FEBS Letters 347 (1994) 105–111

FEBS 14136

Hypothesis

Functional role of a consensus peptide which is common to α-, β-, and γ-tubulin, to actin and centractin, to phytochrome A, and to the TCP1α chaperonin protein

Roy G. Burns*, Christopher D. Surridge

Biophysics Section, Blackett Laboratory, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2BZ, UK

Received 9 May 1994

Abstract

The TRiC (TCP1 Ring Complex) chaperonin complex participates in the functional folding of actin, centractin, α-, β-, γ-tubulin, and phytochrome. Each of the cytoskeletal proteins contain a peptide, RK(A,C,T)F/KRAF, located towards the C-terminus, which is homologous to a TCP1α peptide, while the equivalent phytochrome peptide (RLKAF in certain isoforms) is very similar to the KLRAF peptide of TCP1α. We propose that this TCP1α peptide binds to the nascent polypeptides as they emerge from the ribosome, that this binding restricts the folding pathway, and that the TCP1α peptide is subsequently displaced by the synthesis of the consensus peptide. This hypothesis is strongly supported by the crystallographic structure of actin.

Key words: α-Tubulin; β-Tubulin; γ-Tubulin; Actin; Centractin; Phytochrome A; TCP1α; Consensus peptide; Protein folding

1. Introduction

The α-, β-, and γ-tubulins are each highly conserved, and the three sub-families share 30–35% sequence identity [1]. Similarly, actin and centractin are about 50% identical [2,3]. By contrast, the homologies between members of the tubulin and actin families are restricted to a small number of peptides, most of which (such as a glycine-rich peptide) are common to many nucleotide-binding proteins. There is, however, increasing evidence that the synthesis of natively-folded α-, β-, and γ-tubulin, actin and centractin, and a number of other proteins (including phytochrome [4]) each require TCP1, a constituent protein of a cytoplasmic chaperonin, while a TCP1-like protein has been implicated in the assembly of the hepatitis B virus capsid protein [5].

Synthesis of α- and β-tubulins in an *E. coli* lysate yields assembly-incompetent protein, yet the protein produced in a reticulocyte expression system can be co-assembled with carrier tubulin into bone fida microtubules [6,7]. Reticulocyte component(s) therefore appear to facilitate the functional folding of the nascent tubulin polypeptides. Indeed, the reticulocyte-expressed β-tubulin can be fractionated by either ion exchange or gel filtration chromatography into three components, one of which has an apparent molecular weight in excess of 900 kDa, and monomeric tubulin subunits can be chased from this aggregate component by the inclusion of Mg^{2+}·ATP [7,8]. Similarly, bacterially expressed actin is also assembly-incompetent, yet the functional folding, assayed by assembly, DNase I-binding, or electrophoretic mobility, can be restored by denaturation followed by incubation with Mg^{2+}·ATP and a 19S polymeric aggregate present in reticulocyte lysates [9]. Similar studies have established that the reticulocyte component also facilitates the correct folding of centractin and γ-tubulin [10], and of phytochrome [4]. Furthermore, immature assemblies of the hepatitis B core protein, on expression in a wheat germ translation system, crossreact with a polyclonal antibody raised against a TCP1 peptide, although a TCP1-like protein was apparently not bound to either the expressed monomer or the dimer [5].

The functional component has been purified from reticulocyte lysates, from bovine testis, and tissue culture cells [9,11–13], and in each case it is a toroidal structure consisting of stoichiometric amounts of a number of related proteins. As one of these proteins crossreacts with anti-TCP1 antibodies, the purified ring-like structure has been termed TRiC (TCP1 Ring Complex, [11]). Both the bovine testis TRiC and the purified reticulocyte 19S component (plus additional co-factors) effect the Mg^{2+}·ATP-dependent refolding of guanidinium chloride denatured α- and β-tubulin [11,13,14]. Furthermore, the 900 kDa TRiC complex has been implicated in the folding of actin and tubulin in vivo [15]. In summary, there is persuasive evidence that TRiC functions as a chaperonin in effecting the functional folding of actin, centractin, the tubulins, and phytochrome, and that the wheat germ equivalent may be involved in the assembly of capsids from the hepatitis B core protein. While these

*Corresponding author. Fax: (44) (71) 589 0191.
E-mail: r.burns@ic.ac.uk

SSDI 0014-5793(94)00522-W

activities might relate to any of the constituent TRiC proteins, the antibody studies and a mutation of the yeast TCP1α protein, which affects microtubule-dependent processes [16], strongly suggest the direct involvement of TCP1.

2. Comparison of primary sequences

The 900 kDa TRiC complex contains about nine proteins, six or seven of which are highly homologous [12,17]. Seven mouse TCP1 homologues, termed TCP1α–TCP1η or CCTα–CCTη (for Chaperon Containing TCP1) have been sequenced ([17]; Table 1). Homologues to the mouse TCP1α have been sequenced from rat, CHO, human, *Drosophila*, *S. cerevisiae*, and *Arabdopsis* (see Table 1). More divergent sequences, but which still resemble TCP1α, have been identified in *Avena sativa* and in *S. cerevisiae* (Table 1).

One peptide of the TCP1α family (mouse TCP1α:487–496) is strikingly similar to an α-tubulin peptide (α-Tu:396–406, Table 2). The two peptides are identical at eight of the eleven positions, provided a single gap is inserted into the α-tubulin and two into the TCP1α sequence, and nine out of eleven residues if a conservative (methionine/valine) amino acid substitution is also accepted. The α-tubulin peptide is extremely highly conserved (Fig. 1A), while the TCP1α peptide Leu^{488}–Arg^{493} is conserved in each of the available sequences, making it one of only seven totally conserved peptides of six or more residues.

Inspection of the β- and γ-tubulin peptides equivalent to α-Tu:396–406 show that they are also highly conserved and show homology with the TCP1α peptide (Fig. 1B and C; Table 2). Insertion of two gaps into both the

Table 1
Comparison of the available TCP1 sequences, showing the peptide (TCP1α:486–498 for the mouse sequence) postulated to interact with th nascent polypeptide.

Mouse TCP1α	T D L V A K L R A F H N E	[12, 17-19]
Rat TCP1	T D L V A K L R A F H N E	[24]
CHO TCP1	T D L V A K L R A F H N E	[21]
Human TCP1	T D L V A K L R A F H N E	[20]
Drosophila TCP1	T D L V A K L R S Y H N S	[22]
S. cerevisiae TCP1α	S E L V A K L R S Y H A A	[16, 25]
Arabdopsis TCP1	T E L V A K L R A Y H H T	[23]
Mouse TCP1β	A D L V A Q L R A A H S E	[17]
Mouse TCP1γ	I R L L T S L R A K H T Q	[17, 26]
Mouse TCP1δ	I S T V T E L R N R H A Q	[17]
Mouse TCP1ε	I Q T M T E V R A R Q V K	[17]
Mouse TCP1ζ	Q E T L V K V Q A E H S A	[17]
Avena sativa TCP1	T L T V V K S Q H V K E N	[27]
Mouse TCP1η	T N I L N K L R A R H A Q	[17]
S. cerevisiae TCP1β	S E L V S K L R S S I Y N	[25]

The upper half shows the mouse TCP1α and homologues from rat, CHO cells, human, *Drosophila*, *S. cerevisiae*, and *Arabdopsis*, while the lower half shows the more divergent mouse TCP1β–mouse TCP1η isoforms and the additional sequences from *Avena sativa* and *S. cerevisiae*. Residues which are identical to those of mouse TCP1α are boxed.

Table 2
Paired comparison of the consensus peptides of α-, β-, and γ-tubulin (α-Tu:394–410, β-Tu:384–400, and γ-Tu:394–410), actin (Ac:364–375), centractin (CAc:365–376), and phytochrome A (P:897–905) with TCP1:484–500 of murine TCP1α.

α-Tubulin	K F - D L M Y A K - R . A F V H W Y V G
TCP1α	D S T D L V - A K L R . A F - H N E A Q
β-Tubulin	Q F T - A M F R R - K . A F L H W Y T G
TCP1α	D S T D L V - A K L R . A F - H N E A Q
γ-Tubulin	Q Y - D K L - R K - R n A F L E Q F R K
TCP1α	D S T D - L V A K L R . A F - H N E A Q
Actin	E S G P S I V H R - K . C F
TCP1α	D S T D - L V A K L R . A F - H N E A Q
Centractin	E D G A R S I H R - K . T F
TCP1α	D S T D - L V A K L R . A F - H N E A Q
Phytochrome A	A S E Q T S L K R L K . A F S Y M R H A
TCP1α	D S T D - L V A K L R . A F - H N E A Q

The discussed homologies are restricted to α-Tu:396–406, β-Tu:386–396, γ-Tu:396–406, Ac:369–375, CAc:371–376, P:895–899 and TCP1α:487–496; the flanking residues are shown to illustrate the lack of significant homology elsewhere within the sequences. The heavy boxes identify residues which are identical in each peptide pair, while the light boxes highlight conservative amino acid substitutions. Two types of insertion are shown: (·) has been introduced into α- and β-tubulin, actin, centractin, phytochrome A and TCP1 to align these sequences with γ-tubulin. (–) shows insertions introduced to maximize the homology between each peptide pair. The small number of γ-tubulin sequences leads to some ambiguity in identifying the consensus residue, as indicated by the lower case symbols.

β-tubulin (β-Tu:386–396) and the TCP1α peptides results in identity at four out of eleven positions, and acceptance of conservative amino acid substitutions (lysine/arginine, leucine/alanine, valine/methionine) increases this homology to eight out of eleven residues. The γ-tubulin sequence (γ-Tu:396–406) reveals additional degeneracy: it is necessary to introduce three insertions into both the γ-tubulin and the TCP1α sequences (Table 2). Following this alignment, the γ-tubulin and TCP1α peptides are identical at six out of nine positions. The degeneracy is even more pronounced in *Chlamydomonas rheinhardtii*, in that the γ-sequence contains a unique, four residue insertion (PKQE at γ-Tu:405/406; C Silflow, personal communication).

The common feature of the homologies between TCP1α and the three tubulin sub-families is the motif KRAF or RKAF, supplemented by various N-terminal residues and (in α- and β-tubulin) by a highly conserved, C-terminal histidine. The homology is particularly striking between TCP1α and α-tubulin, less evident in β-tubulin, and least in γ-tubulin. Indeed, the γ-tubulin:TCP1α homology largely depends upon the alignment of the three tubulin sub-families [1]. There is consequently a variable level of degeneracy between TCP1α and the individual members of the tubulin family, although the KRAF or RKAF motif remains a common feature.

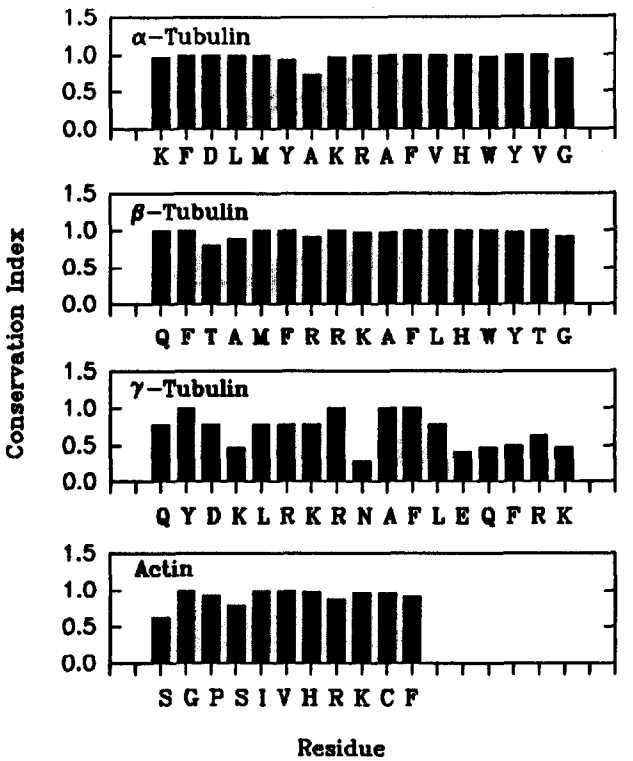

Fig. 1. Comparison of the Conservation Index of α-Tubulin (α-Tu:394–410, 55 sequences), β-Tubulin (β-Tu:384–400, 70 sequences), γ-Tubulin (γ-Tu:394–410, 8 sequences), and actin (Ac: 365–375, 99 sequences). The Conservation Index at each residue position is calculated as $(n_1^2 + n_2^2 + ...)/N^2$, where N is the total number of known sequences of each individual protein and n_1, n_2, ... are the number of these sequences with a specific amino acid at the specified position [33]. This Index, which for an infinite sized sample varies between 0.05 and 1.0, differs from the more conventional calculation of determining the fraction of the sequences with the consensus residue. In particular, it minimizes the contribution of single residue differences, which may be due to a sequencing error, and increases the significance of a specific amino acid substitution when it occurs in several sequences. For instance, the low Conservation Index of α-Tu:A^{400} (0.743) is due to an alanine/serine substitution in 6 out of 55 sequences, plus single lysine and glycine substitutions. The low Conservation Indices of γ-tubulin is due to the availability of only 8 phylogenetically diverse sequences.

The actin sequence lacks this motif. It does however terminate in the highly conserved RKCF C-terminal peptide (Table 2, Fig. 1D). This sufficiently resembles, with its conservative alanine/cysteine substitution, the tubulin peptides to suggest that it is homologous with the TCP1α KLRAF peptide, particularly since it requires the common insertion corresponding to the TCP1α:L^{492}. This homology is increased by considering the three amino acids N-terminal to this motif to six out of seven residues, provided that an additional conservative (leucine/isoleucine) substitution, is accepted (Fig. 1D). Furthermore, as the C-terminal peptide of actin adopts an α-helical conformation, the Ac:H^{371} residue may be quasi-equivalent to the highly conserved histidine residues of α- and β-tubulin (α-Tu:H^{406} and β-Tu:H^{396}), such that it can also contribute to the homology with TCP1α. A similar argument applies to centractin (Table 2), in that its C-terminal peptide is identical to that of actin, except for the CAc:T^{375}/Ac:A^{374} substitution, has a similarly conserved histidine, and in common with the actin and tubulin alignments it is necessary to introduce an insertion corresponding to TCP1α:L^{492}. The presence of a modified motif in phytochrome A will be considered later.

In summary, α-tubulin, β-tubulin, γ-tubulin, actin, and centractin each contain a highly conserved sequence which is homologous to a TCP1α peptide. The consensus peptide, of which the key residues are RK(A,C,T)F or KRAF, will be referred to as the C-Peptide (for consensus/C-terminal peptide), while the quasi-equivalent TCP1α peptide will be referred to as the TCP1α homologue. Using the amino acid compositions of human α-, β-, and γ-tubulin, actin, centractin and TCP1α to calculate the probabilities of these proteins containing the C-Peptide or the TCP1α homologue yield values of 3.15, 2.44, 0.48, 0.54, 3.22, and 4.06×10^{-3}, respectively. Calculation of the probability that the tubulin, actin, and TCP1α families (as represented by the human α-tubulin, actin, and TCP1α proteins) would each contain a related peptide, defined as KR(A/T/C)F or RKAF, yields a value of 184×10^{-9}. This strongly suggests that the common presence of the related peptides is not adventitious.

Inspection of the GenBank protein database with the FindPattern algorithm for proteins containing the RK(A,C,T)F or KRAF peptides yielded 535 sequences (out of 31,808), of which 212 corresponded to either actin or the tubulins. The 323 sequences which encoded neither actin nor tubulin represented 207 different proteins or hypothetical proteins. This compares with the 222 predicted sequences, based upon the frequency of occurrence of the individual amino acids [28] and the total length of the searched sequence (10,792,076 residues after omission of the over-represented actins and tubulins). The motifs are therefore slightly more common in the current database than the predicted frequencies (323 vs. 222), despite the removal of the over-represented, and highly conserved, actins and tubulins.

3. TCP1α and protein folding

Chaperonins, including TRiC, probably assist in protein folding by binding to particular, and in general hydrophobic, peptides and so prevent certain kinetically favourable peptide:peptide interactions which would result in the nascent protein adopting an anomalous tertiary structure. A similar mechanism presumably applies to the chaperonin-mediated assembly of multisubunit complexes. The limitation on the folding pathway is transitory, and is relieved by further extension of the

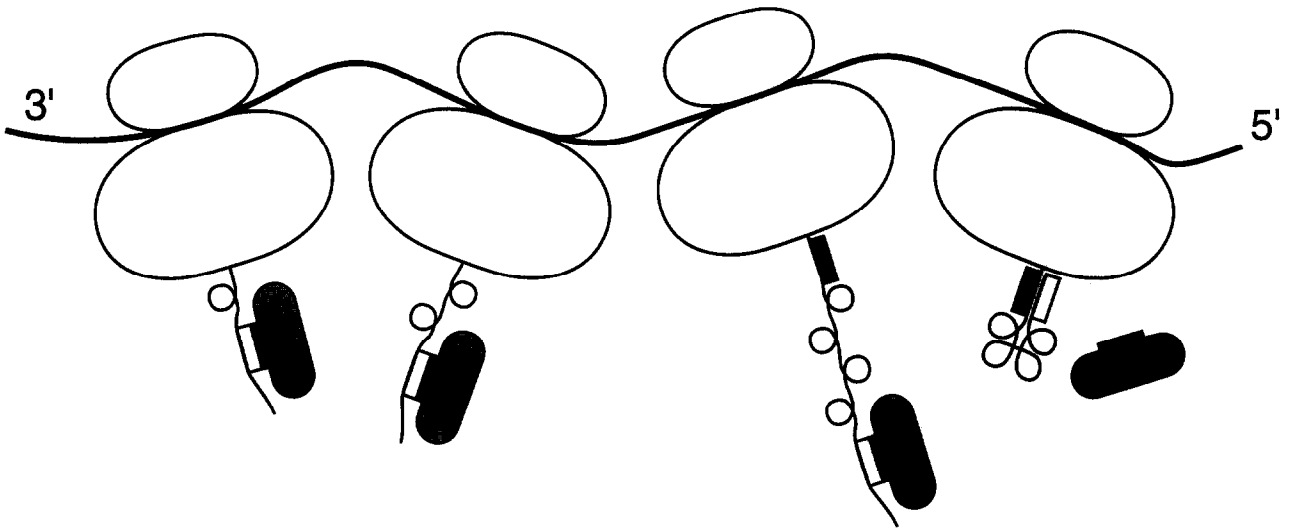

Fig. 2. Schematic of the proposed model, showing an elongating nascent polypeptide chain emerging from four sequential ribosomes. A peptide, located close to the N-terminus of the nascent polypeptide, is identified as the A-Site, and is shown as an open box. This is recognised by the TCP1α peptide (shown as a filled box on the shaded chaperonin complex), which prevents the anomolous folding of the emerging polypeptide (shown in the left-hand two of the ribosomes). As translation of the nascent polypeptide continues, a peptide is synthesized which is termed the C-Peptide and which is shown as a filled box (see third ribosome). The C-Peptide is homologous to the TCP1α peptide, and it displaces the TCP1α from the A-Site, so allowing the emerging polypeptide to adopt its native conformation (shown as a four-leaf clover leaf, right-hand ribosome). The released chaperonin complex is then available to facilitate the folding of another nascent polypeptide chain.

nascent polypeptide from the ribosome, coupled with chaperonin-mediated ATP hydrolysis, which presumably induces a conformational change and the consequential release of the chaperonin from the correctly folded protein. The need to restrict the folding pathway may apply in particular to those proteins which have a complex topology reflecting the functional demands of the folded protein. Consequently, the involvement of the TRiC complex in the folding of tubulin and actin may be determined, in part, by the needs to self-assemble into microtubules and actin filaments, and to undergo a conformational change in response to the assembly-dependent hydrolysis GTP or ATP. Similarly, the role of the TRiC complex in the folding of phytochrome may reflect the structural demands of the photoactivated conformational change.

The identification of the C-Peptide of α-, β-, and γ-tubulin and of actin and centractin suggests a molecular mechanism for releasing of TCP1α from the newly folded polypeptide. We propose that the TCP1α homologue binds to one or more peptides, termed the A-Site, which lie nearer the N-terminus of the nascent polypeptide than the C-Peptide, and that this prevents the anomalous folding of the protein as it emerges from the ribosome. We further propose that when the C-Peptide is subsequently synthesized that the constrained folding pathway positions this peptide such that it competes with the TCP1α homologue, and that this results in the displacement of TCP1α from the correctly folded polypeptide. This model is shown in Fig. 2.

A number of factors may aid the competition between the C-Peptide of the nascent chain and the TCP1α peptide, including a higher affinity of the native C-Peptide for the A-Site than the TCP1α homologue. Significantly, the identified C-Peptides of actin, centractin, α-, β-, and γ-tubulin resemble but are <u>not</u> identical to the TCP1α homologue peptide in that they each lack the residue equivalent to TCP1α:L^{492}.

This model is consistent with the tertiary structure of actin [29–31]. The N-terminus lies within subdomain-1, but only two anti-parallel β-pleated sheets are completed before the formation of subdomain-2. The polypeptide chain then re-enters subdomain-1 and forms a sheet/helix/sheet motif (Ac:105–109, Ac:115–128, Ac:133–138), before exiting to form subdomains-3 and -4. The polypeptide chain then re-enters sub-domain-1 and forms the four C-terminal α-helices (Ac:337–350, Ac:355–357, Ac:361–367, and Ac:370–375). The complexity of this topology highlights why the TRiC, or another, chaperonin may be required in order to restrict the folding pathway.

Inspection of the actin:gelsolin tertiary structure [30] shows that the RKCF-containing Ac:370–375 α-helix abuts a surface formed from residues of the Ac:115–128 α-helix and the Ac:133–138 β-pleated sheet. There are four close associations, defined as being of less that 4Å, between the main chain carboxyls of Ac:K^{373}, C^{374}, F^{375} and Ac:R^{116} or Ac:Y^{133}, and a further eight close associations involving the sidechains (Ac:C^{374} with Ac:R^{116}, Ac:Y^{133} and Ac:V^{134}; Ac:F^{375} with Ac:Y^{169} of subdomain-3; and Ac:H^{371}, see above, with Ac:K^{113}, R^{116}, E^{117}). These close contacts strongly suggest that A-Site is formed from the residues Ac:115–117 and Ac:133–134. Each of the identified A-Site residues

(Ac:K^{113}, R^{116}, E^{117}, Y^{133}, and V^{134}) are highly conserved (Conservation Indices: 0.98, 1.00, 1.00. 0.98, and 0.98, respectively).

Chaperonins are generally considered to protect hydrophobic peptides during protein folding. While the A-Site residues Ac:K^{113}, R^{116} and E^{117} are highly charged, Ac:Y^{133} and Ac:V^{134} lie within an extremely hydrophobic peptide which probably requires protection from the solvent until becoming masked by the subsequent synthesis of the C-terminal C-Peptide. We propose that this protection is effected by the interaction with the TCP1α homologue peptide.

Numerous studies have demonstrated the high chemical reactivity of Ac:C^{374} (see [32]) yet the crystallographic structure shows that this residue is buried and is not exposed to the solvent. The C-Peptide may therefore exhibit some kinetic flexibility, and this may be required to facilitate the release of the TCP1α homologue peptide. The mechanism of this release may also account for the high conservation of Ac:R^{372} despite the structural evidence that it is exposed to the surface and is not in close contact with any other residue.

The sequence homology of actin and centractin strongly suggests that the two proteins have closely related tertiary structures. Consequently, the proposed model probably also applies to the folding of centractin. The CAc:T^{375}/Ac:C^{374} substitution is of particular interest in view of the number of close associations between Ac:C^{374} and the A-Site residues. Significantly, the centractin A-Site contains two conservative substitutions (CAc:F^{133}/Ac:Y^{133}, CAc:I^{134}/Ac:V^{134}), and additional flanking residue differences, which may compensate for the C-Peptide difference. The C-Peptides of the tubulins, which each lie towards the C-terminus of the individual sequences, are similarly proposed to compete for the TCP1α homologue peptide during the folding of the nascent polypeptides, and to interact in the native proteins with peptides equivalent to those forming the A-site of actin. The absence of the tubulin crystallographic structure precludes the direct identification of these additional peptide(s).

The homology between TCP1α and the other proteins focuses on the identified RK(A,C,T)F/KRAF motif (Table 2). The protein-specific variations would indicate that the selective pressures on the individual C-Peptide residues may differ between the three tubulin classes and between actin and centractin, and that the sequences are not determined exclusively by the functional requirements of the interaction with the TCP1α homologue peptide. Indeed, the conservation of the Ac:R^{372}, Ac:K^{373} and Ac:F^{375} (Fig. 1) may be defined in part by their role in specifying the profilin-binding site [30]. Such 'dual selective pressures' may also account for the reported requirement for two unidentified co-factors in the folding of α- and β-tubulin but not of either γ-tubulin or centractin [10,13,14], and undoubtedly contribute to the remarkable conservation of the individual tubulin and actin sequences [32,33].

4. The specificity of TCP1

Only the vertebrate TCP1α sequences contain the KLRAF peptide: *S. cerevisiae*, *Drosophila*, and *Arabdopsis* TCP1α sequences contain the variants KLRSY and KLRAY (Table 1). Assuming that the interaction between the TCP1α homologue peptide and the nascent actin polypeptide mimics that between the C-Peptide and the A-Site of actin, then TCP1α:A^{494} would be anticipated to make sidechain interactions with Ac:R^{116}, Ac:Y^{133} and Ac:V^{134}. The TCP1α:S^{494} substitution of *S. cerevisiae* and *Drosophila* is not compensated by substitution(s) within the actin A-Site. Consequently, subtle differences elsewhere within the non-vertebrate TCP1α sequences may, as discussed for the Ac:C^{374}/CAc:T^{374} substitution, confer a specific advantage to this TCP1α alanine/serine substitution. The *S. cerevisiae*, *Drosophila*, and *Arabdopsis* TCP1α sequences each contain the TCP1α:F^{495} to TCP1α:Y^{495} substitution, which the actin crystallographic structure would imply interacts with Ac:Y^{169} of subdomain 3 until the eventual synthesis of the actin C-Peptide. Significantly, the yeast and one of the Drosophila actin isoforms both contain the complimentary substitution (Ac:Y^{169} to Ac:F^{169}), indicating the apparent conservation of the atomic interactions between the TCP1α homologue and the nascent actin polypeptide.

In vitro and in vivo evidence both highlight the importance of TRiC in the functional folding of proteins of the actin and tubulin families, while the effects of a *Saccharomyces* TCP1α mutation on microtubule-dependent processes indicate that the TCP1α gene product plays an essential role [16].

Inspection of the protein database showed that slightly more sequences contain the TCP1α C-Peptide motif than would be statistically predicted. The proposed role of the C-Peptide motif places it towards the C-terminus of the peptide sequence requiring an interaction with TCP1α for its functional folding (see Fig. 2). The consensus peptide lies at a mean position of 0.506 ± 0.301 from the N-terminus in the 207 identified proteins. The presence of this peptide may not therefore be diagnostic of an involvement of the TCP1α complex in the folding of the specific nascent polypeptide. Alternatively, the TCP1α complex may contribute to the folding of specific structural domains, with the C-Peptide lying towards the C-terminus of the sequence specifying this domain.

The heterogeneity within the C-Peptide residues of α-, β-, and γ-tubulin, and actin and centractin, coupled with the subtle differences between the TCP1α homologous peptide from different organisms, shows that the proposed mechanism by which the TCP1α homologous pep-

tide transiently protects the A-Site is not absolutely conserved. This is amply illustrated by phytochrome, since this is a protein which requires the TRiC complex for its folding [4] but is not identified by the motif screening. Significantly, the sequences of three of the Class A phytochromes contain the peptide RLKAF (Table 2, residues P:895–899). This is strikingly similar to the KLRAF sequence of the TCP1α homologue peptide. The calculated probability of occurrence, using the Avena sativa phytochrome A, Type 4 sequence, is 1.13×10^{-3}. While α-, β-, and γ-tubulins, actin, and centractin each lack the leucine residue corresponding to TCP1α:L^{492}, the phytochrome A sequences lack the C-flanking histidine residue which possibly contributes to the homology between TCP1α and several of the cytoskeletal proteins. The probability that phytochrome A would have the RLKAF and that TCP1α would have the KLRAF peptide is 0.41×10^{-6}. Other Class A phytochromes have related sequences (eg RLKAL and RLKVL in *Arabdopsis thaliana* and *Pisum sativum*) yielding, for the five available phytochrome A sequences, a Conservation Index of 0.84. The Class B and Class C phytochromes are, in common with the observed differences between α-, β-, and γ-tubulins, more divergent. Significantly, the phytochrome A RLK(A,V)(F,L) peptides lie close to the C-terminus (at 0.79), i.e. their location is fully consistent with the proposed model (Fig. 2) of how the TRiC complex effects the functional folding of a nascent polypeptide. By contrast, the hepatitis core protein lacks any peptide which is clearly homologous to this C-Peptide motif. Significantly, the TCP1-like protein appears to be involved in the assembly of the core proteins into a capsid rather than in the folding of the monomeric core subunit [5], such that the C-Peptide motif may be defined by an interface between two monomeric core subunits.

The identification of the C-Peptide in the tubulins, actins and phytochrome A relies upon the homologies with TCP1α. Six other members of the TCP1 (or CCT) family have been identified in mouse (Table 1, [17]), and additional homologues in *Avena sativa* [27] and *S. cerevisiae* [25]. These isoforms each differ from TCP1α within the peptide homologous to the identified C-Peptide (Table 1), which indicates that they are not involved in the folding of the tubulins, actins, or phytochrome A. As each TRiC complex contains multiple subunits, the variety of TCP1 isoforms may result in the formation of a family of TRiC complexes, which range from ones containing a single isoform to ones formed from a mixture of isoforms. Indeed, the complexes, purified from rat and guinea pig brain and testis have differing compositions [34]. The differences in the homologous peptides of the difference TCP1 isoforms (Table 1) may define which particular eucaryotic proteins require which specific TRiC complex for their functional folding. Indeed, TCP1 has been implicated in the folding of firefly luciferase [11], although the efficacy was not been reported: the firefly luciferase sequences lack a peptide resembling the TCP1α homologue. The eucaryotic luciferase sequences do however contain the peptide KRLR, which closely resembles the homologous peptide of mouse TCP1η (KLRAR, Table 1). The KRLR sequence is located close to the C-terminus (at 0.93) and, in common with several other identified C-Peptides, lacks the residue equivalent to the mouse TCP1α:L^{492}. The specific sequences of the individual members of the TCP1 family corresponding to the TCP1α homologous peptide may therefore specify which TRiC complex binds to the nascent polypeptide chain. This may also apply to other chaperonins, including the chaperonin-like protein of the archaebacterium *Sulfobolus* TF55, which shares a 36% identity with TCP1α but which contains the QLRSL sequence in place of the TCP1α KLRAF motif.

Acknowledgements: We wish to thank Dr Paul McLaughlin for generously releasing to us the actin:gelsolin coordinates, and the Medical Research Council for financial support.

References

[1] Burns, R.G. (1991) Cell Motil. Cytoskel. 20, 181–189.
[2] Lees-Miller, J.P., Helfman, D.M. and Schroer, T.A. (1992) Nature 359, 244–246.
[3] Clark, S.W. and Meyer, D.I. (1992) Nature 359, 246–250.
[4] Mummert, E., Grimm, R., Speth, V., Eckerskorn, C., Schiltz, E., Gatenby, A.A. and Schafer, E. (1993) Nature 363, 644–648.
[5] Lingappa, J.R., Martin, R.L., Wong, M.L., Ganem, D., Welch, W.J. and Lingappa, V.R. (1994) J. Cell Biol. 125, 99–111.
[6] Yaffe, M.B., Levison, B.S., Szasz, J. and Sternlicht, H. (1988) Biochemistry 27, 1869–1880.
[7] Yaffe, M.B., Farr, G.W. and Sternlicht, H. (1988) J. Biol. Chem. 263, 16023–16031.
[8] Yaffe, M.B., Farr, G.W., Miklos, D., Horwich, A.L., Sternlicht, M.L. and Sternlicht, H. (1992) Nature 358, 245–248.
[9] Gao, Y., Thomas, J.O., Chow, R.L., Lee, G-H. and Cowan, N.J. (1992) Cell 69, 1043–1050.
[10] Melki, R., Vainberg, I.E., Chow, R.L. and Cowan, N.J. (1993) J. Cell Biol. 122, 1301–1310.
[11] Frydman, J., Nimmesgern, E., Erdjument-Bromage, H., Wall, J.S., Tempst, P. and Hartl, F.U. (1992) EMBO J. 11, 4767–4778.
[12] Lewis, V.A., Hynes, G.M., Zheng, D., Saibal, H. and Willison, K. (1992) Nature 358, 249–252.
[13] Rommelaere, H., van Troys, M., Gao, Y., Melki, R., Cowan, N.J., Vandekerckhove, J. and Ampe, C. (1993) Proc. Natl. Acad. Sci. USA 90, 11975–11979.
[14] Gao, Y., Vainberg, I.E., Chow, R.L. and Cowan, N.J. (1993) Mol. Cell. Biol. 13, 2478–2485.
[15] Sternlicht, H., Farr, G.W., Sternlicht, M., Driscoll, J.K., Willison, K. and Yaffe, M.B. (1993) Proc. Natl. Acad. Sci. USA 90, 9422–9426.
[16] Ursic, D. and Culbertson, M.R. (1991) Mol. Cell. Biol. 11, 2629–2640.
[17] Kubota, H., Hynes, G., Carne, A., Ashworth, A. and Willison, K. (1994) Curr. Biol. 4, 89–99.
[18] Willison, K.R., Dudley, K. and Potter, J. (1986) Cell 44, 727–738.
[19] Kubota, H., Morita, T., Nagata, T., Takemoto, Y., Nozaki, M., Gachelin, G. and Matsushiro, A. (1991) Gene 105, 269–273.
[20] Kirchhoff, C. and Willison, K. (1990) Nucleic Acids Res. 18, 4247.

[21] Ahmad, S. and Gupta, R.S. (1990) Biochim. Biophys. Acta 1087, 253–255.
[22] Ursic, D. and Ganetzky, B. (1988) Gene 68, 267–274.
[23] Mori, M., Murata, K., Kubota, H., Yamamoto, A., Matsushiro, A. and Morita, K. (1992) Gene 122, 381–382.
[24] Morita, T., Kubota, H., Gachelin, G., Nozaki, M. and Matsushiro, A. (1991) Biochim. Biophys. Acta 1129, 96–99.
[25] Miklos, D., Caplan, S., Mertens, D., Hynes, G., Pitluk, Z., Kashi, Y., Harrison- Lvoie, K., Stevenson, S., Brown, C., Barrell, B., Horwich, A.L. and Willison, K. (1994) Proc. Natl. Acad. Sci. USA 91, 2743–2747.
[26] Joly, E.C., Sevigny, G., Todorov, I.T. and Bibor-Hardy, V. (1994) Biochim. Biophys. Acta 1217, 224–226.
[27] Ehmann, B., Krenz, M., Mummert, E. and Schafer, E. (1993) FEBS Lett. 336, 313–316.
[28] King, J.L. and Jukes, T.H. (1969) Science 164, 788–798.
[29] Kabsch, W., Mannherz, H.G., Suck, D., Pai, E.F. and Holmes, K.C. (1990) Nature 347, 37–44.
[30] McLaughlin, P.J., Gooch, J.T., Mannhertz, H.-G. and Weeds, A.G. (1993) Nature 364, 685–692.
[31] Schutt, C., Myslik, J.C., Rozycki, M.D., Goonesekere, N.C.W. and Lindberg, U. (1993) Nature 363, 810–816.
[32] Sheterline, P. and Sparrow, J.C. (1994) Actin. Protein Profile 1, in press.
[33] Burns, R.G. and Surridge, C.D. (1993) in: Microtubules (J. Hyams and C. Lloyd, Eds.) Wiley-Liss, New York, pp. 3–31.
[34] Roobol, A. and Carden, M.J. (1993) J. Neurochem. 60, 2327–2330.