A Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Indices Analysis (CoMSIA) of Anthranilamide Derivatives That Are Multidrug Resistance Modulators

Philippe Labrie,*[,†,‡] Shawn P. Maddaford,[§] Sebastien Fortin,[†] Suman Rakhit,[§] Lakshmi P. Kotra,[#,||] and René C. Gaudreault[†]

Unité des Biotechnologies et de Bioingénierie, Hôpital Saint-François d'Assise, Quebec City, Québec, G1L 3L5, Canada, Faculty of Pharmacy, Université Laval, Pavilion Vandry, Quebec City, Québec, G1K 7P4, Canada, NeurAxon Inc., MaRS Centre, South Tower, 101 College Street, Suite 200, Room 230, Toronto, Ontario, M5G 1L7, Canada, Molecular Design Information Technology Center (MDIT), Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Ontario, M5S 2S2, Canada, and Center for Molecular Design and Preformulations, Toronto General Research Institute, University Health Network, Toronto, Ontario, M5G 1L7, Canada

Received March 1, 2006

In a continuing effort to develop potent and selective modulators of P-glycoprotein (P-gp) activity overcoming the chemoresistance acquired by tumor cells during cancer chemotherapy, we developed 3D quantitative structure−activity relationship (3D QSAR) models using CoMFA and CoMSIA analyses. This study correlates the P-glycoprotein inhibitory activities of 49 structurally related anthranilamide derivatives to several physicochemical parameters representing steric, electrostatic, acceptor, donor, and hydrophobic fields. Both CoMFA and CoMSIA models using three different alignment conformations gave good internal predictions, and their cross-validated r^2 values are between 0.503 and 0.644. These most comprehensive CoMFA and CoMSIA models are useful in understanding the structure−activity relationships of anthranilamide derivatives as well as aid in the design of novel derivatives with enhanced modulation of P-gp activity.

Introduction

A major limitation to the successful chemotherapeutic treatment of cancer is the natural and the acquired resistance of tumor cells to cytotoxic drugs.[1] The overexpression of drug transport proteins such as multidrug-resistance-associated protein (MRP) and P-glycoprotein (P-gp) is a major component involved in multiple drug resistance mechanisms (MDR).[2] As a member of the ATP-binding cassette (ABC) family of ATPase-dependent membrane transporters, P-gp is involved in multidrug resistance mechanisms and therefore is of utmost clinical importance.[3] P-gp was first isolated by Ling and Juliano in 1976 and has been shown to efficiently "pump" substrates out of tumor cells through an ATP-dependent mechanism in a unidirectional fashion.[4] In tumor cells overexpressing P-gp, this results in reduced intracellular drug concentrations, which decreases or abrogates the cytotoxicity of a broad spectrum of antineoplastic agents including 5,12-anthracyclinediones (e.g., doxorubicin), vinca alkaloids (e.g., vincristine), podophyllotoxins (e.g., etoposide), and taxanes (e.g., paclitaxel). Numerous hypotheses on the mechanism of drug transportation by P-gp have been devised; however, they are still debated among experts.[5−9]

The inhibition of P-glycoprotein in clinical oncology is of utmost importance in modern cancer chemotherapy. It is hoped that the use of specific P-gp modulators would restore the therapeutic activity of anticancer agents that are substrates of that transporter because the MDR phenotype has been shown to be reversed by a number of structurally heterogeneous molecules, notably verapamil,[10] cyclosporin A,[11] tacrolimus,[12] quinidine,[13] dihydropyridines (nicardipine),[14] and other drugs.[15,16] Consequently, numerous clinical trials using clinically relevant P-glycoprotein modulator drugs such as verapamil,[17,18] tamoxiphen,[19] progesterone,[20] and cyclosporin A[21] have been tested with limited success mainly because of their intrinsic toxicity or unfavorable pharmacokinetics of the accompanying anticancer drugs. Improved results were obtained using modulators that are specifically directed toward P-gp. The newer agents including cyclosporin D analogues,[22] anthranilamide derivatives (tariquidar),[23,24] acridonecarboxamide derivatives (elacridar),[25] and cyclopropyldibenzosuberane derivatives (zosuquidar)[26] have demonstrated improved P-gp selectivity and pharmacological properties such as the duration of action. Tariquidar (Figure 1), a selective anthranilamide MDR-1 modulator, is the most potent molecule known that can abrogate cell chemoresistance at concentrations ranging from 25 to 80 nM.[27] It is established that MDR modulators share common physicochemical properties such as high lipophilicity and are positively charged at the physiological pH.[6] There are cationic amphiphilic modulators of P-gp that usually possess an aromatic ring system and a basic tertiary nitrogen positioned at a fixed distance from the aromatic system. Several qualitative structure−activity relationships of MDR modulators revealed the importance of the aromatic system and the basic nitrogen[30,31] atom in P-gp inhibitors. Another study of 232 phenothiazines indicated that a molecule bearing an amide carbonyl group and a tertiary amine are necessary to inhibit P-gp.[32] A separate study on 19 propafenone derivatives confirmed the requirement for the carbonyl oxygen, which interacts with the protein via a hydrogen bond to inhibit P-gp.[33]

The first 3D-QSAR of MDR modulators was performed on phenothiazines and related drugs.[34] This was followed by a Hansch-type QSAR studies with propafenone analogues,[35] CoMFA studies of phenothiazines and related drugs,[36] CoMFA studies of propafenone analogues,[37] and simple regression models of propafenone analogues.[38,39] Those models confirmed the importance of hydrogen bond acceptors and donors and the basic nitrogen for P-gp modulators.[38−40]

* To whom correspondence should be addressed. Address: Faculty of Pharmacy, Université Laval, Pavilion Vandry, Quebec City, Québec, G1K 7P4, Canada. Phone: (418) 656-2131, extension 7989. Fax: (418) 525-4372. E-mail: plabrie16@hotmail.com.

† Hôpital Saint-François d'Assise.

‡ Université Laval.

§ NeurAxon Inc.

University of Toronto.

|| University Health Network.

10.1021/jm060239b CCC: $33.50 © 2006 American Chemical Society
Published on Web 12/05/2006

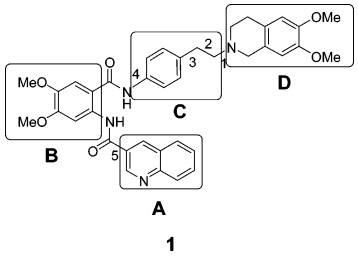


Figure 1. Molecular structure of tariquidar. Different portions of the molecule studied are represented by regions A−D.

3D-QSAR Studies on Anthranilamide Derivatives

In this paper, we report the most comprehensive 3D-QSAR study on a series of 49 tariquidar derivatives.[41] We compared the traditional CoMFA method with the CoMSIA method using anthranilamide P-gp modulators. The latter has been proposed to give better interpretation of results compared to CoMFA method.[42] It is only recently that we can find a comprehensive 3D-QSAR study with anthranilamide derivatives developed by the Wiese and Pajeva group.[43] They used the pharmacophore of Hoechst 33342 to build their model. Hoescht 33342 is known to bind at the H-site of P-gp, similar to **1** (XR9576). There is strong evidence that Hoechst 33342 is a substrate to P-gp spanning its transport and regulatory functions. Although both Hoechst 33342 and **1** share a common binding site, the former is not a P-gp substrate while the latter causes a distinct effect on P-gp and is able to inhibit the ATP hydrolysis and the P-pg function.[24] The 3D-QSAR studies with anthranilamide derivatives developed by the Wiese and Pajeva group have a few common themes between the phenothiazine models and our model. The general structural features of significant importance for anti-MDR activity that are similar for both models are an aromatic ring system substituted by a electroattractor group like methoxyl, trifloromethyl, or chlorine, a tertiary nitrogen atom disposed within an extended side chain, hydrophobicity in the molecule, and the carbonyl group as the acceptor hydrogen group. In the Wiese and Pajeva tariquidar study, 32 anthranilamide derivatives were considered in the training set. It appears that there is no structural diversity in region D and little diversity in the regions B and C, and a single alignment rule was used (Figure 1). In fact, only the tetrahydroquinoline moiety is used in region D. The phenyl group and four molecules with a direct bond to the (alkyl chain) are used in region C, and a few molecules with substitution are used in the anthranilamide core. They obtained a moderate to good average cross-validated coefficient q^2 (0.211−0.794) in CoMFA and CoMSIA models, and they used only one compound as a real external test set. With the above background, we initiated a comprehensive investigation into the structural features of tariquidar, Hoechst 33342, and their derivatives. We included 49 tariquidar derivatives with a variety of functionalities in every region of the molecule (Table 1). It is well-known that the choice of alignment and conformation could influence the final model. Because of several possibilities, we decided to investigate three different conformations and two alignment rules to compare if the choice of the conformation and the alignment would influence the correlation. The choice of an active conformation is an important task prior to building a 3D-QSAR model. The choice of the three different conformations was based on the following parameters: the lowest global energy, extended conformations, and overlap of **1** onto Hoescht 33342. Current studies will help to establish the functional regions that are necessary for the activity of P-gp antagonists.

Computational Details

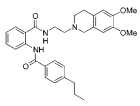
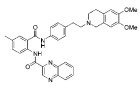
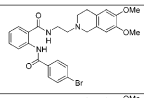
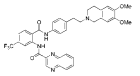
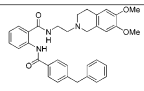
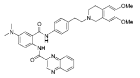
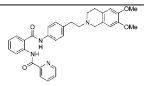
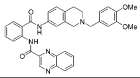
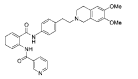
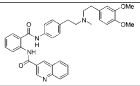
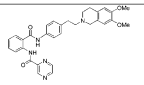
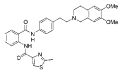
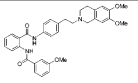
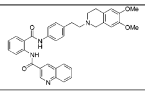
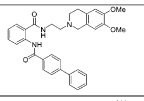
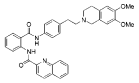
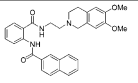
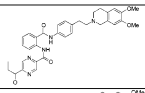
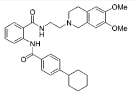
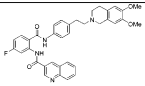
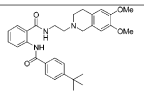
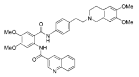
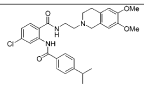
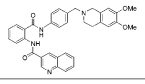
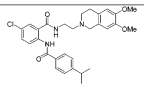
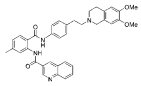
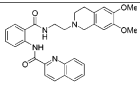
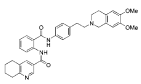
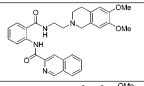
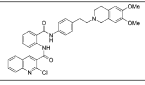
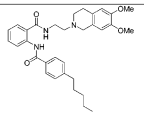
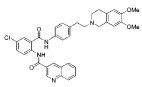
Data Sets and Biological Activity. To ensure that there is consistency in determining the biological activities, the training sets and test sets for the QSAR analyses were taken from the same sample data set consisting of 178 compounds reported by Xenova Group Ltd.[41] The IC_{50} values (μM) reflecting the accumulation of daunorubicin when co-incubated with various compounds on AR 1.0 cells overexpressing P-glycoprotein were considered in building the 3D-QSAR models. Forty-nine compounds were selected for the training set, and 13 compounds were selected for the test set. The latter molecules are different from those of the training set. The molecules of the test set represent 27% of the training set, which is a good ratio to validate a molecular model. The strategy for the selection of the compounds to be included in the test set was a random selection of compounds that exhibited a large range of inhibitory activities. The structure and the IC_{50} values of the compounds chosen to be part of the training and the test set are listed in Table 1.

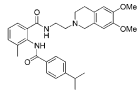
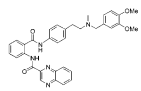
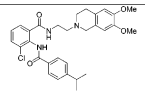
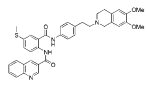
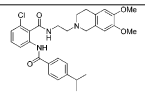
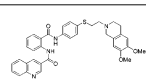
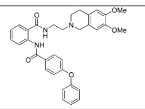
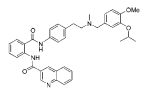
Template Selection, Conformation, and Optimization. In the development of 3D-QSAR models, the choice of the template conformation is important to provide the illustration of a reliable pharmacophore model. Unfortunately, there is no high-resolution X-ray structure of P-gp complexes available for any structure-based drug design efforts. Moreover, anthranilamide derivatives that were used in this study have a flexible molecular structure; thus, the determination of the active or a single conformation is difficult to achieve in the absence of the complex structure. In that context, **1** was selected as a molecular template. This compound was chosen mainly for its importance as a lead structure. In addition, **1** is one of the most potent modulators of P-gp known. In our study, three different conformations of **1** were selected to assess their effect on the determination of a valid 3D-QSAR model. Initial structures were generated using the cleanup procedure within SYBYL and energy-minimized using MAXIMIN2 (Powell method, 2000 iterations, and 0.05 kcal mol^{-1} $Å^{-1}$ energy gradient convergence criteria). It is many times true that the conformation representing the global minimum of the ligand may not bind to the receptor and some degree of torsional freedom is required for the drug to adapt to the receptor binding site to yield a drug−receptor complex of lower energy.[47,48]

The "minimum" energy conformation resulting from a MAXIMIN2 procedure is a good starting point for possible candidate conformations for the compound of interest. It is also important to restrict all possible conformations of the drugs to those that can reasonably be obtained upon binding. Although there are no absolute rules to do that procedure, a ±10 kcal/mol cutoff (difference between the energy of the local minimum conformation and that of a particular chosen conformation) is considered reasonable in CoMFA studies.[49,50]

a. Selection of the Conformations 1 and 2 of Tariquidar. A systematic search on the energy-minimized conformations of **1** was initially undertaken. As shown in Figure 1, five rotatable bonds with an increment of 30° were used to generate 10 259 conformers. The top 50 conformers having the lowest energy were energy-minimized using the Tripos force field software (Powell method, 2000 iterations, and 0.05 kcal mol^{-1} $Å^{-1}$ energy gradient convergence criteria) and optimized using MNDO (full optimization, precise convergence, restricted rotation around the amide bonds−MMOK). Two conformers of **1** were then selected. The first one, designated as 9576_T1, has an energy of 21.167 cal/mol (Figure 2A). The second one, designated as 9576_T2, has an energy of 16.384 kcal/mol

Table 1. Molecular Structures and MDR Reversing Activities of the Molecules Selected for Both the Training and the Test Sets

Code # [a]	Compounds	IC_{50} (μM)	Code # [a]	Compounds	IC_{50} (μM)
t_9297		0.4	t_9543		0.07
t_9380		0.9	t_9561		0.055
t_9442		1	t_9596		0.132
t_9472		0.5	t_9614		0.278
t_9482		0.54	t_9630		0.235
t_9493		0.22	t_9635		0.669
t_9517		0.4	**26** (9544)		0.05
2 (9294)		0.4	**27** (9545)		0.8
3 (9295)		0.39	**28** (9569)		0.93
4 (9331)		1.3	**29** (9571)		0.022
5 (9334)		1.3	**30** (9576)		0.064
6 (9350)		0.8	**31** (9585)		0.04
7 (9354)		0.6	**32** (9588)		0.094
8 (9385)		1.2	**33** (9590)		0.097
9 (9389)		1.8	**34** (9591)		0.425
10 (9395)		1.3	**35** (9593)		0.014

Code # [a]	Compounds	IC_{50} (μM)	Code # [a]	Compounds	IC_{50} (μM)
11 (9398)		1.5	**36** (9612)		0.087
12 (9401)		3	**37** (9615)		0.213
13 (9405)		0.3	**38** (9621)		0.453
14 (9426)		0.69	**39** (9639)		0.108
15 (9427)		0.53	**40** (9642)		0.341
16 (9435)		1.9	**41** (9643)		0.425
17 (9459)		0.65	**42** (9647)		0.033
18 (9460)		1.0	**43** (9648)		0.038
19 (9470)		1.3	**44** (9649)		0.188
20 (9471)		0.2	**45** (9650)		0.061
21 (9492)		1.0	**46** (9651)		0.071
22 (9524)		1.0	**47** (9653)		0.49
23 (9526)		1.4	**48** (9663)		0.175
24 (9538)		0.45	**49** (9665)		0.389
25 (9539)		0.22	**50** (9667)		0.076

[a] Code numbers in parentheses were obtained from the patent WO98/17648 filed by Xenova Ltd. Molecules used in the test set are identified by the prefix "t".

(Figure 2B). The conformer 9576_T1 was the extended conformation with the lowest energy in this family of conformers. We decided to choose an extended conformation of **1** because it is known that a modulator could bind to P-gp in an extended conformation such as propafenone modulator.[32] A few studies have shown a close correlation between the membrane interaction and MDR reversing activity of P-gp inhibitors. These studies have demonstrated that an extended conformation of propafenone was preferable for the activity, suggesting the importance of the membrane-mediated interactions in the MDR reversal.[32] The conformer 9576_T2 is the one close to the global minimum energy. We did not select the lowest energy conformation where the 3,4-tetrahydroisoquinoline moiety is between the benzene ring and the 3-isoquinoline group because it is unrealistic that the molecule could bind in such a closed conformation. Furthermore, it is known that the acceptor group

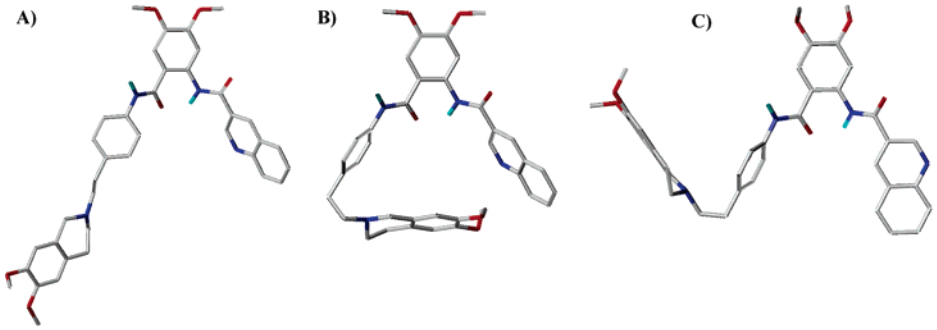


Figure 2. Spatial representation of tariquidar conformers: (A) 9576_T1 conformer; (B) 9576_T2 conformer; (C) 9576_T3 conformer.

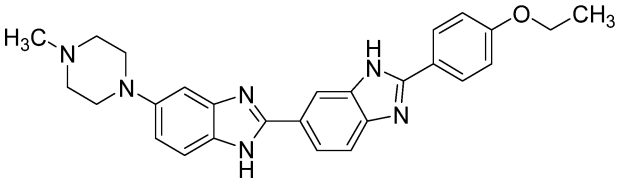


Figure 3. Molecular structure of selected Hoechst 33342 conformers.

on this ring and the amine are important to the activity.[36,38−40] A conformation slightly open to allow additional freedom to the acceptor group and the amine to the protein was selected.

b. Selection of Conformer 3 of Tariquidar. By use of the molecular features of Hoechst 33342, which binds at the H-site of P-gp such as **1**, atoms of the same type (viz. H-donor, H-acceptor, aromatic group) were overlapped. There is no X-ray crystal structure for the bound conformation of Hoechst 33342. Initial structures were generated using the cleanup procedure within SYBYL and energy-minimized using MAXIMIN2 (Powell method, 2000 iterations, and 0.05 kcal mol^{-1} $Å^{-1}$ energy gradient convergence criteria, with charge using the Gasteiger−Hückel method) followed by simulated annealing.[51] Simulated annealing was performed using these parameters: 100 cycles, 2000 K initial temperature for heating during 2000 fs to reach the equilibration, 0 K target temperature for 5000 fs of annealing time, and exponential annealing function. The 100 conformations representing local minima were then optimized using the MNDO method (full optimization, precise convergence) implemented in MOPAC 6. For Hoechst 33342, four clusters of conformations were identified and the lowest energy conformer in each cluster was taken for further analysis. The conformers differed in the position of the benzimidazolyl ring attached to the methylpiperazinyl group. So these conformers are different in the orientation of the nitrogen in the benzimidazolyl rings. One pair of conformers has the nitrogen of both imidazolyl rings either in the same orientation or in opposite directions. Because the heat of formation of each conformer was very close, the conformer exhibiting both imidazolyl nitrogens (−NH and ═N) in both benzimidazolyl rings in opposite directions was chosen. This conformer was used as a template to define the conformation of **1** and is illustrated in Figure 3. Then the conformation of **1** close to the selected conformation of Hoechst 33342 was established using the same procedure described above for Hoechst 33342. After simulated annealing, each conformer was evaluated, and those where both the NH and the carbonyl amide were in the same direction were discarded, leading to the exclusion of 71 conformers. The remaining 21 conformers were then energy-minimized using the Tripos force field (Powell method, 2000 iterations, and 0.05 kcal mol^{-1} $Å^{-1}$ energy gradient convergence criteria, with charge from the Gasteiger−Hückel method) and optimized using MNDO (full optimization, precise convergence, restricted amide bond rotation−MMOK). These 21 conformers were overlapped using the "Fit" procedure onto the previously selected conformer of Hoechst 33342. The structural features of Hoechst 33342 included three aromatic groups, four acceptors, and two donor groups. It has been shown that some of these features were part of the main pharmacophore of Hoechst 33342. Six out of these nine features were selected for further analysis. The following features were used in overlapping the conformers: (1) the oxygen of the ethoxy group of Hoechst 33342 to the methoxyl group of the anthranilamidyl ring of **1**, (2) the benzene ring of Hoechst 33342 to the benzene anthranilamidyl moiety of **1**, (3) nitrogen acceptor atoms (═N−) of Hoechst 33342 to the carbonyl of the amide of **1**, (4) nitrogen donor atoms (−NH) of Hoechst 33342 to the amine of the amide of **1**, (5) the centroid of aromatic ring of the benzimidazolyl of Hoechst 33342 to the centroid of the benzene ring of **1**, and (6) the nitrogen atom of Hoechst 33342 to the basic amine group of **1**. The template is shown in Figure 2, and the best fit is depicted in Figure 4 (rms = 0.587 Å). The energy of the 9576_T3 conformer is 16.116 kcal/mol.

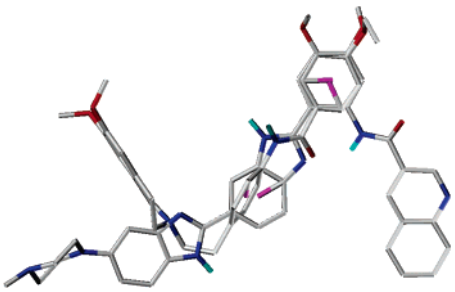


Figure 4. Best fit of the conformation T3 and Hoescht 33342 (conformer T3).

The structures of the other compounds were built based on the selected conformations of compound **1**. All structures were energy-minimized using molecular mechanics (Powell method, 2000 iterations, and 0.05 kcal mol^{-1} $Å^{-1}$ energy gradient convergence criteria), and charges were calculated using the Gasteiger−Hückel method. The geometry optimization was done utilizing MNDO (full optimization, precise convergence, restricted amide bond rotation−MMOK) as implemented in MOPAC as supplied by SYBYL 7.0.

c. Structure Alignment. Two different alignment rules were applied to align the training data set. The first one is Ar−O−N present in all molecules: the benzene ring, the oxygen of the amide near the spacer arm, and the basic nitrogen atom (Figure 5a). The second rule is Ar−Ar−N, which considers the role of the aromatic ring attached to the anthranilamidyl moiety and is applied by fitting the centroid of the benzene, the centroid of the quinolinyl moieties, and the basic nitrogen atom (Figure 5b). In both alignments compound **1** was used as the template

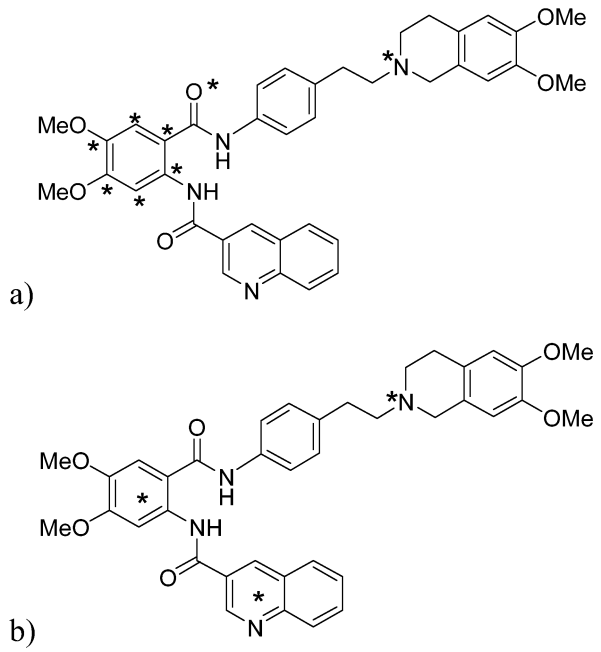


Figure 5. Representation of atoms used in the alignment rules 1 (a) and 2 (b). The asterisk (∗) represents the atoms or centroids that are aligned in the model.

molecule. Command "align database" was used to apply the first alignment rule and the "fit" procedure for the second rule alignment.

CoMFA and CoMSIA. The initial COMFA model was calculated using the SYBYL 7.0 molecular modeling software. For the calculation of charges, the Gasteiger−Hückel method was used as implemented in SYBYL 7.0. For the training compounds set, the CoMFA descriptors—steric (Lennard-Jones 6-12 potential) and electrostatic (Coulombic potential) field energies—were calculated using SYBYL. In general, the following standard characteristics were used to calculate the CoMFA fields: 4.0 Å extension beyond the van der Waals envelopes of the molecules, a distance-dependent dielectric constant ($1/r$), and an sp^3 carbon atom with +1.0 charge serving as the probe atom to the calculate steric and the electrostatic fields. The following standard CoMFA fields were calculated: steric (S), electrostatic (E), and both (B). The effects of changing several parameters were systematically investigated, including dielectric (function as $1/r$ vs constant), grid step size (1−3 Å), probe atom type (H^+, O^{3-}, and $C_{sp3}{}^+$), and the cutoff values for the steric and the electrostatic fields. Some others descriptors were also added to see their effects on the correlation such as molecular weight, dipole moment, molar refractivity, log P, polar volume, and polar surface area.

CoMSIA analysis was performed using the QSAR module in SYBYL 7.0. The five similarity indices in CoMSIA (steric (S), electrostatic (E), hydrophobic (H), H-bond donor (D), and H-bond acceptor (A) descriptors) were calculated using the probe atom $C_{sp3}{}^+$ with a radius of 1 Å and a +1.0 charge placed at the lattice points of the same region of grid as it was used for the CoMFA calculations. CoMSIA similarity indices (A_F) for a molecule j with atom i at a grid point q are calculated by

$$A_{F,kq}(j) = -\sum \omega_{probe,k}\omega_{ik}\, e^{-\alpha r_{iq}^2} \qquad (1)$$

where k represents the following physicochemical properties: steric, electrostatic, hydrophobic, H-bond donor, and H-bond acceptor. A Gaussian type distance dependence was used between the grid point q and each atom i of the molecule. A default value of 0.3 was used as the attenuation factor (R). Here, steric indices are related to the third power of the atomic radii, electrostatic descriptors are derived from atomic partial charges, hydrophobic fields are derived from atom-based parameters,[52] and H-bond donor and acceptor indices are obtained by a rule-based method based on experimental results.[53]

PLS Analysis. The conventional CoMFA and CoMSIA descriptors derived above were used as explanatory variables, and pIC_{50} ($-\log IC_{50}$) values were used as the target variable in PLS regression analyses to derive 3D QSAR models using the implementation in the SYBYL package. The predictive value of the models was evaluated by leave-one-out (LOO) cross-validation with SAMPLS. The cross-validated coefficient, q^2, was calculated using

$$q^2 = 1 - \frac{\sum(Y_{pred} - Y_{actual})^2}{\sum(Y_{actual} - Y_{mean})^2} \qquad (2)$$

where Y_{pred}, Y_{actual}, and Y_{mean} are predicted, actual, and mean values of the target property (pIC_{50}), respectively. $\sum(Y_{pred} - Y_{actual})^2$ is the predictive sum of squares (PRESS). The number of components giving the lowest PRESS value or the optimal number of components (ONC) was used to generate the final PLS regression models. The conventional correlation coefficient r^2 and its standard error, s, were subsequently computed for the final PLS models. CoMFA and CoMSIA coefficient maps were generated by interpolation of the pairwise products between the PLS coefficients and the standard deviations of the corresponding CoMFA or CoMSIA descriptor values. The bootstrapping procedure was used to validate each model. This is a procedure in which n random selections out of the original set of n objects are performed several times (100 times was used to have good statistical information) to simulate different samplings from a larger set of objects. In each run some objects may not be included in the PLS analysis (same method to determine the q^2), whereas some others might be included more than once. Confidence intervals for each term can be estimated from such a procedure, giving an independent measure of the stability of the PLS model.[54−56]

Results and Validation

CoMFA Analysis. A data set of 49 tariquidar derivatives was used with a wide spectrum of activities against AR 1.0 cells overexpressing P-glycoprotein. The data set of 49 modulators was aligned (Figure 6) to derive both the conventional CoMFA and CoMSIA models. Thus, a total of six models were generated with three different conformations of the same template using two alignment rules (alignments 1 and 2). An external test set of 13 compounds was used to determine the accuracy of the model (Figures 7 and 8).

The cross-validated r^2 (q^2) values for the six models relating the accumulation IC_{50} of daunorubicin on MDR cells are shown in the Supporting Information. By use of the default CoMFA settings, which included both steric and electrostatic fields, and by use of the first alignment rule, a cross-validated coefficient (q^2) of 0.490 with two optimum components with 9576_T1, a q^2 of 0.522 with three optimum components with 9576_T2 and a q^2 of 0.527 with five optimum components with 9576_T3 were observed. With the same options, by use of the second alignment rule, a cross-validated coefficient (q^2) of 0.449 with two optimum components with 9576_T1, a q^2 of 0.581 with three optimum components with 9576_T2 and a q^2 of 0.431 with two optimum components with 9576_T3 were observed.

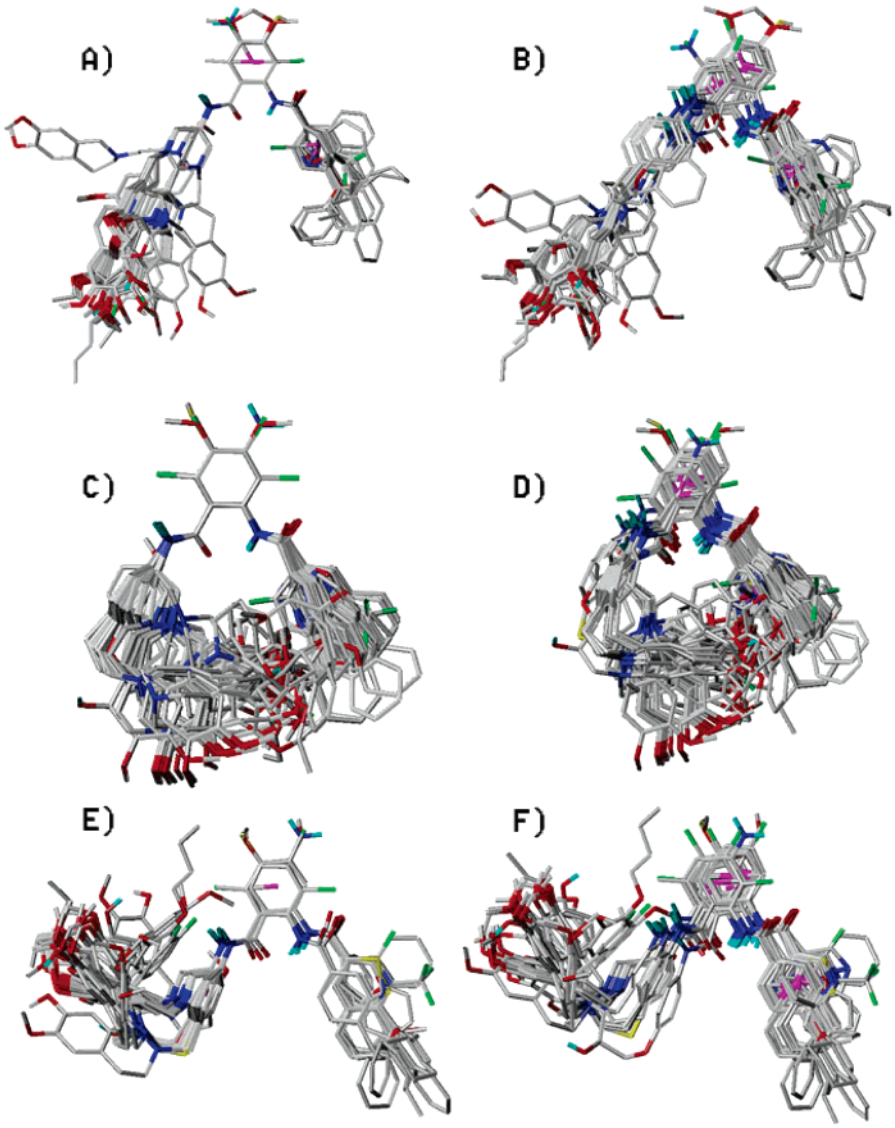


Figure 6. Alignment of P-gp modulator of the training set: (A) 9576_T1 and alignment rule 1; (B) 9576_T1 and alignment rule 2; (C) 9576_T2 and alignment rule 1; (D) 9576_T2 and alignment rule 2; (E) 9576_T3 and alignment rule 1; (F) 9576_T3 and alignment rule 2.

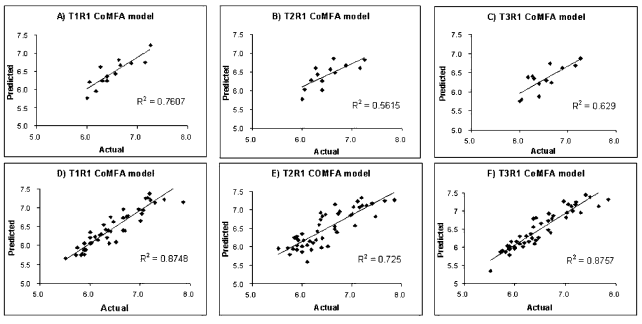


Figure 7. CoMFA predictions for the test (A−C) and the training (D−F) sets for MDR inhibitory activities against AR 1.0 cells overexpressing P-glycoprotein.

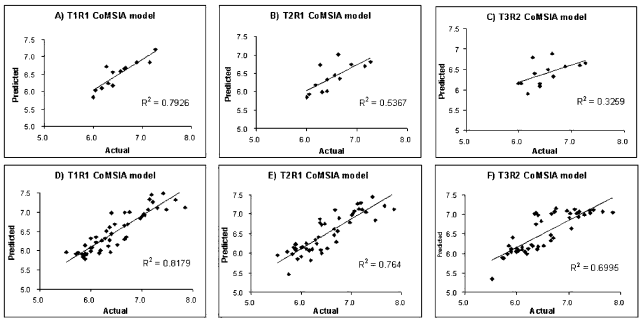


Figure 8. CoMSIA predictions for the test (A−C) and the training (D−F) sets for MDR inhibitory activities against AR 1.0 cells overexpressing P-glycoprotein.

The choice of the CoMFA options described below was based on maximizing the q^2 value.

The statistical parameters associated with all models are shown in Tables 2 and 3. The predicted pIC_{50} values for each training set of compounds and the residual values are given in Tables 4 and 5, respectively. The best model with 9576_T1 was obtained using the following options and alignment rule 1 (T1R1): steric and electrostatic fields with 2 kcal/mol cutoffs, $1/r$ for the dielectric function, 2.0 Å step size, a Csp^{3+} probe atom, and a grid box set at SYBYL's default position. This model had a q^2 value of 0.564 with four components, a conventional r^2 value of 0.874, and a standard error of estimate (SEE) of 0.215. This analysis yielded an $F_{(4,44)}$ value of 76.457. The final cross-validated model utilized 6587 of 6732 actual terms in the analysis. The best model with the 9576_T2 was obtained using the following options and alignment rule 2

Table 2. Statistical Data for QSAR Method with CoMFA and CoMSIA (Alignment 1)

	CoMFA			CoMSIA		
model name:[a]	T1R1	T2R1	T3R1	T1R1	T2R1	T3R1
fields[b]	S, E	E	S, E	S, E, A	E, H	A, D
q^2 [c]	0.564	0.57	0.532	0.502	0.556	0.549
r^2_{CV} [d]	0.559	0.544	0.515	0.537	0.528	0.561
STEP[e]	0.445, 0.420, 0.411, 0.402	0.434, 0.400	0.449, 0.428, 0.424, 0.422	0.444, 0.419, 0.420, 0.412	0.454, 0.407	0.398, 0.395
ONC[f]	4	2	4	4	2	2
SEE[g]	0.215	0.31	0.213	0.259	0.313	0.332
r^2 [h]	0.874	0.726	0.876	0.818	0.721	0.685
F [i]	76.457 (n1 = 4, n2 = 44)	60.838 (n1 = 2, n2 = 46)	77.728 (n1 = 4, n2 = 44)	49.358 (n1 = 4, n2 = 44)	59.454 (n1 = 3, n2 = 46)	49.989 (n1 = 2, n2 = 46)
prob. of $r^2 = 0$	0	0	0	0	0	0
SEE_{boot}	0.183 ± 0.1	0.271 ± 0.119	0.175 ± 0.086	0.224 ± 0.111	0.280 ± 0.130	0.314 ± 0.151
r^2_{boot} [j]	0.909 ± 0.26	0.774 ± 0.52	0.910 ± 0.026	0.855 ± 0.036	0.764 ± 0.059	0.714 ± 0.062
r^2_{pred} [k]	0.761	0.562	0.629	0.793	0.537	0.166
fraction						
S	0.433		0.439	0.557		
E	0.567	1	0.561	0.264	0.581	
H					0.419	
D						0.199D
A				0.179		0.801A

[a] Model name: T1R1 = template T1 with alignment rule 1; T2R1 = template T2 with alignment rule 1; T3R1 = template T3 with alignment rule 1. [b] Fields used: S = steric, E = electrostatic, H = hydrophobicity, A = hydrogen bond acceptor, D = hydrogen bond donor. [c] q^2 = cross-validated correlation coefficient from LOO. [d] r^2_{CV} = cross-validated correlation coefficient. [e] STEP = standard error of prediction. [f] ONC optimal number of components. [g] SEE = standard error of estimate. [h] $F = r^2/(1 - r^2)$. [i] SEE_{boot} = standard error of estimate from bootstrapping. [j] r^2_{boot} = correlation coefficient from bootstrapping. [k] r^2_{pred} = correlation coefficient of the prediction of the test set.

Table 3. Statistical Data for QSAR Method with CoMFA and COMSIA (Alignment 2)

	COMFA			COMSIA		
model name:[a]	T1R2	T2R2	T3R2	T1R2	T2R2	T3R2
field[b]	E	E	E	E	S, E	H, A, D
q^2 [c]	0.507	0.645	0.496	0.506	0.646	0.521
r^2_{CV} [d]	0.525	0.608	0.503	0.504	0.644	0.535
STEP[e]	0.446, 0.408	0.448, 0.382, 0.375	0.461, 0.423, 0.422	0.451, 0.417	0.458, 0.364, 0.357	0.439, 0.404
ONC[f]	2	3	3	2	3	2
SEE[g]	0.346	0.228	0.261	0.36	0.258	0.325
r^2	0.659	0.855	0.81	0.631	0.814	0.7
F [h]	44.499 (n1 = 2, n2 = 46)	88.714 (n1 = 3, n2 = 45)	63.827 (n1 = 3, n2 = 45)	39.334 (n1 = 2, n2 = 45)	65.493 (n1 = 3, n2 = 45)	53.550 (n1 = 2, n2 = 46)
prob. of $r^2 = 0$	0	0	0	0	0	0
SEE_{boot} [i]	0.303 ± 0.122	0.198 ± 0.097	0.219 ± 0.112	0.329 ± 0.154	0.230 ± 0.106	0.290 ± 0.127
r^2_{boot} [j]	0.725 ± 0.058	0.889 ± 0.031	0.860 ± 0.035	0.676 ± 0.061	0.848 ± 0.032	0.750 ± 0.048
r^2_{pred} [k]	0.352	0.036	0.277	0.429	0.091	0.326
fraction						
S					0.184	
E	1	1	1	1	0.816	
H						0.469
D						0.179
A						0.352

[a] Model name: T1R2 = template T1 with alignment rule 2; T2R2 = template T2 with alignment rule 2; T3R2 = template T3 with alignment rule 2. [b] Fields used: S = steric, E = electrostatic, H = hydrophobicity, A = hydrogen bond acceptor, D = hydrogen bond donor. [c] q^2 = cross-validated correlation coefficient from LOO. [d] r^2_{CV} = cross-validated correlation coefficient. [e] STEP = standard error of prediction. [f] ONC = optimal number of components. [g] SEE = standard error of estimate. [h] $F = r^2/(1 - r^2)$. [i] SEE_{boot} = standard error of estimate from bootstrapping. [j] r^2_{boot} = correlation coefficient from bootstrapping. [k] r^2_{pred} = correlation coefficient of the prediction of the test set.

(T2R2): electrostatic fields with 10 kcal/mol cutoffs, 1/r for the dielectric function, 2.0 Å step size, a C_{sp3}^+ probe atom, and a grid box set at SYBYL's default position. This model had a q^2 value of 0.645 for three components, a conventional r^2 value of 0.855, and a SEE of 0.228. The analysis yielded an $F_{(3,45)}$ value of 88.714. The final cross-validated model utilized 3460 of 3465 actual descriptor columns in the analysis. The best final model with the 9576_T3 was obtained using the following options and alignment rule 1 (T3R1): steric fields with 35 kcal/mol cutoffs, electrostatic fields with 1 kcal/mol cutoffs, 1/r for the dielectric function, 2.0 Å step size, a C_{sp3}^+ probe atom, and a grid box set at SYBYL's default position. This model had a q^2 value of 0.532 for four components, a conventional r^2 value of 0.876, and a SEE of 0.213. The analysis yielded an $F_{(4,44)}$ value of 77.728. The final cross-validated model utilized 5296 of 5408 actual descriptor columns in the analysis. The models with the descriptors MD and MW were discarded although the cross-validated coefficient q^2 was higher, the statistical F value was significantly lower, and the non-cross-validated r^2 was comparable to the model without these descriptors (data not show).

A q^2 of 0.5 is generally considered an indication that the model is internally predictive; thus, the q^2 values obtained in the present case are all near that number varying from 0.496 to 0.646. To validate our models, the bootstrapping function was used to determine the error on the r^2 (r^2_{boot}) and the SEE (SEE_{boot}) of the model. This statistical parameter gave an idea of the accuracy of the model. In fact, the best model of each conformation has a small error on the SEE and r^2. T1R1 and T3R1 CoMFA models were the best models obtained with 9576_T1 and 9576_T3. These models have an r^2_{boot} of 0.909 ± 0.26 and 0.910 ± 0.026, respectively. They also have a SEE_{boot} of 0.183 ± 0.1 and 0.175 ± 0.086, respectively. The T2R2 CoMFA model was the best model with 9576_T2 and

Table 4. CoMFA Actual and Predicted Activities for Training Set Molecules (Alignment 1)

		T1R1		T2R1		T3R1	
compd	actual	calcd	residual	calcd	residual	calcd	residual
2	6.39	6.06	0.33	5.98	0.41	6.09	0.3
3	6.41	6.38	0.03	6.23	0.18	6.19	0.22
4	5.89	5.88	0.01	5.92	−0.04	5.96	−0.08
5	5.89	5.89	−0.0018	6.21	−0.33	5.79	0.1
6	6.1	6.23	−0.13	6.05	0.04	6.1	−0.01
7	6.22	6.3	−0.08	6.1	0.12	6.31	−0.09
8	5.92	5.89	0.03	6.18	−0.26	6.15	−0.23
9	5.75	5.94	−0.2	5.79	−0.04	5.91	−0.16
10	5.89	5.76	0.13	6.04	−0.15	5.94	−0.05
11	5.82	5.75	0.07	5.94	−0.11	5.87	−0.05
12	5.52	5.66	−0.14	5.95	−0.43	5.35	0.17
13	6.52	6.09 (outlier)	0.43	6.01	0.52	6.66	−0.13
14	6.16	6.15	0.01	6.15	0.01	6.01	0.15
15	6.28	6.55	−0.28	6.19	0.08	6.15	0.12
16	5.72	5.75	−0.03	5.96	−0.24	5.87	−0.15
17	6.19	6.27	−0.09	5.92	0.26	6.12	0.07
18	6	6.06	−0.06	5.86	0.14	5.98	0.02
19	5.89	6.06	−0.18	6.28	−0.4	6.02	−0.13
20	6.7	6.39	0.31	6.39	0.31	6.41	0.29
21	6	6.35	−0.35	6.35	−0.35	6	−0.0022
22	6	6.2	−0.2	6.01	−0.01	5.96	0.04
23	5.85	5.78	0.08	6.25	−0.39	6.04	−0.19
24	6.66	6.73	−0.07	6.55	0.11	6.48	0.18
25	6.66	6.4	0.26	6.48	0.18	6.73	−0.08
26	7.3	7.14	0.16	7.12	0.18	6.95	0.35
27	6.1	6.23	−0.14	5.59	0.51	6.3	−0.2
28	6.03	6.07	−0.04	6.17	−0.14	6.16	−0.13
29	7.66	7.55	0.11	7.25	0.41	7.13 (outlier)	0.53
30	7.19	7.38	−0.19	7.32	−0.13	7.18	0.01
31	7.4	7.54	−0.14	7.17	0.23	7.45	−0.05
32	7.03	6.84	0.19	7.08	−0.05	6.95	0.08
33	7.01	6.65	0.36	6.88	0.13	6.82	0.19
34	6.37	6.39	−0.02	6.73	−0.36	6.78	−0.41
35	7.85	7.15 (outlier)	0.7	7.27	0.58	7.32 (outlier)	0.53
36	7.06	6.93	0.13	6.57	0.49	7.19	−0.13
37	6.67	6.95	−0.28	7.15 (outlier)	−0.48	6.93	−0.26
38	6.34	6.21	0.13	6.59	−0.24	6.25	0.09
39	6.97	6.96	0.01	6.85	0.12	7.27	−0.3
40	6.47	6.62	−0.15	6.87	−0.41	6.26	0.21
41	6.37	6.4	−0.03	6.92	−0.55	6.56	−0.19
42	7.48	7.22	0.26	6.81 (outlier)	0.67	7.39	0.09
43	7.42	7.63	−0.21	7.17	0.25	7.62	−0.2
44	6.73	6.76	−0.03	6.9	−0.17	6.78	−0.05
45	7.21	7.21	−0.00047	7.12	0.09	7.25	−0.04
46	7.15	7.26	−0.11	7.07	0.08	7	0.15
47	6.31	6.42	−0.11	6.42	−0.11	6.37	−0.06
48	6.76	6.78	−0.02	6.95	−0.19	6.87	−0.12
49	6.41	6.75	−0.34	6.82	−0.41	6.81	−0.4
50	7.12	7.25	−0.13	7.23	−0.11	7.13	−0.01

has an r^2_{boot} of 0.889 ± 0.031 and a SEE_{boot} of 0.198 ± 0.097. Furthermore, the accumulation activity against AR 1.0 cells overexpressing P-glycoprotein for the 13 compounds was predicted from the corresponding external test set. Thirteen compounds were used to validate our CoMFA model, representing 26% of the training set. The models using alignment rule 1 moderately predicted for efficacy of the external test set. These results are not surprising and follow the statistics of every model. In fact, the best prediction of the test set is with the 9576_T1, followed by 9576_T3 and 9576_T2. They respectively have a predictive r^2 (r^2_{pred}) of 0.761, 0.562, and 0.629. The external test set was poorly to moderately predicted with the model using conformations 9576_T1 and 9576_T3 and alignment rule 2. The predictive r^2 of the test set are 0.352 and 0.277, respectively. With 9576_T2, the CoMFA model was not able to predict with accuracy the external test set with an r^2_{pred} of 0.036. Unexpectedly, the CoMFA model having the best the r^2 (0.855) and q^2 (0.645) was the least predictable. It is known that one could have a very good internal prediction and a very poor external prediction.[57,58]

Futhermore, each CoMFA model has one or two outliers in the training set. In general, residual values greater than 2 times the standard error of the residuals generated in the validation procedure are considered outliers.[59] In the T1R1 CoMFA model, the outliers were compounds **13** and **35**. In the T1R2 CoMFA model, the outliers were compounds **29** and **35**. In the T2R1 CoMFA model, the outlier was compound **42**. In the T2R2 CoMFA model, the outliers were compounds **35** and **37**. In the T3R1 CoMFA model, the outliers were compounds **29** and **35**. In the T3R2 CoMFA model, the outliers were compounds **35** and **49**. So compound **35** seems to be a major outlier for almost all models. The explanation for this outlier is hard to define because this molecule has good similarities compared to **1**. The only difference is the dimethoxy groups on the anthranilic moiety are replaced by chlorine at the 3 position of the anthranilic moiety. However, all of those molecules can be outliers for each model with more than 95% confidence. The graphs of the actual pIC_{50} versus the predicted pIC_{50} values for the training set and test set by the conventional CoMFA with 9576_T1, 9576_T2, and 9576_T3

Table 5. CoMFA Actual and Predicted Activities for the Training Set Molecules (Alignment 2)

		T1R2		T2R2		T3R2	
compd	actual	calcd	residual	calcd	residual	calcd	residual
2	6.39	5.93	0.46	6.13	0.26	6.22	0.17
3	6.41	6.31	0.09	6.26	0.15	6.29	0.12
4	5.89	5.89	−0.0017	5.95	−0.06	6.02	−0.13
5	5.89	5.96	−0.08	6.08	−0.2	6.16	−0.27
6	6.1	6.22	−0.12	6.1	−0.0012	6.06	0.04
7	6.22	6.21	0.01	6.06	0.16	6.25	−0.03
8	5.92	5.84	0.08	6.07	−0.15	6.27	−0.35
9	5.75	6.15	−0.4	5.71	0.03	5.74	0.01
10	5.89	5.88	0.01	6.07	−0.18	5.92	−0.03
11	5.82	5.88	−0.06	5.91	−0.09	6.04	−0.22
12	5.52	5.92	−0.39	5.94	−0.42	5.27	0.25
13	6.52	6.17	0.35	6.07	0.45	6.36	0.17
14	6.16	5.74	0.42	6.13	0.03	6.12	0.04
15	6.28	6.18	0.1	6.19	0.09	6.06	0.22
16	5.72	5.86	−0.14	5.9	−0.18	5.87	−0.15
17	6.19	6.2	−0.02	6.1	0.09	6.13	0.06
18	6	6.23	−0.23	6.04	−0.04	6.06	−0.06
19	5.89	6.21	−0.32	6.11	−0.23	6.2	−0.31
20	6.7	6.25	0.45	6.26	0.44	6.25	0.44
21	6	6.23	−0.23	6.11	−0.11	6.19	−0.19
22	6	6.07	−0.07	5.76	0.24	6.11	−0.11
23	5.85	6.36	−0.51	6.06	−0.21	6.29	−0.44
24	6.66	6.39	0.27	6.49	0.16	6.5	0.16
25	6.66	6.27	0.38	6.45	0.21	6.63	0.03
26	7.3	6.84	0.46	7.18	0.12	7.04	0.26
27	6.1	6.28	−0.18	6.21	−0.11	6.11	−0.02
28	6.03	6.37	−0.34	6.23	−0.2	5.79	0.24
29	7.66	6.96 (outlier)	0.7	7.42	0.24	7.14	0.52
30	7.19	6.99	0.21	7.47	−0.27	7.42	−0.23
31	7.4	7.08	0.32	7.38	0.02	7.2	0.2
32	7.03	6.81	0.22	7.18	−0.15	7	0.03
33	7.01	6.63	0.38	6.88	0.13	6.85	0.16
34	6.37	6.52	−0.15	6.63	−0.26	6.78	−0.41
35	7.85	6.9 (outlier)	0.95	7.31 (outlier)	0.54	7.33 (outlier)	0.52
36	7.06	7.06	−0.0011	7.02	0.04	6.69	0.37
37	6.67	6.85	−0.18	7.23 (outlier)	−0.56	7.12	−0.45
38	6.34	6.76	−0.41	6.33	0.01	6.3	0.04
39	6.97	7.12	−0.15	6.52	0.45	7.13	−0.17
40	6.47	7.09	−0.62	6.51	−0.05	6.34	0.12
41	6.37	7.05	−0.68	6.52	−0.14	6.61	−0.24
42	7.48	7.13	0.35	7.46	0.02	7.13	0.35
43	7.42	7.21	0.21	7.52	−0.1	7.22	0.2
44	6.73	7.11	−0.38	6.58	0.15	6.76	−0.03
45	7.21	7.19	0.02	7.2	0.01	7.27	−0.06
46	7.15	7.21	−0.06	7.16	−0.01	7.21	−0.06
47	6.31	6.3	0.01	6.31	0.0025	6.07	0.24
48	6.76	6.9	−0.14	6.68	0.08	6.95	−0.19
49	6.41	6.82	−0.41	6.66	−0.25	6.99 (outlier)	−0.58
50	7.12	7.28	−0.16	7.25	−0.13	7.32	−0.2

models based on the MDR modulator inhibitory activity are shown in Figure 3.

CoMSIA Analysis. Six CoMSIA models, one for each conformation, were generated from the same training sets and the same alignment rules used in CoMFA (alignments 1 and 2). The cross-validated r^2 (q^2) values of the six models, which result from the various CoMSIA options, are shown in the Supporting Information. The statistical parameters associated with all models are shown in Tables 2 and 3. The q^2 values obtained from all CoMSIA models vary from 0.502 to 0.646. By use of all fields from CoMSIA and the first alignment rules, a moderately significant model was observed with a q^2 of 0.479 for four components with 9576_T1, a q^2 of 0.504 for two components with 9576_T2, and a q^2 of 0.52 for two components with 9576_T3. Moderate to high significance models were observed with the second alignment rule. In fact, a q^2 of 0.482 for two components with 9576_T1, a q^2 of 0.631 for four components with 9576_T2, and a q^2 of 0.51 for two components with 9576_T3 were obtained. However, in each model, when individual fields or combinations of the steric, electrostatic, hydrophobic, hydrogen bond acceptor, and hydrogen bond donor were taken into account, better q^2 values were obtained. Indeed, the T1R1 CoMSIA model was the best model with 9576_T1 and used the electrostatic, steric, and hydrogen bond acceptor fields. This model had a q^2 value of 0.502 for four optimum components, a conventional r^2 value of 0.818, and a SEE of 0.259. This yielded an $F_{(4,44)}$ value of 49.358. The T2R2 CoMSIA model was the best model with 9576_T2 and used steric and electrostatic fields. This model had a q^2 value of 0.646 with three optimum components, a conventional r^2 value of 0.814, and a SEE of 0.258. This yielded an $F_{(3,45)}$ value of 65.493. The T3R2 CoMSIA model was the best model with 9576_T3 and used hydrophobicity, hydrogen bond acceptor, and hydrogen bond donor fields and alignment rule 2. This model had a q^2 value of 0.521 for two optimum components, a conventional r^2 value of 0.7, and a SEE of 0.325. This yielded an $F_{(2,46)}$ value of 53.550.

The predicted pIC_{50} values for each training set compounds and the residual values are given in Tables 6 and 7. CoMFA was used to validate our CoMSIA models. The T1R1 CoMSIA

Table 6. CoMSIA Actual and Predicted Activities for the Training Set Molecules (Alignment 1)

		T1R1		T2R1		T3R1	
compd	actual	calcd	residual	calcd	residual	calcd	residual
2	6.39	5.96	0.42	6.08	0.31	6.08	0.31
3	6.41	6.44	−0.03	6	0.4	6.07	0.34
4	5.89	5.93	−0.04	6.08	−0.19	6.08	−0.19
5	5.89	5.78	0.11	6.22	−0.33	6.08	−0.19
6	6.1	6.22	−0.12	6.11	−0.01	6.09	0.01
7	6.22	6.26	−0.04	6.1	0.12	6.09	0.13
8	5.92	5.92	0.005	5.85	0.08	6.11	−0.19
9	5.75	5.95	−0.21	5.46	0.29	5.71	0.04
10	5.89	5.89	−0.01	6.11	−0.22	6.09	−0.2
11	5.82	5.92	−0.1	5.98	−0.16	6.07	−0.25
12	5.52	5.96	−0.44	5.95	−0.43	5.64	−0.12
13	6.52	6.15	0.38	6.12	0.4	6.21	0.31
14	6.16	5.94	0.22	6.07	0.09	6.06	0.1
15	6.28	6.32	−0.04	6.34	−0.06	6.35	−0.07
16	5.72	5.92	−0.2	6.03	−0.31	5.71	0.02
17	6.19	6.01	0.18	5.81	0.38	6.07	0.12
18	6	5.98	0.02	6.15	−0.15	6.09	−0.09
19	5.89	6.14	−0.26	6.23	−0.34	6.25	−0.36
20	6.7	6.35	0.35	6.29	0.41	6.78	−0.09
21	6	6.32	−0.32	6.15	−0.15	6.2	−0.2
22	6	6.07	−0.07	5.91	0.09	6.23	−0.23
23	5.85	5.84	0.02	6.14	−0.28	6.2	−0.34
24	6.66	6.65	0.01	6.62	0.04	6.24	0.42
25	6.66	6.3	0.36	6.46	0.2	6.23	0.43
26	7.3	7.11	0.19	7.05	0.25	6.99	0.31
27	6.1	6.35	−0.26	6.26	−0.17	5.85	0.24
28	6.03	6.08	−0.05	6.16	−0.13	6.15	−0.12
29	7.66	7.32	0.34	7.21	0.45	6.99	0.67
30	7.19	7.45	−0.25	7.29	−0.09	6.98	0.21
31	7.4	7.54	−0.14	7.51	−0.11	6.97	0.43
32	7.03	6.94	0.09	7.02	0.01	6.99	0.04
33	7.01	6.89	0.12	6.97	0.04	7	0.01
34	6.37	6.27	0.1	6.77	−0.4	6.95	−0.58
35	7.85	7.12 (outlier)	0.73	7.13	0.72	6.99	0.86
36	7.06	6.92	0.14	7.07	−0.01	6.46	0.6
37	6.67	6.99	−0.32	7.1	−0.43	6.98	−0.31
38	6.34	6.12	0.23	6.41	−0.07	6.92	−0.57
39	6.97	6.83	0.14	6.78	0.18	7.06	−0.09
40	6.47	6.68	−0.21	6.75	−0.28	6.68	−0.21
41	6.37	6.61	−0.24	6.87	−0.5	6.94	−0.57
42	7.48	7.06	0.42	6.84	0.64	7.13	0.35
43	7.42	7.49	−0.07	7.44	−0.02	7.46	−0.04
44	6.73	6.68	0.05	6.56	0.17	6.97	−0.24
45	7.21	7.27	−0.06	7.11	0.1	6.99	0.22
46	7.15	7.33	−0.18	7.14	0.01	7.06	0.09
47	6.31	6.51	−0.2	6.25	0.06	6.47	−0.16
48	6.76	7	−0.24	6.89	−0.14	7.07	−0.31
49	6.41	6.97	−0.56	6.73	−0.32	6.99	−0.58
50	7.12	7.06	0.06	7.27	−0.15	7.06	0.06

model was the best with 9576_T1 and has an r^2_{boot} of 0.855 ± 0.036 and a SEE_{boot} of 0.224 ± 0.111. The T2R2 and T3R2 CoMSIA models were the best with 9576_T2, and 9576_T3 was obtained with alignment rule 2 with an r^2_{boot} of 0.848 ± 0.032 and 0.75 ± 0.048, respectively. They also have also a SEE_{boot} of 0.230 ± 0.0106 and 0.290 ± 0.127, respectively. Furthermore, the same external test set was used to validate the CoMSIA models. The external test set was predicted poorly for efficacy using the alignment rule 1. These results are not surprising and follow the statistics of every model. In fact, the best predictions of the test set are with 9576_T1, followed by 9576_T2 and 9576_T3. They respectively have a predictive r^2 (r^2_{pred}) of 0.793, 0.537, and 0.166. We are able to predict poorly the external test set with the model using 9576_T1 and 9576_T3 and alignment rule 2. The predictive r^2 of the test set are respectively 0.429 and 0.326. With 9576_T2, the CoMSIA models were not able to predict with accuracy the external test set with an r^2_{pred} of 0.091. Again, unexpectedly, the CoMSIA model that had better statistical results (r^2 of 0.814 and a q^2 of 0.646) could not correctly predict the external test set.

Table 7. CoMSIA Actual and Predicted Activities for the Training Set Molecules (Alignment 2)

		T1R2		T2R2		T3R2	
compd	actual	calcd	residual	calcd	residual	calcd	residual
2	6.39	5.94	0.45	6.05	0.34	6.11	0.28
3	6.41	6.21	0.19	6.16	0.25	6.17	0.24
4	5.89	5.93	−0.04	6.03	−0.14	6.04	−0.16
5	5.89	5.96	−0.07	6.11	−0.22	6.08	−0.19
6	6.1	6.24	−0.15	6.13	−0.03	6.16	−0.06
7	6.22	6.21	0.01	6.08	0.14	6.12	0.1
8	5.92	5.78	0.14	6.11	−0.19	6.4	−0.48
9	5.75	6.15	−0.4	5.62	0.13	5.87	−0.12
10	5.89	5.95	−0.06	6.07	−0.18	6.12	−0.24
11	5.82	5.92	−0.1	5.97	−0.15	5.99	−0.16
12	5.52	5.82	−0.3	6.05	−0.53	5.34	0.18
13	6.52	6.19	0.33	6.08	0.44	6.19	0.34
14	6.16	5.71	0.45	6.13	0.03	6.04	0.12
15	6.28	6.14	0.14	6.23	0.05	6.22	0.05
16	5.72	5.92	−0.2	5.99	−0.27	5.88	−0.16
17	6.19	6.16	0.03	6.02	0.16	5.99	0.19
18	6	6.08	−0.08	5.94	0.06	6.04	−0.04
19	5.89	6.36	−0.47	6.21	−0.32	6.1	−0.22
20	6.7	6.33	0.37	6.26	0.44	6.39	0.31
21	6	6.27	−0.27	6.21	−0.21	6.13	−0.13
22	6	6.21	−0.21	5.86	0.14	6.05	−0.05
23	5.85	6.47	−0.62	6.13	−0.27	6.19	−0.34
24	6.66	6.46	0.2	6.47	0.19	6.33	0.33
25	6.66	6.34	0.32	6.44	0.22	6.47	0.19
26	7.3	6.82	0.48	7.16	0.14	6.99	0.31
27	6.1	6.25	−0.15	6.1	−0.0004	6.09	0.0025
28	6.03	6.27	−0.24	6.28	−0.25	6.07	−0.04
29	7.66	6.91	0.75	7.34	0.32	7.07	0.59
30	7.19	6.97	0.23	7.43	−0.23	6.94	0.25
31	7.4	6.93	0.47	7.4	0.0033	7.1	0.3
32	7.03	6.8	0.23	7.14	−0.11	7.02	0.01
33	7.01	6.71	0.3	7.04	−0.03	6.91	0.1
34	6.37	6.63	−0.26	6.7	−0.33	6.75	−0.38
35	7.85	6.84	1.01	7.26	0.59	7.05	0.8
36	7.06	7.04	0.02	7.04	0.02	6.64	0.42
37	6.67	6.81	−0.14	7.2	−0.53	7.01	−0.34
38	6.34	6.71	−0.36	6.24	0.11	7.01	−0.66
39	6.97	7.14	−0.17	6.55	0.41	7.09	−0.12
40	6.47	7.11	−0.65	6.56	−0.1	6.83	−0.37
41	6.37	7.09	−0.72	6.52	−0.15	7.04	−0.67
42	7.48	7.09	0.39	7.22	0.26	7.11	0.37
43	7.42	7.29	0.13	7.62	−0.2	7.04	0.38
44	6.73	7.11	−0.38	6.56	0.17	7.07	−0.34
45	7.21	7.1	0.11	7.18	0.03	7.07	0.14
46	7.15	7.13	0.02	7.15	−0.002	7.02	0.13
47	6.31	6.34	−0.03	6.14	0.17	6.18	0.13
48	6.76	6.87	−0.11	6.68	0.07	7.16	−0.41
49	6.41	6.8	−0.39	6.79	−0.38	6.98	−0.57
50	7.12	7.28	−0.16	7.17	−0.05	7.14	−0.02

Futhermore, each CoMSIA model has two or three outliers in the training set. In the T1R1 CoMSIA model, compounds **35** and **49** were greater than 2 times the standard error and were defined as outliers. In the T1R2 CoMSIA model, the outliers were compounds **29**, **35**, and **41**. In the T2R1 CoMSIA model, the outliers were compounds **35** and **42**. In the T2R2 CoMSIA model, the outliers were compounds **12**, **35**, and **37**. In the T3R1 CoMSIA model, the outliers were compounds **29** and **35**. In the T3R2 CoMSIA model, the outliers were compounds **35**, **38**, and **41**. Again, compound **35** seems to be a major outlier for almost all models. The graphs of the actual pIC_{50} versus the predicted pIC_{50} values for the training set and test set by the conventional CoMSIA with 9576_T1, 9576_T2, and 9576_T3 models based on the MDR modulator inhibitory activity are shown in Figure 3.

CoMFA and CoMSIA Contours Maps. The q^2 values in the training sets associated with these CoMSIA models are generally similar or superior to those of the CoMFA model. However, in each case the r^2 value are lower compared to

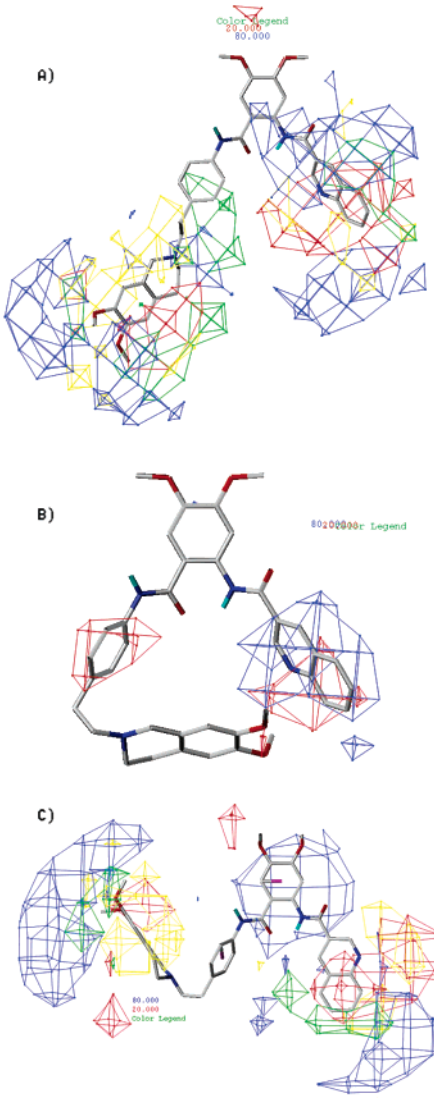


Figure 9. Contour map of steric and electrostatic fields (standard deviation × coefficient) generated with the CoMFA model based on accumulation of daunorubicin on MDR cell. Color coding is as follows. Blue indicates that a positive charge favors high affinity, whereas red indicates that a positive charge does not favor high affinity. Yellow indicates regions where bulky groups decrease activity, whereas green indicates regions where bulky groups increase activity.

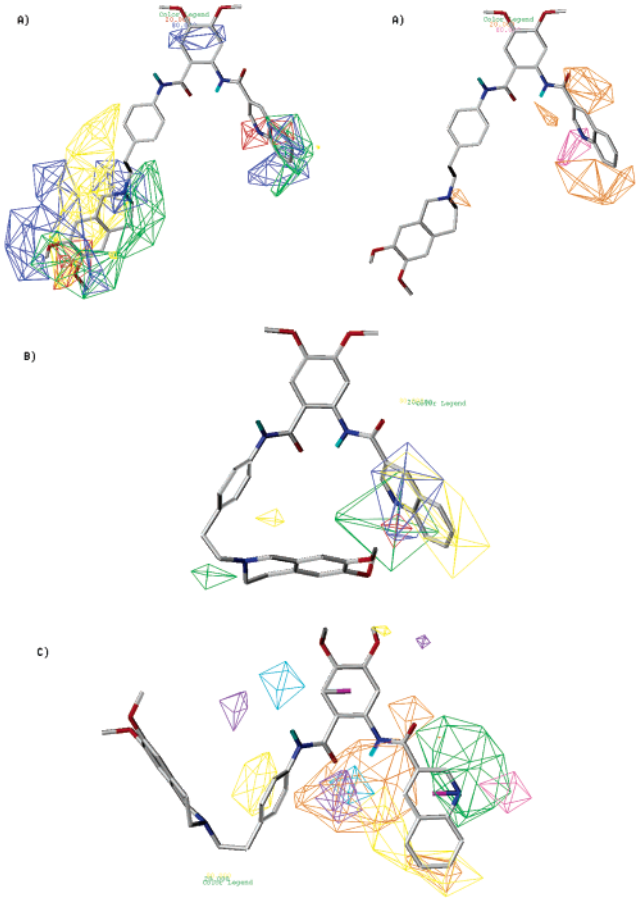


Figure 10. Contour maps generated with the CoMSIA model based on accumulation of daunorubicine in the MDR cell: (A) T1R1 CoMSIA model; (B) T2R1 CoMSIA model; (C) T3R2 CoMSIA model. Color coding is as follows. Blue indicates that a positive charge favors high affinity, whereas red indicates that a positive charge does not favor high affinity. Yellow indicates regions where hydrophobic groups increase the activity, whereas green indicates regions where hydrophobic groups decrease the activity. Magenta indicates that a hydrogen bond acceptor favors high affinity, whereas orange indicates unfavorable contributions of the hydrogen bond acceptor. Cyan indicates that a hydrogen bond donor favors high affinity, whereas purple indicates unfavorable contributions of the hydrogen bond donor. Important note: In the part A, the yellow indicates regions where bulky groups decrease activity, whereas green indicates regions where bulky groups increase activity.

CoMFA models except with alignment rule 1 and 9576_T2, which was very similar. Thus, the CoMFA models have even better internal predictive power than the CoMSIA model. Furthermore, for the predictive r^2 of the external set of every model, CoMFA seems to have better or similar external predictive power than the CoMSIA with alignment rule 1 and to have lower external predictive power when using alignment rule 2. After consideration of both the internal and external predictive power of the models, the best CoMFA and CoMSIA selected for every conformation to construct the stdev*coefficient contour maps were T1R1, T2R1, and T3R1 for the CoMFA and were T1R1, T2R1, and T3R2 for the COMSIA models. For the CoMFA the steric field, the green (sterically favorable) and yellow (sterically unfavorable) contours represent 80% and 20% level contributions. The red (negative charge favorable) and blue (negative charge unfavorable) contours in the CoMFA electrostatic field contours also represent 80% and 20% level contributions (Figure 9), respectively. CoMSIA analyses were also selected to construct contour maps (Figure 10). In the CoMSIA electrostatic field, the red (negative charge favorable) and blue (negative charge unfavorable) contours represent 80% and 20% level contributions, respectively. In the steric field, the green (sterically favorable) and yellow (sterically unfavorable) contours represent 80% and 20% level contributions, respectively. In the hydrogen bond acceptor field, the magenta (favorable) and orange (unfavorable) represent 80% and 20% level contributions, respectively. In the hydrophobicity field, the yellow (hydrophobic favorable) and green (hydrophobic unfavorable) contours represent 80% and 20% level contributions, respectively.

The contours maps of the T1R1 CoMFA and T3R1 CoMFA models were distributed in the entire molecule. In region A of the molecules for both models, the presence of a heteroatom at the 3 position (the amine on the isoquinoline) falls into a negative favorable red region, suggesting that the negative charge was important to the activity. This was confirmed with almost all compounds in the training set. The blue positive region at the ortho position of the amide in both models indicated that a negative charge was not well abided. This was confirmed in the moderate to weak potency of a few compounds

in the training set with an amine heteroatom at that position. For example, compound **9**, **27**, and **47** bear an amine group at that position, while compound **34** has a chlorine atom at that position. Other blue regions, for both models, near the second aromatic ring of the isoquinoline were favorable to a positive charge. In that case, this positive charge could reflect the hydrophobicity of the ring that is necessary to the potency. The negatively charged and favorable red region on top of the second aromatic ring of the isoquinoline, in both models, suggests that negative charge was important to the activity. The same conclusion was also shown with the T1R1 CoMFA model in the A region. Furthermore, in both models, the green region at the top of the isoquinolinyl group indicates that a bulky group may be necessary to increase the potency. However, the yellow contour at the bottom of the isoquinoline in both models suggests that bulky groups are not well accepted. The big favorable electrostatic blue region in both models close to the amide reveals that a positive charge will increase the antagonistic activity. In fact, in the electrostatic region, the influence of the hydrogen bond acceptor carbonyl amide as part of this field could be considered. So the carbonyl amide could be part of the electrostatic fields and help the activity.

In region B of **1**, the small blue positive region in the T1R1 CoMFA model at position 5 of the amide and the big region in the T3R1 CoMFA model indicate that a negative charge was not permitted. This was confirmed to decrease the potency of a few compounds such as 9401 that bear the chlorine atom. Also, this large blue region of the T3R1 CoMFA model indicates the positive influence of the aromatic ring to the activity. The small red region near the dimethoxyl group on the anthranilic moiety in both models indicates that a negative charge is favorable for high affinity to P-gp. This is in concordance with potent molecules such as **1**, **35**, and **50** in the training set, molecules that are negatively charged with the methoxyl group, chlorine atom, and fluorine atom at this position, respectively.

In region C, the green region in the T1R1 CoMFA model near the aromatic ring indicates that bulky groups are well abided to increase the potency. This was confirmed with compounds **43** and **45**. In the T2R1 CoMFA model, the red contours suggest that negative charge could be potentially increase the affinity of P-gp for the benzene ring.

In region D, the presence of the amine on the tetrahydroisoquinolinyl moiety falls into a positive favorable blue region in both models, suggesting that a positive charge was allowed to increase the accessibility of the compounds to the P-glycoprotein pocket. In addition, the amine could be necessary to form a hydrogen bond or it could be protonated to bind to P-gp. These two large blue regions near the dimethoxyl group in both models indicate that a positive charge or a hydrogen bond acceptor was required for potency. In both models, a red region near the dimethoxyl group indicates that a negative charge was allowed to increase the potency. In the T1R1 CoMFA model, the green region near the dimethoxyl moiety reveals that a bulky group at position 3 or 4 on the benzene ring of the tetrahydroisoquinoline group is well tolerated. This was confirmed with compounds **39** and **40**. The green region (sterically favorable) and yellow (sterically unfavorable) contours around the tetrahydroisoquinolinyl moiety indicate that bulky groups could be present to increase the activity and that bulky groups will decrease the activity. This probably means that the binding site was sterically restricted. In the T3R1 CoMFA model, the green region near the dimethoxyl group reveals that the bulky group at position 4 on the benzene ring of the tetrahydroisoquinoline is well tolerated. However, at position 3, a less bulky group could help to increase the activity. This was confirmed with compound **40**. The yellow (sterically unfavorable) contours around the tetrahydroisoquinoline indicate that the bulky group could decrease the activity.

The contours maps of CoMSIA models also indicate features in regions A−D of **1**. In region A, the T1R1 CoMSIA model showing a green favorable region near the second aromatic ring of the isoquinoline indicates that bulky groups could be added to increase the antagonistic potency. This was confirmed with the remarkable difference of the activity between the presence of a single benzene ring compared to the naphthalenyl-like group or a benzenecyclohexyl group in almost all compounds in the training set. At the same place, in both T2R1 and T3R2 CoMSIA models, the favorable hydrophobic yellow contours on the isoquinoline group indicate the importance of the hydrophobic group such as aromatic ring or alkyl group. In the T1R1 CoMSIA models, the blue favorable positive electrostatic indicates that a positive charge increases the potency. The same conclusion using the T1R1 CoMFA model was drawn about the importance of the amine group at the 3 position on the isoquinoline to the activity. In fact, the magenta color shows this position and represents a favorable hydrogen bond acceptor region in the T1R1 and T3R2 CoMSIA models, and a small red negative region shows this position in the T2R1 CoMSIA model. In the T2R1 and T3R2 CoMSIA models, the green unfavorable region near position 3 on the isoquinoline suggests that a hydrophobic group or carbon atom was decreasing the activity of the molecules. In the T1R1 and T3R2 CoMSIA models orange unfavorable hydrogen bond acceptor regions at position 4, 5, or 6 of the isoquinoline group suggest that an acceptor group like amine or oxygen was not well accepted. In fact, compounds **15** and **28** bearing a carbonyl group and compounds such as **14**, **17**, and **18** substituted by an ether group exhibit moderate anti-P-gp activities. Also, in the T1R1 and T3R2 CoMSIA models, an orange unfavorable hydrogen bond acceptor region at position 8 of the isoquinoline suggests that an acceptor group like an aromatic amine group influenced negatively the pharmacological activity, as depicted already by the T1R1 CoMFA model. Furthermore, in the T3R2 CoMSIA model the amide bond near region A of the template falls into an orange unfavorable hydrogen bond acceptor region and a cyan favorable hydrogen bond donor region. Thus, donor groups like carbonyl and acceptor groups like the amine were of penultimate importance for the activity. In the T1R1 CoMSIA model, the small orange region reveals that the hydrogen bond acceptor at position 2 such as amine (e.g., compound **9**) of the isoquinoline decreases the potency of anthralinamides.

In region B of the template in the T3R2 CoMSIA model, the small favorable hydrophobic yellow contours suggest the importance of a hydrophobic group on the acceptor group at position 4 of the anthranilic moiety. Also at the same position and same model, the purple unfavorable hydrogen bond donor region indicates that donor groups decrease the activity. In the T1R1 CoMSIA model, the blue favorable positive electrostatic indicates that a positive charge increases the potency. This probably showed again the importance of a positively charged aromatic group. The amide bond near region B of the template, in the T3R2 CoMSIA model, falls into the favorable hydrogen bond donor area (cyan) and a purple unfavorable hydrogen bond donor region. Thus, acceptor groups like carbonyl and amine were important for the activity.

In region C, T2R1 and T3R2 CoMSIA models show favorable hydrophobic yellow contours, suggesting that the aromatic linker arm was important for the activity. The small unfavorable

hydrogen bond donor purple region between the regions B and C in T3R2 CoMSIA model indicates that a donor group was not tolerated at this place when the group in region D was with a more flexible group.

In region D, the T1R1 CoMFA model conclusion of the region can also be applied. In fact, the presence of the amine on the tetrahydroisoquinoline moiety falls into a positive favorable blue region and orange unfavorable hydrogen bond acceptor regions, suggesting that positive charge was allowed to increase binding to P-glycoprotein and the amine should be protonated to form a hydrogen bond to P-gp. The two blue regions near the dimethoxy group in both models suggest that a positive charge or hydrogen bond acceptor was required for the potency. In the T1R1 CoMSIA model, a red region on the dimethoxy group indicates that negative charge was tolerated to increase the potency. The green region near the dimethoxy reveals that a bulky group at the 3 or 4 position on the benzene ring of the tetrahydroisoquinoline is well tolerated. The sterically favorable green region and yellow sterically unfavorable contours around the tetrahydroisoquinoline indicate that a bulker group in this direction is favored but is not favored in the other.

Few commonalities and differences between the CoMFA and CoMSIA models have been observed after the contour map analysis and can be summarized by the two best models T1R1 CoMFA and CoMSIA. In region A, both CoMFA and CoMSIA show a few commonalities: (1) there is a positive influence on the activity of a heteroatom acceptor group like amine at the 3 position of the isoquinoline; (2) a negative charge at the ortho position of the amide was not well abided; (3) an aromatic bulky group may be necessary to increase the potency. They also have a few differences between both methods. The extra information from the CoMFA model suggests that a bulky group at the end of the isoquinoline is not well accepted. Also, the CoMFA model reveals the importance of the hydrogen bond acceptor carbonyl amide. The CoMSIA model adds more specific information and reveals that an acceptor group like amine or oxygen at position 4, 5, or 6 of the isoquinoline group and an acceptor group at position 8 of the isoquinoline, an aromatic amine group, were not well accepted. In region B, both CoMFA and CoMSIA have no commonalities. They have, however, a few differences between both models. The extra information from the CoMFA model suggests that the negative charge at position 5 of the amide was not well permitted. The CoMSIA model suggests that an aromatic group positively charged is important for the potency. In region C, both CoMFA and CoMSIA have no commonalities. The T1R1 CoMSIA model does not have any contour maps in that region. The T1R1 CoMFA model indicates that bulky groups are well abided to increase the potency. In region D, both CoMFA and CoMSIA share the same common themes: (1) the amine on the tetrahydroisoquinolinyl is important to the activity and could form a hydrogen bond or it could be protonated to bind to P-gp; (2) a positive charge or a hydrogen bond acceptor was required for the potency near the dimethoxyl group; (3) near the dimethoxyl group a negative charge was allowed to increase the potency; (4) a bulky group at position 3 or 4 on the benzene ring of the tetrahydroisoquinoline group is well tolerated; (5) the binding site was probably sterically restricted because of the sterically favorable region and the sterically unfavorable contours around the tetrahydroisoquinolinyl moiety.

After the analysis of these models, we established the functional groups or atoms important to the potency and the binding of **1**. In region A, proton-acceptor groups such as an amine at position 3 in a bulky bis aromatic system were the most important feature for the activity. This observation was also made by Pajeva and Wiese.[48] In addition, the negative charge from an amine or an oxygen atom at position 6 or 7 on the isoquinoline probably increases the potency of **1**. In region B, the presence of a hydrogen bond acceptor group such as a methoxy group or an atom negatively charged like fluorine or chlorine was the favored feature for the activity. Also, the aromatic group could be an important feature if the total electrostatic charge of that group is positive. In region C, the hydrophobic group such as the benzene ring was important to the activity. A similar finding was also suggested by Pajeva and Wiese.[51] In addition, a bulky group near the phenyl ring will probably increase the potency of **1**. In region D, the presence of the acceptor group like the basic amine of the tetrahydroisoquinoline, the presence of the acceptor groups like the oxygen atom of the dimethoxy, and the presence of bulky aromatic groups like the tetrahydroquinoline seem to be the most important features of that region. Specific bulky or less bulky groups at position 3 or 4 on the tetrahydroquinoline will probably help to increase the activity. Also, a negative charge on the aromatic ring of the tetrahydroquinoline will probably increase the activity. Finally, the hydrogen bond acceptor and donor groups of both amides were important for increasing the potency.

Conclusion

In this study, we investigated 3D-QSAR models of anthranilamide MDR modulators. Predictive CoMFA and CoMSIA models were developed for the modulation of P-glycoprotein against cells that overexpress P-glycoprotein using 49 anthranilamide derivatives in the training set taken from a data set of 178 compounds. Three different conformations were used to see the effect of the conformation in the model. On the basis of these three conformations, moderate to good internal predictive 3D-QSAR models were derived. Each model was validated using an external test set of 13 compounds not included in the training set and showed poor to good predictive r^2, between 0.036 and 0.793. Models with the best q^2 did not give good external prediction. T1R1 CoMFA and CoMSIA models were the best models developed so far. T1R1 CoMFA has an r^2_{cv} of 0.559, and T1R1 CoMSIA has an r^2_{cv} of 0.537. The steric, electrostatic, and hydrogen bond acceptor fields were shown to be the most important properties. These fields identified the functional group and atoms possibly related to the bonding and the inhibition of P-gp. We established that an acceptor group such as an amine in position 3 in a bulky bis aromatic system in region A, a hydrogen acceptor group like a methoxyl group or an atom negatively charged like fluorine or chlorine in region B, the aromatic group charged positively in region B, the hydrophobic group like the benzene ring in region C, the acceptor group like the basic amine of the tetrahydroisoquinoline, the acceptor group like the oxygen atom of the dimethoxy, and the bulky aromatic group like the tetrahydroquinoline in region D, and both amide bonds were the most important to be modified to improve the pharmacological activity of anthranilamides as P-glycoprotein antagonists. Furthermore, a few compounds in the test set are the exception. In fact, compound t_9543 is quite potent (70 nM) and does not have an acceptor hydrogen group like a methoxyl group in region B, but the molecule has a methyl group at the 3 position of the anthranilic moiety. Also, test set compounds t_9297, t_9380, and t_9442 did not correspond to all the trends derived from those models. They do not have an aromatic group in region C, and compound t_9297 is more potent (400 nM) than the other two compounds.

It seems that an aliphatic group like a propyl group was better for the potency compared to a biphenyl or bromine.

The conclusion of the Wiese and Pajeva studies with the Free−Wilson analysis and CoMSIA model outlined the presence of a tetrahydroquinoline substructure bonded to the anthranilamide moiety through a phenyl moiety and of a heteroatom in position 3 in a bulky aromatic ring system in A region of **1** to have the most significant impact on anti-MDR activity. The anthranilamide moiety provides hydrogen bond interactions with the protein for inhibition activity. In our studies, we found basically the same idea, but instead of a heteroatom at position 3, we found an acceptor group such as an amine in position 3 in a bulky bis aromatic system in region A. Also, we found that an electroattractor group or acceptor group is necessary for the anthranilamide moiety in region B and the acceptor groups in region D are very important for the activity. A few of these results in this paper are common to other P-gp inhibitors like an aromatic ring system, a basic tertiary nitrogen positioned at a fixed distance from the aromatic system, and an amide carbonyl group that acts as an acceptor and a donor group. However, P-gp inhibitors can bind to different binding sites. So each site will have its own SAR with some common themes.

The derived models in this study explain the observed variance in the activity of anthranilamides. They can help to understand the mechanism of P-gp activity, and they can also provide important insights into structural variations that may lead to the design of new modulators of P-glycoprotein exhibiting high and selective activity.

Acknowledgment. We thank Dr. William Wei for helpful discussions and technical assistance. We thank FRSQ for a studentship for P.L. L.P.K. gratefully acknowledges the Rx&D HRF-CIHR Research Career Award and Premier's Research Excellence Award. An infrastructure grant from the Ontario Innovation Trust provides support for the Molecular Design and Information Technology Centre and is gratefully acknowledged.

Supporting Information Available: Tables of results of the CoMFA and CoMSIA options on cross-validated coefficient $r^2(q^2)$ (SAMPLS) values using templates T1, T2, and T3 with alignment rules 1 and 2. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Young, R. C. Drug resistance: the clinical problem. In *Drug Resistance in Cancer Therapy*; Ozols, R. F., Ed.; Kluwer Academic Publishers: Boston, MA, 1989; pp 1−12.
(2) Robert, J. Multidrug resistance in oncology: diagnostic and therapeutic approaches. *Eur. J. Clin. Invest.* **1999**, *29*, 536−545.
(3) Nussler, V.; Pelka-Fleischer, R.; Zwierzina, H.; Nerl, C.; Beckert, B.; Gullis, E.; Gieseler, F.; Bock, S.; Bartl, R.; Petrides, P. E.; Wilmanns, W. Clinical importance of P-glycoprotein-related resistance in leukemia and myelodysplastic syndromes−first experience with their reversal. *Ann. Hematol.* **1994**, *69* (Suppl. 1), S25−S29.
(4) Juliano, R.; Ling, V. A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochem. Biophys. Acta* **1976**, *455*, 152−162.
(5) Endicott, J.; Ling, V. The biochemistry of P-glycoprotein mediated multidrug resistance. *Annu. Rev. Biochem.* **1989**, *58*, 137−171.
(6) Gottesman, M.; Pastan, I. Biochemistry of multidrug resistance mediated by the multidrug transporter. *Annu. Rev. Biochem.* **1993**, *62*, 385−427.
(7) Abraham, E. H.; Prat, A. G.; Gerweck, L.; Seneveratne, T.; Arceci, R. J.; Kramer, R.; Guidotti, G.; Cantiello, H. F. The multidrug resistance (mdr1) gene product functions as an ATP channel. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 312−316.
(8) Roepe, P. Analysis of the steady-state and initial rate of doxorubicin efflux from a series of multidrug-resistant cells expressing different levels of P-glycoprotein. *Biochemistry* **1992**, *31*, 12555−12564.
(9) Zhu, T. A novel hypothesis for the mechanism of action of P-glycoprotein as a multidrug transporter. *Mol. Carcinog.* **1999**, *25*, 1−13.
(10) Tsuruo, T.; Iida, H.; Tsukagoshi, S.; Sakurai, Y. Overcoming of vincristine resistance in P388 leukemia in vivo and in vitro through enhanced cytotoxicity of vincristine and vinblastine by verapamil. *Cancer Res.* **1981**, *41*, 1967−1972.
(11) Slater, L. M.; Sweet, P.; Stupecky, M.; Gupta, S. Cyclosporin A reverses vincristine and daunorubicin resistance in acute lymphatic leukemia in vitro. *J. Clin. Invest.* **1986**, *77*, 1405−1408.
(12) Arceci, R. J.; Stieglitz, K.; Bierer, B. E. Immunosuppressants FK506 and rapamycin function as reversal agents of the multidrug resistance phenotype. *Blood* **1992**, *80*, 1528−1536.
(13) Tsuruo, T.; Iida, H.; Kitatani, Y; Yokota, K; Tsukagoshi, S.; Sakurai, Y. Effects of quinidine and related compounds on cytotoxic and cellular accumulation of vincristine and adriamycin in drug resistant tumor cells. *Cancer Res.* **1984**, *44*, 4303−4307.
(14) Tsuruo, T.; Iida, H.; Nojiri, M; Tsukagoshi, S.; Sakurai, Y. Circumvention of vincristine and adriamycin resistance in vitro and in vivo by calcium influx blockers. *Cancer Res.* **1983**, *43*, 2905−2910.
(15) Beck, W. T.; Danks, M. K. Characteristics of multidrug resistance in human tumor cells. In *Molecular and Cellular Biology of Multidrug Resistance in Tumor Cells*; Roninson, I. B., Ed.; Plenum Press: New York, 1991; pp 3−46.
(16) Beck, W. Modulators of P-glycoprotein associated multidrug resistance. In *Molecular and Clinical Advances in Anticancer Drug Resistance*; Ozols, R. F., Ed.; Kluwer Academic Publishers: Philadelphia, PA, 1991; pp 151−170.
(17) Benson, A. B.; Trump, D. L.; Koeller, J. M.; Egorin, M. I.; Olman, E. A.; Witte, R.; Tormey, D. C. Phase I study of vinblastine and verapamil given by concurrent iv infusion. *J. Clin. Oncol.* **1985**, *3*, 311−315.
(18) Miller, T. P.; Grogan, T. M.; Dalton, W. S.; Spier, C. M.; Scheper, R. J.; Salmon, S. E. P-glycoprotein expression in malignant lymphomas and reversal of clinical drug resistance with chemotherapy plus high-dose verapamil. *J. Clin. Oncol.* **1991**, *9*, 17−24.
(19) Trump, D. L.; Smith, D. C.; Ellis, P. G.; Rogers, M. P.; Schold, S. C.; Winer, E. P.; Jordan, V. C.; Fine, R. L. High-dose oral tamoxifen, a potential multidrug-resistance-reversal agent: Phase I trial in combination with vinblastine. *J. Natl. Cancer Inst.* **1992**, *84*, 1811−1816.
(20) Christen, R. D.; McClay, E. F.; Plaxe, S. C.; Yen, S. S.; Kim, S.; Kirmani, S; Wilgus, L. L.; Heath, D. D.; Shalinsky, D. R.; Freddo, J. L. Phase I/pharmacokinetic study of high-dose progesterone and doxorubicin. *J. Clin. Oncol.* **1993**, *11*, 2417−242.
(21) Bartlett, N. L;, Lum, B. L.; Fisher, G. A.; Brophy, N. A.; Ehsan, M. N.; Halsey, J.; Sikic, B. I. Phase I trial of doxorubicin with cyclosporin as a modulator of multidrug resistance. *J. Clin. Oncol.* **1994**, *12*, 835−842.
(22) Boekhorst, P. A.; Van, K. J.; Schoester, M.; Sonneveld, P. Reversal of typical multidrug resistance by cyclosporin and its non-immunosuppressive analogue SDZ PSC 833 in Chinese hamster ovary cells expressing mdr1 phenotype. *Cancer Chemother. Pharmacol.* **1992**, *30*, 238−342.
(23) Roe, M.; Folkes, A.; Ashworth, P.; Brumwell, J.; Chima, L.; Hunjan, S.; Pretswell, I.; Dangerfield, W.; Ryder, H.; Charlton, P. Reversal of P-glycoprotein mediated multidrug resistance by novel anthranilamide derivatives. *Bioorg. Med. Lett.* **1999**, *9*, 595−600.
(24) Martin, C.; Berridge, G.; Mistry, P.; Higgins, C.; Charlton, P.; Callaghan, R. The molecular interaction of the high affinity reversal agent XR9576 with P-glycoprotein. *Br. J. Pharmacol.* **1999**, *128*, 403−411.
(25) Hyafil, F.; Vergely, C.; Du Vignaud, P.; Grand-Perret, T. In vitro and in vivo reversal of multiple resistance by GF120918, an acridonecarboxamide derivative. *Cancer Res.* **1993**, *53*, 4595−4602.
(26) Dantzig, A. H.; Shepard, R. L.; Cao, J.; Law, K. L.; Ehlhardt, W. J.; Baughman, T. M.; Bumol, T. F.; Starling, J. J. Reversal of P-glycoprotein-mediated multidrug resistance by a potent cyclopropyldibenzosuberane modulator LY335979. *Cancer Res.* **1996**, *56*, 4171−4179.
(27) Mistry, P.; Stewart, A. J.; Dangerfield, W.; Okiji, S.; Liddle, C.; Bootle, D.; Plumb, J. A.; Templeton, D.; Charlton, P. In vitro and in vivo reversal of P-glycoprotein-mediated multidrug resistance by a novel potent modulator, XR9576. *Cancer Res.* **2001**, *61*, 749−758.
(28) Deleted in proof.
(29) Deleted in proof.
(30) Pearce, H. L.; Safa, A. R.; Bach, N. J.; Winter, M. A.; Cirtain, M. C.; Beck, W. T. Essential features of the P-glycoprotein pharmacophore as defined by a series of reserpine analogs that modulate multidrug resistance. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5128−5132.
(31) Pearce, H. L.; Winter, M. A.; Beck, W. T. Structural characteristics of compounds that modulate P-glycoprotein-associated multidrug resistance. *Adv. Enzyme Regul.* **1990**, *30*, 357−373.

(32) Ramu, A.; Ramu, N. Reversal of multidrug resistance by phenothiazines and structurally related compounds. *Cancer Chemother. Pharmacol.* **1992**, *30*, 165−173.
(33) Chiba, P.; Ecker, G.; Schmid, D.; Drach, J.; Tell, B.; Goldenberg, S.; Gekeler, V. Structural requirements for activity of propafenone-type modulators in P-glycoprotein-mediated multidrug resistance. *Mol. Pharmacol.* **1996**, *49*, 1122−1130.
(34) Pajeva, I. K.; Wiese, M. QSAR and molecular modeling of catamphiphilic drugs able to modulate multidrug resistance in tumors. *Quant. Struct.−Act. Relat.* **1977**, *16*, 1−10.
(35) Salem, M.; Richter, E; Hitzler, M.; Chiba, P.; Ecker, G. Studies on propafenone type modulators of multidrug resistance VIII: synthesis and pharmacological activity of indanone analogues. *Sci. Pharm.* **1996**, *66*, 147−158.
(36) Pajeva, I. K.; Wiese, M. Molecular modeling of phenothiazines and related drugs as multidrug resistance modifiers: a comparative molecular field analysis study. *J. Med. Chem.* **1998a**, *41*, 1815−1826.
(37) Pajeva, I. K.; Wiese, M. A comparitive molecular field analysis of propafenone-type modulators of cancer multidrug resistance. *Quant. Struct.−Act. Relat.* **1998b**, *17*, 301−312.
(38) Ecker, G.; Huber, M.; Schmid, D.; Chiba, P. The importance of a nitrogen atom in modulators of multidrug resistance. *Mol. Pharmacol.* **1999**, *56*, 791−796.
(39) Schmid, D.; Ecker, G.; Kopp, S.; Hitzler, M.; Chiba, P. Structure−activity relationship studies of propafenone analogs based on P-glycoprotein ATPase activity measurements. *Biochem. Pharmacol.* **1999**, *58*, 1447−1456.
(40) Seelig, A. How does P-glycoprotein recognize its substrates? *Int. J. Clin. Pharmacol. Ther*. **1999**, *36*, 50−54.
(41) Ryder, H.; Ashworth, P. A.; Roe, M. J.; Brumwell, J. E.; Hunjan, S.; Folkes, A. J.; Sanderson, J. T.; Williams, S.; Maximen, L. M. Anthranilic acid derivatives as multi drug resistance modulators. WO98/17648, April 30, 1998.
(42) Klebe, G. Comparative molecular similarity indices analysis: CoMSIA. In *3D QSAR in Drug Design*; Kubinyi, H., Folkers, G., Martin, Y. C., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1998; Vol. 3, pp 87−104.
(43) Globisch, C.; Pajeva, I. K.; Wiese, M. Structure−activity relationships of a series of tariquidar analogs as multidrug resistance modulators. *Bioorg. Med. Chem.* **2006**, *5*, 1588−1598.
(44) Ford, J. M.; Prozialeck, W. C.; Heit, W. N. Structural features determining activity of phenothiazines and related drugs for inhibition of cell growth and reversal of multidrug resistance. *Mol. Pharmacol.* **1989**, *35*, 105−115.
(45) Ford, J. M.; Briggeman, E. P.; Pastan, I.; Gottesman, M.; Hait, W. N. Cellular and biochemical characterization of thioxanthenes for reversal of multidrug resistance in human and murine cell lines. *Cancer Res.* **1990**, *50*, 1748−1756.
(46) Ramu, A. In *Resistance to Antineoplastic Drugs*; Kessel, D., Ed.; CRC: Boca Raton, FL, 1989; pp 63−80.
(47) Bush, A.; Martin-Pastor, M. Structure and conformation of complex carbohydrate of glycoproteins, glycolipids, and bacterial polysaccharides C. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 269−293.
(48) Jain, A. N.; Koile, K.; Chapman, D. Compass: predicting biological activities from molecular surface properties, performance comparisons on a steroid benchmark. *J. Med. Chem.* **1994**, *37*, 2315−2327.
(49) Martin, Y. C.; Bures, M. G.; Danaher, E. A.; DeLazzer, J.; Lico, I.; Pavlik, P. A. A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 83−102.
(50) Pearlman, R. S. *3D QSAR in Drug Design. Theory, Methods, and Applications*; ESCOM Science: Leiden, The Netherlands, 1993.
(51) Pajeva, I. K.; Globisch, C.; Wiese, M. Structure−function relationships of multidrug resistance P-glycoprotein. *J. Med. Chem.* **2004**, *47*, 2523−2533.
(52) Viswanadhan, V. N.; Ghose, A. K.; Revankar, G. R.; Robins, R. K. Atomic physichochemical parameters for three dimensional structure directed quantitative structure−activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 163−172.
(53) Klebe, G. The use of composite crystal-field environments in molecular recognition and the de novo design of protein ligands. *J. Mol. Biol.* **1994**, *237*, 212−235.
(54) 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.
(55) Agrafiotis, D. K.; Cedeno, W.; Lobanov, V. S. On the use of neural network ensembles in QSAR and QSPR. *J. Chem. Inf. Comput. Sci*. **2002**, *42*, 903−911.
(56) Wehrens, R.; van der Linden, W. E. Bootstrapping principal component regression models. *J. Chemom.* **1997**, *11*, 157−171.
(57) Kubinyi, H.; Hamprecht, F. A.; Mietzner, T. Three-dimensional quantitative similarity−activity relationships (3D QSiAR) from SEAL similarity matrices. *J. Med. Chem.* **1998**, *41*, 2553−2564.
(58) Golbraikh, A.; Tropsha, A. Beware of q^2. *J. Mol. Graphics Modell.* **2002**, *20*, 269−276.
(59) Bang, S. J.; Cho, S. J. Comparative molecular field analysis (CoMFA) and comparative molecular similarity index analysis (CoMSIA) study of mutagen X. *Bull. Korean Chem. Soc.* **2004**, *25*, 1525−1530.

JM060239B