Correspondence

New tubulins in protozoal parasites

Susan Vaughan, Teresa Attwood, Miguel Navarro, Valerie Scott, Paul McKean and Keith Gull

More than 20 years ago, biochemical analysis of the eukaryotic cell cytoskeleton revealed the major component proteins. The heterodimeric (α/β) protein tubulin was defined as the building block of microtubules, assembled in a polar manner into specifically arranged protofilaments in the microtubule wall [1].

The next two members of the tubulin protein superfamily were both discovered by genetic means — γ tubulin in *Aspergillus* [2] and δ tubulin in *Chlamydomonas* [3]. The γ tubulin is essential for microtubule function and is located in centroso mes and other microtubule-organising centres [4]. The δ tubulin is encoded by the *UNI3* gene in *Chlamydomonas* and a *uni3-1* mutation resulted in flagellar basal bodies that possess doublet rather than triplet microtubules [3]. These four members of the tubulin superfamily can be characterised by their distinct intracellular locations and expression patterns, which are reflected in unique sequence characteristics.

The large number of tubulin sequences available in current databases, coupled with the considerable divergence of those sequences, complicates the task of reliable identification and characterisation of tubulin family members. During the *Saccharomyces cerevisiae* genome project, sequencing revealed the presence of a tubulin gene that was only around 30% identical to the yeast α and β tubulins. This Tub4 protein was conjectured to be a novel tubulin rather than an α, β or γ tubulin [5]. However, subsequent analysis of the completed *S. cerevisiae* genome and molecular and biochemical studies have led to an accepted view that Tub4 is the budding yeast γ tubulin [4]. Consequently, it has been suggested that caution is required in using certain types of sequence analysis methods to classify novel tubulin sequences [6].

Within the genome of the protozoan parasite *Trypanosoma brucei*, the genes for α and β tubulin exist as a cluster of repeated α/β pairs [7]. Recently, we identified the single γ tubulin gene in *T. brucei* [8]. We then conducted a search by PCR and other means for the presence of the *T. brucei* homologue of δ tubulin. To our surprise, after we cloned the *T. brucei* δ tubulin homologue, we also identified two new divergent tubulin-like sequences. The general features can be readily visualised in the automatically generated alignment illustrated in Figure 1. Both of these new sequences are also present within the *T. brucei* genome project databases at the Sanger Centre and TIGR as partial or complete sequences (see Figure 1).

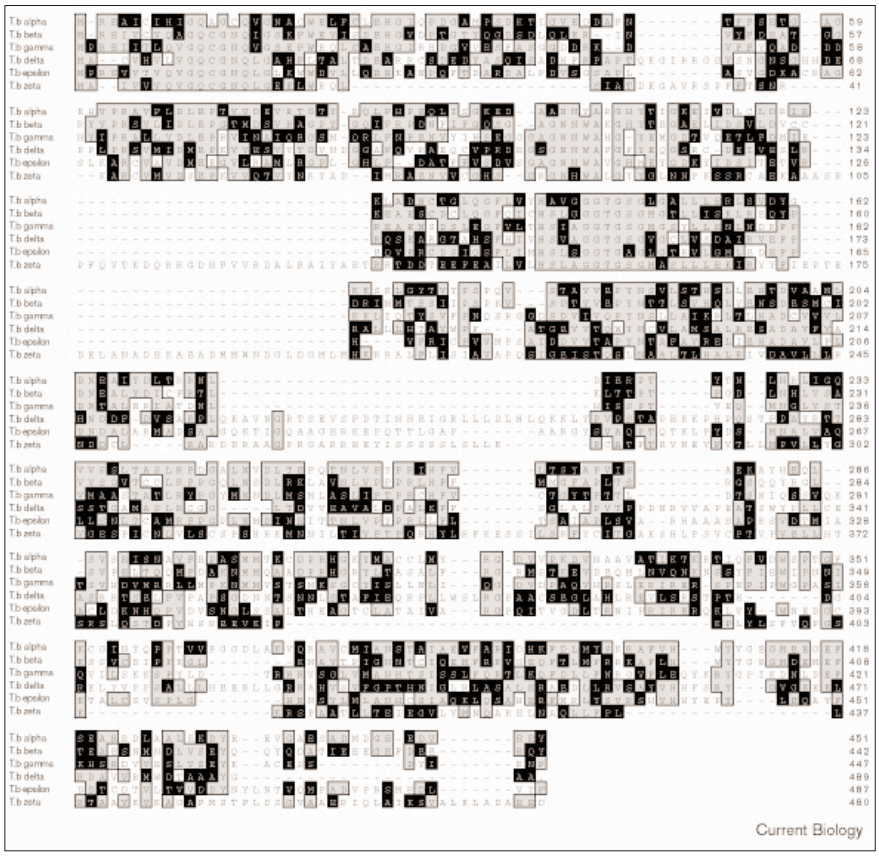
Although it was possible to conclude from sequence database searches using BLAST [9], and from searches of pattern databases such as PROSITE [10], that these sequences were likely to encode tubulins, inference of specific family membership was difficult. In contrast to the *S. cerevisiae* Tub4 debate, all four members of the tubulin superfamily α, β, γ and δ had been identified in *T. brucei*. It seemed possible, therefore, that these other sequences represented new superfamily members. Accordingly, we used the technique of protein fingerprinting used to create the PRINTS pattern database [11] to discover whether a more discriminatory approach might be able to provide analytical tools capable of distinguishing particular tubulin family members. Fingerprinting is a multiple-motif iterative process that commences with sequence alignment and excision of conserved regions. Diagnostic performance is enhanced by iterative database scanning and the motifs 'mature' with each database pass, as more sequences are matched and assimilated into the process. A generic fingerprint was created for the tubulin superfamily, then specific fingerprints for α, β, γ and δ tubulins [12].

We challenged the tubulin fingerprints for their stringency and accuracy in a number of ways. For instance, the γ tubulin fingerprint unequivocally confirmed the previously debated divergent γ tubulins from *S. cerevisiae* and *Caenorhabditis elegans* as γ tubulin members ($E = 9.0 \times 10^{-21}$ and 9.9×10^{-26} respectively). The fingerprint analysis also clearly assigned both of the two novel *T. brucei* sequences as independent new members of the tubulin superfamily. For one of these novel genes, the bioinformatics approach independently confirmed this sequence and the mammalian homologue as ε tubulin, for which biochemical evidence has been recently announced [13].

Bioinformatics analysis of the other novel sequence confirms that it, too, is a member of the tubulin superfamily, but does not belong to an existing grouping, nor to the FtsZ grouping, for which a fingerprint has also been developed [12]. We conclude, therefore, that this form of sequence analysis is likely to prove extremely useful in future analyses of the tubulin superfamily and that the sixth *T. brucei* sequence is likely to represent a new member, ζ (zeta) tubulin.

So far, δ tubulin homologues have now been identified in mammals, green algae and kinetoplastid protozoa, ε tubulin in mammals and kinetoplastid protozoa and ζ tubulin in kinetoplastid protozoa (ζ tubulin is also present in *Leishmania major*;

Figure 1



Amino acid sequence alignment of the six members of the *Trypanosoma brucei* tubulin superfamily. Identical residues are in grey and similar residues in black. Details of the sequences are: *T. brucei* α,β tubulin: amino acid sequence, GenBank accession number K02826; *T. brucei* γ tubulin: amino acid sequence, GenBank accession number Y07591; *T. brucei* δ tubulin: genomic sequence, GenBank accession number AF216742, TIGR clone 41T1; *T. brucei* ε tubulin: genomic sequence, GenBank accession number AF216743, TIGR clone 6E10; *T. brucei* ζ tubulin: genomic sequence, GenBank accession number AF241275, Sanger Centre clone TRYP1.0.2.1892, TIGR clone 22E6.

GenBank accession number AL133468). What might be the reason behind the absence in yeast and certain other organisms of these, and maybe yet other, new tubulins? We look forward to full functional characterisation [6]. However, it is intriguing to note that, at present, possession of these new tubulin genes (δ, ε and ζ) in the genome correlates reasonably well with the expression of a triplet microtubule basal body and a 9 + 2 microtubule axoneme.

Acknowledgements

Sequence data for *T. brucei* chromosome 1 was obtained from The Sanger Centre website at http://www.sanger.ac.uk/Projects/T_brucei/. Sequencing of *T. brucei* chromosome 1 was accomplished as part of the Trypanosoma Genome Network with support by The Wellcome Trust. Sequence data was also obtained from the The Institute for Genomic Research website at http://www.tigr.org. Work in K.G.'s laboratory is funded by The Wellcome Trust and BBSRC; M.N. is a Wellcome Trust International Fellow and S.V. holds a BBSRC studentship; T.A. is a Royal Society University Research Fellow.

References

1. Desai A, Mitchison TJ: **Microtubule polymerisation dynamics.** *Annu Rev Cell Dev Biol* 1997, **13:**83-117.
2. Oakley CE, Oakley BR: **Identification of γ tubulin, a new member of the tubulin superfamily encoded by mipA gene of *Aspergillus nidulans*.** *Nature* 1989, **20:**662-664.
3. Dutcher SK, Trabuco EC: **The UNI3 gene is required for assembly of basal bodies of *Chlamydomonas* and encodes delta-tubulin, a new member of the tubulin superfamily.** *Mol Biol Cell* 1998, **9:**1293-1308.
4. Weiss C, Zheng Y: **γ tubulin complexes and their interaction with microtubule-organising centers.** *Curr Opin Struct Biol* 1999, **9:**250-259.
5. Burns RG: **Identification of two new members of the tubulin family.** *Cell Motil Cytoskel* 1998, **31:**255-258.
6. Schiebel E: **Two new tubulins differ in a split decision.** *Nature Cell Biol* 2000, **2:**E3-E4.
7. Gull K: **The cytoskeleton of Trypanosomatid parasites.** *Annu Rev Microbiol* 1999, **53:**629-655.
8. Scott V, Sherwin T, Gull K: **γ tubulin in trypanosomes: molecular characterisation and localisation to multiple and diverse microtubule organising centres.** *J Cell Sci* 1997, **110:**157-168.
9. Altschul SF, Madden TL, Schaffer AA, Zhang Z, Mill W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25:**3389-3402.
10. Hoffmann K, Bucher P, Falquet L, Bairoch A: **The PROSITE database, its status in 1999.** *Nucleic Acids Res* 1999, **27:**215-219.
11. Attwood TK, Croning MDR, Flower DR, Lewis AP, Mabey JE, Scordis P, Selley JN, Wright W: **PRINTS-S: the database formerly known as PRINTS.** *Nucleic Acids Res* 2000, **28:**225-227.
12. PRINTS website: http://www.bioinf.man.ac.uk/fingerPRINTSCan/
13. Chang P, Stearns T: **δ and ε tubulin: two new human centrosomal tubulin reveal new aspects of centrosome structure and function.** *Nature Cell Biol* 2000, **2:**30-35.

Address: University of Manchester, 2.205 School of Biological Sciences, Stopford Building, Oxford Road, Manchester M13 9PT, UK.
Correspondence: Keith Gull
E-mail: K.Gull@man.ac.uk