Journal of Structural Biology **135,** 157–169 (2001)
doi:10.1006/jsbi.2001.4353, available online at http://www.idealibrary.com on IDEAL®

Gene Duplication and the Evolution of Group II Chaperonins: Implications for Structure and Function

John M. Archibald, Christian Blouin, and W. Ford Doolittle

Program in Evolutionary Biology, Canadian Institute for Advanced Research, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia, Canada

Received January 31, 2001, and in revised form March 27, 2001; published online July 23, 2001

Chaperonins are multisubunit protein-folding assemblies. They are composed of two distinct structural classes, which also have a characteristic phylogenetic distribution. Group I chaperonins (called GroEL/cpn60/hsp60) are present in Bacteria and eukaryotic organelles while group II chaperonins are found in Archaea (called the thermosome or TF55) and the cytoplasm of eukaryotes (called CCT or TriC). Gene duplication has been an important force in the evolution of group II chaperonins: Archaea possess one, two, or three homologous chaperonin subunit-encoding genes, and eight distinct CCT gene families (paralogs) have been described in eukaryotes. Phylogenetic analyses indicate that while the duplications in archaeal chaperonin genes have occurred numerous times independently in a lineage-specific fashion, the eight different CCT subunits found in eukaryotes are the products of duplications that occurred early and very likely only once in the evolution of the eukaryotic nuclear genome. Analyses of CCT sequences from diverse eukaryotic species reveal that each of the CCT subunits possesses a suite of invariant subunit-specific amino acid residues ("signatures"). When mapped onto the crystal structure of the archaeal chaperonin from *Thermoplasma acidophilum*, these signatures are located in the apical, intermediate, and equatorial domains. Regions that were found to be variable in length and/or amino acid sequence were localized primarily to the exterior of the molecule and, significantly, to the extreme tip of the apical domain (the "helical protrusion"). In light of recent biochemical and electron microscopic data describing specific CCT–substrate interactions, our results have implications for the evolution of subunit-specific functions in CCT.

Key Words: **chaperonins; CCT; TriC; thermosome; TF55; gene duplication; evolution; archaea; eukaryotes.**

INTRODUCTION

The chaperonins are a ubiquitous class of molecular chaperone involved in the folding of nonnative proteins (reviewed by Bukau and Horwich, 1998; Ranson *et al.*, 1998; Sigler *et al.*, 1998). Individual chaperonin monomers assemble to form multisubunit complexes that harbor substrates within a central cavity and facilitate protein folding through the hydrolysis of ATP. On the basis of protein sequence similarity, two distantly related types of chaperonins are apparent, and these have distinctive phylogenetic distributions. Group I chaperonins are found in Bacteria and their eukaryotic organellar derivatives, while group II chaperonins are present in Archaea (Archaebacteria) and the eukaryotic cytosol (Willison and Horwich, 1996; Willison and Kubota, 1994). Despite sharing only 20–25% amino acid sequence identity, the basic architecture of group I and group II chaperonins is very similar, and both assemble to form multisubunit oligomers with double-ring quaternary structures (Braig *et al.*, 1994; Ditzel *et al.*, 1998). Both chaperonin types possess equatorial and apical domains linked by a flexible intermediate domain. The equatorial domain is involved in the binding and hydrolysis of ATP and provides most of the intra- and interring contacts, while the apical domain is involved primarily in the binding of substrate (Sigler *et al.*, 1998).

There are also, however, important structural and functional differences between group I and group II chaperonins. In bacteria, chaperonin monomers assemble to form seven-membered rings (Braig *et al.*, 1994), while archaeal and eukaryotic cytosolic chaperonin complexes are composed of eight- or nine-membered rings (reviewed in Gutsche *et al.*, 1999; Klumpp and Baumeister, 1998; Willison and Horwich, 1996). Second, while the cochaperonin GroES/cpn10 serves as a "lid" for the group I chaperonin complex, no such cochaperonin functions in the

1047-8477/01 $35.00

group II chaperonin system. Instead, group II chaperonins possess a built-in lid—an extension of the apical domain called the helical protrusion—that is absent in group I chaperonins. The helical protrusions are thought to seal off the central cavity of the chaperonin oligomer in a fashion analogous to GroES/cpn10 and have also been suggested to play a role in substrate recognition (Horwich and Saibil, 1998; Klumpp and Baumeister, 1998; Klumpp *et al.*, 1997). Finally, group II chaperonins are known to interact with protein cofactors that have no apparent homologs in the bacterial system (Gebauer *et al.*, 1998; Geissler *et al.,* 1998; Siegers *et al.*, 1999; Siegert *et al.*, 2000; Vainberg *et al.*, 1998). The precise role of these cofactors in the protein folding process, and how they interact with the group II chaperonins, is as yet unclear.

Group I and group II chaperonins may also differ in their degree of substrate specificity. While the *Escherichia coli* chaperonin GroEL is known to mediate the folding of a wide range of proteins (Houry *et al.*, 1999), the diversity of characterized substrates that interact with the eukaryotic cytosolic chaperonin CCT (chaperonin containing TCP-1; also known as TriC (TCP-1 ring complex)) seems more limited. The abundant cytoskeletal proteins actin and tubulin are the predominant substrates of CCT (Willison and Horwich, 1996). However, others do continue to be identified (Farr *et al.*, 1997; Feldman *et al.*, 1999; Melki *et al.*, 1997; Won *et al.*, 1998), and the full range of *in vivo* CCT substrates is in fact unknown (Leroux and Hartl, 2000; Willison, 1999).

A striking feature of the group II chaperonins is their hetero-oligomeric composition. While bacterial chaperonins (e.g., GroEL in *E. coli*) are generally homo-oligomeric (constructed from seven identical subunits), the CCT complex in eukaryotes is completely hetero-oligomeric, possessing eight distinct (but clearly homologous) subunits (CCTα, β, γ, δ, ϵ, ζ, η, and θ) whose genes evolved from a common ancestor by gene duplication (Kubota *et al.*, 1994; Willison and Kubota, 1994). Each subunit species is thought to occupy a unique position in each of the eight-membered CCT rings (Liou and Willison, 1997), and recent biochemical and electron microscopic studies have revealed that the binding of actin and tubulin to CCT is both subunit-specific and geometry dependent (i.e., a unique arrangement of CCT subunits is required (Llorca *et al.*, 1999, 2000)). The transition from homo-oligomeric chaperonin rings like those in *E. coli* to completely hetero-oligomeric ones like CCT was therefore an important step in chaperonin evolution.

We have previously attempted to shed light on this transition by studying the evolution of archaeal chaperonins, which exhibit an intermediate degree of complexity. Chaperonins in archaea were first described, as heat shock proteins, in two hyperthermophilic organisms, *Sulfolobus shibatae* and *Pyrodictium occultum* (Phipps *et al.*, 1991, 1993; Trent *et al.*, 1991). The *Sulfolobus* chaperonin showed remarkable amino acid sequence similarity with t-complex polypeptide-1 (TCP-1) (Willison *et al.*, 1986), now known to be the α subunit of CCT (Frydman *et al.*, 1992; Lewis *et al.*, 1992). The chaperonin complex from another archaeon, *Thermoplasma acidophilum*, has been crystallized and is composed of two homologous subunits, α and β (Ditzel *et al.*, 1998). Archaea that possess a single chaperonin gene or three distinct subunit genes have also been identified (Archibald *et al.*, 1999; Gutsche *et al.*, 1999).

From phylogenetic analysis of archaeal and eukaryotic chaperonin protein sequences, we have concluded that (1) hetero-oligomeric chaperonin complexes have arisen multiple times independently in archaea; (2) a completely hetero-oligomeric chaperonin particle (consisting of eight distinct subunit species) evolved very early in eukaryotic evolution, likely in the common ancestor of all eukaryotic cells; and (3) the transition from homo-oligomerism to hetero-oligomerism in eukaryotes and that in various archaeal lineages likely occurred independently of one another (Archibald *et al.*, 1999, 2000). Here we further examine the molecular evolution of archaeal and eukaryotic chaperonins and present new data describing the distribution of slowly evolving subunit-specific "signatures" and variable regions among the different CCT subunits, mapped onto the crystal structure of the *T. acidophilum* chaperonin (Ditzel *et al.*, 1998). These signatures are found in the equatorial, intermediate, and apical domains of the protein. Within the apical domain, a distinct cluster of subunit-specific signatures is located in the putative substrate-binding region of several of the CCTs: these signatures likely mediate the subunit-specific functions described for the binding of actin (Llorca *et al.*, 1999), tubulin (Llorca *et al.*, 2000), and perhaps other CCT substrates.

METHODS

Protein sequence alignments. New archaeal chaperonin protein sequences were identified in the public databases by BLAST (Altschul *et al.*, 1997) and were added manually to an alignment constructed previously (Archibald *et al.*, 1999) based on universally conserved regions. Chaperonin sequences from the crenarchaeote *Pyrobaculum aerophilum* were obtained from the *Pyrobaculum* project homepage (http://www.doe-mbi.ucla.edu/PA/). The final alignment used for phylogenetic analyses contained 40 sequences and, after the removal of ambiguous regions, contained 452 amino acid positions. For CCT, an alignment containing 10 representative archaeal sequences, 5 sequences from each of CCTα, β, γ, δ, ϵ, η, and θ, and 8 CCT ζ sequences was used (53 sequences in total). The alignment contained 355 unambiguously aligned amino acid positions. Both alignments are available from J.M.A. (e-mail: johna@hades.biochem.dal.ca).

Phylogenetic analysis. Phylogenetic trees were inferred using distance-based and maximum-likelihood (ML) methods of tree reconstruction. For distance analyses, the programs PROTDIST (PAM matrices) and FITCH were used in the PHYLIP 3.57 package (Felsenstein, 1995). Statistical support for distance trees was obtained by bootstrapping with 100 resampling replicates, and the majority-rule consensus tree was constructed using CONSENSE (Felsenstein, 1995). For ML analyses, protML was used in MOLPHY (Adachi and Hasegawa, 1996), with the Jones, Taylor, and Thornton amino acid substitution matrix, adjusted to account for the amino acid frequencies observed in the dataset (JTT-F). RELL (resampling estimated log likelihoods) bootstrap values were obtained by quick-add OTU (-q) ML searches of 1000 trees in protML and were used as measures of statistical support for branches in ML trees.

Structural analyses. The crystal structure of the α subunit of the *T. acidophilum* chaperonin (Ditzel *et al.*, 1998) (PDB ID: 1A6D) was visualized using Insight II (MSI, San Diego, CA). Individual amino acid residues as well as specific regions of the structure were highlighted using a script written by C. Blouin (Dalhousie University, Halifax, Nova Scotia, Canada).

RESULTS AND DISCUSSION

Archaeal Chaperonin Evolution: Recurrent Paralogy

The chaperonin complex in *T. acidophilum* (called the thermosome) is composed of two homologous subunits, α and β, that alternate in each of its eight-membered rings (Ditzel *et al.*, 1998; Nitsch *et al.*, 1997). The two subunits share ~60% amino acid sequence identity (Waldmann *et al.*, 1995a,b) and evolved from a common ancestral chaperonin by gene duplication. When compared to chaperonin protein sequences from other archaea, the two subunits are more closely related to each other than to other archaeal chaperonin sequences, suggesting that the gene duplication that produced them occurred after this lineage diverged from other archaea (Archibald *et al.*, 1999). In fact, "lineage-specific" gene duplication (paralogy) is a recurring theme in the evolution of archaeal chaperonins. In light of several newly sequenced archaeal genomes, we reinvestigated the phylogeny of archaeal chaperonins with a more complete dataset than was available previously.

Figure 1 shows the results of a phylogenetic analysis performed with all available chaperonin protein sequences (January 2001). The domain Archaea is composed of two evolutionarily distinct lineages, euryarchaeotes (e.g., *Thermoplasma*) and crenarchaeotes (e.g., *Sulfolobus*). Within euryarchaeotes, lineage-specific duplications of chaperonin genes have occurred in *Methanobacterium thermoautotrophicum*, *Archaeoglobus fulgidus*, and the *Thermococcus/Pyrococcus* clade, in addition to the duplication in the *Thermoplasma* lineage. Recently, the genome of *Thermoplasma volcanium*, an organism closely related to *T. acidophilum*, was completely sequenced (Kawashima *et al.*, 2000). *T. volcanium* also possesses genes encoding two subunits very similar to those in *T. acidophilum*; this indicates that the duplications producing the subunits occurred in their common ancestor, but distinct from duplications in other archaea (Fig. 1). Two separate gene duplications have also occurred within the halophilic archaea (e.g., *Haloferax*). It should be understood that, for the most part, the designation of duplicate subunits as α and β (or a and b) in euryarchaeotes is completely arbitrary. For example, there is no sense in which the α subunit in *T. acidophilum* is more similar to the α subunit of *Ar. fulgidus* than it is to the *T. acidophilum* β subunit.

As is the case for euryarchaeotes, chaperonin se-

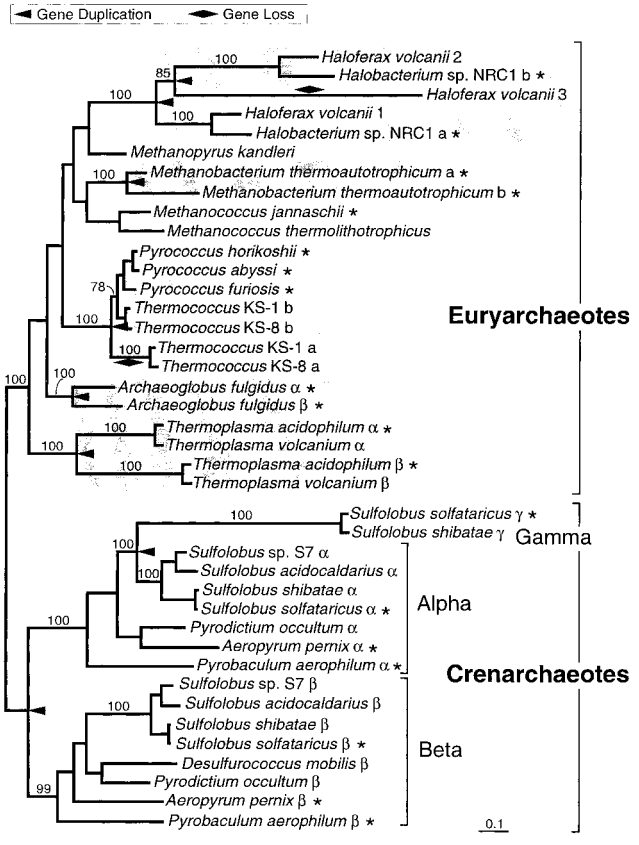


FIG. 1. Phylogenetic analysis of archaeal chaperonins. The tree shown is a ML tree (lnL −15614.72) inferred from a chaperonin protein sequence alignment containing 40 sequences and 452 unambiguously aligned amino acid positions. The two recognized kingdoms within Archaea (euryarchaeotes and crenarchaeotes) are labeled, and inferred gene duplications and gene losses are indicated (see text). Within euryarchaeotes, regions of the tree in which lineage-specific gene duplications have occurred are shaded. For crenarchaeotes, the three different gene/subunit families (α, β, and γ) are indicated. Asterisks appear next to sequences from organisms whose genomes have been completely sequenced. Statistical support values for significant nodes appear above the branches (ML RELL bootstrap values; inferred from a heuristic search of 1000 trees in protML) (Adachi and Hasegawa, 1996) (see text). The scale bar represents the estimated number of amino acid substitutions per site.

quences are available from a wide diversity of crenarchaeotes, and recently, the complete sequences of three crenarchaeal genomes were released (*Aeropyrum pernix* (Kawarabayasi *et al.*, 1999); *Pyrob. aerophilum* (http://www.doe-mbi.ucla.edu/PA/); and *Sulfolobus solfataricus* (http://www.cbr.nrc.ca/sulfhome/)). The phylogenetic tree in Fig. 1 shows that a gene duplication producing α and β subunits occurred in the common ancestor of the *Sulfolobus*, *Pyrodictium*, *Desulfurococcus*, *Aeropyrum*, and *Pyrobaculum* lineages. This can be inferred because the α sequences from these organisms cluster together with high support to the exclusion of the β sequences (Fig. 1). An additional gene duplication has also occurred within *Sulfolobus* evolution, as a third chaperonin subunit (named γ) clusters with the α sequences from four different *Sulfolobus* species to the exclusion of those in *Pyrod. occultum*, *Ae. pernix*, and *Pyrob. aerophilum*.

Where complete genome sequences are available, several instances of archaeal chaperonin subunit loss can also be inferred. For example, each of the *Thermococcus* strains KS-1 and KS-8 has two subunits, "a" and "b" (80.6% identical), whose genes duplicated and diverged prior to the splitting of these two lineages (Fig. 1). Significantly, three completely sequenced *Pyrococcus* genomes (*Pyrococcus furiosis*, *Pyrococcus abyssi*, and *Pyrococcus horikoshii*) reveal the presence of a single chaperonin subunit gene. When compared to the *Thermococcus* proteins, the three *Pyrococcus* sequences appear most similar to the b subunit from the *Thermococcus* strains. This suggests that the duplication and divergence of the a and b chaperonin subunit genes in these lineages occurred in their common ancestor and that subunit a was subsequently lost in the *Pyrococcus* lineage. Chaperonin gene loss can also be inferred for the halophilic euryarchaeote *Halobacterium* sp. NRC-1. The genome of this organism has recently been completely sequenced (Ng *et al.*, 2000) and possesses genes encoding two subunits. However, a closely related halophile, *Haloferax volcanii*, has at least three chaperonin subunits, and the two *Halobacterium* subunits cluster strongly with two of the three *H. volcanii* sequences in phylogenetic analyses (Fig. 1). Again, one can infer that the ancestor of these two organisms had three chaperonin subunits and that in *Halobacterium* sp. NRC-1, the third subunit was lost.

Taken as a whole, phylogenetic analysis of archaeal chaperonins reveals a complex evolutionary pattern. In euryarchaeotes, multiple lineage-specific gene duplications and divergences have occurred, while in crenarchaeotes, a duplication into α and β subunits occurred in their common ancestor. In *Sulfolobus* species, an additional duplication and divergence occurred; we have suggested that the evolution of a third subunit (γ) may have been involved in the transition from eight- to nine-membered chaperonin rings in this lineage (Archibald *et al.*, 1999). Overall, the significance of our results in terms of chaperonin structure and function is that archaeal chaperonins have made the transition from homo-oligomeric chaperonin complexes to hetero-oligomeric ones *at least* six times independently, and duplicate subunits have been lost at least twice (resulting in a reversion from hetero- to homo-oligomerism; Fig. 1).

Eukaryotic Chaperonin Evolution: Deep Paralogy

The pattern of chaperonin subunit evolution in eukaryotes is in stark contrast to that observed in archaea. Biochemical and gene sequence analyses (Frydman *et al.*, 1992; Kubota *et al.*, 1994, 1995b; Lewis *et al.*, 1992; Rommelaere *et al.*, 1993) initially revealed the presence of eight distinct CCT subunits in mammals, and orthologs of the genes encoding these subunits were also found in the yeast *Saccharomyces cerevisiae* (Kubota *et al.*, 1994; Stoldt *et al.*, 1996). Taking advantage of a more comprehensive dataset, we recently showed that all eight CCT subunits were already present in the common ancestor of animals, fungi, plants, and many if not all protists (Archibald *et al.*, 2000). We isolated genes encoding CCT subunits from *Trichomonas vaginalis* and *Giardia lamblia*, representative species of two amitochondriate protist lineages that have been suggested to have diverged early in eukaryotic evolution (Cavalier-Smith, 1987; Hashimoto *et al.*, 1994; Leipe *et al.*, 1993; Sogin *et al.*, 1989; Stiller *et al.*, 1998). At present, there is considerable controversy surrounding the precise nature of the relationships between these organisms and other eukaryotes (for comprehensive review, see Roger, 1999), but it is abundantly clear that they are as evolutionarily distant from animals and fungi as any eukaryotes currently known. Figure 2 shows the results of a phylogenetic analysis of the eight different CCT subunit sequences from a wide range of eukaryotes, rooted with a representative sample of archaeal chaperonins. Most obviously, sequences from *Homo*, the nematode *Caenorhabditis*, *Saccharomyces*, the flowering plant *Arabidopsis*, and the amitochondriate protists *Giardia* and *Trichomonas* cluster together in the different CCT subunit families with extremely high statistical support. This indicates that the gene duplications took place in the common ancestor of these organisms.

While it is thus clear that all the CCT subunits were present early in eukaryotic evolution, ordering the duplication events that produced them has proven difficult. On the basis of amino acid sequence identity, the eight CCT subunits appear to be approximately equally closely related (Kubota *et al.*,

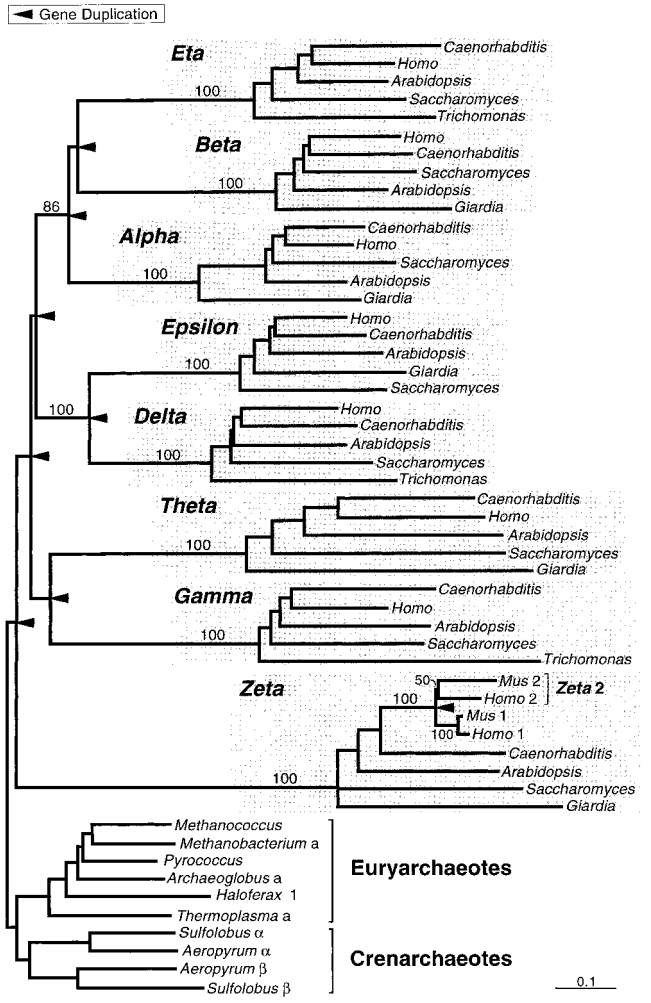


FIG. 2. Phylogeny of eukaryotic CCTs. The tree shown was constructed using the Fitch–Margoliash distance algorithm (Felsenstein, 1995) from an alignment of 53 sequences and 355 unambiguously aligned amino acid sites. The eight different CCT subunit families found in eukaryotes are highlighted in gray. For each CCT subunit family, 5 representative sequences were chosen to represent the full spectrum of eukaryotic diversity: two animals (*Homo* and *Caenorhabditis*) and one from each of fungi (*Saccharomyces*), plants (*Arabidopsis*), and protists (*Giardia* or *Trichomonas*). For CCTζ, the ζ-1 and ζ-2 subunits of *Homo* and *Mus* were also analyzed. The tree is rooted with a representative sample of archaeal chaperonin sequences. Support values for important nodes on the tree are given above the branches and were calculated by bootstrapping with 100 resampling replicates. The scale bar represents the estimated number of substitutions per amino acid site.

1994, 1995a). However, phylogenetic analysis reveals that some of the CCTs do appear to share more recent common ancestry than others. The CCTδ and ϵ subunits appear as a well-supported clade, as do CCTα, η, and β, and were likely the product of more "recent" gene duplications (Fig. 2). Beyond the resolution of these two groups, however, the phylogenetic signal contained in the molecules is insufficient to resolve the exact pattern of subunit diversification. The tree topology shown in Fig. 2 illustrates one possible scenario, but others are also compatible with the data (see Fig. 4b in Archibald *et al.*, 2000). This may be due to the fact that numerous gene duplications (and sequence divergences) occurred in quick succession or that chromosome or whole genome duplications produced a dramatic increase in the number of CCT genes simultaneously. Alternatively, the molecules themselves may simply be "saturated," having accumulated a sufficient number of amino acid substitutions as to make accurate reconstruction of their evolutionary history impossible.

More recently, an additional gene duplication occurred in the evolution of CCT. Two distinct CCTζ subunits (1 and 2) have been characterized in mouse (Hynes *et al.*, 1995; Kubota *et al.*, 1997), and the two sequences share approximately 81% amino acid sequence identity (Kubota *et al.*, 1997). Phylogenetic analysis of these sequences and their human orthologs (Fig. 2) reveals that the CCTζ-1 and CCTζ-2 sequences in mouse and human are more closely related to each other than to CCTζ sequences from other eukaryotes. This suggests that the gene duplication producing them occurred quite recently, likely within mammalian evolution. Interestingly, the CCTζ-2 subunit gene is expressed only in testis, unlike the expression pattern observed for other CCTs, which are expressed in all tissues investigated thus far, including testis (Kubota *et al.*, 1995a, 1997). A specific function for the CCTζ-2 subunit has not been identified, but Kubota *et al.* (1997) suggest that it may mediate the folding of testis-specific proteins, such as α-tubulin isoforms, by substituting for the CCTζ-1 subunit in the CCT complex. Interestingly, genes encoding two different CCTζ subunits are also present in the genome of the flowering plant *Arabidopsis thaliana,* and duplicates of other CCT subunits are also present in mammalian genomes (J. M. Archibald, unpublished results).

Regions of Variability among the Different CCT Subunits

The phylogenetic tree of CCTs in Fig. 2 reveals that, compared to the archaeal chaperonins, the branches leading to the different CCT subunits are remarkably long. In other words, an extreme degree of amino acid sequence divergence has occurred in each of the CCT subunit families. A close examination of the protein sequences shows that some regions of the protein are variable in length as well as sequence (Archibald *et al.*, 2000; Ditzel *et al.*, 1998; Kim *et al.*, 1994; Kubota *et al.*, 1994; Waldmann *et al.*, 1995a). It is significant that these regions do not necessarily correspond to areas of variability *within* a particular CCT subunit family. In fact, many of

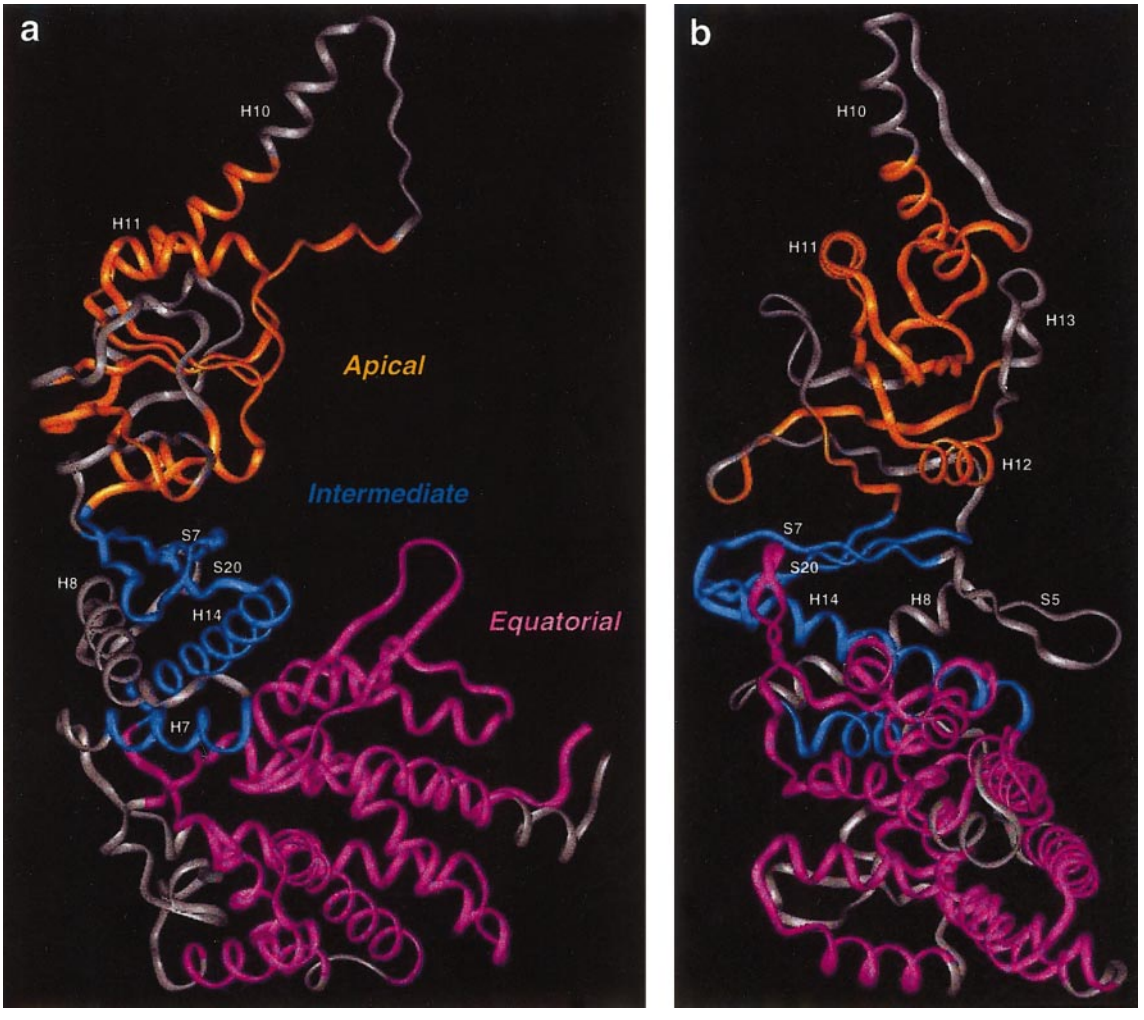
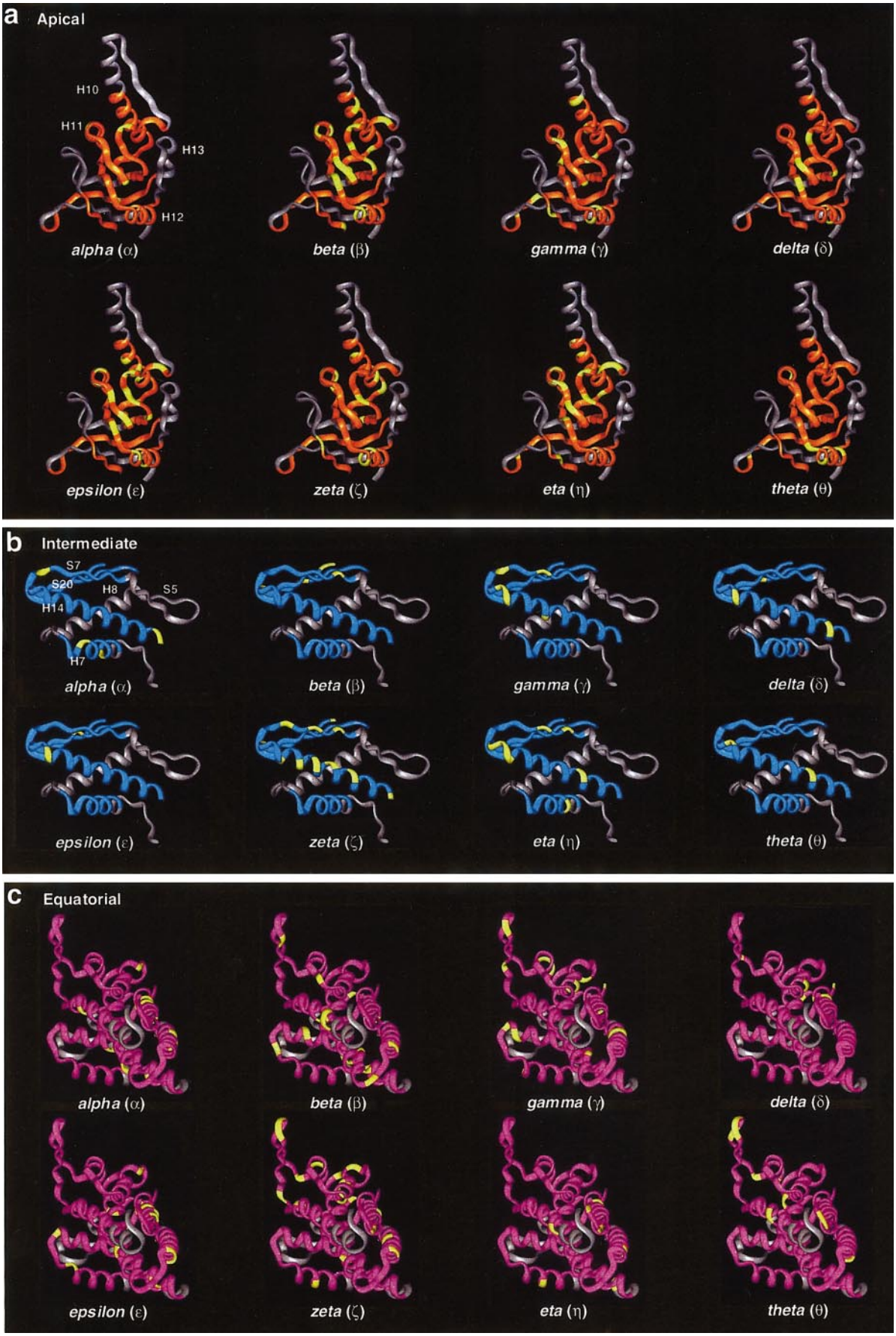


FIG. 3. Regions of variability among the different CCT subunits. Areas of the chaperonin protein sequence that were found to be highly variable in length and/or amino acid sequence between the eight different CCTs were mapped onto the crystal structure of the α subunit of the *Thermoplasma acidophilum* thermosome, the archaeal homolog of CCT (Ditzel *et al.,* 1998). The thermosome consists of two homologous subunits, α and β; since the two subunits are essentially identical in structure (Ditzel *et al.,* 1998), the α subunit was arbitrarily used. (a) Side view. The apical, intermediate, and equatorial domains are labeled and color-coded, and the regions of the molecules found to be variable among the CCT subunits are highlighted in gray. (b) Same structure as in (a), but viewed from within the central cavity of the thermosome. Secondary structural elements are labeled according to Ditzel *et al.* (1998) (H, helix; S, sheet).

the subunits possess insertions that are themselves invariant in sequence across a wide range of eukaryotic species but are not present in any of the other subunits (Archibald *et al.*, 2000; Kim *et al.*, 1994). To investigate what these observations might mean in terms of CCT subunit structure and function, we mapped the variable regions—areas that were highly variable in length and/or amino acid sequence—onto the crystal structure of the α subunit of the *T. acidophilum* thermosome.

The results (Figs. 3a and 3b) show that while regions of variability map to all three "functional" domains, the intermediate and apical domains appear to possess proportionately more variable sequence than the equatorial domain. We found that 38.8, 38.9, and 28.8% of the total amino acid se-

FIG. 4. Subunit-specific "signatures" in individual CCT subunits. Isolated apical (a), intermediate (b), and equatorial (c) domains are shown, one for each of the eight CCT subunits, α, β, γ, δ, ϵ, ζ, η, and θ. The domain orientation is the same as in Fig. 3b. The color-coding matches that shown in Figs. 3a and 3b, and amino acid residues unique to a particular CCT subunit (signatures; see text) are highlighted in yellow. As in Fig. 3, regions found to be variable among the different CCT subunits are highlighted in gray. Secondary structural elements are labeled following Ditzel *et al.* (1998) (H, helix; S, sheet).


a
Apical
H10
H11
H13
H12
alpha (α)
beta (β)
gamma (γ)
delta (δ)
epsilon (ε)
zeta (ζ)
eta (η)
theta (θ)
b
Intermediate
S7
S20
H8
S5
H14
H7
alpha (α)
beta (β)
gamma (γ)
delta (δ)
epsilon (ε)
zeta (ζ)
eta (η)
theta (θ)
c
Equatorial
alpha (α)
beta (β)
gamma (γ)
delta (δ)
epsilon (ε)
zeta (ζ)
eta (η)
theta (θ)

quence in the apical, intermediate, and equatorial domains were variable, and this difference was significant at $P < 0.05$ in a χ^2 test (data not shown). This observation is consistent with that of Kim *et al.* (1994), who found that the equatorial domain, which contains the ATP-binding site, appeared to be the most highly conserved among the different CCT subunits and between the CCTs and *E. coli* GroEL. It is also significant that the variable regions in these domains map predominantly to the exterior of the molecule. A cluster of variable sequence (shaded gray) is clearly evident along most of the "backside" of the molecule (e.g., helix H8; Fig. 4a), corresponding to the outer face of the monomer as it sits in the hetero-oligomer. Regions of variability also map to a loop under helix H11 of the apical domain and to the region surrounding helix H13 (Fig. 4b). In the thermosome, helices H11 and H13 of adjacent subunits contact one another (Ditzel *et al.*, 1998; and see below).

Most interesting is the patch of variability that corresponds to the extreme tip of the apical domain (the helical protrusion). In the *T. acidophilum* thermosome, this region appears to be highly flexible: the analysis of different crystal forms shows that the protrusion has the potential to tilt more than 20° relative to the rest of the apical domain (Klumpp and Baumeister, 1998). Bosch *et al.* (2000) recently crystallized the apical domain of the β subunit and showed that this region adopts a different conformation than that found in the crystal structure of the α subunit apical domain (Klumpp *et al.*, 1997). These authors suggest that sequence variations in the corresponding region of the CCT subunits could produce similar structural differences and that such heterogeneity could mediate subunit-specific roles in substrate recognition. The extreme sequence divergence between the different CCT subunits in this region (Archibald *et al.*, 2000) does make heterogeneity in secondary and tertiary structure particularly likely. However, the exact role of the helical protrusion in the binding of substrate is unclear. While the electron microscopic data available for the CCT–actin complex (Llorca *et al.*, 1999) show actin binding well below the helical protrusions, tubulin seems to interact with a much broader region of the CCT apical domains, including parts of the helical protrusions (Llorca *et al.*, 2000; and see below).

Subunit-Specific Signatures in CCT

In a sequence- and structure-based comparison of group I and group II chaperonins, Kim *et al.* (1994) noted that the different CCT subunits were remarkably divergent, particularly in their apical domains, but that these regions were nevertheless highly conserved within the different subunit classes. We recently investigated this more closely by estimating the rate of amino acid sequence evolution—site by site—across alignments of individual CCT subunits that contained broad taxonomic diversity (Archibald *et al.*, 2000). This method identified numerous residues as being invariant (or nearly so) within a particular subunit family and, more importantly, unique to that subunit.

Here we have mapped the positions of these residues onto the structure of the α subunit of the thermosome (Figs. 4a, 4b, and 4c). These slowly evolving subunit-specific signatures (colored yellow) clearly map to all three domains of the thermosome. As noted previously, some of these signatures are located in the ATP-binding and hydrolysis motifs present in the equatorial domain (Archibald *et al.*, 2000). For example, the CCTθ subunit has the sequence GDGTN, a slight variation on the near-universally conserved motif GDGTT, which forms one of the loops of the ATP-binding pocket (data not shown). In CCTγ, the highly conserved motif NDGAT (found in almost all the CCT subunits) has changed to NDGNA. Surprisingly, multiple alanine replacements of highly conserved residues within the ATP-binding pocket of CCTζ (e.g., GDGTT to AAAAA) had relatively mild effects on cell growth in yeast (Lin *et al.*, 1997), unlike the severe phenotypes observed with mutations in the same region of CCTα (Miklos *et al.*, 1994; Ursic *et al.*, 1994). When the sequences from a wide diversity of eukaryotes are considered, some parts of the ATP-binding domains of the different subunits appear to be less conserved than others (Archibald *et al.*, 2000), suggesting that their ATPase functions may in fact be redundant.

Within the conserved "core" region of the apical domain, many of the CCT subunits possess a cluster of highly conserved, subunit-specific residues on the inside face of the apical domain, just below the helical protrusion (Fig 4a). Significantly, this is precisely the region of the apical domain that appears to be in direct contact with actin and tubulin in CCT-substrate complexes (Llorca *et al.*, 1999, 2000). CCTβ and ϵ, two of the three subunits implicated in the binding of actin (Llorca *et al.*, 1999), have more subunit-specific signatures in this region than the other subunits, while CCTθ has almost none (Fig. 4a). In addition, CCTα and CCTζ have unique insertions in this area, very near helix H11 (data not shown; this region is not indicated as "variable" in Figs. 3 and 4 due to the fact that the CCTα and ζ insertion boundaries are highly conserved with respect to the sequence of the other CCTs and the archaeal chaperonins).

It is interesting that many of the highly conserved subunit-specific amino acids present in the apical domain are charged residues, not hydrophobic residues like those implicated in substrate binding in the *E. coli* GroEL system (Archibald *et al.*, 2000;

Kim *et al.*, 1994; Klumpp *et al.*, 1997; Willison, 1999). This could reflect a fundamentally different mechanism of substrate recognition for the group II chaperonins than that observed for GroEL, one based on hydrogen bonding and electrostatic interactions instead of hydrophobicity (Klumpp *et al.*, 1997; Willison, 1999). However, as noted by Klumpp *et al.* (1997), the helical protrusions of the CCT subunits may contain a sufficient number of hydrophobic residues to mediate substrate binding through hydrophobic interactions. As mentioned previously, the nature of the CCT–tubulin interaction (Llorca *et al.*, 2000) appears consistent with this possibility.

Several areas of the equatorial domain also appear to be "hotspots" for unique amino acid substitutions. For example, all eight subunits possess unique residues in the vicinity of a small turn between helix H14 and strand S20 (Fig. 4b). In the thermosome, this region (together with turn segments between helices H7 and H8, and strands S7 and S8) contacts the equatorial domain of the adjacent subunit (Ditzel *et al.*, 1998). It is thus likely that subunit-specific motifs in these regions (as well as those in the vicinity of helices H11 and H13 in the apical domain; see above) are at least partly responsible for the assembly of a CCT complex with a unique arrangement of subunits. Interestingly, the CCTζ subunit possessed a large number of unique residues in these areas: a string of unique substitutions map to the "outer" surface of helix H14 and to strand S7 (Fig. 4b), both of which face the central cavity in the thermosome structure (Ditzel *et al.*, 1998). The significance of this observation, in terms of chaperonin structure and function, is unclear.

Coevolution of Chaperonin and Substrate?

Willison and co-workers have speculated that the duplication and differentiation of the different CCT subunits early in eukaryotic evolution were concurrent with and facilitated the evolution of the cytoskeleton (see Willison and Horwich, 1996; Willison and Kubota, 1994). Given the overwhelming evidence for subunit-specific interactions between CCT and modern-day actins and tubulins (and perhaps a host of additional cytosolic proteins), it is indeed likely that the cytoskeletal proteins and the various CCT subunits have coevolved with one another (Hartl, 1996; Willison, 1999; Willison and Horwich, 1996). While it is clear that CCT went through intermediate stages of hetero-oligomerism, the process by which such hetero-oligomerism initially arose, and how CCT ultimately became completely hetero-oligomeric, is still a mystery. There are theoretically 5040 possible combinations of subunits in CCT (Liou and Willison, 1997), yet all the available data suggest that a unique arrangement of subunits exists *in vivo* (Liou and Willison, 1997; Llorca *et al.*, 1999, 2000).

Our previous attempt to explain the phylogenetic pattern of archaeal chaperonins led us to propose a model for the origin of hetero-oligomerism in chaperonin complexes that emphasizes coevolution between duplicate *subunits*, as opposed to subunit and substrate (Archibald *et al.*, 1999). Simply put, a series of mutations in the regions of intraring contact of one subunit followed by compensatory changes in a newly evolved duplicate subunit could produce a tendency toward the assembly of an ordered arrangement of subunits within a chaperonin ring (like that observed in the thermosome). Such a process could eventually lead to obligatory hetero-oligomerism, even in the absence of "specialized" roles for the duplicate subunits in protein folding. An important corollary of the fact that hetero-oligomerism evolved multiple times independently in Archaea is that subunit-specific roles in protein folding, *if* they exist, must also have evolved multiple times.

Unfortunately, few data are currently available with which to test this hypothesis: next to nothing is known about the substrates of archaeal chaperonins, and, apart from *T. acidophilum*, very little biochemical and structural information is available on chaperonin complexes from diverse archaeal lineages (Gutsche *et al.*, 1999). The fact that some chaperonin complexes are homo-oligomers and others hetero-oligomers suggests that, for Archaea as a whole, the presence of two (or more) subunits is not a strict requirement for chaperonin function. This is further supported by the fact that duplicate chaperonin genes have occasionally been lost (see above). As suggested previously (Archibald *et al.*, 1999), the extent to which a given lineage can tolerate such loss is likely a function of the degree of sequence divergence between duplicates. Interestingly, the range of similarity between duplicate subunits in archaeal genomes is quite high—between 80.6 and 43.2%.

With respect to the origin of hetero-oligomerism in CCT, the issue becomes a question of which evolved first: subunit-specific roles in protein folding or an ordered arrangement of subunits within a chaperonin ring? Given that CCT–actin and CCT–tubulin interactions are both subunit-specific and geometry-dependent, the evolution of a unique arrangement of the subunits relative to one another would at least have to have been concurrent with the "specialization" of the subunits themselves (Archibald *et al.*, 2000). It is certainly clear that the different CCTs have diverged substantially from one another, not only in their substrate-binding apical domains, but in regions involved in ATP binding and in subunit–subunit contacts (Figs. 3 and 4; Kim *et al.*, 1994).

Genetic studies in yeast support this notion. Lin *et al.* (1997) performed a comprehensive mutational analysis of the CCTζ subunit (CCT6p in *Saccharomyces*) and found this subunit to be sensitive to mutations in the equatorial, intermediate, and apical domains. Presumably, some of these mutations were deleterious, not because of deficiencies in substrate recognition or folding, but because of an inability of the CCTζ subunit to interact with the other subunits to properly assemble the hetero-oligomeric complex.

Llorca *et al.* (1999) showed that actin sits in the central cavity of CCT in a 1,4 configuration. CCTδ interacts with the small domain of actin while either CCTβ or ϵ contacts the large domain. Given this fact, one might expect CCTβ and ϵ to be quite similar. However, if the actin–CCTδ interaction is isomorphous in both arrangements, the CCTβ- and ϵ-binding sites on the large domain of actin would have to be in different locations. Indeed, our phylogenetic analyses show that CCTβ and ϵ are no more closely related to one another than to any of the other subunits (in fact, CCTδ and ϵ appear most closely related; Fig. 2). Further, the two subunits do not appear to share a great degree of sequence similarity in their putative substrate-binding domains, no more than is shared with any other subunit. It therefore seems likely that the actin–CCTβ and actin–CCTϵ interactions are very different. Similarly, no detectable correlation was observed between the various subunits implicated in binding the different domains of tubulin. From the perspective of substrate diversity, it is interesting that α- and β-tubulin appear to interact with CCT in an identical fashion (Llorca *et al.*, 2000), despite the fact that they share only approximately 40% sequence identity (see Ritco-Vonsovici and Willison (2000) for discussion; the diversification of the α-, β-, and γ-tubulin homologs occurred early in eukaryotic evolution (Keeling and Doolittle, 1996; Roger, 1999), similar to the pattern observed for the CCT subunits themselves). The exact nature of the specificity between the apical domains of the CCT subunits and their various substrates is thus likely to be extraordinarily complex, perhaps involving a combination of hydrophobic and nonhydrophobic interactions (Leroux and Hartl, 2000; Ritco-Vonsovici and Willison, 2000).

It is not clear to what extent changes in substrates can affect the evolution of the CCT subunits, but a possible example of "recent" coevolution between CCT and its substrates has been described in cryptomonad algae. Cryptomonads are unusual cells: they possess an additional cytosolic compartment, the result of having acquired their chloroplasts secondarily by engulfing and retaining a photosynthetic eukaryote (Gilson *et al.*, 1997; Zauner *et al.*, 2000). In the cryptomonad *Guillardia theta*, the "nucleomorph" genome (the vestigial nuclear genome of the endosymbiont) encodes orthologs of each of the eight CCT subunits (Archibald *et al.*, 2001), and—despite the fact that microtubules have never been observed in the symbiont—genes for α-, β-, and γ- tubulins (Keeling *et al.*, 1999). Not surprisingly, the tubulins (γ-tubulin in particular) were quite divergent, probably as a result of reduced and/or modified functional constraints in such an unusual cellular environment. Most interesting was the fact that the CCT subunits themselves were equally divergent. As a control, two other nucleomorph-encoded molecular chaperones, hsp70 and hsp90, were found to be remarkably well conserved and were indistinguishable from other eukaryotic cytosolic homologs (Archibald *et al.*, 2001). Interestingly, a detectable homolog of actin was conspicuously absent from the complete sequence of the nucleomorph genome, as were genes for the newly described cofactors of CCT, the GimC subunits. This unique cellular environment may represent a case where extreme divergence in a set of chaperonin substrates (or indeed, loss of a substrate) has resulted in a high degree of divergence in the chaperonin subunits themselves.

CONCLUSION

Gene duplication has been a driving force in group II chaperonin evolution. Phylogenetic analyses of group II chaperonins reveal a complex pattern of gene duplication and gene loss. The eukaryotic chaperonin CCT has evolved into a remarkably sophisticated protein folding "machine," and it is likely that the duplicate CCT subunits have coevolved with one another and with actin, tubulin, and other proteins in the eukaryotic cytosol. While the evolutionary processes that gave rise to the completely hetero-oligomeric CCT complex are obscure, a better understanding of the moderately hetero-oligomeric chaperonins found in diverse archaeal lineages would undoubtedly help.

There are many issues to be addressed. In Archaea, are duplicate subunits always present in the same ring (like the thermosome and CCT), or have multiple distinct homo-oligomeric chaperonin complexes evolved in some lineages? What are the *in vivo* substrates of archaeal chaperonins? Do they interact with a wide range of substrates like GroEL or are they more selective, as appears to be the case for CCT? Will newly discovered CCT substrates prove to have subunit-specific interactions with the chaperonin like those observed for actin and tubulin? Answers to these questions should bring us closer to a more complete understanding of the origin and evolution of the CCT complex in the eukaryotic cytosol and its role in protein folding. The sub-

unit-specific residues that we have identified for CCT should make excellent targets for mutational studies aimed at probing the structure and function of the CCT subunits, in terms of their contribution to substrate binding, ATPase activity, and the overall assembly and function of the CCT complex.

We thank A. Roger, Y. Inagaki, and C. Nesbø for discussion and for critically reviewing the manuscript. M. Leroux is thanked for helpful discussions on chaperonin evolution. This work was supported by a grant awarded to W.F.D. by the Canadian Institutes for Health Research (CIHR). W.F.D. was supported by the Canadian Institute for Advanced Research (Program in Evolutionary Biology) and J.M.A. by a CIHR Doctoral Research Award.

REFERENCES

Adachi, J. M., and Hasegawa, M. (1996) MOLPHY version 2.3: Programs for molecular phylogenetics based on maximum likelihood, *Comput. Sci. Monogr.* **28,** 1–150.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs, *Nucleic Acids Res.* **25,** 3389–3402.

Archibald, J. M., Cavalier-Smith, T., Maier, U., and Douglas, S. (2001) Molecular chaperones encoded by a reduced nucleus: The cryptomonad nucleomorph, *J. Mol. Evol.* **52,** 490–501.

Archibald, J. M., Logsdon, J. M., and Doolittle, W. F. (1999) Recurrent paralogy in the evolution of archaeal chaperonins, *Curr. Biol.* **9,** 1053–1056.

Archibald, J. M., Logsdon, J. M., Jr., and Doolittle, W. F. (2000) Origin and evolution of eukaryotic chaperonins: Phylogenetic evidence for ancient duplications in CCT genes, *Mol. Biol. Evol.* **17,** 1456–1466.

Bosch, G., Baumeister, W., and Essen, L. O. (2000) Crystal structure of the beta-apical domain of the thermosome reveals structural plasticity in the protrusion region, *J. Mol. Biol.* **301,** 19–25.

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

Bukau, B., and Horwich, A. L. (1998) The Hsp70 and Hsp60 chaperone machines, *Cell* **92,** 351–366.

Cavalier-Smith, T. (1987) Eukaryotes with no mitochondria, *Nature* **326,** 332–333.

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

Farr, G. W., Scharl, E. C., Schumacher, R. J., Sondek, S., and Horwich, A. L. (1997) Chaperonin-mediated folding in the eukaryotic cytosol proceeds through rounds of release of native and nonnative forms, *Cell* **89,** 927–937.

Feldman, D. E., Thulasiraman, V., Ferreyra, R. G., and Frydman, J. (1999) Formation of the VHL-elongin BC tumor suppressor complex is mediated by the chaperonin TRiC, *Mol. Cell* **4,** 1051–1061.

Felsenstein, J. (1995) PHYLIP (Phylogeny Inference Package). Distributed by the author, Department of Genetics, University of Washington, Seattle.

Frydman, J., Nimmesgern, E., Erdjument-Bromage, H., Wall, J. S., Tempst, P., and Hartl, F. U. (1992) Function in protein folding of TRiC, a cytosolic ring complex containing TCP-1 and structurally related subunits, *EMBO J.* **11,** 4767–4778.

Gebauer, M., Melki, R., and Gehring, U. (1998) The chaperone cofactor Hop/p60 interacts with the cytosolic chaperonin-containing TCP-1 and affects its nucleotide exchange and protein folding activities, *J. Biol. Chem.* **273,** 29475–29480.

Geissler, S., Siegers, K., and Schiebel, E. (1998) A novel protein complex promoting formation of functional α- and γ-tubulin. *EMBO J.* **17,** 952–966.

Gilson, P. R., Maier, U. G., and McFadden, G. I. (1997) Size isn't everything: Lessons in genetic miniaturisation from nucleomorphs, *Curr. Opin. Genet. Dev.* **7,** 800–806.

Gutsche, I., Essen, L. O., and Baumeister, W. (1999) Group II chaperonins: new TRiC(k)s and turns of a protein folding machine, *J. Mol. Biol.* **293,** 295–312.

Hartl, F. U. (1996) Molecular chaperones in cellular protein folding, *Nature* **381,** 571–579.

Hashimoto, T., Nakamura, Y., Nakamura, F., Shirakura, T., Adachi, J., Goto, N., Okamoto, K., and Hasegawa, M. (1994) Protein phylogeny gives a robust estimation for early divergences of eukaryotes: Phylogenetic place of a mitochondria-lacking protozoan, *Giardia lamblia*, *Mol. Biol. Evol.* **11,** 65–71.

Horwich, A. L., and Saibil, H. R. (1998) The thermosome: Chaperonin with a built-in lid, *Nat. Struct. Biol.* **5,** 333–336.

Houry, W. A., Frishman, D., Eckerskorn, C., Lottspeich, F., and Hartl, F. U. (1999) Identification of in vivo substrates of the chaperonin GroEL, *Nature* **402,** 147–154.

Hynes, G., Kubota, H., and Willison, K. R. (1995) Antibody characterisation of two distinct conformations of the chaperonin-containing TCP-1 from mouse testis, *FEBS Lett.* **358,** 129–132.

Kawarabayasi, Y., Hino, Y., Horikawa, H., Yamazaki, S., Haikawa, Y., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., Kosugi, H., Hosoyama, A., Fukui, S., Nagai, Y., Nishijima, K., Nakazawa, H., Takamiya, M., Masuda, S., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Kikuchi, H., *et al.* (1999) Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, *Aeropyrum pernix* K1, *DNA Res.* **6,** 83–101, 145–152.

Kawashima, T., Amano, N., Koike, H., Makino, S., Higuchi, S., Kawashima-Ohya, Y., Watanabe, K., Yamazaki, M., Kanehori, K., Kawamoto, T., Nunoshiba, T., Yamamoto, Y., Aramaki, H., Makino, K., and Suzuki, M. (2000) Archaeal adaptation to higher temperatures revealed by genomic sequence of *Thermoplasma volcanium*, *Proc. Natl. Acad. Sci. USA* **97,** 14257–14262.

Keeling, P. J., Deane, J. A., Hink-Schauer, C., Douglas, S. E., Maier, U. G., and McFadden, G. I. (1999) The secondary endosymbiont of the cryptomonad *Guillardia theta* contains alpha-, beta-, and gamma-tubulin genes, *Mol. Biol. Evol.* **16,** 1308–1313.

Keeling, P. J., and Doolittle, W. F. (1996) Alpha-tubulin from early-diverging eukaryotic lineages and the evolution of the tubulin family, *Mol. Biol. Evol.* **13,** 1297–1305.

Kim, S., Willison, K. R., and Horwich, A. L. (1994) Cytosolic chaperonin subunits have a conserved ATPase domain but diverged polypeptide-binding domains, *Trends. Biochem. Sci.* **19,** 543–548.

Klumpp, M., and Baumeister, W. (1998) The thermosome: Archetype of group II chaperonins, *FEBS Lett.* **430,** 73–77.

Klumpp, M., Baumeister, W., and Essen, L. O. (1997) Structure of the substrate binding domain of the thermosome, an archaeal group II chaperonin, *Cell* **91,** 263–270.

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.

Kubota, H., Hynes, G., and Willison, K. (1995a) The chaperonin

containing t-complex polypeptide 1 (TCP-1). Multisubunit machinery assisting in protein folding and assembly in the eukaryotic cytosol, *Eur. J. Biochem.* **230,** 3–16.

Kubota, H., Hynes, G., and Willison, K. (1995b) The eighth Cct gene, Cctq, encoding the theta subunit of the cytosolic chaperonin containing TCP-1, *Gene* **154,** 231–236.

Kubota, H., Hynes, G. M., Kerr, S. M., and Willison, K. R. (1997) Tissue-specific subunit of the mouse cytosolic chaperonin-containing TCP-1, *FEBS Lett.* **402,** 53–56.

Leipe, D. D., Gunderson, J. H., Nerad, T. A., and Sogin, M. L. (1993) Small subunit ribosomal RNA+ of *Hexamita inflata* and the quest for the first branch in the eukaryotic tree, *Mol. Biochem. Parasitol.* **59,** 41–48.

Leroux, M. R., and Hartl, F. U. (2000) Protein folding: Versatility of the cytosolic chaperonin TRiC/CCT, *Curr. Biol.* **10,** R260–R264.

Lewis, V. A., Hynes, G. M., Zheng, D., Saibil, H., and Willison, K. (1992) T-complex polypeptide-1 is a subunit of a heteromeric particle in the eukaryotic cytosol, *Nature* **358,** 249–252.

Lin, P., Cardillo, T. S., Richard, L. M., Segel, G. B., and Sherman, F. (1997) Analysis of mutationally altered forms of the Cct6 subunit of the chaperonin from *Saccharomyces cerevisiae, Genetics* **147,** 1609–1633.

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

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

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

Melki, R., Batelier, G., Soulie, S., and Williams, R. C., Jr. (1997) Cytoplasmic chaperonin containing TCP-1: Structural and functional characterization, *Biochemistry* **36,** 5817–5826.

Miklos, D., Caplan, S., Mertens, D., Hynes, G., Pitluk, Z., Kashi, Y., Harrison-Lavoie, K., Stevenson, S., Brown, C., Barrell, B., Horwich, A. L., and Willison, K. (1994) Primary structure and function of a second essential member of the heterooligomeric TCP1 chaperonin complex of yeast, TCP1β, *Proc. Natl. Acad. Sci. USA* **91,** 2743–2747.

Ng, W. V., Kennedy, S. P., Mahairas, G. G., Berquist, B., Pan, M., Shukla, H. D., Lasky, S. R., Baliga, N. S., Thorsson, V., Sbrogna, J., Swartzell, S., Weir, D., Hall, J., Dahl, T. A., Welti, R., Goo, Y. A., Leithauser, B., Keller, K., Cruz, R., Danson, M. J., Hough, D. W., Maddocks, D. G., Jablonski, P. E., Krebs, M. P., Angevine, C. M., and Dale, H. (2000) Genome sequence of *Halobacterium* species NRC-1, *Proc. Natl. Acad. Sci. USA* **97,** 12176–12181.

Nitsch, M., Klumpp, M., Lupas, A., and Baumeister, W. (1997) The thermosome: Alternating alpha and beta-subunits within the chaperonin of the archaeon *Thermoplasma acidophilum, J. Mol. Biol.* **267,** 142–149.

Phipps, B. M., Hoffmann, A., Stetter, K. O., and Baumeister, W. (1991) A novel ATPase complex selectively accumulated upon heat shock is a major cellular component of thermophilic archaebacteria, *EMBO J.* **10,** 1711–1722.

Phipps, B. M., Typke, D., Hegerl, R., Volker, S., Hoffmann, A., Stetter, K. O., and Baumeister, W. (1993) Structure of a molecular chaperone from a thermophilic archaebacterium, *Nature* **361,** 475–477.

Ranson, N. A., White, H. E., and Saibil, H. R. (1998) Chaperonins, *Biochem. J.* **333,** 233–242.

Ritco-Vonsovici, M., and Willison, K. R. (2000) Defining the eukaryotic cytosolic chaperonin-binding sites in human tubulins, *J. Mol. Biol.* **304,** 81–98.

Roger, A. J. (1999) Reconstructing early events in eukaryotic evolution, *Am. Nat.* **154** (Suppl.), S146–S163.

Rommelaere, H., Van Troys, M., Gao, Y., Melki, R., Cowan, N. J., Vandekerckhove, J., and Ampe, C. (1993) Eukaryotic cytosolic chaperonin contains t-complex polypeptide 1 and seven related subunits, *Proc. Natl. Acad. Sci. USA* **90,** 11975–11979.

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

Siegert, R., Leroux, M. R., Scheufler, C., Hartl, F. U., and Moarefi, I. (2000) Structure of the molecular chaperone prefoldin. Unique interaction of multiple coiled coil tentacles with unfolded proteins, *Cell* **103,** 621–632.

Sigler, P. B., Xu, Z., Rye, H. S., Burston, S. G., Fenton, W. A., and Horwich, A. L. (1998) Structure and function in GroEL-mediated protein folding, *Annu. Rev. Biochem.* **67,** 581–608.

Sogin, M. L., Gunderson, J. H., Elwood, H. J., Alonso, R. A., and Peattie, D. A. (1989) Phylogenetic meaning of the kingdom concept: An unusual ribosomal RNA from *Giardia lamblia, Science* **243,** 75–77.

Stiller, J. W., Duffield, E. C., and Hall, B. D. (1998) Amitochondriate amoebae and the evolution of DNA-dependent RNA polymerase II, *Proc. Natl. Acad. Sci. USA* **95,** 11769–11774.

Stoldt, V., Rademacher, F., Kehren, V., Ernst, J. F., Pearce, D. A., and Sherman, F. (1996) The Cct eukaryotic chaperonin subunits of *Saccharomyces cerevisiae* and other yeasts, *Yeast* **12,** 523–529.

Trent, J. D., Nimmesgern, E., Wall, J. S., Hartl, F. U., and Horwich, A. L. (1991) A molecular chaperone from a thermophilic archaebacterium is related to the eukaryotic protein t-complex polypeptide-1, *Nature* **354,** 490–493.

Ursic, D., Sedbrook, J. C., Himmel, K. L., and Culbertson, M. R. (1994) The essential yeast Tcp1 protein affects actin and microtubules, *Mol. Biol. Cell* **5,** 1065–1080.

Vainberg, I. E., Lewis, S. A., Rommelaere, H., Ampe, C., Vandekerckhove, J., Klein, H. L., and Cowan, N. J. (1998) Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin, *Cell* **93,** 863–873.

Waldmann, T., Lupas, A., Kellermann, J., Peters, J., and Baumeister, W. (1995a) Primary structure of the thermosome from *Thermoplasma acidophilum, Biol. Chem. Hoppe Seyler* **376,** 119–126.

Waldmann, T., Nimmesgern, E., Nitsch, M., Peters, J., Pfeifer, G., Muller, S., Kellermann, J., Engel, A., Hartl, F. U., and Baumeister, W. (1995b) The thermosome of *Thermoplasma acidophilum* and its relationship to the eukaryotic chaperonin TRiC, *Eur. J. Biochem.* **227,** 848–856.

Willison, K. R. (1999) Composition and function of the eukaryotic cytosolic chaperonin-containing TCP-1, *in* Bukau, B. (Ed.), Molecular Chaperones and Folding Catalysts, pp. 555–571, Harwood, Amsterdam.

Willison, K. R., Dudley, K., and Potter, J. (1986) Molecular cloning and sequence analysis of a haploid expressed gene encoding t complex polypeptide 1, *Cell* **44,** 727–738.

Willison, K. R., and Horwich, A. L. (1996) Structure and function of chaperonins in archaebacteria and eukaryotic cytosol, *in* Ellis, R. J. (Ed.), The Chaperonins, pp. 107–136, Academic Press, San Diego.

Willison, K. R., and Kubota, H. (1994) The structure, function, and genetics of the chaperonin containing TCP-1 (CCT) in eukaryotic cytosol, *in* Morimoto, R. I., Tissieres, A., and Georgopoulos, C. (Eds.), The Biology of Heat Shock Proteins and Molecular Chaperones, Vol. 26, pp. 299–312, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Won, K. A., Schumacher, R. J., Farr, G. W., Horwich, A. L., and Reed, S. I. (1998) Maturation of human cyclin E requires the function of eukaryotic chaperonin CCT, *Mol. Cell. Biol.* **18,** 7584–7589.

Zauner, S., Fraunholz, M., Wastl, J., Penny, S., Beaton, M., Cavalier-Smith, T., Maier, U. G., and Douglas, S. (2000) Chloroplast protein and centrosomal genes, a tRNA intron, and odd telomeres in an unusually compact eukaryotic genome, the cryptomonad nucleomorph, *Proc. Natl. Acad. Sci. USA* **97,** 200–205.