Journal of Molecular Graphics and Modelling 27 (2008) 497–505

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Identification of tubulin drug binding sites and prediction of relative differences in binding affinities to tubulin isotypes using digital signal processing

Ke Chen [a], J. Torin Huzil [b], Holly Freedman [b], Parameswaran Ramachandran [c], Andreas Antoniou [c], Jack A. Tuszynski [b], Lukasz Kurgan [a,*]

[a] Department of Electrical and Computer Engineering, University of Alberta, ECEFR, 9701 116 Street, Edmonton, AB, Canada T6G 2V4
[b] Department of Oncology, University of Alberta, Cross Cancer Institute, 11560 University Avenue, Edmonton, AB, Canada T6G 1Z2
[c] Department of Electrical and Computer Engineering, University of Victoria, P.O. Box 3055 STN CSC, Victoria, BC, Canada V8W 3P6

ARTICLE INFO

Article history:
Received 8 January 2008
Received in revised form 1 September 2008
Accepted 2 September 2008
Available online 10 September 2008

Keywords:
Tubulin
Isotype
Digital signal processing
Protein hot spot
Drug affinity
Resonant recognition model
Short-time Fourier transform
Characteristic frequency

ABSTRACT

Microtubules are involved in numerous cellular processes including chromosome segregation during mitosis and, as a result, their constituent protein, tubulin, has become a successful target of several chemotherapeutic drugs. In general, these drugs bind indiscriminately to tubulin within both cancerous and healthy cells, resulting in unwanted side effects. However, differences between β-tubulin isotypes expressed in a wide range of cell types may aid in the development of anti-tubulin drugs having increased specificity for only certain types of cells. Here, we describe a digital signal processing (DSP) method that is capable of predicting hot spots for the tubulin family of proteins as well as determining relative differences in binding affinities to these hot spots based only on the primary sequence of 10 human tubulin isotypes. Due to the fact that several drug binding sites have already been characterized within β-tubulin, we are able to correlate hot spots with the binding sites for known chemotherapy drugs. We have also verified the accuracy of this method using the correlation between the binding affinities of characterized drugs and the tubulin isotypes. Additionally, the DSP method enables the rapid estimation of relative differences in binding affinities within the binding sites of tubulin isotypes that are yet to be experimentally determined.

1. Introduction

Microtubules (MTs) are polymeric protein complexes constructed from a heterodimer of two highly homologous proteins known as α- and β-tubulin. The assembly of tubulin heterodimers into a macromolecular MT complex is a tightly regulated and dynamic process [1]. MTs are involved in a broad range of cellular processes, including the maintenance of cellular morphology and active transport of cellular components throughout the cytoplasm [2,3]. One of the most critical roles MTs play, is the formation of the mitotic spindle apparatus, providing the mechanical force required for chromosome separation during mitosis [4]. During mitosis, MTs are essential for the proper assembly of the mitotic spindle apparatus, the organelle responsible for the segregation of aligned chromosomes prior to cell division. As such, MTs have become the target of numerous anti-mitotic agents, including colchicine and antitumor drugs such as the taxanes, epothilones, and Vinca alkaloids. Colchicine, a water-soluble alkaloid binds to the intradimer interface between α- and β-tubulin dimers, presumably disrupting the correct orientation of protofilaments within the MT [5]. Paclitaxel, which continues to be one of the most successful cancer therapeutic agents, has a unique mechanism of action as it binds to and results in the stabilization of MTs within all cells [6,7]. Derivatives of paclitaxel, such as docetaxel, have been synthesized to address the limited solubility of paclitaxel and show increased binding to a site within β-tubulin [8]. Two additional families of similar tubulin-binding drugs are the epothilones and dolastatins, which share a similar mode of action. Through structural studies, both the taxanes and epothilones were shown to bind a unique site within the β subunit of the α/β-tubulin heterodimer [9,10]. Unfortunately, neither of the structures has been able to reveal the precise mechanism of MT stabilization by these drugs. Finally, the Vinca alkaloids include vincristine, vinblastine, and vinorelbine, and are used most commonly in combination chemotherapy

* Corresponding author. Tel.: +1 780 492 5488; fax: +1 780 492 1811.
E-mail address: lkurgan@ece.ualberta.ca (L. Kurgan).

Abbreviations: DSP, digital signal processing; EIIP, electron–ion interaction potential; GDP, guanosine diphosphate; GTP, guanosine triphosphate; IC_{50}, half maximal inhibitory concentration; MTs, microtubules; PDB, protein data bank; RSA, relative solvent accessibility; RRM, resonant recognition model; STDFT, short-time discrete Fourier transform.

1093-3263/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2008.09.001

regimes for the treatment of leukemia's and lymphomas. Vinblastine was shown to bind longitudinally at the inter-tubulin dimer interface, ultimately resulting in the net reduction of polymerized tubulin concentration [11]. Interestingly, at high concentrations vinblastine binds to MTs and results in their depolymerization. However, at low concentrations it is thought to bind to MT tips and suppresses their dynamic instability, leading to MT stabilization [12].

Unfortunately, these anti-tubulin chemotherapeutic agents have a main flaw, namely an array of associated debilitating side effects. This is due to the fact that they bind tubulin indiscriminately, leading to the widespread destruction of both cancerous and healthy cells. However, the expression of several β-tubulin isotypes in a wide range of cell types may provide a foundation for the development of anti-tubulin drug derivatives with increased specificity for particular cancer cells. Recently, the role of β-tubulin isotypes in resistance to anti-mitotic drugs has become a topic of great interest, as the design of isotype-specific drugs would ultimately result in better treatment protocols [13]. Successful chemotherapeutic drugs must take advantage of the differences in the relative vulnerabilities of some target pathways in cancer cells versus normal cells. It is this varied distribution of tubulin isotypes that provides us with a possible link to their role in the polymerization and stability of MTs. Through a search of available protein sequence databases, we have previously identified a total of 10 unique human β-tubulin isotypes, all of which have related distinct amino acid sequences and are generally well conserved [14]. Despite having very similar sequences, specific regions of higher sequence variability have been identified (see Fig. 1). Each β-tubulin isotype has a unique pattern of expression ranging from highly specific expression for classes βIII, βIVa and βVI to constitutive expression for classes βI and βIV [15,16]. Class I β-tubulin is the most commonly expressed isotype in humans, and as such is also the most common isotype found in cancer cells [17]. Both βII and βIII tubulins have been observed at increased levels in human tumors [18–24].

Structural bioinformatics, which is based on predictions from primary protein sequences, has previously been shown to stimulate development of novel strategies for drug development [25]. Similarly, our aim is to develop a computational method that is capable of predicting relative differences in binding affinities to known drug binding sites within different tubulin isotypes, based on their primary amino acid sequences. More specifically, our technique aims at finding interaction sites (also called hot spots) that are located within existing binding sites. Hot spots are defined as a small subset of residues that account for most of the free energy of binding of a protein interface [26]. More precisely, they are usually associated with active sites where mutations cause a significant increase in the binding free energy [27,28]. The experimental detection and analysis of hot spots is currently performed through the use of techniques such as Ala scanning mutagenesis. This procedure is slow and labor-intensive as individual mutant proteins must be purified and analyzed separately [29]. Several computational methods also exist for the prediction of hot spots. One example also utilizes alanine scanning; however this is a virtual mutagenesis technique that uses potential energy functions and thermodynamic cycles to predict the change in the free energy of binding for a theoretical mutation within a protein interface [30]. Another method is based on identifying sequentially or structurally conserved residues with respect to the structure of a set of related protein complexes [31]. Finally, hot spot prediction can also be implemented by identifying chemically complemented polar residues surrounded by a ring of hydrophobic residues [27], thereby identifying interface regions with a well-matched protrusion and crevice [32], and by utilizing non-covalent interactions and structural information [33].

Each of the abovementioned methods requires detailed knowledge of the corresponding protein structure, which, for all but bovine tubulin isotype βII, is unavailable. As a result, we have chosen to use a class of computational prediction methods that searches for the locations of hot spots using only the primary protein sequence. One of the existing methods is based on the observation that active sites are often flanked by Pro [34]. Another, more rigorous prediction method is based on the observation that the selectivity of the protein–target interactions is characterized by the matching of periodicities within the energy distribution of electrons of the interacting molecules. Hot spots are those amino acids that contribute the most to the characteristic frequency of a given protein and, as a result, to the interaction energy with ligands [35]. These residues are generally found clustered within the vicinity of the protein's active sites, and as such, they play a crucial role in determining the structure of the sites [36]. We have chosen to utilize a technique referred to as the digital signal processing (DSP) method [37], which uses "characteristic frequencies" to identify regions of interest within a protein sequence [38]. DSP-based predictions have already been successfully applied in a number of studies investigating proteins such as interleukin-2

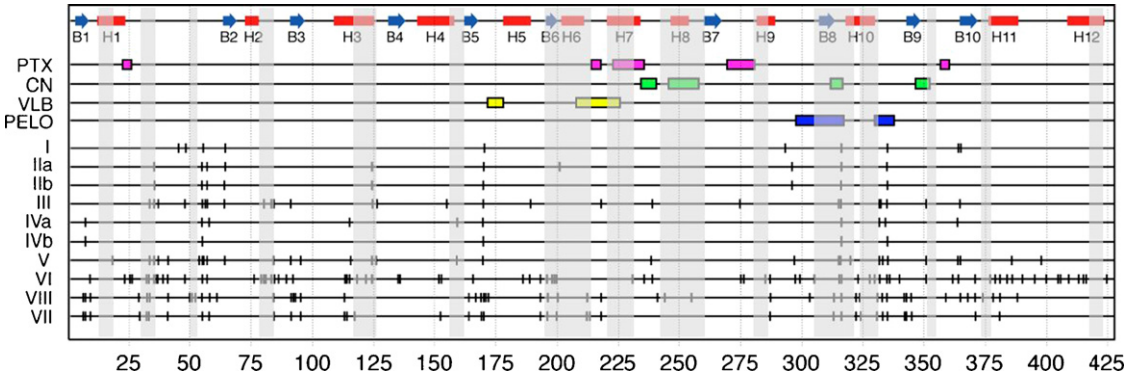

Fig. 1. Comparison of the 10 human β-tubulin sequences used in this study. Here, a schematic representation of the consensus sequence of human tubulin isotypes illustrates regions of dissimilarity. The first row represents the secondary structure of tubulin as determined by Nogales et al. [42]. This is followed by the regions (identified by boxes) within each polypeptide that have been shown to interact directly with paclitaxel (magenta), colchicine (green), vinblastine (yellow) and peloruside (blue). Finally, differences observed between the isotypes and a consensus sequence are represented by vertical ticks, which correspond to residues at which the sequence differs from a consensus sequence. Vertical gray boxes represent the location of all hot spots identified in Fig. 2. The hypervariable region, beginning after residue 425 (known as the C-terminal tail, for which no structural information exists), has not been shown.

[39], SV40 enhancer [40], Ha-ras p21 oncogene product, hemoglobins, myoglobins and lysozymes [41], p21 [36], and cytochrome *C* [37]. Here, we describe a method that predicts the relative differences in binding affinities by combining the characteristic frequency-based method for the prediction of hot spots and a novel scoring function which uses the predicted hot spots to compute a score that approximates the relative binding affinity. Using this method, we have correctly confirmed the locations of all known binding sites within tubulin and have computed the binding affinity scores for all known drug binding sites within 10 different isotypes of β-tubulin. In this way, the DSP computational prediction method can be used to predict high affinity interactions with novel ligands in isotypes or mutants that are yet to be experimentally explored.

2. Methodology

2.1. Identification of tubulin isotypes

Based on our previous examination of over 500 tubulin isotypes across all species in the SwissProt database, we chose a consensus set of 10 human tubulin sequences corresponding to isotypes βI, βIIa, βIIb, βIII, βIVa, βIVb, βV, βVI, βVII and βVIII upon which to perform the DSP analysis [14] (Fig. 1).

2.2. Structural analysis and binding site assignments

A number of structures of tubulin provide us with information as to the specific binding site locations for known chemotherapeutic

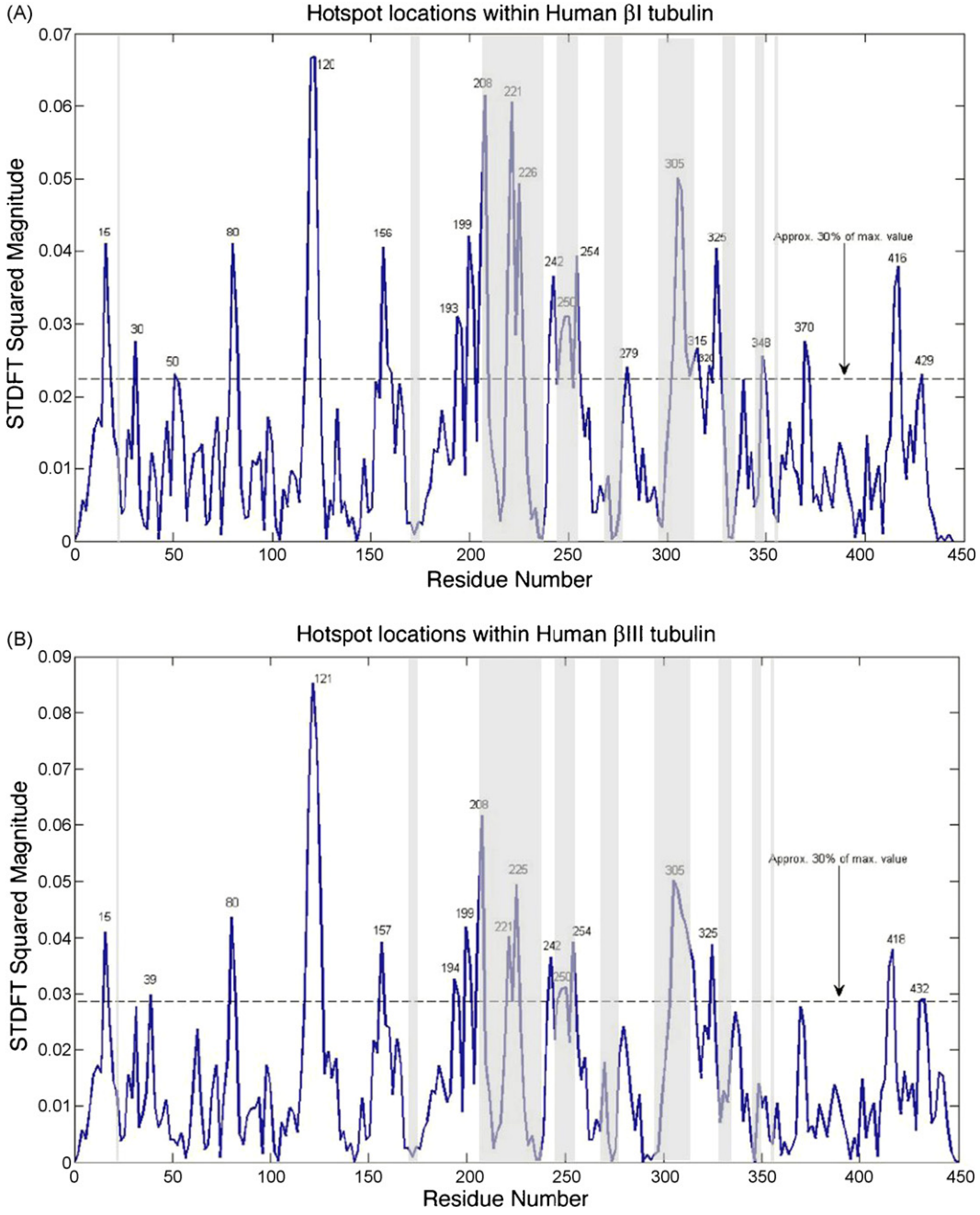

Fig. 2. Representative DSP spectrograms. DSP spectrograms for β-tubulin I (A) and β-tubulin III (B) were prepared using DSP analysis. Predicted hot spots have been chosen to include peaks with amplitudes of above 30% of the maximal amplitude and their immediate neighbors. Vertical gray boxes represent the location of the aggregate drug binding sites as shown in Fig. 1.

agents. The first tubulin structure with PDB identifier 1TUB was crystallized using docetaxel as a stabilizing agent [42]. However, due to difficulties in fitting electron density, this structure contained misalignments and was superseded by 1JFF, in which paclitaxel was subsequently utilized as a stabilizing agent [9]. Two higher resolution tubulin structures, 1SA0 and 1SA1, provide binding locations for colchicine and podophyllotoxin [5], while both the colchicine and vinblastine binding sites are observed in 1Z2B [11]. In addition to these crystallographically determined ligand binding sites, we have also analyzed residues within the putative binding site for peloruside, which was determined through hydrogen–deuterium exchange mass spectrometry and computational docking studies [57]. In the case of the present study, each of the binding surfaces was defined as those residues that were up to 6 Å from the bound ligand.

2.3. Computational prediction of hot spots

The hot spot prediction method involves applying the short-time discrete Fourier transform (STDFT) [43,44] and a model of protein–target interactions known as the resonant recognition model (RRM) [35,38]. First the sequences for a set of proteins (10–15) belonging to the same family, in this case the 10 tubulin isotypes, are converted into numerical sequences using electron–ion interaction potential (EIIP) values [45]. The EIIP value of a given amino acid denotes its average energy of valence electrons, which has been shown to be an optimal amino acid property that can be used in Fourier-based analyses of protein structure/function studies [45]. Next, a discrete Fourier transform (DFT) is computed for each numerical protein sequence and all DFTs are multiplied point-wise. The resulting product is known as the consensus spectrum and it reveals a characteristic frequency for the protein family. The characteristic frequency of a protein represents the frequency of the signal obtained from the energy distribution of the electrons in the molecule. According to the RRM, this signal is periodic and its frequency and phase characteristics are responsible for the selective interaction between a protein and its target(s). For a successful protein–target interaction, both the protein and the target must share the same characteristic frequency but have opposite phase. This corresponds to the fact that a peak in the squared magnitude of the bandpass-filtered protein sequence must match a corresponding trough in the squared magnitude of the bandpass-filtered target sequence. This matching resembles resonance and hence the characteristic frequency is said to provide resonant recognition between a protein and its target. As a result, regions in a protein sequence (using its numerical, EIIP-based, form) where its characteristic frequency is dominant correspond to hot spots.

Hot spots are then determined by using the STDFT, which is a widely used technique that performs time–frequency analysis of numerical sequences. In this technique, the protein sequence of interest is divided into a number of short overlapping sections using a sliding window function. The DFT of each section is then calculated and all these DFTs are consecutively arranged to form the columns of a matrix known as the STDFT matrix. Then, each column of the STDFT matrix is multiplied by the consensus spectrum. Plotting the resulting squared magnitude of the modified STDFT, we obtain a spectrogram in which the locations of the hot spots can be identified by the presence of distinct peaks (Fig. 2). For more details on the STDFT-based hot spot location technique please see Ramachandran et al. [37].

2.4. Hot spot scoring functions

The squared magnitude of the STDFT computed by the hot spot prediction algorithm can be used to quantify significance of a

Table 1
Predicted hot spots for β-tubulin isotypes

β-Tubulin isotype	Predicted hot spots for the below binding sites				
	GDP	Paclitaxel	Colchicine	Vinblastine	Peloruside
I	**15**, 16, 207, 222, 225, **226**	**225**, 227, 280	200, 249, **250**, 253, 314, **315**, 316, 347, **348**, 349	**208**, 220, **221**, 222, 225	304, **305**, 306, 314, **315**, 316
IIa	**15**, 16, 169, 222, **225**, 226	224, **225**	**200**, 249, **250**, 256, 314, **315**, 316	208, 220, **221**, 222, 224, **225**	169, 305, **306**, 307, 314, **315**, 316
IIb	**15**, 16, 169, 222, **225**, 226	224, **225**	**200**, 249, **250**, 256, 314, **315**, 316	208, 220, **221**, 222, 224, **225**	169, 305, **306**, 307, 314, **315**, 316
III	**15**, 16, 207, 222, **225**, 226	224, **225**	200, 249, **250**, 253	**208**, 220, **221**, 222, 224, **225**	304, **305**, 306
IVa	**15**, 16, 207, 222, **225**, 226	224, **225**, 280	200, 249, **250**, 253, 314, **315**, 316, 347, **348**, 349	**208**, 220, **221**, 222, 224, **225**	304, **305**, 306, 314, **315**, 316, 335, **336**, 337
IVb	**15**, 16, 222, **225**, 226	224, **225**, **280**	200, 249, **250**, 253, 314, **315**, 316, 348, **349**, 350	208, 220, **221**, 222, 224, **225**	305, **306**, 307, 314, **315**, 316, 338, **339**, 340
V	15, **16**, 207, 222, **225**, 226	224, **225**, 280, 357, **358**, 359	200, 249, **250**, 253, **312**, 313	**208**, 220, **221**, 222, 224, **225**	298, **299**, 300, 305, **306**, 307, 311, **312**, 313, 336, **337**, 338
VI	15, **16**, 207, 222, **225**, 226	224, **225**, 358, **359**, 360	**200**, 249, **250**, 253, 312, 368	**208**, 220, **221**, 222, 224, 225	304, **305**, 306, 310, **311**, 312
VII	207, 225, 252	223, **224**, 225, 280, 318	200, 239, **240**, 249, **250**, 252, **253**, 316, 347, **348**, 349	**208**, 219, **220**, 221, **224**, 225	304, **305**, 306, 316
VIII	207, **225**, 226	224, **225**, 364, **280**, **318**, 364	200, 249, **250**, 253, 347, **348**, 349	**208**, 224, **225**	304, **305**, 306, 334, **335**, 336

Binding sites were defined by residues that fell within 6 Å of bound ligands. Bold-faced numbers denote significant peaks on the corresponding spectrograms, while roman font denotes positions of the immediate peak neighbors.

particular hot spot with respect to a given binding site [37]. Additionally, although a peak in the spectrogram denotes an individual residue, the actual hot spot may also encompass several neighboring residues. The STDFT divides the protein sequence into a number of overlapping segments and processes these segments individually. Due to the fact that the segments are overlapping, the number of output values from the DSP is often bigger than the total number of residues. To correct for this discrepancy, we can simply rescale the *x*-axis of the generated plots to the correct range. Due to this post-processing, an exact one-to-one mapping between the amino acids and the output values cannot be obtained. In effect, this technique scans a small group of residues at a time to detect the strength of the characteristic frequency; it does not scan each amino acid separately. Since the hot spot prediction algorithm identifies small stretches rather than individual amino acids as hot spots, it is a small group of amino acids clustered around a peak that is taken as the hot spot.

2.5. Binding affinity estimations

Following the identification of possible hot spots, we are then able to compute relative binding affinity for a given binding site. We note that simply using the number of hot spots in a given site to quantify the relative binding affinity of that site is not sufficient (some binding sites may include several predicted hot spots, but have a low squared STDFT magnitude). Therefore, we have chosen to utilize both the number of the predicted hot spots together with their corresponding squared magnitude of the STDFT to develop two scoring functions. The first is an inter-binding-score that for a given binding site can be used to quantify the relative differences in potential ligand binding affinity across different isotypes and the other is an intra-binding score that reflects relative binding affinity of different binding sites within the same isotype.

The inter-binding score is based on the sum of squared STDFT magnitudes of all hot spots located within a given site:

$$\text{inter-binding-score}_i = \sum_{k \in \text{BS}_i} \text{sm}_k \times 100$$

where sm_k is the STDFT squared magnitude of the kth residue that includes a peak and the two neighboring amino acids while BS_i denotes the ith binding site. The magnitudes are summed only for the peaks and two neighboring residues that constitute the binding site, the squared magnitude of the neighbors are assumed equal to

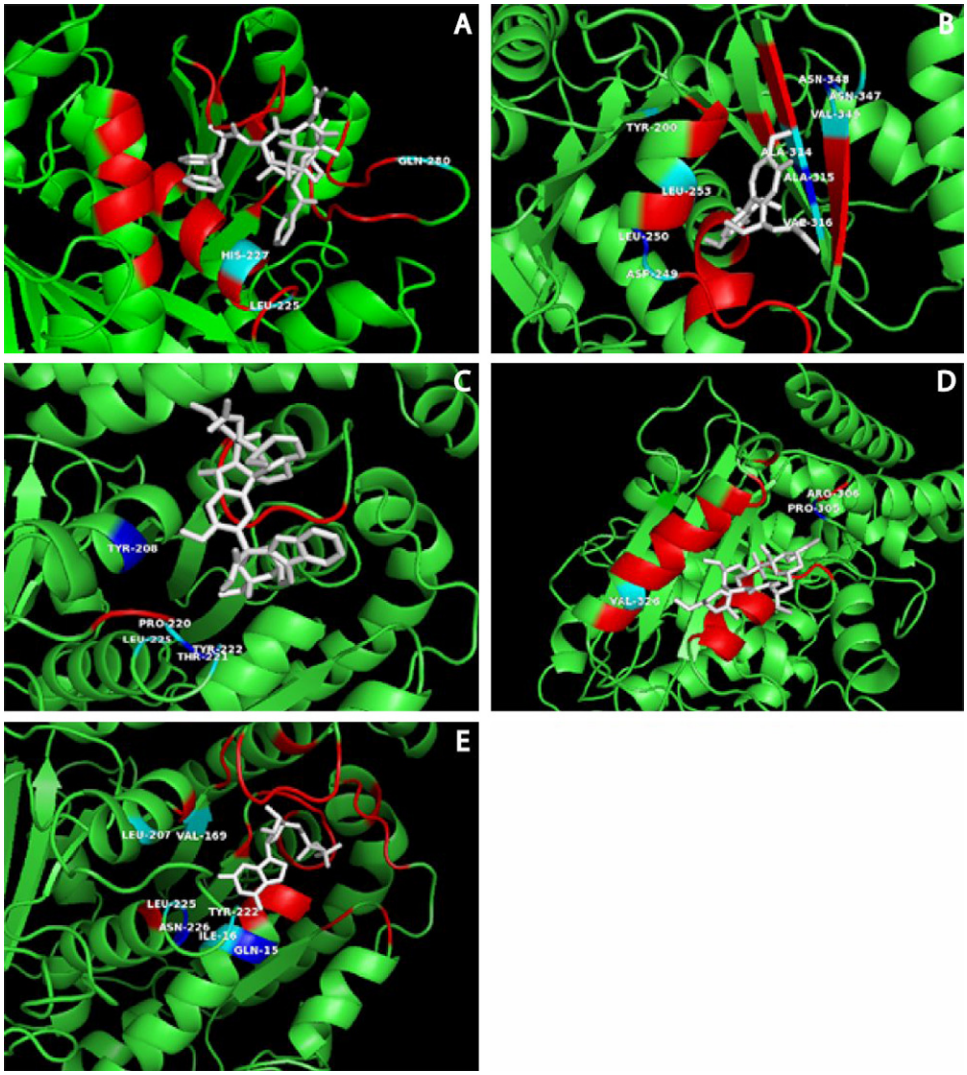

Fig. 3. The five ligands in their corresponding β-tubulin-binding site with the predicted hot spots. Drug positions within tubulin have been obtained from their respective PDB coordinates. In all panels, red cartoons indicate those residues within 6 Å of the ligands, blue cartoons with residue labels denote the predicted hot spots, while cyan denotes closest neighbors of the hot spots. (A) Binding site for paclitaxel, (B) colchicine, (C) vinblastine, (D) peloruside and (E) GDP.

the magnitude of the corresponding peak, and the sum is multiplied by 100 for convenience in order to obtain integer values.

While the above score is capable of expressing relative differences for the same site across proteins, the cumulative value may provide misleading results when evaluating binding across different sites (a site with a few strong interactions, i.e., large magnitudes, could be overshadowed by a site with large number of weaker interactions).

The intra-binding score is defined as the average score over all hot spots in a given site, i.e., the cumulative value is divided by the number of hot spots, which expresses the overall strength of the binding for a given site:

$$\text{intra-binding-score}_i = \sum_{k \in \text{BS}_i} \frac{\text{sm}_k}{N_i} \times 100$$

where N_i denotes the total number of hot spots (the peaks and two neighbors) found in the ith binding site.

This score should not be used to evaluate differences within a given site, i.e., it would provide a larger value for a site with strong hot spots (large magnitudes) when compared with the same site that includes the same strong hot spots and additional weak hot spots (small magnitudes). Finally, we note that the values obtained for both of these scores depend not only on the protein under investigation (and the resulting predicted hot spots), but also on the binding conformation of the specific ligands. This is because they are computed using only the hot spots found within the binding surface that is defined using a previously docked ligand.

3. Results and discussion

Using the DSP method, we have produced spectrograms for 10 identified human β-tubulin isotypes. However, due to the high degree of similarity between them, we have only shown two representative examples (Fig. 2). We have selected the spectrogram of βI tubulin as it is the most prevalent tubulin isotype, and the spectrogram of βIII tubulin as it is the most promising target for cancer chemotherapy development due to its over-expression in cancerous cells [18–24]. Each of the predicted hot spots from Fig. 2 has been chosen to include residues and their neighbors with peak amplitudes of at least 30% of the maximal amplitude (Table 1) [37]. The gray boxes shown in Fig. 2 correspond to the location of the aggregate binding sites, which allows for direct comparison between the two isotypes with respect to differences in the predicted hot spots that are located in the binding sites. This information could be coupled with the annotation of the predicted hot spots for the 10 considered human β-tubulin sequences, which is shown using gray bars in Fig. 1. We observed that several hot spots are predicted in locations where the two isotypes differ, which is denoted by vertical ticks on the corresponding two rows (I and III) which overlap with the gray shading. In order to facilitate the computation of binding affinity scores, the peaks from the histograms have been grouped by the corresponding binding sites for paclitaxel, colchicine, vinblastine, peloruside, and GDP. Within the β-tubulin monomer, the paclitaxel binding site is located on the face that contacts the MT lumen, while vinblastine, peloruside, and GDP bind at the inter-dimer interface between α- and β-tubulin heterodimers. Colchicine binds within the intradimer interface, presumably interfering with tubulin dimer assembly into MTs (Fig. 3). In each of these cases, there are several residues located within each binding site that the DSP method has identified as a potential hot spot. However, there seems to be no particular spatial order of the detected residues, other than perhaps that they are not located in the centers of the binding sites. For instance, the paclitaxel, colchicine, and peloruside binding sites show hot spots that flank the binding site, giving its general contour (Fig. 3).

In general, we observed that the predicted hot spots were distributed throughout the entire primary tubulin amino acid sequence. However, almost all of the hot spots that were identified as significant were restricted to residues located on the surface of tubulin, as identified by relative solvent accessibility (RSA) values. Fig. 4 compares the distribution of RSA values along the entire tubulin sequence, when constrained to the residues that compose the binding sites for GDP, paclitaxel, colchicine, vinblastine, and peloruside, and for the predicted hot spots. Among the predicted hot spots only seven residues, which corresponds to about 13% of all predictions, are buried (RSA value equals 0). Similar fraction of buried residues is observed among the amino acids that constitute the five binding sites. This was an interesting result and validation of this technique for identifying druggable sites, due to the fact that no information regarding the tertiary structure of tubulin was included in the analysis, yet almost 87% of the identified hot spots are restricted to the protein surface. At the same time, we observed that although some hot spots were associated with typically druggable sites such as grooves and pockets, other predictions were located on ridges and flats. Additionally, those surface hot spots that do not directly correlate to drug binding sites are located within a narrow band on the exterior surface of tubulin as it would be observed in a MT structure (Fig. 5). This is an interesting observation because this is the surface of β-tubulin that is most capable of making direct contact with MT associated proteins

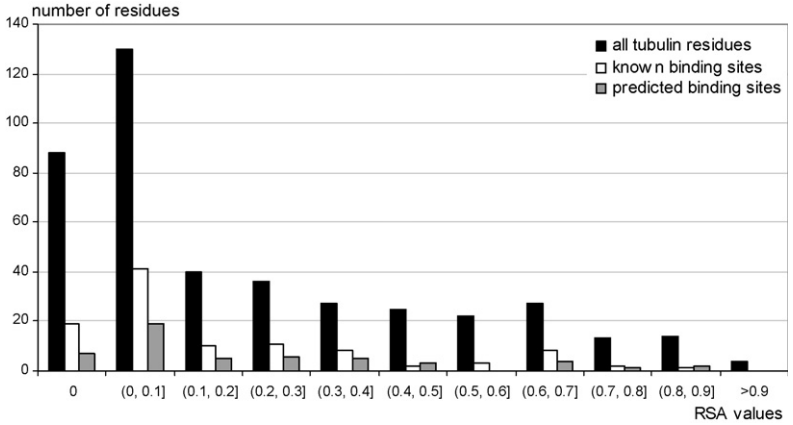

Fig. 4. Distribution of RSA values for the entire β-tubulin sequence, the residues that compose GDP, paclitaxel, colchicine, vinblastine, and peloruside binding sites, and for the predicted hot spots. The RSA values were binned into 0.1 wide intervals with the exception of the deeply buried residues, i.e., those with RSA = 0, which were shown separately.

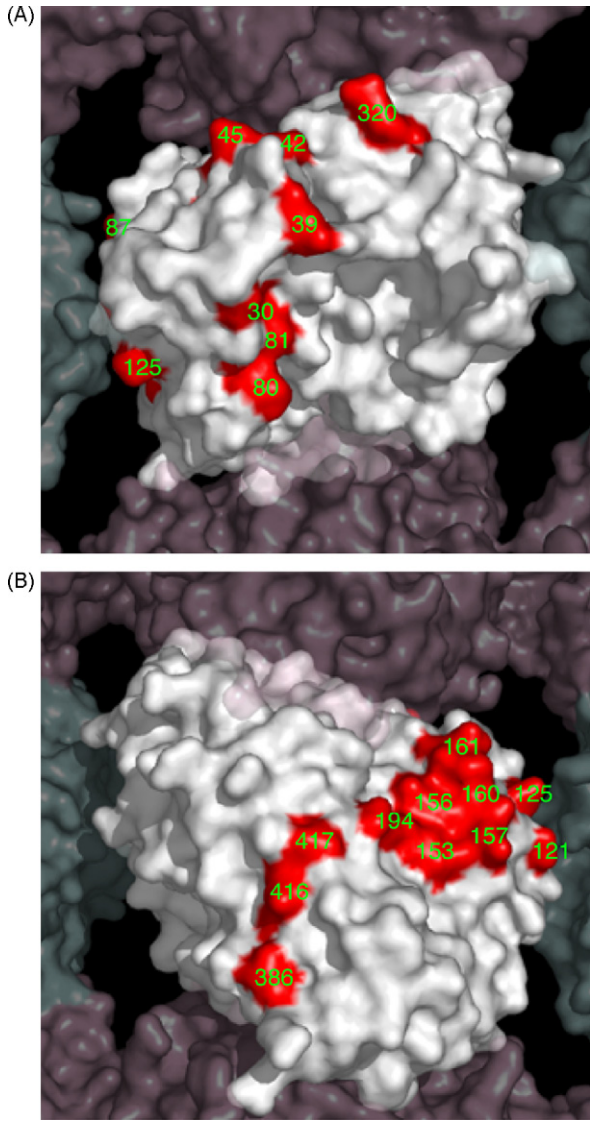

Fig. 5. Surface map of unassigned hot spots in all human β-tubulin isoforms. A solvent accessible surface was drawn onto the β-tubulin monomer (center) obtained by Nogales et al. [9] (PDB identifier 1JFF). This figure shows a single β-tubulin monomer within an MT containing 13 protofilaments, which was reconstructed as described by Li et al. [59]. Adjacent α-tubulin surfaces are colored pink and adjacent β-tubulin surfaces are colored cyan. Unassigned surface hotpots are colored red. Panel A illustrates the MT from the luminal surface. Panel B illustrates a 180° rotation about the *y*-axis to show the exterior surface of the microtubule.

(MAPs). Additionally, the majority of these "unassigned" hot spots that are not located on the surfaces of tubulin correlate with either their lateral or longitudinal interaction surfaces. This observation implies a conservation of function, during MT assembly, across all isotypes within these surface regions.

When we compare the βI and βIII spectrograms (Fig. 2), we note that several hot spots are shared between these isotypes (15, 80, 120/121, 156/157, 193/194, 199, 208, 221, 225/226, 242, 250, 254, 305, 325, 416/418, and 429/432), while some peaks are present only in the spectrogram of βI tubulin (30, 279, 315, 348, and 370). However, positions 30 and 370 are not associated with known binding sites, while 279, 315, and 348 are located within the paclitaxel and colchicine binding sites. This unique distribution of hot spots indicates a difference within the binding sites of these two proteins and may provide a reference point from which to construct drugs that differentiate between the βI and βIII isotypes.

Using the predicted hot spots in Table 1, we then computed inter-binding scores for the ligand binding sites across each isotype (Table 2). Interestingly, we observed a correlation between the order of known binding affinities to isotypes and the differences in the values of the proposed scores. We note that although our scores are quantitative, the known binding affinities for each ligand across isotypes can be expressed only in a qualitative fashion, as it is virtually impossible to normalize individual results.

For the paclitaxel site, we have based our relative values for binding affinity on the sensitivity of MT stabilization effects of each drug [46,47]. Using this measure, the order of isotype sensitivity to paclitaxel is βII > βIV > βIII, which differs from the order of hot spot scores although the three scores are all quite close. Our inter-binding scores indicate that βIV should bind paclitaxel more strongly than either βIII or βII, the latter having approximately the same binding affinity to paclitaxel, a result that qualitatively agrees with experimental results [46].

For colchicine and vinblastine, the experimentally determined order for binding affinity is loosely βIV > βII > βIII and βIV ≈ βII ≫ βIII, respectively. This trend correlates with our predicted inter-binding scores (see Table 2). On-rate measurements for colchicine [48] indicate that βIV > βII > βIII binding is preferred, while dynamic stability of the MT complex, as measured by the decay of binding, indicates that βIII > βII, or shows a reverse order to that of the weaker isotypes. Our inter-binding scores favor the dynamic order of isotypes, which tends to be a more complete set of experimental data. Vinblastine binding to tubulin isotypes has been quantitatively assessed using the half maximal inhibitory concentration (IC_{50}) technique for MT depolymerization and by measuring tubulin polymerization rates under various conditions of vinblastine concentration [46,47]. The IC_{50} measurements suggest that preferred binding affinity is βIV ≈ βII ≫ βIII. In essence, vinblastine binds most strongly to βIV, almost as strongly to βII, and much less strongly to βIII. It should be kept in mind, however, that these binding values were not measured directly in these studies. Nonetheless, our inter-binding scores match exactly this order of isotype affinities for vinblastine.

Peloruside is a recently discovered macrolide that appears to occupy a site distinct from that of the taxanes and functions synergistically with other MT stabilizing agents [49,50]. Peloruside is an attractive chemotherapy drug, as it retains activity in multidrug resistance lines and is able to block growth in cell lines resistant to paclitaxel and epothilone A arising from induced mutations in β-tubulin [49,51]. While there is currently no published data available on the location of the peloruside binding site on tubulin, we have recently performed an analysis using ligand–receptor docking, coupled with deuterium exchange mass spectrometry to identify a putative peloruside binding site on the surface of tubulin [57]. In this study, the lowest energy peloruside conformation was identified as interacting with residues 296–304 and 333–342 within β-tubulin. Using our hot spot analysis within this region, we can make a prediction that binding affinities among the three isotypes tested should be βIV > βII = βIII. This result remains to be experimentally verified. Interestingly, the peloruside isotype inter-binding affinity ordering is very close to that observed for the paclitaxel intra-binding affinity.

The binding of guanosine triphosphate (GTP) to tubulin affects the stability of MTs, as only tubulin dimers with both the α- and β-tubulin sites bound with GTP are considered assembly competent. Conversely, the hydrolysis of GTP into GDP at the exchangeable site within β-tubulin leads to MT disassembly. Previously, MT dynamicity assays measured the assembly and disassembly rates of the three tubulin isotypes and found that the on-rates, as well as

Table 2
Ligand binding scores over all binding sites and tubulin isotypes

Binding site	β-Tubulin isotype									
	I	IIa	IIb	III	IVa	IVb	V	VI	VII	VIII
Inter										
GDP	30	27	27	28	30	24	30	30	14	16
Paclitaxel	12	10	10	10	12	12	18	16	22	19
Colchicine	32	23	23	14	32	32	22	21	44	25
Vinblastine	29	34	34	28	34	34	34	34	28	16
Peloruside	14	14	14	14	18	18	22	22	16	15
Intra										
GDP	5.0	4.5	4.5	4.7	5.0	4.8	5.0	5.0	4.7	5.3
Paclitaxel	4.0	5.0	5.0	5.0	4.0	4.0	3.0	3.2	4.4	3.8
Colchicine	3.2	3.3	3.3	3.5	3.2	3.2	3.7	3.5	4.0	3.6
Vinblastine	5.8	5.7	5.7	4.7	5.7	5.7	5.7	5.7	4.7	5.3
Peloruside	4.0	3.9	3.9	5.0	3.7	3.3	3.5	5.5	5.0	4.0

The intra-binding affinity scores represent comparisons between binding sites within each isotype (columns). The inter-binding affinity scores represent comparisons for each binding site across all isotypes (rows). Each score was computed based on STDFT squared magnitude values for hot spots, which include each predicted peak and its two neighboring residues that belong to a given binding site.

overall dynamicity were ordered according to βIII > βIV > βII, suggesting that βIII is the most dynamic and βII is the least dynamic of these three isotypes [52]. Additionally, measurement of MT decay rates at 0 °C for MTs composed of purified tubulin isotypes also found that βIII produced MTs that were more stable than those composed of βII [53]. Our own intra-binding affinity predictions for the GTP site suggests that βIII is more stable than βIVb but less than βIVa and should also be more stable than either βIIa or βIIb. The same conclusion also applies to βIVa but not βIVb. In other words, the results from our analysis are a bit more subtle, due to the distinction between the *a* and *b* variants of both βII and βIV. Moreover, there is an additional complication due to the fact that GTP binding affinity determines both the on- and off-rate for tubulin dimers, each of which affects a different process (dynamicity and stability). Furthermore, more recent data on the assembly dynamics of isotypically purified tubulin [58] indicate much smaller differences between the isotypes that has been reported in earlier publications.

As we are unable to directly compare the intra-binding affinity scores between isotypes, we cannot comment on their relative magnitudes and how this relates to drug binding between different sites on tubulin. We can, however, compare the qualitative trends for drug binding affinities within the binding sites of each individual isotype. In general, we can see from Table 2 that each of the intra-binding site affinity scores follows the overall trend of GDP > vinblastine > colchicine > peloruside > paclitaxel. While this general trend is observed for all of the isotypes, the GDP and colchicine binding-site affinity scores deviate from this trend in the βVII and βVIII isotypes. This general trend for drug binding affinity correlate with experimental binding affinities that demonstrate the greatest binding affinity for vinblastine [11], followed by peloruside [50] and paclitaxel [54], with the lowest binding affinities reported for colchicine [55]. While we are unable to definitively comment on the differences in binding affinity scores across isotypes, the observation that differences in MT isotype composition [55] alter colchicine binding kinetics is supported by the variability in our intra-binding site scores.

4. Conclusions

Through the use of DSP analysis of human tubulin isotype sequences, we can report on a number of encouraging results regarding known tubulin-binding ligands. In this study we have demonstrated that the binding affinities of tubulin-binding ligands qualitatively correlate between previously reported experimental results and our present computational predictions. In addition, we have also identified several interesting residues that may lead to the identification of putative binding sites and which have not yet been assigned to any known tubulin-binding compounds or protein interaction surfaces. It will be of great interest to experimentally examine these residues in future research aimed at discovering new chemotherapy agents that target tubulin and, more specifically, tubulin isotypes. The computational prediction methods that are outlined here can also be used to predict possible high affinity regions within the binding sites of other proteins. The key feature of the proposed method is the fact that it requires only the knowledge of the corresponding primary structures, which would allow applications into proteins with unsolved tertiary structures.

The above results could be attributed to the performed DSP analysis that draws upon the biochemical properties of a protein's sequence, namely the EIIP sequence, which signifies the average energy of each residue's valence electrons. Since ionization potential correlates with the strength with which attractive Van der Waals energies can bind atoms from molecules in the surrounding environment, e.g., from a ligand, the performed DSP-based prediction of hot spots may relate to the identification of amino acids positioned to form such contacts. This point of view would link the success of DSP predictions to thermodynamic binding computations indicating that favorable Van der Waals contacts at a binding interface provide the largest contributions to hot spot interactions [56].

Acknowledgements

This research was supported by grants from NSERC, Alberta Cancer Board, The Allard Foundation and MITACS. HF is grateful for the post-doctoral fellowship funding from PIMS and The Avadh Bhatia Foundation. KC would like to acknowledge the support provided by the Alberta Ingenuity Student Scholarship program.

References

[1] K.C. Chou, C.T. Zhang, G.M. Maggiora, Solitary wave dynamics as a mechanism for explaining the internal motion during microtubule growth, Biopolymers 34 (1994) 143–153.
[2] M.E. Pichichero, C.J. Avers, The evolution of cellular movement in eukaryotes: the role of microfilaments and microtubules, Subcell. Biochem. 2 (1973) 97–105.
[3] J.S. Hyams, H. Stebbings, The mechanism of microtubule associated cytoplasmic transport. Isolation and preliminary characterisation of a microtubule transport system, Cell Tissue Res. 196 (1979) 103–116.

[4] H.W. Wang, E. Nogales, Nucleotide-dependent bending flexibility of tubulin regulates microtubule assembly, Nature 435 (2005) 911–915.
[5] R.B. Ravelli, B. Gigant, P.A. Curmi, I. Jourdain, S. Lachkar, A. Sobel, M. Knossow, Insight into tubulin regulation from a complex with colchicine and a stathmin-like domain, Nature 428 (2004) 198–202.
[6] M.C. Wani, H.L. Taylor, M.E. Wall, P. Coggon, A.T. McPhail, Plant antitumor agents. VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*, J. Am. Chem. Soc. 93 (1971) 2325–2327.
[7] P.B. Schiff, S.B. Horwitz, Taxol stabilizes microtubules in mouse fibroblast cells, Proc. Natl. Acad. Sci. U.S.A. 77 (1980) 1561–1565.
[8] I. Ringel, S.B. Horwitz, Studies with RP 56976 (taxotere): a semisynthetic analogue of taxol, J. Natl. Cancer Inst. 83 (1991) 288–291.
[9] J. Lowe, H. Li, K.H. Downing, E. Nogales, Refined structure of alpha beta-tubulin at 3.5 A resolution, J. Mol. Biol. 313 (2001) 1045–1057.
[10] J.H. Nettles, H. Li, B. Cornett, J.M. Krahn, J.P. Snyder, K.H. Downing, The binding mode of epothilone A on alpha, beta-tubulin by electron crystallography, Science 305 (2004) 866–869.
[11] B. Gigant, C. Wang, R.B. Ravelli, F. Roussi, M.O. Steinmetz, P.A. Curmi, A. Sobel, M. Knossow, œStructural basis for the regulation of tubulin by vinblastine, Nature 435 (2005) 519–522.
[12] R.J. Toso, M.A. Jordan, K.W. Farrell, B. Matsumoto, L. Wilson, Kinetic stabilization of microtubule dynamic instability in vitro by vinblastine, Biochemistry 32 (1993) 1285–1293.
[13] C.A. Burkhart, M. Kavallaris, S. Band Horwitz, The role of beta-tubulin isotypes in resistance to antimitotic drugs, Biochim. Biophys. Acta 1471 (2001) 01–9.
[14] J.T. Huzil, R.F. Luduena, J. Tuszynski, Comparative modelling of human beta-tubulin isotypes and implications for drug binding, Nanotechnology 17 (2006) S90–S100.
[15] N.J. Cowan, L. Dudley, Tubulin isotypes and the multigene tubulin families, Int. Rev. Cytol. 85 (1983) 147–173.
[16] H.C. Jensen-Smith, R.F. Luduena, R. Hallworth, Requirement for the betaI and betaIV tubulin isotypes in mammalian cilia, Cell Motil. Cytoskeleton 55 (2003) 213–220.
[17] K.F. Sullivan, D.W. Cleveland, Identification of conserved isotype-defining variable region sequences for four vertebrate beta tubulin polypeptide classes, Proc. Natl. Acad. Sci. U.S.A. 83 (1986) 4327–4331.
[18] C.A. Scott, C.C. Walker, D.A. Neal, C.E. Harper, R.A. Bloodgood, K.D. Somers, S.E. Mills, L.I. Rebhun, P.A. Levine, Beta-tubulin epitope expression in normal and malignant epithelial cells, Arch. Otolaryngol. Head Neck Surg. 116 (1990) 583–589.
[19] S. Mozzetti, C. Ferlini, P. Concolino, F. Filippetti, G. Raspaglio, S. Prislei, D. Gallo, E. Martinelli, F.O. Ranelletti, G. Ferrandina, G. Scambia, Class III beta-tubulin overexpression is a prominent mechanism of paclitaxel resistance in ovarian cancer patients, Clin. Cancer Res. 11 (2005) 298–305.
[20] L. Prasannan, D.E. Misek, R. Hinderer, J. Michon, J.D. Geiger, S.M. Hanash, Identification of beta-tubulin isoforms as tumor antigens in neuroblastoma, Clin. Cancer Res. 6 (2000) 3949–3956.
[21] S. Ranganathan, D.W. Dexter, C.A. Benetatos, A.E. Chapman, K.D. Tew, G.R. Hudes, Increase of beta(III)- and beta(IVa)-tubulin isotopes in human prostate carcinoma cells as a result of estramustine resistance, Cancer Res. 56 (1996) 2584–2589.
[22] J.H. Dozier, L. Hiser, J.A. Davis, N.S. Thomas, M.A. Tucci, H.A. Benghuzzi, A. Frankfurter, J.J. Correia, S. Lobert, Beta class II tubulin predominates in normal and tumor breast tissues, Breast Cancer Res. 5 (2003) R157–R169.
[23] R.E. Ferguson, C. Taylor, A. Stanley, E. Butler, A. Joyce, P. Harnden, P.M. Patel, P.J. Selby, R.E. Banks, Resistance to the tubulin-binding agents in renal cell carcinoma: no mutations in the class I beta-tubulin gene but changes in tubulin isotype protein expression, Clin. Cancer Res. 11 (2005) 3439–3445.
[24] M. Haber, C.A. Burkhart, D.L. Regl, J. Madafiglio, M.D. Norris, S.B. Horwitz, Altered expression of M beta 2, the class II beta-tubulin isotype, in a murine J774.2 cell line with a high level of taxol resistance, J. Biol. Chem. 270 (1995) 31269–31275.
[25] K.C. Chou, Structural bioinformatics and its impact to biomedical science, Curr. Med. Chem. 11 (2004) 2105–2134.
[26] T. Clackson, J.A. Wells, A hot spot of binding energy in a hormone–receptor interface, Science 267 (1995) 383–386.
[27] A.A. Bogan, K.S. Thorn, Anatomy of hot spots in protein interfaces, J. Mol. Biol. 280 (1998) 1–9.
[28] I.S. Moreira, P.A. Fernandes, M.J. Ramos, Hot spots—a review of the protein–protein interface determinant amino-acid residues, Proteins (2007).
[29] W.L. DeLano, Unraveling hot spots in binding interfaces: progress and challenges, Curr. Opin. Struct. Biol. 12 (2002) 14–20.
[30] S. Huo, I. Massova, P.A. Kollman, Computational alanine scanning of the 1:1 human growth hormone–receptor complex, J. Comput. Chem. 23 (2002) 15–27.
[31] B. Ma, T. Elkayam, H. Wolfson, R. Nussinov, Protein–protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 5772–5777.
[32] X. Li, O. Keskin, B. Ma, R. Nussinov, J. Liang, Protein–protein interactions: hot spots and structurally conserved residues often locate in complemented pockets that pre-organized in the unbound states: implications for docking, J. Mol. Biol. 344 (2004) 781–795.
[33] Y. Gao, R. Wang, L. Lai, Structure-based method for analyzing protein–protein interfaces, J. Mol. Model. 10 (2004) 44–54.
[34] R.M. Kini, H.J. Evans, Prediction of potential protein–protein interaction sites from amino acid sequence. Identification of a fibrin polymerization site, FEBS Lett. 385 (1996) 81–86.
[35] I. Cosic, Macromolecular bioactivity: is it resonant interaction between macromolecules?—theory and applications, IEEE Trans. Biomed. Eng. 41 (1994) 1101–1114.
[36] E. Pirogova, Q. Fang, M. Akay, I. Cosic, Investigation of the structural and functional relationships of oncogene proteins, Proc. IEEE 90 (2002) 1859–1867.
[37] P. Ramachandran, A. Antoniou, P.P. Vaidyanathan, Identification and location of hot spots in proteins using the short-time discrete Fourier transform, in: Proceedings of the 38th Asilomar Conference on Signals, Systems, and Computers, vol. 2, 2004, pp. 1656–1660.
[38] E. Pirogova, G.P. Simon, I. Cosic, Investigation of the applicability of dielectric relaxation properties of amino acid solutions within the resonant recognition model, IEEE Trans. Nanobiosci. 2 (2003) 63–69.
[39] I. Cosic, M. Pavlovic, V. Vojisavljevic, Prediction of "hot spots" in interleukin-2 based on informational spectrum characteristics of growth-regulating factors. Comparison with experimental data, Biochimie 71 (1989) 333–342.
[40] I. Cosic, D. Nesic, Prediction of 'hot spots' in SV40 enhancer and relation with experimental data, Eur. J. Biochem. 170 (1987) 247–252.
[41] I. Cosic, M.T. Hearn, 'Hot spot' amino acid distribution in Ha-ras oncogene product p21: relationship to guanine binding site, J. Mol. Recognit. 4 (1991) 57–62.
[42] E. Nogales, S.G. Wolf, K.H. Downing, Structure of the alpha beta tubulin dimer by electron crystallography, Nature 391 (1998) 199–203.
[43] A.V. Oppenheim, R.W. Schafer, Discrete-time Signal Processing, Prentice Hall Inc., New Jersey, 1999.
[44] M. Portnoff, Time-frequency representation of digital signals and systems based on short-time Fourier analysis, IEEE Trans. Acoust. Speech Signal Process. 28 (1980) 55–69.
[45] J. Lazovic, Selection of amino acid parameters for Fourier transform-based analysis of proteins, Comput. Appl. Biosci. 12 (1996) 553–562.
[46] W.B. Derry, L. Wilson, I.A. Khan, R.F. Luduena, M.A. Jordan, Taxol differentially modulates the dynamics of microtubules assembled from unfractionated and purified beta-tubulin isotypes, Biochemistry 36 (1997) 3554–3562.
[47] I.A. Khan, R.F. Luduena, Different effects of vinblastine on the polymerization of isotypically purified tubulins from bovine brain, Invest. New Drugs 21 (2003) 3–13.
[48] A. Banerjee, A. D'Hoore, Y. Engelborghs, Interaction of desacetamidocolchicine, a fast binding analogue of colchicine with isotypically pure tubulin dimers alpha beta II, alpha beta III, and alpha beta IV, J. Biol. Chem. 269 (1994) 10324–10329.
[49] S.L. Mooberry, G. Tien, A.H. Hernandez, A. Plubrukarn, B.S. Davidson, Laulimalide and isolaulimalide, new paclitaxel-like microtubule-stabilizing agents, Cancer Res. 59 (1999) 653–660.
[50] E. Hamel, B.W. Day, J.H. Miller, M.K. Jung, P.T. Northcote, A.K. Ghosh, D.P. Curran, M. Cushman, K.C. Nicolaou, I. Paterson, E.J. Sorensen, Synergistic effects of peloruside A and laulimalide with taxoid site drugs, but not with each other, on tubulin assembly, Mol. Pharmacol. 70 (2006) 1555–1564.
[51] D.E. Pryor, A. O'Brate, G. Bilcer, J.F. Diaz, Y. Wang, Y. Wang, M. Kabaki, M.K. Jung, J.M. Andreu, A.K. Ghosh, P. Giannakakou, E. Hamel, The microtubule stabilizing agent laulimalide does not bind in the taxoid site, kills cells resistant to paclitaxel and epothilones, and may not require its epoxide moiety for activity, Biochemistry 41 (2002) 9109–9115.
[52] D. Panda, H.P. Miller, A. Banerjee, R.F. Luduena, L. Wilson, Microtubule dynamics in vitro are regulated by the tubulin isotype composition, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 11358–11362.
[53] P.M. Schwarz, J.R. Liggins, R.F. Luduena, Beta-tubulin isotypes purified from bovine brain have different relative stabilities, Biochemistry 37 (1998) 4687–4692.
[54] J.F. Diaz, I. Barasoain, J.M. Andreu, Fast kinetics of taxol binding to microtubules. Effects of solution variables and microtubule-associated proteins, J. Biol. Chem. 278 (2003) 8407–8419.
[55] A. Banerjee, R.F. Luduena, Distinct colchicine binding kinetics of bovine brain tubulin lacking the type III isotype of beta-tubulin, J. Biol. Chem. 266 (1991) 1689–1691.
[56] V. Lafont, M. Schaefer, R.H. Stote, D. Altschuh, A. Dejaegere, Protein–protein recognition and interaction hot spots in an antigen–antibody complex: free energy decomposition identifies "efficient amino acids", Proteins 67 (2007) 418–434.
[57] J.T. Huzil, J.K. Chik, G.W. Slysz, H. Freedman, J. Tuszynski, R.E. Taylor, D.L. Sackett, D.C. Schriemer, A unique mode of microtubule stabilization induced by peloruside A, J. Mol. Biol. 378 (2008) 1016–1030.
[58] V. Rezania, O. Azarenko, M.A. Jordan, H. Bolterauer, R.F. Ludueña, J.T. Huzil, J.A. Tuszynski, Microtubule assembly of isotypically purified tubulin and its mixtures, Biophys. J. 95 (4) (2008) 1993–2008.
[59] H. Li, D.J. DeRosier, W.V. Nicholson, E. Nogales, K.H. Downing, Microtubule structure at 8 angstrom resolution, Structure 10 (10) (2002) 1317–1328.