Relating Molecular Flexibility to Function: A Case Study of Tubulin

Ozlem Keskin,*† Stewart R. Durell,‡ Ivet Bahar,§ Robert L. Jernigan,‡ and David G. Covell*

*Computational Technologies Laboratory, Screening Technologies Branch, Developmental Therapeutics Program, National Cancer Institute–Frederick, National Institutes of Health, Frederick, Maryland 21702 USA; †College of Arts and Sciences, Department of Chemistry, Koc University, Rumelifeneri Yolu, 80910 Sariyer, Istanbul, Turkey; ‡Molecular Structure Section, Laboratory of Experimental and Computational Biology, Division of Basic Sciences, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892 USA; and §Center for Computational Biology and Bioinformatics, and Department of Molecular Genetics and Biochemistry, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15213 USA

ABSTRACT Microtubules (MT), along with a variety of associated motor proteins, are involved in a range of cellular functions including vesicle movement, chromosome segregation, and cell motility. MTs are assemblies of heterodimeric proteins, $\alpha\beta$-tubulins, the structure of which has been determined by electron crystallography of zinc-induced, pacilitaxel-stabilized tubulin sheets. These data provide a basis for examining relationships between structural features and protein function. Here, we study the fluctuation dynamics of the tubulin dimer with the aim of elucidating its functional motions relevant to substrate binding, polymerization/depolymerization and MT assembly. A coarse-grained model, harmonically constrained according to the crystal structure, is used to explore the global dynamics of the dimer. Our results identify six regions of collective motion, comprised of structurally close but discontinuous sequence fragments, observed only in the dimeric form, dimerization being a prerequisite for domain identification. Boundaries between regions of collective motions appear to act as linkages, found primarily within secondary-structure elements that lack sequence conservation, but are located at minima in the fluctuation curve, at positions of hydrophobic residues. Residue fluctuations within these domains identify the most mobile regions as loops involved in recognition of the adjacent regions. The least mobile regions are associated with nucleotide binding sites where lethal mutations occur. The functional coupling of motions between and within regions identifies three global motions: torsional and wobbling movements, en bloc, between the α- and β-tubulin monomers, and stretching longitudinally. Further analysis finds the antitumor drug pacilitaxel (TaxotereR) to reduce flexibility in the M loop of the β-tubulin monomer; an effect that may contribute to tightening lateral interactions between protofilaments assembled into MTs. Our analysis provides insights into relationships between intramolecular tubulin movements of MT organization and function.

INTRODUCTION

Although numerous proteins are now structurally characterized by crystallographic and nuclear magnetic resonance studies, relatively little is known about their dynamic properties. Crystallographic studies have, however, provided snapshots of multiple conformational states for many enzymes; these states are typically induced by associations with either small ligands or other macromolecules. There is considerable interest in obtaining relationships between dynamic properties and molecular function. Recent advances in spectroscopic techniques permit detailed examinations of individual atomic motions (Weiss, 1999) and the global motions associated with folding and unfolding (Arrington and Robertson, 2000). Large-scale molecular motions are associated with, for example, motor proteins known to exist as large macromolecular aggregates. Previously, a number of computational studies based on applications of classical mechanics and molecular simulations (Brooks et al., 1995) were reported. Although, in principle, these computational methods hold the promise of defining relationships between molecular motions and biological function, there has actually been only limited success toward this goal, especially for slow large-scale motions of large proteins or proteins existing in multimeric assemblies. To address this limitation, we have developed a computational approach, derived from classical mechanics, which can be applied to large macromolecular dynamics. Here, we will examine the global motions associated with the complex polymers known as microtubules (MTs).

MTs are used for chromosomal segregation during cell division, for intracellular transport, and for a diverse range of specialized cellular functions. MTs represent an important target for cancer therapy, inasmuch as tubulin binding agents represent a therapeutic class of compounds with broad activity against solid and hematologic neoplasias (Chabner and Longo, 1996). Their mechanisms of action are thought to occur by directly targeting MTs and affecting polymerization, which is modulated partly by exchangeable ligand nucleotides, specific MT-associated proteins and chemotherapeutic agents. Based on the atomic structure of the α/β-tubulin heterodimer, a simple model is developed to examine the range of motions predicted for tubulin as it appears in its monomeric and dimeric forms. Our analysis will specifically examine dynamic protein movements within the α/β-tubulin dimer in the context of MT function and mechanism of action for taxoid-based chemotherapeutics.

Submitted October 3, 2001 and accepted for publication March 19, 2002.

Address reprint requests to Ozlem Keskin or David Covell, Ft. Detrick Bldg 430, Room 215, Frederick, MD 21702. Tel.: 301-846-5785; Fax: 301-846-5762; E-mail: covell@helix.nih.gov or okeskin@ku.edu.tr.

0006-3495/02/08/663/18 $2.00

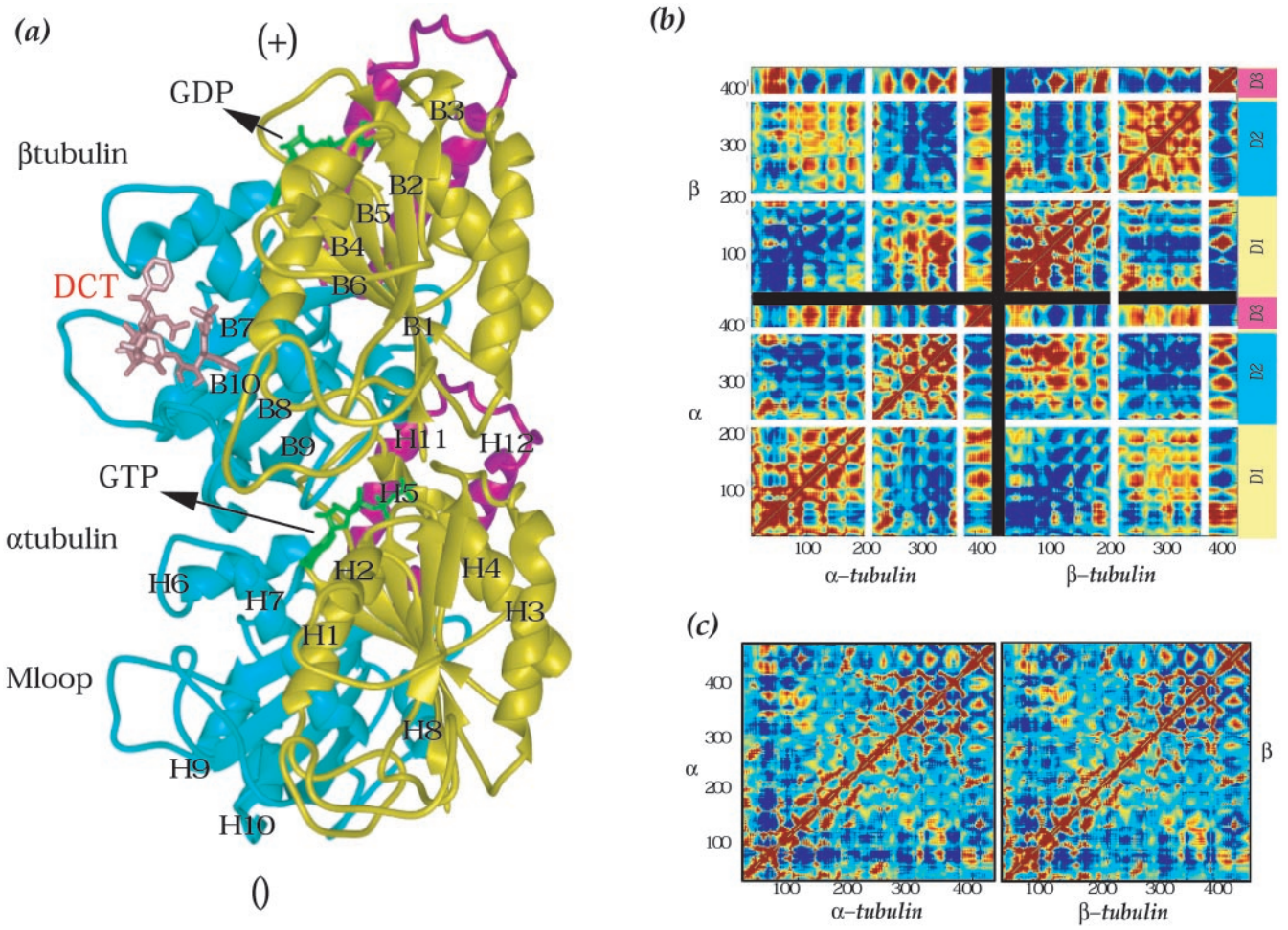

FIGURE 1 (*a*) Ribbon diagrams of the $\alpha\beta$ tubulin dimmer (Nogales et al., 1998a). The protein consists of two subunits: GTP-bound α-tubulin ($-$ end), and GDP-bound β-tubulin ($+$ end). The two monomers have 40% sequence identity, and their structures are similar except for a few differences in the loops. Each monomer contains a pair of central β-sheets surrounded by α-helices. The secondary structural elements are labeled H1–H12 (helices), and B1–B10 (β-strands). Each monomer consists of three domains: nucleotide-binding domain having a typical Rossmann fold (D1, residues 1–205, *yellow*), drug-binding domain (D2, residues 206–381, *blue*) and MT-associated protein-binding domain (D3, residues 382–440, *magenta*). These are also called the N-terminal, intermediate, and C-terminal domains, respectively. The anticancer drug pacilitaxel (DCT), bound to β-tubulin, is displayed in pink; GTP and GDP in *green*. Correlation map between residue fluctuations in the dimer (*b*) and monomers (*c*) calculated over all $3N$-6 modes. The two axes refer to residue indices, and the colors indicate different strengths and types of inter-residue orientational correlation functions (Eq. 9). Blue and red regions correspond to negatively (opposite-direction) and positively (same-direction) correlated motions, respectively. The color-coded domains are also displayed on the right-hand side of the upper map. The monomeric sequences are separated by black lines for clarity in the upper map. The white lines enclose the three functional domains, to emphasize their strong intradomain coupling. (*c*) Comparison with the map for (*left*) α- and (*right*) β-tubulins in monomeric forms shows that intradomain and interdomain couplings become substantially enhanced in the dimeric form. These increased couplings induced in the dimeric forms may have important functional implications (see text).

TUBULIN STRUCTURE

Figure 1 *a* displays the ribbon diagram of the $\alpha\beta$-tubulin dimer obtained at a resolution of 3.7 Å by electron crystallography of zinc-induced, pacilitaxel-stabilized tubulin sheets (Nogales et al., 1998a), (PDB code 1tub). The overall folds of the α- and β-monomers are nearly identical, composed of two β-sheets of six and four strands, flanked by 12 α-helices. Structure-based sequence alignment (Nogales et al., 1998b) shows the differences between the α- and β-monomers to consist of two gaps in the β-monomer at sequence positions 45–46 and 361–368 of the α-tubulin. The larger gap, corresponding to a longer loop in α-tubulin, is the specific pacilitaxel-binding site in β-tubulin.

Three functional domains have been assigned by Nogales and coworkers (1998a) to each tubulin monomer. The N-terminal, intermediate, and C-terminal domains comprise residues 1–205, 206–381, and 382–440. These contain the respective nucleotide, pacilitaxel, and MT-associated proteins binding sites; and are colored yellow, blue, and ma-

TABLE 1 Functionally important segments in tubulin (Nogales et al., 1998b)*

Elements†	Residues	Function	Flexibility‡
T1	9–11	Nucleotide binding, interaction with the phosphates (α/β: Gly-10, Gln-11)	R
H1	12–23	Nucleotide binding, interaction with guanine (β: Cys-12, Gln-15, Ile-16)	R
Loop H1–B2	24–63	Lateral contacts	F
H2–B3	79–89	Lateral contacts	F
T2	69–71	Interaction with β and γ phosphates of nucleotides (β: Asp-69, Leu-70, Glu-71)	R
T3	98–101	Interaction with γ phosphate (β: Gly-98, Ala-99, Gly-100, Asn-101)	R
H3	114–132	Longitudinal contacts	F
T4	140–145	Signature (Glycine rich) loop. Backbone interactions with α and β phosphates	R
H4–T5	157–161	Lateral contacts	F
T5 and B5	162–181	Ribose binding loop: (β: Val-171, Ser-178, Asp-179); interaction with H11 (α/β: 384–396) and the monomer upstream	R
T6	205–206	H-bonds with guanine base (β: Asn-206)	R
H7	226–239	Hydrophobic contact of conserved residues (β: Tyr-224) with guanine	R
T7	240–252	Nucleotide contacts (α/β: 239–251)	R
H10	327–341	Longitudinal and lateral contacts	F
M loop	272–288	Lateral contacts (with H3, C terminus of H2–B3 loop, and H1–B2 loop in the adjacent protofilament)	F
C-terminal	381–440	MAP-binding domain	F
	α: Glu-254 β: Lys-254	Close proximity to the phosphates of the neighboring tubulin monomer	R
	α: Lys-40	Acetylation site (α-subunit)	F

*See Figure 1 *a*.
†T, B and H refer to the loops, β-sheets, and α-helical secondary structures, respectively.
‡R and F refer to segments exhibiting the least and most highly fluctuating regions, respectively.

genta in Fig. 1 *a*, respectively. Although dividing tubulin's monomers into three segments, these domain definitions do not precisely define the binding pockets for tubulin's ligands, because, for example, the bound nucleotide is known to interact with the helix H7 of the intermediate domain, and paclitaxel interacts with portions of the N-terminal domain (Nogales et al., 1998a). The N-terminal domain forms a Rossmann fold, which is typical of nucleotide-binding proteins. Helices H1 and H2 (using, hereafter, the secondary-structure labeling of Nogales et al., 1998a) are on one side of the sheet, whereas helices H3, H4, and H5 are on the other. Strands B1–B3 and their long connecting loop are located adjacent to helices H1–H2, overall forming the peptide segment 1–95. These loops can accommodate amino-acid substitutions to form tubulin isoforms (Burns and Surridge, 1993, Luduena, 1998). The intermediate domain contains a mixed β-sheet and five surrounding α-helices, among which H8 lies at the longitudinal interface between monomers. The loop connecting B7 and H9 in α-tubulin also makes strong lateral contacts between protofilaments in the assembled MT (shown as M-loop in Fig. 1 *a*). The loop connecting H11 and H12 is believed to be important for interactions with the next monomer along the protofilament (Nogales et al., 1998a).

MTs are polar structures with one end (+) capable of growth by addition of a dimeric subunit, whereas the opposite end (−) loses subunits when not stabilized by being embedded into the centrosome (Alberts et al., 1994). The β-tubulin monomer is located at the plus (+) end of the MT, the site where polymerization and depolymerization with the next dimer occurs, whereas α-tubulin forms the minus (−) end. GTP binds to α- and β-tubulins. GTP bound to β-tubulin is both exchangeable and is hydrolyzed to GDP when an additional subunit is added to the MT (Desai and Mitchison, 1997).

Sequence identity between the two monomers of α/β-tubulin is 40%. Sequence is conserved within selected regions of secondary structure, and in the loops making the lateral and longitudinal contacts between protofilaments of the assembled MT. Likewise, the loops involved in nucleotide binding are highly conserved, including, in particular, the signature motif of tubulins, a glycine-rich loop GGGTGSG. Tubulin shares sequence homology with the protein FtsZ (filamentation temperature sensitive protein Z), a major cytoskeletal protein involved in bacterial cell division, also belonging to the GTPase family (Nogales et al., 1998b). Based on sequence alignment with FtsZ, Nogales et al. (1998b) proposed the segments listed in Table 1 to be critical for function. This list includes regions making lateral and longitudinal contacts between and within protofilaments (Nogales et al., 1999) and segments with specific binding functions. Residue numbers in this table refer to the numbering of α-tubulin. Our analysis will reexamine sequence conservation within the tubulin family and explore relationships between conserved regions and those residues identified here to play an instrumental role in coordinating the collective motions of the molecule.

IDENTIFICATION OF COLLECTIVE MOTIONS

The slow, large-amplitude motions, also referred to as global motions, are conceived to be essential for protein

function (Amadei et al., 1993; Bahar et al., 1998; Hinsen, 1998; de Groot et al., 1998). It is known that the low-frequency modes make major contributions to thermal conformational fluctuations (Brooks et al., 1995; Go, 1990). Such motions can influence the interactions of proteins with other molecules and its environment. Higher frequency fluctuations, in contrast, are more localized, involve only a few residues, and can play an important role in signal transmission, enzyme reactions, and other internal processes. One standard technique for studying protein dynamics, and, in particular, low-frequency domain motions, is normal mode analysis (NMA) (Case, 1994; Hinsen et al., 1999; Thomas et al., 1996). Comparisons of the low-frequency normal modes with large amplitude fluctuations observed in molecular dynamics simulations of proteins in their native state indicate common mechanisms of conformational fluctuations, consistent with the robustness of the global motions for a given molecular architecture (Amadei et al., 1993; Thomas et al., 1996; Hayward et al., 1997; Kitao and Go, 1999; Doruker et al., 2000).

Recent coarse-grained studies of vibrational motions in globular proteins have provided significant insights into the dynamics of folded proteins (Bahar et al., 1997; ben Avraham, 1993; Tirion, 1996). Ben Avraham (1993) found that the density distribution of the slow modes follows a characteristic, universal curve when expressed as a function of frequency. The existence of such a universal curve was attributed to underlying structural similarities between proteins (ben Avraham, 1993). This behavior is also consistent with the dominance of nonspecific, generic, inter-residue interactions that stabilize native protein structures (Bahar and Jernigan, 1997). In the normal mode analysis of G-actin, Tirion (1996) adopted a single-parameter Hookean potential for the pairwise interactions between atoms, and was able to satisfactorily reproduce the vibrational dynamics predicted when detailed atomic potentials are used. Along the same line, Bahar et al. (1997) proposed a network model for folded proteins in which the interactions between residues in close proximity are replaced by linear springs. This so-called Gaussian Network Model (GNM) assumes that a protein in its folded state is equivalent to a fully elastic network, along the lines of the elasticity theory of random polymer networks (Flory, 1976). The equilibrium fluctuations of C^{α} atoms for several proteins as predicted by the GNM were in close agreement with x-ray crystallographic measurements (Bahar et al., 1997; Haliloglu et al., 1998; Demirel et al., 1998; Bahar et al., 1999; Keskin et al., 2000). Recent applications of the GNM have shown the use of the method in finding the dynamic features of the systems from their folded structures. Some of these proteins or complexes include tRNA-synthetase (Bahar and Jernigan, 1998), HIV-1 reverse transcriptase (Bahar et al., 1999), HIV-1 protease (Bahar et al., 1998), apomyoglobin (Haliloglu et al., 1998), cofactor binding fragment of CysB, the lysine/arginine/ornithine binding protein, the enzyme porphobilinogen deaminase, and ribose binding protein (Keskin et al., 2000).

In principle, GNM is equivalent to an NMA with three major simplifying assumptions: identical force constants for all interacting residues, coarse graining of the protein structure, usually taken at one point per residue, and absence of directional effects; all fluctuations being assumed to be isotropic. GNM thus predicts the relative amplitudes of fluctuations but cannot address the directionality of these motions. In the present study, we extend the GNM approach to include the anisotropic effects on fluctuation dynamics. We will decompose molecular motions into a series of $3N$-6 modes (for a protein of N residues) and focus on the lowest frequency modes. The fluctuations near the native state will be shown to be easily obtainable from a Hessian matrix based on the harmonic potential energies of a bead-and-spring model. This general analytical method can be used to explore the intrinsic fluctuations and the collective motions of biomolecular assemblies. We have recently applied the method to GroEL/GroES complex with 8015 residues to study its dynamic features (Keskin et al., 2002).

THEORY

In general, the total intramolecular potential V of a protein of N residues may be expressed as a series expansion in the fluctuations ΔR_i of individual residue positions as

$$V = V_0 + \sum_{i=1}^{N} \left(\frac{\partial V}{\partial \Delta \mathbf{R}_i}\right)_0 \Delta \mathbf{R}_i + \frac{1}{2}\sum_{j=1}^{N}\sum_{i=1}^{N} \left(\frac{\partial^2 V}{\partial \Delta \mathbf{R}_i \partial \Delta \mathbf{R}_j}\right)_0 \Delta \mathbf{R}_i \Delta \mathbf{R}_j + \cdots, \quad (1)$$

where the superscript zero refers to a reference state, here the crystal form. Inasmuch as the crystal structure represents an equilibrium state (or an energy minimum), the first derivative of V with respect to residue positions is zero, i.e., $(\partial V/\partial \Delta \mathbf{R}_i)_0 = 0$. Also, setting the reference state potential $V_0 = 0$, the total energy is approximated as

$$V = \frac{1}{2}\sum_{i}\sum_{j}\left(\frac{\partial^2 V}{\partial \Delta \mathbf{R}_i \partial \Delta \mathbf{R}_j}\right)_0 \Delta \mathbf{R}_i \Delta \mathbf{R}_j, \quad (2)$$

by neglecting the higher-order terms in Eq. 1, as appropriate for small fluctuations. This equation can be rewritten in terms of the Hessian matrix ($\mathbf{H}$) of the second derivatives of V as

$$V = \tfrac{1}{2}[\Delta \mathbf{R}]^T \mathbf{H} [\Delta \mathbf{R}]. \quad (3)$$

Here $\Delta \mathbf{R}$ represents the $3N$-dimensional vector composed of the N fluctuation vectors $\Delta \mathbf{R}_i$, and the superscript T designates its transpose.

A schematic representation of the model used in the computations here and in recent studies (Atilgan et al.,

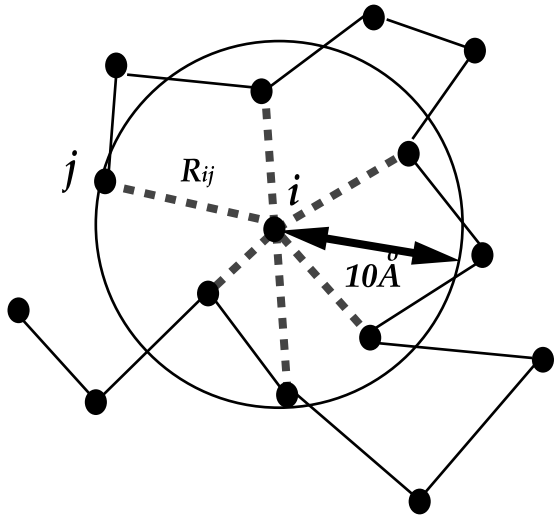

FIGURE 2 Schematic representation of the coarse-grained network model used in this study. The protein is represented by a virtual backbone (*solid lines*), each interaction site corresponding to an α-carbon. Residue pairs are subject to a harmonic force provided that they are within a cutoff distance range of 10 Å around the residue *i*. Dotted lines designate the residues interacting with that residue.

2001; Keskin et al., 2002) is given in Fig. 2. The dots correspond to the residues, each conveniently identified by its C^{α} atom coordinates in the crystal structure. $\mathbf{R}_{ij}$ designates the vector pointing from the *i*th residue along the sequence. Only the residues within a selected cutoff distance (R_c) in the neighborhood of a central residue are assumed to interact. This decreases the number of interacting pairs from its maximum of $N(N-1)/2$ to a number proportional to N. Knowledge-based studies find a cutoff distance of 6.5 Å for interacting pairs within a first-interaction shell (ben Avraham, 1993; Miyazawa and Jernigan, 1985). Here we adopted a relatively longer cutoff distance, $R_c = 10$ Å, to include a somewhat broader range of interactions given by

$$V = \frac{1}{2}\gamma \sum_i \sum_j (\Delta\mathbf{R}_i - \Delta\mathbf{R}_j) \cdot (\Delta\mathbf{R}_i - \Delta\mathbf{R}_j) * (R_c - R_{ij}), \quad (4)$$

where $\cdot$ designates the dot product, and $(R_c - R_{ij})$ is the Heaviside step function. $(R_c - R_{ij})$ assumes the value 1 if its argument is positive (i.e., $R_{ij} \le R_c$) and 0, otherwise. Eq. 4 selects nonbonded residue neighbors in the first coordination shell around a central residue and the first and second bonded neighbors along the backbone, to account for chain connectivity (Bahar et al., 1998). The nonbonded interactions are defined by the folded state, whereas bonded interactions play a dominant role in defining the overall molecular machinery and topology. The potential V thus accounts for interactions between all close residue pairs.

The $3N \times 3N$ Hessian matrix can be readily calculated from the above potential. It is conceptually easier to conceive of it as an $N \times N$ matrix of 3×3 superelements defined as $\mathbf{H}_{ij} = \partial^2 V/\partial\Delta\mathbf{R}_i\partial\Delta\mathbf{R}j$. Using the classical statistical mechanical theory of networks (Flory, 1976), it can be shown that the mean-square fluctuations of junctions scale with the inverse of $\mathbf{H}$ as

$$\langle(\Delta R_i)^2\rangle = kT \,\mathrm{trace}[\mathbf{H}^{-1}]_{ii}. \quad (5)$$

Here $[\mathbf{H}^{-1}]_{ii}$ is the *i*th diagonal superelement (3×3 matrix) of the inverse $\mathbf{H}^{-1}$, and "trace" designates the sum of the diagonal elements of this 3×3 matrix. Eq. 5 will be used in the present study for calculating the mean-square fluctuations of individual residues; whereas the fluctuations in the global motions will be extracted by a mode analysis. It should be noted that the determinant of $\mathbf{H}$ is zero. Therefore, $\mathbf{H}$ cannot be inverted directly. $\mathbf{H}^{-1}$ is an approximation calculated from the $3N$-6 nonzero eigenvalues, λ_m, and the corresponding eigenvectors, $\mathbf{u}_m$, of $\mathbf{H}$. A major limitation with full atomic NMA is the large computational time, here we use a coarse-grained approximation and reduce the dimensions of $\mathbf{H}$ to $3N \times 3N$ where N is the number of residues, whereas, in full atomic calculations, the Hessian matrix has a dimension of $3n \times 3n$ where n is the number of atoms in the system. Extremely large proteins can therefore be examined with this approximation. The computational time for the calculation of $3N$ ($3 \times 867 = 2601$) normal modes for tubulin with the present model requires around 2 h on a SGI R10000. It would require almost 2 weeks to compute the full atomic normal modes on a Cray SVI-4/96–96 Supercomputer.

Information about global dynamics is obtained by decomposing the motions into a series of modes, and concentrating on the modes at the slowest (largest amplitude) end of the spectrum. To elucidate the mechanism of these modes, $\mathbf{H}$ is rewritten as the product of three matrices, the diagonal matrix of nonzero eigenvalues λ_m ($1 \le m \le 3N$-6), the matrix $\mathbf{U}$ of the corresponding eigenvectors $\mathbf{u}_m$, and the transpose of $\mathbf{U}$, i.e.,

$$\mathbf{H} = \mathbf{U}\Lambda\mathbf{U}^{-1} = [\mathbf{u}_1\mathbf{u}_2\mathbf{u}_3 \ldots \mathbf{u}_{3N-6}] \times \mathrm{diag}(\lambda_1\lambda_2\lambda_3 \ldots \lambda_{3N-6})[\mathbf{u}_1\mathbf{u}_2\mathbf{u}_3 \ldots \mathbf{u}_{3N-6}]^T. \quad (6)$$

We note that $\mathbf{U}$ is an orthonormal matrix, i.e., $\mathbf{U}^T = \mathbf{U}^{-1}$, and the inverse of $\mathbf{H}$ can be readily written as $\mathbf{H}^{-1} = \mathbf{U}\Lambda^{-1}\mathbf{U}^{-1}$. The latter can alternatively be written as a sum of $3N$-6 matrices of size $3N \times 3N$, each representing the contribution of a single internal mode,

$$\mathbf{H}^{-1} = \sum_m [\lambda_m^{-1}\mathbf{u}_m\mathbf{u}_m^T]. \quad (7)$$

This equation provides a simple means of decomposing the dynamics into a series of modes. For example, the mean

square fluctuation of residue i driven by the mth mode of motion is calculated using

$$[(\Delta \mathbf{R}_i)^2]_m = 3kT\ \mathrm{Tr}[\lambda_m^{-1}\mathbf{u}_m\mathbf{u}_m^T]_{ii}. \quad (8)$$

The orientational cross-correlations (Fig. 1, *b* and *c*) between the fluctuations of C^α are found by normalizing the cross-correlations according to

$$C(i,j) = \frac{\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle}{\{\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_i \rangle \langle \Delta \mathbf{R}_j \cdot \Delta \mathbf{R}_j \rangle\}^{1/2}}, \quad (9)$$

where $\langle \Delta \mathbf{R}_I \cdot \Delta \mathbf{R}_j \rangle$ is calculated from $\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle = kT$ trace$[\mathbf{H}^{-1}]_{ij}$, by suitable choice of indices in Eq. 5. The positive and negative limits of $C(i, j)$ are 1 and -1, and correspond to pairs of residues exhibiting fully correlated (same direction, same sense) and fully anticorrelated (same direction, opposite sense) motions, respectively. Zero correlation refers to uncorrelated, or orthogonal motions.

The most dominant mode motions and shapes are very robust regardless of the resolution of the crystallographic protein structures. These modes give information on the global dynamics of proteins that are the most cooperative motions and consequently reflect mostly the overall shape. Explorations of the sensitivity of our results with respect to coordinate positions revealed that coarser-grained models based one site per 2, 5, and 10 residues still conserve the global mode motions and shapes found with models based on one point per α carbon. Further studies based on uniformly distributed perturbations to residue coordinates, in the range of 0.5 to 3 Å per position, find that the most dominant three modes remain significantly correlated ($r > 0.83$, $p < 0.01$) with calculations based on the reported coordinate positions themselves. Additional attempts to define each residue's local neighborhood, based on establishing the network of connectivities between residues from contacting Hessian matrix calculated from inter-residue buried surfaces area yielded no significant differences from the present results. Taken together, these results suggest the robustness of our method irrespective of the precise details of residue position.

In summary, ANM identifies the mechanisms of motions of a protein structure in its native state. Motions provided by these calculations are deterministic (unique) as explained in the preceding paragraphs, in contrast to ensembles of motions obtained from molecular dynamics. The information of the global motions of tubulin here is found with a coarse-grained model at the expense of specificity loss at the atomic level.

RESULTS AND DISCUSSION

Analysis of domain motions

Correlations between residue fluctuations describe those parts of the structure that move collectively, as a unified group, and how these regions move with respect to one another. The correlations obtained over all modes of motion in the present study are illustrated in Fig. 1 *b*, for the tubulin dimer, where the axes represent the sequences for α- and β-tubulin. Blue and red colors in this map indicate negatively and positively correlated regions, respectively, whereas yellow refers to regions with uncorrelated fluctuations. For convenience, a vertical bar has been placed along the right-hand side of the map to indicate regions D1–D3 of each monomer, color-coded according to the regions shown in the ribbon diagram in Fig. 1 *a*.

Three highly correlated regions are observed for each of the tubulins in the dimeric form. These regions are enclosed in white outlines on the map for visual clarity. The partitioning of the structure into three such regions on the basis of its fluctuation dynamics is reasonably consistent with the equilibrium structure found by Nogales et al. (1998a). Additionally, the relative motions between domains finds that the N-terminal, intermediate, and C-terminal domains of the α-subunit are negatively correlated (or anticorrelated) with their counterparts in the β-subunit. This means that each tubulin monomer shows symmetrical, opposite direction, movements with respect to the dimer interface. Meanwhile within each separated monomer, the N- and C-termini exhibit some tendency to move in the same direction (except for the peptide segment 10–80) whereas the intermediate domain of the molecule undergoes mostly opposite direction motions.

The concerted motions observed for the dimer are substantially absent when analyses are performed separately on each monomer. Inspection of the correlation maps based on fluctuation calculations for individual monomers, as shown in Fig. 1 *c*, does not reveal the distinctive blocks indicative of correlated motions of the three regions found in the dimer. Consistent with the observation that the α- and β-tubulin monomers are structurally similar, their correlation maps are nearly identical. The noteworthy exception being peptide segment 45–55 in β-tubulin, having more concerted movements than the corresponding residues in α-tubulin. Therefore, strong intradomain and interdomain couplings are observed in the dimer, whereas, in the monomers, residues are rather correlated only with their near neighbors along the backbone chain. Local connectivity appears as the main force controlling these motions in the monomeric form, whereas in the dimer, long-range (nonbonded) effects become important. This result strongly suggests that dimerization enhances the coherent movement of functional domains in tubulin.

Inspection of the correlation map in Fig. 1 *b* clearly suggests that similar patterns of motion are observed between distantly separated residues along the protein sequence. These collective motions between sequence-separated residues are the basis for the patchwork appearance of the correlation map in Fig. 1 *b*. In an effort to group together residues with the most similar correlated patterns, average

linkage hierarchical cluster analysis has been performed on the calculated correlations. Using the MATLAB Statistics tools to calculate the Minkowski distances among all rows of the correlation matrix, a Ward's-based cluster tree was determined. The results of this cluster analysis, as shown in Fig. 3, identify six regions of motion within the tubulin dimer. The selection of six regions was based on repeated hierarchical clustering of randomly perturbed tubulin coordinates and selection of the number of clusters necessary for at least 90% overlap between cluster memberships. In other words, the clustering was repeated with slightly perturbed coordinates to determine the most robust set of cluster boundaries. In this analysis, fewer clusters lead to improved membership overlap, whereas more clusters substantially degrade the fraction of residues common to each cluster. Residues defining cluster boundaries are not affected by the number of clusters; except insofar as residues at the boundaries between regions disappear when smaller regions are combined into a larger region, for fewer clusters, or new boundary residues appear if more clusters are selected. The ribbon diagram in Fig. 3 a, identifies these six regions, by color, while preserving the tubulin view displayed in Fig. 1 *a*. Figure 3 *b* displays the correlation map shown in Fig. 1 *b*, but reordered to reflect the cluster memberships for the six regions. These regional descriptions divide the tubulin dimer and the companion correlation map into distinct parts according to their collective motions. The bar at the right of the correlation map is colored according to the regions identified in the ribbon structure. The two smallest regions, S1 (*red*) and S2 (*yellow*) consist of parts at the tail (−) and head (+) of the tubulin dimer, respectively, and appear at the lower left corner of the reorganized correlation map. These two regions exhibit strong positively correlated motions within and between each other. The largest region, S6 (*dark blue*), consists of residues involved in the dimer interface, and appears at the upper right of the correlation map. Like regions S1 and S2, this region exhibits positively correlated motions within itself, indicative of collective motion at the dimer interface, while also sharing positively correlated motions with portions of its neighboring regions, S5 (*light blue*) and S4 (*dark green*). S6 is strongly negatively correlated with regions S1, S2, and S3 to suggest an overall motion for the tubulin dimer in which the central region at the dimer interface moves opposite to the motions observed at the head and tail of the dimer. The regions separating the head and tail from the interfacial part exhibit a complex range of motions that are both positively and negatively correlated within themselves and with the other regions. These regions appear to act as a buffer region between the correlated motions of regions S1, S2, and S3 with region S6. Comparison of these six regions with the three functional domains identified in Fig. 1 *a* indicates considerable overlap, however portions of D1–D3 now appear in more than one of the six regions, S1–S6.

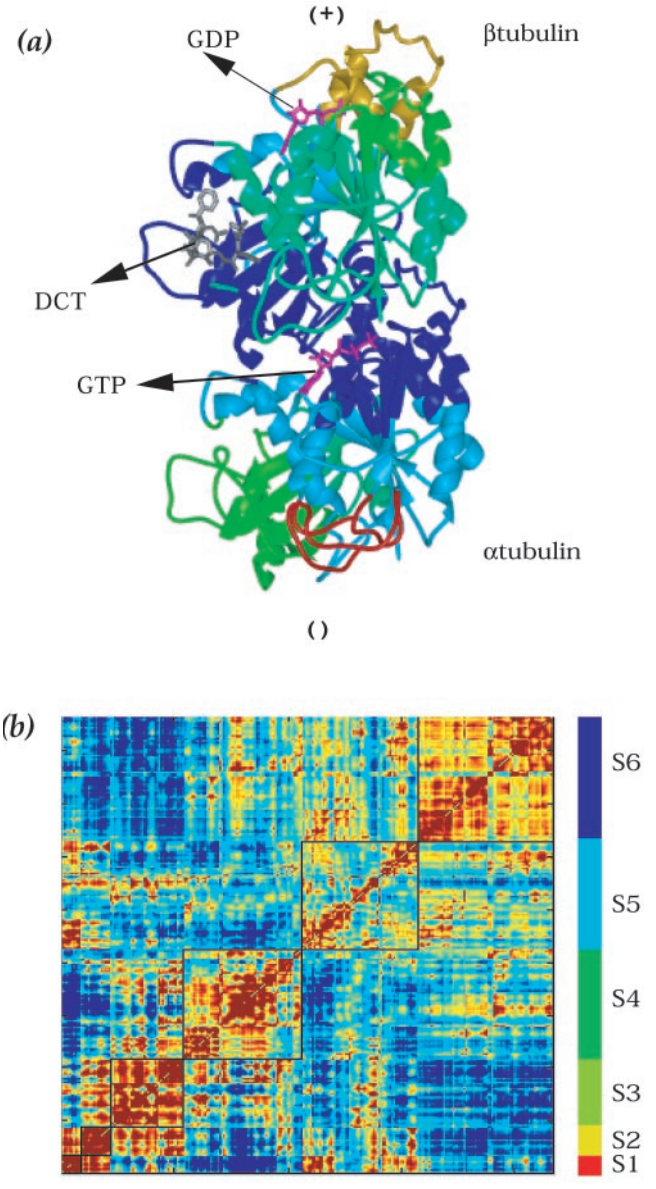

FIGURE 3 (*a*) Ribbon diagram illustrating the similar correlated motions within the dimer. Pacilitaxel and the nucleotides are colored in gray and magenta, respectively. Six regions of motion are colored differently. Each monomer exhibits similar types of motions (*dark blue*) at the dimer interface. The six regions were obtained by hierarchical clustering of the residue correlation map shown in Fig. 1 *b*. (*b*) The reordered correlation map. The regional boundaries are separated by black lines. The bar on the right-hand side of the map identifies the coloring of the dimer ribbon shown in (*a*) and shows region-specific labels S1–S6.

Relationship between domains and secondary structure

The identification of the six regions of motion according to tubulin's sequence and its associated secondary structure is shown in Fig. 4. In contrast to the contiguous domain definitions of Nogales et al. (1998a), the six regions identified here, some of which appear as nearby structural neigh-

bors in Fig. 3 *a*, comprise both near and distant sequence neighbors. The location of secondary-structure elements along the sequence finds a very close correspondence between region boundaries, as indicated by the locations of the changes in color along the sequence, and secondary structure. Of the 22 secondary-structure elements identified in α-tubulin, all but two of them, B6 and H9, have regional boundaries within or at the termini of their secondary structures, whereas, except for the N-terminal region of β-tubulin, all but three of its secondary structures, H8, H9, and B9 are associated with regional boundaries. Further inspection of Fig. 4 reveals a close relationship between secondary structure and regions of minimal residue fluctuation. The symbols ($*$) and ($\wedge$) below the tubulin sequence in Fig. 4, represent the least and most fluctuating residues in tubulin, based on the analysis presented later in the text. The noteworthy observation, here, is that all secondary structures, with the exception of H10 in α- and β-tubulin, are associated with fluctuation minima. Conversely, both H10 and B3, and most of the solvent-exposed loops in each monomer correspond to regions of highest fluctuations.

To explore apparent relationships between the tubulin sequence and either its secondary-structure elements or our regional boundaries, sequence analysis was performed on the tubulin family. By using both the α- and β-sequences as reference probes, a FASTA query of the nonredundant database of protein sequences (Genetic Computer Group (GCG), Wisconsin Package, version 9) identified 1055 different sequences. Extracting only those nonidentical sequences with lengths within ± 50 residues and having sequence identity greater than 30% with α-tubulin, 511 tubulin sequences were identified. Multiple alignments of these sequences were conducted using the PILEUP utility in the GCG package. The sequence similarity of these alignments was completed using the PLOTSIMILARITY utility in GCG, based on the BLOSSUM62 scoring matrix. The most highly conserved sequence positions are identified in magenta colored capital letters along the sequences shown in Fig. 4, *A* and *B*. Consistent with the previous analysis of Erickson (1998) the most conserved regions involve the nucleotide binding sites (Nogales et al., 1998b), identified as boxed resides in Fig. 4, and the signature recognition loop, GGGTGSG; all of which are associated with regions corresponding to fluctuation minima. A catalog of the correspondence between regional boundaries and conserved sequence positions shows that only the regional boundaries near H9, H10, and B10 of both α- and β-tubulin lack nearby regions of sequence conservation. The majority of conserved regions are found at positions of nucleotide binding; most of which appear at fluctuation minima.

Despite the absence of an apparent relationship between sequence conservation and secondary structure, it may be reasonable to speculate that the overall folded geometry of secondary-structure elements might be conserved within the tubulin family, and this may be essential for tubulin function. Our finding is that regional boundaries occur within secondary structures, which themselves are at minimally fluctuating positions, to suggest a coupling between super-secondary structure and tubulin's global motions. Additional inspection of amino acid types around our regional boundaries finds a high incidence of hydrophobic residues. The coexistence of conserved super-secondary architecture, which lacks sequence conservation, aside from the tendency to position hydrophobic residues at regional boundaries, is consistent with the concept that secondary structures may actively choreograph global motions. In this view, the secondary structures may act as anchor positions, surrounded by more highly fluctuating amino acid segments.

Functional implications of regional movements

Although the structure of each monomer is quite similar, the six regions of motion shown in Figs. 3 and 4 appear at different locations in each monomer. In general, the motions of the interface for both α- and β-tubulin are very similar, largely consisting of S6 and S5, the exception being the loop region S4 (proceeding H1 in β-tubulin), which extends outward from the dimer interface. This inverse ordering for domains of motion between monomers is consistent with the notion of polarity, or the head-to-tail arrangement of dimers in the growing protofilament. The observation of the positively correlated motions of loop regions S1 and S2 could play a role in polymerization.

An additional aspect of these regional assignments is supportive evidence for their correlated motions. One example occurs at the dimer interface, where interfacial residues of both monomers are expected to exhibit similar motions. Clearly our assignment of interfacial residues to region S6 supports this claim; a result in distinct contrast with the domains assignments of Nogales et al. (1998a), where all three domains participate in the dimer interface.

FIGURE 4 Secondary structure of tubulin's sequence colored according to its six regions of motion. (*A*) and (*B*) display the sequences of α- and β-tubulin, respectively. The symbols ($*$) and ($\wedge$) below the residue names represent the most constrained and flexible regions of the sequence, respectively, according to the mean square fluctuation curves displayed in Fig. 6, *a* and *b*. The residues corresponding to the mutationally important yeast-tubulin residues (Reijo et al., 1994; Richards et al., 2000) are colored orange. Highly conserved residues within a family of 511 tubulin sequences are represented in magenta upper case letters. The nucleotide-binding residues (Nogales et al., 1998b) are boxed. The regional colors are the same as in Fig. 3. The solvent-exposed residues are identified with a line above the residue names. A sequential numbering scheme is used for β-tubulin. Note that the crystal structure (Nogales et al., 1998a) has gaps at sequence positions 45–46 and 361–368.

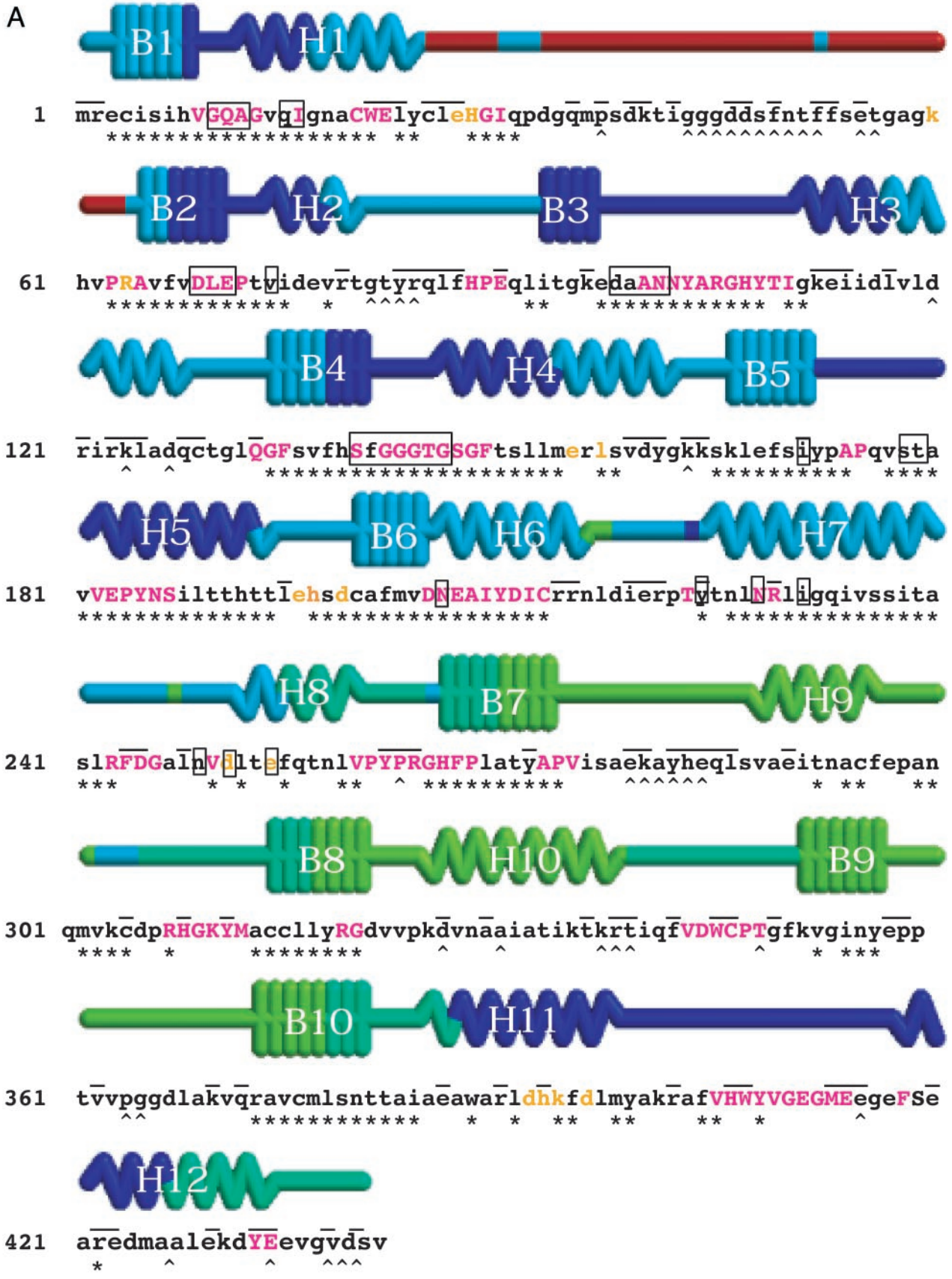
A
B1
H1
1 mrecisihVGQAGvgIgnaCWElyclеHGIqpdgqmpsdktigggddsfntffsetgagk
B2
H2
B3
H3
61 hvPRAvfvDLEPtvidevrtgtyrqlfHPEqlitgkedaANNYARGHYTIgkeiidlvld
B4
H4
B5
121 rirkladqctglQGFsvfhSfGGGTGSGFtsllmerlsvdygkksklefsiypAPqvsta
H5
B6
H6
H7
181 vVEPYNSiltthttlehsdcafmvDNEAIYDICrrnldierpTytnlNRligqivssita
H8
B7
H9
241 slRFDGalhVdltefqtnlVPYPRGHFPlatyAPVisaekayheqlsvaeitnacfepan
B8
H10
B9
301 qmvkcdpRHGKYMaccllyRGdvvpkdvnaaiatiktkrtiqfVDWCPTgfkvginyepp
B10
H11
361 tvvpggdlakvqravcmlsnttaiaeawarldhkfdlmyakrafVHWYVGEGMEegeFSe
H12
421 aredmaalekdYEevgvdsv

FIGURE 4

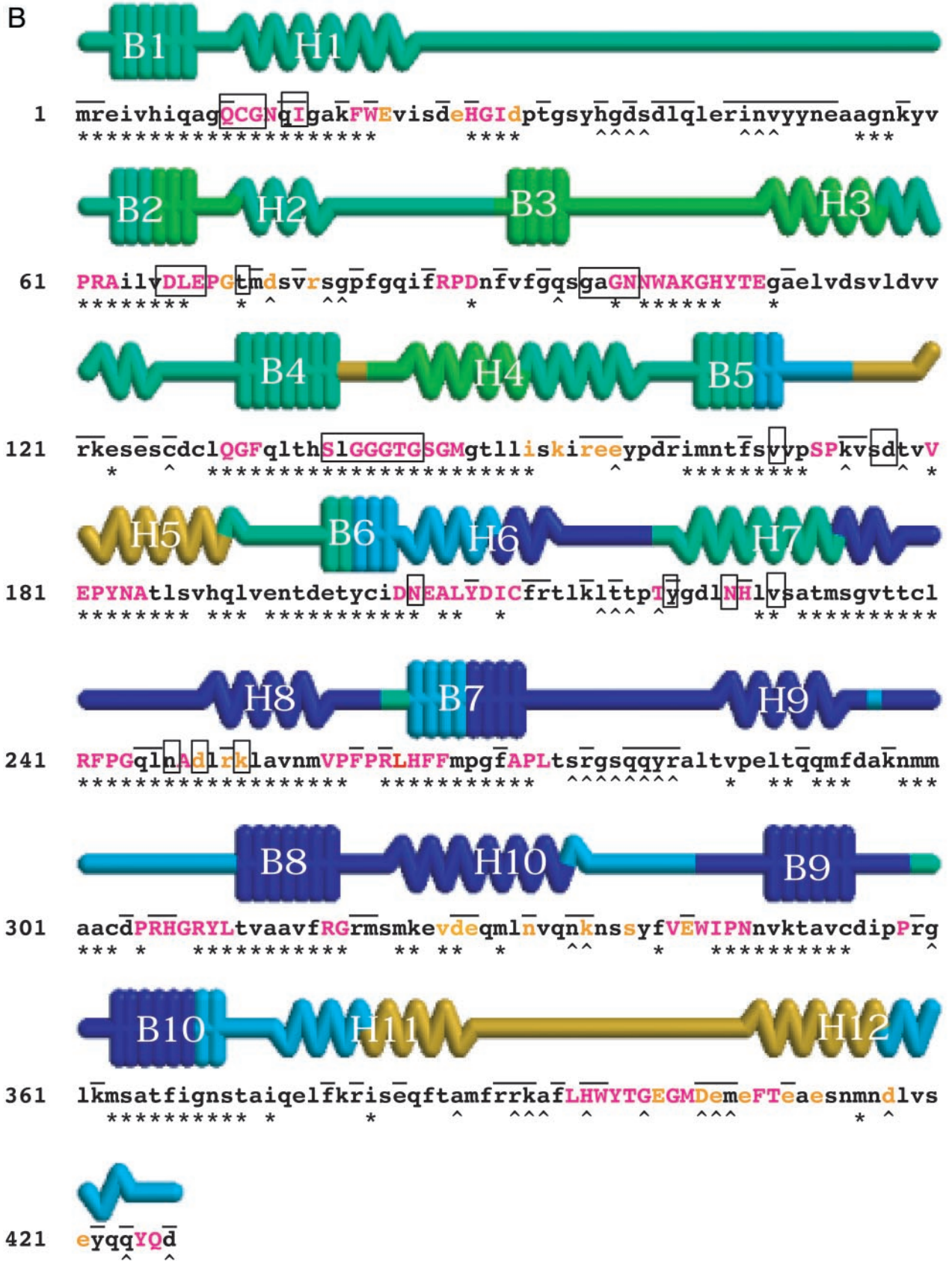

FIGURE 4. Continued.

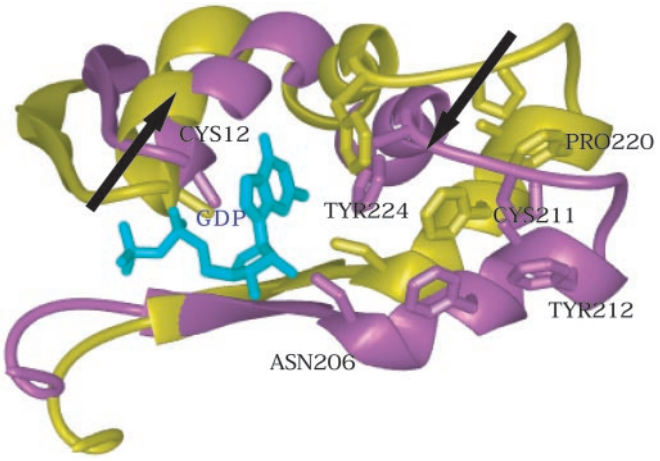

FIGURE 5 Ribbon diagram illustrating the anticorrelated motions between two putative cross-linking sites β:Cys-12 and the peptide segment 206–224, including the two residues, β:Tyr-224 and β:Asn-206, making contacts with guanine (Table 1). The two fluctuating conformations (*magenta* and *yellow*) are displayed to indicate the details of the motions near the nucleotide-binding site. The opposite-direction motions of the two segments, shown by arrows, may provide a functional flexibility for ligand binding.

Further support involves the main binding sites of colchicine within the β-subunit. Colchicine is believed to bind at residues Cys-356 and Cys-241 (Linse and Mandelkow, 1998) and the region around residues 216–243 (Uppluri et al., 1993). All of these regions are found in the S6 region and would be expected to exhibit similarly correlated motions.

GTP interacts with all of our six regions of motion (see boxed residues in Fig. 4) to produce a highly dynamic binding pocket. The Cys-12 residue in β-tubulin (β:Cys-12) can be experimentally cross-linked to GTP (Bai et al., 1999). Bai et al. demonstrated with direct photoaffinity labeling experiments that the peptide β:206–224 (β:Cys-211) might be a second site involved in cross-linking of nucleotide. As indicated in Fig. 4 *B*, H1 (residues 11–24) of β-tubulin appears in region S4. In contrast, peptides 206–211 and 212–220 appear in regions S5 and S6, respectively. Figure 5 shows the calculated fluctuations near the nucleotide-binding site for the slowest mode of motion. In this motion, the nucleotide binding loop remains flexible within the β-subunit. Although this loop motion is relatively local and small in comparison to the opposite rotations of each subunit, it does not occur independently of the remainder of the structure and reflects the details of the entire highly cooperative structure. Its precise motions are consistent with their role in ligand capture. In Fig. 4 *B*, the putative range of motions for peptide segments 206–224 and 10–14 are displayed in magenta and yellow, respectively. We note, in particular, that the positions of residues β:Tyr-224 and Asn-206, known to make contacts with guanine (Table 1), move in the opposite direction with respect to Cys-12 to open and close the backbone around the nucleotide. Although little is known about flexibilities associated with the binding pocket for nonexchangeable (N-site) GTP bound to α-tubulin, our results find flexibilities comparable to the E-site GDP binding pocket. Evidence supporting this flexibility is found in the experiments of Bai et al. (1998). Using ultraviolet irradiation, they found that N-site GTP rapidly binds to α:Cys-295. Given the large distance between α:Cys-295 and N-site GTP, these authors postulate that a large conformational change would be needed for photoreaction between these sites. Our results place the N-site in the interfacial domain of motion, S6, and the α:Cys-295 site in region S3. Inspection of the relative motions between these two regions, in Fig. 3, finds them to be negatively correlated. Although we cannot predict the absolute magnitude of these motions, our finding that these regions in α-tubulin have a tendency to move toward each other is consistent with shortening their separation distance, and, possibly, facilitating photoreaction.

The role of residue flexibility in tubulin recognition and microtubule assembly

The mean square fluctuations of residues, which were used to establish the correlations displayed in Figs. 1 and 3, also identify the relative displacements of residues within regions of the tubulin dimer. In Fig. 6, we display the distribution of mean square fluctuations of C^{α} atoms, $\langle \Delta R_i^2 \rangle$, as a function of residue index i for the α- and β-monomers of the tubulin dimer in parts (*a*) and (*b*), respectively. The same figure also shows ribbon diagrams of the dimeric structure from two different views, color-coded according to the relative sizes of residue movements. The segments subject to the largest and smallest amplitude fluctuations are colored red and blue, respectively, with the intermediate fluctuations colored spectrally within this color range. Previously, we had shown a strong correspondence between our calculated fluctuations and crystallographic temperature factors (Bahar et al., 1997, 1998, 1999; Keskin et al., 2000). Although temperature factors were not provided with the electron crystallographic data of tubulin, our calculations can be used to predict the flexibilities of peptides in the tubulin dimer. In general, regions having the highest mobilities are located in solvent-accessible loops, many of which are identified to be involved in contacts with adjacent dimers in the protofilament or with accessory proteins. Among the regions of highest flexibility are residues α:43–53, β:37–41, and β:51–53, belonging to the loop connecting secondary structural elements H1 and B2 and known to be involved in lateral contacts between adjacent protofilaments (Table 1). Likewise, the M loop residues 280–285 are distinguished by their extremely high flexibility, also consistent with their known role in recognizing lateral neighbors (Nogales et al., 1999). Several additional maxima include the loop residues 410–413, 221–225, 336–340,

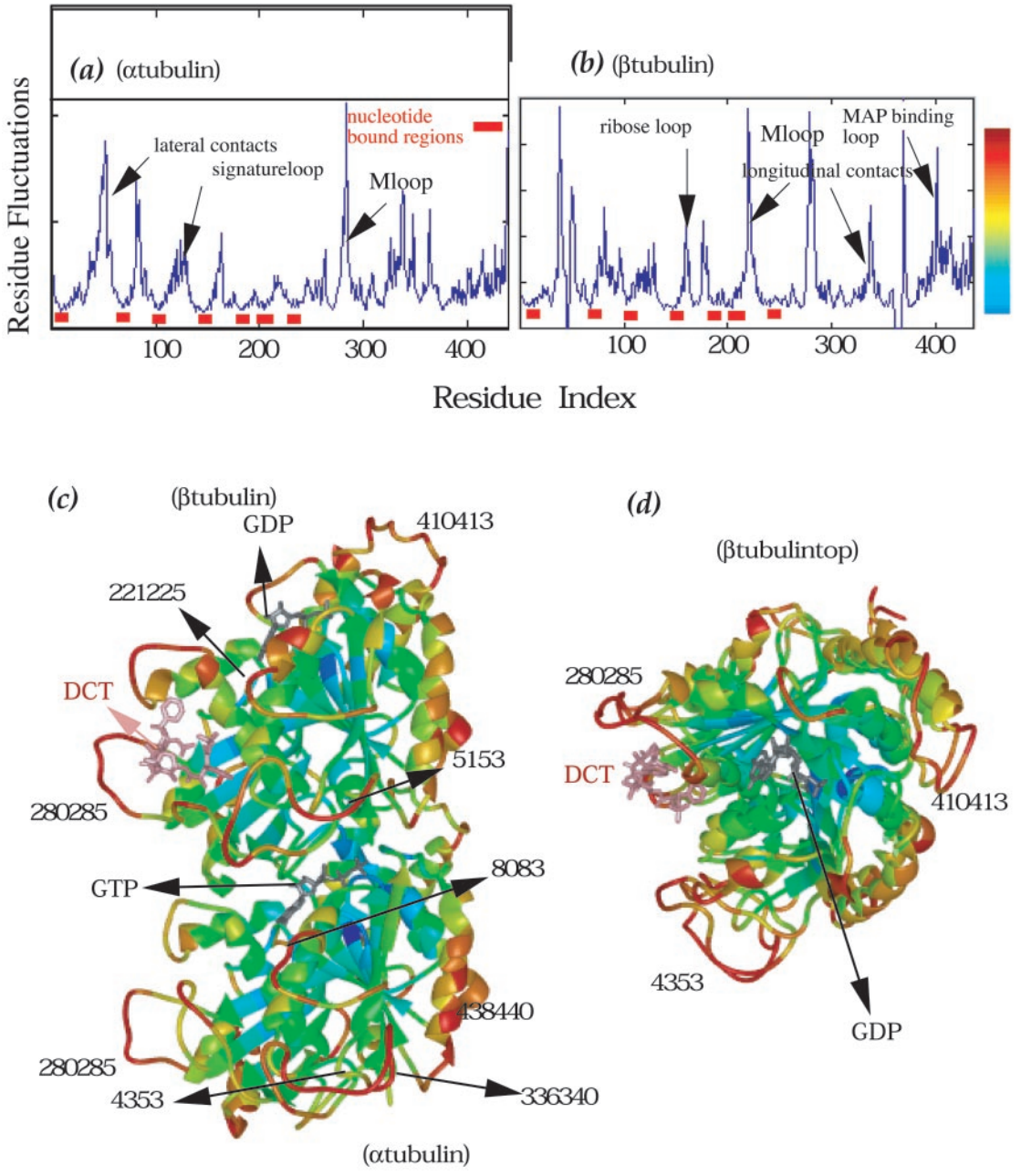

FIGURE 6 Residue flexibilities calculated for (*a*) α-tubulin, (*b*) β-tubulin monomers in the dimer, and (*c* and *d*) the ribbon diagram of the dimer colored on the basis of the relative amplitudes of fluctuations of individual residues shown in two different views. A blue-to-red color spectrum is used to represent different levels of flexibilities, where the smallest motions are in is blue and the highest in red, as shown in the color bar on the right of the figures. (*c*) The dimer in the same orientation as in Fig. 1 *a*. (*d*) Top view of the dimer. A longitudinal internal axis is observed to be the least flexible part of the dimer. The most mobile regions (peaks in *a* and *b*) are colored red and marked on the ribbon diagrams in *c* and *d*. Maximal flexibility (*red*) is observed almost exclusively in the loop regions involved in lateral and longitudinal contacts for protofilament and microtubule assembly.

common in α- and β-monomers, and thought to be engaged in recognition of the neighboring dimeric subunit along the longitudinal direction of the protofilament (Nogales et al., 1999). Finally, α:80–83 belongs to the H2–B3 loop involved in lateral contacts between adjacent protofilaments, and two local peaks around α/β:159–162 are near the ribose binding peptide 162–181 (Table 1).

Noteworthy differences in the fluctuations observed between α- and β-tubulin relate to their dimerization interface. For example, the large fluctuations observed around residues 172–181 of β-tubulin are not observed in α-tubulin. This region in α-tubulin makes structural contacts with β-tubulin, presumably to rigidify the dimer interface. Consistent with this premise, the corresponding loop in α-tubu-

TABLE 2 Lethal mutation sites and conserved residues from mutation studies*

Lethal mutations in α/β-tubulins	Flexibility	Conservation	Region Boundary	Reference
E22A, E27A, D31A	R	+	−	(Reijo et al., 1994)
G71W, D74A, R77A	R	+	+	(β-tubulin)
I152F, R156K, E157A, E158A	R	−	+	
K154A, R156A	R	−	+	
E157A, E158A	R	−	+	
D249A, R251A, K252A	R	−	−	
E327A, D328A, E329A†	R	−	−	
K332A-K336A	F	−	+	
D339A, E343A	R	+	+	
E401A, D404A, E407A†	F	+	+	
E405A, E407†	F	+	+	
E410A, E412A†	F	−	+	
D417A, E421A	F	−	+	
E431A, D432A†‡				
E27A, H28A	R	+	+	(Richards et al., 2000)
K61A, R65A	R	+	+	(α-tubulin)
E156A, E157A†	R	−	+	
E197A, H198A, D200A	R	−	+	
D252A, E255A	R	−	+	
D393A, R394A, K395A, D397A	R	−	+	

*The top and bottom groups give the mutations in alanine-scanning mutagenesis (Reijo et al., 1994; Richards et al., 2000). Each set represents one mutation experiment. The second column indicates whether the mutational data set corresponds to a minimum (R) or a maximum (F) in Fig. 6, *a* and *b*. The third column provides information about the conservation of these sites among 550 tubulin sequences. In the fourth column, (−) sign represents that the residues are further than three residues from regional boundaries, (+) otherwise are close to the regional boundaries according to Fig. 4.
†The dominant lethal mutations.
‡The positions of these residues are not provided in the PDB.

lin looses flexibility upon polymerization. Our calculations on the individual monomeric forms of α- and β-tubulins further support this suggestion. Residues of α:172–179, 217–224, 389–413, β:241–262, 320–331, and 340–351 are also highly flexible. Each of these regions involves interfacial residues (Nogales et al., 1999), and its calculated flexibilities are significantly suppressed upon dimerization. These results are consistent with the notion that the intrinsic, predimerization flexibility may contribute to efficient mutual recognition of the α and β monomers. Finally, we note that Lys-40 in α-tubulin is frequently acetylated, and is one of the loops that protrude into the inside of the MT (Downing and Nogales 1998a,b). This residue is observed to enjoy a high flexibility in the present analysis, which could be a requirement for its efficient reactivity. A summary of the fluctuations corresponding to these functionally important sites is given in the last column of Table 1.

Many regions of minimal fluctuations are closely identified with the sequence-conserved nucleotide binding sites in α- and β-tubulin. The nucleotide binding sites are highlighted by boxes in Fig. 4, and involve residues in the loops between B1–H1, B2–H2, B3–H3 and the glycine-rich B4–H4 loop. Additional regions outside the GTP binding pocket are also found to have low mobilities. Mutational studies of tubulin function suggest the importance of residues at these locations. The alanine scanning mutagenesis data obtained for yeast β-tubulin (Reijo et al., 1994) identifies three dominant-lethal mutation sites at 327–329; 401, 404, 407; and 431–432, and eleven recessive-lethal mutations, all displayed in orange letters in Fig. 4 *B*. The lethal mutations in yeast α-tubulin (Richards et al., 2000) are also colored orange in Fig 4 *A*. As noted earlier, the symbols (∗) and (/\) below the tubulin sequence in Fig. 4 represent the least and most flexible regions, respectively, according to the mean square fluctuations shown in Fig. 6, *A* and *B*. Examination of these lethal mutation sites indicate a close correspondence with the fluctuation minima. In fact, nearly all lethal mutations occur at fewer than three residue positions away from the minima in the global fluctuation curve. In summary, Table 2 lists the mutation sites in the first column, and qualitative measures of flexibility based on our calculations (F, flexible; R, rigid) in the second column. In the third column a + indicates residues near the putative domain linkage sites identified in Fig. 4. Most of the mutational sites correspond to minima and are also close to our putative linkage sites between domains of motion. When the global, large-scale modes of motions are dissected into their slowest modes, there is a remarkable agreement with regions proposed by Nogales et al. (1998b) to be critically important (Table 1) for tubulin function. Nearly all minima in the slowest mode of motion curve coincide with the peptide segments listed in Table 1 to be important for drug

or nucleotide binding. The remaining few segments listed in Table 1 belong to another category; those providing intermolecular lateral contacts for MT assembly, as discussed in the preceding section. These key residue positions apparently function to maintain binding sites within tubulin's structure necessary for ligand interaction.

Taken together, these results are consistent with prior GNM analyses (Bahar et al., 1998; Demirel et al., 1998; Bahar et al., 1999; Keskin et al., 2000; Bahar and Jernigan, 1998), which proposed that minima in the global-mode shapes generally coincide with biologically active sites, such as catalytic sites in enzymes or, in the case of tubulin, nucleotide binding sites. Maxima, in contrast, correspond to segments distinguished by their enhanced mobilities, often implicated in substrate recognition. Evidently, residues that are detected as minima in the three slowest modes of motion and as lethal-mutation sites in experiments or as the conserved residues from the sequence analysis can be viewed as critical sites for the biological function of tubulins, and thereby MTs.

Relationship between dominant modes of motion and MT assembly

Low-frequency modes dominate the overall global dynamics. We have examined the cumulative contribution of the lowest frequency modes in the observed dynamics. The first five modes account for ~40% of the overall dynamics, and the first 30 modes correspond to ~70%. These values are obtained by the ratio of individual eigenvalues to the sum of the eigenvalues.

In addition to providing suggestions about the role of selected residues in tubulin function, our results can also be used to suggest ideas about the dominant mechanism of motions that could be implicated in MT assembly. The dominant mode of motion is found to be the rigid body rotation of each monomer in opposite directions, as illustrated in Fig. 7 *a*, where the blue and pink colors represent the α- and β-tubulin chains, respectively. Schematic images have been added to the right of this figure for clarity. The rotation axis of the slowest mode coincides with the longitudinal axis of the tubulin dimer. In this motion, the innermost regions closest to the longitudinal axis are the most rigid parts of the molecule, acting as the rotation axis, and the extent of displacements increases with radial distance from this central axis. This mode may contribute to the formation and optimization of lateral contacts between tubulin dimmers as they appear in protofilaments of MTs.

The next dominant mode is illustrated in Fig 7 *b*. In this case, the monomers are observed to have opposite sense rigid body rotations about the central out-of-plane (or radial) axes. This motion involves a counterclockwise rotation of the α-monomer about an axis perpendicular to the plane of the paper, accompanied by a clockwise rotation of the β-monomer, and vice versa on the right side. These two

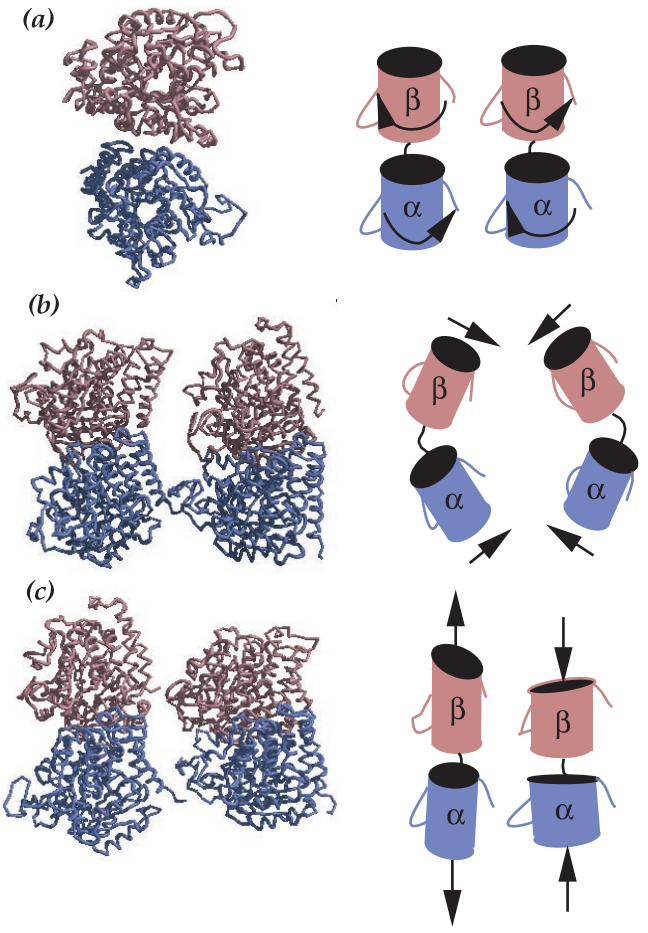

FIGURE 7 Mechanisms of the dominant modes of motion. Two extrema of the fluctuating conformations are shown by the ribbon diagrams (*left*) and schematic diagrams (*right*), for the slowest modes (*a–c*). Mode 1 is an opposite-sense rigid-body rotation of the monomers about the longitudinal axes. (*a*) displays the top views of the α- and β-monomers, rotated by the action of mode 1. (*b*) The next two slowest modes induce again rigid-body rotations of the individual monomers, but this time around out-of-plane axes centered in each monomer, thus giving rise to an overall wobbling of the two monomers. This is succeeded by a similar perpendicular wobbling mode, indicating an approximate cylindrical symmetry to the wobbling when these two modes are considered. (*c*) An overall extension/contraction of the dimer along with elongation and shortening.

rotations describe a wobbling or rolling of the two monomers over each other, giving rise to a slight shift in the positions of interfacial contacting residues. This mode appears to be facilitated by the relatively less compact packing of interfacial residues.

The next dominant mode (not shown) is similar to the second mode, except for a 90° rotation about the cylindrical axis, making it the third of the orthogonal (approximately) rotation axis. The second and third modes are, in a sense, equivalent, or complement each other, as would be expected from the approximate cylindrical shape of the dimeric structure. A weighted average of the fluctuations induced in the second and third modes indicates that these modes may be

involved in longitudinal interactions (Nogales et al., 1999) related possibly to polymerization. It is noteworthy that the exposed loops and the C-terminal segments have the largest displacements in the second mode.

Experimental electron micrographs show that, upon rapid disassembly of MTs, the protofilaments peel off and curl. In addition, tubulin-GDP can polymerize into closed rings of ~14–16 dimers. Thus, the tubulin-GDP state is postulated to exist in a curved conformation, in contrast to the straight conformation for the tubulin-GTP state. A conformational change of this type occurring at the dimer interface may produce a kink between dimmers (Downing and Nogales 1998a,b,c), which may be facilitated by the contributions of these second and third modes.

A further dominant mode (Fig. 7 *c*) reveals an overall stretching/contraction of the molecule along its longitudinal axis. This motion is coupled with the inward and outward movement of the M-loop about the same axis and with the elongation and shortening of the dimer. These coupled motions can be viewed as an overall compression/expansion, which may aid the formation of lateral contacts between the protofilaments of the MT. Additional slow modes will not be discussed here, other than to note that they are likewise indicative of similar breathing motions. Most of the residues identified to make lateral contacts (Nogales et al., 1999) emerge as maxima in the weighted-average fluctuation curve of these modes (not shown).

Effect of pacilitaxel binding on tubulin flexibility

The taxoids paclitaxel and docetaxel are increasingly used in the treatment of human cancers (Rao et al., 1992). The mechanism of action of these drugs is to bind β-tubulin and affect the stabilization of abnormal arrays of MTs in cells leading to an interference of normal tubulin function, including formation of an effective mitotic spindle. The detailed mechanism of taxoid stabilization remains controversial: binding to tubulin could either strengthen the weak lateral interactions between tubulin subunits (Andreu et al., 1992; Howard and Timasheff, 1988), or improve the longitudinal bonds of MTs as proposed by Dye et al. (1992). One of the most significant differences between the α- and β-monomers of tubulin is the deletion of eight-loop residues in the β-monomer. The electron crystallographic structure shows that pacilitaxel occupies the deleted site in the β-monomer, near the sequences β:1–31 and 217–231. The main interaction of the taxane ring with tubulin is at Leu-275, at the beginning of the B7–H9 loop. The mutation of Phe-272 to Val causes partial pacilitaxel resistance in a cell culture line (Giannakakou et al., 1997). Phe-272 sits on the β-strand at the back of the taxoid binding site, and its side chain interacts with the drug.

The presently examined electron crystallographic structure includes a molecule of pacilitaxel, whose structure has been solved by x-ray crystallography (Gueritte-Voegelein et al., 1991), bound to the β-monomer. For an assessment of the effect of drug binding on the fluctuation dynamics of the dimer, we constructed a taxoid-free model. This was done by energy minimization of the crystallographic structure in the absence of drug, using the CHARMM energy minimization algorithm with parameter set 22 (Brooks et al., 1983). The same procedure was also repeated starting from the original, pacilitaxel-bound structure. We observed small local structural changes for the side chains near the substrate upon energy minimization, but no significant changes could be observed on a global scale.

Figure 8 shows the difference in the calculated residue fluctuations for all modes of motion between the holo and apo models. A dramatic decrease is observed in the flexibility of the peptide loops β:272–285 and β:214–224, and lesser reductions in loop β:35–44, and segment β:351–378 for the holo form, all of which are in close proximity to bound pacilitaxel. These regions are shown as blue in Fig. 8. The segment β:272–285 includes the M loop, which appears to be important for lateral contacts between protofilaments, and thereby its apparent decreased flexibility in the presence of pacilitaxel could contribute to drug-enhanced stabilization of MTs.

In addition to these effects in the immediate neighborhood of the substrate-binding site, we observe some long-range (nonlocal) effects induced by pacilitaxel binding, although these are relatively small motions. For example, there is an enhancement in the flexibility of the α-tubulin loops α:32–62 and α:276–286 (i.e., the M loop of α-tubulin) and of the C-terminus of the α-monomer, α:434–440, upon pacilitaxel binding. There is also an increase in the mobility of the C-terminal domain of the β-monomer. The ribbon diagram in the right panel shows these regions with enhanced fluctuations in red. In general, the α-monomer, or the minus end of the dimer becomes more mobile, whereas the β-monomer, or the plus end of the dimer, becomes less mobile at its intersubunit interface. The binding of pacilitaxel to β-tubulin appears to induce opposite sense mobilities between the plus and minus ends of the dimer. The induced mobility of these regions may reflect the increase in internal energy indirectly induced by reducing the entropy of the molecular at the taxol binding site.

The observed reduced flexibility of the M loop in β-tubulin is consistent with the suggestions of Downing and Nogales (1998a) and Amos and Lowe (1999) regarding the role of this loop in the stabilization of lateral contacts. This reduced flexibility appears to be directly related to the increased packing density in its close neighborhood, as a result of the occupation of an otherwise empty site by the substrate (pacilitaxel), and a simultaneous reconfiguration of the originally flexible M loop. It is possible that this apparent change in tubulin mobility contributes significantly to the stabilization of intramolecular and intermolecular contacts. In this view, lateral contacts made by the M loop become more stable, making MT disassembly more

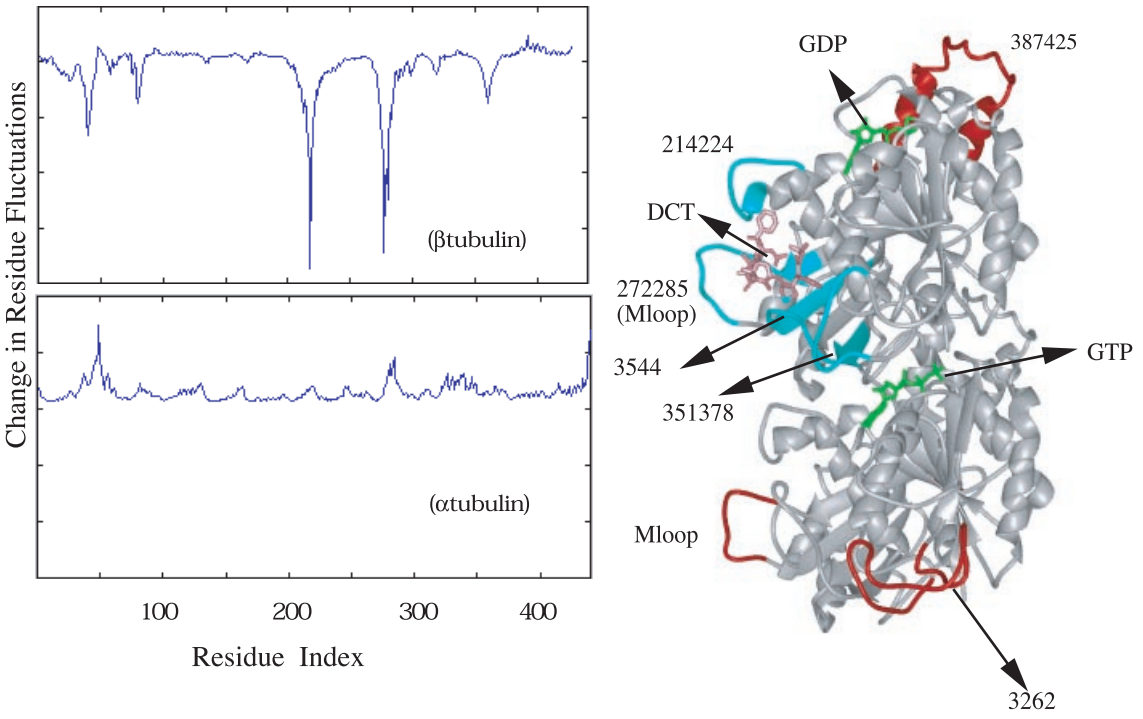

FIGURE 8 Effect of pacilitaxel binding on the flexibility of the α- and β-tubulin residues in the α/β-dimer. The changes in fluctuations induced by pacilitaxel binding are displayed, for the β-monomer (*upper left curve*) and α-monomer (*lower left curve*). A severe suppression in the mobilities of the loops contacting the drug is observed in β-tubulin, along with an enhanced flexibility in certain regions in α-tubulin, indicated on the right-most ribbon diagram. Regions with enhanced mobilities are colored red, and those with suppressed mobilities are colored blue. (GDP, GTP are colored green and pacilitaxel is colored pink.)

difficult. Taken together, these effects may offer explanations for the mechanism of action for taxoids, based primarily on tubulin's enhanced stability and loss of flexibility in the taxoid-bound form (Wilson and Jordan, 1995).

CONCLUSION

Understanding the structural and functional importance of macromolecular flexibility continues to be a widely studied question. An increasing number of structural studies indicates that the functions of many proteins are intrinsically linked to their ability to change conformation in a precise manner. Here we have presented an approach that is grounded in mechanics, with a treatment that is both robust and sufficiently simple to permit the examination of relatively large multimeric structures and their motions.

In MTs, the marginal stability of the structure—or the coupled flexibility of different domains, which becomes operative only upon dimerization of tubulins—seems to be integrally linked to its role in polymerization/depolymerization, and MT assembly. Our results indicate that loops involved in the recognition of each monomeric subunit are highly flexible. Collective motions are confined to six regions, composed of structurally close but discontinuous residue positions. Boundaries between regions of collective motions appear to act as linkages found primarily within minimally fluctuating secondary-structure elements that exhibit poor sequence conservation, but with a tendency for hydrophobic residues at these regional boundaries. Regions of lowest mobility coincide with both secondary structures and lethal-mutation sites. Regions involved in the recognition of each monomeric subunit are found in highly flexible loops. The antitumor drug pacilitaxel is found to stabilize the M loop of the β- tubulin monomer, an effect that might enhance lateral interactions between protofilaments. Our analysis demonstrates that relationships between experimental and computational results can provide a framework for examination of structure–function associations for relatively large protein complexes.

The authors are thankful to the staff of Advanced Biomedical Computing Center (ABCC), Frederick Cancer Research & Development Center for access to their computer facilities. Special acknowledgment is made for the assistance of Dr. N. Pattabiraman, Dr. J. Collins and Mr. G. Smythers of the ABCC Staff. Additional thanks are extended to Dr. E. Hamel for a careful reading of this manuscript. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This work was funded in part by grants from the National Institutes of Health (NO1-CO-56000 and the U.S. Army Breast Cancer Research Project (DAMD17-98-1-8323).

REFERENCES

Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. 1994. Molecular Biology of the Cell, 3rd ed., Garland Pub., New York.

Amadei, A., A. B. Linssen, and H. J. Berendsen. 1993. Essential dynamics of proteins. *Proteins.* 17:412–425.

Amos, L. A., and J. Lowe. 1999. How taxol stabilises microtubule structure. *Chem. Biol.* 6:R65–R69.

Andreu, J. M., J. Bordas, J. F. Diaz, D. A. Garcia, R. Gil, F. J. Medrano, E. Nogales, E. Pantos, and E. Towns-Andrews. 1992. Low resolution structure of microtubules in solution. Synchrotron x-ray scattering and electron microscopy of taxol-induced microtubules assembled from purified tubulin in comparison with glycerol and MAP-induced microtubules. *J. Mol. Biol.* 226:169–184.

Arrington, C. B., and A. D. Robertson. 2000. Microsecond to minute dynamics revealed by EX1-type hydrogen exchange at nearly every backbone hydrogen bond in a native protein. *J. Mol. Biol.* 296: 1307–1317.

Atilgan, A. R., S. R. Durell, R. L. Jernigan, M. C. Demirel, O. Keskin, and I. Bahar, 2001. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* 80:505–515.

Bahar, I., and R. L. Jernigan. 1997. Inter-residue potentials in globular proteins and the dominance of highly specific hydrophilic interactions at close separation. *J. Mol. Biol.* 266:195–214.

Bahar, I., and R. L. Jernigan. 1998. Vibrational dynamics of transfer RNAs: comparison of the free and synthetase-bound forms. *J. Mol. Biol.* 281:871–884.

Bahar, I., A. R. Atilgan, M. C. Demirel, and B. Erman. 1998. Vibrational dynamics of folded proteins: significance of slow and fast motions in relation to function and stability. *Phys. Rev. Lett.* 80:2733–2736.

Bahar, I., B. Erman, R. L. Jernigan, A. R. Atilgan, and D. G. Covell. 1999. Collective motions in HIV-1 reverse transcriptase: examination of flexibility and enzyme function. *J. Mol. Biol.* 285:1023–1037.

Bahar, I., A. R. Atilgan, and B. Erman. 1997. Direct evaluation of thermal fluctuations in proteins using a single parameter harmonic potential. *Folding Des.* 2:173–181.

Bai, R., K. Choe, J. B. Ewell, N. Y. Nguyen, and E. Hamel. 1998. Direct photoaffinity labeling of cysteine-295 of alpha-tubulin by guanosine 5′-triphosphate bound in the nonexchangeable site. *J. Biol. Chem.* 273: 9894–9897.

Bai, R., J. B. Ewell, N. Y. Nguyen, and E. Hamel. 1999. Direct photoaffinity labeling of cysteine 211 or a nearby amino acid residue of beta-tubulin by guanosine 5′-diphosphate bound in the exchangeable site. *J. Biol. Chem.* 274:12710–12714.

ben Avraham, D. 1993. Diffusion-limited three-body reactions in one dimension. *Phys. Rev. Lett.* 71:3733–3735.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy minimization and dynamic calculations. *J. Comput. Chem.* 4:187–217.

Brooks, B. R., D. Janezic, and M. Karplus. 1995. Harmonic analysis of large systems .1. Methodology. *J. Comput. Chem.* 16:1522–1542.

Burns, R. G., and C. D. Surridge. 1993. Tubulin: conservation and structure. *In* Microtubules. J. S. Hyams and C. W. Lloyd, editors. Wiley, New York. 3–32.

Case, D. A. 1994. Normal mode analysis of protein dynamics. *Curr. Opin. Struct. Biol.* 4:285–290.

Chabner, B. A., and D. L. Longo. 1996. Cancer Chemotherapy and Biotherapy: Principles and Practice. Lippincot-Raven, Philadelphia, PA.

de Groot, B. L., S. Hayward, D. M. van Aalten, A. Amadei, and H. J. Berendsen. 1998. Domain motions in bacteriophage T4 lysozyme: a comparison between molecular dynamics and crystallographic data. *Proteins.* 31:116–127.

Demirel, M. C., A. R. Atilgan, R. L. Jernigan, B. Erman, and I. Bahar. 1998. Identification of kinetically hot residues in proteins. *Protein Sci.* 7:2522–2532.

Desai, A., and T. J. Mitchison. 1997. Microtubule polymerization dynamics. *Annu. Rev. Cell Dev. Biol.* 13:83–117.

Doruker, P., A. R. Atilgan, and I. Bahar. 2000. Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: application to alpha-amylase inhibitor. *Proteins.* 40:512–524.

Downing, K. H., and E. Nogales. 1998a. Tubulin structure: insights into microtubule properties and functions. *Curr. Opin. Struct. Biol.* 8:785–791.

Downing, K. H., and E. Nogales. 1998b. New insights into microtubule structure and function from the atomic model of tubulin. *Eur. Biophys. J.* 27:431–436.

Downing, K. H., and E. Nogales. 1998c. Tubulin and microtubule structure. *Curr. Opin. Cell Biol.* 10:16–22.

Dye, R. B., P. F. Flicker, D. Y. Lien, and R. C. Williams, Jr. 1992. End-stabilized microtubules observed in vitro: stability, subunit, interchange, and breakage. *Cell Motil. Cytoskeleton.* 21:171–186.

Erickson, H. P. 1998. Atomic structures of tubulin and FtsZ. *Trends Cell Biol.* 8:133–137.

Flory, P. J. 1976. Statistical thermodynamics of random networks. *Proc. R. Soc. Lond. Ser. A.* 351–380.

Genetic Computer Group, Wisconsin Package. Version 9.0. GCG, Madison, WI.

Giannakakou, P., D. L. Sackett, Y. K. Kang, Z. Zhan, J. T. Buters, T. Fojo, and M. S. Poruchynsky. 1997. Paclitaxel-resistant human ovarian cancer cells have mutant beta-tubulins that exhibit impaired paclitaxel-driven polymerization. *J. Biol. Chem.* 272:17118–17125.

Go, N. 1990. A theorem on amplitudes of thermal atomic fluctuations in large molecules assuming specific conformations calculated by normal mode analysis. *Biophys. Chem.* 35:105–112.

Gueritte-Voegelein, F., D. Guenard, F. Lavelle, M. T. Le Goff, L. Mangatal, and P. Potier. 1991. Relationships between the structure of taxol analogues and their antimitotic activity. *J. Med. Chem.* 34:992–998.

Haliloglu, T., I. Bahar, and B. Erman. 1998. Gaussian dynamics of folded proteins. *Phys. Rev. Lett.* 79:3090–3093.

Hayward, S., A. Kitao, and H. J. Berendsen. 1997. Model-free methods of analyzing domain motions in proteins from simulation: a comparison of normal mode analysis and molecular dynamics simulation of lysozyme. *Proteins.* 27:425–437.

Hinsen, K. 1998. Analysis of domain motions by approximate normal mode calculations. *Proteins.* 33:417–429.

Hinsen, K., A. Thomas, and M. J. Field. 1999. Analysis of domain motions in large proteins. *Proteins.* 34:369–382.

Howard, W. D., and S. N. Timasheff. 1988. Linkages between the effects of taxol, colchicine, and GTP on tubulin polymerization. *J. Biol. Chem.* 263:1342–1346.

Keskin, O., R. L. Jernigan, and I. Bahar. 2000. Proteins with similar architecture exhibit similar large-scale dynamic behavior. *Biophys. J.* 78:2093–2106.

Keskin, O., I. Bahar, D. Flatow, D. G. Covell, and, R. L. Jernigan. 2002. Molecular mechanisms of chaperonin GroEL-GroES function. *Biochemistry.* 41:491–501.

Kitao, A., and N. Go. 1999. Investigating protein dynamics in collective coordinate space. *Curr. Opin. Struct. Biol.* 9:164–169.

Linse, K., and E. M. Mandelkow. 1988. The GTP-binding peptide of beta-tubulin. Localization by direct photoaffinity labeling and comparison with nucleotide-binding proteins. *J. Biol. Chem.* 263:15205–15210.

Luduena, R. F. 1998. Multiple forms of tubulin: different gene products and covalent modifications. *Int. Rev. Cytol.* 178:207–275.

Miyazawa, S., and R. L. Jernigan. 1985. Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation. *Macromolecules.* 18:534–552.

Nogales, E., S. G. Wolf, and K. H. Downing. 1998a. Structure of the alpha beta tubulin dimer by electron crystallography. *Nature.* 391:199–203.

Nogales, E., K. H. Downing, L. A. Amos, and J. Lowe. 1998b. Tubulin and FtsZ form a distinct family of GTPases. *Nat. Struct. Biol.* 5:451–458.

Nogales, E., M. Whittaker, R. A. Milligan, and K. H. Downing. 1999. High-resolution model of the microtubule. *Cell.* 96:79–88.

Rao, S., S. B. Horwitz, and I. Ringel. 1992. Direct photoaffinity labeling of tubulin with taxol. *J. Natl. Cancer Inst.* 84:785–788.

Reijo, R. A., E. M. Cooper, G. J. Beagle, and T. C. Huffaker. 1994. Systematic mutational analysis of the yeast beta-tubulin gene. *Mol. Biol. Cell.* 5:29–43.

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.

Thomas, A., M. J. Field, L. Mouawad, and D. Perahia. 1996. Analysis of the low frequency normal modes of the T-state of aspartate transcarbamylase. *J. Mol. Biol.* 257:1070–1087.

Tirion, M. M. 1996. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* 77:1905–1908.

Uppuluri, S., L. Knipling, D. L. Sackett, and J. Wolff. 1993. Localization of the colchicine-binding site of tubulin. *Proc. Natl. Acad. Sci. U.S.A.* 90:11598–11602.

Weiss, S. 1999. Fluorescence spectroscopy of single biomolecules. *Science.* 283:1676–1683.

Wilson, L., and M. A. Jordan. 1995. Microtubule dynamics: taking aim at a moving target. *Chem. Biol.* 2:569–573.