The evolution of the structure of tubulin and its potential consequences for the role and function of microtubules in cells and embryos

JACK A. TUSZYNSKI[1,*], ERIC J. CARPENTER[1], J. TORIN HUZIL[1], WOJTEK MALINSKI[1], TYLER LUCHKO[1]
and RICHARD F. LUDUEÑA[2]

[1]*Department of Physics, University of Alberta, Edmonton, Alberta, Canada and* [2]*Department of Biochemistry, MSC 7760, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA*

ABSTRACT **This paper discusses the results of homology modeling and resulting calculation of key structural and physical properties for close to 300 tubulin sequences, including α, β, γ, δ and ε -tubulins. The basis for our calculations was the structure of the tubulin dimer published several years ago by Nogales *et al*. (1998), later refined to 3.5 Å resolution by Löwe *et al*. (2001). While, it appears that the α, β and γ-tubulins segregate into distinct structural families, we have found several differences in the physical properties within each group. Each of the α, β and γ- tubulin groups exhibit major differences in their net electric charge, dipole moments and dipole vector orientations. These properties could influence functional characteristics such as microtubule stability and assembly kinetics, due to their effects on the strength of protein-protein interactions. In addition to the general structural trends between tubulin isoforms, we have observed that the carboxy-termini of α and β-tubulin exists in at least two stable configurations, either projecting away from the tubulin (or microtubule) surface, or collapsed onto the surface. In the latter case, the carboxy-termini form a lattice distinctly different from that of the well-known A and B lattices formed by the tubulin subunits. However, this C-terminal lattice is indistinguishable from the lattice formed when the microtubule-associated protein tau binds to the microtubule surface. Finally, we have discussed how tubulin sequence diversity arose in evolution giving rise to its particular phylogeny and how it may be used in cell- and tissue-specific expression including embryonal development.**

KEY WORDS: *tubulin, microtubule, isoform, homology model, dipole*

Introduction

Microtubules (MTs) are cylindrical protein filaments found in all eukaryotes and are critically involved in a variety of cellular processes including cellular motility, cytoplasmic transport and mitosis (Hyams, 1994). Their component protein, tubulin forms a heterodimer that is composed of two homologous polypeptides, designated α and β tubulin (Postingl, 1981; Krauhs, 1981). Each individual tubulin heterodimer is 8 nm in length and interacts laterally and longitudinally to form protofilaments and then MTs (Downing and Nogales, 1998). In addition to microtubules, whose protofilament numbers vary betweens 8 and 17, under special experimental conditions tubulin can also give rise to polymorphic assemblies that include sheets, macrotubes, ribbons and several additional, exotic structures (Unger, 1990).

Microtubules undergo cycles of rapid growth and disassembly in a process known as dynamic instability which has been extensively studied both *in vivo* and *in vitro* (Mitchison and Kirschner, 1984). This instability appears to be critical for proper MT function, especially during mitosis (Kirschner and Schulze, 1986). The assembled tubulin dimer hydrolyzes a bound guanosine triphosphate (GTP) at its exchangeable nucleotide binding site to GDP; the kinetics of this process in β tubulin is critical in regulating dynamic instability by affecting the loss of the "lateral cap" (the top tubulin layer in the growing MT that stabilizes the entire structure).

In addition to forming MTs, tubulin interacts with a large number of ancillary proteins. Some of these, such as tektin, may play structural roles, while others like *tau* or MAP2 are known as microtubule-associated proteins (MAPs) are involved in MT assembly, stability and may also mediate interactions with other

---

*Address correspondence to: Dr. J. A. Tuszynski. Department of Physics, University of Alberta, Edmonton, Alberta, T6G 2J1, Canada. Fax: +1-780-492-0714.
e-mail: jtus@phys.ualberta.ca

proteins. Yet another subset of proteins, such as kinesin and dynein, are motor proteins that move cargoes, such as cellular vesicles, along MTs as part of active intracellular transport (Kikkawa, 2001; Wang and Sheetz, 2000).

An interesting feature of tubulin is its presence as multiple isoforms in eukaryotic organisms and its numerous post-translational modifications, such as phosphorylation, acetylation, detyrosination and polyglutamylation (Banerjee, 2002; Luduena, 1998). In addition to the ubiquitous α and β-tubulin, MTs often require the presence of an additional tubulin protein,'γ-tubulin, for correct assembly as a templating protein (Erickson, 2000; Luduena, 1998). Two additional tubulins, designated δ and ε, have also been characterized and although their role in MT assembly remains uncertain (Vaughan, 2000), models of their function in MT assembly have been proposed (Inclan and Nogales, 2001).

At the molecular level, the role of tubulin is extremely complex and seems to be related to the structural variations observed between α and β isoforms (Richards, 2000). The existence and distribution of numerous homologous forms of α and β-tubulin provides a link to the structure of tubulin and the polymerization and stability properties of MTs. It has been demonstrated that apparently small overall differences in the structures and therefore binding energies and chemical affinities between different tubulins may translate into significant deviations in the growth and catastrophe rates for MTs. The three-dimensional structure of bovine brain tubulin has been determined crystallographically, resulting in several atomic structures that are available in the Protein Data Bank, entries 1TUB (Nogales, 1998) and 1JFF (Lowe, 2001) being typical examples. Using these structures, it is possible to computationally predict the structures of related tubulin sequences with some degree of accuracy. We have created homology models of nearly 300 different tubulins, representing α and β-tubulins, as well as γ, δ and ε-tubulins from animals, plants, fungi and protists. For most of the resulting tubulin structures, we have been able to estimate the magnitudes and orientations of their dipole moments, charge distributions, surface areas and volumes. The magnitudes and orientations of the dipole moment of the tubulin dimer may play a role in MT assembly and stability due to their contribution to the total energy via dipole-dipole interactions. In addition, we have been able to generate plausible conformations for the carboxy-terminal region of each tubulin monomer. The C-termini of α- and β-tubulin were not resolved in the original crystallographic structures of tubulin due to their flexibility and possibly sample inhomogeneity (Nogales, 1998). The importance of the carboxy-termini is highlighted by the fact that they are the site of most of tubulin's post-translational modifications, that they bind to MAPs and that most of the differences among tubulin isoforms cluster here. We have observed that the carboxy-termini can either project out to a considerable distance from the MT or collapse onto the surface. Our results raise the possibility that the conformational transitions in the C-termini of tubulin may play a significant role in kinesin and dynein movement and that they may have a novel role in ion transport along MTs, especially in the axoplasm of neurons.

Tubulin's diversity

Gene orthologs are homologous proteins with amino acid sequences that have diverged due to mutations accumulated since their separation by speciation events (Fitch, 2000). The variations that occur can be neutral: implying that they are irrelevant to the process of natural selection, crucial: meaning that they adapt the function of a protein to a given selective pressure, or can be somewhere in between. In general, a weak but statistically significant correlation is found between this sequence variability and a protein's solvent accessible surface area (Rodionov and Blundell, 1998). Interestingly, an apparent conservation of the three-dimensional pattern of conserved amino acids has been observed in several families of structurally homologous proteins that points to the existence of a folding nucleus (Mirny and Shakhnovich, 1999). Previous reports agree that buried residues within a protein tend to evolve more slowly, but most of these analyses have been restricted to globular enzymes. Interestingly, the mapping of evolutionary rates onto the sequence of the α/β tubulin dimer demonstrate the opposite pattern (Roger, 1996). The relative rates of evolution at sites within the core of the tubulin proteins are faster than the rates of sites at the surface. The large number of intermolecular interactions between tubulins and other proteins apparently restrains the sites at the protein surface, relative to the constraints imposed by side-chain packing
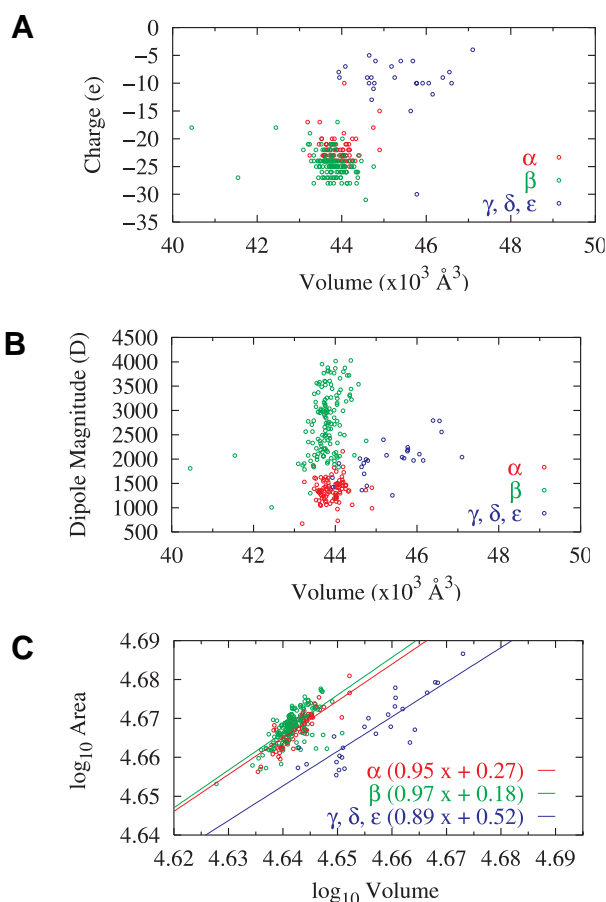

Fig. 1. Illustration of the distribution of physical properties of the monomer models. *Each ring corresponds to one model and is color-coded to show three groups within the library. Three pairs of dependencies are shown:* **(A)** *charge vs. volume,* **(B)** *dipole moment magnitude vs. volume and* **(C)** *log surface area vs. log volume.*
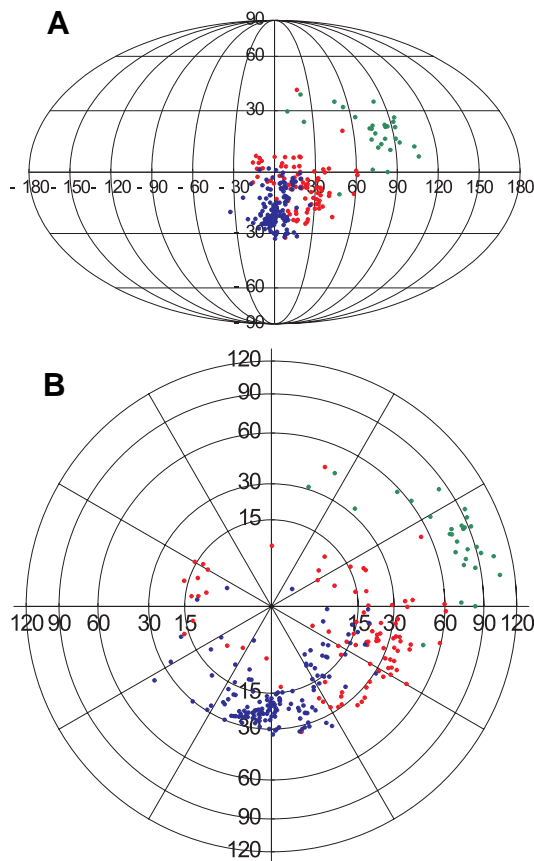
Fig. 2. Distribution of center-of-mass dipole moment orientations within the monomer library. Color-coding is as in Fig. 1. *The center of each plot corresponds to a dipole moment normal to a cylindrical microtubule.* **(A)** *An equal area (Mollweide) projection with the vertical axis representing angle with the protofilament axis and the horizontal axis angle with the protofilament-radial plane.* **(B)** *A polar projection plot, using a nonlinear radial scale for the angle for the normal position; the vertical axis corresponding to the microtubule axis.*

within the core of the proteins. This is consistent with the common assumption that co-evolution of many interacting proteins leads to a high degree of sequence constraint within the proteins where these interactions occur.

Raff *et al*. (1997) emphasized overwhelming evidence in favor of the idea that FtsZ is a homolog of tubulin, the ubiquitous eukaryotic cytoskeletal protein involved in many essential cellular processes including mitosis. Despite only limited primary sequence homology centered on a GTP-binding motif termed the 'tubulin signature sequence', the recently solved crystal structures of FtsZ and tubulin show extensive structural similarity throughout the proteins. In addition, FtsZ, like tubulin, binds and hydrolyzes GTP and assembles into protofilaments that have structures similar to those within MT's. This assembly is GTP-dependent and disassembly occurs when the GTP supply is exhausted, suggesting that FtsZ polymers, like MTs, are dynamically unstable. Tubulin and FtsZ share a common fold of two domains connected by a central helix (Nogales, 1998). Structure-based sequence alignment shows that common residues localize in the nucleotide-binding site and regions that interact with the

nucleotide of the next subunit in the protofilament, suggesting that tubulin and FtsZ use similar contacts to form filaments. Surfaces that would make lateral interactions between protofilaments or interact with motor proteins are, however, different.

Most eukaryotic organisms express multiple tubulin genes, with each distinct tubulin isoform having an essential function. It has been shown, that in the testis of *Drosophila* and the touch neurons of *Caenorhabditis elegans*, not every tubulin molecule is interchangeable. For example, the *mec7* and *mec12* tubulins are required for assembly of specialized MTs that contain 15 protofilament. It has been demonstrated that γ tubulin is required for the initiation of MT assembly. Additionally, four new members of the tubulin family have recently been identified (Dutcher, 2001). The first, δ-tubulin was identified as a flagellar assembly mutant in the green algae *Chlamydomonas reinhardti* and has also been found in *Chlamydomonas Euplotes*, human, rat, *Xenopus*, mouse, dog and Trypanosomids. Second, η-tubulin was discovered due to a mutation in *Paramecium* that results in a basal body duplication defect. The ε and ζ-tubulins, were discovered using bioinformatics techniques, however their cellular functions have yet to be determined. ε-tubulin is present in *Chlamydomonas*, human, mouse, rat, *Xenopus* and Trypanosomids. ζ-tubulin has been found in *Trypanosoma* and *Leishmania* and as an expressed sequence tag in *Xenopus laevis*. Interesingly, the δ, ε, ζ and η-tubulins do not appear to have a ubiquitous distribution throughout all eukaryotic organisms, as is the case for the α and β tubulins. The genomes of *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Aradbidopsis thaliana* and *C. elegans* do not have clear homologs to the δ, ε, ζ and η-tubulins. The absence of these tubulins from fungi and plants can be correlated with the lack of centrioles and basal bodies. The high degree of divergence among the new tubulins in both the same and different families suggests that their functions are not as constrained by their primary sequences as those of the α, β and γ-tubulins. The elucidation of the biological function of these new members of the tubulin superfamily will hopefully resolve their roles within the cell.

Early sequence comparisons for tubulin gave valuable insights into the function of α and β-tubulin (Little, 1981; Little and Seehaus, 1988; Burns, 1991; Burns and Surridge, 1994). Alignments between the orthologs and paralogs of α and β-tubulin are unambiguous due to the highly conserved amino acid sequences between these two proteins. Recently, Kuchnir-Fygenson *et al.* (2004) analyzed the homology within the tubulin family using multiple sequence alignment techniques. They compared members of the tubulin super-family from human, *Drosophila*, *Chlamydomonas*, *Trypanosoma* and *Paramecium*. Within this subset, it was shown that all the α-tubulins are 89–95% similar to one another while the β-tubulins were only 88–94% similar. It was also demonstrated that the γ-tubulins have similarities ranging from 72% to 94%, while the three δ-tubulins were only 47–57% similar. The two complete sequences for the ε-tubulin are only 58% similar. A portion of the ε-tubulin gene from *Chlamydomonas* is 79% similar to the human gene over the 87 amino acids within the region that is most conserved between humans and *Trypanosoma*. The predicted *Xenopus* sequences for ζ and η-tubulin, 89 and 97 amino acids long, are 60% and 61% similar to the *Trypanosoma* and *Paramecium* sequences, respectively. Among the 21 full-length tubulin sequences, only 7 amino acids are

conserved and include the the GDP/GTP-binding site, an aspartate within helix 7 and a proline near the M-loop that interacts with the H3 helix within a tubulin from the adjacent protofilament within the MT.

By examining sequence alignments between α- and β-tubulin, Kuchnir-Fygenson *et al.* (2004) identified residues that differ significantly in variability. Most of these residues were shown to be clustered around the nucleotide-binding pocket, where the greatest functional differences between the two types of tubulin exist. The remaining residues associated with large differences in variability are found in the N-terminal loop between helix (H) 1 and beta sheet (S) 2. The statistical distribution of residue variability in both α and β-tubulins is strongly peaked at low values, with >50% of residues scoring in the bottom 10% of the variability range. In the structure-based alignment, the correlation coefficient for variability between the tubulins is $R(\alpha/\beta)= 0.42$. This alignment involves two gaps in β-tubulin: a small one in the disordered N-terminal loop (β, 39–40) and a larger one in the loop between S9 and S10 (β, 362–365). Three distinct regions are misaligned to accommodate multiple residues with large differences in variability between the paralogs. These misaligned regions include a total of approximately 100 residues located in helices H1 and H4, sheets S4, S5 and S9, two turns (T4 and T5) and the disordered N-terminal loop (L1). Among these, there are 40 positions that appear to be of particular functional importance where homologous residues differ in variability by more than a standard deviation from the mean. Half of these are clustered around the nucleotide-binding pocket, four are clustered around the taxol binding site on β-tubulin and one participates in lateral binding between MT protofilaments. Of the 20 residues near the nucleotide-binding pocket, 16 are more variable in the «N-site», which binds GTP without catalyzing its hydrolysis and less variable in the «E-site», which hydrolyzes GTP as dimers assemble into protofilaments. Of the four residues that are less variable in the N-site, the tyrosine residue at position 172 in α-tubulin is particularly interesting as a target for directed mutagenesis because it interacts directly with the nucleotide. Because analogous differences in variability were so plausibly connected with functional differences in the other misaligned regions, Kuchnir-Fygenson (2004) predicted that these residues in the N-terminal loop have a role in making the biochemical functions of α and β-tubulin distinct. It is suspected that the functional distinction is related to GTP hydrolysis and that the tenuously structured glycines within β-tubulin are involved in a hydrolysis-driven conformational change that eventually results in a catastrophe event.

We conclude this introductory survey on tubulin diversity stating that while it is tantalizing to speculate about a direct relationship between tubulin structure and function, the precise molecular basis of the properties of tubulin is still not well understood. This is in part because tubulin's highly flexible C-terminal conformation (Keskin, 2002) makes it difficult to crystallize this region.

Tubulin phylogeny

Phylogenetics is the science of estimating and analyzing evolutionary relationships. Phylogenetic relationships among micro-organisms are especially difficult to discern. DNA, RNA and proteins can be considered to be 'information molecules' since they retain a record of an organism's evolutionary history. A logical approach then is to compare nucleic acid or protein sequences from different organisms and estimate the evolutionary relationships based on the degree of homology between the sequences. The nucleotide or amino acid differences within a gene can therefore reflect the evolutionary distance between two organisms. Closely related organisms will exhibit fewer sequence differences than distantly related organisms. Through the molecular approach in determining phylogenetic relationships, the differences are readily quantifiable as sequences from different organisms can be compared and the number of differences computed. These data are often expressed in the form of phylogenetic trees, where the positions and lengths of the "branches" depict the relatedness of sequences between organisms. Interestingly, phylogenetic trees that are produced from different gene sequences may produce different topologies. Possible reasons for these inconsistencies are: (a) disparities in evolutionary rates among lineages, (b) uneven taxonomic sampling, (c) single explosive radiation of major eukaryotic taxa, (d) horizontal DNA transfer between species and (e) gene paralogies (i.e., duplications) and gene conversion. The first two phenomena result in a long-branch attraction artifact in which many slowly evolving sequences cluster to the exclusion of a few rapidly evolving sequences. The long branches that are far apart in the lower portion of the eukaryotic branch may be a result of the experimental procedure. In addition, events like horizontal DNA transfer and gene duplications complicate the analysis of molecular phylogenetic data. Some of these problems are resolved by combining data into consensus trees, for example, by combining protein data from actin, α-tubulin and β-tubulin (Roger, 1996).

Wilson and Borisy (1997) stated that although highly conserved, the variable regions within α and β-tubulins show divergence from other α and β-tubulins within the same species, but can show conservation between different species. Such conservation raises the question of whether the diversity in tubulin sequence can mediate the diversity in MT organization. Recent studies probing the function of β-tubulin isoforms in axonemes of insects suggest that tubulin structure, through interactions with extrinsic proteins, can direct the architecture and supra-molecular organization of MTs. The tubulin gene family includes, at least, three highly conserved subfamilies, α and β and γ-tubulin, that

Fig. 3. Comparison amongst human -tubulin sequences. *Each of the horizontal lines represents one sequence and has boxes at horizontal locations corresponding to residues at which the sequence differs from a consensus sequence. Note the presence of a hypervariable region, beginning after residue 425.*
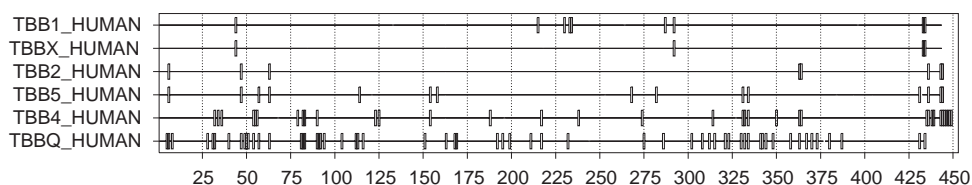
arose from a series of gene duplications early in eukaryotic evolution (Edlind *et al.*, 1996; Keeling and Doolittle, 1996). Currently, α and β-tubulin genes have the widest taxonomic representation and have been identified in early diverging eukaryotic lineages (Burns, 1996; Edlind, 1996; Keeling and Doolittle, 1996). Cleveland *et al.* (1980) concluded that the human genome contains approximately 14 copies of the α and β-tubulin genes.

Adoutte *et al.* (1984) produced a survey of the electrophoretic behavior of the tubulins of 23 species (mostly protists) as well as their reactivity towards 4 anti-tubulin antibodies. This allowed a rough evaluation of evolutionary relatedness between the various groups of protists analyzed. The results indicated a striking agreement between a number of published phylogenies. Gaertig *et al.* (1993) cloned and sequenced the two β-tubulin genes of the ciliated protozoan *Tetrahymena thermophila*. These two genes encode identical 443 amino acid peptides which are 99.7% identical to the β-tubulin proteins of *T. pyriformis* and 95% identical to the human β1-tubulin. *T. thermophila* contains only one α-tubulin gene, implying that all of the diverse MT structures in this unicellular organism must be formed from a single α and a single β-tubulin heterodimer. Additionally, Gaertig *et al.* (1993) also carried out a phylogenetic analysis of 84 complete β-tubulin peptide sequences. This analysis supports the following two hypotheses regarding β-tubulin evolution and function: (a) that multifunctional β-tubulins are under greater evolutionary constraint than β-tubulins present in specialized cells or in cells with very few MT-related functions, which can evolve rapidly; and (b) that cells which form axonemes maintain a homogeneous population of tubulins. Luduena (1998) reviewed the distribution and possible functional significance of the various forms of tubulin. In analyzing the differences among the tubulin isoforms that are encoded by different genes, it was shown that some appear to have no functional significance, while some increase the overall
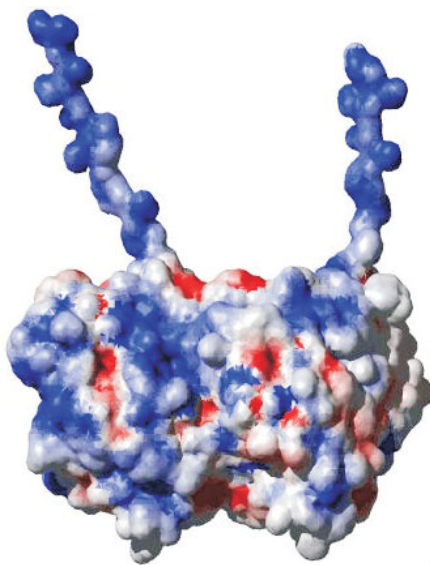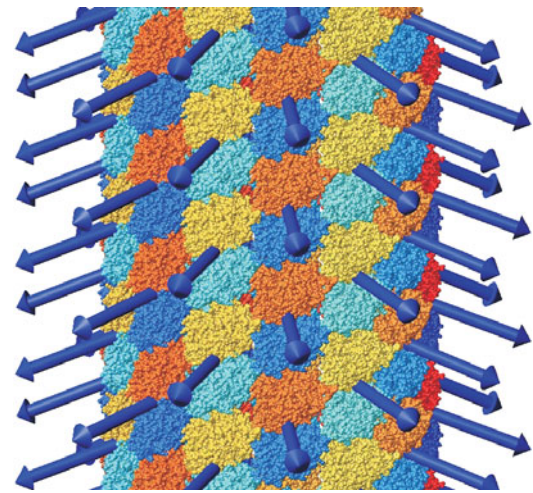

Fig. 5. **Dipole orientation of dimers in a microtubule.** *The arrows indicate the orientation of the permanent dipole moment with respect to the centers-of-mass of individual tubulin-tubulin dimers with respect to the surface of a microtubule. Note that dipole moments point into the dimers. Figure prepared using MolMol (Koradi, 1996).*

adaptability of the organism to environmental challenges and yet others appear to perform specific functions including formation of particular organelles and interactions with specific proteins (Keeling, 1998).

Keeling and Doolittle (1996) sequenced genes from a number of protozoa and constructed phylogenetic trees of α and β-tubulins. These trees were consistent with each other, but inconsistent with other molecular phylogenies. The agreement between α and β-tubulin trees could arise only by the co-adaptation of one molecule to variants of the other as a result of their intimate steric association in MTs. Therefore, these trees may not provide independent support for the phylogenetic results. Relationships between the protist tubulins were also examined by constructing trees of all three types. These trees were found to be of limited value for determining the position of the root within each subfamily because of the great interfamily distances, but they confirmed the classification of all known genes into three monophyletic subfamilies. Divergent genes from *C. elegans* and *S. cerevisiae* representing the novel classes of δ and ε-tubulin were found to be specifically related to γ-tubulins from animals and fungi respectively and therefore are seen as rapidly evolving orthologs of γ-tubulin.

Edgcomb *et al.* (2001) elucidated the relationships among jakobid flagellates and other early-diverging eukaryotic lineages by characterizing α and β-tubulin genes from four jakobids. Tubulin gene phylogenies were in general agreement with mitochondrial gene phylogenies and ultrastructural data. The final representation of major eukaryotic groups was pared down to the alignments of 42 α and 39 β-tubulin sequences that were amenable to more rigorous phylogenetic analyses. A variety of traditional eukaryotic groups that are well established on the basis of morphology or molecular phylogeny were also recovered in either α or β-tubulin trees. In addition, several «higher- order» eukaryotic groupings were observed, such as fungi plus microsporidia, animals plus fungi and ciliates plus apicomplexa (alveolates). Unlike rRNA trees, tubulin phylogenies do not show many inde-

Fig. 4. **A map of the electric potential on the surface of a tubulin dimer with C-termini tails.** *Red regions represent positive charge, while blue regions indicate negative charge. The intensity of coloring indicates the local surface strength and polarity of the field. Figure prepared using MolMol (Koradi et al., 1996).*

pendent deep-branching lineages leading to a «crown» radiation of eukaryotes but, instead, a deep split occurs within eukaryotes separating two robust groupings, a plant-protist «superclade» and an animals-plus-fungi group, with several protist lineages intervening. It is likely that the extreme conservation of both tubulin genes leaves few phylogenetically informative sites, especially for groups in the plant-protist superclade (Keeling, 1999; 2000).

Homology modeling

The presence of multiple forms of tubulin provides us with the opportunity to create a model library from which we can estimate physical characteristics for a large number of tubulin sequences. Once three-dimensional protein structures are determined, it becomes possible to use techniques, such as homology modeling to predict structures of proteins with similar sequences (Chothia and Lesk, 1986). Homology modeling utilizes several structural motifs from template proteins and pieces them together to form the final model. A scoring function then assesses both the sequence identity between the target sequence and template and the overall quality of the template that is being considered. The scores are ranked and the fold with the best score is assumed to be the one adopted by the target sequence. A computer program uses alignment of the sequences with known related structures to obtain spatial restraints that the output structure must satisfy. Additional restraints derived from statistical studies of representative protein and chemical structures are also used to ensure a physically plausible result. Missing regions are predicted by simulated annealing of a molecular mechanics model. Since the 3D structures of tubulin obtained from crystallography lack the extreme carboxy-termini, we used this capability to create structure within this region of the protein. Our experience with molecular mechanics models of these regions suggests that the residues are too mobile for such structures to be anything more than one of many possible conformations. It is our approximation that any models with gross errors will be insignificant when comparisons are made amongst several models.

We have performed homology modeling on a set tubulin sequences, containing α and β-tubulins from animals, plants, fungi and protists in the Uniprot Knowledgebase. As an initial step, the Swiss-Prot database Release 40.3 of 08-Nov-2002 (Boeckmann, 2003) (available at http://www.expasy.org/sprot/) was searched for tubulin amino acid sequences. A search using the keyword «tubulin» was manually filtered to separate actual tubulin sequences from those of other tubulin-related proteins. This provided some 290 sequences, representing a wide range of species. Of these, 27 sequences are annotated as being fragmentary leaving 263 complete tubulin monomer sequences. Of particular interest were the 15 human sequences obtained in this process. We believe that this core library is large enough to adequately sample the diversity of tubulin variation, therefore no attempt to systematically add to this set from other data sources has been made, although select sequences of interest from other sources have also been studied.

Appendix A summarizes all the tubulin sequences used in this study. The table names the source organism and for each α, β, γ, δ and ε, gives the name used in the database. It is important to relate the biochemical data encapsulated by the amino acid
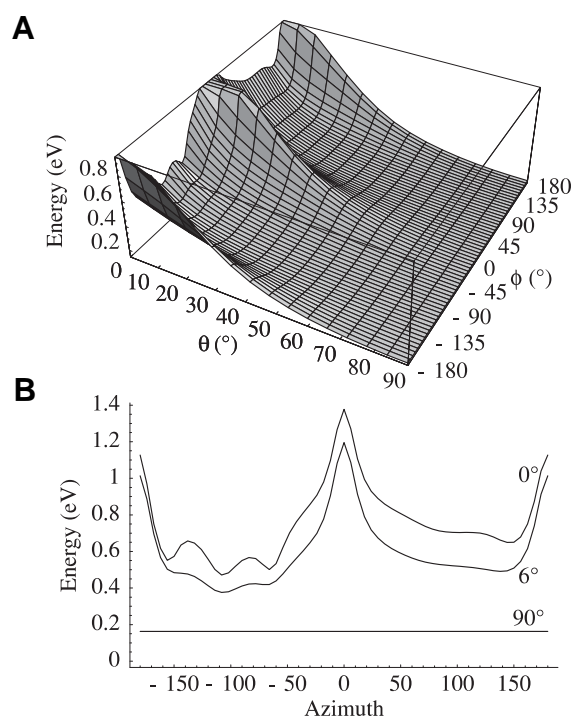

A

B

Fig. 6. Graph of energy level versus position of rigid C-terminal tail in a simplified electrostatic model. (A) *Energy of system as function of azimuth (ϕ) and declination (θ). **(B)** The energy levels from part (A) for declinations of 0°, 6° and 90° from top to bottom respectively. Note energy minima for sticking straight out ($\theta = 90°$) and for lying on the surface ($\theta = 0°$). See Priel* et al. *(2005).*

sequence to the biologically relevant information presented in this table. Fragmentary sequences were processed like all the others, however the results from such sequences must be reviewed before accepting data from them and were filtered out when appropriate in the data sets presented here. In some circumstances it may be appropriate to replace missing regions with a standard sequence to estimate results for these fragments.

Currently, a total of three structures, providing a total of four templates are available in the Brookhaven Protein Databank upon which we can construct homology models of tubulin. The first and most obvious choice, was the initial Nogales structure of the α/β tubulin heterodimer (PDB identifier 1TUB) (Nogales, 1998). However, this structure was solved at a relatively low resolution and contains numerous omissions and misalignments due to difficulties in density fitting. A refinement of the 1TUB structure, PDB identifier 1JFF, was subsequently deposited with a slightly better resolution at 3.5 Å (Lowe, 2001). Along with increased resolution, the 1JFF structure also addressed a number of misalignment errors within the 1TUB structure and therefore makes a better choice for a homology modeling template. However, the 1JFF structure contains a large, nine-residue gap within the α tubulin M loop (serine 277 to leucine 286). Fortunately, this gap was not present in the 1TUB structure, providing us with a template structure for this region. Shortly after we began modeling the tubulin isoforms, a new structure of tubulin complexed with stathmin was released (PDB identifier 1SA0) (Ravelli, 2004). This structure was solved at 3.58 Å and contained most of the a

loop that was not present in the 1JFF structure. The presence of the complete missing a loop in 1TUB and partial loop in 1SAO provided us with adequate templates for modeling this section of the α-tubulins. A fourth template structure from which to build the tubulin homology models was available to us. This structure was that of the *E. coli* cell division protein 1FSZ (Lowe and Amos, 1998). Unfortunately, these five template structures are not entirely unique, the 1TUB and 1JFF being largely the same models.

When generating models of the α tubulin monomers, only the FtsZ and the two α-structures were input to Modeller as templates (Sanchez and Sali, 2000). Similarly, when modeling β-monomers only the FtsZ and the two β structures were used. In other cases all five structures were input. Models were also constructed without including the FtsZ structure, a putative prokaryotic homologue of tubulin (Erickson, 1997; 1998). These were judged to be insignificantly different. After constructing several test models we determined that the presence of extended loops within 1FSZ, that were absent from all of the tubulin sequences tended to alter the alignment scores and produce alignment anomalies between the tubulin alignments. It was therefore decided to eliminate the FtSZ structure from our templates and because of the high sequence homology between the β and α-tubulins we chose to use only 1JFF as our template.

Results

With the resulting library of structural tubulin models, various computational estimates of physical properties of the different tubulins may be made. These include the volume, surface area, net charge and dipole moments. All volumes and surface areas reported herein are calculated using a Connolly solvent accessible surface. Similarly, all dipole moments reported are with respect to the protein center-of-mass or alternatively a geometric mean of the model co-ordinates which is expected to be a good approximation of the former. We performed these calculations on the model structures, typically using analysis tools within the Gromacs (Lindahl, 2001) molecular dynamics package (version 3.1.4). Because only a single δ and a single ε sequence are present and these somewhat resemble γ, it is convenient to consider three groups: α, β and γ\δ\ε. We show some scatter plots of these three groups below in Figs. 1 and 2.

We also analyzed the properties of the C-terminal projection. We first needed to define this region. We used Clustal W (version 1.82) (Thompson, 1994) in order to obtain a multiple sequence alignment amongst the peptides. Such an alignment is illustrated in Fig. 3 which highlights variations in amino acid sequences. The multiple alignment then allows rapid identification of corresponding residues in all of sequences. Because this alignment is done amongst different sequences and with different parameters than the alignments performed with Modeller it may be that slight differences in the alignments have occurred. Studying these may be instructive in understanding some of the uncertainties associated with the Modeller outputs. However, we have ignored this possibility and defined residue 436 to be the first in the tail. This corresponds well to the beginning of the hypervariable region and the last contact of the backbone with the globular structure in 1TUB and 1JFF. From this definition, we can then readily read and analyze the C-terminal projection sequences for each tubulin species.

The library of tubulin structures including α, β and γ from animals, plants, fungi and protists, as well as δ and ε from humans, were analyzed by molecular mechanics to determine their net charges, dipole moment components, dipole orientations, volumes, surface areas and the lengths and charges of their C-termini. The results of our computations in this regard are shown in Appendix B.

Fig. 1A shows a scatter diagram of the net charge/volume ratios of the different tubulins. This plot is striking in that the net charge on the β-tubulins is by far the greatest ranging between 17 and 32 elementary charges (e) depending on the particular β tubulin with an average value in this case at approximately 25 e. Next come the α tubulins with net charges varying between 10 and 25 elementary charges. Finally, the γ, δ and ε tubulins have a much lower value of the net charge; in some cases being very close to zero. There appears to be little if any correlation between the size of a protein and its charge, although the γ-tubulins are noticeably larger in volume. Further, it should be kept in mind, that the charge on a tubulin dimer will be neutralized in solution due to the presence of counter-ions which almost completely screen the net charge. This was experimentally determined for tubulin by the application of an external electric field; the resulting value of an unscreened charge of approximately 0.2 e per monomer was found (Stracke, 2002). What is, however, of great interest in connection with polymerization of tubulin into MT's and with protein-protein and drug-protein binding is the actual distribution of charges on the surface of tubulin. Fig. 4 illustrates this for the Downing-Nogales structure with plus signs indicating the regions of positive charge and minus signs negatively charged locations. This figure shows C-termini in two very upright positions. Each of the different tubulins will show differences in this regard but lack of space prevents us from showing in detail where the differences occur for specific monomers.
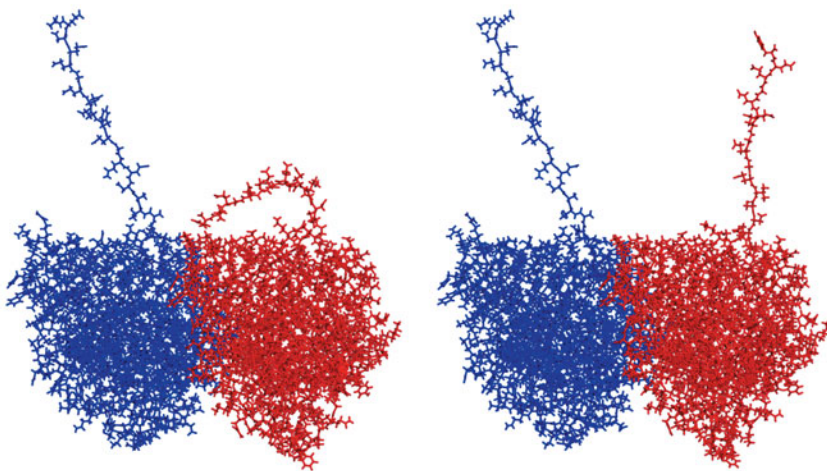

Fig. 7. Examples of C-termini in up and down positions. *Although only the C-terminus on an α - monomer is shown in a lowered state, both C-termini can be in either state.*

Fig. 1C shows the logarithm of surface area against the logarithm of volume for the different tubulins. Note the high level of correlation between the tubulin family and the surface-volume functional dependence. It is worth noting that the γ, δ and ε-tubulins are significantly more compact than the α tubulins, which are slightly more compact than the β-tubulins. We also show a best-fit line to these scatter plots for each of the three families. Note that the α and β families have a very similar slope with a value close to the unity that is indicative of cylindrical symmetry in the overall geometry. The γ, δ and ε-tubulins are characterized by a lower value of the slope indicating a greater amount of spherical symmetry for which surface-to-volume dependence scales with a power of 2/3.

As seen in Fig. 1B, α-tubulins have relatively low dipole moments about their centers-of-mass ranging between 1,000 and
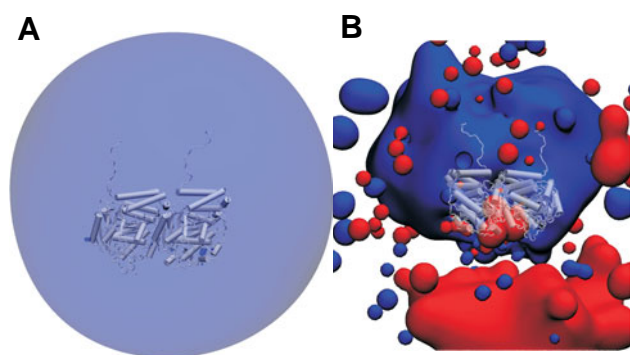
Fig. 8. A tubulin dimer in vacuum (A) *is seen principally as a monomer from the point of view of electrostatic interactions;* **(B)** *it is mainly dipolar when surrounded by counter ions and water. Figure prepared with VMD (Humphrey, 1996).*

2,000 debye, while the β-tubulins are very high in this regard with the corresponding values ranging between 1,000 and 4,000 debye and with the average value close to 3,000 debye. The γ-, δ- and ε-tubulins have dipole moments whose values are in between those two ranges. There is clearly some correlation between the magnitude of the dipole moment and volume for β- and γ-tubulins while the α tubulins are so compact that no correlation is readily apparent. In Fig. 2 we have illustrated the important aspect of dipole organization for tubulin, namely its orientation. Fig. 2A shows a Mollweide projection of dipole orientation in tubulin. The center of the diagram corresponds to the orientation perpendicular to the cylinder surface in a MT. The northerly direction represents the direction along the protofilament towards the plus-end while the southerly direction points towards the minus-end. Points along the horizontal axis are oriented tangentially to the MT surface. We conclude from this diagram and its magnification around the center region in Fig. 2 that both α- and β-tubulins orient their dipole moments in a direction that is close to being perpendicular to the MT surface. Notably, the same cannot be said about the dipole orientation for γ- tubulin, whose dipole orientation is largely tangential to the cylindrical surface of the MT formed out of its subunits.

The C-terminal tails (residues beyond the H-12 helix) display a largely random configuration, in stark contrast to the rest of the monomer. This is due to the inherent flexibility of this region of the sequence and the *in vacuo* simulation techniques employed by

Modeler. While the variety of configurations was not unexpected they were included for completeness and do not detract from the rest of the model. Furthermore, they reflect the conformational flexibility of these regions. When tubulin is modeled as a dimer or monomer in solution, implicit or explicit, that simulates physiological conditions we see a variety of conformations for the C-terminal tails. While conformations that point away from the MT dominate (as would be entropically expected) the tails will also occasionally make contact with what would be the surface of the MT. Generally speaking, the C-terminal tails sweep out a large volume of space over several nanoseconds. Our models show that only α- and β-tubulins have C-terminal tails that project outwards from the tubulin, due to their high negative charges. Fig. 6 shows the energy levels of different orientations of the C-termini in a toy model with rigid tails and simplified electrostatics and suggests that there is relatively little energetic difference between the tails projecting straight outward from the rest of the tubulin and lying on the surface of tubulin in certain energy minima. While other surface states are quite strongly disfavored, the surface minima occur in positions where their negative charges form electrostatic bonds with certain positively charged residues on the tubulin surface. This variation is illustrated by Fig. 7. In a more realistic model this would need to be extended to consider different C-terminal structures and surface charges. This aspect is still underway and it will be extended to investigate the effects of post-translational modifications such as phosphorylation and polyglutamation on the equilibrium C-terminal states.

Discussion

The structures of α and β-tubulins are known to be quite similar, being nearly indistinguishable at approximately 6 Å (Li, 2002) despite only a 40% amino acid homology. Since the sequences within the α or β-tubulin family are more similar to each other than to those sequences belonging to the other families of tubulins, it is reasonable to believe that any given sequence should produce a structure very similar to another member of a given family. Further support for this comes from the published structures of Nogales (1998) and Lowe (2001) which are of a porcine sequence, but which were fit to structural data from an inhomogeneous bovine sample. Accordingly, by substituting appropriate amino acid side chains and properly adjusting other residues to accommodate insertions and deletions in the sequence, crystallographic structures can be used as a framework to produce model structures with different sequences with a high degree of confidence. Indeed, we have demonstrated that homology modeling produced a set of models that are essentially indistinguishable from one another. Therefore, homology modeling may not be an appropriate test, as it attempts to move the resulting models towards the templates. The universal convergence that we see here only proves that the structure can accommodate all of the tubulin sequences that we have investigated. While homology modeling or threading may be the best way to generate a large number of models, additional methods are required to test the stability of the resulting conformations obtained from the homology modeling. Having said that, it is still of interest to compare the results of our calculations and find possible consequences of the differences in the physical characteristics between various tubulin isoforms.

The dipole moment of tubulin could play a role in MT assembly and in other processes such as interactions with neighboring proteins and ions, espcially when the dipole moments can be ordered leading to collective effects (Brown and Tuszynski, 1999) or even dynamical propagating states (Trpisova and Tuszynski, 1997). This could be instrumental in the docking process of molecules to tubulin and in the proper steric configuration of a tubulin dimer as it approaches a MT for binding. An isolated (in vacuum) tubulin dimer has an electric field dominated by its net charge as seen in the nearly spherical isopotential surface in Fig. 8A. In contrast, a dimer, shown in Fig.8B, surrounded by water molecules and counter-ions, as is physiologically relevant, has an isopotential surface with two lobes much like the dumbbell shape of a mathematical dipole. In a MT, the individual dimers have near parallel dipole moments relative to their nearest neighbors; this is illustrated in Fig. 5. The greater the individual dipole of each of its units, the less stable the MT will generally be since these near parallel dipole-dipole interactions are repulsive, tending to push apart the MT. Note that the strength of the interaction potential is proportional to the square of the dipole moment; hence MT structures formed from tubulin units with larger dipole moments should be more prone to undergo disassembly catastrophes (Trpisova and Tuszynski, 1997) compared to those MTs that contain low dipole moment tubulins. For organisms that express more than one type of tubulin isoform in the same cell, one can conceive that MT dynamic behavior could be regulated by altering the relative amounts of the different isoforms according to their dipole moments. However, these effects are moderated by local interactions between two dimers, which produces a more complicated electrostatic field than the assembly of dipoles described here would suggest; for details see (Baker, 2001).
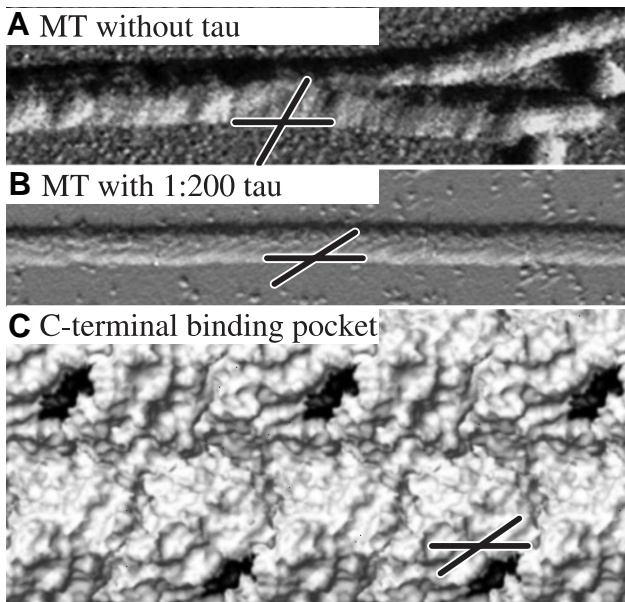

A MT without tau

B MT with 1:200 tau

C C-terminal binding pocket

Fig. 9. Microtubule surface patterns show different angular structures. *STM observations of MTs* **(A)** *without tau MAPs and* **(B)** *with tau as reported by (Makrides, 2003).* **(C)** *A view of a microtubule surface with a C-terminal binding site observed in molecular dynamics simulations highlighted. The three angle markers are at 65°, 32.5° and 35°, respectively.*

In terms of surface/volume ratios, α and β-tubulins are the least compact, while γ, δ and ϵ are the most compact. There is abundant evidence that both α and β have variable conformations. This is attested to by their interaction with drugs and is consistent with the dynamic instability of MTs. In contrast, there is no evidence of γ, δ and ϵ participating in dynamic instability nor is there any theoretical reason to imagine such flexibility. It is reasonable to postulate that a less compact structure may have a more flexible conformation.

Isoform composition has a demonstrable effect on MT assembly kinetics (Panda, 1994; Banerjee and Kasmala, 1998). This could be due to changes in the electrostatics of tubulin, in particular, its net charge. This property although significantly screened by counter-ions does affect MT assembly by influencing dimer-dimer interactions over relatively short distances (approximately 5 nm) as well as the kinetics of assembly. These short-range interactions have been recently studied by (Sept, 2003) by calculating the energy of protofilament-protofilament interactions. These authors concluded from their work that the two types of MT lattices (A and B lattices differ in the relative positions of the protofilaments relative to each other) correspond to the local energy minima. However, the matters can be significantly more complicated since not only does tubulin exist as multiple isoforms, but the protein can also undergo various post-translational modifications, such as phosphorylation, acetylation, detyrosination and polyglutamylation (Banerjee, 2002; Luduena, 1998).

Our models predict that the C-terminal tails of α and β can readily adopt the two extreme conformations: either projecting outwards from the tubulin (and the MT surface) or lying on the surface, albeit such that their charged residues can form electrostatic bonds with complimentary charges on the surface. The state of the C-terminus (upright, down, or in intermediate states; see Fig. 7) is easily influenced by local ion concentrations including pH. This conformational complexity has many implications. First, a projecting C-terminus could play a major role in signaling. The fact that tubulin isoforms differ markedly in their C-termini (Sackett, 1995) suggests that specific sequences may mediate the functional roles of the isoforms. These sequences would be readily available for interactions with other proteins in a projecting C-terminus (Sarkar, 2001). Second, the C-termini are the sites of many of the post-translational modifications of tubulin, i.e. polyglutamylation, polyglycylation, detyrosination/tyrosination, removal of the penultimate glutamatic acid and phosphorylation of serine and tyrosine (Redeker, 1998). It is known that the C-termini are essential to normal MT function (Duan, 2002); a projecting C-terminus would be easily accessible to enzymes that effect these modifications and the modification could influence the likelihood of the C-terminus changing conformation. In addition, if the modification plays a role in signaling, then the signal would be readily available in a projecting C-terminus, as mentioned above. Third, projecting C-termini would automatically create spacing between MT's. It is known that MT's are never closely packed and are surrounded by what is sometimes referred to as an «exclusion zone» (Dustin, 1984). This is a region of space around them that is strongly unfavorable to the presence of other MT's in the vicinity. Although MAP's likely play a role in establishing such spacing, electrostatic repulsion among C-terminal ends might also influence this as well. The C-termini are the major sites of binding of the MAP's to tubulin. A projecting C-terminus may
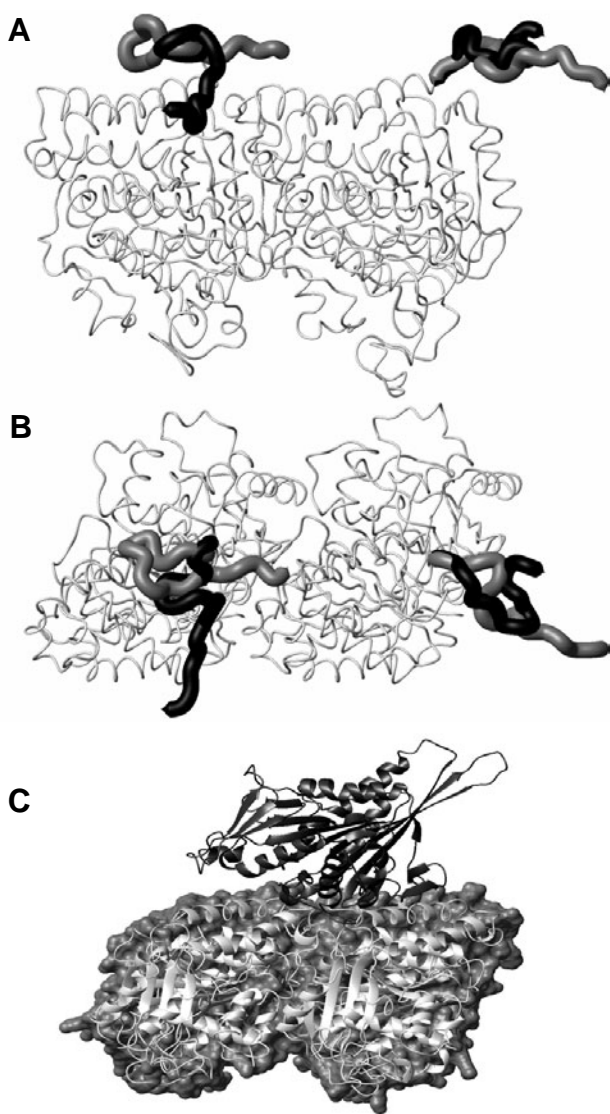
Fig. 10. Microtubule surface decorations. *Two states of the C-terminal tail are shown:* **(A)** *view tangential to the microtubule surface,* **(B)** *at right angles looking in from the outside of a microtubule. For comparison,* **(C)** *the tubulin-kinesin motor domain complex. This figure is from PDB file 1IA0 (Kikkawa, 2001) and shows two monomers from adjacent dimers with an -monomer on the left and a -monomer on the right.*

facilitate MAP binding and, conversely, MAP binding could influence the conformation of the C-terminus. Evidence for this is provided by the work of Makrides *et al.* (2003) who showed that when the MAP tau binds to MT's, it triggers a structural change on the MT surface whereby a structural element, presumably tau, lies on the surface of the MT forming a lattice whose alignment angle is much more acute than that of the lattice of tubulin subunits. This lattice is apparently superimposed on top of the normal MT (A or B) lattice. The orientations of the C-termini when they are lying on the surface of the MT form nearly the same kind of pattern that Makrides *et al.* (2003) observed, a striking confirmation of the potential accuracy of our modeling. The observed angles are listed in the caption of Fig. 9. These results raise the possibility

that the orientation of the C-termini of the α and β-subunits determines the arrangement of tau molecules on the MT.

In addition, the state of the C-termini could mediate how motor proteins such as kinesin or dynein bind to and move on MT's. Our models show that kinesin binds preferentially to upright C-termini and not to C-termini lying on the surface of the MT (see Fig. 10). Very minor changes in the local ionic environment or the pH could halt the processive motion of a two-headed kinesin (Wriggers and Schulten, 1998) by collapsing the C-termini. One can postulate that the proportion of C-termini that are in the upright conformation in a given portion of the MT could determine the actual rate of kinesin movement. It is likely that such arguments could apply to other motor proteins as well. One might imagine that the very fine coordination of movements that occurs in processes such as mitosis could be influenced or even caused by the conformational state of the C-termini in particular areas of the MT. Finally, one can speculate that the C-termini could collapse in waves that could simultaneously couple to a wave of ions that could polarize or depolarize a membrane. This could be a form of MT signaling that has not yet been considered in the literature. A quantitative model of ionic wave transmission coupled to coordinated motion of the C-termini of dendritic MT's has been recently developed by Priel *et al.* (2005).

The role and function of MTs in embryos has been discussed in depth by Gordon (1999). Based on important studies of the multi-gene families of tubulin (Lewis and Cowan, 1990), it can be concluded that all tubulins in the same organism appear able to participate in all major functions. Nonetheless, some tubulins are produced primarily in certain tissues at specific stages of development. Interestingly, tissue-specific tubulins are often expressed at lower evolutionary levels in some other tissues. To complicate matters more, nearly identical tubulins may have entirely different patterns of expression. However, some tubulins differ in their ability to participate in dynamic equilibrium phenomena, their role in MT nucleation as well as sperm formation. Particularly intriguing is the fact that tubulins can even differ from one part of the cell in an embryo to another and during mitosis. This is suggestive of a specialized role of tubulin isoform in cell splitter effects. In particular, the role of tubulin isoforms was outlined by Gordon (1999) concluding with the following observations:

1. Different tubulin isoforms are involved in cortical rotation.

2. Tubulins in nematodes may vary between cell types.

3. The origin of tubulin isoforms coincides with the origin of new cell types and specifically with their cell state splitters (Kube-Granderath and Schliwa, 1998).

4. Some tubulin isoforms nevertheless occur in multiple tissues (Lu, 1998).

5. There is a highly probable and very intriguing role of tubulin isoforms in brain complexity and its origin.

6. Tubulin isoforms could be used to distinguish otherwise histologically equivalent cell types.

Two hypotheses regarding multiple tubulin gene sequences have been advanced, both of which may have some factual basis. First, individual tubulin genes may encode functionally divergent polypeptides conferring unique properties to the final MT that is polymerized by the cell. Alternatively, multiple polypeptides may themselves be functionally equivalent but represent the products of duplicated genes that have evolved to possess different regulatory sequences for activation of transcription during alternative

programs of differentiation.

While this paper attempted to shed light on the gross biophysical features distinguishing between different tubulins, we have been able to relate such differences only to the presence of the major sub-families. Very subtle differences in the amino-acid composition within each sub-family while still elude our ability to relate structure and function precisely, may contain significant changes in tubulin's functional domains. Our quest to uncover these relationships is still ongoing.

References

ADOUTTE, A, CLAISSE, M AND CANCE, J. (1984) Tubulin evolution: an electrophoretic and immunological analysis, *Orig. Life*. 13:177-82.

BAKER, N.A., SEPT, D., JOSEPH, S., HOLST, M.J. AND MCCAMMON, J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA*. 98:10037-41.

BANERJEE, A. AND KASMALA, L. T. (1998). Differential assembly kinetics of α-tubulin isoforms in the presence of paclitaxel. *Biochem Biophys Res Commun*, 245: 349-51.

BANERJEE, A. (2002). Coordination of posttranslational modifications of bovine brain α-tubulin. polyglycylation of δ2 tubulin. *J. Biol. Chemistry*, 277:46140-46144.

BOECKMANN, B., BAIROCH, A., APWEILER, R., BLATTER, M.C., ESTREICHER, A., GASTEIGER, E., MARTIN, M. J., MICHOUD, K., O'DONOVAN, C., PHAN, I., PILBOUT, S. AND SCHNEIDER M. (2003), The SWISS-PROT Protein Knowledgebase and its Supplement TrEMBL in 2003, *Nucleic Acids Res,* 31: 365.

BROWN JA. AND TUSZYNSKI JA. (1999) A review of the ferroelectric model of microtubules. *Ferroelectrics*. 220:141-155.

BURNS, R.G. (1991) α-, β- and γ-Tubulins: Sequence comparisons and structural constraints. *Cell Motil. Cytoskeleton*. 20: 181–189.

BURNS, R.G. and SURRIDGE, C.D. (1994). Tubulin: Conservation and structure. In Microtubules (eds. J.S. Hyams and C.W. Lloyd), pp. 3–32. Wiley-Liss, New York.

CARPENTER, E. J., HUZIL, T.J. LUCHKO, T., TUSZYNSKI, J.A. and LUDUENA, R.F. (2005) Homology Modeling of Tubulin Isotypes and their Influence on the Biophysical Properties of Microtubules, submitted to *Biophysical Journal*.

CHOTHIA, C. AND LESK A.M. (1986) The relation between the divergence of sequence and structure in proteins, *EMBO J.* 5: 823-826.

CLEVELAND, D. W. AND SULLIVAN, K. F. (1985) Molecular biology and genetics of tubulin. *Ann. Rev. Biochem.* 54: 331-365.

DELANO W., (2002). The PyMOL Molecular Graphics System.

DOWNING, K. H. AND NOGALES, E. (1998). Tubulin and microtubule structure. *Curr. Opin. Cell Biology*. 10:16-22.

DUAN, J. AND GOROVSKY, M. A. (2002). Both carboxy-terminal tails of α- and β-tubulin are essential, but either one will suffice. *Current Biology*. 12:313

DUSTIN, P. (1984) Microtubules (Springer-Verlag, Berlin).

DUTCHER, S.K. (2001). The tubulin fraternity: α to η. Curr. *Opinion Cell Biol.* 13: 49-54.

EDGCOMB, V.P., ROGER, A.J., SIMPSON, A. G. B. KYSELA D. T. and SOGIN M. L. (2001) Evolutionary Relationships Among «Jakobid» Flagellates as Indicated by α- and -Tubulin Phylogenies, *Molecular Biology and Evolution*. 18:514-522.

EDLIND, T. D., LI, J., VISVESVARA, G. S., VODKIN, M. H. MCLAUGHLIN, G. L. AND KATIYAR, S. K. (1996). Phylogenetic analysis of -tubulin sequences from amitochondrial protozoa. *Mol. Phylogenet. Evol.* 5:359–367.

ERICKSON, H.P. (1998). Atomic structures of tubulin and FtsZ. *Trends in Cell Biology*. 8:133-137.

ERICKSON, H.P. (1997). FtsZ, a tubulin homolog, in prokaryote cell division. *Trends Cell Biol.* 7:362-367.

ERICKSON, H. P. (2000). γ-tubulin nucleation: template or protofilament. *Nat Cell Biol.* 2(6), E93-96.

FITCH, W.M. (2000). Homology: A personal view on some of the problems. *Trends Genet.* 16: 227–231.

GAERTIG J., THATCHER T.H., MCGRATH K.E., CALLAHAN R.C. AND GOROVSKY M.A. (1993) Perspectives on tubulin isotype function and evolution based on the observation that Tetrahymena thermophila microtubules contain a single α- and -tubulin. *Cell Motil Cytoskeleton*. 25:243-53.

GORDON, R. (1999) The Hierarchical Genome & Differentiation Waves, (World Scientific, Singapore).

HUMPHREY, W., DALKE, A. and SCHULTEN, K., (1996) VMD - Visual Molecular Dynamics, *J. Molec. Graphics*, 14:33-38.

HYAMS, J.S. AND LLOYD C.W., eds. (1994) Microtubules, Wiley-Liss, New York.

INCLAN, Y. F. AND NOGALES, E. (2001). Structural models for the self-assembly and microtubule interactions of γ-, δ- and ε-tubulin. *J. Cell Science*. 114:413-422.

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 α-, β- and γ-tubulin genes. *Mol. Biol. Evol.* 16:1308–1313.

KEELING, P. J., DEANE, J. A. AND MCFADDEN, G. I. (1998). The phylogenetic position of α- and -tubulins from the Chlorarachnion host and Cercomonas (Cercozoa). *J. Eukaryot. Microbiol.* 45:561–570.

KEELING, P. J. AND DOOLITTLE, W. F. (1996). α-tubulin from early-diverging eukaryotic lineages and the evolution of the tubulin family. *Mol. Biol. Evol.* 13:1297–1305.

KEELING, P. J., LUKER, M. A. AND PALMER, J. D. (2000). Evidence from -tubulin phylogeny that microsporidia evolved from within the fungi. *Mol. Biol. Evol.* 17:23–31.

KESKIN, O., DURELL, S. R., BAHAR, I., JERNIGAN, R. L. AND COVELL, D. G. (2002). Relating molecular flexibility to function: a case study of tubulin. *Biophys. Journal*. 83:663-680.

KIKKAWA, M., SABLIN, E.P., OKADA, Y., YAJIMA, H., FLETTERICK, R.J. and HIROKAWA, N. (2001), Switch-Based Mechanism of Kinesin Motors, *Nature*. 411: 439.

KIRSCHNER, M. AND SCHULZE, E. (1986). Morphogenesis and the control of microtubule dyanmics in cells. *J Cell Sci Suppl.* 5, 293-310.

KORADI, R. BILLETER, M. and WUTHRICH, H. (1996) MOLMOL: A Program for Display and Analysis of Macromolecular Structures, *Journal of Molecular Graphics*. Feb.: 51–55.

KRAUHS, E., LITTLE, M., KEMPF, T., HOFER-WARBINEK, R., ADE, W. AND POSTINGL, H. (1981). Complete amino acid sequence of β-tubulin from porcine brain. *Proc. Natl. Acad. Sci. USA* 78: 4156–4160.

KUCHNIR FYGENSON, D. NEEDLEMAN, D. J. and SNEPPEN, K. (2004). Variability-based sequence alignment identifies residues responsible for functional differences in α and β tubulin, *Protein Science*. 13:25-31.

KUBE-GRANDERATH, E. AND SCHLIWA, M. (1998). Unusual tubulins for unusual cells. *Protist*. 149:123–126.

LEWIS, S. A. and COWAN, N. J. (1990). Microtubule Proteins, chapter Tubulin Genes: Structure, Expression, Regulation, pages 37-66. CRC Press, Inc.,Boca Raton, Florida.

LI, H., DEROSIER, D. J., NICHOLSON, W. V., NOGALES, E. AND DOWNING, K. H. (2002). Microtubule structure at 8 A resolution. *Structure*. 10:1317- 1328.

LINDAHL, E., HESS, B. AND VAN DER SPOEL, D. (2001). GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Mod.* 7: 306-317.

LITTLE, M., KRAUHS, E. AND PONSTINGL, H. (1981). Tubulin sequence conservation. *Biosystems*. 14: 239–246.

LITTLE, M. and SEEHAUS, T. (1988). Comparative analysis of tubulin sequences.

Comparative Biochemistry & Physiology - B: Comparative Biochemistry. 90:655-670.

LOWE, J., LI, H., DOWNING, K. H. AND NOGALES, E. (2001). Refined structure of α β-tubulin at 3.5 a resolution. *J Mol. Biol.* 313:1045-1057.

LOWE, J. AND AMOS, L. A. (1998). Crystal structure of the bacterial cell-division protein FtsZ. *Nature.* 391: 203-06.

LU, Q., MOORE, G. D., WALSS, C. AND LUDUENA, R. F. (1998). Structural and functional properties of tubulin isotypes. *Advances in Structural Biology.* 5:203-227.

LUDUENA, R. F. (1998). Multiple forms of tubulin: different gene products and covalent modifications. *Int Rev Cytol.* 178: 207-275.

MAKRIDES, V., SHEN, T. E., BHATIA, R., SMITH, B. L., THIMM, J., LAL, R. AND FEINSTEIN, S. C. (2003). Microtubule-dependent oligomerization of tau: implications for physiological tau function and tauopathies. *J. Biol. Chem.* 278: 33298-33304.

MARTI-RENOM, M.A. STUART, A. FISER, A. SANCHEZ, R. MELO, F. and SALI. A. (2000) Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 29, 291-325.

MIRNY, L.A. and SHAKHNOVICH, E.I. (1999). Universally conserved positions in protein folds: Reading evolutionary signals about stability, folding kinetics and function. *J. Mol. Biol.* 291: 177–196.

MITCHISON, T. and KIRSCHNER, M. (1984). Dynamic instability of microtubule growth. *Nature.* 312: 237–242.

NOGALES, E. (2001). Structural insights into microtubule function. Ann. Rev. *Biophys. Biomol. Struct.* 30: 397–420.

NOGALES E., DOWNING K.H., AMOS L.A. AND LOWE, J. (1998) Tubulin and FtsZ form a distinct family of GTPases. *Nat Struct Biol.* 6:451-458.

NOGALES, E., WOLF, S. G. AND DOWNING, K. H. (1998). Structure of the α tubulin dimer by electron crystallography. *Nature.* 391: 199-203.

PANDA, D., MILLER, H. P., BANERJEE, A., LUDUENA, R. F. AND WILSON, L. (1994). Microtubule dynamics in vitro are regulated by the tubulin isotype composition. *Proc. Natl. Acad. Sci. USA.* 91:11358-11362.

POSTINGL, H., KRAUHS, E., LITTLE, M. AND KEMPF, T. (1981) Complete amino acid sequence of α-tubulin from porcine brain. *Proc. Natl. Acad. Sci. USA.* 78: 2757–2761.

PRIEL, A., WOOLF, N. AND J.A. TUSZYNSKI, (2005), Molecular Dynamics of C-termini in Tubulin: Implications for Transport to Active Synapses, submitted to *Eur. Biophys. J.*

RAFF, E.C., FACKENTHAL, J.D., HUTCHENS, J.A., HOYLE, H.D. AND TURNER, F.R. (1997). Microtubule Architecture Specified by a β -Tubulin Isoform, *Science.* 275: 70.

RAVELLI R. B. G., GIGANT, B., CURMI, P. A., JOURDAIN, I., LACHKAR, S., SOBEL, A. AND KNOSSOW, M. (2004) Insight into tubulin regulation from a complex with colchicine and a stathmin-like domain*. Nature.* 428: 198-202.

REDEKER, V., ROSSIER, J. AND FRANKFURTER, A. (1998). Posttranslational modifications of the c-terminus of α-tubulin in adult rat brain: α 4 is glutamylated at two residues. *Biochemistry.* 37:14838-14844.

RICHARDS, K. L. ANDERS, K. R., NOGALES, E., SCHWARTZ, K., DOWNING, K. H. AND BOTSTEIN, D. (2000). Structure-function relationships in yeast tubulins. *Molecular Biology of the Cell.* 11:1887-1903.

RODIONOV, M.A. AND BLUNDELL, T.L. (1998) Sequence and structure conservation in a protein core. *Proteins.* 33: 358–366.

ROGER, A. J. (1996) Studies on the phylogeny and gene structure of early-branching eukaryotes. Ph.D. thesis, Dalhousie University, Halifax, Nova Scotia, Canada

SACKETT, D. L. (1995). Subcellular Biochemistry, volume 24 of Proteins: Structure, Function and Engineering, chapter Structure and Function in the Tubulin Dimer and the Role of the Acidic Carboxyl Terminus. Plenum Press, New York.

SANCHEZ, R. AND SALI, A. (2000) Comparative protein structure modeling, *Methods Mol Biol.* 143: 97-129.

SARKAR, T., MANNA, T., BHATTACHARYYA, S., MAHAPATRA, P., PODDAR, A., ROY, S., PENA, J., SOLANA, R., TARAZONA, R. AND BHATTACHARYYA, B. (2001). Role of the carboxy-termini of tubulin on its chaperone-like activity. *Proteins.* 44:262-269.

SEPT, D., BAKER, N. A. AND MCCAMMON, J. A. (2003). The physical basis of microtubule structure and stability. *Protein Sci.* 12: 2257-61.

STRACKE, R. BOEHM, K.J. WOLLWEBER, L. UNGER E. and TUSZYNSKI J.A. (2002) Analysis of the migration behavior of single microtubules in electric fields, *Biochemistry and Biophysics Research Communications.* 293: 602-609.

THOMPSON, J. D., HIGGINS, D. G. AND GIBSON, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673-80.

TRPISOVA B. AND TUSZYNSKI J.A. (1997) Possible link between guanosine 5' triphosphate hydrolysis and solitary waves in microtubules. *Physical Review A.* 55:3288-3305.

UNGER, E., BOHM, K. J. AND VATER, W. (1990). Structural diversity and dynamics of microtubules and polymorphic tubulin assemblies. *Electron Microsc Rev.* 3: 355-95.

VAUGHAN S., ATTWOOD T., NAVARRO M., SCOTT V., MCKEAN P. AND GULL K. (2000) New tubulins in protozoal parasites. *Curr. Biol.* 10: R258-R259.

WANG, Z. and SHEETZ, M. P. (2000). The c-terminus of tubulin increases cytoplasmic dynein and kinesin processivity. *Biophysical Journal.* 78:1955-1964.

WILSON, P.G. AND BORISY, G.G. (1997). Evolution of the Multi-Tubulin Hypothesis *BioEssays.* 19: 451-454.

WRIGGERS W. and SCHULTEN, K. (1998) Nucleotide-Dependent Movements of the Kinesin Motor Domain Predicted by Simulated Annealing, *Biophys J.* 75: 646.

Appendix A

Tubulin sequences

Achlya klebsiana	β TBB_ACHKL
Acremonium coenophialum	β TBB_ACRCO
Ajellomyces capsulata (Histoplasma capsulatum)	α TBA_AJECA; β TBB_AJECA
Anemia phyllitidis (Fern)	α TBA1_ANEPH, TBA2_ANEPH; β TBB1_ANEPH, TBB2_ANEPH, TBB3_ANEPH; γ TBG_ANEPH
Arabidopsis thaliana (Mouse-ear cress)	α TBA1_ARATH, TBA2_ARATH, TBA3_ARATH, TBA6_ARATH;
	β TBB1_ARATH, TBB2_ARATH, TBB4_ARATH, TBB5_ARATH, TBB6_ARATH, TBB7_ARATH, TBB8_ARATH, TBB9_ARATH;
	γ TBG2_ARATH
Aspergillus flavus	β TBB_ASPFL
Aspergillus parasiticus	β TBB_ASPPA
Avena sativa (Oat)	α TBA_AVESA; β TBB1_AVESA
Babesia bovis	β TBB_BABBO
Blepharisma japonicum	α TBA_BLEJA
Bombyx mori (Silk moth)	α TBA_BOMMO; β TBB_BOMMO
Botrytis cinerea (Botryotinia fuckeliana)	β TBB_BOTCI
Brugia pahangi	β TBB1_BRUPA
Caenorhabditis briggsae	β TBB7_CAEBR
Caenorhabditis elegans	α TBA2_CAEEL, TBA8_CAEEL; β TBB2_CAEEL, TBB4_CAEEL, TBB7_CAEEL; γ TBG_CAEEL
Candida albicans (Yeast)	α TBA_CANAL; β TBB_CANAL; γ TBG_CANAL
Cephalosporium acremonium (Acremonium chrysogenum)	β TBB_CEPAC
Chlamydomonas incerta	β TBB_CHLIN
Chlamydomonas reinhardtii	α TBA1_CHLRE, TBA2_CHLRE; β TBB_CHLRE; γ TBG_CHLRE
Chlorella vulgaris	α TBA_CHLVU
Chondrus crispus (Carragheen)	β TBB1_CHOCR
Cicer arietinum (Chickpea) (Garbanzo)	β TBB_CICAR
Cochliobolus heterostrophus (Drechslera maydis)	γ TBG_COCHE
Colletotrichum gloeosporioides (Anthracnose fungus) (Glomerella cingulata)	β TBB1_COLGL, TBB2_COLGL
Colletotrichum graminicola (Anthracnose fungus) (Glomerella graminicola)	β TBB1_COLGR, TBB2_COLGR
Cyanophora paradoxa	β TBB1_CYAPA
Daucus carota (Carrot)	β TBB1_DAUCA, TBB2_DAUCA
Dictyostelium discoideum (Slime mold)	α TBA_DICDI; β TBB_DICDI
Drosophila erecta (Fruit fly)	β TBB2_DROER
Drosophila melanogaster (Fruit fly)	α TBA1_DROME, TBA2_DROME, TBA3_DROME, TBA4_DROME; β TBB2_DROME, TBB3_DROME; γ TBG2_DROME
Ectocarpus variabilis	β TBB5_ECTVR, TBB6_ECTVR
Eimeria tenella	β TBB_EIMTE
Eleusine indica (Goosegrass)	α TBA1_ELEIN, TBA2_ELEIN, TBA3_ELEIN; β TBB1_ELEIN, TBB2_ELEIN, TBB3_ELEIN, TBB4_ELEIN
Emericella nidulans (Aspergillus nidulans)	α TBA1_EMENI, TBA2_EMENI; β TBB1_EMENI, TBB2_EMENI; γ TBG_EMENI
Entamoeba histolytica	α TBA1_ENTHI; γ TBG_ENTHI
Epichloe typhina	β TBB_EPITY
Erysiphe graminis (subsp. hordei) (Grass mildew)	β TBB_ERYGR
Erysiphe pisi (Pea powdery mildew)	β TBB2_ERYPI
Euglena gracilis	α TBA_EUGGR; β TBB_EUGGR
Euplotes aediculatus	γ TBG_EUPAE
Euplotes crassus	β TBB_EUPCR; γ TBG2_EUPCR
Euplotes focardii	β TBB_EUPFO
Euplotes octocarinatus	α TBA_EUPOC; β TBB_EUPOC; γ TBG2_EUPOC
Euplotes vannus	α TBA_EUPVA
Gadus morhua (Atlantic cod)	β TBB1_GADMO
Gallus gallus (Chicken)	α TBA1_CHICK, TBA2_CHICK, TBA3_CHICK, TBA4_CHICK, TBA5_CHICK, TBA8_CHICK;
	β TBB1_CHICK, TBB2_CHICK, TBB3_CHICK, TBB4_CHICK, TBB5_CHICK, TBB6_CHICK, TBB7_CHICK
Geotrichum candidum (Oospora lactis)	β TBB1_GEOCN, TBB2_GEOCN
Giardia lamblia (Giardia intestinalis)	β TBB_GIALA
Gibberella fujikuroi (Bakanae and foot rot disease fungus) (Fusarium moniliforme)	β TBB_GIBFU
Glycine max (Soybean)	β TBB1_SOYBN, TBB2_SOYBN, TBB3_SOYBN
Haemonchus contortus (Barber pole worm)	α TBA_HAECO
Haliotis discus (Abalone)	β TBB_HALDI
Homarus americanus (American lobster)	α TBA1_HOMAM, TBA2_HOMAM, TBA3_HOMAM; β TBB1_HOMAM, TBB2_HOMAM
Homo sapiens (Human)	α TBA1_HUMAN, TBA2_HUMAN, TBA4_HUMAN, TBA6_HUMAN, TBA8_HUMAN;
	β TBB1_HUMAN, TBB2_HUMAN, TBB4_HUMAN, TBB5_HUMAN, TBBQ_HUMAN, TBBX_HUMAN;
	γ TBG1_HUMAN, TBG2_HUMAN; δ TBD_HUMAN; ε TBE_HUMAN
Hordeum vulgare (Barley)	α TBA1_HORVU, TBA2_HORVU, TBA3_HORVU; β TBB_HORVU
Leishmania mexicana	β TBB_LEIME
Leptomonas seymouri	α TBA_LEPSE
Lupinus albus (White lupine)	β TBB1_LUPAL, TBB2_LUPAL
Lymnaea stagnalis (Great pond snail)	β TBB_LYMST
Lytechinus pictus (Painted sea urchin)	α TBA_LYTPI; β TBB_LYTPI
Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm)	β TBB1_MANSE
Mus musculus (Mouse)	α TBA1_MOUSE, TBA2_MOUSE, TBA3_MOUSE, TBA6_MOUSE, TBA8_MOUSE; γ TBG1_MOUSE, TBG2_MOUSE
Mycosphaerella graminicola (Septoria tritici)	α TBA_MYCGR
Mycosphaerella pini (Dothistroma pini)	β TBB_MYCPJ
Naegleria gruberi	α TBA_NAEGR; β TBB_NAEGR
Neurospora crassa	α TBA1_NEUCR, TBA2_NEUCR; β TBB_NEUCR; γ TBG_NEUCR
Notophthalmis viridescens (Eastern newt) (Triturus viridescens)	α TBA_NOTVI
Notothenia coriiceps neglecta (Black rockcod) (Yellowbelly rockcod)	β TBB1_NOTCO
Octopus dofleini (Giant octopus)	α TBA_OCTDO; β TBB_OCTDO
Octopus vulgaris (Octopus)	α TBA_OCTVU

Onchocerca gibsoni	β TBB_ONCGI
Oncorhynchus keta (Chum salmon)	α TBA_ONCKE
Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri)	α TBAT_ONCMY
Oryza sativa (Rice)	α TBA1_ORYSA; β TBB1_ORYSA, TBB2_ORYSA, TBB3_ORYSA; γ TBG2_ORYSA
Oxytricha granulifera	α TBA_OXYGR
Paracentrotus lividus (Common sea urchin)	α TBA1_PARLI; β TBB_PARLI
Paramecium tetraurelia	β TBB1_PARTE
Patella vulgata (Common limpet)	α TBA2_PATVU
Pelvetia fastigiata	α TBA1_PELFA, TBA2_PELFA
Penicillium digitatum	β TBB_PENDI
Pestalotiopsis microspora	β TBB_PESMI
Phaeosphaeria nodorum (Septoria nodorum)	β TBB_PHANO
Physarum polycephalum (Slime mold)	α TBAD_PHYPO, TBAE_PHYPO, TBAN_PHYPO; β TBB1_PHYPO, TBB2_PHYPO
Physcomitrella patens (Moss)	γ TBG_PHYPA
Phytophthora cinnamomi	β TBB_PHYCI
Picea abies (Norway spruce) (Picea excelsa)	α TBA_PICAB
Pisum sativum (Garden pea)	α TBA1_PEA; β TBB1_PEA, TBB2_PEA, TBB3_PEA
Plasmodium berghei yoelii	α TBA_PLAYO
Plasmodium falciparum	α TBA_PLAFK; β TBB_PLAFA, TBB_PLAFK; γ TBG_PLAFO
Plasmodium falciparum (isolate K1 / Thailand)	α TBA_PLAFK; β TBB_PLAFK
Plasmodium falciparum (isolate NF54)	γ TBG_PLAFO
Pleurotus sajor-caju (Oyster mushroom)	β TBB_PLESA
Pneumocystis carinii	α TBA1_PNECA, TBAA_PNECA; β TBB_PNECA
Polytomella agilis	β TBB_POLAG
Porphyra purpurea	β TBB1_PORPU, TBB2_PORPU, TBB3_PORPU, TBB4_PORPU
Prunus dulcis (Almond) (Prunus amygdalus)	α TBA_PRUDU
Pseudopleuronectes americanus (Winter flounder) (Pleuronectes americanus)	β TBB_PSEAM
Rattus norvegicus (Rat)	β TBB1_RAT
Reticulomyxa filosa	γ TBG_RETFI
Rhynchosporium secalis	β TBB_RHYSE
Saccharomyces cerevisiae (Baker's yeast)	α TBA1_YEAST, TBA3_YEAST; β TBB_YEAST; γ TBG_YEAST
Schizophyllum commune (Bracket fungus)	α TBAA_SCHCO, TBAB_SCHCO; β TBB_SCHCO
Schizosaccharomyces japonicus	γ TBG_SCHJP
Schizosaccharomyces pombe (Fission yeast)	α TBA1_SCHPO, TBA2_SCHPO; β TBB_SCHPO; γ TBG_SCHPO
Solanum tuberosum (Potato)	β TBB1_SOLTU, TBB2_SOLTU
Sordaria macrospora	α TBA_SORMA
Strongylocentrotus purpuratus (Purple sea urchin)	β TBB_STRPU
Stylonychia lemnae	α TBA1_STYLE, TBA2_STYLE; β TBB_STYLE
Sus scrofa (Pig)	α TBA_PIG; β TBB_PIG
Tetrahymena pyriformis	α TBA_TETPY; β TBB_TETPY
Tetrahymena thermophila	α TBA_TETTH; β TBB_TETTH
Thalassiosira weissflogii (Marine diatom)	β TBB_THAWE
Torpedo marmorata (Marbled electric ray)	α TBA_TORMA
Toxoplasma gondii	α TBA_TOXGO; β TBB_TOXGO
Trichoderma viride	β TBB1_TRIVI,'TBB2_TRIVI
Triticum aestivum (Wheat)	α TBA_WHEAT; β TBB1_WHEAT, TBB2_WHEAT,—TBB3_WHEAT, TBB4_WHEAT, TBB5_WHEAT
Trypanosoma brucei rhodesiense	α TBA_TRYBR; β'TBB_TRYBR
Trypanosoma cruzi	α TBA_TRYCR; β TBB_TRYCR
Ustilago violacea (Smut fungus) (Microbotryum violaceum)	γ TBG_USTVI
Venturia inaequalis	β TBB_VENIN
Volvox carteri	α TBA1_VOLCA; β TBB1_VOLCA
Xenopus laevis (African clawed frog)	α TBA_XENLA; β'TBB2_XENLA, TBB4_XENLA; γ TBG_XENLA
Zea mays (Maize)	α TBA1_MAIZE,'TBA2_MAIZE, TBA3_MAIZE, TBA4_MAIZE, TBA5_MAIZE, TBA6_MAIZE; β TBB1_MAIZE, TBB2_MAIZE, TBB3_MAIZE, TBB4_MAIZE, TBB5_MAIZE, TBB6_MAIZE, TBB7_MAIZE,'TBB8_MAIZE; γ TBG1_MAIZE, TBG2_MAIZE, TBG3_MAIZE

Appendix B

Summary of computation results

Each of the 290 Swiss-Prot Sequences is listed below with some results. Fragmentary sequences are marked with an initial "F". Charges are reported in fundamental charge (electronic) units.

	Name	|M|	Net Charge	Volume (Å³)	Area (Å²)	Charge	Size	C-Terminal Tail Sequence
	TBA1_ANEPH	1.344	-22	43.723	46.120	-8	16	GAESGEGEEGDEGEEY
	TBA1_ARATH	1.574	-24	43.726	46.097	-9	15	GGEGAEDDDEEGDEY
F	TBA1_CHICK	1.904	-21	40.490	43.082	-8	16	GVDSVEGEGEEEGEEY
	TBA1_CHLRE	1.042	-21	43.643	45.934	-6	16	GAESAEGAGEGEGEEY
	TBA1_DROME	1.225	-22	44.031	46.824	-6	15	GMDSGDGEGEGAEEY
	TBA1_ELEIN	1.407	-24	43.861	46.749	-9	16	GAEFDEGEEGDEGDEY
	TBA1_EMENI	1.467	-24	44.070	46.434	-7	14	ASDSLEEEGEEVEY
	TBA1_ENTHI	730	-10	44.061	46.461	-3	14	ALDNTIEGESMTAQ
	TBA1_HOMAM	1.502	-22	44.167	46.824	-7	16	GMDSADGEDIEGGDEY
	TBA1_HORVU	1.321	-23	43.591	45.827	-8	15	GAEGADDEGDEGDDY
	TBA1_HUMAN	1.449	-24	44.250	47.174	-8	16	GVDSVEGEGEEEGEEY
	TBA1_MAIZE	1.323	-24	43.835	46.652	-9	16	GAEFDEGEDGDEGDEY
	TBA1_MOUSE	1.470	-24	44.263	47.102	-8	16	GVDSVEGEGEEEGEEY
	TBA1_NEUCR	970	-19	44.052	46.358	-5	13	AGDYNDVDVDAEY
	TBA1_ORYSA	1.625	-24	43.648	45.940	-8	15	GAEGADDENDDGEDY
	TBA1_PARLI	2.160	-25	44.184	46.804	-9	17	GVDSADAEGEEEEGDEY
	TBA1_PEA	1.344	-23	43.568	45.724	-8	17	GAESGDGDDDGLGEEEY
	TBA1_PELFA	1.681	-24	43.907	46.568	-9	18	GAETADGDGEEEEFGEEY
	TBA1_PNECA	1.352	-20	44.335	47.012	-6	14	GQDSIEGEIMEEEY
	TBA1_SCHPO	1.411	-22	44.895	47.968	-7	16	GQDSMDNEMYEADEEY
	TBA1_STYLE	1.531	-23	43.243	45.451	-7	15	GIEIVEGEGEEEGME
	TBA1_VOLCA	1.063	-21	43.630	45.981	-6	16	GAESAEGAGEGEGEEY
	TBA1_YEAST	1.083	-22	43.874	46.462	-5	11	GADSYAEEEF
F	TBA2_ANEPH	1.640	-15	35.461	37.487	-6	13	AAEGVDEPEGDDY
	TBA2_ARATH	1.632	-23	43.766	46.803	-9	15	GAEGGDDEDDEGEEY
	TBA2_CAEEL	1.373	-22	43.891	46.319	-7	15	GADSNEGGEEEGEEY
	TBA2_CHICK	1.140	-25	43.774	46.365	-6	14	ATDLFEDENEAGDS
	TBA2_CHLRE	1.146	-22	43.601	45.661	-6	16	GAESAEGAGEGEGEEY
	TBA2_DROME	1.300	-21	44.117	46.892	-6	14	GIDSTTELGEDEEY
	TBA2_ELEIN	1.444	-21	43.843	45.941	-8	12	GAEVEEDDEEEY
	TBA2_EMENI	1.570	-21	44.173	46.890	-6	14	AGDSLDMEGEEAEY
	TBA2_HOMAM	1.387	-20	44.252	47.078	-8	16	GIDTADGEDDEEANDY
	TBA2_HORVU	1.542	-24	43.706	46.254	-9	16	GAEFDDGEDGDEGDEY
	TBA2_HUMAN	1.494	-23	44.046	46.631	-7	15	GVDSVEAEAEEGEEY
	TBA2_MAIZE	1.353	-24	43.670	46.060	-9	16	GAEFDEGEEGDDGDEY
	TBA2_MOUSE	1.526	-24	44.189	46.902	-8	16	GVDSVEGEGEEEGEEY
	TBA2_NEUCR	1.189	-22	43.970	46.398	-6	14	AADSMEGEDVEAEY
	TBA2_PATVU	1.727	-24	44.206	46.802	-8	17	GVDSVEGEGEEEGGEEY
	TBA2_PELFA	1.823	-25	43.972	46.730	-9	18	GAETAEGEGEEEDFGEEY
	TBA2_SCHPO	1.057	-23	44.414	47.084	-6	14	GQDSMEVDYMEEEY
	TBA2_STYLE	1.868	-21	43.463	45.795	-7	15	GIETAEGEGEEEGME
	TBA3_ARATH	1.455	-23	43.768	46.341	-9	15	GAEGGDDEEDEGEDY
F	TBA3_CHICK	1.425	-11	31.862	34.077	-3	12	GRNSADGGEFEE
	TBA3_DROME	1.126	-22	44.025	46.744	-6	15	GMDSGDGEGEGAEEY
	TBA3_ELEIN	1.341	-24	43.623	45.927	-8	15	GAEGADDEGDEGEDY
	TBA3_HOMAM	1.742	-24	44.024	46.425	-8	16	GVDSADAEGEEEGEEY
	TBA3_HORVU	1.229	-24	43.774	46.615	-9	16	GAEFDEGEDGDEGDEY
	TBA3_MAIZE	1.600	-20	43.523	45.861	-8	15	AAEGGSDDGDEEEEY
	TBA3_MOUSE	1.494	-23	44.046	46.631	-7	15	GVDSVEAEAEEGEEY
	TBA3_YEAST	1.380	-20	43.773	46.394	-3	9	GADSYAEEF
F	TBA4_CHICK	1.816	-19	31.763	34.085	-7	12	GTDSFEDENDEE
	TBA4_DROME	1.362	-18	44.750	46.802	-8	16	GLDNAEEGGDEDFDEF
	TBA4_HUMAN	1.486	-24	44.006	46.802	-8	13	GIDSYEDEDEGEE
F	TBA4_MAIZE	786	-13	5.653	6.442	-9	15	AEFDEGEDGDEGDEY
	TBA5_CHICK	1.423	-24	44.001	46.787	-8	13	GLDSYEDEEEGEE
	TBA5_MAIZE	1.259	-24	43.664	46.181	-8	15	GAEGADDEGDEGDDY
	TBA6_ARATH	1.472	-23	43.549	45.981	-9	15	GAEGGDDEDDEGEEY
	TBA6_HUMAN	1.230	-23	44.020	46.936	-7	14	GADSDEGDEGEEY
	TBA6_MAIZE	1.304	-24	43.616	45.962	-8	15	GAEGADDEGDEGDDY
	TBA6_MOUSE	1.262	-23	44.005	46.878	-7	14	GADSAEGDDEGEEY
	TBA8_CAEEL	1.483	-21	44.092	46.452	-6	14	SRDTADLEEENDEF
F	TBA8_CHICK	1.420	-17	31.941	34.148	-7	14	GTDSMDGEDEGEEY
	TBA8_HUMAN	1.326	-24	44.109	46.847	-7	14	GTDSFEEENEGEEF
	TBA8_MOUSE	1.109	-23	44.094	46.772	-7	14	GTDSFEEENEGEEF
F	TBA_AJECA	800	-11	40.916	42.810	0	0	
F	TBAA_PNECA	701	0	21.164	22.926	0	0	
	TBAA_SCHCO	1.160	-20	43.529	46.458	-6	14	GMDSADAEEEAAEY
	TBA_AVESA	672	-17	43.193	45.318	-3	12	GAEGADDEVTRG
F	TBA_BLEJA	135	-17	4.939	5.727	-8	18	GIETAEAEGEEEGYGEEL
	TBA_BOMMO	1.189	-23	44.003	46.588	-6	15	GMDSAEGEGEGAEEY
	TBAB_SCHCO	1.389	-17	43.480	46.447	-6	13	GTDSAEEEGEY
	TBA_CANAL	1.610	-20	43.827	46.383	-6	12	GTDSFPEEEEEY
	TBA_CHLVU	1.279	-23	43.800	46.512	-8	16	GAESAEADGEDEGEEY
	TBA_DICDI	988	-15	44.898	47.487	-6	13	SASTEGEEQEEEY
	TBAD_PHYPO	1.513	-22	43.833	46.203	-6	15	GAESSEAGGDEEGEY

Name			Net	Volume	Area	C-Terminal Tail		
		|M|	Charge	(Å³)	(Å²)	Charge	Size	Sequence
	TBA_PLAFK	1.639	-22	44.160	46.869	-8	18	GIESNEAEGEDEGYEADY
F	TBA_PLAYO	1.414	-12	19.400	20.787	-7	15	GIETNDGEGEDEGYE
	TBA_PRUDU	1.362	-23	43.612	46.258	-8	15	GAESAEGEDDEGDDY
	TBA_SORMA	1.057	-23	43.781	46.691	-6	14	AADSMEGEEVEAEY
	TBA_TETPY	1.189	-21	43.728	46.142	-6	14	GIETAEGEGEEEGY
	TBA_TETTH	1.116	-21	43.758	46.335	-6	14	GIETAEGEGEEEGY
	TBAT_ONCMY	1.411	-23	44.043	46.641	-7	15	GVDSVEGEAEEGEEY
	TBA_TORMA	1.334	-24	44.319	47.358	-8	16	GVDSVEGEGEEEGEEY
	TBA_TOXGO	1.562	-23	44.098	46.709	-8	18	GIETAEGEGEEEGYGDEY
	TBA_TRYBR	1.334	-24	43.868	46.477	-8	16	GAESADMDGEEDVEEY
	TBA_TRYCR	1.425	-25	43.758	46.172	-8	16	GAESADMEGEEDVEEY
	TBA_WHEAT	1.561	-24	43.805	46.562	-9	16	GAEFDEGEDGDEGDEY
	TBA_XENLA	1.313	-23	43.943	46.479	-7	14	GADSADAEDEGEEY
F	TBB1_ANEPH	1.688	-21	43.331	45.950	-8	18	QDATAEREGEYEEDYDEA
	TBB1_ARATH	3.129	-27	43.752	46.147	-12	21	QDATADEEDEYDEEEEQVYES
	TBB1_AVESA	3.182	-25	38.101	41.157	-13	23	QDATADEEGEYEDEEEDLQAEDM
	TBB1_BRUPA	1.961	-26	43.981	46.705	-11	23	QDATADEEGDLQEGESEYIEQEE
	TBB1_CHICK	2.149	-25	43.815	46.865	-11	20	QDATADEQGEFEEEGEEDEA
	TBB1_CHOCR	1.895	-27	43.978	45.919	-13	26	GEATADGVEGYEEEGYENDHPEDDEE
	TBB1_COLGL	3.581	-24	43.617	45.527	-13	22	QDAGMDDDEAEEAYEEEEPVEE
	TBB1_COLGR	2.349	-24	43.341	45.418	-12	22	QDAGMDDEYGEEYEDEAPAEEE
	TBB1_CYAPA	2.061	-25	43.704	46.639	-11	22	QDATAEEEGEGDEEEAEGEAAA
	TBB1_DAUCA	3.080	-17	31.338	33.360	-12	22	QDATADEEEYYEDEEEEEAQGM
	TBB1_ELEIN	3.102	-26	43.750	46.610	-12	20	QDATAEDEEEYEDEEEEMAA
	TBB1_EMENI	2.478	-23	43.751	46.675	-11	22	QDASISEGEEEYAEEEIMEGEE
	TBB1_GADMO	1.950	-25	43.818	47.122	-11	20	QDATAEEEGEFEEEGEEELA
	TBB1_GEOCN	1.914	-24	43.809	46.274	-11	22	QNATVDDEDMEYEDELPLEDEM
	TBB1_HOMAM	3.186	-24	44.266	45.948	-11	22	QEATADDEAEFEEEGEVEGEYA
	TBB1_HUMAN	2.504	-25	43.765	46.854	-11	19	QDATAEEEEDFGEEAEEEA
	TBB1_LUPAL	3.531	-25	43.898	46.734	-12	22	QDATADEDGYEYEDEEEEIGEEA
	TBB1_MAIZE	2.840	-25	43.777	46.781	-12	21	QDATADEEGEYEDEEEGDLQD
	TBB1_MANSE	1.831	-25	44.083	46.838	-12	22	QEATADEDAEFDEEQEQEIEDN
	TBB1_NOTCO	3.004	-25	43.698	46.443	-11	21	QDATAEEEGEFEEEGEYEDGA
	TBB1_ORYSA	2.216	-25	43.757	46.832	-12	19	QDATADEEYEDEEEEAEAE
	TBB1_PARTE	2.083	-24	43.491	46.266	-9	17	QDATAEEEGEFEEEGEQ
F	TBB1_PEA	3.760	-26	44.209	46.988	-13	25	QDATADEDEYGEEEGDEEEYGQHDI
	TBB1_PHYPO	1.794	-23			-9	42	QDATIDDEEGGEEEEGGAEEEARQRKHYVIDYVPSVCVILIR
F	TBB1_PORPU	2.434	-28	43.887	47.046	-11	32	EAATVEGEEEEDAYAEGAVVNGDQSYEDQYAA
	TBB1_RAT	3.037	-25	43.856	46.824	-11	20	QDATADEQGEFEEEEGEDEA
	TBB1_SOLTU	3.446	-26	43.921	45.964	-12	23	QDATADDEEEEYDDEAADDHQYES
	TBB1_SOYBN	2.056	-22	43.717	46.392	-10	20	QDATAVDDHEDEDEDEAMAA
	TBB1_TRIVI	1.906	-21	43.239	45.387	-10	19	QDASADDGEEYEEDAPMEE
	TBB1_VOLCA	2.371	-24	43.623	46.596	-10	18	QDASAEEEGEFEGEEEEN
	TBB1_WHEAT	2.331	-25	44.053	47.377	-12	20	QDATADEEGEYEEEEELEQE
	TBB2_ANEPH	1.809	-18	40.451	43.580	-8	18	QDATAEPEGXYEEDYDEA
	TBB2_ARATH	4.027	-27	44.380	47.533	-14	25	QDATADEEGDYEDEEEGEYQQEEEY
	TBB2_CAEEL	3.151	-24	44.042	46.516	-10	25	QEATAEDDVDGYAEGEAGETYESEQ
	TBB2_CHICK	2.944	-24	43.791	46.642	-11	20	QDATADEQGEFEEEGEEDEA
	TBB2_COLGL	2.737	-24	43.772	46.801	-13	22	QDAGVDEEEEEYEEEAPLEEEV
	TBB2_COLGR	2.964	-24	43.776	46.726	-13	22	QDAGVDEEEEEYEDDAPLEEEV
	TBB2_DAUCA	1.849	-25	43.469	46.734	-9	19	QDATAEEDDYDDGEGSTGD
	TBB2_DROER	3.591	-25	43.757	46.469	-12	21	QEATADEEGEFDEDEEGGGDE
	TBB2_DROME	3.031	-26	43.647	46.257	-12	21	QEATADEEGEFDEDEEGGGDE
	TBB2_ELEIN	2.516	-26	44.115	47.287	-13	23	QDATADEDGEYEDELDGQEEEDM
	TBB2_EMENI	2.477	-22	43.740	46.549	-10	24	QEATVSDGEGAYDAEEGEAYEQEE
	TBB2_ERYPI	2.621	-22	43.844	46.800	-11	22	QDASISEGEEEYEEEQQLENEE
	TBB2_GEOCN	2.995	-28	44.193	46.317	-12	26	QNAGVDEDEELMDHEEYADEGVEDFN
	TBB2_HOMAM	2.078	-24	44.467	45.943	-12	22	QEATADDEAEFEEEGEVEGEYD
	TBB2_HUMAN	2.311	-25	43.874	47.064	-11	20	QDATAEEEGEFEEEAEEEVA
	TBB2_LUPAL	4.016	-26	44.007	46.759	-13	23	QDATADEDGYEYEDEEEVGEEDA
	TBB2_MAIZE	2.952	-23	43.628	46.573	-10	19	QDATADEEADYEEEAAAE
	TBB2_ORYSA	2.060	-25	44.025	47.077	-12	22	QDATADEEGEYEDEEQQEADDM
	TBB2_PEA	3.660	-28	44.120	47.264	-14	24	QDATAEEDEYEEEEEDYHQHDEM
	TBB2_PHYPO	3.796	-24	44.198	47.050	-12	29	QEASVDDEAMEDDAEAEGGAGQNEAVEEF
	TBB2_PORPU	2.073	-27	41.546	44.677	-13	23	QDATAEEEGEYDEDEDDEGGDYA
	TBB2_SOLTU	2.778	-26	44.047	46.135	-12	24	QDATADDEEEYDDEAADDHHQYES
	TBB2_SOYBN	3.324	-26	44.355	47.559	-14	24	QDATADEYEEEEEEEEFAQHDM
	TBB2_TRIVI	2.655	-24	43.739	46.060	-12	21	QEAGIDEEEEYEDEAPMEAEE
	TBB2_WHEAT	3.606	-27	43.865	46.565	-13	22	QDATSDEEGEYEDEDQEPEEDM
	TBB2_XENLA	2.320	-24	43.640	46.526	-10	18	QDATADEQGEFEEEEDEA
F	TBB3_ANEPH	2.202	-9	24.028	25.945	-8	18	QDASAEPXXEQEEDYEEA
	TBB3_CHICK	3.524	-26	43.756	46.491	-12	20	QDATAEEEGEFEEEAEEEAE
	TBB3_DROME	3.220	-23	44.396	46.320	-10	23	QEATADEFDPEVNQEEVEGDCI
	TBB3_ELEIN	3.004	-27	43.974	47.142	-13	21	QDATAEEEEEYDDDEEEEVAA
	TBB3_MAIZE	3.339	-25	43.486	46.041	-12	20	QDATAEEYDEEEQDGEEEHD
	TBB3_ORYSA	3.258	-27	43.797	46.374	-13	22	QDATADDEEEDYGDEEEDEVAA
F	TBB3_PEA	3.520	-27	43.323	46.649	-14	24	QDATADEEGEYEDEEEEEPEHGYE
	TBB3_PORPU	3.074	-26	43.530	46.185	-13	20	QEASADDEADEFDEEEGDEE
F	TBB3_SOYBN	1.142	-9	40.339	43.199	0	0	
	TBB3_WHEAT	3.487	-28	43.671	46.343	-13	20	QDATADEEEEYDEEEEEEAA
	TBB4_ARATH	2.867	-25	43.751	46.536	-11	19	QDATAGEEEYEEEEEEYET
	TBB4_CAEEL	2.493	-24	43.684	46.649	-10	19	QEATADDEGEFDEHDQDVE
	TBB4_CHICK	3.244	-24	44.049	46.491	-11	24	QDATAEEEGEMYEDDEEEESEQGAK
	TBB4_ELEIN	2.724	-25	43.906	46.994	-12	21	QDATADEEGDYEDEDEALHDE
	TBB4_HUMAN	2.494	-25	44.223	47.074	-11	25	QDATAEEEGEMYEDDEEESEAQGPK
	TBB4_MAIZE	2.580	-24	43.757	46.284	-11	20	QDATAEEYEEEEHDGEEEHA

	Name	|M|	Net Charge	Volume (Å³)	Area (Å²)	Charge	Size	C-Terminal Tail Sequence
	TBB8_ARATH	3.721	-25	44.295	47.021	-13	24	QDATADEEEGYEYEEDEVEVQEEQ
	TBB8_MAIZE	2.389	-25	43.822	47.070	-11	20	QDATADEEAEYEDEEAIQDE
	TBB9_ARATH	2.221	-27	43.541	46.710	-12	19	QDATVGEEEYEEDEEEEEA
	TBB_ACHKL	3.371	-27	43.582	46.093	-12	21	QDATAEEEGEFDEDEEMDEMM
	TBB_ACRCO	3.306	-25	44.030	47.131	-13	23	QDAGIDEEEEEYEEEAPVDEPLE
	TBB_AJECA	2.016	-17	43.900	46.112	-6	19	QDASISEGEDEYFDYAWAM
	TBB_ASPFL	3.877	-23	43.818	46.347	-12	23	QDASISEGEEEYLEEEEPLEHEE
	TBB_ASPPA	3.766	-23	43.974	46.808	-12	23	QDASISEGEEEYLEEEEPLEHEE
	TBB_BABBO	1.296	-22	43.390	46.335	-7	16	QEATIDDDADDMVNDY
	TBB_BOMMO	3.873	-25	44.236	47.190	-14	25	QDATIDQEFEDEEEVEEQNDDSDEQ
	TBB_BOTCI	2.912	-22	43.734	46.687	-11	22	QDASISEGEEEYYEEEVPIEGEE
	TBB_CANAL	2.998	-27	43.650	45.896	-11	24	QEASIDEEELEYADEIPLEDAAME
	TBB_CEPAC	3.089	-24	43.756	46.590	-13	22	QDAGIDEEEEEYEEELPLEGEE
	TBB_CHLIN	2.565	-24	43.441	46.047	-10	18	QDASAEEEGEFEGEEEEA
	TBB_CHLRE	2.107	-24	43.581	46.513	-10	18	QDASAEEEGEFEGEEEEA
	TBB_CICAR	3.368	-26	44.094	46.346	-13	22	QDAIAEEEDEYEEEGEEQYDEQ
	TBB_DICDI	2.376	-25	44.756	46.369	-12	25	SNQETEEDGGEYQEEHEEHEEQAEN
	TBB_EIMTE	2.763	-24	43.849	46.691	-11	24	QDATAEEEGEFDEEEGVMDAEGAA
	TBB_EPITY	2.713	-24	43.982	47.088	-13	23	QDAGIDEEEEEYEEEAPVDEPLE
	TBB_ERYGR	3.174	-21	43.521	46.022	-10	21	QEASISEGEEEYPEEVSNEEE
	TBB_EUGGR	2.075	-28	43.339	45.941	-11	20	QDATVEEEGEFDEEEDVEQY
	TBB_EUPCR	2.969	-26	43.733	46.318	-11	21	QDATAEEEGEYVEDEDEMDGM
	TBB_EUPFO	2.381	-23	43.737	46.622	-11	19	QDATAEEEGEFDDEEEMDV
	TBB_EUPOC	2.789	-25	43.474	46.099	-10	19	QDATAEEEGEMDEEEGAME
	TBB_GIALA	2.850	-24	43.973	47.134	-12	21	QEAGVDEGEEFEEEEDFGDEQ
	TBB_GIBFU	3.764	-24	43.718	46.451	-13	21	QDAGIDEEEEEYEEELPEGEE
F	TBB_HALDI	1.426	-5	33.790	36.441	0	0	
	TBB_HORVU	3.816	-27	43.828	46.711	-13	22	QDATADEEGEYEDEDQEAEDDM
	TBB_LEIME	2.182	-25	43.641	46.103	-10	18	QDATVEEEGEFDEEEEAY
F	TBB_LYMST	1.499	-16	11.211	12.653	-12	24	QDATAEDEGEFDEEEAEGEGQEYA
F	TBB_LYTPI	3.329	-13	17.813	19.768	-12	21	QDATAEEEGEFDEEEGDEEAA
	TBB_MYCPJ	3.541	-22	43.590	46.322	-11	22	QEASVSEGEEEYDEEAPLEGEE
	TBB_NAEGR	3.441	-26	44.385	47.524	-12	26	QDATAEEEGEFDENEGAEGEEQPADY
	TBB_NEUCR	3.951	-24	43.679	46.401	-13	22	QDAGVDEEEEEYEEEAPLEGEE
	TBB_OCTDO	2.060	-23	44.106	46.618	-8	22	QEARSTDSDEYDNEEYYNQQEE
	TBB_ONCGI	2.425	-21	43.866	46.451	-8	19	QDASADDELNETIEQAETE
	TBB_PARLI	3.785	-26	43.883	46.679	-13	22	QDATAEEEGEFDEEEEGDEEAA
	TBB_PENDI	2.715	-21	43.815	46.891	-10	22	QEASVSEGEEEYLAEDIVDEEV
	TBB_PESMI	2.968	-24	43.723	46.551	-13	21	QDAGVDEEEEEYEEEPLPEDE
	TBB_PHANO	2.991	-21	43.778	46.470	-11	22	QEASISEGEEEYDEEAPLEAEE
	TBB_PHYCI	2.274	-24	43.659	46.213	-9	19	QDGTAEEEGEFDEDEEWMR
	TBB_PIG	2.481	-25	43.854	47.042	-11	20	QDATADEQGEFEEEGEEDEA
	TBB_PLAFA	2.213	-28	43.732	46.620	-11	20	QDATAEEEGEFEEEEGDVEA
	TBB_PLAFK	2.142	-27	43.685	46.607	-11	20	QDATAEEEGEFEEEEGDVEA
	TBB_PLESA	3.119	-25	44.047	46.981	-12	21	QDATADEEEGEYEEEPAEEEQ
	TBB_PNECA	1.903	-22	43.093	45.489	-9	17	EIAGVDEEVELDDEIET
	TBB_POLAG	2.708	-24	43.429	45.899	-10	18	QDASAEEEGEFGEEEEN
	TBB_PSEAM	2.177	-25	43.784	46.877	-11	20	QDATAEEEGEGEEEGDEEVA
	TBBQ_HUMAN	1.005	-18	42.440	44.994	-3	10	QDATAEGEGV
	TBB_RHYSE	3.601	-21	43.740	46.423	-11	22	QDASISEGEEEYEEEAPMEPEE
	TBB_SCHCO	2.109	-25	43.970	46.855	-11	20	QDATVEEEGEYEEEVIEDQE
	TBB_SCHPO	2.331	-27	43.447	45.914	-11	23	QEAGIDEGDEDYEIEEEKEPLEY
F	TBB_STRPU	3.410	-18	29.276	31.526	-13	22	QDATAEEEGEFDEEEEGDEEAA
	TBB_STYLE	2.563	-24	43.277	45.763	-10	17	QDATAEDEEEMDEEQME
	TBB_TETPY	2.046	-24	43.558	46.340	-10	18	QDATAEEEGEFEEEEGEN
	TBB_TETTH	1.987	-25	43.553	46.384	-10	18	QDATAEEEGEFEEEEGEN
	TBB_THAWE	2.776	-23	43.338	45.963	-10	18	QDATADEEGEFDEDEMEG
	TBB_TOXGO	2.957	-27	43.888	46.653	-12	24	QDATAEEEGEFDEEEGEMGAEEGA
	TBB_TRYBR	1.836	-24	43.476	45.924	-9	17	QDATIEEEGEFDEEEQY
	TBB_TRYCR	1.960	-25	43.397	45.909	-9	17	QDATIEEEGEFDEEEQY
	TBB_VENIN	3.561	-22	43.550	46.257	-11	22	QEASVSEGEEEYDEEAPLEGEE
	TBBX_HUMAN	2.968	-24	43.586	46.372	-11	19	QDATAEEEDFGEEAEEEA
	TBB_YEAST	3.539	-31	44.569	47.246	-13	32	QEATVEDDEEVDENGDFGAPQNQDEPITENFE
	TBD_HUMAN	1.360	-5	44.651	45.580	0	0	
	TBE_HUMAN	855	-6			0	18	DATKNMPVQDLPRLSIAM
	TBG1_HUMAN	1.838	-10	44.645	45.231	-1	13	TRPDYISWGTQEQ
	TBG1_MAIZE	2.100	-10	46.059	46.111	-4	28	ESPDYIKWGMEDPGEANVVAALDSKLVV
	TBG1_MOUSE	1.995	-11	44.752	45.707	-1	13	TRPDYISWGTQEQ
	TBG2_ARATH	2.558	-10	46.599	47.774	-4	33	ESPDYIKWGMEDPGQLMTGEGNASGVADPKLAF
	TBG2_DROME	1.969	-6	44.800	45.402	-1	18	TQIDYPQWSPAVEASKAG
	TBG2_EUPCR	2.039	-15	45.633	46.882	-4	29	EKMDYINRGKDDEDMDYDPRAPPNFRPIE
	TBG2_EUPOC	2.242	-10	45.772	47.629	-5	29	EKMDYINWGSDDDDMQFDPREPPKFSNIQ
	TBG2_HUMAN	1.697	-13	44.707	45.747	-2	13	TQPDYISWGTQEQ
	TBG2_MAIZE	1.981	-12			-5	28	ESPDYIKWGMEDPGEANVAADLDSKLVV
	TBG2_MOUSE	1.447	-10	44.771	45.967	-1	13	TRPDYISWGTQEQ
	TBG2_ORYSA	1.972	-12	46.151	46.463	-4	28	ESPDYIKWGMEDAGEANVAAALDSKLVV
F	TBG3_MAIZE	2.091	-9	41.587	42.200	-4	28	ESPDYIKWGMEDPGEANVVAALDSKLVV
	TBG_ANEPH	2.793	-9	46.392	47.491	-3	30	ESADYIKWGMEDRSKTLSADGTMDLSLPSS
	TBG_CAEEL	1.420	-9	43.945	45.973	1	10	VQKDYLTRGL
	TBG_CANAL	2.711	-23			-13	38	KEITYLDDDDEDDLEDGDGGGGGNGNGYNNIDDADMGI
	TBG_CHLRE	2.019	-6	45.685	46.544	-1	13	ESADYIQRQMMAS
F	TBG_COCHE	1.096	-2	26.055	27.658	0	0	
	TBG_EMENI	2.009	-9	44.603	46.275	-4	17	ERENYLDPDAGKDEVGV
	TBG_ENTHI	1.255	-6	45.398	46.350	0	26	ESIEYSHYSYPIHKYLKGIAKEKIAE
	TBG_EUPAE	2.163	-10	45.767	47.109	-5	29	EKMDYINWGNDDDDMQFDPREPPKFSNIQ
	TBG_NEUCR	2.077	-9	45.255	46.778	-4	23	EDANYLNPELGENASADTDKRMA

Name	Net		Volume	Area	C-Terminal Tail		
	|M|	Charge	(Å³)	(Å²)	Charge	Size	Sequence
TBG_PHYPA	2.785	-8	46.549	47.782	-2	34	ESADYIKWGMEDRGKQVSGEGNTSGTVDSRVGAS
TBG_PLAFO	2.397	-7	45.179	46.542	-2	11	ERDDYFTNTYI
TBG_RETFI	2.039	-4	47.100	48.599	1	44	QTSDYINWGMKQQQSQISQKESSSLANENGNGANNKPGKSAMAL
TBG_SCHJP	1.912	-7	44.087	45.524	-2	8	EQPEYLSM
TBG_SCHPO	1.618	-8	43.930	45.423	-2	8	EDPNYLSL
TBG_USTVI	1.970	-10	45.915	47.039	-2	30	EKADYIDYGAGPGYVKGEDRREKGREAVEG
TBG_XENLA	1.913	-9	44.698	45.368	0	13	TRPDYISWGTQDK
TBG_YEAST	2.193	-30	45.778	47.349	-13	33	EQDSYLDDVLVDDENMVGELEEDLDADGDHKLV