PHILOSOPHICAL TRANSACTIONS
OF THE ROYAL SOCIETY B
BIOLOGICAL SCIENCES

The substrate specificity of eukaryotic cytosolic chaperonin CCT

Journal:	*Philosophical Transactions B*
Manuscript ID	RSTB-2017-0192.R2
Article Type:	Review
Date Submitted by the Author:	n/a
Complete List of Authors:	Willison, Keith; Imperial College London
Issue Code (this should have already been entered but please contact the Editorial Office if it is not present):	ALLOSTERY
Subject:	Biochemistry < BIOLOGY
Keywords:	protein folding, chaperonin, actin, tubulin, WD40 propeller

Phil. Trans. R. Soc. B. article template

PHILOSOPHICAL TRANSACTIONS B

Phil. Trans. R. Soc. B.
doi:10.1098/not yet assigned

The Substrate Specificity of Eukaryotic Cytosolic Chaperonin CCT

Keith R Willison

Department of Chemistry, Imperial College London, South Kensington Campus, London SW7 2AZ, UK
ORCID ID:0000-0002-1353-1471

Keywords: Chaperonin CCT, protein folding, actin, tubulin, 7-bladed WD40 propellers, APC/C

Abstract

The cytosolic chaperonin CCT (Chaperonin Containing TCP-1) is an ATP-dependent double-ring protein machine mediating the folding of members of the eukaryotic cytoskeletal protein families. The actins and tubulins are obligate substrates of CCT because they are completely dependent on CCT activity to reach their native states. Genetic and proteomic analysis of the CCT interactome in the yeast *Saccharomyces cerevisiae* revealed a CCT-network of ~300 genes and proteins involved in many fundamental biological processes. We classified network members into sets such as substrates, CCT co-factors and CCT-mediated assembly processes. Many members of the 7-bladed propeller family of proteins are commonly found tightly bound to CCT isolated from human and plant cells and yeasts. The anaphase promoting complex (APC/C) co-factor propellers, Cdh1p and Cdc20p, are also obligate substrates since they both require CCT for folding and functional activation. *In vitro* translation analysis, in prokaryotic and eukaryotic cell extracts, of a set of yeast propellers demonstrates their highly differential interactions with CCT and GroEL. Individual propeller proteins have idiosyncratic interaction modes with CCT because they emerged independently with neo-functions many times throughout eukaryotic evolution. We present a toy model in which cytoskeletal protein biogenesis and folding flux through CCT couples cell growth and size control to time dependent cell cycle mechanisms. (200 words)

Introduction

The cytosolic chaperonin CCT is a 1-MDa multi-subunit protein complex which functions in cytoskeletal protein folding in all eukaryotes (Willison; 1999; Valpuesta et al; 2002, 2005). CCT stands for Chaperonin Containing Tailless complex polypeptide 1 (TCP-1), and is also named TRiC (TCP-1 Ring Complex). CCT is a Group II chaperonin and like all chaperonins has a double-toroid shape enclosing a central cavity and can, upon binding and hydrolysis of ATP, assist in the folding of non-native proteins or in the assembly of substrates with their respective partners. CCT is constructed from eight protein subunits, encoded by eight homologous genes, positioned in each ring in a precise arrangement (Dekker et al; 2011a, Kalisman and Levitt; 2012). Each CCT subunit consists of three domains: the equatorial domain harbouring the ATP binding site, an intermediate domain, and the apical domain involved in substrate binding, as first identified in the X-ray crystal structure of the Group I chaperonin GroEL (Braig et al; 1994). The sequence differences between the eight CCT subunits are located mainly in their apical domains (Pappenberger et al; 2002, Dekker et al; 2011a) which provide the substrate binding specificity of each subunit (Kubota et al; 1994). The selective nature of the CCT machinery and its individual subunits was revealed in cryo-EM maps of CCT-actin (Llorca et al; 1999) and CCT-tubulin complexes (Llorca et al; 2000, 2001) and through direct biochemical analyses of actin binding sites (Hynes and Willison; 2000, McCormack et al; 2001, Neirynck et al; 2006). CCT shows sequential allosteric behaviour in its ATP-binding and hydrolysis regime (Horovitz and Willison; 2005, Gruber et al; 2017) and its actin folding regime (Llorca et al;

*Author for correspondence (keith.willison@imperial.ac.uk).

2001, Dekker et al; 2011a, Stuart et al; 2011). Our long-standing view is that CCT actively folds a very restricted group of non-cytoskeletal proteins which includes many members of the 7-bladed WD40-propeller repeat protein family. The core set of obligate CCT substrates, together with their co-factors and other binding partners, is involved in ancient and fundamental biological processes including transcription, chromatin modification and secretion. Figure 1 shows the core CCT-interactome in pictorial form to aid the reader grasp the depth and breadth of the system.

Willison (2018) has reviewed the CCT structure and allosteric mechanism and its role in normal biological processes and in disease. New data on the WD40 interactions with CCT in *Saccharomyces cerevisiae* CCT and mutagenesis of CCT-binding motifs in the Cdh1 co-activator of APC/C are presented and reviewed in this paper. We then calculate the folding capacity of the CCT pool and the variance in CCT complex copy number in growing yeast and propose that access to CCT interaction may be a mechanism through which proteins can detect the instantaneous flux of actin and tubulin synthesis via CCT-binding site availability manifested through their individual effective association rate constants.

The CCT Interactome

All eight CCT genes are essential in yeast (Dekker et al; 2008) and even minor perturbations of any of their ATP binding site motifs (Amit et al; 2010) and ATP-hydrolysis kinetics (Shimon et al; 2008) result in loss of viability and severe growth defects which include abnormally large cell sizes and defective cell shapes. Although the genetic and physical CCT-interactome comprises ~300 members in yeast (Dekker et al; 2008) most of the components are not actively folded by CCT but are co-factors and ancillary genetic effectors. The reason for the existence of such a large and complex network is that CCT has become coupled to many fundamental biological processes because it co-evolved with, and facilitated the novel biophysical properties of the emerging cytoskeleton in the primitive eukaryotic cell (Kubota et al; 1994, Willison 1999, Dekker et al; 2011b). Recent independent computational analysis (Aswathy et al; 2016) and global experiments in yeast concur with the view that CCT is highly selective towards the cytoskeleton in its network specificity (Rizzolo et al; 2017).

We produced the comprehensive yeast CCT 'interactome' by employing mass spectrometric proteomics experiments using CCT-3CBP and CCT-6CBP tagged complexes and a Synthetic Genetic Array (SGA) screen for synthetic lethal interactions with a temperature-sensitive *cct1-2* allele (Dekker et al; 2008). Using the combined results of these integrative approaches, a large number of novel CCT physical and genetic interactors were identified, defining several functional interaction networks of CCT. Integral to the proteomics experiment was the focus on ATP-dependent release from CCT, which was taken as a potential indicator of interacting proteins that are folding substrates. Before subjecting CCT to LC-ESI-MS, the chaperonin, whilst bound to calmodulin-resin through the affinity tag in the CCT3 or the CCT6 subunit, was washed with either a buffer containing ATP or a plain buffer. The rationale behind this was that CCT, which is still active when bound to calmodulin-resin, will fold and release its substrates only in the presence of ATP, and the released substrates that are subsequently washed away will not pass through the mass spectrometer. We identified many ATP-elutable CCT interacting proteins, of which the most abundant matches included known CCT interacting proteins, namely the folding substrates actin and tubulin, and their respective folding regulators Plp2p and Plp1p (Stirling et al; 2007, McCormack et al; 2009). Dekker et al (2008) only identified one WD40 protein in the physical interaction set but subsequent higher resolution LC-ESI-MS analysis in our group, using more advanced mass spectrometers, routinely recovers the following avid CCT-binding WD40 proteins: Vid27p, Sec31p, Spt8p, Sec13p and Cdc55p. It is difficult to estimate the coverage of the MS analysis in our work but the dataset is certainly incomplete for many technical reasons such as variable biochemical stringency of CCT interactions and low cellular

abundance of some partner proteins such as the APC/C activators, Cdh1p and Cdc20p, which are dependent on CCT for their folding and maturation *in vivo* (Camasses et al; 2003) and *in vitro* (Passmore et al; 2003). Arava et al (2003) calculated protein synthesis rates from 5701 mRNAs in yeast; Actp is synthesised at a rate of 2.2 proteins per second whereas Cdh1p and Cdc20p are synthesised 100x and 366x more slowly. It also is unsurprising that we did not detect CCT-bound Cdh1p and Cdc20p because SILAC mass spectrometry fails to recover these two proteins from over 4000 proteins identified (de Godoy et al; 2008). In the early global protein interaction screen in yeast the 16 CCT-binding WD40 proteins identified were all over-expressed as high copy number baits supporting the suggestion that many of the WD40 interactors are indeed low copy number proteins (Ho et al; 2002, Valpuesta et al; 2002). The core evolutionarily conserved WD40 proteins are components of the Type II phosphatase system (Cdc55p) and TOR complex (Lst8), RNA polymerase II and chromatin modification (Taf5p), RNA processing (U3 snoRNP), secretion (Coatomer I and II) cell cycle control through protein degradation control (Cdh1p and Cdc20p) as defined in Dekker et al (2008) and sketched in Figure 1.

7-bladed WD40 repeat protein family

The WD-40 repeat proteins contain multiple, tandem copies of a ~40 amino acid repeat which forms a blade composed of four anti-parallel β-sheets (Figure 2). The WD-repeat begins with a glycine-histidine (GH) dipeptide and ends with the tryptophan-aspartic acid (WD) dipeptide, hence the name (Neer et al; 1994; Smith et al; 1999). Individual WD-repeats show great sequence variation at every residue position and even the GH and WD dipeptides are not invariant (Smith et al; 1999) as observed for Tup1p in the alignment in Figure 2. In addition, the number of tandem repeats differs and variable-sized insertions are found both within and between them, which makes the bioinformatic identification and characterisation of WD-repeat proteins very difficult. For example, Ski8p, a protein regulating mRNA degradation, had been predicted to contain five WD repeats in its primary sequence but crystallographic analysis revealed seven WD repeats within the sequence and the two extra repeats are so divergent from the predicted WD-consensus sequence that they cannot be detected from the protein's primary sequence (Madrona and Wilson; 2004). Although there are examples of WD-proteins having 8-blades, such as Cdc4p (Orlicky et al; 2003) and Sif2p (Cerna and Wilson; 2005) the most common arrangement is for seven tandem WD-repeats to form the circularized β-propeller structure as in Tup1p (Figure 2). A newly constructed database of WD40-repeat proteins (WDSPdb) lists 83 yeast proteins containing more than 6 repeats (Wang et al; 2015). We have constructed a 30-member sub-family of homologous 7-bladed WD40 proteins in *Saccharomyces cerevisiae* using sequential Psi-BLAST homology searching, Phyre analysis and hand alignment (Table 1). The members of this family in *Saccharomyes cerevisiae* participate in many different cellular functions but the common property of the WD40 propeller domains is that they mediate protein-protein interactions, often as components of large protein complexes. The β-sheets of the propeller structure provide a stable platform for protein-protein interaction via the loops that connect them. Structure/function analysis of Doa1p shows that its propeller binds ubiquitin as do those from Cop1p, Sec27p and Taf5p (Pashkova et al; 2010) and the APC/C co-factors. No enzyme activity has yet been attributed to any propeller domain.

Early studies on folding of WD40-repeat proteins from the group of Neer lead to the 'Velcro' model of folding (Garcia-Higuera et al; 1996) so named because it was noticed that the 7th blade of the propeller structure is composed of three contiguous β-strands located at the C-terminus of the sequence but the fourth is encoded at the N-terminus of the repeats. The two segments associate together in 3D space and interact to form the Velcro cap and thus close the propeller fold. Figure 2 shows structural details of the Tup1p Velcro cap. Bringing these four strands together in space to complete folding must be under careful kinetic control and it is likely that different WD40 folds have evolved with different folding speeds and therefore different chaperone dependencies due to kinetic

partitioning. Kinetic partitioning of proteins due to alterations of local folding rates in their folding pathways is likely to be responsible for GroE-dependent versus GroE-independent folding behaviour and accounts for the fact that single point mutations can convert MetK from *Ureaplasma urealyticum*, a bacterium with no GroE chaperonin system, into an obligate substrate of *Escherichia coli* GroEL/ES (Ishimoto et al; 2014).

When I first made the CCT-WD40 connection it was obvious that the CCT ring had the required shape and dimensions to facilitate the binding of the propeller (Valpuesta et al; 2002). Recent cryo-electron microscopy structures have shown that the human $G\beta_1$ propeller is near native and bound to the CCT6, CCT3, CCT1, CCT4 side of the ring of insect CCT (Plimpton et al; 2015). If higher resolution structures can demonstrate that the role of CCT in WD40 propeller folding turns out to be the one of permitting the two ends of the repeat to interact effectively to form blade 7 it would be analogous to our mechanistic model for the role of CCT in actin folding in which the manipulation of the C-terminus of the actin intermediate is the second critical reaction step preceding release of folded actin (Stuart et al; 2011, Dekker et al; 2011a, Olshina et al; 2016, Willison; 2018).

In vitro interactions between CCT and WD40 repeat proteins

In order to probe different WD40 repeat proteins for their CCT interaction behavior 15 yeast proteins (Table 2) from the core 30-member 7-bladed WD40 set (Table 1) and the 8-bladed Cdc4p, were tested. They were tested for CCT and GroEL binding using standard *in vitro* translation time course systems (McCormack et al; 2001, Pappenberger et al; 2006). The results are a clear demonstration of multiple modes of interaction with CCT (Table 2 and Figure 3). Some proteins do not bind CCT or GroEL whereas others bind both chaperonins strongly (sample assay data in Figure 3). Although 14 of these 16 proteins showed some degree of binding to CCT, only 5 demonstrated very strong interaction with it. These strong CCT interactors are Cdc55p, Vid27p, Taf5p, Cdh1p and Cdc20p each of which has been identified in other genetic and physical studies (Table 1 and Figure 1). It is striking that these proteins accumulate on CCT over time unlike actin which has a demonstrable precursor-product relationship in the rabbit reticulocyte lysate system because it is being actively processed and folded by CCT (McCormack et al; 2001). This reinforces a common role for CCT in the assembly, rather than folding, of these WD40 proteins with elements of their respective protein complexes, as is clearly the case for the Gβ-Gγ heterodimer assembly process (Plimpton et al; 2015). The stalling of the folding of WD40 propellers due to the blade 7 completion problem permits CCT to intervene in the process and provide a holding platform until binding partners arrive.

CCT and co-activators of the Anaphase Promoting Complex (APC/C)

Cdh1p and Cdc20p are co-activators of the Anaphase Promoting Complex/Cyclosome (APC/C), an ubiquitin ligase that has a prominent role in regulating cell cycle progression. The APC/C controls mitosis by regulating of entry into anaphase, progression through anaphase, exit of mitosis and G1 phase, by degrading sequentially different cell cycle proteins. It is a 1.5 MDa protein complex and contains 13 core subunits, and 3 related adaptors, two of which are Cdh1p and Ccd20p (Yu; 2007). The adaptor proteins regulate the APC/C's activity by binding to the complex to deliver target substrates for ubiquitination. The interaction of Cdh1p and Cdc20p with APC/C allows APC/C's ubiquitinase activity. The WD40 repeat domains of Cdh1 and Cdc20p have been shown to be critical in substrate recognition (Kraft et al; 2005, Burton et al; 2005).

CCT is the only known folding pathway for the production of functional and biologically active Cdh1p and Cdc20p *in vivo* (Camasses et al; 2003) and in *vitro* (Passmore et al; 2003). Camasses et al; (2003) mapped the CCT-interaction sites of Cdc20p to the propeller blades 3, 4 and 5. Some new data which maps CCT-binding sites in Cdh1p have been obtained. It was first shown that the Cdc20p

and Cdh1p propellers are interchangeable for CCT interaction by swapping them over and finding that the hybrid proteins are processed normally by rabbit CCT. Consequently mutations in evolutionarily conserved residues were screened for CCT interaction by *in vitro* translation time courses. Figure 4 shows the results. There are consequential CCT-interaction sites in blades 4 (N405A, D406A, N407A) and 5 (S436A, P437A). Disrupting the histidine (H490) of the tetrad motif in blade 6 arrests processing. Mutations in the 'velcro' interaction strand were discovered in blade 7D, R253A and V254A, which cause complete arrest of CCT processing of the full-length protein. This strongly suggests the important activity of CCT towards propellers is to facilitate the joining of the two ends of the protein domain together.

Connecting cytoskeletal protein biogenesis and folding flux to cellular networks

Proteins interact with the CCT system for reasons that include folding, for the actins and tubulins, and just-in-time assembly processes (Dekker; 2010). Trying to locate a newly synthesized protein about to leave one of the 200,000 ribosomes without a high-abundance and/or high-affinity detection system is a difficult task due to the law of mass action. For a protein such as Cdh1p, with a rate of synthesis 100 times less than actin, when considering the bulk cell synthesis rate not the translation rate, the ability to interact with one of the 3000 CCT complexes is effectively a 'concentrating' process that moves a nascent Cdh1 protein from one of 200,000 to one of 3000 points in cytoplasmic space and is readily accomplished because of the very high affinity that the Cdh1 folding intermediate has for CCT (Passmore et al; 2003). As is the case with actin, whose native state has no affinity for CCT, the Cdh1 protein can fold or assemble with other partners on CCT and then leave, never to return because its native state or assembled state has relinquished its affinity for CCT. It is likely that there are many variations on this theme controlled by kinetic partitioning and other time dependent processes such as chemical modifications. A further detection enhancement will occur if a binding partner also has independent binding affinity for separate CCT subunits because the combined affinities become multiplicative through addition of logarithmic association rates constants. This effect is observed for the yeast CCT-ACT1-PLP2 system where the PLP2 co-factor increases the rate of native actin production by 30-fold (McCormack et al; 2009).

We suggest that the availability of unoccupied and therefore unemployed CCT complexes will be a function of the flux through the CCT system of its two most abundant co-dependent partners, actin and tubulin. The number of available CCT binding sites can be used as a metric for any binding partner to register the instantaneous values of dActin/dt + dTubulin/dt through the copy number of the empty CCT complexes; physically manifested as the effective association rate constant. In *Saccharomyces cerevisiae* the number of available CCT complexes will be in the low thousands since CCT is present at between 3000 (Biochemical analysis; MacCormack et al; 2009) and 6000 (Mass spectrometry; Brownridge et al; 2013) copies per yeast cell. Using mass spectrometry analysis of pure CCT-substrate complexes (Dekker et al; 2008) and calculations our laboratory estimates that 50-60% of CCT is occupied with actin and tubulin folding (Figure 5A). This flux counting mechanism is probably the reason for actin and CCT subunit genes being amongst the most haplo-insufficient genes for growth in yeast YPD medium (Deutschbauer et al; 2005), as are the exosome and U3 snoRNP subunits which are components of the CCT-core (Figure 1H and J). Haplo-insufficiency is loss of fitness due to absence of one gene copy in a diploid. When gene copy number is increased in aneuploid yeast strains (Dephoure et al; 2014) many components of protein complexes show dosage compensation, including all the 137 haplo-insufficient ribosomal subunits. However, CCT subunit levels are not attenuated and continue to maintain the tightest distribution of all the 84 protein complexes analyzed. A further genetic test examines epistatic interactions between CCT and other genes. Epistasis is a measurement of the fitness of pairs of double mutants compared to wild-type and single mutants and has been measured for 61 gene pairs in Saccharomyces cerevisiae (He et

al; 2010). CCT2 is strongly positively epistatic with ACT1, NUS1, RPC10 and RP55). A direct way to measure the variance in CCT subunit levels is to use the protein signals obtained by comparing haploid versus diploid yeast in datasets from mass spectrometric SILAC experiments (de Goody et al; 2008). Figure 5B shows a plot of haploid versus diploid protein abundance ratios for CCT and prefoldin complexes and actin and tubulin and PLP2. The standard deviation of the eight CCT subunits (mean = 0.994) is 2.8 times less than the six prefoldin subunits (mean = 1.017) and 6.5 times less than the 27 member WD40 protein set (values in Table 1 – excluding Vid27p).

The tighter the number distributions of the main substrates, actin and tubulin, and the recipient CCT complex pool the more accurate and less noisy and more sensitive the flux measurement system becomes. In this respect it is also significant that actin, tubulin and CCT subunits sit in the centre of the Gaussian mean of the protein abundance distribution in yeast (Fig 5C) and protein half-life distribution (Fig 5D) because they will have median concentrations throughout the cytoplasm and median half-lives which are well in excess of the 3-hour cell cycle time for yeast growing in YPD medium at 30^{o}C. In conclusion, according to this model CCT is not only a physical gatekeeper between folding space and native space but also a device which can help transmit information about the state of the cell to the many elements of the essential protein complexes which bind it.

Discussion

All our experiments concerning the CCT substrate spectrum in mammalian cells and yeast teach us that CCT has a highly-restricted set of substrates. Other laboratories disagree with us and originally put the figure at 9-15% of all cytosolic proteins (EMBO J 1999). The most recent study suggests 7% of cytosolic proteins are substrates based up binding of *in vitro* denatured, pulse-chase-labelled total cell extracts to chaperonins *in vitro* (Yam et al; 2008). Our laboratory has performed pulse chase analysis of newly translated polypeptides in mammalian cells *in vivo* which never recovers more than 1% of the total protein counts bound to CCT (Sternlicht et al; 1993, Grantham and Willison, 2006). The labelled proteins that are bound are dominated by actin and tubulin signals. It is becoming even more difficult to sustain the view that CCT is a general cytosolic chaperone in the light of global gene and protein network analysis of the CCT interactome in yeast (Dekker et al; 2008, Rizzolo et al; 2017) which points to the central involvement of CCT in cytoskeletal function and several essential cellular networks involved in chromatin remodeling, secretory pathways, phosphatase activity and cell cycle regulation via control of the folding of the APC/C regulators. It is remarkable how in each network, always one or more CCT-dependent WD40 repeat proteins are involved (Figure 1).

Acknowledgments

Elizabeth McCormack provided technical assistance.

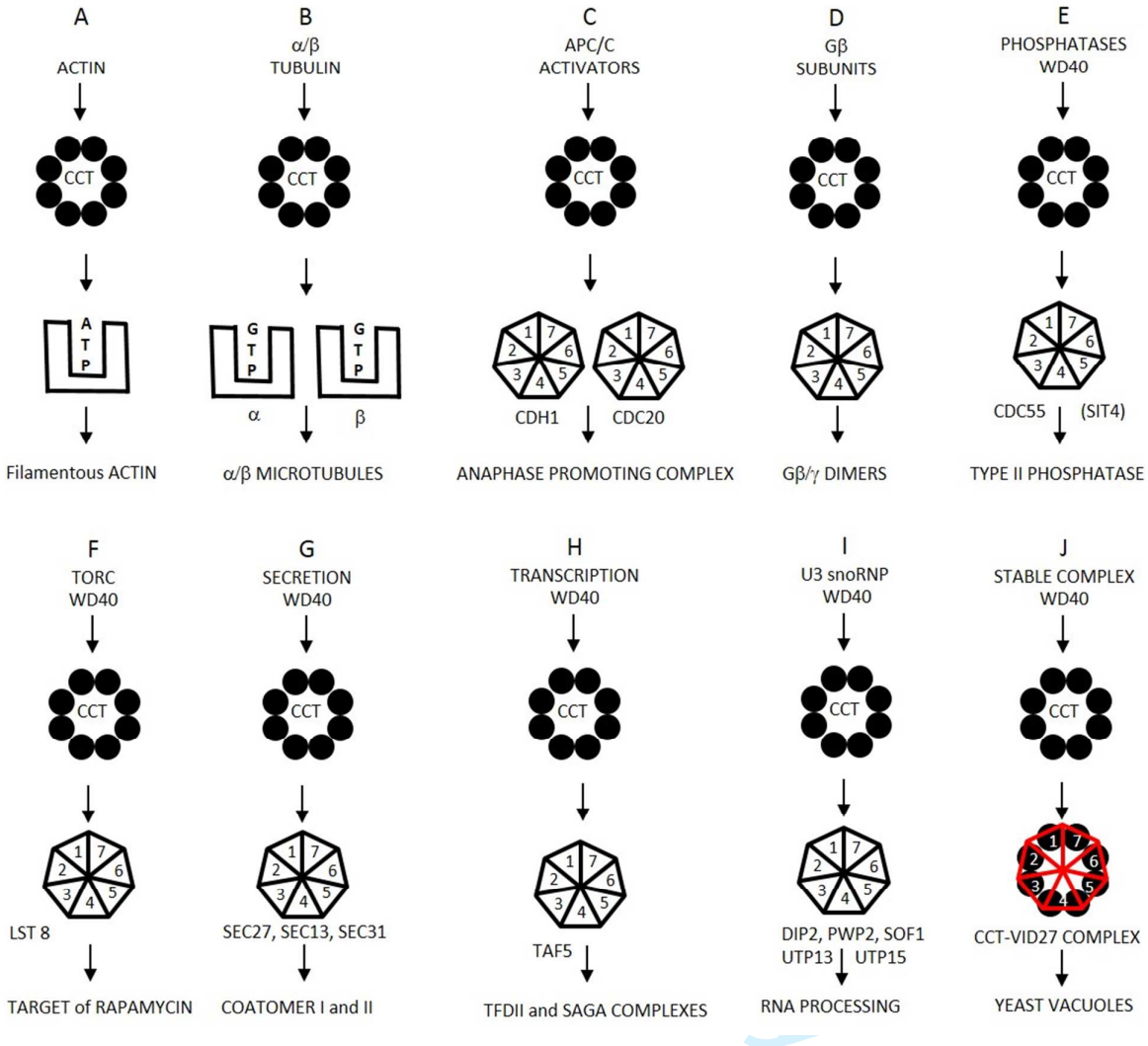
Competing Interests

'I have no competing interests.'

Funding

This work was funded by Cancer Research UK and an EPSRC Platform Grant.

FIGURE 1. Cartoons of schemes for CCT action upon its substrates and binding complex formation

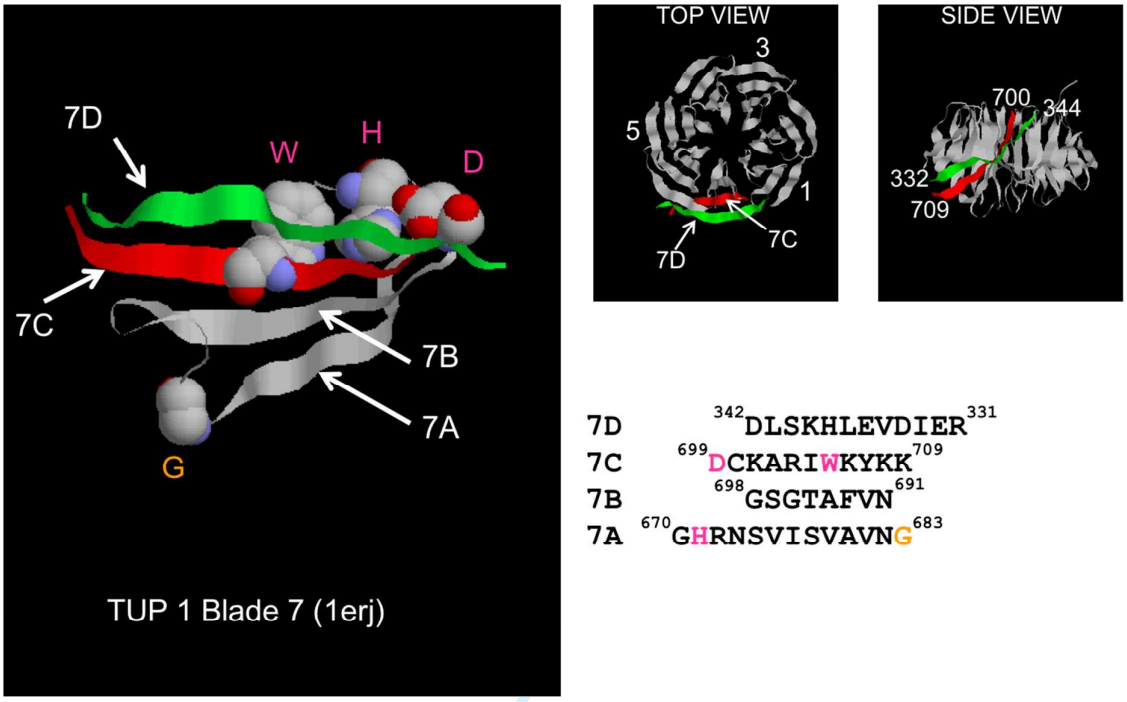


LEGEND Figure 1.

A) Direct folding to native state: Actin filaments (F-actin) are composed of a single monomer species (G-actin) and CCT activity yields assembly competent monomers. It is not yet known if ATP loading is concomitant with folding or whether nucleotide-free actin is released into the 2-5 mM ATP pool present in cell cytoplasm and then equilibrates with free ATP.
B) Folding and assembly: Tubulin filaments are composed of two monomer species (α-tubulin and β-tubulin) which both bind GTP. GTP concentration is low in cells (50μM) and CCT may well play a direct role in GTP loading of β-tubulin. There is strong evidence for post-CCT-dependent folding and assembly step(s), involving several co-factor proteins, which produce filament-assembly competent tubulin monomers.
C) Production of the APC activators: Cdc20p and Cdh1p propeller proteins are CCT-dependent.
D) Assembly of G-protein complexes (Willardson and Howlett; 2007).
E) Type II phosphatase complex – CCT interaction conserved from yeast to humans
F) TORC complex: WD40 subunits
G) Secretory complexes ER and Golgi

H) Transcription factor and several HDAC complexes interact with CCT (Dekker et al; 2008).
I) RNA processing
J) Holding activity of the *Saccharomyces cerevisiae* Vid27p propeller protein involved in autophagy.

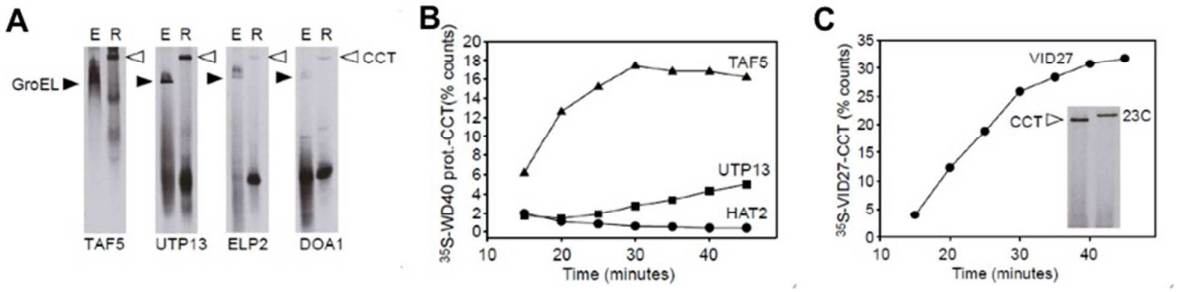
FIGURE 2. The Velcro cap



LEGEND Figure 2.

Structural model of the propeller of TUP1 (PDB: 1erj). TUP1 is a 713-amino acid residue protein with a single 7-bladed propeller. The left-hand panel shows a ribbon diagram with the W, D and H residues of the WD40 repeat highlighted in pink which interact to form a structural element connecting blade 7A and 7C across which blade 7D binds like a "velcro" cap (Neer et al; 1994). The two right-hand panels show top and side views of the propeller; a conical frustum. The amino acid sequences of the interaction region between the β-sheets are shown below with the four sequences aligned to reflect the registration of the blades in 3D space (K. R. Willison unpublished analysis).

FIGURE 3. Screening WD-40-repeat proteins for CCT interaction

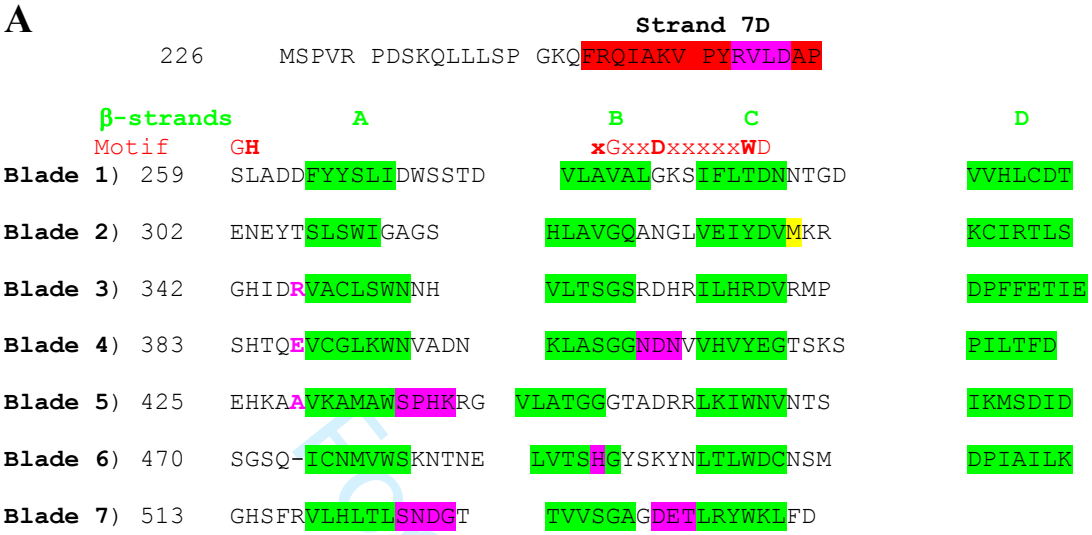


LEGEND Figure 3.

Panel (A). 6% native PAGE analysis of TAF5, UTP13, ELP2 and DOA1 after *in vitro* translation in *E. coli* lysate (E) or rabbit reticulocyte lysate (R) *in vitro* transcription/translation system to detect GroEL (filled arrowhead) or CCT interaction (open arrowhead) respectively.
Panel (B). Graphs of time courses of CCT-bound protein by 6% native PAGE analysis of rabbit reticulocyte *in vitro* translation assays of TAF5, UTP13 and HAT2.
Panel (C). 6% native PAGE analysis of time course of rabbit reticulocyte *in vitro* translation assay of VID27. Inset image of sample at the end of the time course showing interaction of CCT with Vid27p which can be shifted with the anti-TCP1 antibody 23C, demonstrating specificity of CCT binding. Vid27p is the most avid CCT binding protein in yeast (Aloy et al; 2004, Dekker et al; 2008, Rizzolo et al; 2017). Vid27 is a little studied, non-essential, WD40 repeat protein most closely related to BUB3 and possibly involved in vacuolar protein degradation and is found in yeasts and plants but not in higher eukaryotes.

Figure 4. Mapping CCT-binding residues in the Cdh1p propeller

A



B

Mutations	Residue numbers	Blade location	CCT interaction
CDH1 full-length	N405A, D406A, N407A	Blade 4 conserved	strong arrest/slow processing
CDH1 full-length	S436A and P437A	Blade 5 conserved	strong arrest/slow processing
CDH1 full-length	H438A and K439A	Blade 5 conserved	slight arrest/slow processing
CDH1 full-length	S524A and N525A	Blade 7 conserved	slight arrest/slow processing
CDH1 full-length	D526A and G527A	Blade 7 conserved	Arrested
CDH1 full-length	D536A, E537A, T538A	Blade 7 conserved	slight arrest/slow processing
CDH1 full-length	R253A and V254A	Strand 7D conserved	Arrested
CDH1 full-length	L255Aand D256A	Strand 7D conserved	slightly arrested
CDH1 M226-R566	R346A	Blade 3 conserved	Processed
CDH1 M226-R566	E387A	Blade 4 conserved	strong arrest/slow processing
CDH1 M226-R566	A429S	Blade 5 conserved	slight arrest/slow processing
CDH1 M226-R566	H490R	Blade 6 tetrad	Arrested
CDH1 M226-R566	D526A and G527A	Blade 7 conserved	Arrested
CDH1 M226-R566	D536A, E537A, T538A	Blade 7 conserved	Processed

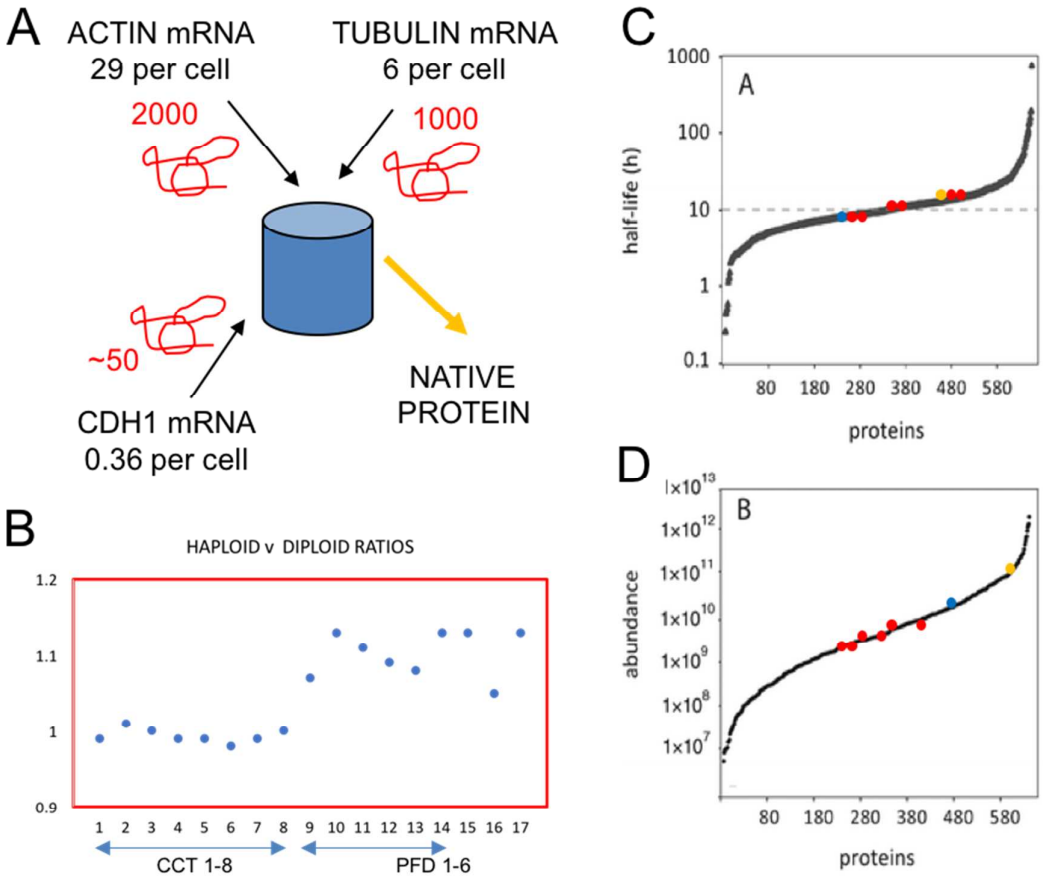
LEGEND Figure 4.

Mutants in the Cdh1p propeller (amino acid residues 226-566) were screened for strength of binding and degree of processing by CCT using *in vitro* translation/native gel analysis. Six mutations were tested in the context of the propeller domain only.

A. The β-strands in the blades are highlighted in green. Motif: the red bold letters highlight the residues (GH-G-D-WD) forming the hydrogen-bonding network ('structural tetrad') which stabilizes the blades. The M332A mutation, coloured yellow, is included to prevent internal translation initiation at this methionine which reduces the accuracy of quantitation in these assays (McCormack et al; 2009); the equivalent residue in human CDH1 is alanine. The point mutants which arrest the processing are highlighted in magenta.

B. List of 14 mutants and their CCT interaction behaviour. Strong arrest indicates that >15% of the counts become associated with CCT; slow processing means that the appearance of the folded Cdh1 (either full-length or the propeller domain) is retarded compared to wild-type control.

FIGURE 5. Yeast CCT Capacity



Legend Figure 5.

Number estimates for the components of this yeast model are from the following sources. Ribosomes (Warner 1999). Actin, Act1p (Garrels database), CCT complex (Willison laboratory biochemistry and Uniprot database). Calculated mRNA copy numbers per cell based upon 15000 transcripts per cell growing at 30°C - ACT1:28.97, CDH1: 0.36, TUB2:5.94, TCP1: 1.81, CCT2: 1.56, CCT3:1.62, CCT4:2.16, CCT5:2.24, CCT6: 2.61, CCT7: 2.23, CCT8: 1.7, CDH1:0.36 (Arava et al; 2003).

REFERENCE LIST

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.

Arava, Y., Wang, Y., Storey, J.D., Liu, C.L., Brown, P.O., and Herschlag, D. (2003). Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **100**, 3889-3894.

Aswathy, N., Pullepu, D., Kabir, M.A. (2016). The interactome of CCT complex – A computational analysis. *Computational Biology and Chemistry* **64**, 396–402

Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D.C., Joachimiak, A., Horwich, A.L., Sigler, P.B. (1994). The crystal-structure of the bacterial chaperonin GroEl at 2.8 A˚. *Nature* **371**, 578–586.

Brownridge, P., Lawless, C., Payapilly, A.B., Lanthaler, K., Holman, S.W., Harman, V.M., Grant, C.M., Beynon, R.J. and Hubbard, S.J. (2013). Quantitative analysis of chaperone network throughput in budding yeast. *Proteomics* **13**, 1276-1291. DOI 10.1002/pmic.201200412.

Burton, J. L., V. Tsakraklides, et al. (2005). Assembly of an APC-Cdh1-substrate complex is stimulated by engagement of a destruction box. *Mol Cell* **18**, 533-542.

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.

Cerna, D., and Wilson, D.K. (2005). The structure of Sif2p, a WD repeat protein functioning in the SET3 corepressor complex. *J Mol Biol* **351,** 923-935.

Chaudhuri, I., Soding, J., and Lupas, A.N. (2008). Evolution of the β-propeller fold. *Proteins* **71**, 795-803.

de Godoy, L.M.F., Olsen, J.V., Cox, J., Nielsen, M.L., Hubner, N.C., Frohlich, F., Walther, T.C and Mann, M. (2008). Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* **455**, 1251-1254.

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

Dekker, C. (2010). On the role of the chaperonin CCT in the just-in-time assembly process of APC/C^{Cdc20}. *Febs Letters* **584**, 477-481.

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

Dekker, C., Willison, K.R., and Taylor, W.R. (2011b). On the evolutionary origin of chaperonins. *Proteins* **79**, 1172-1192.

Dephoure, N., Hwang, S., O'Sullivan, C., Dodgson, S.E., Gygi, S.P., Amon, A., and Torres, E.M. (2014) Quantitative proteomic analysis reveals posttranslational responses to aneuploidy in yeast. eLife **3,** e03023. DOI: 10.7554/eLife.03023.

Deutschbauer, A.M., Jaramillo, D.F., Proctor, D.F, Kumm, J., Hillenmeyer, M.E., Davis, R.W., Nislow, C., and Giaever, G. (2005) Mechanisms of Haploinsufficiency Revealed by Genome-Wide Profiling in Yeast. *Genetics* **169**, 1915–1925.

Feldman, R.M., Correll, C.C., Kaplan, K.B., and Deshaies, R.J. (1997). A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p. *Cell* **91**, 221-230.

Grantham, J., Brackley, K. I., and Willison, K.R. (2006). Substantial CCT activity is required for cell cycle progression and cytoskeletal organization in mammalian cells, *Exp Cell Res* **312**, 2309–2324.

Gruber, R., Levitt, M., and Horovitz, A. (2017). Sequential allosteric mechanism of ATP hydrolysis by the CCT/TRiC chaperone is revealed through Arrhenius analysis. *Proc Natl Acad Sci* USA **14**, 5189–5194.

He, X., Qian, W., Wang, Z., Lil, Y., and Zhang, J (2010). Prevalent positive epistasis in *Escherichia coli* and *Saccharomyces cerevisiae* metabolic networks. *Nature Genetics* **42**, 272-276.

Ho, Y., A. Gruhler, et al. (2002). Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* **415**, 180-183.

Horovitz, A., Willison, K.R. (2005). Allosteric regulation of chaperonins. *Curr Opin Struct Biol* **15**, 646–651.

Hynes, G. M., and Willison, K. R. (2000) Individual subunits of the eukaryotic cytosolic chaperonin mediate interactions with binding sites located on subdomains of beta-actin. *J Biol Chem* **275**, 18985–18994

Ishimoto, T., Fujiwara, K., Niwa, T., and Taguchi, H. (2014). Conversion of a chaperonin GroEL-independent protein into an obligate substrate. *J Biol Chem* **289**, 32073-32080.

Kraft C, Vodermaier HC, Maurer-Stroh S, Eisenhaber F, Peters J-M. (2005). The WD40 propeller domain of Cdh1 functions as a destruction box receptor for APC/C substrates. *Molecular Cell* **18**, 543–553.

Kubota, H., Hynes, G., Carne, A., Ashworth, A., and Willison, K. (1994). Identification of six Tcp-1-related genes encoding divergent subunits of the TCP-1-containing chaperonin. *Curr. Biol* **4**, 89-99.

Llorca, O., McCormack, E.A., Hynes, G., Grantham, J., Cordell, J., Carrascosa, J.L., Willison, K.R., Fernandez, J.J., 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., 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., Martın-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

Madrona, A. Y. and D. K. Wilson (2004). The structure of Ski8p, a protein regulating mRNA degradation: Implications for WD protein structure. *Protein Sci* **13**, 1557-65.

McCormack, E. A., Llorca,O., Carrascosa,J.L., Valpuesta,J.M., and Willison, K. R. (2001) *J. Struct. Biol.* **135,** 198 –204

McCormack, E.A., Altschuler, G.M., Dekker, C., Filmore, H., 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.

Neer, E.J., Schmidt, C.J., Nambudripad, R., and Smith, T.F. 1994. The ancient regulatory-protein family of WD-repeat proteins. *Nature* **371,** 297-300.

Neirynck, K., Waterschoot, D., Vandekerckhove, J., Ampe, C., and Rommelaere, H. (2006) Actin interacts with CCT via discrete binding sites: a binding transition-release model for CCT-mediated actin folding. *J. Mol. Bio*l. **355**, 124–138.

Olshina, M.A., Baumann, A., Willison, K.R., and Jake Baum (2016) Plasmodium actin is incompletely folded by heterologous protein-folding machinery and likely requires the native Plasmodium chaperonin complex to enter a mature functional state. *FASEB JOURNAL* **30**, 405-416.

Orlicky, S., Tang, X., Willems, A., Tyers, M. and Sicheri, F. (2003) Structural basis for phosphodependent substrate selection and orientation by the SCFCdc4 ubiquitin ligase. *Cell* **112**, 243-256.

Pappenberger, G., Wilsher, J. A., Roe, S. M., Counsell, D. J., Willison, K.R., and Pearl, L. H. (2002). Crystal structure of the CCTγ apical domain: Implications for substrate binding to the eukaryotic cytosolic chaperonin. *J Mol Biol* **318**, 1367-1379.

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/γ subunit. *J Mol Biol* **360**, 484-496.

Pashkova, N., Gakhar, L., Winistorfer, S.C., Yu, L., Ramaswamy, S., Piper, R.C. (2010) WD40 Repeat Propellers Define a Ubiquitin-Binding Domain that Regulates Turnover of F Box Proteins. *Mol Cell* **40***,* 433–443.

Passmore, L., McCormack, E. A., Au, S. W. N., Paul, A., Willison, K. R., Harper, J. W., and Barford, D. (2003). Doc1 mediates the activity of the anaphase-promoting complex by contributing to substrate recognition. *EMBO J* **22**, 786-796.

Plimpton, R.L., Cuéllar, J., Lai, C.W.L., Aoba, T., Makaju, A., Franklin, S., Mathis, A.D., Prince, J.T., Carrascosa, J.L., Valpuesta, J.M., and Willardson, B.M (2015). Structures of the Gβ-CCT and

PhLP1–Gβ-CCT complexes reveal a mechanism for G-protein β-subunit folding and Gβγ dimer assembly. *Proc. Natl. Acad. Sci. USA* **112**, 2413–2418.

Rizzolo, K., Huen, J., Kumar., 16 co-authors and Houry, W.A. (2017) Features of the Chaperone Cellular Network Revealed through Systematic Interaction Mapping. *Cell Reports* **20**, 2735-2748.

Shimon, L., Hynes, G.M., McCormack, E.A., Willison, K.R., and 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, 469-477.

Smith, T.F., Gaitatzes, C., Saxena, K., and Neer, E.J. 1999. The WD repeat: a common architecture for diverse functions. *Trends Biochem Sci* **24,** 181-185.

Stirling, P.C., Srayko, M., Takhar, K.S., Pozniakovsky, A., Hyman, A.A., 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**, 336–2345.

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. *J Biol Chem* **286**, 178–184

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

Vaisman, N., Tsouladze, A., Robzyk, K., Ben-Yehuda, S., Kupiec, M., and Kassir, Y. (1995). The role of Saccharomyces cerevisiae Cdc40p in DNA replication and mitotic spindle formation and/or maintenance. *Mol Gen Genet* **247,** 123-136.

Valpuesta, J.M., Martin-Benito, J., Gomez-Puertas, P., Carrascosa, J.L., and Willison, K.R. (2002). Structure and function of a protein folding machine: the eukaryotic cytosolic chaperonin CCT. *FEBS Lett* **529,** 11-16.

Valpuesta, J. M., Carrascosa, J. L. and Willison, K. R. (2005) Structure and Function of the Cytosolic Chaperonin CCT, in Protein Folding Handbook (eds J. Buchner and T. Kiefhaber), Wiley-VCH Verlag GmbH, Weinheim, Germany. doi: 10.1002/9783527619498.ch54

Wang, Y., Hu, X.J., Wu, X.H., Ye, Z.Q, and Wu, Y.D. (2015) WDSPdb* a database for WD40-repeat proteins. *Nuc. Acids. Res* **43**, 339-344 (doiorg/10.1093/nan/gkv1023)

Warner, J.R. (1999) The economics of ribosome biosynthesis in yeast. *TIBS* **24**, 437-440.

Willardson, B.M. and Howlett, A.C. (2007). Function of transducin-like proteins in G protein signalling and chaperone-assisted protein folding. *Cellular Signaling*. **19**, 2417-2427.

Willison, K. R. (1999). Composition and Function of the Eukaryotic Cytosolic Chaperonin Containing TCP-1. In *Molecular Chaperones and Folding Catalysts. Regulation, Cellular Function and Mechanisms*., B. Bukau, ed.: Harwood Academic Publishers, pp. 555-571.

Willison, K.R. (2018). Biochem J – commissioned manuscript in preparation.

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

Yu, H. (2007) Cdc20: a WD40 activator for a cell cycle degradation machine. *Mol Cell* **27**, 3-

TABLE 1. List of the core 7-bladed WD40 repeat containing proteins in yeast

This set was constructed using sequential Psi-BLAST homology searching, Phyre analysis and hand alignment – full alignment file is available on request. Gene name, protein complex, length aa and haploid v diploid are from *Saccharomyces cerevisiae*. The UNIPROT entry list shows the genes from *Encephalitis cuniculi* parasite which a eukaryote with a minimal (~1990 genes) and compacted genome; all the WD40 proteins are detected by homology except TUP1.

Number	Gene name	Protein complex	UNIPROT	Length aa	HAPLOD v DIPLOID
1	COP1	Coatomer I	Q8SSJ2	1201	1
2	SEC27	Coatomer I	Q8SRA6	889	0.96
3	SEC13	Coatomer II	Q8SQT8	297	1.05
4	SEC31	Coatomer II	Q8SRF6	1273	0.97
5	ASC1	Ribosome	Q8SRB0	319	0.93
6	RSA4	Ribosome	Q8SW59	515	1.01
7	TAF5	TFIID	Q8SQS4	798	1.12
8	ELP2	Pol II elongation	Q8SSL8	788	1.02
9	HIR1	Histone repressor	Q8SSG4	875	1.13
10	HAT2	Acetyltransferase	Q8SRK1	401	1.03
11	SWD2	Set1p complex	Q8SVQ1	329	1.15
12	DOA1	Ubiquitin	Q8SQK6	715	0.98
13	CAC2	CAF1-p60	Q8SRA5	468	1.12
14	TUP1	Transcription	ND	713	0.91
15	CDC55	PP2A	Q8SRX4	526	1.08
16	CDC20	APC/C	Q8SS21	610	not detected
17	CDH1	APC/C	Q8SVZ6	566	not detected
18	PRP46	RNA splicing	Q8STZ5	451	1.04
19	PFS2	PolyA factor	Q8SW96	465	1.03
20	GLE2	Nuclear Pore (NPC)	Q8SRM6	365	0.98
21	DIP2	U3 snoRNP (SSU)	Q8SW87	943	0.98
22	PWP2	U3 snoRNP (SSU)	Q8SSO7	923	1.02
23	SOF1	U3 snoRNP (SSU)	Q8SSI4	489	1.02
24	UTP13	U3 snoRNP (SSU)	Q8SVM7	817	0.99
25	UTP15	U3 snoRNP (SSU)	Q8SU33	513	1.01
26	TIF34	Initiation Factor 3	Q8SR77	347	0.98
27	RRB1	Ribosome biogenesis	Q8SRW1	512	0.95
28	CIA1	Fe/S protein assembly	Q8STN6	330	1.03
29	YL149	Hypothetical	Q8STM2	730	1.12
30	VID27	Vacuole	Q8STM0	782	1.32

TABLE 2. Screening WD40-repeat proteins for CCT interaction

Fifteen yeast proteins predicted to contain 7-blade WD40 repeats, by 3DPSSM structure based sequence alignment (Kelley et al; 2000) and CDC4, an 8-bladed WD40 (Orlicky et al; 2003), non-CCT interacting, protein functional in *cct1-2* mutant cells at 37°C (Camasses et al; 2003) were analysed for interaction with CCT (column 3) and GroEL (column 4) by quantitative in vitro translation and native gel analysis. The extent of binding to CCT or GroEL was determined by calculating the fraction bound to chaperonin compared to total protein synthesized (weak = < 5%, medium = 5-15%, strong = >15%). Results from independent studies are shown in column 5 and referenced in column 6 ((1. Ho et al; 2002), (2. Camasses et al; 2003), (3. Passmore et al; 2003), (4. Gavin et al; 2002), (5. Aloy et al; 2004)). nd indicates not done, nt indicates not previously studied.

Protein	Protein Complex / Function	CCT binding	GroEL binding	CCT binding	Ref
		this study	this study	published work	
ASC1	Ribosome subunit/signalling	weak	weak	nt	
CAC2	CAF1-p60 / Chromatin	medium	yes	nt	
CDC4	Skip1 / Ubiquitin	weak	yes	no	2
CDC20	APC / Ubiquitin E3	strong	yes	yes + *in vivo*	1,2
CDC55	PP2A regB / phosphatase	strong	nd	yes + *in vivo*	1,3
CDH1	APC / Ubiquitin E3	strong	yes	yes + *in vivo*	1,2,3
DOA1	Ubiquitin	weak	no	nt	
ELP2	Pol II elongation / Transcription	negative	no	nt	
HAT2	HistoneAT TypeB / Acetylation	weak	no	nt	
SEC13	Coatomer II / Secretion	weak	no	nt	
SEC27	Coatomer I / Secretion	medium	nd	yes	1
TAF5	TFIID / Transcription	strong	yes	yes	1,3
TIF34	Initiation Factor 3 / Translation	negative	no	nt	
TUP1	Tup1-Ssn6 / Transcription repression	medium	nd	nt	
UTP13	U3 snoRNA (SSU) / Translation	medium	yes	nt	
VID27	Vacuole/protein degradation	Strong	nd	yes	4,5

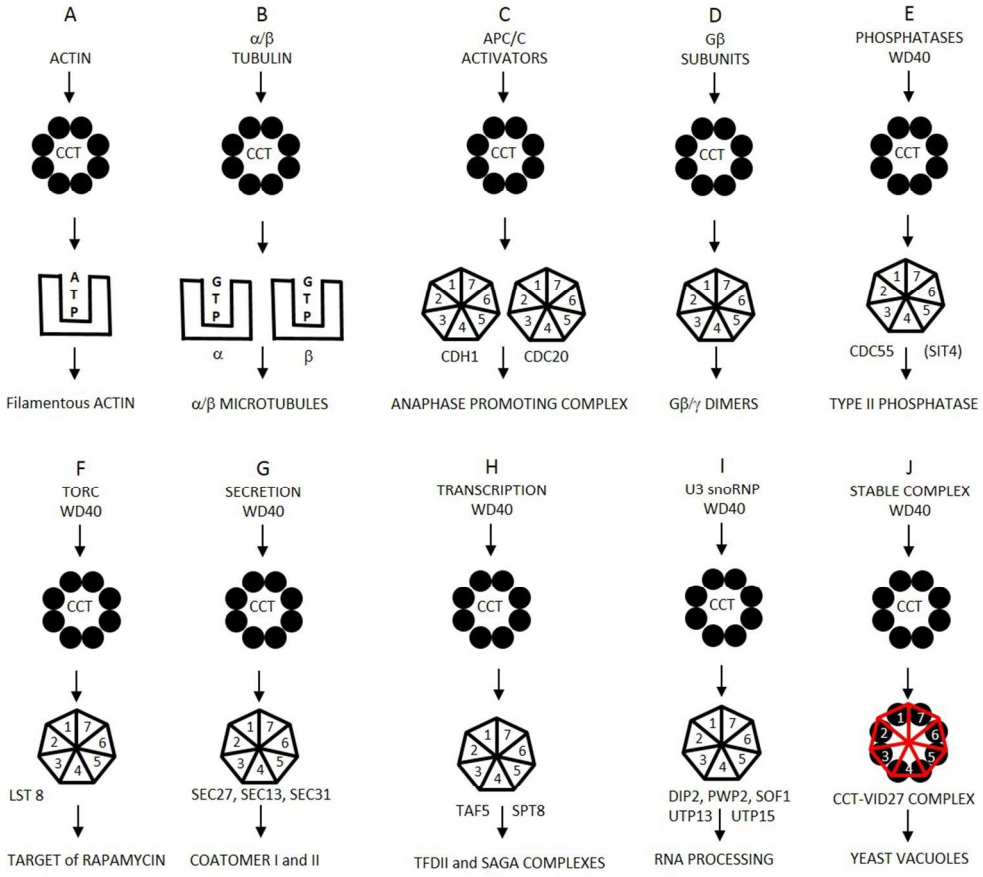


FIGURE 1. Cartoons of schemes for CCT action upon its substrates and binding complex formation

190x169mm (150 x 150 DPI)

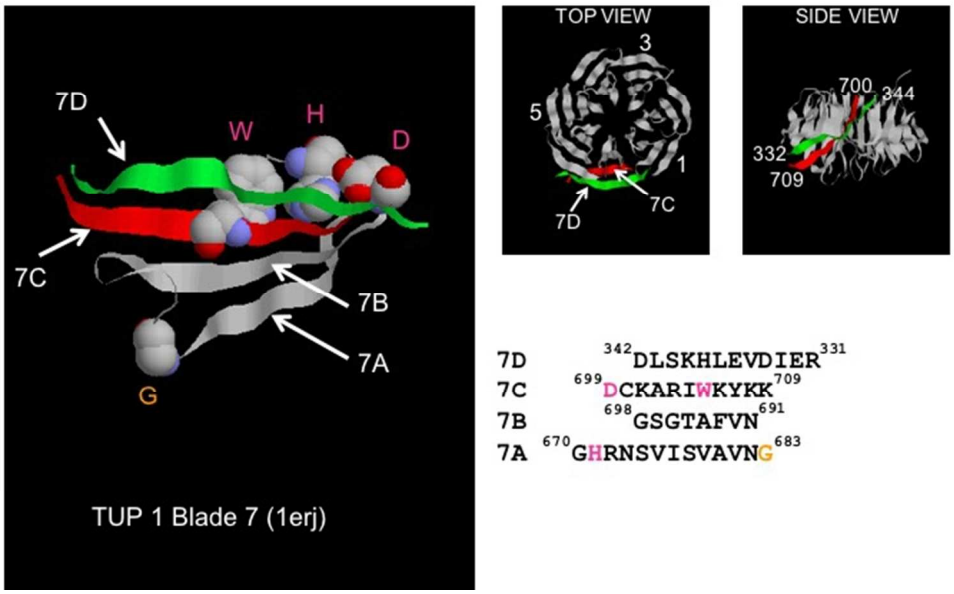


FIGURE 2. The Velcro cap

246x146mm (72 x 72 DPI)

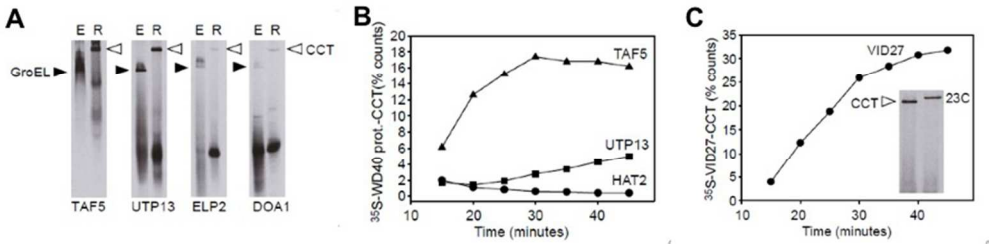
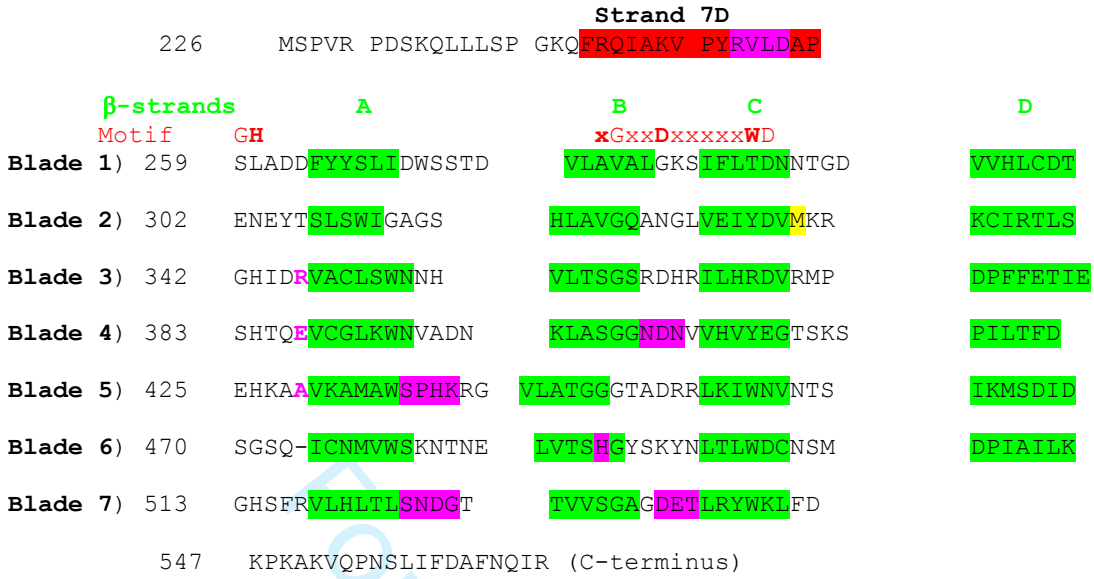


FIGURE 3. Screening WD-40-repeat proteins for CCT interaction

315x80mm (72 x 72 DPI)

A



B

Deletion analysis	Residue Number	Location	CCT interaction
CDH1 full-length	N405A,D406A,N407A	Blade 4 conserved	strong arrest/slow processing
CDH1 full-length	S436A and P437A	Blade 5 conserved	strong arrest/slow processing
CDH1 full-length	H438A and K439A	Blade 5 conserved	slight arrest/slow processing
CDH1 full-length	S524A and N525A	Blade 7 conserved	slight arrest/slow processing
CDH1 full-length	D526A and G527A	Blade 7 conserved	arrested
CDH1 full-length	D536A,E537A,T538A	Blade 7 conserved	slight arrest/slow processing
CDH1 full-length	R253A and V254A	Strand 7D conserved	arrested
CDH1 full-length	L255Aand D256A	Strand 7D conserved	slightly arrested
CDH1 M226-R566	R346A	Blade 3 conserved	processed
CDH1 M226-R566	E387A	Blade 4 conserved	strong arrest/slow processing
CDH1 M226-R566	A429S	Blade 5 conserved	slight arrest/slow processing
CDH1 M226-R566	H490R	Blade 6 tetrad	arrested
CDH1 M226-R566	D526A and G527A	Blade 7 conserved	arrested
CDH1 M226-R566	D536A, E537A, T538A	Blade 7 conserved	processed

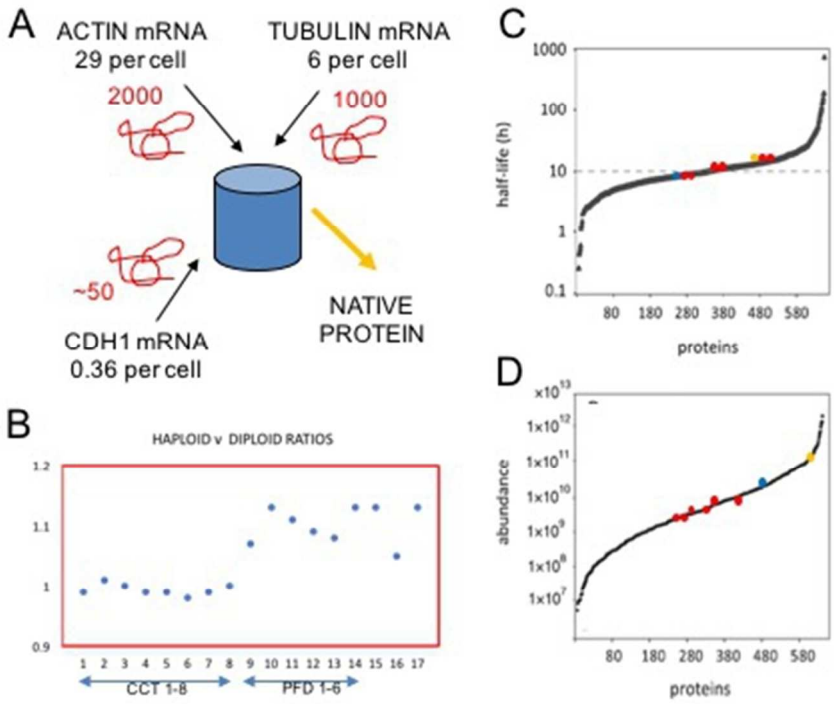


FIGURE 5. Yeast CCT Capacity

152x127mm (72 x 72 DPI)