Interactions of subunit CCT3 in the yeast chaperonin CCT/TRiC with Q/N-rich proteins revealed by high-throughput microscopy analysis

Michal Nadler-Hollya,1, Michal Brekerb,1, Ranit Gruber, Ariel Aziac,d, Melissa Gymrek, Miriam Eisenstein,e, Keith R. Willisonl, Maya Schuldinerb,2, and Amnon Horovitza,2

Departments of aStructural Biology, bMolecular Genetics, and cChemical Research Support, Weizmann Institute of Science, Rehovot 76100, Israel; dThe Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan 52900, Israel; eWhitehead Institute for Biomedical Research, Cambridge, MA 02142; and Department of Chemistry, Institute of Chemical Biology, Imperial College London, London SW7 2AZ, United Kingdom

Edited* by George H. Lorimer, University of Maryland, College Park, MD, and approved October 2, 2012 (received for review June 12, 2012)

The eukaryotic chaperonin containing t-complex polypeptide 1 (CCT/TRiC) is an ATP-fueled machine that assists protein folding. It consists of two back-to-back stacked rings formed by eight different subunits that are arranged in a fixed permutation. The different subunits of CCT are believed to possess unique substrate binding specificities that are still mostly unknown. Here, we used high-throughput microscopy analysis of yeast cells to determine changes in protein levels and localization as a result of a Glu to Asp mutation in the ATP binding site of subunits 3 (CCT3) or 6 (CCT6). The mutation in subunit CCT3 was found to induce cytoplasmic foci termed P-bodies where mRNAs, which are not translated, accumulate and can be degraded. Analysis of the changes in protein levels and structural modeling indicate that P-body formation in cells with the mutation in CCT3 is linked to the specific interaction of this subunit with Gln/Asn-rich segments that are enriched in many P-body proteins. An in vitro gel-shift analysis was used to show that the mutation in subunit CCT3 interferes with the ability of CCT to bind a Gln/Asn-rich protein aggregate. More generally, the strategy used in this work can be used to unravel the substrate specificities of other chaperone systems.

molecular chaperones | polyQ proteins | protein mis-folding | protein aggregation | high-content analysis

Chaperonins are ATP-dependent protein-folding machines that are present in all kingdoms of life. They consist of two back-to-back stacked oligomeric rings with a cavity at each end, where protein substrate binding and folding take place (for reviews see refs. 1 and 2). The chaperonins can be divided into two groups: group I, found in the bacterial cytoplasm (e.g., GroEL in Escherichia coli), mitochondria, and chloroplasts; and group II found in archaea and the eukaryotic cytosol. Numerous studies have shown that the group I chaperonin GroEL can facilitate the folding of a large number of different proteins in vitro, although its role in the cell is much more limited (for review see ref. 3). The group II eukaryotic chaperonin containing t-complex polypeptide 1 (CCT/TRiC) also seems to have a specialized role in vivo in the folding of actin (4), tubulin (5), and other essential proteins, including regulators of cell division and cytoskeleton formation (6–8), despite seeming to possess broad binding specificity (6). The list of members in the CCT interactome is, however, not yet fully established, and the conditions for entry into this exclusive club remain poorly understood.

An important distinction between group I and group II chaperonins is that the former consist of homo-oligomeric rings, whereas the latter usually consist of hetero-oligomeric rings that contain two or three different subunits in the case of many archaeal chaperonins and eight different subunits in the case of CCT. The eight subunits of CCT are arranged in a defined permutation (9), and the orientation of the two rings of CCT with respect to each other is also fixed (10). CCT’s hetero-oligomeric structure has enabled a combinatorial mode of protein substrate binding to specific subunits (8, 11) to coevolve as found in the case of actin, for example, which binds to particular subunits on both sides of the ring (11). Hence, additional challenges, in the case of CCT, are (i) to characterize the (potentially overlapping) interactions of different combinations of subunits and (ii) to establish which properties of a substrate determine its dependence on a specific subset of CCT subunits for efficient folding. The hetero-oligomeric structure of CCT is also most likely responsible for its sequential intraring allostery (12) that differs from the concerted mechanism of GroEL (13). It has been suggested that sequential allostery facilitates domain-by-domain substrate release in a temporal order that is determined by the specific subunits to which the substrate is bound (12).

To obtain further insight into the functional significance of CCT’s hetero-oligomeric structure, we recently generated (14) a set of Saccharomyces cerevisiae yeast strains in which an aspartic acid, which is conserved in the ATP binding site of all eight subunits (and also across species), is mutated to glutamic acid in each subunit of CCT in turn. In vivo analyses of these CCT mutant yeast strains revealed surprisingly large phenotypic differences between them in, for example, growth rates, heat and cold sensitivities, and morphologies. The phenotypic differences between the strains are likely to be due, in part, to misfolding of substrates that require the mutated subunit for efficient folding (Fig. L4). These yeast strains, therefore, provide an excellent model system for gaining insight into the substrate specificities of the different subunits. Here we describe progress in this direction that was attained by high-throughput microscopy screening (15, 16) of a library generated by crossing strains with a mutation in CCT3 or CCT6 with a library of strains expressing different yeast ORFs fused to GFP (17) (Fig. 1B). Our results show that each of the mutations in CCT produced a unique effect on the yeast proteome and, more specifically, that subunit CCT3 has a stronger interaction in vivo with Q/N-rich proteins than subunit CCT6. In particular, subunit CCT3 is found to interact with Q/N-rich proteins that form cytoplasmic foci termed P-bodies, where mRNAs that are not being translated accumulate and can be degraded (18). More broadly, our results demonstrate that unraveling...
The unique effects of the mutations were next characterized by examining the changes they caused in protein abundance. We found that each of the mutations affected the levels of several tens of proteins. For example, 46 and 18 proteins were found to be respectively up- and down-regulated significantly (P < 0.01) in the MA3 vs. wild-type strain (Dataset S1). A plot of the fold changes in protein levels in MA6 against the fold changes of the corresponding proteins in MA3 (Fig. 3A) revealed that three of the nine proteins that change most in MA3 but significantly less in MA6 (Pub1, New1, and Dbp2) are Q/N-rich and involved in P-body formation. This finding prompted us to examine whether the fold changes in levels of Q/N-rich proteins are, in general, larger in their absolute values in MA3 relative to MA6. We therefore calculated the SDs of the fold change values in MA3 and MA6 for the list of yeast Q/N-rich proteins [compiled previously (25)] and compared them with the SDs of many groups of the same size that contain randomly selected proteins that are not included in the list of Q/N-rich proteins. It may be seen in Fig. 3B that the SD of the fold change values of the Q/N-rich proteins in MA3 is significantly larger than the SD of the groups of randomly selected proteins in MA3 that are not Q/N-rich. By contrast, the SD of the fold change values of the Q/N-rich proteins in MA6 is only marginally larger than the SD of the groups of randomly selected proteins (Fig. 3B). These results indicate that the levels of Q/N-rich proteins are affected (decreased or increased) more by the mutation in subunit CCT3 than by the corresponding mutation in subunit CCT6.

Given that changes in protein levels can be due to a direct interaction with the mutated subunit or to an indirect cellular cause, we next asked whether the proteins (including those that are not Q/N-rich) whose levels are found to change significantly in MA3 but not MA6 (Fig. 3A) have more interactions with Q/N-rich proteins than randomly chosen proteins. We found that eight of these nine proteins have at least one physical or genetic interaction with Q/N-rich proteins, whereas in randomly chosen groups of nine proteins a significantly smaller number of only four to five proteins, on average, have physical or genetic interactions with Q/N-rich proteins (Fig. 3C). Taken together, these results and those in Fig. 3B led us to hypothesize that subunit CCT3 has a specialized role in folding Q/N-rich proteins.

**Structural Modeling Reveals a Potential Binding Site in Subunit CCT3 for Q/N-Rich Regions.** Given the above-described evidence obtained from the high-throughput microscopy screening approach for the interaction of subunit CCT3 with Q/N-rich proteins, we next asked whether there is any structural evidence for such an interaction. Structural data for the entire CCT complex is currently available only at 3.8-Å resolution (10), but fortunately the crystal structure of a yeast CCT complex has been solved at higher resolution (20). We found that the structure of subunit CCT3 contains regions that interact with Q/N-rich regions.

---

**Fig. 1.** Scheme showing the approach put forward in this work for screening effects of mutations in CCT on substrate folding and localization. (A) The mutation in the ATP binding sites of subunits CCT3 or CCT6 (in red) can interfere with substrate binding and cause misfolding. Substrates can also misfold after binding to the mutant CCT or be released only partially. Misfolding can lead to aggregation or degradation that may be reflected in changes in protein levels and/or localization. (B) Strains with the mutation in CCT were crossed with a library of strains expressing endogenous yeast proteins fused to GFP. The resulting strains, each expressing wild-type (control) or mutant CCT and an endogenous yeast protein fused to GFP, were subjected to high-throughput microscopy screening to determine changes in protein levels and localization.
structure of the apical domain of subunit CCT3 from mouse has been solved at 2.2-Å resolution (26). This structure, as well as that of the apical domain of GroEL (27) as a control, were analyzed using an algorithm that detects energetically favorable anchoring spots for amino acid side chains on protein surfaces (28). The analysis shows that the apical domain of GroEL contains relatively few Q/N anchoring spots and that most of them are not located in the substrate binding area (Fig. 4). By contrast, the apical domain of subunit CCT3 is found to contain a large number of Q/N anchoring spots, many of which are located in the probable substrate binding area. Particularly remarkable is the observation that the spacing of six of the anchoring spots in the substrate binding area of subunit CCT3 is ideal for binding a Q/N-rich segment in a regular β-strand conformation (Fig. 4).

Given the model of a binding site in CCT3 for a Q/N-rich segment (Fig. 4), we sought to uncover the entire set of yeast proteins that have the potential to bind directly to the anchoring spots in our predicted CCT3 binding surface. To this end, a comprehensive scan of the 6,583 yeast proteins found in the UniProt database was performed for segments of seven residues that contain five or more Q/N. Overall, 737 proteins were found to contain such segments. Of the 27 proteins in the MA3 strain that aggregate or whose levels change by a factor of at least two, six were found to contain at least one such Q/N-rich segment \((P \text{ value } 0.025, \text{Fischer exact test)}\) (Fig. 5). By contrast, only one of the 26 proteins, New1, which changed in the MA6 strain, was found to contain such a segment \((P \text{ value } 0.8, \text{Fischer exact test)}\) (Fig. 5). However, New1 was also found to change in MA3, thus indicating a more general explanation for this change.

In Vitro Binding Assays Demonstrate the Importance of Subunit CCT3 in Binding Q/N-Rich Regions. To further test whether subunit CCT3 interacts with Q/N-rich regions, we established an in vitro native gel-shift assay for the binding of CCT to aggregates of Q/N-rich proteins. Using this assay, we characterized the binding of purified CCT complexes to aggregates of Lsm4-EYFP, a Q/N-rich protein that is involved in P-body formation. Wild-type CCT was found to have the highest affinity for these aggregates, followed by CCT(CCT6D89E) and then CCT(CCT3D91E) with the lowest affinity (Fig. 6). The lower affinity for the aggregates of subunits CCT3 and CCT6 with the mutations may be due to the larger fraction of time they spend in an ATP-bound state. Strikingly, the rank order of affinities found in the in vitro assays mirrors the stronger interaction in vivo with Q/N-rich proteins of CCT3 compared with CCT6 found using the high-throughput microscopy screening.

Conclusions
In this study we have shown, using high-throughput microscopy screening, that subunit CCT3 has a stronger interaction in vivo with Q/N-rich proteins than subunit CCT6. Our results do not rule out, however, the interaction of other CCT subunits, such as CCT1 (22), with Q/N-rich proteins, and they also do not imply that CCT3 is involved in the folding of Q/N-rich proteins only. A subset of Q/N-rich proteins involved in mRNA decay seems to be particularly affected by the mutation in subunit CCT3 (Figs. 2 and 5), although many other Q/N-rich proteins are also affected (Fig. 3B). Given that P-body formation can be induced by depletion of one of its components (18), the link between P-body formation and CCT that is found here might be due to misfolding of a CCT3-dependent substrate (e.g., Pub1 whose level in MA3 is found to be lower than in wild-type by a factor of ~10). Our experimental findings are supported by a computational structural analysis of the apical domain of CCT3 that revealed a putative binding site for six Q/Ns in a β-strand conformation (Fig. 4).
4). Such a Q/N binding site is not observed in the apical domain of GroEL that is also found to contain much fewer anchoring spots for Q/N. This structural difference between the apical domains of CCT3 and GroEL may reflect the fact that eukaryotes (including yeast) contain more Q/N-rich proteins than mesophilic bacteria such as E. coli (Fig. 7 and ref. 25). Consequently, it is not surprising that binding sites for Q/N-rich segments have evolved in eukaryotic but not prokaryotic chaperonins. We anticipate that using the approach outlined here for high-content screening of genetically engineered libraries will facilitate elucidating the specificities of other cellular folding machineries and their components.

Methods

Synthetic Genetic Arrays. The S. cerevisiae (S288C) Y6545 strain was used to build the SGA query strains (Δhis3 Δleu2, lys2+ Δmet15, Δura3, Δcan:MFA1pr-leu2, Δlyp1, cyh2, Tef2pr-mCherry::URA) with the mutations D91E and D89E in subunits CCT3 and CCT6, respectively. Introduction of the mutations in CCT3 and CCT6 was performed by homologous recombination as previously described.

Fig. 4. Mapping of anchoring spots for Q/N on the surfaces of the apical domain of GroEL (Left) and CCT3 (Center and Right). The solvent accessible surfaces of the apical domains of CCT3 (26) and GroEL (27) are shown with the substrate binding sites highlighted in brown. The terminal CH₂CONH₂ groups of the anchors with ΔG ≤ −4 kcal/mol are shown, with the C, N and O atoms in green, blue, and red, respectively. Arrows in Center highlight two stretches of approximately equally spaced anchors. (Right) A model of a poly-Q β-strand bound to CCT3 and occupying the long stretch of anchoring sites. Figure prepared with Chimera software (33, 34).

Fig. 5. Graphical representation of the most Q/N-rich segments of seven residues in the proteins that aggregate or whose levels increase or decrease by a factor of two in the MA3 and MA6 strains. Light gray, dark gray, and black designate Q, N, and all other residues, respectively.
described (14), and the resulting strains are designated here as MA3 and MA6. A library of haploid S. cerevisiae strains, each expressing an endogeous locus yeast protein tagged in-frame at its C terminus to GFP (17), was crossed with these MA3 and MA6 query strains. These diploid strains were subjected to sporulation and selection on SD (– arg, – lys, – ura, – leu, – his, + S-AEC, + canavanine, + G418) plates for haploid strains that carry the mutation in CCT and express a given GFP-tagged protein. SI Methods provides more details.

High-Throughput Fluorescence Microscopy. Microscopic screening was performed using an automated microscopy setup as previously described (15). Cells were moved from agar plates into liquid in 384-well polystyrene growth plates using the RoTor arrayer. Liquid cultures were grown over-

Gel-Shift Assays. Equal amounts (2 µg) of wild-type or mutant CCT complexes that were purified as before (29) were mixed with increasing amounts of extract containing aggregates of Lsm4 fused to EYFP obtained from cells harboring the pAG424GAL expression plasmid of this protein (30, 31). The reaction mixtures were incubated for 5 min on ice and then separated by native gel electrophoresis. The amounts of free and aggregate-bound CCT were determined by Western blotting using a mouse monoclonal anti-pol-

Mapping Anchoring Spots. ANCHORSMAP (28) was used to detect preferred binding locations of Asn and Gln residues on the surfaces of the apical domains of CCT3 [Protein Data Bank (PDB) code 1gml] and GroEL (PDB code 1kid). ANCHORSMAP detects surface cavities adequate for binding single amino acid side chains, scatter thousands of excited amino acid probes in the vicinity of these cavities, optimizes their positions, and accurately estimates the ∆G of binding, taking into consideration that the amino acid is part of a hypothetical protein. Anchoring spots with ∆G ≤ −4 kcal/mol were considered in subsequent analyses because such low energy anchoring spots were found to often correspond to experimentally detected hot spots (28). Solvent accessibility was calculated with NACCESS v2.1 (32).

ACKNOWLEDGMENTS. We thank Dr. Simon Alberti for the collection of plasmids expressing Q/N-rich proteins fused to EYFP and Dr. Roy Parker for the Edc3-mCherry plasmid. This work was supported by the Kahn Center for Systems Biology grant (to M.B.) and by Grant 153/08 of the Israel Science Foundation (to A.H.). Part of this work was carried out while K.R.W. was a Weston Visiting Professor at the Weizmann Institute. A.H. is an incumbent of the Carl and Dorothy Bennett Professorial Chair in Biochemistry.


