions identified in both prokaryotic and eukaryotic protein-folding pathways, until now no stable complexes between Hsp60s and Hsp70s had been described. In this paper, we show for the first time the formation of a highly specific and stable complex between their eukaryotic representatives CCT and Hsc70. Moreover, we show that such a stable complex does not form between their prokaryotic counterparts GroEL and DnaK, despite the fact that cooperation between these chaperones has been observed<sup>7</sup>. Binding assays with deletion mutants revealed that the N-terminal ATP binding domain of Hsc70, rather than its C-terminal protein binding domain, interacts with CCT, whereas EM analyses specifically identified the  $\beta$ -subunit of CCT to be involved in the interaction. Docking of the crystal structure of the Hsc70<sub>NBD</sub> in the three-dimensional reconstruction of the complex identified the highly conserved loop V<sub>189</sub>GAERNV<sub>195</sub> as the putative binding site in Hsc70. A Hsc70 mutant in which this loop was replaced with the corresponding sequence in its homologous prokaryotic counterpart DnaK resulted in a substantial reduction in complex formation, whereas, surprisingly, a DnaK mutant with the Hsc70<sub>189–195</sub> loop inserted did form stable complexes with CCT.

What then could be the role of the complex between CCT and Hsc70? The observation that a stable complex is not formed between their prokaryotic counterparts GroEL and DnaK suggests that this step of the protein-folding pathway has evolved toward a more complex mechanism in eukaryotes (Fig. 6a). Whereas in prokaryotes this interaction seems to be nonspecific, the formation of a stable and

highly specific complex in eukaryotes may provide particular advantages. For example, the specific interaction between the Hsc70<sub>189–195</sub> loop and the CCT $\beta$  subunit may induce conformational changes in either of the partners, which could make their cooperation more efficient. Unfortunately, however, the resolution of our current structural analysis does not allow us to confirm that such conformational changes indeed occur. Another interesting issue concerns how the interaction between the two chaperones is modulated. Our experiments reveal that the interaction between the two chaperones is an ATP-independent process, a behavior that is common to other CCT cochaperones such as prefoldin<sup>21</sup> and PhLP<sup>22</sup>.

The existence of a processive folding machinery formed by CCT and Hsc70 has been suggested previously<sup>11,38</sup>. This machinery could serve to maintain newly synthesized

proteins in a protected environment, thus avoiding the release of these non-native intermediates into the cytosol, which might result in their irreversible aggregation. Docking of the atomic structure of Hsc70 $\Delta$ C into the three-dimensional reconstruction of the CCT–Hsc70<sub>NBD</sub> complex strengthens this notion, because the Hsc70<sub>PBD</sub> would be facing the chaperonin cavity, close to the substrate binding region of CCT on the inner side of its apical domains<sup>19,20,25</sup> (Fig. 6b). It is therefore tempting to speculate that the interaction of substrate-bound Hsc70 with CCT $\beta$  would place the unfolded polypeptide close to the substrate binding domains of CCT $\beta$  or nearby subunits, which have been shown to have an important role in binding unfolded proteins such as actin<sup>20</sup> and tubulin<sup>19</sup>. According to the sequential model of conformational changes proposed for CCT<sup>5,18</sup>, CCT $\beta$  would be one of the last subunits to undergo the conformational change, and in this way it is tempting to speculate that Hsc70 would liberate its substrate at the end of the sequential changes, thereby ensuring its correct processing. Release of the unfolded substrate from Hsc70<sub>PBD</sub>, perhaps facilitated by other factors or by induced conformational changes in Hsc70 or CCT, would then result in substrate trapping by CCT. In that way, the concerted action of both chaperones would prevent exposure of the unfolded substrate to the bulk solvent, as has already been observed<sup>38</sup>, thereby making the folding process more efficient. This higher efficiency can even be observed *in vitro*, as the combined folding activity of CCT and Hsc70 for unfolded luciferase is higher than the activity of the two chaperones on their own. This increase in efficiency does not take place in the presence of the CCT binding-defective mutant Hsc70 $\Delta$ C<sub>DnaK189–192</sub>, which has a similar folding activity to Hsc70 $\Delta$ C. This reinforces the notion of a collaboration between CCT and Hsc70 in the folding of certain proteins that relies on the formation of a stable complex between the two chaperones.

## METHODS

**Protein preparation.** We purified bovine CCT from soluble extracts of bovine testis as described<sup>21</sup>. Full-length Hsc70 and the heat-shock isoform Hsp70 were obtained from STRESSGENE; the various Hsc70 mutants were overexpressed in *E. coli* and purified as described<sup>16</sup>. DnaK and DnaK<sub>NBD</sub> were expressed and purified as described<sup>39</sup>. For construction of DnaK<sub>Hsc70-185–190</sub>, we generated two fragments by PCR, corresponding to DnaK residues 1–181 and 181–638. In the latter, an XbaI site was introduced by a silent single-base replacement, and

DnaK T<sub>185</sub>GNRIT<sub>190</sub> residues were exchanged by the Hsc70 sequence corresponding to V<sub>189</sub>GAERNV<sub>195</sub> (underlined), using the following primer: 5'-GCTCTAGATAAAGGTGTGGGCGCGGAACGTAACGTTGCGGTTTATGACCTGGGT-3'. Fragments were accordingly digested and introduced into NcoI/HindIII sites of pTrc99A (Amersham) vector.

**Preparation and isolation of N-terminal and C-terminal domains of Hsc70.** We purified Hsc70<sub>NBD</sub> and Hsc70<sub>PDB</sub> domains as described<sup>40</sup>, with some modifications. After digestion with papain, the solution was loaded onto a 1-ml HisTrap HP column (GE Healthcare). A linear gradient over ten column volumes from 0 M to 0.15 M imidazole in 10 mM Tris, pH 7.2, was applied to the column. Whereas the Hsc70<sub>PDB</sub> domain was collected in the pass-through, the Hsc70<sub>NBD</sub> domain was eluted at 0.15 M imidazole. The corresponding fractions were pooled and concentrated using an AMICON centrifugal filter unit (Millipore), and the Hsc70 fragments were analyzed by SDS-PAGE.

**Formation of CCT–Hsc70 complexes.** We generated complexes between CCT and the various forms of Hsc70 by co-incubating purified bovine CCT and Hsc70 at 30 °C at a molar ratio of 1:10 for 20 min. This material was then applied to a 2-ml Superose 6 gel-filtration column run in 10 mM Tris-HCl, pH 7.4, 50mM NaCl, 5mM MgCl<sub>2</sub> and 1mM DTT. Fractions emerging from the column were analyzed by SDS-PAGE and western blotting.

**Binding assay between chaperonins and Hsp70/DnaK chaperones.** We carried out the binding of the chaperonins CCT or GroEL to the various Hsp70 forms (Hsc70, Hsp70, DnaK and the various mutants described above) in the following way. The chaperonin and the Hsp70 chaperone were incubated at a 1:10 molar ratio for 30 min at 30 °C. Aliquots were then subjected to a native electrophoresis using 4% (w/v) acrylamide gels run at 90–100 V for 4–5 h. The chaperonin bands, clearly separated in the gel from the unbound Hsp70 chaperone, was stained with Bio-Safe Coomassie (BIO-RAD) and subsequently excised, dehydrated at 95° for five minutes and rehydrated with the denaturing loading electrophoresis buffer. Samples were boiled for 5 min and loaded onto a 10% (w/v) SDS-PAGE. After electrophoresis, gels were blotted electrophoretically for 50 min to a 0.2-mm nitrocellulose membrane (BIO-RAD) using a Trans-Blot SD Transfer Cell (BIO-RAD). The nitrocellulose membrane was subsequently blocked by washing the membrane with WB buffer (3% (w/v) skimmed milk, 0.05% (v/v) Tween-20 in PBS buffer) for 1 h at room temperature. When needed, the membrane was then incubated with a 1:1000 dilution of mouse monoclonal anti-Hsc70 antiserum (Stressgen) followed by a 1:5000 dilution of a sheep-anti-mouse horseradish peroxidase-conjugated secondary antibody (GE Healthcare). Immunoblots were developed with the ECL Plus chemiluminescence reagent (GE Healthcare) and visualized with a Storm 860 PhosphorImager (Amersham Biosciences). The band intensities were quantified using Quantity One software (BIO-RAD).

**ATPase activity.** ATPase activities were measured by adding 20 µg of the various Hsc70 mutants in 10 mM Tris-HCl, pH 8.0, 50 mM KCl, 5 mM MgCl<sub>2</sub>, 1 mM DTT and 2 mM ATP to 25 µl containing 0.3% (v/v) of  $\gamma$ -<sup>32</sup>P ATP (3,000 Ci mM<sup>-1</sup>; 10 mM mCi ml<sup>-1</sup>). Aliquots of 2 µl were then taken at 5 min, 10 min and 15 min and mixed with 2 µl of 50 mM EDTA. Finally, 2 µl aliquots were spotted in duplicate on PEI-cellulose TLC plates and resolved in 0.5 M phosphate buffer. TLC plates were developed by phosphorimage analysis using a Storm 840 Imager (GE Healthcare) and Quantity One software (BIO-RAD).

**Luciferase refolding assay.** We denatured 2.5 µM luciferase (Sigma) in 6 M guanidine hydrochloride (GdnHCl), 50 mM Tris-HCl, pH 7.7, 10 mM DTT for 1 h at 25 °C and diluted 100-fold into 50 mM Tris-HCl, 20 mM KCl, 20 mM MgCl<sub>2</sub>, 10 mM DTT, 0.7 mg ml<sup>-1</sup> BSA, pH 7.7. Dilution was carried out in the absence or presence of CCT, Hsc70ΔC, the CCT binding defective Hsc70ΔC<sub>DnaK189–192</sub> (5:1 chaperone:luciferase molar ratio) or a combination of the two types of chaperones (2.5:2.5:1 CCT:Hsc70:luciferase molar ratio). Reactivation was initiated by addition of 5 mM ATP. Luciferase activity was

measured after 90 min at 25 °C, using the luciferase assay system (Promega E1500) in a Microplate Luminometer Orion (Berthold).

**Electron microscopy.** For EM, 5 µl aliquots of the CCT–Hsc70<sub>NBD</sub> complex were applied to glow-discharged carbon grids for 1 min. We then stained them for 1 min with 2% (w/v) uranyl acetate. Images were recorded at 0° tilt in a JEOL 1200EX-II electron microscope, operated at 100 kV, on Kodak SO-163 film at 60,000× nominal magnification.

For cryo-EM, 5 µl aliquots of the same solution were applied to glow-discharged Quantifoil 1.2 µm holey carbon grids for 1 min, blotted for 3 s and frozen rapidly in liquid ethane at –180 °C. Low-dose images (<10e<sup>-2</sup> A<sup>-2</sup>) of complexes were taken on a FEI Tecnai G<sup>2</sup> FEG200 electron microscope at 200 kV using a Gatan side-entry cryo-holder with a nominal magnification of 62,000× and 1.8–3.0 µm underfocus.

**Image processing.** Micrographs were digitized in a Zeiss SCAI scanner with a sampling window corresponding to 3.5 Å pixel<sup>-1</sup> for negatively stained samples and 3.2 Å pixel<sup>-1</sup> for vitrified samples. Two-dimensional classification of the negative-stain data was performed using maximum-likelihood procedures<sup>41</sup>. For the three-dimensional reconstruction of the frozen-hydrated data, an initial refinement with 14,325 images was performed with the EMAN package<sup>42</sup>, using as a first model an eight-fold symmetrized CCT structure. After several rounds of refinement without symmetry imposition, an extra mass protruding from one of the CCT rings appeared. This structure was calculated from 4,097 images selected by EMAN and was subsequently used for a further classification procedure using three-dimensional maximum-likelihood techniques<sup>24</sup>. After 14 iterations of ML3D classification, standard XMIPP utilities<sup>43</sup> were used to flatten the solvent density in the refined maps, which were subsequently submitted to ten additional iterations of ML3D classification. During all these iterations, local eight-fold symmetry was imposed inside a cylindrically shaped area, including most of the CCT density but excluding the density attributed to Hsc70<sub>NBD</sub>. For the three-dimensional reconstruction procedures, the XMIPP utilities<sup>43</sup> and the projection-matching protocols as implemented in the SPIDER package<sup>44</sup> were used. Docking of the atomic structure of the Hsc70<sub>NBD</sub> domain was carried out manually and optimized using COLACOR, an off-lattice correlation maximizer distributed with Situs 2.2, based on the a local optimization of COLORES<sup>45</sup>. Docking of Hsc70ΔC into the three-dimensional reconstruction of the CCT–Hsc70<sub>NBD</sub> complex was performed superimposing the atomic structure of Hsc70ΔC<sup>16</sup> (PDB 1YUW) onto the docked Hsc70<sub>NBD</sub> using the RMS criteria of the  $\alpha$ -carbon chain.

**Sequence alignment and phylogenetic tree.** Homologous sequences belonging to the Hsc70/Hsp70/DnaK/Hsp110/SSE1 family of proteins were obtained from protein databases using BLAST<sup>46</sup> and aligned using ClustalW<sup>47</sup> and T-COFFEE<sup>48</sup> algorithms. The obtained phylogeny has been inferred using maximum-likelihood methods applied to their complete multiple-sequence alignment. The highest probability model was selected using algorithm Prottest<sup>49</sup>, and the tree was built through the PHYML program<sup>50</sup> by using the most probable evolutive model and evaluating the resulting tree topology reliability by nonparametric bootstrapping (1,000 replicates).

*Note: Supplementary information is available on the Nature Structural & Molecular Biology website.*

#### ACKNOWLEDGMENTS

We thank C. Rodríguez for technical assistance and the Barcelona Supercomputing Center, Centro Nacional de Supercomputación, for providing computer resources. This work was supported by grants CSD2006-00023, BFU2007-62382/BMC and GEN2003-20642-C09-06 from the Spanish Ministry of Education (J.M.V.) and BFU2007-64452 (A.M.). This work was also funded by the EU-grant '3D repertoire' (LSHG-CT-2005-512028) and by the grant RGP63/2004 from the Human Frontiers Scientific Program.

#### AUTHOR CONTRIBUTIONS

J.C. purified proteins and did part of the EM and image processing; J.M.-B. did part of the image processing and the docking analysis; S.H.W.S. did part of the image processing; R.S. generated several Hsc70 mutants; F.M. and A.M. generated several DnaK mutants; E.L.-V. and P.G.-P. did the homology analysis and designed some experiments; J.L.C. did part of the electron microscopy and designed some experiments; J.M.V. did the cryo-EM and devised some experiments.



Published online at <http://www.nature.com/nsmb/>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>

1. Mogk, A., Bukau, B. & Deuerling, E. Cellular functions of cytosolic *E. coli* chaperones. in *Molecular Chaperones in the Cell*. (ed. Lund, P.) 1–34 (Oxford Univ. Press, Oxford, 2001).
2. Martín-Benito, J. *et al.* Divergent substrate-binding mechanisms reveal an evolutionary specialization of eukaryotic prefoldin compared to its archaeal counterpart. *Structure* **15**, 101–110 (2007).
3. Bukau, B. & Horwich, A.L. The hsp70, hsp60 chaperone machines. *Cell* **92**, 351–366 (1998).
4. Braig, K. *et al.* The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* **371**, 578–586 (1994).
5. Valpuesta, J.M., Carrascosa, J.L. & Willison, K.R. Structure and function of the cytosolic chaperonin CCT. in *Protein Folding Handbook* (eds. Buchner, J. & Kiefhaber, T.) 725–755 (Wiley-VCH, Weinheim, 2005).
6. Craig, E.A. & Huang, P. Cellular functions of Hsp70 chaperones. in *Protein Folding Handbook* (eds. Buchner, J. & Kiefhaber, T.) 490–515 (Wiley-VCH, Weinheim, 2005).
7. Langer, T. *et al.* Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. *Nature* **356**, 683–689 (1992).
8. Lewis, V.A., Hynes, G.M., Zheng, D., Saibil, H. & Willison, K. T-complex polypeptide-1 is a subunit of a heteromeric particle in the eukaryotic cytosol. *Nature* **358**, 249–252 (1992).
9. Frydman, J., Nimmesgern, E., Ohtsuka, K. & Hartl, F.U. Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones. *Nature* **370**, 111–117 (1994).
10. Hynes, G., Sutton, C.W. & Willison, K.R. Peptide mass fingerprinting of chaperonin-containing TCP-1 (CCT) and copurifying proteins. *FASEB J.* **10**, 137–147 (1996).
11. Melville, M.W., McClellan, A.J., Meyer, A.S., Darveau, A. & Frydman, J. The Hsp70 and TriC/CCT chaperone systems cooperate *in vivo* to assemble the Von Hippel-Lindau tumor suppressor complex. *Mol. Cell. Biol.* **23**, 3141–3151 (2003).
12. Behrends, C. *et al.* Chaperonin TRiC promotes the assembly of polyQ expansion proteins into nontoxic oligomers. *Mol. Cell* **23**, 887–897 (2006).
13. Tam, S., Geller, R., Spiess, C. & Frydman, J. The chaperonin TRiC controls polyglutamine aggregation and toxicity through subunit-specific interactions. *Nat. Cell Biol.* **8**, 1155–1162 (2006).
14. Kitamura, A. *et al.* Cytosolic chaperonin prevents polyglutamine toxicity with altering the aggregation state. *Nat. Cell Biol.* **8**, 1163–1170 (2006).
15. Chou, C.C. *et al.* Crystal structure of the C-terminal 10-kDa subdomain of Hsc70. *J. Biol. Chem.* **278**, 30311–30316 (2003).
16. Jiang, J., Prasad, K., Lafer, E.M. & Sousa, R.M. Structural basis of interdomain communication in the Hsc70 chaperone. *Mol. Cell* **20**, 513–524 (2005).
17. Liou, A.K.F. & Willison, K.R. Elucidation of the subunit orientation in CCT (chaperonin containing TCP1) from the subunit composition of micro-complexes. *EMBO J.* **16**, 4311–4316 (1997).
18. Rivenzon-Segal, D., Wolf, S.G., Shimon, L., Willison, K.R. & Horowitz, A. Sequential ATP-induced allosteric transitions of the cytoplasmic chaperonin containing TCP-1 revealed by EM analysis. *Nat. Struct. Mol. Biol.* **12**, 233–237 (2005).
19. Llorca, O. *et al.* Eukaryotic type II chaperonin CCT interacts with actin through specific subunits. *Nature* **402**, 693–696 (1999).
20. Llorca, O. *et al.* Eukaryotic chaperonin CCT stabilizes actin and tubulin folding intermediates in open quasi-native conformations. *EMBO J.* **19**, 5971–5979 (2000).
21. Martín-Benito, J. *et al.* Structure of eukaryotic prefoldin and of its complexes with unfolded actin and the cytosolic chaperonin CCT. *EMBO J.* **21**, 6377–6386 (2002).
22. Martín-Benito, J. *et al.* Structure of the complex between the cytosolic chaperonin CCT and phospho-tyrosine protein. *Proc. Natl. Acad. Sci. USA* **101**, 17410–17415 (2004).
23. Hynes, G., Sutton, C.W. & Willison, K.R. Peptide mass fingerprinting of chaperonin-containing TCP-1 (CCT) and copurifying proteins. *FASEB J.* **10**, 137–147 (1996).
24. Scheres, S.H.W. *et al.* Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat. Methods* **4**, 27–29 (2007).
25. Liu, Q. & Hendrickson, W.A. Insights into Hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1. *Cell* **131**, 106–120 (2007).
26. Dragovic, Z., Broadley, S.A., Shomura, Y., Bracher, A. & Hartl, F.U. Molecular chaperones of the Hsp110 family act as nucleotide exchange factors of Hsp70s. *EMBO J.* **25**, 2519–2528 (2006).
27. Raviol, H., Sadlish, H., Rodríguez, F., Mayer, M.P. & Bukau, B. Chaperone network in the yeast cytosol: Hsp110 is revealed as an Hsp70 nucleotide exchange factor. *EMBO J.* **25**, 2510–2518 (2006).
28. Shaner, L., Sousa, R. & Morano, K.A. Characterization of Hsp70 binding and nucleotide exchange by the yeast Hsp110 chaperone Sse1. *Biochemistry* **45**, 15075–15084 (2006).
29. Mayer, M.P. & Bukau, B. Regulation of Hsp70 chaperones by co-chaperones. in *Protein Folding Handbook* (eds. Buchner, J. & Kiefhaber, T.) 516–562 (Wiley-VCH, Weinheim, 2005).
30. Xu, Z., Horwich, A.L. & Sigler, P.B. The crystal structure of the asymmetric GroEL-GroES-(ADP)<sub>7</sub> chaperonin complex. *Nature* **388**, 741–750 (1997).
31. Geissler, S., Siegers, K. & Schiebel, E. A novel protein complex promoting formation of functional  $\alpha$ - and  $\gamma$ -tubulin. *EMBO J.* **17**, 952–966 (1998).
32. Vainberg, I.E. *et al.* Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin. *Cell* **93**, 863–873 (1998).
33. López-Fanarraga, M., Avila, J., Guasch, A., Coll, M. & Zabala, J.C. Postchaperonin tubulin folding cofactors and their role in microtubule dynamics. *J. Struct. Biol.* **135**, 219–229 (2001).
34. Lukov, G.L., Hu, T., McLaughlin, J.N., Hamm, H.E. & Willardson, B.M. Phosducin-like protein acts as a molecular chaperone for G protein  $\beta\gamma$  dimer assembly. *EMBO J.* **24**, 1965–1975 (2005).
35. Stirling, P.C. *et al.* PhLP3 modulates CCT-mediated actin and tubulin folding via ternary complexes with substrates. *J. Biol. Chem.* **281**, 7012–7021 (2006).
36. Gebauer, M., Melki, R. & Gehring, U. The chaperone cofactor Hop/p60 interacts with the cytosolic chaperonin-containing TCP-1 and affects its nucleotide exchange and protein folding activities. *J. Biol. Chem.* **273**, 29475–29480 (1998).
37. Dittmar, K.D. & Pratt, W.B. Folding of the glucocorticoid receptor by the reconstituted Hsp90-based chaperone machinery. The initial Hsp90.p60.hsp70-dependent step is sufficient for creating the steroid binding conformation. *J. Biol. Chem.* **272**, 13047–13054 (1997).
38. Thulasiraman, V., Yang, C.F. & Frydman, F. *In vivo* newly translated polypeptides are sequestered in a protected folding environment. *EMBO J.* **18**, 85–95 (1999).
39. Moro, F., Fernandez, V. & Muga, A. Interdomain interaction through helices A and B of DnaK peptide binding domain. *FEBS Lett.* **533**, 119–123 (2003).
40. Montgomery, D., Jordan, R., McMacken, R. & Freire, E. Thermodynamic and structural analysis of the folding/unfolding transitions of *Escherichia coli* molecular chaperone DnaK. *J. Mol. Biol.* **232**, 680–692 (1993).
41. Scheres, S.H.W. *et al.* Maximum-likelihood multi-reference refinement for electron microscopy images. *J. Mol. Biol.* **348**, 139–149 (2005).
42. Ludtke, S.J., Baldwin, P.R. & Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128**, 82–97 (1999).
43. Sorzano, C.O.S. *et al.* XMIPP: a new generation of an open-source image processing package for electron microscopy. *J. Struct. Biol.* **148**, 194–204 (2004).
44. Frank, J. *Three-dimensional Electron Microscopy of Macromolecular Assemblies* 182–246 (Academic, San Diego, 1996).
45. Chacon, P. & Wrigger, W. Multi-resolution contour-based fitting of macromolecular structures. *J. Mol. Biol.* **317**, 375–384 (2002).
46. Altschul, S.F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
47. Thompson, J.D., Higgins, D.G. & Gibson, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
48. Notredame, C., Higgins, D.G. & Heringa, J. T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217 (2000).
49. Abascal, F., Zardoya, R. & Posada, D. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**, 2104–2105 (2005).
50. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704 (2003).