

## Conformational Analysis of the Carboxy-Terminal Tails of Human $\beta$ -Tubulin Isoforms

Tyler Luchko,<sup>\*,†</sup> J. Torin Huzil,<sup>‡</sup> Maria Stepanova,<sup>†</sup> and Jack Tuszynski<sup>\*,‡</sup>

<sup>\*</sup>Department of Physics, University of Alberta, Edmonton, Alberta T6G 2G7, Canada; <sup>†</sup>National Institute for Nanotechnology NRC, Edmonton, Alberta T6G 2M9, Canada; and <sup>‡</sup>Department of Oncology, University of Alberta, Edmonton, Alberta T6G 1Z2, Canada

**ABSTRACT** Several isoforms of the structural protein tubulin have been characterized. Their expression offers a plausible explanation for differences regarding microtubule function. Although sequence variation between tubulin isoforms occurs throughout the entire protein, it is the extreme carboxy-terminal tails (CTTs) that exhibit the greatest concentration of differences. In humans, the CTTs range in length from 9 to 25 residues and because of a considerable number of glutamic acid residues, contain over 1/3 of tubulin's total electrostatic charge. The CTTs are believed to be highly disordered and their precise function has yet to be determined. However, their absence has been shown to result in altered microtubule stability and a reduction in the interaction with several microtubule-associated proteins (MAPs). To characterize the role that CTTs play in microtubule function, we examined the global conformational differences within a set of nine human  $\beta$ -tubulin isoforms using replica exchange molecular dynamics simulations. Through the analysis of the resulting configuration ensembles, we quantified differences such as the CTTs sequence influence on overall flexibility and average secondary structure. Although only minor variations between each CTT were observed, we suggest that these differences may be significant enough to affect interactions with MAPs, thereby influencing important properties such as microtubule assembly and stability.

### INTRODUCTION

Microtubules (MTs) are hollow cylinders constructed from linear chains of the protein tubulin. Tubulin is highly conserved throughout the entire eukaryotic kingdom, and the two main classes,  $\alpha/\beta$ , are expressed from multiple genes, producing several isoforms with seemingly identical functions (1–3). At the cellular level, it has been hypothesized that MT stability is regulated by subtle variations observed between  $\alpha$  and  $\beta$  isoforms (4–6). There is ~80%–95% sequence identity between isoforms; however, the extreme carboxy-terminal tails (CTTs) exhibit considerable differences, having only 50%–60% identity in this region (7,8). Early phylogenetic comparisons of the vertebrate  $\beta$ -tubulin families identified the CTTs, along with an internal variable domain, as the primary isoform defining features of the  $\beta$ -tubulin protein (9).

The importance of the CTTs has been demonstrated through their removal using limited proteolytic cleavage by subtilisin. After cleavage, the critical tubulin concentration required for polymerization was found to be ~50 times lower than that for intact tubulin (10). This rate was shown to decrease further through the addition of MT-stabilizing compounds such as paclitaxel (11). Proteolyzed tubulin also exhibits altered protofilament-bending, resulting in the formation of sheets, bundles of twisted filaments, rings, unstructured aggregates, or MTs with reversed polarity (10,12). Finally, the presence of the highly charged CTTs is

thought to obstruct tubulin/tubulin interactions, regulating the rate and conformation of MT assembly through unfavorable electrostatic interactions (13). This hypothesis is supported by the observation that divalent cations, or the substitution of acidic residues, results in an increased rate of MT assembly and a decreased rate of MT disassembly (14,15).

In addition to tubulin interactions within the MT itself, the functional stability of MTs *in vivo* has been shown to depend on interactions with MT-associated proteins (MAPs) through interactions with the CTTs. For example, the correct assembly of the kinetochore and the integrity of the mitotic spindle is dependent on the ability of the Dam1 complex to bind the CTTs of  $\beta$ -tubulin (16). Interactions between the CTTs and MAP2 or MAP  $\tau$  have also been shown to result in the stabilization of MT structures *in vivo* (17,18). Interactions between kinesin and MTs may also be significantly modulated through direct contacts with the CTT (19) or through the indirect modulation of kinesin's ability to bind ATP (20). Recent evidence also suggests that the CTTs may play a role in apoptosis, as interactions between MTs and the proapoptotic and antiapoptotic members of the Bcl-2 family are also affected by the presence of the CTT regions of  $\alpha$ - and  $\beta$ -tubulin (21), an observation that may offer a plausible explanation for the previously observed interactions between Bcl-2 and paclitaxel (22).

Although the presence of intact CTTs is seemingly essential for proper MT assembly and function, little is known about their structure and how this may affect interactions with other proteins. Several early studies have examined possible conformations of the CTTs using NMR and circular dichroism (CD) spectroscopy and demonstrated that there was

Submitted June 13, 2007, and accepted for publication October 22, 2007.

Address reprint requests to Jack Tuszynski, Room 3336, Cross Cancer Institute, 11560 University Avenue, Edmonton AB T6G 1Z2, Canada. Tel.: 780-432-8906; Fax: 780-432-8892; E-mail: jtus@phys.ualberta.ca.

Editor: Michael Edidin.

inherent disorder within the CTTs (23,24). Although a region of increased helicity toward the amino terminus was identified, a finding that was subsequently confirmed by electron crystallographic analysis, this has provided little additional information about CTT structure, as the last 10 residues of  $\alpha$ -tubulin and the last 18 residues of  $\beta$ -tubulin were not visualized due to lack of density (25). A possible explanation for this lack of density is the nonhomogenous presence of tubulin isotypes in the MT preparations used in the crystallographic analysis. However, homogenous samples of  $\alpha\beta$ II and  $\alpha\beta$ III tubulin failed to improve the quality of density maps, and Nogales et al. (25) concluded that their lack of resolution was indeed due to disorder in this region.

Fortunately, molecular modeling provides us with the unique ability to examine conformations of the CTTs at a level of detail experimental analysis is unable to yet provide. However, even modeling has its limitations in this regard, where a persistent problem with low temperature protein-folding simulations is that of obtaining adequate sampling. This problem exists because simulations generally become trapped in one of a large number of local energy minima. Several generalized ensemble algorithms, based on non-Boltzmann probability weight factors, are capable of overcoming this problem by introducing a random walk in energy space (26). However, it is often not a trivial matter to determine the non-Boltzmann weight factors, and for this reason we chose replica exchange molecular dynamics (REMD) as a method to examine possible conformations of the tubulin CTTs (27,28). REMD utilizes a large number of parallel simulations at different temperatures, with exchanges between trajectories attempted periodically using Metropolis criteria. All replicas of the system then perform a random walk through temperature space and, as a consequence, also through energy space. A replica may therefore overcome an energy barrier through exchange with replicas at higher temperatures. This allows configurations to be sampled at a given temperature on timescales not otherwise possible while still maintaining a thermodynamically consistent ensemble of configurations. Here, we discuss the creation of nine models of human  $\beta$ -tubulin CTT peptides using REMD and their analysis for relative conformational flexibility within the ensemble.

## METHODS AND MATERIALS

### Peptide construction and system configuration

Based on our previous examination of  $\beta$ -tubulin isotypes, we chose a set of nine peptides corresponding to the consensus CTT sequences of  $\beta$ I (GI:34222261),  $\beta$ II (GI:68299771 and GI:42476191),  $\beta$ III (GI:50592995),  $\beta$ IVa (GI:21361321),  $\beta$ IVb (GI:68051719),  $\beta$ V (GI:14210535),  $\beta$ VI (GI:41152077),  $\beta$ VII (TUBB4Q) (GI:55770867), and  $\beta$ VIII (TUBB8) (GI:42558278) (Table 1) (8). Accession numbers correspond to Entrez Nucleotide expressed mRNA sequence IDs contained only within the human genome. Using the crystallographic structure of tubulin (Protein Data Bank (PDB) ID: 1JFF) as a guide, we selected the conserved DA[TK]A motif that is positioned at the extreme end of helix H12 as the initiation point for the

**TABLE 1** ClustalW multiple sequence alignment of the CTT sequences

Isotype	Sequence	Length	Charge
I	DATAEEEEED-FGEEAEEEA	18	-12
II	DATADEQGE-FEEEEGEDEA	19	-12
III	DATAEEEGEMYEDDEEESEAQGPK	24	-12
IVa	DATAEEEGEF-EEEEAEVEA	18	-11
IVb	DATAEEEGE-FEEEEAEVEA	19	-12
V	DATANDGEEAFEDDEEEIDG	20	-12
VI	DAKAVLEED-EEVTEEAEMEPEDKGH	25	-11
VII	DATAEGGGV	9	-3
VIII	DATAEEEEED-EEYAEEVEA	18	-12

The length and net charge at pH 7.0 was determined for each peptide, including the C-terminal cap.

construction of all the peptides (29). This position marks the beginning of the structurally undefined region of the CTT in the structure and the domain examined using NMR and CD spectroscopy (23,24). All nine CTTs were prepared identically in an extended conformation, having the N-terminal charge neutralized by capping with an acetyl group using the PyMol v0.99 residue and fragment builder facility (30). The conventional protonation state at pH 7.0 was used for all residues, with a focus on His, which was protonated on the  $\epsilon$  nitrogen throughout.

### Parameterization and model preparation

All calculations utilized the AMBER99 force field (31,32), which was applied using PDB2GMX in GROMACS 3.3 (33). Each peptide was placed in a rhombic-dodecahedron unit cell consisting of  $\sim$ 2800 TIP3P waters (34). Sodium and chlorine counterions were added to the most energetically favorable locations (as determined using GENION) such that the net charge of the system was neutralized and a final ion concentration of 100 mM was established. For all molecular dynamics (MD) and REMD calculations, rigid bonds were maintained using the LINCS algorithm for the peptide and SETTLE for waters. Particle Mesh Ewald (PME) was used for electrostatic interactions with a cutoff of 0.8 nm and a Lennard-Jones cutoff of 1.0 nm. Constant pressure was maintained by allowing the box size to fluctuate isotropically. Each system was minimized using a low-memory Broyden-Fletcher-Goldfarb-Shanno minimizer in GMXRUN until machine precision was achieved. This was followed by 20 ps of heating and 1.1 ns of equilibration.

### REMD

To achieve a transition probability of 0.3, 43 target temperatures between 273 and 382 K were selected. Each replica was then heated/cooled to its target temperature over 50 ps and simulated without exchange for 500 ps, followed by 500 ps of REMD. Production REMD runs consisted of 10 ns of dynamics for each replica with exchanges attempted every 5 ps. Energy and conformational snapshots were saved every 1 and 10 ps, respectively. For each isotype, this produced an ensemble of 43,000 conformations and an aggregate simulation time of 430 ns over all temperatures.

### Cluster analysis

A cluster analysis of the resulting REMD conformations was used to determine preferred conformations and relative populations of each peptide. The GROMOS clustering algorithm (39), implemented in G\_CLUSTER, was used for this purpose. All  $C\alpha$  atoms were root mean-square deviation (RMSD) fit to the extended starting structure, and an RMSD cutoff was set for the  $C\alpha$  atoms; for each structure from the ensemble, all structures within

this cutoff were assigned as neighbors. The structure with the most neighbors was then identified as the center of a cluster and was removed from the pool with all its neighbors. This process was repeated until all structures had been assigned to a cluster. The RMSD cutoff was chosen such that the largest cluster contained 50% plus 1% of the structures.

## Principal component analysis

As MD simulations tend to produce immense quantities of data, principal component analysis (PCA) is a powerful mathematical tool used to detect correlations in MD trajectories (40,41). To perform PCA, all of the conformations from the REMD ensemble were RMSD least squares fit to the reference structures, effectively removing all rotations and translations. Then, for non-mass-weighted PCA, the covariance matrix can be calculated as

$$\sigma_{ij} = \langle (r_i - \langle r_i \rangle)(r_j - \langle r_j \rangle) \rangle,$$

where  $r_1 \dots r_{3N}$  are the Cartesian coordinates of the  $N$  atoms used in the analysis and  $\langle \dots \rangle$  denotes the ensemble average. The resulting matrix can then be diagonalized, and the resulting  $3N$ -dimensional eigenvectors,  $\mathbf{v}_i$ , are organized in descending order of eigenvalues,  $\lambda_i$ . Eigenvalues represent the variance along their associated eigenvector, and the larger the eigenvalue the more significant the correlated motion.

A PCA of all nine isotopes of  $\beta$ -tubulin was performed at 311 K using the positions of all C $\alpha$  atoms. To mimic the CTTs bound at their N-terminus, the backbone atoms of the first three amino acids of the ensemble were RMSD fit to the reference structure for the isotype at 311 K, producing an average structure (Fig. 1). This was essential for the physiological relevance of the PCA calculation and the considerable motion of the CTTs. Eigenvalues of different isotopes cannot be directly compared, as different numbers of atoms were used in the covariant analysis. This can be overcome by normalizing the eigenvalues by the number of atoms used in the analysis:

$$\sigma_i = \sqrt{\frac{\lambda_i}{N}},$$

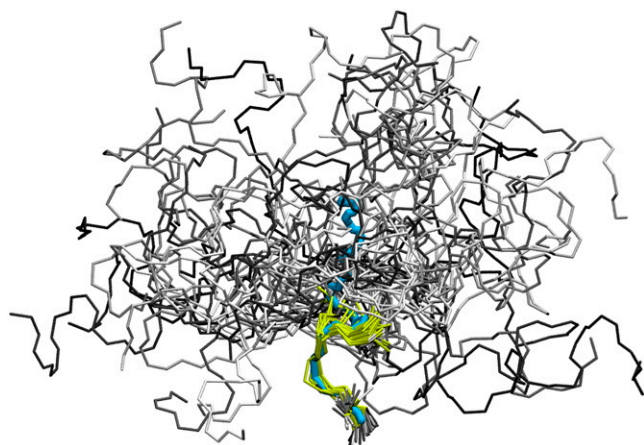


FIGURE 1 Human  $\beta$ -tubulin CTT sequence alignment and consensus. Illustration of reference and average structures used for PCA. The N, C, and C $\alpha$  atoms from the three N-terminal residues of each conformation from the  $\beta$ I ensemble were RMSD fit to a representative structure backbone. The backbone atoms of 50 randomly selected conformers of the ensemble are shown in black, gray, and white to delineate conformers in no specific order. The first four residues of each peptide are shown in yellow. The average structure, used for subsequent PCA calculations, is shown in light blue. Image produced using VMD (63).

where  $\sigma_i$  is the root normalized eigenvalue representing the standard deviation along the  $i$ th eigenvector.

## Sequence alignments

Sequence alignments of CTTs were performed with the default values using the European Bioinformatics Institute Clustal server (<http://www.ebi.ac.uk/clustalw/>), with the exception that the extended gap penalty was increased to a value of 0.5 (Table 1).

## Motif identification

Structural motifs within the CTT ensemble were identified by performing pairwise alignment of each CTT sequence using the ClustalW alignment facility in MacVector (MacVector, Cary, NC). Motifs were identified as those sequences containing at least four consecutive identical residues. A structural similarity search was then performed using only sequence runs of four or more identical or similar residues. Structural similarity was determined by calculating the  $\rho_{sc}$  between all pairs of conformations from the ensembles of each isotype to create a  $1001 \times 1001$  matrix. Here,  $\rho$  is defined as

$$\rho(A, B) = \frac{2 \times \text{RMSD}(A, B)}{\left( R_{\text{gyr}}^2(A) + R_{\text{gyr}}^2(B) - \text{RMSD}^2(A, B) \right)^{1/2}},$$

where  $A$  and  $B$  are structures with an equal number of points,  $R_{\text{gyr}}$  is the radius of gyration, and RMSD stands for the RMSD of the two structures. To obtain  $\rho_{sc}$ , we translate the centroid of each structure to the origin and rotate it such that the principal axes of inertia lie on the coordinate axes. Then, for both structures, we satisfy the condition that  $R_{\text{gyr}}=1$  by scaling each axis independently such that each contributes equally to the radius of gyration. The  $\rho_{sc}$  provides a length-independent measure of conformational similarity of two structures (42). Structures with a  $\rho_{sc}$  of 0 are considered identical; those with a factor of 2 are considered maximally dissimilar. Mirror images have a value of  $\sqrt{2}$ , whereas structures with a factor of 0.3–0.5 indicate visual similarity. The fraction of conformational pairs with a  $\rho_{sc}$  of  $<0.3$  is the fraction of time that regions within the two isotopes are structurally similar. See Table 3 for regions that showed higher than 50%  $\rho_{sc}$  similarity between isotopes.

## RESULTS

### Amino acid composition of CTT consensus sequences

The absolute length of each human  $\beta$ -tubulin CTT ranges from 9 to 25 residues; however, their typical length falls between 18 and 20 residues (Table 1). The exceptions are  $\beta$ VII, which contains only nine residues, and  $\beta$ III/ $\beta$ VI, which have 24 and 25 residues, respectively. Although the overall amino acid composition of each CTT is quite similar, there are some notable factors to consider. When we examine all nine human CTT sequences in aggregate, the residue that occurs most frequently is Glu (42%). The occurrence of all other residues drops significantly from this value: Ala (18%), Asp (12%), Gly (8%), Thr (5%), Val (4%), and Phe (3%). The  $\beta$ VIII CTT contains no Gly, the  $\beta$ IVa/b CTTs have a proportionally low Asp content, and, finally, the  $\beta$ VII CTT has a reduced Glu content, which may simply be a result of its length. Interestingly, the remainder of the amino acids

occurs exclusively in the  $\beta$ III (Lys, Met, Pro, Gln, Ser, and Tyr),  $\beta$ V (Ile, Asn), and  $\beta$ VI (His, Lys, Leu, and Pro) CTTs at frequencies of 1% or less. Finally, Cys, Trp, and Arg are absent from all the CTT sequences.

## REMD and completeness of sampling

A notable exclusion from these simulations is the tubulin protein itself, a factor that will undoubtedly influence the results as they have been explained here. The decision to exclude the bulk tubulin was made due to its large system size and the computational resources that would be required to explore the conformational space of an entire tubulin dimer. Therefore, simulations of only the CTTs become a problem of protein folding, and adequate sampling is critical to the proper interpretation of the results. To increase sampling efficiency at different temperatures, we used replica exchange MD (REMD), which does not provide information on the dynamics of the system. However, as our ultimate interest is in the overall conformations of the CTTs, the loss of dynamics data was an acceptable compromise to gain increased sampling.

The completeness of sampling can be determined by calculating the normalized overlaps between two different parts of an MD trajectory. Such overlaps can indicate whether or not both parts are spanning over the conformational space equally and not diffusing to new parts (43–45). Subspace overlap between two sets of  $n$  orthonormal vectors  $\mathbf{v}_1, \dots, \mathbf{v}_n$  and  $\mathbf{w}_1, \dots, \mathbf{w}_n$  is defined as

$$\text{overlap}(\mathbf{v}, \mathbf{w}) \equiv \frac{1}{n} \sum_{i=1, j=1}^{n, n} (\mathbf{v} \times \mathbf{w})^2.$$

When the overlap has a value of 1, the sets  $\mathbf{v}$  and  $\mathbf{w}$  can be considered to span the same subspace. This measure, however, does not account for the magnitudes of the eigenvalues, meaning that differences between all eigenvectors contribute equally. Furthermore, when two or more eigenvalues are equal, the corresponding eigenvectors are random, causing a random variation in the subspace overlap number. An alternative method, suggested by Hess (45), is the normalized overlap, defined as

$$d = \sqrt{\text{tr}\left(\left(\sqrt{C_1} - \sqrt{C_2}\right)^2\right)}$$

$$\text{normalized overlap}(C_1, C_2) = 1 - \frac{d}{\sqrt{\text{tr}(C_1) + \text{tr}(C_2)}},$$

where  $d$  is the difference between the covariant matrices,  $C_1$  and  $C_2$ , and  $\text{tr}$  is the trace. Here, if the overlap is 0, then the two sets are considered to be orthogonal, whereas an overlap of 1 indicates that the matrices are identical. Covariant analysis of the trajectories from each isotype at 311 K, chronologically divided into thirds, was performed using the same procedure used for the PCA. The subspace and

**TABLE 2** PCA overlap of CTT peptides

PCA overlap at 311 K for N-terminus fit						
Isotype	Subspace overlap			Normalized overlap		
	1st vs. 2nd	1st vs. 3rd	2nd vs. 3rd	1st vs. 2nd	1st vs. 3rd	2nd vs. 3rd
I	0.93	0.95	0.95	0.86	0.84	0.86
II	0.90	0.87	0.92	0.84	0.77	0.78
III	0.94	0.91	0.95	0.82	0.78	0.86
IVA	0.92	0.84	0.91	0.77	0.62	0.75
IVB	0.93	0.89	0.89	0.78	0.76	0.83
V	0.92	0.89	0.89	0.75	0.78	0.76
VI	0.88	0.88	0.94	0.72	0.71	0.84
VII	0.96	0.91	0.98	0.70	0.56	0.76
VIII	0.92	0.86	0.94	0.79	0.76	0.85

The subspace overlap consists of the first 10 eigenvectors only. For both overlap calculations, the ensemble was divided into thirds and all are compared to each other.

normalized overlaps calculated between each of these thirds are reported in Table 2. The high overlap between the thirds indicates that each part of the simulation is sampling approximately the same conformational space, and it is unlikely that there are unexplored regions missed earlier in the run. Although not a guarantee of complete equilibrium sampling, we concluded that the overlap using both of the above mentioned methods is acceptable and that adequate sampling within all of these systems has been obtained.

## Clustering and secondary structure

Clustering each CTT peptide can provide a representation of probable folded conformations; however, a bell-shaped  $\rho_{sc}$  distribution about a representative structure demonstrates that there is actually no native folded conformation for most of the isotypes (Fig. 2). The only exception was the  $\beta$ VII CTT, which exhibited a significant population of folded structures with a  $\rho_{sc} < 0.5$  and a second population of unfolded structures with a  $\rho_{sc} > 0.5$ . These results recapitulate previous observations that the CTTs are extremely flexible, and any structures within them are transient in nature and cannot be captured by the clustering methodology used in this study. Therefore, we felt that a more appropriate analysis was to consider the ensemble average of secondary structures as calculated from the entire 10 ns of ensemble conformations at 311 K by STRIDE (46) (Fig. 3). These results demonstrated that, whereas the  $\rho_{sc}$  distribution showed no native folded conformations, many of the CTTs contain regions that are either  $\alpha$ - or 3–10-helical at least 40% of the time.

## Motif identification

Having established the presence of transient secondary structures within the ensembles, we performed pairwise alignments of all the CTT sequences to identify potential sequence

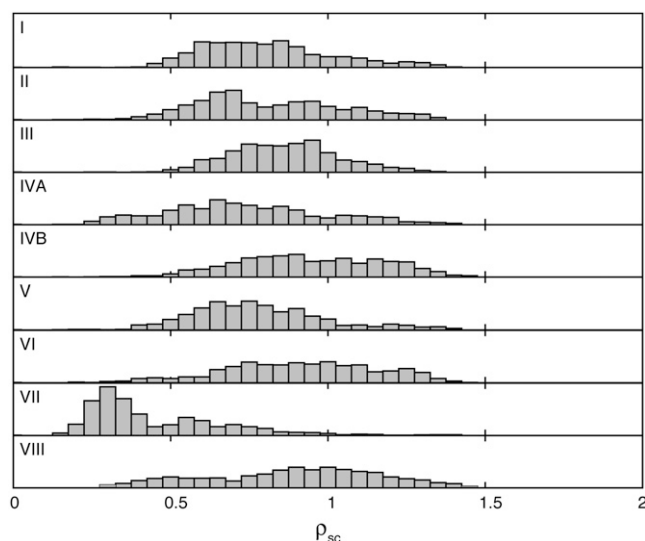


FIGURE 2  $\rho_{sc}$  distributions of CTTs about representative structures. The large spread in the histograms of the  $\rho_{sc}$  distributions for each CTT indicates the lack of a single folded conformation. A distinct population of 67% conformers have a  $\rho_{sc} < 0.5$  for the  $\beta$ VII isotype, suggesting a distinct folded conformation for this CTT.

motifs with which to correlate our modeling results. A  $\rho_{sc}$  matrix was then calculated for all the conformations across each isotype at 311 K. Only those motifs having a high fraction of low  $\rho_{sc}$  were determined to be significantly similar ( $\rho_{sc}$  of 0.3 or less) (Table 3). Through this analysis, we identified two motifs that showed both sequence and conformational similarities. To visually compare the motifs, the conformations of the aligned subsequences were clustered as described in Materials and Methods. Illustrated in Fig. 4 are RMSD alignments of representative structures (those with the most neighbors) from each isotype containing the respective motif. The first of these motifs was identified as a probable casein kinase-2 (CK-2) binding motif at the N-terminal end of the peptide (Fig. 4 A). This motif was also independently identified using a Prosite search for motifs within each of the CTT sequences (47) (not shown). The second motif was determined to correspond to a previously identified MAP2 binding motif found within  $\alpha$ -tubulin (Fig. 4 B) (48). We should note that although a common conformation for each of the motifs across isotypes has been identified, these are not stable folds; and depending on the isotype and residues included in the search, anywhere between 1% and 69% of the ensemble structures have a  $\rho_{sc} < 0.3$  when compared to the motif conformation.

## PCA

Not only is the secondary structure of each CTT significant, but their flexibility is also of critical importance when attempting to understand how they are able to conform when interacting with MAPs. Because of the complex interplay of

many factors, sophisticated statistical analysis of conformational ensembles, such as PCA, appears to be the most appropriate computational tool to characterize the flexibility. PCA was performed on the ensemble at 311 K, with the backbone atoms of the first three residues RMSD fit to a reference structure (Fig. 1). Resulting eigenvectors were ordered by descending eigenvalues, which represent the variance of the motion along the principal components. In Fig. 5, the six largest root normalized eigenvalues are shown for each isotype. Except for  $\beta$ VII and  $\beta$ IVa, to a lesser extent, three components can be identified with prominent eigenvalues of comparable magnitude, which is significantly higher than the magnitude of other eigenvalues. The three components with the largest eigenvalues represent correlated motions of the peptide fragments with the most significant standard deviations of the motion along the corresponding orthogonal directions. Since the standard deviations shown in Fig. 5 were normalized for the number of residues, they can be employed as a universal measure for comparison of conformational motions in different CTT peptides.

## Two-dimensional projections of the REMD conformation ensemble

A more detailed representation of the conformational motions in peptides is provided by projections of the ensemble of conformations onto the planes spanned by the most important principal components (Fig. 6). Here, the spatial distributions of occupancies of the various conformational states are shown over the planes spanned by the first and second and by the first and third principal components. A representative isotype,  $\beta$ V, illustrates that both distributions are smooth and without evidence of considerable clustering. Similar behavior exists in all isoforms considered except for  $\beta$ IVa and  $\beta$ VII, which show significant clustering of conformational occupancy, with maxima at (0.1,0) and (−0.2,0), respectively. The widths of the distributions shown in Fig. 6 are not directly dependent on the length of the peptide, as the eigenvalues have been normalized by the number of residues. The reduced width of  $\beta$ VII, therefore, indicates its reduced mobility. The bin size of the histogram has been chosen in proportion to the width of the distribution; so the maxima seen for  $\beta$ IVa and  $\beta$ VII are not artifacts of the sampling but represent regions of increased occupancy. Since the occupancy of a conformational state is inversely proportional to the free energy of that state (49,50), the lack of clustering in the other isotypes suggests that the global minima are broad or that there are generally no significant local minima. In contrast,  $\beta$ VII has a significant basin of attraction, indicative of a folded conformation. This observation is consistent with the preceding cluster analysis, which identified a well-defined folded state in the  $\beta$ VII isoform only. Although less defined, the occupancy distribution for  $\beta$ IVa also exhibits a preferred folded conformation, demonstrating that the PCA methodology is more sensitive to transient structural motifs than the

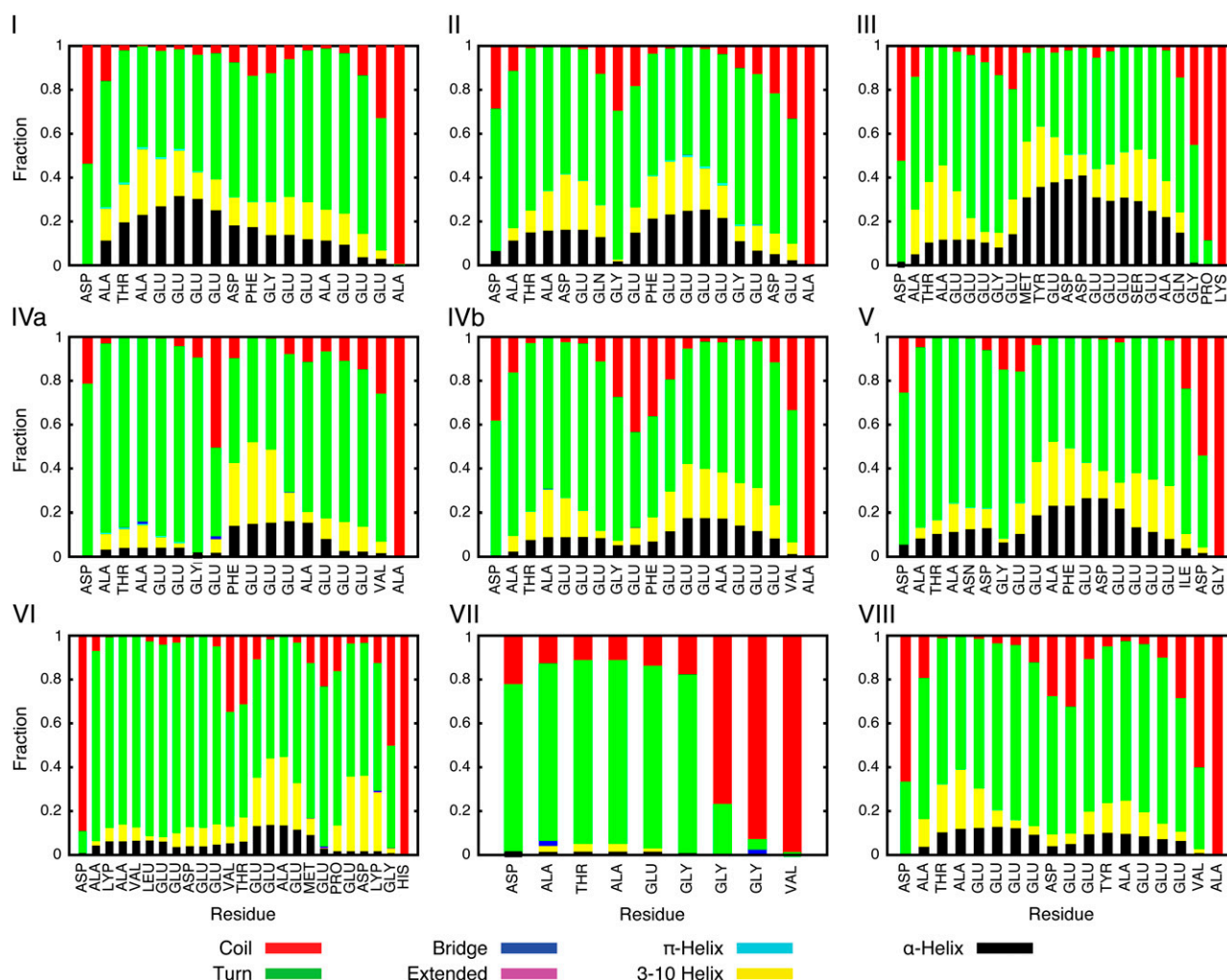


FIGURE 3 Time average of each type of secondary structure for each CTT. The total fractional secondary structure content was determined as an average over the entire 10 ns MD simulation at 311 K. The fractional time averages of each type of secondary structure are stacked to sum up to one. Although experimental studies of tubulin suggest that the CTTs are unstructured, these results suggest that many of the CTTs contain a significant amount of transient helix.

clustering analysis. One more noteworthy feature is that the two-dimensional distributions over the first and second components are largely similar to those over the first and third components, suggesting a significant level of symmetry among principal components.

### Relative CTT flexibility

In addition to information regarding the motion of each CTT peptide, PCA provides the ability to compare relative flexibilities. As we studied only the nine human  $\beta$ -tubulin isoforms, it is not possible to comment on the flexibility of the CTT peptides in an absolute sense. However, there have been a few studies discussing the flexibility of small peptides. Although they use a different approach to calculating the flexibility, Ma et al. made a survey of 28 short peptides and determined that the native helical structures of the peptides were more flexible than random or disordered conformations, in agreement with our observations (Fig. 7 B)

(51). Unfortunately, their methodology, calculating the vibrational free energy, does not allow comparison of flexibility between peptides. Here, as a measure of the relative flexibility of each peptide, the eigenvalues of principal components were compared. To characterize the flexibility, we employed the value  $\sigma_{12} = \sqrt{\sigma_1^2 + \sigma_2^2}$ , where  $\sigma_1$  and  $\sigma_2$  are the normalized eigenvalues for the first and the second principal components, respectively. The magnitude of the PCA eigenvalue parameter  $\sigma_{12}$  can be viewed as the mean square width of the two-dimensional REMD configuration ensembles shown in Fig. 6, and thus  $\sigma_{12}$  can be directly interpreted as the flexibility of the peptide associated with its major correlated motions. Although the CTTs are characterized by three principal components (Fig. 5), the considerable symmetry between the second and third component demonstrated in Fig. 6 allows the use of only two components out of three to quantify the flexibility of CTTs. The correlations between the normalized distance (the distance between the N- and C-termini C $\alpha$ s, divided by the number of residues),



**TABLE 3** Sequence motif identification

CK2 Motif						
Isotype	Range	Sequence	Isotype	Range	Sequence	Similarity
III	1–9	DATAEEEEGE	IVb	1–9	DATAEEEEGE	72.56
II	1–8	DATADEQG	III	1–8	DATAEEEG	76.64
II	1–8	DATADEQG	IVb	1–8	DATAEEEG	62.97
II	1–5	DATAD	VII	1–5	DATAE	61.4
IVa	1–5	DATAE	VII	1–5	DATAE	64
II	1–4	DATA	V	1–4	DATA	64.1
IVb	1–4	DATA	V	1–4	DATA	62.34
V	1–4	DATA	VI	1–4	DAKA	51.26
V	1–4	DATA	VII	1–4	DATA	67.19
V	1–4	DATA	VIII	1–4	DATA	52.9
VI	1–4	DATA	VII	1–4	DATA	52.7
MAP2 Motif						
Isotype	Range	Sequence	Isotype	Range	Sequence	Similarity
II	10–14	FEEDD	III	11–15	YEDDE	60.89
II	10–14	FEEDD	V	11–15	FEDEE	71.61
IVa	11–14	EEAE	VI	14–17	EEAE	71.87
I	12–15	EEAE	VI	14–17	EEAE	78.71
IVb	12–15	EEAE	VI	14–17	EEAE	80.93
I	15–18	EEEA	III	15–18	EEES	67.62
IVa	14–17	EEEV	V	15–18	EEEI	69.12
IVb	15–18	EEEV	V	15–18	EEEI	85.56
V	15–18	EEEI	VIII	14–17	EEEV	55.65

Pairwise alignments of each CTT sequence identified several similar regions between peptides. Structural similarity between sequences was determined by calculating the  $\rho_{sc}$  between all pairs of conformations. Those pairs with high conformational and sequence similarity were used to characterize the motif conformations for the CK-2 and MAP2 domains (Fig. 4).

the time-averaged helical content, and the average clustering parameter  $\langle \rho_{sc} \rangle$  with the value  $\sigma_{12}$ , which we use as the major measure of flexibility, can be seen in Fig. 7.

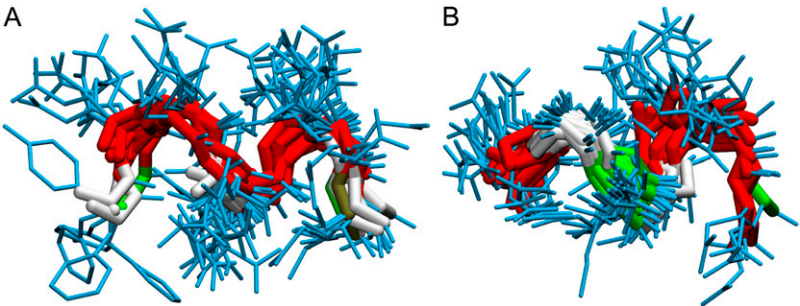
# DISCUSSION

The CTT's involvement in MT regulation is critical; it is essential to understand the conformational differences adopted by different tubulin isotypes. We suggest that the sequence variability within the CTTs may have arisen as a mechanism to conserve the overall tubulin structure to maintain proper MT assembly, while still providing a flexible, solvent-exposed region of increased variability for interactions with proteins that affect MT function (see alignment in Table 1). This is both an attractive and reasonable hypothesis, as the role of intrinsically disordered proteins in protein interactions has been implicated as a

mechanism to enhance specificity (52). The following discussion addresses three main points about the CTTs and their interactions with MAPs: first, the relative flexibilities of each of the CTT peptides; second, the presence of transient structure within these peptides; and finally, how this may influence protein interactions. We will discuss this possibility further when examining CTT flexibility in the context of MAP interactions, in particular interactions with the CK-2 and MAP2 binding motifs that we identified in this work.

# CTT flexibility and secondary structure

Accurate flexibility and secondary structure measurements are critical when the goal is to propose mechanisms for MAP interactions with the large number of possible CTT conformations on the MT surface. The overall distribution of amino acids within the CTT will obviously have an effect on their



**FIGURE 4** Sequence motif identification. High scoring sequence pairs from Table 3 were RMSD fit to illustrate the overall fold and position of side chains in the structural motifs. In both images, red residues are acidic, white residues are nonpolar, green residues are polar, and all side chains are blue. Image created using VMD (63).

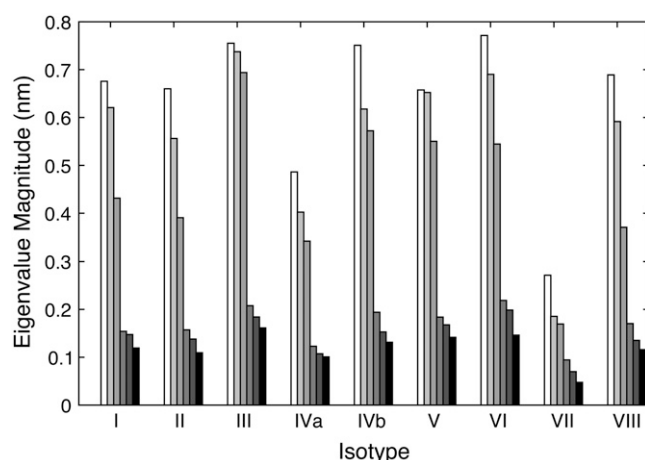


FIGURE 5 Root mean-square normalized eigenvalues for each isotype. The magnitudes of the first six root normalized eigenvalues (white to black) are shown for each isotype and represent the standard deviation of the motion along the eigenvector normalized for the overall length of the peptide. The relative size of the root-normalized eigenvalues indicates the relative flexibilities of the individual peptides.

flexibility and result in secondary structure. For example, the overall helical propensity of  $\beta$ II,  $\beta$ III,  $\beta$ IVa,  $\beta$ IVb, and  $\beta$ V is interrupted by the presence of Gly (Fig. 3). Additionally, regions with uninterrupted stretches of Glu also tend to produce regions with greater helical propensity, providing an argument for the prevalence of this residue within the CTT sequences. This observation can be compared with the results of Roe et al. for simulations performed on decaalanine (53). Although shorter than all but the  $\beta$ VII peptide, the overall  $\alpha$ - and 3–10-helical content within decaalanine was similar to the helical content observed here. However, through the analysis of the distribution of secondary structure along the CTT sequences, we note that our results differ from that of Roe et al. in that the helical content of the CTTs drops near the middle of the sequence for most isotypes

whereas for decaalanine the central residues have a maximal amount of helix.

Three characteristics emerge which are representative of the behaviors of CTTs: the end-to-end distance, normalized by the number of residues, the PCA-based eigenvalues, and the averaged helical content (Fig. 7, A and B). To test the applicability of PCA as a measure of relative flexibility, we compared PCA-based flexibilities with the average parameter of structural similarity  $\langle\rho_{sc}\rangle$ , which we computed from the distributions in Fig. 2 (Fig. 7 C). Although less detailed than PCA, the distributions of conformations over the parameter  $\rho_{sc}$  provide an alternative statistical characterization of the CTTs configurations, and thus they can be expected to correlate with the PCA-based results. The correlation between the PCA-based flexibilities and  $\langle\rho_{sc}\rangle$  is more pronounced than any other correlation that we investigated. Most of the CTTs show a clear proportionality between  $\sigma_{12}$  and  $\langle\rho_{sc}\rangle$ , and the variability of  $\langle\rho_{sc}\rangle$  that corresponds to similar  $\sigma_{12}$  is  $<15\%$ . This similarity of results from two fundamentally different statistical tools demonstrates that these statistical methodologies are the best suited to characterize CTTs' flexibility. Interestingly, rather than the expected exponential decay, the magnitudes of the first three eigenvalues are of similar magnitude and significantly larger than the remaining eigenvalues, indicating an isotropy of occupancies in the three-dimensional space (Fig. 5).

More simply, the CTTs are flexible enough for the ensemble of their configurations to be spherically symmetric with respect to an immobilized base of amino acids (Fig. 1). Since these three eigenvalues are significantly larger than all the others, three-dimensional symmetry is likely reached through a few highly flexible bonds within the peptides that allow rotations of large angles, thereby generating nearly spherically symmetric distributions of occupancies. In Fig. 6 this is illustrated by a comparison of the histogram for  $\beta$ V, which is representative of a significant level of symmetry typical for most isotypes, with the less symmetric ones for

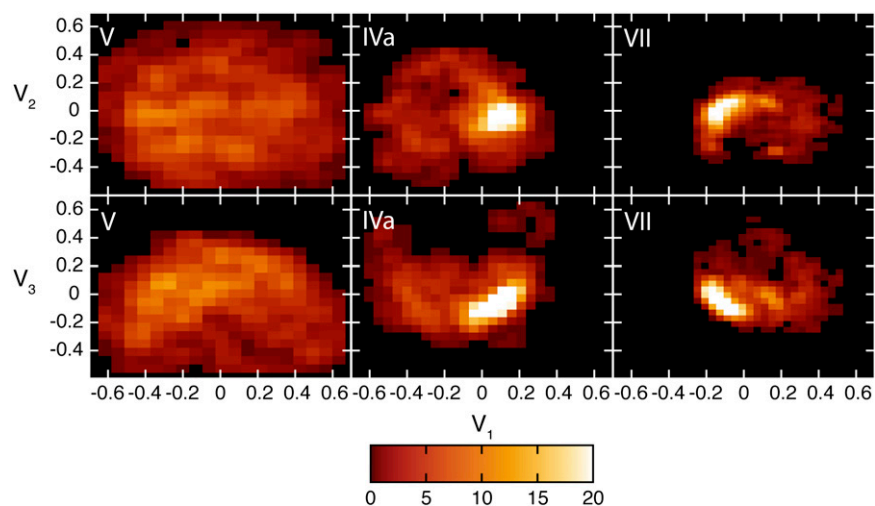


FIGURE 6 Projections of the ensemble of conformations onto the planes of the three most important principal components. The first and second principal components (upper row) and the first and third principal components (lower row) are plotted on the x and y axes, respectively, for  $\beta$ V,  $\beta$ IVa, and  $\beta$ VII. The histograms represent the occupancies of the corresponding conformational states, with lighter colors indicating more frequently visited areas. The same coloring scheme is used for all isotypes and is capped at 20 counts.  $\beta$ V is typical of the remaining isotypes in the width of the distribution and lack of a significant maxima.



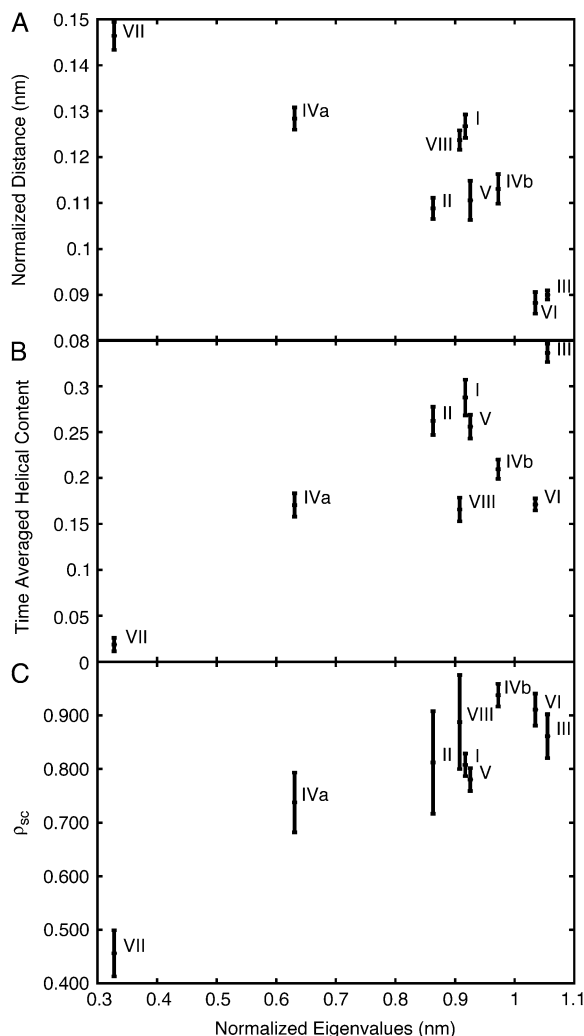


FIGURE 7 Peptide's flexibility indicators. The end-to-end distance normalized by the number of residues (A), time average helicity (B), and  $\langle \rho_{sc} \rangle$  (C) of each CTT peptide at 311 K plotted as functions of  $\sigma_{12}$ . The plots demonstrate the correlation between each of the above values and the major flexibility indicator,  $\sigma_{12}$ . In *b*,  $\beta VI$  can be excluded from the trend due to the presence of a Pro, which disrupts the secondary structure but contributes to the flexibility of the peptide. The error in all three plots is the standard error calculated using box averaging (64).

$\beta IVa$  and  $\beta VII$ . As the magnitudes of the eigenvalues shift to an exponential decay, spherical symmetry is lost and correlated motions begin to appear. At the same time, an individual CTT may preserve relatively stable motifs contained within some secondary structure. In terms of the PCA results presented in Fig. 5, these motifs are represented by the fourth, fifth, and higher eigenvalues. The fact that these components, although significantly less pronounced than the first three, still have considerable nonzero eigenvalues indicates that the conformational motifs within the CTTs are of transient nature and subject to variability over the trajectory. This is consistent with our observation of a significant amount of transient helicity (Fig. 3), indicated by a substantial amount of 3–10-helix.

Considering the PCA eigenvalue parameter  $\sigma_{12}$  as the measure of flexibility, it is evident that the  $\beta III$  and  $\beta VI$  CTTs are the most flexible of all CTTs. These are the longest two sequences, which may account for some of their flexibility, but interestingly they also have the shortest end-to-end distance per residue. The cause of these shared structural properties differs in the two cases: whereas the  $\beta III$  CTT has among the highest degree of secondary structure, the  $\beta VI$  CTT has among the lowest. The  $\beta VI$  CTT differs from all the other fragments, as it contains a Pro near the C-terminus of the peptide. This Pro breaks the helical structure of the fragment and, at the same time, does not facilitate an extended conformation. This accounts for the lack of secondary structure but is in contrast to  $\beta III$ , which lies predominantly in the  $\alpha$ -helical region of the  $\Phi/\Psi$  distribution. When accounting for the lone Pro and omitting the data for  $\beta VI$ , the correlation between helicity and flexibility becomes greater. The  $\beta II$ ,  $\beta IVb$ , and  $\beta V$  CTTs show the same spring-like flexibility of  $\beta III$ , exhibiting a similar helical content and moderate end-to-end distance per residue. These CTTs also contain Gly at the same position that breaks the helical content of the fragment. In contrast, the  $\beta IVa$  and  $\beta VII$  CTTs displayed the least amount of flexibility. In the case of  $\beta IVa$ , one reason for this reduction in flexibility may be the early occurrence of Gly in its sequence, which stiffens this region of the peptide by decreasing the helicity. This decrease in helicity in the first seven residues, compared to  $\beta IVb$ , can be seen in Fig. 3.  $\beta IVa$  also contained the most stable DA[TK]AEE motif (data not shown), which resulted in the largest normalized distance of all the common isotopes.

The observation that peptides containing a significant amount of transient helical conformations are more flexible than an extended one may seem counterintuitive. The correlation between the normalized distance and flexibility in Fig. 7 A is most easily understood in terms of entropy. Take the following example: the freely jointed, or ideal chain, polymer model states that compact forms of the polymer contain more states that are accessible (i.e., have greater entropy) than do extended states (54). Thus, an entropic force drives the polymer or polypeptide into compact conformations. A greater number of accessible states suggests that the chain is more flexible and that extended states are more rigid. Although the ideal chain has no internal energy and is free to collapse, a peptide fragment does contain internal energy that can stiffen the peptide. The source of the force that causes the peptides to be in a rigid, extended state is the peptide backbone. Individual residues that induce a compact conformation through helical propensity or residues that induce a turn (such as proline) provide flexibility to the structure of the peptide while reducing the normalized distance. However, there is not a complete one-to-one correspondence between the flexibility and the normalized distance. For example,  $\beta I$  and  $\beta VIII$  have a normalized distance comparable to  $\beta IVa$  but a  $\sigma_{12}$  closer to the more compact  $\beta V$ . Although the normalized distance of the peptide is a strong indicator of flexibility,

the details of the compact structure, such as the degree of helicity, also contribute. We are not the first to observe this phenomenon in peptides; Ma et al. calculated the vibrational free energy of 28 short peptides in native helical and random extend states and determined that the native helical structures of the peptides were more flexible in all cases (51).

It is becoming clear that each  $\beta$ -tubulin isotype has a unique pattern of expression ranging from specific for  $\beta$ II,  $\beta$ III,  $\beta$ IVa, and  $\beta$ VI to constitutive for  $\beta$ I and  $\beta$ IV (56,57). The  $\beta$ I isotype is the most commonly expressed in humans and, as such, is also the most common isotype found in cancer cells (58). Our observations also suggest that the CTT of this isotype shows “typical” behavior in our analysis. It is difficult to correlate our results to biological function; however, a clear pattern illustrating  $\beta$ III as a statistical outlier has emerged. The  $\beta$ III CTT was one of the most flexible peptides and contained a large amount of transient secondary structure. This is interesting, as  $\beta$ III-tubulin has been observed at increased levels in human tumors and implicated in the development of drug resistance to standard chemotherapy treatments (59–61). In addition to the altered expression of  $\beta$ III, the alteration of tubulin regulatory proteins such as MAP4 has also been implicated in changes to MT dynamics and the development of drug resistance (62). Because several MAPs have now been shown to interact directly with the CTTs (16–19,21), the differences that we observed here may have a significant impact on their affinity to the surface of an MT.

### Motifs and MAP interactions

Using the conformational similarities for each CTT peptide, in combination with sequence comparisons, we identified two distinct transient structural motifs (Table 3). Several similar regions were observed as significant conformations throughout the trajectory files, indicating that they may be common motifs that could potentially be recognized by MAPs. The first motif was identified as a putative CK-2 phosphorylation site for which the consensus motif has been identified as [ST]XX[DE] (47). All of the CTTs, with the exception of  $\beta$ VI and  $\beta$ VII, were shown to contain this motif within the conserved DA[TK]A sequence. The second motif identified was the MAP2 binding site, for which the consensus sequence has previously been characterized as EEAEVE (48). Only the CTT of  $\beta$ I,  $\beta$ IVa, and  $\beta$ IVb strictly contain this motif; however, most of the other CTTs also contain a similar motif (Table 3). The conformations identified for  $\beta$ I,  $\beta$ IVa,  $\beta$ IVb, and  $\beta$ VI most commonly form a transient helix that is between four and five residues in length. In  $\beta$ VIII, there seems to be an increasing amount of unstructured loop associated with this region, which is most likely a result of the displacement of the amino end due to the presence of a bulky Tyr residue. Both the CK-2 and MAP2 motifs display a significantly higher degree of helical propensity when compared to the entire ensemble of CTT

conformations (Fig. 3). Interestingly, with the exception of  $\beta$ VI and  $\beta$ VII, the increased helical propensity of the proposed CK-2 domain suggests that the H12 helix may extend several additional residues farther than those that have already been observed in the crystal structures of tubulin.

Finally, the motifs identified here are not stable structures but rather commonly occurring, transient conformations. For example, in the “stiffest” isotype,  $\beta$ IVa, the CK-2 DATA motif is occupied 69% of the time whereas the MAP2 motif EEVE is occupied 17% of the time (Table 3). The two motifs are simultaneously folded in the ensemble of CTT trajectories only 12% of the time. As the sequences contained within the two motifs are extended to include DATAEE and AEEVEA, respectively, this drops to 2% (Table 3). Rather, the shared similarity of the folded states indicates that these conformations are individually inducible, particularly when binding to their respective substrates. As we are dealing with a subset of all CTTs, we were unable to determine accurately if any additional structural or sequence motifs exist in the tails. A more accurate determination of CTT motifs would require a larger survey of CTTs across all identified  $\alpha$ - and  $\beta$ -tubulin isotypes (see Table 1).

### CONCLUSIONS

The role of the unstructured tubulin CTTs, particularly concerning MT stability and interactions with MAPs, is a question that has yet to be conclusively answered. Here we have shown, using a combination of REMD and PCA, that although the experimental structures of tubulin suggest the CTTs are unstructured, many contain a significant amount of transient conformations that are linked by a few highly rotatable bonds. We believe that the role of the transient secondary structure, coupled with increased CTT flexibility, may be twofold. First, global flexibility coupled with weak helical tendencies could provide a mechanism that allows the CTTs to search conformational space and hence easily bind to other proteins. This could provide a convenient mechanism for MAPs to dock to the MT surface. Second, increased flexibility could also enhance binding affinities to MAPs as a result of their ability to conform to a specific binding site. The presence of well-defined structural motifs within the CTTs may, consequently, play an important role in binding specificity, where the presence of smaller domains provides a common structural signal for binding, while the global flexibility of the CTT is maintained.

The role of  $\beta$ -tubulin CTTs in MT function is appealing, since subtle differences in a small stretch of amino acids could have profound consequences on MAP interactions. Unfortunately, it is extremely challenging at this time to elicit tight correlations between the results presented here and their biological consequences, as there is very little experimental data regarding CTTs explicitly in the context of tubulin isotypes. As interest in the properties of tubulin isotypes and their potential role in targeted chemotherapy

treatments increases, we anticipate a significant body of data to become available soon. As a result, a clear understanding of CTT structure and function may provide the means by which we can begin to rationally design and develop novel drugs or peptide mimetics that target CTT binding sites, providing better selectivity for chemotherapies.

All calculations were performed on the Westgrid Glacier cluster and the Center of Excellence in Integrated Nanotools cluster. We thank Dr. Richard Luduena for his inspiration and insightful commentary during the assembly of this manuscript.

This research was directly funded through grants from the National Science and Engineering Research Council (NSERC), Mathematics of Information Technology and Complex Systems, and the Canadian Space Agency (under the contract PWGSC/CSA 9F007-052743/001/ST) awarded to J.A.T. Additional financial support from the Allard Foundation, Technology Innovations of Rochester, NY, and Oncovista of San Antonio, TX, is also gratefully acknowledged. This research was also supported in part by grant No. W81XWH-05-1-0238 from the U.S. Army/Department of Defense to Richard F. Luduena (University of Texas Health Science Center at San Antonio, San Antonio, TX). T.L. acknowledges financial support from the Province of Alberta, NSERC, National Research Council, and the University of Alberta.

## REFERENCES

- Hall, J. L., L. Dudley, P. R. Dobner, S. A. Lewis, and N. J. Cowan. 1983. Identification of two human  $\beta$ -tubulin isotypes. *Mol. Cell. Biol.* 3:854–862.
- Lewis, S. A., M. E. Gilmartin, J. L. Hall, and N. J. Cowan. 1985. Three expressed sequences within the human  $\beta$ -tubulin multigene family each define a distinct isotype. *J. Mol. Biol.* 182:11–20.
- Lu, Q., and R. F. Luduena. 1994. In vitro analysis of microtubule assembly of isotypically pure tubulin dimers. Intrinsic differences in the assembly properties of  $\alpha\beta$  II,  $\alpha\beta$  III, and  $\alpha\beta$  IV tubulin dimers in the absence of microtubule-associated proteins. *J. Biol. Chem.* 269: 2041–2047.
- Richards, K. L., K. R. Anders, E. Nogales, K. Schwartz, K. H. Downing, and D. Botstein. 2000. Structure-function relationships in yeast tubulins. *Mol. Biol. Cell.* 11:1887–1903.
- Panda, D., H. P. Miller, A. Banerjee, R. F. Luduena, and L. Wilson. 1994. Microtubule dynamics in vitro are regulated by the tubulin isotype composition. *Proc. Natl. Acad. Sci. USA.* 91:11358–11362.
- Fackenthal, J. D., F. R. Turner, and E. C. Raff. 1993. Tissue-specific microtubule functions in *Drosophila* spermatogenesis require the  $\beta$  2-tubulin isotype-specific carboxy terminus. *Dev. Biol.* 158:213–227.
- Carpenter, E. J., J. T. Huzil, R. F. Luduena, and J. A. Tuszynski. 2006. Homology modeling of tubulin: influence predictions for microtubule's biophysical properties. *Eur. Biophys. J.* 36:35–43.
- Huzil, J. T., R. F. Luduena, and J. Tuszynski. 2006. Comparative modelling of human  $\beta$  tubulin isotypes and implications for drug binding. *Nanotechnology.* 17:S90–S100.
- Fackenthal, J. D., J. A. Hutchens, F. R. Turner, and E. C. Raff. 1995. Structural analysis of mutations in the *Drosophila*  $\beta$  2-tubulin isoform reveals regions in the  $\beta$ -tubulin molecular required for general and for tissue-specific microtubule functions. *Genetics.* 139:267–286.
- Sackett, D. L., B. Bhattacharyya, and J. Wolff. 1985. Tubulin subunit carboxyl termini determine polymerization efficiency. *J. Biol. Chem.* 260: 43–45.
- Knipling, L., J. Hwang, and J. Wolff. 1999. Preparation and properties of pure tubulin S. *Cell Motil. Cytoskeleton.* 43:63–71.
- Peyrot, V., C. Briand, and J. M. Andreu. 1990. C-terminal cleavage of tubulin by subtilisin enhances ring formation. *Arch. Biochem. Biophys.* 279:328–337.
- Bhattacharyya, B., D. L. Sackett, and J. Wolff. 1985. Tubulin, hybrid dimers, and tubulin S. Stepwise charge reduction and polymerization. *J. Biol. Chem.* 260:10208–10216.
- Mejillano, M. R., E. T. Tolo, R. C. J. Williams, and R. H. Himes. 1992. The conversion of tubulin carboxyl groups to amides has a stabilizing effect on microtubules. *Biochemistry.* 31:3478–3483.
- Wolff, J., D. L. Sackett, and L. Knipling. 1996. Cation selective promotion of tubulin polymerization by alkali metal chlorides. *Protein Sci.* 5:2020–2028.
- Westermann, S., A. Avila-Sakar, H. W. Wang, H. Niederstrasser, J. Wong, D. G. Drubin, E. Nogales, and G. Barnes. 2005. Formation of a dynamic kinetochore-microtubule interface through assembly of the Dam1 ring complex. *Mol. Cell.* 17:277–290.
- Serrano, L., J. Avila, and R. B. Maccioni. 1984. Controlled proteolysis of tubulin by subtilisin: localization of the site for MAP2 interaction. *Biochemistry.* 23:4675–4681.
- Rodionov, V. I., F. K. Gyoeva, A. S. Kashina, S. A. Kuznetsov, and V. I. Gelfand. 1990. Microtubule-associated proteins and microtubule-based translocators have different binding sites on tubulin molecule. *J. Biol. Chem.* 265:5702–5707.
- Kikkawa, M., E. P. Sablin, Y. Okada, H. Yajima, R. J. Fletterick, and N. Hirokawa. 2001. Switch-based mechanism of kinesin motors. *Nature.* 411:439–445.
- Skinotis, G., J. C. Cochran, J. Muller, E. Mandelkow, S. P. Gilbert, and A. Hoenger. 2004. Modulation of kinesin binding by the C-termini of tubulin. *EMBO J.* 23:989–999.
- Knipling, L., and J. Wolff. 2006. Direct interaction of Bcl-2 proteins with tubulin. *Biochem. Biophys. Res. Commun.* 341:433–439.
- Rodi, D. J., R. W. Janes, H. J. Sanganee, R. A. Holton, B. A. Wallace, and L. Makowski. 1999. Screening of a library of phage-displayed peptides identifies human bcl-2 as a taxol-binding protein. *J. Mol. Biol.* 285:197–203.
- Sugiura, M., R. B. Maccioni, J. R. Cann, E. J. York, J. M. Stewart, and G. Kotovych. 1987. A proton magnetic resonance and a circular dichroism study of the solvent dependent conformation of the synthetic tubulin fragment Ac tubulin,  $\alpha$  (430–441) amide and its interaction with substance-P. *J. Biomol. Struct. Dyn.* 4:1105–1117.
- Otter, A., and G. Kotovych. 1988. The solution conformation of the synthetic tubulin fragment Ac-tubulin- $\alpha$ (430–441)-amide based on two-dimensional ROESY experiments. *Can. J. Chem.* 66:1814–1820.
- Nogales, E., S. G. Wolf, and K. H. Downing. 1998. Structure of the  $\alpha\beta$  tubulin dimer by electron crystallography. *Nature.* 391:199–203.
- Mitsutake, A., Y. Sugita, and Y. Okamoto. 2001. Generalized-ensemble algorithms for molecular simulations of biopolymers. *Biopolymers.* 60:96–123.
- Seibert, M. M., A. Patriksson, B. Hess, and D. van der Spoel. 2005. Reproducible polypeptide folding and structure prediction using molecular dynamics simulations. *J. Mol. Biol.* 354:173–183.
- Okabe, T., M. Kawata, Y. Okamoto, and M. Mikami. 2001. Replica-exchange Monte Carlo method for the isobaric-isothermal ensemble. *Chem. Phys. Lett.* 335:435–439.
- Nogales, E., S. G. Wolf, I. A. Khan, R. F. Luduena, and K. H. Downing. 1995. Structure of tubulin at 6.5 Å and location of the taxol-binding site. *Nature.* 375:424–427.
- Delano, W. L. 2002. The PyMOL Molecular Graphics System. <http://www.pymol.org>.
- Sorin, E. J., and V. S. Pande. 2005. Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. *Biophys. J.* 88: 2472–2493.
- Wang, J., P. Cieplak, and P. A. Kollman. 2000. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* 21:1049–1074.
- Lindahl, E., B. Hess, and D. van der Spoel. 2001. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* 7:306–317.

34. Mahoney, M. W., and W. L. Jorgensen. 2000. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* 112:8910–8922.
35. Reference deleted in proof.
36. Reference deleted in proof.
37. Reference deleted in proof.
38. Reference deleted in proof.
39. Daura, X., K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, and A. E. Mark. 1999. Peptide folding: when simulation meets experiment. *Angew. Chem. Int. Ed.* 38:236–240.
40. Garcia, A. E. 1992. Large-amplitude nonlinear motions in proteins. *Phys. Rev. Lett.* 68:2696–2699.
41. Amadei, A., A. B. Linssen, and H. J. Berendsen. 1993. Essential dynamics of proteins. *Proteins*. 17:412–425.
42. Maiorov, V. N., and G. M. Crippen. 1995. Size-independent comparison of protein three-dimensional structures. *Proteins*. 22:273–283.
43. Amadei, A., B. L. de Groot, M. A. Ceruso, M. Paci, A. Di Nola, and H. J. Berendsen. 1999. A kinetic model for the internal motions of proteins: diffusion between multiple harmonic wells. *Proteins*. 35: 283–292.
44. Hess, B. 2000. Similarities between principal components of protein dynamics and random diffusion. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics*. 62:8438–8448.
45. Hess, B. 2002. Convergence of sampling in protein simulations. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 65:031910–031920.
46. Frishman, D., and P. Argos. 1995. Knowledge-based protein secondary structure assignment. *Proteins*. 23:566–579.
47. Falquet, L., M. Pagni, P. Bucher, N. Hulo, C. J. Sigrist, K. Hofmann, and A. Bairoch. 2002. The PROSITE database, its status in 2002. *Nucleic Acids Res.* 30:235–238.
48. Paschal, B. M., R. A. Obar, and R. B. Vallee. 1989. Interaction of brain cytoplasmic dynein and MAP2 with a common sequence at the C terminus of tubulin. *Nature*. 342:569–572.
49. Grubmüller, H. 1995. Predicting slow structural transitions in macromolecular systems: conformational flooding. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics*. 52:2893–2906.
50. Kosztin, I., B. Barz, and L. Janosi. 2006. Calculating potentials of mean force and diffusion coefficients from nonequilibrium processes without Jarzynski's equality. *J. Chem. Phys.* 124:64106.
51. Ma, B., C. Tsai, and N. Ruth. 2000. A systematic study of the vibrational free energies of polypeptides in folded and random states. *Biophys. J.* 79:2739–2753.
52. Sugase, K., H. J. Dyson, and P. E. Wright. 2007. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature*. 447:1021–1025.
53. Roe, D. R., A. Okur, L. Wickstrom, V. Hornak, and C. Simmerling. 2007. Secondary structure bias in generalized Born solvent models: comparison of conformational ensembles and free energy of solvent polarization from explicit and implicit solvation. *J. Phys. Chem. B*. 111:1846–1857.
54. Plischke, M., and B. Birger. 2003. *Equilibrium Statistical Physics*. World Scientific, Singapore/USA.
55. Reference deleted in proof.
56. Cowan, N. J., and L. Dudley. 1983. Tubulin isotypes and the multigene tubulin families. *Int. Rev. Cytol.* 85:147–173.
57. Jensen-Smith, H. C., R. F. Luduena, and R. Hallworth. 2003. Requirement for the  $\beta$ I and  $\beta$ IV tubulin isotypes in mammalian cilia. *Cell Motil. Cytoskeleton*. 55:213–220.
58. Sullivan, K. F., and D. W. Cleveland. 1986. Identification of conserved isotype-defining variable region sequences for four vertebrate  $\beta$  tubulin polypeptide classes. *Proc. Natl. Acad. Sci. USA*. 83:4327–4331.
59. Ranganathan, S., D. W. Dexter, C. A. Benetatos, A. E. Chapman, K. D. Tew, and G. R. Hudes. 1996. Increase of  $\beta$ (III)- and  $\beta$ (IVa)-tubulin isotypes in human prostate carcinoma cells as a result of estramustine resistance. *Cancer Res.* 56:2584–2589.
60. Kavallaris, M., C. A. Burkhart, and S. B. Horwitz. 1999. Antisense oligonucleotides to class III  $\beta$ -tubulin sensitize drug-resistant cells to Taxol. *Br. J. Cancer*. 80:1020–1025.
61. Mozzetti, S., C. Ferlini, P. Concolino, F. Filippetti, G. Raspaglio, S. Prislei, D. Gallo, E. Martinelli, F. O. Ranelletti, G. Ferrandina, and G. Scambia. 2005. Class III  $\beta$ -tubulin overexpression is a prominent mechanism of paclitaxel resistance in ovarian cancer patients. *Clin. Cancer Res.* 11:298–305.
62. Kavallaris, M., A. S. Tait, B. J. Walsh, L. He, S. B. Horwitz, M. D. Norris, and M. Haber. 2001. Multiple microtubule alterations are associated with Vinca alkaloid resistance in human leukemia cells. *Cancer Res.* 61:5803–5809.
63. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph* 14:33–38, 27–28.
64. Hess, B. 2002. Determining the shear viscosity of model liquids from molecular dynamics simulations. *J. Chem. Phys.* 116:209–217.