J Mol Model (2005) 11: 48–54
DOI 10.1007/s00894-004-0220-y

ORIGINAL PAPER

Suzanne L. Cunningham · Albert R. Cunningham
Billy W. Day

CoMFA, HQSAR and molecular docking studies of butitaxel analogues with β-tubulin

Received: 5 May 2004 / Accepted: 8 November 2004 / Published online: 23 December 2004

Abstract Results from biochemical analyses for a series of 20 butitaxel analogues, paclitaxel and docetaxel were used to build two- and three-dimensional quantitative structure-activity relationship (QSAR) models in order to investigate the properties associated with microtubule assembly and stabilization. A comparative molecular field analysis (CoMFA) model was built using steric and electrostatic fields. The CoMFA model yielded an r^2 of 0.943 and a cross-validated r^2 (i.e. q^2) of 0.376. Hologram quantitative structure-activity relationship (HQSAR) modeling of these same data generated an r^2 of 0.919 and a q^2 of 0.471. Contour maps used to visualize the steric and electrostatic contributions associated with activity or lack thereof were, as expected, localized to the varied position of the taxane system. Each analogue was docked successfully into a model of β-tubulin derived from previously determined cryoelectron microscopy analyses of the tubulin α/β heterodimer. All analogues superimposed well with paclitaxel bound to the protein, as well as with each other. Defining the variable region of each structure as the ligand and docking it separately into the paclitaxel site revealed a modest correlation ($r^2 = 0.53$) between activity and docking energy of all the compounds in the dataset. When only the butitaxel derivatives were considered, the correlation increased to 0.61. The mathematical models derived here provide information for the future development of taxanes.

Keywords Taxanes · CoMFA · CoMSIA · FlexiDock

S. L. Cunningham
Department of Environmental and Occupational Health,
University of Pittsburgh, Pittsburgh PA, 15261, USA

S. L. Cunningham · A. R. Cunningham
Department of Environmental Studies,
Louisiana State University, Baton Rouge LA,
70803, USA

B. W. Day
Departments of Pharmaceutical Sciences and of Chemistry,
University of Pittsburgh, Pittsburgh PA, 15261, USA

S. L. Cunningham (✉)
1285 Energy, Coast and Environment Building,
Baton Rouge LA, 70803, USA
E-mail: scunni3@lsu.edu
Tel.: +1-225-5784253
Fax: +1-225-5784286

Introduction

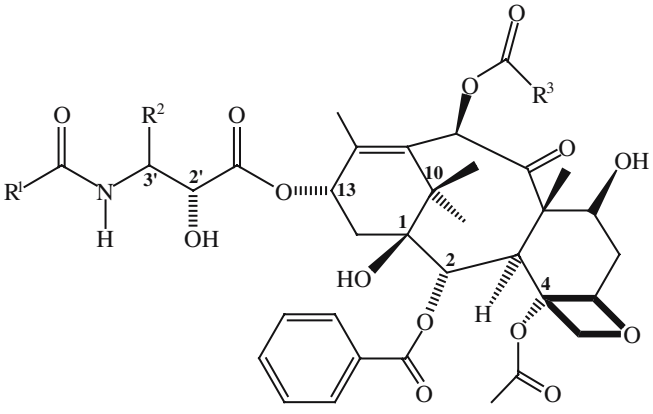
Microtubules, cytoskeletal polymers made up of tubulin heterodimers, are integral to essential cell functions [1]. They are involved in cell division, motility and shape, and provide the rail system for intracellular transport [2]. When microtubule stability is compromised, cell division may be altered or blocked, often leading to a block of the cell cycle at G2/M and apoptosis [3]. The essential functions of tubulin/microtubules in the cell cycle have made several tubulin-interactive agents clinically-useful as anti-cancer drugs [4].

Microtubule assembly results from the GTP hydrolysis-driven process of tight, non-covalent interactions between heterodimers of α- and β-tubulin. The result is tubes (microtubules), which consist normally of thirteen apparent protofilaments, a result of the three-start, left-handed helical structure of the lateral associations between heterodimers. Heterodimers continually add at the growing end and leave the shrinking end of the microtubule. During interphase, an array of microtubules emanates from the single centrosome of the cell and provides tracks for motor proteins to deliver cargo amongst the nuclear envelope, organelles and the cell periphery. As the cell cycle progresses, the centrosome duplicates and the cytoskeleton is broken down. The mitotic spindle is then formed by reassembly of the microtubules beginning at the two centrosomes, and it is the bipolar mitotic spindle that is responsible for separation of the sister chromatids into two identical daughter cells. This time of dramatically increased dynamics of growth and shrinkage of microtubules is the target of tubulin-interactive anticancer agents [3].

Taxanes are a group of antimitotic agents including paclitaxel (Taxol) and docetaxel (Taxotere) that are classified as microtubule stabilizers or hypernucleators. These chemicals, like other mechanistically-related natural products (discodermolide, epothilones, eleutherobins, sarcodictyins, laulimalides and peloruside), bind β-tubulin in microtubules, causing alterations in lateral contacts between protofilaments. This reorganization alters the protofilament number, stabilizes the polymer against shrinkage, and ultimately causes bundling of microtubules. One major end result is interference with proper mitotic spindle formation and the dynamics of the microtubules within it, leading to difficulties in, or prevention of, cell division. Cells treated with microtubule stabilizers are unable to proceed through G2/M of the cell cycle. Therefore, the use of taxanes in cancer treatment has become widespread since the 1990's, when paclitaxel and then docetaxel were found to show good activity against some solid tumors [5]. Recent goals in taxane development have been aimed to increasing potency and aqueous solubility [6].

A vast amount of chemical structure and biological activity data exists for taxanes. Significant loss or gain in activity is associated with substitutions at C1, C2 and C4 of the baccatin nucleus (see structure shown in Table 1). Specifically, removal of the hydroxyl group at C1 has been shown to reduce activity modestly. The benzoyloxy group at C2 is essential to maintaining activity while substituting acyl groups at C2 and C4 yields increased overall activity. Opening the oxetane ring reduces the activity of taxanes, as does removal of the acetate at C4. In the ester emanating from the baccatin ring system, the N′-acyl group is required for maintaining activity [7]. Ojima et al. [8, 9] found that modification at C10 and alkenyl or alkyl group replacement of the 3′ phenyl, as well as substituted

Table 1 Biological activities used for QSAR analyses and computed docking energies for paclitaxel, docetaxel and butitaxel analogues



Compound	R1	Microtubule assembly assay	B16 melanoma cytotoxicity	Docking energy (kcal mol^{-1})
		(EC_{50}/EC_{50}(paclitaxel))		
1 Paclitaxel	R^1 = Ph, R^2 = Ph, R^3 = Ac	1.00	1.00	−66.82
2 Docetaxel	R^1 = tert-BuO, R^2 = Ph, R^3 = H	0.45	0.41	−53.43
3 Butitaxel	R^1 = Ph, R^2 = tert-Bu, R^3 = H	1.80	7.50	−65.85
4 tert-BuO	R^1 = tert-BuO, R^2 = tert-Bu, R^3 = H	0.38	0.40	−67.68
5	Ethoxy	1.40	3.40	−53.12
6	Propoxy	1.70	1.50	−52.44
7	Butoxy	0.86	2.10	−58.21
8	Hexyloxy	2.30	12.00	−55.76
9	Isopropoxy	0.84	1.20	−54.17
10	Cyclopropyl	0.95	2.10	−65.25
11	Cyclobutyl	0.67	3.00	−61.12
12	Cyclopentyl	1.50	1.40	−60.03
13	1-Ethylpropyl	1.70	26.00	−65.49
14	Cyclohexyl	1.50	3.80	−53.34
15	2-Furyl	1.50	2.10	−63.29
16	3-Furyl	1.60	2.30	−47.72
17	2-Thienyl	0.42	0.53	−64.29
18	3-Thienyl	0.85	1.20	−68.68
19	5-methyl-2-thienyl	2.40	32.00	−56.54
20	3-methyl-2-thienyl	1.40	1.30	−69.22
21	2-thienyl-2-ethenyl	2.70	2.70	−12.75
22	2-Thienylmethyl	1.90	3.30	−57.02
23	1-Methyl-2-pyrrolyl	0.96	1.20	−58.57

Data obtained from Ali et al. [10]

benzoyl groups at C2, showed increased potency against drug-resistant cancer cell lines.

Ali et al. [10] synthesized a series of 20 butitaxel derivatives and tested them for microtubule assembly enhancing activity and cytotoxicity against B16 melanoma cells. These compounds, derived from replacement of the 3′-phenyl group of paclitaxel, showed excellent in vitro cytotoxicity and improved water solubility. Here, we used the structural and activity data of the butitaxels to develop two- and three-dimensional quantitative structure-activity relationship (QSAR) models in order to investigate the chemical properties that are associated with biological activity. Because the data came from a congeneric series of compounds and the tests were performed within the same laboratory, thus eliminating interlaboratory variation, these structures served as excellent probe molecules through which to study the chemical properties associated with activity and ligand-target interactions at the paclitaxel binding pocket of β-tubulin. Through the use of comparative molecular field analysis (CoMFA), partial least squares (PLS) and hologram QSAR (HQSAR) modeling, models yielding insight into the differences in potency based upon chemical structure were developed. Additionally, a docking procedure incorporating data from the paclitaxel binding region of β-tubulin was used to investigate the molecular interactions between the ligands and their target.

Materials and methods

Data source and structures

All structure-activity data for this study were taken directly from the report of Ali et al. [10]. Chemical structures and potencies for the microtubule polymerization and B16 melanoma cytotoxicity assays are presented in Table 1. The coordinates of α/β tubulin bound with taxane from the high resolution electron crystallographic analyses of Lowe et al. [11] were used as the receptor and starting ligand structures in the molecular docking study. The structures for butitaxel analogues, paclitaxel and docetaxel are available through the corresponding author.

CoMFA modeling

All models were generated using Tripos SYBYL version 6.8 on a Silicon Graphics Fuel workstation running the IRIX 6.5 operating system. Since the data for each assay was continuous, a PLS regression approach was implemented for the QSAR modeling. This method is a variation on principal components regression. PLS was considered superior because of its ability to handle multivariate regression with a greater number of independent variables than compounds. This method has proven accurate even when covariates in the model are correlated [12].

All structural models were assigned Gasteiger-Hückel charges and minimized with the Tripos Force Field version 5.2. Once minimized, all models were aligned to the template of paclitaxel minus hydrogen atoms and substituents R^1–R^3 (see Table 1). Proper alignment of structures is critical for obtaining valid CoMFA results. Moreover, since CoMFA assumes that each structure exhibits activity at the same active site of the protein, it is vital that all compounds be aligned in a pharmacologically active orientation [13].

A variety of possible descriptor fields were examined for the dataset. These included CoMFA and comparative molecular similarity indices analysis (CoMSIA), steric, electrostatic, and hydrogen bond donor and acceptor fields, along with dipole moment and ClogP descriptors. ClogP is a fragment-based estimation of a compound's lipophilicity and equates to the logarithm of the compound's octanol:water partition coefficient. This value was calculated by the Tripos SYBYL implementation of the BioByte ClogP program [12].

Comparative molecular field analysis measures the steric and electrostatic fields around each molecule and relates these measurements to each molecule's activity. Each measurement is taken at repeated intervals throughout a three-dimensional lattice framework. The values are derived from interactions of a probe atom with a specific size and charge at each point. The standard 2 Å lattice frame and sp^3 carbon probe atom with a $+1$ charge and distance-dependent dielectric constant were chosen.

The steric and electrostatic CoMSIA field is an aggregate field that, like CoMFA, includes steric and electrostatic components. Overall, CoMFA and CoMSIA both predict chemical activity well, which makes it prudent to consider both fields in a QSAR analysis.

The fields generated in CoMSIA are reported to be less sensitive to changes in the alignment of molecules in the database than CoMFA. By utilizing Gaussian type distance dependence within the molecular similarity indices, CoMSIA provides smoother contour maps than does CoMFA [12].

HQSAR modeling

Hologram QSAR is a relatively new two-dimensional computational technique that uses a fragmenting approach that relates substructural components of compounds to their biological activity. Each compound is split into a set of unique fragments, forming a molecular hologram, or fingerprint, of the compound. All possible fragments are encoded, including linear, overlapping and branched fragments. Additionally, the molecular hologram records the number of times each fragment is present within the chemical structure. Although this method is two-dimensional in nature, it also utilizes three-dimensional information such as chirality and molecular hybridization [13]. HQSAR has been shown to perform very well and provide comparable results to other QSAR techniques [12].

Fingerprints are generated first by converting the chemical structure to a binary bit string (a "key"), which uniquely identifies each fragment of four to seven atoms in length. Each fragment is then mapped to a pseudo-random integer. In this way, each fragment is assigned a unique random number. This number is used to place, or hash, the fragment's information into the appropriate bin for ease of retrieval later. Storage space is minimized because the chemical structure is a binary string stored in a table. The molecular hologram goes one step further and keeps track of the number of times the fragment appears in the structure definition.

Varying hologram lengths, which are the number of bins used to represent the structure, are utilized in HQSAR. Because hashing data repeatedly will inevitably produce collisions, or two sets of data hashed to the same bin, the hologram lengths are prime numbers. Using prime lengths increases the chances of resolving the collision and reassigning the collided fragment to another bin.

Once structural information is encoded into the molecular hologram, HQSAR runs a PLS analysis on the matrix of compounds and their corresponding holograms. The PLS model uses the potency value as its dependent variable and the molecular hologram as descriptor variables. PLS coefficients are generated for each bin's contribution to the model, and then each individual atom's contribution can be assessed by dividing the coefficient by the number of atoms in the fragment.

Docking studies using FlexiDock

The FlexiDock module of SYBYL was used to perform a docking study of the butitaxel analogues. FlexiDock utilizes a genetic algorithm to determine the optimal geometrical configuration of the ligand model in torsional space. This algorithm calculates energy scores produced from the Tripos Force Field for each docking solution [14]. FlexiDock defines the active site based upon user input. For this study, the active site was defined as all amino acid residues within 3 Å of the atomic coordinates of paclitaxel. Models of each compound were then rotated manually and aligned into the receptor. Electrostatic charges and hydrogens were added, and the program iteratively defined the 20 best-fit docking models based upon the information given [14]. All ligand models were docked successfully into the crystal structure of β-tubulin at the paclitaxel binding pocket described by Lowe et al. [11].

Results and discussion

Comparative molecular field analysis

All compounds in the dataset were aligned to a template structure, paclitaxel, using the SYBYL database align utility. Once the structures were aligned, descriptors were defined for the dataset. The CoMFA field was derived from the Tripos Standard electrostatic and steric fields with the standard value of 30 kcal mol^{-1} chosen as optimal. The best PLS CoMFA model was obtained by using region focusing to define the active region of the butitaxel analogues more narrowly. Region focusing applies weights to the lattice points in the CoMFA region in order to increase their contribution to the overall CoMFA model. Discriminant power was used to weight the lattice points, where each weight represents the contribution to the model components [12]. All compounds except for the 2-thienyl derivative fit the model. A reliable QSAR model could not be generated when the 2-thienyl analogue **17** was included in the analysis. Analogue **17** was the second most potent of all of the compounds and more closely aligned conformationally with paclitaxel than with butitaxel. The Gasteiger-Hückel charges of **17** were also more similar to paclitaxel than butitaxel, indicating that **17** would sterically and electrostatically behave more closely to paclitaxel. Since compound **17** was a biological outlier as described by Ali et al. [10] (i.e. analogue **17** was the most potent of all substituted thienyl compounds tested, with cytotoxicity twice that of paclitaxel) and did not conformationally align to butitaxel, these factors may underlie the failure of the CoMFA model derived here to describe its activity. Individual analysis of compound **17** revealed that the R^1 substituent was oriented away from the *tert*-butyl group at R^2, while these two substituents were in closer proximity for other potent analogues. Contour maps were analyzed to compare the steric and electrostatic contributions of **17** with the structurally similar 3-thienyl derivative **18**. The models of the two compounds had very similar electrostatic and steric regions (data not shown) and no significant differences were observed. Overall, however, the computational differences between **17** and the other analogues argued for its exclusion from the model.

Once analogue **17** was removed from the analysis, a CoMFA model with five components and an r^2 of 0.943 for the microtubule polymerization assay potency data was derived. The crossvalidated r^2 (i.e. q^2) was 0.376.

Table 2 PLS CoMFA, HQSAR and cross-validation results for butitaxel analogues

CoMFA	PLS components	r^2	q^2	s	F	P
	5	0.943	0.376	0.061	52.674	<0.001
HQSAR	Components	r^2	s	q^2	s	Hologram length
	6	0.923	0.074	0.446	0.197	53

The 2-thienyl analogue was removed from analyses for reasons cited in the text

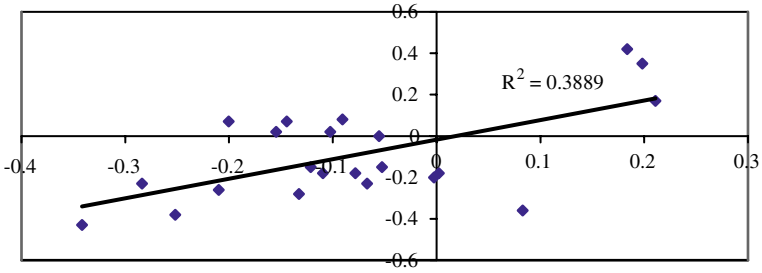


Fig. 1 Observed versus calculated values for butitaxel analogues derived from manual drop-one validation

This model proved to be the optimal model of those tested. Results of the model are shown in Table 2.

A manual drop-one validation was then performed for the final CoMFA model; i.e. each compound's activity was predicted from a model that was generated by excluding it. Linear regression analysis of the predicted versus actual activities gave an r^2 of 0.389. Although the model failed to predict activities of the compounds accurately, it did segregate the most from the least potent derivatives. This exercise also supported the results of the internal SYBYL cross-validation for this model. A plot of the observed versus calculated values for each compound is shown in Fig. 1.

Contour maps for the microtubule stabilizing data were also generated. Figure 2 shows the contour map of the CoMFA PLS model with the most potent analogue, the *tert*-butoxy derivative **4**, embedded in the map. Figure 3 shows the least potent analogue, 2-thienyl-2-ethenyl **21**, embedded. Areas of blue indicate positive electrostatic charge potential associated with increased activity, while red indicates negative charge and increased activity. Yellow regions indicate areas of steric hindrance to activity, while green regions indicate a steric contribution to potency. Analysis of each map showed that the potency level was influenced by both the electrostatic and steric activity of the substituent at R^1. Comparing models for the *tert*-butoxy **4** and 2-thienyl-2-ethenyl **21** analogues, it should be noted that the R^1 substituent of the more potent compound, analogue **4**, was positioned much closer to the sterically favored area and further away from the negatively charged area. This indicated that the steric contribution of the *tert*-butoxy group was favorable to increased activity in the microtubule assembly assay and that this substituent possessed minimal negative charge contribution to the overall activity. In contrast, Fig. 3 shows that the 2-thienyl-2-ethenyl R^1 substituent of analogue **21** pointed more towards the negatively charged and sterically unfavorable areas, indicating that this substituent sterically and electrostatically hindered activity. Additionally, the R^2 substituent of analogue **21** was slightly further away from the positively charged area, again indicating that this compound would be less potent. Overall, the analysis revealed that the model's contribution was 75% steric and 25% electrostatic.

A statistically significant CoMFA model could not be derived from the B16 melanoma cytotoxicity data, even when the approach outlined above for the microtubule assembly data, dropping analogue **17** and using a region-focused descriptor, was used. The q^2 values obtained in the exercises with the cytotoxicity data never exceeded 0.05.

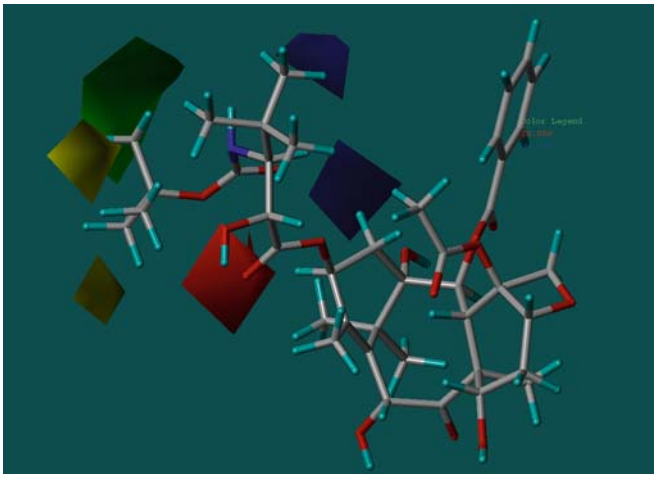


Fig. 2 CoMFA standard deviation × coefficient contour plot. *Green contours* indicate areas where bulky groups increase activity while *yellow contours* indicate areas where bulky groups decrease activity. *Red contours* indicate regions where negatively charged groups increase activity while *blue contours* indicate regions where negatively charged groups decrease activity. The most potent analogue, tert-butoxy **4**, is embedded to demonstrate its affinity for the sterically favored region

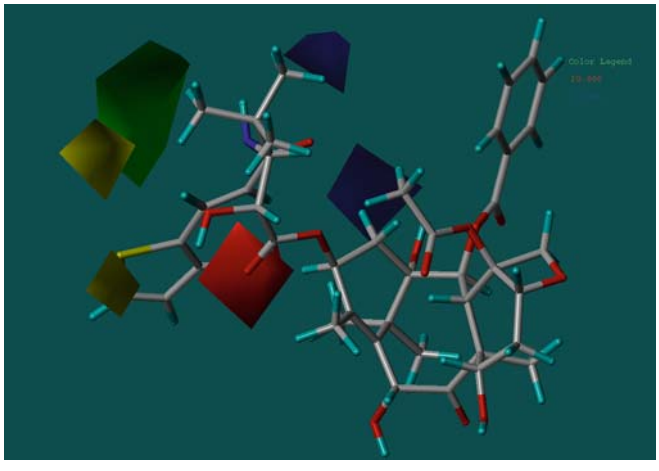


Fig. 3 CoMFA standard deviation × coefficient contour plot. The least potent analogue, 2-thienyl-2-ethenyl **21**, is embedded to demonstrate its affinity for the sterically unfavoured and negatively charged areas

HQSAR modeling

As with the CoMFA model, a HQSAR model could not be obtained when analogue **17** was included in the analysis. Once **17** was removed from consideration, however, a HQSAR with r^2 of 0.919 and q^2 of 0.471 with six components for the microtubule polymerization assay was generated. Several models were fit using different hologram lengths. A length of 97 bins was found to be optimal. These results were comparable to the results obtained in the CoMFA model. Numerical HQSAR results are shown in Table 2.

When the activity of the 2-thienyl analogue **17** was predicted using this HQSAR model, the potency was calculated to be −0.211 units [units = $-\log(EC_{50}/EC_{50(paclitaxel)})$]. The reported experimental value was 0.377 units. The largest difference between predicted and actual value for the other compounds was 0.251 units. Therefore, this discrepancy clearly indicated that the 2-thienyl analogue **17** was an outlier and again supported its removal from the learning set in order to derive a valid HQSAR model. Using the B16 melanoma cytotoxicity data, HQSAR models could be derived with q^2 values no greater than 0.10. Varying hologram lengths and removal of compound **17** from the HQSAR learning set did not improve the results.

Modeling of B16 melanoma cytotoxicity data

Comparative molecular field analysis and HQSAR models could not be generated from the cytotoxicity data. Linear regression analysis of the two datasets revealed no relationship between the biological endpoints measured in each assay ($r^2 = 0.275$). Therefore, only 27.5% of the variation in the cytotoxicity data could be explained by the microtubule interference in the cell. We speculate that the remaining variation may be caused by a multiplicity of mechanisms and other biological factors for which the CoMFA and HQSAR models developed here could not explain. This variation was most likely responsible for the failure of these modeling techniques to describe the cytotoxicity data.

FlexiDock analyses

A docking study of the butitaxel structures was completed using the SYBYL FlexiDock module. In the first series of docking runs, all ligands were minimized utilizing aggregates. For each structure, the R^1 substituent in the respective molecular model was minimized while the remainder of the structure was aggregated and held static. This produced 23 ligands conformationally similar to butitaxel for the purposes of docking. Twenty solutions were generated for each analogue. The FlexiDock runs generated very similarly-oriented solutions in the paclitaxel binding pocket. The conformation with the lowest total energy score was chosen as the optimally docked ligand.

Investigation of each model for docked ligands revealed that the models successfully positioned ligands into the paclitaxel binding pocket, but docking energies and potency did not correlate. Since the dataset of ligands was highly congeneric, it was concluded that discrimination between activity and docking energy would not be evident when computed with the entire structure of such similar ligands.

We therefore next considered only the variable portion, R^1, of each compound as the ligand. FlexiDock analyses were generated on this new set of truncated ligands. As before, the lowest energy solution was determined to be optimal. The results are shown in Table 1.

All 23 of the truncated molecular models were successfully docked into the paclitaxel binding pocket. Figure 4 shows docked solutions for each member of the dataset in the paclitaxel binding pocket on the Connelly surface of β-tubulin and demonstrates the relatively close alignment among the series of compounds.

A low docking energy score (i.e. ease of docking) modestly correlated with activity in the microtubule assembly assay. For all structures in the dataset, the correlation between potency and docking energy was 0.53. Although the 2-thienyl analogue **17** had to be removed from the CoMFA and HQSAR models, it did dock successfully in the paclitaxel binding site. When models for the parent compound butitaxel, as well as paclitaxel and docetaxel, were removed, the correlation increased to 0.61. The rationale for removal of these compounds from this analysis includes their structural differences from the other compounds at either or both the R^2 and R^3 positions.

Paclitaxel binds in a hydrophobic pocket near the M-loop and the site of lateral interactions of β-tubulin [2]. Three functional domains are evident in the proposed model of tubulin, namely the GTP-binding, motor protein/MAP-binding and drug binding domains. These domains, part of the very compact tubulin structure,

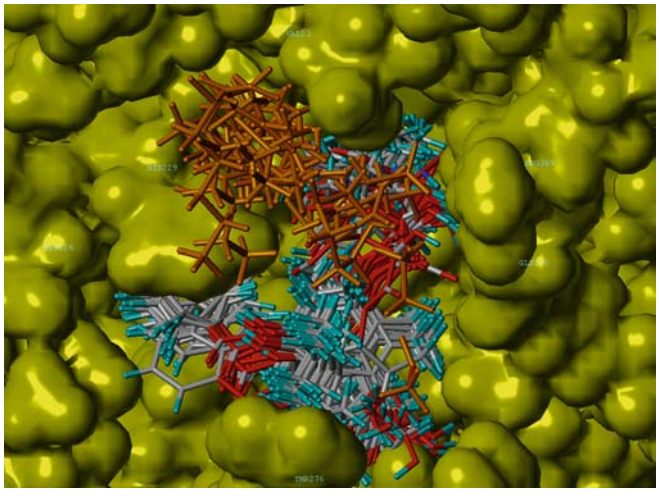


Fig. 4 Structures of the 21 butitaxel analogues, paclitaxel and docetaxel docked into a Connelly surface representation of the paclitaxel binding site of β-tubulin

interact with each other intimately. Therefore, the effects of nucleotides, drugs and additional proteins in the cell on tubulin are inherently linked [15].

Models of the paclitaxel binding site of β-tubulin suggest that the N′ benzoyl group can be substituted with marginal effect on the binding capacity [16]. The C2-benzoyl side chain, however, is necessary for binding [7]. Several residues of β-tubulin make direct contact with paclitaxel. In helix H1, Val23 has hydrophobic contact with the phenyl moieties of N′ (R^1 in Table 1) and C3′ (R^2), and Asp26 hydrogen bonds with the amide nitrogen on the side chain. The phenyl portion of the C2-benzoyl group has hydrophobic contact with Leu217 and Leu219 in the H6-H7 loop as well as His229 and Leu230 in the core helix. Ala233 and Ser236 make contact with the C3′ phenyl group. Hydrophobic interactions also occur with C3′ phenyl group and Phe272. The baccatin ring system binds to β-tubulin at Pro274, Leu275, Thr276, Ser277 and Arg278. The pocket also includes residues Pro360, Arg369, Gly370 and Leu371. Thr276 contacts the oxetane ring, a required component for binding and activity of taxane derivatives [16].

In the FlexiDock analyses, docetaxel and butitaxel docked solutions matched very well with the docked paclitaxel solution. Analysis of the less potent butitaxel derivatives revealed a difference in the orientation of the R^1 substituent. While still in the paclitaxel binding pocket, the less potent analogues were further away from the pocket defined by the residues Ala233, Ser236, Val23, His229 and Asp26. The 2-thienyl-2-ethenyl derivative **21** was oriented closer to residue Glu22 and further away from the center of the binding pocket. The distance from individual atoms of **21** and residue Val23 ranged from 4.4–6.6 Å. The hexyloxy analogue **8** docked towards the outer edge of the binding pocket. Its atoms were 6.8–9.2 Å away from Val23. In contrast, potent analogues such as 2-thienyl **17** and *tert*-butoxy **4** docked well inside the pocket, interacting with the residues mentioned previously and their atoms' distances from Val23 never exceeded 6.1 Å. Their docked solutions superimposed closely to those of pacitaxel, butitaxel and docetaxel.

Conclusions

Comparative molecular field analysis and HQSAR models for 20 butitaxel analogues were developed using microtubule assembly activity data previously published by Ali et al. [10]. There was agreement between the two models in terms of predictive potential and overall r^2, which was 0.943 for the CoMFA model and 0.923 for the HQSAR analysis. The 2-thienyl analogue **17** was a statistical outlier in both of the models, but was successfully docked into the paclitaxel binding pocket. There was a wide variation between the predicted and actual potency values for compound **17**. Because its inclusion would compromise the predictive potential of the models, it was excluded from CoMFA and HQSAR analyses. Individual analysis of 2-thienyl analogue **17** revealed that it was extremely potent, but was more conformationally similar to paclitaxel than to butitaxel. Further investigation revealed that the R^1 substituent in the molecular model of compound **17** oriented away from the *tert*-butyl group at R^2, while other potent compounds had R^1 and R^2 in closer proximity. CoMFA contour maps revealed no significant differences in the steric and electrostatic properties of compound **17** and other butitaxel derivatives that were included in the final model.

The coordinates of β-tubulin as determined by high resolution electron crystallography were used to successfully dock each of the analogues into the paclitaxel binding pocket. By defining the ligand as only the variable portion of the compound at R^1, we were able to align all structures well, indicating that the docking procedure employed was faithful to the known interactions of the baccatin system of paclitaxel with tubulin. Docking energy did, but only weakly, correlate with the experimentally-determined activity of the compound.

Analysis of model parameters and contour maps revealed that the factor that most influenced the activity of these compounds was the steric potential of the R^1 substituent. Analogues with this substituent positioned closer to the sterically favorable area and away from the negatively charged area were more potent and promising analogues for anticancer applications.

Acknowledgement This work was supported by a grant from the National Institutes of Health (CA78039).

References

1. Meurer-Grob P, Kasparian J, Wade RH (2001) Biochemistry 40:8000–8008
2. Nogales E (2000) Annu Rev Biochem 69:277–302
3. Nicolaou KC, Hepworth D, King NP, Finlay MR (1999) Pure Appl Chem 71:989–997
4. Desai A, Mitchison TJ (1997) Annu Rev Cell Dev Biol 13:83–117
5. Esteva FJ, Valero V, Pusztai L, Boehnke-Michaud L, Buzdar AU, Hortobagyi GN (2001) Oncologist 6:133–146
6. Barboni L, Lambertucci C, Appendino G, Vander Velde DG, Himes RH, Bombardelli E, Wang M, Snyder JP (2001) J Med Chem 44:1576–1587
7. Kingston DI (2000) J Nat Prod 63:726–734
8. Ojima I, Wang T, Miller ML, Lin S, Borella CP, Geng X, Pera P, Bernacki RJ (1999) Bioorg Med Chem Lett 9:3423–3428
9. Ojima I, Slater JC, Michaud E, Kuduk SD, Bounaud PY, Vrignaud P, Bissery MC, Veith J, Pera P, Bernacki RJ (1996) J Med Chem 39:3889–3896
10. Ali SM, Hoemann MZ, Aube J, Georg GI, Mitscher LA, Jayasinghe LR (1997) J Med Chem 40:236–241
11. Lowe J, Li H, Downing KH, Nogales E (2001) J Mol Biol 313:1045–1057
12. Tripos (2001) SYBYL ligand based design. Tripos, St. Louis
13. Avery MA, Alvim-Gaston M, Rodrigues CR, Barreiro EJ, Cohen FE, Sabnis YA, Woolfrey JR (2002) J Med Chem 45:292–303
14. Tripos (2000) SYBYL receptor based design. Tripos, St. Louis
15. Nogales E, Wolf SG, Downing KH (1998) Nature 391:199–203
16. Lowe J, Li H, Downing KH, Nogales E (2001) J Mol Biol 313:1045–1057