Structural and functional studies on the eukaryotic chaperonin TRiC/CCT and its cooperating chaperone Hgh1

Leonie Mönkemeyer

aus

Brühl

2019

Acknowledgements

Table of Contents

1 Summary

Chaperonins are ubiquitous and essential molecular chaperones present in all domains of life. The eukaryotic cytosolic chaperonin TRiC is a large hetero-oligomeric complex of almost 1 MDa, which consists of two stacked rings of eight paralogous subunits each. TRiC uses conformational cycling controlled by ATP hydrolysis to assist in the folding and maturation of approximately 10% of cytosolic proteins. Although TRiC was discovered in the early 1990s, many aspects of its mechanism of function and its role in the molecular chaperone network remain to be clarified.

In the present work, the TRiC subunit topology in the complex was elucidated by a new approach, which combines chemical crosslinking and mass spectrometry to identify molecular contacts (Leitner et al., 2012). Crosslinked inter-subunit peptides were mapped onto the structure of the paralogous archaeal chaperonin thermosome. Possible relative subunit arrangements were inferred based on the distance restraints imposed by the employed crosslinker. Independent and self-consistent data sets for yeast and bovine TRiC lead to an unambiguous assignment of the TRiC subunit topology. The complex has overall two-fold symmetry with two homo-typic contacts between the rings. Importantly, based on the new topology, the TRiC complex exhibits functional asymmetry with a segregation of subunits with net positive and negative cavity surface charges and high and low ATP hydrolysis activity into opposing half-rings, respectively.

Substrate folding by the chaperonin is relying on transient encapsulation within the TRiC cavity, which can accommodate proteins up to 70 kDa. The present work presents experimental evidence for partial encapsulation of over-sized substrates by TRiC (Russmann et al., 2012). Folding of model substrates such as fusion proteins of actin, an obligate chaperonin substrate, and green fluorescent proteins (27 kDa) and the natural multi-domain substrate hSnu114 (109 kDa) via transient TRiC encapsulation was analyzed by protease protection. These experiments suggest that TRiC can mediate folding of large proteins by

segmental or domain-wise encapsulation. In the case of hSnu114, a structural homologue of eukaryotic elongation factor 2 (eEF2), selective encapsulation of C-terminal fragments with up to ∼35 kDa was found, presumably reflecting a stepwise folding mechanism.

Finally, we elucidated the function of Hgh1/FAM203, a conserved eukaryotic protein of 45 kDa associated with TRiC in human cells. We found that the orthologous Hgh1 acts – in collaboration with TRiC– as a specialized chaperone in the biogenesis of the multi-domain protein eEF2 in budding yeast. In the absence of Hgh1, a substantial fraction of newly synthesized Eft, the yeast orthologue of eEF2, is degraded or aggregates, indicating increased Eft misfolding. We solved the crystal structure of Hgh1 and analyzed the interaction of wildtype and mutant Hgh1 with Eft. These experiments revealed that Hgh1 is an armadillo repeat protein that binds via a bipartite interface to the central domain III of Eft, which hydrogen-deuterium exchange experiments demonstrated to be the most dynamic domain in mature Eft. Hgh1 binding to Eft folding intermediates prevents aberrant interactions and recruits TRiC to the complex. TRiC likely facilitates the folding of the domains VI and V in Eft, which are C-terminal to domain III. These domains, however, fold stably only in the context of the full-length protein, likely after the N-terminal domains G, G' and II have assumed their native structure. Eft folding is completed upon dissociation of TRiC and Hgh1.

2 Introduction

2.1 What is a protein?

Proteins are biologically active linear polymers of amino acids connected by peptide bonds. After synthesis at the ribosome, the elongated polypeptide chain has to fold into a defined three-dimensional structure, its native state, in order to be functional. Proteins are abundant in all domains of life and carry out essential tasks including signaling, transport and metabolism.

2.2 Protein Folding

Christian Anfinsen pioneered studies on protein folding in the 1950s (Anfinsen, 1973). He found that a small denatured protein will spontaneously refold to its native state after removal of the denaturing agent. He concluded that the native state of a protein is its conformation of lowest free energy, which is solely determined by its amino acid sequence and which can be reached without help of any additional external factors. Multiple weak, non-covalent interactions guide the folding process (Brockwell and Radford, 2007). Initially, mainly hydrophobic interactions drive chain collapse and burial of hydrophobic side chains inside the protein core thereby restricting the conformational space that must be searched upon folding (Dinner et al., 2000). This hydrophobic collapse is followed by more subtle rearrangements optimizing hydrogen-bonding, van-der-Waals interactions and salt bridges. For many proteins folding occurs fast at timescales of milliseconds or less (Brockwell and Radford, 2007; Dobson et al., 1998). This indicates that the folding process must be directed and kinetically driven, rather than a simple scanning of all possible conformations (Levinthal, 1968). Instead of following a precise folding path, proteins are thought to explore funnel shaped potential energy landscapes (Figure 1) (Dobson et al., 1998).
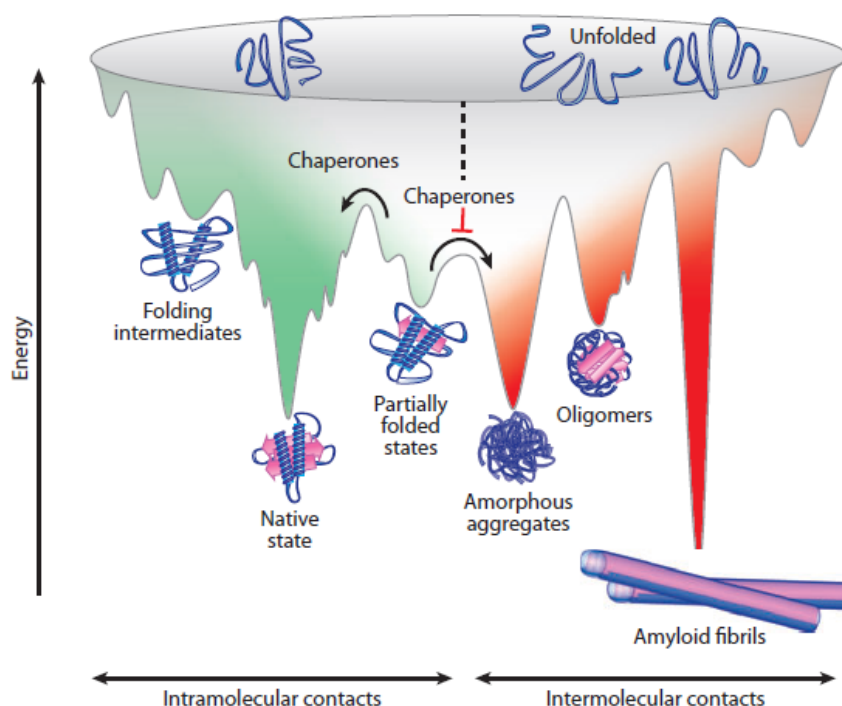
Figure 1: Folding or aggregation? Potential energy landscape showing possible fates of an unfolded polypeptide chain traveling downhill. Productive intramolecular interaction leading to the native state shown in green. Unproductive non-native interactions leading to aggregation shown in red. Molecular chaperones promote productive interactions and prevent non-native contacts. Figure adapted from (Kim et al., 2013).

For larger proteins containing several domains and complex folds, the energy landscape towards the thermodynamically favored native state is often rugged and presents local energy minima. The presence of these energy wells, which can be interpreted as kinetically trapped folding intermediates, slow down the folding process as they are at least transiently populated along the way to the native state (Brockwell and Radford, 2007; Dinner et al., 2000). A folding intermediate presents non-native interactions, which must either be resolved prior to correct folding or lead to misfolding. Partially folded or misfolded proteins typically expose hydrophobic or unstructured residues which may lead to aggregation (Figure 1) (Chiti and Dobson, 2006). Even natively folded proteins may tend to aggregate upon subtle changes in the physiological environment since their folded states are often only marginally stable and represent a compromise between thermodynamic stability and conformational flexibility (Gershenson et al., 2014).

Aggregation is largely driven by hydrophobic forces and is concentration dependent. Although aggregates are thought to form mostly amorphous structures, some non-native proteins also give rise to highly structured amyloid fibrils which are found in the context of many neurodegenerative diseases (Figure 1) (Chiti and Dobson, 2006).

2.3 Molecular chaperones

Anfinsen's findings of spontaneous protein folding without external aid apply stringently only for small single-domain proteins. Larger and/or multi-domain proteins often fold inefficiently and need assistance by molecular chaperones to fold efficiently and on a biologically relevant time scale. Moreover, the physiological conditions of protein folding inside a living cell hardly compare to the situation in Anfinsen's test tube (Gershenson and Gierasch, 2011). First, the situation in the cytosol is ultimately more complex because many different polypeptides fold at the same time. In addition, one major difference is that *in vivo* proteins are synthesized by ribosomes in a vectorial manner – a process called translation. Ribosome-associated (nascent) polypeptide chains cannot assume their native conformations until an independent folding unit, a domain (ca. 50-300 amino acids), has emerged from the ribosomal exit tunnel (Zhang and Ignatova, 2011). Finally, the high protein concentration in the cytosol (300-400 mg/ml) leads to molecular crowding and excluded volume effects favoring the accumulation of misfolded states and aggregation (Ellis and Minton, 2006).

A complex network of molecular chaperones exists, which prevents aberrant interactions of non-native chains, folding intermediates and misfolded states, and has a critical role in maintaining the integrity of the cellular proteome (recently reviewed by (Balchin et al., 2016)). A molecular chaperone can be defined as any protein interacting, stabilizing or helping another protein to attain its functional state without being part of its final structure (Hartl, 1996). Many chaperones are referred to as heat shock proteins (Hsps)

because they are upregulated upon heat stress. The Hsps are divided into evolutionary conserved families according to their molecular weight: Hsp40s, Hsp60s, Hsp70s, Hsp90s, Hsp100s and the small Hsps. Chaperones recognize exposed hydrophobic segments, a characteristic feature of non-native proteins in general, explaining their broad substrate selectivity. These interactions maintain non-native proteins soluble and in a folding-competent state. ATP and specific co-chaperones regulate cycles of substrate binding and release in the ATP-dependent molecular chaperones Hsp60, Hsp70, Hsp90 and Hsp100. Molecular chaperones may not only prevent intermolecular aggregation, but also prevent or reverse intramolecular misfolding. Beyond their role in de-novo protein folding, chaperones are also involved in many other aspects of proteome maintenance, such as macromolecular complex assembly, protein transport and degradation, aggregate dissociation and refolding of stress-denatured proteins (Hartl and Hayer-Hartl, 2002; Kim et al., 2013).

The general chaperone pathways for de-novo protein folding are conserved in all domains of life (i.e. in bacteria, archaea and eukarya) (Figure 2). Folding usually initiates co-translationally in the cytosol as soon as the nascent chain protrudes from the ribosome exit tunnel. The exit tunnel restricts the conformational space of the growing polypeptide chain but is large enough to allow the formation of α-helices or small tertiary structure elements (Wilson and Beckmann, 2011). In polysomes, neighboring exit tunnels point away from one another so as to minimize interactions between growing nascent chains that would lead to aggregation (Brandt et al., 2010; Brandt et al., 2009). Trigger factor (TF) in bacteria and nascent-chain-associated complex (NAC) in eukarya are the first chaperones to meet the nascent chains, respectively. Subsequently, members of the canonical Hsp70 family (DnaK in prokaryotes, Hsp70 in eukarya) may bind the nascent chain, however without directly contacting the ribosome (Calloni et al., 2012). An exception is the ribosome-associated complex (RAC) in budding yeast, consisting of Ssz1 (a non-canonical, specialized Hsp70 homologue) and zuotin (a specialized J-domain protein that binds the ribosome), which cooperates with the ribosome-binding Hsp70 isoforms Ssb1/2 (Preissler and Deuerling, 2012). Prefoldin,

which is absent in bacteria, is another chaperone that can bind nascent chains co-translationally. Nascent chain-binding chaperones delay chain compaction and prevent non-native interactions until a nucleus for productive folding is generated.
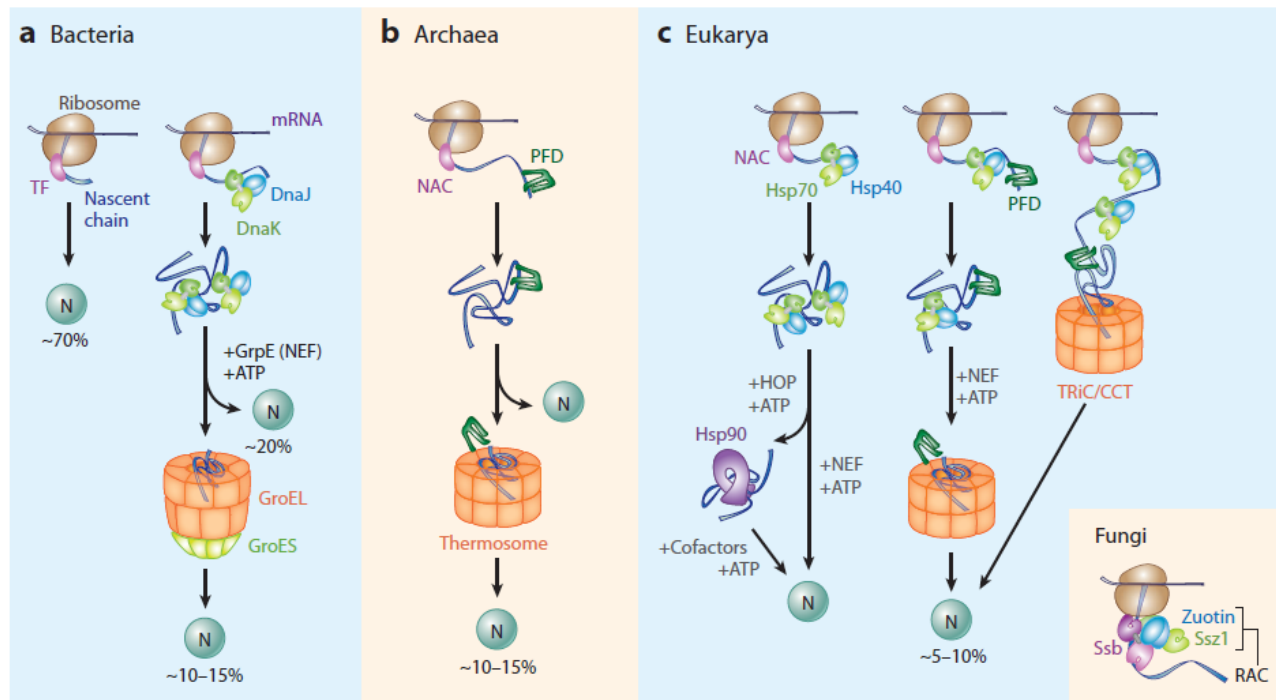
Figure 2: Molecular chaperones involved in de-novo folding in the cytosol. a) In bacteria, interaction with trigger factor (TF) is sufficient for folding of ca 70% of the proteome. Another 20% fold to completion using the Hsp70 system compromising DnaK (Hsp70), DnaJ (Hsp40) and nucleotide exchange factor GrpE. Finally, 10-15% of the newly made proteins are substrate of the chaperonin GroEL with its co-chaperone GroES. b) In archaea, nascent-chain-associated complex (NAC) is the ribosome-bound chaperone. Prefoldin (PFD) may bind clients co-translationally and facilitate their transfer to a chaperonin called thermosome. c) In eukarya, nascent chains interact with NAC or ribosome-associated complex (RAC) in fungi (insert). Downstream of the ribosome, the Hsp70 system and PFD receive clients assisting their folding or providing access to the Hsp90 system or to the chaperonin TRiC/CCT (Tcp1 containing Ring Complex /Chaperonin Complex containing Tcp1). TRiC/CCT may bind nascent chains co-translationally and support domain-wise folding. Figure adapted from (Kim et al., 2013).

The different chaperone families interact directly or cooperate with specific adapter proteins (e.g. Hop/Sti1, which connects Hsp70 and Hsp90) to ensure that the unfolded client protein does not undergo premature aggregation (Figure 2) (Hartl and Hayer-Hartl, 2002; Kirschke et al., 2014). The Hsp70 system

thereby serves as a central hub receiving clients co-translationally, promoting their folding directly or keeping them in a folding-competent state (Calloni et al., 2012; Mayer, 2013). Hsp70 then permits access of the substrates to the more specialized downstream chaperones such as Hsp90 or the chaperonins (also referred to as Hsp60) or alternatively relieves the cell from terminally misfolded proteins by providing a connection to the degradation system (e.g. via the E3 ubiquitin-ligase CHIP, which directly interacts with Hsp70 and Hsp90) (Zhang et al., 2015). In many archaeal species which lack the Hsp70 system, prefoldin is thought to take over some of these tasks (Hartl and Hayer-Hartl, 2002).

2.4 The Hsp90 chaperone system

Hsp90 is an essential key regulator of protein homeostasis (proteostasis), both under normal conditions and upon stress. Indeed, Hsp90 is one of the most abundant proteins in the eukaryotic cytosol under physiological conditions (Schopf et al., 2017). Budding yeast and vertebrates contain two Hsp90 genes named *HSC82* and *HSP82* in *S. cerevisiae* and *HSP90α* and *HSP90β* in *H. sapiens*, respectively (Chen et al., 2006). Hsc82 and Hsp90β are constitutively expressed, while Hsp82 and Hsp90α are heat inducible. The function of Hsp90 can be described as a conformational regulator, which evolved to control protein function and activity. It does so by facilitating the formation of specific active conformations in the case of client kinases (Boczek et al., 2015), by helping in the assembly of multiprotein complexes such as the kinetochore complex (Kitagawa et al., 1999) and by promoting ligand binding to receptors, such as steroid hormone receptors (SHR) (Kirschke et al., 2014).

Hsp90 is a homodimer and dimerization is essential for its function *in vivo* (Wayne and Bolon, 2007). The Hsp90 monomer is composed of an N-terminal Nucleotide Binding Domain NTD (ca 25 kDa), a Middle Domain MD (ca 40 kDa) and a C-terminal Dimerization Domain CTD (ca 12 kDa). A flexible and charged linker connects NBD and MD. The complex conformational cycle of Hsp90 (Figure 3) is regulated by binding and release of nucleotides, clients and cofactors (Balchin et al., 2016). In absence of nucleotide, the

chaperone adopts a v-shaped open conformation. Binding of ATP to the NTD leads to closure of a lid-segment that locks the nucleotide in place. Further conformational changes induce NTD dimerization and then association with the MD, leading to the formation of a twisted closed state. This conformation is capable of ATP hydrolysis. After ATP hydrolysis and nucleotide release, Hsp90 reverts to the open state (Balchin et al., 2016; Schopf et al., 2017).
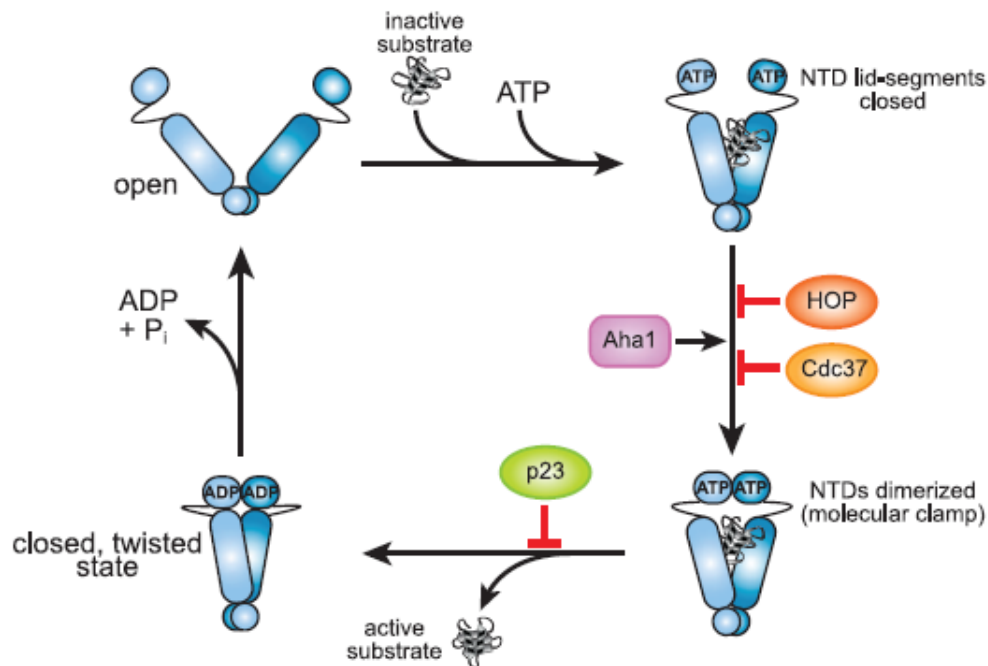
Figure 3: The conformational cycle of Hsp90. Hsp90 functions as a dimer. Each protomer consists of an N-terminal domain (NTD), a middle domain (MD) and a C-terminal dimerization domain (CTD). The Hsp90 dimer adopts an open conformation in the absence of nucleotide binding. ATP binding induces closure of lid segments and then NTD dimerization. Association of the dimerized NTDs with the MD, which contributes catalytic residues, induces ATP hydrolysis in the closed twisted state. After nucleotide release, Hsp90 reverts to the open conformation. The progression through the conformational cycle converts the bound substrate from an inactive to an active conformation. Adapted from (Balchin et al., 2016).

Eukaryotic Hsp90 cooperates with a large set of cochaperones that regulate ATPase activity and client recruitment. This is in stark contrast to the bacterial homologue of Hsp90, HtpG, which functions independently of cofactors (Ratzke et al., 2012). Cochaperones act at various stages along the Hsp90 cycle

and in some cases form mixed complexes with the chaperone (Li et al., 2011). While some cochaperones compete for identical binding sites, other cochaperones can bind simultaneously or synergistically.

The cochaperones Hop and Cdc37 stabilize the open conformation of the Hsp90 dimer, which inhibits ATP hydrolysis and facilitates client binding. Hop (Sti1 in yeast) organizes the transfer of clients from Hsp70 to Hsp90 (Kirschke et al., 2014). Cdc37 is a specific cofactor for kinase clients (Boczek et al., 2015). In contrast to Hop and Cdc37, Aha1 and its homologue Hch1 accelerate ATP hydrolysis by facilitating the transition to the closed state (Armstrong et al., 2012; Koulov et al., 2010). The cochaperone p23 (Sba1 in yeast) acts in the later stages of the chaperone cycle, by stabilizing NTD dimerization and inhibiting ATP hydrolysis (Li et al., 2011). p23 is important for steroid hormone receptor (SHR) maturation. Many additional Hsp90 cochaperones contain tetratricopeptide repeat (TPR) domains that recognize the sequence motif MEEVD at the Hsp90 C-terminus. Some of these cochaperones, such as Cyp40 (Cpr6 and Cpr7 in yeast), contain additional peptidyl-prolyl-isomerase (PPIase) domains (Mayr et al., 2000). Yeast contains three essential Hsp90 cochaperones: Cdc37, Cns1 and Sgt1. While the function of Cdc37 for maturation of kinases is well established, the roles of Sgt1 and Cns1 are less clear. Sgt1 seems to be involved in yeast kinetochore assembly. The function of Cns1 partially overlaps with that of the PPIase cochaperone Cpr7. Cns1 overexpression can rescue slow growth and reduced Hsp90 activity observed upon CPR7 deletion. Moreover, both cochaperones seem to be able to interact directly with one another and with the Hsp90 dimer. It was speculated, that both chaperones are involved in a common essential function that remains to be defined(Tesic et al., 2003).

2.5   The chaperonins

The multi-subunit cylindrical chaperonins are unique among the chaperone families because they can enclose a single substrate molecule inside their central cavity, allowing it to fold unimpaired by aggregation (Hartl, 1996; Hayer-Hartl et al., 2016). Chaperonin substrates have been identified in bacteria,

archaea and eukaryotes, revealing a similar fraction (between 5% and 15%) of the total cellular protein content to interact with the respective chaperonins (Dekker et al., 2011; Hirtreiter et al., 2009; Kerner et al., 2005; Yam et al., 2008). In all domains of life, several essential proteins are among the potential chaperonin substrates, explaining the general necessity for chaperonins. Chaperonins can be divided in two distantly related groups (Kim et al., 2013). Group I chaperonins exist in the bacterial cytosol and the inner compartment of organelles likely derived from endosymbiontic bacteria, i.e. the mitochondrial matrix (Hsp60) and the chloroplast stroma (Cpn60). Group II chaperonins occur in archaea (thermosome) and in the eukaryotic cytosol (TRiC/CCT) (Horwich et al., 2007). Cytosolic group I chaperonins as well as mitochondrial Hsp60s are thermally inducible heat shock proteins (Hartl, 1996). Archaeal chaperonins are also members of the heat shock regulon, being transcribed under all conditions but strongly up-regulated upon heat shock (Gutsche et al., 1999). Thermosomes are highly abundant and can constitute up to 1-2% of total cell protein under basal conditions (Horwich et al., 2007). In contrast to the archaeal thermosomes, TRiC is less abundant (Horwich et al., 2007) and not upregulated upon stress, which is consistent with a role in the de-novo folding of a discrete subset of client proteins.

All chaperonins share a common general architecture of two rings of seven to nine 60 kDa subunits, stacked back-to-back. Sequence and fold of the chaperonin subunits are highly conserved, consisting of three nested domains, called equatorial, intermediate and apical domain (Figure 6B+C). The equatorial domain harbors the ATP binding pocket and mediates most of the contacts within a ring and all of the contacts between rings (Braig et al., 1994; Ditzel et al., 1998). The apical domains contain the binding site for the substrate protein. Chaperonins are "molecular machines", which employ ATP hydrolysis to power extensive conformational changes that switch the chaperonin from an open substrate receptive state to a closed substrate encapsulated state.

Substrate recognition by the chaperonins occurs in the open state at a location of the apical domain facing towards the central cavity. In thermosomes, this region exposes hydrophobic residues, similar to the

substrate binding site in bacterial chaperonins (Gomez-Puertas et al., 2004). Thus, both thermosomes and group I chaperonins appear to recognize substrates by exposed hydrophobic segments (Douglas et al., 2011; Horwich et al., 2007; Pereira et al., 2010). Consistently, thermosomes can successfully fold or prevent aggregation of several GroEL model substrates *in-vitro* (such as citrate synthase, GFP or rhodanese) (Iizuka et al., 2004; Kusmierczyk and Martin, 2003). The putative substrate binding sites in TRiC seem to rely on a mixture of hydrophobic as well as charged and hydrophilic interactions and may allow for a more elaborate substrate recognition mechanism (Joachimiak et al., 2014). The best-characterized TRiC substrates are the cytoskeletal proteins actin and tubulin. Notably, both actin and tubulin are strictly dependent on TRiC and cannot be folded by any other chaperone system (Tian et al., 1995).

2.5.1   The group I chaperonin GroEL

The paradigm for group I chaperonins is GroEL from the bacterium *Escherichia coli*. The 800 kDa GroEL complex is built from two staggered homo-heptameric rings and cooperates with the cochaperone GroES, which consists of a homo-heptameric ring of ~10 kDa subunits (Hsp10). GroES forms a dome-shaped lid which cycles on and off the ends of the GroEL cylinder (Figure 4) (Saibil et al., 2013).

The individual GroEL subunits are composed of an equatorial ATPase domain, an intermediate hinge domain and an apical substrate binding domain forming the entrance to the GroEL cavity of ~45 Å width (Figure 4) (Saibil et al., 2013; Walter and Buchner, 2002). Substrate binding occurs via hydrophobic residues lining a surface cleft. Two or more apical domains interact with a substrate that is in a collapsed compact state without tertiary structure, termed a "molten globule" (Elad et al., 2007). The binding to GroEL prevents aggregation of such species, while folding depends on encapsulation of isolated substrate

by binding of the GroES cochaperone (Hartl and Hayer-Hartl, 2002; Hayer-Hartl et al., 2016; Tang et al., 2006).
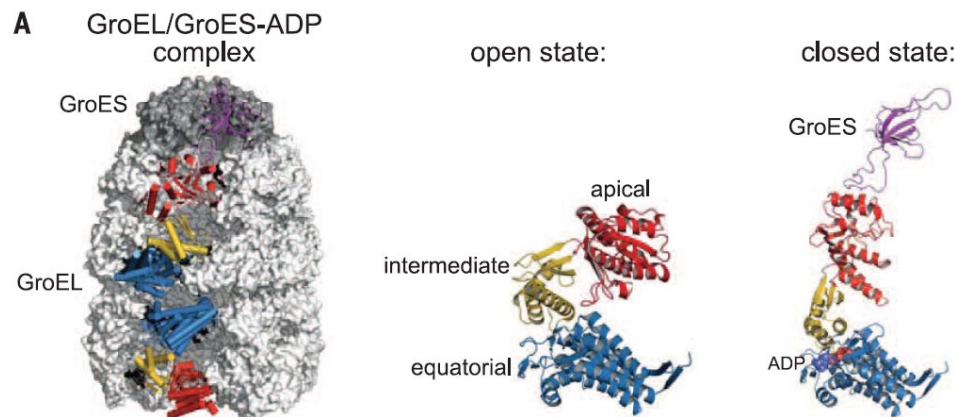
Figure 4: Structure of the bacterial chaperonin GroEL/ES. Crystal structure of the asymmetrical GroEL/ES complex with one subunit of both rings highlighted in color. Structure of individual subunits are shown next to the complex. Direct comparison of open and closed state shows dramatic structural reorganization of GroEL in the course of its functional cycle. Equatorial, intermediate and apical domains are depicted in blue, yellow and red, respectively. GroES is highlighted in purple. Figure adapted from (Balchin et al., 2016).

Folding occurs during a complex functional cycle, which is regulated by the binding and hydrolysis of ATP in the GroEL subunits (Figure 5). The Michaelis-Menten curve shows positive cooperativity in ATP hydrolysis within one ring, but negative cooperativity between the two rings, suggesting that only one ring is active at a time. GroES binding slows ATP hydrolysis. The conventional model of GroEL function thus proposes that the two rings function alternatingly as folding chambers in a two-stroke mechanism (Hartl, 1996; Hayer-Hartl et al., 2016; Horwich et al., 2007). First, 7 ATP molecules bind cooperatively to a substrate-loaded GroEL ring. Next, the GroES heptamer closes the cavity and stabilizes major conformational changes in GroEL (Saibil et al., 2013). As a result, the cavity volume increases from ~85000 $Å^3$ (GroEL cavity alone) to ~175000 $Å^3$ (GroEL-GroES cavity) and the physical properties of the inner GroEL wall change from hydrophobic to hydrophilic (Hayer-Hartl et al., 2016; Xu et al., 1997). The cavity volume is sufficiently large to encapsulate a 60 kDa protein. GroES inserts its "mobile loops" into the substrate

binding site of GroEL, releasing the unfolded substrate from the cage wall into the cavity where it may undergo conformational rearrangements required to reach its native state. Folding can proceed inside the closed and cis-ring (the ring now active) until all ATP are hydrolyzed to ADP (~6 sec at 25°C and ~2 sec at 37°C) (Hayer-Hartl et al., 2016). Finally, binding of ATP to the opposite trans-ring sends an allosteric signal to the cis-ring inducing dissociation of GroES (Horwich et al., 2007). This opens the cavity so that the substrate can exit. In case of incomplete folding or misfolding, the substrate may rebind to undergo a subsequent round of encapsulation. Whether the two folding chambers strictly follow the sequential mechanism described above or may be active at the same time is a question of ongoing research (Hayer-Hartl et al., 2016; Taguchi, 2015).
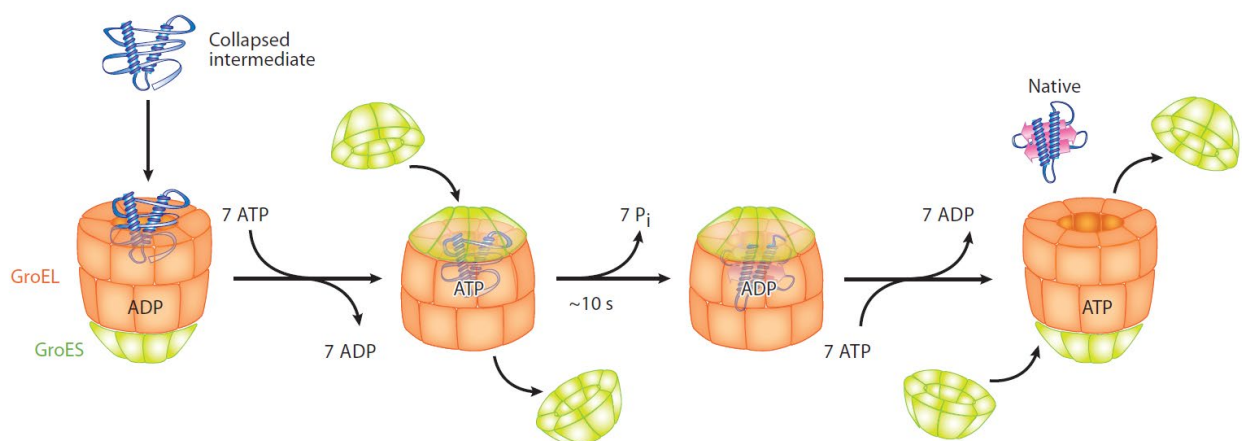
Figure 5: Functional cycle of GroEL/ES. Substrate binds to empty GroEL cis-ring as compact folding intermediate. Cooperative binding of 7 ATP to precedes the closure of the cis-cavity by GroES-Heptamer. Substrate is released into the folding chamber and is free to fold for the time required to hydrolyze all ATP to ADP. ATP binding to the opposite trans-ring dissociates GroES and releases the substrate regardless of its folding state. Figure adapted from (Kim et al., 2013).

Approximately 250 GroEL substrates have been identified in *E. coli*, many comprising domains with complex α+β or α/β topologies (e.g. TIM barrel fold) (Houry et al., 1999; Kerner et al., 2005). These folds contain many non-local, long-range interactions and would fold slowly in free solution, populating

kinetically trapped and aggregation prone intermediates for prolonged times. However, which features precisely render a substrate GroEL-dependent is not completely understood (Hayer-Hartl et al., 2016).

Improvement of substrate folding by GroEL can be mechanistically explained by the chaperonin acting like a passive "Anfinsen cage", which prevents aggregation by isolating the unfolded substrate from the crowded environment thereby providing "infinite dilution" (Ellis, 1994; Hayer-Hartl et al., 2016). This would increase the yield of the folding reaction. For several model substrates, however, in addition to increased folding yields, an acceleration of folding is also observed, which points to a more active role of GroEL in modulating the complex energy landscapes of folding (Chakraborty et al., 2010; Georgescauld et al., 2014; Gupta et al., 2014; Tang et al., 2006). One explanation for this rate acceleration would be that steric confinement inside the GroEL cavity prevents formation of non-productive, expanded folding intermediates, thereby reducing their entropy (Chakraborty et al., 2010) and guiding them along paths avoiding kinetic traps (Georgescauld et al., 2014). In agreement with this, reducing the conformational space of a model protein by the introduction of disulfide bonds lead to a similar acceleration of folding as its confinement inside the GroEL cavity (Chakraborty et al., 2010). In addition to the cavity volume, also the flexible C-termini of the GroEL subunits and the negatively charged residues of the cage wall are critical for the observed acceleration (Hayer-Hartl et al., 2016). The charges may induce a local network of ordered water molecules which in turn increase the efficiency of hydrophobic collapse (England et al., 2008). Additionally, acceleration of folding might result from iterative annealing, which describes a remodeling of folding intermediates upon initial binding. According to this model, the ATP-dependent movements of GroEL apical domains exert a stretching force onto the substrate which might break non-native contacts in kinetically trapped intermediates (Lin et al., 2008). However, the significance of iterative annealing remains unclear because a single round of encapsulation shows the same rate and yield of folding for some model substrates (Chakraborty et al., 2010; Tang et al., 2006). In general, the distinct

mechanistic principles are not mutually exclusive but may contribute to different extent to the folding of different substrates.

2.5.2 Architecture of group II chaperonins

Group II chaperonins share the common chaperonin architecture of two rings arranged back-to-back and enclosing two folding chambers. The individual rings however comprise not seven subunits like GroEL but eight or in some archaea nine protomers (Gutsche et al., 1999). Importantly, most group II chaperonins are hetero-oligomeric with the archaeal thermosomes having up to three paralogous subunits, culminating in the eukaryotic Tcp1 containing ring complex (TRiC, also referred to as CCT for chaperonin complex containing Tcp1) being composed of eight distinct subunits. Compared to group I chaperonins, group II chaperonins have a ~30 amino acid insertion in the apical domain, which forms an iris-like lid in the closed conformation and renders group II chaperonins independent of a GroES-like co-chaperone (Figure 6). Furthermore, the inter-ring contacts are different between the two chaperonin groups. While the two rings of group I chaperonins are arranged in a staggered fashion where one subunit is contacting two subunits of the opposite ring (Braig et al., 1994), the rings are organized in-phase with subunits right on top of one another in the case of group II chaperonins (Ditzel et al., 1998). These different inter-ring contacts suggest a divergent inter-ring signaling mechanism for the two chaperonin groups. Rather recently, a third group of chaperonins has been described (Techtmann and Robb, 2010). These group III chaperonins are found in certain bacteria. However, they share the general structural features of archaeal chaperonins, including eight instead of seven-membered rings and an in-built lid.
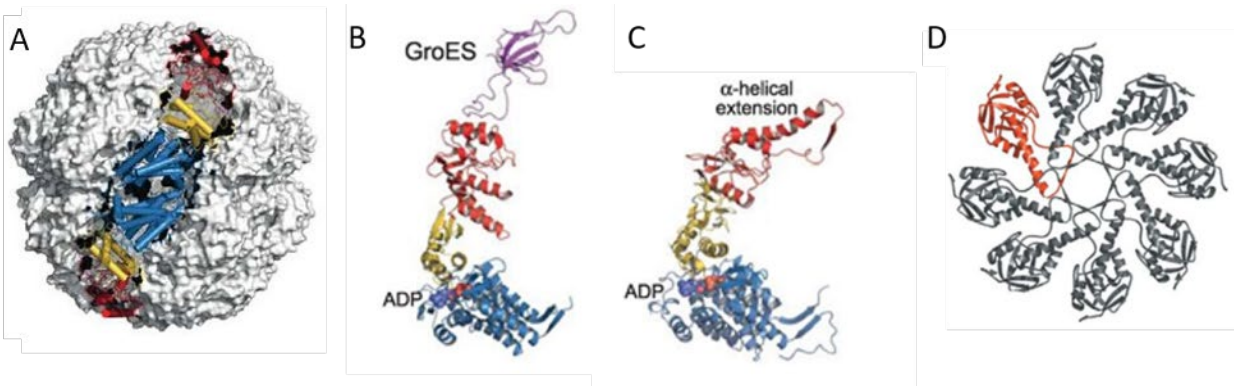
Figure 6: Group II chaperonins. A) Crystal structure of the closed TRiC complex shown in space-filling representation. One subunit in each ring is depicted schematically, highlighting the inter-ring contact sites. B+C) comparison of isolated subunits of GroEL/ES (B) and TRiC (C) both in the closed state. The group II chaperonin subunit has an α-helical extension missing in GroEL. (D) Arrangement of apical domains in the closed form of the Thermoplasma acidophilum thermosome structure. A top view is shown. The α-helical extensions form an iris-like lid. A+B+C adapted from (Balchin et al., 2016), D adapted from (Spiess et al., 2004)

2.5.3 Structures of group II chaperonins in different nucleotide states

The first group II chaperonin structure was the hetero-oligomeric thermosome of the archaeon *Thermoplasma acidophilum*, which was solved in the closed conformation (Ditzel et al., 1998) (Figure 6). In the structure, the chaperonin cavities are closed by formation of a mixed 8-stranded β-barrel around the apical pore at both ends of the chaperonin. The tight interactions of the iris-like lid stabilize the closed conformation and at the same time control inter-ring communication (Reissmann et al., 2007). The lids close over two cavities of ~ 130 000 Å³ each, large enough to encapsulate a polypeptide of up to ~50 kDa (Ditzel et al., 1998).

In the following years, many structural studies were executed on archaeal thermosomes as well as on the eukaryotic TRiC complex, which lead to well-resolved X-ray and cryo-EM structures in different nucleotide states (Cong et al., 2010; Cong et al., 2012; Dekker et al., 2011; Douglas et al., 2011; Huo et al., 2010; Munoz et al., 2011; Pereira et al., 2010; Zhang et al., 2011). In the nucleotide-free apo-state, the group II

chaperonins adopt an open conformation, where the apical domains do not contact one another (Figure 7) (Douglas et al., 2011; Huo et al., 2010; Pereira et al., 2010; Reissmann et al., 2007). Binding of ATP induces a half-closed conformation, caused by a ~45° rigid body rotation of the apical domains leading to the so-called pre-hydrolysis state (Zhang et al., 2011). Full closure requires ATP hydrolysis, and –in contrast to GroEL– non-hydrolyzable ATP analogs are not sufficient to induce stable lid formation (Meyer et al, 2003). The fully closed group II chaperonin complex can only be generated upon stabilization of the trigonal bi-pyramidal transition-state of ATP hydrolysis, usually provided by incubation with ADP or ATP and aluminum fluoride (AlFx) (Chabre, 1990; Melki et al, 1997; Meyer et al, 2003).

The chaperonin of *Methanococcus maripaludis* (MmCpn) serves as a valuable model for group II chaperonins because its homo-oligomeric nature enables the simple introduction of mutations or deletions. Figure 7 shows structures of a lid-less version of the chaperonin of MmCpn in ATP-free, ATP-bound and ATP-hydrolysis states solved by cryo-EM (Zhang et al., 2011).
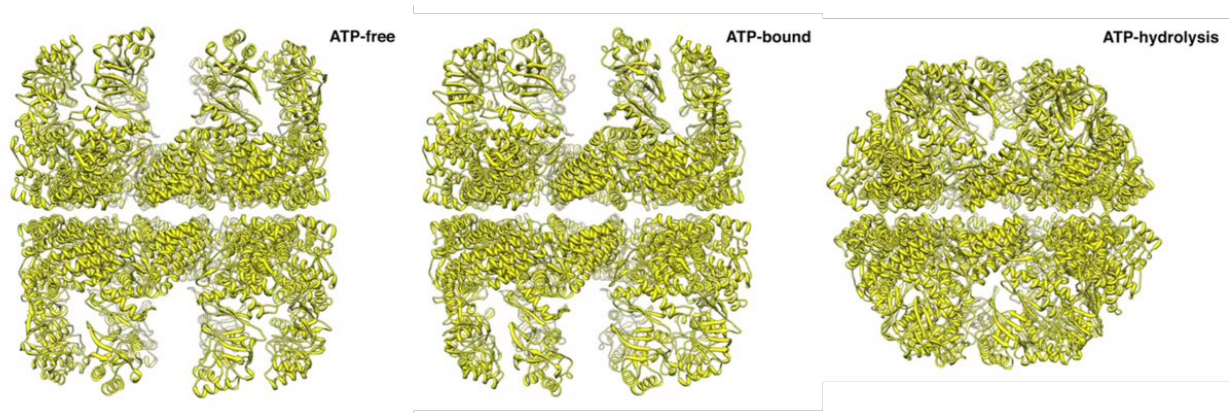
Figure 7: The chaperonin of M. maripaludis (MmCpn) in different nucleotide states, namely from left to right nucleotide-free state, ATP-bound state and ATP-hydrolysis state. The ATP-bound state shows a slight inward tilt of the apical domains compared to the ATP-free state, however the chaperonin clearly remains in an open conformation. Only the ATP-hydrolysis state is fully closed. Structures taken from (Zhang et al., 2011).

2.5.4 Structure of the eukaryotic TRiC complex

Structure determination of the highly complex TRiC chaperonin was a challenging endeavor, but at the outset crystal structures of TRiC in the open state (Munoz et al., 2011) and in the closed state (Dekker et al., 2011) as well as a high resolution cryo-EM structure of closed TRiC (Cong et al., 2010) were available (Figure 8). In general, the structures are similar to thermosome structures in the open and closed states (Skjaerven et al., 2015). However, the hetero-oligomeric nature of TRiC is reflected by certain asymmetries. The TRiC crystal structure in the open state for example shows all apical domains of one ring adopting different conformations (Munoz et al., 2011). In the second ring, six apical domains are disordered due to structural flexibility. The apical domains are not separated as they are in the MmCpn structure but certain apical domains contact each other (Figure 8A). Moreover, symmetry analysis of high-resolution cryo-EM data suggest, that during the transition to the closed state, neighboring apical domains undergo pairwise association (Cong et al., 2012).

Similarly, the closed state of TRiC exhibits asymmetric features. The iris-like lid, which seals the cavities, does not form a perfect circle but a slightly asymmetric arrangement (Dekker et al., 2011). In contrast to other chaperonin structures, the N-terminus of one subunit (CCT4) is located on the outside of the cavity and the subsequent linker to the first secondary structure element threads through the inter-ring interface in the crystallographic model (Dekker et al., 2011). Importantly, the N-termini of all other subunits reside inside the cavity, as do the N-termini of thermosome and GroEL subunits.
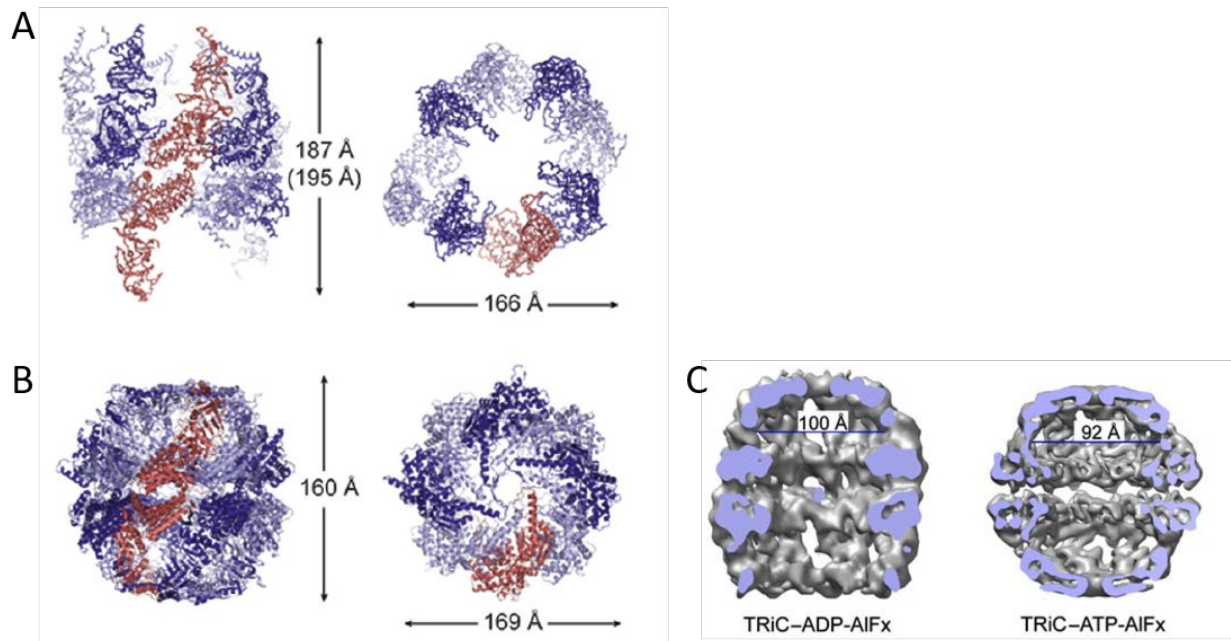
Figure 8: Overview of TRiC structures. A) TRiC crystal structure in the open state at 5.5 Å resolution. B) TRiC crystal structure in the closed state at 3.8 Å resolution. C) Symmetry-free cryo-EM reconstruction of asymmetric ADP-AlFx state and symmetric ATP-AlFx state at 10-14 Å resolution. A+B adapted from (Skjaerven et al., 2015), C adapted from (Cong et al., 2012).

Structural analysis of group II chaperonins mostly revealed complexes with both rings in the same conformation, leading to symmetrically open, half-open or closed complexes. However, Cong et al. succeeded in following TRiC along its ATP-driven conformational cycle by symmetry-free cryo-EM reconstructions (Cong et al, 2012). In addition to the states characterized for MmCpn earlier, they also analyzed an ADP-state and an ADP-AlFx state. While TRiC remains open upon binding of ADP, incubation with ADP and AlFx lead to an asymmetric TRiC complex with one ring open and one ring closed (Figure 8C). Incubation with ATP and AlFx lead to the fully closed complex. Further evidence for asymmetric group II chaperonin complexes analogous to the GroEL/ES bullet complex comes mainly from low resolution EM reconstructions and small-angle X-ray scattering (SAXS) data (Clare et al., 2008; Llorca et al., 1999b; Meyer et al., 2003; Schoehn et al., 2000). Biochemical studies show evidence for negative inter-ring cooperativity,

which suggests that asymmetric complexes might be the physiologically relevant species (Kafri et al., 2001; Reissmann et al., 2007).

2.5.5 Allosteric transitions in the functional cycle of group II chaperonins

The transitions between the different nucleotide states describe the chaperonin's functional cycle, which is strictly dependent on synchronized movements of the individual chaperonin subunits. Consequently, chaperonins are highly allosteric molecular machines. The subunits of one ring form a functional unit coupled by positive cooperativity in ATP binding (Horovitz et al., 2001; Reissmann et al., 2007). Beyond this, there is negative allosteric regulation across the rings, which forces the two hemispheres of the chaperonin to function alternately similar to a 'two-stroke' engine (Bigotti et al., 2006; Kafri et al., 2001) (Figure 9). This special allosteric behavior is observed in both group I and group II chaperonins (Bigotti et al., 2006; Horovitz et al., 2001; Kafri et al., 2001; Reissmann et al., 2007) and is referred to as nested cooperativity.
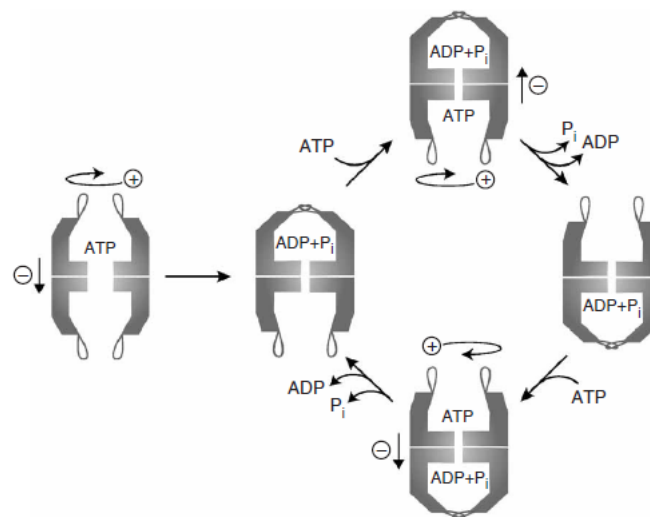
Figure 9: The functional cycle of group II chaperonins. Eight ATP bind cooperatively to one ring of the chaperonin (+) while binding to the second ring is inhibited (-). Subsequent ATP hydrolysis in the cis ring leads to ring closure. Re-

opening of the lid appears to be the rate-limiting step of the conformational cycle (Bigotti et al., 2006), probably because it involves dissociation of the very stable β-barrel at the apex. Due to negative inter-ring cooperativity, ATP hydrolysis in the trans-ring can only proceed once the products of ATP hydrolysis (ADP + Pi) have left the cis-ring. Figure adapted from (Reissmann et al., 2007).

Although the allosteric behavior is similar for both chaperonin groups, there are major differences in how ATP binding and hydrolysis regulate the functional cycle. In GroEL, ATP binding causes a clockwise rotation and upwards tilt of the apical domain leading to the conformation stabilized by GroES binding and consequently to closure of the cavity. These movements can only be realized without steric clashes when all GroEL apical domains move simultaneously in a concerted manner (Ma et al., 2000). In group II chaperonins, the apical domains move in a counterclockwise rotation followed by a tilt of the entire subunit towards the center of the cavity (Zhang et al., 2010a; Zhang et al., 2011). This large rocking motion of the equatorial domain is only possible when the subunits are arrangemend in phase. In GroEL, the staggered arrangement more-or-less locks the equatorial domains in place and major conformational rearrangements are observed mostly in the intermediate and apical domains (Ma et al., 2000). Finally, structures of the open state group II chaperonin show no structural constraints to non-concerted intra-ring conformational changes. Instead, at least for TRiC there is experimental evidence for a sequential mechanism for cavity closure (Lin and Sherman, 1997; Munoz et al., 2011; Rivenzon-Segal et al., 2005).

The in-built lid has a special function in regulating the functional cycle of group II chaperonins. Deletion of the helical protrusion in MmCpn leads to loss of both intra-ring and inter-ring cooperativity (Reissmann et al., 2007). A lid-less mutant of TRiC hydrolyzes ATP and binds unfolded model substrates like wildtype, however, there is no productive folding (Reissmann et al., 2007). Taken together, these results indicate that the helical protrusions do not simply close the cavity, but are deeply integrated into the allosteric network of group II chaperonins. They couple the ATP hydrolysis reaction to productive substrate folding, overall exhibiting similar functions to GroES in the functional cycle of group I chaperonins.

2.5.6 Evolution of subunit specialization in group II chaperonins

The present diversity of group II chaperonin subunits probably arose by a process of gene duplication and differentiation. In the archaeal lineages, gene duplication, loss and conversion was a frequent event and lead to appearance of multiple chaperonin subunits (Archibald et al., 2001). Defined hetero-oligomer formation is thought to be enabled by a mutation in the intra-ring subunit interface, followed by compensatory mutations in the adjacent subunit. Importantly, the archaeal apical domains, which are responsible for substrate recognition, are subject to significant gene conversion, arguing for a general client selectivity shared by all chaperonin subunits. All fully sequenced archaeal genomes encode at least one thermosome subunit and sometimes additional genes for group I chaperonins are present (Lund, 2011). Interestingly, a few group II chaperonin genes are also found in bacterial genomes.

Similar gene duplication processes lead to the formation of the eukaryotic TRiC complex. Multiple, rapid gene duplication events occurred early in the eukaryotic evolution and gave rise to eight paralogous subunits (Archibald et al., 2001; Kubota et al., 1995). The paralogous subunits are described by the Greek letters α, β, γ, δ, ε, ζ, η and θ in mammalian TRiC and by the numbers 1-8 in yeast TRiC (Kubota et al., 1995; Stoldt et al., 1996). Interestingly, the degree of sequence conservation between the paralogous subunits of one species (~ 30% sequence identity) (Kubota et al., 1995) is much lower than the sequence conservation of orthologous subunits (~ 95% within mammals, ~60% between mammals and yeast) (Kim et al., 1994). Among the paralogous subunits, the equatorial domains are well conserved while most sequence divergence occurs in the apical domains (Kim et al., 1994). Taken together, this suggests non-redundant functions of the individual subunits in the hetero-oligomeric ring, probably specialization in substrate binding by the apical domains. All eight TRiC subunits are individually essential in budding yeast and probably occupy a fixed position in the octameric ring (Stoldt et al., 1996). As many obligate TRiC substrates are only found in eukaryotes, it is very likely that TRiC and its substrates co-evolved (Archibald et al., 2001).

2.5.7 Arrangement of subunits in the TRiC complex

The order of the subunits in the TRiC complex was unknown at the outset of this project. The crystallographic as well as cryo-EM data on TRiC were of insufficient resolution for a precise sequence assignment. Since the backbone traces for all subunits are highly similar, the individual subunits cannot be assigned based on the current structural data alone.

Several groups proposed subunit arrangements for TRiC. The original proposal for the intra-ring subunit order was based on 2D-electrophoresis and western-blot analysis of TRiC "micro-complexes" of 2-3 subunits, under the assumption that TRiC would first split into single rings (Liou and Willison, 1997). The proposed arrangement was the basis for many subsequent structural studies (Llorca et al., 2000; Llorca et al., 1999a; Martin-Benito et al., 2004; Rivenzon-Segal et al., 2005). Martin-Benito et al. (Martin-Benito et al., 2007) determined cryo-EM structures of TRiC in complex with subunit-specific antibodies and proposed that the inter-ring interface would contain no homotypic contacts, i.e. no two subunits of the same kind contacting each other (Figure 10A). This model was challenged by subsequent cryo-EM and X-ray structures of TRiC at 4.0 Å and 3.8 Å resolution, respectively (Cong et al., 2010) (Dekker et al., 2011). Both models propose two homotypic (i.e. between two identical subunits) inter-ring contacts, but involving different pairs of identical subunits (Figure 10B+C). While the crystal structure followed the original proposal for the intra-ring arrangement, the model based on the cryo-EM data had a completely new intra- and inter-ring configuration.

A   B   C
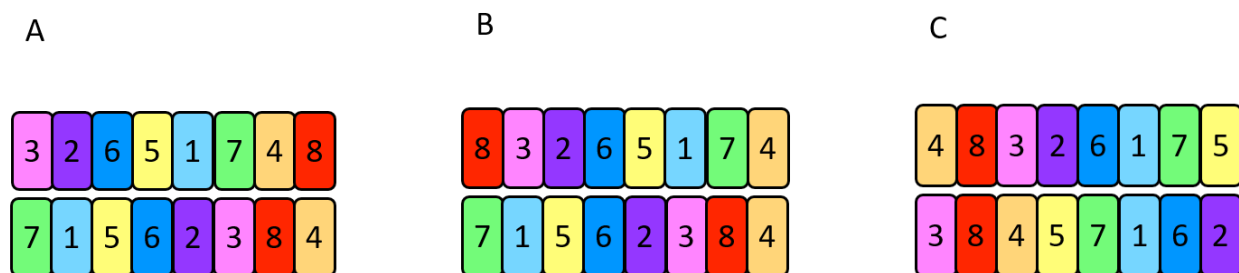
Figure 10: Schematic representation of proposed TRiC arrangements A) Subunit order in the initial model. The intra-ring order was based on the data from Liou and Willison, 1997 (Liou and Willison, 1997), and the inter-ring register on the cryoEM density of TRiC-antibody complexes by Martin-Benito et al., 2007 (Martin-Benito et al., 2007). B) Subunit order in the crystallographic model by Dekker et al., 2011 (Dekker et al., 2011), with homo-typic contacts of subunits 6-6 and 4-4. C) Newer model based on single-particle cryoEM and chemical crosslinking, which was proposed by Cong et al., 2010 (Cong et al., 2010). Here the homo-typic contacts are formed between subunits 8-8 and 1-1.

2.5.8   Substrate binding and recognition by TRiC

While substrate recognition in GroEL and thermosomes relies on hydrophobic interactions, a more elaborated substrate recognition mechanism was proposed for the eukaryotic TRiC complex. In TRiC, the putative substrate binding sites are located in the apical domain, in a groove between helices 10 and 11 (Figure 11) (Joachimiak et al., 2014; Pereira et al., 2010; Spiess et al., 2006), which corresponds to the substrate binding site identified in GroEL between the helices H and I (Chen and Sigler, 1999). The eight paralogous TRiC subunits show great sequence divergence and comprise subunits with a mixture of hydrophobic as well as charged and hydrophilic residues (Gomez-Puertas et al., 2004; Joachimiak et al., 2014; Pappenberger et al., 2002; Spiess et al., 2006; Yebenes et al., 2011). The differential layout of the substrate binding sites lead to the hypothesis, that specific TRiC subunits interact with specific motifs exposed by the substrate molecules. Remarkably, exchanging the substrate binding motifs between different TRiC subunits suffices to confer specificity for a given substrate motif (Spiess et al., 2006). For stable substrate binding, several simultaneous interactions of this kind appear to be required, as seen in the EM structures of TRiC in complex with actin and tubulin (Llorca et al., 2000; Llorca et al., 1999a).

Among all subunits, CCT2 harbors the most nonpolar classical 'chaperone-like' binding surface, and was found to be crosslinked to several full-length substrates (Joachimiak et al., 2014). The unique sequence divergence observed for TRiC apical domains expands the range of possible recognition motifs beyond the simple hydrophobic pattern found for thermosomes and group I chaperonins and allows for recognition of structurally diverse substrates (Spiess et al., 2004).

Another region of TRiC interacting with the substrate is the so-called sensor loop (Munoz et al., 2011). This loop is located deep inside the cavity at the top of the equatorial domain and connects to the ATP binding pocket. This is consistent with the structures of TRiC in complex with the substrates tubulin, Vid27 and 2ABG (Gavin et al., 2006; Herzog et al., 2012; Munoz et al., 2011), which show chaperonin-substrate interactions at the bottom of the cavity. The sensor loop forms a beta sheet together with the N- and C-terminal segments of the neighboring chaperonin subunit (Figure 11). The N- and C-termini were suggested to be involved in protein folding by GroEL (Tang et al., 2006) and thermosomes (Bergeron et al., 2009; Zhang et al., 2010a). It was hypothesized that via this sensor loop substrate and ATP binding could be coordinated (Munoz et al., 2011) or – in an even more intricate model – that conformational changes by ATP hydrolysis could be transmitted onto the substrate thereby actively promoting its folding (Llorca et al., 2001; Stuart et al., 2011). Indeed, FRET (Fluorescence Resonance Energy Transfer) measurements show substantial structural rearrangements of actin not only upon actin binding to TRiC but also upon ATP binding to TRiC (Villebeck et al., 2007). Remarkably, these rearrangements are not seen with GroEL, which can bind actin but does not support its folding (Balchin et al., 2018; Tian et al., 1995). Thus, TRiC and GroEL influence the folding landscape of their substrates in different ways.
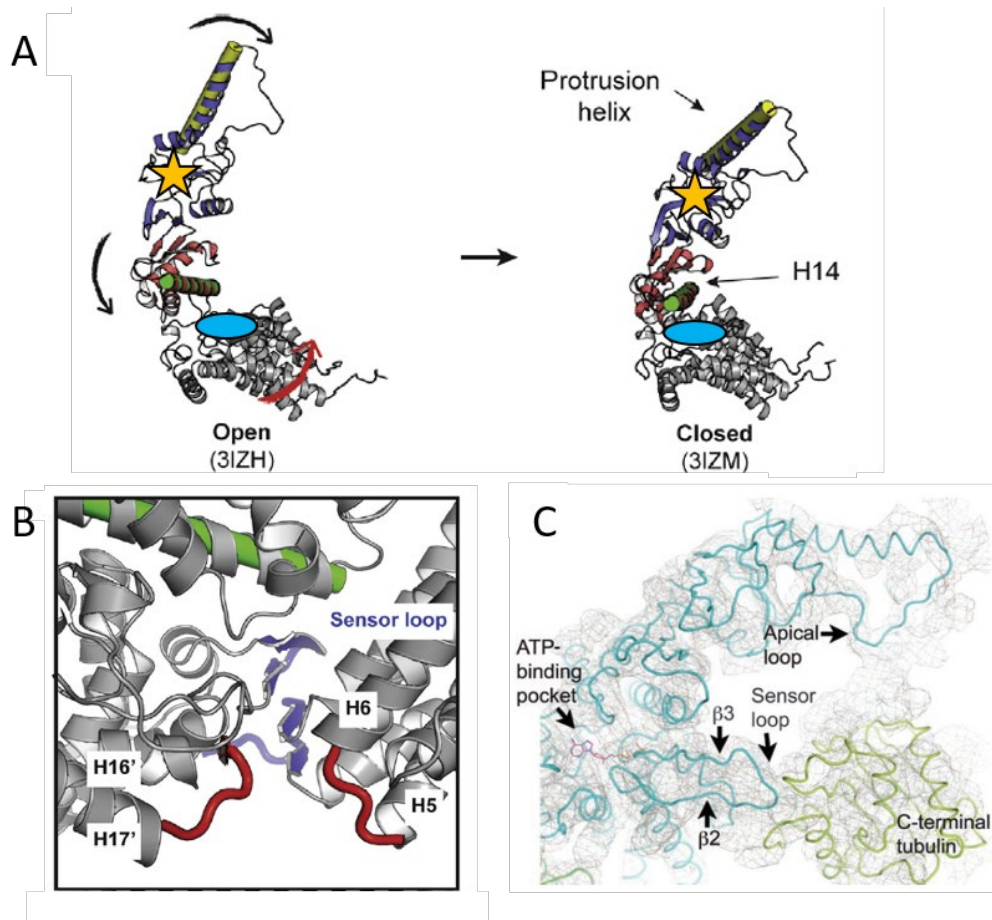
Figure 11: Details of the group II chaperonin subunit. A) Structure of isolated subunit of MmCpn in open and closed state. Approximate positions of the substrate binding site and the ATP binding site are indicated as orange star (below the protrusion helix) and blue oval (in the equatorial domain below helix 14 (H14)), respectively. B) Subunit interface in the equatorial domain indicating how the ß-strands of the sensor loop interact with N- and C-termini of the neighboring subunit (blue). C) Cryo-EM reconstitution of TRiC-tubulin complex. The TRiC subunit is contacting the substrate molecule with its apical domain and the sensor loop. Also the proximity of sensor loop and ATP binding pocket is evident. A+B modified from (Skjaerven et al., 2015), C adapted from (Munoz et al., 2011).

There is controversy about the degree of folding which substrates have acquired before interacting with TRiC. TRiC-bound actin is highly protease sensitive (Meyer et al., 2003) and the TRiC binding motifs of VHL protein (von Hippel Lindau tumor suppressor) become buried in the protein core upon folding (Feldman et al., 2003), which argues for a largely unstructured state of these substrates. On the other hand, especially EM structures of TRiC-substrate-complexes suggest that substrates have acquired a compact

structural state (Balchin et al., 2018; Munoz et al., 2011; Yebenes et al., 2011) and are bound to TRiC in a quasi-native conformation (Llorca et al., 2000). Analysis of this question is complicated by the fact, that *in-vivo*, TRiC cooperates with a multitude of cellular factors which may stabilize substrate conformations not existing in reconstituted systems.

2.5.9   TRiC substrates

Although TRiC was initially believed to serve exclusively in actin and tubulin folding (Sternlicht et al., 1993), numerous non-cytoskeletal TRiC substrates are known today (Spiess et al., 2004). They are involved in various cellular processes including cell cycle regulation (e.g. cyclin E1, CDC20, CDH1), protein trafficking (e.g. myosin heavy chain) and chromatin modification (e.g. by histone deacetylases HDAC3). While GroEL substrates are generally of lower than average abundance (Kerner et al., 2005), some TRiC substrates, e.g. actin but also eEF2 (eukaryotic elongation factor 2), are among the most highly expressed eukaryotic proteins - even though TRiC itself is only present at moderate levels in budding yeast (10 times more eEF2 than TRiC; TRiC present in 1.8×10^4 copies per cell) (Kulak et al., 2014).

Proper TRiC function also has important implications for health and disease. For instance, the TRiC substrate proteins p53, von Hippel Lindau tumor suppressor (VHL), and STAT3 link TRiC to cancer (Feldman et al., 2003; Lopez et al., 2015). In addition, TRiC function is associated with neurodegenerative diseases as TRiC binds the N-terminal fragment of huntingtin, the protein which is mutated in Huntington's disease. In model systems TRiC modulates huntingtin aggregation properties and reduces its cytotoxicity (Behrends et al., 2006; Tam et al., 2006). Finally, several viral proteins require TRiC for folding or virus assembly. Consequently, downregulation of TRiC impairs replication of several important human pathogens including hepatitis C virus and HIV (Lopez et al., 2015).

Although the list of potential TRiC substrates keeps expanding, a clear structural or sequence feature that marks stringent TRiC substrates remains elusive. One major structural motif shared by several TRiC substrate proteins is a WD40 domain, which consists of β-blades forming a propeller structure (Lopez et al., 2015; Spiess et al., 2004; Yam et al., 2008). Among others, this element can be found in the TRiC substrates telomerase co-factor TCAB1, in the cell cycle regulators CDC20 and CDH1 and in the Gß subunit of G-protein signaling complexes (Camasses et al., 2003; McLaughlin et al., 2002) (Freund et al., 2014). Apparently, TRiC assistance is required to form the final ß-propeller structure, which is entirely composed of anti-parallel ß-strands. ß-sheet-rich structures have notorious folding difficulties because correct interactions between non-contiguous sequences need to be established accurately. This requires long search times, giving rise to long-lived folding intermediates, which are prone to aggregation and misfolding (Plaxco et al., 1998). TRiC may facilitate folding of these complicated structures by preventing kinetically trapped non-productive states, which in this special case may include preventing strand-swapping between domains (Spiess et al., 2004).

Another mechanistic aspect of TRiC folding is the release of the folded substrate, which in many cases is coupled to the availability of binding partners. Many TRiC substrates are subunits of homo- or hetero-oligomeric complexes (Yam et al., 2008). Functionality of the substrate is not only dependent on its proper folding but also on its correct incorporation into a higher order assembly, as it is the case for tubulin, CDC20 or VHL (Camasses et al., 2003; Feldman et al., 2003; Lopez-Fanarraga et al., 2001). In these cases, TRiC prevents the premature release of unassembled components and thereby constitutes a major quality control mechanism for quaternary interactions.

Finally, an important characteristic of the ensemble of putative TRiC substrates is its great variety in sizes. Numerous TRiC substrate proteins are larger than 70 kDa and these are typically multi-domain proteins (Dekker et al., 2008; Yam et al., 2008). The TRiC cavity has roughly the same dimensions as the GroEL/ES cavity and is ~180,000 $Å^3$ large (Cong et al., 2012). This is large enough to accommodate substrates up to

70 kDa molecular weight. The majority of GroEL substrates are smaller than 50 kDa and likely can be fully encapsulated (Houry et al., 1999; Kerner et al., 2005). Only a small subset of GroEL substrates is larger than 70 kDa and apparently, these interactions are either non-productive for folding (Kerner et al., 2005) or originate from cycles of binding and release without encapsulation by GroES (Chaudhuri et al., 2001). In contrast, TRiC might function by enclosing only those domains that critically depend on it for folding. The closed lid of the TRiC complex has a central opening of approximately 5 Å, which might be wide enough to accommodate an extended linker sequence connecting structured domains. A domain-wise folding mechanism could also act co-translationally. In this context, TRiC might delay folding or prevent misfolding of an early domain until later domains critically required for proper folding of the protein as a whole have emerged from the ribosome. Indeed, TRiC binding to nascent chains has been reported (Etchells et al., 2005; Frydman et al., 1994). However, clear evidence whether partial encapsulation can occur and whether this results in productive folding is missing.

2.5.10 Co-chaperones and regulators of TRiC activity

As a molecular chaperone, TRiC is tightly integrated into the cellular network of protein quality control factors, serving as a network hub. TRiC receives unfolded or misfolded proteins and promotes their folding directly. In addition, TRiC might hand over potentially aggregation-prone intermediates to appropriate downstream systems, which may include transfer to the ubiquitin-proteasome system (Balchin et al., 2016; Buchberger et al., 2010; Guerrero et al., 2008). Obviously, this task demands for numerous physical interactions with non-substrate proteins. Upstream chaperones can modulate substrate presentation to the chaperonin and adjust substrate flux to the needs of the cell. For instance, Hsp70 (Cuellar et al., 2008; Melville et al., 2003) and the hexameric jellyfish-like prefoldin complex (Geissler et al., 1998; Vainberg et al., 1998) were shown to physically interact with the chaperonin. Prefoldin acts as a holdase, which

stabilizes folding intermediates against aggregation. Hydrophobic patches at the tips of six long coiled-coil "tentacles" bind unfolded clients (Lundin et al., 2004). These tips are also thought to make contact to the chaperonin, probably handing over the bound substrate from the prefoldin cavity into the cylindrical chaperonin (Martin-Benito et al., 2002). Yeast mutants lacking functional prefoldin exhibit defects in actin and tubulin folding (Geissler et al., 1998; Siegers et al., 1999). In contrast, the loss of prefoldin is embryonically lethal in the nematode worm *Caenorhabditis elegans* (Lundin et al., 2008), and prefoldin knock-out mice die five weeks after birth due to various abnormalities caused by cytoskeletal defects (Cao et al., 2008). Together, TRiC and prefoldin are part of an important folding pathway for cytoskeletal proteins and possibly other substrates.

Some TRiC interactors serve as co-chaperones that cooperate with TRiC to ensure optimal client processing. An example for assistance downstream of TRiC are the five tubulin folding cofactors CoA-E which are required for the assembly of α-tubulin and ß-tubulin (both monomers are TRiC substrates) into α/ß-tubulin hetero-dimers, which are the building block of microtubules (Lopez-Fanarraga et al., 2001). Another group of TRiC co-chaperones are the phosducin-like proteins (PhLPs), which can be subdivided in three homologous families called PhLP1, PhLP2 and PhLP3 (Willardson and Howlett, 2007). All PhLPs have been shown to interact with the TRiC chaperonin (McLaughlin et al., 2002; Stirling et al., 2006; Stirling et al., 2007). PhLP1 seems to be primarily involved in the incorporation of the Gß subunit into heterotrimeric G-proteins in metazoans, while PhLP2 and PhLP3 function in actin and tubulin biogenesis (Willardson and Howlett, 2007). Cryo-EM reconstitution of the PhLP1-TRiC complex suggest a similar interaction as observed for the prefoldin-TRiC complex (Martin-Benito et al., 2004). The two domains of PhLP1 contact two juxtaposed apical domains in the chaperonin ring, thereby spanning one end of the cylinder, possibly facilitating the exchange of a bound substrate. In the yeast *S. cerevisiae*, the PhLP2 homologue Plp2 is essential and temperature-sensitive alleles show defects in TRiC-regulated processes such as actin and

tubulin function and cell cycle regulation (Stirling et al., 2007). Deletion of the PhLP3 homologue Plp1 seems to primarily interfere with tubulin biogenesis (Lacefield and Solomon, 2003; Stirling et al., 2006).

2.6 The uncharacterized protein Fam203/Hgh1

A recent genome-wide interaction screen in HeLa cells (Hein et al., 2015) identified the uncharacterized protein Fam203 as a TRiC interactor. Fam203 exhibited binding dynamics similar to PhLPs, and was thus suggested to perform a regulatory function. Homologues of the 45 kDa Fam203 are present throughout the eukaryotic tree, which probably indicates a conserved function. The yeast homologue Hgh1 is a protein of intermediate overall abundance (Kulak et al., 2014). Hgh1 was shown to physically interact with the TRiC subunit Cct6 and the eukaryotic elongation factor 2 homologue Eft in a proteomics screen (Gavin et al., 2006). The interaction between Hgh1 and Eft was confirmed in an independent proteome-wide study (Krogan et al., 2006). Eft is the yeast orthologue of eEF2, an essential and highly abundant elongation factor (see chapter 2.7). In *S. cerevisiae*, Eft is encoded by two genes, *EFT1* and *EFT2,* and simultaneous deletion of both genes is lethal (Perentesis et al., 1992).

While deletion of *HGH1* does not impair growth (Rodriguez-Pena et al., 1998), deletion of both *HGH1* and *EFT2* causes a synthetic growth defect, implying a functional relation between Eft and Hgh1 (Costanzo et al., 2010). Deletion of *HGH1* triggers a mild heat shock response, indicative of increased protein-folding stress in the cytosol (Brandman et al., 2012) . Moreover, the absence of Hgh1 renders yeast cells sensitive to the Hsp90 inhibitor Macbecin (McClellan et al., 2007), and the combined deletion of *HGH1* and components of the Hsp90 machinery, such as *CPR7*, *HCH1*, *HSC82*, *HSP82* and *STI1*, causes a synthetic growth defect (Brandman et al., 2012; Costanzo et al., 2010). Hgh1 interacts directly with the essential Hsp90 cochaperone Cns1 (Gavin et al., 2006; Schlecht et al., 2012; Tarassov et al., 2008). Thus, Hgh1

appears be tightly linked to the cellular chaperone network (Rizzolo et al., 2017; Rizzolo et al., 2018). However, the precise function of Fam203/Hgh1 remains unclear.

2.7 eEF2: Eukaryotic elongation factor 2

During the ribosomal elongation cycle of mRNA translation, amino acids are covalently linked to the growing polypeptide chain in the specific sequence encoded by the sequence of three-nucleotide codons in the mRNA chain. The following amino acid is delivered to the ribosomal A-site by elongation factor 1 (EF1) in the form of an aminoacyl-tRNA. Correct codon-anticodon matching is communicated to the ribosomal peptidyl transferase center (PTC), which catalyzes peptide bond formation between the last residue in the peptidyl-tRNA and the newly docked aminoacyl-tRNA. As a result, the extended polypeptide chain is transferred from the P-site to the A-site and an empty tRNA is left behind in the P-site. Ratcheting of the ribosomal subunits then leads to a hybrid-state, where the acceptor ends of the tRNAs contact the E and P-sites and the anticodon loops contact the P and A-sites, respectively. Resolution of this hybrid state to the canonical E and P-sites is called translocation and requires the activity of elongation factor 2 (EF2). Binding of EF2 is thought to stabilize the hybrid-state, while conformational changes induced by GTP hydrolysis unlock the ribosome. As a result, the empty tRNA has now reached the E-site where it can exit the ribosome and the elongated peptidyl-tRNA chain has moved back to the P-site, resulting in an empty A-site. Translation can proceed with the next cycle of elongation (for details see reviews (Voorhees and Ramakrishnan, 2013) and (Dever and Green, 2012)).

EF2 is a 93 kDa multi-domain GTPase that is essential and highly conserved in sequence and structure. For each elongation cycle of the growing nascent chain, one EF2 molecule binds to the ribosome and hydrolyzes one GTP. Bacterial EF-G, archaeal aEF2 and eukaryotic eEF2 share similar overall structure composed of six structured domains (G, G', II, III, IV and V) with the GTP-binding pocket being located in

the first domain (Noble and Song, 2008) (Figure 12). During the translocation step, EF-G undergoes extensive conformational changes (Lin et al., 2015). Likewise, substantial conformational differences are found between the crystal structures of yeast eEF2 alone and in complex with the antibiotic Sordarin (Jorgensen et al., 2003).
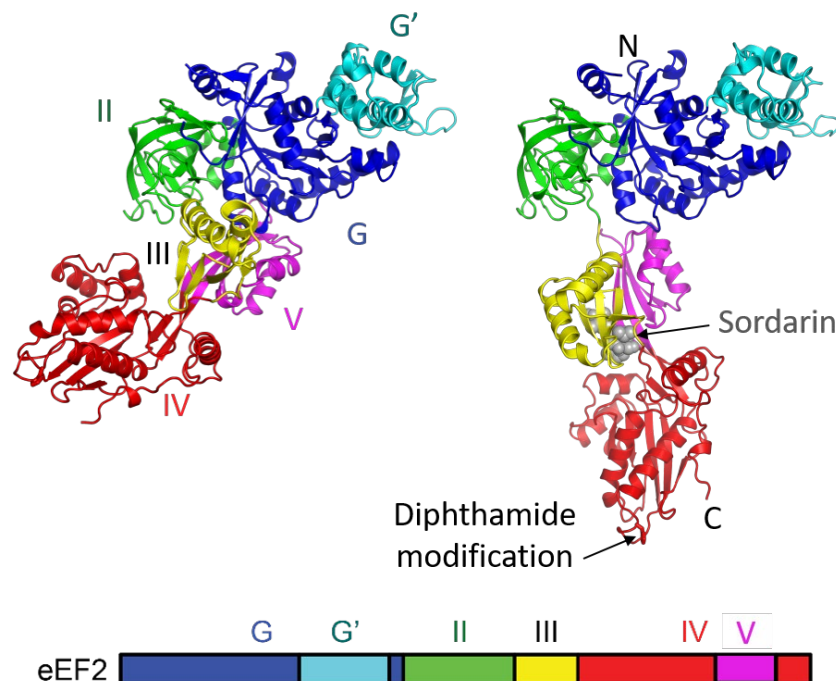
Figure 12: Crystal structures of yeast eEF2 in the apo-state (left) and sordarin-bound state (right). The structures are aligned by their N-terminal GTP-binding modules, comprising domains G, G' and II (dark blue, cyan and green). The domain structure is color-coded and shown schematically below. Figure modified from (Jorgensen et al., 2003).

In contrast to its prokaryotic homologue EF-G, archaeal aEF2 and eukaryotic eEF2 both carry a characteristic and unique posttranslational modification called diphthamide. A conserved histidine residue at the tip of domain IV (H699 in yeast, H715 in mammals) (Figure 12) is modified by the step-wise and coordinated action of seven conserved proteins called Dph1-Dph7 (Su et al., 2013). In archaea, diphthamide group synthesis initiates on fully folded elongation factor 2 (Zhang et al., 2010b). Cryo-EM reconstitutions of eEF2-ribosome complexes show that the tip of domain IV points to the ribosomal decoding center, suggesting that the diphthamide modification improves translation fidelity (Spahn et al.,

2004). Indeed, yeast mutants missing diphthamide show increased sensitivity to translation inhibitors and especially increased -1 frameshifting (Ortiz et al., 2006). However, the diphthamide modification in eEF2 is not strictly required for cell growth (Kimata and Kohno, 1994; Phan et al., 1993). The diphthamide group is the cellular target for several bacterial toxins, such as exotoxin A from *P. aeruginosa* and diphtheria toxin from *C. diphtheriae*. These toxins inactivate eEF2 by transferring the ADP-ribosyl moiety from NAD^+ onto the diphthamide imidazole, which results in inhibition of mRNA translation and cell death (Yates et al., 2006). Furthermore, activity of eEF2 in mammals is negatively regulated via phosphorylation by eEF2 kinase, enabling reversible inactivation for translational regulation (Kenney et al., 2014).

3   Aim of this study

Chaperonins are ubiquitous and essential in all domains of life. While the bacterial GroEL/ES chaperonin system has been well characterized, fundamental aspects of its eukaryotic counterpart, TRiC, remained unexplored.

To resolve the discrepancies between the models for the arrangement of the eight subunits in the TRiC rings, we first set out to determine the subunit topology in TRiC by the novel approach of combining chemical crosslinking with mass spectroscopy (XL-MS). Furthermore, we addressed the open question of how TRiC may promote folding of substrates too large to be encapsulated entirely inside the TRiC cavity. Finally, we explored the structure and function of the putative TRiC cochaperone, Fam203/Hgh1.

4 Results / Publications

4.1   Paper I: The molecular architecture of the eukaryotic chaperonin TRiC/CCT.

The Molecular Architecture of the Eukaryotic Chaperonin TRiC/CCT

Alexander Leitner,[1,8] Lukasz A. Joachimiak,[2,8] Andreas Bracher,[4,8] Leonie Mönkemeyer,[4,8] Thomas Walzthoeni,[1,5,8] Bryan Chen,[2] Sebastian Pechmann,[2] Susan Holmes,[3] Yao Cong,[6] Boxue Ma,[6] Steve Ludtke,[6] Wah Chiu,[6] F. Ulrich Hartl,[4] Ruedi Aebersold,[1,7,*] and Judith Frydman[2,*]

[1]Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, 8093 Zurich, Switzerland
[2]Department of Biology
[3]Department of Statistics
Stanford University, Stanford, CA 94305, USA
[4]Department of Cellular Biochemistry, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany
[5]PhD Program in Molecular Life Sciences, University of Zurich/ETH Zurich, 8057 Zurich, Switzerland
[6]National Center for Macromolecular Imaging; Verna and Marrs McLean Department of Biochemistry and Molecular Biology Baylor College of Medicine, Houston, TX 77030, USA
[7]Faculty of Science, University of Zurich, CH-8006 Zurich, Switzerland
[8]These authors contributed equally to this work
*Correspondence: aebersold@imsb.biol.ethz.ch (R.A.), jfrydman@stanford.edu (J.F.)
DOI 10.1016/j.str.2012.03.007

SUMMARY

TRiC/CCT is a highly conserved and essential chaperonin that uses ATP cycling to facilitate folding of approximately 10% of the eukaryotic proteome. This 1 MDa hetero-oligomeric complex consists of two stacked rings of eight paralogous subunits each. Previously proposed TRiC models differ substantially in their subunit arrangements and ring register. Here, we integrate chemical crosslinking, mass spectrometry, and combinatorial modeling to reveal the definitive subunit arrangement of TRiC. In vivo disulfide mapping provided additional validation for the crosslinking-derived arrangement as the definitive TRiC topology. This subunit arrangement allowed the refinement of a structural model using existing X-ray diffraction data. The structure described here explains all available crosslink experiments, provides a rationale for previously unexplained structural features, and reveals a surprising asymmetry of charges within the chaperonin folding chamber.

INTRODUCTION

The eukaryotic chaperonin TRiC/CCT (hereafter, TRiC) is essential for cell survival, employing ATP hydrolysis to fold ∼10% of the proteome (Yam et al., 2008), including many essential proteins, such as cytoskeletal components and cell cycle regulators (Hartl et al., 2011; Spiess et al., 2004). The folding of many of these substrates is strictly dependent on TRiC. The TRiC subunits are related to the simpler archaeal chaperonin, the thermosome (Ditzel et al., 1998; Pereira et al., 2010; Shomura et al., 2004). Most thermosomes and TRiC consist of two 8-membered rings that are stacked back-to-back. Many archaeal species

have just one thermosome gene (Zhang et al., 2010). In stark contrast, the eukaryotic complex consists of eight different but related subunits (CCT1 to CCT8), all of which are essential in yeast. The subunit specialization occurred very early in eukaryote evolution (Archibald et al., 2001) and is conserved to such an extent that the sequence identity between orthologous mammalian and yeast subunits of the same type is nearly 60%, whereas the sequence identity between paralogous subunits in the same organism is only about 30%. Each of the eight TRiC subunits may differ in substrate specificity; as a result, nonnative polypeptides engage the chaperonin through combinatorial interaction with selected subunits (Feldman et al., 2003; Llorca et al., 2001; Muñoz et al., 2011; Spiess et al., 2006). This mode of recognition dictates the topology of bound substrates, thereby influencing their folding trajectory (Douglas et al., 2011).

The original proposition for the TRiC subunit arrangement came from a western blot analysis of low-molecular-weight subcomplexes found in very low amounts in crude mammalian cell extracts (Liou and Willison, 1997). Similar electrophoretic mobility was used to infer neighbors in the intact complex. Although these low abundance entities were never characterized further, they remain the foundation for a large body of structural work on TRiC (Llorca et al., 1999, 2000; Martín-Benito et al., 2004, 2007; Rivenzon-Segal et al., 2005), including the recent crystal structure of the closed conformation (Dekker et al., 2011). Under the assumption that the fragmentation was always preceded by dissociation into single rings, the incomplete data (subunit θ was apparently not part of any microcomplex) were consistent with the proposed arrangement, CCT 6-5-1-7-4-8-3-2 (i.e., TCP ζ-ε-α-η-δ-θ-γ-β). Later electron microscopy (EM) studies of TRiC with bound subunit-specific antibodies seemed to confirm this arrangement (Martín-Benito et al., 2007). Because of the complexity of the problem, the data employed was sparse, and the assignment of the subunits was only possible under far-reaching assumptions. The inherent ambiguity of the antibody decoration approach is underscored by the inability to predict the correct interring register, even from three-dimensional (3D)
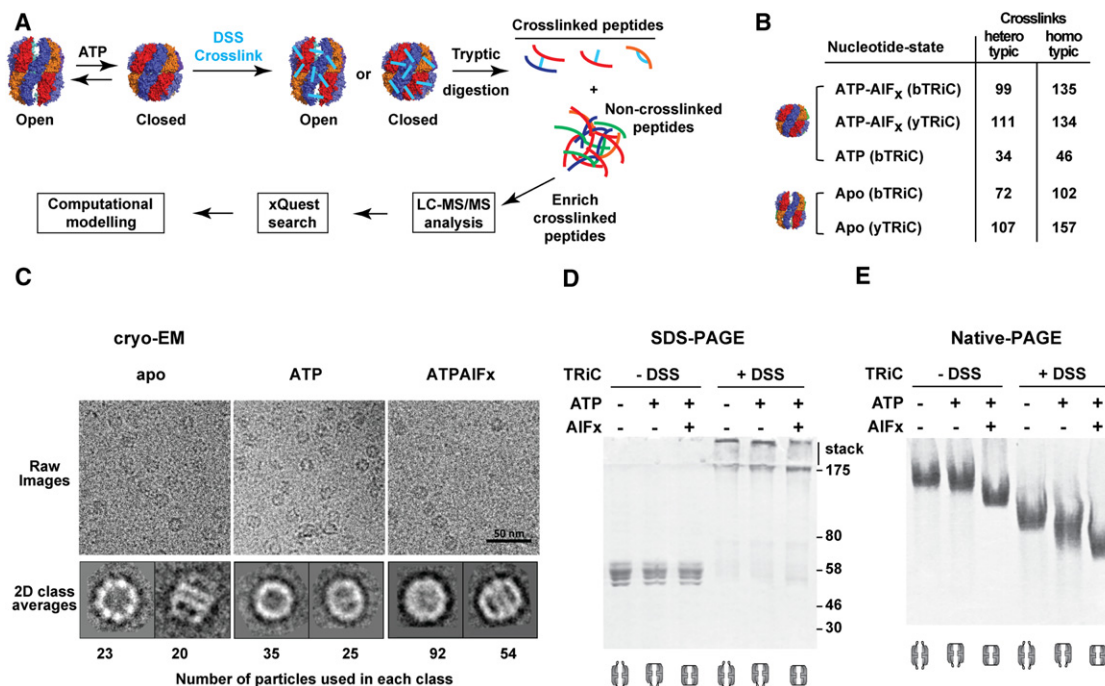
Figure 1. Mass Spectrometry Analysis of Crosslinked TRiC Yields Specific Intersubunit Crosslinks

(A) TRiC was incubated with or without nucleotide to generate the desired conformational state, treated with crosslinking reagent, and proteolyzed to generate an ensemble of crosslinked and noncrosslinked peptides. Crosslinked peptides were chromatographically enriched and analyzed by LC-MS-MS. The identity of the peptides and anchor lysine residues was determined using xQuest (Rinner et al., 2008). Validated crosslinks were used for TRiC model building.

(B) Summary of crosslinks identified using TRiC purified from two different species, bovine (bTRiC) and yeast (yTRiC).

(C) Cryo-EM imaging evidence for the structural integrity of crosslinked TRiC in the apo (left), ATP (middle), and ATP+AIF$_x$ (right) states. (Top and bottom panels) Representative cryo-EM images and corresponding characteristic top and side views of the reference-free 2D class averages of the crosslinked TRiC; numbers of raw particle images used to derive the averages are indicated.

(D and E) SDS- (D) and native-PAGE (E) analysis of bTRiC in indicated nucleotide states without (lanes 1–3) or with (lanes 4–6) crosslinking.

See also Figure S1 and Table S2.

reconstructions of such complexes (Martín-Benito et al., 2007). Unfortunately, the quality of the subsequent electron microscopy and X-ray crystallographic data was not sufficient to unequivocally establish the correct subunit arrangement (Cong et al., 2010; Dekker et al., 2011; Martín-Benito et al., 2007).

Understanding the architecture and detailed mechanism of large multisubunit complexes is commonly limited by this inability to obtain high-resolution structural information. In the absence of atomic resolution data, orthogonal structural information is needed for accurate interpretation. An emerging structure determination technique that has the potential to obtain a highly redundant three-dimensional map of constraints is crosslinking coupled with mass-spectrometry (XL-MS; reviewed in Leitner et al., 2010; Rappsilber, 2011). In this approach, the native protein complex is incubated with a crosslinking reagent capable of forming specific covalent bonds with exposed and frequently occurring side chains. Most commonly, amine-reactive reagents, such as disuccinimidyl suberate (DSS), for crosslinking of lysine residues are used, although a variety of reagents have been introduced (Petrotchenko and Borchers, 2010). Next, the complex is proteolytically digested and subjected to MS analysis for identification of the crosslinked peptides (Figure 1A). The crosslinked anchor sites provide a comprehensive three-dimensional map as a framework for molecular modeling. Previously, the application of the XL-MS approach had been limited to individual proteins and small complexes (reviewed by Sinz, 2006). Recent advances in MS instrumentation and the development of more powerful analysis software have permitted the application of XL-MS to a number of increasingly complex assemblies (Bohn et al., 2010; Chen et al., 2010; Maiolica et al., 2007; Schulz et al., 2007). Multisubunit complexes studied by XL-MS include the 26S proteasome (Bohn et al., 2010; Lasker et al., 2012), eukaryotic RNA polymerases (Chen et al., 2010), and the ribosome (Lauber and Reilly, 2011).

We used the XL-MS approach to investigate the order and orientation of the 16 subunits in the 1 MDa complex TRiC/CCT. Structural data of TRiC has been obtained at near-residue resolution, 4.0 and 3.8 Å, by single-particle averaging cryo-electron microscopy (cryo-EM) and X-ray crystallography (Cong et al., 2010; Dekker et al., 2011). The derived models agree in that both rings have a specific subunit order and that the two rings are related by 2-fold symmetry, creating two homomeric contacts across the equator. However, the proposed subunit orders completely disagree (CCT 6-5-1-7-4-8-3-2 vs. CCT 8-4-5-7-1-6-2-3 for Dekker et al., 2011 and Cong et al., 2010, respectively). Here, we resolve this issue by the orthogonal XL-MS approach and present the definite model for the TRiC/CCT structure.

RESULTS

Crosslinking Tandem Mass Spectrometry Approach

Our experimental strategy (Figure 1A) exploited recent advances in chemical crosslinking combined with mass spectrometry (Rinner et al., 2008) to identify residues in close spatial proximity in functionally competent TRiC/CCT complexes. These distance constraints then guided the selection of the most likely subunit arrangement by molecular modeling. The number of distance constraints was maximized by applying this strategy to TRiC purified from two evolutionary distant organisms, *Bos taurus* and *Saccharomyces cerevisiae* (bTRiC and yTRiC). At the peptide level, the complexes from each species are expected to yield virtually unrelated tryptic cleavage products. Furthermore, approximately 40% of the surface lysine positions available for crosslinking are scrambled between the bovine and yeast orthologs, resulting in an improved sampling of the subunit surfaces (Table S1 available online).

The conformation of nucleotide-free TRiC is highly heterogeneous, resulting in greater structural ambiguity. ATP hydrolysis leads to a more compact state, whereby a built-in lid closes over the central TRiC folding chamber (Meyer et al., 2003). To facilitate the subsequent modeling analysis, TRiC was crosslinked following incubation with ATP or ATP+AlF$_x$; both conditions induce the closed conformation for which highly reliable structural models derived from archaeal chaperonins exist (Ditzel et al., 1998; Pereira et al., 2010; Shomura et al., 2004). Native protein complexes were incubated with two different isotopically labeled forms of DSS (Müller et al., 2001), which crosslinks exposed primary amino groups found in lysine side chains and polypeptide N-termini. The complex was then digested with trypsin, and samples enriched for crosslinked peptides (Leitner et al., 2012) were analyzed by capillary liquid chromatography tandem mass spectrometry (LC-MS-MS), and the resulting complex fragment ion spectra were assigned to the corresponding peptide sequences using xQuest (Rinner et al., 2008; Figure 1A). Under our experimental conditions, the extent of lysine modification approached saturation. For example, yTRiC has a total of 334 lysines, and of these, 151 were involved in crosslinks in the corresponding ATP-AlFx data set. Furthermore, many crosslinks were identified by multiple peptide pairs. Overall, we identified 997 peptide pairs across all experiments with an estimated false discovery rate (FDR) of less than 5% (Figure 1B and Table S2). They consisted of 423 *heterotypic* crosslinks, that is, crosslinks between different subunits in the TRiC complex, and 574 *homotypic* crosslinks, that is, crosslinks within the same subunit or between two identical subunits. Of the 423 heterotypic crosslinks, 302 mapped to likely ordered parts of the subunit homology models; these were used for determining the overall topology of the complex (see Mapping of the crosslinks onto structural model section and Figure S2A). The remainder mapped primarily to the unstructured N- and C-terminal tails (Figure S2A).

Integrity of the Complex during Crosslinking

To verify that the complex integrity was not affected by crosslinking, we assessed the conformation of crosslinked and native TRiC by EM and gel electrophoresis (Figures 1C–1E and S1). bTRiC samples incubated with or without ATP or ATP+AlF$_x$ were analyzed before and after DSS treatment. Two-dimensional class averages of cryo-EM single particles of TRiC indicated that the conformations before (Cong et al., 2010) and after crosslinking were virtually indistinguishable at low resolution (Figure 1C, bottom panel). Thus, TRiC integrity was not detectably compromised by crosslinking. SDS-PAGE of DSS-crosslinked TRiC yielded high-molecular-weight species consistent with full crosslinking of all TRiC subunits (Figure 1D). DSS-treated TRiC migrated as a single band in native gels, indicating the stabilization of a coherent complex population (Figure 1E). The faster migration of DSS-treated TRiC is expected due to the overall reduction in charge by the crosslinker. Of note, the ATP and ATP+AlF$_x$-induced closed states exhibited a characteristic mobility shift, consistent with the cryo-EM analysis. Similar results were obtained for yTRiC (Figure S1). We conclude that that the crosslinks identified in this study are derived from structurally intact chaperonin complexes.

Mapping of the Crosslinks onto a Structural Model

The identified intermolecular crosslinks were next employed as spatial constraints to derive the most likely TRiC/CCT subunit arrangement (Figures 2 and S2). Homology models were first generated for each of the eight subunits using the crystal structure of the related archaeal chaperonin from *Methanococcus maripaludis* in the nucleotide-bound state (Pereira et al., 2010). The crosslinked lysine positions obtained in the ATP and ATP+AlFx-induced states were then mapped onto the homology models. Of note, only heterotypic crosslinks that mapped to ordered parts of the structure were used in the subsequent calculations to evaluate the compatibility of different geometries between two different subunits [i.e., crosslinks involving residues in loops of unclear conformation and flexible tails were discarded (see Experimental Procedures and Figure S2A)]. Importantly, identical results were obtained using other archaeal group II chaperonin structures as templates (see Application of XL-MS analysis to the dynamic open state of TRiC; Figure S3). For each pair of crosslinked subunits, the fifteen possible pairwise orientations in the hexadecamer were generated (Figure S2B), and the respective lysine distances were calculated (Figures S2C and S2D). The contour length between two C$_\alpha$ atoms of DSS-crosslinked lysines is approximately 24 Å (Müller et al., 2001). We applied a slightly longer C$_\alpha$-C$_\alpha$ distance cutoff of 30 Å to account for protein dynamics and potential model inaccuracies (see also Application of XL-MS analysis to the dynamic open state of TRiC; Figure S3). We also checked whether these crosslinks were physically possible, eliminating any crosslinks that would traverse the protein core. For the complexes of both species, the same unique TRiC/CCT subunit order, namely CCT 6-8-7-5-2-4-1-3 (Figure 2A; i.e., TCP ζ-θ-η-ε-β-δ-α-γ), was obtained. Both rings are related by 2-fold symmetry, as predicted by previous structural analysis, with CCT6/ζ and CCT2/β engaging in homotypic interring contacts. This subunit arrangement, determined by XL-MS, was thus independently determined from two unrelated data sets for TRiC, from two evolutionarily distant species (Figures 2A and 2B). Of note, the heterotypic crosslinked peptides were different in yTRiC and bTRiC; this likely reflects the variability of surface exposed lysines in the two TRiC complexes (Figures 2C and 2D). The set of unambiguous crosslinks was complete for the closed
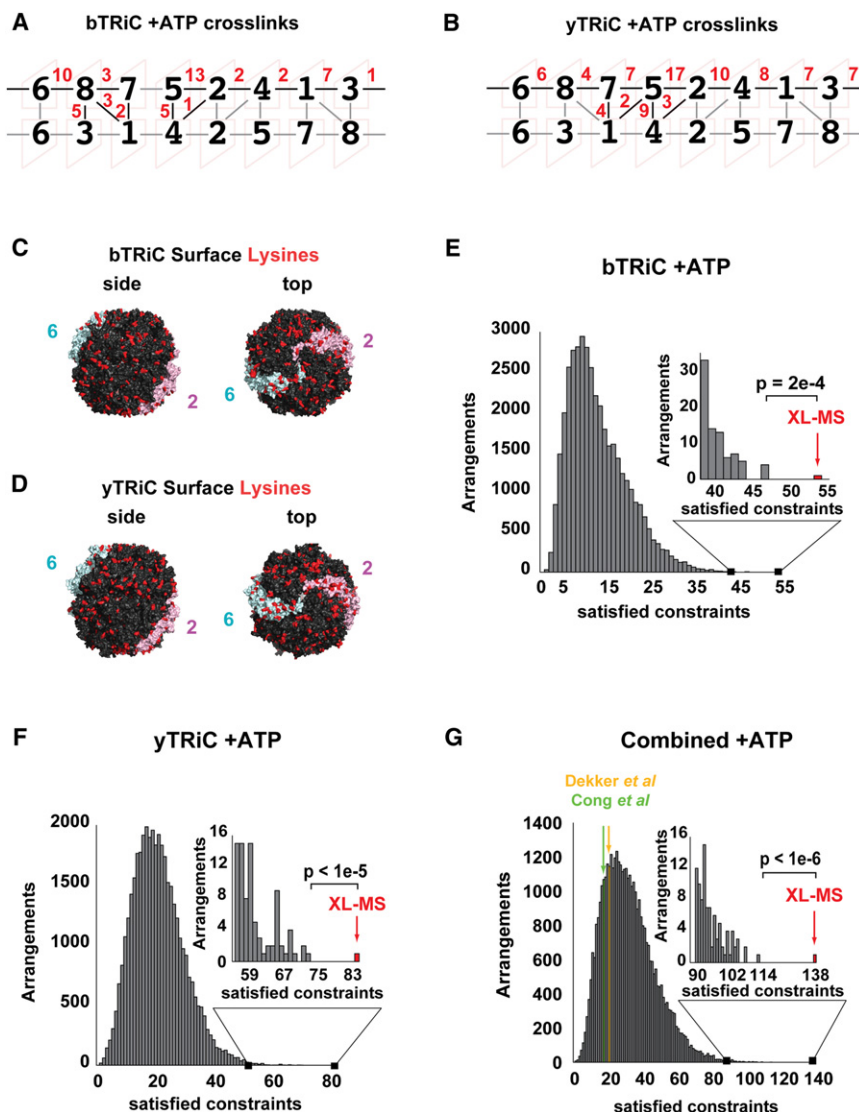
Figure 2. Mass Spectrometry-Derived Constraints Reveal the TRiC Subunit Arrangement

(A and B) Subunit arrangement for (A) bTRiC and (B) yTRiC derived from data sets for the closed state. CCTx subunits are shown as black numbers. The total number of heterotypic crosslinks supporting this arrangement is denoted in red.

(C and D) Surface representation of the bTRiC and yTRiC complexes, showing the surface distribution of lysines (shown in red; see also Table S1); CCT2 (cyan) and CCT6 (pink) are highlighted for orientation.

(E–G) Combinatorial analysis of the heterotypic crosslinking constraints. A histogram showing the distribution of numbers of constraints satisfying the 30 Å cutoff in each conceivable arrangement for closed bTRiC (E), closed yTRiC (F), and the combined data sets (G). (Inset) Right tail of the distribution. The XL-MS arrangement satisfies the largest number of constraints (indicated by red arrow), which are 54 of 64 and 84 of 102 crosslinks for the bTRiC and yTRiC closed-state data sets, respectively; that is, 138 of a total 166 for the combined closed-state data sets. The XL-MS p-value indicates statistical significance over the second-best arrangement. The previously proposed arrangements (Cong et al., 2010; Dekker et al., 2011) are consistent with only 17 (green) and 23 (yellow) of the 166 crosslinks in the combined bTRiC and yTRiC closed-state data sets.

See also Figure S2 and Tables S1 and S3.

conformation of yTRiC. Every directional intraring neighbor-pair relationship and the interring register were established by individual crosslinks (Figure 2B). For bTRiC, only one intraring neighbor pair (CCT5-CCT7) relationship was not directly established by crosslinks (Figure 2A). In case of the closed conformation data set of yTRiC, each intraring subunit contact was established by at least four different crosslinked peptide pairs. Thus, a wrong assignment of any individual neighbor-pair relationship at an FDR of 5% is highly unlikely (probability 6.25×10^{-6} or less). This shows that the assignment must be correct beyond a reasonable doubt.

Combinatorial Analysis of Distance Constraints

The statistical significance of the arrangement determined by XL-MS as the unique solution to the experimental distance constraints was further investigated by an unbiased combinatorial approach that determined the number of constraints satisfied for each of the 40,320 possible subunit arrangements (see Experimental Procedures for details). This approach explicitly

evaluated the ambiguity of several plausible pairs of subunit orientations satisfying a given distance cutoff (see Experimental Procedures for details; Figure 2 and Table S3). The distribution of arrangements satisfying these constraints is shown for both the individual (Figures 2E and 2F) and the combined closed TRiC data sets (Figure 2G) and demonstrates that the arrangement determined by XL-MS is the only subunit ordering that can explain the majority of the heterotypic crosslinks, satisfying 85% (Figure 2E) and 82% (Figure 2F) of the crosslinks for the individual data sets and 83% for the combined data set (Figure 2G). The secondary solutions (see Table S3 and Experimental Procedures for details) are significantly worse than the XL-MS determined arrangement; indeed, the correctness of the XL-MS determined arrangement is statistically significant, relative to the second-best arrangement, with p-values of 2×10^{-4} and $<10^{-5}$, respectively, for the bovine and yeast data sets. Combining the yTRiC and bTRiC data increased the statistical significance of the XL-MS determined arrangement (p-value $<10^{-6}$; Figure 2G) with respect to the second-best arrangement. Importantly, the previously proposed TRiC subunit arrangements (Cong et al., 2010; Dekker et al., 2011) explain only a minor fraction (10% and 13%, respectively) of the observed crosslinks (Figure 2G) and thus are essentially incompatible with our extensive crosslink data set.
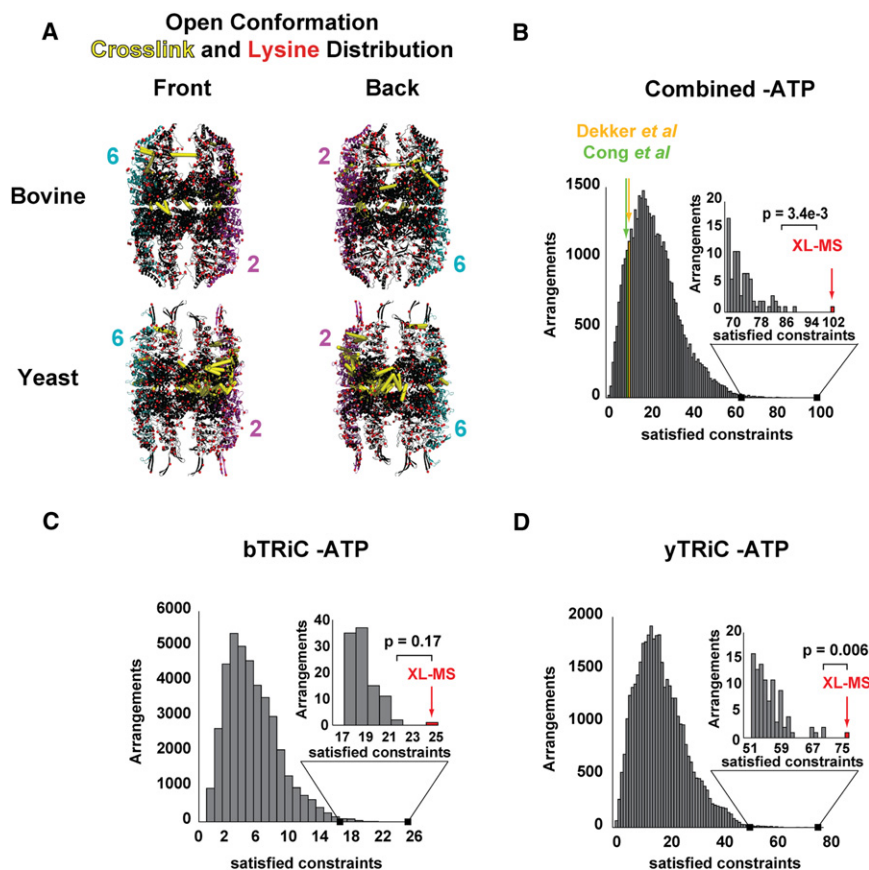
Figure 3. Global Analysis of Mass Spectrometry-Derived Constraints for TRiC in the Open Conformation

(A) Mapping the crosslinked lysines (yellow lines) onto open-state models of bTRiC or yTRiC (colored as in Figures 2C and 2D). The crosslinks preferentially map to the equatorial domains, consistent with increased flexibility of the apical domains in the open state.

(B–D) Combinatorial analysis of heterotypic crosslinking constraints from open conformation data. The number of constraints satisfying the 36 Å cutoff in each conceivable arrangement is shown as a histogram for (B) combined open bTRiC and yTRiC, (C) open bTRiC, and (D) open yTRiC data sets. (Inset) Right tail of the distribution. The XL-MS arrangement satisfies the largest number of constraints (indicated by red arrow); for the three respective data sets these are 102 of 136 (combined), 25 of 36 (bTRiC), and 77 of a total of 100 (yTRiC). The p-value indicates statistical significance of XL-MS over the second-best arrangement. The previously proposed arrangements (Cong et al., 2010; Dekker et al., 2011) are consistent with only ten (green) and eleven (yellow) of the 136 crosslinks in the combined bTRiC and yTRiC closed-state data sets.
See also Figure S3 and Table S3.

Application of XL-MS Analysis to the Dynamic Open State of TRiC

To assess whether the XL-MS and modeling strategy can be applied to structurally less well-defined complexes, we next analyzed crosslinks obtained for the more flexible open state of TRiC without nucleotide using the coordinates of the open state of Mm-Cpn as a model (Pereira et al., 2010; Figure 3). For both bTRiC and yTRiC, a similar number of identified peptide pairs was obtained as in the closed state (Figures 1E and S2A), but fewer constraints passed the 30 Å distance cutoff, particularly for the highly dynamic apical domains (Figure 3A, yellow lines). To account for the increased flexibility of the open state and the lower confidence level of available structural models, the distribution of matching crosslinks over the considered models was computed using a 36 Å distance cutoff (Figures 3B–3D). This analysis also yielded the XL-MS determined arrangement as the best solution, satisfying 75% of the crosslinks (p-values 3.4×10^{-3}, 6.1×10^{-3}, and 0.17 for the combined yTRiC and bTRiC data sets, respectively; see Figures 3B–3D, Table S3, and the Experimental Procedures for details), highlighting the power of our cross-species strategy to model the subunit topology even for structurally flexible, less-characterized complexes. As shown below (Figure S8), these larger distances likely reflect inadequacies of our initial homology model.

To systematically explore how the choice of template and distance cutoff influences our analysis, we next computed the number of satisfied constraints as a function of distance using the different available group II chaperonin structures as templates (Figure S3; Ditzel et al., 1998; Pereira et al., 2010; Shomura et al., 2004). For the closed data sets, this analysis indicated a clear convergence between 24 and 30 Å (Ditzel et al., 1998; Pereira et al., 2010; Shomura et al., 2004; Figures S3B–S3F). Notably, the quality of the optimal arrangement was not sensitive to the exact structural group II chaperonin template employed to build the models (Figure S3). For longer distance cutoffs, the number of satisfied constraints approached the total number of constraints but decreased the discrimination between the optimal arrangement and the median of random solutions (data not shown), supporting our choice of distance cutoff (Figures S3B–S3F).

The Refined XL-MS Structural Model

Prior attempts to generate an accurate structural model for TRiC/CCT were confounded by the low resolution of available cryo-EM and X-ray data. The previous cryo-EM model was based on the visual analysis of density features in the apical domains (Cong et al., 2010). Reanalysis of these cryo-EM data (Cong et al., 2010) with more quantitative and statistical procedures (see Table S4) suggests that the quality of the map suffices for rough backbone tracing but lacks the resolvability to distinguish the highly similar TRiC subunits, so this previous interpretation has to be revoked. The interpretation of X-ray diffraction data from the closed conformation suffers from model bias since no experimental phases are available. We refined a structural model representing the XL-MS determined subunit arrangement against these X-ray diffraction data,

Table 1. Crystallographic Refinement Statistics and Model Geometry

Model	3P9D + 3P9E	XL-MS (Refmac, NCS, No TLS)	XL-MS (Refmac, NCS, TLS)
Resolution limits	30 – 3.8	30 – 3.8	30 – 3.8
R_{work}/R_{free}	0.3178/ 0.3513	0.2696/ 0.3279	0.2568/ 0.3046
Figure of merit	0.672	0.715	0.751
Number of atoms			
Protein	110,444	119,056	119,056
Ligand/ion	784	1,024	1,024
Water	7	0	0
Average B factors			
Protein (Å^2)	141	125	139
Ligand /ion (Å^2)	130	103	123
Water (Å^2)	43	–	–
rmsds			
Bonds (Å)	0.012	0.007	0.007
Angles (°)	0.986	1.052	1.068
Ramachandran plot			
% Preferred (Coot)	85.8	89.5	90.1
% Outliers (Coot)	4.68	3.16	2.89
Number non-Proline *cis* peptides	184	0	0

To allow a fair comparison with the original model (Dekker et al., 2011; PDB codes 3P9E and 3P9D), the XL-MS model was also refined without TLS B-factor parameterization (middle column). The statistics for the Dekker model were determined by Refmac using the default values from CCP4i.

carefully avoiding overt model bias (see the Supplemental Experimental Procedures for details; Dekker et al., 2011). Our final XL-MS structural model has clearly improved refinement statistics and model geometry compared to the published model based on the original subunit topology (Table 1). Strikingly, unanticipated features of the refined XL-MS-based structure provide a rationale for several crosslinks mapping to regions not included in the homology model. Indeed, the refined XL-MS-derived structure could explain approximately 94% of heterotypic (Figures 4A and 4B) and 97% of homotypic (data not shown) crosslinks, according to the 30 Å criterion. This is much better than the thermosome-based homology models. Thus, the XL-MS-based structure explains virtually all experimentally obtained crosslinks; the fraction of outliers corresponds to the 5% FDR for the MS assignment.The XL-MS-derived structure is also more plausible with regard to TRiC sequence features. Our structure accounts for several large insertions unique to individual TRiC subunits, which are well-defined in the electron density. For instance, CCT6 has a unique 10-residue insertion after helix α8 (residues 282–291), which elongates this helix by two turns (Figures 4C and 4D). This feature is clearly discernible in unbiased difference maps (Figures 4D and S4A). The XL-MS model furthermore explains structurally defined distinctive insertions in CCT4 (residues 291–295 and 371–374), CCT1 (residues 341–345 and 484–

495), and CCT6 (residues 481–485; Figures 4C, 4D, and S4B and not shown). In the construction of the Dekker model, these aberrant density features, which are clearly present in the map, had been mostly ignored (Figure S4).

Another striking finding of our model is that most of the N-termini preceding strand β1 are resolved in the density. This revealed two unexpected features, which were validated by crosslinking data. First, we find in our model that CCT4 is the single subunit that has an outward pointing N-terminal density in the map (Figure 5A). In contrast, CCT5 was the corresponding subunit with an outward pointing N-terminus in the original model (Dekker et al., 2011). Strikingly, CCT4 is the only CCT subunit that has a conserved proline at the N-terminal junction to helix α1 (Figures 5B, 5C, and S5A). This provides an evolutionary and structural rationale for why CCT4 is the only CCT subunit with an outward-pointing N-terminus, explaining the aberrant density (Figure 5A). In contrast, CCT5 has a glycine at this position, as do most other TRiC subunits and archaeal subunits (Figure 5C), all of which have inward pointing N-termini (Figure 5A). Of note, the outward conformation of the CCT4 N-terminus is strongly corroborated by a series of crosslinks within our data set, establishing contacts of K12 and K14 to residues on the complex exterior (Figure 5D). These crosslinks are incompatible with an inward-facing N-terminus but are entirely consistent with the subunit docking and the CCT4 sequence data. Similarly, crosslinks between the N-terminus of CCT5 and residues on the cavity walls support the location of the CCT5 N-terminus inside the complex (Figure 5E). Altogether, these observations ascertain the validity of the XL-MS model.

The XL-MS-based TRiC structure also provides unanticipated insights into interring interactions between the N-termini of CCT1 and CCT8. In the crystal structure there is an extensive direct interaction between the N-termini of the CCT8 subunits across the equator (Figure S5B). Perhaps these unique structural features help to correctly establish the subunit topology in TRiC by stabilizing the ring-ring interface. They might also contribute to allosteric rearrangements during the functional cycle. The extensive interactions between the CCT8 N-termini are consistent with previous crosslinking and 2D gel data (Cong et al., 2010), which had suggested direct contacts between CCT8 subunits (Figures S5C and S5D). Indeed, all the crosslinks observed in Cong et al., 2010, which by themselves are insufficient to unambiguously determine the correct arrangement, are fully consistent and explained by the XL-MS architecture.

In Vivo Validation of XL-MS Architecture Using Disulfide Mapping

To independently validate the intraring subunit order and interring subunit register determined by XL-MS, we next employed in vivo near-neighbor disulfide engineering (Figures 6 and S6). The XL-MS-determined arrangement predicts that subunits CCT2 and CCT6 form interring homotypic contacts (Figures 2A, 2B, and 6A). Previous models predict homotypic contacts for either CCT4 and CCT6 (Dekker et al., 2011) or CCT1 and CCT8 (Cong et al., 2010; Figure 6A). We engineered cysteine pairs at residues predicted to be proximal (C_α-C_α < 6 Å) in a homotypic interring interface, thus permitting disulfide bond
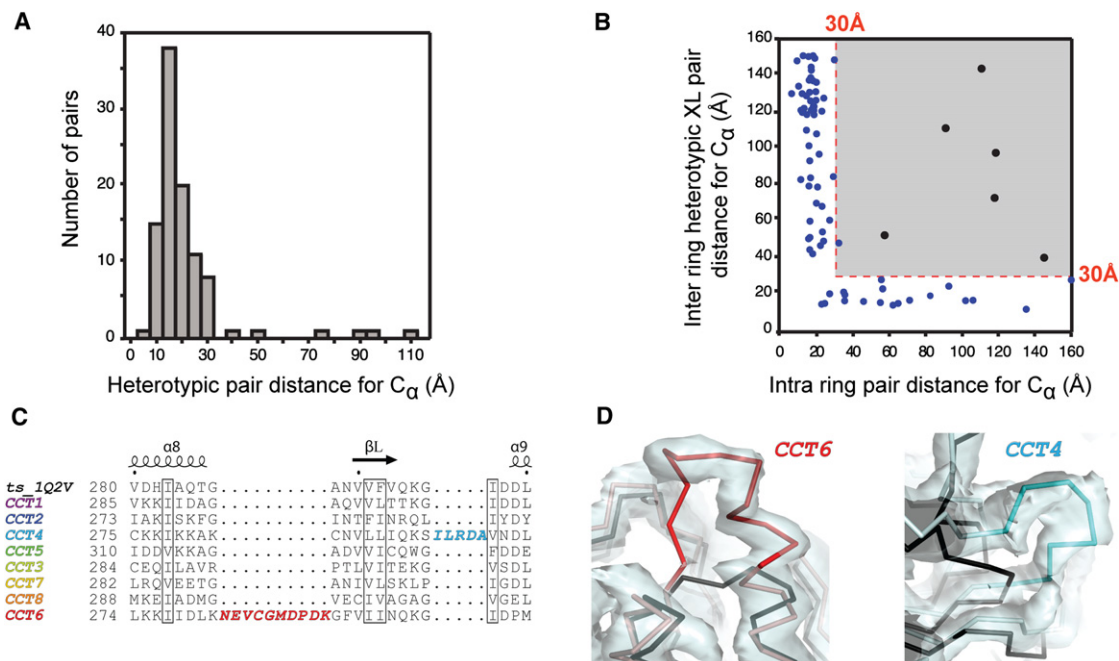
Figure 4. Crossvalidation of Crystal Structure and Crosslink Data for yTRiC

(A) Distance distribution for the closed-state yTRiC heterotypic crosslink data set. The median heterotypic Cα-Cα crosslink distance in the model is 16.4 Å.

(B) Heterotypic crosslink Cα pair distances for inter- and intraring subunit pairings observed in the refined XL-MS-based crystal structure. The crosslinks compatible with the XL-MS arrangement are highlighted in blue; crosslinks mapping to the gray box exceed the cutoff.

(C) Alignment showing unique insertions in yTRiC subunits CCT6 and CCT4.

(D) Unbiased 2Fo-Fc electron density for these insertions at 1 σ. The thermosome structure is shown in black for comparison.

See also Figure S4 and Table S4.

formation (Figures 6B and 6C). Importantly, the yTRiC interring interface is otherwise free of cysteines. The CCTx-(Cys)$_2$ genes supported normal growth of yeast lacking the corresponding wild-type gene (Figure S6A). Disulfide crosslinking of TRiC obtained from CCTx-(Cys)$_2$ cells was induced by oxidation with CuCl$_2$ (Figures 6D and S6D). As predicted by the XL-MS-based model, disulfide-crosslinked dimers occurred in a time- and oxidant-dependent manner only in TRiC from CCT2-(Cys)$_2$ and CCT6-(Cys)$_2$ cells (Figures 6E and 6F). No such dimers were observed for CCT1-(Cys)$_2$, CCT4-(Cys)$_2$, and CCT8-(Cys)$_2$ (Figures 6F, 6H, and 6I), indicating that these subunits do not form homotypic contacts in TRiC. In conjunction with the wealth of evidence from the crosslinking distance constraints and crystallographical analysis, this orthogonal in vivo approach definitively validates the XL-MS-derived arrangement as the correct topology of TRiC across eukaryotes.

DISCUSSION

Previous attempts to define the TRiC topology have been mired in controversy because of the pseudosymmetry of the complex and confounded by methodological limitations. To resolve this long-standing problem, we developed and applied a crosslinking tandem mass spectrometry approach to generate two complete and self-consistent sets of constraints to model the topology of the eukaryotic chaperonin TRiC. These data unambiguously assign the intraring subunit order in the TRiC complex and inval-

idate the previously proposed arrangements. Importantly, the XL-MS-derived model is also consistent with previous crosslinking data (Figure S5) and likely compatible with the subunit spacing derived from 3D cryo-EM reconstructions of TRiC decorated with antibodies (Martín-Benito et al., 2007).

Importantly, the prior models of TRiC are entirely incompatible with our data, because their subunit orders diverge significantly from ours (Cong et al., 2010; Dekker et al., 2011). Figure 7 shows the crosslinks obtained from the closed conformation of yTRiC or bTRiC mapped onto the three respective final structure models. It is evident that, whereas the XL-MS model explains ~95% of the obtained crosslinks, only a small fraction of the crosslinks fit the previous models. The few consistent intersubunit crosslinks locate close to the apical pore, where all eight subunits meet, that is, these ambiguous crosslinks fit to the majority of conceivable subunit topologies. In contrast, XL-MS data is consistent with the previously reported crosslinking data from Cong et al., 2010, which alone cannot discriminate between the Cong et al., 2010, and XL-MS-derived models (Figure S5).

The subunit docking into the density of the original crystallographic yTRiC model seemed to be corroborated by antibody binding to a FLAG epitope fused to the exposed N-terminus of CCT5 in the presence of ATP (Dekker et al., 2011). However, yeast has an anomalously long CCT5 N-terminal peptide that could easily reach out from the cavity through the apical opening (Figure S7). Because pore closure in TRiC is not stringently induced by the addition of only ATP, it allows transient exposure
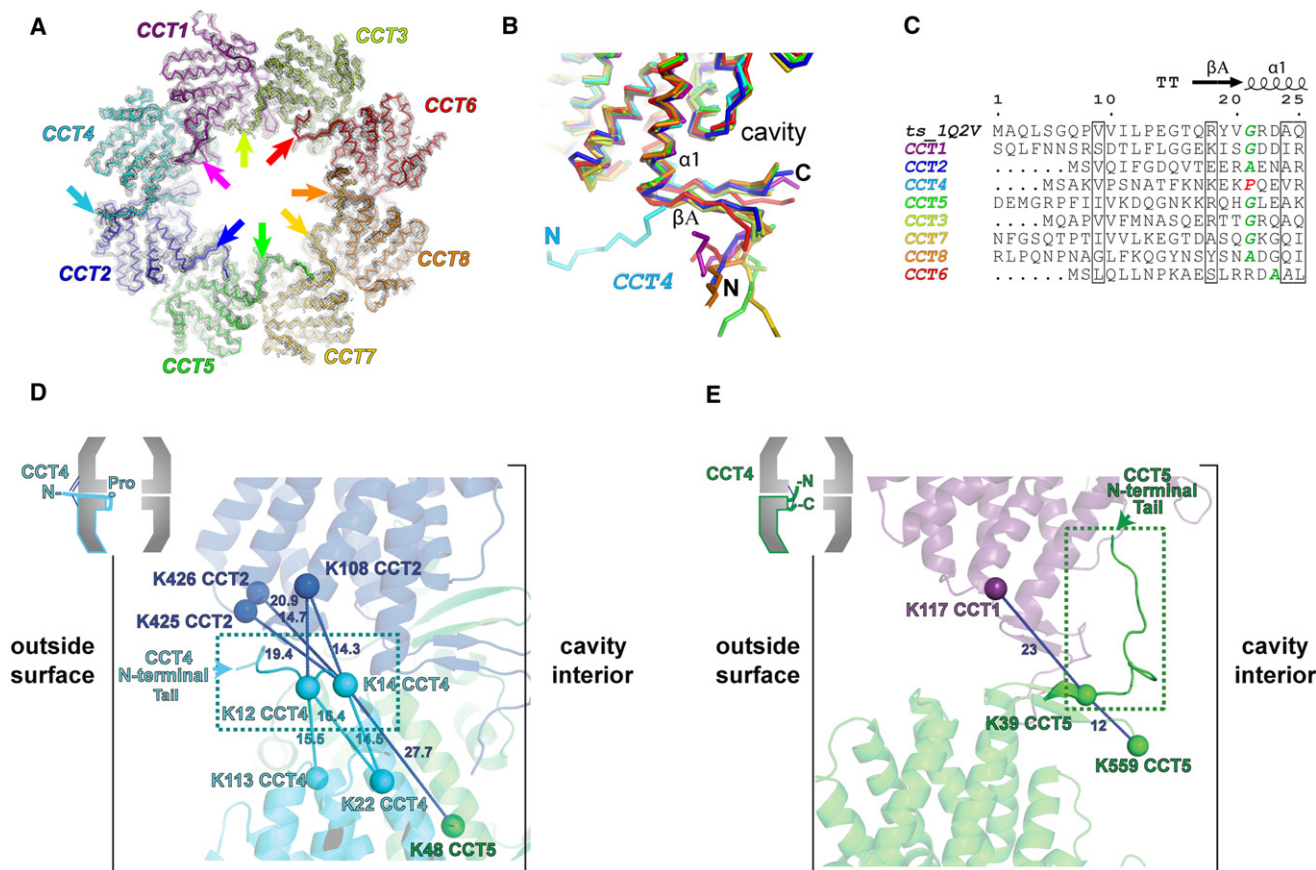
Figure 5. Features of the TRiC Crystal Structure Model Based on the XL-MS Subunit Order

(A) Electron density for XL-MS crystal structure model. The view from the equator shows the cavity of one ring. The final 2Fo-Fc density at 1.5 σ is shown as meshwork. The N-terminal β strands of TRiC subunits 1, 2, 3, 5, 6, 7 and 8 are highlighted by arrows. The N-terminus of CCT4 (cyan) is inserted between CCT4 and CCT2. Please note that side-chain density is hardly visible at all and thus cannot be used for sequence docking.

(B) Superposition of the yeast TRiC subunits, highlighting the aberrant CCT4 geometry at the N-terminus (cyan).

(C) Alignment of the N-terminal sequences of the thermosome and the yTRiC subunits. The junction residue between βA and α1 is shown in italics, highlighting residues compatible (green) or incompatible (red) with the thermosome geometry. The sharp transition is also facilitated by small helix residues facing the β strands, as observed in CCT6. Numbering and secondary structure elements refer to the thermosome structure (PDB code 1Q2V; Shomura et al., 2004).

(D) Validation of the CCT4 N-terminus geometry by crosslinking. The location of the CCT4 N-terminal tail (dashed box) is corroborated by specific crosslinks to residues on the outside surface. The backbones of CCT2, CCT4, and CCT5 are shown in blue, cyan, and green, respectively. The Cα atoms of lysines are shown as spheres, and crosslinks in between lysine Cα atoms are shown by dashed lines. The distance between lysine Cα's is denoted in Å.

(E) Localization of the CCT5 N-terminus in the cavity by crosslinking. The location of the CCT5 N-terminal tail (dashed box) is corroborated by specific crosslinks to residues on the cavity surface. CCT1 and CCT5 are indicated in magenta and green, respectively.

See also Figure S5.

to the antibody, which would explain the reported experimental result. Our crosslinking data on the closed conformation of yTRiC unambiguously show that the N-terminus of CCT4 is located on the exterior surface of the complex, close to the equator of the complex (Figure 5D), whereas the N-terminal segment of CCT5 was involved in crosslinks to the interior (Figure 5E). Taken together with the conserved proline in the CCT4 N-terminus, this provides strong evidence for the XL-MS model and against the Dekker subunit docking.

The XL-MS-derived model of the eukaryotic chaperonin uncovers unexpected structural features instrumental to understand its function. Strikingly, it shows that the conserved and highly charged surface of the closed chamber of TRiC has a conspicuous segregation of positive and negative charges

contributed by subunits CCT5-2-4 and CCT3-6-8, respectively, and results in a bipolar distribution within the folding chamber (Figures 8A and 8B). The high conservation of the inner surface suggests functional importance in the folding of encapsulated substrate proteins (Figure 8C). Indeed, the bacterial chaperonin GroEL has a negatively charged chamber that is critical for folding (Tang et al., 2008). In comparison, the charge patterning on the outside surface of TRiC is less conserved (Figures 8D–8F). The least conservation within the chamber occurs at the interface between the positive and negative hemispheres, likely reflecting interspecies variation in the charge asymmetry boundaries (see arrow in Figure 8C).

An interesting feature that is shared between the EM and X-ray structures of the open TRiC conformations is pairwise
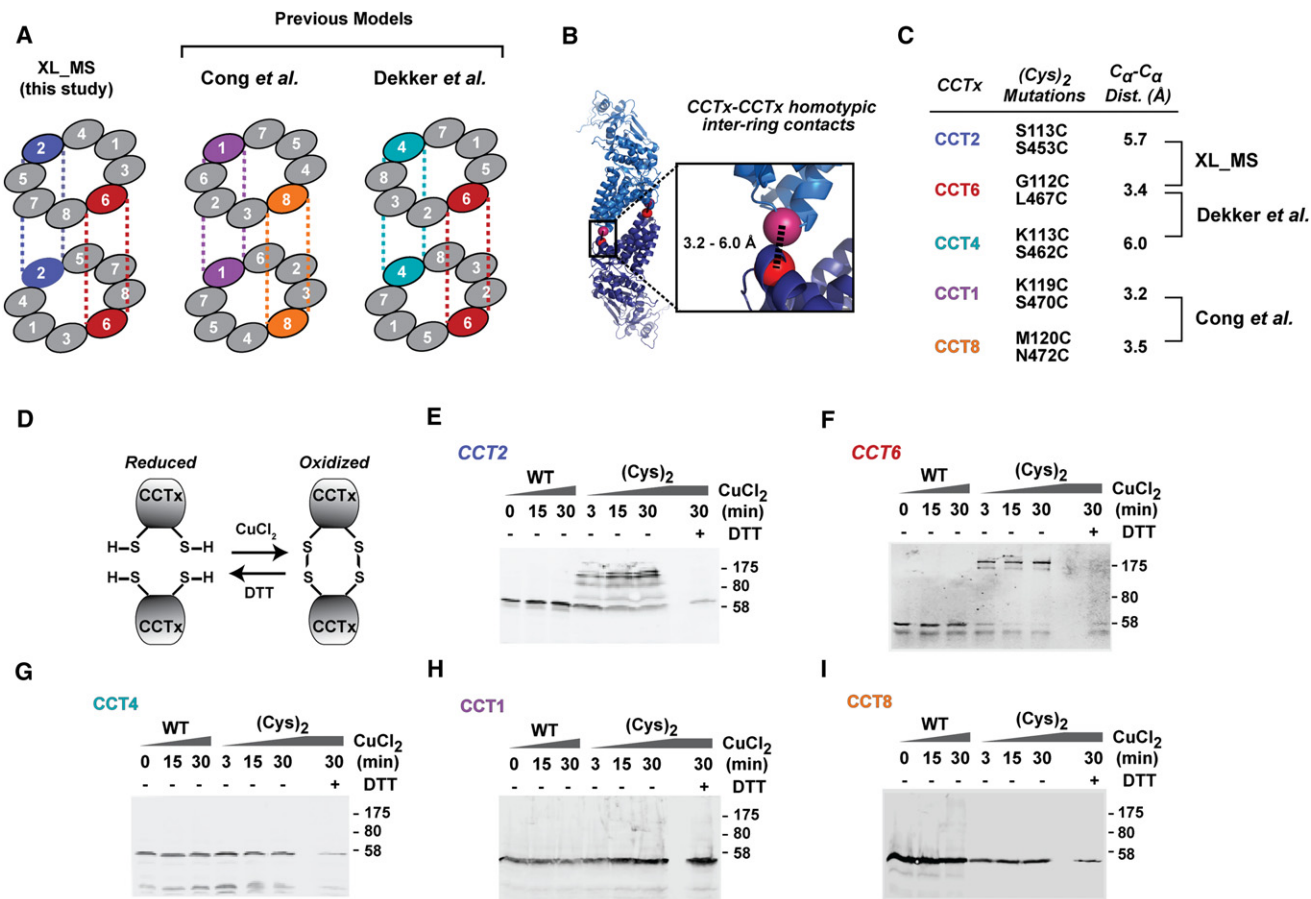
Figure 6. In Vivo Validation of the Interring Register Using Disulfide Crosslinking

(A) Different TRiC models predict distinct pairs of homotypic contacts. XL_MS (this study) proposes CCT2 and CCT6 interring pairs; previous studies proposed CCT1 and CCT8 (Cong et al., 2010) or CCT4 and CCT6 pairs (Dekker et al., 2011).

(B) Model of the interring interface highlighting residues substituted by cysteines for disulfide bond formation.

(C) Summary of relevant cysteine replacements and inter-Cys distances. All CCTx-$(Cys)_2$ subunits support wild-type growth (Figure S6).

(D) Near-neighbor disulfide mapping. Symmetrically related cysteine pairs will form disulfide bonds under oxidizing conditions ($CuCl_2$), which are reversed with the reducing agent DTT.

(E–I) Incubation under oxidizing conditions reveals that subunits CCT2-$(Cys)_2$ and CCT6-$(Cys)_2$ form DTT-sensitive disulfide dimers, whereas WT subunits and the $(Cys)_2$ variants of subunits CCT4, CCT1, and CCT8 do not.

See also Figure S6.

association of the apical domains, yielding a 4-fold pseudosymmetry (Cong et al., 2011; Muñoz et al., 2011; Figure S8). This is also apparent in our open conformation data sets. In the yTRiC data set, we find multiple crosslinks between the apical domains of CCT1-3 (four crosslinks), CCT6-8 (two crosslinks), CCT7-5 (three crosslinks), and CCT2-4 (six crosslinks) but only one or no crosslinks for the other apical intraring pairs. The pattern is less pronounced in the bTRiC open-state data set. These open-state apical domain contacts may help propagate allosteric rearrangements throughout the ring (Reissmann et al., 2007; Rivenzon-Segal et al., 2005).

In the light of the XL-MS-derived topology, earlier data on CCT-substrate and CCT-cofactor complexes will have to be re-interpreted (Dekker et al., 2011; Llorca et al., 1999, 2000; Muñoz et al., 2011; Cuéllar et al., 2008; Martín-Benito et al., 2004). Here, we examine only the crystallographic information on tubulin

binding (Muñoz et al., 2011). The position of the 2-fold interring axis cannot be directly derived from the crystal structure of the TRiC-tubulin complex because of extensive disorder in one ring (Muñoz et al., 2011). However, comparison with the EM structure of TRiC in the open conformation (Cong et al., 2011) suggests that the subunit with the most retracted apical domain orients perpendicular to the axis (subunit 3 in Cong et al., 2011, chain G in Muñoz et al., 2011), that is, should be assigned either CCT1 or CCT7, and consequently the tubulin density sits on top of the axis. The reported crosslink between tubulin and the C terminus of CCT2 (Muñoz et al., 2011) suggests that tubulin interacts with the equatorial domains of TRiC subunits CCT5-2-4 and the aberrant apical domain belongs to CCT7 (Figure S8B). Interestingly, tubulin appears to bind near the negatively charged region of the cavity. In contrast, we could not detect meaningful density for actin in the cavity of the
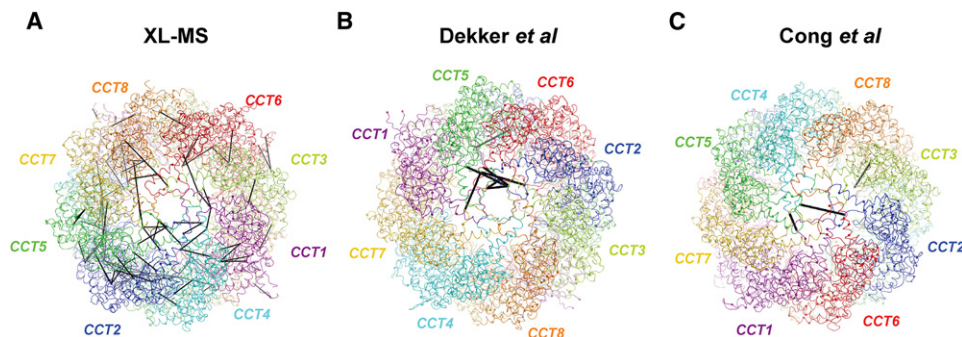
Figure 7. Consistency of TRiC Structural Models with Crosslinking Data

Heterotypic crosslinks obeying the 30 Å criterion were mapped onto ribbon representations of the XL-MS (A), Dekker et al., 2011 (B), and Cong et al., 2010 (C) structural models of TRiC.

See also Figure S7 and Movie S1.

closed-state crystal structure, unlike previously reported (Dekker et al., 2011). This suggests that TRiC-associated actin present in the crystal may be poorly ordered.

The unequivocal solution to the TRiC/CCT topology will prove critical to understand its assembly, mechanism, and allosteric regulation. The XL-MS-derived model reveals a surprising degree of asymmetry in this ring-shaped chaperonin, for the surface properties of the chamber, and probably also for allosteric transitions and substrate binding. The conserved hetero-oligomeric structure of TRiC provides the structural basis for these asymmetric features. This study highlights the power of mass-spectrometry-guided approaches to facilitate structural modeling of hetero-oligomeric complexes. Accurate model building of many large dynamic macromolecular complexes using data from X-ray crystallography and cryo-EM alone is often extremely difficult. The successful application to the challenging case of the pseudosymmetrical TRiC/CCT suggests that XL-MS, in combination with low-resolution structural data and computational modeling can reveal the topology of other complexes, even if they consist of highly homologous subunits.

EXPERIMENTAL PROCEDURES

bTRiC was purified as described previously (Feldman et al., 2003); yTRiC was affinity-purified using His$_6$- and Strep-tagged Plp2p, followed by Heparin affinity and Superose-6 size exclusion chromatography. DSS-treated TRiC complexes were characterized by SDS-PAGE, native-PAGE, and cryo-EM 2D class averages to confirm the structural integrity of the crosslinked complex. DSS-crosslinked TRiC samples were treated with trypsin, enriched for crosslinked peptides by size exclusion chromatography, and analyzed by tandem mass spectrometry. Crosslinked peptides were identified by xQuest (Rinner et al., 2008). The anchor lysine residues were mapped onto homology models of bTRiC and yTRiC subunits arranged in all pairwise subunit combinations (representing 15 possible spatial orientations), and Cα-Cα distances were computed. The distance matrix was used to evaluate all possible arrangements of the hexadecameric complex and deduce the best arrangement. A parametric bootstrap test was used to evaluate the significance of the best with respect to the second-best arrangements as simulated according to a binomial distribution function. Plasmids of the indicated yTRiC subunits containing introduced cysteine pairs (Cys)$_2$ at putative homotypic interface contacts were inserted in the respective cctxΔ by plasmid shuffling; the corresponding TRiC complexes were tested for the formation of specific disulfide bonds using SDS-PAGE and western blot. The XL-MS topology model was refined against the deposited crystal structure factors (Dekker

et al., 2011) using Refmac (Murshudov et al., 1997). For manual model editing, Coot was employed (Emsley and Cowtan, 2004). Ortholog CCT sequences were retrieved from the National Center for Biotechnology Information (Sayers et al., 2009) and aligned using ClustalW (Thompson et al., 1994); the conservation scores were calculated using Rate4site (Pupko et al., 2002), mapped onto the XL-MS structure using Consurf (Ashkenazy et al., 2010), and visualized using Pymol (http://www.pymol.org).

ACCESSION NUMBERS

Atomic coordinates for the refined XL-MS-derived structure of the yeast TRiC complex have been deposited in the Protein Data Bank (http://www.pdb.org; PDB codes 4D8Q and 4D8R, respectively). The structure factor amplitudes were previously deposited under the accession code 3P9D (Dekker et al., 2011).

SUPPLEMENTAL INFORMATION

Supplemental Information includes eight figures, four tables, Supplemental Experimental Procedures, and one movie and can be found with this article online at doi:10.1016/j.str.2012.03.007.

Figure 8. Chemical Properties of the XL-MS-Derived TRiC Chaperonin Structure

Analysis of the electrostatic charge distribution of yeast (A and D) and bovine (B and E) TRiC complexes.

(A and B) The folding chamber for yTRiC and bTRiC reveals a striking asymmetry of charged residues on the inside of the cavity, where subunits CCT1-CCT3-CCT6-CCT8 are positively charged (blue) and subunits CCT7-CCT5-CCT2-CCT4 are neutral (white) or acidic (red).

(D and E) In contrast, the outside surface of yTRiC and bTRiC shows moderate conservation of charged residues.

(C and F) Surface conservation of TRiC. The similarity scores from aligning each 100 orthologous sequences were mapped onto the yTRiC structure. A color gradient from green to red indicates decreasing conservation. The internal cavity surface is strikingly conserved. Interestingly, interfacial regions between pairs of subunits (CCT4/CCT1 and CCT7/CCT8) are less conserved as indicated by arrows. Consistent with the charge variability between bTRiC and yTRiC, the outside surface of the TRiC complex is not highly conserved across orthologs.

See also Figure S8.

REFERENCES

Archibald, J.M., Blouin, C., and Doolittle, W.F. (2001). Gene duplication and the evolution of group II chaperonins: implications for structure and function. J. Struct. Biol. *135*, 157–169.

Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res. *38* (Web Server issue), W529–W533.

Bohn, S., Beck, F., Sakata, E., Walzthoeni, T., Beck, M., Aebersold, R., Förster, F., Baumeister, W., and Nickell, S. (2010). Structure of the 26S proteasome from Schizosaccharomyces pombe at subnanometer resolution. Proc. Natl. Acad. Sci. USA *107*, 20992–20997.

Chen, Z.A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kamenski, T., Rasmussen, M., Lariviere, L., Bukowski-Wills, J.C., Nilges, M., et al. (2010).

Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. EMBO J. *29*, 717–726.

Cong, Y., Baker, M.L., Jakana, J., Woolford, D., Miller, E.J., Reissmann, S., Kumar, R.N., Redding-Johanson, A.M., Batth, T.S., Mukhopadhyay, A., et al. (2010). 4.0-A resolution cryo-EM structure of the mammalian chaperonin TRiC/CCT reveals its unique subunit arrangement. Proc. Natl. Acad. Sci. USA *107*, 4967–4972.

Cong, Y., Schröder, G.F., Meyer, A.S., Jakana, J., Ma, B., Dougherty, M.T., Schmid, M.F., Reissmann, S., Levitt, M., Ludtke, S.L., et al. (2011). Symmetry-free cryo-EM structures of the chaperonin TRiC along its ATPase-driven conformational cycle. EMBO J. *31*, 720–730.

Cuéllar, J., Martín-Benito, J., Scheres, S.H., Sousa, R., Moro, F., López-Viñas, E., Gómez-Puertas, P., Muga, A., Carrascosa, J.L., and Valpuesta, J.M. (2008). The structure of CCT-Hsc70 NBD suggests a mechanism for Hsp70 delivery of substrates to the chaperonin. Nat. Struct. Mol. Biol. *15*, 858–864.

Dekker, C., Roe, S.M., McCormack, E.A., Beuron, F., Pearl, L.H., and Willison, K.R. (2011). The crystal structure of yeast CCT reveals intrinsic asymmetry of eukaryotic cytosolic chaperonins. EMBO J. *30*, 3078–3090.

Ditzel, L., Löwe, J., Stock, D., Stetter, K.O., Huber, H., Huber, R., and Steinbacher, S. (1998). Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. Cell 93, 125–138.

Douglas, N.R., Reissmann, S., Zhang, J., Chen, B., Jakana, J., Kumar, R., Chiu, W., and Frydman, J. (2011). Dual action of ATP hydrolysis couples lid closure to substrate release into the group II chaperonin chamber. Cell 144, 240–252.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132.

Feldman, D.E., Spiess, C., Howard, D.E., and Frydman, J. (2003). Tumorigenic mutations in VHL disrupt folding in vivo by interfering with chaperonin binding. Mol. Cell 12, 1213–1224.

Hartl, F.U., Bracher, A., and Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. Nature 475, 324–332.

Lasker, K., Förster, F., Bohn, S., Walzthoeni, T., Villa, E., Unverdorben, P., Beck, F., Aebersold, R., Sali, A., and Baumeister, W. (2012). Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. Proc. Natl. Acad. Sci. USA 109, 1380–1387.

Lauber, M.A., and Reilly, J.P. (2011). Structural analysis of a prokaryotic ribosome using a novel amidinating cross-linker and mass spectrometry. J. Proteome Res. 10, 3604–3616.

Leitner, A., Walzthoeni, T., Kahraman, A., Herzog, F., Rinner, O., Beck, M., and Aebersold, R. (2010). Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. Mol. Cell. Proteomics 9, 1634–1649.

Leitner, A., Reischl, R., Walzthoeni, T., Herzog, F., Bohn, S., Förster, F., and Aebersold, R. (2012). Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. Mol. Cell. Proteomics 11, M111.014126.

Liou, A.K., and Willison, K.R. (1997). Elucidation of the subunit orientation in CCT (chaperonin containing TCP1) from the subunit composition of CCT micro-complexes. EMBO J. 16, 4311–4316.

Llorca, O., McCormack, E.A., Hynes, G., Grantham, J., Cordell, J., Carrascosa, J.L., Willison, K.R., Fernandez, J.J., and Valpuesta, J.M. (1999). Eukaryotic type II chaperonin CCT interacts with actin through specific subunits. Nature 402, 693–696.

Llorca, O., Martín-Benito, J., Ritco-Vonsovici, M., Grantham, J., Hynes, G.M., Willison, K.R., Carrascosa, J.L., and Valpuesta, J.M. (2000). Eukaryotic chaperonin CCT stabilizes actin and tubulin folding intermediates in open quasi-native conformations. EMBO J. 19, 5971–5979.

Llorca, O., Martín-Benito, J., Gómez-Puertas, P., Ritco-Vonsovici, M., Willison, K.R., Carrascosa, J.L., and Valpuesta, J.M. (2001). Analysis of the interaction between the eukaryotic chaperonin CCT and its substrates actin and tubulin. J. Struct. Biol. 135, 205–218.

Maiolica, A., Cittaro, D., Borsotti, D., Sennels, L., Ciferri, C., Tarricone, C., Musacchio, A., and Rappsilber, J. (2007). Structural analysis of multiprotein complexes by cross-linking, mass spectrometry, and database searching. Mol. Cell. Proteomics 6, 2200–2211.

Martín-Benito, J., Bertrand, S., Hu, T., Ludtke, P.J., McLaughlin, J.N., Willardson, B.M., Carrascosa, J.L., and Valpuesta, J.M. (2004). Structure of the complex between the cytosolic chaperonin CCT and phosducin-like protein. Proc. Natl. Acad. Sci. USA 101, 17410–17415.

Martín-Benito, J., Grantham, J., Boskovic, J., Brackley, K.I., Carrascosa, J.L., Willison, K.R., and Valpuesta, J.M. (2007). The inter-ring arrangement of the cytosolic chaperonin CCT. EMBO Rep. 8, 252–257.

Meyer, A.S., Gillespie, J.R., Walther, D., Millet, I.S., Doniach, S., and Frydman, J. (2003). Closing the folding chamber of the eukaryotic chaperonin requires the transition state of ATP hydrolysis. Cell 113, 369–381.

Müller, D.R., Schindler, P., Towbin, H., Wirth, U., Voshol, H., Hoving, S., and Steinmetz, M.O. (2001). Isotope-tagged cross-linking reagents. A new tool in mass spectrometric protein interaction analysis. Anal. Chem. 73, 1927–1934.

Muñoz, I.G., Yébenes, H., Zhou, M., Mesa, P., Serna, M., Park, A.Y., Bragado-Nilsson, E., Beloso, A., de Cárcer, G., Malumbres, M., et al. (2011). Crystal

structure of the open conformation of the mammalian chaperonin CCT in complex with tubulin. Nat. Struct. Mol. Biol. 18, 14–19.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. 53, 240–255.

Pereira, J.H., Ralston, C.Y., Douglas, N.R., Meyer, D., Knee, K.M., Goulet, D.R., King, J.A., Frydman, J., and Adams, P.D. (2010). Crystal structures of a group II chaperonin reveal the open and closed states associated with the protein folding cycle. J. Biol. Chem. 285, 27958–27966.

Petrotchenko, E.V., and Borchers, C.H. (2010). Crosslinking combined with mass spectrometry for structural proteomics. Mass Spectrom. Rev. 29, 862–876.

Pupko, T., Bell, R.E., Mayrose, I., Glaser, F., and Ben-Tal, N. (2002). Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. Bioinformatics 18 (Suppl 1), S71–S77.

Rappsilber, J. (2011). The beginning of a beautiful friendship: cross-linking/ mass spectrometry and modelling of proteins and multi-protein complexes. J. Struct. Biol. 173, 530–540.

Reissmann, S., Parnot, C., Booth, C.R., Chiu, W., and Frydman, J. (2007). Essential function of the built-in lid in the allosteric regulation of eukaryotic and archaeal chaperonins. Nat. Struct. Mol. Biol. 14, 432–440.

Rinner, O., Seebacher, J., Walzthoeni, T., Mueller, L.N., Beck, M., Schmidt, A., Mueller, M., and Aebersold, R. (2008). Identification of cross-linked peptides from large sequence databases. Nat. Methods 5, 315–318.

Rivenzon-Segal, D., Wolf, S.G., Shimon, L., Willison, K.R., and Horovitz, A. (2005). Sequential ATP-induced allosteric transitions of the cytoplasmic chaperonin containing TCP-1 revealed by EM analysis. Nat. Struct. Mol. Biol. 12, 233–237.

Sayers, E.W., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., DiCuccio, M., Edgar, R., Federhen, S., et al. (2009). Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 37 (Database issue), D5–D15.

Schulz, D.M., Kalkhof, S., Schmidt, A., Ihling, C., Stingl, C., Mechtler, K., Zschörnig, O., and Sinz, A. (2007). Annexin A2/P11 interaction: new insights into annexin A2 tetramer structure by chemical crosslinking, high-resolution mass spectrometry, and computational modeling. Proteins 69, 254–269.

Shomura, Y., Yoshida, T., Iizuka, R., Maruyama, T., Yohda, M., and Miki, K. (2004). Crystal structures of the group II chaperonin from Thermococcus strain KS-1: steric hindrance by the substituted amino acid, and inter-subunit rearrangement between two crystal forms. J. Mol. Biol. 335, 1265–1278.

Sinz, A. (2006). Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. Mass Spectrom. Rev. 25, 663–682.

Spiess, C., Meyer, A.S., Reissmann, S., and Frydman, J. (2004). Mechanism of the eukaryotic chaperonin: protein folding in the chamber of secrets. Trends Cell Biol. 14, 598–604.

Spiess, C., Miller, E.J., McClellan, A.J., and Frydman, J. (2006). Identification of the TRiC/CCT substrate binding sites uncovers the function of subunit diversity in eukaryotic chaperonins. Mol. Cell 24, 25–37.

Tang, Y.C., Chang, H.C., Chakraborty, K., Hartl, F.U., and Hayer-Hartl, M. (2008). Essential role of the chaperonin folding compartment in vivo. EMBO J. 27, 1458–1468.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Yam, A.Y., Xia, Y., Lin, H.T., Burlingame, A., Gerstein, M., and Frydman, J. (2008). Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies. Nat. Struct. Mol. Biol. 15, 1255–1262.

Zhang, J., Baker, M.L., Schröder, G.F., Douglas, N.R., Reissmann, S., Jakana, J., Dougherty, M., Fu, C.J., Levitt, M., Ludtke, S.J., et al. (2010). Mechanism of folding chamber closure in a group II chaperonin. Nature 463, 379–383.

Supplemental Information

The Molecular Architecture

of the Eukaryotic Chaperonin TRiC/CCT

Alexander Leitner, Lukasz A. Joachimiak, Andreas Bracher, Leonie Mönkemeyer, Thomas Walzthoeni, Bryan Chen, Sebastian Pechmann, Susan Holmes, Yao Cong, Boxue Ma, Steve Ludtke, Wah Chiu, F. Ulrich Hartl, Ruedi Aebersold, and Judith Frydman

<u>Inventory of Supplemental Information</u>

Supplemental Figures

A

Native-PAGE (Coomassie stain)

B

SDS-PAGE (Coomassie stain)

C

EM (negative stain)

-DSS, -ATP/AlFx, 30'

+DSS, -ATP/AlFx, 30'

Figure S1. Yeast TRiC/CCT structural integrity is not compromised by DSS crosslinking, related to Figure 1. (A) Non-denaturing gradient PAGE analysis of yeast TRiC/CCT in presence or absence of ATP+AlF$_x$ during exposure to DSS crosslinking reagent. DMSO was used as negative control. Incubation with DSS results in increased electrophoretic mobility, probably because each modification of a lysine residue removes a positive charge from TRiC. In complex with ATP+AlF$_x$, TRiC undergoes a compaction to the closed conformation, resulting in considerably increased electrophoretic mobility. Protease digestion and mass spectrometry analyses were carried out after exposure to DSS for 30 min under the same conditions. (B) SDS-PAGE analysis of the same samples. The eight TRiC/CCT subunits form a stack of bands at ~60 kDa apparent molecular weight. Treatment with DSS diminished the intensity of these bands. The crosslinking products mostly exceeded the size range and were not resolved at the top of the gel. Selected molecular weight marker bands are indicated. (C) Negative stain electron microscopy (EM) analysis of crosslinked TRiC/CCT. Samples of apo-TRiC with or without DSS treatment for 30 min were analyzed by uranyl acetate negative stain EM. The scale bar indicates 20 nm. No gross morphological differences were observed between samples, suggesting continuing structural integrity. Close inspection of the grid further showed no evidence for frequent crosslinking between complexes.

Figure S2. Scheme for derivation of distance constraints from crosslinked peptides, related to Figure 2**. (A) Only heterotypic crosslinks that map to structured parts of the models are used in the subsequent modeling (shown in **bold** for each dataset). Heterotypic crosslinks involved in at least one flexible tail are shown in *italics*. (B) Illustration of the 15 possible orientations for two crosslinked subunits. Subunit "A" (orange) is positionally fixed while "b" (green) is placed at each of the remaining 15 positions. Subunit "b" is in the same position as "A" only in homotypic crosslinks (faint first option). (C) A Flow scheme summarizing the general analysis procedure that converts crosslinks to distance constraints (see Methods for detail). (D) Scheme illustrating a simplified scenario to derive constraints from mass spectrometry crosslinking data. The crosslinked peptides were mapped onto the correct subunit and the distance was computed for all pairwise orientations of two subunits. For simplicity we illustrate the local nearest neighbor distances (five pairwise orientations). Using 30 Å as a general cut-off guide, we determine the preferred orientation for each crosslinked peptide by comparing the distances for all orientations. The preferred spatial pairings based on these data were then combined to generate a model consistent with the constraints.
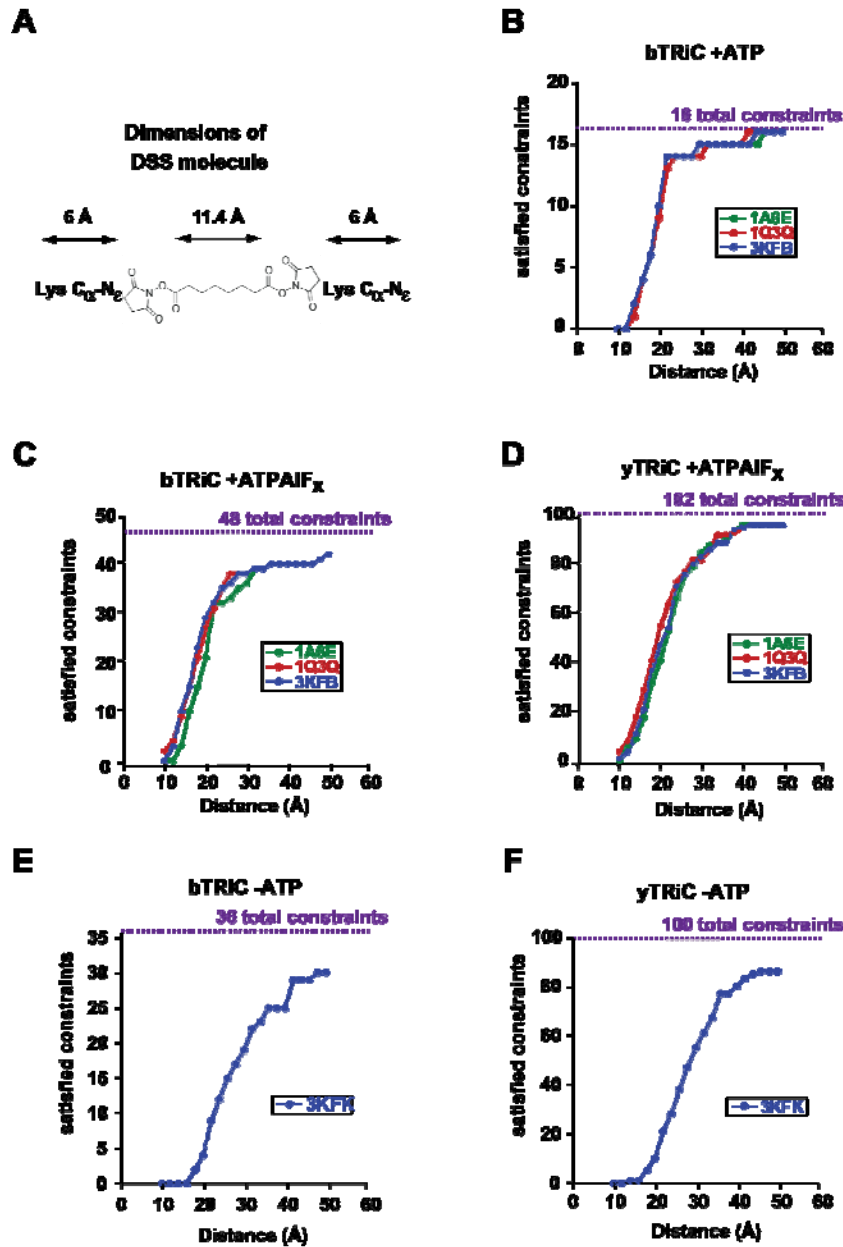
Figure S3. Satisfied constraint profiles for each dataset saturate at a distance consistent with the physical crosslinker length, related to Figure 3. (A) Schematic illustration of the physical CαLys-DSS-CαLys length. Satisfied constraint profiles for: (B) bovine TRiC (+ATP), (C) bovine TRiC (+ ATP+AlF$_x$), (D) yeast TRiC (+ ATP+AlF$_x$), (E) bovine TRiC (-ATP) and (F) yeast TRiC (-ATP). The satisfied constraint profile is independent of alignment template used (3KFB, 1A6E and 1Q3Q (Ditzel et al., 1998; Pereira et al., 2010; Shomura et al., 2004), shown in blue, green and red, respectively). The profile saturates at a value that approximates the extended crosslink length of 30 Å,

which takes into account for the local structural dynamics and approaches the total number of crosslinks (purple dotted line).
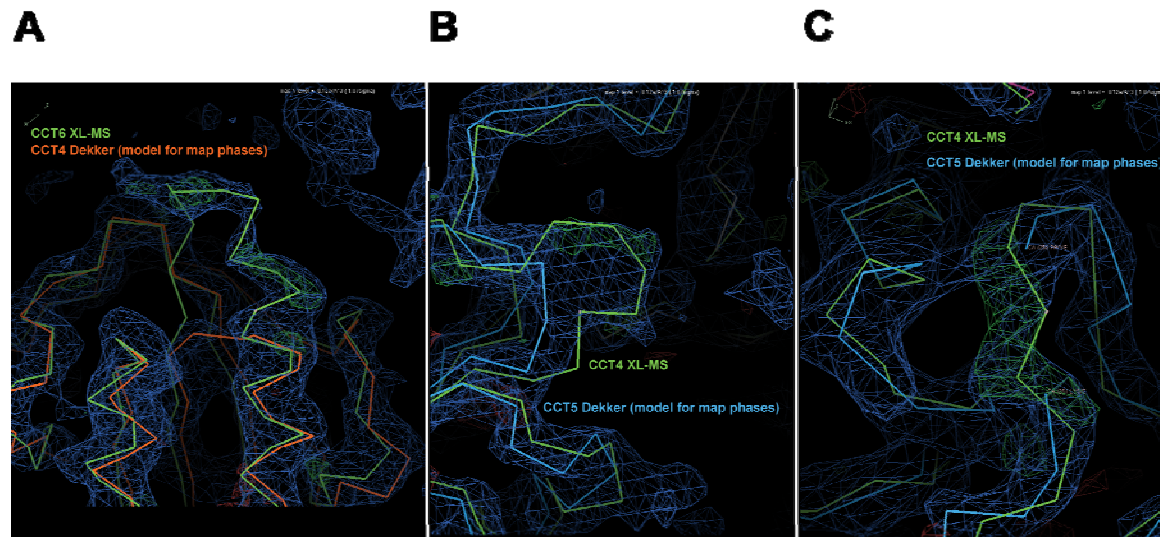
Figure S4. Agreement of the XL-MS model with features of the Dekker density, related to Figure 4**.** (A-C) Refmac was used to calculate weighted electron density maps with phases derived from the Dekker model. 2Fo-Fc density at 1.1 σ is shown in blue; positive Fo-Fc difference density at 3 σ indicating missing segments is shown in green. The XL-MS and Dekker model are shown as backbone traces as indicated. Panels A and B correspond to the portions used for subunit docking that are shown in Figure 5B of the main text. In the portion shown in panel C, the Liou/Dekker topology requires a five residue insertion between residues 255 and 261 present in CCT5 (loose ends), while there is strikingly well-defined density indicating a short connection, as predicted for the XL-MS topology. There are many more examples for a rather uncritical attitude of the creators of the original model towards aberrant density features with respect to the Liou/Dekker topology.
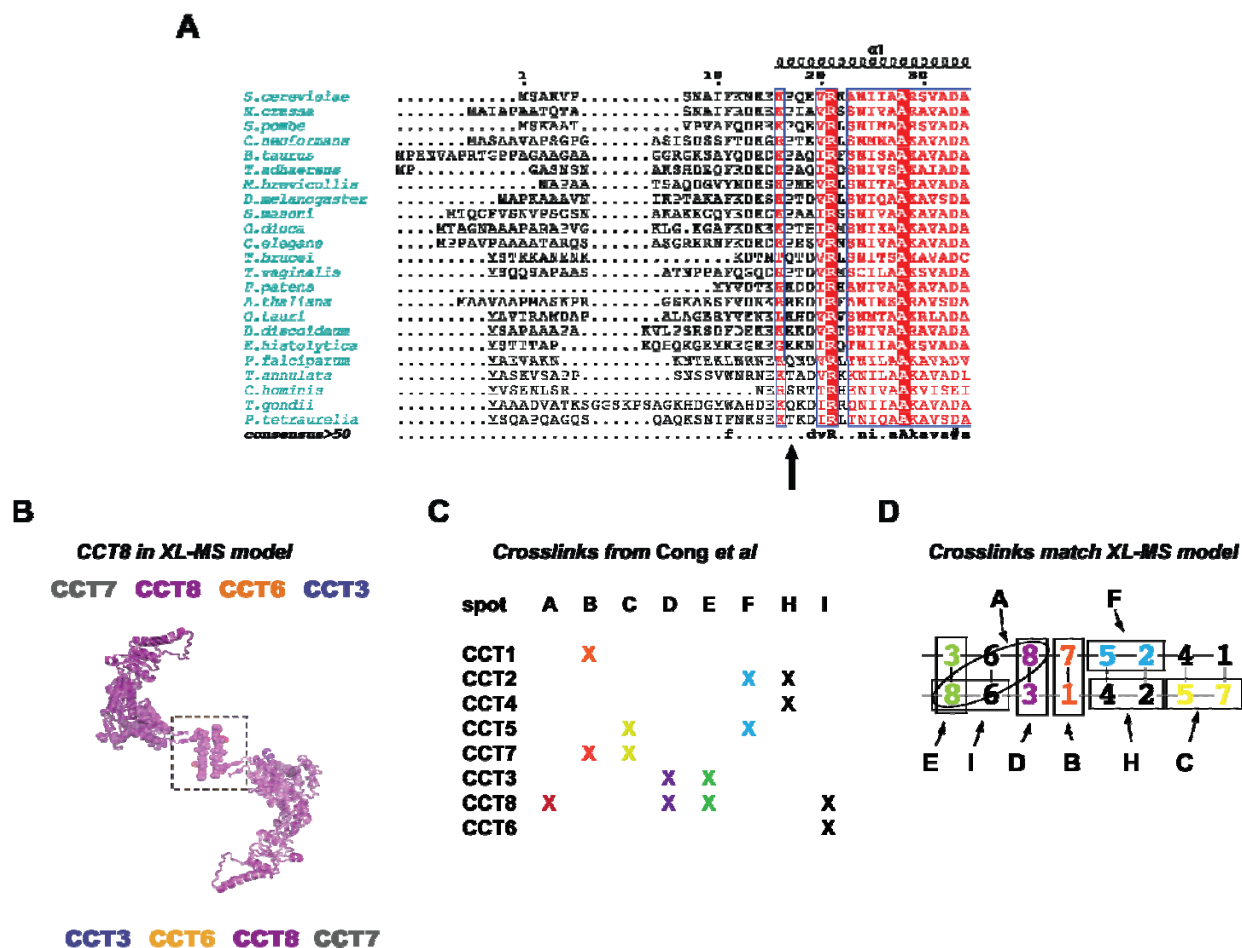
Figure S5. XL-MS structural features are consistent with existing and previous crosslinking data (Cong et al) , related to Figure 5**.** (A) Sequence alignment of CCT4 sequences highlights the conservation of the proline residue (indicated by an arrow) responsible for the outward-facing N-terminus. Numbering and secondary structure elements refer to the TRiC model. The conserved residue Phe11 makes a hydrophobic contact in the channel between TRiC subunits. (B) The CCT8 N-termini in the XL-MS model form inter-ring interactions. A hemisphere of the XL-MS structural model containing subunits CCT7, CCT8, CCT6 and CCT3 is shown in surface and colored grey, magenta, orange and blue, respectively. CCT8 tails are shown as spacefill (dashed box). (C-D) As described in Cong et al, bTRiC complexes were formaldehyde treated and separated on a 2D gel, spots were excised from the gel and analyzed by mass spectrometry to identify the TRiC subunits present in each spot as shown. This approach, however, did not reveal the specific site of crosslink. The eight identified crosslinked pairings (spots A-I) in this formaldehyde-based experiment are fully consistent with the XL-MS model explaining all (8 unique pairings) observed crosslinks

even the observed CCT-CCT8 crosslink, which is explained by the direct contact of the N-terminal tails of CCT8 in the XL-MS structural model.
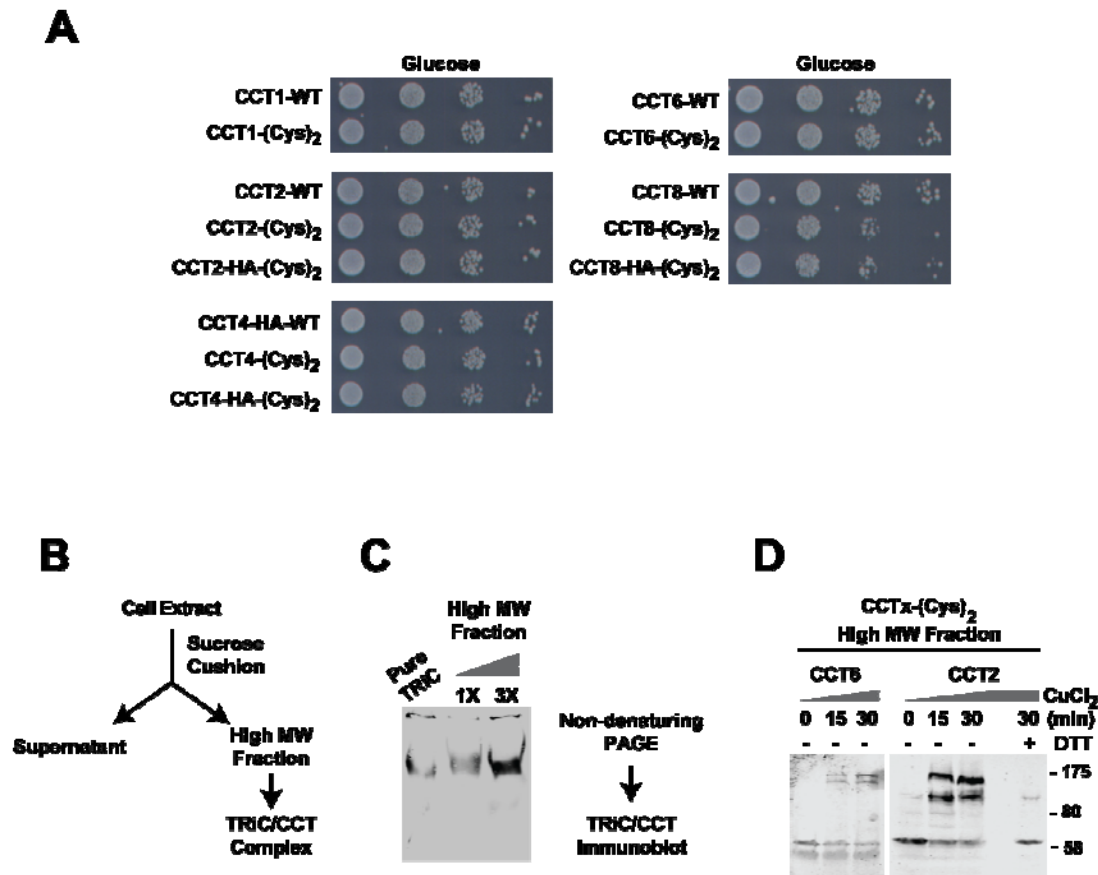
Figure S6. CCTx (Cys)$_2$ subunits are incorporated into the TRiC complex and support viability, related to Figure 6. (A) The CCTx (Cys)$_2$ mutants are able to complement the deletion of the respective wild-type gene. (B) Scheme to isolate CCTx (Cys)$_2$-containing TRiC. (C) High molecular weight fractions were analyzed on a non-denaturing PAGE gel. The genetically modified CCTx (Cys)$_2$ containing subunit is incorporated into the TRiC complex. (D) The sucrose cushion samples were treated with an oxidizing reagent and time points were collected at 0, 15 and 30 min and 30 min in the presence of DTT. The samples were analyzed on an SDS-PAGE gel and monitored for the formation of a higher molecular weight crosslinked species.

```
                                              TT   βA   00000000000000000  α1
                                                   →
ts_1Q2V    1  ...................MAQLSGQPVVILPEGTQRYVGRDAQRINILAARIIAET
CCT1       1  ..................MSQLFNNSRDTLFLGGEKISGDDIRNQNVLATMAVANV
CCT2       1  ...................MSVQIFGDQVTEERAENARLSAFVGAIAVGDL
CCT4       1  .................MSAKVPSNATFKNKEKPQEVRKANIIAARSVADA
CCT5       1  MAARPQQFPMEMPDLSNAIVAQDEMGRPFIIVKDQGNKKRQHGLEAKKSHILAARSVASI
CCT3       1  ..................MQAPVVFMNASQERTTGRQAQISNITAAKAVADV
CCT7       1  ...............MNFGSQTPTIVVLKEGTDASQGKGQIISNINACVAVQEA
CCT8       1  ...............MSLRLPQNPNAGLFKQGYNSYSNADGQIIKSIAAIRELHQM
CCT6       1  ...................MSLQLLNPKAESLRRDAALRVNVTSAEGLQSV
consensus>50  ...........................i.a...v.d.


              η1    TT   βB   TT   βC   α2        α3          Ω
ts_1Q2V   39  VRTTLGPKGMDKMLVDSLG.DIVVTNDCATILDKIDLQHPAAKMNVEVAKTQDKEAGDGT
CCT1      40  VKSSLGPVGLDKMLVDDIG.DFTVTNDGATILSLLDVQHPAGKILVELAQQQDREIGDGT
CCT2      33  VKSTLGPKGMDKLLQSASSNTCMVTNDGATILKSIPLDNPAARVLVNISKVQDDEVGDGT
CCT4      35  IRTSLGPKGMDKMIKTSRG.EIIISNDGHTILKQMAILHPVARMLVEVSAAQDSEAGDGT
CCT5      61  IKTSLGPRGLDKILISPDG.EITITNDGATILSQMELDNEIARLLVQLSKSQDDEIGDGT
CCT3      35  IRTCLGPKAMLKMLLDPMG.GLVLTNDGHAILREIDVAHPAAKSMLELSRTQDEEVGDGT
CCT7      40  LKPTLGPLGSDILIVTSNQ.KTTISNDGATILKLLDVVHPAARTLVDISRAQDAEVGDGT
CCT8      42  CLTSMGPCGRNKIIVNHLG.KIIITNDAATMIRELDIVHPAVKVLVMATEQQDIDMGDGT
CCT6      33  LETNLGPKGTLKMLVDGAG.NIKLTRDGKVLITEMQIQSPTAVLLARAAAAQDEITGDGT
consensus>50  ....§GP.g.dkm.....g....itnDg.tiL...d...p.ak.lve....Qd.e.GDGT


              α4              α5                  βD      Ω....Ω
ts_1Q2V   98  TTAVVIAGELLR.KAEELLDQNTHPSTITKGYALAAEKAQEILDEIAIRVDPDDE....E
CCT1      99  TSVVIIASELLKRANELVKN.KIHPTTIITGFRVALREAIRFINEVLSTSVDTLG..KE
CCT2      93  TSVTVLSAELLREAEKLIDQSKIHPQTIIEGYRLASAAALDALTKAAVDNSHDKT.MFRE
CCT4      94  TSVVILTGALLGAAERLLNKG.IHPTIIADSFQSAAKRSVDILLEMCEKVSLSDR....E
CCT5     120  TGVVVLASALLDQALELIQKG.IHPIKIANGFDEAAKLAISKLEETCDDISASNDELFRD
CCT3      94  TTVIILAGEILAQCAPYLIEKNIHPVIIIQALKKALTDALEVIKQVSKEVDVEND....A
CCT7      99  TSVTILAGELMKEAKPFLEEG.ISSHLIMKGYRKAVSLAVEKINELAVDITSEKS.SGRE
CCT8     101  NLVMILAGELLNVSEKLISMG.LSAVEIIQGYNMARKFTLKELDEMVVGEITDKN..DKN
CCT6      92  TTVVCLVGELLRQAHRFIQEG.VHPRILITDGPEIARKESMKFLDEFKISKTNLSN..DRE
consensus>50  t.v.il..el§...........ihp..I..gy..a.........e.............e


              α6        η2         α7      βE      βF    η3   βG
ts_1Q2V  153  TLLKIAATSITGKNAESHKELLAKLAVEEAVKQVAEKK.......DGKYVVDLDNIKFEKK
CCT1     155  TLINIAKTSMSSKIIGADSDPFSNMVVDALLAVKTQNSKG......EIKYPVKAVNVLKA
CCT2     152  DLIHIAKTTLSSKILSQDKDHFAELATNAILRLKGS............TNLEHIQIIKI
CCT4     149  QLVRAASTSLSSKIVSQYSSPLAPLAVDSVLKISDENSK........NVDLNDIRLVKK
CCT5     179  FLLRAAKTSLGSKIVSKDHDRFAEMAVEAVINVMDKDRK........DVDFDLIKMQGR
CCT3     150  AMKKLIQASIGTKYVIHWSEKMCELALDAVKTVRKDLGQTVEGEPNFEIDIKRYVRVEKI
CCT7     157  LLERCARTAMSSKLIHNNADFFVKMCVDAVLSLDRND...........LDDKLIGIKKI
CCT8     158  ELLKMIKPVISSKRYGS.EDILSELVSEAVSHVLPVAQQAG....EIPYFNVDSIRVVKI
CCT6     149  FLLQVARSSLLTKVDADLTEVLTPIVTDAVLSVYDAQADN........LDLHMVEIMQM
consensus>50  .§....a.....K......d....l..§av..v................d...l...k.


              η4   βH   βI     TT   βJ  TT    βK
ts_1Q2V  206  AGEGVEESELVRGVVIDKEVVHPRMPKRVEN......AKIALINEALEVKRTETDAKINI
CCT1     209  HGKSATESLLVPGYALNCTVASQ....AMPKRIAGGNVKIACLDLNLQKARMAMGVQINI
CCT2     199  LGGKLSDSFLDEGFILAKKFGNN.....QPKRIENAKILIANTTLDTDRVKIFG.TKFKV
CCT4     200  VGGTIDDTEMIDGVVLTQTAIKSAGGPTRKEK.....AKIGLIQFQISPPAPDTENNIIV
CCT5     230  VGGSISDSKLINGVILDKDFSHPQMPKCVLPKEGSDGVKIAILTCPFEPPKPKTKHKLDI
CCT3     210  PGGDVLDSRVLKGVLLNKDVVHPKMSRHIEN......PRVLLDCPLEYKKGESQTNIEI
CCT7     205  PGGAMEESLFINGVAFKKTFSYAGF.EQQPKKFNNPKILSLNVELELKAEKDN..AEVRV
CCT8     213  MGGSLSNSTVIKGMVFNREPEGHVKSLSEDKK....BKVAVFTCPLDIANTETKGTVLL
CCT6     200  QHLSPKDITFIRGLVLDHGGRHPDMPTRVKN......AYVLILNVSLEYEKTEVNSGFFY
consensus>50  .g....§s..i.G.....................i.........e..k........i


              α8                          βL          α9
ts_1Q2V  260  TSPDQLMSFLEQEEKMLKDMVDHIAQTG.........ANVVFVQKG.....IDDLAQHY
CCT1     265  DDPEQLEQIRKREAGIVLERVEKIIDAG.........AQVVLTTKG.....IDDLCLKE
CCT2     253  DSTAKLAQLEKAEREKMKNKIAKISKFG.........INTFINRQL.....IYDYPEQL
CCT4     255  NDYRQMDKILKERAYLLNICKKIEKAK..........CNVLLIQKSILRDAVNDLALHF
CCT5     290  SSVEEYQRLQTYEQDKFKEMIDDVKKAG.........ADVVICQWG.....FDDEANHL
CCT3     264  EKEEDWNRILQIEEEQVQLMCEQILAVR.........PTLVITEKG.....VSDLAQHY
CCT7     262  EHVEDYQAIVDAEWQLIFEKLRQVEETG.........ANIVLSKLP.....IGDLATQF
CCT8     268  HNAQEMLDFSKGEEKQIDAMMKEIADMG.........VECIVAGAG.....VGELALHY
CCT6     254  SSADQRDKLAASERKFVDAKLKKIIDLKNEVCGMDPDKGFVLINQKG.....IDPMSLDV
consensus>50  ...eq......E.........!........n.v.........i.dl....
```

Figure S7.p2

```
                          βM              α10                    βN      η5            η6  βO          βP
                        2020    →    22222222222          →     202          202  →           →    →  T..

ts_1Q2V   305  LAKYGIMAVRRVKKSDMERLAKATGAKIVTNVEDLT......PEDLGYAEVVERRKLA..
CCT1      310  FVEAKIMGVRRCKKEDLRRIARATGATLVSSMSNLEGEETFESSYLGLCDEVVQAKFS..
CCT2      298  FTDLGINSIEHADFEGVERLALVTGGEVVSTFDEPS......KCRLGECDVIERIMLG..
CCT4      305  LSKLNIMVVKDIEREEIEFLSKGLGCKPDADIELFT......EDRLGSADLVEFIDSDGS
CCT5      335  LLQNDLPAVRWVGGQELEHIAISTNGRIVPRFQDLS......KDRLGTCSRIYEQEFG.T
CCT3      309  LLKGGCSVLRRVKKSDNNRIARVTGATIVNRVEDLK......ESDVGTNCGLFKVEMIG.
CCT7      307  FADRNIFCAGRVSADDMNRVIQAVLGSIQSTTSDIK......PEELGTCALFEEMQIG..
CCT8      313  LNRYGILVLKVPSKFELRRLCRVCGATPLPRLGAPT......PEELGLVETVKTMEIGGD
CCT6      309  FAKHNILALRRAKRRNHERLQLVTGGEAQNSVEDLS......PQILGFSGLVYQETIG..
consensus>50   .....i.........d.e......g........d..........1G....v.e......

                          βQ              βR              α11                          βS
                     .T  →         →         →         22222222222222222222222      →         l

ts_1Q2V   357  .GENMIFVEGCKNPKAVTILIRGGTEEVIDEVERALEDAVKVVRDVME..DGAVLPAGGA
CCT1      368  .DDECILIRGTSKHSSSSIILRGANDYSLDEMERSLHDSLSVVRRTLE..SGNVVPGGGC
CCT2      350  .EQPFLKFSGCKAGEACTIVLRGATDQTIDEAERSLHDALSVLSQTTK..ETRTVLGGGC
CCT4      359  KIVRVTGIRNNNARPTVSVVIRGANNMIIDETERSLHDALCVIRCLVK..ERGLIAGGGA
CCT5      388  TKDRMLIIEQSKETKTVTCFVRGSNKMIVDEAERALHDSLCVVRNLVK..DSRVVYGGGA
CCT3      362  .DEYFSFLDNCKEEPKACTIMLRGGSKDILNEIDRNLQDAMAVARNVML..SPSLSPGGGA
CCT7      359  .SERYNLFQGCPQARTCTILLRGGAEQVIAEVERSLHDAIMIVKRALQ..NKLIVAGGGA
CCT8      367  .RVTVFKQEQGEISRTSTIILRGATQNNIDDIERAIDDGVAAVRGLMKPSGGKLLPGAGA
CCT6      361  .EEKFTYVTENTDPKSCTILIKGSTHYALAQTKDAVRDGLRAVANVLK..DKNIIPGAGA
consensus>50   ..e.............ti..rG......dg.er.l.D...vv...........ggGa

                          α12            ..00  α13              α14           α15
                     2222222222222...00  2222222222222   222222222   2222222222

ts_1Q2V   414  PEIELAIRLDEYAK...QVGGKEALAIENFADALKIIPKTLAENAGLDTVEMLVKVISEH
CCT1      425  VEAALNIYLDNFAT...TVGSREQLAIAEFARALLIIPKTLAVNAAKDSSELVAKLRSYH
CCT2      407  AEMVMSRAVDTEAQ...NIDGKKSLAVEAFARALRQLPTILADNAGFDSSELVSKLRSSI
CCT4      417  PEIEISRRLSKEAR...SMEGVQAFIRQEIFASALEVIPTTLAENAGLNSIKVVTELRSKH
CCT5      446  AEVTMSIAVSEEAD...KQRGIDQYAFRGFAQALDTIPMTLAENSGLDPIGTLSTLKSKQ
CCT3      419  TEMAVSVKIAEKAK...QLEGIQQWPYQAVADAMECIPRTLIQNAGGDPIRLLSQLRAKH
CCT7      416  TEMEVSRCLRDYSK...TIAGKQQMIINAFARALEVIPRQLCENAGFDAIEILNKLRIAH
CCT8      426  TEIELISRITKYGE...RTPGLLQLAIKQFAVAFEVVPRTLAETAGLDVNEVLPNLYAAH
CCT6      418  FYIALSRYLRSANMNKLGAKGKTKTGIEAFAEALLVIPKTLVKNSGFDPLDVLAMVEDEL
consensus>50   .e....................g.........FA.AI..iP.tL.enag.#.......l....

                          βT              βU  α16       βV         α17
                     202          →             →    0000     →   202202202222222222

ts_1Q2V   471  KNR...........GLGIGIDVF...EGKPADMLEKGIIEPLRVKKQAIKSASEAAIMI
CCT1      482  AASQMAKPEDVKRRSYRNYGLDLI...RGKIVDEIHAGVLEPTISKVKSLKSALEACVAI
CCT2      464  YNG...........ISTSGLDLN...NGTIADMRQLGIVESYRLKRAVVSSASEAAEVL
CCT4      474  ENG...........ELNDGISVR...RSGTTNTYEEHILQPVLVSTSAITLASECVKSI
CCT5      503  LKE.........KISNIGVDCL...GYGSNDMKELFVVDPFIGKKQQILLATQLCRMI
CCT3      476  AQG...........NFTTGIDGD...KGKIVDMVSYGIWEPEVIKQQSVKTAIESACLL
CCT7      473  SKG...........EKWYGVVFE...TENIGDNFAKFVWEPALVKINALNSATEATNLI
CCT8      483  NVTEPG....AVKTDHLYKGVDIDGESDEGVKDIREENIYDMLATKKFAINVATEAATTV
CCT6      478  DDAQDSD.......ETRYVGVDLN...IGDSCDPTIEGIWDSYRVLRNAITGATGIASNL
consensus>50   ...............Gvd.........#.....I.#....k...i..A.e.....

                          βW
                     l  →

ts_1Q2V   516  LRIDDVIAAKATKPEGGQGGGMPGGMGGMDMGM
CCT1      539  LRIDTMITVDPEPPKEDPHDH............
CCT2      509  LRVDNIIRARPRTANRQHM...............
CCT4      519  LRIDDIAFSR.......................
CCT5      549  LKIDNVITSGKDEY...................
CCT3      521  LRVDDIVSGVRKQE...................
CCT7      818  LSVDETITNKGSESANAGMMPPQGAGRGRGMPM
CCT8      539  LSIDQIIMAKKAGGPRAPQGPRPGNWDQED...
CCT6      528  LLCDELLRAGRSTLKETPQ..............
consensus>50   L..iDd.i.........................
```

Figure S7. Sequence alignment of yeast CCT paralogs and the thermosome sequence, related to Figure 7. Numbering and secondary structure elements refer to the thermosome model(Shomura et al., 2004).

10

Figure S8. Schematic representation of substrate interaction surfaces from previous structural structures, related to Figure 8. Presumed subunit order in (A) the symmetry-free cryoEM map of open TRiC (Cong et al., 2011) and a schematic (B) of the X-ray crystal structure of the TRiC-tubulin complex (Munoz et al., 2011). Only the apical domains are shown, highlighting the four-fold pseudo symmetry. Red lines and numbers denote inter-apical domain crosslinks in the datasets to the open forms of yTRiC (red) and bTRiC (blue).

Supplemental Tables

Table S1: Lysine variability between orthologous TRiC subunits indicates approximately 40% of the lysines in each sequence are positionally variable, related to Figure 2**.**

subunit	Lysine Positions					[e] %variable
	[a]Total bovine	[b]Total yeast	Conserved positions	[c]variable bovine	[d]variable yeast	
CCT1	41	39	22	19	17	45
CCT2	41	39	25	16	14	38
CCT3	39	35	24	15	11	35
CCT4	39	39	20	19	19	49
CCT5	48	44	30	18	14	35
CCT6	34	47	25	9	22	38
CCT7	39	39	27	12	12	31
CCT8	41	36	18	23	18	53

[e]Calculated as the ratio between the variable lysine positions (sum of [c]bovine and [d]yeast) and all lysine positions (sum of bovine[a] and yeast[b]).

Table S2: Overview of all assigned crosslinks, related to Figure 1**.**

See attached excel file

Table S3. Statistical analysis of crosslinking datasets, related to Figures 2 and 3**.**

Dataset[a]	n[b]	Arrngmt. rank	Tp[c]	Min	1st. Qu.	Median	Mean	3rd. Qu.	Max	p-value
Closed bovine	60	1	50	41	47	49	48.5	50	55	2e-4
		2	44	35	42	43	43.4	45	50	
Closed yeast	97	1	79	68	75	77	76.9	78	85	<1e-5
		2	68	57	66	68	67.5	69	78	
Closed combined	157	1	129	113	123	125	125.4	127	137	<1e-6
		2	112	98	109	111	110.9	113	126	
Open bovine	34	1	23	17	22	23	22.9	24	29	0.17
		2	21	14	20	21	21.2	22	27	
Open yeast	95	1	72	61	69	71	70.6	72	80	6e-3
		2	66	56	64	66	65.6	67	75	
Open combined	129	1	95	81	92	94	93.65	95	104	<3.4e-3
		2	87	76	85	87	86.84	89	96	

[a]Error distributions for the best and 2nd best arrangement (see arrangement rank) were calculated according to a binomial distribution where the true independent crosslinks[b] (n) are detected as constraints fulfilled with a probability pT=0.85 and pF=0.10. The distribution was centered at the tp values.

[b]number of independent constraints.

[c]true positives.

Table S4. Fitting scores comparing the four refined Rosetta models to the cryo-EM map (EMDB accession number 5148), related to Figure 4**.**

Models	-FRC	MPE	-CCC	Chimera
XL-MS-1	-0.658	0.857	-0.538	0.704
XL-MS-2	-0.660	0.860	-0.541	0.706
Cryo-EM	-0.662	0.850	-0.541	0.705
Random	-0.669	0.846	-0.543	0.723

Three EMAN2 fitting scores including FRC (Fourier Ring Correlation), MPE (Mean Phase Error), and CCC (Cross Correlation Coefficient) in addition to the Chimera similarity fitting score were calculated. Other than the Chimera score, lower value of the scores indicates better fit between the model and the cryo-EM map. XL-MS-1 and XL-MS-2 refer to the two models following the same XL-MS ordering, but related to each other with a 180º in-plane rotation due to the degeneracy of the two-fold symmetry axis. Cryo-EM and Random refer to the model following the original cryo-EM ordering and a random CCT subunit ordering, respectively. In each scoring criteria, the difference is within the fitting error bars. These imply that the cryo-EM map is not well resolved enough to derive a unique atomic model.

The proposed XL-MS ordering also conflicts with the cryo-EM structure-based ordering(Cong et al., 2010). At the time, it appeared that the density provided sufficient

constraints for an unambiguous subunit assignment and the resulting model was consistent with a limited set of cross-linking data. The conflict presented by the XL-MS data prompted several additional attempts at full atom modeling of all subunits. We tested the cryo-EM and the new XL-MS derived orderings, as well as a random ordering to serve as a baseline. While these new results show that the cryo-EM based ordering still produces a marginally better score than the new proposed ordering, more significant is that the random ordering produces a score which is marginally better than either of the proposed orderings. The relative similarity of the scores for all of the refined models suggests that the cryo-EM map simply does not have sufficient resolution to discriminate between the various orderings, and the previously published map-derived ordering likely results from an over-interpretation of the available data. This is not to say that the cryo-EM map is in any way inaccurate, simply that the subunit homology is too strong to permit unambiguous assignment based on a 4 Å resolution map alone.

Supplementary Movie 1, related to Figure 7.

Movie showing the location of the crosslinks within one ring and between the two rings in the closed conformation of yeast TRiC. A sideview of the lower half of TRiC is shown, and the complex is slowly revolving around the pseudo-eightfold axis (vertical). The subunit backbones from the crystal structure of yeast TRiC are shown as coils. The α subunit CCT1 is indicated in purple, the β subunit CCT2 in blue, the γ subunit CCT3 in yellow-green, the δ subunit CCT4 in cyan, the ε subunit CCT5 in green, the ζ subunit CCT6 in red, the η subunit CCT7 in yellow and the θ subunit CCT8 in orange. Crosslinked lysines below the 30 Å distance cut-off are represented as dotted lines. Within the ring, heterotypic and homotypic crosslinks are shown in pink and grey, respectively. At the ring-ring interface, heterotypic and homotypic crosslinks are shown in white and pale yellow, respectively.

Supplemental Experimental Procedures

Purification of TRiC from B. taurus and S. cerevisiae

Bovine TRiC was purified as described previously (Feldman et al., 2003). TRiC from *S. cerevisiae* was purified using a four-step chromatographic method, which takes advantage of its strong interaction with Plp2p. All steps were performed at 4 °C. Yeast cells were suspended in buffer YN (50 mM HEPES-NaOH pH 7.4, 200 mM NaCl, 5 mM MgCl$_2$, 10 % glycerol, 20 mM imidazole, 10 mM β-mercaptoethanol, 1 mM PMSF, and Complete protease inhibitor cocktail (Roche Diagnostics, Freiburg, Germany)) and lyzed using a bead beater (BioSpec Products, Bartlesville, OK, USA). Cell debris was removed by sequential centrifugation steps at 3000 g and 18600 g. The clear supernatant was subsequently incubated for 1 h with Ni^{2+}-Chelating Sepharose (GE Healthcare, München, Germany) beads, which were decorated with Plp2p bearing a combined 6xHis and Strep-tag II. The beads were collected in a column, washed with a step gradient of buffer YN containing increasing amounts of imidazole (20 / 50 / 100 mM). Bound TRiC was then eluted together with Plp2p using buffer YN containing 250 mM imidazole. The Plp2p-TRiC containing fractions were merged and applied to a Strep-Tactin column (IBA). This column was extensively washed with buffer YS (50 mM HEPES-NaOH pH 7.4, 200 mM NaCl, 1 mM EDTA, 10 % glycerol, and 1 mM DTT). Bound proteins were subsequently eluted with buffer YS containing 2.5 mM desthiobiotin. To separate Plp2p from TRiC, the protein-containing fractions were applied on Heparin resin (GE Healthcare). Unbound Plp2p was washed away with

buffer YH (20 mM HEPES-NaOH pH 7.4, 10 mM MgCl$_2$, 10 mM CaCl$_2$, 10 % glycerol, 10 mM β-mercaptoethanol). TRiC was eluted with a linear salt gradient from 200 mM to 1000 mM NaCl in buffer YH. As final step, the TRiC-containing fractions were subjected to size exclusion chromatography on Superose 6 (GE Healthcare), which was developed with buffer YSEC containing 10 mM HEPES-NaOH pH 7.4, 200 mM NaCl, 5 mM MgCl$_2$, 10 % glycerol and 1 mM DTT. The TRiC-containing fractions were merged, concentrated to 2 g l^{-1} by ultrafiltration and snap-frozen in liquid nitrogen for storage at -80 °C.

Non-denaturing PAGE and SDS-PAGE of crosslinked samples

A purified 1 μM TRiC sample was incubated in 30 mM HEPES-KOH pH 7.4, 100 mM KCl, 10 mM MgCl$_2$, 10% glycerol and 1 mM DTT in the presence of 1 mM ATP, 1 mM EDTA or 1 mM ATP/5 mM NaF and 5 mM AlNO$_3$ for 20 minutes at 30 °C. The DSS crosslinker (dissolved at 25 mM in dimethyl formamide) was subsequently added to each reaction to an initial 1 mM concentration and then incubated for another 30 minutes at 35 °C. Excess DSS was quenched by addition of ammonium hydrogen carbonate to 50 mM and incubated at 37 °C for 30 min. For the SDS-PAGE analysis the samples were spun, mixed with 4X SDS sample buffer, boiled for 5 minutes and resolved on a 12% SDS-PAGE gel. For the non-denaturing PAGE the samples were spun, mixed with native gel sample buffer and loaded onto a 4% non-denaturing gel.

CryoEM 2D analysis of crosslinked TRiC samples

The cross-linked TRiC in the apo, ATP and ATP+AlF$_x$ states were frozen in vitreous ice on 400 mesh R1.2/1.3 Quantifoil grids, respectively using a procedure previously described (Cong et al., 2011). All the images were collected on a Gatan 4kx4k CCD camera at ~60,000 microscope magnification in the JEM2010 (for apo and ATP+AlF$_x$ states) and JEM2200FS (for ATP state) electron microscopes. 3780, 2319, and 8941 particle images were boxed out for apo, ATP, and ATP+AlF$_x$ state, respectively (Figure 1C), using e2boxer.py from EMAN2 (Tang et al., 2007). Furthermore, we carried out the reference-free 2D image analysis on the boxed-out particle images using refine2d.py from EMAN1 (Ludtke et al., 1999). The resulting class averages are shown in Figure 1C. Comparing with the normal TRiC complex in the apo and nucleotide bound states (Cong and Ludtke, 2010; Cong et al., 2011), this analysis demonstrates that the cross-linked TRiC can still form the normally appearing TRiC complex.

Crosslinking, sample processing and LC-MS/MS analysis.

Both bovine and yeast TRiC complexes were crosslinked in both the open state (in absence of ATP) as well in a closed state in the presence of ATP, aluminum and fluoride ions (AlF$_x$ state, prepared according to (Meyer et al., 2003)). In addition bovine TRiC was crosslinked in the presence of ATP alone. In all cases, TRiC preparations were crosslinked at a total protein concentration of 1 – 1.5 mg ml^{-1} using 50 – 100 µg starting material. The crosslinking buffer was 20 mM HEPES-KOH (pH 7.4) containing 50 mM NaCl, 5 mM MgCl$_2$, 0.1 mM EDTA, and 1 mM DTT (for bovine TRiC) or 100 mM potassium acetate, 5 mM magnesium acetate, and 1 mM DTT for yeast TRiC. The crosslinking reaction was initiated by adding disuccinimidyl suberate stock solution (25

mM DSS-d_0 and –d_{12}, Creative Molecules) in DMF to a final concentration of 1 mM.

Samples were incubated at 37 °C for 30 – 60 minutes. Excess reagent was quenched

by addition of ammonium hydrogen carbonate to 50 mM and incubation at 37 °C for 30

min.

After the quenching step, samples were evaporated to dryness in a vacuum centrifuge

and resuspended in 8 M urea. Proteins were reduced with 2.5 mM TCEP (37 °C, 30

min) and alkylated with 5 mM iodoacetamide (30 min, room temperature, protected from

light). The sample solutions were diluted to 1 M urea with 50 mM ammonium hydrogen

carbonate and trypsin (Promega) was added at an enzyme-to-substrate ratio of 1:50.

Proteolysis was carried out at 37 °C overnight followed by acidification with formic acid

to 2% (v/v). Samples were then purified by solid-phase extraction using Sep-Pak tC18

cartridges (Waters) according to standard protocols.

Samples were fractionated by size exclusion chromatography (SEC) on a Superdex

Peptide column as described elsewhere (Leitner, 2012). Two fractions collected from

SEC were evaporated to dryness and reconstituted in water/acetonitrile/formic acid

(95:5:0.1, v/v/v) to a final concentration of approximately 0.5 µg µl^{-1}. 2 µl each were

injected for duplicate LC-MS/MS analyses on an Eksigent 1D-NanoLC-Ultra HPLC

system coupled to a Thermo LTQ Orbitrap XL system. Peptides were separated on self-

packed New Objective PicoFrit columns (11 cm x 0.075 mm I.D.) containing Magic C_{18}

material (Michrom, 3 um particle size, 200 Å pore size) at a flow rate of 300 nl min^{-1} and

using the following gradient. 0-5 min = 5 %B, 5-95 min = 5-35 %B, 95-97 min = 35-95

%B and 97-107 min = 95 %B, where A = (water/acetonitrile/formic acid, 97:3:0.1) and B

= (acetonitrile/water/formic acid, 97:3:0.1). The mass spectrometer was operated in

data-dependent mode by selecting the five most abundant precursor ions (m/z 350-1600, charge state 3+ and above) from a preview scan and subjecting them to collision-induced dissociation (normalized collision energy = 35%, 30 ms activation). Fragment ions were detected at low resolution in the linear ion trap. Dynamic exclusion was enabled (repeat count 1, exclusion duration 30 s).

Data analysis of mass spectrometry data

Thermo .raw files were converted into the open .mzXML format using msconvert (proteowizard.sourceforge.net) and analyzed using an in-house version of xQuest (Rinner et al., 2008). Spectral pairs with a precursor mass difference of 12.075321 Da were extracted and searched against the respective FASTA databases containing the eight subunits of TRiC.

xQuest settings were as follows: Maximum number of missed cleavages (excluding the crosslinking site) = 2, peptide length = 5-50 amino acids, fixed modifications = carbamidomethyl-Cys (mass shift = 57.021460 Da), mass shift of the light crosslinker = 138.068080 Da, mass shift of mono-links = 156.078644 and 155.096428 Da, MS[1] tolerance = 10 ppm, MS[2] tolerance = 0.2 Da for common ions and 0.3 Da for crosslink ions, search in ion-tag mode.

Post-search manual validation and filtering was performed using the following criteria: xQuest score > 16, mass error between -4 and +7 ppm, %TIC > 10, and a minimum peptide length of six amino acids. In addition, at least four assigned fragment ions (or at least three contiguous fragments) were required on each of the two peptides in a crosslink.

Homology model building

Comparative homology models were built from parent PDB's detected by PDB-BLAST and aligned by the K*Sync alignment method (Chivian and Baker, 2006). Loop regions were assembled from fragments and optimized to fit the aligned template structure. We then used the Rosetta Software package for iterative backbone optimization and loop modeling routines to generate the final models. However, in specific instances we were unable to successfully find acceptable torsional angles to close the loops in some of the models. Generally, the structural template that was used to construct each of the CCT models was based on a homologous thermosome group II chaperonin (PDB codes: 3KFB, 1A6E or 1Q3Q) (Ditzel et al., 1998; Pereira et al., 2010; Shomura et al., 2004) whose sequence identity to each of the eight CCT subunits ranges between 30-50%, as such the corresponding domain confidence score for each model is high (3) and is proportional to the E-value (conf= -log(e-val)(Chivian and Baker, 2006). For this sequence identity regime the homology models we have generated should capture the structural features that are conserved amongst chaperonins especially in regions of higher sequence similarity such as the active site.

Generation of pairwise subunit orientations and distance constraints

The coordinates for the experimentally determined archeal chaperonin crystal structures were acquired from the RSCB and the corresponding symmetry mates were generated to build an intact holo chaperonin structure using a symmetry Python pymol plugin. This was carried out for a series of experimentally determined templates crystallized in different nucleotide states (PDB codes: 1Q3Q/3KFK/3KFB/1A6E) (Ditzel et al., 1998;

Pereira et al., 2010; Shomura et al., 2004). We then set a single subunit in the holo

chaperonin structure as a reference and generated all exhaustive pairs of homology

models of bovine and yeast CCT subunits (CCT1-CCT8) between the reference

orientation and all of its 15 neighbors by aligning the Cα coordinates of each homology

model to the coordinates of the respective subunit in the intact holo-complex. Because

the chaperonin subunits exhibit significant structural and sequence similarity we carried

out Cα-atom structural alignments using a Python plugin in Pymol. Next, we mapped the

peptide fragments with the identified lysine residue from each crosslinked peptide pair

to the corresponding sequence on each homology model and computed a $Lys_{cctx}C\alpha$-

$Lys_{ccty}C\alpha$ distance for all 15 orientations for each inter-subunit crosslink (16 possible

orientations for the intra-subunit crosslinks to evaluate homotypic contacts between the

two rings) to generate a distance matrix for an entire crosslink dataset. In instances

where crosslinked peptide fragments map to regions not modeled in the CCT subunit

homology models (i.e. regions significantly different from the structural template used to

generate the homology model such as loops, C-terminal and N-terminal tails), these

crosslinks were not further evaluated. This analysis was carried out using a series of

different crystal structures of archaeal chaperonins as alignment templates; the resulting

distance dependence on the constraints is largely independent of template. Homotypic

crosslinks were used to validate the inter ring register by identifying homotypic contacts

across the two rings that fall below the 30 Å (closed state) or 36 Å (open state)

threshold using the intra ring order identified in XL-MS. The homotypic crosslink data,

both in the combined closed and open state, were evaluated for consistency for intra

subunit distances and inter subunit distances consistent with homotypic subunit

contacts across the two rings (data not shown). Using the refined crystal structure of yeast TRiC based on the XL-MS arrangement, we additionally computed how many heterotypic and homotypic crosslinks are consistent with the model and what is the crosslink distance distribution.

Global analysis of crosslink derived constraints

We enumerated all possible double-ringed arrangements based on two assumptions: 1) total number of eight membered rings (containing 8 unique subunits) can be exhaustively described using 7! unique combinations due to the periodicity of the rings and 2) in any given complex the two rings are identical resulting in 8 possible ways of assembling a unique ring arrangement. These assumptions reduce the number of possible hexadecameric complexes to 8! unique arrangements. The spatial relationship of each subunit within an arrangement is assigned using a numerical descriptor and each arrangement is evaluated in the context of the derived distance matrix. Each spatial arrangement for a given crosslinked-peptide pair that falls below a set distance threshold is counted as a satisfied constraint and if it falls above it is counted as a violated constraint. The number of satisfied constraints is tallied for each arrangement and represented in a histogram. Using this approach if a particular crosslink satisfies two separate spatial orientations these are treated separately and count as satisfied constraints in the two arrangements, thus the contribution of crosslinks that satisfy more than one spatial orientation is reduced.

Statistical analyses

The simulation data were generated according to a binomial distribution where the true independent crosslinks are detected as constraints fulfilled with a probability $pT = 0.85$ and false positives occur at a rate $pF = 0.10$ and evaluated using a parametric bootstrap test. Under this very conservative model with a high false positive bias, we generate error distributions of the number of observed constraints for the best and second best arrangement for each independent dataset utilizing only independent constraints and assuming these conservative probabilities ($pT = 0.85$ and $pF = 0.10$). Comparing the error distributions for the best and the second best arrangement for each dataset we can evaluate the p-value to assess the probability that the XL-MS solution (i.e. one that fulfills the most constraints) is significant.

Disulfide engineering

Plasmid Generation: Wild type CCT1-HA, CCT2-HA, CCT4-HA, CCT6-HA and CCT8-HA were cloned into a pCu426 plasmid containing the uracil (URA3) auxotrophic marker (Tam et al., 2006). The CCTx(Cys)$_2$ mutants were generated using QuikChange (Stratagene) from wild type plasmids (LEU) with endogenous promoters: pAB CCT1, pAB CCT2, pAB CCT4, pAB CCT6, pAB CCT8 (Kabir et al., 2005). The following double cysteine mutants were generated: CCT1(Cys)$_2$: K119C/S470C, CCT2(CYS)$_2$: S113C/S453C, CCT4(Cys)2: K113C/S462C, CCT6(Cys)$_2$: G112C/L467C and CCT8(Cys)$_2$: M120C/N472C.

Strain Construction: pCu CCT1-HA (URA3), pCu CCT2-HA (URA3), pCu CCT4-HA (URA3), pCu CCT6-HA (URA3) and pCu CCT8-HA (URA3) plasmids were transformed

into the respective heterozygous CCTx haploid deletion strain (background MATa/α his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 met15Δ0/MET15 lys2Δ0/LYS2 ura3Δ0/ura3Δ0; Winzeler 1999) before sporulation and tetrad dissection. Verified haploid strains containing the pCu CCTx (URA3) and deleted for the corresponding *cctx* gene were then transformed with a pAB CCTx (LEU) plasmid containing the appropriate CCTx(Cys)$_2$ mutation. Counter-selection against wild type pCu CCTx with plates containing 5-fluoroorotic Acid (5'FOA) –LEU allowed the mutant plasmids to be expressed without wild type CCTx present. In addition, counter-selection was also confirmed by the absence of growth when replica-plated onto –URA plates after 5'FOA counter-selection.

Viability Assays: Yeast were resuspended in water to an OD$_{600}$ of 0.08 and serially diluted 10-fold four times before spotting 10 µl of each dilution onto appropriate media. All plates contained glucose as the primary sugar source. The plates were incubated at 30°C for 48 hours before scanning.

Disulfide crosslinking: 25 ml culture of yeast was harvested at OD$_{600}$ = 1.0, pelleted and washed with water. The rinsed pellet was resuspended in 500 µl of Lyticase Buffer (1.2 M sorbitol, 50 mM Tris-HCl pH 8.0, 0.5 mM MgCl$_2$, 30 mM DTT), incubated in 25 $^{\circ}$C for 15 minutes and spun at 6000 rpm for 1 minute. The pellet was resuspended with 500µl Lyticase Buffer and 100 µl lyticase (ref) and incubated at 30°C for 30 minutes. The reaction was spun at 6000 rpm for 1 minute and washed twice with Lyticase Buffer. The final pellet was resuspended with 150 µl of chilled Lysis Buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.1% Triton-X100, and protease cocktail (100 µg/ml Pepstatin A, 75 µg/ml Leupeptin, 10 mM Benzamidine, 100 µg/ml Aprotinin, 2 mM AEBSF). The reactions were vortexed three times for 30 seconds, spun at 14,000 rpm

for 10 min at 4°C and the supernatant was used in the subsequent crosslinking reactions.

Yeast lysates were treated with 250 uM CuCl$_2$ and incubated at 2 5°C. Throughout the 30 minute time course, aliquots were removed and the reactions were quenched using a final concentration of 10 mM sodium iodoacetate and 50 mM N-ethyl maleimide. The crosslinked samples were mixed with 6X SDS sample buffer (without reducing agent), incubated at 95 °C for 3 min and spun at 13000 rpm for 5 minutes and resolved on a 10% SDS-PAGE. For the reduced sample control, a final concentration of 100 mM DTT was added and the samples were incubated at 25 ^{o}C for 10 minutes, followed by the addition of 6X SDS sample buffer, incubated at 95 °C for 3 min and spun at 13000 rpm for 5 minutes prior to loading. The 10% SDS-PAGE gels were subsequently transferred onto a 0.2 μm nitrocellulose membrane and blocked with 3% BSA in TBS-Tween. CCT1 was detected with a rabbit polyclonal antibody. CCT2-HA, CCT4-HA, and CCT8-HA, were detected with a mouse HA-antibody (Covance MMS-150R). CCT6 was detected with a rabbit polyclonal antibody (kind gift from Dr. Sherman). The primary antibodies were detected with a goat anti-mouse IRDye800 or goat anti-rabbit IRDye700 secondary antibody, scanned and processed with an Odyssey imager using software from LI-COR Biosciences.

Model refinement against the cryo-EM density map

We tested the XL-MS ordering against the cryo-EM density map (Cong et al., 2010). Starting from the homology models (Booth et al., 2008) for each of the eight subunits, we generated four models of the entire TRiC complex: one following our original cryo-

EM derived subunit-ordering (Cong et al., 2010), two based on the XL-MS ordering because of the degeneracy of the two-fold symmetry axis, and the last one from a random CCT subunit ordering. For each ordering, the all-atom homology model of each individual subunit of the complex was refined with the cryo-EM density restraints by utilizing the Rosetta relax protocol incorporating the density-fitting function (DiMaio et al., 2009). From the 16 generated models, the one with the best Rosetta score was chosen to present that specific subunit. Here the Rosetta score incorporates both the stereochemistry and the fit-to-density scores. This leads to a refined model for each ordering composed by the top-ranking model for individual subunits.

Crystal structure refinement

For refinement of a model based on the XL-MS topology against the X-ray diffraction data from the Willison lab (Dekker et al., 2011), we started from a consensus model based on the thermosome crystal structure (PDB code 1Q2V (Shomura et al., 2004)). In this model, loop regions deviating in length between TRiC subunits and thermosome were deleted, and divergent residues were modeled as alanine. Positive difference density was apparent in all nucleotide binding pockets. Initial refinement with Refmac (Murshudov et al., 1997) yielded a model that agreed with the location of the homotypic contacts in the crystal lattice and the non-crystallographic symmetry between the 4 octameric rings in the asymmetric unit from Dekker et al (Dekker et al., 2011) as judged by the subunit-to-subunit residual mean square deviations determined with Lsqman (Kleywegt, 1994)(<1 Å for identical vs. ~2 Å for homologous subunits). Using this information, a Swiss-Model (Arnold et al., 2006) homology model of the octameric ring

29

having the XL-MS topology was superposed on the preliminary model, and refined with Refmac using non-crystallographic symmetry (NCS, using the MEDIUM setting in Refmac) restraints between the four copies of each subunit. The refinement statistics were slightly better for one of the two possible orientations of the XL-MS topology, with CCT2 (XL-MS) matching CCT6 (Dekker et al.). This initial homology model did not contain those loops, which differed in length from the thermosome structure. The difference density for these loops was largely consistent with the loop lengths expected for our model topology. Larger loops tended to be poorly ordered. After an initial round of manual model editing with Coot and a further refinement cycle, strong positive difference density for an alternative, outward facing arrangement of the N-terminus of CCT4 became apparent. Further distinctive features that were not in the initial thermosome model, and that agree with the backbone of the Dekker model, are the N-terminus of CCT8, which reaches toward the opposite ring, and a large insertion in CCT1, residues 484-495. In subsequent refinement cycles, we additionally used the Translation-Libration-Screw (TLS) option of Refmac (Winn et al., 2003), defining individual TRiC subunits as TLS groups. Portions that were not discernible in the 2Fo-Fc maps were not included into the model. Non-glycine residues facing solvent channels without discernible sidechain density were modeled as alanine. *Cis* peptide bonds were disallowed. In the Refmac refinement, the weighting term (keyword "MATRIX") was manually adjusted to 0.004. Otherwise, the default parameters from the CCP4i GUI were used. The final statistics for our model are shown in the last column of Supplementary Table 4). For the calculation of the unbiased omit maps, residues 282-293 and 289-296 in CCT6 and CCT4, respectively, were deleted in the model. After

applying random coordinate shifts up to 0.2 Å using PDBSET, these models were refined with REFMAC, and weighted 2Fo-Fc and Fo-Fc maps were calculated. The final model coordinates were deposited to the PDB database under accession codes 4D8R and 4D8Q.

Conservation score calculations

Bovine and yeast sequences for TRiC subunits were acquired from NCBI and aligned using ClustalW (Thompson et al., 1994). We also generated a large set of homologous sequences using the bovine sequence for each of the eight subunits as a query in a BLAST search (Altschul et al., 1990), these sequence lists were then curated to include a similar organism distribution for each paralog with the help of the KEGG Orthology database (Mao et al., 2005). To compute the per residue conservation scores we employed a Bayesian method implemented in Rate4site (Pupko et al., 2002) and mapped onto the structures using Consurf (Ashkenazy et al., 2010). In brief an MSA file containing the query sequences (ie. 100 orthologous CCT sequences) and a template structural model were used as input (homology model above). The conservation scores derived using this metric correspond to the sites evolutionary rate. The output conservation scores are normalized so that the average score corresponds to 0 with a standard deviation of 1. The lowest score represents the most conserved position in a protein, however, it does not necessarily indicate absolute conservation, but rather indicates that this position is the most conserved in this specific protein calculated using a specific MSA. To visually illustrate the surface conservation across

orthologs, we used pymol to color code the conservation values onto the XL-MS

refined yeast structure.

Supplemental References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J Mol Biol *215*, 403-410.

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics *22*, 195-201.

Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res *38*, W529-533.

Booth, C.R., Meyer, A.S., Cong, Y., Topf, M., Sali, A., Ludtke, S.J., Chiu, W., and Frydman, J. (2008). Mechanism of lid closure in the eukaryotic chaperonin TRiC/CCT. Nat Struct Mol Biol *15*, 746-753.

Chivian, D., and Baker, D. (2006). Homology modeling using parametric alignment ensemble generation with consensus and energy-based model selection. Nucleic Acids Res *34*, e112.

Cong, Y., Baker, M.L., Jakana, J., Woolford, D., Miller, E.J., Reissmann, S., Kumar, R.N., Redding-Johanson, A.M., Batth, T.S., Mukhopadhyay, A*., et al.* (2010). 4.0-A resolution cryo-EM structure of the mammalian chaperonin TRiC/CCT reveals its unique subunit arrangement. Proc Natl Acad Sci U S A *107*, 4967-4972.

Cong, Y., and Ludtke, S.J. (2010). Single particle analysis at high resolution. Methods Enzymol *482*, 211-235.

Cong, Y., Schroder, G.F., Meyer, A.S., Jakana, J., Ma, B., Dougherty, M.T., Schmid, M.F., Reissmann, S., Levitt, M., Ludtke, S.L., *et al.* (2011). Symmetry-free cryo-EM structures of the chaperonin TRiC along its ATPase-driven conformational cycle. The EMBO journal.

Dekker, C., Roe, S.M., McCormack, E.A., Beuron, F., Pearl, L.H., and Willison, K.R. (2011). The crystal structure of yeast CCT reveals intrinsic asymmetry of eukaryotic cytosolic chaperonins. The EMBO journal *30*, 3078-3090.

DiMaio, F., Tyka, M.D., Baker, M.L., Chiu, W., and Baker, D. (2009). Refinement of protein structures into low-resolution density maps using rosetta. J Mol Biol *392*, 181-190.

Ditzel, L., Lowe, J., Stock, D., Stetter, K.O., Huber, H., Huber, R., and Steinbacher, S. (1998). Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. Cell *93*, 125-138.

Feldman, D.E., Spiess, C., Howard, D.E., and Frydman, J. (2003). Tumorigenic mutations in VHL disrupt folding in vivo by interfering with chaperonin binding. Mol Cell *12*, 1213-1224.

Kabir, M.A., Kaminska, J., Segel, G.B., Bethlendy, G., Lin, P., Della Seta, F., Blegen, C., Swiderek, K.M., Zoladek, T., Arndt, K.T., *et al.* (2005). Physiological effects of unassembled chaperonin Cct subunits in the yeast Saccharomyces cerevisiae. Yeast *22*, 219-239.

Kleywegt, G.J., and Jones, T.A. (1994). A super position. CCP4/ESF-EACBM Newsletter on Protein Crystallography *31*, 9-14.

Leitner, A. (2012). Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. Mol Cell Proteomics *in press*.

Ludtke, S.J., Baldwin, P.R., and Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. J Struct Biol *128*, 82-97.

Mao, X., Cai, T., Olyarchuk, J.G., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. Bioinformatics *21*, 3787-3793.

Meyer, A.S., Gillespie, J.R., Walther, D., Millet, I.S., Doniach, S., and Frydman, J. (2003). Closing the folding chamber of the eukaryotic chaperonin requires the transition state of ATP hydrolysis. Cell *113*, 369-381.

Munoz, I.G., Yebenes, H., Zhou, M., Mesa, P., Serna, M., Park, A.Y., Bragado-Nilsson, E., Beloso, A., de Carcer, G., Malumbres, M.*, et al.* (2011). Crystal structure of the open conformation of the mammalian chaperonin CCT in complex with tubulin. Nat Struct Mol Biol *18*, 14-19.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr *53*, 240-255.

Pereira, J.H., Ralston, C.Y., Douglas, N.R., Meyer, D., Knee, K.M., Goulet, D.R., King, J.A., Frydman, J., and Adams, P.D. (2010). Crystal structures of a group II chaperonin reveal the open and closed states associated with the protein folding cycle. J Biol Chem *285*, 27958-27966.

Pupko, T., Bell, R.E., Mayrose, I., Glaser, F., and Ben-Tal, N. (2002). Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. Bioinformatics *18 Suppl 1*, S71-77.

Rinner, O., Seebacher, J., Walzthoeni, T., Mueller, L.N., Beck, M., Schmidt, A., Mueller, M., and Aebersold, R. (2008). Identification of cross-linked peptides from large sequence databases. Nat Methods *5*, 315-318.

Shomura, Y., Yoshida, T., Iizuka, R., Maruyama, T., Yohda, M., and Miki, K. (2004). Crystal structures of the group II chaperonin from Thermococcus strain KS-1: steric

hindrance by the substituted amino acid, and inter-subunit rearrangement between two crystal forms. J Mol Biol *335*, 1265-1278.

Tam, S., Geller, R., Spiess, C., and Frydman, J. (2006). The chaperonin TRiC controls polyglutamine aggregation and toxicity through subunit-specific interactions. Nat Cell Biol *8*, 1155-1162.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. J Struct Biol *157*, 38-46.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res *22*, 4673-4680.

Winn, M.D., Murshudov, G.N., and Papiz, M.Z. (2003). Macromolecular TLS refinement in REFMAC at moderate resolutions. Methods Enzymol *374*, 300-321.

4.2 Paper II: Folding of large multidomain proteins by partial encapsulation in the chaperonin TRiC/CCT.

Folding of large multidomain proteins by partial encapsulation in the chaperonin TRiC/CCT

Florian Rüßmann[a], Markus J. Stemp[a,1], Leonie Mönkemeyer[a], Stephanie A. Etchells[a,2], Andreas Bracher[a,3], and F. Ulrich Hartl[a,b,3]

[a]Department of Cellular Biochemistry, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany; and [b]Center for Integrated Protein Science Munich, 81377 Munich, Germany

The eukaryotic chaperonin, TRiC/CCT (TRiC, TCP-1 ring complex; CCT, chaperonin containing TCP-1), uses a built-in lid to mediate protein folding in an enclosed central cavity. Recent structural data suggest an effective size limit for the TRiC folding chamber of ~70 kDa, but numerous chaperonin substrates are substantially larger. Using artificial fusion constructs with actin, an obligate chaperonin substrate, we show that TRiC can mediate folding of large proteins by segmental or domain-wise encapsulation. Single or multiple protein domains up to ~70 kDa are stably enclosed by stabilizing the ATP-hydrolysis transition state of TRiC. Additional domains, connected by flexible linkers that pass through the central opening of the folding chamber, are excluded and remain accessible to externally added protease. Experiments with the physiological TRiC substrate hSnu114, a 109-kDa multidomain protein, suggest that TRiC has the ability to recognize domain boundaries in partially folded intermediates. In the case of hSnu114, this allows the selective encapsulation of the C-terminal ~45-kDa domain and segments thereof, presumably reflecting a stepwise folding mechanism. The capacity of the eukaryotic chaperonin to overcome the size limitation of the folding chamber may have facilitated the explosive expansion of the multidomain proteome in eukaryotes.

folding cage | molecular chaperone | Snu114 homolog | 116 kDa U5 small nuclear ribonucleoprotein component

It is now widely accepted that many newly synthesized polypeptides require assistance from molecular chaperones to reach their folded states efficiently and at a biologically relevant time scale. The discovery and mechanistic analysis of the chaperonins, a class of molecular chaperones forming 800 to 1,000-kDa double-ring structures, was important in shaping this view (1–4). The chaperonins are essential, ATP-regulated macromolecular machines that function as nano-cages for single protein molecules to fold in isolation, unimpaired by aggregation.

Two distantly related groups of chaperonins are distinguished (4–6): members of group I, also called Hsp60s, occur in eubacteria, mitochondria, and chloroplasts. They have seven ~60-kDa subunits per ring and cooperate with detachable, lid-shaped cofactors (Hsp10 proteins), which function in closing and opening the Hsp60 folding chamber. Group II chaperonins are found in archaea and the eukaryotic cytosol. Their rings are usually homo- or heterooligomers of eight subunits. Unlike the Hsp60s, they function independently of Hsp10 cofactors and instead have an in-built lid consisting of finger-like extensions protruding from the apical domains of the chaperonin subunits.

The paradigm group I chaperonin system is the homooligomeric GroEL protein of *Escherichia coli* and its lid-factor GroES (reviewed in refs. 4, 6, and 7). The apical domains of the GroEL subunits expose hydrophobic amino acid residues, which recognize solvent-exposed hydrophobic regions in unfolded substrate proteins. Binding of ATP and GroES causes an extensive conformational rearrangement, resulting in the substrate protein to be displaced into the central cavity and encapsulated under the hood of GroES. The hydrophobic binding regions on GroEL are buried in the complex with GroES, and the inner surface of the chaperonin cavity becomes hydrophilic, providing an environment permissive for folding. Encapsulated protein is released when GroES dissociates from GroEL after ~10 s, a reaction that is regulated by the GroEL ATPase.

According to cryo-EM and crystal structures, the GroEL-GroES folding cage can accommodate proteins of up to ~70-kDa molecular mass (8, 9). Indeed, the majority of bona fide GroEL substrates are smaller than 50 kDa (10–12), consistent with an average size of soluble *E. coli* proteins of ~35 kDa (Fig. S1). Only a small number of proteins >70 kDa have been reported to interact with GroEL, but these interactions were either nonproductive for folding (12) or used a mechanism of GroEL binding and release without encapsulation by GroES (13). In contrast to prokarya, eukaryotic cells contain a substantially greater number and fraction of multidomain proteins (14), enabling increased structural and functional diversity as well as more complex regulation. Approximately 70% of all eukaryotic proteins are predicted to contain two or more domains (15). Compared with bacteria, the average polypeptide size is increased to ~55 kDa, and 25% of cytosolic proteins in yeast exceed 70 kDa in size (Fig. S1) (16). Interestingly, the eukaryotic group II chaperonin TRiC/CCT (TRiC, TCP-1 ring complex; CCT, chaperonin containing TCP-1) has been shown to interact with numerous proteins >70 kDa, both in mammalian cells and yeast (17, 18) (Fig. S1). However, the folding chamber of TRiC is equivalent in size to that of GroEL-GroES (19–21). How then does TRiC assist the folding of proteins that exceed the apparent size limit of its cavity?

The TRiC chaperonin complex consists of eight distinct, paralogous subunits per ring, which are arranged in a precise orientation (22, 23). These subunits have been highly conserved during evolution from a simpler archaeal precursor (24). They differ in their apical binding regions and are thought to adapt the eukaryotic chaperonin to a range of substrates (25, 26), including many essential components such as the cytoskeletal proteins actin and tubulin, and cell cycle regulators (18, 27–30). Interestingly, many of these proteins have complex domain topologies with a pronounced β-sheet propensity (18), a property that is also characteristic of the substrates of archaeal group II chaperonins (31). Similar to GroEL-GroES, TRiC alternates between open

and closed conformations in an ATP-regulated manner (19–21, 32–36). In contrast to GroEL, however, the closed conformation is induced by formation of the ATP hydrolysis transition state, which can be artificially stabilized by addition of ATP·AlF$_x$ (33). Lid closure is mediated by a camera iris-like rearrangement of the apical subunit extensions (37), inducing formation of an eight-stranded β-barrel structure at the apical pore. The resulting cavity would be large enough to enclose proteins of up to ~70 kDa (20, 21) but not the numerous larger substrates. TRiC might thus encapsulate only those domains of modular proteins that critically require it for folding. The central opening in the closed TRiC complex has a diameter of ~5 Å and thus could accommodate extended linker sequences connecting structured domains in many multidomain proteins.

Here we investigated whether such partial encapsulation of multidomain proteins by TRiC can occur, and how it is accomplished. We explored the functional size of the TRiC cavity using fusion proteins composed of the TRiC-dependent protein actin and variants of GFP. Depending on the sequence context of the actin "domain", these proteins fold to the native state in a TRiC-dependent manner. In the presence of ATP·AlF$_x$, protein fragments up to ~70 kDa are protected from protease digestion by TRiC, indicating transient encapsulation by the chaperonin during the folding reaction. We show further that a physiological

multidomain substrate, hSnu114, which exceeds the chaperonin cavity size considerably, is partially encapsulated. In this case, TRiC selects C-terminal segments of the protein up to ~45 kDa for encapsulation, whereas the N-terminal ~65 kDa of the protein are excluded from the cavity.

Results

TRiC Mediates the Folding of Actin in the Context of Fusion Proteins. The folding of the 42-kDa cytoskeletal protein actin is strictly TRiC-dependent, as shown both in vitro and in vivo (28, 38–40). Native actin binds to DNase I (39, 41, 42), providing a convenient folding assay. To investigate whether actin folding by TRiC can occur when actin is part of larger multidomain proteins, we generated fusion proteins of actin with GFP or with GFP and blue fluorescent protein (BFP) (Fig. 1A). The three "domains," actin (A), GFP (G), and BFP (B), were arranged in different orders connected by flexible glycine-alanine-serine linker sequences of 15–18 residues. A BFP-GFP construct (BG) served as a control protein.

To investigate the folding of the actin fusion proteins, we expressed the respective constructs in vitro in rabbit reticulocyte lysate (RRL) in the presence of [^{35}S]-methionine ([^{35}S]-Met). RRL contains the components of the Hsp70 chaperone system (Hsc70/Hsp40) as well as TRiC and its cofactor prefoldin in functional form (39). Fractionation of RRL by centrifugation
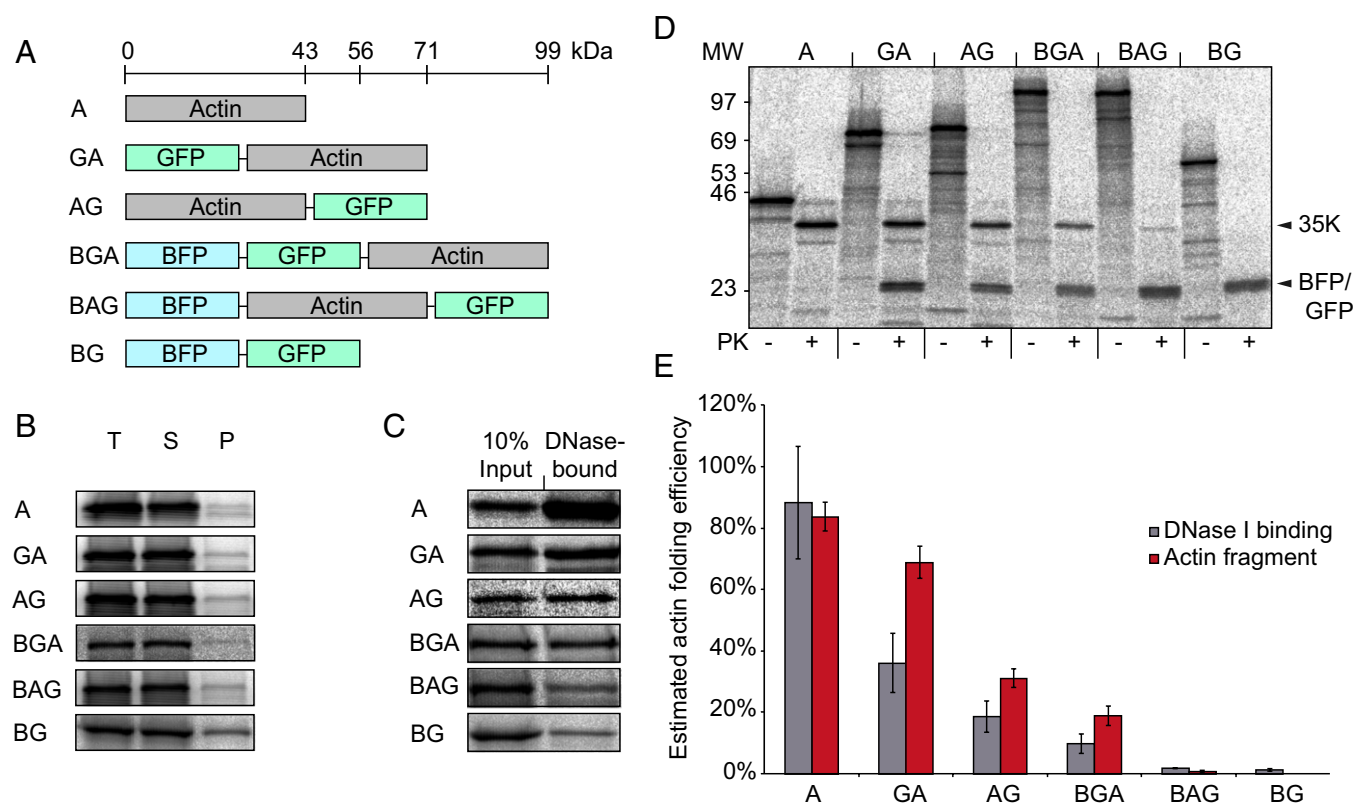

Fig. 1. TRiC-mediated folding of model multidomain proteins. (*A*) Domain structure of actin-fusion constructs. A, actin; GA, GFP-actin; AG, actin-GFP; BGA, BFP-GFP-actin; BAG, BFP-actin-GFP; BG, BFP-GFP. (*B*) Solubility of in vitro translation products. Actin-fusion proteins were synthesized in RRL (90 min, 30 °C) in the presence of [^{35}S]-Met. Total translation reactions (T) were separated by centrifugation into soluble (S) and pellet (P) fractions and analyzed by SDS/PAGE and fluorography. Sections of gels containing the full-length proteins are shown. (*C*) Retention of actin-fusion constructs on DNase I beads. Adjacent lanes show total translation reactions (10% of input material) and the fraction bound to DNase I. The fraction of DNase I-bound BG (~1% of total) represents nonspecific binding. (*D*) SDS/PAGE and fluorography of translation reactions expressing the proteins indicated. Reactions were analyzed after incubation without or with PK (83 μg/mL) for 10 min on ice. The PK-resistant actin fragment of ~35 kDa derived from folded actin and the BFP/GFP domain are indicated. The positions of molecular weight (MW) markers are indicated in kDa. (*E*) Quantification of actin folding yield. The fraction of DNase I bound full-length protein (gray) and the relative intensity of the protease-resistant actin fragment (red) were quantified by densitometry. Averages ± SD from at least three independent experiments are shown. The intensity of the actin fragment (residues 48–375 of actin) relative to full-length actin was calculated by taking the number of methionine residues in actin and GFP/BFP into account.

showed that the full-length proteins of all fusion constructs were soluble (Fig. 1*B*). Retention on DNase I affinity beads after translation was used to estimate the fraction of natively folded actin in the context of the different fusion proteins (Fig. 1*C*). Additionally, we estimated the efficiency of actin folding according to the abundance of a C-terminal protease-resistant actin fragment of ~35 kDa (residues 48 to the actin C terminus, residue 375), produced by digestion with proteinase K (PK) (Fig. 1*D*). This fragment is only generated upon digestion of native or native-like actin (43) but not upon digestion of folding-deficient actin carrying the mutation G150P (44) (Fig. S2).

The vast majority of actin was properly folded upon expression in RRL, as judged by DNase I binding of full-length protein and production of the ~35-kDa fragment (Fig. 1*E*). The actin domain of the fusion proteins GA, AG, and BGA also folded, albeit with lower efficiency than actin alone. This is consistent with the reported coassembly of GFP-actin and actin-GFP fusion proteins with wild-type actin into cytoskeletal filaments in various eukaryotic cells (45), including yeast (46) and neurons (47). The estimated folding yield for the fusion proteins was generally higher when judged by the intensity of the protease-resistant ~35-kDa fragment, suggesting that the fusion partners hindered access of the actin moiety to the DNase I beads (Fig. 1 *D* and *E*). (Note that this effect is not due to the production of the ~35-kDa fragment from prematurely terminated polypeptide chains, because for GA these would not contain complete, foldable actin.) Earlier studies suggested that correct localization of the actin C terminus close to the N terminus (48) occurs late in the folding process and requires flexibility of the terminal segments (49, 50). The lower folding efficiency of AG compared with GA may thus be due to the GFP moiety reducing the mobility of the actin C terminus. Consistent with this interpretation, no significant actin folding was observed for the BAG fusion construct, although the fusion protein BGA still folded with detectable efficiency.

To confirm that actin folding of the fusion proteins was TRiC-dependent, we depleted TRiC from RRL by complex formation with immobilized human phosducin-like protein 1, a known high-affinity interactor of TRiC (51). As expected, retention of actin and GA on DNase I beads was lost or strongly reduced when TRiC was depleted by ~90% but was restored upon readdition of purified bovine TRiC (Fig. S3).

These results demonstrate that TRiC is able to assist actin folding in the context of an N-terminal or C-terminal fusion protein. Folding is prevented, however, when both chain termini of actin are joined to fluorescent protein domains.

Partial Protein Encapsulation by TRiC. To analyze the mechanism by which TRiC enables the folding of actin segments in the context of the fusion proteins, we established a protease protection assay to measure protein encapsulation in the TRiC cavity. The closed form of TRiC is resistant to PK. It is generated during ATP hydrolysis and can be stabilized by AlF$_x$ (in the presence of ATP) to capture the transition state of ATP hydrolysis (33, 52) (Fig. 2 *A* and *B*). In contrast, the apical domains of the TRiC subunits are cleaved by PK in the open state, giving rise to a characteristic pattern of 25- to 37-kDa fragments on SDS/PAGE (33, 53) (Fig. 2*B*). The resulting complex remains assembled under non-denaturing conditions but is unable to encapsulate and fold actin, indicating that folding requires lid closure (33, 37). The open and closed conformations of TRiC can be separated by native PAGE, where the more compact, closed conformation migrates faster (Fig. 2*C*).

Analysis of actin translation reactions by native PAGE showed [^{35}S]-Met labeled product comigrating with TRiC in the high molecular weight region of the gel (Fig. 2*D*). Native actin migrates in the low molecular weight region as a diffuse band (54). The presence of full-length actin in the TRiC complex was confirmed by excising the band from the native gel and
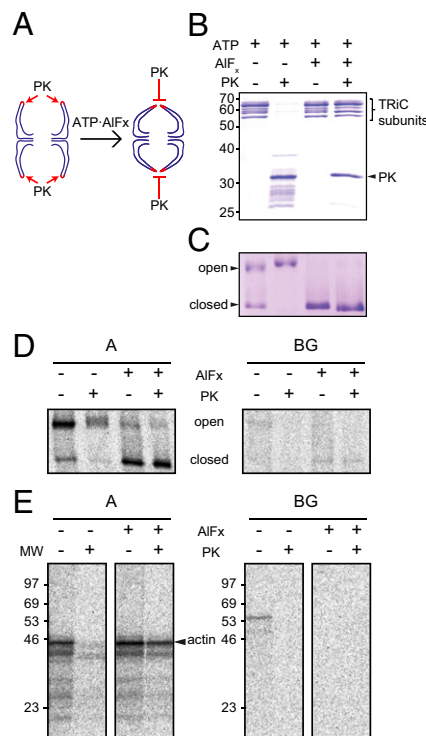

Fig. 2. Stabilizing the closed state of TRiC. (*A*) Principle of protease protection assay. Binding of ATP·AlF$_x$ stabilizes TRiC in the closed state. The apical domains of the TRiC subunits, shown in red, are sensitive to proteolytic cleavage by PK in the open state, but not in the closed state. Note that protease cleavage in the apical domains does not affect complex assembly (33), as shown in *C*. (*B*) Protease protection of TRiC in the closed conformation. Purified bovine TRiC was incubated with or without ATP·AlF$_x$ (60 min at 30 °C), followed by treatment with PK as described in *Materials and Methods*. Reactions were analyzed by SDS/PAGE and Coomassie staining. (*C*) Analysis of TRiC by native PAGE and Coomassie staining. Samples were treated as in *B*. Open TRiC migrates more slowly than the closed form. The TRiC complex remains assembled after PK cleavage. (*D*) (*Left*) Native PAGE analysis of TRiC-bound [^{35}S]-Met labeled actin. (*Right*) [^{35}S]-Met labeled BFP-GFP fusion protein (BG) served as a control for the specificity of TRiC binding. Actin and BG were synthesized in RRL as in Fig. 1. Translation reactions were incubated with or without ATP·AlF$_x$ and subjected to PK treatment before native PAGE. Radiolabeled protein bound to open or closed TRiC was visualized by fluorography. (*E*) Reanalysis of TRiC-bound [^{35}S]-Met actin and BG by SDS/PAGE. The regions of the native gel in *D* containing open and closed TRiC were excised and reanalyzed by SDS/PAGE and fluorography. The positions of molecular weight (MW) markers are indicated in kDa.

reanalyzing it by SDS/PAGE (Fig. 2*E*). Whereas in the untreated translation reaction actin migrated with the open form of TRiC, addition of ATP·AlF$_x$ converted the complex into the closed form (Fig. 2*D*). Upon incubation of open TRiC with PK, actin was no longer detectable, indicating the degradation of the TRiC-bound protein. As reported previously (33), in the presence of ATP·AlF$_x$, TRiC-bound actin was resistant to protease degradation, indicating that the protein was stably encapsulated inside the TRiC cavity (Fig. 2 *D* and *E*). Some incomplete translation products of actin, which are unable to fold, also accumulated on TRiC and were protease protected in the closed complex. BFP-GFP (BG) essentially did not interact with TRiC (Fig. 2 *D* and *E*), indicating that binding of the actin fusion proteins to TRiC was mediated by actin.

Next, we analyzed the encapsulation of the actin fusion proteins by TRiC, as outlined in Fig. 3*A*. Protease treatment of complete translation reactions in the absence of ATP·AlF$_x$
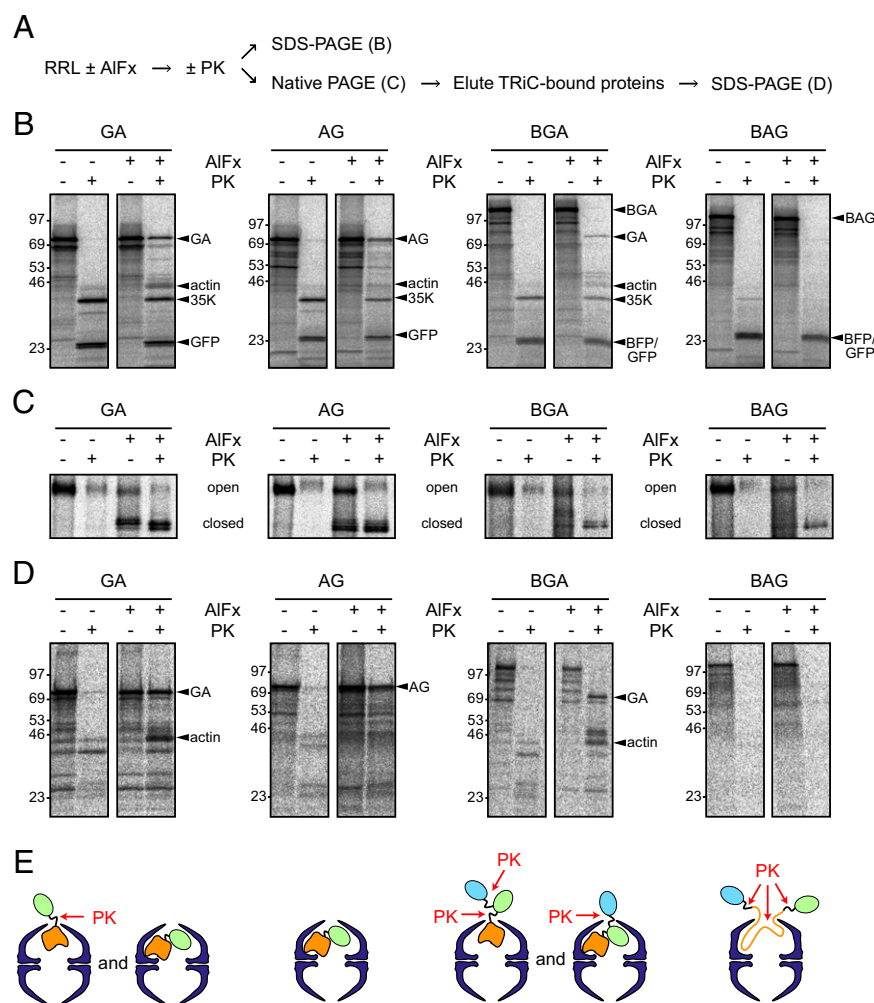
Fig. 3. Encapsulation of actin fusion proteins by TRiC. (*A*) Flow diagram of the experiment. Translation reactions were treated with or without ATP·AlF$_x$, followed by incubation with or without PK and analysis by SDS/PAGE or native PAGE and fluorography. TRiC-containing bands were excised from the native gel and reanalyzed by SDS/PAGE. (*B*) SDS/PAGE and fluorography of complete translation reactions of GA, AG, BGA, and BAG fusion proteins. The position and size of the full-length proteins as well as of protease-protected fragments are indicated. In the presence of ATP·AlF$_x$, additional protease-protected bands are observed. The positions of molecular weight markers are indicated in kDa. (*C*) Native PAGE analysis of TRiC-bound actin fusion proteins. Translation reactions were treated as in *B*. The regions of native PAGE gels containing the TRiC complex are shown. (*D*) Reanalysis of TRiC-bound translation products by SDS/PAGE. Complexes of TRiC with actin fusion proteins were excised from native PAGE gels and separated by SDS/PAGE. The positions of molecular weight markers are indicated in kDa. (*E*) Proposed topology of TRiC-bound actin fusion proteins in the presence of ATP·AlF$_x$. Red arrows show accessibility to protease. Note that in case of TRiC-bound BAG, the TRiC cavity fails to close completely.

resulted in cleavage of the full-length fusion proteins and the production of protease-resistant fragments of ~35 kDa, derived from folded actin, and of ~25 kDa, representing folded BFP or GFP (Fig. 3*B*). When the translation reactions were incubated with ATP·AlF$_x$ before PK treatment, a fraction of the full-length fusion proteins GA and AG was resistant against proteolysis (Fig. 3*B*), suggestive of encapsulation by TRiC. However, the BGA and BAG full-length products were degraded even in the presence of ATP·AlF$_x$. A protected band corresponding in size to GA was generated from BGA but not from BAG (Fig. 3*B*). Furthermore, protected fragments around 45 kDa were observed when actin was the C-terminal domain of the fusion protein (GA and BGA).

Folding intermediates of all actin fusion proteins fractionated with the open form of TRiC on native PAGE and were largely protease-sensitive (Fig. 3*C*). GA and AG also comigrated with the closed form of TRiC in the presence of ATP·AlF$_x$. This material was not diminished by protease treatment, indicative of efficient encapsulation (Fig. 3*C*). In contrast, TRiC-bound BGA and BAG did not migrate as a uniform band in the presence of ATP·AlF$_x$. Treatment with protease after incubation with ATP·AlF$_x$ reduced the amount of TRiC-associated radio-labeled protein derived from BGA and BAG. The protease-protected material remaining showed the typical migration of the closed TRiC complex, suggesting that segments of the large fusion proteins exposed on the outside of the chaperonin had been removed (Fig. 3*C*).

Reanalysis of the TRiC-bound material from the native gel by SDS/PAGE confirmed that the TRiC-bound actin fusion proteins were degraded in the absence of ATP·AlF$_x$ (Fig. 3*D*). Importantly, in the presence of ATP·AlF$_x$, most of the TRiC-bound GA and AG was protected against protease degradation (Fig. 3*D*), confirming that the ATP·AlF$_x$-mediated stabilization of GA and AG observed in the complete translation reaction (Fig. 3*B*) was due to encapsulation by TRiC. In addition, substantial amounts of a protease-protected band corresponding in size to actin were detected in case of GA, suggesting that in a fraction of TRiC:GA complexes only actin (perhaps with various lengths of linker segments) was encapsulated and not the complete GA fusion protein (Fig. 3*E*). Selective encapsulation of the actin moiety was not observed with AG, coinciding with the lower folding efficiency of actin in AG (Fig. 1 *D* and *E*). The TRiC-bound full-length proteins of BGA and BAG remained PK-sensitive in the presence of ATP·AlF$_x$, resulting in protected fragments of ~70 kDa and ~45 kDa in case of BGA (Fig. 3*D*). Thus, cleavage had occurred in the linker regions between BFP and GFP and between GFP and actin, respectively (Fig. 3*E*). In contrast, no defined protected fragments were generated from BAG, demonstrating that this construct can neither be fully nor partially encapsulated inside the TRiC cavity. The BAG-derived radio-labeled material that remained associated with closed TRiC after protease treatment (Fig. 3*C*) apparently represented small peptide fragments that were not resolved by SDS/PAGE (Fig. 3*D*). The cleaved GFP and BFP domains were not detected

because they were removed by PK and no longer migrated with the protease-treated chaperonin on the native gel.

Sizing the TRiC Cavity. The size of the fragments of the GA and BGA fusions protected by TRiC matched the estimated holding capacity of the closed TRiC cavity of ~70 kDa (20). However, we could not rule out that the folded, protease-resistant GFP domain might protrude from the cavity, shielding the flexible interdomain linker region from protease access. In this scenario the volume capacity of the TRiC cavity may be overestimated.

To address this possibility, we generated a version of GA containing a protease-sensitive mutant of GFP, lacking residues 3–38. This N-terminally truncated GFP (ΔN-GFP) (Fig. S4A) is predicted to be defective in forming the stable GFP β-barrel and hence should be degraded when protruding from the TRiC cavity. Indeed, the ~25-kDa band corresponding to protease-resistant GFP was absent in the digestion pattern of ΔN-GA (Fig. S4B). Formation of the protease-resistant actin fragment of ~35 kDa and DNase I binding were strongly diminished compared with GA, suggesting a negative effect of ΔN-GFP on actin folding (Fig. S5). However, when ATP·AlF$_x$ was added to the translation reaction, protection of full-length ΔN-GA from PK digestion was nevertheless observed (Fig. S4B). TRiC-bound ΔN-GA comigrated with the open and closed forms of the chaperonin (Fig. S4C). Excision of the TRiC-bands from the native gel and reanalysis by SDS/PAGE confirmed that full-length ΔN-GA was indeed accommodated in a protease-protected topology by the closed TRiC complex (Fig. S4D). Thus, TRiC is capable of encapsulating a complete protein of the size of ~70 kDa. Of note, the presence of proteolytic fragments in the presence of ATP·AlF$_x$ (Fig. S4D) suggests that the fully occupied TRiC cavity is either dynamic or cannot be locked in the closed conformation, providing transient access to added protease. This phenomenon becomes apparent when the enclosed protein is unable to fold, as in the case of ΔN-GA.

Partial Encapsulation of an Authentic Multidomain Substrate by TRiC. Numerous substrate proteins of yeast and mammalian TRiC exceed the 70-kDa size limit (17, 18) (Fig. S1). We expressed selected homologs of these proteins from human cDNA libraries in RRL (Table S1). The 109-kDa spliceosomal U5 subunit, hSnu114, also known as snRNP116, exhibited high expression levels and robust binding to TRiC, comparable to the actin fusion proteins, and thus was analyzed in detail. hSnu114 is a close sequence homolog of translation elongation factor 2 (eEF2) (55), which has a complex, nested multidomain structure (56) (Fig. 4A): domains G′ and V in eEF2 are inserted into domains G and IV, respectively, resulting in a linear arrangement of four globular domains with α/β structure. Compared with eEF2, hSnu114 has an additional N-terminal extension of 112 predominantly acidic residues, which is predicted to be unstructured.

To probe the folding of hSnu114, we separately translated the full-length protein (amino acids 1–972), the N-terminal part (hSnu114 [1–580]) comprising the unstructured N terminus and domains G, G′, and II, and a C-terminal part (hSnu114 [581–972]) comprising domains III, IV, and V (Fig. 4A). PK digestion of complete translation reactions containing full-length hSnu114 (109 kDa) resulted in the formation of protease-resistant fragments of ~50, 55, and 95 kDa (Fig. 4B). A similar pattern of ~50- and 55-kDa fragments was observed for hSnu114 [1–580], whereas hSnu114 [581–972] was completely degraded. Thus, whereas the N-terminal 580-aa segment apparently contains stable, autonomously folded domains, the C-terminal 392-aa residues of hSnu114 are structured only in the context of the full-length protein. The largest protease-resistant fragment of ~95 kDa is consistent with the removal of an unstructured N-terminal extension, as predicted. Addition of ATP·AlF$_x$ before PK treatment did not significantly change the fragmentation pattern for full-length
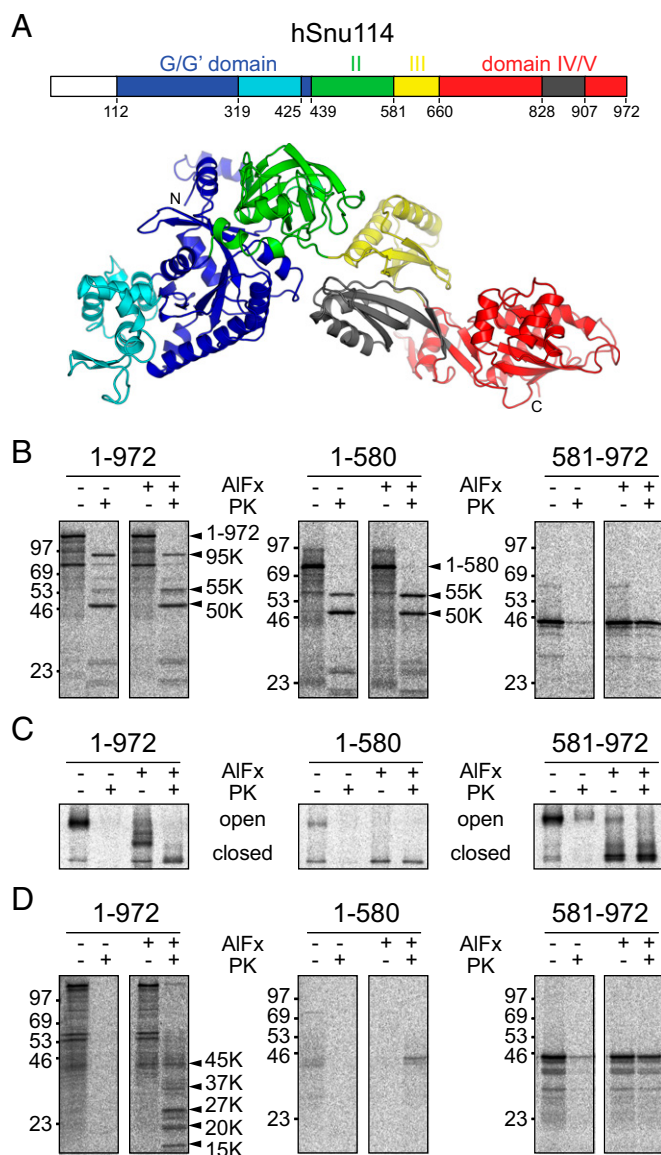

Fig. 4. Interaction of the large substrate protein hSnu114 with TRiC. (A) Structural model of hSnu114. (Upper) Predicted domain structure of hSnu114, based on the crystal structure of the homologous yeast protein, eEF2 (56). The domain nomenclature of eEF2 is used, and residue numbers corresponding to domain boundaries are indicated. (Lower) An hSnu114 structural model from the Swiss-Model Repository in ribbon representation (69). (B) Full-length hSnu114 [1–972] and fragments hSnu114 [1–580] and hSnu114 [581–972] were synthesized in RRL. Complete translation reactions were incubated with or without ATP·AlF$_x$, followed by protease protection assay and analysis by SDS/PAGE and fluorography, as described in Fig. 3B. The position and size of the full-length translation products as well as of protease-protected fragments are indicated in kDa (K). The positions of molecular weight markers are indicated on the left as in Fig. 1D. (C) Native PAGE analysis of TRiC-bound hSnu114 proteins from translation reactions in B (as in Fig. 3C). The region of the native gel containing the TRiC complex is shown. Analysis by fluorography. (D) SDS/PAGE analysis of TRiC-bound hSnu114 proteins. TRiC-hSnu114 protein complexes were excised from native gels in C and reanalyzed by SDS/PAGE and fluorography, as described in Fig. 3D.

hSnu114 and hSnu114 [1–580] when complete translation reactions were analyzed (Fig. 4B). In contrast, hSnu114 [581–972] was completely protease-protected upon addition of ATP·AlF$_x$, suggesting tight association with TRiC (Fig. 4B). Analysis of the TRiC complex by sequential native PAGE and SDS/PAGE

confirmed that full-length hSnu114 and hSnu114 [581–972] bound efficiently to TRiC, whereas hSnu114 [1–580] bound only weakly (Fig. 4 C and D). Indeed, the ~45-kDa hSnu114 [581–972] was completely encapsulated and protease protected by TRiC in the presence of ATP·AlF$_x$ (Fig. 4D). Surprisingly, however, PK treatment of full-length hSnu114 bound to TRiC in the presence of ATP·AlF$_x$ did not result exclusively in the protected fragment of ~45 kDa that was observed with hSnu114 [581–972]. Instead, a series of fragments with ~15-, 20-, 27-, 37-, and 45-kDa molecular mass were trapped inside TRiC (Fig. 4D), which suggests that the C-terminal segment has different folding properties when in context with the TRiC-independent N-terminal domains of the full-length protein. Interestingly, as in the case of the fusion construct BGA (Fig. 3C), protease treatment resulted in faster migration of the closed TRiC:hSnu114 complex in the native gel (Fig. 4C). This suggests that portions of hSnu114 protruding from the cage increase the hydrodynamic radius of the intact complex.

To confirm that the fragments of full-length hSnu114 encapsulated by TRiC were indeed derived from the C-terminal part of the protein, we expressed versions of hSnu114 with either N-terminal or C-terminal HA-epitope tags, HA-hSnu114 and hSnu114-HA, respectively. Consistent with the disorder predicted for the N terminus of hSnu114, HA-tagged fragments were absent in HA-hSnu114 translation reactions after protease treatment in the presence of ATP·AlF$_x$, although initially full-length HA-hSnu114 and incomplete translation products were clearly detectable by anti-HA Western blotting (Fig. 5A). In contrast, we detected HA-reactive protease-protected fragments of ~15, 20, 27, and 37 kDa upon Western blot of an equally treated translation reaction of the C-terminally tagged hSnu114-HA (Fig. 5A). The same band pattern was observed when the TRiC-containing band was excised from native PAGE (Fig. 5B) and reanalyzed by Western blotting (Fig. 5C). Thus, TRiC encapsulates and protects C-terminal segments of hSnu114 presumably derived from domains IV and V (Fig. 5D). These fragments seem to have a strong TRiC binding motif in common, which is likely located in the C-terminal 140 residues, as based on the smallest fragments allowing encapsulation (Fig. 5D). This region contains multiple β-strands, consistent with the finding that many TRiC substrates are rich in β-sheet structure (18) (Fig. 5D).

To test whether the C-terminal 140 residue segment of hSnu114 competes with hSnu114 [581–972] for TRiC binding, we constructed a soluble fusion protein of maltose binding protein (MBP) with hSnu114 residues 830–972. Addition of this protein to the translation displaced hSnu114 [581–972] from TRiC and resulted in its aggregation (Fig. S6). Displacement was ATP-dependent, demonstrating that hSnu114 [581–972] is actively cycling on the chaperonin and apparently incompetent to fold in the absence of the N-terminal domains of hSnu114.

Discussion

Adaptation of TRiC to the Folding of Multidomain Proteins.

The shift toward larger multidomain proteins in eukaryotes must have imposed evolutionary constraints both on the folding proteins as well as the cellular folding machinery. This is due, at least in part, to the need of preventing adjacent, concomitantly folding domains from engaging in aberrant interactions with one another (14, 57). Independent domain folding is facilitated by the sequential emergence of folding units from the ribosome and their cotranslational folding and requires chaperone assistance in many cases (14, 58, 59). Specifically, long-lived domain folding intermediates may need to be shielded by chaperones to facilitate the folding of adjacent modules (Fig. 6). Our present results suggest that the eukaryotic chaperonin TRiC/CCT plays an important role in multidomain protein folding by enclosing parts of proteins too large to be encapsulated in their entirety. This may allow the efficient folding of certain domain topologies within the specialized physical environment of the chaperonin cavity and reduce
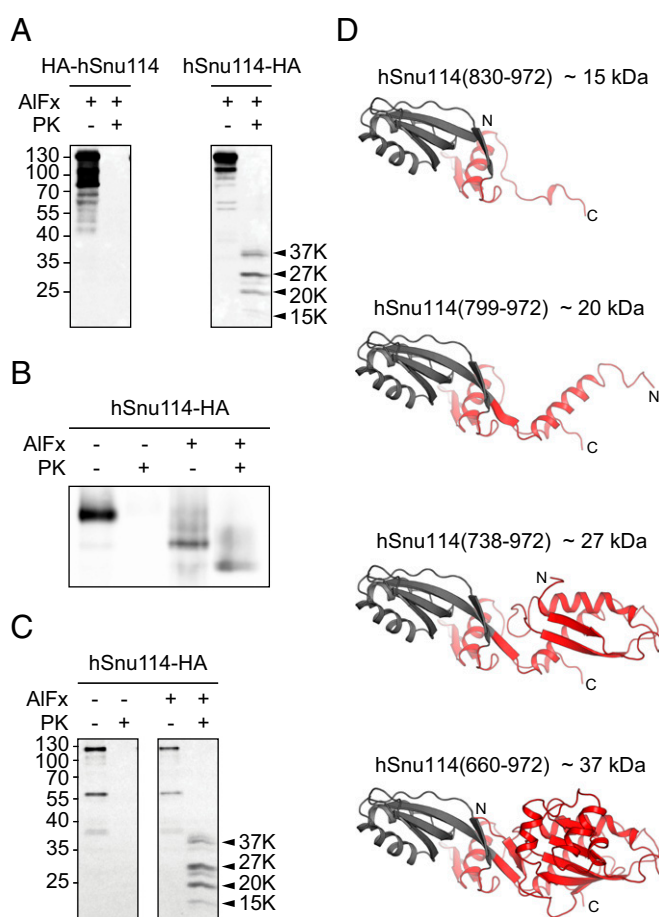
Fig. 5. TRiC recognizes and encapsulates C-terminal segments of hSnu114. (A) Western blot analysis of HA-tagged fragments of hSnu114 protease-protected by TRiC encapsulation. HA-hSnu114 and hSnu114-HA were translated in RRL in the absence of radiolabeled amino acid. PK treatment was performed in the presence of ATP·AlF$_x$ (see Fig. 3B), followed by SDS/PAGE and Western blotting with anti-HA antibody. Protease-protected fragments of hSnu114-HA are indicated. The positions of proteolytic fragments and molecular weight markers are indicated as in Fig. 4B. (B) Native PAGE analysis of TRiC:hSnu114-HA complexes. The experiment was performed as in Fig. 3C, except that analysis was by anti-HA Western blotting. (C) Anti-HA Western blot analysis of TRiC-bound translation products. TRiC-bound material was excised from native PAGE gels and reanalyzed by SDS/PAGE (see Fig. 3D). (D) Putative structures of hSnu114-HA fragments encapsulated by TRiC. C-terminal proteolytic fragments observed in A and C were mapped on the structural model (Fig. 4A). Note that the hSnu114 sequence contains additional 17 amino acid residues at the C terminus, which were not included in the model. The largest fragment corresponds approximately to domains IV (red) and V (gray) (Fig. 4A).

possible interdomain interference during folding. Partial protein encapsulation is also consistent with reports that the eukaryotic chaperonin can interact with nascent polypeptide chains and mediate cotranslational folding in cooperation with the Hsp70 chaperone system (59, 60).

TRiC and other group II chaperonins use helical protrusions from their apical domains that serve as a built-in lid, providing an alternative to the detachable ring-shaped cofactors used by the group I chaperonins (61, 62). Instead, closure in group II chaperonins occurs in an iris-like manner, whereby the tips of the helical protrusions join, resulting in the formation of an apical pore, wide enough for a single peptide strand (Fig. S7). There is genetic as well as structural evidence for a sequential or step-wise movement of the individual TRiC subunits during lid
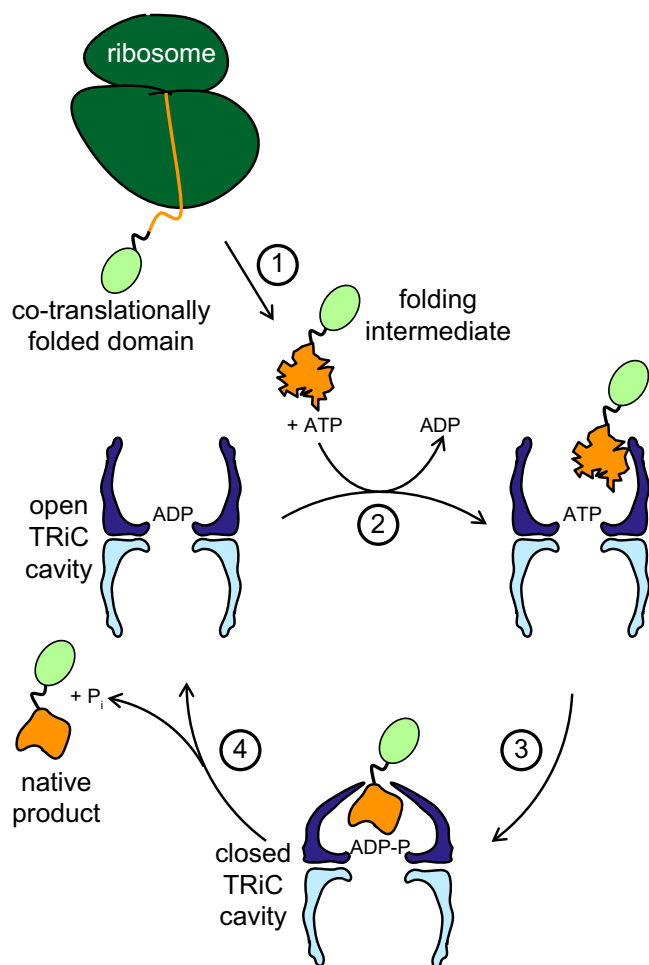
Fig. 6. Folding of multidomain proteins by TRiC. Hypothetical model for the folding of a two-domain protein having a TRiC-dependent C-terminal domain, analogous to the GFP-actin fusion protein. Step 1, the N-terminal domain is TRiC-independent and folds cotranslationally, presumably with the aid of ribosome-associated chaperones and the Hsp70 system (59, 70). Step 2, the folding intermediate containing a nonnative C-terminal domain is stabilized against aggregation by Hsp70 or prefoldin and is posttranslationally transferred to TRiC. Alternatively, TRiC may interact cotranslationally with a chaperonin-dependent domain (59, 60). Step 3, the C-terminal domain is encapsulated by TRiC upon ATP hydrolysis and is induced to fold in the specialized physical environment of the chaperonin cavity (23). The iris-like closing mechanism allows the flexible interdomain linker to protrude through the apical pore. Step 4, upon opening of the TRiC cavity, the substrate is either released as successfully folded, native protein or enters another folding cycle.

closure (35, 36, 63, 64). Thus, a flexibly attached folded domain, itself not interacting with TRiC (such as GFP in the actin fusions), might be excluded from the cavity, adopting a topology in which the extended linker traverses the central channel (Fig. 6). Cotranslational folding of the GFP domain would then explain why the selective encapsulation of actin was more efficient in the fusion construct GA than in AG (Fig. 3D). It is conceivable that the pore exhibits some degree of flexibility. For example, detachment of the helical protrusion from one of the TRiC subunits might allow an expansion, without compromising the geometry needed for sequential ATP hydrolysis within the ring. A recent cryo-EM analysis of intermediates of the TRiC conformational cycle suggests that one subunit, tentatively designated a3, might function as such a "hatch" (20). Although an expanded pore might allow the encapsulation of an internal

segment of a multidomain protein, this was not observed with actin fused between fluorescent protein domains. TRiC may thus preferentially assist the folding of C-terminal or N-terminal protein domains that tend to populate kinetically trapped intermediates (Fig. 6). Whether the compositionally less complex group II chaperonins of archaea already support domain-wise protein encapsulation remains to be addressed, but it is conceivable that this capacity was acquired only more recently in evolution, along with the increase in the number of paralogous chaperonin subunits.

The ability of TRiC to interact with parts of larger protein structures may also help to explain its role in inhibiting the growth of fibrillar aggregates formed by N-terminal fragments of polyglutamine expanded huntingtin (65–67). In this case, TRiC may cap the ends of protofibrils, thereby blocking growth by monomer addition (65).

Partial Encapsulation of the Authentic TRiC Substrate hSnu114. Our data indicate that the TRiC-dependent region of the 109-kDa protein hSnu114 is located close to the C terminus. Notably, rather than a single, well-defined domain, we found a series of C-terminal fragments with approximate sizes of 15, 20, 27, 37, and 45 kDa to be encapsulated by TRiC. The homology model of hSnu114 predicts domain boundaries roughly 45, 37, and 15 kDa from the C terminus, suggesting that TRiC preferably encloses segments beginning at interdomain linkers. Such flexible sequences should be compatible with the steric constraints imposed by the apical pore of TRiC (Fig. S7). The N-terminal domains of hSnu114, comprising residues 1–580, may use folding assistance by the Hsp70 chaperone system, which is known to functionally cooperate with TRiC (59, 68).

Notably, the longest of the TRiC-encapsulated fragments was unable to fold when removed from the context of the full-length protein and was retained by the chaperonin. This suggests that the N-terminal domains form a scaffold necessary to stabilize the structure of the C-terminal region upon its release from TRiC. Surprisingly, the two most prominent C-terminal fragments of 20 and 27 kDa do not coincide with the predicted domain structure of hSnu114. This may suggest that the apical pore of TRiC can select cryptic internal motifs, which either become structured only in the context of the fully folded domain or occur in loop connections at the surface. The detection of a series of C-terminal fragments protected by the TRiC cavity may thus reflect hSnu114 molecules at different stages of a folding process in which C-terminal segments are added stepwise onto a scaffold formed by the N-terminal domains. A global analysis of TRiC-protected protein fragments by proteomics may provide insight into the mechanisms by which the chaperonin assists multidomain protein folding.

Materials and Methods

Detailed experimental procedures are provided in *SI Materials and Methods*. Briefly, TRiC substrates were expressed in vitro using a coupled transcription-translation RRL system in the presence of [^{35}S]-Met. Folding of actin was monitored using retention on immobilized DNase I and by following the formation of a PK-resistant 35-kDa fragment of actin in SDS/PAGE. The closed conformation of TRiC was induced by addition of AlF$_x$ [final concentrations 30 mM KF and 5 mM Al(NO$_3$)$_3$] (33), and substrate exposure was probed with 2.9 μM PK (10 min on ice) followed by SDS/PAGE analysis. TRiC-bound material was isolated by native PAGE using gels containing 6 mM KF and 1 mM Al(NO$_3$)$_3$ and reanalyzed by SDS/PAGE.

1. Ellis RJ (2003) Protein folding: Importance of the Anfinsen cage. *Curr Biol* 13(22): R881–R883.
2. Gershenson A, Gierasch LM (2011) Protein folding in the cell: Challenges and progress. *Curr Opin Struct Biol* 21(1):32–41.
3. Hartl FU (1996) Molecular chaperones in cellular protein folding. *Nature* 381(6583): 571–579.
4. Horwich AL, Fenton WA, Chapman E, Farr GW (2007) Two families of chaperonin: Physiology and mechanism. *Annu Rev Cell Dev Biol* 23:115–145.
5. Frydman J (2001) Folding of newly translated proteins in vivo: The role of molecular chaperones. *Annu Rev Biochem* 70:603–647.
6. Hartl FU, Bracher A, Hayer-Hartl M (2011) Molecular chaperones in protein folding and proteostasis. *Nature* 475(7356):324–332.
7. Saibil HR, Ranson NA (2002) The chaperonin folding machine. *Trends Biochem Sci* 27 (12):627–632.
8. Clare DK, Bakkes PJ, van Heerikhuizen H, van der Vies SM, Saibil HR (2009) Chaperonin complex with a newly folded protein encapsulated in the folding chamber. *Nature* 457(7225):107–110.
9. Xu Z, Horwich AL, Sigler PB (1997) The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. *Nature* 388(6644):741–750.
10. Fujiwara K, Ishihama Y, Nakahigashi K, Soga T, Taguchi H (2010) A systematic survey of in vivo obligate chaperonin-dependent substrates. *EMBO J* 29(9):1552–1564.
11. Houry WA, Frishman D, Eckerskorn C, Lottspeich F, Hartl FU (1999) Identification of in vivo substrates of the chaperonin GroEL. *Nature* 402(6758):147–154.
12. Kerner MJ, et al. (2005) Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell* 122(2):209–220.
13. Chaudhuri TK, Farr GW, Fenton WA, Rospert S, Horwich AL (2001) GroEL/GroES-mediated folding of a protein too large to be encapsulated. *Cell* 107(2):235–246.
14. Netzer WJ, Hartl FU (1997) Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* 388(6640):343–349.
15. Jacob E, Horovitz A, Unger R (2007) Different mechanistic requirements for prokaryotic and eukaryotic chaperonins: a lattice study. *Bioinformatics* 23(13):i240–i248.
16. Netzer WJ, Hartl FU (1998) Protein folding in the cytosol: Chaperonin-dependent and -independent mechanisms. *Trends Biochem Sci* 23(2):68–73.
17. Dekker C, et al. (2008) The interaction network of the chaperonin CCT. *EMBO J* 27(13): 1827–1839.
18. Yam AY, et al. (2008) Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies. *Nat Struct Mol Biol* 15 (12):1255–1262.
19. Cong Y, et al. (2010) 4.0-A resolution cryo-EM structure of the mammalian chaperonin TRiC/CCT reveals its unique subunit arrangement. *Proc Natl Acad Sci USA* 107(11): 4967–4972.
20. Cong Y, et al. (2012) Symmetry-free cryo-EM structures of the chaperonin TRiC along its ATPase-driven conformational cycle. *EMBO J* 31(3):720–730.
21. Dekker C, et al. (2011) The crystal structure of yeast CCT reveals intrinsic asymmetry of eukaryotic cytosolic chaperonins. *EMBO J* 30(15):3078–3090.
22. Kalisman N, Adams CM, Levitt M (2012) Subunit order of eukaryotic TRiC/CCT chaperonin by cross-linking, mass spectrometry, and combinatorial homology modeling. *Proc Natl Acad Sci USA* 109(8):2884–2889.
23. Leitner A, et al. (2012) The molecular architecture of the eukaryotic chaperonin TRiC/CCT. *Structure* 20(5):814–825.
24. Archibald JM, Blouin C, Doolittle WF (2001) Gene duplication and the evolution of group II chaperonins: implications for structure and function. *J Struct Biol* 135(2): 157–169.
25. Amit M, et al. (2010) Equivalent mutations in the eight subunits of the chaperonin CCT produce dramatically different cellular and gene expression phenotypes. *J Mol Biol* 401(3):532–543.
26. Spiess C, Miller EJ, McClellan AJ, Frydman J (2006) Identification of the TRiC/CCT substrate binding sites uncovers the function of subunit diversity in eukaryotic chaperonins. *Mol Cell* 24(1):25–37.
27. Frydman J, et al. (1992) Function in protein folding of TRiC, a cytosolic ring complex containing TCP-1 and structurally related subunits. *EMBO J* 11(13):4767–4778.
28. Gao Y, Thomas JO, Chow RL, Lee GH, Cowan NJ (1992) A cytoplasmic chaperonin that catalyzes beta-actin folding. *Cell* 69(6):1043–1050.
29. Spiess C, Meyer AS, Reissmann S, Frydman J (2004) Mechanism of the eukaryotic chaperonin: Protein folding in the chamber of secrets. *Trends Cell Biol* 14(11):598–604.
30. Yaffe MB, et al. (1992) TCP1 complex is a molecular chaperone in tubulin biogenesis. *Nature* 358(6383):245–248.
31. Hirtreiter AM, et al. (2009) Differential substrate specificity of group I and group II chaperonins in the archaeon *Methanosarcina mazei*. *Mol Microbiol* 74(5):1152–1168.
32. Booth CR, et al. (2008) Mechanism of lid closure in the eukaryotic chaperonin TRiC/CCT. *Nat Struct Mol Biol* 15(7):746–753.
33. Meyer AS, et al. (2003) Closing the folding chamber of the eukaryotic chaperonin requires the transition state of ATP hydrolysis. *Cell* 113(3):369–381.
34. Muñoz IG, et al. (2011) Crystal structure of the open conformation of the mammalian chaperonin CCT in complex with tubulin. *Nat Struct Mol Biol* 18(1):14–19.
35. Rivenzon-Segal D, Wolf SG, Shimon L, Willison KR, Horovitz A (2005) Sequential ATP-induced allosteric transitions of the cytoplasmic chaperonin containing TCP-1 revealed by EM analysis. *Nat Struct Mol Biol* 12(3):233–237.
36. Shimon L, Hynes GM, McCormack EA, Willison KR, Horovitz A (2008) ATP-induced allostery in the eukaryotic chaperonin CCT is abolished by the mutation G345D in CCT4 that renders yeast temperature-sensitive for growth. *J Mol Biol* 377(2):469–477.
37. Reissmann S, Parnot C, Booth CR, Chiu W, Frydman J (2007) Essential function of the built-in lid in the allosteric regulation of eukaryotic and archaeal chaperonins. *Nat Struct Mol Biol* 14(5):432–440.
38. Chen X, Sullivan DS, Huffaker TC (1994) Two yeast genes with similarity to TCP-1 are required for microtubule and actin function in vivo. *Proc Natl Acad Sci USA* 91(19): 9111–9115.
39. Stemp MJ, Guha S, Hartl FU, Barral JM (2005) Efficient production of native actin upon translation in a bacterial lysate supplemented with the eukaryotic chaperonin TRiC. *Biol Chem* 386(8):753–757.
40. Vinh DB, Drubin DG (1994) A yeast TCP-1-like protein is required for actin function in vivo. *Proc Natl Acad Sci USA* 91(19):9116–9120.
41. Kabsch W, Mannherz HG, Suck D, Pai EF, Holmes KC (1990) Atomic structure of the actin:DNase I complex. *Nature* 347(6288):37–44.
42. Lazarides E, Lindberg U (1974) Actin is the naturally occurring inhibitor of deoxyribonuclease I. *Proc Natl Acad Sci USA* 71(12):4742–4746.
43. Kuznetsova IM, et al. (1999) Effect of self-association on the structural organization of partially folded proteins: Inactivated actin. *Biophys J* 77(5):2788–2800.
44. McCormack EA, Llorca O, Carrascosa JL, Valpuesta JM, Willison KR (2001) Point mutations in a hinge linking the small and large domains of beta-actin result in trapped folding intermediates bound to cytosolic chaperonin CCT. *J Struct Biol* 135(2): 198–204.
45. Yoon Y, Pitts K, McNiven M (2002) Studying cytoskeletal dynamics in living cells using green fluorescent protein. *Mol Biotechnol* 21(3):241–250.
46. Doyle T, Botstein D (1996) Movement of yeast cortical actin cytoskeleton visualized in vivo. *Proc Natl Acad Sci USA* 93(9):3886–3891.
47. Fischer M, Kaech S, Knutti D, Matus A (1998) Rapid actin-based plasticity in dendritic spines. *Neuron* 20(5):847–854.
48. Kabsch W, Holmes KC (1995) The actin fold. *FASEB J* 9(2):167–174.
49. Lee JY, Duan L, Iverson TM, Dima RI (2012) Exploring the role of topological frustration in actin refolding with molecular simulations. *J Phys Chem B* 116(5):1677–1686.
50. Neirynck K, Waterschoot D, Vandekerckhove J, Ampe C, Rommelaere H (2006) Actin interacts with CCT via discrete binding sites: A binding transition-release model for CCT-mediated actin folding. *J Mol Biol* 355(1):124–138.
51. McLaughlin JN, et al. (2002) Regulatory interaction of phosducin-like protein with the cytosolic chaperonin complex. *Proc Natl Acad Sci USA* 99(12):7962–7967.
52. Melki R, Cowan NJ (1994) Facilitated folding of actins and tubulins occurs via a nucleotide-dependent interaction between cytoplasmic chaperonin and distinctive folding intermediates. *Mol Cell Biol* 14(5):2895–2904.
53. Szpikowska BK, Swiderek KM, Sherman MA, Mas MT (1998) MgATP binding to the nucleotide-binding domains of the eukaryotic cytoplasmic chaperonin induces conformational changes in the putative substrate-binding domains. *Protein Sci* 7(7): 1524–1530.
54. Frydman J, Hartl FU (1996) Principles of chaperone-assisted protein folding: Differences between in vitro and in vivo mechanisms. *Science* 272(5267):1497–1502.
55. Fabrizio P, Laggerbauer B, Lauber J, Lane WS, Lührmann R (1997) An evolutionarily conserved U5 snRNP-specific protein is a GTP-binding factor closely related to the ribosomal translocase EF-2. *EMBO J* 16(14):4092–4106.
56. Jørgensen R, et al. (2003) Two crystal structures demonstrate large conformational changes in the eukaryotic ribosomal translocase. *Nat Struct Biol* 10(5):379–385.
57. Wright CF, Teichmann SA, Clarke J, Dobson CM (2005) The importance of sequence diversity in the aggregation and evolution of proteins. *Nature* 438(7069):878–881.
58. Frydman J, Erdjument-Bromage H, Tempst P, Hartl FU (1999) Co-translational domain folding as the structural basis for the rapid de novo folding of firefly luciferase. *Nat Struct Biol* 6(7):697–705.
59. Frydman J, Nimmesgern E, Ohtsuka K, Hartl FU (1994) Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones. *Nature* 370 (6485):111–117.
60. Etchells SA, et al. (2005) The cotranslational contacts between ribosome-bound nascent polypeptides and the subunits of the hetero-oligomeric chaperonin TRiC probed by photocross-linking. *J Biol Chem* 280(30):28118–28126.
61. Zhang J, et al. (2010) Mechanism of folding chamber closure in a group II chaperonin. *Nature* 463(7279):379–383.
62. Zhang J, et al. (2011) Cryo-EM structure of a group II chaperonin in the prehydrolysis ATP-bound state leading to lid closure. *Structure* 19(5):633–639.
63. Lin P, Sherman F (1997) The unique hetero-oligomeric nature of the subunits in the catalytic cooperativity of the yeast Cct chaperonin complex. *Proc Natl Acad Sci USA* 94(20):10780–10785.
64. Reissmann S, et al. (2012) A gradient of ATP affinities generates an asymmetric power stroke driving the chaperonin TRIC/CCT folding cycle. *Cell Rep* 2(4):866–877.
65. Behrends C, et al. (2006) Chaperonin TRiC promotes the assembly of polyQ expansion proteins into nontoxic oligomers. *Mol Cell* 23(6):887–897.
66. Kitamura A, et al. (2006) Cytosolic chaperonin prevents polyglutamine toxicity with altering the aggregation state. *Nat Cell Biol* 8(10):1163–1170.
67. Tam S, Geller R, Spiess C, Frydman J (2006) The chaperonin TRiC controls polyglutamine aggregation and toxicity through subunit-specific interactions. *Nat Cell Biol* 8(10):1155–1162.
68. Cuéllar J, et al. (2008) The structure of CCT-Hsc70 NBD suggests a mechanism for Hsp70 delivery of substrates to the chaperonin. *Nat Struct Mol Biol* 15(8):858–864.
69. Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T (2009) The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res* 37(Database issue):D387–D392.
70. Kramer G, Boehringer D, Ban N, Bukau B (2009) The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nat Struct Mol Biol* 16(6):589–597.

Supporting Information

Rüßmann et al. 10.1073/pnas.1218836109

SI Materials and Methods

Plasmids and Constructs. Actin fusion constructs were expressed in rabbit reticulocyte lysate (RRL) from a pET22b-based vector under the control of a T7 promoter using the actin sequence of *Saccharomyces cerevisiae*, Act1p. All actin fusion constructs carried a C-terminal 6xHis-tag. The actin, blue fluorescent protein (BFP), and GFP domains were connected by flexible 15- to 18-aa linkers using the sequence SGSAASAAGAGEAAA (with minor modifications). The GFP domains were based on the GFPcycle3 sequence (1). BFP was lacking the cycle3 mutations and carried the chromophore mutation Y66H (2). Protease-sensitive GFP (ΔN-GFP) was created by deletion of amino acids 3–38.

Plasmids encoding putative TRiC substrates (3, 4) were obtained from Mammalian Gene Collection (MGC) human verified full-length cDNA libraries (imaGenes) and expressed under the control of T7 (MYH1, THNSL1), T3 (EEF2, MMS19, MUC17), or SP6 (TSHZ3, EFTUD2, UBA1, LARS, SKIV2L, KIF13A, FASN, USP10, ATP6V0A1, COPB2, GANAB, SMARCC2, NUP153, DYNC1H1) promoters in coupled RRL transcription–translation reactions (Promega).

HA-hSnu114 and hSnu114-HA were expressed from a pET22b-based vector under the control of a T7 promoter and carried an HA tag (YPYDVPDYA) at the N or C terminus, respectively. HA-hSnu114 carried a C-terminal and hSnu114-HA an N-terminal FLAG-tag (DYKDDDDK) in addition.

Human phosducin-like protein 1 (PhLP1) carrying an N-terminal 6xHis-Tag was expressed from a pProEx-HTb construct.

MBP-hSnu114(830-972) was expressed from a pMal-c2-based vector.

In Vitro Transcription–Translation Reactions.
Proteins were expressed in vitro in coupled RRL transcription–translation reactions using TNT coupled reticulocyte lysate (Promega). The reactions were performed in the presence of 10 μCi L-[^{35}S]-Met (Perkin-Elmer) and 1 μg plasmid DNA per 50 μL reaction for 60–90 min at 30 °C. Translation was terminated before DNase I binding by incubation with 10 U/mL apyrase (Sigma-Aldrich) for 2 min at 30 °C to ensure that no further translation or TRiC-mediated folding occurred during the DNase I binding reaction. Apyrase treatment was omitted before protease protection assays.

Solubility of translation products was analyzed by centrifugation of total translation lysate at $20,000 \times g$ for 30 min at 4 °C. Total, pellet, and supernatant fractions were analyzed by SDS/PAGE and fluorography.

Protein Purification.
TRiC was purified from bovine testis essentially as previously described (5, 6), with the following modifications. Pooled fractions from HiTrap heparin (GE Healthcare) chromatography were further separated on a HiLoad Superdex 200 (GE Healthcare) gel filtration column. Fractions containing the 1 MDa TRiC complex were pooled and concentrated to 4–5 mg protein/mL using Vivaspin (Sartorius) centrifugal concentrators.

Human PhLP1 was expressed in *Escherichia coli* BL21 cells and purified using a HisTrap Ni Sepharose column (GE Healthcare). Imidazole was removed using a HiTrap desalting column. The final protein concentration was ~11 mg/mL, determined by absorbance at 280 nm.

MBP-hSnu114(830–972) was expressed in *E. coli* and purified using Amylose Resin (New England Biolabs). The final protein concentration was ~19 mg/mL, determined by absorbance at 280 nm.

DNase I Binding Assay.
Translation reactions were diluted 10-fold with buffer [1× PBS, 5 mM EDTA, 1 mM DTT, 0.1% Tween 20, and 1× Complete Protease Inhibitor Mixture (Roche)] before incubation with DNase I (Roche) immobilized on CNBr-activated Sepharose 4B (GE Healthcare) for 120 min at 4 °C. After washing the resin once with 500 μL buffer W1 [50 mM Tris·HCl (pH 7.4), 150 mM NaCl, 5 mM EDTA, 1% Triton X-100], twice with 500 μL buffer W2 [50 mM Tris·HCl (pH 7.4), 500 mM NaCl, 5 mM EDTA, and 1% Triton X-100] and once with 500 μL buffer W3 (1× PBS) in a mini spin column, bound protein was eluted with 2× SDS-loading buffer [4% (wt/vol) SDS, 20% (vol/vol) glycerol, 5% (vol/vol) β-mercaptoethanol, and 120 mM Tris·HCl (pH 6.8)] by incubation at 95 °C for 2 min followed by centrifugation at $16,100 \times g$ for 2 min.

Estimating Actin Folding Efficiency from 35K Fragment.
Actin fusion proteins were transcribed and translated in RRL in the presence of L-[^{35}S]-Met at 30 °C for 90 min. Translation was stopped by incubation with 110 μM cycloheximide (Sigma-Aldrich) at 30 °C for 2 min. Proteinase K (PK) digest was carried out for 10 min on ice with a final PK concentration of 83 μg/mL. Protease action was stopped by incubation with 10 mM phenylmethane sulfonylfluoride. Samples were analyzed by SDS/PAGE and fluorography. Quantification of band intensities was done using AIDA software version 4.15 (Raytest). The actin folding efficiency was estimated from the ratio of the intensities of the 35K fragment and the corresponding full-length protein corrected for the number of methionines contained in the respective polypeptide.

Depletion of TRiC from Reticulocyte Lysate.
PhLP1 (650 μg) was incubated with 500 μL PBS containing 40 mg/mL Dynabeads TALON (Dynal Biotech) for 2 h at 4 °C. After washing three times with 1.5 mL PBS, the beads were incubated with 130 μL coupled RRL for 2 h at 4 °C. Efficiency of depletion was analyzed by Western blotting using an antibody specific for CCTα (Stressgen). Depleted RRL was used either directly for transcription and translation or after supplementation with 1 μM purified bovine TRiC.

Protease Protection by TRiC.
Proteins were transcribed and translated in RRL in the presence of L-[^{35}S]-Met as described above (60 min at 30 °C). Samples were either treated with ATP·AlF$_x$ [for a 50-μL reaction: 1.5 μL of 1 M KF, 1 μL of 250 mM Al(NO$_3$)$_3$, and 0.5 μL of 100 mM ATP] or mock treated (for a 50-μL reaction: 2.5 μL H$_2$O and 0.5 μL of 100 mM ATP) for 30–60 min at 30 °C (7). PK digest was carried out for 10 min on ice with a final PK concentration of 83 μg/mL. Protease action was stopped by addition of PMSF to a final concentration of 10 mM.

Native PAGE.
Proteins were separated by 5–13% (wt/vol) acrylamide gradient native PAGE essentially as previously described (8). Native gels were run as Clear Native PAGE at 4 °C for 16–18 h, initially at 100 V. The voltage was increased to 200 V when the sample had entered the stacking gel. Including low concentrations of AlF$_x$ [6 mM KF, 1 mM Al(NO$_3$)$_3$] into the gel matrix and running buffers resulted in stabilization of the closed conformation of TRiC during electrophoresis.

Protein Extraction from Native PAGE Slices.
TRiC bands were visualized by Coomassie staining (identified using purified TRiC

standard) and excised. The slices were incubated in 150 μL 2×
SDS loading buffer for 2 h at 68 °C with constant shaking at
1,400 rpm, followed by incubation at 95 °C for 10 min. The sample
was then centrifuged at 20,000 × g for 10 min. Proteins were
precipitated from the supernatant by incubation with sodium de-
oxycholate and trichloroacetic acid on ice. Precipitated proteins

were pelleted by centrifugation, washed once with ice-cold ace-
tone, pelleted again by centrifugation, and finally resuspended in
2× SDS loading buffer and analyzed by SDS/PAGE.

Bioinformatic Methods. Protein structures were displayed using
PyMOL (www.pymol.org).

1. Crameri A, Whitehorn EA, Tate E, Stemmer WP (1996) Improved green fluorescent
   protein by molecular evolution using DNA shuffling. *Nat Biotechnol* 14(3):315–319.
2. Heim R, Prasher DC, Tsien RY (1994) Wavelength mutations and posttranslational
   autoxidation of green fluorescent protein. *Proc Natl Acad Sci USA* 91(26):12501–12504.
3. Dekker C, et al. (2008) The interaction network of the chaperonin CCT. *EMBO J* 27(13):
   1827–1839.
4. Yam AY, et al. (2008) Defining the TRiC/CCT interactome links chaperonin function to
   stabilization of newly made proteins with complex topologies. *Nat Struct Mol Biol*
   15(12):1255–1262.

5. Frydman J, et al. (1992) Function in protein folding of TRiC, a cytosolic ring complex
   containing TCP-1 and structurally related subunits. *EMBO J* 11(13):4767–4778.
6. Villebeck L, et al. (2007) Conformational rearrangements of tail-less complex polypeptide
   1 (TCP-1) ring complex (TRiC)-bound actin. *Biochemistry* 46(17):5083–5093.
7. Meyer AS, et al. (2003) Closing the folding chamber of the eukaryotic chaperonin
   requires the transition state of ATP hydrolysis. *Cell* 113(3):369–381.
8. Schägger H, von Jagow G (1991) Blue native electrophoresis for isolation of
   membrane protein complexes in enzymatically active form. *Anal Biochem* 199(2):
   223–231.

Fig. S1. Size distribution of substrate proteins of bacterial and eukaryotic chaperonins. (*A*) Size distributions of the *E. coli* and *S. cerevisiae* proteomes. (*B*) Size
distributions of the interactors of *E. coli* GroEL (1) and *S. cerevisiae* TRiC (2).

1. Kerner MJ, et al. (2005) Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell* 122(2):209–220.
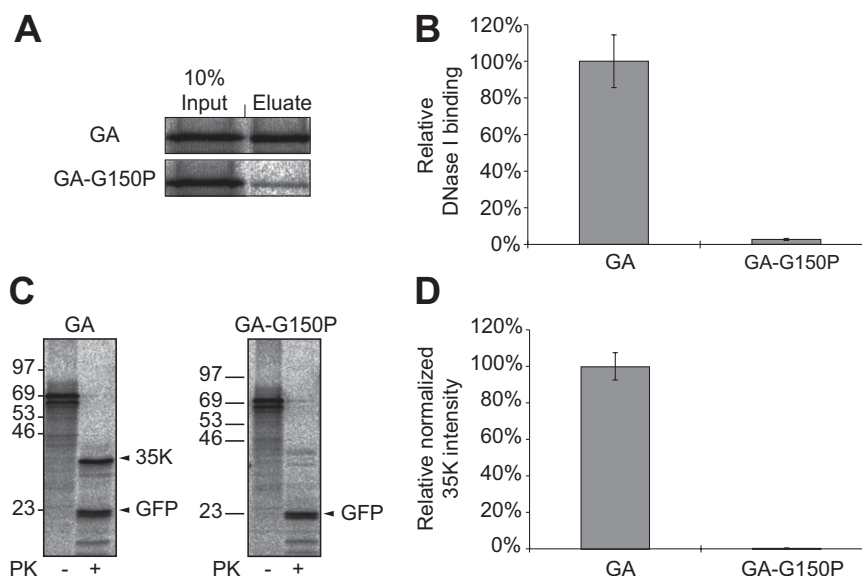2. Dekker C, et al. (2008) The interaction network of the chaperonin CCT. *EMBO J* 27(13):1827–1839.

Fig. S2. Folding of GFP-actin(G150P) fusion protein. (*A*) Retention of actin-fusion constructs, GFP-actin (GA), and GFP-actin(G150P) (GA-G150P) on DNase I beads. Total translation reactions (10% of input material) and the fraction bound to DNase I (Eluate) were analyzed by SDS/PAGE and fluorography (Fig. 1*C*). (*B*) Quantification of actin folding yield. The fraction of DNase I bound full-length translation product was quantified by densitometry. Averages ± SD from three independent experiments are shown. DNase I binding of GA is set to 100%. (*C*) SDS/PAGE and fluorography of translation reactions containing GA and GA-G150P before and after treatment with PK. The PK-resistant fragment of 35K derived from actin and the PK-resistant GFP domain are indicated. (*D*) Quantification of the relative intensity of the PK-resistant actin fragment of 35K. Error bars represent SD values from three independent experiments. The relative intensity of the 35K fragment of GA is set to 100%.
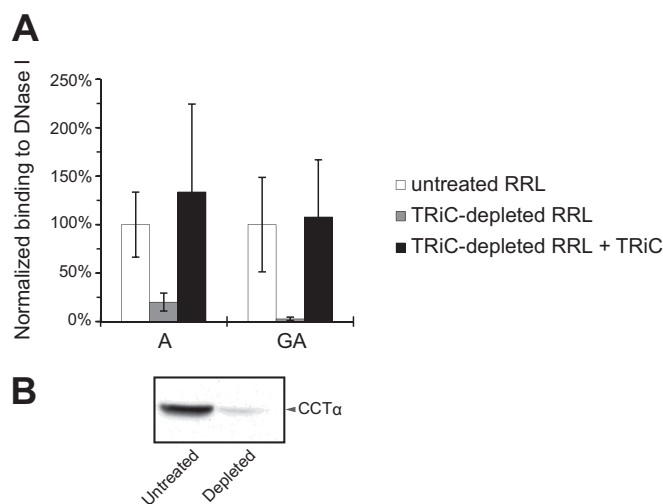

Fig. S3. TRiC-dependence of actin folding. (*A*) Folding of actin and actin fusion protein is TRiC-dependent. Folding efficiency of actin (A) and GFP-actin (GA) was measured by DNase I retention assay as in Fig. 1 and Fig. S2. The fraction of full-length translation product retained on DNase I beads is shown for unmodified RRL, TRiC-depleted RRL, and TRiC-depleted RRL supplemented with purified bovine TRiC. Error bars represent SD values for three independent experiments. (*B*) Depletion of TRiC from RRL confirmed by Western blotting with anti-CCTα antibody.

Fig. S4. Probing the size of the TRiC cavity. (A) Schematic representation of GFP-actin and ΔN-GFP-actin fusion proteins. Sequence elements of GFP missing in ΔN-GFP are indicated in yellow in the ribbon diagram of the GFP crystal structure. (B) Translation reactions as in Fig. 3B were incubated with or without ATP•AlF$_x$ and subjected to PK treatment, as indicated. Analysis by SDS/PAGE and fluorography is shown. (C) Native PAGE analysis of TRiC-bound ΔN-GA. Translation reactions were treated as in B. The region of the native gel containing the TRiC complex is shown and was analyzed by fluorography. (D) SDS/PAGE analysis of TRiC-bound ΔN-GA translation products. The TRiC:ΔN-GA complexes were excised from native gel in C and reanalyzed by SDS/PAGE/fluorography, as described in Fig. 3D.

Fig. S5. Folding of ΔN-GFP-actin fusion protein. (*A*) SDS/PAGE and fluorography of translation reactions containing actin fusion proteins GFP-actin (GA) and ΔN-GFP-actin (ΔN-GA) before and after treatment with PK. The PK-resistant fragment of 35K derived from actin and the PK-resistant GFP domain produced from GA are indicated. (*B*) Quantification of the relative intensity of the PK-resistant actin fragment of 35K. Error bars represent SD values of three independent experiments. The relative intensity of the 35K fragment of GA is set to 100%. (*C*) Quantification of folding yield by DNase I binding. The fraction of full-length actin-fusion constructs GA and ΔN-GA retained by DNase I beads was quantified by densitometry as in Fig. 1*E*. Averages ± SD of three independent experiments are shown. DNase I binding of GA is set to 100%.
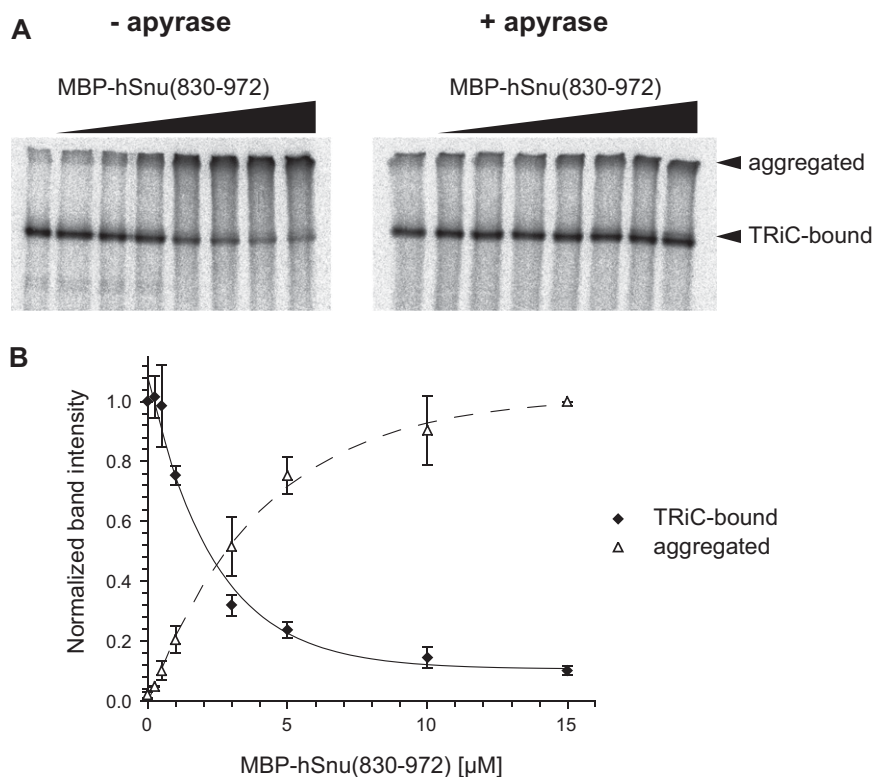
Fig. S6. ATP-dependent cycling of C-terminal hSnu114 fragment on TRiC. (*A*) Displacement of hSnu114 [581–972] from TRiC by MBP-hSnu(830–972) as visualized by native PAGE. hSnu114 [581–972] was synthesized in RRL in the presence of [^{35}S]-Met as in Fig. 4*A*. Translation reactions were stopped by addition of cycloheximide (*Left*; - apyrase) or apyrase (*Right*; + apyrase) and 5 min later increasing concentrations of purified MBP-hSnu(830–972) (0–15 μM) were added. After 30 min, reactions were analyzed by native PAGE and fluorography. The positions of TRiC-bound and aggregated proteins are indicated. (*B*) Quantification of TRiC-bound and aggregated hSnu114 [581–972] from reactions minus apyrase in *A* by densitometry. TRiC-bound hSnu114 [581–972] in the absence of MBP-hSnu(830–972) competitor and aggregated hSnu114 [581–972] in the presence of the highest concentration of competitor were set to 1. Error bars represent SD values from three independent experiments.
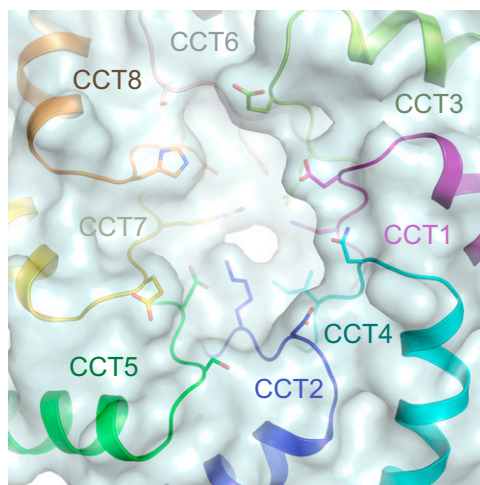
Fig. S7. The apical pore in the crystal structure of the closed TRiC complex. The peptide backbone of TRiC from *S. cerevisiae* [Protein Data Bank code 4D8Q (1)] is shown in ribbon representation enveloped by the molecular surface. Side chains in the pore are shown in stick representation. Note that several flexible side chains were disordered in the crystal and are thus shown in an arbitrary conformation. The lumen in the closed crystal structure would be wide enough for the peptide backbone of a flexible linker region but is likely to display substantial structural flexibility.

1. Leitner A, et al. (2012) The molecular architecture of the eukaryotic chaperonin TRiC/CCT. *Structure* 20(5):814–825.

Table S1.   Putative multidomain substrate proteins of TRiC (1, 2)

Gene name	Protein name	Construct	MW (kDa)
EEF2	eEF2	Full-length	95
TSHZ3	Teashirt 3 (Tsh3)	Full-length	99
EFTUD2	snRNP116/hSnu114	Full-length	109
MMS19	MET18	Full-length	113
UBA1	ubiquitin-activating enzyme E1	Full-length	118
LARS	Leucyl tRNA synthetase	Full-length	135
SKIV2L	Superkiller viralicidic activity 2-like protein	Full-length	138
KIF13A	Kinesin member 13A	Full-length	199
MYH1	Myosin	Full-length	223
FASN	Fatty acid synthase	Full-length	273
THNSL1	Aro1	Full-length	83
USP10	Ubp3	Full-length	87
ATP6V0A1	Stv1	Full-length	96
COPB2	Sec27	Full-length	102
GANAB	Rot2	ΔN (1–97)	96
SMARCC2	Swi3	Full-length	133
NUP153	Nup1	Full-length	154
MUC17	Sla1	ΔN (1–3270)	128
DYNC1H1	Dyn1	ΔN (1–3658)	112

1. Dekker C, et al. (2008) The interaction network of the chaperonin CCT. *EMBO J* 27(13):1827–1839.
2. Yam AY, et al. (2008) Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies. *Nat Struct Mol Biol* 15(12): 1255–1262.

4.3 Paper III: Chaperone function of Hgh1 in the biogenesis of Eukaryotic Elongation Factor 2 (submitted for publication)

Chaperone Function of Hgh1 in the Biogenesis of Eukaryotic Elongation Factor 2

Leonie Mönkemeyer, Courtney L. Klaips, David Balchin, Roman Körner, F. Ulrich Hartl*
and Andreas Bracher*#

Department of Cellular Biochemistry, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

Lead Contact

* Correspondence to bracher@biochem.mpg.de and uhartl@biochem.mpg.de

SUMMARY

Eukaryotic elongation factor 2 (eEF2) is an abundant and essential component of the translation machinery. The biogenesis of this 93 kDa multi-domain protein is assisted by the chaperonin TRiC/CCT. Here we show in yeast cells that the highly conserved protein Hgh1 (FAM203 in humans) is a chaperone that cooperates with TRiC in eEF2 folding. In the absence of Hgh1, a substantial fraction of newly-synthesized Eft, the yeast ortholog of eEF2, is degraded or aggregates. We solved the crystal structure of Hgh1 and analyzed the interaction of wild-type and mutant Hgh1 with Eft. These experiments revealed that Hgh1 is an armadillo repeat protein that binds to the dynamic central domain III of Eft via a bipartite interface. Hgh1 binding recruits TRiC to the C-terminal Eft module and prevents unproductive interactions of domain III, allowing efficient folding of the N-terminal GTPase module. Eft folding is completed upon dissociation of TRiC and Hgh1.

KEYWORDS

HIGHLIGHTS

- Eukaryotic elongation factor 2 (eEF2) is a chaperone-dependent multi-domain protein.

- Hgh1 functions as a chaperone for eEF2 and cooperates with the chaperonin TRiC.

- The crystal structure shows that Hgh1 is a highly conserved armadillo repeat protein.

- Hgh1 binds to domain III of eEF2 and recruits TRiC for eEF2 folding.

INTRODUCTION

Eukaryotic cells invest extensively in their protein translation machinery. Ribosomal proteins and eukaryotic elongation factors 1 and 2 (eEF1 and eEF2) account for ~13, 1.8 and 1.5% of total protein mass in *S. cerevisiae*, respectively (Liebermeister et al., 2014). eEF2, the eukaryotic ortholog of bacterial elongation factor EF-G, is a highly conserved GTPase that controls the translocation of mRNA during translation via GTP hydrolysis (Dever and Green, 2012). For each elongation cycle of the growing nascent chain, one eEF2 molecule binds to the ribosome and hydrolyzes one GTP. eEF2 is a protein of 93 kDa consisting of six domains (Figure 1A). The consecutive N-terminal GTPase domains G, G' and II form a rigid block, while the C-terminal domains III–V undergo conformational rearrangements during the functional cycle (Jørgensen et al., 2003; Murray et al., 2016; Voorhees et al., 2014). At the apex of domain IV, eEF2 is post-translationally modified with a diphthamide group (Figure 1A). This modification is also found in the archaeal elongation factor (Zhang et al., 2010); it serves to improve translation fidelity, but is not essential under normal growth conditions (Murray et al., 2016).

Folding of eEF2 is assisted by the eukaryotic cytosolic chaperonin TRiC/CCT (for TCP-1 ring complex/chaperonin containing TCP-1) (Dekker et al., 2008; Yam et al., 2008). TRiC, a hexadecameric double-ring complex of ~1 MDa with ATPase activity, assists in the folding and maturation of ~10% of cytosolic proteins (Thulasiraman et al., 1999), including the abundant cytoskeletal proteins actin and tubulin. TRiC substrates fold upon transient encapsulation in the chaperonin cavity, thereby avoiding aggregation of folding intermediates (Lopez et al., 2015). Proteomic analysis and genomic approaches identified 136 TRiC substrates in yeast (Dekker et al., 2008), with the eEF2 ortholog, Eft, being the TRiC client of highest cellular abundance (Kulak et al., 2014). Insights into the complex folding mechanism of eEF2 may be obtained from

the recent analysis of a close structural homolog, the splicesosome subunit hSnu114 (37% sequence identity with eEF2). Like eEF2, hSnu114 (109 kDa) exceeds the size limit of the TRiC chamber (~70 kDa), and folding was shown to occur by domain-wise encapsulation (Rüßmann et al., 2012). While an N-terminal fragment of hSnu114 comprising domains G, G' and II folded independently of chaperonin, the C-terminal segment (domains III, IV and V) required encapsulation by TRiC, but reached stable structure only upon interaction with the N-terminal region in the context of the full-length protein (Rüßmann et al., 2012).

The activity of TRiC in protein folding is supported by co-chaperones, which associate with the chaperonin. These include the hexameric prefoldin complex (GimC in yeast) and the phosducin-like proteins (PhLPs), which were proposed to aid the folding and maturation of the cytoskeletal proteins actin and tubulin, and the G-protein β-subunit, respectively (Lukov et al., 2005; McCormack et al., 2009; McLaughlin et al., 2002; Siegers et al., 1999; Stirling et al., 2006; Stirling et al., 2007). In addition, the human protein FAM203A/B was recently identified as a potential TRiC cofactor, based on its high abundance in pulldown fractions of TRiC (Hein et al., 2015). Homologs of the 45 kDa FAM203A/B are present throughout the eukaryotic tree, suggesting a conserved function. The yeast homolog, Hgh1, is a protein of intermediate overall abundance and was shown to interact with Eft and the TRiC subunit Cct6 in a proteomic screen (Gavin et al., 2006). The interaction between Hgh1 and Eft was confirmed in an independent proteome-wide study (Krogan et al., 2006). Notably, in yeast the essential Eft protein is encoded by two genes, *EFT1* and *EFT2*, and deletion of either gene results in reduced Eft levels without an apparent growth defect (Perentesis et al., 1992; Veldman et al., 1994). While deletion of *HGH1* does not impair growth (Rodriguez-Pena et al., 1998), deletion of both *HGH1* and *EFT2* causes a synthetic growth defect, suggesting a functional relation between Eft and Hgh1

(Costanzo et al., 2010). Deletion of *HGH1* triggers a mild heat shock response, indicative of increased protein-folding stress in the cytosol (Alford and Brandman, 2018; Brandman et al., 2012). Moreover, *hgh1Δ* cells are sensitive to the Hsp90 inhibitor Macbecin (McClellan et al., 2007), and the combined deletion of *HGH1* and components of the Hsp90 machinery, such as *CPR7*, *HCH1*, *HSC82, HSP82* and *STI1*, causes a synthetic growth defect (Costanzo et al., 2010; Kuzmin et al., 2018; McClellan et al., 2007). Hgh1 interacts directly with the essential Hsp90 cochaperone Cns1 (Gavin et al., 2006; Schlecht et al., 2012; Tarassov et al., 2008). Thus, Hgh1 appears to have an integral function in the cellular chaperone network (Rizzolo et al., 2017; Rizzolo et al., 2018).

Here we show that Hgh1 serves as a chaperone in Eft folding. In its absence, a substantial fraction of newly-synthesized Eft misfolds and is either degraded or aggregates. Hgh1, Eft and TRiC form a ternary complex. Binding of Hgh1 to the structurally dynamic central domain III of Eft facilitates the interaction between Eft and TRiC. The crystal structure of Hgh1 revealed an armadillo repeat fold with conserved surface areas close to the N-terminus and at the concave face of the solenoid. Mutation of both sites abolished the interaction of Hgh1 with Eft, and expression of mutant Hgh1 failed to complement the slow-growth phenotype of the *eft2Δhgh1Δ* double deletion strain. Together our results suggest that binding of Hgh1 masks domain III of Eft, avoiding the formation of misfolded species, and recruits TRiC to the C-terminal Eft module.

RESULTS

Hgh1 is required for efficient Eft folding

Deletion of the *HGH1* gene in *S. cerevisiae* did not result in a growth defect (Figure 1B), consistent with previous findings (Rodriguez-Pena et al., 1998). To explore the genetic relationship of *HGH1* and *EFT2* in more detail, we deleted both genes simultaneously. Unlike the single-deletion strains, *eft2Δhgh1Δ* cells exhibited a diminished growth rate (Costanzo et al., 2010). This effect was particularly clear at the lower growth temperature of 20°C (Figure 1B). Deletion of *HGH1* caused a ~35% reduction in the level of Eft protein compared to wild-type (WT) cells (Figure 1C). While Eft was reduced by ~60% in *eft2Δ* cells (Figure 1C) without causing a discernible growth defect, *eft2Δhgh1Δ* cells contained only ~25% of Eft relative to WT (Figure 1C), a level at which Eft is apparently limiting for growth (Figure 1B). Consistent with a role of Hsp90 in Eft folding, inhibition of Hsp90 with the specific inhibitor Macbecin reduced Eft to 62% in WT and to 26% in *hgh1Δ* cells (Figure 1D). Thus, the reported growth defect of *hgh1Δ* cells upon Hsp90 inhibition (McClellan et al., 2007) may be caused by insufficient Eft protein (Figure 1D).

Fractionation of *hgh1Δ* cells showed that a substantial fraction of Eft was insoluble, indicative of misfolding and aggregation (Figure 1E). Thus, the decrease of total Eft in *hgh1Δ* cells might be due to degradation of misfolded protein. Preexistent Eft was stable for hours upon cycloheximide (CHX) shutoff, in line with earlier reports (Belle et al., 2006; Christiano et al., 2014) (Figure 1F). Stability was preserved in *hgh1Δ* cells (Figure 1F), suggesting that Hgh1 is required for efficient Eft folding upon synthesis, rather than for conformational maintenance of Eft that has already folded.

eEF2 is a large multi-domain protein that is likely to fold in a domain-wise fashion. It consists of two structurally independent modules, an N-terminal GTPase module, residues 1–482, comprising domains G, G' and II, and a C-terminal module of residues 483–842, comprising domains III, IV and V (Figure 1A). We overexpressed the corresponding fragments (Eft-N and Eft-III+C, respectively) with C-terminal HA-tags, as well as a construct consisting of residues 566–842 (Eft-C), excluding domain III (Figure 2A). Eft-N was produced mostly in soluble form (Figure 2B). In contrast, Eft-III+C and Eft-C were insoluble (Figure 2B). Thus, Eft-III+C is unable to fold independently, as observed previously for the corresponding fragment of the Eft homolog hSnu114 (Rüßmann et al., 2012), suggesting overall similar folding properties. Absence of Hgh1 in *hgh1Δ* cells had little influence on the solubility of the expressed Eft truncation mutants, whereas full-length Eft-HA was mostly insoluble in *hgh1Δ* cells (Figure 2C). Hgh1 may prevent misfolding of full-length Eft by blocking unproductive inter-domain interactions during folding.

Eft was previously shown to be a substrate of the TRiC chaperonin (Dekker et al., 2008; Yam et al., 2008). To determine which part of Eft interacts with TRiC, we expressed the Eft truncation constructs either in WT yeast or in a strain in which TRiC subunit 2 (Cct2) contains an internal calmodulin-binding affinity tag (*cct2*-Int) (Pappenberger et al., 2006). Calmodulin affinity pulldowns and immunoblotting revealed a specific association of the folding-incompetent Eft-III+C and Eft-C constructs with TRiC (Figure 2D, lanes 7 and 8), reminiscent of the interaction pattern of C-terminal fragments of hSnu114 with TRiC (Rüßmann et al., 2012). Eft-N and full-length Eft were non-specifically precipitated and were present in the TRiC pulldown at levels similar to the background control (Figure 2D, lanes 1, 2, 5 and 6).

Hgh1, Eft and TRiC from a ternary complex

To determine whether Hgh1, Eft and TRiC form a ternary complex, we expressed Hgh1-FLAG and Eft2-HA under control of their natural promoters and analyzed their interactors using immunoprecipitation and quantitative label-free mass spectrometry (MS). Hgh1 pulldown with anti-FLAG coprecipitated TRiC and Eft, as shown by immunoblotting (Figure 3A). Mass spectrometric analysis, with anti-MYC pulldown as control, demonstrated a highly significant enrichment of all eight TRiC subunits in the anti-Hgh1 precipitate (Figure 3B). Eft had an intensity similar to the TRiC subunits. Hgh1 thus interacts, directly or indirectly, with Eft and TRiC. Besides the eukaryotic chaperonin, other cytosolic chaperones, including Hsp70 (Ssa1) and Hsp40 (Ydj1) as well as the AAA+ proteins Rvb1 and Rvb2 were enriched in the Hgh1 pulldown. Ribosomal proteins were also enriched in the Hgh1 pulldown, while Hsp90 was below the significance threshold. In contrast, pulldown fractions of Eft showed no significant enrichment of TRiC and Eft, suggesting that only the small fraction of newly-synthesized Eft interacts with these chaperones (Figure S1).

To determine whether Hgh1 interacts with Eft directly, we purified Hgh1-FLAG upon expression in *E. coli* and Eft2-HA upon expression in *S. cerevisiae*. The proteins did not stably interact in co-precipitation experiments. However, incubation with the lysine-reactive chemical crosslinker disuccinimidyl suberate (DSS) revealed a crosslink product of ~180 kDa on SDS-PAGE, which corresponded to covalently linked Hgh1-FLAG and Eft2-HA, as shown by immunoblotting (indicated by arrowheads in Figure 4A). This result is consistent with Hgh1 interacting with a small fraction of non-native Eft that is in equilibrium with the native protein under the in vitro conditions. Indeed, purified Hgh1-FLAG, when used at concentrations exceeding endogenous Hgh1, also co-precipitated Eft from cell lysate (Figure 4A, lane 7).

Interestingly, the co-precipitated material also exhibited the characteristic band pattern of TRiC, suggesting the presence of a ternary complex of Hgh1, Eft and TRiC (Figure 4A, lane 7; compare Figure 3A, left). DSS crosslinking of this material produced an additional anti-FLAG positive band of ~120 kDa (indicated by asterisk in Figure 4A, lane 16), presumably corresponding to crosslink products of Hgh1-FLAG with TRiC subunits.

We further analyzed the interactions between purified Hgh1, Eft and TRiC by native gel electrophoresis (native-PAGE). TRiC migrates on native-PAGE as a high molecular weight complex, as detected by immunoblotting with anti-Tcp1 antibody (Figure 4B) (Leitner et al., 2012). Hgh1-FLAG or Eft2-HA alone did not detectably interact with TRiC (Figure 4B, lanes 6 and 11). However, when all three proteins were present, both Hgh1-FLAG and Eft-HA co-migrated with TRiC (Figure 4B, lanes 8 and 12), indicating formation of a ternary complex. Thus, Hgh1 mediates the interaction of Eft with TRiC.

Hgh1 binds a partially unfolded conformation of domain III of Eft

To probe the binding site(s) for Hgh1 on Eft, we first investigated the influence of known Eft ligands on complex formation. The formation of crosslinking products between purified Hgh1 and Eft was used as a readout. Binding of guanine nucleotide to domain G (Figure 1A) is known to influence the conformation of eEF2 switch region I (amino acids 25–73), which is partially disordered in available crystal structures (Bartish and Nygård, 2008). Addition of GDP, GTP or the non-hydrolysable analog GMP-PNP was of minimal effect on Eft–Hgh1 complex formation (Figure 5A), suggesting that Hgh1 does not interact with the regulatory elements of the nucleotide binding pocket in domain G. In contrast, complex formation was abolished in the presence of the fungal antibiotic Sordarin, which binds to the interface between domains III and

V (Figure 1A, lanes 5 and 10). This suggested that Hgh1 interacts with Eft at or near domains III and V.

Next, we analyzed the molecular interface between Hgh1 and Eft using Eft truncation constructs. Truncations containing the C-terminal module (domains IV and V) were insoluble (Figure 2B). However, the N-terminal module of Eft (Eft-N; residues 1–482) could be purified in soluble form upon expression in yeast cells (Figure 2B). As judged from the circular dichroism (CD) spectrum, Eft-N is natively folded (Figure S2A). Of note, soluble expression of Eft-N in *E. coli* failed, suggesting that folding of this module depends on eukaryotic chaperones. The fragment corresponding to domain III (Eft-III; residues 483–565) was expressed in *E. coli* and purified from inclusion bodies. His$_6$-tagged Eft-III was soluble upon dilution from denaturant, but was not natively folded (Figure S2B), presumably due to the presence of hydrophobic segments that are shielded upon interface formation with the other domains (Figure S2C). Eft-N did not interact with Hgh1, as judged by the crosslinking assay (Figure 5B, lane 13). However, we identified a clear crosslinking product between Hgh1 and Eft-III (Figure 5B, lane 14), blue arrowhead). Together with the finding that Sordarin prevented binding of Hgh1 to Eft, these results point to domain III as a major binding site of Hgh1.

To understand how Hgh1 interacts with Eft, we analyzed the structural dynamics of Eft using hydrogen/deuterium exchange combined with mass spectrometry (H/DX). Backbone amide hydrogens are protected from exchange when involved in stable secondary structure, buried in the core of a folded protein or at a protein-protein interface (Engen and Smith, 2001; Wales and Engen, 2006). The degree of deuterium incorporation therefore correlates with structural flexibility. H/DX measurements were performed for Eft in the absence and presence of Sordarin or Hgh1 (Figure S3A). Sequence coverage was near-complete for all conditions.

Analysis of unbound Eft showed that most regions of Eft that are structured in the crystal lattice (PDB 1N0V) exchange deuterium slowly with the solvent (blue color in Figure 5C). However, we observed increased rates of H/D exchange for peptides mapping to domain III of Eft (red color in Figure 5C), indicating greater structural dynamics in this region than anticipated from the available structural data (Jørgensen et al., 2003). In the presence of Sordarin, which binds to the interface between domains III and V, deuterium incorporation into domain III was greatly diminished, consistent with structural stabilization of this domain (Figure 5D). Since Sordarin inhibits binding of Hgh1 (Figure 5A, lane 10), these data suggest that Hgh1 selectively recognizes a structurally dynamic conformation of domain III of Eft.

Addition of Hgh1 in a 2-fold molar excess over Eft resulted in reduced deuterium incorporation into peptides 522–538 and 536–540, located in domain III of Eft (Figures 5E, F and S3B). This is consistent with a direct interaction between Hgh1 and domain III of Eft, and supports the crosslinking data (Figure 5B). We also observed significant protection of a short peptide in domain II comprising residues 433–438 (Figures 5E, F), suggesting that Hgh1 makes additional surface contacts in this region.

Taken together, these results indicate that Hgh1 recognizes a non-native conformation of domain III in Eft. This conformation would be present during biogenesis of Eft, but is only transiently populated by the native protein at equilibrium, explaining the inefficient interaction between native Eft, Hgh1 and TRiC in vitro (Figure 4).

Structure and mutational analysis of Hgh1

To gain insight into the mechanism of Hgh1 function, we determined the crystal structure of Hgh1. Limited proteolysis with proteinase-K (Figure S4A), followed by mass spectrometry,

showed that the 31 C-terminal residues of the 394 residue protein were protease-sensitive and presumably unstructured, in agreement with their low sequence complexity and high negative charge (20 Asp/Glu residues). We obtained two crystal forms for Hgh1(1–363), which diffracted to 2.33 and 3.0 Å resolution, respectively, and solved the phase problem by Se–SAD at 2.7 Å resolution (Table S1). The experimental electron density map was readily interpretable (Figure S4B). Both crystal forms contain four independent copies of Hgh1 (Figure S4CD), which are conformationally very similar, with root mean square deviations of 0.06–0.53 Å for the matching Cα positions 4–357 (Figure S4EF). Hgh1 has a curved tubular shape with dimensions of 40 x 50 x 80 Å (Figure 6AB). The α-helical solenoid protein comprises four imperfect armadillo repeats, followed by two non-canonical 3-helix repeats. A pair of helices caps the N-terminal end of the solenoid. The helix repeats contain three substantial insertions layered onto the convex surface of the solenoid (Figure 6A, top). Insertion I, residues 131–156, containing a short β-hairpin, follows on helix H9 in armadillo repeat 3; insertion II, residues 203–209, is situated after helix H12 in armadillo repeat 4; and insertion III, residues 266–294, is placed after helix H16 and consists of two short helices. The contact areas between the insertions and the solenoids are largely hydrophobic, and thus these regions appear to be stably structured.

Sequence homologs of Hgh1 exist throughout the eukaryotic tree (Figure S5). Two surface areas are conserved in all homologs, suggestive of protein-protein interaction interfaces (Figure 6B). One of these regions is located near the N-terminus forming the groove between helices H1 and H2; it comprises Glu7 and the hydrophobic residues Leu8, Phe11, Val19 and Ala23. The second conserved area is also mainly hydrophobic and extends along the concave face of the solenoid, with residues Arg197, Lys236, Asn237 and Phe240 forming a highly conserved cluster (Figure S4GH). In animal and fungal sequences, an additional area of high

surface conservation is found close to the C-terminus at the groove between helices H19 and H20, including residues Tyr326, Arg330, His333 and Val349.

To functionally validate these putative contact regions, we introduced the triple-mutations E7S/L8A/F11A (MutN), K236A/N237H/F240A (MutM) or Y326A/R330A/H333A (MutC) and purified the mutant proteins as soluble C-terminally FLAG-tagged proteins (Figure 6C). In contrast to WT Hgh1, Hgh1-MutN did not form a detectable crosslinking product with Eft in vitro. Hgh1-MutM crosslinked less efficiently than WT. In contrast, Hgh1-MutC interacted like the WT protein (Figure 6D). As expected, the interaction with Eft was also abolished with the mutant Hgh1-MutN+M, combining the mutations in MutN and MutM (Figure 6D). Note that Hgh1-MutN+M is stably folded at physiological temperature, as indicated by CD measurements (Figure S6AB). These results indicate that the spatially separated N-terminal and middle regions of Hgh1 both participate in binding to Eft sites. Pulldown of Hgh1-FLAG from cell extracts showed further that the interaction with TRiC correlated with the ability of the Hgh1 variants to interact with Eft. Essentially no interaction with TRiC or Eft was detected with Hgh1-MutN, and Hgh1-MutM showed reduced affinity for both TRiC and Eft. Mutation of site C did not affect the interaction between Hgh1 and TRiC or Eft2 (Figure 6E). This confirms that synergistic interactions between all three proteins underlie the formation of the ternary Eft-Hgh1-TRiC complex.

Next, we asked whether the physical interaction between Hgh1 and Eft is required for biological function in vivo. WT and mutant Hgh1 proteins were expressed with C-terminal HA-tags in the *eft2Δhgh1Δ* strain, and cell growth monitored (Figure 6F; Figure S6C). Expression of *hgh1-MutN* partially suppressed the growth defect of *eft2Δhgh1Δ* cells, while expression of *hgh1-MutM* and *hgh1-MutC* restored normal growth (Figure 6F). In contrast, *Hgh1*-MutN+M

failed to complement the growth defect (Figure 6F). These results demonstrate that the interactions between Hgh1 and Eft observed by biochemical analysis in vitro are critical in vivo. The ability of Hgh1 to interact with non-native Eft is thus needed to restore concentrations of functional Eft to levels sufficient for normal growth.

DISCUSSION

Eukaryotic translation elongation factor 2, eEF2, is a highly abundant, 93 kDa protein with an intricate six-domain structure. Our data demonstrate that the conserved protein Hgh1 serves as a chaperone in the folding of the yeast eEF2 ortholog, Eft. Hgh1 binds mainly to the structurally dynamic domain III of Eft and appears to prevent aberrant intramolecular interactions that would otherwise slow the folding of the flanking domains. Additionally, Hgh1 recruits the chaperonin TRiC to the C-terminal Eft module. In the absence of Hgh1, a large fraction of Eft misfolds, resulting in aggregation or degradation (Figure 7).

The rather hydrophobic but structurally dynamic domain III of Eft is located between the N-terminal GTPase module and the C-terminal module of Eft that binds the ribosomal decoding-center. Structural flexibility of this domain is probably functionally relevant, as the bacterial ortholog EF-G undergoes substantial conformational changes during polypeptide elongation on the ribosome (Lin et al., 2015). Consistent with this, domain III is also the binding site for the translocation inhibitors Fusidic acid and Sordarin. While functionally important, the flexibility of domain III is likely to interfere with the correct folding of the adjacent N- and C-modules, which are large (480 and 277 amino acids, respectively) and have complex folds that are stabilized by long-range interactions (Figure 1A).

We propose that Hgh1 interacts with domain III during translation of Eft (Figure 7, step 1), consistent with the enrichment of ribosomal proteins in the Hgh1 pulldown fraction (Figure 3B). The armadillo repeat structure of Hgh1 appears to be well suited for binding of extended hydrophobic substrate sequences at its concave face, similar to many other armadillo repeat proteins (Reichen et al., 2014). Hgh1 has an additional hydrophobic groove close to its N-terminus, suggesting that discontinuous binding motifs in domain III may be recognized. This is in agreement with our H/DX data, showing protection by Hgh1 of residues 522–540. Indeed, in an extended conformation, this segment would be long enough to connect the binding sites in Hgh1. Binding of Hgh1 may fulfill a dual role in Eft folding: It prevents domain III from interfering with the folding of the N-module (Figure 7, step 1) and recruits TRiC for the folding of the C-terminal module (Figure 7, step 2). TRiC may recognize elements in domains III–V of Eft, which become exposed in the complex with Hgh1, in analogy to binding of C-terminal domains of hSnu114 by the chaperonin (Rüßmann et al., 2012). Our earlier experiments with this structural homolog of Eft revealed the encapsulation by TRiC of C-terminal fragments up to 37 kDa (Rüßmann et al., 2012). Folding of the N-module of Eft (domains G, G' and II), must be completed first, because the C-terminal domains III–V cannot fold stably in isolation. Folding of the N-module might be facilitated by the Hsp70 components Ssa1 and Ydj1, which were found to associate with Hgh1 in vivo. A further involvement of the Hsp90 system in N-module folding might explain the reduced Eft levels in presence of the Hsp90 inhibitor Macbecin. The reported interaction of Hgh1 with Cns1 may serve to recruit Hsp90 to the N-domain ((Gavin et al., 2006; Schlecht et al., 2012; Tarassov et al., 2008) and accompanying manuscript by Schopf et al.). TRiC-assisted folding of the C-module results in release of chaperonin (Figure 7, step 3) and structure formation in domain III finally induces Hgh1 dissociation, generating native Eft (Figure

7, step 4). While it is an intriguing possibility that Hgh1, by binding to domain III, coordinates the action of Hsp90 and TRiC in Eft folding, the two chaperone systems may also provide alternative and partially redundant chaperone pathways.

In the absence of Hgh1, Eft folding becomes inefficient, leading to the build-up of folding intermediates, which are either cleared by degradation or aggregate. Because of the sheer mass of new Eft chains that are produced per cell duplication, these species are expected to engage a substantial fraction of the cellular chaperone arsenal. This extra burden on the cytosolic proteostasis machinery may explain the heat shock response observed in *hgh1Δ* cells (Brandman et al., 2012). In the absence of Hgh1, the load on the Hsp90 machinery increases, as evidenced by the sensitivity of *hgh1Δ* cells to the Hsp90 inhibitor Macbecin (McClellan et al., 2007) and the synthetic growth defects upon combined loss of Hgh1 and components of the Hsp90 machinery (Costanzo et al., 2010). *HGH1* also shows strong genetic interactions with genes encoding subunits Cog3 and Cog5–Cog8 of the Conserved Oligomeric Golgi (COG) complex (Costanzo et al., 2010), a vesicle tethering complex that requires Hsp90 for assembly (McClellan et al., 2007). Thus, the functional cooperation of Hgh1 and the Hsp90 system might extend beyond Eft biosynthesis.

Whether FAM203 plays a role similar to Hgh1 in mammalian cells, remains to be confirmed. Interestingly, Hgh1/FAM203 homologs appear restricted to eukarya. It seems possible that eubacteria and archaea express proteins functionally equivalent to Hgh1 for the efficient folding of Eft orthologs EF-G and EF-2.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

A.B. conceived the project. L.M. and A.B. planned, performed and analyzed the biochemical experiments with input from F.U.H.. C.L.K. prepared yeast strains. R.K. conducted the MS experiments and processed the proteomics data. D.B. performed the H/DX experiments and analyzed the data. L.M. and A.B. solved the crystal structures and interpreted the data. A.B. and F.U.H. participated in data interpretation, and wrote the paper with contributions from other authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*, 213-221.

Alford, B.D., and Brandman, O. (2018). Quantification of Hsp90 availability reveals differential coupling to the heat shock response. J Cell Biol.

Bartish, G., and Nygård, O. (2008). Importance of individual amino acids in the Switch I region in eEF2 studied by functional complementation in S. cerevisiae. Biochimie *90*, 736-748.

Belle, A., Tanay, A., Bitincka, L., Shamir, R., and O'Shea, E.K. (2006). Quantification of protein half-lives in the budding yeast proteome. Proc Natl Acad Sci U S A *103*, 13004-13009.

Brandman, O., Stewart-Ornstein, J., Wong, D., Larson, A., Williams, C.C., Li, G.W., Zhou, S., King, D., Shen, P.S., Weibezahn, J.*, et al.* (2012). A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. Cell *151*, 1042-1054.

Brockhauser, S., Ravelli, R.B., and McCarthy, A.A. (2013). The use of a mini-kappa goniometer head in macromolecular crystallography diffraction experiments. Acta Crystallogr D Biol Crystallogr *69*, 1241-1251.

Christiano, R., Nagaraj, N., Frohlich, F., and Walther, T.C. (2014). Global proteome turnover analyses of the Yeasts S. cerevisiae and S. pombe. Cell Rep *9*, 1959-1965.

Collaborative Computational Project, N. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr *50*, 760-763.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L., Toufighi, K., Mostafavi, S., *et al.* (2010). The genetic landscape of a cell. Science *327*, 425-431.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol *26*, 1367-1372.

Dekker, C., Stirling, P.C., McCormack, E.A., Filmore, H., Paul, A., Brost, R.L., Costanzo, M., Boone, C., Leroux, M.R., and Willison, K.R. (2008). The interaction network of the chaperonin CCT. EMBO J *27*, 1827-1839.

Dever, T.E., and Green, R. (2012). The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harb Perspect Biol *4*, a013706.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.

Engen, J.R., and Smith, D.L. (2001). Investigating protein structure and dynamics by hydrogen exchange MS. Anal Chem *73*, 256A-265A.

Evans, P. (2006). Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr *62*, 72-82.

Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr *69*, 1204-1214.

French, G., and Wilson, K. (1978). On the treatment of negative intensity observations. Acta Cryst Sect A *34*, 517-525.

Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dumpelfeld, B.*, et al.* (2006). Proteome survey reveals modularity of the yeast cell machinery. Nature *440*, 631-636.

Gouet, P., Courcelle, E., Stuart, D.I., and Metoz, F. (1999). ESPript: multiple sequence alignments in PostScript. Bioinformatics *15*, 305-308.

Guttman, M., Wales, T.E., Whittington, D., Engen, J.R., Brown, J.M., and Lee, K.K. (2016). Tuning a High Transmission Ion Guide to Prevent Gas-Phase Proton Exchange During H/D Exchange MS Analysis. J Am Soc Mass Spectrom *27*, 662-668.

Hein, M.Y., Hubner, N.C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I.A., Weisswange, I., Mansfeld, J., Buchholz, F.*, et al.* (2015). A human interactome in three quantitative dimensions organized by stoichiometries and abundances. Cell *163*, 712-723.

Iacob, R.E., Murphy, J.P., 3rd, and Engen, J.R. (2008). Ion mobility adds an additional dimension to mass spectrometric analysis of solution-phase hydrogen/deuterium exchange. Rapid Commun Mass Spectrom *22*, 2898-2904.

Jørgensen, R., Ortiz, P.A., Carr-Schmid, A., Nissen, P., Kinzy, T.G., and Andersen, G.R. (2003). Two crystal structures demonstrate large conformational changes in the eukaryotic ribosomal translocase. Nat Struct Biol *10*, 379-385.

Kabsch, W. (2010). XDS. Acta Crystallogr D Biol Crystallogr *66*, 125-132.

Kleywegt, G.T., and Jones, T.A. (1994). A super position. CCP4/ESF-EACBM Newsletter on Protein Crystallography *31*, 9-14.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P.*, et al.* (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637-643.

Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nat Methods *11*, 319-324.

Kuzmin, E., VanderSluis, B., Wang, W., Tan, G., Deshpande, R., Chen, Y., Usaj, M., Balint, A., Mattiazzi Usaj, M., van Leeuwen, J.*, et al.* (2018). Systematic analysis of complex genetic interactions. Science *360*.

Leitner, A., Joachimiak, L.A., Bracher, A., Monkemeyer, L., Walzthoeni, T., Chen, B., Pechmann, S., Holmes, S., Cong, Y., Ma, B.*, et al.* (2012). The molecular architecture of the eukaryotic chaperonin TRiC/CCT. Structure *20*, 814-825.

Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of protein investment in cellular functions. Proc Natl Acad Sci U S A *111*, 8488-8493.

Lin, J., Gagnon, M.G., Bulkley, D., and Steitz, T.A. (2015). Conformational changes of elongation factor G on the ribosome during tRNA translocation. Cell *160*, 219-227.

Lopez, T., Dalton, K., and Frydman, J. (2015). The Mechanism and Function of Group II Chaperonins. J Mol Biol *427*, 2919-2930.

Lukov, G.L., Hu, T., McLaughlin, J.N., Hamm, H.E., and Willardson, B.M. (2005). Phosducin-like protein acts as a molecular chaperone for G protein betagamma dimer assembly. EMBO J *24*, 1965-1975.

McClellan, A.J., Xia, Y., Deutschbauer, A.M., Davis, R.W., Gerstein, M., and Frydman, J. (2007). Diverse cellular functions of the Hsp90 molecular chaperone uncovered using systems approaches. Cell *131*, 121-135.

McCormack, E.A., Altschuler, G.M., Dekker, C., Filmore, H., and Willison, K.R. (2009). Yeast phosducin-like protein 2 acts as a stimulatory co-factor for the folding of actin by the chaperonin CCT via a ternary complex. J Mol Biol *391*, 192-206.

McLaughlin, J.N., Thulin, C.D., Hart, S.J., Resing, K.A., Ahn, N.G., and Willardson, B.M. (2002). Regulatory interaction of phosducin-like protein with the cytosolic chaperonin complex. Proc Natl Acad Sci U S A *99*, 7962-7967.

Mumberg, D., Müller, R., and Funk, M. (1995). Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. Gene *156*, 119-122.

Murray, J., Savva, C.G., Shin, B.S., Dever, T.E., Ramakrishnan, V., and Fernandez, I.S. (2016). Structural characterization of ribosome recruitment and translocation by type IV IRES. Elife *5*.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of Macromolecular Structures by the Maximum-Likelihood Method. Acta Crystallogr D Biol Crystallogr *53*, 240-255.

Pappenberger, G., McCormack, E.A., and Willison, K.R. (2006). Quantitative actin folding reactions using yeast CCT purified via an internal tag in the CCT3/gamma subunit. J Mol Biol *360*, 484-496.

Perentesis, J.P., Phan, L.D., Gleason, W.B., LaPorte, D.C., Livingston, D.M., and Bodley, J.W. (1992). Saccharomyces cerevisiae elongation factor 2. Genetic cloning, characterization of expression, and G-domain modeling. J Biol Chem *267*, 1190-1197.

Rappsilber, J., Ishihama, Y., and Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Anal Chem *75*, 663-670.

Reichen, C., Hansen, S., and Plückthun, A. (2014). Modular peptide binding: from a comparison of natural binders to designed armadillo repeat proteins. J Struct Biol *185*, 147-162.

Rizzolo, K., Huen, J., Kumar, A., Phanse, S., Vlasblom, J., Kakihara, Y., Zeineddine, H.A., Minic, Z., Snider, J., Wang, W.*, et al.* (2017). Features of the Chaperone Cellular Network Revealed through Systematic Interaction Mapping. Cell Rep *20*, 2735-2748.

Rizzolo, K., Kumar, A., Kakihara, Y., Phanse, S., Minic, Z., Snider, J., Stagljar, I., Zilles, S., Babu, M., and Houry, W.A. (2018). Systems analysis of the genetic interaction network of yeast molecular chaperones. Mol Omics *14*, 82-94.

Rodriguez-Pena, J.M., Cid, V.J., Sanchez, M., Molina, M., Arroyo, J., and Nombela, C. (1998). The deletion of six ORFs of unknown function from Saccharomyces cerevisiae chromosome VII reveals two essential genes: YGR195w and YGR198w. Yeast *14*, 853-860.

Rüßmann, F., Stemp, M.J., Mönkemeyer, L., Etchells, S.A., Bracher, A., and Hartl, F.U. (2012). Folding of large multidomain proteins by partial encapsulation in the chaperonin TRiC/CCT. Proc Natl Acad Sci U S A *109*, 21208-21215.

Schlecht, U., Miranda, M., Suresh, S., Davis, R.W., and St Onge, R.P. (2012). Multiplex assay for condition-dependent changes in protein-protein interactions. Proc Natl Acad Sci U S A *109*, 9213-9218.

Siegers, K., Waldmann, T., Leroux, M.R., Grein, K., Shevchenko, A., Schiebel, E., and Hartl, F.U. (1999). Compartmentation of protein folding in vivo: sequestration of non-native polypeptide by the chaperonin-GimC system. EMBO J *18*, 75-84.

Sikorski, R.S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics *122*, 19-27.

Stirling, P.C., Cuellar, J., Alfaro, G.A., El Khadali, F., Beh, C.T., Valpuesta, J.M., Melki, R., and Leroux, M.R. (2006). PhLP3 modulates CCT-mediated actin and tubulin folding via ternary complexes with substrates. J Biol Chem *281*, 7012-7021.

Stirling, P.C., Srayko, M., Takhar, K.S., Pozniakovsky, A., Hyman, A.A., and Leroux, M.R. (2007). Functional interaction between phosducin-like protein 2 and cytosolic chaperonin is essential for cytoskeletal protein function and cell cycle progression. Mol Biol Cell *18*, 2336-2345.

Tarassov, K., Messier, V., Landry, C.R., Radinovic, S., Serna Molina, M.M., Shames, I., Malitskaya, Y., Vogel, J., Bussey, H., and Michnick, S.W. (2008). An in vivo map of the yeast protein interactome. Science *320*, 1465-1470.

Thulasiraman, V., Yang, C.F., and Frydman, J. (1999). In vivo newly translated polypeptides are sequestered in a protected folding environment. EMBO J *18*, 85-95.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods *13*, 731-740.

Vagin, A.A., and Isupov, M.N. (2001). Spherically averaged phased translation function and its application to the search for molecules and fragments in electron-density maps. Acta Crystallogr D Biol Crystallogr *57*, 1451-1456.

Veldman, S., Rao, S., and Bodley, J.W. (1994). Differential transcription of the two Saccharomyces cerevisiae genes encoding elongation factor 2. Gene *148*, 143-147.

Voorhees, R.M., Fernandez, I.S., Scheres, S.H., and Hegde, R.S. (2014). Structure of the mammalian ribosome-Sec61 complex to 3.4 A resolution. Cell *157*, 1632-1643.

Wach, A., Brachat, A., Pöhlmann, R., and Philippsen, P. (1994). New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. Yeast *10*, 1793-1808.

Wales, T.E., and Engen, J.R. (2006). Hydrogen exchange mass spectrometry for the analysis of protein dynamics. Mass Spectrom Rev *25*, 158-170.

Wisniewski, J.R., Zielinska, D.F., and Mann, M. (2011). Comparison of ultrafiltration units for proteomic and N-glycoproteomic analysis by the filter-aided sample preparation method. Anal Biochem *410*, 307-309.

Yam, A.Y., Xia, Y., Lin, H.T., Burlingame, A., Gerstein, M., and Frydman, J. (2008). Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies. Nat Struct Mol Biol *15*, 1255-1262.

Young, C.L., Raden, D.L., Caplan, J.L., Czymmek, K.J., and Robinson, A.S. (2012). Cassette series designed for live-cell imaging of proteins and high-resolution techniques in yeast. Yeast *29*, 119-136.

Zhang, Y., Zhu, X., Torelli, A.T., Lee, M., Dzikovski, B., Koralewski, R.M., Wang, E., Freed, J., Krebs, C., Ealick, S.E.*, et al.* (2010). Diphthamide biosynthesis requires an organic radical generated by an iron-sulphur enzyme. Nature *465*, 891-896.

FIGURE LEGENDS

Figure 1. Characterization of the *hgh1*Δ deletion strain

(A) Structural model and domain structure of eEF2. The structural model is a composite of the cryoEM structure of a ribosome-bound eEF2:GMP-PCP complex (pdb code 5IT7) (Murray et al., 2016) and the crystal structure of the *S. cerevisiae* eEF2 ortholog, Eft, in complex with the anti-fungal compound Sordarin (pdb code 1N0U) (Jørgensen et al., 2003). The ligands are shown in space-filling representation. The domain structure is color-coded and shown schematically below. The diphthamide group is shown in stick representation.

(B) Synthetic growth defect upon deletion of *HGH1* and *EFT2*. Dilution series of the parent WT strain BY4741 and *hgh1*Δ, *eft2*Δ and *eft2*Δ*hgh1*Δ cells were spotted onto YPD agar and cultivated for 48 h at 20°C.

(C) Loss of *HGH1* causes reduction in Eft levels. Protein levels of Eft were analyzed by immunoblotting in the yeast strains indicated. Cells were grown at 30°C. Phosphoglycerate kinase (Pgk1) served as a loading control. Eft levels were quantified by densitometry. Averages with standard deviations (SD) from three independent experiments are shown as % of WT control.

(D) Reduction of Eft levels upon Hsp90 inhibition. WT and *hgh1*Δ cells were grown for 3 h after addition of 40 μM Macbecin or DMSO alone, and Eft and Pgk1 levels analyzed by immunoblotting. Averages from three independent experiments with SD are shown.

(E) Partial insolubility of Eft in *hgh1*Δ cells. Log-phase WT and *hgh1*Δ cells were lysed by bead milling. Lysates were fractionated by centrifugation and Eft protein in total (T), soluble (S) and

pellet fraction (P) analyzed by immunoblotting. Bar graph shows the levels of insoluble Eft

relative to total. Averages from three independent experiments with SD are shown.

(F) Stability of mature Eft in absence of *HGH1*. The levels of Eft in WT and *hgh1*Δ cells were

monitored by immunoblotting at the indicated time points after inhibition of protein synthesis by

0.1 mM cycloheximide (CHX). Averages from three independent experiments with SD are

shown on the right.

Figure 2. Solubility of Eft truncation constructs and interaction with TRiC

(A) Schematic representations of Eft and Eft truncation constructs.

(B–C) Solubility of Eft truncation constructs in WT (B) and *hgh1*Δ cells (C). The indicated

constructs were expressed with a C-terminal HA-tag at 30°C under control of the *Gal* promoter.

Protein amounts in total (T), soluble (S) and pellet (P) fractions were analyzed by

immunoblotting with anti-HA antibody. Pgk1 served as loading control.

(D) Association of Eft truncation constructs with TRiC. Eft constructs were expressed either in a

strain harboring affinity-tagged TRiC (*cct2*–Int), or the parent WT strain. Soluble cell lysates

were subjected to affinity chromatography, followed by immunoblotting with anti-HA antibody.

Asterisk marks a cross-reactive band. Note that Eft-N binds non-specifically to the

immunobeads.

Figure 3. Analysis of Hgh1 interactors

(A) SDS-PAGE analysis of proteins interacting with Hgh1-FLAG. Cells expressing Hgh1-

FLAG and Eft2-HA under control of their natural promoters were grown to late log-phase

(OD_{600} ~2). Soluble lysate was passed over anti-FLAG or anti-HA affinity resin, with anti-MYC

resin serving as background control. The anti-FLAG eluate was analyzed by SDS-PAGE and Coomassie staining and immunoblotting with anti-FLAG, anti-HA or anti-TCP1 antibodies. Representative results are shown.

(B) Volcano plot representation of Hgh1-FLAG interactors. The components of the anti-FLAG immunoprecipitates from three independent experiments were analyzed by label-free quantification using anti-MYC as background control. The proteins to the left and right of the solid lines were significantly depleted or enriched in the FLAG precipitates, respectively. Dark blue color indicates the eight subunits of the TRiC complex. Other chaperones are shown in purple. Ribosomal proteins are marked in cyan. Hgh1 and Eft are shown in green and red, respectively. See also Figure S1.

Figure 4. Formation of a ternary Hgh1:Eft:TRiC complex

(A) Detection of a complex between Hgh1-FLAG and native Eft-HA by chemical crosslinking with DSS. A Coomassie blue-stained SDS-PAGE gel and a corresponding immunoblot probed with anti-FLAG and anti-HA antibodies are shown. Hgh1-FLAG and Eft-HA at 2 μM each were present when indicated (lanes 1-6, 9-14 and 17-22, respectively). Proteins captured from yeast lysate with a large excess of recombinant Hgh1-FLAG were analyzed in lanes 7, 8, 15, 16 and 23, 24. The Hgh1-FLAG:Eft-HA crosslink product is indicated by arrowheads. Asterisks mark putative crosslinking products with single TRiC subunits.

(B) Interactions of TRiC with Hgh1-FLAG and Eft-HA. Purified TRiC (2 μM hexadecamer) was mixed with Hgh1-FLAG and Eft-HA (each 2 μM) as indicated, and subsequently analyzed by native-PAGE and immunoblotting with anti-Tcp1, anti-FLAG and anti-HA antibodies. The position of the Hgh1-FLAG:Eft2-HA:TRiC complex is indicated by arrowheads.

Figure 5. Hgh1 binds the structurally dynamic domain III of Eft

(A) Effect of Eft ligands on the interaction with Hgh1. Equimolar mixtures of Hgh1 and Eft (10 µM each) in presence of GTP, GDP, GMP-PNP or Sordarin (2 mM each) were treated with DMSO (left) or DSS crosslinker (right). Reactions were analyzed by SDS-PAGE and Coomassie staining. Arrowhead indicates position of the crosslinking product. A representative gel is shown.

(B) Interaction of Hgh1 and Eft truncation constructs, Eft-N and Eft-III. The domain structure of Eft and Eft truncation constructs is indicated on top, together with a 3D-model of Eft in space-filling mode. Proteins (2 µM each) were treated with DMSO (left) or DSS (right). Coomassie blue-stained SDS-PAGE gels are shown. Arrowheads in pink and cyan indicate crosslinking products of Eft and Eft-III with Hgh1, respectively. See also Figure S2.

(C) Structural dynamics of Eft. Deuterium incorporation into Eft during 10-1000 s exposure to deuteration buffer was analyzed by H/DX, followed by pepsin digest and mass spectrometry of peptides. Fractional deuterium incorporation into Eft peptides is mapped onto the crystal structure of Eft (Jørgensen et al., 2003). Increasing deuteration is shown as a gradient from blue to red. Dark grey color indicates missing sequences. Data for 300 s exposure to deuterium are shown. See also Figure S3.

(D) Reduced structural dynamics of domain III in the Eft-Sordarin complex. Eft was incubated with 0.2 mM Sordarin and processed as above. Differential deuterium incorporation relative to Eft alone is mapped onto the structure. Blue and red colors indicate decreased and increased deuterium incorporation, respectively. Bound Sordarin is shown in purple. Data for 300s exposure to deuterium are shown.

(E) Altered structural dynamics of Eft domain III in presence of Hgh1. Differential deuterium incorporation of Eft (2 μM) in presence of Hgh1 (4 μM) relative to Eft alone is shown for 300 s of deuterium exposure. Blue and red colors indicate decreased and increased deuterium incorporation, respectively. Peptides protected by Hgh1 are indicated.

(F) Deuterium uptake kinetics for peptides 433-438 (domain II), 522-538 (domain III) and 536-540 (domain III), representing putative Hgh1 interaction sites. Time traces for Eft alone, the Eft-Sordarin complex and Eft/Hgh1 are shown in black, pink and green, respectively.

Figure 6. Structure and mutational analysis of Hgh1

(A) Ribbon representation of the Hgh1 crystal structure. Two orthogonal views are shown. The three-helix repeat motifs are indicated in blue, cyan, green, yellow, orange and red, respectively. The N-terminal capping helices are shown in purple. Secondary structure elements and the three insertions (Ins1-3) into the helical solenoid structure are indicated. See also Figure S4.

(B) Surface conservation in FAM203 family proteins. The same views as in panel (A) are shown. A cyan-white-magenta color gradient indicates increasing surface conservation, based on the similarity score from the sequence alignment in Figure S5. Highly conserved residues are indicated.

(C) Location of mutation sites in the Hgh1 structure. Residues replaced with Ala in the mutant constructs MutN, MutM and MutC are indicated in purple, gold and red, respectively.

(D) Interaction of Hgh1 mutant proteins with Eft as detected by crosslinking. Equimolar mixtures (10 μM) of Hgh1 mutant protein and Eft were treated with DMSO (left) or DSS (right). Reactions were analyzed by SDS-PAGE and Coomassie staining. Arrowhead indicates position of crosslinking products. See also Figure S6.

(E) Interaction of Hgh1-FLAG mutant proteins with TRiC. Purified mutant proteins were used in large (~6-fold) excess to capture Eft and TRiC from yeast lysate by anti-FLAG affinity chromatography. The eluted material was analyzed by SDS-PAGE and immunoblotting with anti-FLAG, anti-Tcp1 and anit-Eft.

(F) Complementation of growth defect of *eft2Δhgh1Δ* cells with mutant Hgh1 constructs. WT and mutant Hgh1 were expressed in *eft2Δhgh1Δ* cells under control of the ADH promoter. The growth phenotype was analyzed by dilution series on selective agar medium. See also Figure S6.

Figure 7. Hypothetical model for Hgh1 function in Eft biogenesis

(1) Hgh1 (green) binds cotranslationally to domain III of Eft folding intermediates. The N-domain of Eft may fold in a Hsp90-dependent manner (see accompanying paper by Schopf et al.). (2) Upon completion of Eft synthesis, Hgh1 recruits TRiC to the C-terminal domain of Eft. (3) The C-domain folds upon encapsulation in the TRiC cavity, followed by Eft release from TRiC. (4) The folded C-terminal domain stabilizes domain III, causing Hgh1 release and completing Eft folding. Hypothetical model based on data from this study as well as (Rüßmann et al., 2012).

STAR Methods

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be

fulfilled by the Lead Contact, Andreas Bracher (bracher@biochem.mpg.de).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Strains used in this study were *S. cerevisiae* BY4741 and YPH499 (Sikorski and Hieter, 1989).

The genotypes of the strains and their mutant derivatives are listed in KEY RESOURCES

TABLE. Cells were grown in YPD or synthetic complete (SC) medium at 30°C, unless

otherwise noted. The respective media contained either 2% glucose (YPD, SC), or 1%

raffinose/2% galactose (SCRaf/Gal). The amount of yeast cells used in the various experiments

was defined as $OD_{600} \cdot mL$ units (1 $OD_{600} \cdot mL$ cells corresponds to the amount of yeast cells

present in 1 mL of a culture with an OD_{600} of 1).

METHODS DETAILS

Molecular cloning

Plasmids used in this study are listed in the KEY RESOURCES TABLE. Primers used for

plasmid construction are listed in Table S2. DNA cloning was performed by DNA ligation using

T4 DNA ligase (New England Biolabs). PCR was performed using Pfu DNA Polymerase

(Promega) or Kod DNA Polymerase (Novagen).

A DNA fragment encoding His_6-Hgh1 was amplified from yeast genomic DNA of strain

YPH499 (see KEY RESOURCES TABLE) by PCR using primers Hgh1-F and Hgh1-R and

cloned into the EheI and HindIII sites of pProEx-HtB (Life Technologies) to generate pProEx-HtB-Hgh1. The amino acid sequence of the encoded His6-TEV tag is MSYYHHHHHHDYDIPTTENLYFQGALRGG.

A DNA fragment encoding His$_6$-Hgh1–FLAG was amplified from pProEx-HtB-Hgh1 by PCR using primers Hgh1-F and Hgh1-FLAG-R and cloned into the EheI and HindIII sites of pProEx-HtB (Life Technologies) to generate pProEx-HtB-Hgh1-FLAG.

A DNA fragment encoding His$_6$-Hgh1(1–363) was amplified from pProEx-HtB-Hgh1 by PCR using primers Hgh1(1-363)-F and Hgh1(1-363)-F and cloned into the EheI and HindIII sites of pProEx-HtB to generate pProEx-HtB-Hgh1(1–363).

A DNA fragment encoding His$_6$-Hgh1 was amplified from pProEx-HtB-Hgh1 by PCR using primers His-Hgh1overexpression-F and Hgh1-R and cloned into the SmaI and HindIII sites of p416gal (Mumberg et al., 1995) to generate p416gal-His$_6$-TEV-Hgh1.

A DNA fragment encoding His$_6$-Eft was amplified from yeast genomic DNA of strain YPH499 (see KEY RESOURCES TABLE) by PCR using primers His-Eft-F and His-Eft-R and cloned into the EheI and XhoI sites of p416gal-His$_6$-Hgh1 to generate p416gal-His$_6$-TEV-Eft.

A DNA fragment encoding His$_6$-Eft-N was amplified from p416gal-His$_6$-Eft by PCR using primers His-Eft-F and His-Eft(1-482)-R and cloned into the EheI and XhoI sites of pProEx-HtB to generate pProEx-HtB-Eft(1-482).

A DNA fragment encoding His$_6$-Eft-III was amplified from p416gal-His$_6$-Eft by PCR using primers Eftforw483 and Eftrev565 and cloned into the EheI and XhoI sites of pProEx-HtB to generate pProEx-HtB-Eft(483-565).

A DNA fragment encoding Eft-HA was amplified from p416gal-His$_6$-Eft by PCR using primers Eft-HA-F and Eft-HA-R and cloned into the SmaI and XhoI sites of p416gal (Mumberg et al., 1995) to generate p416gal-Eft-HA.

A DNA fragment encoding His$_6$-Eft-N was amplified from p416gal-His$_6$-Eft by PCR using primers Eft-HA-F and His-Eft(1-482)-R and cloned into the EheI and XhoI sites of p416gal-His$_6$-Eft to generate p416gal-His6-TEV-Eft(1-482).

A DNA fragment encoding Eft-N-HA was amplified from p416gal-His$_6$-Eft by PCR using primers Eft-HA-F and Eft482HA-R and cloned into the SmaI and XhoI sites of p416gal (Mumberg et al., 1995) to generate p416gal-Eft(1-482)-HA.

A DNA fragment encoding Eft-III+C-HA was amplified from p416gal-His$_6$-Eft by PCR using primers Eft2forw483 and Eft-R and cloned into the SmaI and MfeI sites of p416gal-Eft-HA to generate p416gal-Eft(483-842)-HA.

A DNA fragment encoding Eft-C-HA was amplified from p416gal-His$_6$-Eft by PCR using primers Eft2forw566 and Eft-HA-R and cloned into the SmaI and XhoI sites of p416gal-Eft-HA to generate p416gal-Eft(566-842)-HA.

A DNA fragment encoding Hgh1-HA was amplified from pProEx-HtB-Hgh1 by PCR using primers Hgh1-HA-F and Hgh1-HA-R and cloned into the XbaI and HindIII sites of p415adh (Mumberg et al., 1995) to generate p415adh-Hgh1-HA.

A DNA fragment encoding Hgh1mutN-FLAG (point mutations E7S/L8A/F11A) was amplified from pProEx-HtB-Hgh1 by PCR using primers Hgh1mutN-F and Hgh1-R and cloned into the EheI and HindIII sites of pProEx-HtB-Hgh1 to generate pProEx-HtB-Hgh1mutN-FLAG.

A DNA fragment encoding Hgh1mutN-HA was amplified from pProEx-HtB-Hgh1 by PCR using primers Hgh1mutN-HA-F and Hgh1-HA-R and cloned into the XbaI and HindIII sites of p415adh (Mumberg et al., 1995) to generate p415adh-Hgh1mutN-HA.

The point mutations K236A/N237H/F240A were introduced into pProEx-HtB-Hgh1-FLAG following the QuikChange protocol (Stratagene) using the primers Hgh1mutM-F and Hgh1mutM-R to generate pProEx-HtB-Hgh1mutM-FLAG.

The point mutations K236A/N237H/F240A were introduced into p415adh-Hgh1-HA following the QuikChange protocol (Stratagene) using the primers Hgh1mutM-F and Hgh1mutM-R to generate p415adh-Hgh1mutM-HA.

The point mutations Y326A/R330A/H333A were introduced into pProEx-HtB-Hgh1-FLAG following the QuikChange protocol (Stratagene) using the primers Hgh1mutC-F and Hgh1mutC-R to generate pProEx-HtB-Hgh1mutC-FLAG.

The point mutations Y326A/R330A/H333A were introduced into p415adh-Hgh1-HA following the QuikChange protocol (Stratagene) using the primers Hgh1mutN-F and Hgh1mutN-R to generate p415adh-Hgh1mutC-HA.

A DNA fragment encoding Hgh1mutN+M-FLAG (point mutations E7S/L8A/F11A/K236A/N237H/F240A) was amplified from pProEx-HtB-Hgh1mutM-FLAG by PCR using primers Hgh1mutN-F and Hgh1-R and cloned into the EheI and HindIII sites of pProEx-HtB-Hgh1 to generate pProEx-HtB-Hgh1mutN+M-FLAG.

A DNA fragment encoding Hgh1mutN+M-HA was amplified from p415adh-Hgh1mutM-HA by PCR using primers Hgh1mutN-HA-F and Hgh1-HA-R and cloned into the XbaI and HindIII sites of p415adh (Mumberg et al., 1995) to generate p415adh-Hgh1mutN+M-HA.

To generate the internally tagged variant of *CCT2*, *CCT2*–Int, we first generated the plasmid pInt to facilitate the cloning procedure. The DNA sequence bearing two multiple cloning sites, MCS1 and MCS2, flanking the insertion sequence coding for a StrepII tag, the Calmodulin Binding Protein and a His$_8$ tag, followed by the *trp1* selection marker, DNA1, was synthesized by Geneart AG. This construct was cloned into the EheI and HindIII sites of pProEx-HtB (Life Technologies) to generate pInt. DNA sequences for the N- and C-terminal fragments of *CCT2*, amino acids 1–362 and 363–527, respectively, were amplified by PCR using the primer pairs CCT2N-F / CCT2N-R and CCT2C-F / CCT2C-R and inserted into MCS1 and MCS2 of pInt, using the *Bam*HI / *Not*I and *Xho*I / *Pst*I restriction site pairs, yielding the plasmid pInt-CCT2.

All constructs were verified by DNA sequencing.

Yeast strains

Yeast strains used in this study are listed in the KEY RESOURCES TABLE. The *hgh1Δ* and *eft2Δ* strains were created by transforming strain BY4741 with a PCR-generated cassette using pFA6aKanMX4 (Wach et al., 1994) or pAG60 as a template and primers hgh1 KO-F/hgh1 KO-R or eft2KO-F/eft2KO-R, respectively. Transformants were selected by growth on media containing 300 μg mL^{-1} G418 or media lacking uracil, as appropriate, and confirmed by genomic PCR using external primers.

The *hgh1*::*hgh1*-FLAG strain was created by transforming strain BY4741 with a PCR-generated cassette using pFA6a-6GLY-FLAG-His3MX6 (Addgene) as a template and primers gen_Hgh1-FLAG-F/gen_Hgh1-FLAG-R. Transformants were selected by growth on media lacking histidine.

The *eft2*::*eft2*-HA and *hgh1*::*hgh1*-FLAG *eft2*::*eft2*-HA strains were created by transforming strains wild-type BY4741 and BY4741 *hgh1*::*hgh1*-FLAG with a PCR-generated cassette using pCY 3140-02 (Young et al., 2012) as a template and primers gen_Eft2-HA-F/gen_Eft2-HA-R.

Transformants were selected by growth on the respective selective media containing 300 µg mL^{-1} Hygromycin B.

The YPH499 *cct2*::*CCT2*–Int-*trp1* strain was created by transforming the wild-type strain YPH499 with a PCR-generated cassette using the plasmid pInt-CCT2 as a template and primers CCT2int-F/CCT2int-R. Transformants were selected by growth on media lacking tryptophan.

Preparation of cell extracts for immunoblot analysis

Yeast strains were cultured at 30°C, unless noted otherwise. Log-phase cells (OD$_{600}$ less than 1.0) were used for experiments, except where indicated.

Denaturing lysis

Yeast culture (6.4 OD$_{600}$ ·mL) was treated at a 10:1 ratio with trichloroacetic acid (TCA). The cells were transferred into 2 mL-siliconized tubes (Sigma) and flash-frozen in liquid nitrogen for storage. The thawed samples were centrifuged at 20,000 x g for 8 min at RT. The pellet was washed with 1 mL acetone (–20°C). The dry residue was re-suspended in 100 µL BB1 buffer (1% SDS, 6 M urea, 50 mM Tris-HCl pH 7.5, 1 mM EDTA) and 100 µL of 1 mm-glass beads were added. Cells were lysed by bead-milling with a MP-Beadbeater24 (2 x 40 s at 6.0 m s^{-1}, intermitted cooling with ice). Subsequently, 1 mL IP buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% Triton X-100, 5 mM EDTA, 10 mM β-mercaptoethanol (β-ME)) supplemented with 1

mM phenylmethylsulfonyl fluoride (PMSF) and Complete protease inhibitor cocktail (Roche) were added. After centrifugation at 20,800 x g for 15 min at 4°C, the supernatant was collected.

Native lysis and fractionation

Yeast culture (13 $OD_{600} \cdot mL$) was harvested by centrifugation at 5250 x g for 15 min. After transfer into 2 mL-siliconized tubes (Sigma), the cells were re-suspended in 500 µl IP buffer supplemented with 125 U benzonase, 2 mg L^{-1} puromycin·2 HCl, 1 mM PMSF and Complete protease inhibitor cocktail. After addition of 300 µL 1mm-glass beads, the mixture was cooled on ice. The cells were lysed by bead-milling with a MP-Beadbeater24 (4 x 20 s at 6.0 m s^{-1}, with intermittent cooling). Cell debris and glass beads were removed by sedimentation at 500 x g for 1 min at 4°C. The lysate (50 µL) was split into soluble and pellet fractions by centrifugation at 20,800 x g for 15 min at 4°C. The pellet was washed with 100 µL IP buffer and subsequently dissolved in 50 µL BB1 buffer.

SDS-PAGE

Protein samples were separated by electrophoresis on NuPAGE 4%–12% Bis-Tris SDS gels (Invitrogen) using NuPAGE MOPS SDS running buffer (Invitrogen) at 200 V.

Native PAGE

Protein samples were separated by electrophoresis on NuPAGE 3%–8% Tris-acetate protein gels (Invitrogen) using NPage buffer (50 mM Tris and 38 mM glycine) at 150 V at 4°C.

Immunoblotting

Proteins were transferred from polyacrylamide gels to nitrocellulose (GE Healthcare) membranes in blotting buffer (25 mM Tris, 192 mM glycine, 20% methanol) at 250 mA. Membranes were washed in TBST buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Tween-20) and blocked with 3% bovine serum albumine in TBST buffer for 1 h at room-temperature. Membranes were incubated with primary antibodies in TBST buffer overnight at 4°C and washed three times with TBST. Membranes were then incubated with secondary antibody for 1 h at room-temperature and again washed 3 times with TBST. Horseradish peroxidase (HRP)-conjugated (Bio-Rad, Dako or Sigma-Aldrich) secondary antibodies were used. Luminata Classico (Merck) was used as substrate of HRP. The ImageQuant LAS 4000 mini (GE Healthcare) was used for signal detection. Images were analyzed using AIDA software (Raytest).

Immune precipitation and label-free quantification by MS

The yeast strain yLM003 (BY4741 *eft2*::*eft2*-HA *hgh1*::FLAG-*hgh1*) was grown in 4 L YPD medium at 30°C in shaking flask culture to an OD_{600} of approximately 2.0–2.5. Cells were harvested by sedimentation at 4000 x g for 25 min at 4°C, yielding ~20 g wet cell mass. The cell pellet was re-suspended in 40 mL ice-cold TRiC standard buffer (50 mM HEPES-NaOH, pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 10% glycerol) supplemented with 100 U benzonase, 0.2 mM 1,4-dithiothreitol (DTT), 1 mM PMSF and Complete protease inhibitor cocktail. Subsequently, 60 g 1mm-glass beads were added to the ice-cooled cell suspension. Cells were lysed by bead-milling in an ice-cooled Microspec bead beater. Cellular debris was removed by centrifugation at 2000 x g and 40,000 x g for 20 min and 1 h, respectively, all at 4°C. The supernatant was split into three equal fractions, which were incubated with 250 µl anti-FLAG, anti-HA or anti-Myc affinity

resin, respectively, for 1 h at 4°C on a rotating wheel. Subsequently, the resins were collected

into chromatography columns and washed with 20 column volumes (CV) TRiC standard buffer.

For protein elution, 1 CV 0.1 M glycine-HCl pH 3.0 was added, incubated at RT for 5 min and

the flow-through collected into a tube containing 0.1 CV of 0.5 M Tris-HCl pH 8.0 and 1.5 M

NaCl. This elution step was repeated three times. The resulting fractions 1–4 were analyzed by

SDS-PAGE. Fractions 2 and 3 were used for mass spectrometry (MS) analysis.

Sample preparation for mass spectrometry

Proteins were digested using the filter-aided sample preparation (FASP) method, essentially as

described (Wisniewski et al., 2011) without further fractionation. After tryptic digestion,

peptides were desalted using homemade columns containing C18 Empore disks (Rappsilber et

al., 2003). Peptides were eluted with 1% formic acid in 70% acetonitrile (ACN) and dried in a

vacuum concentrator.

LC-MS/MS and data analysis

Tryptic peptides were dissolved in 5% formic acid solution and analyzed by nanoLC-MS/MS

using an EASY-nLC 1200 nano liquid chromatography system (Thermo Fisher) coupled to a Q-

Exactive HF mass spectrometer (Thermo Fisher). Samples were injected onto a home-made 25

cm silica reversed-phase capillary column (New Objective) packed with 1.9-μm ReproSil-Pur

C18-AQ (Dr. Maisch GmbH). Samples were loaded on the column by the nLC autosampler at a

flow rate of 0.5 μL per minute. No trap column was used. Peptides were separated by a 190-min

gradient of 5–30% between buffer RPA (0.1% formic acid in water) and buffer RPB (0.1%

formic acid in 80% acetonitrile) at a flow rate of 300 nL min[-1]. MS/MS analysis was performed

with standard settings using cycles of one high resolution (60000 FWHM setting) MS scan followed by MS/MS scans of the 15 most intense ions with charge states of 2 or higher at a resolution setting of 15000 FWHM. Protein identification and SILAC based quantitation was performed with MaxQuant (version 1.3.0.5) using default settings (Cox and Mann, 2008). The UNIPROT *S. cerevisiae* (strain ATCC 204508 / S288c) database (version 2014-04-06) was used for protein identification. MaxQuant uses a decoy version of the specified UNIPROT database to adjust the false discovery rates for proteins and peptides below 1%. Volcano Plots were calculated using Perseus software (version 1.6.0.7) (Tyanova et al., 2016) with FDR and s0 settings of 0.05 and 2.0, respectively.

Protein purification procedures

All protein purification steps were performed at 4°C unless otherwise indicated. Protein concentrations in the final preparations were determined by absorbance at 280 nm. Purified protein samples were concentrated by ultrafiltration and snap-frozen in liquid nitrogen for storage at –80°C.

His$_6$-tagged Hgh1 constructs

The respective pProEx-HtB plasmid was transformed into *Escherichia coli* Bl21 codon+ RIL cells. Cells were grown in lysogenic broth (LB) medium at 37°C to an $OD_{600} = 0.7$ and induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 4 h at 37°C (or overnight at 20°C). The harvested cells were re-suspended in ice-cold buffer H (20 mM HEPES-NaOH pH 7.4, 150 mM NaCl, 10 mM β-ME), which was supplemented with 10 mM imidazole and 1 mM PMSF. The cells were lysed by ultrasonication on ice. After removal of cell debris by centrifugation, the supernatant was applied to Ni^{2+}-chelating Sepharose equilibrated in buffer H.

The column was washed with a step gradient of buffer H containing increasing amounts of imidazole (10 / 30 / 100 mM). The bound protein was eluted with 3 CVs buffer H containing 250 mM imidazole. This was followed by overnight cleavage of the His$_6$ moiety at 4°C with His$_6$-TEV protease. After transfer into buffer H containing 10 mM imidazole using a desalting column, the material was passed over the Ni-chelating Sepharose column and the flow-through collected. Finally, the concentrated flow-through was subjected to size exclusion chromatography (SEC) on Superdex 200 (GE Healthcare) in buffer H containing 1 mM DTT instead of β-ME.

TRiC

TRiC from *S. cerevisiae* was purified using a three-step chromatographic method modified from (Leitner et al., 2012), which takes advantage of the strong interaction of TRiC with Plp2. Yeast cells were suspended in buffer YN (50 mM HEPES-NaOH pH 7.4, 200 mM NaCl, 5 mM MgCl$_2$, 10% glycerol) containing 20 mM imidazole, 10 mM β-ME, 1 mM PMSF, and Complete protease inhibitor cocktail, and lysed using a bead beater (BioSpec Products, Bartlesville, OK, USA). Cell debris was removed by sequential centrifugation steps at 3000 x g and 18,600 x g. The clear supernatant was subsequently incubated for 1 h with Ni^{2+}-chelating Sepharose beads, which were decorated with Plp2 bearing a His$_6$-tag. The beads were collected in a column, and washed with a step gradient of buffer YN containing increasing amounts of imidazole (20/50/100 mM). Bound TRiC was then eluted together with Plp2 using buffer YN containing 250 mM imidazole. To separate Plp2 and substrates from TRiC, the protein-containing fractions were supplemented with 1 mM ATP and applied on Heparin resin (GE Healthcare). Unbound proteins were eluted with 5 CVs of buffer YH (20 mM HEPES-NaOH pH 7.4, 10 mM MgCl$_2$, 10 mM CaCl$_2$, 10% glycerol, 10 mM β-ME) supplemented with 100 mM NaCl and 1 mM ATP. Note that this procedure failed to remove all contaminating substrates as judged by SDS-PAGE. TRiC was

eluted with a linear salt gradient from 200 mM to 1000 mM NaCl in buffer YH. As a final step, the TRiC-containing fractions were subjected to SEC on Superose 6 (GE Healthcare), which was developed with buffer YSEC containing 10 mM HEPES-NaOH pH 7.4, 200 mM NaCl, 5 mM MgCl$_2$, 10% glycerol and 1 mM DTT. The TRiC-containing fractions were merged, concentrated to 2 g L^{-1} by ultrafiltration and snap-frozen in liquid nitrogen for storage at –80°C.

Eft-HA

The yeast strain yLM003 (BY4741 *eft2*::*eft2*-HA) was grown in 3 L YPD medium to an OD$_{600}$ of 7.5. The cells were harvested by centrifugation, resulting in a cell pellet of 43 g wet weight. The cells were re-suspended in 100 mL buffer Q (20 mM Tris-HCl pH 7.6, 5 mM MgCl$_2$, 10% glycerol, 1 mM DTT) supplemented with 300 mM KCl, 1 mM PMSF and Complete protease inhibitor cocktail and lysed by bead milling. After adjusting pH 7.4 with 1 M Tris base, the lysate was cleared by sequential centrifugation steps at 24,000 x g and 186,000 x g for 20 min and 1 h, respectively. The supernatant was diluted 1:1 with buffer Q and this mixture dialyzed against buffer Q containing 0.1 mM PMSF. The dialysate was cleared by ultracentrifugation at 186,000 x g for 20 min. The supernatant was loaded onto a DEAE-Sepharose (GE Healthcare) column (80 mL CV) equilibrated with buffer Q supplemented with 50 mM KCl (Q50). Unbound material was eluted with 3 CVs Q50. Eft-HA was eluted with a linear gradient (10 CVs) from 50 mM to 350 mM KCl in buffer Q. The pooled fractions containing Eft-HA were dialyzed against buffer Q supplemented with 30 mM KCl. The dialysate was applied to a MonoQ (GE Healthcare) column (20 mL CV). The column was washed with buffer Q50 until the absorbance at 280 nm returned to the baseline, followed by a wash with 3 CVs of buffer Q supplemented with 100 mM KCl. Two peaks of Eft-HA were eluted with a linear gradient (10 CVs) from 50 mM to 300 mM KCl in buffer Q, at conductivities of 19.5–20.5 mS cm^{-1} and 22.3–23.3 mS cm^{-1},

respectively. The latter peak was further purified on Heparin Sepharose (GE Healthcare) using the same buffer system and SEC on Superdex 200 (GE Healthcare) in buffer YSEC (50 mM HEPES-NaOH, pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, 1 mM DTT).

His$_6$-Eft-N

BY4741 cells were transformed with p416Gal-His$_6$-Eft(1-482). The transformed cells were grown on SC Δura medium containing 2% Glucose. His$_6$-Eft-N expression was induced by overnight growth on 6 L SC Δura medium containing 2% galactose and 1% raffinose. The harvested cells were re-suspended in 100 mL buffer E (50 mM Tris-HCl pH 7.6, 300 mM KCl) supplemented with 10 mM imidazole, 10 mM β-ME, 1 mM PMSF and Complete protease inhibitor cocktail. The cells were lysed by bead milling, and cell debris was removed by centrifugation steps at 2000 x g and 24,000 x g. The lysate was applied to Ni^{2+}-chelating Sepharose. The column was washed with a step gradient of imidazole in buffer E (10 mM/50 mM/80 mM). His$_6$-Eft-N was eluted with buffer E containing 250 mM imidazole. His$_6$-Eft-N was further purified by SEC on Superdex 200 using buffer EMQ (20 mM Tris-HCl pH 7.6, 30 mM KCl, 5 mM $MgCl_2$, 10% glycerol, 1 mM DTT) and anion exchange chromatography on MonoQ using a gradient from 30 to 300 mM KCl in buffer EMQ.

His$_6$-Eft-III

The plasmid pProEx-HtB-Eft(483-565) was transformed into *E. coli* Bl21 (DE3) cells. Cells were grown in 2 L LB medium at 37°C to $OD_{600} = 0.7$ and induced with 1 mM IPTG for 3 h at 37°C. The cell pellet was re-suspended in 70 mL buffer L (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM β-ME) augmented with 25 mg lysozyme, 175 U benzonase and 1 mM PMSF. Cells were lysed by ultrasonication. The insoluble material was collected after sedimentation at 40,000 x g for 25 min at 4°C. The inclusion bodies were washed twice by homogenization in 25 mL buffer W (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5% Triton X-100, 10 mM β-ME) and

sedimentation at 40,000 x g. Subsequently, the inclusion bodies were washed in the same manner with buffer L. Finally, the pellet was homogenized in 10 mL buffer RS (8 M guanidinium-HCl pH 8.0, 10 mM β-ME), followed by centrifugation at 55,000 x g for 30 min at 4°C. His_6-Eft-III in the supernatant was isolated by ion affinity chromatography under denaturing conditions on Ni^{2+}-NTA Agarose (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

Pulldown of Hgh1-FLAG from cell lysate

The yeast strain yLM003 (BY4741 Eft2::Eft2-HA) was grown in 6 L YPD medium to an OD_{600} of 7.0. Yeast cells (80 g) were suspended in 100 mL ice-cold buffer TSB (50 mM HEPES-NaOH pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 10% glycerol) supplemented with 1 mM PMSF, 0.5 mM DTT and Complete protease inhibitor cocktail. After bead mill lysis, cell debris was removed by centrifugation. Purified Hgh1–FLAG (10 mg) was added to the lysate, followed by incubation for 30 min at 4°C. Subsequently, anti-FLAG affinity resin (10 mL) was added, followed by further incubation (30 min). The resin was collected into a gravity-flow column and washed with 20 CVs buffer TSB supplemented with 0.5 mM DTT. For recovery of the bound protein, the resin was incubated with 1 CV buffer TSB containing 0.5 mM DTT and 0.1 g L^{-1} FLAG peptide (Sigma) for 10 min, followed by draining and collecting of the eluate. This step was repeated three times.

Chemical crosslinking

Equimolar mixtures of Hgh1-FLAG/Eft-HA or Hgh1-FLAG/Eft-HA/TRiC in presence or absence of Eft ligands (2 mM) in buffer XL (20 mM HEPES-NaOH pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT) were prepared. These were mixed at a 25:1 ratio with either a solution

of 25 mM disuccinimidyl suberate (DSS) in DMSO or DMSO alone, followed by incubation at 25°C for 30 min. Reactions were quenched by addition of 150 mM NH4HCO3. Samples were analyzed by SDS-PAGE or native-PAGE as indicated.

Hydrogen/deuterium exchange–mass spectrometry

Sample preparation

Eft was prepared at 2 µM in HDX buffer (50 mM HEPES-NaOH pH 7.4, 150 mM NaCl, 5 mM MgCl2, 1 mM tris(2-carboxyethyl)phosphine (TCEP)) with or without 200 µM Sordarin or 4 µM Hgh1. To initiate the deuterium exchange reaction, 5 µL protein was added to 45 µL deuteration buffer (HDX buffer prepared in D2O) and incubated for different times (10, 30, 100, 300 or 1000 s) at 20°C before quenching the reaction by addition of 50 µL ice cold quench buffer (100 mM sodium phosphate pH 2.1, 10 mM TCEP, 3 M guanidine HCl) to a final pH of 2.6.

Peptide Mass Analysis and Data Processing

Quenched samples were injected into a Waters ACQUITY UPLC M-class with H/DX via a 50 µL sample loop. Proteins were digested using an Enzymate BEH-pepsin column (Waters) at a flow rate of 200 µL min[-1] and temperature of 20°C. Peptides were trapped and desalted for 3 min at 100 µL min[-1] before transfer to a 1.0 x 100 mm ACQUITY UPLC peptide CSH C18 column (Waters) held at 0°C. Peptides were eluted over 7 min with an 8-40% acetonitrile gradient in 0.1% formic acid, pH 2.5. Mass analysis was performed on a Waters Synapt G2Si. T-wave ion mobility was used as an orthogonal peptide separation step between the UPLC and mass spectrometer (Iacob et al., 2008). Ion guide settings were adjusted to minimize gas-phase back exchange as described previously (Guttman et al., 2016). Peptides were identified by analyzing MS[E] data for 4-5 undeuterated control experiments using PLGS (Waters). Mass spectra were processed in DynamX (Waters) and peak selection was manually verified for all peptides. All

experiments were performed under identical conditions. Deuterium levels were therefore not corrected for back exchange and are reported as relative (Wales and Engen, 2006). Experiments were performed in triplicate, including at least two separate preparations of protein-protein or protein-ligand complexes where applicable. Errors of mass measurements were <0.2 Da.

Circular dichroism spectroscopy

Far-UV CD spectra as well as thermal transitions of proteins were measured with a Jasco J-715 spectrometer equipped with a Peltier-thermostat using 0.1 cm cuvettes. Wavelength scans were recorded at 4°C, temperature scans at the indicated wavelength applying a temperature gradient of 60°C h^{-1}. The proteins were analyzed at the indicated concentration in a buffer containing 50 mM potassium phosphate pH 8.0.

Crystallization

Initial crystals of Hgh1(1-363) were obtained with the help of the MPIB Crystallization Facility by the sittting-drop vapor diffusion method using the Index crystallization screen (Hampton Research, Aliso Viejo, CA) at 4°C by mixing 100 nL sample with 100 nL reservoir.

The final crystals of spacegroup P1 grew in a hanging drop vapor diffusion setup at 4°C from a mixture of 1 µL Hgh1(1-363) at 8.2 mg mL^{-1} and 1 µL precipitant containing 10% PEG-3350, 0.2 M MgCl$_2$, 0.1 M Tris-HCl pH 8.5 and 20 mM DTT. For cryoprotection, crystals were incubated in 17.5% PEG-3350, 7.5% glycerol, 0.2 M MgCl$_2$, 0.1 M Tris-HCl pH 8.5 and 20 mM DTT and 20% PEG-3350, 15% glycerol, 0.2 M MgCl$_2$, 0.1 M Tris-HCl pH 8.5 and 20 mM DTT for each 10 min, respectively, before flash-cooling in liquid nitrogen.

Crystals of spacegroup *C*2 grew in a hanging drop vapor diffusion setup at 4°C from a mixture of 1 µL Hgh1(1-363) and 1 µL precipitant containing 16% PEG-3350, 0.1 M Na_3-citrate and 20 mM DTT. Crystals were stepwise transferred into 20% PEG-3350, 15% glycerol, 0.1 M Na_3-citrate and 20 mM DTT, followed by flash-cooling in liquid nitrogen.

Structure solution and refinement

X-ray diffraction data were collected by the oscillation method at beamlines ID29 and ID23-1 at European Synchrotron Radiation Facility (ESRF), Grenoble, France. In order to maximize data completeness in spacegroup *P*1, data were collected at two kappa angles, employing the minikappa goniometer head (Brockhauser et al., 2013).

The data were integrated and scaled with XDS (Kabsch, 2010). Pointless (Evans, 2006), Aimless (Evans and Murshudov, 2013) and Truncate (French and Wilson, 1978), as implemented in the CCP4i interface (Collaborative Computational Project, 1994), were used to convert the data to CCP4 format.

The structure of selenomethionine-labelled Hgh1(1-363) was solved by Se-single-wavelength anomalous diffraction (SAD) using the AutoSol pipeline as implemented in Phenix (Adams et al., 2010). Manual model building was performed with Coot (Emsley and Cowtan, 2004). The model was refined with Phenix (Adams et al., 2010). Anisotropic motions were modelled using Translation-Libration-Screw (TLS) parametrization. The crystals of native protein in spacegroup *P*1 were isomorphic. The final models contain four copies of Hgh1(1-363) each.

The crystal structure of Hgh1(1-363) in space group *C*2 was solved by molecular replacement with Molrep (Vagin and Isupov, 2001). Coot was employed for manual model building (Emsley and Cowtan, 2004). The model was refined with Refmac5 using local non-crystallographic

symmetry (NCS) restraints and TLS parametrization (Murshudov et al., 1997). The final model contains four copies of Hgh1(1-363). Residues facing solvent channels with disordered sidechains were modelled as alanines.

Structure analysis

Coordinates were aligned with Lsqkab and Lsqman (Kleywegt and Jones, 1994). Figures were generated with the programs Pymol (http://www.pymol.org) and ESPript (Gouet et al., 1999).
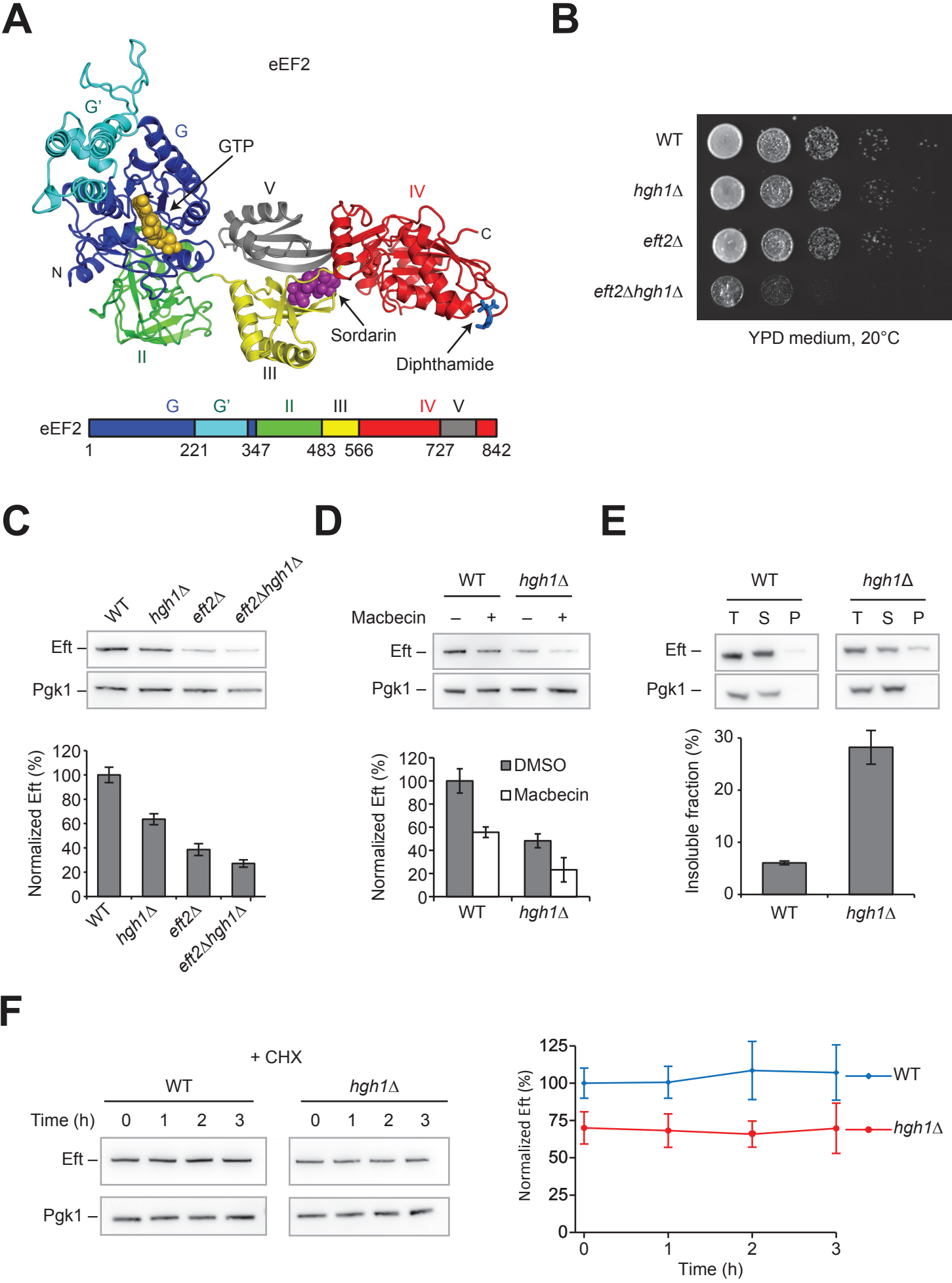
QUANTIFICATION AND STATISTICAL ANALYSIS

The relative levels of Eft in Figures 1C-F were quantified using AIDA software. Error bars represent standard deviation (SD) from three independent experiments.

P values for volcano plots in Figures 3 and S1 were calculated by Student's t test for proteins, which were detected by LS-MS/MS in three biological repeats.

DATA AND SOFTWARE AVAILABILITY

Accession Numbers

Coordinates and structure factors reported in this manuscript have been deposited in the Protein Data Bank with accession codes 6HB1, 6HB2 and 6HB3. The full gels of Figures 1C-F, 2B-D, and 6E can be viewed at Mendeley Data (URL: https://doi.org/10.17632/sw9sxts73z.1).
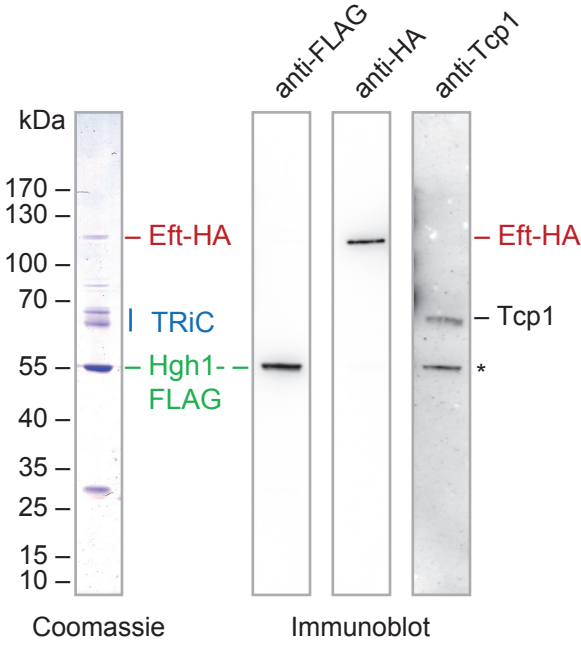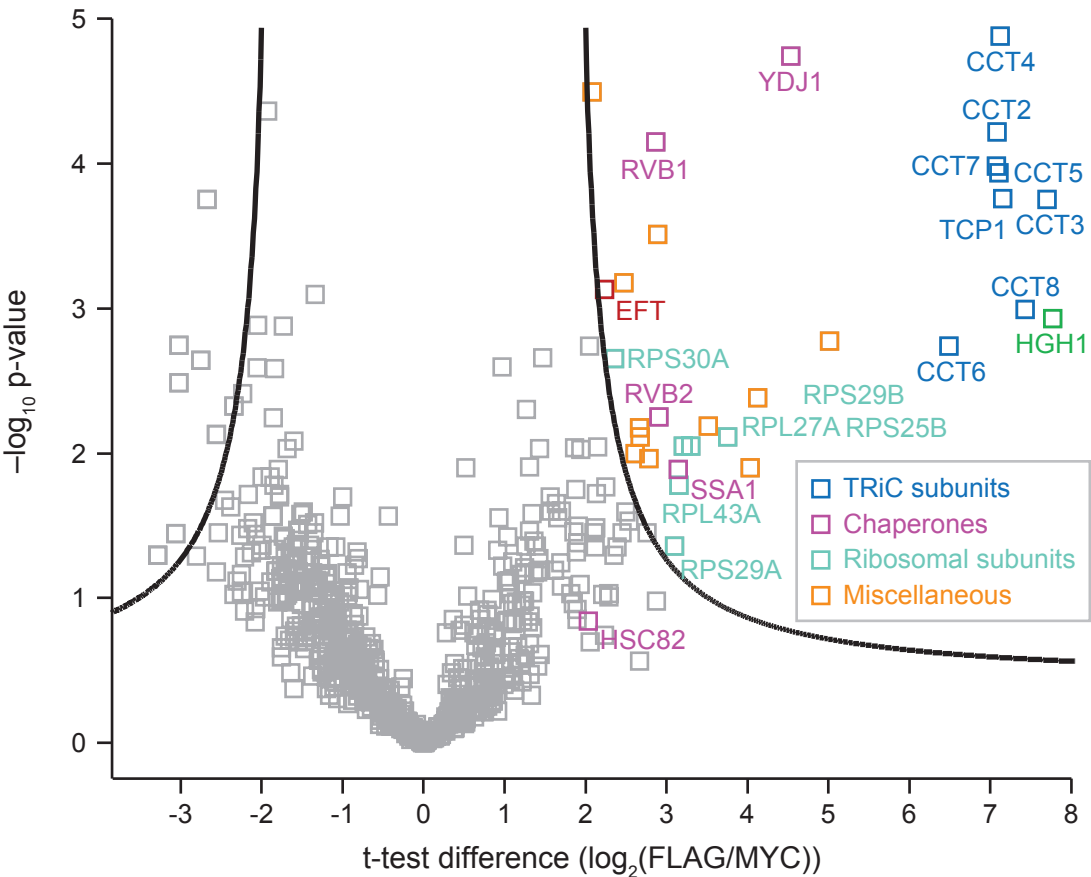
A

Input | + DSS

Lane 1 2 3 4 5 | 6 7 8 9 10

B

Hgh1 / DSS

Lane 1–14

C
Eft
300 s

Fraction H/D exchange
0   0.25   0.5   0.75

D
Difference:
Eft·Sordarin – Eft
300 s

Sordarin

H/D exchange difference
-0.25   0   0.25

E
Difference:
Eft / Hgh1 – Eft
300 s

433–438
536–540
522–538

H/D exchange difference
-0.25   0   0.25

F

Peptide 433–438
Peptide 522–538
Peptide 536–540

Relative deuterium uptake (Da)
Deuteration time (s)

Eft
Eft·Sordarin
Eft+Hgh1
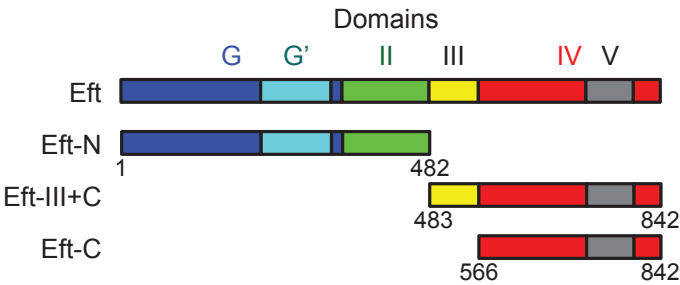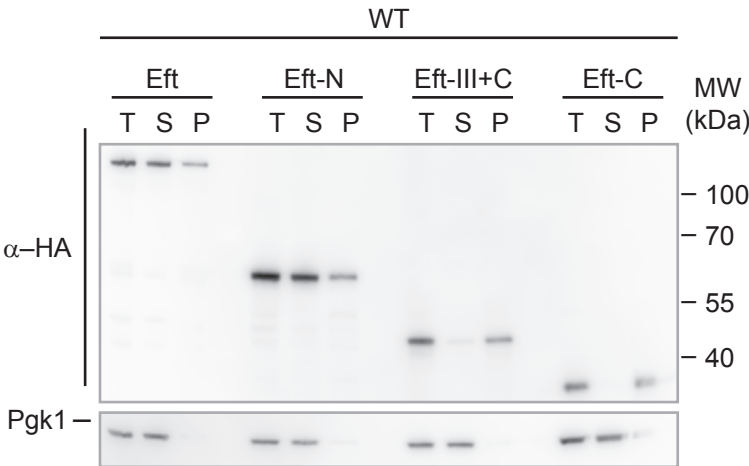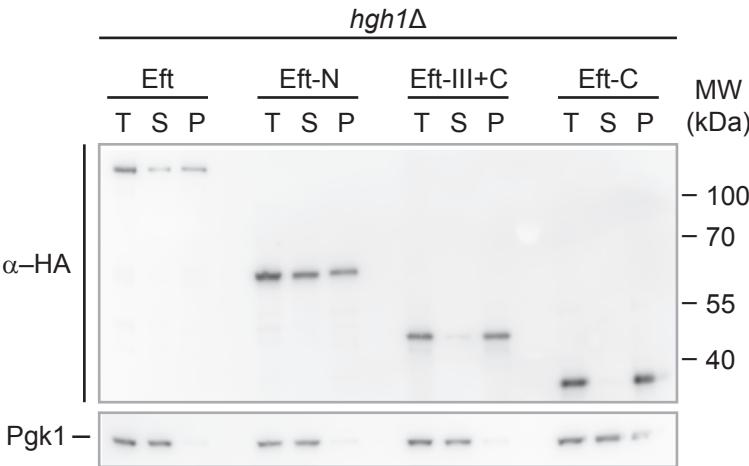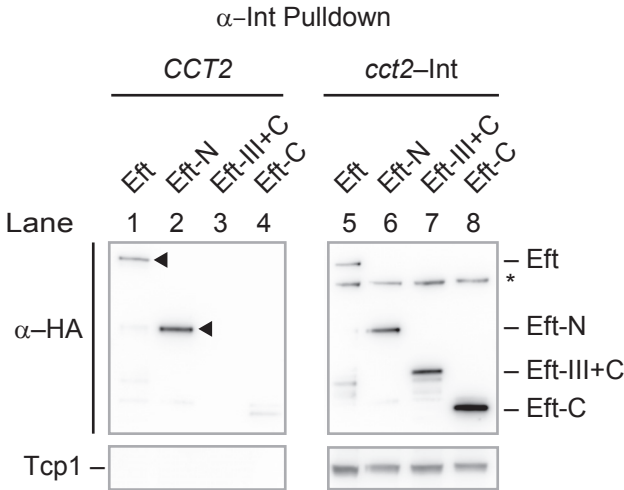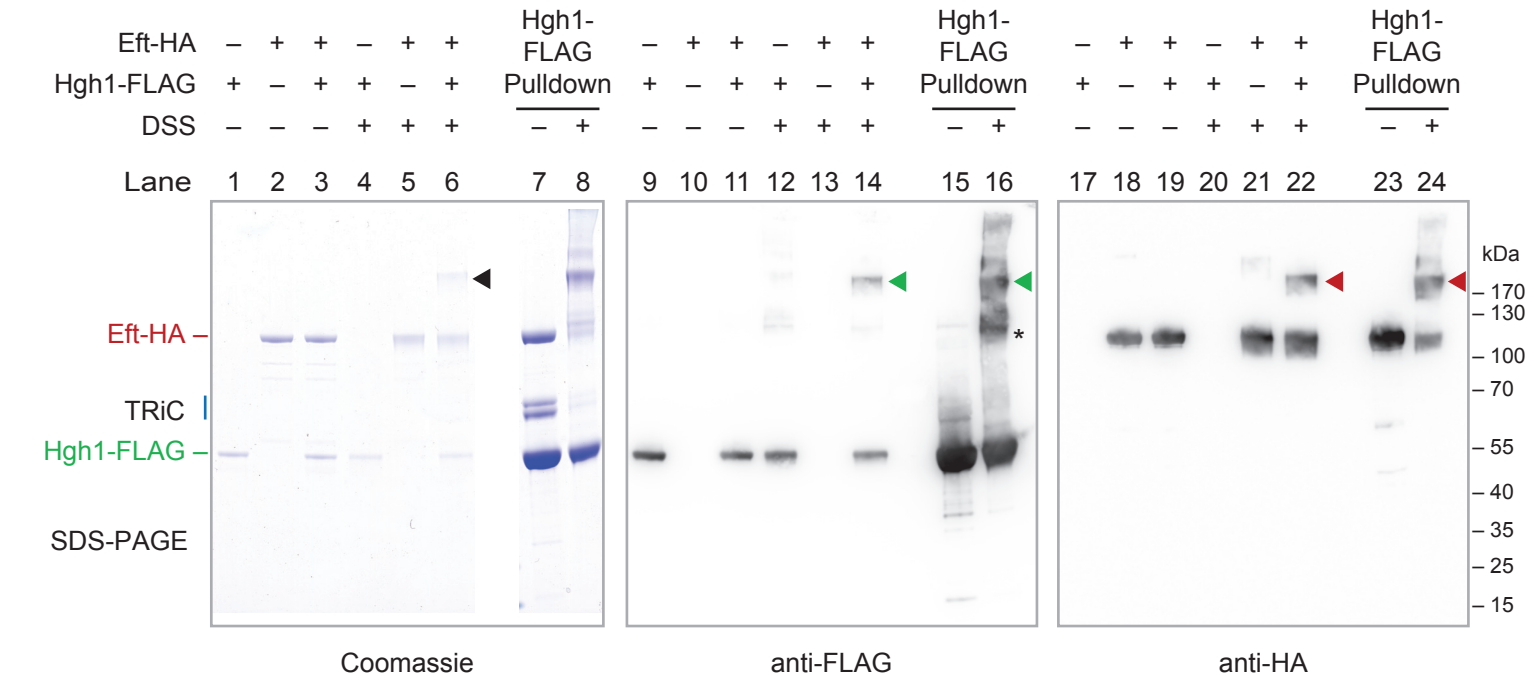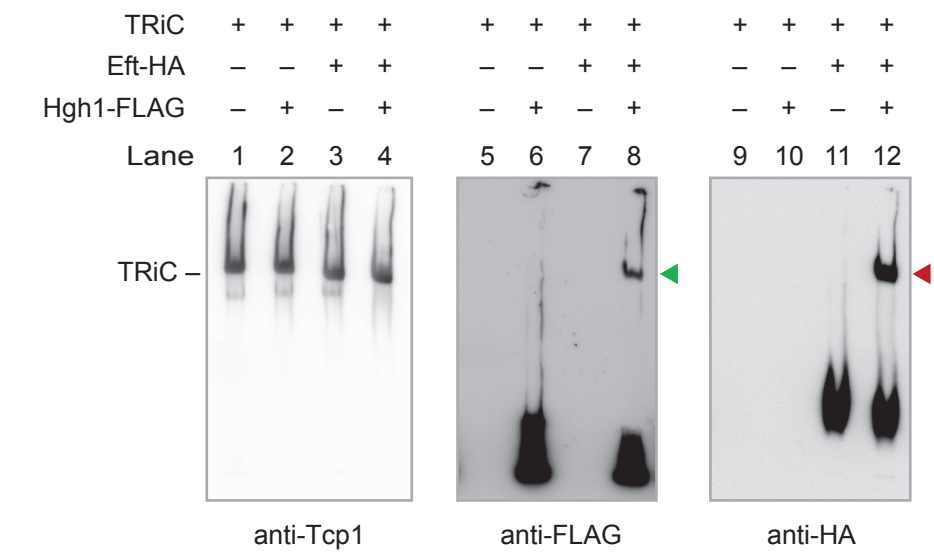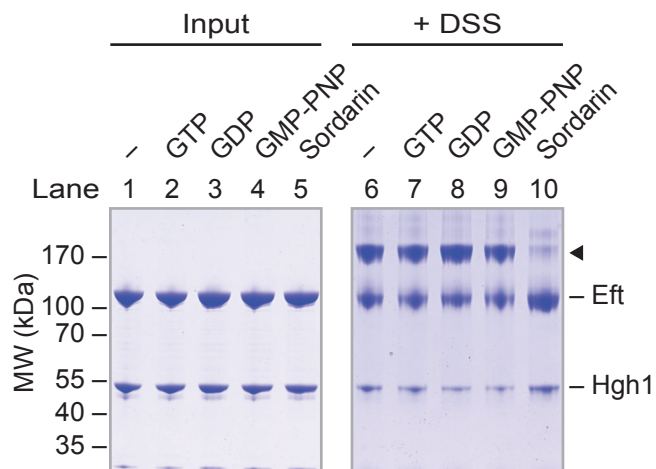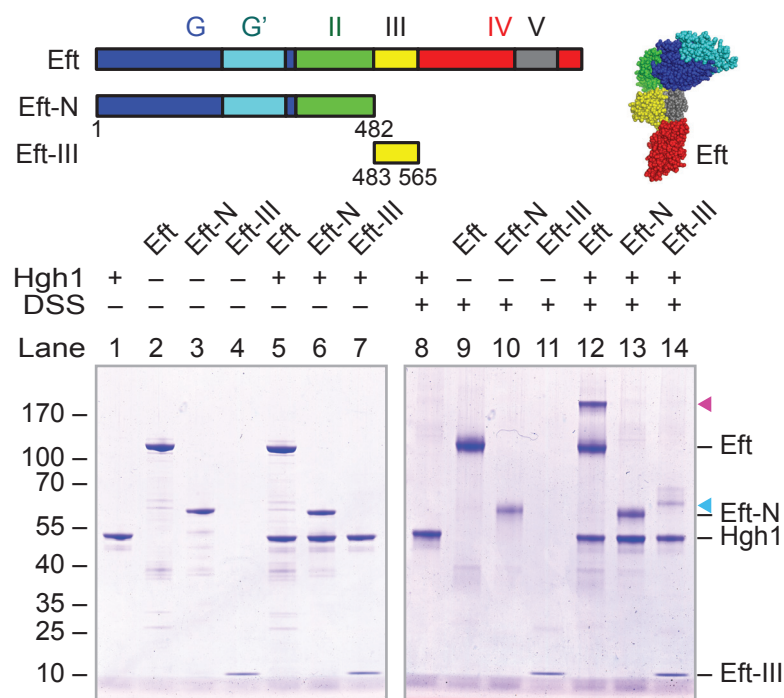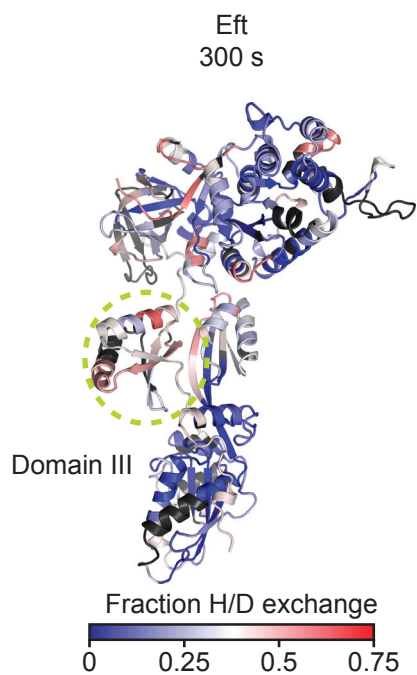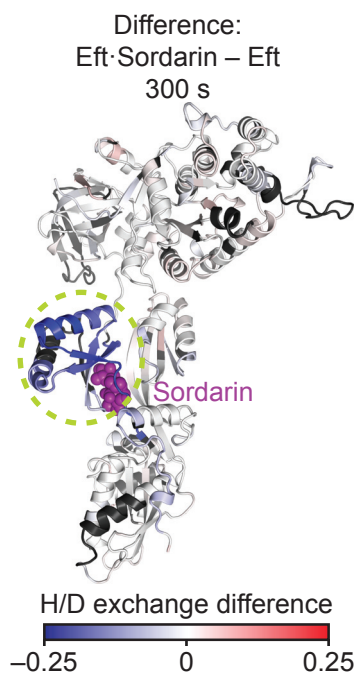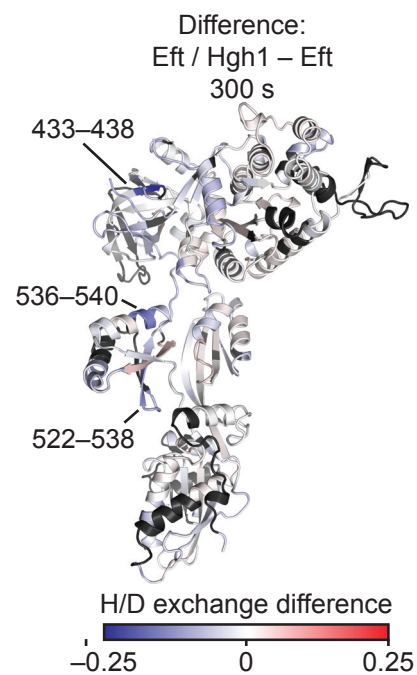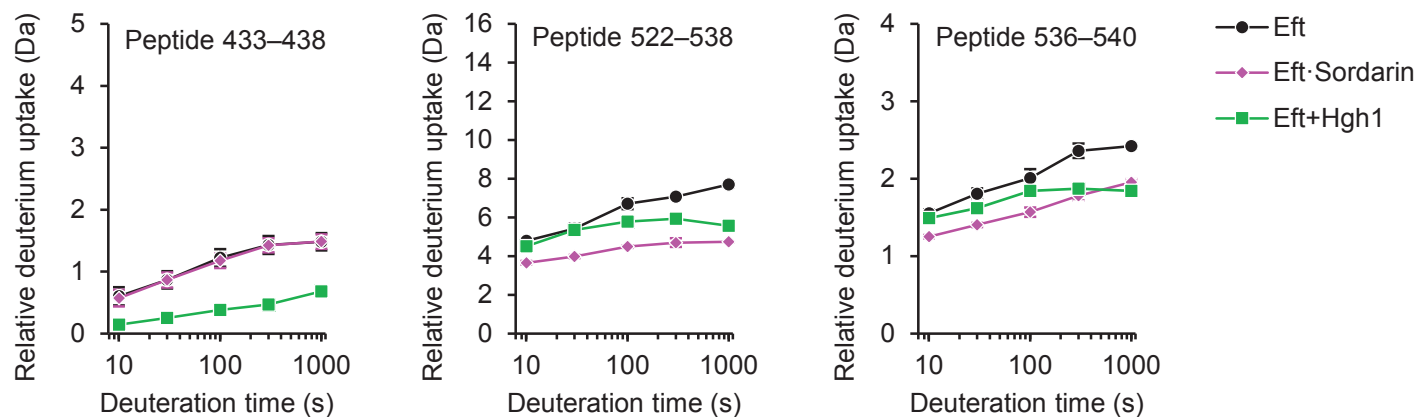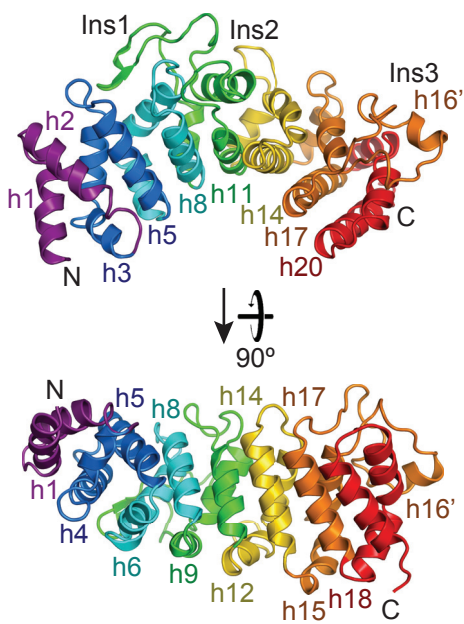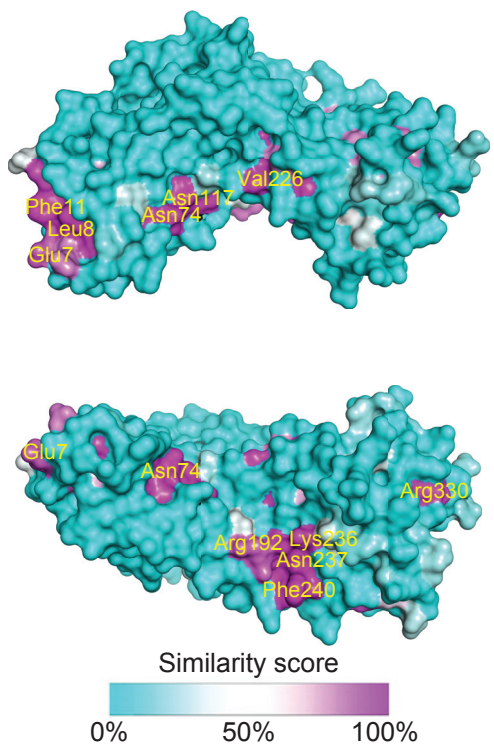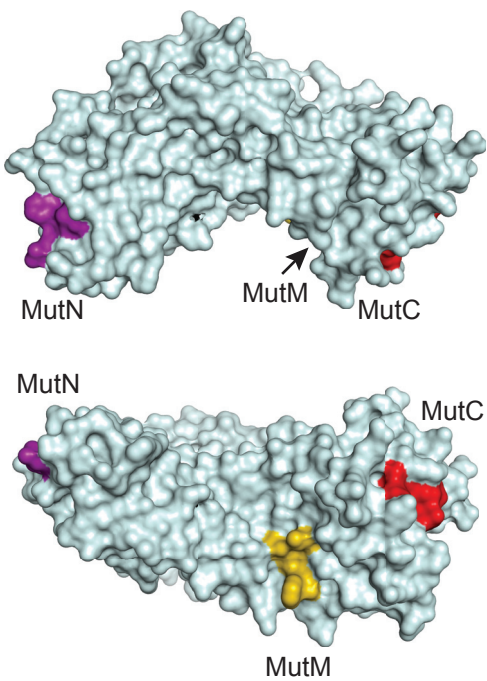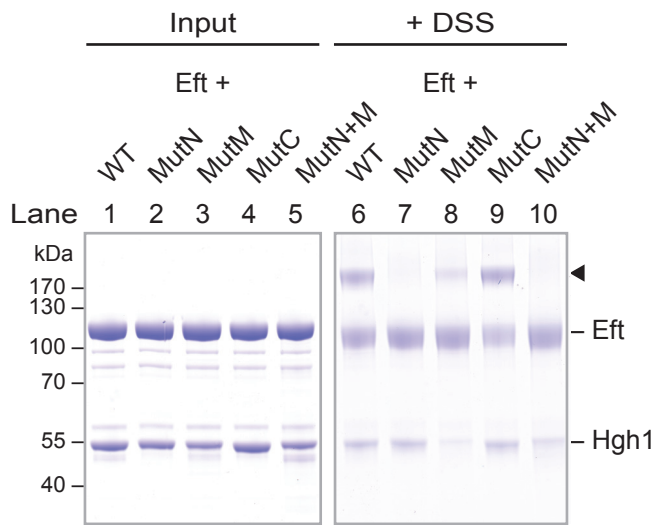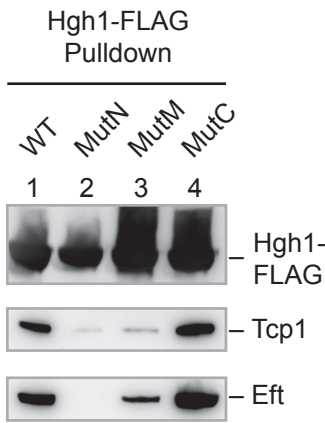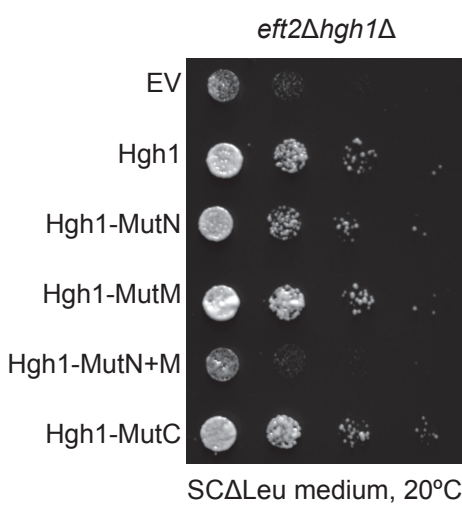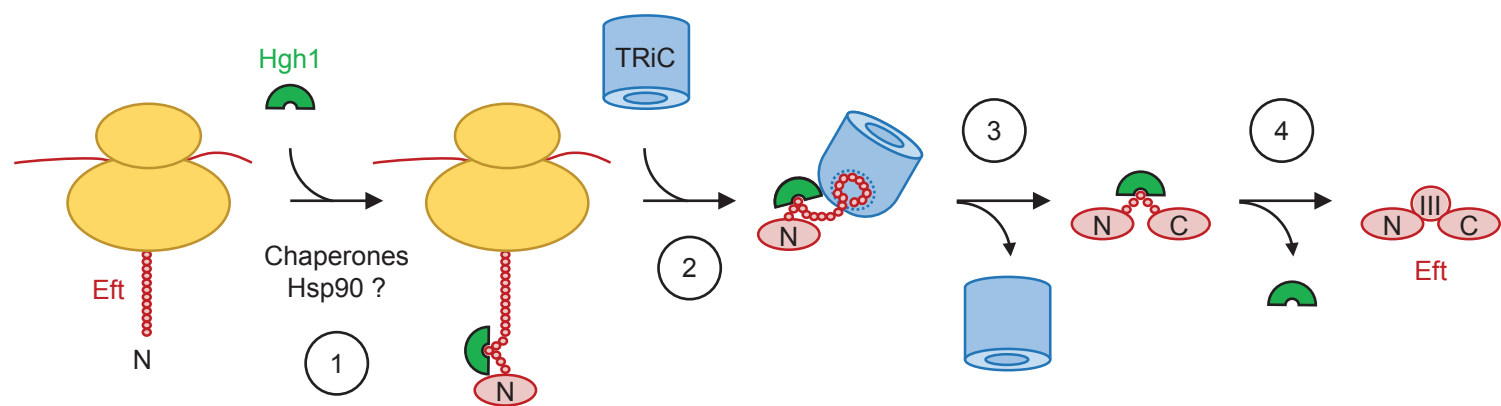
KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-Rabbit IgG (whole molecule)-Peroxidase antibody produced in goat	Sigma-Aldrich	Cat. #A6154
Goat anti-mouse IgG HRP conjugated	Dako	Cat. #P0447
Mouse monoclonal anti-FLAG	Sigma-Aldrich	Cat. #F3165
Mouse monoclonal anti-Pgk1	Life Technologies	Cat. #459250
Rabbit anti-Hgh1 antiserum	This study	N/A
Rabbit anti-Tcp1	{Behrends, 2006 #136}	–
Rabbit polyclonal anti-eEF2	Aviva	Cat. #ARP58457_P050
Rat monoclonal anti-HA High Affinity, 3F10	Merck	Cat. #ROAHAHA
Bacterial and Virus Strains		
Escherichia coli Bl21 codon+ RIL	Agilent	Cat. #NC9122855
Chemicals, Peptides, and Recombinant Proteins		
Benzonase	Novagen	Cat. #707416
Complete protease inhibitor cocktail, EDTA-free	Roche	Cat. #05056489001
Cycloheximide	Sigma-Aldrich	Cat. #C4859
Disuccinimidyl suberate (DSS) d0/d12	Creative Molecules	Cat. #001S
EheI (SfoI) restriction enzyme	Thermo Fisher Scientific	Cat. #ER0441
FLAG peptide	Sigma-Aldrich	Cat. #F3290
Geneticin selective antibiotic (G418 sulfate)	Thermo Fisher Scientific	Cat. #11811031
GMP-PNP Na salt	Jena Bioscience	Cat. #NU-401-50
Hygromycin B	Calbiochem	Cat. #400051
Kod hot start DNA polymerase	Novagen	Cat. #71086-3
Macbecin I	Tocris	Cat. #3061
NuPAGE 3-8% Tris-Acetate Protein Gels, 1.0 mm, 10-well	Invitrogen	Cat. #EA0375BOX
NuPAGE 4-12% Bis-Tris Protein Gels, 1.0 mm, 10-well	Invitrogen	Cat. #NP0321BOX
NuPAGE 4-12% Bis-Tris Protein Gels, 1.0 mm, 15-well	Invitrogen	Cat. #NP0323BOX
Pfu DNA polymerase	Promega	Cat. # M7745
Puromycin dihydrochloride	Gibco BRL	Cat. #A11138
Sordarin	Sigma-Aldrich	Cat. #S1442
T4 DNA ligase	New England Biolabs	Cat. # M0202S
TCEP	Thermo Fisher Scientific	Cat. #20491
Critical Commercial Assays		
Bio-Rad Protein Assay Dye Reagent Concentrate	Bio-Rad	Cat. #500-0006
Luminata Classico	Merck	Cat. #WBLUC0500
Deposited Data		
Coordinates and structure factors: Hgh1(1-363) Se-SAD	This study	PDB: 6HBB
Coordinates and structure factors: Hgh1(1-363), crystal form I	This study	PDB: 6HBA
Coordinates and structure factors: Hgh1(1-363) , crystal form II	This study	PDB: 6HBC

Immunoblot data: (Mendeley Data)	This study	doi:10.17632/sw9sxts73z.1
Experimental Models: Organisms/Strains		
BY4741 MATα *his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0*	EUROSCARF	http://www.euroscarf.de
BY4741 *hgh1Δ::KanMX4*	This study	yCLK14
BY4741 *eft2Δ::URA3*	This study	yCLK17
BY4741 *eft2Δ::URA3 hgh1Δ::KanMX4*	This study	yCLK19
BY4741 *hgh1*::FLAG-*hgh1-HIS3*	This study	yLM001
BY4741 *eft2*::*eft2*-HA-*hphMX4*	This study	yLM002
BY4741 *eft2*::*eft2*-HA-*hphMX4 hgh1*::*hgh1*-FLAG-*HIS3*	This study	yLM003
YPH499 *MATa ura3-52 lys2-801_amber ade2-101_ochre trp1-Δ63 his3-Δ200 leu2-Δ1*	(Sikorski & Hieter, 1989)	–
YPH499 *cct2*::*CCT2*-Int-*TRP1*	This study	yLM004
Oligonucleotides		
See Table S2	See Table S2	See Table S2
Recombinant DNA		
pProEx-HtB	Life Technologies	–
pProEx-HtB-Hgh1	This study	N/A
pProEx-HtB-Hgh1-FLAG	This study	N/A
pProEx-HtB-Hgh1mutN-FLAG	This study	N/A
pProEx-HtB-Hgh1mutM-FLAG	This study	N/A
pProEx-HtB-Hgh1mutN+M-FLAG	This study	N/A
pProEx-HtB-Hgh1mutC-FLAG	This study	N/A
pProEx-HtB-Hgh1(1-363)	This study	N/A
pProEx-HtB-Eft(1-482)	This study	N/A
pProEx-HtB-Eft(483-565)	This study	N/A
pInt	This study	N/A
p416gal	(Mumberg *et al.*, 1995)	–
p416gal-His6-TEV-Eft	This study	N/A
p416gal-His6-TEV-Eft(1-482)	This study	N/A
p416gal-Eft-HA	This study	N/A
p416gal-Eft(1-482)-HA	This study	N/A
p416gal-Eft(483-842)-HA	This study	N/A
p416gal-Eft(566-842)-HA	This study	N/A
p416gal-His6-TEV-Hgh1	This study	N/A
p415adh	(Mumberg *et al.*, 1995)	–
p415adh-Hgh1-HA	This study	N/A
p415adh-Hgh1mutN-HA	This study	N/A
p415adh-Hgh1mutM-HA	This study	N/A
p415adh-Hgh1mutN+M-HA	This study	N/A
p415adh-Hgh1mutC-HA	This study	N/A
pFA6aKanMX4	{Wach, 1994 #135}	–
pAG60	Addgene	Cat. #35128
pFA6a-6GLY-FLAG-His3MX6	Addgene	Cat. #20750
pCY 3140-02	{Young, 2012 #103}	–
Software and Algorithms		

AIDA software v4.27.039	Raytest	https://www.raytest.com/downloads/aida-imageanalyzer.html
Aimless	{Evans, 2013 #71}	http://www.ccp4.ac.uk/html/aimless.html
CCP4i	{Collaborative Computational Project, 1994 #77}	http://www.ccp4.ac.uk/ccp4i_main.php
Clustal Omega	{Sievers, 2011 #78}	https://www.ebi.ac.uk/Tools/msa/clustalo/
Coot	{Emsley, 2004 #69}	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
DSSP	{Kabsch, 1983 #110}	–
DynamX 3.0	Waters	N/A
ESPript	{Gouet, 1999 #76}	http://espript.ibcp.fr/ESPript/ESPript/
Lsqman	{Kleywegt, 1994 #73}	http://xray.bmc.uu.se/usf/lsqman_man.html
MaxQuant v1.3.0.5	(Cox & Mann, 2008)	http://www.coxdocs.org/doku.php?id=maxquant:common:download_and_installation#download_and_installation_guide
MolProbity	{Chen, 2010 #131}	https://www.phenix-online.org/documentation/reference/molprobity_tool.html
Molrep	{Vagin, 2001 #75}	http://www.ccp4.ac.uk/html/molrep.html
Perseus v1.6.0.7	{Tyanova, 2016 #139}	http://www.coxdocs.org/doku.php?id=perseus:start
Phenix	{Adams, 2010 #66}	https://www.phenix-online.org/
PLGS 3.0	Waters	N/A
Pointless	{Evans, 2006 #70}	http://www.ccp4.ac.uk/html/pointless.html
ProtScale		https://web.expasy.org/protscale
Pymol		http://www.pymol.org
Refmac5	{Murshudov, 1997 #74}	http://www.ccp4.ac.uk/html/refmac5.html
Spectra Manager v2.12.00		https://jascoinc.com/products/spectroscopy/spectroscopy-software/
Truncate	{French, 1978 #72}	http://www.ccp4.ac.uk/html/truncate.html

XDS	{Kabsch, 2010 #68}	http://xds.mpimf-heidelberg.mpg.de/
Other		
ACQUITY UPLC peptide CSH C18 column, 1.7 μm, 2.1.0 x 150 mm	Waters	Cat#186006938
ACQUITY UPLC peptide CSH Vanguard pre-column	Waters	Cat#186006939
Anti-c-MYC agarose	Pierce	Cat. #20168
Anti-DYKDDDDK G1 affinity resin	Genscript	Cat. #L00432
Anti-HA agarose resin	Pierce	Cat. #26181
Calmodulin Sepharose 4B	GE Healthcare	Cat. #17052901
Chelating Sepharose Fast Flow	GE Healthcare	Cat. #17-0575-01
DEAE-Sepharose Fast Flow	GE Healthcare	Cat. #17070901
Enzymate pepsin column	Waters	Cat#186007233
HiTrap Heparin HP 5 x 5 ml	GE Healthcare	Cat. #17040703
Index HT	Hampton Research	Cat. #HR2-134
Mono Q HR16/10	GE Healthcare	Cat. #17050601
Ni-NTA Agarose	Qiagen	Cat. #30230
Superdex 200	GE Healthcare	Cat. #28990944
Superose 6	GE Healthcare	Cat. #17517201

Supplementary Information

Chaperone Function of Hgh1 in the Biogenesis of Eukaryotic Elongation Factor 2.

Leonie Mönkemeyer, Courtney L. Klaips, David Balchin, Roman Körner, F. Ulrich Hartl and Andreas Bracher.

Max-Planck Institute of Biochemistry, Department of Cellular Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany.
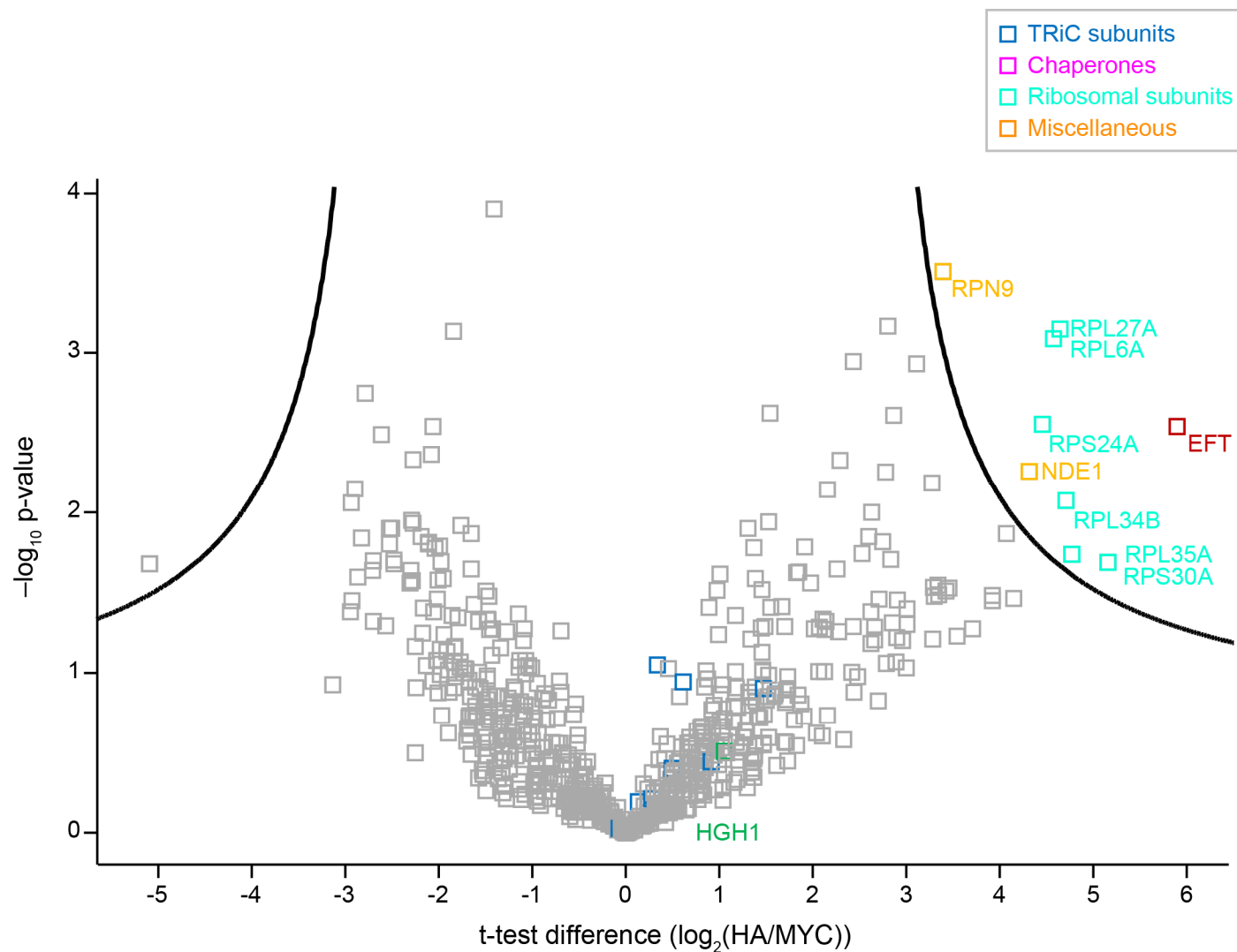
Supplementary Figures

Figure S1. Mass-spectrometric analysis of Eft-interacting proteins. Related to Figure 3. A Volcano plot representation of the Eft2-HA interactors is shown. The components of the anti-HA and anti-MYC immune precipitates from three independent experiments were analyzed by label-free quantification. The proteins to the left and right of the solid lines were significantly depleted or enriched in the anti-HA precipitates, respectively. Enriched ribosomal subunits are marked in cyan. TRiC subunits and Hgh1 are poorly enriched. These are highlighted in dark blue and green color, respectively.

A

His$_6$-Eft-N

Secondary structure composition

	Helix	Strand	Turn	Disorder
Eft structure	0.35	0.21	0.11	0.34
CD deconvolution	0.34	0.17	0.22	0.27

B

His$_6$-Eft-III

Secondary structure composition

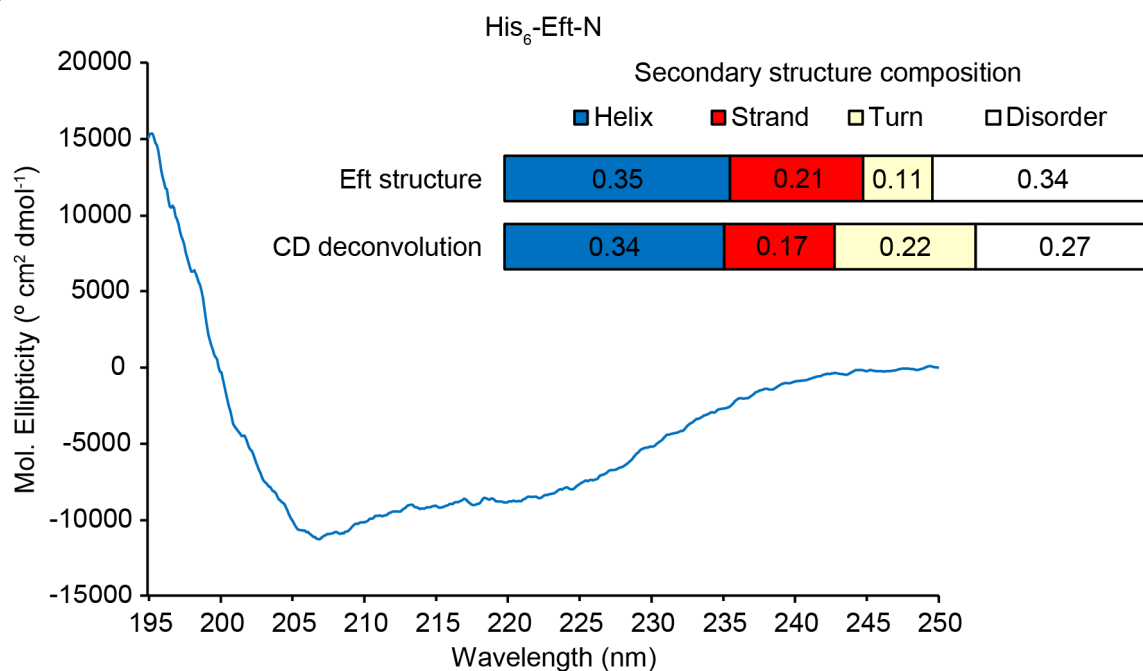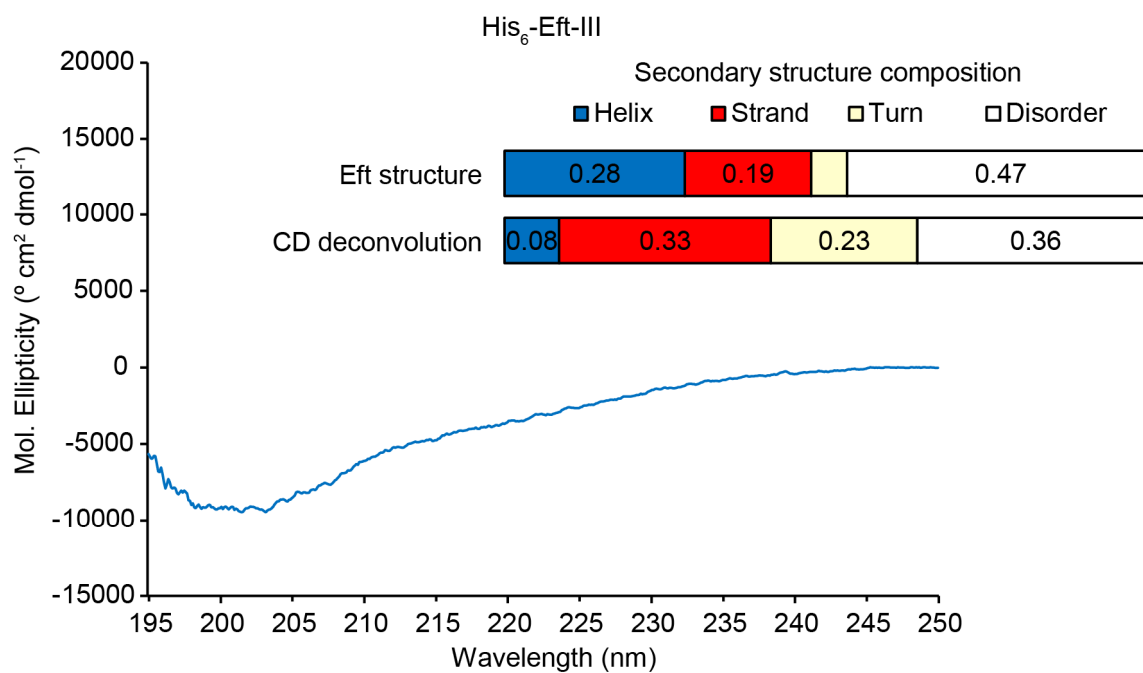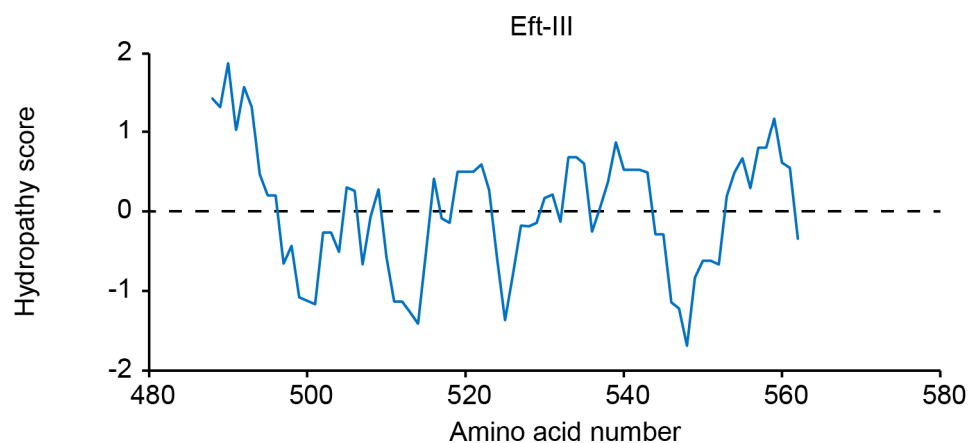|	Helix	Strand	Turn	Disorder
Eft structure	0.28	0.19		0.47
CD deconvolution	0.08	0.33	0.23	0.36

C

Eft-III

3

Figure S2. Properties of purified Eft truncation proteins. Related to Figure 5. (A) CD spectrum and secondary structure composition of His6-tagged Eft-N. The spectrum was recorded at 4°C. The diagrams show the secondary structure compositions as derived from the Eft crystal structure (pdb code 1N0U {Jørgensen, 2003 #54}) and the CD spectrum, respectively. The fractions of helix and sheet structure are based on the secondary structure assignment in the Eft crystal structure using the program DSSP {Kabsch, 1983 #110}; CD spectrum deconvolution was performed with the program CONTIN {Provencher, 1982 #115} as implemented in Spectra Manager (Jasco). (B) CD spectrum and secondary structure element composition of His6-tagged Eft-III. The spectrum was recorded at 4°C. The secondary structure assignment was as above. (C) Hydropathy plot for Eft-III. The hydropathy score for the Eft-III sequence was calculated using a gliding frame of nine residues {Kyte, 1982 #113}. Segments above the dashed line are predominantly hydrophobic. The plot was generated with the ProtScale program. (https://web.expasy.org/protscale).

A

X A M V A F T V D Q M R S L M D K V T N V R N M S V I A H V D H G K S T L T D S L V Q R A G I I S A A K A G E A R F T D T R K D E Q E R G I T I K S T A I S L Y S E M S D E D V K E I K Q K T D G N S F L I N L I D S P G H V D F S S E V T A A L R V T D G A L V V V D T I E G V C V Q
  5   10   15   20   25   30   35   40   45   50   55   60   65   70   75   80   85   90   95   100   105   110   115   120   125   130   135   140

T E T V L R Q A L G E R I K P V V V I N K V D R A L L E L Q V S K E D L Y Q T F A R T V E S V N V I V S T Y A D E V L G D V Q V Y P A R G T V A F G S G L H G W A F T I R Q F A T R Y A K K F G V D K A K M M D R L W G D S F F N P K T K K W T N K D T D A E G K P L E R A F N M F I L
     145   150   155   160   165   170   175   180   185   190   195   200   205   210   215   220   225   230   235   240   245   250   255   260   265   270   275   280

D P I F R L F T A I M N F K K D E I P V L L E K L E I V L K G D E K D L E G K A L L K V V M R K F L P A A D A L L E M I V L H L P S P V T A Q A Y R A E Q L Y E G P A D D A N C I A I K N C D P K A D L M L Y V S K M V P T S D K G R F Y A F G R V F A G T V K S G Q K V R I Q G P N Y
     285   290   295   300   305   310   315   320   325   330   335   340   345   350   355   360   365   370   375   380   385   390   395   400   405   410   415   420

V P G K K D D L F I K A I Q R V V L M M G R F V E P I D D C P A G N I I G L V G I D Q F L L K T G T L T T S E T A H N M K V M K F S V S P V V Q V A V E V K N A N D L P K L V E G L K R L S K S D P C V L T Y M S E S G E H I V A G T G E L H L E I C L Q D L E H D H A G V P L K I S P
     425   430   435   440   445   450   455   460   465   470   475   480   485   490   495   500   505   510   515   520   525   530   535   540   545   550   555   560

P V V A Y R E T V E S E S S Q T A L S K S P N K H N R I Y L K A E P I D E E V S L A I E N G I I N P R D D F K A R A R I M A D D Y G W D V T D A R K I W C F G P D G N G P N L V I D Q T K A V Q Y L H E I K D S V V A A F Q W A T K E G P I F G E E M R S V R V N I L D V T L H A D A I
     565   570   575   580   585   590   595   600   605   610   615   620   625   630   635   640   645   650   655   660   665   670   675   680   685   690   695   700

H R G G G Q I I P T M R R A T Y A G F L L A D P K I Q E P V F L V E I Q C P E Q A V G G I Y S V L N K K R G Q V V S E E Q R P G T P L F T V K A Y L P V N E S F G F T G E L R Q A T G G Q A F P Q M V F D H W S T L G S D P L D P T S K A G E I V L A A R K R H G M K E E V P G W Q E Y
     705   710   715   720   725   730   735   740   745   750   755   760   765   770   775   780   785   790   795   800   805   810   815   820   825   830   835   840

Y D K L

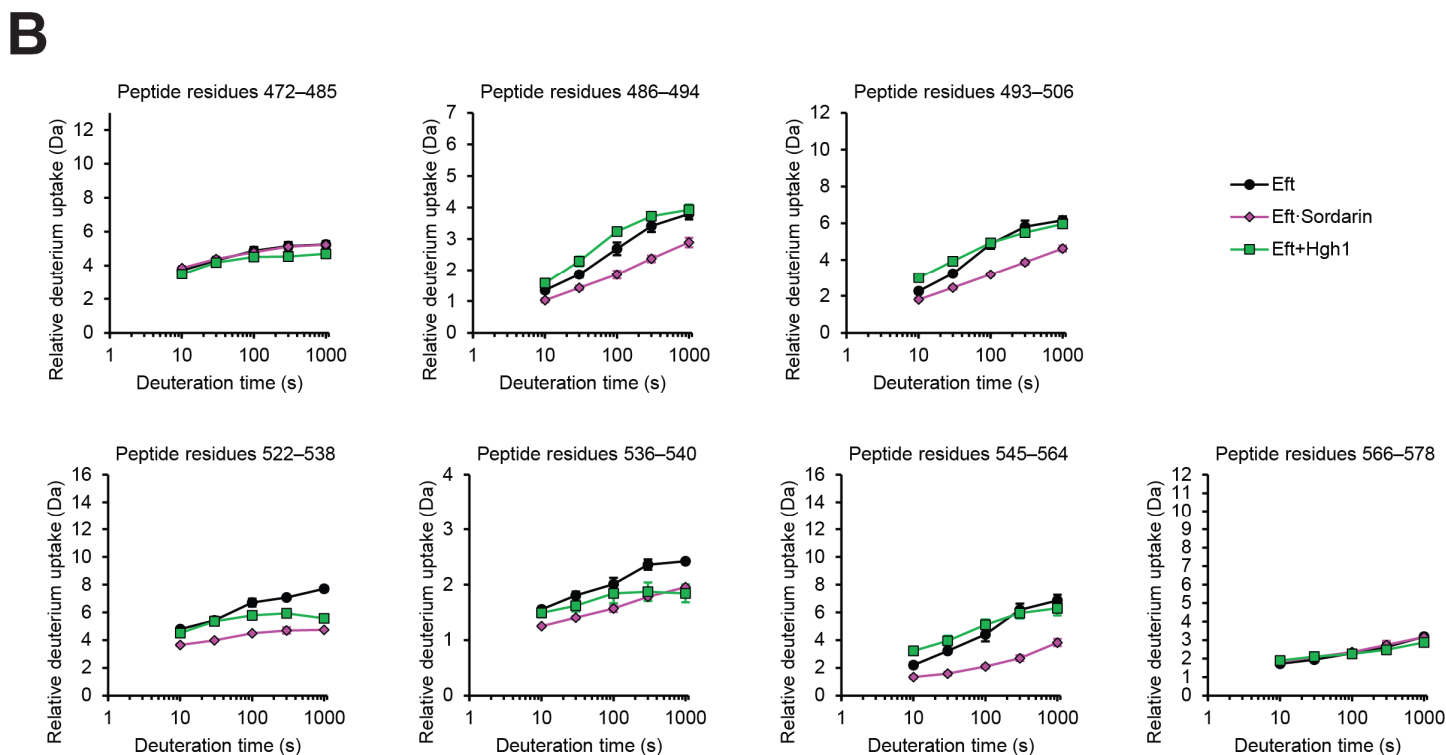Total: 256 Peptides, 87.9% Coverage, 3.44 Redundancy

B

Figure S3. H/DX Analysis of Eft. Related to Figure 5. (A) Peptide coverage of Eft in H/DX experiments. The detected pepsin fragments of Eft are shown as blue bars underneath the Eft sequence. (B) Deuterium uptake kinetics for selected peptides covering Eft domain III. Peptides for the adjacent regions, residues 472–485 and 566–578, are shown for comparison. The time traces for Eft alone, the Eft-Sordarin complex and Eft/Hgh1 are shown in black, pink and green, respectively.
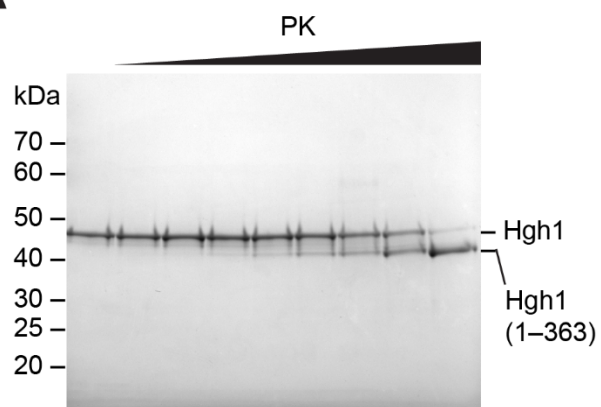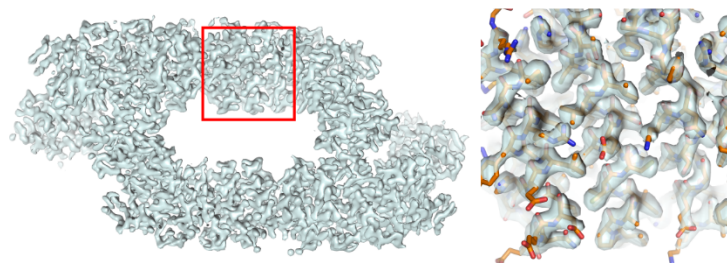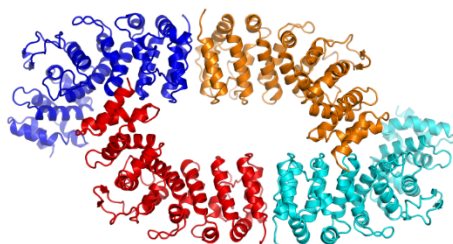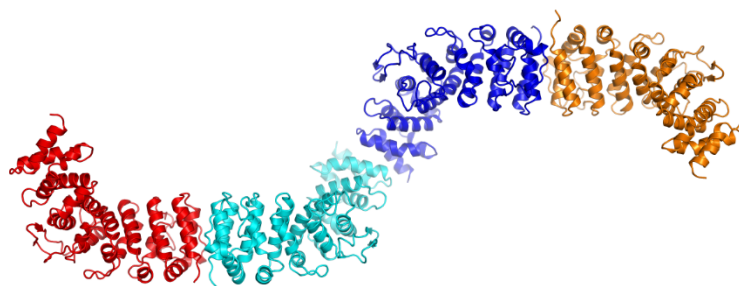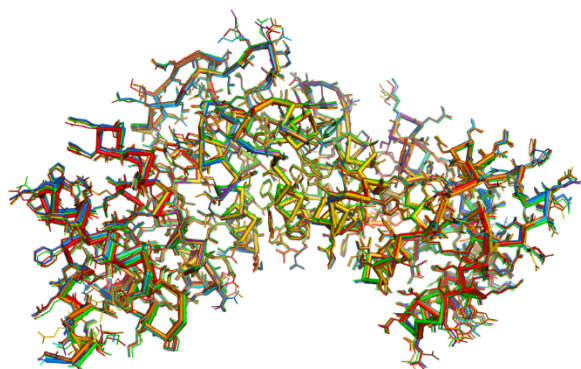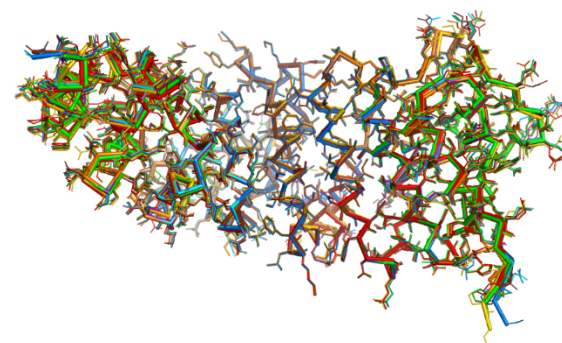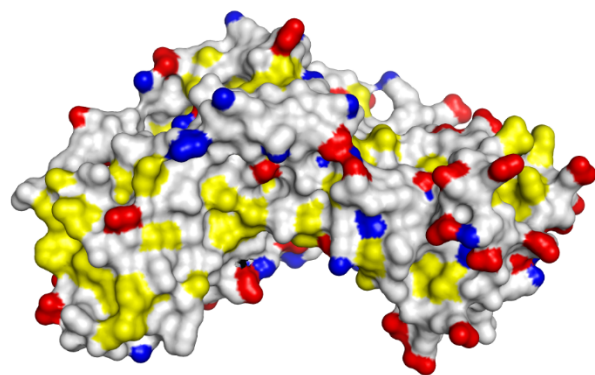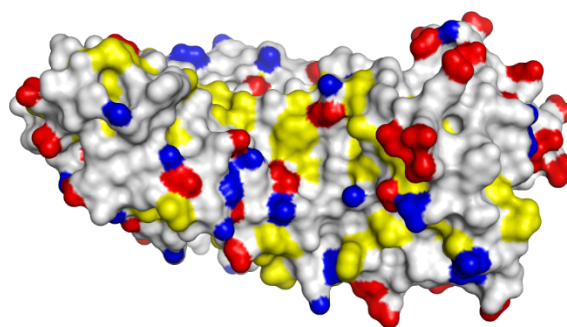
Figure S4. Crystal structure of Hgh1. Related to Figure 6. (A) Identification of structured core of Hgh1 by limited proteolysis. Hgh1 at 3 µM was treated with increasing concentrations of Proteinase-K (0, 0.04, 0.08, 0.16, 0.32, 0.64, 1.28, 2.52 and 5.12 µg mL^{-1}) for 60 min on ice. The protease reaction was stopped by addition of 4 µM phenylmethylsulfonyl fluoride. Reactions were analyzed by SDS-PAGE and Coomassie staining. The band at approximately 40 kDa in the rightmost lane was identified as Hgh1(1–363) by MS analysis. (B) Experimental electron density map. The map calculated with the Se-SAD phases after density modification is shown in isosurface representation at 1.5 σ. The final model of selenomethionine labeled Hgh1 is superposed. On the left, the complete asymmetric unit is shown. On the right, the boxed area is magnified. (C–D) Asymmetric units of the crystals of spacegroups P1 (C) and C2 (D), respectively. The protein chains are shown in ribbon representation. Each chain has a unique color. The contact between the Hgh1 C-termini is observed in both crystal forms. (E–F) Superposition of the independent Hgh1 copies in the crystal lattices. The eight molecules are shown as wireframe models. The same views as in Figure 6A are shown. (G–H) Surface properties of Hgh1. Hgh1 is shown in surface representation with hydrophobic sidechains, negatively and positively charged groups highlighted in yellow, blue and red, respectively. The same views as in Figure 6A are shown.

Multiple sequence alignment (helix annotations: h1–h20, α19, β1, β2 shown above sequences).

Block 1 (h1–h5):

```
 1 S.cerevisiae       1  MTSQLNELVEFLH.SPQPAVRQIAIDNLVGFSAGPT..SKV.FKNDSYRPIKDIIKMIMDPEHGTRVIIQQGVTILVNLS
   S.pombe            1  .MSELIELVGFLH.DQNPQVRMLAVQHLLPYTARNHPQFSI.WFHNDFEPVKDLKALLLKDK....PQIASQAVTALVNVS
   N.crassa           1  MPTELELVGFIA.SPNANIRLLATENLVPFSASQP..AL.FKTDELTPIKHLKFLIRDH....PKIAEHALTCLINLT
 2 S.moellendorffii   1  MAEELDELLMFLS.SPAPNLRAAVDIIEGLTGSDEGIAKLSTKLSI..LLPALVQLLAGN....KARSGSMSRALVNLS
   O.sativa           1  MADELDELIGFLS.SPQPNVRGAAAGIVRGLTGDADGLRALSARADR..ALPALLRLLASA.GGELGTGGAAADSLVNLS
   A.thaliana         1  MVTELELVEFLS.SPSPPVKKAAVEIVSGLTGSEEGLQSLSKYSEI..LLPSLSQLLNES....KEVSEPAAQALVNLS
 3 P.infestans        9  QVKSVEELVSFLS.HPRPELRKSASSLLNMTASDQGMFQL.LLKFKFDVVQALCRVVSDM....RPIAEDAIKALINLT
   D.discoideum       1  .MDQLPELVPFLL.EPKSEIKLLALQHLLGVSDNQEARDIL.KSTQ...IINNCIKLITD..SNHVVVRHALTILINLC
   M.brevicollis      1  MADEVSELLGFMSLDTRLDVRSAAVDLVLGLTVSQDKNVFF.KTNRE..FVPALCTIVKED..RQVVIVKDAVAALINLA
 4 D.melanogaster     1  .METVKELVQFMQPNQRLDLKAVALTHVLGTSGSEGKSAI.LSLDE..MLMAIFGLTFD...ANQTVAKDAVLSLINLT
   T.adhaerens        1  MESQIKEIAQFLTLQTRPDIKLPAIEFIAGLTGSSETNTLI.ARNTT..IIDAAVNLTKD..EIENVADAAYITLINVT
   H.sapiens         18  PEAEVVKLLPFLAPGARADLQAAAVRHVLALTGCGPGRALL.AGQAA..LLQALMELA....PASAPARDARALVNLA
   consensus>50         ...e.el...Fl.....d.....A......t.........l..d......i..a...L!Nl.
```

Block 2 (h6–β2):

```
 1 S.cerevisiae      77  EDKLVRNIILSDDK......KFLKFLVWKIVDL...TNPNADIMCILLSNLAKD.DGILAVLNIKRNSSGEEVDDGLKLA
   S.pombe           74  QNEKVRK.VLMD.D......EFLQLIFSIVTNP...LHGLADLSCMLLCNLAKE.EDFARILDMQVPLR.
   N.crassa          72  AEKDVLEYVATD.D......RFLGILLGLLVEP...EEVNANLMAMLLANLTKW.DGLKSLLDRKQPAP.
 2 S.moellendorffii  74  QDQEVGEMLLELG......AVASAMDQIGKQG...QDVDKLMVMLLVNLTTLEPGARLLCEEGDEK.
   O.sativa          77  QDGALSARLVSLG......AVVAAMDVVAKRGGEQPGLARSLVMLLANLTQVDSGVAALLQVGDEK.
   A.thaliana        74  QKCELAKKMIQMG......LIKVAMDMLYKP...ESCITRLLVMLLVNLTQLDDGVSSLLQIDDEK.
 3 P.infestans       83  AANPAACERALK.Y......DLLNRVMTQVEDS...EWRLTDYSMMLLANVTTTPEGAKALLGYDDKA.
   D.discoideum      72  QDTDMLNDIVKKN......IVPRLVDGTTDT...KNKMSEIFAMLLSNVTHTKEGCLSLMQCGKELE.
   M.brevicollis     76  HDEDLASIVIREG......IVEATTKHILDP...EETYPDPFCLLLVNLTRHNEAAERLCHHTVDE.
 4 D.melanogaster    74  SEEEAAIKVFQLAKQLQPPFAIVEVAAKEITNE...QSDLADPWSMVLSNLTRVESLVHEILDTL.
   T.adhaerens       75  GDDHAANYLCKKD......YVQSLLLYALKK...TAKYADKACMILCNLTRLSQGCDTTMEILQSQD.
   H.sapiens         90  ADPGLHETLLAADP......GLPARLMGRALDP...QWPWAEEAAAALANLSREPAPCAALMAALAAAE.
   consensus>50         .d.............d...mlL.Nlt.........ld.
```

Block 3 (h10–h14):

```
 1 S.cerevisiae     147  ALNKEVFKSLRAMDCLMDCFVKGYD.KKL.......TKY..ASFNYLAFFFADISRFKLGRMYFIEEQEYDGVVPISKLL
   S.pombe          131  ....EFSLSKNIIDQLMDLFVKGTD.HGI.......NEY..ANFDFLANVFADMTRFERGRKYFTTLQEYDHVIPASKLV
   N.crassa         130  ....EKLKSDDRVINQLMDLFVKGAD.GTY.......NKH..ADYDYLSYVFADLSKHEEIRHYFVNRQDYDDVIPINKIK
 2 S.moellendorffii 131  .......LKGLHMSRLVRLFSRTPE........EGDDAYEHVGAILVNVSRTDLGRKLMLDT..KRG...LLKQV
   O.sativa         137  .......MQGLYVAKLVRSFCRSSS.E........SE.EEDIFEHVASILVNISKVEAGRRILMEP..KRG...LLKQI
   A.thaliana       131  .......MHGLHIMKLVRSFCRSSG........ETADDQFEHVGSILVNISKTEDGRKLLLEP..KRR...LLKQI
 3 P.infestans      141  .TDTTLALRERKIRTLTNAFLEGEP.EPDGVESSTGEPKWDDEYQYVANILANISQLEQGRDFLLKMRQSTSLAG...AL
   D.discoideum     130  .........AFFIMKLVQVLTMDSNQEDY......FKS...TKNNWIVNILFNLNQ.....KLD...IFKEIL
   M.brevicollis    133  .......KPVVHELVRVFC...LADKF.....NEN...GNLDHLANVLFNITQTSEGRRLVMARTNEGC...VIQRLL
 4 D.melanogaster   137  .......RDDHTLPRLAKAFA...QLDYN.....KKK..AKLHYLAPIFCNLTQVSRGRELCCHR..KYE...LLEKLL
   T.adhaerens      133  .........GRKIFNQLLDAFF...C.SGY.....NKH...ATYHHVAAILANVTQVQAARRWILDK..DKL...IIQKCL
   H.sapiens        150  .......PADSGLERLVRALC....T.PGY.....NAR...APLHYLAPLLSNLSQRPAARAFLLDP..DRC...VVQRLL
   consensus>50         ................L...f.............d....i..#......gR.
```

Block 4 (h13–h17):

```
 1 S.cerevisiae     217  VFTEKYDAKVRREGVASTIKNSLFDSETHERLLKDEKINLLPYILLPIASAKDSEIDEEDMFNLPDELQ..LLPEDKERD
   S.pombe          197  VFTEHK.SLLRRTGVAAIIKNISFDIPFQKVLMDEEGINVLPYLLLPLAGPE.ELSEEDMDGMFDELQ..LLPDDKKRE
   N.crassa         197  VFTEHK.SDIRRKGVASIIKNAAFDIPSHPAFLDEDDVNILPYILLPIMGGE..VYPEDEALEMLTDLQ..LLPPDKQRD
 2 S.moellendorffii 186  LRQCDSKSLVRRRGVIGTIRNCCFEADTQIQNLLSLSELLWPCLILPAAGSK..IYSEEERSKMPPELAV.PLSIERQAE
   O.sativa         194  IRQSDSTNQLRKKGVASTIRNCCFEADTQIQNLLSLAEYIWPALLLPVAGSK..IYSEEDRSKMPPELAN.ALSHEREAV
   A.thaliana       187  IRQFDSTNQLRKKGVAGTIRNCCFEAKNQLQNILLISEFLWPALLLPVAGSK..TYSEQDVAKMPPELGS.ALSIEREPV
 3 P.infestans      216  LPQLKSPNNVIKRRGILGIIRNCCFDTDNHFYLYDQLDI..PTHMMVLLAGPE..DEDDDKMGMNPVVY..SQGDKKWRD
   D.discoideum     188  PLVRHE.NVIKRRGILGIIRNCCYSEQHHDYLISPE.VDILTKLCLPIRGND..KLDDDDLVGLHIDLHNSSLPIGNERD
   M.brevicollis    190  PFTQYVDSRIRRGGVIGALRNCCFEYQDHEWLLSPE.VDLLPYLLLPLAGPE..EFDDEDMDGMPDDLQ..VLPPDDKKRD
 4 D.melanogaster   194  PFASFEGSVRRGGTIGILKNVCFDTVYHDVILNEQ..SSILVAILQPLCGPE..EFSDEDNELLPIELQ..YLPESKTRE
   T.adhaerens      188  PYTQYYDNAIRRGGLIAVIKNCCFESDCNEWLLGEE.IDILPHLLLPLAGPE..ELKQEETDRLPLDLQ..YLGEDKVRE
   H.sapiens        206  PLTQYPDSSVRRGGVVGTLRNCCFEHRHHEWLLGPE.VDILPFLLLPLAGPE..DFSEEEMERLPVDLQ..YLPPDKQRE
   consensus>50         .........vrr.Gv...N.c%e......l.........p.lllp.ag......###..$..el....l..d...re
```

Block 5 (h18–α19):

```
 1 S.cerevisiae     295  PIPAITCCHLESILLLCTTHAGREYLRDKSVYPLVRELHKNVE....NEDIGELCYRIVNMLMRGEPGAGAVEEMPSK..
   S.pombe          272  PDHFIMKTLVETLVLLTATREGREHMRRRKVYPIIRELHLNVD....DEEIREVCDQLVQMLVRDEAPEELEHI......
   N.crassa         272  PDSSNILTHLETLTLLTTTRAGRDLMRSISVYPIVRETHLRVN....DEAVQDACERLVNVLMRDEAEGEEIEGGAAT..
 2 S.moellendorffii 263  PDSQIRREALDALYLLIKQEAGFRAFWELGGADVLKSGYENEE....DPEVMDKYEQLGSLVRTDTHALLFFSLHART..
   O.sativa         271  ENSEIRQQALEAIYMIVLQDEGRRAFWSVNGPRILQVGYEDEE....DPKVMEAYELIGSLLVGKGDDEQDQEQG.....
   A.thaliana       264  TDPDIRVQTLEAIYLILIEEAGRRAFWSVNGPRILQLGYEYEE....DPKAMRAYEQVGSLLVEEAGGPEESLEP.....
 3 P.infestans      290  HDRLVRLAAVDCLLLLCTTRNGRKELRRKKVYPIIRDAHLVES....DEEIGDQIYKLVDFLIRDEEGDEPDWDEVRA..
   D.discoideum     264  QDRECRKMVVESLLFLTGTKKGRVSMRTAKVYPILNFNFET.....EEELRDNVEKVVELIIRDEEGDPTPEEIEQM..
   M.brevicollis    265  MDADIRMMLLEALLQLAATRGMREQMREQNRNKVYPILREYHKWEP....LDDVRDACEQVVQLLINEEHDLA.PNLH.EVNV
 4 D.melanogaster   269  EDPDLRKMLLECLLQLCSTRRSREILRSRGVYEILREYHKWEAKVGKDSDCLLACENVVDILIKKEEEIGLDNYKTEVEV
   T.adhaerens      263  PEADLRKMLIESLMQLCATKPCRMYIREKMAYAILHELHQWDK....NEDVLSTCEKLIQVLIGDEPEAGMEDLR.NVTV
   H.sapiens        281  PDADIRKMLVEAIMLLTATRGRPQQVRDQGAYLILREHSWEP....EPDVRTACEKLIQVLIGDEPERGMENLL.EVQV
   consensus>50         .d...r....#.....l..t..gr...r.....y.i.re....e.....d.dv....e..vq.li.de.
```

Block 6:

```
 1 S.cerevisiae     369  ...NAEEEEEEESE.EEDDDEEDEIVEVA
   S.pombe          342  .................QDNPPDEDDVIVEVD
   N.crassa         346  ......A.............EEEDDIVEV.
 2 S.moellendorffii 337  ...Y..QTQTHSS.............
   O.sativa         342  ......GQKP.................
   A.thaliana       335  .........TKAES...............
 3 P.infestans      364  ...KSFAAEQETKT.......GENGAVVQPP
   D.discoideum     338  ..NKKQKLEDEDAQFET.......DEI...
   M.brevicollis    339  DPTIAKKLDEIKEIISSGQGPEVELVPEPE
 4 D.melanogaster   349  PAEQSEKFVQEDAAYVK......SLLD...
   T.adhaerens      338  PDNYSITDKSEAIVKK.........
   H.sapiens        356  PEDVEQQLQQLDCREQE..QLERELAPEPW
   consensus>50         ...........................
```

8

Figure S5. Alignment of Hgh1/Fam203 sequences. Related to Figure 6. Amino acid sequences of a representative set of Hgh1/Fam203 homologs were aligned using the EBI Clustal-Ω server {Sievers, 2011 #78}. Secondary structure elements for Hgh1 from Saccharomyces cerevisiae are indicated above the sequences. The Hgh1 helical repeat structure is indicated by the coloring of secondary structure elements. The sequences from fungi (1), plants (2), protists (3) and animals (4) are grouped separately. Similar residues are shown in red and identical residues in white on a red background. Blue frames indicate homologous regions. The consensus sequence is shown at the bottom. Asterisks below the sequence indicate mutations in Hgh1 analyzed in this study (Figure 6C). The Uniprot accession codes for the sequences are: P48362, *Saccharomyces cerevisiae*; Q10498, *Schizosaccharomyces pombe*; Q7SA64, *Neurospora crassa*; D8SRW6, *Selaginella moellendorffii*; Q338F8, *Oryza sativa*; Q9M9T2, *Arabidopsis thaliana*; D0NXG5, *Phytophthora infestans*; Q76NW7, *Dictyostelium discoideum*; A9UZU4, *Monosiga brevicollis*; Q9VBG6, *Drosophila melanogaster*; B3RJZ9, *Trichoplax adhaerens*; Q9BTY7, *Homo sapiens*.
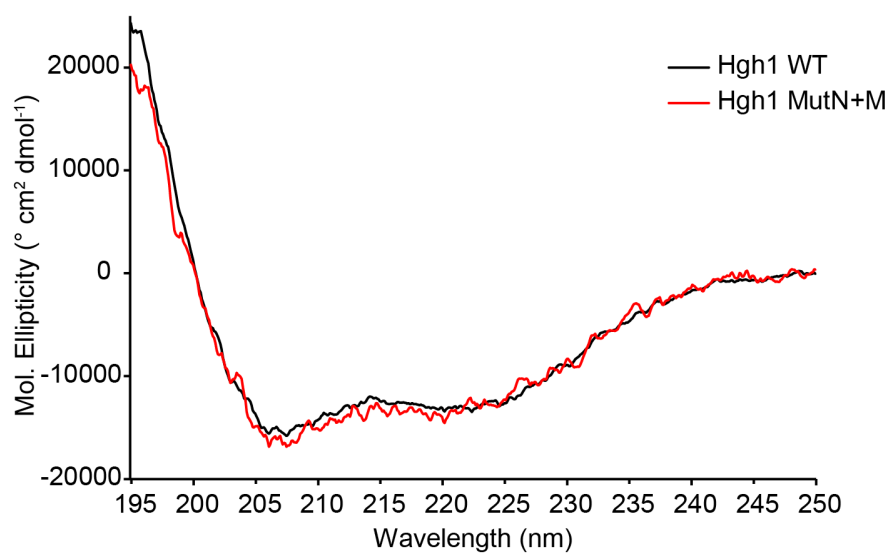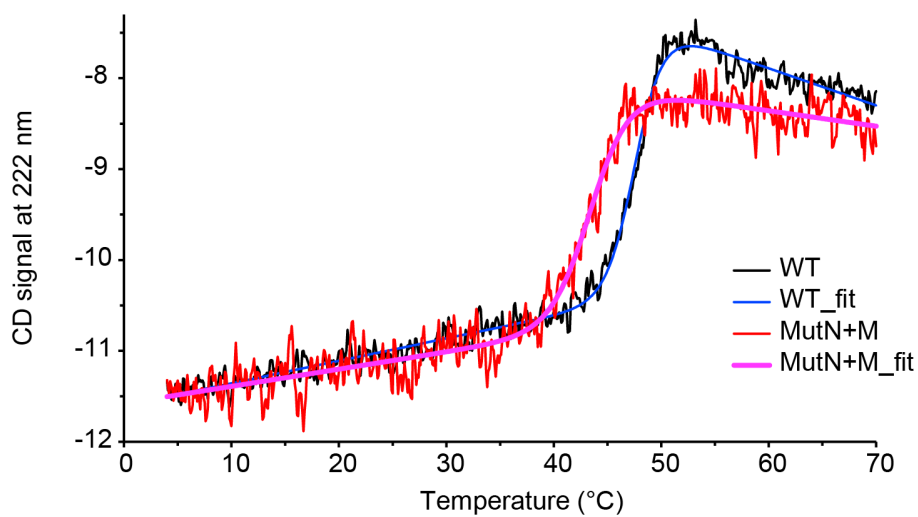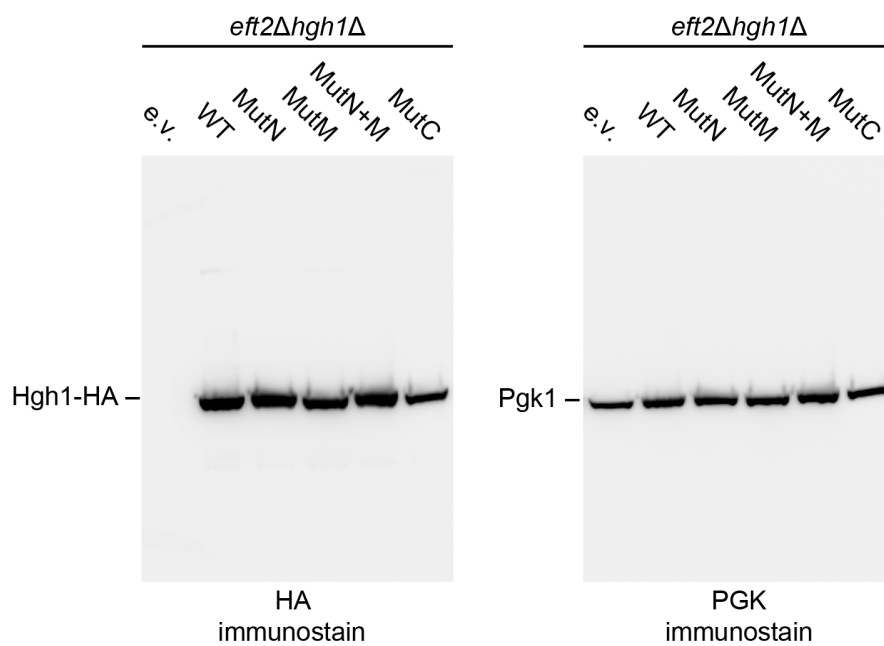
Figure S6. Analysis of the Hgh1 mutant proteins. Related to Figure 6. (A) CD spectra of wildtype Hgh1 and Hgh1(MutN+M) proteins. The spectra were recorded at 4°C. (B) Melting curves of wildtype Hgh1 and Hgh1(MutN+M) proteins. The CD signal at 222 nm wavelength was monitored. Temperature was increased at a rate of 1°C min^{-1}. The fitted values for the melting points of Hgh1 and Hgh1(MutN+M) were 47.6 and 43.5°C, respectively. (C) Complementation of growth phenotype of *eft2Δhgh1Δ* cells with mutant Hgh1 constructs. WT and mutant Hgh1 were expressed in *eft2Δhgh1Δ* cells under control of the ADH promoter. The protein level of Hgh1-HA was analyzed by immunoblotting. Phosphoglycerate kinase (Pgk1) served as a loading control.

SUPPLEMENTARY TABLES

Table S1. Crystallographic Data Collection and Refinement Statistics

Dataset	Se-SAD	Crystal form I	Crystal form II
Space group	$P1$	$P1$	$C2$
Cell dimensions			
a, b, c (Å)	80.4, 87.2, 88.7	80.2, 86.4, 88.4	191.8, 69.6, 206.9
α, β, γ (°)	118.1, 106.0, 99.5	118.45, 105.8, 99.2	90, 93.8, 90
Wavelength (Å)	0.97895	0.97895	1.00800
Resolution (Å)	44.75 – 2.7 (2.78 – 2.7)*	44.20 – 2.33 (2.37 – 2.33)	47.96 – 3.0 (3.09 – 3.0)
R_{merge}	0.193 (1.528)	0.108 (0.911)	0.104 (0.772)
$I/\sigma I$	13.3 (2.7)	11.5 (1.9)	9.8 (1.5)
Completeness (%)	99.1 (93.9)	98.8 (91.2)	99.3 (95.0)
Redundancy	13.8 (11.6)	6.9 (4.8)	3.4 (3.5)
Refinement			
Resolution (Å)	30 – 2.7	30 – 2.33	30 – 3.0
No reflections	104053	79858	52041
R_{work} / R_{free}	0.190 / 0.214	0.177 / 0.226	0.220 / 0.230
Number of atoms			
Protein	23137	23165	11489
Cl⁻ ion	1	3	-
Water	72	403	-
B-factors			
Protein	68.7	63.5	73.9
Cl⁻ ion	51.6	126.6	-
Water	53.2	50.4	-
R.m.s. deviations			
Bond length (Å)	0.004	0.003	0.005
Bond angles (°)	0.759	0.683	0.954

* Values in parenthesis for outer shell.

RESOURCE TABLES

Table S2. Oligonucleotides

Name	DNA Sequence
hgh1 KO-F	AATTAAAAATAGCTCAAAAAAATCAACAAAAAATTGAGTGcagctgaagcttcgtacgc
hgh1 KO-R	TATACAATTCTATGCTATGTGAACTGTCCTTGAAAGTGACgcataggccactagtggatctg
eft2 KO-F	ACTCAAAGACCACAAACACAAACTATAACATAATTGCAAGcagctgaagcttcgtacgc
eft2 KO-R	AACTGAAAAAGTTAAATAATTAAAAATTGTTTAACCATTCgcataggccactagtggatctg
Hgh1-F	phospho-GCCCtCCGCGGtggaATGACTTCACAATTGAATGAATTAGTGGAATTTCTGC
Hgh1-R	CCCaagcttCTAGGCCACTTCGACAATTTCATCTTCTTCG
Hgh1-FLAG-R	CCCaagcttCTActtgtcatcgtcatccttgtaatcGGCCACTTCGACAATTTCATCTTCTTCG
Hgh1mutN-F	phospho-GCCTCTAGA ATGACTTCACAATTGAATtcagcaGTGGAAgctCTGCATTCACCACAA
Hgh1mutN-HA-F	P-GCCTCTAGA ATGACTTCACAATTGAATtcagcaGTGGAAgctCTGCATTCACCACAA
Hgh1mutM-F	GTCGCTTCCACAATTgcacATTCCCTAgctGATTCTGAAACTCAT
Hgh1mutM-R	ATGAGTTTCAGAATCAGCTAGGGAATGTGCAATTGTGGAAGCGAC
Hgh1mutC-F	AGAGACAAGTCCGTTgctCCACTAGTGgctGAGCTGgctAAAAACGTTGAGAAT
Hgh1mutC-R	ATTCTCAACGTTTTTAGCCAGCTCAGCCACTAGTGGAGCAACGGACTTGTCTCT
His-Hgh1 overexpression-F	phospho-gggATGTCGTACTACCATCACCATCACCATC
Hgh1(1-363)-F	phospho-GCCCGCGGtgga ATGACTTCACAATTGAATGAATTAGTGGAATTTCTGC
Hgh1(1-363)-R	CCCaagctt CTA TTCCACTGCTCCCGCACCTGG
Hgh1-HA-F	ctag tctaga ATGACTTCACAATTGAATGAATTAGTGGAATTTCTGC
Hgh1-HA-R (p415adh)	CCCaagcttCTA AGCGTAATCTGGAACATCGTATGGGTA GGCCACTTCGACAATTTCATCTTCTTCG
His-Eft-F	phospho-gccATGGTTGCTTTCACTGTTGACCAAATGCG
His-Eft-R	CCGctcgagTTACAATTTGTCGTAATATTCTTGCCAGCCTGG
His-Eft(1-482)-R	CCG CTCGAG TTA TTTCATGACCTTCATGTTGTGAGCAGTTTCAC
Eft-T56D-F	ggtgaagctcgtttcGACgataccagaaaggat
Eft-T56D-R	atcctttctggtatcGTCgaaacgagcttcacc
Eft-HA-F	phospho-ggg ATGGTTGCTTTCACTGTTGACCAAATGCG
Eft-HA-R	CCGCTCGAGTTAAGCGTAATCTGGAACATCGTATGGGTACAATTTGTCGTAATATTCTTGCCAGCCTGG
gen_Hgh1-FLAG-F	GAAGAAGATGACGACGAAGAAGATGAAATTGTCGAAGTGGCC GGGGGAGGCGGGGGTGGA
gen_Hgh1-FLAG-R	GTATTATACAATTCTATGCTATGTGAACTGTCCTTGAAAGTGAC GAATTCGAGCTCGTTTAAAC
gen_Eft2-HA-F	GGTATGAAGGAAGAAGTTCCAGGCTGGCAAGAATATTACGACAAATTG GGTGACGGTGCTGGTTTA
gen_Eft2-HA-R	CGACAAAAACTGAAAAAGTTAAATAATTAAAAATTGTTTAACCATTC ATCGATGAATTCGAGCTCG
CCT2int-F	GACCAAGAACCGCCAATAGACAACATATG TCCATGGAAAAGAGAAG
CCT2int-R	GGCGGTTTATTATTTTCGAGGTATATCCTAAATATG TACGACTCACTATAGGG
Eft2forw483	phospho-gccatgTTCTCTGTCTCTCCAGTTGTGCAAGTCGC
Eft2rev565new	gggggGCTCGAGttaTTCTCTGTAAGCGACAACTGGTGGGG
Eft482HA-R	CCGCCG CTCGAG TTA AGCGTAATCTGGAACATCGTATGGGTA TTTCATGACCTTCATGTTGTGAGCAGTTTCAC
Eft2forw566	phospho-gccatgACTGTTGAAAGTGAATCTTCTCAAACTGCTTTGTCC

CCT2N-F	cgc ggatcc ATGAGTGTACAGATATTTGGAGACCAGGTTAC
CCT2N-R	aagaaaaaa gcggccgc c ACCAGCTTTGCAACCACTAAACTTCAAGAATG
CCT2C-F	ccg ctcgag c GAAGCTTGTACCATCGTCCTAAGAGGT
CCT2C-R	aaaa ctgcag TTACATATGTTGTCTATTGGCGGTTCTTGGTC
DNA1	TTAATTAAGGCGCCCCATATGGGATCCGGTGATCGATCTGCGGCCGCCTCCGGAATTCAAG CAGGCCTCGGTAAGACGAGCTCAACTAGTCAAGGGTCGACTGGGAGCGGGAGCGGATGGA GCCACCCGCAGTTCGAAAAAGGTTCTGGTAAGAGAAGATGGAAAAAGAATTTCATAGCCG TCTCAGCAGCCAACCGCTTTAAGAAAATCTCATCCTCCGGGGCACTTGGGTCCGGGCATC ATCACCACCATCATCACCACGGCTCGGGGTCTCGAGGAGCATGCGGTACCTCTAGATTCG AAGCTAGCTAGCTGCAGGGAACGATCATTCACTATATATATATCAATTTATATATACGTA TGTGTAATTGAAGAAAGATACGTTTTTTCCTCTATTGAGAGGCCTGCTGGATGAATAGCT TTACCTTTTCTAAATCCTTGATACCATCAGTCTCTACTCCTCCACTTACATCGACACCAA TCGCATTTGGTAACATATTAATGGCAACAGAAACGTTATCAGGATTCAATCCACCAGCGA TAATGAATTTTATCTCGGGATGACTTGCAGACCAACTGGAAATTGCACTCCAATTCAATT TCTCACCAGTGCCACCTTCACCAGAATCGAACAACGTCAGCACATTGTCTACGTGTTCAT ACAGGTCCAGTAGTAATTCACAATCCTGTGGGAACTGGAACCTCTTAATGATTGGAATTG AAGATGGGATCAAAGATCTGTATTCTTTAATATCTTCATCTCCATGTAATTGTATCACAT CCAAATTATATTCGTGGTACAGTTGAAGGACATCATCAACGGACTGATTTCTAAACACCC CGACCAATTTAGTACCTTTCACGTTCTCTTGTTGGTGAACAGCAGTTGAAATACCTTTCG CAACAGATGAGTCAATGGTTCTTTTCCTACCGGGAACACAAATGATACCTAAGTAATCAG CACCATCATCCACAGCAGTCTTTGCAGCTTCAACGGTTTGCAAACCACACACTTTAACGA GCATCACAGTAGATTTCGCAAGAGGGTAGCTCGCTCAGAGTACCCAAGTAAATGATTAGT AAACTGATGTTTGATAGTTCAATTTTTCAATGAAATAACCTTATATTAAAATTGATATTA CTATTATACAAAAATAAAGAATAAAGGATTTGAGTTTATACATAAAATACCATTATTATT TGTTCAGTGAGAGATACCGGGGGTATATGGGATGTGTGTAGTGATACCATGCAATCATGTA TCAAACATGGGCCCGGTACCCAATTCGCCCTATAGTGAGTCGTAAAGCTTGGCGCGCC

5 Discussion

The eukaryotic chaperonin TRiC is likely the most complex molecular chaperone known, which makes it difficult to define its exact mechanism and its critical role in the proteostasis network of the eukaryotic cell. Research over the past decades provided insight into many aspects of chaperonin structure and function and described the interactions of TRiC with substrates and other components of the proteostasis network (Dekker et al., 2011; Dekker et al., 2008; Douglas et al., 2011; Joachimiak et al., 2014; Munoz et al., 2011; Reissmann et al., 2007; Yam et al., 2008). The present work reveals the correct TRiC subunit topology by applying a new approach combining chemical crosslinking and mass spectrometry (Leitner et al., 2012). Moreover we provided experimental evidence for partial encapsulation of over-sized substrates by TRiC, explaining how folding of large multi-domain proteins exceeding the cavity size can be assisted by the eukaryotic chaperonin (Russmann et al., 2012). Finally, we elucidated the structure and function of Hgh1, which – in collaboration with TRiC – acts as a specialized chaperone in the biogenesis of the abundant elongation factor eEF2.

5.1 The subunit topology of the TRiC complex

The high structural similarity of the eight paralogous TRiC subunits posed a problem in defining the arrangement of the subunits in the complex. Intermediate resolution (5.5 – 3.8 Å) structures of the TRiC complex were available (Cong et al., 2010; Dekker et al., 2011; Munoz et al., 2011), revealing the tertiary structure, but the details in side chain density required for unequivocal identification of individual subunits were missing. We solved this problem by an orthogonal approach of combining chemical crosslinking (XL) with mass spectrometry (MS) analysis, which has been developed over the last decade (Leitner et al., 2016; Liu and Heck, 2015). The XL-MS approach provided residue-specific information on crosslinked peptides, which could be used to define the subunit topology based on the structural model

of the paralogous thermosome complex and known distance restraints for the employed crosslinker (Shomura et al., 2004). Two independent and self-consistent data sets for yeast and bovine TRiC were obtained, which conclusively provided an unambiguous assignment of the TRiC subunit topology.
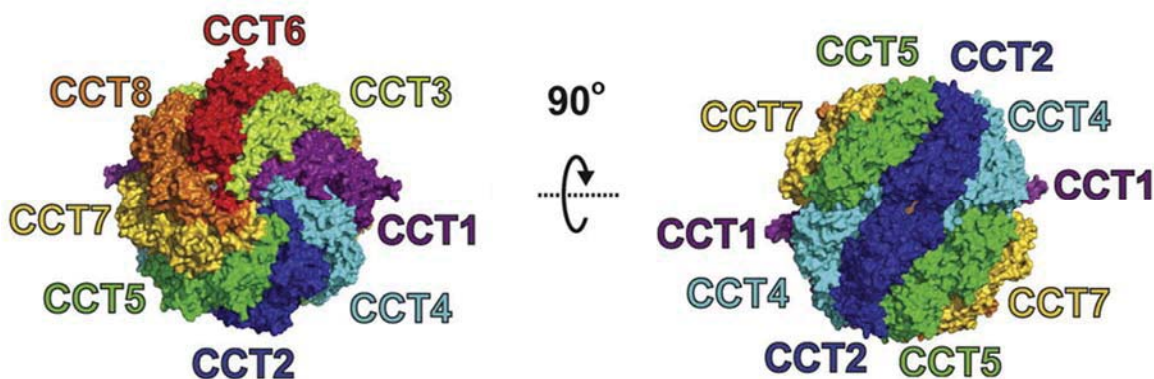

Figure 13: Subunit order of the TRiC complex according to Leitner et al. Left panel shows top view of closed TRiC structure indicating the subunit order in each ring: CCT 6-8-7-5-2-4-1-3. In the right panel, the complex is turned by 90° to show the homotypic contact of CCT2 subunits across the rings. The second homotypic contact involving subunits CCT6 is hidden on the back of this representation. Overall, the TRiC structure has 2-fold symmetry. Figure modified from (Leitner et al., 2012)

According to our model, the subunit order in each ring is CCT 6-8-7-5-2-4-1-3. Moreover, the TRiC complex exhibits two homotypic contacts across its equator (CCT2 and CCT6) and has overall two-fold symmetry (Figure 13), consistent with available cryo-EM (Cong et al., 2010) and crystallographic data (Dekker et al., 2011). Surprisingly, our subunit arrangement is very different from all previously proposed models and these earlier models in turn are entirely incompatible with our crosslinking data. However, our model is corroborated by previous crosslinking data (Cong et al., 2010) and the subunit spacing derived from cryo-EM reconstitutions (Martin-Benito et al., 2007). This shows that previous data were simply insufficient for a definitive topology assignment. Importantly, other studies independently confirmed the subunit arrangement in the model of TRiC (Herzog et al., 2012; Kalisman et al., 2012; Kalisman et al., 2013; Zang et al., 2016). Moreover, because of its well-defined topology, the TRiC complex has become a model system for optimizing crosslinking procedures and protocols (Leitner et al., 2014; Walzthoeni et al., 2015).

5.1.1 The TRiC complex exhibits striking functional asymmetry

The current model reveals a striking segregation of positively and negatively charged surfaces within the folding chamber of TRiC. This bipolar charge distribution is highly conserved in the individual TRiC subunit sequences across the eukaryotic tree. Positive and negative surface charges are contributed by the subunits CCT5-2-4 and CCT3-6-8, respectively. This coincides with the two-fold symmetry axis generated by the homotypic contacts in subunits CCT2 and CCT6 (Figure 14). Hence, the entire particle is formed by two hemispheres with segregated surface charges arranged in phase. The subunit surfaces at the junction of the two hemispheres show less conservation, which suggests increased interspecies variations. The charge separation inside the cavity might be functionally important in affecting folding trajectories. Indeed, the negative charges inside the GroEL folding cavity were shown to be relevant for productive folding (Tang et al., 2008). Although mutant GroEL with zero net charge inside the cavity surface was able to bind and encapsulate a model protein like wildtype chaperonin, there was no acceleration of folding (Chakraborty et al., 2010; Tang et al., 2006).

A second layer of asymmetry becomes apparent when analyzing the ATP binding affinity of the different TRiC subunits. Systematic mutation of residues important for ATP binding and hydrolysis in the different TRiC subunits suggested that these activities are critical in CCT1, CCT2, CCT4 and CCT7 but not needed in subunits CCT6 and CCT8 for survival of budding yeast under stress conditions (Amit et al., 2010). Apparently, only four of the eight subunits in TRiC bind and hydrolyze ATP under physiological conditions (Jiang et al., 2011; Reissmann et al., 2012). The subunits CCT5 and CCT4 were found to have the highest ATP hydrolysis activity, while the subunits CCT2 and CCT1 showed intermediate ATP hydrolyzing activity. Crosslinking with a photo-activated ATP analog showed that subunits CCT3-6-8-7 are poor ATP binders (Reissmann et al., 2012). Strikingly, the subunits with high and low ATP hydrolysis activity, CCT5-2-4-1 and CCT3-6-8-7, form contiguous, juxtaposed blocks in the rings of the present TRiC model, similarly to the segregation of charges (Figure 14). High- and low-affinity hemispheres are mostly in register between the

rings, staggered only by one subunit. The asymmetric ATP utilization of the different TRiC subunits is conserved from yeast to mammals (Reissmann et al., 2012).
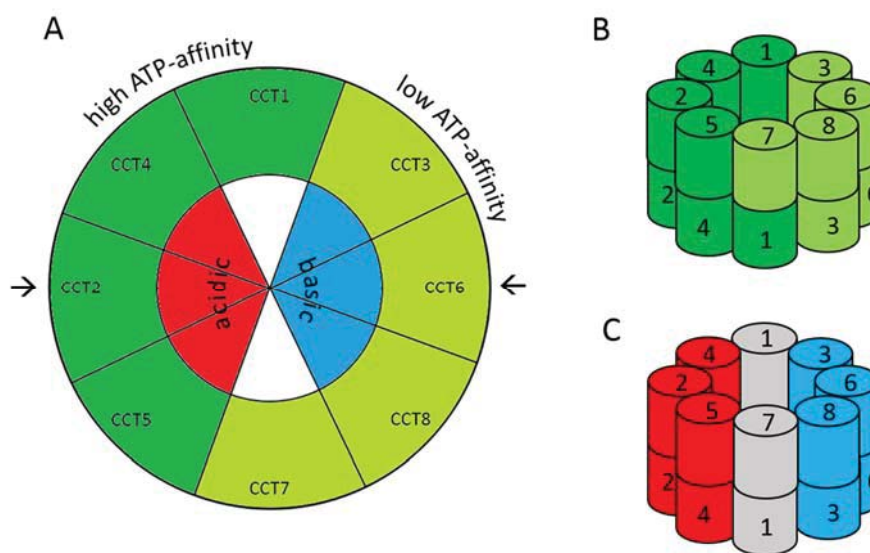

Figure 14: Asymmetry in the TRiC chaperonin. A) Schematic representation showing subunits with high and low ATP-affinity in dark and light green, respectively, and subunits with acidic and basic surface facing the cavity in red and blue, respectively. Arrows indicate homotypic symmetry contacts of the TRiC subunit arrangement. B+C) Schematic representation showing the hemispheres according to ATP-affinity (B) and surface charge (C) in the context of the assembled two-ring chaperonin. Panel A modified from (Skjaerven et al., 2015).

Kalisman et al. suggested that the low ATP-affinity hemisphere might be primarily involved in substrate binding (Kalisman et al., 2013). Relying on a computational algorithm analyzing crystallographic R-values, they assign residual densities in the crystal structures of open and closed TRiC as bound tubulin and actin, respectively, which contact the TRiC subunits CCT3-6-8 and hence the lobe with low ATP affinity. They conclude that TRiC is partitioned into an ATP-binding and a substrate binding lobe (Kalisman et al., 2013).

In the meantime, however, Joachimiak et al. analyzed several substrates binding to TRiC apical domains and found many of them interacting with the subunit CCT2 (Joachimiak et al., 2014). Furthermore, the Box1 motif of VHL was earlier described to bind to the apical domain of CCT1 (Spiess et al., 2006). Binding

of Actin to TRiC was analyzed in detail by Balchin et al. and revealed initial binding of non-native actin to the consecutive TRiC subunits CCT4-2-5-7-8, comprising subunits of the high and low ATP-affinity hemispheres (Balchin et al., 2018). An exclusive distinction of ATP and substrate binding lobes therefore seems unlikely.

The TRiC subunit topology is the basis for the segregation of enzyme activity and surface properties in the particle. The hetero-oligomeric nature generates chemical and functional asymmetries absent in other chaperonins. Possibly, this may be the source for the unique ability of TRiC to fold substrates which cannot be folded by any other chaperonin or chaperone system (Tian et al., 1995).

5.1.2 Implications for TRiC function

The hetero-oligomeric nature of TRiC distinguishes it from other chaperonins, enabling the described asymmetries in ATP utilization and net surface charge distribution, which are likely fundamental components of its folding mechanism. We can now integrate these important findings into an improved model for the functional cycle of TRiC (Figure 15).
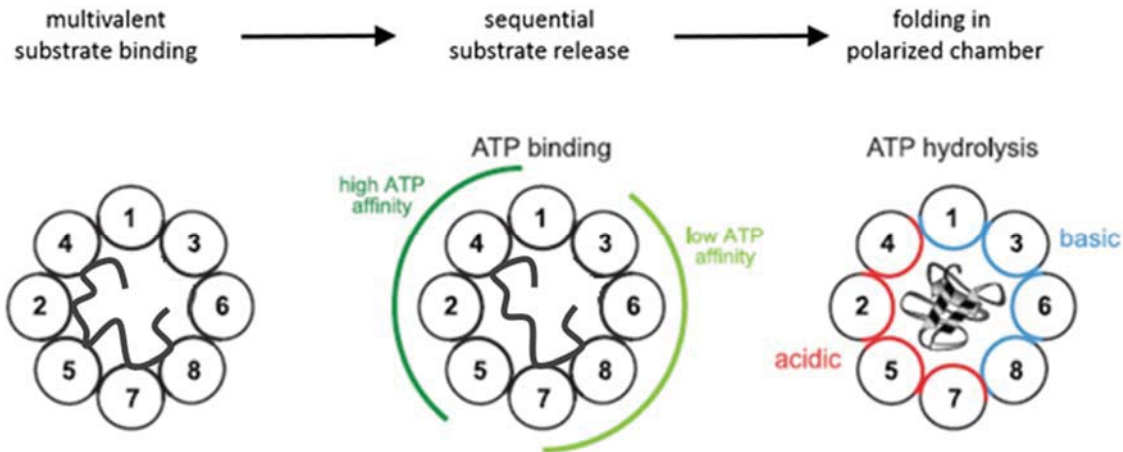
Figure 15: TRiC folding mechanism. A non-native substrate (e.g. actin) binds asymmetrically to selective TRiC apical domains (CCT4-2-5-7-8) thereby adopting a distinct topology. Differential ATP binding affinities in the individual TRiC subunits generate an asymmetric power stroke, which trigger sequential ring closure concomitantly with sequential substrate release (from subunits CCT2 and CCT5). After full encapsulation, the substrate folds to completion guided by a charge-segregated asymmetric folding chamber. Figure modified from (Balchin et al., 2016) with information about actin folding from (Balchin et al., 2018)

The process is initiated by the binding of a folding substrate to TRiC in a defined topology. The eight paralogous TRiC subunits differ in their substrate recognition properties (Joachimiak et al., 2014). Specifically, the apical domains provide a modular set of binding surfaces integrating polar and hydrophobic contributions that allow a combinatorial recognition of substrate polypeptides. Most probably, multiple low-affinity contacts to different apical domains are necessary for stable substrate binding. This is in agreement with early cryo-EM reconstitutions of TRiC-substrate complexes (Llorca et al., 2000; Llorca et al., 1999a) and with the recent analysis of the TRiC-actin complexes by hydrogen-deuterium exchange (H/D-X) (Balchin et al., 2018). The distribution of subunit-specific substrate interactions imposes a global topology on the TRiC-bound polypeptide.

After stable substrate binding, the asymmetric ATP utilization of the subunits in the hetero-oligomeric ring initiates a highly asymmetric power stroke (Reissmann et al., 2012). ATP hydrolysis is initiated in the high-ATP-affinity lobe and then propagates through the entire ring, suggesting a sequential or partial ring

closure mechanism, which was already proposed based on earlier observations (Lin and Sherman, 1997; Rivenzon-Segal et al., 2005). The subunit CCT4 is most sensitive towards mutations of its ATP binding pocket and was assigned the highest ATP affinity (Reissmann et al., 2012). This CCT4 subunit may therefore initiate the power stroke. Of note, in our newly refined model of the TRiC structure, the N-termini of the CCT4 subunits are located outside the cavity (Leitner et al., 2012). The N- and C-termini of other subunits are forming mixed four-stranded ß-sheets lining the inner wall of the cavity. These ß-sheets also involve the proximal sensor loop, which communicates with the respective ATP-binding site at its base (Munoz et al., 2011). It is conceivable that ATP-induced conformational changes are transmitted through this tight interaction network and that the deviation at subunit CCT4 may impose the direction of the propagation of ring closure (CCT4→CCT1→CCT3→…). CCT4 is the direct neighbor of CCT2, which was shown to bear the most hydrophobic substrate-binding site and may serve as a general anchor side for unfolded substrates, while specificity is provided by additional TRiC interactions (Joachimiak et al., 2014). Initiation of ring closure at the subunit CCT4 would cause its substrate-release-loop to insert into the substrate binding site of CCT2 and eject the substrate bound there (Douglas et al., 2011). Thereby, the most hydrophobic part of the bound substrate would be released first. Indeed, this is in agreement with the recent findings about TRiC-mediated actin folding, which show a partial release of actin from subunits CCT2 and CCT5 upon ATP binding (Balchin et al., 2018).

Upon ATP-induced lid-closure, the substrate is released into the cavity (Douglas et al., 2011). As ring-closure probably proceeds sequentially, it is likely that also the individual TRiC-substrate interactions are disrupted sequentially. This may allow certain subdomains to fold, while others are still bound to the chaperonin, thereby critically influencing the folding pathway of the substrate. As the binding mode to TRiC for a given substrate is likely very specific, substrate encapsulation becomes a highly defined and customized process. This is further emphasized by the charge segregation inside the folding chamber,

which provides different environmental effects on the folding trajectory depending on when and where a substrate is released (Leitner et al., 2012).

We only begin to understand the manifold levels at which TRiC can promote productive folding for structurally diverse and topologically complex substrates. However, revealing TRiC's inherent asymmetric properties decisively advanced our understanding of its complex folding mechanism.

5.2   Multi-domain protein folding

An essential aspect of TRiC-dependent protein folding is transient substrate encapsulation inside the central cavity (Reissmann et al., 2007). However, the volume of this cavity is limited. Yet, several large substrates exist which clearly exceed the capacity of the chamber (Spiess et al., 2004). A possible solution for TRiC-assisted folding of these large substrates is partial encapsulation of TRiC-dependent domains.

Domain-wise protein folding is of special importance for eukaryotic cells, in which roughly 70% of all proteins are multi-domain proteins (Han et al., 2007). A typical domain spans 50-300 amino acids (Netzer and Hartl, 1998), which fit the cavity size very well in case of a globular fold. Domain-wise protein folding is probably facilitated by vectorial protein synthesis at the ribosome. While compact domains may fold spontaneously, especially long-lived domain folding intermediates require shielding from aberrant interactions provided by chaperones (Frydman et al., 1994). Our results suggest that the eukaryotic chaperonin TRiC plays an important role in multi-domain protein folding. We demonstrate that with model substrates such as fusion proteins of actin and green fluorescent proteins and the natural multi-domain substrate hSnu114 (109 kDa), encapsulation of selected domains by TRiC occurs while the remainder of the protein protrudes through the oculus of the iris-like lid (Russmann et al., 2012).

5.2.1 Partial encapsulation by sequential closure of the iris-like lid

TRiC belongs to the group II chaperonins and therefore has an in-built shutter mechanism formed by helical protrusions, which adopt an iris-like lid structure upon ATP-hydrolysis. Furthermore, the asymmetric ATP utilization in the hetero-oligomeric TRiC rings strongly supports a sequential mechanism of cavity closure (Lin and Sherman, 1997; Reissmann et al., 2012; Rivenzon-Segal et al., 2005). These features are well compatible with a process of partial encapsulation: The apical domains in the TRiC subunits move sequentially towards the central axis, thereby enclosing a selected (sub)-domain, while an inter-domain linker polypeptide threads the apical pore. A similar scenario would not be possible for a system using a detachable GroES-like lid.

The precise boundaries of the encapsulated fragment do not seem to be strictly defined. Protease protection experiments with the natural substrate hSnu114 revealed multiple encapsulation products with sizes between 15 kDa and 45 kDa. Also for the destabilized ΔN-GFP-actin fusion proteins, heterogeneity of the TRiC-encapsulated fragments was observed. The segment being ultimately encapsulated is probably the result of a complex interplay between conformational dynamics of the TRiC complex and of the substrate. The TRiC complex displays a high degree of structural flexibility not only in its open state (Munoz et al., 2011) but probably also along its transition to the well-defined closed state. Cryo-EM reconstitutions of several intermediates in the conformational cycle of TRiC suggest highly independent movements of individual TRiC subunits (Cong et al., 2012; Zang et al., 2016). Furthermore, the not-yet folded substrate is assumed to be structurally highly dynamic. Depending on the accessibility of binding motifs on TRiC and on the substrate at any given time, the topology of the bound segment may change. Beyond that, the segments of the multi-domain protein, which do not require TRiC for folding, will proceed to their folded states thereby affecting the flexibility of the inter-domain linkers, which has to thread through the oculus of the lid for successful partial encapsulation. All these factors may contribute to heterogeneity in the degree of encapsulation. Importantly, this does not mean that random

segments will be encapsulated, but that a range of defined encapsulation products is possible. Partial encapsulation requires the apical pore to exhibit a certain degree of flexibility. In a partial encapsulation event, the iris may not always close completely, but an external folded domain may act as a plug to seal the cavity enough to prevent protease access to the encapsulated segment. Beyond this, the protease-sensitive helical protrusions of TRiC subunits (Reissmann et al., 2007) are protected from digestion suggesting that the individual subunits adopt a conformation similar to the closed structures.

5.2.2 Positional effects on TRiC-dependent domains

Partial encapsulation by TRiC may occur post-translationally or co-translationally, depending on the position of the TRiC-dependent domain in the substrate. Co-translational folding of an N-terminal domain by TRiC is supported by reports on binding of TRiC to nascent chains (Etchells et al., 2005; Frydman et al., 1994). The encapsulation of a domain surrounded by other domains or located at the C-terminus seems to be strongly influenced by the domain context. Although all encapsulation products of hSnu144 were shown to contain the C-terminal peptide, the fragment length differed depending on whether the entire protein or the fragment containing the C-terminal domains III-VI was processed. Moreover, the efficiency of actin-folding was dependent on the position of the actin segment in the fusion protein: The actin segment at the C-terminal end folded more efficiently than at the N-terminal end. Our results suggest that the N-terminal protein segments significantly affect the encapsulation process. A folded N-terminal domain may restrict the flexibility of TRiC-substrate complexes, thereby facilitating the selective binding of the segment containing the TRiC-dependent domain. Beyond this, a folded N-terminal domain may serve as a scaffold for progressive assembly of further parts of a multi-domain protein. Of note, folding of actin failed when it was placed between TRiC-independent GFP domains in the fusion protein. This suggest that selection of internal domains for partial encapsulation is difficult. If such substrates exist, internal

TRiC-dependent domains would require highly specialized inter-domain linkers, which would need to simultaneously traverse the narrow entrance pore of the folding chamber.

5.2.3 The folding of the natural substrate hSnu114

The 109 kDa protein hSnu114, alternatively referred to as snRPN116 (Yam et al., 2008), is a regulatory subunit of the human spliceosome (Frazer et al., 2008). It displays high sequence and structural homology to eukaryotic elongation factor 2 (eEF2) (Fabrizio et al., 1997). Recently, high-resolution cryoEM structures of the spliceosome became available, which show the yeast homolog of hSnu114, Snu114 (Yan et al., 2015). hSnu114 has a complex nested multi-domain structure (Figure 16) (Jorgensen et al., 2003). In contrast to eEF2, hSnu114 has an N-terminal extension of 112 residues, which was predicted to be unstructured outside of the context of the spliceosome. Initially identified as a TRiC interactor in a high-throughput screen (Yam et al., 2008), stable binding of hSnu114 to TRiC in rabbit reticulocyte lysate was confirmed. Next, the folding behavior of hSnu114 was analyzed in detail by applying combinations of TRiC encapsulation and protease protection assays.
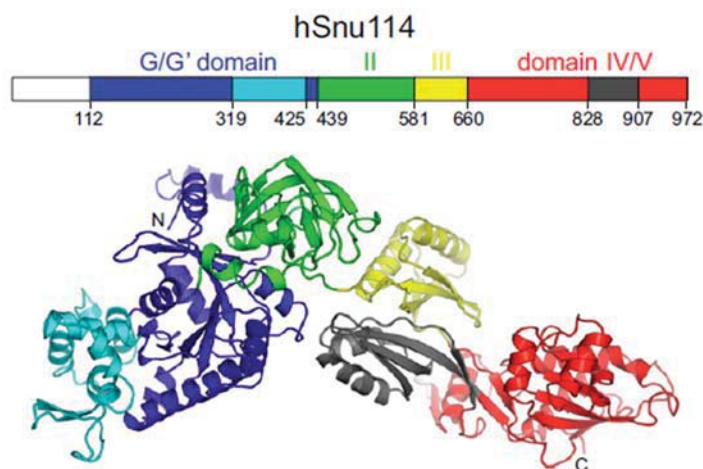

Figure 16: Structural model for hSnu114. Upper: Predicted domains of hSnu114 based on homology to eEF2. The domains are named according to eEF2 nomenclature. Residue numbering refers to hSnu114 sequence. Lower:

The data confirm that the 112-residue N-terminal extension is largely unstructured, as it is readily digested upon addition of protease. The following N-terminal part up to residue 580 apparently contains domains that fold efficiently in rabbit reticulocyte lysate. Although these N-terminal domains are resistant to protease digest without encapsulation, some of the N-terminal construct hSnu114 [1–580] co-migrates with closed TRiC, indicating that it transiently interacts with the chaperonin. The C-terminal construct hSnu114 [581–927] is highly sensitive to digestion unless protected via encapsulation by TRiC. hSnu114 [581–972] is apparently incompetent to fold stably outside the context of the full-length protein, and thus permanently cycles on and off the chaperonin. During partial encapsulation of the full-length hSnu114 [1–972], C-terminal segments of various sizes are protected from proteolysis. The size of the selected C-terminal fragments roughly matches predicted domain boundaries. This is consistent with mainly flexible inter-domain linkers being able to meet the steric constraints imposed by the apical pore. A common strong TRiC binding motif in hSnu114 must be located in the C-terminal 140 residues, as judged from the smallest encapsulated fragment. The set of multiple C-terminal fragments observed may reflect progressive stages of folding with subsequent folding units being added to the N-terminal folding scaffold.

5.3   Chaperone function of Hgh1 in the biogenesis of eEF2

Hgh1 is the yeast homologue of Fam203, a conserved eukaryotic protein of 45 kDa, which was proposed to regulate TRiC function (Hein et al., 2015). Large-scale physical and genetic interaction screens had linked Hgh1 to the yeast orthologue of eukaryotic elongation factor 2, Eft (Costanzo et al., 2010; Gavin et al., 2006; Krogan et al., 2006). Eft itself was consistently found as TRiC substrate (Dekker et al., 2008; Yam et al., 2008) and constitutes the TRiC client of highest cellular abundance in the yeast cytosol (Kulak et al., 2014). Our results demonstrate that Hgh1 is a specific chaperone for de-novo folding of Eft. Hgh1 binds

to domain III of Eft and thereby appears to prevent aberrant interactions. Moreover, Hgh1 recruits TRiC to the Eft C-terminal domains, which require TRiC interaction for folding. In the absence of Hgh1, a significant portion of Eft misfolds, resulting in aggregation or premature degradation.

5.3.1 Structure and dynamics of Eft

Eft is a 93 kDa highly abundant multi-domain protein. It consists of six domains, which are arranged in a complex nested topology (Figure 17). Domains G' and V are inserted into domains G and IV, respectively, so that during folding residues far apart in the sequence have to come into close proximity to establish correct contacts.
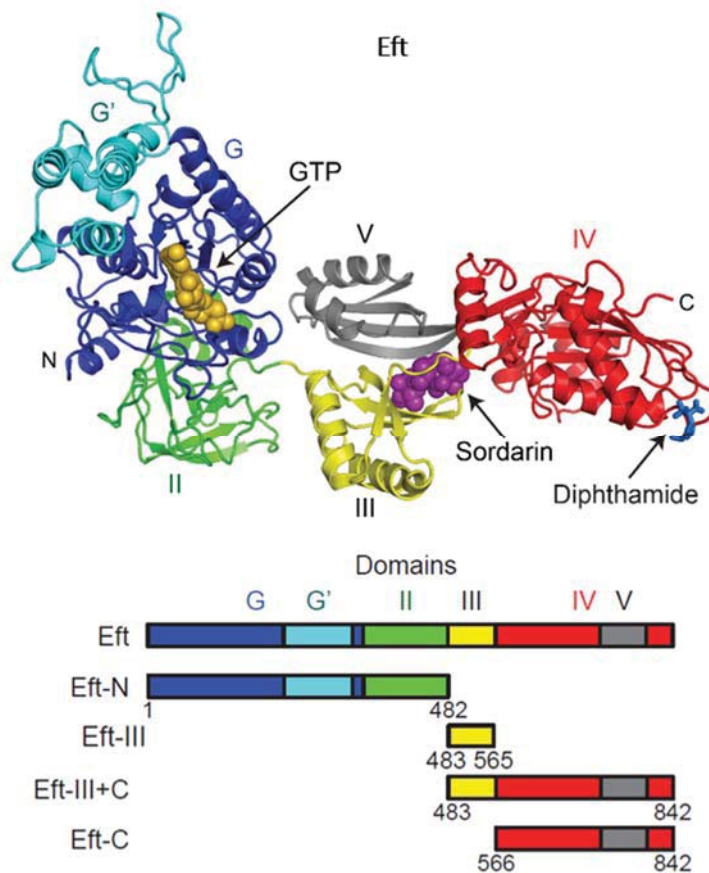
Figure 17: Structure and domain organization of Eft. The Eft model is a composite of the cryoEM structure of a ribosome-bound eEF2:GMP-PCP complex (Murray et al., 2016) and the crystal structure of the S. cerevisiae eEF2 orthologue, Eft, in complex with the anti-fungal compound sordarin (Jorgensen et al., 2003). The domain structure is color-coded and shown schematically below. Additionally, the Eft truncation constructs used in the study are shown. Figure modified from (Monkemeyer et al., 2018).

The six domains of Eft form two rigid blocks: The N-terminal module, Eft-N (residues 1-482), compromises domains G, G', and II and contains the GTP binding site. The C-terminal module, Eft-C (residues 566-842), compromises domains IV and V and reaches into the decoding center upon ribosome binding. Upon translation, large structural rearrangements of the C-module relative to the N-module occur. The N- and C-modules are connected via domain III of Eft (Eft-III). Our HD-X measurements revealed that domain III (residues 483-565) is structurally more dynamic than the other domains. Moreover, we find that Eft-III is not natively folded in isolation. This small domain (82 amino acids) contains several hydrophobic residue

segments, which may form interfaces with other domains in the full-length protein. Flexibility of domain III connecting the rather rigid modules Eft-N and Eft-C might be functionally important. Stabilization of this domain, as induced by binding of Sordarin (Figure 17) or fusidic acid, inhibits Eft function. While being functionally important to allow large conformational changes required in the molecule, the intrinsic flexibility of domain III may interfere with folding of the topologically complex flanking modules Eft-N and Eft-C.

5.3.2   Hgh1 binds to Eft domain III

We find that Hgh1 binds to domain III of Eft, as judged by our crosslinking and hydrogen-deuterium exchange (HD-X) results. This interaction probably occurs co-translationally, as soon as domain III of Eft emerges from the ribosomal exit tunnel. This is consistent with the enrichment of ribosomal proteins upon Hgh1-FLAG pull-down. The structure of Hgh1, which we solved by X-ray crystallography, reveals an armadillo repeat fold. Proteins with this curved α-helical topology have been found to interact with an extended peptide of their interaction partner via hydrophobic residues along their concave face (Reichen et al., 2014). The concave face in Hgh1 harbors a conserved surface area that contributes to the interaction with Eft, as shown by mutation. Likewise, a conserved surface groove close to the Hgh1 N-terminus contributes to Eft binding. Our HD-X analysis indicates that the Hgh1 binding site on Eft comprises residues 522-540, located in Eft-III. In an extended conformation, this peptide would be long enough to contact both binding sites in Hgh1. Of note, in the HD-X measurement, Eft sequence 522-540 is represented by several smaller peptides, which show distinct deuterium incorporation rates. Some residues appear strongly protected, while others are more dynamic. This is consistent with a bi-partite binding interface on Hgh1.

Interestingly, domain III seems to transiently unfold in mature Eft enabling binding of Hgh1 to functional Eft. *In vitro* this Hgh1-Eft complex is unstable and was detected only by crosslinking. *In vivo*, this interaction is presumably negligible because cellular levels of Hgh1 (5.6 x 10^3 copies per cell) are low in comparison to Eft (1.8 x 10^5 copies per cell) (Kulak et al., 2014). This suggests that the cellular level of Hgh1 is carefully regulated. Indeed, Hgh1 overexpression seemed to reduce cellular fitness in a large-scale survey (Yoshikawa et al., 2011).

The binding of Hgh1 to the hydrophobic and intrinsically flexible domain Eft-III may stabilize Eft folding intermediates and prevent pre-mature unproductive interactions. This may serve to keep the complex modules Eft-N and Eft-C apart and facilitate progression to their native states. Additionally, we suggest that Hgh1 recruits further members of the chaperone network to assist in the folding of Eft.

5.3.3   Hgh1 facilitates the recruitment of molecular chaperones to Eft folding intermediates

Hgh1 may serve a dual function in Eft folding: It might prevent domain III from interfering with the folding of the N- and C-terminal modules in Eft and at the same time recruit additional chaperones to Eft folding intermediates. The TRiC complex seems to be recruited to elements in the C-terminal Eft domains. Importantly, the truncated constructs Eft-C and Eft-III+C cannot fold stably in isolation but remain TRiC-bound or aggregate. This is analogous to the behavior observed for the structural homologue hSnu114 (Russmann et al., 2012). Analysis of hSnu114 revealed that C-terminal fragments up to 37 kDa were encapsulated by TRiC and that – as in the case of Eft - the C-terminal module would only fold stably in the context of the full-length protein.

To allow efficient structural Eft maturation, the N-module probably must completes folding first. However, the N-module also seems to require assistance by eukaryotic chaperones. Our attempts to express Eft-N in *E. coli* failed to produce soluble protein. The Hsp70 system may be involved in the folding of the N-

terminal domains, as the yeast Hsp70, Ssa1, and its cofactor, Ydj1, were found to associate with Hgh1 *in vivo*. Another candidate for folding the N-module is Hsp90. The activity of the Hsp90 system seems to be important for Eft folding, because cellular Eft levels are significantly reduced upon treatment of cells with the specific Hsp90 inhibitor Macbecin. A direct interaction between the essential Hsp90 cochaperone Cns1 and Hgh1 may help recruiting Hsp90 to Eft folding intermediates (Gavin et al., 2006; Schlecht et al., 2012; Tarassov et al., 2008). Cns1 is a multi-domain protein that interacts via its TPR domain with the C-terminal EEVD peptide of Hsp90 (Hainzl et al., 2004). Both these interactions appear to be conserved between the human orthologues TTC4 (Cns1), Fam203 (Hgh1) and Hsp90 (Huttlin et al., 2017; Huttlin et al., 2015; Kristensen et al., 2012).

Taking these findings together, a model emerges (Figure 18) in which Hgh1 assists Eft folding by coordinating the action of Hsp90 and TRiC. Hgh1 and TRiC would shield domain III and the C-module of Eft, respectively, until the N-module is natively folded. In the next step, TRiC assists the folding of the C-module, followed by release of the chaperonin. Structure formation in domain III finally triggers the release of Hgh1. Hsp90 and TRiC might thus process Eft folding intermediates sequentially. Although this is a plausible scenario, both chaperone systems may also provide alternative or partially redundant folding pathways.
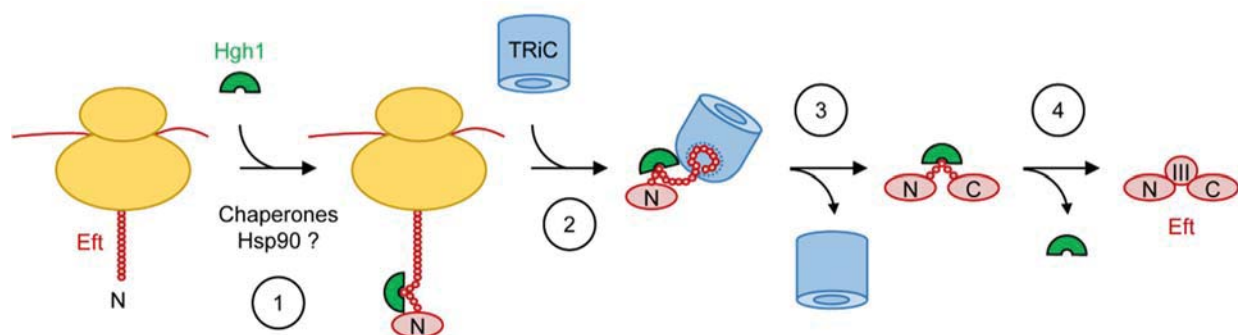

Figure 18: Model for Hgh1-assisted folding of Eft. Hgh1 binds to domain III, probably co-translationally (1). After Eft synthesis is completed, TRiC is recruited to the C-module of Eft. (2). In this state, domain III and the C-module in Eft are effectively shielded to facilitate folding of the N-module unimpaired by aberrant intramolecular interactions,

probably with assistance of Hsp90. TRiC leaves the complex after successful folding of the C-module (3). Finally, domain III adopts its native fold and Hgh1 is released (4). Figure adapted from (Monkemeyer at al., 2018).

5.3.4   Absence of Hgh1 leads to Eft misfolding

In the absence of Hgh1, the folding efficiency of Eft is reduced. Eft folding intermediates accumulate, which eventually aggregate or are cleared by degradation. Importantly, *HGH1* is not an essential gene and cells tolerate loss of *HGH1* - and the resulting decrease in Eft levels - without apparent growth defect. Thus cells appear to produce Eft in excess to what is required for optimal growth in culture. Indeed, Eft makes up 1.5% of total protein mass in *S. cerevisiae* (Liebermeister et al., 2014). Given the complexity of Eft folding, alternative folding pathways may exist to ensure the allocation of sufficient Eft capacity required for viability.

Inefficient folding of the highly abundant Eft protein due to absence of Hgh1 probably constitutes a considerable stress to the cell, especially during rapid cell growth. A substantial part of the cytosolic chaperones might be occupied with de-novo synthesized Eft. This folding stress may be the cause of the heat shock response observed upon deletion of *HGH1* (Brandman et al., 2012). Several observations suggest that deletion of *HGH1* increases the cellular dependence on Hsp90 (Alford and Brandman, 2018). First, cells deleted for *HGH1* are sensitive to the Hsp90 inhibitor Macbecin (McClellan et al., 2007). Additionally, the combined deletion of *HGH1* and several components of the Hsp90 system, such as *CPR7*, *HCH1*, *HSC82*, *HSP82* and *STI1*, causes a synthetic growth defect (Costanzo et al., 2010). Deletion of *HGH1* furthermore displays negative genetic interactions with subunits of the Conserved Oligomeric Golgi (COG) complex (Costanzo et al., 2010), a vesicle tethering complex that requires Hsp90 for assembly (McClellan et al., 2007). This suggests that Hsp90 capacity becomes limiting for cell growth when Hgh1 is absent. Whether the functional cooperation between Hgh1 and Hsp90 extends to other processes than Eft biogenesis remains to be established.

6   Conclusion and Outlook

Understanding the architecture of the TRiC complex helped in understanding the asymmetric properties of the chaperonin cavity and provided new insights into the folding mechanism of TRiC. Very recently, Balchin et al. used HD-X to follow the conformational progression of actin during TRiC-mediated folding (Balchin et al., 2018). This analysis confirmed that non-native actin binds to several TRiC subunits in an extended conformation, and that partial release of these contacts - induced by an asymmetric ATP utilization - is important to direct the folding pathway. This is the first comprehensive analysis revealing how the unique features of TRiC contribute to successful folding of an obligate substrate. The extent to which this mechanism is specific to actin folding remains to be investigated.

Crosslinking coupled to mass spectrometry is a method with great potential to reveal interactions in large complexes. Details of the instrumental requirements and analysis tools are now publicly available and the method has already been used to analyze the topology of ribosomes, proteasomes and chromatin remodelers (Leitner et al., 2016). Furthermore, the method can now be extended to the analysis of membrane proteins, of protein networks and their changes upon perturbations or even to whole proteomes.

We characterized the function of Hgh1, a chaperone that cooperates with TRiC in the biogenesis of Eft. An independent unpublished study by Schopf et al. suggests that Hgh1 moreover recruits Cns1 and Hsp90 to Eft folding intermediates (personal communication). Although our study revealed important insights into Eft folding, we do not yet fully understand the folding pathway of this complex protein, especially regarding the relative contributions of TRiC and the Hsp90 system. It will be interesting to see whether Fam203 in mammalian cells functions in Eft folding in a manner similar to Hgh1.

More generally, although multi-domain proteins like Eft comprise a large fraction of the known eukaryotic proteomes, their biogenesis pathways remain virtually unexplored. *In vitro* translation with reconstituted

components would probably be best suited to understand Eft folding in more detail. The PURE system

containing the prokaryotic translation machinery (Shimizu et al., 2006) could be augmented with TRiC and

Hgh1 and possibly other eukaryotic factors that participate in Eft folding, such as Cns1 and Hsp90. Binding

of GTP and Sordarin to the N-terminal and C-terminal Eft module, respectively, could serve as a read out

for proper folding of these segments. Similarly, monitoring the interaction of Eft with the ribosome,

phosphorylation by eEF2 kinase and modification by the diphthamide machinery might provide useful

probes of the Eft folding status.

7   References

Alford, B.D., and Brandman, O. (2018). Quantification of Hsp90 availability reveals differential coupling to the heat shock response. J Cell Biol.

Amit, M., Weisberg, S.J., Nadler-Holly, M., McCormack, E.A., Feldmesser, E., Kaganovich, D., Willison, K.R., and Horovitz, A. (2010). Equivalent mutations in the eight subunits of the chaperonin CCT produce dramatically different cellular and gene expression phenotypes. J Mol Biol *401*, 532-543.

Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. Science *181*, 223-230.

Archibald, J.M., Blouin, C., and Doolittle, W.F. (2001). Gene duplication and the evolution of group II chaperonins: Implications for structure and function. J Struct Biol *135*, 157-169.

Armstrong, H., Wolmarans, A., Mercier, R., Mai, B., and LaPointe, P. (2012). The co-chaperone Hch1 regulates Hsp90 function differently than its homologue Aha1 and confers sensitivity to yeast to the Hsp90 inhibitor NVP-AUY922. PLoS One *7*, e49322.

Balchin, D., Hayer-Hartl, M., and Hartl, F.U. (2016). In vivo aspects of protein folding and quality control. Science *353*, aac4354.

Balchin, D., Milicic, G., Strauss, M., Hayer-Hartl, M., and Hartl, F.U. (2018). Pathway of Actin Folding Directed by the Eukaryotic Chaperonin TRiC. Cell.

Behrends, C., Langer, C.A., Boteva, R., Bottcher, U.M., Stemp, M.J., Schaffar, G., Rao, B.V., Giese, A., Kretzschmar, H., Siegers, K.*, et al.* (2006). Chaperonin TRiC promotes the assembly of polyQ expansion proteins into nontoxic oligomers. Mol Cell *23*, 887-897.

Bergeron, L.M., Shis, D.L., Gomez, L., and Clark, D.S. (2009). Small molecule inhibition of a Group II chaperonin: pinpointing a loop region within the equatorial domain as necessary for protein refolding. Archives of biochemistry and biophysics *481*, 45-51.

Bigotti, M.G., Bellamy, S.R., and Clarke, A.R. (2006). The asymmetric ATPase cycle of the thermosome: elucidation of the binding, hydrolysis and product-release steps. J Mol Biol *362*, 835-843.

Boczek, E.E., Reefschlager, L.G., Dehling, M., Struller, T.J., Hausler, E., Seidl, A., Kaila, V.R., and Buchner, J. (2015). Conformational processing of oncogenic v-Src kinase by the molecular chaperone Hsp90. Proc Natl Acad Sci U S A *112*, E3189-3198.

Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D.C., Joachimiak, A., Horwich, A.L., and Sigler, P.B. (1994). The Crystal-Structure of the Bacterial Chaperonin Groel at 2.8-Angstrom. Nature *371*, 578-586.

Brandman, O., Stewart-Ornstein, J., Wong, D., Larson, A., Williams, C.C., Li, G.W., Zhou, S., King, D., Shen, P.S., Weibezahn, J.*, et al.* (2012). A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. Cell *151*, 1042-1054.

Brandt, F., Carlson, L.A., Hartl, F.U., Baumeister, W., and Grunewald, K. (2010). The Three-Dimensional Organization of Polyribosomes in Intact Human Cells. Mol Cell *39*, 560-569.

Brandt, F., Etchells, S.A., Ortiz, J.O., Elcock, A.H., Hartl, F.U., and Baumeister, W. (2009). The Native 3D Organization of Bacterial Polysomes. Cell *136*, 261-271.

Brockwell, D.J., and Radford, S.E. (2007). Intermediates: ubiquitous species on folding energy landscapes? Curr Opin Struc Biol *17*, 30-37.

Buchberger, A., Bukau, B., and Sommer, T. (2010). Protein quality control in the cytosol and the endoplasmic reticulum: brothers in arms. Mol Cell *40*, 238-252.

Calloni, G., Chen, T., Schermann, S.M., Chang, H.C., Genevaux, P., Agostini, F., Tartaglia, G.G., Hayer-Hartl, M., and Hartl, F.U. (2012). DnaK Functions as a Central Hub in the E. coli Chaperone Network. Cell Rep *1*, 251-264.

Camasses, A., Bogdanova, A., Shevchenko, A., and Zachariae, W. (2003). The CCT chaperonin promotes activation of the anaphase-promoting complex through the generation of functional Cdc20. Mol Cell *12*, 87-100.

Cao, S., Carlesso, G., Osipovich, A.B., Llanes, J., Lin, Q., Hoek, K.L., Khan, W.N., and Ruley, H.E. (2008). Subunit 1 of the prefoldin chaperone complex is required for lymphocyte development and function. Journal of immunology *181*, 476-484.

Chakraborty, K., Chatila, M., Sinha, J., Shi, Q.Y., Poschner, B.C., Sikor, M., Jiang, G.X., Lamb, D.C., Hartl, F.U., and Hayer-Hartl, M. (2010). Chaperonin-Catalyzed Rescue of Kinetically Trapped States in Protein Folding. Cell *142*, 112-122.

Chaudhuri, T.K., Farr, G.W., Fenton, W.A., Rospert, S., and Horwich, A.L. (2001). GroEL/GroES-mediated folding of a protein too large to be encapsulated. Cell *107*, 235-246.

Chen, B., Zhong, D., and Monteiro, A. (2006). Comparative genomics and evolution of the HSP90 family of genes across all kingdoms of organisms. BMC Genomics *7*, 156.

Chen, L., and Sigler, P.B. (1999). The crystal structure of a GroEL/peptide complex: plasticity as a basis for substrate diversity. Cell *99*, 757-768.

Chiti, F., and Dobson, C.M. (2006). Protein misfolding, functional amyloid, and human disease. Annual Review of Biochemistry *75*, 333-366.

Clare, D.K., Stagg, S., Quispe, J., Farr, G.W., Horwich, A.L., and Saibil, H.R. (2008). Multiple states of a nucleotide-bound group 2 chaperonin. Structure *16*, 528-534.

Cong, Y., Baker, M.L., Jakana, J., Woolford, D., Miller, E.J., Reissmann, S., Kumar, R.N., Redding-Johanson, A.M., Batth, T.S., Mukhopadhyay, A.*, et al.* (2010). 4.0-A resolution cryo-EM structure of the mammalian chaperonin TRiC/CCT reveals its unique subunit arrangement. Proc Natl Acad Sci U S A *107*, 4967-4972.

Cong, Y., Schroder, G.F., Meyer, A.S., Jakana, J., Ma, B., Dougherty, M.T., Schmid, M.F., Reissmann, S., Levitt, M., Ludtke, S.L.*, et al.* (2012). Symmetry-free cryo-EM structures of the chaperonin TRiC along its ATPase-driven conformational cycle. EMBO J *31*, 720-730.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L., Toufighi, K., Mostafavi, S.*, et al.* (2010). The genetic landscape of a cell. Science *327*, 425-431.

Cuellar, J., Martin-Benito, J., Scheres, S.H., Sousa, R., Moro, F., Lopez-Vinas, E., Gomez-Puertas, P., Muga, A., Carrascosa, J.L., and Valpuesta, J.M. (2008). The structure of CCT-Hsc70 NBD suggests a mechanism for Hsp70 delivery of substrates to the chaperonin. Nat Struct Mol Biol *15*, 858-864.

Dekker, C., Roe, S.M., McCormack, E.A., Beuron, F., Pearl, L.H., and Willison, K.R. (2011). The crystal structure of yeast CCT reveals intrinsic asymmetry of eukaryotic cytosolic chaperonins. EMBO J *30*, 3078-3090.

Dekker, C., Stirling, P.C., McCormack, E.A., Filmore, H., Paul, A., Brost, R.L., Costanzo, M., Boone, C., Leroux, M.R., and Willison, K.R. (2008). The interaction network of the chaperonin CCT. EMBO J *27*, 1827-1839.

Dever, T.E., and Green, R. (2012). The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harbor perspectives in biology *4*, a013706.

Dinner, A.R., Sali, A., Smith, L.J., Dobson, C.M., and Karplus, M. (2000). Understanding protein folding via free-energy surfaces from theory and experiment. Trends Biochem Sci *25*, 331-339.

Ditzel, L., Lowe, J., Stock, D., Stetter, K.O., Huber, H., Huber, R., and Steinbacher, S. (1998). Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. Cell *93*, 125-138.

Dobson, C.M., Sali, A., and Karplus, M. (1998). Protein folding: A perspective from theory and experiment. Angew Chem Int Edit *37*, 868-893.

Douglas, N.R., Reissmann, S., Zhang, J., Chen, B., Jakana, J., Kumar, R., Chiu, W., and Frydman, J. (2011). Dual action of ATP hydrolysis couples lid closure to substrate release into the group II chaperonin chamber. Cell *144*, 240-252.

Elad, N., Farr, G.W., Clare, D.K., Orlova, E.V., Horwich, A.L., and Saibil, H.R. (2007). Topologies of a substrate protein bound to the chaperonin GroEL. Mol Cell *26*, 415-426.

Ellis, R.J. (1994). Molecular Chaperones - Opening and Closing the Anfinsen Cage. Curr Biol *4*, 633-635.

Ellis, R.J., and Minton, A.P. (2006). Protein aggregation in crowded environments. Biol Chem *387*, 485-497.

England, J., Lucent, D., and Pande, V. (2008). Rattling the cage: computational models of chaperonin-mediated protein folding. Curr Opin Struc Biol *18*, 163-169.

Etchells, S.A., Meyer, A.S., Yam, A.Y., Roobol, A., Miao, Y., Shao, Y., Carden, M.J., Skach, W.R., Frydman, J., and Johnson, A.E. (2005). The cotranslational contacts between ribosome-bound nascent polypeptides and the subunits of the hetero-oligomeric chaperonin TRiC probed by photocross-linking. The Journal of biological chemistry *280*, 28118-28126.

Fabrizio, P., Laggerbauer, B., Lauber, J., Lane, W.S., and Luhrmann, R. (1997). An evolutionarily conserved U5 snRNP-specific protein is a GTP-binding factor closely related to the ribosomal translocase EF-2. EMBO J *16*, 4092-4106.

Feldman, D.E., Spiess, C., Howard, D.E., and Frydman, J. (2003). Tumorigenic mutations in VHL disrupt folding in vivo by interfering with chaperonin binding. Mol Cell *12*, 1213-1224.

Frazer, L.N., Nancollis, V., and O'Keefe, R.T. (2008). The role of Snu114p during pre-mRNA splicing. Biochem Soc Trans *36*, 551-553.

Freund, A., Zhong, F.L., Venteicher, A.S., Meng, Z., Veenstra, T.D., Frydman, J., and Artandi, S.E. (2014). Proteostatic control of telomerase function through TRiC-mediated folding of TCAB1. Cell *159*, 1389-1403.

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.

Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dumpelfeld, B.*, et al.* (2006). Proteome survey reveals modularity of the yeast cell machinery. Nature *440*, 631-636.

Geissler, S., Siegers, K., and Schiebel, E. (1998). A novel protein complex promoting formation of functional alpha- and gamma-tubulin. EMBO J *17*, 952-966.

Georgescauld, F., Popova, K., Gupta, A.J., Bracher, A., Engen, J.R., Hayer-Hartl, M., and Hartl, F.U. (2014). GroEL/ES Chaperonin Modulates the Mechanism and Accelerates the Rate of TIM-Barrel Domain Folding. Cell *157*, 922-934.

Gershenson, A., and Gierasch, L.M. (2011). Protein folding in the cell: challenges and progress. Curr Opin Struc Biol *21*, 32-41.

Gershenson, A., Gierasch, L.M., Pastore, A., and Radford, S.E. (2014). Energy landscapes of functional proteins are inherently risky. Nat Chem Biol *10*, 884-891.

Gomez-Puertas, P., Martin-Benito, J., Carrascosa, J.L., Willison, K.R., and Valpuesta, J.M. (2004). The substrate recognition mechanisms in chaperonins. Journal of molecular recognition : JMR *17*, 85-94.

Guerrero, C., Milenkovic, T., Przulj, N., Kaiser, P., and Huang, L. (2008). Characterization of the proteasome interaction network using a QTAX-based tag-team strategy and protein interaction network analysis. Proc Natl Acad Sci U S A *105*, 13333-13338.

Gupta, A.J., Haldar, S., Milicic, G., Hartl, F.U., and Hayer-Hartl, M. (2014). Active Cage Mechanism of Chaperonin-Assisted Protein Folding Demonstrated at Single-Molecule Level. J Mol Biol *426*, 2739-2754.

Gutsche, I., Essen, L.O., and Baumeister, F. (1999). Group II chaperonins: New TRiC(k)s and turns of a protein folding machine. J Mol Biol *293*, 295-312.

Hainzl, O., Wegele, H., Richter, K., and Buchner, J. (2004). Cns1 is an activator of the Ssa1 ATPase activity. The Journal of biological chemistry *279*, 23267-23273.

Han, J.H., Batey, S., Nickson, A.A., Teichmann, S.A., and Clarke, J. (2007). The folding and evolution of multidomain proteins. Nat Rev Mol Cell Biol *8*, 319-330.

Hartl, F.U. (1996). Molecular chaperones in cellular protein folding. Nature *381*, 571-579.

Hartl, F.U., and Hayer-Hartl, M. (2002). Protein folding - Molecular chaperones in the cytosol: from nascent chain to folded protein. Science *295*, 1852-1858.

Hayer-Hartl, M., Bracher, A., and Hartl, F.U. (2016). The GroEL-GroES Chaperonin Machine: A Nano-Cage for Protein Folding. Trends Biochem Sci *41*, 62-76.

Hein, M.Y., Hubner, N.C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I.A., Weisswange, I., Mansfeld, J., Buchholz, F.*, et al.* (2015). A human interactome in three quantitative dimensions organized by stoichiometries and abundances. Cell *163*, 712-723.

Herzog, F., Kahraman, A., Boehringer, D., Mak, R., Bracher, A., Walzthoeni, T., Leitner, A., Beck, M., Hartl, F.U., Ban, N.*, et al.* (2012). Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. Science *337*, 1348-1352.

Hirtreiter, A.M., Calloni, G., Forner, F., Scheibe, B., Puype, M., Vandekerckhove, J., Mann, M., Hartl, F.U., and Hayer-Hartl, M. (2009). Differential substrate specificity of group I and group II chaperonins in the archaeon Methanosarcina mazei. Mol Microbiol *74*, 1152-1168.

Horovitz, A., Fridmann, Y., Kafri, G., and Yifrach, O. (2001). Review: allostery in chaperonins. J Struct Biol *135*, 104-114.

Horwich, A.L., Fenton, W.A., Chapman, E., and Farr, G.W. (2007). Two families of chaperonin: Physiology and mechanism. Annu Rev Cell Dev Bi *23*, 115-145.

Houry, W.A., Frishman, D., Eckerskorn, C., Lottspeich, F., and Hartl, F.U. (1999). Identification of in vivo substrates of the chaperonin GroEL. Nature *402*, 147-154.

Huo, Y., Hu, Z., Zhang, K., Wang, L., Zhai, Y., Zhou, Q., Lander, G., Zhu, J., He, Y., Pang, X.*, et al.* (2010). Crystal structure of group II chaperonin in the open state. Structure *18*, 1270-1279.

Huttlin, E.L., Bruckner, R.J., Paulo, J.A., Cannon, J.R., Ting, L., Baltier, K., Colby, G., Gebreab, F., Gygi, M.P., Parzen, H.*, et al.* (2017). Architecture of the human interactome defines protein communities and disease networks. Nature *545*, 505-509.

Huttlin, E.L., Ting, L., Bruckner, R.J., Gebreab, F., Gygi, M.P., Szpyt, J., Tam, S., Zarraga, G., Colby, G., Baltier, K.*, et al.* (2015). The BioPlex Network: A Systematic Exploration of the Human Interactome. Cell *162*, 425-440.

Iizuka, R., So, S., Inobe, T., Yoshida, T., Zako, T., Kuwajima, K., and Yohda, M. (2004). Role of the helical protrusion in the conformational change and molecular chaperone activity of the archaeal group II chaperonin. The Journal of biological chemistry *279*, 18834-18839.

Jiang, Y., Douglas, N.R., Conley, N.R., Miller, E.J., Frydman, J., and Moerner, W.E. (2011). Sensing cooperativity in ATP hydrolysis for single multisubunit enzymes in solution. Proc Natl Acad Sci U S A *108*, 16962-16967.

Joachimiak, L.A., Walzthoeni, T., Liu, C.W., Aebersold, R., and Frydman, J. (2014). The structural basis of substrate recognition by the eukaryotic chaperonin TRiC/CCT. Cell *159*, 1042-1055.

Jorgensen, R., Ortiz, P.A., Carr-Schmid, A., Nissen, P., Kinzy, T.G., and Andersen, G.R. (2003). Two crystal structures demonstrate large conformational changes in the eukaryotic ribosomal translocase. Nature structural biology *10*, 379-385.

Kafri, G., Willison, K.R., and Horovitz, A. (2001). Nested allosteric interactions in the cytoplasmic chaperonin containing TCP-1. Protein science : a publication of the Protein Society *10*, 445-449.

Kalisman, N., Adams, C.M., and Levitt, M. (2012). Subunit order of eukaryotic TRiC/CCT chaperonin by cross-linking, mass spectrometry, and combinatorial homology modeling. Proc Natl Acad Sci U S A *109*, 2884-2889.

Kalisman, N., Schroder, G.F., and Levitt, M. (2013). The crystal structures of the eukaryotic chaperonin CCT reveal its functional partitioning. Structure *21*, 540-549.

Kenney, J.W., Moore, C.E., Wang, X., and Proud, C.G. (2014). Eukaryotic elongation factor 2 kinase, an unusual enzyme with multiple roles. Advances in biological regulation *55*, 15-27.

Kerner, M.J., Naylor, D.J., Ishihama, Y., Maier, T., Chang, H.C., Stines, A.P., Georgopoulos, C., Frishman, D., Hayer-Hartl, M., Mann, M.*, et al.* (2005). Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. Cell *122*, 209-220.

Kim, S., Willison, K.R., and Horwich, A.L. (1994). Cystosolic Chaperonin Subunits Have a Conserved Atpase Domain but Diverged Polypeptide-Binding Domains. Trends Biochem Sci *19*, 543-548.

Kim, Y.E., Hipp, M.S., Bracher, A., Hayer-Hartl, M., and Hartl, F.U. (2013). Molecular chaperone functions in protein folding and proteostasis. Annu Rev Biochem *82*, 323-355.

Kimata, Y., and Kohno, K. (1994). Elongation factor 2 mutants deficient in diphthamide formation show temperature-sensitive cell growth. The Journal of biological chemistry *269*, 13497-13501.

Kirschke, E., Goswami, D., Southworth, D., Griffin, P.R., and Agard, D.A. (2014). Glucocorticoid Receptor Function Regulated by Coordinated Action of the Hsp90 and Hsp70 Chaperone Cycles. Cell *157*, 1685-1697.

Kitagawa, K., Skowyra, D., Elledge, S.J., Harper, J.W., and Hieter, P. (1999). SGT1 encodes an essential component of the yeast kinetochore assembly pathway and a novel subunit of the SCF ubiquitin ligase complex. Mol Cell *4*, 21-33.

Koulov, A.V., LaPointe, P., Lu, B., Razvi, A., Coppinger, J., Dong, M.Q., Matteson, J., Laister, R., Arrowsmith, C., Yates, J.R., 3rd*, et al.* (2010). Biological and structural basis for Aha1 regulation of Hsp90 ATPase activity in maintaining proteostasis in the human disease cystic fibrosis. Molecular biology of the cell *21*, 871-884.

Kristensen, A.R., Gsponer, J., and Foster, L.J. (2012). A high-throughput approach for measuring temporal changes in the interactome. Nat Methods *9*, 907-909.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P.*, et al.* (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637-643.

Kubota, H., Hynes, G., and Willison, K. (1995). The Chaperonin Containing T-Complex Polypeptide-1 (Tcp-1) - Multisubunit Machinery Assisting in Protein-Folding and Assembly in the Eukaryotic Cytosol. Eur J Biochem *230*, 3-16.

Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nat Methods *11*, 319-324.

Kusmierczyk, A.R., and Martin, J. (2003). Nucleotide-dependent protein folding in the type II chaperonin from the mesophilic archaeon Methanococcus maripaludis. The Biochemical journal *371*, 669-673.

Lacefield, S., and Solomon, F. (2003). A novel step in beta-tubulin folding is important for heterodimer formation in Saccharomyces cerevisiae. Genetics *165*, 531-541.

Leitner, A., Faini, M., Stengel, F., and Aebersold, R. (2016). Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. Trends Biochem Sci *41*, 20-32.

Leitner, A., Joachimiak, L.A., Bracher, A., Monkemeyer, L., Walzthoeni, T., Chen, B., Pechmann, S., Holmes, S., Cong, Y., Ma, B.*, et al.* (2012). The molecular architecture of the eukaryotic chaperonin TRiC/CCT. Structure *20*, 814-825.

Leitner, A., Joachimiak, L.A., Unverdorben, P., Walzthoeni, T., Frydman, J., Forster, F., and Aebersold, R. (2014). Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. Proc Natl Acad Sci U S A *111*, 9455-9460.

Levinthal, C. (1968). Are There Pathways for Protein Folding. J Chim Phys Pcb *65*, 44-+.

Li, J., Richter, K., and Buchner, J. (2011). Mixed Hsp90-cochaperone complexes are important for the progression of the reaction cycle. Nat Struct Mol Biol *18*, 61-66.

Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of protein investment in cellular functions. Proc Natl Acad Sci U S A *111*, 8488-8493.

Lin, J., Gagnon, M.G., Bulkley, D., and Steitz, T.A. (2015). Conformational changes of elongation factor G on the ribosome during tRNA translocation. Cell *160*, 219-227.

Lin, P., and Sherman, F. (1997). The unique hetero-oligomeric nature of the subunits in the catalytic cooperativity of the yeast Cct chaperonin complex. Proc Natl Acad Sci U S A *94*, 10780-10785.

Lin, Z., Madan, D., and Rye, H.S. (2008). GroEL stimulates protein folding through forced unfolding. Nat Struct Mol Biol *15*, 303-311.

Liou, A.K., and Willison, K.R. (1997). Elucidation of the subunit orientation in CCT (chaperonin containing TCP1) from the subunit composition of CCT micro-complexes. EMBO J *16*, 4311-4316.

Liu, F., and Heck, A.J. (2015). Interrogating the architecture of protein assemblies and protein interaction networks by cross-linking mass spectrometry. Curr Opin Struct Biol *35*, 100-108.

Llorca, O., Martin-Benito, J., Grantham, J., Ritco-Vonsovici, M., Willison, K.R., Carrascosa, J.L., and Valpuesta, J.M. (2001). The 'sequential allosteric ring' mechanism in the eukaryotic chaperonin-assisted folding of actin and tubulin. EMBO J *20*, 4065-4075.

Llorca, O., Martin-Benito, J., Ritco-Vonsovici, M., Grantham, J., Hynes, G.M., Willison, K.R., Carrascosa, J.L., and Valpuesta, J.M. (2000). Eukaryotic chaperonin CCT stabilizes actin and tubulin folding intermediates in open quasi-native conformations. EMBO J *19*, 5971-5979.

Llorca, O., McCormack, E.A., Hynes, G., Grantham, J., Cordell, J., Carrascosa, J.L., Willison, K.R., Fernandez, J.J., and Valpuesta, J.M. (1999a). Eukaryotic type II chaperonin CCT interacts with actin through specific subunits. Nature *402*, 693-696.

Llorca, O., Smyth, M.G., Carrascosa, J.L., Willison, K.R., Radermacher, M., Steinbacher, S., and Valpuesta, J.M. (1999b). 3D reconstruction of the ATP-bound form of CCT reveals the asymmetric folding conformation of a type II chaperonin. Nature structural biology *6*, 639-642.

Lopez-Fanarraga, M., Avila, J., Guasch, A., Coll, M., and Zabala, J.C. (2001). Review: postchaperonin tubulin folding cofactors and their role in microtubule dynamics. J Struct Biol *135*, 219-229.

Lopez, T., Dalton, K., and Frydman, J. (2015). The Mechanism and Function of Group II Chaperonins. J Mol Biol *427*, 2919-2930.

Lund, P. (2011). Insights into chaperonin function from studies on archaeal thermosomes. Biochem Soc T *39*, 94-98.

Lundin, V.F., Srayko, M., Hyman, A.A., and Leroux, M.R. (2008). Efficient chaperone-mediated tubulin biogenesis is essential for cell division and cell migration in C. elegans. Developmental biology *313*, 320-334.

Lundin, V.F., Stirling, P.C., Gomez-Reino, J., Mwenifumbo, J.C., Obst, J.M., Valpuesta, J.M., and Leroux, M.R. (2004). Molecular clamp mechanism of substrate binding by hydrophobic coiled-coil residues of the archaeal chaperone prefoldin. Proc Natl Acad Sci U S A *101*, 4367-4372.

Ma, J., Sigler, P.B., Xu, Z., and Karplus, M. (2000). A dynamic model for the allosteric mechanism of GroEL. J Mol Biol *302*, 303-313.

Martin-Benito, J., Bertrand, S., Hu, T., Ludtke, P.J., McLaughlin, J.N., Willardson, B.M., Carrascosa, J.L., and Valpuesta, J.M. (2004). Structure of the complex between the cytosolic chaperonin CCT and phosducin-like protein. Proc Natl Acad Sci U S A *101*, 17410-17415.

Martin-Benito, J., Boskovic, J., Gomez-Puertas, P., Carrascosa, J.L., Simons, C.T., Lewis, S.A., Bartolini, F., Cowan, N.J., and Valpuesta, J.M. (2002). Structure of eukaryotic prefoldin and of its complexes with unfolded actin and the cytosolic chaperonin CCT. EMBO J *21*, 6377-6386.

Martin-Benito, J., Grantham, J., Boskovic, J., Brackley, K.I., Carrascosa, J.L., Willison, K.R., and Valpuesta, J.M. (2007). The inter-ring arrangement of the cytosolic chaperonin CCT. EMBO reports *8*, 252-257.

Mayer, M.P. (2013). Hsp70 chaperone dynamics and molecular mechanism. Trends Biochem Sci *38*, 507-514.

Mayr, C., Richter, K., Lilie, H., and Buchner, J. (2000). Cpr6 and Cpr7, two closely related Hsp90-associated immunophilins from Saccharomyces cerevisiae, differ in their functional properties. The Journal of biological chemistry *275*, 34140-34146.

McClellan, A.J., Xia, Y., Deutschbauer, A.M., Davis, R.W., Gerstein, M., and Frydman, J. (2007). Diverse cellular functions of the Hsp90 molecular chaperone uncovered using systems approaches. Cell *131*, 121-135.

McLaughlin, J.N., Thulin, C.D., Hart, S.J., Resing, K.A., Ahn, N.G., and Willardson, B.M. (2002). Regulatory interaction of phosducin-like protein with the cytosolic chaperonin complex. Proc Natl Acad Sci U S A *99*, 7962-7967.

Melville, M.W., McClellan, A.J., Meyer, A.S., Darveau, A., and Frydman, J. (2003). The Hsp70 and TRiC/CCT chaperone systems cooperate in vivo to assemble the von Hippel-Lindau tumor suppressor complex. Molecular and cellular biology *23*, 3141-3151.

Meyer, A.S., Gillespie, J.R., Walther, D., Millet, I.S., Doniach, S., and Frydman, J. (2003). Closing the folding chamber of the eukaryotic chaperonin requires the transition state of ATP hydrolysis. Cell *113*, 369-381.

Munoz, I.G., Yebenes, H., Zhou, M., Mesa, P., Serna, M., Park, A.Y., Bragado-Nilsson, E., Beloso, A., de Carcer, G., Malumbres, M.*, et al.* (2011). Crystal structure of the open conformation of the mammalian chaperonin CCT in complex with tubulin. Nat Struct Mol Biol *18*, 14-19.

Murray, J., Savva, C.G., Shin, B.S., Dever, T.E., Ramakrishnan, V., and Fernandez, I.S. (2016). Structural characterization of ribosome recruitment and translocation by type IV IRES. Elife *5*.

Netzer, W.J., and Hartl, F.U. (1998). Protein folding in the cytosol: chaperonin-dependent and -independent mechanisms. Trends Biochem Sci *23*, 68-73.

Noble, C.G., and Song, H. (2008). Structural studies of elongation and release factors. Cellular and molecular life sciences : CMLS *65*, 1335-1346.

Ortiz, P.A., Ulloque, R., Kihara, G.K., Zheng, H., and Kinzy, T.G. (2006). Translation elongation factor 2 anticodon mimicry domain mutants affect fidelity and diphtheria toxin resistance. The Journal of biological chemistry *281*, 32639-32648.

Pappenberger, G., Wilsher, J.A., Roe, S.M., Counsell, D.J., Willison, K.R., and Pearl, L.H. (2002). Crystal structure of the CCTgamma apical domain: implications for substrate binding to the eukaryotic cytosolic chaperonin. J Mol Biol *318*, 1367-1379.

Pereira, J.H., Ralston, C.Y., Douglas, N.R., Meyer, D., Knee, K.M., Goulet, D.R., King, J.A., Frydman, J., and Adams, P.D. (2010). Crystal structures of a group II chaperonin reveal the open and closed states associated with the protein folding cycle. The Journal of biological chemistry *285*, 27958-27966.

Perentesis, J.P., Phan, L.D., Gleason, W.B., LaPorte, D.C., Livingston, D.M., and Bodley, J.W. (1992). Saccharomyces cerevisiae elongation factor 2. Genetic cloning, characterization of expression, and G-domain modeling. The Journal of biological chemistry *267*, 1190-1197.

Phan, L.D., Perentesis, J.P., and Bodley, J.W. (1993). Saccharomyces cerevisiae elongation factor 2. Mutagenesis of the histidine precursor of diphthamide yields a functional protein that is resistant to diphtheria toxin. The Journal of biological chemistry *268*, 8665-8668.

Plaxco, K.W., Riddle, D.S., Grantcharova, V., and Baker, D. (1998). Simplified proteins: minimalist solutions to the 'protein folding problem'. Curr Opin Struct Biol *8*, 80-85.

Preissler, S., and Deuerling, E. (2012). Ribosome-associated chaperones as key players in proteostasis. Trends Biochem Sci *37*, 274-283.

Ratzke, C., Nguyen, M.N., Mayer, M.P., and Hugel, T. (2012). From a ratchet mechanism to random fluctuations evolution of Hsp90's mechanochemical cycle. J Mol Biol *423*, 462-471.

Reichen, C., Hansen, S., and Pluckthun, A. (2014). Modular peptide binding: from a comparison of natural binders to designed armadillo repeat proteins. J Struct Biol *185*, 147-162.

Reissmann, S., Joachimiak, L.A., Chen, B., Meyer, A.S., Nguyen, A., and Frydman, J. (2012). A gradient of ATP affinities generates an asymmetric power stroke driving the chaperonin TRIC/CCT folding cycle. Cell Rep *2*, 866-877.

Reissmann, S., Parnot, C., Booth, C.R., Chiu, W., and Frydman, J. (2007). Essential function of the built-in lid in the allosteric regulation of eukaryotic and archaeal chaperonins. Nat Struct Mol Biol *14*, 432-440.

Rivenzon-Segal, D., Wolf, S.G., Shimon, L., Willison, K.R., and Horovitz, A. (2005). Sequential ATP-induced allosteric transitions of the cytoplasmic chaperonin containing TCP-1 revealed by EM analysis. Nat Struct Mol Biol *12*, 233-237.

Rizzolo, K., Huen, J., Kumar, A., Phanse, S., Vlasblom, J., Kakihara, Y., Zeineddine, H.A., Minic, Z., Snider, J., Wang, W.*, et al.* (2017). Features of the Chaperone Cellular Network Revealed through Systematic Interaction Mapping. Cell Rep *20*, 2735-2748.

Rizzolo, K., Kumar, A., Kakihara, Y., Phanse, S., Minic, Z., Snider, J., Stagljar, I., Zilles, S., Babu, M., and Houry, W.A. (2018). Systems analysis of the genetic interaction network of yeast molecular chaperones. Mol Omics *14*, 82-94.

Rodriguez-Pena, J.M., Cid, V.J., Sanchez, M., Molina, M., Arroyo, J., and Nombela, C. (1998). The deletion of six ORFs of unknown function from Saccharomyces cerevisiae chromosome VII reveals two essential genes: YGR195w and YGR198w. Yeast *14*, 853-860.

Russmann, F., Stemp, M.J., Monkemeyer, L., Etchells, S.A., Bracher, A., and Hartl, F.U. (2012). Folding of large multidomain proteins by partial encapsulation in the chaperonin TRiC/CCT. Proc Natl Acad Sci U S A *109*, 21208-21215.

Saibil, H.R., Fenton, W.A., Clare, D.K., and Horwich, A.L. (2013). Structure and Allostery of the Chaperonin GroEL. J Mol Biol *425*, 1476-1487.

Schlecht, U., Miranda, M., Suresh, S., Davis, R.W., and St Onge, R.P. (2012). Multiplex assay for condition-dependent changes in protein-protein interactions. Proc Natl Acad Sci U S A *109*, 9213-9218.

Schoehn, G., Hayes, M., Cliff, M., Clarke, A.R., and Saibil, H.R. (2000). Domain rotations between open, closed and bullet-shaped forms of the thermosome, an archaeal chaperonin. J Mol Biol *301*, 323-332.

Schopf, F.H., Biebl, M.M., and Buchner, J. (2017). The HSP90 chaperone machinery. Nat Rev Mol Cell Biol *18*, 345-360.

Shimizu, Y., Kuruma, Y., Ying, B.W., Umekage, S., and Ueda, T. (2006). Cell-free translation systems for protein engineering. FEBS J *273*, 4133-4140.

Shomura, Y., Yoshida, T., Iizuka, R., Maruyama, T., Yohda, M., and Miki, K. (2004). Crystal structures of the group II chaperonin from Thermococcus strain KS-1: steric hindrance by the substituted amino acid, and inter-subunit rearrangement between two crystal forms. J Mol Biol *335*, 1265-1278.

Siegers, K., Waldmann, T., Leroux, M.R., Grein, K., Shevchenko, A., Schiebel, E., and Hartl, F.U. (1999). Compartmentation of protein folding in vivo: sequestration of non-native polypeptide by the chaperonin-GimC system. EMBO J *18*, 75-84.

Skjaerven, L., Cuellar, J., Martinez, A., and Valpuesta, J.M. (2015). Dynamics, flexibility, and allostery in molecular chaperonins. FEBS letters *589*, 2522-2532.

Spahn, C.M., Gomez-Lorenzo, M.G., Grassucci, R.A., Jorgensen, R., Andersen, G.R., Beckmann, R., Penczek, P.A., Ballesta, J.P., and Frank, J. (2004). Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA translocation. EMBO J *23*, 1008-1019.

Spiess, C., Meyer, A.S., Reissmann, S., and Frydman, J. (2004). Mechanism of the eukaryotic chaperonin: protein folding in the chamber of secrets. Trends in cell biology *14*, 598-604.

Spiess, C., Miller, E.J., McClellan, A.J., and Frydman, J. (2006). Identification of the TRiC/CCT substrate binding sites uncovers the function of subunit diversity in eukaryotic chaperonins. Mol Cell *24*, 25-37.

Sternlicht, H., Farr, G.W., Sternlicht, M.L., Driscoll, J.K., Willison, K., and Yaffe, M.B. (1993). The t-complex polypeptide 1 complex is a chaperonin for tubulin and actin in vivo. Proc Natl Acad Sci U S A *90*, 9422-9426.

Stirling, P.C., Cuellar, J., Alfaro, G.A., El Khadali, F., Beh, C.T., Valpuesta, J.M., Melki, R., and Leroux, M.R. (2006). PhLP3 modulates CCT-mediated actin and tubulin folding via ternary complexes with substrates. The Journal of biological chemistry *281*, 7012-7021.

Stirling, P.C., Srayko, M., Takhar, K.S., Pozniakovsky, A., Hyman, A.A., and Leroux, M.R. (2007). Functional interaction between phosducin-like protein 2 and cytosolic chaperonin is essential for cytoskeletal protein function and cell cycle progression. Molecular biology of the cell *18*, 2336-2345.

Stoldt, V., Rademacher, F., Kehren, V., Ernst, J.F., Pearce, D.A., and Sherman, F. (1996). Review: The Cct eukaryotic chaperonin subunits of Saccharomyces cerevisiae and other yeasts. Yeast *12*, 523-529.

Stuart, S.F., Leatherbarrow, R.J., and Willison, K.R. (2011). A two-step mechanism for the folding of actin by the yeast cytosolic chaperonin. The Journal of biological chemistry *286*, 178-184.

Su, X., Lin, Z., and Lin, H. (2013). The biosynthesis and biological function of diphthamide. Critical reviews in biochemistry and molecular biology *48*, 515-521.

Taguchi, H. (2015). Reaction Cycle of Chaperonin GroEL via Symmetric "Football" Intermediate. J Mol Biol *427*, 2912-2918.

Tam, S., Geller, R., Spiess, C., and Frydman, J. (2006). The chaperonin TRiC controls polyglutamine aggregation and toxicity through subunit-specific interactions. Nature cell biology *8*, 1155-1162.

Tang, Y.C., Chang, H.C., Chakraborty, K., Hartl, F.U., and Hayer-Hartl, M. (2008). Essential role of the chaperonin folding compartment in vivo. EMBO J *27*, 1458-1468.

Tang, Y.C., Chang, H.C., Roeben, A., Wischnewski, D., Wischnewski, N., Kerner, M.J., Hartl, F.U., and Hayer-Hartl, M. (2006). Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. Cell *125*, 903-914.

Tarassov, K., Messier, V., Landry, C.R., Radinovic, S., Serna Molina, M.M., Shames, I., Malitskaya, Y., Vogel, J., Bussey, H., and Michnick, S.W. (2008). An in vivo map of the yeast protein interactome. Science *320*, 1465-1470.

Techtmann, S.M., and Robb, F.T. (2010). Archaeal-like chaperonins in bacteria. Proc Natl Acad Sci U S A *107*, 20269-20274.

Tesic, M., Marsh, J.A., Cullinan, S.B., and Gaber, R.F. (2003). Functional interactions between Hsp90 and the co-chaperones Cns1 and Cpr7 in Saccharomyces cerevisiae. The Journal of biological chemistry *278*, 32692-32701.

Tian, G., Vainberg, I.E., Tap, W.D., Lewis, S.A., and Cowan, N.J. (1995). Specificity in chaperonin-mediated protein folding. Nature *375*, 250-253.

Vainberg, I.E., Lewis, S.A., Rommelaere, H., Ampe, C., Vandekerckhove, J., Klein, H.L., and Cowan, N.J. (1998). Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin. Cell *93*, 863-873.

Villebeck, L., Moparthi, S.B., Lindgren, M., Hammarstrom, P., and Jonsson, B.H. (2007). Domain-specific chaperone-induced expansion is required for beta-actin folding: a comparison of beta-actin conformations upon interactions with GroEL and tail-less complex polypeptide 1 ring complex (TRiC). Biochemistry *46*, 12639-12647.

Voorhees, R.M., and Ramakrishnan, V. (2013). Structural basis of the translational elongation cycle. Annu Rev Biochem *82*, 203-236.

Walter, S., and Buchner, J. (2002). Molecular chaperones - Cellular machines for protein folding. Angew Chem Int Edit *41*, 1098-1113.

Walzthoeni, T., Joachimiak, L.A., Rosenberger, G., Rost, H.L., Malmstrom, L., Leitner, A., Frydman, J., and Aebersold, R. (2015). xTract: software for characterizing conformational changes of protein complexes by quantitative cross-linking mass spectrometry. Nat Methods *12*, 1185-1190.

Wayne, N., and Bolon, D.N. (2007). Dimerization of Hsp90 is required for in vivo function. Design and analysis of monomers and dimers. The Journal of biological chemistry *282*, 35386-35395.

Willardson, B.M., and Howlett, A.C. (2007). Function of phosducin-like proteins in G protein signaling and chaperone-assisted protein folding. Cellular signalling *19*, 2417-2427.

Wilson, D.N., and Beckmann, R. (2011). The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling. Curr Opin Struc Biol *21*, 274-282.

Xu, Z.H., Horwich, A.L., and Sigler, P.B. (1997). The crystal structure of the asymmetric GroEL-GroES-(ADP)(7) chaperonin complex. Nature *388*, 741-750.

Yam, A.Y., Xia, Y., Lin, H.T., Burlingame, A., Gerstein, M., and Frydman, J. (2008). Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies. Nat Struct Mol Biol *15*, 1255-1262.

Yan, C., Hang, J., Wan, R., Huang, M., Wong, C.C., and Shi, Y. (2015). Structure of a yeast spliceosome at 3.6-angstrom resolution. Science *349*, 1182-1191.

Yates, S.P., Jorgensen, R., Andersen, G.R., and Merrill, A.R. (2006). Stealth and mimicry by deadly bacterial toxins. Trends Biochem Sci *31*, 123-133.

Yebenes, H., Mesa, P., Munoz, I.G., Montoya, G., and Valpuesta, J.M. (2011). Chaperonins: two rings for folding. Trends Biochem Sci *36*, 424-432.

Yoshikawa, K., Tanaka, T., Ida, Y., Furusawa, C., Hirasawa, T., and Shimizu, H. (2011). Comprehensive phenotypic analysis of single-gene deletion and overexpression strains of Saccharomyces cerevisiae. Yeast *28*, 349-361.

Zang, Y., Jin, M., Wang, H., Cui, Z., Kong, L., Liu, C., and Cong, Y. (2016). Staggered ATP binding mechanism of eukaryotic chaperonin TRiC (CCT) revealed through high-resolution cryo-EM. Nat Struct Mol Biol *23*, 1083-1091.

Zhang, G., and Ignatova, Z. (2011). Folding at the birth of the nascent chain: coordinating translation with co-translational folding. Curr Opin Struc Biol *21*, 25-31.

Zhang, H.Q., Amick, J., Chakravarti, R., Santarriaga, S., Schlanger, S., McGlone, C., Dare, M., Nix, J.C., Scaglione, K.M., Stuehr, D.J.*, et al.* (2015). A Bipartite Interaction between Hsp70 and CHIP Regulates Ubiquitination of Chaperoned Client Proteins. Structure *23*, 472-482.

Zhang, J., Baker, M.L., Schroder, G.F., Douglas, N.R., Reissmann, S., Jakana, J., Dougherty, M., Fu, C.J., Levitt, M., Ludtke, S.J.*, et al.* (2010a). Mechanism of folding chamber closure in a group II chaperonin. Nature *463*, 379-383.

Zhang, J., Ma, B., DiMaio, F., Douglas, N.R., Joachimiak, L.A., Baker, D., Frydman, J., Levitt, M., and Chiu, W. (2011). Cryo-EM structure of a group II chaperonin in the prehydrolysis ATP-bound state leading to lid closure. Structure *19*, 633-639.

Zhang, Y., Zhu, X., Torelli, A.T., Lee, M., Dzikovski, B., Koralewski, R.M., Wang, E., Freed, J., Krebs, C., Ealick, S.E.*, et al.* (2010b). Diphthamide biosynthesis requires an organic radical generated by an iron-sulphur enzyme. Nature *465*, 891-896.