ELSEVIER

Journal of Molecular Graphics and Modelling 21 (2003) 263–272

Journal of Molecular Graphics and Modelling

www.elsevier.com/locate/JMGM

Investigation of structural requirements of anticancer activity at the paclitaxel/tubulin binding site using CoMFA and CoMSIA

Mohd N. Islam, Yuqing Song, Magdy N. Iskander*

Department of Medicinal Chemistry, Victorian College of Pharmacy, Monash University, 381 Royal Parade, Parkville 3052, Vic., Australia

Received 14 April 2002; accepted 15 August 2002

Abstract

CoMFA and CoMSIA analysis were utilized in this investigation to define the important interacting regions in paclitaxel/tubulin binding site and to develop selective paclitaxel-like active compounds. The starting geometry of paclitaxel analogs was taken from the crystal structure of docetaxel. A total of 28 derivatives of paclitaxel were divided into two groups—a training set comprising of 19 compounds and a test set comprising of nine compounds. They were constructed and geometrically optimized using SYBYL v6.6. CoMFA studies provided a good predictability ($q^2 = 0.699$, $r^2 = 0.991$, PC = 6, S.E.E. = 0.343 and $F = 185.910$). They showed the steric and electrostatic properties as the major interacting forces whilst the lipophilic property contribution was a minor factor for recognition forces of the binding site. These results were in agreement with the experimental data of the binding activities of these compounds. Five fields in CoMSIA analysis (steric, electrostatic, hydrophobic, hydrogen-bond acceptor and donor properties) were considered contributors in the ligand–receptor interactions. The results obtained from the CoMSIA studies were: $q^2 = 0.535$, $r^2 = 0.983$, PC = 5, S.E.E. = 0.452 and $F = 127.884$. The data obtained from both CoMFA and CoMSIA studies were interpreted with respect to the paclitaxel/tubulin binding site. This intuitively suggested where the most significant anchoring points for binding affinity are located. This information could be used for the development of new compounds having paclitaxel-like activity with new chemical entities to overcome the existing pharmaceutical barriers and the economical problem associated with the synthesis of the paclitaxel analogs. These will boost the wide use of this useful class of compounds, i.e. in brain tumors as the most of the present active compounds have poor blood–brain barrier crossing ratios and also, various tubulin isotypes has shown resistance to taxanes and other antimitotic agents.

Keywords: Anticancer agents; Cancer chemotherapy; CoMFA; CoMSIA; Microtubules; Paclitaxel; Paclitaxel analogs; Paclitaxel binding site; Pharmacophore; Tubulin ligands

1. Introduction

Microtubules (MT), the dynamic pipe-like protein composed of α-, β-heterodimers, are integral components of the mitotic spindle which are involved in cell division. Their functions and biophysical properties have made α- and β-tubulin the subject of intense study. Interfering with normal MT dynamics, for example, by the addition of tubulin ligands, can cause cell distress and affect MT stability and function, including mitosis, cell motion and intracellular organelle transport. Tubulin is considered an important pharmacological target in cancer chemotherapy as its polymerization and depolymerization are essential for cell division [1]. The unique dynamic characteristics of MTs and the relationship between MTs and cellular mitosis have created great interest in designing microtubule-targeted drugs for anticancer therapy, especially after the successful introduction of taxanes into clinical oncology and the widespread use of the vinca alkaloids, vincristine and vinblastine. These compounds inhibit cell mitosis by binding to the protein tubulin in the mitotic spindle and preventing polymerization into the MTs. This mode of action is also shared with other natural agents, e.g. colchicine and podophylotoxin. Recently, the taxane class of antimicrotubule anticancer agents has become the most important member of the chemotherapeutic agents in use. This class of compounds binds specifically to a domain distinct from those of colchicine and vinca alkaloids and inhibit depolymerization of polymerized tubulin [2–4]. The first cytotoxic marketed taxoid compound was paclitaxel (1) Fig. 1. It was isolated from the bark of the Pacific yew tree *Taxus brevifolia* Nutt. (Taxaceae) [5] and recently from the hazelnut tree and fungi living on these trees [6]. This drug has been totally synthesized [7,8] and at present is extensively used in

* Corresponding author. Tel.: +61-3-9903-9545; fax: + 61-3-9903-9582.
E-mail address: magdy.iskander@vcp.monash.edu.au (M.N. Iskander).

1093-3263/02/$ – see front matter © 2002 Elsevier Science Inc. All rights reserved.
PII: S1093-3263(02)00162-6

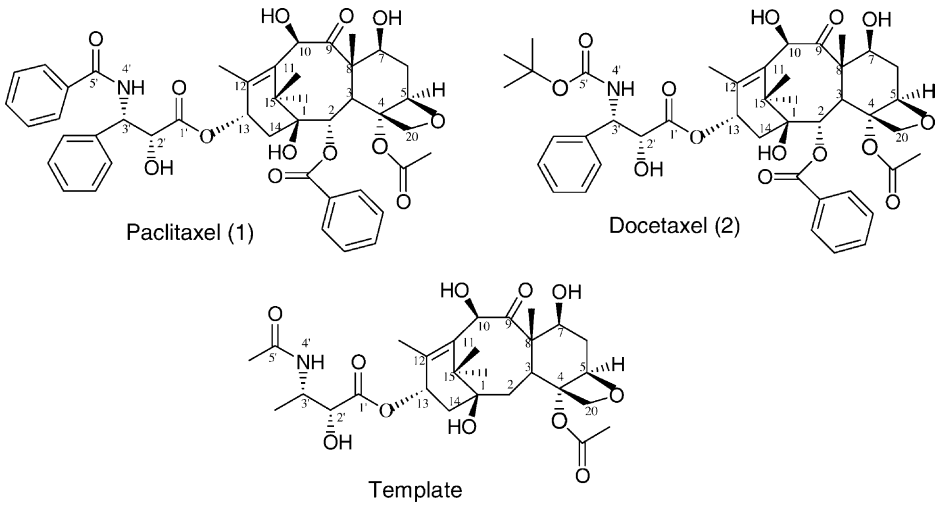


Fig. 1. Structures of paclitaxel, docetaxel and a template.

the treatment of breast and ovarian cancers. It is also used against skin, lung, and brain and neck carcinomas [9]. However, there are major problems associated with the clinical use of these naturally occurring compounds: (a) there are only very small amounts of these compounds present in the plants and (b) various tubulin isotypes has shown resistance to taxanes and other antimitotic agents. Therefore, there is an urgent need to design and develop new antimitotic agents.

Structural modifications of paclitaxel (1) (Fig. 1), at C10 and C2 have resulted in second generation taxoids [10]. Structural variations along the upper part of the paclitaxel molecule comprising of C6–C12, appears to have less impact on the bioactivity profile of the compound. The lower part, comprising of C14 and C1–C5 appears to be a region which is crucial to the activity of paclitaxel, as structural changes have lead to changes in the activity. Opening of the oxetane ring was associated with a decrease in activity. It was found that it serves as a hydrogen-bond acceptor and a rigid lock on the taxoid skeleton [11,12].

Extensive Structure Activity Relationship studies demonstrated that the A-ring side chain at C13 with a C2′–OH and C2–benzoyl group are essential for both the cytotoxicity and stabilization of MTs exhibited by paclitaxel [13–16] taxoid skeleton [11]. Although the C4 acetyl group has no significant role, it may be responsible for the defined conformation of paclitaxel. The C1–OH group makes a significant contribution to the overall bioactivity [11]and the C7–OH is not essential for biological activity [17]. Recently, it has been reported that bioactive taxanes could be obtained by modifying C5 side chain of the core taxane structure [18].

To date, more than 350 taxoids have been isolated, developed and characterized [19,20] but very few are as potent as paclitaxel. Poor water solubility and very low yield of paclitaxel led researchers to design and develop new analogs of taxol that are more cytotoxic. Docetaxel (2) (Fig. 1), with minimal structural modifications at the C13 side chain and C10 substitution showed more water solubility and was more potent than paclitaxel [21] whereas baccatin, a new taxoid without C13 side chain, has no significant cytotoxicity. So, C13 side chain could be a target for developing new analogues. Synthesis of second generation taxoids by appropriate modification at C3′, C2 and acylation at C10 produced taxoids with extremely high potency against drug resistant cancer cells [22–24]. Design of new taxanes or taxane like antimitotic agents has been directed towards solving the limitations of both solubility and resistance in these compounds.

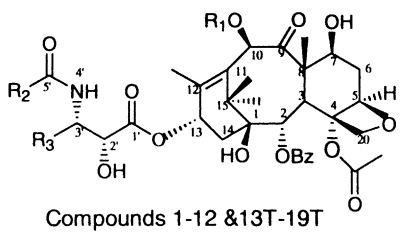
The aim of this investigation is to determine the most essential structural properties for antimitotic activity of selected anticancer agents based on the paclitaxel/tubulin binding site. The obtained information will then be utilized in our research program for the design paclitaxel-like compounds. The new compounds are expected to be as active and selective but with a simpler chemical structure.

2. Materials and methods

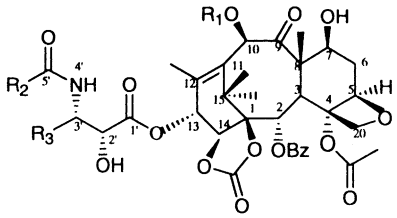
2.1. Biological data

All biological data on paclitaxel and its analogs were obtained from the one source [25]and were estimated under the same experimental conditions on MCF-7 (breast) cell line. The potency was defined as relative activity, which is calculated as [IC_{50}(paclitaxel)/IC_{50}(analogue)]; IC_{50} values expressed in nM concentration. The values of relative activity were used as the dependent variable in the structural analysis. Compounds having inhibitory concentration range from 0.13 nM up to 9.30 nM were subjected in this study of which compounds 1 and 2 (Fig. 1) are already used commercially as anticancer agents. In the present study, the structures of 19 compounds (compounds 1–12 and 22–28, Table 1) served as a training set and nine additional compounds (compounds

Table 1
Paclitaxel, its analogs and their biological activities investigated for CoMFA analysis



Compounds 1-12 &13T-19T



Compounds 20T-21T& 22-28

Compounds	R_1	R_2	R_3	BIOAC (IC_{50} nM)	RELAC	LOGP
1	Ac	Phenyl	Phenyl	1.70	1.00	6.93
2	H	*t*-BuO	Phenyl	1.00	1.70	6.55
3	Ac	*t*-BuO	4-F-C_6H_4	0.42	4.05	7.24
4	CH_3 CH_2CO	*t*-BuO	$CH_2CH(CH_3)_2$	0.35	4.86	7.57
5	H	*t*-BuO	$CH_2CH(CH_3)_2$	4.00	2.35	6.40
6	H	*t*-BuO	$CH{=}C(CH_3)_2$	0.55	3.09	6.74
7	CH_3 CH_2CO	*t*BuO	$CH{=}C(CH_3)_2$	0.18	9.44	7.92
8	Cyclopropane-CO	*t*-BuO	$CH{=}C(CH_3)_2$	0.20	8.50	7.75
9	Ac	*t*-BuO	$CH_2CH_2CF_3$	9.30	0.18	6.03
10	CH_3COO–	*t*-BuO	$CH_2CH(CH_3)_2$	0.33	5.15	7.01
11	CH_3	*t*-BuO	$CH_2CH(CH_3)_2$	0.28	6.07	6.90
12	Ac	phenyl	4-F-C_6H_4	5.10	0.33	6.98
13T	Ac	*t*-BuO	$CH{=}C(CH_3)_2$	0.60	2.83	7.25
14T	$(CH_3)_2$ NCO	*t*-BuO	$CH{=}C(CH_3)_2$	0.13	13.17	7.28
15T	CH_3 CH_2NHCO	*t*-BuO	$CH_2CH(CH_3)_2$	0.31	5.48	7.28
16T	CH_3 NHCO	*t*-BuO	$CH{=}C(CH_3)_2$	0.40	4.25	7.09
17T	$(CH_3)_2$ CHNHCO	*t*-BuO	$CH{=}C(CH_3)_2$	0.46	3.69	7.97
18T	$(CH_3CH_2)_2$ NCO	*t*-BuO	$CH{=}C(CH_3)_2$	0.20	8.50	7.95
19T	CH_3OCO	*t*-BuO	$CH{=}C(CH_3)_2$	0.14	12.14	7.35
20T	Cyclopropane-CO	*t*-BuO	$CH_2CH(CH_3)_2$	2.70	0.63	8.76
21T	$(CH_3)_2$ NCO	*t*-BuO	$CH{=}C(CH_3)_2$	1.70	1.00	8.25
22	H	*t*-BuO	2-Furyl	0.50	3.40	6.96
23	H	*t*-BuO	$CH{=}C(CH_3)_2$	0.50	3.40	8.00
24	Ac	*t*-BuO	$CH_2CH(CH_3)_2$	1.10	1.55	8.39
25	H	*t*-BuO	$CH_2CH(CH_3)_2$	3.00	0.56	7.65
26	Ac	*t*-BuO	$CH{=}C(CH_3)_2$	3.30	0.51	8.74
27	H	*t*-BuO	$CH_2CH_2CH_3$	4.20	0.40	7.30
28	Ac	*t*-BuO	$CH_2CH_2CH_3$	4.60	0.36	8.05

BIOAC: biological activity, RELAC: relative activity and T: compounds for test set.

13T–21T, Table 1) were used as test set to evaluate the predictive ability of the model obtained in this experiment. The compounds in the test set have been randomly selected from [25] and not been included in the training set during study.

3. Computational methods

3.1. Molecular modeling

Molecular modeling studies were performed using the SYBYL v6.6 Software package, Silicon Graphic workstation. Molecular structures were built using the SKETCH option in SYBYL [26] starting from the crystal structure of docetaxel (1TUB.pdb, E. Nogales, S.G. Wolf, K.H. Downing, 1997, Tubulin αβ dimer, electron defraction). Geometry optimization was carried out using MAXIMIN molecular mechanics and Tripos force field supplied within SYBYL, with convergence criterion set at 0.05 kcal/(Å mol). No significant conformational changes were observed during optimization process. An underlying assumption in the structural analysis is that all molecules in the data set bind to the same receptor in a similar way. For that reason, the optimized geometries were superimposed by atom based alignment to a template of crystal structure. All values were filled with valence and Gustiger–Huckel charges were calculated for each compound. Random conformational search was also carried out to recognize a reasonably low energy conformation of each compound.

3.2. CoMFA analysis

Comparative molecular field analysis [27] of these molecules was carried out on the steric and electrostatic fields using the default values. A three dimensional cubic lattice, with a 2 Å grid spacing, was generated automatically

around these molecules in order to ensure that the grid extended the molecular dimensions by 4 Å in all directions. Threshold column filtering of 2.0 kcal/mol was set to hasten the analysis and reduce the amount of noise. The steric and electrostatic fields were calculated separately for each molecule using sp^3 carbon atom probe with a charge of 1 (default probe atom in SYBYL) and energy cut off values of 30 kcal/mol were selected for both steric and electrostatic field. The probe atom was placed at each lattice point and their steric and electrostatic interactions with each atom in the molecule were computed using CoMFA standard scaling. log *P* values were calculated by ACD/CHEM [28] sketch program.

3.3. PLS analysis

Initial PLS analysis were carried out using Leave-One-Out option (cross-validated) to obtain the optimal number of components to be used in the subsequent final analysis. A subset of CoMFA field sample points falling with a standard deviation of $\leq$2.0 kcal/mol was used to run PLS regression analysis. Finally, non-cross-validated analysis was performed using the optimal number of previously identified components and was employed to analyze the result of CoMFA.

3.4. Predictive ability

The overall predictive ability of this analysis was evaluated by the term q^2 which was calculated according to the following equation:

$$q^2 = \frac{(\mathrm{SSY} - \mathrm{PRESS})}{\mathrm{SSY}},$$

where SSY represents the variance of the biological activities of molecules around the mean value and PRESS is the prediction error sum of squares derived from the leave-one-out method. The uncertainty of prediction is defined as

$$\mathrm{SPRESS} = \left[\frac{\mathrm{PRESS}}{(n-k-1)}\right]^{1/2},$$

where k is the number of variables in the model and n the number of compounds used in the study.

3.5. CoMSIA analysis

Comparative molecular similarity index analysis (CoMSIA) was performed to evaluate hydrophobic, H bond donor and acceptor property of a molecule by employing the standard options of SYBYL [29]. The steric, electrostatic, hydrophobic, H-bond acceptor and H-bond donor fields were calculated separately using sp^3 carbon atom probe with a charge of +1 provided in SYBYL v6.6. Threshold column filtering of 1.0 kcal/mol was used during this analysis. Similar to CoMFA, a data table has been constructed from similarity indices [30] calculated at the intersections of a regularly spaced lattice (2 Å spacing). Similarity indices $A_{\mathrm{F,K}}$ between the compounds of interest and a probe atom have been calculated according to the following equation:

$$A^q_{\mathrm{F,K}}(j) = \sum_{i=1}^{n} w_{\mathrm{probe},k}\, w_{ik} \mathrm{e}^{-\alpha r^2_{iq}},$$

where q is the grid point for molecule j, with i the summation index over all atoms of the molecule j; w_{ik} the actual value of physicochemical property k of atom i; $w_{\mathrm{probe},k}$ indicates probe atom with charge +1, radius 1 Å, hydrophobicity +1, H-bond donor and acceptor property +1; α is the attenuation factor; r_{iq} is the mutual distance between probe atom at grid point q and atom i of the test molecule. The default value of α is 0.3. Larger values result in a steeper Gaussian function, and an increasing attenuation of the distance-dependent effects of molecular similarity.

4. Results and discussions

4.1. CoMFA analysis

The statistical parameters of CoMFA analysis of 19 compounds are summarized in Table 2. CoMFA analysis using steric, electrostatic fields and log *p*, gave cross-validated q^2 value of 0.566 with an optimized components of 4, the value of the cross-validated co-relation co-efficient is $r^2 = 0.920$ (model 1a), which is relatively acceptable. Using steric, electrostatic and log *p* contribution in different combination, e.g. steric–electrostatic, steric–log *p*, and electrostatic–log *p*, produced lesser predictive powers (Table 2, models 1b–d). After omitting the most outlier compound 6 (although biologically active, the residual is 1.63), model 2a gave a good predictive ability, q^2 value of 0.657 and $r^2 = 0.988$ with six components and models 2–d are also noticeable. Further omitting the second outlier compound 27 (biologically less active compound, the residual is −1.28), the obtained model 3a gave significantly good predictive ability, q^2 value of 0.699 and $r^2 = 0.991$ with six-component contributed by all fields. Other models like 3b–d (Table 2) are also significant. Analyzing different models obtained it is found that the models containing all contributing fields result in better predictive powers than those of models containing only two fields in different combinations. It can be concluded that all fields, e.g. steric, electrostatic and lipophilic contribution play the key role for biological activity of these compounds although steric and electrostatic contributions are prominent. The reasons for different results might be that these compounds adopt a different binding conformation at the receptor-binding site. These predictive results have been obtained, indicating a good statistical co-relation of the fields. The results also revealed that the resultant CoMFA model had a fair predictive ability (Table 3) and supports biological data of these compounds. The quality

Table 2
Results of CoMFA analysis

Model	Results of PLS analysis					Field		log *p*
	q^2	PC	r^2	S.E.E.	*F*	Steric	Electrostatic	
1a	0.566	4	0.920	0.875	42.124	0.530	0.345	0.125
1b	0.473	3	0.878	1.067	35.970	0.710	0.290	–
1c	0.527	2	0.790	1.356	30.028	0.896	–	0.104
1d	0.520	5	0.904	1.061	24.522	–	0.892	0.108
2a	0.657	6	0.988	0.393	157.121	0.629	0.253	0.118
2b	0.568	6	0.990	0.363	175.551	0.614	0.386	–
2c	0.644	6	0.991	0.343	196.483	0.899	–	0.101
2d	0.548	5	0.914	0.999	25.650	–	0.889	0.111
3a	0.699	6	0.991	0.343	185.910	0.520	0.362	0.119
3b	0.577	6	0.993	0.311	227.760	0.598	0.402	–
3c	0.665	6	0.989	0.375	155.602	0.890	–	0.110
3d	0.566	5	0.908	1.051	21.824	–	0.893	0.107

Model 1: all 19 compounds in the training set; model 2: 18 compounds in the training set (omitting the most outlier compound 6); model 3: 17 compounds in the training set (omitting the most outlier compounds 6 and 27). *Note*: q^2, the leave-one-out (LOO) cross-validated r^2 value, *n*, number of optimal components; r^2, no-cross-validated regression co-efficient; S.E.E., standard error of estimate; *F*, *F*-statistic for analysis.

Table 3
Actual and predicted biological activities of test set compounds

Compound number	Biological activity		Residual
	Actual	Predicted	
13T	2.83	2.80	0.03
14T	1.00	0.98	0.02
15T	5.48	5.50	−0.02
16T	4.25	4.29	−0.04
17T	3.69	3.67	0.02
18T	8.50	8.50	0.00
19T	12.14	12.11	0.03
20T	0.63	0.66	−0.06
21T	1.00	0.98	0.02

of the CoMFA models is represented in the Fig. 2, which shows plots of observed biological activity versus activity predicted from the best CoMFA model (model 3a), using steric, electrostatic and log *p*. This model showed the highest cross-validated q^2 value, best SEP and used a modest number of PLS components. This model suggested that contribution from steric fields (52%) is higher than that of electrostatic (36.2%) and lipophilic (11.9%) fields.

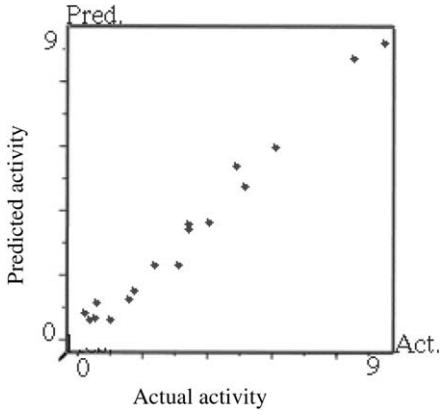
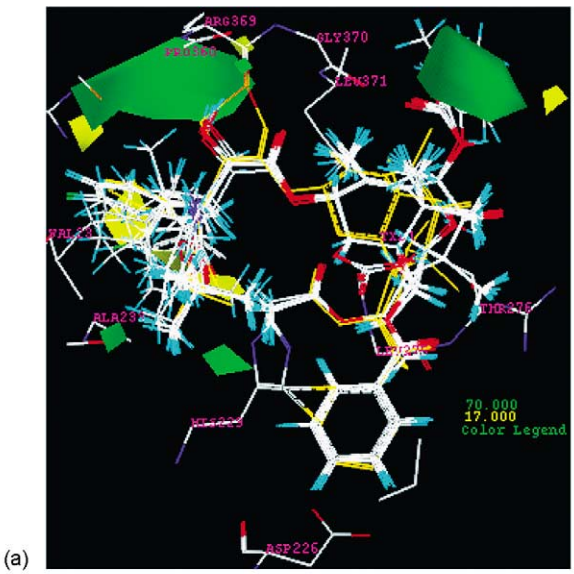
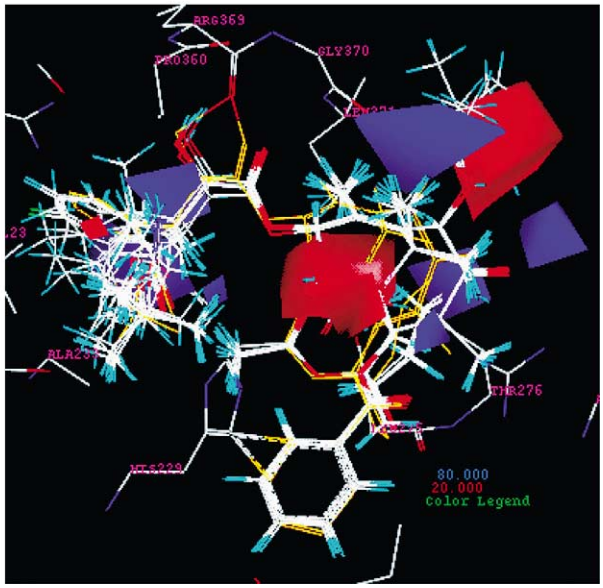


Fig. 2. Actual versus predicted activity for best CoMFA model.

The CoMFA steric and electrostatic fields for the analysis are presented as contour maps inFig. 3a and b where the potent biological activity is displayed in the maps. In general, the color polyhedra surrounded lattice points where the QSAR strongly associated changes in compound field values with changes in biological potency. Green polyhedra surrounded regions where more bulk is favorable for increasing potency, while yellow polyhedra surrounded regions where less bulk is good. Red and blue contours show regions of desirable negative and positive electrostatic interactions, respectively. Considering the ligand–receptor interaction, amino acid residues surrounding the ligand docetaxel crystal structure in the binding region were merged into both figures. A large region of green contour (Fig. 3a), around the C2′ and C3′ suggest that more potent analogs may be obtained by introducing bulky substituents to this region. Presence of this contour in the vicinity of C3′ phenyl and 4-F-phenyl groups of compounds 1, 2 and 3, respectively, close to Arg320, Asp26, Val 23 and Ser 236 residue in the binding pocket, is notice worthy. Incorporation of bulkier groups will enhance antimitotic activity. Compounds 1(paclitaxel) and 2 (docetaxel) are used clinically as anticancer drugs. The C3′-2-methyl prop-1-enyl substituent of biologically active compounds 7, 8 and 24 and C3′-2-methyl propyl side chain of compound 25 have been overlapped with this contour, which is situated in a large pocket surrounded by residue Pro360, that may interact with the bulky groups. Other compounds having the same bulky groups are near about this contour map indicating addition of bulky groups at C3′ position may enhance the activity. Another region of green polyhedra near C10 position indicated that an additional bulky atom in this site would increase activity but the dimension of this substituent



(a)



(b)

Fig. 3. (a) Contour maps of CoMFA steric std∗coeff map are shown in green (80% contribution) refer to sterically favored regions; yellow (20% contribution) indicate disfavored areas. Superimposed are the aligned paclitaxel analogs with docetaxel crystal structure bound to key residues in the active site of tubulin receptor. Areas contoured by green polyhedra correspond to regions where steric bulk is favorable for affinity and areas contoured by yellow polyhedra indicate disfavored for steric bulk. (b) Contour maps of CoMFA electrostatic field. Electrostatic std∗coeff contour map are shown in red (80% contribution) indicate regions where negatively charged substituents are favored and blue contours (20% contribution) refer to regions where negatively charged substituents are disfavored. Greater values of bio-activity measurement are co-related with: more positive charge near blue and more negative charge near red.

should not be too large. Compounds 3, 4, 7 and 11 containing acetyl, propanoyl, and methyl substituents, respectively, at C10 position, lying very close to this contour, agrees with their reported biological activity, which supports the result of this experiment. This is also in agreement with the CoMFA result that the methyl or propyl substitution at this position produced better activity than that of hydrogen atom. Two small green contours embedded in a large pocket of tubulin, near methyl groups of *t*-BuO and acetate side chain at C4 position, also indicated the possibility of enhancing activity by adding bulky groups. These results thus reveal the importance of the steric field of molecules contributing biological activity through contour maps. Sterically unfavorable region (yellow) lying around C3′ position surrounded by Val23 and Ala233 reflects the possibly detrimental effect of introducing further bulky groups in this region.

The electrostatic contour map (Fig. 3b) shows a region of red polyhedra around C10, C1 and C14 position indicated that electron rich groups are beneficial to the activity. Most compounds have electron rich groups at C10 position, which implies that addition of bulky groups containing electronegative atoms at C10 would enhance activity. Biologically active compound 10, with an electronegative group $CH_3COO–$ at C10 position situated completely inside the red polyhedra supported our CoMFA result. It may be summarized that not too bulky group and relatively high negative charges at the C10 position will enhance activity. C1 and C14 position of all compounds covered by a large red contour indicated that the addition of electronegative atoms would enhance activity. Compounds 22–28, known as 1, 14-carbonate taxoids which have more electronegative carbonate groups connected between C1 and C14 are situated inside the contour map. Although compounds 25–28 are biologically less active than compounds 22–24(which are highly active, Table 1), this contour supports the experimental data. Although there is no specific information regarding to the biological activity of C9 and C7 position, which is very close to Thr276 (distance 2.40 Å), blue contours around these positions indicates that incorporation of electropositive side chain will enhance binding to receptor. This finding could lead to a new generation of paclitaxel analogs by introducing more electropositive atoms in these positions. Another blue contours near C16 methyl and ethyl or propyl group of bulky side chain at C10, two blue contours, very close to C2′, C3′ and around the side chain of C3′ indicated more affinity could be obtained by introducing more positive charge. Considering the facts aforementioned, it can be confirmed that both the steric and electrostatic interactions are important for inhibitory activity of these compounds.

4.2. CoMSIA analysis

To estimate the hydrophobic contributions of these antimitotic agents, hydrophobic similarity index fields were constructed by using CoMSIA, which cannot be calculated by Lennard–Jones and Coulombic fields encoded in CoMFA. The CoMSIA analysis was performed employing the standard options of SYBYL v6.6. Applying the model 3a of CoMFA (Table 2), CoMSIA study gave good predictive ability, $q^2 = 0.535$ and $r^2 = 0.983$, PC = 5, S.E.E. = 0.458 and $F = 124.410$ (model C, Table 4) that supports the results

Table 4
Results of CoMSIA analysis

Model	A	B	C
q^2	0.414	0.475	0.535
r^2	0.811	0.972	0.983
PC	3	5	5
S.E.E.	1.246	0.570	0.458
F	25.693	83.650	124.410
Contribution			
Steric	0.056	0.133	0.130
Electrostatic	0.429	0.286	0.286
Hydrophobic	0.171	0.176	0.170
H-acceptor	0.135	0.169	0.166
H-donor	0.210	0.236	0.247

Model A: all 19 compounds in the training set; model B: 18 compounds in the training set (omitting compound 6); model C: 17 compounds in the training set (omitting compounds 6 and 27).

obtained from CoMFA. A summary of the analysis obtained using the best CoMFA and CoMSIA models are shown in (Table 5). Surprisingly, the steric contribution of CoMSIA was drastically less than that obtained by CoMFA study and models (data not shown) without H-bond acceptor or donor fields, resulted in less predictive power. H-bond acceptor and donor properties of receptor may also be a factor for activity of the compounds investigated. However, they produced very good hydrophobic contribution (17%). Contour maps of hydrophobic, H-bond donor and acceptor regions have been presented in Fig. 4a–c, respectively. Contour maps of CoMSIA hydrophobic field (Fig. 4a) contribution reflect that regions of yellow polyhedra are good for increasing the potency by introducing hydrophobic substituents, while hydrophilic substituents are beneficial to the activity at the region of white areas. A large yellow region overlapped by Pro 360 residue, enveloped the majority of the saturated and unsaturated aliphatic groups at C3′ position. This highlights the fact that the more hydrophobic derivatives exhibit enhanced binding affinities for this receptor. This region is situated in a large pocket surrounded by Arg 320, Asp 26, Ser 236, Val 23 and Ala 233 where more hydrophobic groups may be accumulated. Two white contours around the methyl groups of C3′ side chain of all compounds, aromatic ring of compounds 1, 2 and *t*-BuO groups of all compounds suggested that addition of hydrophilic groups predict enhanced activity.

Table 5
Summary of the analysis results obtained using the best CoMFA and CoMSIA models

Method	CoMFA	CoMSIA
q^2	0.699	0.535
r^2	0.991	0.983
PC	6	5
S.E.E.	0.343	0.458
F	185.910	124.410
Contribution		
Steric	0.52	0.130
Electrostatic	0.362	0.286
log *p*	0.119	–
Hydrophobic	–	0.170
H-acceptor	–	0.166
H-donor	–	0.247

Note: q^2, the leave-one-out (LOO) cross-validated r^2 value; PC, number of optimal components; r^2, no-cross-validated regression co-efficient; S.E.E., standard error of estimate; *F*, *F*-statistic for analysis.

In principle, contour maps hydrogen-bond donor and acceptor fields Fig. 4a and b) highlighted areas where ligands and putative hydrogen partners in the receptor site could form hydrogen bonds that significantly influence binding affinity. Basically, the acceptor field contains information about where hydrogen-bond donating groups should be on the receptor whereas the donating field depicts where hydrogen-bond acceptor groups should be located on the receptor. CoMSIA H-bond acceptor field (Fig. 4c) shows the regions where H-bond donors on the receptor are predicted to enhance (magenta 80% favored) and disfavor (red 20%) binding of ligands with the receptor. The magenta contours enclose the area where donor hydrogen on the receptor are expected to enhance activity. This area, close to Gly 370 and Leu 371, enveloped all carbonyl oxygen and all methyl groups attached to these carbonyl groups, with the exception that the C10 position of compounds 14, 26 and 28 indicated hydrogen-bond donors on receptor are predicted to enhance activity. The cyclopropane ring of compound 8 is outside this contour. One large red contour surrounding C1 OH, the carbonyl oxygen of C2 benzoyl group of all compounds and carbonyl oxygen of compounds 22–28, indicated that hydrogen-bond donors on receptors are disfavored for binding. Although in the binding pocket this contour did not appear to overlap any residues that are hydrogen-bond donors, there is evidently a correlation between the affinities of the compounds in the training set and a hydrogen-bond acceptor in this region. Two small red contours near *t*-BuO methyl groups and aromatic ring and aliphatic groups of compounds 1, 25, 8, and 7, respectively, indicated that hydrogen-bond donors on receptors are detrimental for binding. Contour maps of CoMSIA H-bond donor field (Fig. 4b) shows five areas of cyan (80% favored) indicating that acceptors on the receptor is predicted to favor binding and one purple (20% disfavored) area near C3′ phenyl ring, where an acceptor on the receptor is predicted to disfavor binding. One large cyan region near C1′ and C2′ position, very close to Arg 369, Gly 370 and Pro 360 indicated that acceptors on these residues might enhance the activity. Other cyan regions near methyl group of C17, around acetyl group of C10 (close to Glu 281) indicated the same. One purple contour embedded in a pocket covering by Asp 26 and Pro 360, overlapped by C3′ aromatic ring and other aliphatic groups of most of the compounds, indicated unfavorable affinity. The contour that disfavors acceptors on the receptor site is attributable to compounds with low affinity that have hydrogen-bond donor groups in the vicinity.

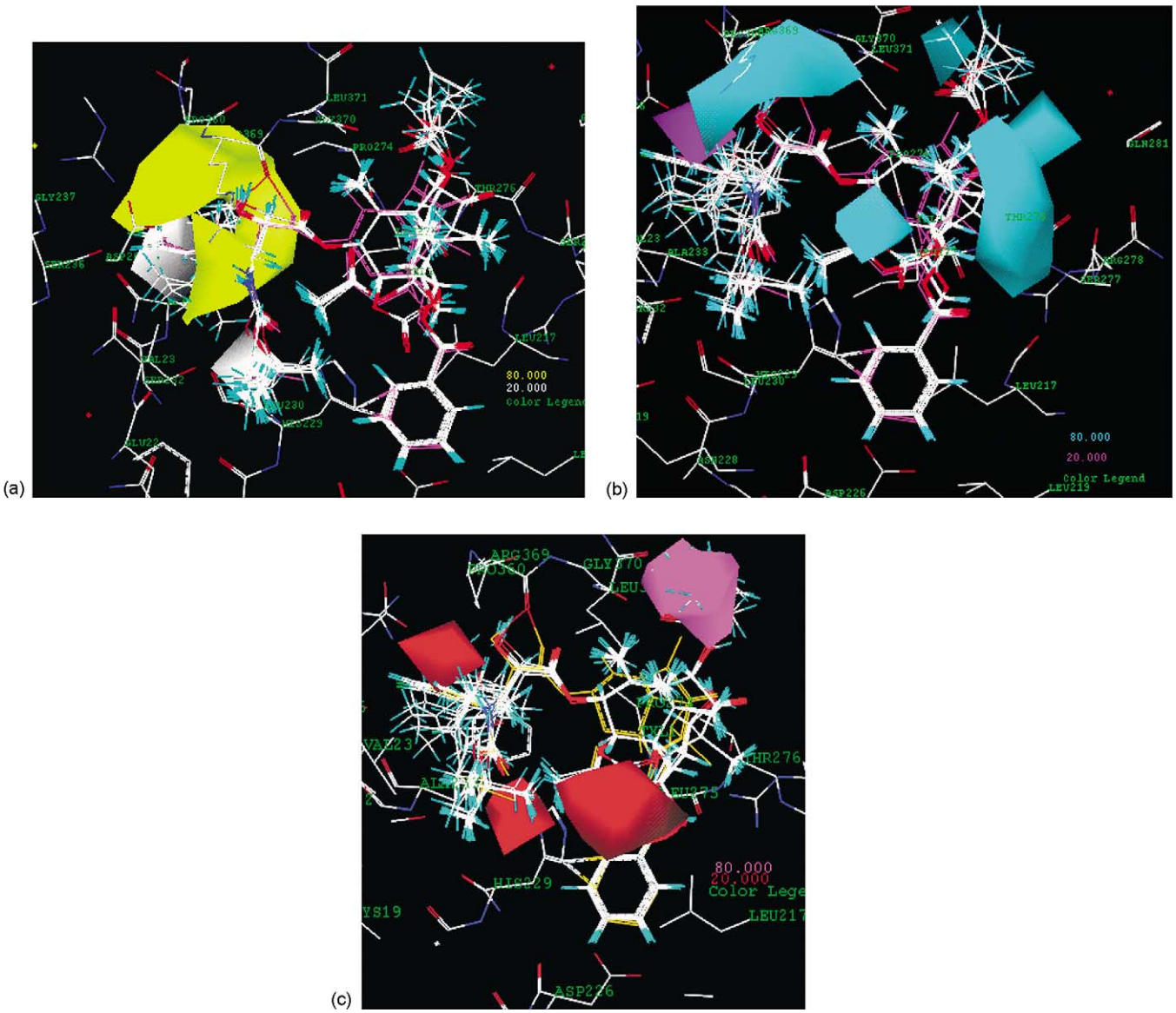


Fig. 4. (a) Contour maps of hydrophobic field of CoMSIA analysis: Yellow color region (80% favored) indicates that hydrophobes are predicted to enhance biological activity whereas the region of white (20% disfavored) color represents the enhancement of biological activity by adding hydrophilic groups. (b) Contour maps of CoMSIA H-bond acceptor field. The magenta contour maps are H-bond donor groups on the receptor favoring binding. The red contours are H-bond donor groups on the receptor disfavoring binding. (c) Contour maps of CoMSIA H-bond donor field. Five areas of cyan (80% favored) indicate where acceptors on the receptor are predicted to favor binding and one purple (20% disfavored) is where an acceptor on the receptor is predicted to disfavor binding. The contour that disfavors acceptors on the receptor site is attributable to compounds with low affinity that have hydrogen-bond donor groups in the vicinity.

4.3. Prediction for the compounds in the test set

The highest predictive power model (model 3a, Table 2) was used to predict the inhibitory activities of the compounds in the test set (Table 1). The comparison of the observed and predicted biological activities of the test set by model 3a is given in Table 2 which shows clearly the usefulness of the model for the prediction of the activities of the compounds that are not included in the training set. The residual values (Table 3) indicated that CoMFA model 3a could accurately predict the activities of all compounds in the test set.

4.4. Pharmacophore model

Using the results obtained by CoMFA, CoMSIA, we have proposed a putative pharmacophore models (Fig. 5a and b) accommodating docetaxel, the most promising antimitotic agent. This model explains the key structural requirements for the activity of paclitaxel analogues. The key structural components may be (i) C10–OH group and (ii) C3′–phenyl groups which are essential for biological activity. Although CoMFA model did not reveal an urgent need of C2 benzoyl group, SAR of other reports [13–16]

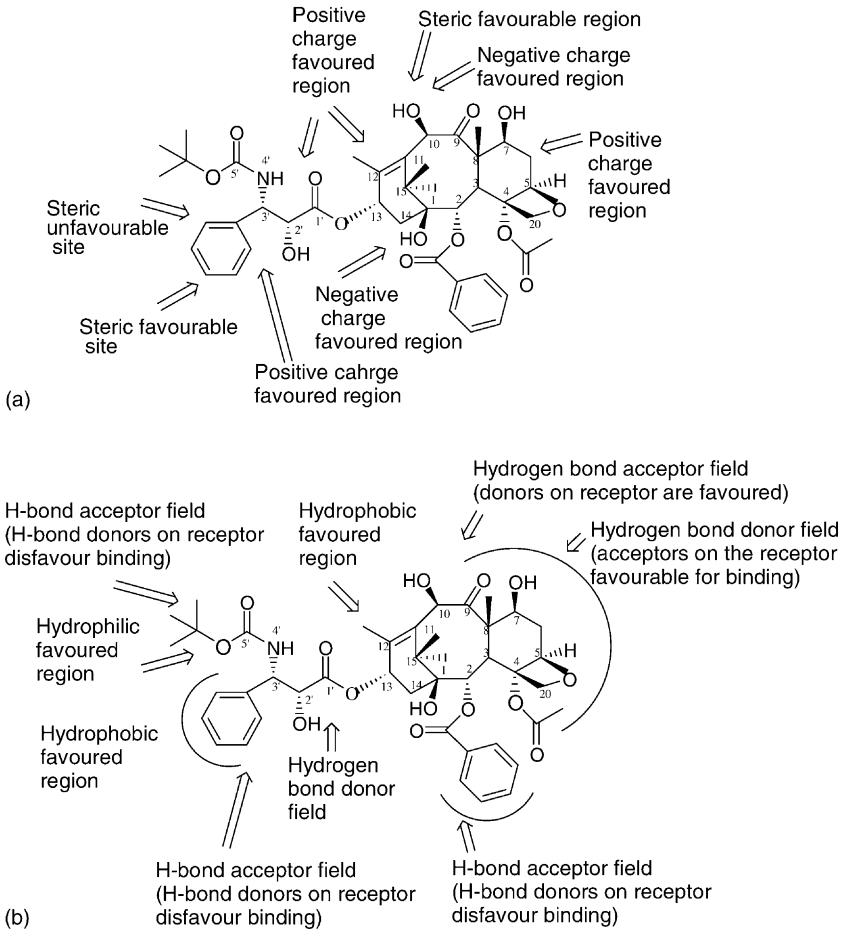


Fig. 5. (a) Pharmacophore model from the CoMFA study. (b) Pharmacophore model from the CoMSIA study.

suggested that it is crucial for both binding and biological activity.

5. Conclusion

The analysis of 19 paclitaxel analogs by both CoMFA and CoMSIA methods provides a good predictive model for the design of new series of active analogs of paclitaxel. In this study, the developed models are reasonable based on both statistical significance and predictive ability. The CoMFA contour maps show that steric and electrostatic fields are more prominent than lipophilic field in contributing towards binding as well as biological activity. These contour maps also show good consistency with ligand–receptor interactions and structural requirements of paclitaxel analogs. The CoMSIA analysis describes the steric, electrostatic, hydrophobic and H-bonding properties of paclitaxel analogs, however, these newly introduced hydrogen bonding fields did not improve the predictive power of the analysis. Paclitaxel and its analogs are very complex and conformationally flexible compounds. Although the result of this study provides a real life 3D-QSAR application, the indicated property features of this analysis, i.e. 3D-QSAR may provide a helpful guideline to design and predict the affinity of novel compounds with enhanced inhibitory activities.

Acknowledgements

This work was supported in part by The Australian Postgraduate Awards (APA) and Monash Graduate Scholarship (MGS). We would like to Mrs. Jan Iskander for proof reading the manuscript.

References

[1] M.A. Jordan, I. Wilson, Curr. Opin. Cell Biol. 10 (1998) 123.
[2] E.K. Rowinsky, L.A. Cazenave, R.C. Donehower, J. Natl. Cancer Inst. 82 (1990) 1247.
[3] N.L. Seibel, G.H. Reaman, Invest. New Drugs 14 (1996) 49.
[4] P.B. Schiff, S.B. Horwitz, Proc. Natl. Acad. Sci. U.S.A. 77 (1980) 1561.
[5] M.C. Wani, H.L. Taylor, M.E. Wall, P. Coggon, A.T. McPhail, J. Am. Chem. Soc. 93 (1971) 2325.
[6] R.F. Service, Science 288 (2000) 27.

[7] R.A. Holton, H.B. Kim, C. Somoza, F. Liang, R.J. Biediger, P.D. Boatman, M. Shindo, C.C. Smith, S. Kim, H. Nadizadeh, Y. Suzuki, C. Tao, P. Vu, S. Tang, P. Zhang, K.K. Murthi, L.N. Gentile, J.H. Liu, J. Am. Chem. Soc. 116 (1994) 1597.
[8] K.C. Nicolaou, Z. Yang, J.J. Liu, H. Ueno, P.G. Nantermet, R.K. Guy, C.F. Claiborne, J. Renaud, E.A. Couladouros, K. Paulvannan, E.J. Sorensin, Nature 367 (1994) 630.
[9] K.D. Miller, G.W. Sledge Jr., Cancer Invest. 17 (1999) 121.
[10] L. He, G.A. Orr, S.B. Horwitz, DDT 6 (2001) 1153.
[11] D.G.I. Kingston, J. Nat. Prod. 63 (2000) 726.
[12] G. Appendino, B. Dannieli, J. Jackupovic, E. Belloro, G. Scambia, E. Bomberdelli, Tetrahedron Lett. 38 (1997) 4273.
[13] J. Parness, D.G.I. Kingston, R.C. Powel, C. Harracksing, S.B. Horwitz, Biochem. Biophys. Res. Commun. 105 (1982) 1082.
[14] V. Guritte, D. Guenard, F. Lavalle, L.M.T. Goff, L. Mangatal, P. Potier, J. Med. Chem. 34 (1991) 992.
[15] G.I. Georg, G.C.B. Harriman, D.G. Vandervelde, T.C. Boge, Z.S. Cheruvallath, ACS Symp. Ser. 583 (1995) 217.
[16] M.Z. Hoemann, D.G. Vandervelde, G.I. Georg, L.R. Jayasinghe, J. Org. Chem. 60 (1995) 2918.
[17] R. Matrovic, R.N. Saicic, J. Chem. Soc., Perkin Trans. 1 (2000) 5965.
[18] O.L. Zamir, J.H. Wu, Anti-Cancer Drug Des. 15 (2000) 73.
[19] E. Baloglu, D.G.I. Kingston, J. Nat. Prod. 62 (1999) 1448.
[20] G. Appendiano, Nat. Prod. Rep. 12 (1995) 349.
[21] I. Ringel, S.B. Horwitz, J. Natl. Cancer Inst. 83 (1991) 288.
[22] I. Ojima, J. Slater, E. Michaud, S.D. Kuduk, P.Y. Bounaud, P. Vrignaud, M.C. Bissery, J. Veith, P. Pera, R.J. Bernacki, J. Med. Chem. 39 (1996) 3889.
[23] I. Ojima, S.D. Kuduk, P. Pera, J.M. Veith, R.J. Bernacki, J. Med. Chem. 40 (1997) 279.
[24] I. Ojima, J. Slater, S.D. Kuduk, C.S. Takeuchi, R.H. Gimi, C.M. Sun, Y.H. Park, P. Pera, J.M. Veith, R.J. Bermacki, J. Med. Chem. 40 (1997) 267.
[25] I. Ojima, S. Lin, T. Wang, Curr. Med. Chem. 6 (1999) 942.
[26] SYBYL v6.6, Tripos Inc., 1699 South Hanley Road, St. Louis, MO 63144, USA.
[27] D.R. Cramer III, D.E. Paterson, J.D. Bunce, J. Am. Chem. Soc. 110 (1988) 5959.
[28] ACD version 3.00, 1997, http://www.acdlabs.com.
[29] R. Kiyama, Y. Tamura, F. Watanabe, J. Med. Chem. 42 (1999) 1723; H. Matter, W. Schwab, D. Barbier, J. Med. Chem. 42 (1999) 1908.
[30] G. Klebe, U. Abraham, T.J. Mietzner, J. Med. Chem. 37 (1994) 4130.