nature structural & molecular biology

Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies

Alice Y Yam[1,4], Yu Xia[2,4], Hen-Tzu Jill Lin[3], Alma Burlingame[3], Mark Gerstein[2] & Judith Frydman[1]

Folding within the crowded cellular milieu often requires assistance from molecular chaperones that prevent inappropriate interactions leading to aggregation and toxicity. The contribution of individual chaperones to folding the proteome remains elusive. Here we demonstrate that the eukaryotic chaperonin TRiC/CCT (TCP1-ring complex or chaperonin containing TCP1) has broad binding specificity *in vitro*, similar to the prokaryotic chaperonin GroEL. However, *in vivo*, TRiC substrate selection is not based solely on intrinsic determinants; instead, specificity is dictated by factors present during protein biogenesis. The identification of cellular substrates revealed that TRiC interacts with folding intermediates of a subset of structurally and functionally diverse polypeptides. Bioinformatics analysis revealed an enrichment in multidomain proteins and regions of β-strand propensity that are predicted to be slow folding and aggregation prone. Thus, TRiC may have evolved to protect complex protein topologies within its central cavity during biosynthesis and folding.

Eukaryotic cells contain several distinct chaperone families that together promote protein folding[1,2]. Misregulation of this process leads to misfolding and aggregation events that are linked to multiple pathological disorders[3,4]. It is thought that proteins differ widely in their chaperone requirements[2]. It is unclear, however, what features of a folding polypeptide, if any, determine its interaction with specific chaperones. Determining whether different chaperones evolved to meet the folding requirements of specific classes of substrates is central to understanding the logic of cellular protein folding and assembly. Addressing this possibility requires a better understanding of which types of proteins require a given chaperone. However, the cellular substrates of most eukaryotic chaperones have not yet been defined.

The essential chaperonin TRiC/CCT is distinguished from other chaperones by its unique ring-shaped architecture, which gives rise to a central cavity that serves as a folding chamber for substrate polypeptides[5,6]. It is not known why some proteins require the ring-shaped TRiC to fold whereas others can reach their native states with the assistance of simpler chaperone systems. Indeed, the cellular function of TRiC remains ill-defined and controversial. TRiC was originally proposed to be highly specialized to recognize a few cytoskeletal proteins through specific sequence elements[7]. However, the recent identification of additional TRiC substrates has called into question this original idea[8–12]. Here we have determined the principles of substrate selection by TRiC and defined the subset of cellular proteins that interact with this chaperonin in eukaryotic cells using a combination of experimental and computational analyses.

RESULTS

Principles of TRiC substrate selection

TRiC is part of a chaperone network linked to protein synthesis[13] and has been shown to facilitate folding of newly translated proteins *in vivo*[9,14]. Previous studies established that TRiC interacts transiently with a subset of cellular proteins during biogenesis[14]. We thus examined the flux of newly translated proteins through TRiC in mammalian cells using a previously established pulse-chase analysis technique, whereby newly made proteins are specifically labeled with ^{35}S-methionine during a brief pulse, and folding and maturation occurs during the chase period[14]. We isolated newly synthesized polypeptides interacting with TRiC using immunoprecipitation with antibodies against TRiC subunits β and ε (**Fig. 1a** and **Supplementary Fig. 1** online). Two-dimensional PAGE analysis showed that, soon after translation, many newly made proteins associated with TRiC. Following a period of chase, these proteins were dissociated, as expected for chaperone substrates, which should be released upon completion of folding (**Fig. 1a**, left and middle). We also observed chaperonin complex assembly during the time course of the chase, whereby the β and ε subunits associated with the remaining TRiC subunits[14] (**Fig. 1a**, middle). MS analysis of TRiC-interacting proteins

[1]Department of Biology and BioX Program, E200A James Clark Center, 318 Campus Drive, Stanford University, Stanford, California 94043, USA. [2]Department of Molecular Biophysics & Biochemistry, PO Box 208114, Yale University, New Haven, Connecticut 06520, USA. [3]Department of Pharmaceutical Chemistry, 600 16th Street, MC 2240, Genentech Hall, Suite N472A, University of California at San Francisco, San Francisco, California 94143, USA. [4]Present addresses: Novartis Vaccines and Diagnostics, Inc., 4560 Horton Street, M/S 4.3, Emeryville, California 94608, USA (A.Y.Y.); Department of Chemistry, 44 Cummington Street, Boston University, Boston, Massachusetts 02215, USA (Y.X.). Correspondence should be addressed to J.F. (jfrydman@stanford.edu).

Received 12 May; accepted 16 October; published online 16 November 2008; doi:10.1038/nsmb.1515

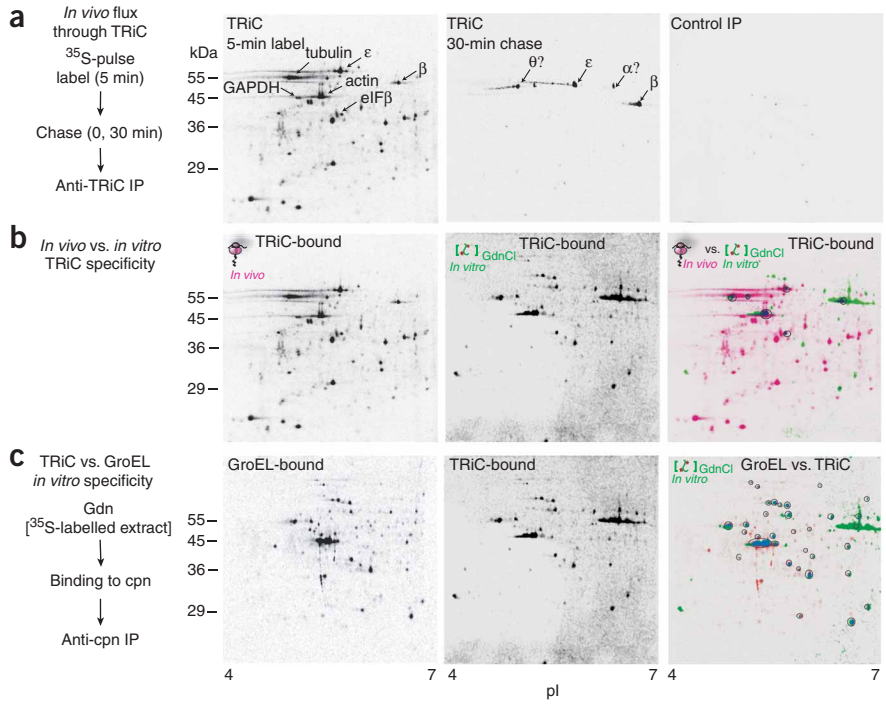

Figure 1 Principles of TRiC substrate selection in the eukaryotic cytosol. (**a**,**b**) Human fibroblast cells (TSA-201) were pulse labeled with ^{35}S-methionine for 5 min, followed by a 0 or 30 min chase. Total soluble protein was immunoprecipitated with anti-TRiC antibodies or a nonimmune antibody control, and the precipitates were separated on two-dimensional gels. (**b**) TRiC recognizes different proteins *in vitro* and *in vivo*. Two-dimensional gels of *in vitro* (green) and *in vivo* (magenta) TRiC-bound proteins were compared. The two gels were merged with overlapping spots circled and shown in blue (right). (**c**) A denatured ^{35}S-labeled cytosolic extract was diluted into buffer containing either GroEL or TRiC. Bound proteins were assessed by immunoprecipitation with anti-chaperonin (cpn) antibodies (EL or TRiC), and separated by two-dimensional gel electrophoresis. Gel images of GroEL-bound (red) and TRiC-bound (green) proteins were merged with overlapping spots circled and shown in blue (right).

identified only highly abundant substrates, namely the WD repeat–containing translation initiation factor-3β and glyceraldehyde-3-phosphate dehydrogenase (GAPDH), in addition to the known TRiC substrates actin and tubulin (**Supplementary Table 1** online). Thus, identification of low-abundance cellular substrates of TRiC required alternative genome-wide approaches.

To better define the principles that govern TRiC substrate specificity, we next examined what determines association of cellular proteins with TRiC. In principle, chaperonin-substrate interactions may be solely determined by the presence of specific TRiC binding motifs in the substrates, such as sequence elements, that distinguish them from the rest of the proteome. A prediction of this model is that the *in vitro* substrate specificity of TRiC toward denatured cytosolic proteins will mirror that observed *in vivo*. Accordingly, we compared the subset of cellular proteins that bind TRiC upon translation *in vivo* (**Fig. 1b**, left) with those proteins binding TRiC when the same ^{35}S-labeled extract is denatured and presented to the chaperonin *in vitro* (**Fig. 1b**, middle). The sets of eukaryotic proteins interacting with TRiC *in vivo* and *in vitro* were markedly different. Whereas actin and tubulin were prominently bound in either condition, examination of the merged gels revealed less than 10% overlap between the proteins selected by TRiC *in vitro* and *in vivo* (**Fig. 1b**, right). We conclude that TRiC does not select its substrates based solely on the presence of specific sequence motifs, as was proposed from *in vitro* studies[15,16]. Instead, TRiC substrate selection in the cell is strongly dependent on the context of translation, where both co-translational folding events and cooperating chaperone systems may affect the conformation of *de novo* folding intermediates[11,17,18].

TRiC was proposed to be a highly specific chaperone, in contrast to the prokaryotic chaperonin GroEL, which shows broad specificity[7]. To further define the substrate-recognition principles of TRiC, we next determined whether TRiC recognizes a more restricted range of proteins when compared to GroEL[19]. To compare the cellular proteins recognized directly by TRiC and GroEL, we presented denatured ^{35}S-labeled cytosolic proteins to purified GroEL or TRiC and then immunoprecipitated the chaperonin-substrate complexes (**Fig. 1c** and **Supplementary Fig. 2** online). A large fraction of cellular proteins, including actin and tubulin, was recognized by both chaperonins. Comparison of the GroEL and TRiC-bound protein spectra revealed that the recognition specificities of these chaperonins were markedly similar, with more than 80% overlap between the protein sets recognized by either chaperonin (**Fig. 1c**, right). Given the well-established affinity of GroEL for hydrophobic substrate determinants, this result indicates that hydrophobicity is a strong component of TRiC substrate recognition, consistent with previous findings[8,9,20,21]. Notably, the similar binding specificity of both chaperonins suggests that TRiC retains the capacity to recognize a wide breadth of proteins and is not intrinsically a highly specific chaperonin. We conclude that TRiC possesses broad recognition specificities, yet in the cell it interacts with only a defined set of substrates.

Genomic screen for TRiC substrates during biosynthesis

Our finding that TRiC substrate selection in the cell is determined in the context of protein biosynthesis raises the question of what features distinguish proteins recognized by TRiC from the wide spectrum of potential interactors during biogenesis. Accordingly, we adapted a genome-wide approach that allowed us to identify physiologically relevant TRiC substrates by detecting chaperone interactions in the context of translation.

We used a screening approach that monitored which proteins interact with TRiC during translation of cDNA expression pools in a cell-free mammalian translation system[22]. This approach, termed small pool expression cloning (SPEC), allowed us to detect TRiC-substrate interactions in a physiologically relevant context (**Fig. 2a**). Notably, SPEC presents several unique advantages for the identification of physiological chaperone substrates. First, these translation lysates contain the full complement of chaperones and translation components required to fold most cytosolic proteins to an active state[23]. TRiC-substrate interactions are thus examined in the context of protein biosynthesis, whereby co-translational folding can occur in the presence of physiological levels of upstream chaperones and folding cofactors. Second, proteins are translated at low levels, in the picomolar

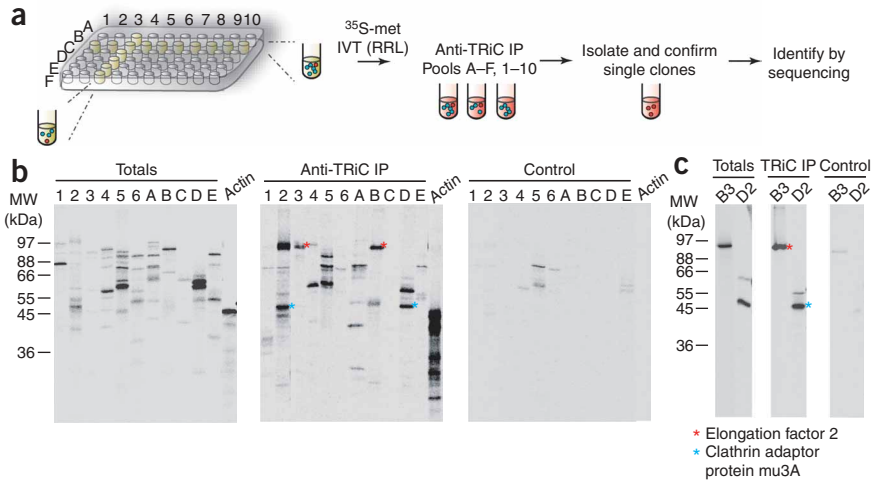

Figure 2 Screening for the TRiC interactome in the context of translation. (**a**) Small pool expression cloning screen. Bacterial cDNA clones were cultured in 96-well plates and pooled in rows and columns, from which plasmid DNA was purified. The orthogonal cDNA pools were expressed in rabbit reticulocyte lysate (RRL) in the presence of ^{35}S-methionine and immunoprecipitated with anti-TRiC antibodies. (**b**,**c**) Total expressed cDNA pools in RRL (Totals) and TRiC-bound proteins (anti-TRiC IP) were analyzed by SDS-PAGE and phosphorimaging (**b**). Immunoprecipitations with nonimmune antibodies demonstrated the specificity of the TRiC immunoprecipitations (Control). Actin expression and TRiC binding served as a positive control throughout. Red and blue asterisks indicate TRiC-interacting candidate clones present in orthogonal pools. Interactions were confirmed by expression of single candidate clones B3 and D2 in RRL, followed by anti-TRiC immunoprecipitation (**c**). cDNAs were identified by sequencing from upstream promoter sequences.

range, and thus interactions with the chaperonin can be observed without overexpression of the substrates[24]. Furthermore, the lack of bias toward detection of abundant proteins in the SPEC approach permits the identification of chaperone interactions for rare, low-abundance proteins, unlike large-scale immunoprecipitation reactions.

We carried out a high-throughput screen using single mouse cDNAs arrayed in a 96-well plate format. Orthogonal cDNA pools were made along rows and columns for each plate and analyzed by *in vitro* translation in rabbit reticulocyte lysate in the presence of ^{35}S-methionine (**Fig. 2a**). The presence of TRiC substrates encoded by cDNAs in the pools was detected by co-immunoprecipitation with the chaperonin followed by SDS-PAGE and phosphorimaging analysis (**Fig. 2b**). We isolated the cDNAs encoding candidate substrates expressed in two overlapping pools and confirmed their interaction with TRiC before using DNA sequencing to identify them (**Fig. 2c**). To simplify the screen, we subtracted the abundant housekeeping mRNAs from the cDNA library, thus increasing the proportion of lower-abundance transcripts surveyed[25]. Of note, the identification of α-tubulin among the TRiC-interacting substrates demonstrated that housekeeping proteins were still represented in the library, thus providing a good representation of transcripts in the cell (**Supplementary Table 1**).

In all, we surveyed 2,600 clones and isolated more than 100 TRiC-interacting proteins. The screen was not conducted to saturation but was concluded upon the repeated identification of several clones. From these data, we anticipate that 6–7% of all cytosolic proteins interact with TRiC. Notably, this analysis is in agreement with previous estimates from pulse-chase analysis[14] and proteomics studies[26]. Because only high-affinity interactions are detected by immunoprecipitation, the predicted number of TRiC-interacting proteins might be underestimated. Assuming that the eukaryotic cytosol contains approximately 4,500 proteins, based on estimates from the yeast *Saccharomyces cerevisiae*, our analysis predicts that on the order of 300 cytosolic proteins interact with TRiC *in vivo*.

Distinct *in vivo* kinetics of substrate flux through TRiC

Our screen revealed that a wide range of proteins interact with TRiC upon translation, most of which are conserved across eukaryotes (**Supplementary Table 1**). Accordingly, we used *S. cerevisiae* to validate the *in vivo* TRiC interactions, taking advantage of the extensive annotation of the yeast genome. Initially, we isolated candidates carrying an N-terminal GST-tag[27] by affinity purification and assessed their association with TRiC using immunoblot analysis. Notably, all of the yeast homologs of the SPEC-derived candidates surveyed bound to TRiC *in vivo*, but neither GST alone nor GST-tagged to Ypt6, a control GTP binding protein, associated with the chaperonin (**Fig. 3a**). This supports the idea that the screen identified bona fide TRiC substrates and that the interaction with TRiC is evolutionarily conserved across eukaryotic organisms.

One hallmark of chaperones that facilitate *de novo* folding is that they associate transiently with newly made proteins and dissociate during the course of polypeptide folding and maturation. As such, we examined the *in vivo* flux of SPEC-derived TRiC substrates through the chaperonin in yeast cells. Pulse-chase analysis followed by anti-TRiC immunoprecipitation indicated that, as in mammalian cells, a range of yeast proteins transiently associate with TRiC early in their biogenesis and dissociate during the chase, consistent with release upon folding (**Fig. 3b**, 15 min).

To observe the flux of single substrates through the chaperonin, we subjected TRiC immunoprecipitations to a second round of immunoprecipitation against a unique tag in the substrates (**Fig. 3c,d**). We observed transient association kinetics for the SPEC-identified substrates, indicating that these proteins transit through the chaperonin during their biogenesis (**Fig. 3d**). Notably, we observed two types of association kinetics among TRiC substrates. One class of proteins, including actin and the small purine nucleoside phosphorylase (PNP1), bound early to TRiC and rapidly dissociated (**Fig. 3d,e**). Elongation factor 2 (EFT2) also bound rapidly to TRiC, but dissociated slowly from the chaperonin (**Fig. 3d,e**). Another set of substrates showed much slower kinetics of association with TRiC. For instance, the mRNA export factor MEX67, the homotetrameric citrate synthase CIT2, the multidomain protein arginyl tRNA synthetase RRS1 (also known as ArgRS) and the WD40 G_β protein STE4 did not bind to TRiC at the earliest chase times, but instead their binding seemed to peak later in the time course (**Fig. 3d,e**). Notably, the TRiC-association kinetics were independent of substrate size because the larger EFT2 associated rapidly with TRiC whereas the smaller STE4 did not. It is tempting to speculate that the slower TRiC binding kinetics arises from the action of cooperating upstream chaperones delaying binding to TRiC during translation[13,18]. Notably, all the slower binding proteins share the common property of belonging to larger protein complexes, raising the possibility that these substrates accumulate on TRiC before their release into oligomeric assemblies. In

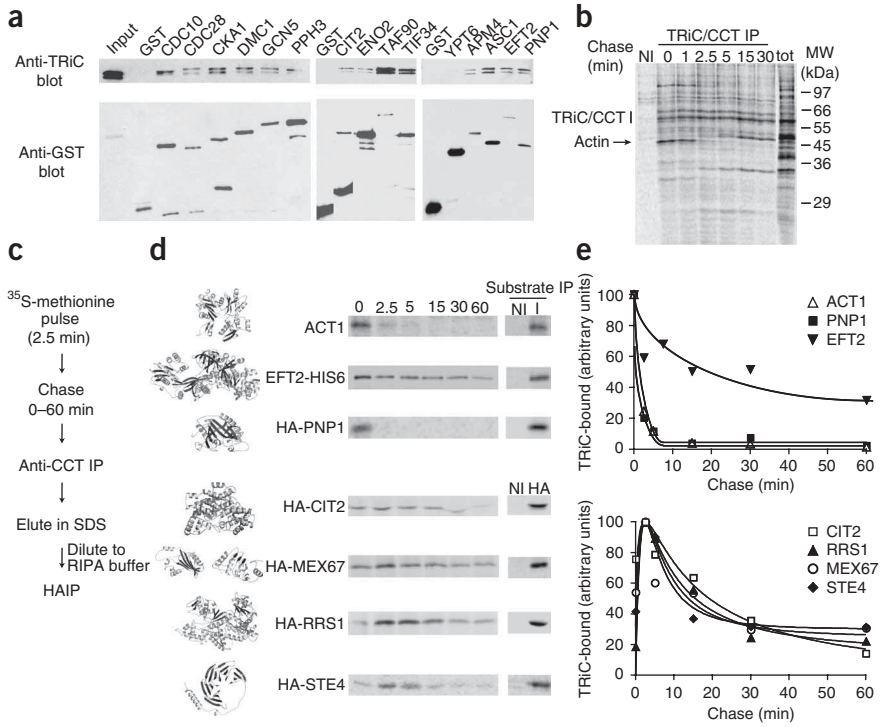

Figure 3 *In vivo* analysis of the TRiC interactome reveals distinct kinetics of substrate flux through the chaperonin. (**a**) GST fusion proteins were expressed in *S. cerevisiae* behind a galactose-inducible promoter and affinity purified with glutathione Sepharose. Purified proteins were separated by SDS-PAGE and subjected to western blot analysis with anti-TRiC or anti-GST antibodies. (**b**,**c**) The flux of newly synthesized proteins binding to TRiC was assessed by pulse-chase analysis in wild-type yeast. ^{35}S-labeled cells were harvested at the indicated time points, lysed and immunoprecipitated (IP) with anti-TRiC antibodies. NI, nonimmune; tot, total lysate. (**c**–**e**) The flux of individual proteins through TRiC was assessed by pulse-chase analysis of wild-type yeast expressing tagged substrates. TRiC-bound proteins from the pulse-chase analysis were eluted with SDS and subjected to a second immunoprecipitation with either nonimmune antibodies (NI) or antibodies recognizing tagged substrates (I). Structural models for each of the substrates were generated by multiple sequence alignment with homologous proteins using SWISS-MODEL[50]. TRiC-bound substrates isolated in the pulse-chase were quantitated with ImageQuant software (version 5.2, Amersham Biosciences) and the kinetics of their association plotted over time (**e**).

support of this idea, several WD40-containing F-box proteins translated in a cell-free system accumulated on TRiC in the absence but not in the presence of their oligomeric partner protein Skp1 (**Supplementary Fig. 3** online). We conclude that a genome-wide screen for the TRiC interactome identified bona fide TRiC substrates that flux through the chaperonin during biogenesis *in vivo*.

Physical and structural properties of the TRiC interactome

We next sought to define what distinguishes the relatively small fraction of TRiC-interacting proteins from the cytosolic proteome at large. Having identified a large set of proteins that interact with TRiC in a physiologically relevant context (herein referred to as the TRiC interactome; **Supplementary Table 1**), we applied a bioinformatics approach to identify common features among TRiC substrates that may underlie their physiological association with TRiC. Proteins from the interactome spanned a breadth of cellular processes, most of which are conserved across eukaryotic organisms. Notably, 40% of the TRiC interactome comprises essential genes, more than twice the proportion of essential genes in the yeast genome[28]. Of note, the frequency of TRiC substrates belonging to oligomeric protein complexes was also enriched in the TRiC interactome to 90%, compared to less than 50% of all yeast proteins in the MIPS complex database[29]. These two observations may be linked, as proteins that belong to oligomeric complexes are more highly connected in the protein network and, thus, are disproportionately essential[30]. We hypothesize that the eukaryotic chaperonin assists in folding proteins belonging to oligomeric assemblies and may serve as a reservoir to stabilize them against aggregation or degradation before complex formation.

Previous biochemical experiments using actin and tubulin led to the proposal that TRiC selects its substrates by specific recognition of a set of polar sequence elements[15,16]. Accordingly, we analyzed the TRiC interactome to search for the presence of specific sequence motifs that may confer TRiC binding. Sequence analysis of the TRiC interactome failed to reveal statistically significant consensus sequences, nor did we find an enrichment in previously proposed TRiC binding sequence elements[15,16]. This is consistent with the finding that TRiC is not a highly sequence-specific chaperonin (**Fig. 1c**).

In principle, physical and/or structural properties could influence protein folding characteristics and contribute to chaperonin binding. A comparison of the size distribution of TRiC-interacting proteins to the naturally occurring distribution of cytosolic proteins in the *S. cerevisiae* proteome (**Fig. 4a**) or the murine proteome (**Supplementary Fig. 4** online) demonstrated a statistically significant enrichment for proteins ranging in size between 40 kDa and 75 kDa, in good agreement with results obtained using pulse-chase approaches[14] (**Figs. 1** and **4a**). Of note, the chaperonin chamber is predicted to accommodate polypeptides in this size range[31]. As a single folding domain is approximately 25–30 kDa[1], this finding suggests that the TRiC interactome consists mostly of multidomain proteins. Notably, our screen also identified many larger substrates, raising the possibility that some substrates are not completely encapsulated during folding. In support of this idea, the 100-kDa protein myosin is known to be an obligate substrate of TRiC[32]. Furthermore, the bacterial chaperonin GroEL similarly supports the folding of several large proteins[33].

The proteins in the TRiC interactome showed an enrichment in hydrophobic sequences (**Fig. 4b**). This finding could be explained by substrate size, as larger proteins retain correspondingly larger hydrophobic cores. Alternatively, hydrophobic interdomain contacts within substrate proteins could contribute to the higher proportion of hydrophobic sequences. Inappropriate interactions among these hydrophobic surfaces, for instance, in domain swapping, could complicate the folding of multidomain proteins[34]. Previous observations that TRiC can recognize hydrophobic determinants[8,9,20,21] (**Fig. 1c**) suggest that the chaperonin may prevent these inappropriate interactions.

Analysis of the structural propensities of the TRiC interactome revealed the most striking property shared among TRiC substrates.