Journal of Molecular Graphics and Modelling 27 (2009) 930–943

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

Applying linear interaction energy method for binding affinity calculations of podophyllotoxin analogues with tubulin using continuum solvent model and prediction of cytotoxic activity

Md. Afroz Alam, Pradeep Kumar Naik *

Department of Bioinformatics and Biotechnology, Jaypee University of Information Technology, Waknaghat, Solan 173215, Himachal Pradesh, India

ARTICLE INFO

Article history:
Received 15 August 2008
Received in revised form 7 February 2009
Accepted 9 February 2009
Available online 20 February 2009

Keywords:
Podophyllotoxin
Free energy of binding
Binding affinity
Docking
Surface generalized Born continuum salvation model
Linear interaction energy

ABSTRACT

Podophyllotoxin and its analogues have important therapeutic value in the treatment of cancer, due to their ability to induce apoptosis in cancer cells in a proliferation-independent manner. These ligands bind to colchicine binding site of tubulin near the α- and β-tubulin interface and interfere with tubulin polymerization. The binding free energies of podophyllotoxin-based inhibitors of tubulin were computed using a linear interaction energy (LIE) method with a surface generalized Born (SGB) continuum solvation model. A training set of 76 podophyllotoxin analogues was used to build a binding affinity model for estimating the free energy of binding for 36 inhibitors (test set) with diverse structural modifications. The average root mean square error (RMSE) between the experimental and predicted binding free energy values was 0.56 kcal/mol which is comparable to the level of accuracy achieved by the most accurate methods, such as free energy perturbation (FEP) or thermodynamic integration (TI). The squared correlation coefficient between experimental and SGB–LIE estimates for the free energy for the test set compounds is also significant (R^2 = 0.733). On the basis of the analysis of the binding energy, we propose that the three-dimensional conformation of the A, B, C and D rings is important for interaction with tubulin. On the basis of this insight, 12 analogues of varying ring modification were taken, tested with LIE methodology and then validated with their experimental potencies of tubulin polymerization inhibition. Low levels of RMSE for the majority of inhibitors establish the structure-based LIE method as an efficient tool for generating more potent and specific inhibitors of tubulin by testing rationally designed lead compounds based on podophyllotoxin derivatization.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Microtubules are involved in a wide range of cellular functions and are critical to the life cycle of the cell. Composed of alternating α- and β-protofilaments, microtubules are highly dynamic macromolecular assemblies that are organized in a polar, spatial and temporal cell cycle specific manner. The organization is regulated by numerous factors including the intrinsic ability of microtubule subunits, tubulin heterodimers, to form non-equilibrium, dynamic polymers. The α- and β-tubulins rapidly assemble and disassemble to meet the cell's needs [1,2]. Since inhibition of tubulin polymerization or blockage of microtubule disassembly increases the number of cells in metaphase arrest, microtubules are attractive molecular targets for anticancer therapeutics. Small molecules have been shown to bind at four major drug binding sites on tubulin: the vinca, taxane, colchicine and peloruside A [3–5].

Among the plethora of physiological activities and potential medicinal and agricultural applications, the antineoplastic and antiviral properties of podophyllotoxin congeners and their derivatives are arguably the most eminent from a pharmacological perspective. Podophyllotoxin is an antitumor lignan mainly found in the plants *Podophyllum hexandrum* and *Podophyllum peltatum*. Since the discovery of the therapeutic properties of podophyllotoxin, new findings related to its activities, its mechanism of action and pharmacological properties have been unveiled. Structure–activity relationships (SAR) have shown that podophyllotoxin analogues preferentially inhibit tubulin polymerization, which leads to arrest of the cell cycle in the metaphase [6,7]. Different derivatives of podophyllotoxin have been demonstrated to bind to the colchicine site, as shown by the fact that podophyllotoxin has been reported to compete with colchicine for the binding site in tubulin [8] and its affinity is double than that of colchicine. These compounds including colchicine affect cancer and normal cells alike and lead to the appearance of adverse side effects [9].

* Corresponding author. Tel.: +91 1792 239227; fax: +91 1792 245362.
E-mail address: pknaik73@rediffmail.com (P.K. Naik).

1093-3263/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2009.02.003

Following binding of podophyllotoxin, the GTP hydrolyzing capacity of tubulin is inhibited, but colchicine stimulates an assembly-independent GTPase activity directed at the exchangeable site-bound GTP [10]. Podophyllotoxin binds to β-tubulin at its interface with α-tubulin resulting in inhibition of tubulin polymerization. This binding mode was recently confirmed by the determination of a 4.20 Å X-ray structure of α- and β-tubulins complexed with podophyllotoxin (PDB_ID:1SA1), showing that podophyllotoxin also binds at the colchicine site [11].

While podophyllotoxin has played a central role in elucidating the physical properties and biological functions of tubulin and microtubules, its high toxicity has limited its therapeutic application [12]. Although colchicine site agents share a general toxicity, the promise to discover therapeutically useful analogues has fueled continued research. Over the years, a large number of natural and synthetic analogues of podophyllotoxin have been identified as colchicine site inhibitors. Since a wide variety of molecular scaffolds are available for optimization, this diversity presents a significant challenge to determining the essential features for activity. A rational approach for the discovery of a pharmaceutically acceptable, economically viable activity model awaits development of a predictive quantitative structure–activity relationship. With the advent of parallel synthesis methods and technology, we might expect the number of podophyllotoxin analogues to be tested to grow dramatically. Combinatorial methods could also be envisioned as a semi-rational approach to this discovery strategy. One method of orchestrating these strategies is to make use of linear interaction energy (LIE) models for the rapid prediction and virtual prescreening of cytotoxic activity. The linear interaction energy approximation is a way of combining molecular mechanics calculations with experimental data to build a model scoring function for the evaluation of ligand–protein binding free energies. The LIE method [13] is a semi-empirical model that has become widely used to predict protein–ligand binding affinities. In LIE, the free ligand in water and the solvated protein–ligand complex are simulated and from these two calculations the ligand surrounding electrostatic and van der Waals (vdw) energies are collected. The binding free energy is then evaluated as proposed by Åqvist [13]. A continuum solvation model was developed based on the proposed LIE method by adding continuum electrostatic ligand–water interaction energies by using an equivalent form of equation [14]. However, the proposed generalized Born (GB)–LIE method overestimates the change in solvation energy and this is caused by consistent underestimation of the effective Born radii in the protein–ligand complex [14]. To further assess the usefulness of continuum models for estimating binding free energies, more accurate GB models should be carried out. The LIE method has been applied on a number of protein–ligand systems with promising results producing small errors on the order of 1 kcal/mol for free energy prediction [15]. This approach could then be applied to larger sets of inhibitors and contribute to fast and efficient ligand design. At present, a linear interaction energy method for rational design of podophyllotoxin analogues for tubulin polymerization inhibition has not been determined.

The availability of structural information on tubulin facilitates understanding the structure–activity relationships for tubulin polymerization inhibition. In this study, we have applied a structure-based linear interaction energy method implementing a surface generalized Born (SGB) continuum model for solvation to build a binding affinity model for estimating the binding free energy for a diverse set of podophyllotoxin analogues with tubulin. The magnitude of free energy changes upon binding of inhibitors to tubulin directly correlates with the experimental potency of these inhibitors; hence, fast and accurate estimation of binding free energies provides a means to screen the compound libraries for lead optimization and for generating more potent and specific inhibitors of tubulin by testing rationally designed lead compounds based on podophyllotoxin derivatization.

2. Materials and methods

2.1. LIE methodology

The LIE method employs experimental data on binding free energy values for a set of ligands (referred as training set) to estimate the binding affinities for a set of novel compounds. The method is based on the linear response approximation (LRA), which dictates that binding free energy of a protein–ligand system is a function of polar and non-polar energy components that scale linearly with the electrostatic and van der Waals interactions between a ligand and its environment. The free energy of binding (FEB) for the complex is derived from considering only two states: (1) free ligand in the solvent and (2) ligand bound to the solvated protein. The conformational changes and entropic effects pertaining to unbound receptor are taken into account implicitly and only interactions between the ligand and either the protein or solvent are computed during molecular mechanics calculations. Among the various formulations of the LIE methodology developed in the past, the SGB–LIE method [15] has been shown to be 1 order of magnitude faster than the methods based on explicit solvent with the same order of accuracy. In the LIE method,

$$\Delta G_{\text{bind}} = \alpha\langle\Delta U_{\text{ele}}\rangle + \beta\langle\Delta U_{\text{vdw}}\rangle + \gamma\langle\Delta_{\text{SASA}}\rangle \quad (1)$$

where $\langle\Delta U_{\text{ele}}\rangle$ and $\langle\Delta U_{\text{vdw}}\rangle$ denotes the average change in the electrostatic and van der Waals interaction energy of the ligand in the free and bound states, respectively, and $\langle\Delta_{\text{SASA}}\rangle$ is the change in the solvent accessible surface area (SASA) of the ligand. The α, β, and γ terms are adjustable parameters that need to be determined by fitting the experimental data on the training set compounds. The SGB–LIE method also offers better accuracy in treating the long-range electrostatic interactions. However, the SGB–LIE method used in this studied is based on the original formulation proposed by Jorgensen [15] and implemented in *Liaison* (Schrödinger, Inc. Portland, OR, USA) using the OPLS-2005 force field. A novel feature of Liaison is that the simulation takes place in implicit (continuum) rather than explicit solvent, hence the name Liaison, for Linear Interaction Approximation in Implicit Solvation. The explicit-solvent version of the methodology was first suggested by Åqvist and Hansson [16], based on approximating the charging integral in the free-energy-perturbation formula with a mean-value approach, in which the integral is represented as half the sum of the values at the endpoints, namely the free and bound states of the ligand. The empirical relationship used by Liaison is shown below:

$$\Delta G_{\text{bind}} = \alpha(\langle U^{\text{b}}_{\text{ele}}\rangle - \langle U^{\text{f}}_{\text{ele}}\rangle) + \beta(\langle U^{\text{b}}_{\text{vdw}}\rangle - \langle U^{\text{f}}_{\text{vdw}}\rangle) + \gamma(\langle U^{\text{b}}_{\text{cav}}\rangle - \langle U^{\text{f}}_{\text{cav}}\rangle) \quad (2)$$

Here "⟨" and "⟩" represent the ensemble average, b represents the bound form of the ligand, f represents the free form of the ligand, and α, β and γ are the coefficients. U_{ele}, U_{vdw} and U_{cav} are the electrostatic, van der Waals and cavity energy terms in the SGB continuum solvent model. The cavity energy term, U_{cav}, is proportional to the exposed surface area of the ligand. Thus, the difference: $\langle U^{\text{b}}_{\text{cav}}\rangle - \langle U^{\text{f}}_{\text{cav}}\rangle$ measures the surface area lost by contact with the receptor. The energy terms involved can be computed using energy minimization, molecular dynamics, or Monte Carlo calculations. In the SGB model of solvation, there is no explicit van der Waals or electrostatic interaction between the solute and solvent. The contribution for net free energy of solvation comes from two energy terms, namely, reaction field energy (U_{rxn}) and

cavity energy (U_{cav}): $U_{SGB} = U_{rxn} + U_{cav}$. The cavity and reaction field energy terms implicitly take into account the van der Waals and the electrostatic interactions, respectively, between the ligand and solvent. The application of the SGB–LIE method for a given protein–ligand system essentially involves computing four energy components, i.e., the van der Waals and Coulombic energy between the ligand and protein and the reaction field and cavity energy between the ligand and continuum solvent. The total electrostatic energy in the SGB–LIE method is the sum of Coulombic and reaction field energy terms.

2.2. Computational details

Preparation of receptor and ligands was done using the Schrödinger package from Schrödinger Inc. [17]. All the calculations for the SGB–LIE method were performed in the Liaison package from Schrödinger Inc. [18]. The Liaison module performs LIE calculations in the OPLS force field with a residue-based cutoff of 15 Å. The OPLS force field was also used for charge assignment and all energy calculations.

2.3. Receptor preparation

The X-ray structure of the complex between podophyllotoxin and tubulin protein (PDB_ID:1SA1) has been used as initial structure in the preparation of podophyllotoxin binding site. After manual inspection and cleaning of structure we retained a complex composed of protein chains α and β and podophyllotoxin ligand. Hydrogen was added to the model automatically via the Maestro interface [19] leaving no lone pair and using an explicit all-atom model. All the water molecules were removed from the complex. The multi step Schrödinger's Protein preparation tool (PPrep) has been used for final preparation of protein. PPrep neutralizes side chains that are not close to the binding cavity and do not participate in salt bridges [19]. This step is then followed by restrained minimization of co-crystallized complex, which reorients side chain hydroxyl groups and alleviates potential steric clashes. Progressively weaker restraints (tethering force constants 3, 1, 0.3, 0.1) were applied to non-hydrogen atoms only. The complex structure was energy minimized using OPLS_2005 force field and the conjugate gradient algorithm, keeping all atoms except hydrogen fixed. The minimization was stopped either after 1000 steps or after the energy gradient converged below 0.01 kcal/mol. The energy-minimized receptor structure was subsequently used for docking of podophyllotoxin analogues and SGB–LIE calculations.

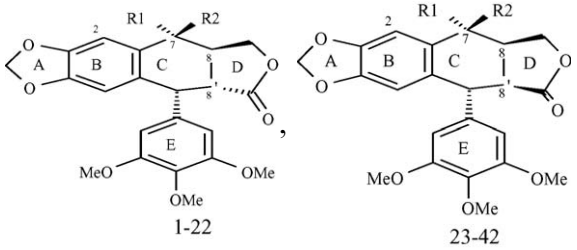
2.4. Preparation of ligands

Podophyllotoxin is well known for its antitumor activity. However, the clinical application of it and its analogues in the treatment of cancer has been limited by severe toxic side effects during administration of the drugs [20,21]. With a view to achieving greater therapeutic efficiency many podophyllotoxin analogues have been isolated and via molecular manipulation, a large number of semisynthetic derivatives have been synthesized. However, new findings related to their activities, mechanism of action and pharmacological properties have been unexplored. A total of 112 podophyllotoxin analogues were used in the study and were taken from various sources belonging to different ring modifications. For better interpretation all these compounds were divided into following four sublibraries.

Sublib-I commonly known as tetralinelactones consist of 29 compounds (1–29) (Table 1a). These molecules were rationally designed as functional mimics of natural podophyllotoxin with the goal of simplifying the chemical synthesis and improving the cytotoxic activity. Structural modification mainly introduced varying radicals at position 7 in podophyllotoxin scaffold. Reports have been made of compounds with oxygenated substituents in the form of ethers, esters and diverse nitrogen radicals [22–26].

Sublib-II contains compounds (30–70) (Table 1b) known as non-lactonic tetralines. Structural modifications in this group include the opening of the lactone ring (D-ring) in podophyllotoxin scaffold, to give rise to compounds with different degrees of oxidation at positions C-9 and C-9′ [23,26]. In general these molecules lacking a lactone ring.

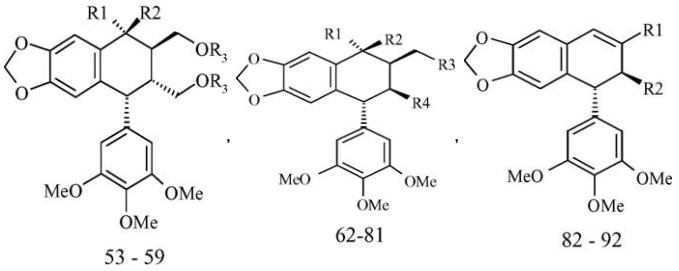
Table 1a
Podophyllotoxin derivatives (tetraline lactones) with cytotoxic activities against P-388 cell line used in the work.

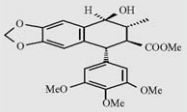
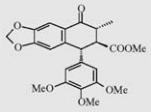


.

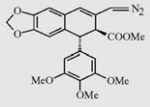
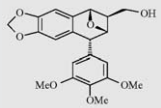
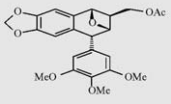
Analogue	R1	R2	Experimental IC_{50}	Analogue	R1	R2	Experimental IC_{50}
1	OH	H	0.012	16	H	H	0.10
2	H	H	0.010	17	H	H(2–OMe)	0.23
3	H	H(2–OMe)	0.01	18	OH	H	6.0
4	OH	H(4′–OH)	0.027	19	OAc	H	0.55
5	OAc	H	0.625	20	OAc	H(2–OMe)	1.02
6	OMe	H	0.06	21	OMe	H	0.12
7	H	OH	0.06	22	H	OH(2–OMe)	0.11
8	H	Ac	0.05	23	H	OAc	0.44
9	H	OMe	0.06	24	H	OAc(2–OMe)	0.51
10	H	Cl	0.6	25	H	OMe	0.12
11	Cl	H	0.6	26	H	$H\Delta^7$	0.013
12	=O		1.8	27	=O		12.0
13	=N–OH		2.3	28	=N–OH		2.3
14	=N–OAc		2.1	29	=N–OMe		2.3
15	=N–OMe		0.2				

Table 1b
Podophyllotoxin derivatives (non-lactonic tetralines) with cytotoxic activities against P-388 cell line used in the work.



Analogue	R1	R2	R3	Experimental IC_{50}	Analogue	Structure	Experimental IC_{50}
30	OH	H	H	1.2	35		23.3
31	H	OH	H	12.0			
32	H	OMe	H	11.6			
33	H	OMe	Ac	9.7			
34	OMe	H	Ac	9.7	36		3.5

Analogue	R1	R2	R3	R4	Experimental IC_{50}	Analogue	R1	R2	R3	R4	Experimental IC_{50}
37	H	H	OH	COOMe	0.058	47	H	OMe	OAc	CH_2OAc	9.7
38	H	H	OAc	COOMe	0.21	48	H	OH	OH	CH_2OH	47.9
39	H	H	OAc	CH_2OAc	5.14	49	H	OH	OH	COOMe	1.1
40	OH	H	OH	CH_2OH	23.9	50	=O		OH	COOMe	5.63
41	OH	H	OH	COOMe	0.22	51	=O		OAc	COOMe	0.20
42	OAc	H	OAc	CH_2OAc	7.4	52	=N–OH		OAc	COOMe	2.0
43	OAc	H	OAc	COOMe	1.1	53	H	H	CHO	COOMe	2.34
44	OMe	H	OH	CH_2OH	23.2	54	H	H	=N–OMe	COOMe	2.30
45	OMe	H	OAc	CH_2OAc	19.4	55	H	H	=N–OMe	COOMe	10.94
46	H	OMe	OH	CH_2OH	11.6	56	H	H	=N–allyl	COOMe	2.5

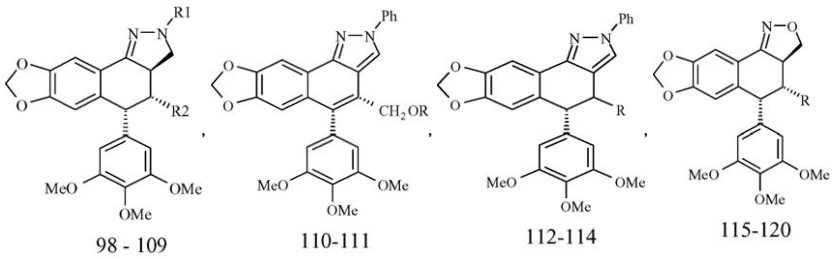
Analogue	R1	R2	Experimental IC_{50}	Analogue	R1	R2	Experimental IC_{50}
57	CH_2OH	COOMe	0.02	64	CH=N–OH	COOMe	2.27
58	CHO	CH_2OH	0.25				
59	CHO	COOMe	0.23	65	CH=N–OMe	COOMe	0.22
60	CH=N–NH2	COOMe	0.57	66		COOMe	0.20
61	CH=N–NH–CH_2CF_3	COOMe	0.48	67		CH_2OH	1.00
62	CH=N–NH–Ph	COOMe	1.94	68			0.57
63	CH=N–NH–Ph	CH_2OH	1.02				5.66
69			6.25	70			

Sublib-III also includes a group of lignans (71–84) (Table 1c) that have heterocyclic rings fused to the cyclolignan skeleton. This group is commonly called as pyrazolignans [23,25–27] and isoxazolignans [26,28] and they were obtained by reacting podophyllotoxin with differently substituted hydrazines and hydroxylamines.

Sublib-IV contains 28 compounds (85–112) (Table 1d) commonly known as aza-podophyllotoxin analogues. The preparation of this group of compounds requires selective chemical manipulation of the two aromatic rings (B- and E-rings) of the podophyllotoxin scaffold. These molecules are readily prepared from anilines, benzaldehydes and tetronic acid or 2,3-cyclopentanedione in good to excellent yield and have also shown better cytotoxic activity [29].

All these podophyllotoxin analogues were built from the scaffolds by different ring modification and substitution of functional groups as mentioned in Tables 1a–1d. We used ISIS Draw 2.3 software for sketching structures and converting them to their 3D representation by using ChemSketch 3D viewer of ACDLABS 8.0. LigPrep [19] was used for final preparation of ligands from libraries. LigPrep is a utility of Schrödinger software suit that combines tools for generating 3D structures from 1D (Smiles) and 2D (SDF) representation, searching for tatutomers and steric isomers and performing a geometry minimization of ligands.

Table 1c
Podophyllotoxin derivatives (pyrazolignans and isoxazolignan) with cytotoxic activities against P-388 cell used in the work.



Analogue	R1	R2	Experimental IC_{50}	Analogue	R1	R2	Experimental IC_{50}
71	Ph	COOH	1.9	74	*m*-NO_2Ph	COOMe	4.5
72	Ph	CH_2OH	4.1	75	Me	COOMe	5.6
73	Ph	CH_2OAc	4.7	76	$COCH_3COOMe$	COOMe	21

Analogue	R	Experimental IC_{50}	Analogue	R	Experimental IC_{50}
77	H	10	81	COOMe	23
78	CHO	21	82	COOMe(4′–OH)	12
79	CH_2Ac	2.2	83	CH_2OH	2.6
80	COOH	2.2	84	CHO	2.4

The ligands were minimized by means of molecular mechanics force fields (MMFFs) with default setting. Each of these compounds had associated in vitro cytotoxicity values (IC_{50} values reported in μM) against cell line P388. Studied on in vitro cytotoxicity of podophyllotoxin and its analogues were reported mostly on P388 cell line. The reason being due to its resistance to anticancer drug vinorelbine [30]. P388 is a murine leukemia cell line. Out of the seven β-tubulin isotype classes; class I was the major β-tubulin isotype (60–72%), followed by class III (11.3–11.7%) while β-tubulin classes IVa + IVb were the least abundant (1.2–1.7%) of total β-tubulin in P388 cell line [31].

2.5. Docking of the ligands

All the ligands were docked to the tubulin receptor using Glide version 4.0. After ensuring that protein and ligands are in correct form for docking, the receptor-grid files were generated using grid-receptor generation program, using van der Waals scaling of the receptor at 0.4. The default size was used for the bounding and enclosing boxes was generated at the centroid of the tubulin binding site by selecting the bound podophyllotoxin ligand. The ligands were docked initially using the "standard precision" method and further refined using "extra precision" Glide algorithm. For the ligand docking stage, van der Waals scaling of the ligand was set at 0.5. Of the 50,000 poses that were sampled, 4000 were taken through minimization (conjugate gradients 1000) and the 30 structures having the lowest energy conformations were further evaluated for the favorable Glide docking score. A single best conformation for each ligand was considered for further analysis.

2.6. LIE calculations

The docked complex corresponding to each analogue was transported to the Liasion package for subsequent SGB–LIE calculations. Sampling technique such as molecular dynamics (MD) has been used for LIE conformation space sampling in the present work. The system was initially heated to 300 K for 5 ps and then subjected to a MD simulation for 25 ps. A residue-based cutoff of 12 Å was set for the non-bonding interactions. The non-bonded pair list was updated every 10 fs. The time integration step of 1.0 fs and sampling LIE energies every 10 steps was used. During the MD simulations, all the residues of the receptor beyond 12 Å from the bound ligand were frozen. Similarly, the average LIE energies for the ligand were obtained using the OPLS-2005 force field. The average LIE energy terms were used for building binding affinity model and free energy estimation for podophyllotoxin analogues. The α, β and γ LIE fitting parameters were determined based on Gaussian elimination method using Matlab 6.5 as described by Thomas and Finny [32] and by fitting the experimental data on the training set compounds.

In order to explore the reliability of the proposed model we used the cross validation method. Prediction error sum of squares (PRESS) is a standard index to measure the accuracy of a modeling method based on the cross validation technique. The r^2_{cv} was calculated based on the PRESS and SSY (sum of squares of deviations of the experimental values from their mean) using following formula:

$$r^2_{cv} = 1 - \frac{\text{PRESS}}{\text{SSY}} = 1 - \frac{\sum_{i=1}^{n} (y_{exp} - y_{pred})^2}{\sum_{i=1}^{n} (y_{exp} - \bar{y})^2}$$

where y_{exp}, y_{pred} and $\bar{y}$ are the predicted, observed and mean values of the cytotoxic activities of the podophyllotoxin analogues. The cross validation analysis performed by using the leave one out (LOO) method in which one compound removed from the data set and its activity predicted using the model derived from the rest of the data points. The cross-validated correlation coefficient (q^2) that resulted in optimum number of components and lowest standard error of prediction were considered for further analysis and calculated using following equations:

$$q^2 = 1 - \frac{\sum_y (y_{pred} - y_{observed})^2}{\sum_y (y_{observed} - y_{mean})^2}$$

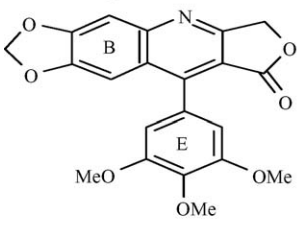
$$\text{PRESS} = \sum_y (y_{predicred} - y_{observed})^2$$

where y_{pred}, $y_{observed}$ and y_{mean} are the predicted, observed and mean values of the cytotoxic activities of the podophyllotoxin analogues and PRESS is the sum of the predictive sum of squares. The predictive ability of the models is expressed by the r^2 predictive value, which is analogous to cross-validated r^2 (q^2).

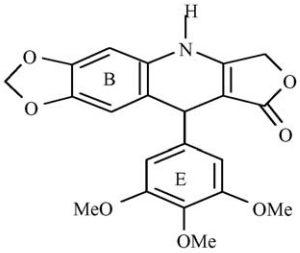
$$r^2_{pred} = \frac{\text{SD} - \text{PRESS}}{\text{SD}}$$

Table 1d
Aza-podophyllotoxin derivatives with cytotoxic activities against P-388 cell line used in the work.

Aza-Analogue



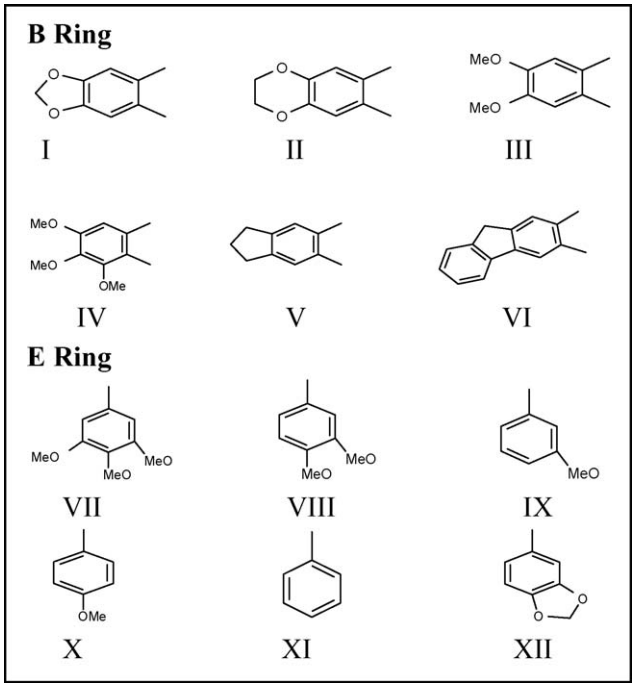
Modification 1



Modification 2

.

Substitution of B- and E-rings at 1 and 2 analogues:							
Modification 1				Modification 2			
Analogue	B-ring	E-ring	Experimental IC_{50}	Analogue	B-ring	E-ring	Experimental IC_{50}
85	I	VII	100	99	I	VII	0.0018
86	II	VII	80	100	II	VII	0.0017
87	III	VII	100	101	III	VII	4.9
88	III	VIII	39	102	III	VIII	0.76
89	III	XII	2.0	103	III	XII	0.77
90	IV	VII	29	104	IV	VII	2.6
91	V	VII	100	105	V	VII	0.0041
92	VI	VII	63	106	VI	VII	0.92
93	I	VIII	40	107	I	VIII	0.048
94	I	IX	100	108	I	IX	0.0053
95	I	X	100	109	I	X	0.13
96	I	XI	60	110	I	XI	0.0053
97	I	XII	100	111	I	XII	0.030
98	I	VII	71	112	I	VII	0.028



3. Results and discussions

The original crystal structure of tubulin–podophyllotoxin complex PDB_ID:1SA1 (PDB ID:1SA1) was used to validate the Glide-XP docking protocol. This was done by moving the co-crystallized podophyllotoxin ligand outside of active site and then docking it back into the active site. The top 10 configurations after docking were taken into consideration to validate the result (Table 2). The RMSD was calculated for each configuration in comparison to the co-crystallized podophyllotoxin and the value was found to be in between 0.02 and 0.85 Å. Whereas the RMSD value calculated out of 10 accepted poses for each configuration was found in between 0.59 and 1.33 Å. This revealed that the docked configurations have similar binding positions and orientations within the binding site and are similar to the crystal structure. The best-docked structure, which is the configuration

Table 2
The RMSD and docking score from the docking simulation of 10 lowest configurations of co-crystal podophyllotoxin in tubulin protein (1SA1).

Configuration	Glide score	ΔG_{score}[a]	RMSD (Å)[b]	RMSD (Å)[c]
1	−10.26	0	0.85	0.60
2	−10.20	−0.06	0.02	0.86
3	−9.80	−0.46	0.68	1.33
4	−9.72	−0.54	0.57	1.26
5	−9.50	−0.76	0.04	0.67
6	−9.25	−1.01	0.04	0.67
7	−8.78	−1.48	0.80	0.59
8	−8.47	−1.79	0.13	1.02
9	−7.87	−2.39	0.03	0.79
10	−7.72	−2.54	0.07	0.90

[a] $\Delta G_{score} = E_i - E_{lowest}$.
[b] RMSD = RMSD between docked and crystallographic podophyllotoxin structure.
[c] RMSD = RMSD between docked poses corresponding to each configuration.

with the lowest Glide score, is compared with the crystal structure and is shown in Fig. 1. These docking results illustrate that the best-docked podophyllotoxin complex agrees well with its crystal structure and that Glide-XP docking protocol successfully reproduces the crystal tubulin–podophyllotoxin complex. The binding modes of five superimposed ligands from each class within podophyllotoxin binding site are given in Fig. 2(a–d). In this figure we can observe that all the ligands were well fitted the defined binding pocket.

We have applied the SGB–LIE method to a training set of 76 podophyllotoxin analogues to build a binding affinity model that was then used to compute the free energy of binding and predicted pIC_{50} for a test set of 36 analogues. Further the SGB–LIE model developed was validated using 12 new podophyllotoxin analogues for which the experimental tubulin polymerization inhibition was known. The training set for building the binding affinity model was comprised of four subsets of podophyllotoxin analogues as mentioned in Tables 1a–1d. For all the four subsets included in the training set the experimental IC_{50} values against the cell lines P388 are available. With the wide range of difference between the IC_{50} values and the large diversity in the structures, the combined set of 76 ligands is ideal to be considered as a training set, as the set does not suffer from bias, due to the similarity of the structures. Also, the training set containing 76 analogues contains enough data points not to suffer from over parameterization by the LIE model. Training set compounds were docked into the colchicine binding site of tubulin protein and the SGB–LIE calculations were performed using the Liaison module. The simulations were performed both for the ligand-free and ligand-bound state. The various interaction energy terms described in the methods were collected and are presented in Tables 3a–3d. The largest contribution for the binding energy comes from the van der Waals interactions. This is obvious as the podophyllotoxin analogues used in the study are mostly lipophilic molecules that interact favorably with a binding cavity lined with hydrophobic residues. The hydrophobic center that is located in the middle of trimethoxyphenyl moiety of podophyllotoxin is surrounded by Leu β242, Ala β250, Leu β255, Ala β316, Val β318 and Ile β378 residues (Fig. 2(a–d)). The cavity energy term in the bound state is smaller (1.45–2.07 kcal/mol) than in the free state (3.18–7.62 kcal/mol) for all the compounds, as there is less energy penalty for creating a cavity in solvent when part of the ligand is buried into the hydrophobic binding site. The reaction field energy term in the free state lies in a very narrow range (−22.41 to −28.73 kcal/mol) for all compounds, but it varies in a wide range in the bound state (−6.37 to −31.33 kcal/mol) as the solvent accessible surface area varies with ligand structure in the bound form. The energy values in Tables 3a–3d were used to fit Eq. (2) using the Gaussian

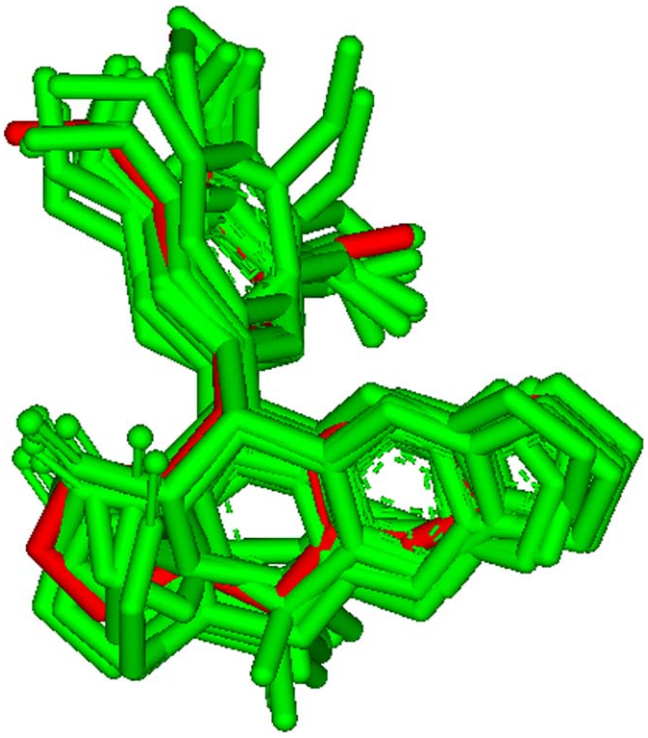


Fig. 1. Superposition of all the docked configurations of podophyllotoxin on crystal structure (red-stick). RMSD (heavy atoms) = 0.02–0.85 Å.

elimination method. The values obtained for the three fitting parameters, α, β and γ are – 0.141, −0.093 and −1.071, respectively. The large value of the cavity energy term signifies the fact that binding is largely driven by the ligand's ability to bury itself in the binding cavity, which is understandable given that most of the ligands are highly hydrophobic in nature. Even though the R value is low, vdw interactions contribute significantly toward the free energy of binding due to the large magnitude of the vdw interaction term. In Tables 3a–3d, the experimental free energy values obtained from the $RT\ln IC_{50}$ and the free energy values estimated using SGB–LIE fitting parameters are presented. The root mean square error (RMSE) between the experimental values and the values obtained by the fit was 0.48 kcal/mol, which is an indicator of the robustness of the fit. The quality of the fit can also be judged by the value of the squared correlation coefficient (r^2), which was 0.871 for the training set. Fig. 3 graphically shows the quality of fit. The statistical significance of the SGB–LIE model is evaluated by the correlation coefficient r, standard error s, F-test value, significance level of the model P, leave-one-out cross-validation coefficient q^2 and predictive error sum of squares PRESS.

$$\Delta G_{bind} = (-0.141)\langle U_{ele}\rangle + (-0.093)\langle U_{vdw}\rangle + (-1.07)\langle U_{cav}\rangle \quad (3)$$

(n = 76, r^2 = 0.871, r^2_{pred} = 0.864, s = 0.598, F = 166.8, P = 0.0001, q^2 = 0.865, PRESS = 28.05)

The SGB–LIE model developed in this study is statistically (q^2 = 0.865, r^2 = 0.871, F = 166.77) best fitted and consequently used for prediction of cytotoxic activities (pIC_{50}) of training and test sets of molecules as reported in Tables 3a–3d and 4. The predicted activity calculated from free energy of binding is satisfactory with small deviation compared with experimental activity of training and test sets of molecules. The calculated free energy of binding represents the experimental activity well. Theoretically, FEB can be partitioned into several components: vdw, electrostatic and solvent accessible surface area [13]. In this

Table 3a
Average electrostatic (ele), van der Waals (vdw) and cavity (cav) energy terms as well as binding affinity model calculations for the first Training subset inhibitors (tetralinelactone podophyllotoxin analogues) using SGB–LIE method.

Ligand	$\langle U_{ele}\rangle$ (kcal/mol)[a]	$\langle U_{vdw}\rangle$ (kcal/mol)[a]	$\langle U_{cav}\rangle$ (kcal/mol)[a]	$pIC_{50,expt}$[b]	$\Delta G_{bind,expt}$ (kcal/mol)[c]	$\Delta G_{bind,LIE}$ (kcal/mol)[d]	$pIC_{50,pred}$[e]
1	11.7	−41.4	3.8	1.921	−2.6	−1.9	1.413
2	12.4	−42.0	4.3	2.002	−2.7	−2.5	1.810
3	12.0	−44.7	4.2	2.002	−2.7	−2.0	1.483
5	10.5	−52.5	3.8	0.198	−0.3	−0.7	0.542
7	10.5	−48.5	4.1	1.217	−1.7	−1.3	0.977
8	13.7	−43.6	3.8	1.298	−1.8	−2.0	1.471
9	9.0	−35.9	2.4	1.217	−1.7	−0.5	0.363
11	11.4	−47.4	3.8	0.220	−0.3	−1.3	0.960
13	10.4	−49.2	2.1	−0.359	0.5	0.8	−0.581
15	13.9	−57.4	3.8	0.697	−0.9	−0.7	0.509
16	10.4	−53.4	4.2	0.997	−1.4	−1.0	0.755
18	10.6	−63.8	2.3	−0.777	1.1	1.9	−1.391
19	12.9	−58.9	4.2	0.257	−0.3	−0.8	0.618
21	11.5	−53.7	3.5	0.917	−1.2	−0.4	0.269
23	11.0	−57.8	4.0	0.359	−0.5	−0.5	0.355
25	11.5	−54.2	3.7	0.917	−1.2	−0.6	0.449
26	11.4	−49.4	3.6	1.892	−2.6	−0.8	0.623
28	10.5	−56.2	2.9	−0.359	0.5	0.6	−0.479
29	7.1	−58.6	3.4	−0.359	0.5	0.8	−0.593

[a] $\langle U_{ele}\rangle$, $\langle U_{vdw}\rangle$ and $\langle U_{cav}\rangle$ energy terms represents the ensemble average of the energy terms calculated as the difference between bound and free state of ligands and its environment.
[b] pIC_{50} refers to the experimental predicted cytotoxic activity using P388 cell line and is calculated as $pIC_{50} = -\log IC_{50}$.
[c] $\Delta G_{bind,expt}$ refers to free energy of binding for tubulin inhibition and is computed using the relationship: $\Delta G_{binding} \approx -2.303RTpIC_{50,expt}$, where 298 K is used in the work for temperature T.
[d] $\Delta G_{bind,LIE}$ refer to the absolute free energy values obtained using SGB–LIE method.
[e] $pIC_{50,\ pred}$ refers to predicted cytotoxic activity of ligands and is estimated using the relationship: $pIC_{50pred} = -(\Delta G_{bind,LIE}/2.303RT)$.

Table 3b
Average electrostatic (ele), van der Waals (vdw) and cavity (cav) energy terms as well as binding affinity model calculations for the second Training subset inhibitors (non-lactonic tetralines podophyllotoxin analogues) using SGB–LIE method.

Ligand	$\langle U_{ele}\rangle$[a] (kcal/mol)	$\langle U_{vdw}\rangle$[a] (kcal/mol)	$\langle U_{cav}\rangle$[a] (kcal/mol)	$pIC_{50,expt}$[b]	$\Delta G_{bind,expt}$[c] (kcal/mol)	$\Delta G_{bind,LIE}$[d] (kcal/mol)	$pIC_{50,pred}$[e]
30	8.3	−44.4	2.8	−0.161	0.2	−0.0	0.022
32	6.8	−54.0	2.1	−1.151	1.6	1.8	−1.335
33	9.0	−61.8	3.1	−0.836	1.1	1.2	−0.854
34	7.9	−52.6	2.2	−0.953	1.3	1.4	−1.040
36	11.8	−57.2	2.8	−0.616	0.8	0.7	−0.482
37	14.3	−46.7	4.4	1.012	−1.4	−2.4	1.744
38	11.8	−42.4	3.7	0.719	−1.0	−1.8	1.288
40	10.3	−51.0	1.7	−1.012	1.4	1.4	−1.053
41	11.4	−44.8	4.3	0.924	−1.3	−2.1	1.548
42	6.6	−51.4	2.2	−0.968	1.3	1.5	−1.097
44	9.2	−48.0	1.8	−0.851	1.2	1.2	−0.867
45	7.6	−48.5	1.8	−0.990	1.3	1.5	−1.098
46	8.0	−57.0	2.2	−1.181	1.6	1.8	−1.340
48	9.2	−56.8	1.7	−1.364	1.9	2.1	−1.545
49	11.8	−50.7	3.3	0.015	−0.0	−0.5	0.356
51	12.5	−46.8	4.0	0.653	−0.9	−1.7	1.221
52	8.3	−47.0	2.8	−0.257	0.3	0.1	−0.100
54	11.0	−55.5	2.8	−0.557	0.8	0.6	−0.426
55	11.4	−59.4	2.1	−1.181	1.6	1.7	−1.242
56	12.3	−54.1	2.2	−0.763	1.0	0.9	−0.663
57	14.0	−51.9	4.7	1.489	−2.0	−2.2	1.632
59	14.5	−52.1	4.4	0.763	−1.0	−1.9	1.428
60	13.3	−46.4	3.3	0.323	−0.4	−1.1	0.808
62	11.1	−59.2	2.8	−0.733	1.0	0.9	−0.656
64	12.1	−54.2	2.7	−0.477	0.6	0.4	−0.290
65	14.8	−51.8	4.8	1.034	−1.4	−2.4	1.789
66	13.2	−53.7	4.2	0.462	−0.6	−1.3	0.989
68	14.8	−56.2	4.3	0.528	−0.7	−1.5	1.122
69	12.7	−59.5	2.6	−0.777	1.1	0.9	−0.669
70	6.9	−55.4	2.4	−1.012	1.4	1.6	−1.143

[a] $\langle U_{ele}\rangle$, $\langle U_{vdw}\rangle$ and $\langle U_{cav}\rangle$ energy terms represents the ensemble average of the energy terms calculated as the difference between bound and free state of ligands and its environment.
[b] pIC_{50} refers to the experimental predicted cytotoxic activity using P388 cell line and is calculated as $pIC_{50} = -\log IC_{50}$.
[c] $\Delta G_{bind,expt}$ refers to free energy of binding for tubulin inhibition and is computed using the relationship: $\Delta G_{binding} \approx -2.303RTpIC_{50,expt}$, where 298 K is used in the work for temperature T.
[d] $\Delta G_{bind,LIE}$ refer to the absolute free energy values obtained using SGB–LIE method.
[e] $pIC_{50,\ pred}$ refers to predicted cytotoxic activity of ligands and is estimated using the relationship: $pIC_{50pred} = -(\Delta G_{bind,LIE}/2.303RT)$.

Table 3c
Average electrostatic (ele), van der Waals (vdw) and cavity (cav) energy terms as well as binding affinity model calculations for the third training subset inhibitors (pyrazolignans and isoxazolignans podophyllotoxin analogues) using SGB–LIE method.

Ligand	$\langle U_{ele}\rangle$[a] (kcal/mol)	$\langle U_{vdw}\rangle$[a] (kcal/mol)	$\langle U_{cav}\rangle$[a] (kcal/mol)	$pIC_{50,expt}$[b]	$\Delta G_{bind,expt}$[c] (kcal/mol)	$\Delta G_{bind,LIE}$[d] (kcal/mol)	$pIC_{50,pred}$[e]
71	11.8	−47.1	2.4	−0.565	0.8	0.2	−0.117
73	5.6	−50.8	2.7	−0.726	1.0	1.0	−0.725
74	9.7	−53.0	3.2	−0.660	0.9	0.1	−0.095
75	9.5	−55.3	3.4	−0.726	1.0	0.1	−0.089
77	11.2	−65.4	3.4	−0.909	1.2	0.8	−0.592
78	5.8	−50.6	2.1	−1.012	1.4	1.6	−1.198
80	4.4	−44.5	2.7	−0.623	0.8	0.6	−0.477
82	9.6	−63.2	2.8	−0.924	1.3	1.5	−1.112
84	10.8	−47.3	2.3	−0.653	0.9	0.4	−0.328

[a] $\langle U_{ele}\rangle$, $\langle U_{vdw}\rangle$ and $\langle U_{cav}\rangle$ energy terms represents the ensemble average of the energy terms calculated as the difference between bound and free state of ligands and its environment.
[b] pIC_{50} refers to the experimental predicted cytotoxic activity using P388 cell line and is calculated as $pIC_{50} = -\log IC_{50}$.
[c] $\Delta G_{bind,expt}$ refers to free energy of binding for tubulin inhibition and is computed using the relationship: $\Delta G_{binding} \approx -2.303RTpIC_{50,expt}$, where 298 K is used in the work for temperature T.
[d] $\Delta G_{bind,LIE}$ refer to the absolute free energy values obtained using SGB–LIE method.
[e] $pIC_{50,\ pred}$ refers to predicted cytotoxic activity of ligands and is estimated using the relationship: $pIC_{50pred} = -(\Delta G_{bind,LIE}/2.303RT)$.

Table 3d
Average electrostatic (ele), van der Waals (vdw) and cavity (cav) energy terms as well as binding affinity model calculations for the fourth training subset inhibitors (aza-podophyllotoxin analogues) using SGB–LIE method.

Ligand	$\langle U_{ele}\rangle$[a] (kcal/mol)	$\langle U_{vdw}\rangle$[a] (kcal/mol)	$\langle U_{cav}\rangle$[a] (kcal/mol)	$pIC_{50,expt}$[b]	$\Delta G_{bind,expt}$[c] (kcal/mol)	$\Delta G_{bind,LIE}$[d] (kcal/mol)	$pIC_{50,pred}$[e]
85	6.1	−60.6	3.0	−1.951	2.7	1.6	−1.150
87	6.7	−54.6	2.3	−2.017	2.7	1.6	−1.202
89	4.9	−48.2	3.3	−0.521	0.7	0.2	−0.164
91	5.1	−53.8	2.5	−1.936	2.6	1.6	−1.174
92	8.9	−54.0	2.2	−1.826	2.5	1.4	−1.023
94	7.7	−48.6	2.4	−1.239	1.7	0.9	−0.632
95	6.3	−50.3	2.8	−1.085	1.5	0.7	−0.542
97	4.4	−66.5	3.6	−2.061	2.8	1.7	−1.245
99	12.9	−41.8	4.3	2.420	−3.3	−2.6	1.895
100	11.9	−36.8	5.0	2.706	−3.7	−3.6	2.629
102	5.2	−52.3	4.5	0.345	−0.5	−0.7	0.486
103	1.8	−53.9	4.9	0.271	−0.4	−0.5	0.368
105	9.8	−45.9	4.6	1.635	−2.2	−2.0	1.494
107	8.2	−57.1	5.1	0.939	−1.3	−1.3	0.987
108	8.6	−37.0	4.6	2.405	−3.3	−2.7	2.008
110	9.4	−42.4	4.9	2.303	−3.1	−2.7	1.961
111	8.0	−48.9	5.6	2.185	−3.0	−2.5	1.870
112	8.5	−46.5	4.9	1.782	−2.4	−2.1	1.577

[a] $\langle U_{ele}\rangle$, $\langle U_{vdw}\rangle$ and $\langle U_{cav}\rangle$ energy terms represents the ensemble average of the energy terms calculated as the difference between bound and free state of ligands and its environment.
[b] pIC_{50} refers to the experimental predicted cytotoxic activity using P388 cell line and is calculated as $pIC_{50} = -\log IC_{50}$.
[c] $\Delta G_{bind,expt}$ refers to free energy of binding for tubulin inhibition and is computed using the relationship: $\Delta G_{binding} \approx -2.303RTpIC_{50,expt}$, where 298 K is used in the work for temperature T.
[d] $\Delta G_{bind,LIE}$ refer to the absolute free energy values obtained using SGB–LIE method.
[e] $pIC_{50,\ pred}$ refers to predicted cytotoxic activity of ligands and is estimated using the relationship: $pIC_{50pred} = -(\Delta G_{bind,LIE}/2.303RT)$.

study the SASA energy term has been replaced by the cavity energy term as proposed by Zhou et al. [15].

Satisfied with the robustness of the binding affinity model developed using the training set, we applied the LIE model to the podophyllotoxin analogues comprising the test set. The test set includes 36 compounds categorized into four subgroups as mentioned above in Tables 1a–1d. The analogues comprising the test set were also obtained from different sources [29,33]. Since the experimental values of IC_{50} for these inhibitors are already available, this set of molecules provides an excellent data set for testing the prediction power of the SGB–LIE method for new ligands. Table 4 presents the free energy values estimated for the 36 test compounds for which experimental IC_{50} values were available to enable the accuracy check. The free energy values were estimated based on optimized SGB–LIE parameters α, β and γ from the training set. The overall RMSE between the experimental and predicted free energy of binding values was 0.56 kcal/mol which is comparable to the level of accuracy achieved by the most accurate methods such as free energy perturbation. The squared correlation coefficient between experimental and SGB–LIE estimates for the free energy for the test set compounds is also significant ($R^2 = 0.733$). The estimated free energy values for the test set ligands are plotted against the experimental data in Fig. 4. There is a close match between the experimental and LIE free energy values of the ligands in the test set. The predicted cytotoxic activity estimated based on LIE free energy is also very close to experimental cytotoxic activity for the test set (Table 4).

To evaluate the accuracy of the SGB–LIE estimation for tubulin polymerization inhibition potencies, we have taken a separate data set called as validation set consisting of 12 analogues of podophyllotoxin (Table 5). Colchicine and its two structural derivatives were also taken in the validation set (Table 5) in view of that these compounds also binds to tubulin in the same binding site. Their experimental activity and chemical structures were obtained from literature [34,35]. The experimental activity (IC_{50} value) of these compounds obtained from in vitro study of tubulin

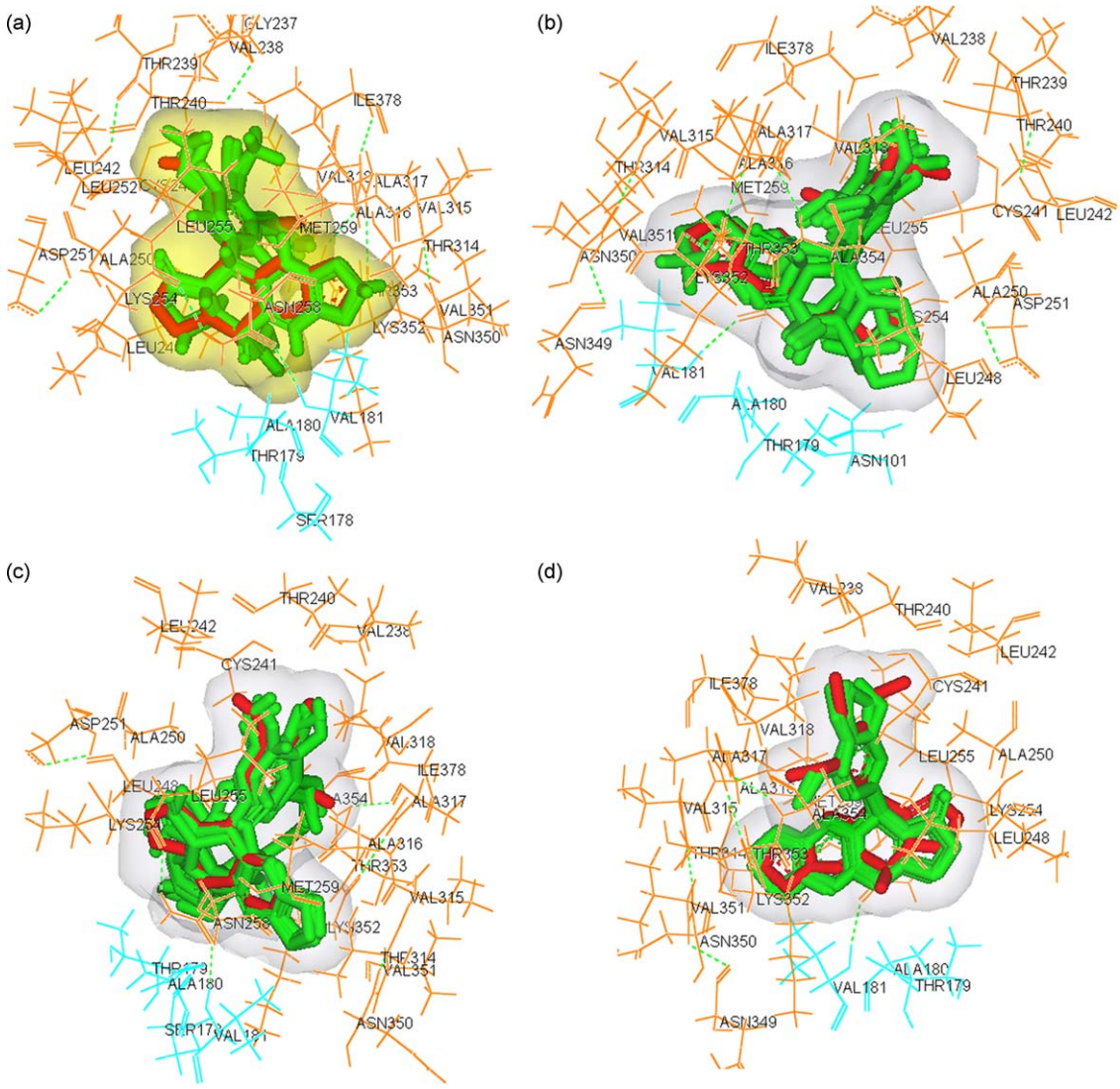


Fig. 2. (a–d). Superposition of podophyllotoxin analogues (five analogues) belonging to (a) tetraline lactones, (b) non-lactonic tetralines, (c) pyrazoline and isoxazoline derivatives and (d) aza-podophyllotoxin derivatives within binding site of tubulin along with the co-crystal podophyllotoxin (red color).

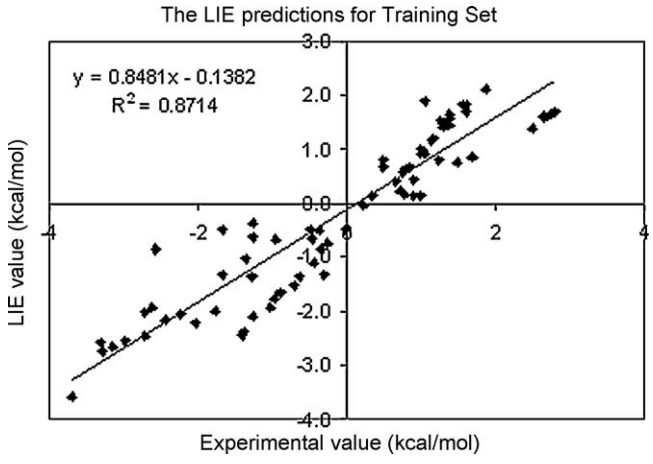


Fig. 3. Free energy values estimated by the SGB–LIE method for 76 podophyllotoxin analogues comprising the training set plotted against corresponding experimental data. The RMS error is 0.481 kcal/mol between the two data sets for 76 ligands studied here.

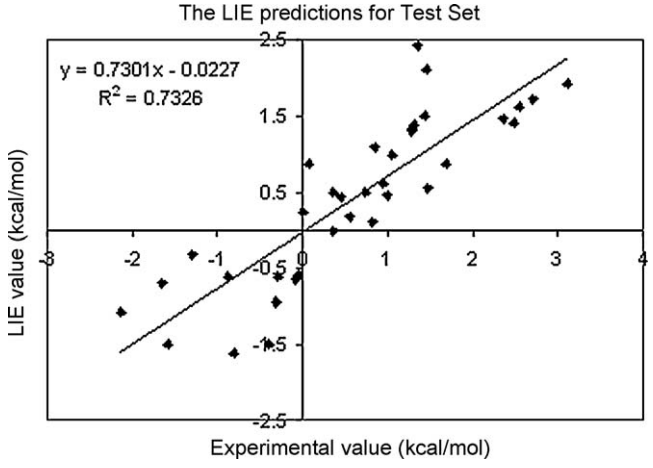


Fig. 4. Free energy values estimated by the SGB–LIE method for 36 podophyllotoxin analogues comprising the test set plotted against corresponding experimental data. The RMS error is 0.561 kcal/mol between the two data sets for 76 ligands studied here.

Table 4
Average electrostatic (ele), van der Waals (vdw) and cavity (cav) energy terms as well as binding affinity model calculations for the test set using SGB–LIE method.

Ligand	$\langle U_{ele}\rangle$[a] (kcal/mol)	$\langle U_{vdw}\rangle$[a] (kcal/mol)	$\langle U_{cav}\rangle$[a] (kcal/mol)	$pIC_{50,expt}$[b]	$\Delta G_{bind,expt}$[c] (kcal/mol)	$\Delta G_{bind,LIE}$[d] (kcal/mol)	$pIC_{50,pred}$[e]
4	11.0	−49.5	3.8	1.569	−2.1	−1.1	0.797
6	9.2	−44.9	3.3	1.217	−1.7	−0.7	0.512
10	10.9	−53.3	3.7	0.220	−0.3	−0.6	0.454
12	13.7	−53.7	2.4	−0.257	0.3	0.5	−0.366
14	11.6	−49.2	2.3	−0.323	0.4	0.4	−0.315
17	11.9	−46.7	3.0	0.638	−0.9	−0.6	0.447
20	10.6	−47.2	2.5	−0.007	0.0	0.2	−0.169
22	10.2	−52.2	3.5	0.961	−1.3	−0.3	0.238
24	13.8	−50.9	4.0	0.293	−0.4	−1.5	1.097
27	12.4	−62.7	3.3	−1.078	1.5	0.6	−0.408
31	10.3	−54.4	2.0	−1.049	1.4	1.5	−1.095
35	9.4	−54.1	2.2	−0.961	1.3	1.4	−1.009
39	10.3	−61.7	3.1	−0.763	1.0	1.0	−0.717
43	13.8	−53.0	3.4	0.066	−0.1	−0.6	0.477
47	10.0	−53.3	2.1	−0.939	1.3	1.3	−0.966
50	10.3	−46.9	1.5	−0.939	1.3	1.3	−0.948
53	12.4	−50.4	2.3	−0.535	0.7	0.5	−0.363
58	14.3	−49.0	3.9	0.587	−0.8	−1.6	1.187
61	13.5	−58.8	4.2	0.227	−0.3	−0.9	0.689
63	13.5	−58.5	3.8	0.037	−0.0	−0.6	0.428
67	12.4	−56.5	3.3	−0.257	0.3	−0.0	0.013
72	1.8	−42.8	2.5	−0.623	0.8	1.1	−0.790
76	10.6	−63.8	1.9	−0.983	1.3	2.4	−1.777
79	7.8	−49.3	2.7	−0.689	0.9	0.6	−0.446
81	9.1	−60.4	2.0	−1.071	1.5	2.1	−1.545
83	10.4	−44.8	2.4	−0.601	0.8	0.1	−0.073
86	5.2	−63.6	3.0	−2.288	3.1	1.9	−1.405
88	7.4	−59.0	2.8	−1.819	2.5	1.4	−1.030
90	5.9	−56.2	3.3	−1.232	1.7	0.9	−0.640
93	2.8	−52.7	2.8	−1.731	2.4	1.5	−1.073
96	3.6	−49.5	2.2	−1.987	2.7	1.7	−1.251
98	2.7	−47.8	2.3	−1.870	2.5	1.6	−1.187
101	2.0	−61.1	4.6	−0.733	1.0	0.5	−0.337
104	2.1	−46.0	3.5	−0.403	0.5	0.2	−0.137
106	7.7	−52.1	2.7	−0.051	0.1	0.9	−0.630
109	6.3	−52.0	5.1	1.159	−1.6	−1.5	1.099

[a] $\langle U_{ele}\rangle$, $\langle U_{vdw}\rangle$ and $\langle U_{cav}\rangle$ energy terms represents the ensemble average of the energy terms calculated as the difference between bound and free state of ligands and its environment.
[b] pIC_{50} refers to the experimental predicted cytotoxic activity using P388 cell line and is calculated as $pIC_{50} = -\log IC_{50}$.
[c] $\Delta G_{bind,expt}$ refers to free energy of binding for tubulin inhibition and is computed using the relationship: $\Delta G_{binding} \approx -2.303RTpIC_{50,expt}$, where 298 K is used in the work for temperature T.
[d] $\Delta G_{bind,LIE}$ refer to the absolute free energy values obtained using SGB–LIE method.
[e] $pIC_{50,\ pred}$ refers to predicted cytotoxic activity of ligands and is estimated using the relationship: $pIC_{50pred} = -(\Delta G_{bind,LIE}/2.303RT)$.

Table 5
The validation set along with their experimental activity expressed as the IC50 value for tubulin polymerization inhibition (TPI).

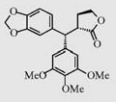
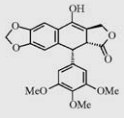
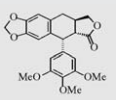
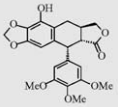
Ligand	Name	Structure	Experimental activity
1	G4		>50 M
2	Dehydropodophyllotoxin		>25 μM
3	Deoxypodophyllotoxin		0.5 μM
4	β-Peltatin		0.7 μM

Table 5 (*Continued*)

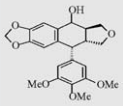
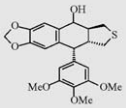
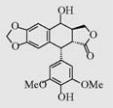
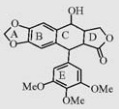
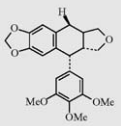
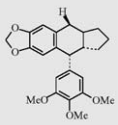
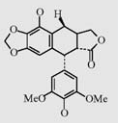
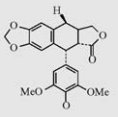
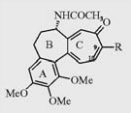
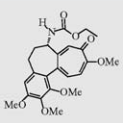
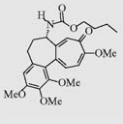
Ligand	Name	Structure	Experimental activity
5	Anhydropodophyllol		1.0 μM
6	Podophyllotoxin cyclic sulfide		10 μM
7	4′-Demethylpodophyllotoxin		0.5 μM
8	Podophyllotoxin		0.6 μM
9	Deoxypodophyllotoxin cyclic ether		0.8 μM
10	Deoxypodophyllotoxin cyclopentane		5.0 μM
11	α-Peltatine		0.5 μM
12	4′-Demethyldeoxypodophyllotoxin		0.2 μM
13	Colchicine		2.4 μM
14	3-(Ethoxycarbonyl)-3-demethylthiocolchicine		1.4 μM
15	3-(Butoxycarbonyl)-3-demethylthiocolchicine		1.4 μM

Table 6
LIE fitting, free energy values (ΔG_{bind} kcal/mol) and predicted potencies (pIC_{50}), obtained from the SGB–LIE method and experimental data for the validation set.

Ligand	$\langle U_{ele}\rangle$ (kcal/mol)[a]	$\langle U_{vdw}\rangle$ (kcal/mol)[a]	$\langle U_{cav}\rangle$ (kcal/mol)[a]	TPI_{expt} (pIC_{50} value)[b]	$\Delta G_{bind,expt}$ (kcal/mol)[c]	$\Delta G_{bind,LIE}$ (kcal/mol)[d]	TPI_{pred} (pIC_{50} value)[e]
1	13.4	−46.3	4.3	−7.699	10.5	−2.2	1.588
2	14.0	−56.3	4.1	−1.398	1.9	1.2	−0.905
3	11.8	−60.1	3.3	0.301	−0.4	−0.1	0.072
4	12.1	−61.2	3.6	0.155	−0.2	0.1	−0.078
5	10.8	−58.5	3.5	0.000	0.0	0.2	−0.117
6	8.3	−47.3	3.5	−1.000	1.4	0.9	−0.638
7	11.8	−63.7	3.5	0.301	−0.4	−0.0	0.007
8	11.6	−56.2	3.2	0.222	−0.3	−0.3	0.241
9	10.8	−57.4	4.0	0.097	−0.1	−0.4	0.292
10	10.7	−55.7	2.5	−0.699	0.9	1.0	−0.711
11	11.5	−60.7	4.0	0.301	−0.4	−0.2	0.185
12	14.1	−63.4	4.2	0.699	−0.9	−0.6	0.429
13	13.6	−53.4	6.0	−0.380	0.5	−3.4	2.474
14	15.1	−67.1	6.0	−0.146	0.2	−2.4	1.751
15	8.1	−56.3	6.1	−0.146	0.2	−2.5	1.809

[a] $\langle U_{ele}\rangle$, $\langle U_{vdw}\rangle$ and $\langle U_{cav}\rangle$ energy terms represents the ensemble average of the energy terms calculated as the difference between bound and free state of ligands and its environment.
[b] pIC_{50} refers to the experimental predicted IC_{50} value for TPI and is calculated as $pIC_{50} = -\log IC_{50}$.
[c] $\Delta G_{bind,expt}$ refers to binding free energy for tubulin–analogue interaction and is computed using the relationship: $\Delta G_{binding} \approx -2.303RTpIC_{50,expt}$, where 298 K is used in the work for temperature T.
[d] $\Delta G_{bind,LIE}$ refer to the absolute binding free energy values obtained using SGB–LIE method.
[e] $pIC_{50,pred}$ refers to predicted IC_{50} value for TPI based on SGB–LIE method and is estimated using the relationship: $pIC_{50pred} = -(\Delta G_{bind,LIE}/2.303RT)$.

polymerization inhibition (TPI). For all the compounds excluding colchicine and its two derivatives, SGB–LIE predictions produce exactly the same trend for tubulin polymerization inhibition, even though the exact magnitudes of these values do not match very well to experimental values (Table 6). Podophyllotoxin competitively inhibit the binding of colchicine to tubulin [36], implying that it bind to tubulin at the same site. The structural feature of podophyllotoxin that share with colchicine is the trimethoxyphenyl moiety. For colchicine and podophyllotoxin, it has been suggested that the binding sites for the two drugs do not completely overlap, with the trimethoxyphenyl rings of the agents binding in the same site on the tubulin heterodimer [37,38]. Harr et al. [34] suggested that the trimethoxyphenyl rings of the two drugs were situated in different regions of space, nearly orthogonal to each other. This revealed that these rings may bind to different regions of tubulin at the colchicine binding site. The RMSE between the experimental and predicted binding free energy was 1.32 kcal/mol. For compound G4 the RMSE is more than 1.29 kcal/mol. Excluding G4 from the data set the RMSE for the rest of the 11 compounds is 0.29 kcal/mol, which means that the SGB–LIE modeling was able to predict the binding free energy of the 11 compounds within 0.29 kcal/mol, which is comparable to the level of accuracy achieved by the most accurate methods, such as free energy perturbation.

4. Conclusion

We have demonstrated that the SGB–LIE method can be applied to estimate the binding free energy with a high level of accuracy for a diverse set of podophyllotoxin analogues with tubulin. The magnitude of free energy changes upon binding of these analogues to tubulin have directly correlated with the experimental potency of these inhibitors. Despite the limitation imposed by the insufficient sampling inherent in the energy minimization protocol, the method has reproduced experimental data with reasonably small error for the majority of podophyllotoxin analogues. Using LIE methodology, we have been able to verify the experimental observation that derivatized podophyllotoxin compounds, with their C-ring removed (as in G4) or unsaturated (as in dehydropodophyllotoxin), have inhibition potencies reduced. When the C-ring's substituent was removed as in deoxypodophyllotoxin or substituted to B-ring as in β-peltatin, the resulting analogues were still a potent inhibitor. This indicated that the three-dimensional conformation of the C-ring and the resulting conformational influence on the D-ring is important for interaction with tubulin. This concurs with the finding that sterioisomers like epipodophyllotoxin are much less potent. The decreased potency of lactone D-ring analogues was also usually predicted by SGB–LIE model. Few analogues with modifications on the E-ring have been tested in vitro for TPI. Removal of the 4′-methyl to give the phenol results in a small increase in potency. An increase in potency is also seen when the C-ring hydroxyl is moved to ring B: α-peltatin is slightly more potent than β-peltatin. The influences of these structural modifications were correctly predicted by SGB–LIE model developed in the study. However, the SGB–LIE predictions could not produce exactly the same trend of tubulin polymerization inhibition for the colchicine and two of its structural derivatives. This is obvious as the mode of interaction of colchicine is different at the colchicine binding site of tubulin than that of podophyllotoxin. It was suggested that the trimethoxyphenyl rings of the two drugs were bind to different regions of tubulin at the colchicine binding site. Podophyllotoxin is well known for its antitumor activity. It has better tubulin polymerization inhibition in comparison to colchicine. However, the clinical application of it and its analogues in the treatment of cancer has been limited by severe toxic side effects during administration of the drugs. With a view to achieving greater therapeutic efficiency many podophyllotoxin analogues have been isolated and via molecular manipulation, a large number of semisynthetic derivatives have been synthesized. However, new findings related to their activities, mechanism of action and pharmacological properties have been unexplored. The interaction of colchicine with tubulin is 'irreversible' and temperature-sensitive. Podophyllotoxin binds faster than colchicine and the binding is reversible and less temperature-sensitive which makes them more useful in the field of cancer therapy. The temporal and reversible binding of podophyllotoxin with tubulin overcomes the problem of inhibiting the cell multiplication of normal cell. Most of the toxic effects of the podophyllotoxin and its derivatives are due to their scant selectivity between cancer and normal cells.

Moreover, the SGB–LIE method is able to predict the binding free energy and cytotoxic activity of rationally designed podophyllotoxin congeners with relative success. The difference in the exact magnitudes of estimated vs. experimental free energy of

binding for compounds in the training set, test set and validation set may be due to the limitations imposed by inadequate sampling and force field parametrization. In addition, the calculation of absolute binding free energy from experimental IC_{50} values for cytotoxicity obtained from the in vitro cell line is only an approximation. In practical the IC_{50} value of a drug molecule is dependent upon a number of factors including solubility, membrane permeability, p-glycoprotein activity against the compound, etc. However, the SGB–LIE model developed is able to predict the binding energy of the validation set quite accurately in comparison to the binding kinetics in vitro. This may be the fact that tubulin is the most potential target for podophyllotoxin. Further, the strong relationship between the experimental and predicted FEB could be established by in vitro studies of all these podophyllotoxin analogues with isolated tubulin. A detailed study on the SARs for podophyllotoxin analogues can throw light on the moieties and functional groups important in determining the inhibition potency. The close estimation of inhibition potencies of a wide range of structural derivatives for podophyllotoxin establishes the SGB–LIE methodology as an efficient tool for screening novel compounds with very different structures. The mechanism of action of any drug is very important in drug development. Generally, the drug compound binds with a specific target, a receptor, to mediate its effects. Therefore, suitable drug–receptor interactions are required for high activity. Understanding the nature of these interactions is very significant and theoretical calculations, in particular the SGB–LIE method, seem to be a proper tool for gaining such understanding. The results obtain will give information on how the chemical structure of the drug should be modified to achieve suitable interactions and for the rapid prediction and virtual prescreening of anti-tumor activity. This will lead to new proposals regarding possible improvements to the therapeutic indices of podophyllotoxins. Compared to the empirical methods, such as scoring function approaches, the LIE method is more accurate due to the semiempirical approach adopted in which experimental data are used to build the binding affinity model. The SGB–LIE method seems promising when compared to the FEP or thermodynamic integration (TI) methods in achieving comparable accuracy with must faster speed even for structurally very different ligands.

References

[1] K.H. Downing, E. Nogales, Tubuline structure: insight into microtubule properties and functions, Curr. Opin. Struct. Biol. 8 (1998) 785–791.
[2] K.H. Downing, E. Nogales, New insight into microtubules structure and function from the atomic model of tubuline, Eur. Biophys. J. 27 (1998) 431–436.
[3] A. Jordan, J.A. Hadfield, N.J. Lawrence, A.T. McGown, Tubuline as a target of anticancer drugs: agents which interact with the mitotic spindle, Med. Res. Rev. 18 (1998) 259–296.
[4] G.M. Cragg, D.J. Newman, A tale of two tumor targets: topoisomerase I and tubuline. The Wall and Wani contribution to cancer chemotherapy, J. Nat. Prod. 67 (2004) 259–296.
[5] 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.
[6] J.A. Snyder, R.J. McIntosh, Biochemistry and physiology of microtubules, Ann. Rev. Biochem. 45 (1976) 699–720.
[7] R.L. Margolis, L. Wilson, Opposite end assembly and disassembly of microtubules at steady state in vitro, Cell 13 (1978) 1–8.
[8] F. Cortese, B. Bhattacharyya, P. Wolff, Podophyllotoxin as a probe for the colchicine binding site of tubulin, J. Biol. Chem. 252 (1977) 1134–1140.
[9] D. Ayres, C. Loike, J.D. In Lignans, Chemical Biological and Clinical Properties, Cambridge University Press, Cambridge, 1990.
[10] C.M. Lin, E. Hamel, Effects of inhibitors of tubulin polymerization on GTP hydrolysis, J. Biol. Chem. 256 (1981) 9242–9245.
[11] R.B.G. 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.
[12] J.M. Schilstra, R.S. Martin, M.P. Bayley, The effect of podophyllotoxin on microtubule dynamics, J. Biol. Chem. 264 (1989) 8827–8834.
[13] J. Åqvist, C. Medina, J.E. Samuelsson, A new method for predicting binding affinity in computer-aided drug design, Protein Eng. 7 (1994) 385–391.
[14] J. Carlsson, M. Andér, M. Nervall, J. Åqvist, Continuum solvation models in the linear interaction energy method, J. Phys. Chem. B 110 (24) (2006) 12034–12041.
[15] R. Zhou, R.A. Frienser, A. Ghosh, R.C. Rizzo, W.L. Jorgensen, R.M. Levy, New linear interaction method for binding affinity calculations using a continuum solvent model, J. Phys. Chem. B. 105 (2001) 10388–10397.
[16] T. Hansson, J. Åqvist, Estimation of binding free energies for HIV proteinase inhibitors by molecular dynamics simulations, Protein Eng. 8 (1995) 1137–1145.
[17] FirstDiscovery 2.7, 2.7 ed.; Schrödinger Inc.: Portland, 2004.
[18] Liaison, version 4.0, Schrödinger, LLC, New York, NY, 2005.
[19] Schrodinger L. L.C., http://www.schrodinger.com (accessed: 24. 04. 2007).
[20] I. Jardine, in: J.M. Cassady, J.D. Douras (Eds.), Anticancer Agents on Natural Product Models, Academic Press, New York, 1980, , Chapter 9.
[21] S.G. Weiss, W. Tin-Wa, A. von Wartburg, H.J. Stahelin, Potential anticancer agents II: antitumor and cytotoxic lignans from *Linum album* (Linaceae), J. Pharm. Sci. 64 (1975) 95–98.
[22] A. San Feliciano, J.M. Miguel del Corral, M. Gordaliza, M.A. Castro, Acetylated lignans form *Juniperus sabina*, Phytochemistry 28 (1989) 659–660.
[23] A. San Feliciano, M. Gordaliza, J.M. Miguel del Corral, M.A. Castro, M.D. García-Grávalos, P. Ruiz- Lázaro, Antineoplastic and antiviral activities of some cyclolignans, Planta Med. 59 (1993) 246–249.
[24] J.M. Miguel del Corral, M. Gordaliza, M.A. Castro, L.J. Morales, J.L. López, A. San Feliciano, Methyl ethers of podophyllotoxin-related cyclolignans, J. Nat. Prod. 58 (1995) 870–877.
[25] J.C. Doré, C. Viel, N. Pageot, M. Gordaliza, M.A. Castro, J.M. Miguel del Corral, A. San Feliciano, Multivariate analysis approach to antineoplastic and antiviral structure-activity relationships to a series of podophyllotoxins, J. Pharm. Belg. 51 (1996) 9–18.
[26] M. Gordaliza, G.T. Faircloth, M.A. Castro, J.M. Miguel del Corral, M.L. López-Vázquez, A. San Feliciano, Immunosuppressive cyclolignans, J. Med. Chem. 39 (1996) 2865–2868.
[27] M. Gordaliza, J.M. Miguel del Corral, M.A. Castro, M.L. López-Vázquez, A. San Feliciano, M.D. García Grávalos, A. Carpy, Synthesis and evaluation of pyrazolignans. A new class of cytotoxic agents, Bioorg. Med. Chem. 3 (1995) 1203–1210.
[28] M. Gordaliza, J.M. Miguel del Corral, M.A. Castro, M.D. García-Grávalos, H. Broughton, A. San Feliciano, Cytotoxic cyclolignans related to podophyllotoxin, II Farmaco 56 (2001) 297–304.
[29] Y. Hitotsuyanagi, M. Fukuyo, K. Tsuda, M. Kobayashi, A. Ozeki, H. Itokawa, K. Takeya, 4-Aza-2, 3-dehydro-4-deoxypodophyllotoxin: simple aza-podophyllotoxin analogues possessing potent cytotoxicity, Bioorg. Med. Chem. Lett. 10 (2006) 315–317.
[30] M. Marty, P. Fumoleau, A. Adenis, Y. Rousseau, Y. Merrouche, G. Robinet, I. Senac, C. Puozzo, Oral vinorelbine pharmacokinetics and absolute bioavailability study in patients with solid tumors, Ann. Oncol. 12 (11) (2001) 1643–1649.
[31] A. Aggarwal, A. Kruczynski, A. Frankfurter, J.J. Correia, S. Lobert, Murine leukemia P388 vinorelbine-resistant cell lines are sensitive to vinflunine, Invest. New Drugs 26 (2008) 319–330.
[32] G.B. Thomas, R.L. Finny, In Calculus and Analytic Geometry, 9th ed., Addison-Wesley, 2001.
[33] M. Gordaliza, M.A. Castro, J.M. Miguel del Corral, A. San Feliciano, Antitumor properties of podophyllotoxin and related compound, Curr. Pharm. Des. 6 (2000) 1811–1839.
[34] E.T. Haar, H.S. Rosenkranz, E. Hamel, B.W. Day, Computational and molecular modeling evaluation of structural for tubuline polymerization inhibition by colchicine site agent, Bioorg. Med. Chem. 10 (1996) 1659–1671.
[35] J.D. Lokie, C.F. Brewer, H. Sternlicht, W.J. Gensler, S.B. Horwitz, Structure activity study of the inhibition of microtubule assembly in vitro by podophyllotoxin and its congeners, Cancer Res. 38 (1978) 2688–2693.
[36] S.B. Hastie, Interactions of colchicine with tubulin, Pharm. Ther. 51 (1991) 377–401.
[37] J.M. Andreu, S.N. Timasheff, Interaction of tubulin with single ring analogues of colchicine, Biochemistry 21 (3) (1982) 534–543.
[38] J.M. Andreu, S.N. Timasheff, Conformational states of tubulin liganded to colchicine, tropolone methyl ether, and podophyllotoxin, Biochemistry 21 (1982) 6465–6476.