molecular pharmaceutics

articles

Overcoming Tumor Drug Resistance with C2-Modified 10-Deacetyl-7-propionyl Cephalomannines: A QSAR Study

Corwin Hansch and Rajeshwar P. Verma*

Department of Chemistry, Pomona College, 645 North College Avenue, Claremont, California 91711

Received August 14, 2008; Revised Manuscript Received January 13, 2009; Accepted March 4, 2009

Abstract: The microtubule-stabilizing taxanes such as paclitaxel and docetaxel are the two most important anticancer drugs currently used in clinics for the treatment of various types of cancers. However, the major common drawbacks of these two drugs are drug resistance, neurotoxicity, substrate for drug transporter P-gp, cross-resistance with other chemotherapeutic agents, low oral bioavailability, and no penetration in the blood−brain barrier (BBB). These limitations have led to the search for new taxane derivatives with improved biological activity. In the present paper, we discuss the quantitative structure−activity relationship (QSAR) studies on a series of C2-modified 10-deacetyl-7-propionyl cephalomannines (**IV**) with respect to their binding affinities toward β-tubulin and cytotoxic activities against both drug-sensitive and drug-resistant tumor cells, in which resistance is mediated through either P-gp overexpression or β-tubulin mutation mechanisms, by the formulation of five QSARs. Hydrophobicity and molar refractivity of the substituents (π_X and MR_X) are found to be the most important determinants for the activity. Parabolic correlations in terms of MR_X (eqs 2 and 4) are encouraging examples in which the optimum values of MR_X are well-defined. We believe that these two QSAR models may prove to be adequate predictive models that can help to provide guidance in design and synthesis, and subsequently yield very specific cephalomannine derivatives (**IV**) that may have high biological activities. On the basis of these two QSAR models, 10 cephalomannine analogues (**IV-21** to **IV-30**) are suggested as potential synthetic targets. Internal (cross-validation (q^2), quality factor (Q), Fischer statistics (F), and Y-randomization) and external validation tests have validated all the QSAR models.

Keywords: Cephalomannines; drug-resistant; hydrophobicity; molar refractivity; QSAR; β-tubulin

Introduction

Cancer has become a major public health problem in the United States and various parts of the world. Currently, it is responsible for about 25% of deaths in the United States. A very recent estimate, conducted by the American Cancer Society, indicates that at least 1 437 180 new cancer cases and 565 650 deaths from cancers are expected to occur in the United States in 2008.[1]

It has been established that tubulin-binding agents are among the most important drugs used for the treatment of a wide variety of cancers.[2] The microtubule-stabilizing taxanes such as paclitaxel (PTX, Taxol; **I**) and docetaxel (DTX, Taxotere; **II**) are the two most important anticancer drugs currently used in clinics for the treatment of various types of cancers, including breast, colon, lung, ovarian, and

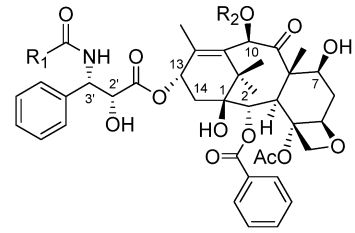
* Corresponding author. Mailing address: Department of Chemistry, Pomona College, 645 N. College Ave., Claremont, CA 91711. Telephone: 909-607-4249. Fax: 909-607-7726. E-mail: rverma@pomona.edu.

(1) Jemal, A.; Siegel, R.; Ward, E.; Hao, Y.; Xu, J.; Murray, T.; Thun, M. J. Cancer statistics, 2008. *Ca Cancer J. Clin.* **2008**, *58*, 71–96.

(2) Rowinsky, E. K.; Tolcher, A. W. Antimicrotubule Agents. In *Cancer Principles and Practice of Oncology*; DeVita, V. T. J., Hellman, S., Rosenberg, S. A., Eds.; Lippincott, Williams and Wilkins: Philadelphia, PA, 2001; pp 431−447.

10.1021/mp800138w CCC: $40.75
Published on Web 03/31/2009

prostate.[3−6] These drugs bind to the β-subunit of tubulin polymer in a stoichiometric ratio and promote tubulin polymerization. This phenomenon disrupts tubulin polymerization dynamics, leading to cell cycle arrest and ultimately cell death by apoptosis.[7−10] However, the major common drawbacks of these two drugs are drug resistance, neurotoxicity, substrate for drug transporter P-gp, cross-resistance against different chemotherapeutic agents (such as vinblastin, etoposide, and doxorubicin), low oral bioavailability, and no penetration in blood−brain barrier (BBB). The two main mechanisms involved in the resistance to taxanes are (a) the expression of the multidrug resistance (MDR) phenotype and (b) the alterations of their cellular target, namely, the tubulin/microtubule system.[11,12] To prevent these disadvantages and improve the clinical application of taxanes, several new agents have entered into the clinical trials. At present, a large number of microtubule-stabilizing agents are in various stages of clinical trials including discodermolides, eleutherobin, epothilones, laulimalides, peloruside, sarcodictyins, and taxane derivatives that have unique features to bind the taxane site of β-tubulin.[12−14]



R_1 = Ph, R_2 = Ac; Paclitaxel (PTX, Taxol; **I**)

R_1 = ^{t}BuO, R_2 = H; Docetaxel (DTX, Taxotere; **II**)

R_1 = $CH_3CH{=}C(CH_3)$, R_2 = Ac; Cephalomannine (NSC 318735, Taxol B; **III**)

Cephalomannine (**III**) is a natural congener of paclitaxel (**I**), which was isolated in the 1970s from leaves, stems, and roots of *Taxus wallichiana* Zucc.[15] It is a compound of great interest due to its similarity in both structure and antitumor spectrum to those of paclitaxel. In addition, it is also known that some of the 3′-*N*-acyl analogues of **III** exhibit slightly better activity than paclitaxel itself.[16] In the present paper, we demonstrate the quantitative structure−activity relationship (QSAR) studies on a series of C2-modified 10-deacetyl-7-propionyl cephalomannines (**IV**) with respect to their cytotoxic activities against both drug-sensitive and drug-resistant tumor cells in which resistance is mediated through either P-gp overexpression or β-tubulin mutation mechanisms. Binding affinities of these cephalomannines (**IV**) toward β-tubulin has also been used to develop a QSAR model to understand their chemical−biological interactions in mathematical terms.

Since the advent of the QSAR methodology (one of the well-developed areas in computational chemistry) about 46 years ago,[17] it has become increasingly helpful in understanding many aspects of chemical−biological interactions in drug-design process and pesticide research as well as in the areas of toxicology.[18−22] This method is useful in

(3) Verma, R. P.; Hansch, C. Taxane analogues against breast cancer: a quantitative structure-activative relationship study. *ChemMedChem* **2008**, *3*, 642–652.

(4) Roh, E. J.; Kim, D.; Lee, C. O.; Choi, S. U.; Song, C. E. Structure−activity relationship study at the 3′-*N*-position of paclitaxel: Synthesis and biological evaluation of 3′-*N*-acyl-paclitaxel analogues. *Bioorg. Med. Chem.* **2002**, *10*, 3145–3151.

(5) Ojima, I.; Slater, J. C.; Kuduk, S. D.; Takeuchi, C. S.; Gimi, R. H.; Sun, C.-M.; Park, Y. H.; Pera, P.; Veith, J. M.; Bernacki, R. J. Syntheses and Structure−Activity Relationships of Taxoids Derived from 14β-Hydroxy-10-deacetylbaccatin III. *J. Med. Chem.* **1997**, *40*, 267–278.

(6) Arbuck, S. G.; Blaylock, B. A. Taxol: Clinical Results and Current Issues in Development. In *Taxol: Science and Applications*; Suffness, M., Ed.; CRC Press: Boca Raton, FL, 1995; pp 379−415.

(7) Hansch, C.; Verma, R. P. Understanding tubulin/microtubule-taxane interactions: A Quantitative Structure−Activity Relationship Study. *Mol. Pharmaceutics* **2008**, *5*, 151–161.

(8) Wang, T.-H.; Wang, H.-S.; Song, Y.-K. Paclitaxel-induced cell death: where the cell cycle and apoptosis come together. *Cancer* **2000**, *88*, 2619–2628.

(9) Blagosklonny, M. V.; Fojo, T. Molecular effects of paclitaxel: myths and reality (a critical review). *Int. J. Cancer* **1999**, *83*, 151–156.

(10) Schiff, P. B.; Fant, J.; Horwitz, S. B. Promotion of microtubule assembly in vitro by taxol. *Nature (London)* **1979**, *277*, 665–667.

(11) Galletti, E.; Magnani, M.; Renzulli, M. L.; Botta, M. Paclitaxel and docetaxel resistance: Molecular mechanisms and development of new generation taxanes. *ChemMedChem* **2007**, *2*, 920–942.

(12) Kuppens, I. E. L. M. Current state of the art of new tubulin inhibitors in the clinic. *Curr. Clin. Pharmacol.* **2006**, *1*, 57–70.

(13) He, L.; Orr, G. A.; Horwitz, S. B. Novel molecules that interact with microtubules and have functional activity similar to taxol. *DDT* **2001**, *6*, 1153–1164.

(14) Pineda, O.; Farràs, J.; Maccari, L.; Manetti, F.; Botta, M.; Vilarrasa, J. Computational comparison of microtubule-stabilising agents laulimalide and peloruside with taxol and colchicine. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4825–4829.

(15) Miller, R. W.; Powell, R. G.; Smith, C. R., Jr.; Arnold, E.; Clardy, J. Antileukemic Alkaloids from *Taxus wallichiana* Zucc. *J. Org. Chem.* **1981**, *46*, 1469–1474.

(16) Baloglu, E.; Hoch, J. M.; Chatterjee, S. K.; Ravindra, R.; Bane, S.; Kingston, D. G. I. Synthesis and biological evaluation of C-3′NH/C-10 and C-2/C-10 modified paclitaxel analogues. *Bioorg. Med. Chem.* **2003**, *11*, 1557–1568.

(17) Hansch, C.; Maloney, P. P.; Fujita, T.; Muir, R. M. Correlation of biological activity of phenoxyacetic acids with Hammett substituent constants and partition coefficients. *Nature (London)* **1962**, *194*, 178–180.

(18) Hansch, C.; Leo, A. In *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*; American Chemical Society: Washington, DC, 1995.

(19) Bermudez-Saldana, J. M.; Cronin, M. T. D. Quantitative structure-activity relationships for the toxicity of organophosphorus and carbamate pesticides to the rainbow trout Onchorhynchus mykiss. *Pest Manage. Sci.* **2006**, *62*, 819–831.

(20) Hansch, C.; Telzer, B. R.; Zhang, L. Comparative QSAR in toxicology: examples from teratology and cancer chemotherapy of aniline mustards. *Crit. Rev. Toxicol.* **1995**, *25*, 67–89.

elucidating the mechanisms of chemical−biological interaction in various biomolecules, particularly enzymes, membranes, organelles, and cells, as well as in virus, bacteria, and human.[18,23−32] It has also been utilized for the evaluation of absorption, distribution, metabolism, and excretion (ADME) phenomena in many organisms and whole animal studies.[33,34] The QSAR approach employs extra-thermodynamically derived and computational-based descriptors to correlate biological activity in isolated receptors, in cellular systems, and in vivo. Four standard molecular descriptors are routinely used in QSAR analysis: electronic, hydrophobic, steric, and topological indices. These descriptors are invaluable in helping to delineate a large number of receptor−ligand interactions that are critical to biological processes. The quality of a QSAR model depends strictly on the type and quality of the data, and not on the hypotheses, and is valid only for the compound structure analogues to those used to build the model. QSAR models can stand alone to augment other computational approaches or can be examined in tandem with equations of a similar mechanistic genre to establish their authenticity and reliability.[35]

Materials and Methods

IC_{50} and K_d data of cephalomannine derivatives (**IV**) have been collected from the literature[36] and converted into molar concentration. IC_{50} is the molar concentration of a cephalomannine derivative (**IV**) that inhibits 50% of growth of both drug-sensitive and drug-resistant tumor cells in which resistance is mediated through either P-gp overexpression or β-tubulin mutation mechanisms.[36] K_d is the equilibrium binding constant of a cephalomannine derivative (**IV**) to β-tubulin at 35 °C; log $1/IC_{50}$ and log $1/K_d$ are the subsequent dependent variables that define the biological parameters for QSAR development. Physicochemical descriptors are autoloaded, and multiregression analyses (MRA) are used to derive the QSAR by using the C-QSAR program.[37] Selection of descriptors is made on the basis of permutation and correlation matrices among the descriptors in order to avoid collinearity problems. Details about the C-QSAR program, the search engine, the choice of parameters, and their use in the development of QSAR models have already been discussed.[38,39] The descriptors used in this paper have been discussed previously in detail along with their application.[18] Briefly, *C* log *P* is the calculated partition coefficient of a compound in *n*-octanol/water and is a measure of its hydrophobicity, whereas π is the hydrophobic parameter for substituents only. CMR is the calculated molar refractivity for the whole molecule, which is calculated from the Lorentz−Lorenz equation: $[(n^2 - 1)/(n^2 + 2)](MW/\delta)$, where n is the refractive index, MW is the molecular weight, and δ is the density of the substance.[18,40−42] Molar refractivity (MR) is dependent on both the volume and polarizability. It

(21) Schultz, T. W.; Hewitt, M.; Netzeva, T. I.; Cronin, M. T. D. Assessing applicability domains of toxicological QSARs: definition, confidence in predicted values, and the role of mechanisms of action. *QSAR Comb. Sci.* **2007**, *26*, 238–254.

(22) Selassie, C. D.; Garg, R.; Kapur, S.; Kurup, A.; Verma, R. P.; Mekapati, S. B.; Hansch, C. Comparative QSAR and the Radical Toxicity of Various Functional Groups. *Chem. Rev.* **2002**, *102*, 2585–2605.

(23) Winkler, D. A. The role of quantitative structure-activity relationships (QSAR) in biomolecular discovery. *Briefings Bioinf.* **2002**, *3*, 73–86.

(24) Verma, R. P.; Hansch, C. Matrix metalloproteinases (MMPs): Chemical-biological functions and QSARs. *Bioorg. Med. Chem.* **2007**, *15*, 2223–2268.

(25) Fujikawa, M.; Nakao, K.; Shimizu, R.; Akamatsu, M. QSAR study on permeability of hydrophobic compounds with artificial membranes. *Bioorg. Med. Chem.* **2007**, *15*, 3756–3767.

(26) Verma, R. P.; Hansch, C.; Selassie, C. D. Comparative QSAR studies on PAMPA/modified PAMPA for high throughput profiling of drug absorption potential with respect to Caco-2 cells and human intestinal absorption. *J. Comput.-Aided. Mol. Des.* **2007**, *21*, 3–22.

(27) Kodithala, K.; Hopfinger, A. J.; Thompson, E. D.; Robinson, M. K. Prediction of skin irritation from organic chemicals using membrane - interaction QSAR analysis. *Toxicol. Sci.* **2002**, *66*, 336–346.

(28) Verma, R. P. Cytotoxicity of heterocyclic compounds against various cancer cells: a quantitative structure-activity relationship study. *Top. Heterocycl. Chem.* **2007**, *9*, 53–86.

(29) Garg, R.; Gupta, S. P.; Gao, H.; Babu, M. S.; Debnath, A. K.; Hansch, C. Comparative Quantitative Structure−Activity Relationship Studies on Anti-HIV Drugs. *Chem. Rev.* **1999**, *99*, 3525–3601.

(30) Verma, R. P.; Hansch, C. Combating the Threat of Anthrax: A Quantitative Structure−Activity Relationship Approach. *Mol. Pharmaceutics* **2008**, *5*, 745–759.

(31) Verma, R. P.; Kurup, A.; Mekapati, S. B.; Hansch, C. Chemical-biological interactions in human. *Bioorg. Med. Chem.* **2005**, *13*, 933–948.

(32) Hall, L. H.; Hall, L. M. QSAR modeling based on structure-information for properties of interest in human health. *SAR QSAR Environ. Res.* **2005**, *16*, 13–41.

(33) Hansch, C.; Leo, A.; Mekapati, S. B.; Kurup, A. QSAR and ADME. *Bioorg. Med. Chem.* **2004**, *12*, 3391–3400.

(34) Khan, M. T. H.; Sylte, I. Predictive QSAR modeling for the successful predictions of the ADMET properties of candidate drug molecules. *Curr. Drug Discovery Technol.* **2007**, *4*, 141–149.

(35) Selassie, C. D.; Mekapati, S. B.; Verma, R. P. QSAR: then and now. *Curr. Top. Med. Chem.* **2002**, *2*, 1357–1379.

(36) Yang, C.-G.; Barasoain, I.; Li, X.; Matesanz, R.; Liu, R.; Sharom, F. J.; Yin, D.-L.; Díaz, J. F.; Fang, W.-S. Overcoming tumor drug resistance with high-affinity taxanes: A SAR study of C2-modified 7-acyl-10-deacetyl cephalomannines. *ChemMedChem* **2007**, *2*, 691–701.

(37) *C-QSAR Program*; BioByte Corp.: Claremont, CA; www.biobyte.com

(38) Hansch, C.; Hoekman, D.; Leo, A.; Weininger, D.; Selassie, C. D. Chem-Bioinformatics: Comparative QSAR at the Interface between Chemistry and Biology. *Chem. Rev.* **2002**, *102*, 783–812.

(39) Verma, R. P.; Hansch, C. Development of QSAR models using C-QSAR program: a regression program that has dual databases of over 21,000 QSAR models. *Nat. Protoc.*, http://dx.doi.org/10.1038/nprot.2007.125.

(40) Lorentz, H. A. *The Theory of Electrons*, 2nd ed.; Dover: New York, 1952.

(41) Partington, J. R. *An Advanced Treatise on Physical Chemistry*; Longmans, Green & Co. Ltd: New York, 1953; Vol. IV: Physico-Chemical Properties.

Table 1. Biological and Physicochemical Parameters Used to Derive QSARs 1, 2, and 3

		log 1/K_d (eq *1*)			log 1/IC_{50} (eq *2*)			log 1/IC_{50} (eq *3*)				
compd	*X*	obsd	pred	Δ	obsd	pred	Δ	obsd	pred	Δ	π_X	MR_X
IV-1	Ph	6.62	6.28	0.34	7.88	7.28	0.60	5.91	5.81	0.10	2.02	2.51
IV-2[a,c]	3-N_3-Ph	9.00	7.04	1.96	8.52	8.17	0.35	7.77	6.81	0.96	2.87	3.26
IV-3[a]	3-OMe-Ph	7.72	6.69	1.03	8.08	7.56	0.52	6.80	6.35	0.45	2.20	3.13
IV-4	3-CHO-Ph	5.84	6.38	−0.54	7.16	7.00	0.16	5.82	5.95	−0.13	1.60	3.01
IV-5	3-CH_2OH-Ph	ND[d]	6.20	ND[d]	5.77	6.38	−0.61	ND[d]	5.71	ND[d]	0.98	3.13
IV-6	3-Cl-Ph	7.21	6.84	0.37	7.88	8.10	−0.22	6.56	6.55	0.01	2.73	3.00
IV-7	3-Br-Ph	7.08	7.06	0.02	7.54	8.17	−0.63	6.71	6.84	−0.13	2.88	3.29
IV-8[c]	3-I-Ph	7.24	7.46	−0.22	7.52	7.97	−0.45	6.61	7.36	−0.75	3.14	3.82
IV-9[b]	3-CN-Ph	6.59	6.31	0.28	7.82	6.86	0.96	5.68	5.86	−0.18	1.45	2.99
IV-10	3-Me-Ph	6.69	6.73	−0.04	7.81	7.90	−0.09	6.22	6.41	−0.19	2.52	2.97
IV-11	2-thiophene	6.38	6.06	0.32	7.21	6.87	0.34	5.49	5.53	−0.04	1.74	2.32
IV-12	3-thiophene	5.99	6.06	−0.07	6.97	6.87	0.10	5.53	5.53	0.00	1.74	2.32
IV-13	2-chromone	5.32	5.31	0.01	4.85	5.06	−0.21	ND[d]	4.54	ND[d]	−0.09	2.30
IV-14	CH=C(Me)$_2$	5.63	5.86	−0.23	6.02	6.46	−0.44	4.99	5.27	−0.28	1.83	1.91
IV-15	CH_2C(Me)=CH_2	5.86	5.81	0.05	6.01	6.32	−0.31	5.40	5.20	0.20	1.81	1.83
IV-16	CH=CHMe	5.45	5.44	0.01	5.52	5.24	0.28	4.82	4.72	0.10	1.43	1.44
IV-17[b]	CH2CH(Me)$_2$	5.62	5.94	−0.32	5.31	6.64	−1.33	ND[d]	5.37	ND[d]	2.09	1.86
IV-18	2,4-di-F-Ph	6.24	6.23	0.01	7.09	7.15	−0.06	5.73	5.75	−0.02	1.87	2.54
IV-19[a−c]	2,4-di-Cl-Ph	6.40	7.22	−0.82	6.99	8.16	−1.17	6.16	7.05	−0.89	3.01	3.49
IV-20	2,5-di-OMe-Ph	6.94	6.96	−0.02	7.60	6.95	0.65	6.82	6.71	0.11	2.01	3.75

[a] Not used in the derivation of QSAR 1. [b] Not used in the derivation of QSAR 2. [c] Not used in the derivation of QSAR 3. [d] Not determined.

can be used for a substituent or for the whole molecule. MR is thus a means of characterizing the bulk and polarizability of a substituent or compound. Although it contains no information about the shape, it has found considerable use in biological QSAR, where intermolecular effects predominate. MR is scaled at 0.1 to make it equiscalar with π.[18,42] The successful use of MR in QSAR studies can be seen in various publications.[3,7,18,24,29,42] The indicator variable *I* is assigned the value of 1 or 0 for special features with special effects that cannot be parametrized, and has been explained wherever used.

In QSAR equations, *n* is the number of data points, *r* is the correlation coefficient between observed values of the dependent and the values calculated from the equation, r^2 is the square of the correlation coefficient and represents the goodness of fit, q^2 is the cross-validated r^2 (a measure of the quality of the QSAR model), and *s* is the standard deviation. The cross-validated r^2 (q^2) is obtained by using a leave-one-out (LOO) procedure as described by Cramer et al.[43] *Q* is the quality factor, for which $Q = r/s$ (where *r* is the correlation coefficient and *s* is the standard deviation). Chance correlation, due to the excessive number of descriptors (which also increases the *r* and *s* values), can, therefore, be detected by the examination of the *Q* value. High values of *Q* indicate the high predictive power of the QSAR models and the lack of "overfitting". *F* represents the Fischer statistics (Fischer ratio), $F = fr^2/[(1 - r^2)m]$, where *f* is the number of degrees of freedom $[f = n - (m + 1)]$, *n* is the number of data points, and *m* is the number of variables. The *F* value is actually the ratio between explained and unexplained variance for a given number of degrees of freedom. Thus, it indicates a true relationship, or the significance level for MLR models (within parentheses, the figure in each equation refers to the literature *F*-value at 99% level[44]). The modeling was taken to be optimal when *Q* reached a maximum together with *F*, even if slightly nonoptimal *F* values have normally been accepted. A significant decrease in *F* with the introduction of one additional variable (with increasing *Q* and decreasing *s*) could mean that the new descriptor is not as significant as expected; that is, its introduction has endangered the statistical quality of the combination. However, the statistical quality could be improved by the introduction of a more convincing descriptor.[45,46] Compounds were deemed to be outliers on the basis of their deviation between observed and predicted activities from the equation (obsd − pred > 2*s*, where *s* is the standard deviation).[47−51] Outliers are those compounds which have unexpected biological activities and are unable to fit in a QSAR model. The presence of outliers is not only

(42) Selassie, C. D. In *Burger's Medicinal Chemistry and Drug Descovery: History of Quantitative Structure−Activity Relationships*, 6th ed.; Abraham, D. J., Ed.; John Wiley & Sons: New York, 2003; Vol. 1: Drug Discovery.

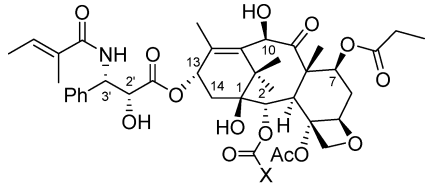
(43) Cramer, R. D., III.; Bunce, J. D.; Patterson, D. E.; Frank, I. E. Cross validation, Bootstrapping and partial least squares compared with multiple regression in conventional QSAR studies. *Quant. Struct.−Act. Relat.* **1988**, *7*, 18–25.

(44) Bennett, C. A.; Franklin, N. L. *Statistical Analysis in Chemistry and the Chemical Industry*. John Wiley & Sons: New York, 1967; pp 708−709; 5th printing.

due to the possibility that the molecules may act by different mechanisms or interact with the receptor in different binding modes but also due to the intrinsic noise associated with both the original data and methodological aspects involved in the construction of a QSAR model.[52−54] Each regression equation includes 95% confidence limits for each term in parentheses.

Results and Discussion

1. Binding Affinity of 2-OCOX-10-Deacetyl-7-propionyl Cephalomannines (IV) to β-Tubulin.[36] (Table 1)



IV

$$\log 1/K_d = 0.40(\pm 0.23)\pi_X + 0.56(\pm 0.25)MR_X + 4.06(\pm 0.60)$$
$$n = 16,\ r^2 = 0.848,\ s = 0.269,\ q^2 = 0.792,\ Q = 3.424,\ F_{2,13} = 36.263(6.701) \quad (1)$$
$$\pi_X \text{ vs } MR_X,\ r = 0.495$$
range in $\log 1/K_d = 5.32 - 9.00$
outliers: X = 3-N_3-Ph; 3-OMe-Ph; 2, 4-di-Cl-Ph

In this equation, π_X and MR_X are the calculated hydrophobicity and molar refractivity of the X substituents, respectively. Positive π_X suggests that binding affinity of cephalomannine derivatives (**IV**) to β-tubulin increases with an increase in hydrophobicity of X substituents. This may be due to the interaction between X substituents of **IV** and the hydrophobic pocket of the protein (β-tubulin). It may correspond to the recent pharmacophore studies on paclitaxel,

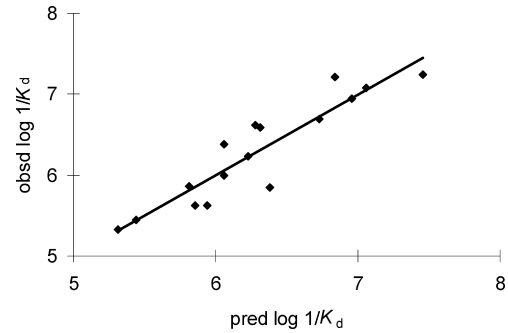


Figure 1. Plot of observed versus predicted log $1/K_d$ from eq 1.

which suggested that the phenyl ring of the 2-benzoate group is close to His227/His229 of β-tubulin and is flanked by a "hydrophobic basin" composed of Phe272, Ala273, Ser374 (the position of this residue is replaced by Ala in the majority of the human β-tubulin isoform), and other residues, thus leading to strong hydrophobic interactions between paclitaxel (cephalomannines **IV** in the present case) and β-tubulin.[36,55] The increase in the molar refractivity of X substituents (MR_X) also increases the tubulin binding affinity (positive coefficient). Thus, the binding affinity of cephalomannine derivatives (**IV**) to β-tubulin depends on both the hydrophobic and molar refractivity descriptors of their X substituents. A comparison between observed and predicted values of log $1/K_d$ for cephalomannine analogues (**IV**) used in the development of eq 1 is shown in Figure 1.

2. Cytotoxicity of 2-OCOX-10-Deacetyl-7-propionyl Cephalomannines (IV) against A2780 (Parental Ovarian Carcinoma Cells).[36] (Table 1)

$$\log 1/IC_{50} = 0.97(\pm 0.38)\pi_X + 4.16(\pm 2.85)MR_X - 0.71(\pm 0.54)MR_X^2 - 0.69(\pm 3.76)$$
$$n = 17,\ r^2 = 0.836,\ s = 0.462,\ q^2 = 0.661,\ Q = 1.978,\ F_{3,13} = 22.089(5.739) \quad (2)$$
optimum $MR_X = 2.95(2.64 - 4.12)$
π_X vs MR_X, $r = 0.514$
range in $\log 1/IC_{50} = 4.85 - 8.52$
outliers: X = 3-CN-Ph; $CH_2CH(Me)_2$; 2, 4Cl-Ph

This is a parabolic correlation in terms of MR_X, which suggests that the cytotoxic activities of cephalomannine derivatives (**IV**) first increase with an increase in the molar refractivity of X substituents up to an optimum MR_X value

(45) Pogliani, L. Modeling with Special Descriptors Derived from a Medium-Sized Set of Connectivity Indices. *J. Phys. Chem.* **1996**, *100*, 18065–18077.
(46) Pogliani, L. From Molecular Connectivity Indices to Semiempirical Connectivity Terms: Recent Trends in Graph Theoretical Descriptors. *Chem. Rev.* **2000**, *100*, 3827–3858.
(47) Selassie, C. D.; Kapur, S.; Verma, R. P.; Rosario, M. Cellular Apoptosis and Cytotoxicity of Phenolic Compounds: A Quantitative Structure−Activity Relationship Study. *J. Med. Chem.* **2005**, *48*, 7234–7242.
(48) Verma, R. P.; Hansch, C. Cytotoxicity of Organic Compounds against Ovarian Cancer Cells: A Quantitative Structure−Activity Relationship Study. *Mol. Pharmaceutics* **2006**, *3*, 441–450.
(49) Verma, R. P.; Hansch, C. Understanding human rhinovirus infections in terms of QSAR. *Virology* **2007**, *359*, 152–161.
(50) Hansch, C.; Verma, R. P. 20-(*S*)-Camptothecin analogues as DNA topoisomerase I inhibitors: A QSAR study. *ChemMedChem* **2007**, *2*, 1807–1813.
(51) Verma, R. P.; Hansch, C. Investigation of DNA-binding properties of organic molecules using quantitative structure−activity relationship (QSAR) models. *J. Pharm. Sci.* **2008**, *97*, 88–110.
(52) Verma, R. P.; Hansch, C. An approach toward the problem of outliers in QSAR. *Bioorg. Med. Chem.* **2005**, *13*, 4597–4621.
(53) Polanski, J.; Bak, A.; Gieleciak, R.; Magdziarz, T. Modeling Robust QSAR. *J. Chem. Inf. Model.* **2006**, *46*, 2310–2318.
(54) Ekins, S.; Mestres, J.; Testa, B. In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. *Br. J. Pharmacol.* **2007**, *152*, 9–20.
(55) Snyder, J. P.; Nettles, J. H.; Cornett, B.; Downing, K. H.; Nogales, E. The binding conformation of Taxol in β-tubulin: A model based on electron crystallographic density. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5312–5316.

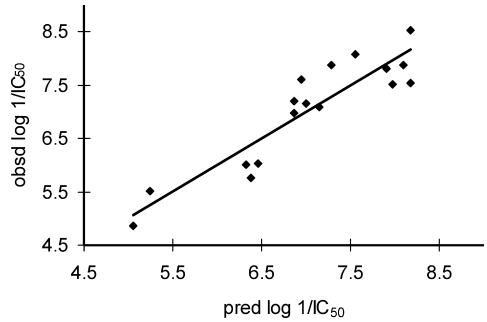


Figure 2. Plot of observed versus predicted log $1/IC_{50}$ from eq 2.

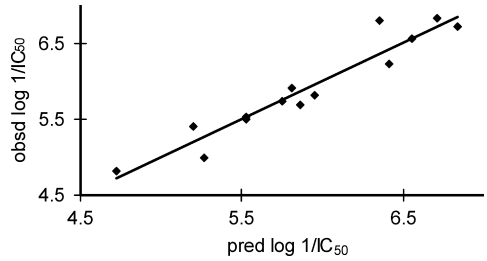


Figure 3. Plot of observed versus predicted log $1/IC_{50}$ from eq 3.

of 2.95 and then decrease. The increase in the hydrophobicity of X substituents (π_X) further increases the cytotoxic activities of the compounds (**IV**) as evidenced by the positive coefficient of π_X (0.97). Thus, the cytotoxicities of cephalomannine derivatives (**IV**) correlate with π_X (linear fashion) followed by MR_X (parabolic fashion). A comparison between observed and predicted values of log $1/IC_{50}$ for cephalomannine analogues (**IV**) used in the development of eq 2 is shown in Figure 2. This is not surprising to obtain a high correlation between cytotoxicities to A2780 cells and binding affinities to β-tubulin of cephalomannine derivatives (**IV**) (log $1/IC_{50}$ vs log $1/K_d$, $r = 0.840$), because these compounds are the microtubule-stabilizing agents.

3. Cytotoxicity of 2-OCOX-10-Deacetyl-7-propionyl Cephalomannines (IV) against A2780/AD10 (Multidrug-Resistant P-gp Overexpression Ovarian Carcinoma Cells).[36] (Table 1)

$$\begin{aligned}&\log 1/IC_{50} = 0.53(\pm 0.32)\pi_X + \\ &\qquad 0.73(\pm 0.23)MR_X + 2.91(\pm 0.62)\\ &n = 14,\ r^2 = 0.918,\ s = 0.202,\ q^2 = 0.873,\\ &\qquad Q = 4.743,\ F_{2,11} = 61.573(7.206)\\ &\pi_X \text{ vs } MR_X,\ r = 0.517\\ &\text{range in log } 1/IC_{50} = 4.82 - 7.77\\ &\text{outliers: } X = 3\text{-}N_3\text{-Ph};\ 3\text{-I-Ph};\ 2,4\text{-di-Cl-Ph}\end{aligned} \quad (3)$$

This is a linear equation in terms of π_X and MR_X, which suggests that the molecule with highly hydrophobic and bulkier or more polarizable X substituent will display more cytotoxic activity against A2780/AD10 cells. A comparison between observed and predicted values of log $1/IC_{50}$ for cephalomannine analogues (**IV**) used in the development of eq 3 is shown in Figure 3.

4. Cytotoxicity of 2-OCOX-10-Deacetyl-7-propionyl Cephalomannines (IV) against 1A9 (Parental Ovarian Carcinoma).[36] (Table 2)

$$\begin{aligned}&\log 1/IC_{50} = 1.20(\pm 0.41)\pi_X + 5.58(\pm 2.75)MR_X - \\ &\quad 0.94(\pm 0.52)MR_X^2 - 0.80(\pm 0.66)I_{HAL} - 2.86(\pm 3.69)\\ &n = 17,\ r^2 = 0.880,\ s = 0.436,\ q^2 = 0.728,\\ &\qquad Q = 2.151,\ F_{4,12} = 22.000(5.412)\\ &\text{optimum } MR_X = 2.97(2.74 - 3.51)\\ &\pi_X \text{ vs } MR_X,\ r = 0.486\\ &\text{range in log } 1/IC_{50} = 4.82 - 8.72\\ &\text{outliers: } X = \text{3-Me-Ph};\ CH_2CH(Me)_2;\ 2,4\text{-di-Cl-Ph}\end{aligned} \quad (4)$$

According to this equation, the cytotoxic activities of cephalomannine derivatives (**IV**) against 1A9 cells correlate with MR_X (in a parabolic fashion; optimum $MR_X = 2.97$) followed by π_X and I_{HAL}. The increase in the hydrophobicity of X substituents (π_X) increases the cytotoxic activities of the compounds (**IV**) as suggested by its positive coefficient (1.20). I_{HAL} is an indicator variable that pinpoints the unusual activity of the halogen containing X substituents. The indicator variable (I_{HAL}) takes the value of 1 and 0 for the presence and absence of halogen in the X substituent, and its negative coefficient suggests that the presence of halogen containing X substituents will be detrimental to the activity. A comparison between observed and predicted values of log $1/IC_{50}$ for cephalomannine analogues (**IV**) used in the development of eq 4 is shown in Figure 4.

5. Cytotoxicity of 2-OCOX-10-Deacetyl-7-propionyl Cephalomannines (IV) against PTX22 (Paclitaxel-Resistance Tumor Cells with β-Tubulin Mutation).[36] (Table 2)

$$\begin{aligned}&\log 1/IC_{50} = 1.71(\pm 0.74)\pi_X + 0.85(\pm 0.63)MR_X - \\ &\qquad 1.01(\pm 0.92)I_{HAL} + 0.82(\pm 1.95)\\ &n = 14,\ r^2 = 0.823,\ s = 0.541,\ q^2 = 0.623,\\ &\qquad Q = 1.677,\ F_{3,10} = 15.499(6.552)\\ &\pi_X \text{ vs } MR_X,\ r = 0.861\\ &\text{range in log } 1/IC_{50} = 4.85 - 8.72\\ &\text{outliers: } X = \text{3-I-Ph};\ \text{3-Me-Ph};\ 2,4\text{-di-Cl-Ph}\end{aligned} \quad (5)$$

According to this equation, an increase in the hydrophobicity and molar refractivity of X substituents (π_X and MR_X) will increase the cytotoxic activities of compounds (**IV**) as evidenced by their positive coefficients (1.71 and 0.85). Indicator variable (I_{HAL}) with negative coefficient suggests that the presence of halogen containing X substituents will be detrimental to the activity. Although a very high correlation between π_X and MR_X ($r^2 = 0.741$), both parameters were used in the development of eq 5, because the removal of either parameter significantly decreased the predictive ability

Table 2. Biological, Physicochemical, and Structural Parameters Used to Derive QSARs 4 and 5

		log 1/IC_{50} (eq *4*)			log 1/IC_{50} (eq *5*)					
compd	*X*	obsd	pred	Δ	obsd	pred	Δ	π_X	MR_X	I_{HAL}
IV-1	Ph	7.80	7.67	0.13	6.36	6.40	−0.04	2.02	2.51	0
IV-2	3-N_3-Ph	8.72	8.81	−0.09	8.72	8.49	0.23	2.87	3.26	0
IV-3	3-OMe-Ph	8.24	8.05	0.19	8.14	7.21	0.93	2.20	3.13	0
IV-4	3-CHO-Ph	7.28	7.36	−0.08	6.02	6.10	−0.08	1.60	3.01	0
IV-5	3-CH_2OH-Ph	5.86	6.60	−0.74	4.85	5.14	−0.29	0.98	3.13	0
IV-6	3-Cl-Ph	8.19	7.93	0.26	7.29	7.02	0.27	2.73	3.00	1
IV-7	3-Br-Ph	7.60	8.01	−0.41	6.80	7.52	−0.72	2.88	3.29	1
IV-8[b]	3-I-Ph	7.57	7.75	−0.18	6.64	8.41	−1.77	3.14	3.82	1
IV-9	3-CN-Ph	7.91	7.19	0.72	6.49	5.83	0.66	1.45	2.99	0
IV-10[a,b]	3-Me-Ph	7.48	8.47	−0.99	6.36	7.64	−1.28	2.52	2.97	0
IV-11	2-thiophene	7.25	7.13	0.12	5.70	5.76	−0.06	1.74	2.32	0
IV-12	3-thiophene	7.15	7.13	0.02	5.52	5.76	−0.24	1.74	2.32	0
IV-13	2-chromone	4.82	4.91	−0.09	ND[c]	2.60	ND[c]	−0.09	2.30	0
IV-14	CH═C(Me)$_2$	6.05	6.57	−0.52	5.10	5.56	−0.46	1.83	1.91	0
IV-15	CH_2C(Me)═CH_2	5.97	6.38	−0.41	5.42	5.46	−0.04	1.81	1.83	0
IV-16	CH═CHMe	5.40	4.96	0.44	ND[c]	4.48	ND[c]	1.43	1.44	0
IV-17[a]	CH_2CH(Me)$_2$	5.10	6.78	−1.68	ND[c]	5.96	ND[c]	2.09	1.86	0
IV-18	2,4-di-F-Ph	7.04	6.71	0.33	5.60	5.15	0.45	1.87	2.54	1
IV-19[a,b]	2,4-di-Cl-Ph	6.93	8.00	−1.07	5.59	7.90	−2.31	3.01	3.49	1
IV-20	2,5-di-OMe-Ph	7.62	7.30	0.32	6.78	7.42	−0.64	2.01	3.75	0

[a] Not used in the derivation of QSAR 4. [b] Not used in the derivation of QSAR 5. [c] Not determined.

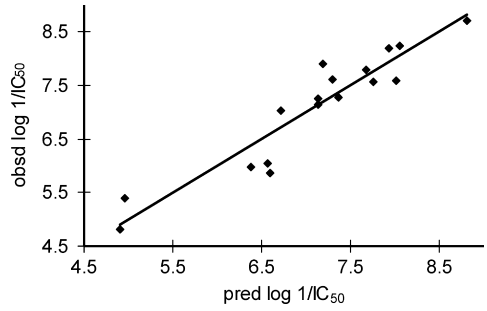


Figure 4. Plot of observed versus predicted log 1/IC_{50} from eq 4.

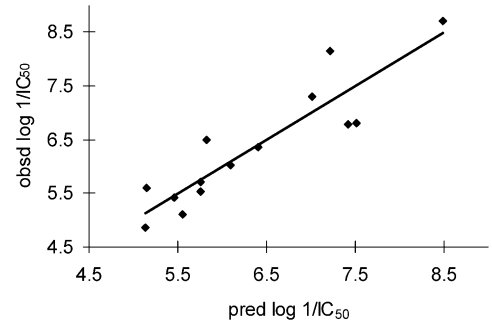


Figure 5. Plot of observed versus predicted log 1/IC_{50} from eq 5.

of the model (see eqs 5a and 5b, where the same outliers were considered as that of eq 5).

$$\log 1/IC_{50} = 1.96(\pm 0.93)\pi_X - 1.00(\pm 1.19)I_{HAL} + 2.68(\pm 1.78)$$
$$n = 14,\ r^2 = 0.664,\ s = 0.711,\ q^2 = 0.504,\ Q = 1.146,\ F_{2,11} = 10.869(7.206) \quad (5a)$$

$$\log 1/IC_{50} = 1.21(\pm 1.09)MR_X + 0.04(\pm 1.43)I_{HAL} + 2.98(\pm 3.06)$$
$$n = 14,\ r^2 = 0.357,\ s = 0.984,\ q^2 = 0.055,\ Q = 0.607,\ F_{2,11} = 3.054(7.206) \quad (5b)$$

Thus, the development of QSAR 5 with three parameters, that is, π_X, MR_X, and I_{HAL}, is justified. A comparison between observed and predicted values of log 1/IC_{50} for cephalomannine analogues (**IV**) used in the development of eq 5 is shown in Figure 5.

Validation of QSAR Models

QSAR model validation becomes an essential part in the development of statistically robust models, because the real utility of a QSAR model is in its ability to predict accurately the modeled property for new compounds not present in the data set. A comparison of the regression coefficients/statistics of QSARs 1−5 is shown in Table 3. Criteria of validation for the QSAR models have already been discussed previously.[56−61] The following approaches have been used to validate QSAR models (eqs 1−5):

Internal Validation: (1) Fraction of the Variance (r^2). According to the literature, the predictive QSAR model must have $r^2 > 0.6$.[60,61] The values of r^2 for these five QSAR

(56) Jaworska, J. S.; Comber, M.; Auer, C.; Van Leeuwen, C. J. Summary of workshop on regulatory acceptance of (Q)SARs for human health and environmental endpoints. *Environ. Health Perspect.* **2003**, *111*, 1358–1360.

Table 3. Comparison of the Regression Coefficients and the Statistics Obtained from the Multiregression Analyses (MRA) Process for QSARs 1−5

			regression coefficients					statistical parameters				
QSAR no.	system	n	π_X	MR_X	MR_X^2	I_{HAL}	intercept	r^2	q^2	s	Q	F^a
1	β-tubulin	16	0.40	0.56			4.06	0.848	0.792	0.269	3.424	36.263(6.701)
2	A2780 cells	17	0.97	4.16	−0.71		−0.69	0.836	0.661	0.462	1.978	22.089(5.739)
3	A2780/AD10 cells	14	0.53	0.73			2.91	0.918	0.873	0.202	4.743	61.573(7.206)
4	1A9 cells	17	1.20	5.58	−0.94	−0.80	−2.86	0.880	0.728	0.436	2.151	22.000(5.412)
5	PTX22 cells	14	1.71	0.85		−1.01	0.82	0.823	0.623	0.541	1.677	15.499(6.552)

[a] The figures within parentheses refer to the literature F-value at 99% level.

Table 4. Y-Randomization Data for QSARs 1−5

	NOR-1[a]		NOR-2		NOR-3		NOR-4		NOR-5	
QSAR no.	r^2	q^2	r^2	q^2	r^2	q^2	r^2	q^2	r^2	q^2
1	0.418	0.145	0.103	−0.310	0.579	0.280	0.095	−0.450	0.429	0.089
2	0.504	0.198	0.449	0.198	0.462	0.262	0.448	0.026	0.464	0.211
3	0.090	−0.389	0.268	−0.238	0.355	−0.332	0.089	−0.443	0.217	−0.275
4	0.528	0.046	0.477	0.154	0.489	0.226	0.477	0.003	0.492	0.198
5	0.610	0.246	0.362	−0.074	0.513	0.275	0.374	0.050	0.359	−0.130

[a] NOR = number of Y-randomization.

models (eqs 1−5) range from 0.823 to 0.918, which suggests that these QSAR models explain 82.3−91.8% of the variance in the data.

(2) Cross-Validation Test. According to the literature, the predictive QSAR model must have $q^2 > 0.5$.[60,61] The cross-validated r^2 (q^2) values for these QSAR models range from 0.623 to 0.873.

(3) Standard Deviation (s). The smaller the value of s, the better the QSAR model. The values of s for these QSAR models range from 0.202 to 0.541.

(4) Quality Factor (Q). High values of Q (1.677−4.743) for these QSAR models suggest their high predictive power.

(5) Fischer Statistics (F). The larger the value of F, the greater the probability that the QSAR model is significant. The F-values for these QSAR models range from 15.499 to 61.573, which are statistically significant at the 99% level.

(6) Y-Randomization Test. In this test, the dependent-variable vector (**Y** vector) is randomly shuffled, and a new QSAR model is developed using the original independent variable matrix. The process is repeated several times. It is expected that the resulting QSAR models should have low r^2 and low q^2 values. This is a widely used technique to ensure the robustness of a QSAR model. The statistical data of r^2 and q^2 for five runs are listed in Table 4 (eqs 1−5). The poor values of r^2 and q^2 in the Y-randomization test confirm the robustness of these QSAR models.[3,7,47,61−63]

(7) Lack of Overfitting. A QSAR model overfits if it includes more descriptors than required. The lack of overfitting for all these QSAR models (eqs 1−5) was checked by using the following conditions: (i) number of data points/number of descriptors ≥ 4, (ii) high values of Q, (iii) comparison of the original QSAR model with that of fewer descriptors, and (iv) Y-randomization test (Table 4). These tests suggest that the high r^2 values of QSAR models (eqs 1−5) are not due to the chance correlation or overfitting.[7,64]

External Validation: Selection of the Training and Test Sets. The original data set of QSAR model (eq 1) was divided into training [$n = 12$ (75%)] and test [$n = 4$ (25%)] sets in a random manner. The QSAR model for this training set was generated by using the same descriptors as those of

(57) Eriksson, L.; Jaworska, J.; Worth, A. P.; Cronin, M. T. D.; McDowell, R. M.; Gramatica, P. Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs. *Environ. Health Perspect.* **2003**, *111*, 1361–1375.

(58) Cronin, M. T. D.; Walker, J. D.; Jaworska, J. S.; Comber, M. H. I.; Watts, C. D.; Worth, A. P. Use of QSARs in international decision-making frameworks to predict ecologic effects and environmental fate of chemical substances. *Environ. Health Perspect.* **2003**, *111*, 1376–1390.

(59) Cronin, M. T. D.; Jaworska, J. S.; Walker, J. D.; Comber, M. H. I.; Watts, C. D.; Worth, A. P. Use of QSARs in international decision-making frameworks to predict health effects of chemical substances. *Environ. Health Perspect.* **2003**, *111*, 1391–1401.

(60) Golbraikh, A.; Tropsha, A. Beware of q^2! *J. Mol. Graphics Modell.* **2002**, *20*, 269–276.

(61) Tropsha, A.; Gramatica, P.; Gombar, V. K. The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models. *QSAR Comb. Sci.* **2003**, *22*, 69–77.

(62) Wold, S.; Eriksson, L. In *Chemometrics Methods in Molecular Design: Statistical Validation of QSAR Results*; van de Waterbeemd, H., Ed.; VCH: Weinheim, Germany, 1995; pp 309−318.

(63) Melzig, M. F.; Tran, G. D.; Henke, K.; Selassie, C. D.; Verma, R. P. Inhibition of neutrophil elastase and thrombin activity by caffeic acid esters. *Pharmazie* **2005**, *60*, 869–873.

(64) Zhang, S.; Golbraikh, A.; Oloff, S.; Kohn, H.; Tropsha, A. A novel automated lazy learning QSAR (ALL-QSAR) approach: method development, applications, and virtual screening of chemical databases using validated ALL-QSAR models. *J. Chem. Inf. Model.* **2006**, *46*, 1984–1995.

Table 5. Random Selection Pattern of the Test Set Compounds as well as the Regression Coefficients and Statistical Parameters of the QSAR for Their Respective Training Set Compounds Obtained from the Division of the Original QSAR Models (eqs *1*−*5*)

			regression coefficients					statistical parameters			
QSAR no.[a]	test set compd	training set compd	π_X	MR_X	MR_X^2	I_{HAL}	intercept	r^2	q^2	s	R_{pred}^2
1	**IV-9**, **IV-11**, **IV-13**, **IV-16**	rest of the compd ($n = 12$)	0.60	0.48			3.77	0.846	0.694	0.270	0.679
2	**IV-1**, **IV-2**, **IV-7**, **IV-14**	rest of the compd ($n = 13$)	0.99	3.69	−0.62		0.09	0.865	0.633	0.431	0.766
3	**IV-6**, **IV-14**, **IV-16**, **IV-20**	rest of the compd ($n = 10$)	0.60	0.59			3.16	0.855	0.722	0.213	0.937
4	**IV-9**, **IV-13**, **IV-16**, **IV-20**	rest of the compd ($n = 13$)	1.52	8.88	−1.57	−0.72	−7.71	0.929	0.542	0.292	0.565
5	**IV-3**, **IV-5**, **IV-9**, **IV-20**	rest of the compd ($n = 10$)	1.55	0.96		−0.88	0.75	0.904	0.589	0.417	0.660

[a] QSAR for the respective training set of the original QSAR.

Table 6. Predicted Biological Activities (log $1/K_d$ or log 1/C) Obtained from QSARs 1−5 along with Physicochemical Parameters of the Proposed Molecules

compd	X	C log P	π_X	MR_X	log K_d (eq *1*)	log 1/C (eq *2*)	log 1/C (eq *3*)	log 1/C (eq *4*)	log 1/C (eq *5*)
IV-21	3-CH_2OCH_3-Ph	4.61	1.82	3.59	6.79	6.92	6.50	7.27	6.96
IV-22	3-CH_2COOCH_3-Ph	4.49	1.70	4.09	7.02	6.18	6.80	6.31	7.18
IV-23	3-$NHCH_3$-Ph	4.86	2.07	3.34	6.76	7.35	6.45	7.80	7.18
IV-24	3-$COCH_3$-Ph	4.39	1.60	3.47	6.64	6.81	6.29	7.13	6.48
IV-25	3-$NHCOCH_3$-Ph	4.30	1.51	3.84	6.81	6.35	6.52	6.55	6.65
IV-26	3,4-$(OMe)_2$-Ph	4.69	1.90	3.75	6.91	6.85	6.65	7.17	7.23
IV-27	NHPh	4.62	1.83	2.88	6.40	7.22	5.98	7.63	6.37
IV-28	$NHSO_2Ph$	4.25	1.46	3.75	6.74	6.41	6.42	6.63	6.48
IV-29	NHCOPh	4.51	1.71	3.38	6.63	6.98	6.29	7.34	6.60
IV-30	CH_2Ph	5.02	2.23	2.97	6.62	7.62	6.26	8.12	7.14

eq 1 and validated on the basis of their statistics (acceptance criteria: $r^2 > 0.6$ and $q^2 > 0.5$). The predictive capacity of this model is judged from their predictive R^2 (R_{pred}^2) value, which was calculated by eq 6:

$$R_{pred}^2 = 1 - \frac{\sum (Y_{pred(test)} - Y_{test})^2}{\sum (Y_{test} - \bar{Y}_{training})^2} \qquad (6)$$

In eq 6, $Y_{pred(test)}$ and Y_{test} are the respective predicted and observed activities of the test set compounds, and $\bar{Y}_{training}$ is the observed mean activity of the training set compounds. Similarly, the QSAR models (eqs 2−5) were also divided into training and test sets for their external validation. A random selection pattern of the test sets as well as the regression coefficients and the statistical parameters of their respective training sets for all these QSAR models (eqs 1−5) are given in Table 5.

New Molecule Prediction

QSARs 2 and 4 (parabolic correlations in terms of MR_X) are encouraging examples in which the optimum values of MR_X are well-defined. We believe that these two models may prove to be adequate predictive models for providing guidance in design and synthesis, and for yielding very specific compounds (**IV**) that may have high anti-ovarian cancer activity with fewer side effects and superior pharmacological properties. On the basis of these two QSAR models, 10 cephalomannine analogues (**IV-21** to **IV-30**) are suggested as potential synthetic targets. The activities of these proposed compounds are also predicted by using the rest of the three QSAR models (eqs 1, 3, and 5). All the proposed cephalomannine derivatives fulfill the following three specific conditions: (i) increased π_X with log $P \leq 5.00$, (ii) $MR_X \leq 4.00$, and (iii) no halogen atom. In the proposed molecules, most of the compounds have $X =$ 3-substituted phenyl group. This is according to the previous observations, 2-*ortho*- and 2-*para*-substituted benzoates; especially the latter exhibit much lower activity than 2-*meta*-substituted benzoates of paclitaxel. The poor activity of 2-*para*-substituted taxanes originates from a steric effect that decreases the hydrophobic interaction between the phenyl ring and adjacent amino acid residues. The increased activity of 2-*meta*-substituted taxanes may be attributed due to the rotation about the bond between the aryl ring and the carbonyl group.[65,66] The predicted log $1/K_d$ and log $1/IC_{50}$ values of the proposed cephalomannine analogues (**IV-21** to **IV-30**) obtained from QSARs 1−5 are given in Table 6 along with their physicochemical parameters.

(65) Kingston, D. G. I.; Chaudhary, A. G.; Chordia, M. D.; Gharpure, M.; Gunatilaka, A. A. L.; Higgs, P. I.; Rimoldi, J. M.; Samala, L.; Jagtap, P. G.; Giannakakou, P.; Jiang, Y. Q.; Lin, C. M.; Hamel, E.; Long, B. H.; Fairchild, C. R.; Johnston, K. A. Synthesis and Biological Evaluation of 2-Acyl Analogues of Paclitaxel (Taxol). *J. Med. Chem.* **1998**, *41*, 3715–3726.

(66) Chaudhary, A. G.; Gharpure, M. M.; Rimoldi, J. M.; Chordia, M. D.; Kingston, D. G. I.; Grover, S.; Lin, C. M.; Hamel, E.; Gunatilaka, A. A. L. Unexpectedly Facile Hydrolysis of the 2-Benzoate Group of Taxol and Syntheses of Analogs with Increased Activities. *J. Am. Chem. Soc.* **1994**, *116*, 4097–4098.

Table 7. Contributions of Different Physicochemical and Structural Parameters in eqs 1−5

QSAR no.	QSAR models	*n*	r^2	q^2	*s*
1-1	$\log 1/K_d = 0.66\pi_X + 5.03$	16	0.577	0.418	0.432
1-2	$\log 1/K_d = 0.77MR_X + 4.27$	16	0.686	0.621	0.372
2-1	$\log 1/IC_{50} = 1.08\pi_X + 4.92$	17	0.671	0.599	0.609
2-2	$\log 1/IC_{50} = 0.98MR_X + 4.34$	17	0.403	0.277	0.820
2-3	$\log 1/IC_{50} = 3.40MR_X - 0.45MR_X^2 + 1.30$	17	0.447	0.289	0.817
2-4	$\log 1/IC_{50} = 0.88\pi_X + 0.45MR_X + 4.07$	17	0.733	0.596	0.568
3-1	$\log 1/IC_{50} = 1.06\pi_X + 3.78$	14	0.543	0.427	0.456
3-2	$\log 1/IC_{50} = 0.92MR_X + 3.45$	14	0.819	0.768	0.287
4-1	$\log 1/IC_{50} = 1.06\pi_X + 5.07$	17	0.570	0.472	0.739
4-2	$\log 1/IC_{50} = 1.06MR_X + 4.18$	17	0.420	0.290	0.859
4-3	$\log 1/IC_{50} = 3.96MR_X - 0.54MR_X^2 + 0.54$	17	0.475	0.323	0.846
4-4	$\log 1/IC_{50} = 0.67I_{HAL} + 6.92$	17	0.072	−0.110	1.086
4-5	$\log 1/IC_{50} = 1.24\pi_X - 0.56I_{HAL} + 4.87$	17	0.605	0.486	0.734
4-6	$\log 1/IC_{50} = 1.04MR_X + 0.10I_{HAL} + 4.22$	17	0.421	0.216	0.888
4-7	$\log 1/IC_{50} = 3.96MR_X - 0.55MR_X^2 + 0.12I_{HAL} + 0.56$	17	0.477	0.271	0.876
4-8	$\log 1/IC_{50} = 0.81\pi_X + 0.60MR_X + 3.90$	17	0.673	0.485	0.667
4-9	$\log 1/IC_{50} = 0.97\pi_X + 5.25MR_X - 0.89MR_X^2 - 2.09$	17	0.812	0.615	0.525
5-1	$\log 1/IC_{50} = 1.56\pi_X + 3.26$	14	0.559	0.380	0.780
5-2	$\log 1/IC_{50} = 1.21MR_X + 2.97$	14	0.357	0.162	0.942
5-3	$\log 1/IC_{50} = 0.28I_{HAL} + 6.28$	14	0.011	−0.291	1.168
5-4	$\log 1/IC_{50} = 1.96\pi_X - 1.00I_{HAL} + 2.68$	14	0.664	0.504	0.711
5-5	$\log 1/IC_{50} = 1.21MR_X + 0.04I_{HAL} + 2.98$	14	0.357	0.055	0.984
5-6	$\log 1/IC_{50} = 1.31\pi_X + 0.84MR_X + 1.42$	14	0.716	0.506	0.654

Overview

An analysis of our QSAR models (eqs 1−5) reveals a number of interesting points. The most important of these are π_X and MR_X. A positive linear correlation between the activity and π_X is found in all five equations. The coefficient associated with π_X ranges from 0.40 (eq 1) to 1.71 (eq 5). These data suggest that activity might be improved by increasing hydrophobicity of *X* substituents. A positive linear correlation with MR_X is found in three equations (eqs 1, 3, and 5) with coefficients 0.56−0.85, which suggests that the activity might be improved by increasing molar refractivity of *X* substituents. A parabolic correlation with MR_X is present in two equations (eqs 2 and 4), which suggests that the activity is optimal for a particular value of MR_X. The optimal values of MR_X for these two equations are 2.95 and 2.97, respectively.

In this paper, eq 1 is for the binding affinity of 2-OCOX-10-deacetyl-7-propionyl cephalomannines (**IV**) to β-tubulin, eqs 2 and 4 for the cytotoxicity against sensitive cancer cells (A2780 and 1A9), and eqs 3 and 5 for the cytotoxicity against resistant cancer cells (A2780/AD10 and PTX22). On comparison of QSAR models for sensitive cancer cells (A2780 and 1A9), we can see that eq 2 is very similar to that of eq 4 (parabolic correlation in terms of MR_X followed by π_X). It is also interesting to note that both equations have the same optimum MR_X (2.95 and 2.97) and similar range of activity ($\log 1/IC_{50}$ = 4.85−8.54 and 4.82−8.72). However, this is not surprising because parental ovarian carcinoma 1A9 is a clone of A2780 and acting with the very same mechanism. In the case of QSAR models for sensitive and resistant cancer cells (A2780 and A2780/AD10), we can see that eq 2 is not similar to eq 3. The coefficients of π_X in eq 2 and 3 are significantly different as well as the different MR_X term. This suggests that the mechanism for the cytotoxicity against sensitive cancer cells (A2780) is different from that of resistant cancer cells (A2780/AD10). Similarly, eqs 2 and 5, eqs 3 and 4, and eqs 4 and 5 are also different from each other with respect to π_X coefficients and MR_X term. On the basis of these comparisons among eqs 2−5, we can say that the mechanism for the cytotoxicity against sensitive cancer cells (eqs 2 and 4) is different from that for the resistant cancer cells (eqs 3 and 5). Of course, the common target is the β-tubulin because there is a good correlation between the cytotoxicity of cephalomannines (**IV**) against all four cancer cell lines and their binding affinity to β-tubulin. On the other hand, the mechanism is also different between two QSAR models 3 and 5, respectively, for the resistant cancer cells A2780/AD10 (MDR P-gp-overexpression) and PTX22 (paclitaxel-resistant with β-tubulin mutation). This can be judged by comparing the coefficients of π_X in both eqs 3 and 5 that are significantly different, although the indicator variable (I_{HAL}) is present only in eq 5 and not in eq 3.

To extend our understanding toward the importance of comparative QSARs, we further explored the contributions of different parameters in all five eqs 1−5 (see Table 7). From QSAR nos. 1-1 and 1-2, one can see that MR_X is the most important contributor to the binding affinity of cephalomannines (**IV**) to β-tubulin with an r^2 value of 0.686, but π_X is also the important contributor (r^2 = 0.577). Thus, neither steric (MR_X) nor hydrophobic (π_X) nature alone is the dominant factor in the activity. QSAR nos. 2-1 (r^2 = 0.671) and 4-1 (r^2 = 0.570) suggest that π_X is the most

Table 8. Correlations among the Binding Affinity of 2-OCOX-10-Deacetyl-7-propionyl Cephalomannines (**IV**) to β-Tubulin and their Cytotoxic Activities against Different Tumor Cell Lines

QSAR no.	QSAR Models	*n*	r^2	q^2	*s*
7	$\log 1/K_d = 0.74 \log 1/IC_{50(A2780)} + 1.32$	19	0.705	0.604	0.503
8	$\log 1/K_d = 1.08 \log 1/IC_{50(A2780/AD10)} + 0.09$	17	0.881	0.825	0.310
9	$\log 1/K_d = 0.70 \log 1/IC_{50(1A9)} + 1.61$	19	0.718	0.619	0.492
10	$\log 1/K_d = 0.80 \log 1/IC_{50(PTX22)} + 1.59$	16	0.900	0.839	0.276
11	$\log 1/IC_{50(A2780)} = 0.88 \log 1/IC_{50(A2780/AD10)} + 1.96$	17	0.685	0.593	0.463
12	$\log 1/IC_{50(A2780)} = 0.60 \log 1/IC_{50(A2780/AD10)} + 4.24MR_X - 0.73MR_X^2 - 2.17$	17	0.885	0.795	0.301
13	$\log 1/IC_{50(A2780)} = 0.94 \log 1/IC_{50(1A9)} + 0.43$	20	0.985	0.982	0.131
14	$\log 1/IC_{50(A2780)} = 0.64 \log 1/IC_{50(PTX22)} + 3.22$	17	0.758	0.673	0.390
15	$\log 1/IC_{50(A2780/AD10)} = 0.73 \log 1/IC_{50(1A9)} + 0.72$	17	0.690	0.603	0.434
16	$\log 1/IC_{50(A2780/AD10)} = 0.64 \log 1/IC_{50(PTX22)} + 2.05$	16	0.810	0.748	0.319
17	$\log 1/IC_{50(A2780/AD10)} = 0.52 \log 1/IC_{50(PTX22)} + 0.45\pi_X + 1.80$	16	0.908	0.868	0.230
18	$\log 1/IC_{50(A2780/AD10)} = 0.49 \log 1/IC_{50(PTX22)} + 0.45MR_X + 1.74$	16	0.907	0.837	0.231
19	$\log 1/IC_{50(A2780/AD10)} = 0.46 \log 1/IC_{50(PTX22)} + 0.30\pi_X + 0.29MR_X + 1.67$ [π_X vs MR_X; $r = 0.496$]	16	0.939	0.856	0.195
20	$\log 1/IC_{50(1A9)} = 0.69 \log 1/IC_{50(PTX22)} + 2.95$	17	0.810	0.744	0.360

important contributor to the cytotoxic activities of cephalomannines (**IV**) against sensitive cancer cells (A2780 and 1A9). On the other hand, the most important contributor to the cytotoxicity of cephalomannines (**IV**) against the resistant cancer cells A2780/AD10 (MDR P-gp-overexpression) is MR_X (QSAR no. 3-2, $r^2 = 0.819$), but in the case of resistant cancer cells PTX22 (paclitaxel-resistant with β-tubulin mutation) the most important contributor is π_X (QSAR no. 5-1, $r^2 = 0.559$). The contradictory results for the resistance cancer cells suggest further the different resistance mechanism between eq 3 (MDR P-gp-overexpression) and eq 5 (paclitaxel-resistant with β-tubulin mutation). This situation may be explained by the fact that PTX22 cells have mutations in β-tubulin gene at the position of Ala 364→Thr. This amino acid lies on the floor of the deep hydrophobic pocket surrounding paclitaxel, and changes to them confer paclitaxel resistance. On the other hand, P-gp is able to pump many hydrophobic molecules, including taxoids, out of the cell.[36,55] On comparison between the cytotoxicity of cephalomannines (**IV**) against sensitive and resistant cancer cells (A2780 and A2780/AD10), one can see that π_X is the most important contributor against A2780 cells (QSAR no. 2-1, $r^2 = 0.671$) while the most important contributor in the case of resistant cancer cells A2780/AD10 is MR_X (QSAR no. 3-2, $r^2 = 0.819$). This suggests further that the mechanism for cytotoxicity against sensitive cancer cells (A2780) is different from that of resistant cancer cells (A2780/AD10).

Activity−activity relationships are also good methods for understanding the sensitive/resistance mechanism. Correlations among the binding affinity of cephalomannines (**IV**) to β-tubulin and their cytotoxic activities against different tumor cell lines are given in Table 8 (QSAR nos. 7−20). The high correlations among the binding affinity of cephalomannines (**IV**) to β-tubulin and their cytotoxicies against sensitive/resistance cancer cells in QSAR nos. 7−10 ($r^2 =$ 0.705, 0.881, 0.718, and 0.900) suggest that the binding affinity to β-tubulin is the main driving force for entry of cephalomannines (**IV**) into both sensitive and resistance cancer cells. The correlation between the cytotoxicity of A2780 and A2780/AD10 (QSAR no. 11, $r^2 = 0.685$) represents the drugs and P-gp interaction. By adding a significant MR_X^2 term in QSAR no. 11 resulting in the development of QSAR no. 12 ($r^2 = 0.885$) confirms the difference between eqs 2 and 3; that is, the mechanism for the cytotoxicity against sensitive cancer cells (A2780) is different from that of resistant cancer cells (A2780/AD10; MDR P-gp-overexpression). There is a very high correlation between the cytotoxicity of two sensitive cancer cell lines A2780 and 1A9 (QSAR no. 13, $r^2 = 0.985$, Figure 6), because 1A9 is actually a clone of A2780 cells and acting by the same mechanism. QSAR nos. 14 and 15 represent the comparison between the cytotoxicity of sensitive and resistance cancer cells. There is a good correlation between the cytotoxicity of the resistance cancer cells A2780/AD10 and PTX22 (QSAR no. 16, $r^2 = 0.810$, Figure 7). It may be due to the high correlation between the binding affinity of cephalomannines (**IV**) to β-tubulin and cytotoxicies against these resistance cancer cells as shown in QSAR nos. 8 and 10 ($r^2 = 0.881$ and 0.900, respectively). By the addition of a significant parameter, that is, π_X, MR_X, and the both in

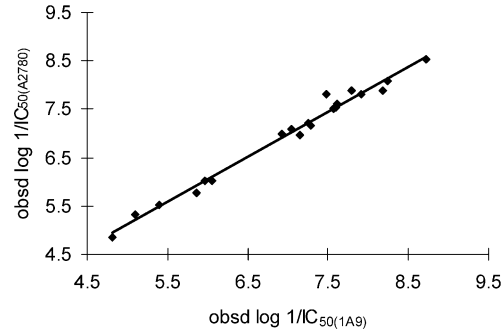


Figure 6. Plot of log $1/IC_{50(A2780)}$ versus log $1/IC_{50(1A9)}$ (see Table 8; QSAR no. 13).

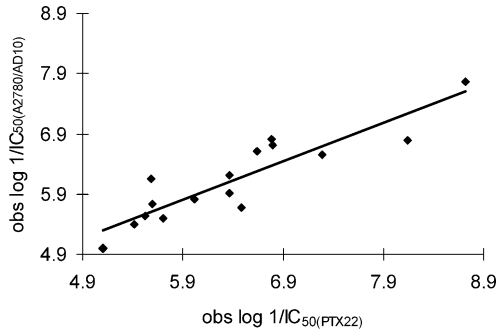


Figure 7. Plot of log $1/IC_{50(A2780/AD10)}$ versus log $1/IC_{50(PTX22)}$ (see Table 8; QSAR no. 16).

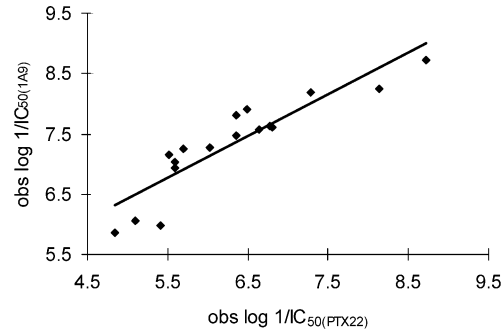


Figure 8. Plot of log $1/IC_{50(1A9)}$ versus log $1/IC_{50(PTX22)}$ (see Table 8; QSAR no. 20).

QSAR no. 16, resulting in the development of QSAR nos. 17, 18, and 19 ($r^2 = 0.908$, 0.907, and 0.939, respectively) confirms the difference between eqs 3 and 5. The difference between eqs 3 and 5 further suggests that the resistance mechanism is different between eq 3 (MDR P-gp-overexpression) and eq 5 (paclitaxel-resistant with β-tubulin mutation). There is also a good correlation between the cytotoxicity of 1A9 and PTX22 (QSAR no. 20, $r^2 = 0.810$, Figure 8), because π_X is the most important determinant for the cytotoxicity of **IV** against both cancer cell lines.

Conclusion

The microtubule-stabilizing taxanes such as paclitaxel and docetaxel are the two most important anticancer drugs currently used in clinics for the treatment of various types of cancers. However, the success of these two drugs has been tempered by the development of various undesired side effects as well as multidrug resistance (MDR), cross-resistance with other chemotherapeutic agents, dose limiting toxicities, neurotoxicity, low oral bioavailability, and very limited CNS penetration, primarily due to P-gp mediated efflux at the blood−brain barrier (BBB). These limitations have led to the search for new taxane derivatives with improved biological activity. We believe that the QSAR paradigm may be helpful in the design and development of novel tubulin binding molecules as new anticancer agents, which will be expected to achieve improvements in anticancer activity, toxicity profile, pharmacology, and drug formulation. Our QSAR results suggest that π_X and MR_X are the two most important determinants for the activity of cephalomannine derivatives (**IV**). Parabolic correlations in terms of MR_X (eqs 2 and 4) are encouraging examples in which the optimum values of MR_X are well-defined. We believe that these two models may prove to be adequate predictive models for providing guidance in design and synthesis, and for yielding very specific compounds (**IV**) that may have high anti-ovarian cancer activity with fewer side effects and superior pharmacological properties. On the basis of these two QSAR models, 10 cephalomannine analogues (**IV-21** to **IV-30**) are suggested as potential synthetic targets. The predicted biological activities of these proposed molecules obtained from all the QSAR models (eqs 1−5) are given in Table 6.

MP800138W