Available online at www.sciencedirect.com
SCIENCE DIRECT
Bioorganic & Medicinal Chemistry
Bioorganic & Medicinal Chemistry 14 (2006) 1588–1598

Structure–activity relationships of a series of tariquidar analogs as multidrug resistance modulators

Christoph Globisch,[a] Ilza K. Pajeva[a,b] and Michael Wiese[a,*]

[a]*Institute of Pharmacy, University of Bonn, An der Immenburg 4, 53121 Bonn, Germany*
[b]*Centre of Biomedical Engineering, Bulgarian Academy of Sciences, Acad. G. Bonchev Str. Bl. 105, 1113 Sofia, Bulgaria*

Received 2 June 2005; revised 4 October 2005; accepted 6 October 2005
Available online 22 November 2005

Abstract—Tariquidar (XR9576) analogs, modulators of cancer multidrug resistance (MDR), were subjected to QSAR and 3D-QSAR analyses. The structural features contributing to anti-MDR activity were identified by the Free-Wilson analysis and pharmacophore search using Hoechst 33342 as a template. 3D-QSAR CoMFA and CoMSIA models were derived and tested. The best models yielded an external predictivity of 0.66–0.75 squared correlation coefficient and outlined HB-acceptor, steric, and hydrophobic fields as the most important 3D properties. On the basis of the QSAR and 3D-QSAR analyses it was suggested that the strong inhibitory potency of the compounds studied is related to the presence of a bulky aromatic ring system with a 3rd positioned heteroatom toward the anthranilamide nucleus in the opposite end of the tetrahydroquinoline group. The results can help in directing the rational design of new generations of potent P-glycoprotein MDR modulators.

1. Introduction

Multidrug resistance (MDR) remains a significant problem for effective cancer chemotherapy. Although multiple mechanisms can be involved in MDR, the most widely registered form is the typical MDR, associated with the function of the ABC transporter P-glycoprotein (P-gp). Since the first report of MDR overcoming by the calcium channel blocker verapamil,[1] a lot of effort has been invested in finding potent and specific P-gp inhibitors, called also MDR modulators. Nowadays the third generation of MDR modulators is under investigation including tariquidar (XR9576), zosuquidar (LY335979), laniquidar R1010933, and others.[2] XR9576 is one of the most potent MDR modulating agents described to date. It achieves a complete reversal of MDR at low concentrations (25–80 nM) and holds a long duration of activity in a panel of resistant tumor cells.[3,4] XR9576 shows nearly a 100-fold difference to verapamil in inhibiting rhodamine uptake and about a 1000-fold difference for calcein AM uptake in MDR human plasma membrane vesicles.[5] The compound has also been shown to bind to P-gp with high affinity and to have no effect on the MDR associated protein MRP.[6] Recently, a series of compounds based on the anthranilamide nucleus of XR9576 has been synthesized and tested for anti-MDR activity.[7] These MDR modulators are reported to bind to the same binding site of P-gp as its substrate Hoechst 33342 does.[8]

In parallel with the search of new and effective MDR modulators, many studies have been undertaken to elucidate the structure–activity relationships of different classes of MDR reversing agents. They deal mostly with representatives of the first generation of reversing agents and employ either empirical or classical QSAR and 3D-QSAR approaches.[9] Recently published studies on pharmacophore identification of P-gp related drugs[10–13] point to hydrophobic and hydrogen bond (HB)-acceptor interactions as the most likely to be involved in the ligand binding to P-gp. At the same time approaches like CoMSIA[14] allow us, additionally to the CoMFA steric and electrostatic fields, to estimate the role of the drug hydrophobic and HB properties in the 3D-QSAR.[15]

In this paper, we report a comprehensive study on a series of 32 anthranilamide derivatives including XR9576 (Table 1) by QSAR and 3D-QSAR methods. The structural features, contributing significantly to anti-MDR

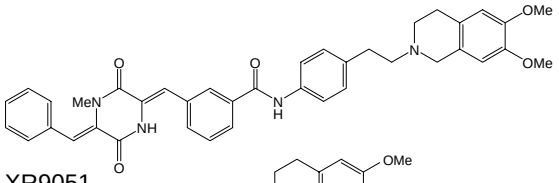
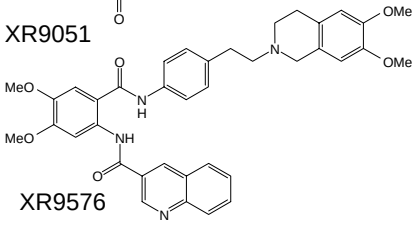
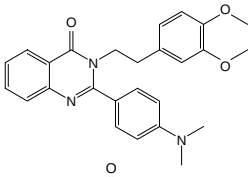
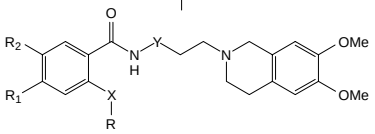
Abbreviations: MDR, multidrug resistance; P-gp, P-glycoprotein; HB, hydrogen bond; CoMFA, comparative molecular field analysis; CoMSIA, comparative molecular similarity indices analysis; LOO, leave-one-out cross-validation.

Keywords: P-glycoprotein; Multidrug resistance; Comparative molecular similarity indices analysis; Comparative molecular field analysis.

* Corresponding author. Tel.: +49 228 735213; fax: +49 228 737929; e-mail: mwiese@uni-bonn.de

0968-0896/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2005.10.058

Table 1. Structures, activity, and calculated pK_a values of the compounds studied

No.	Structure					IC_{50} (nM)[a]	pIC_{50}[b]	pK_a[c]
1	 XR9051					220 ± 60	6.66	7.85
2	 XR9576					38 ± 18	7.42	7.84
3						5000	5.30	3.74
								
	X	Y	R	R_1	R_2			
4	NHCO	Direct bond	4-*i*-Propylphenyl	H	H	1100 ± 150	5.96	7.24
5	NHCO	Direct bond	4-Cyclohexylphenyl	H	H	1600 ± 300	5.80	7.24
6	NHCO	Direct bond	4-Dimethylaminobenzyl	H	H	2400 ± 650	5.62	7.24
7	NHCO	Direct bond	Phenyl	H	H	6800 ± 1800	5.17	7.24
8	NHCO	Phenyl	Phenyl	H	H	250 ± 25	6.60	7.84
9	NHCO	Phenyl	4-*i*-Propylphenyl	H	H	850 ± 370	6.07	7.84
10	CO	Phenyl	Phenyl	H	H	1000 (n = 1)	6.00	7.84
11	O	Phenyl	Phenyl	H	H	2700 ± 350	5.57	7.84
12	S	Phenyl	Phenyl	H	H	2300 (n = 1)	5.64	7.84
13	Direct bond	Phenyl	Phenyl	H	H	1700 (n = 1)	5.77	7.85
14	NHCO	Phenyl	3-Thiophenyl	H	H	730 ± 270	6.14	7.84
15	NHCO	Phenyl	2-Furanyl	H	H	3300 ± 800	5.48	7.84
16	NHCO	Phenyl	3-Furanyl	H	H	920 ± 520	6.04	7.84
17	NHCO	Phenyl	2-Pyrazinyl	H	H	204 ± 10	6.69	7.84
18	NHCO	Phenyl	2-(5-Methylpyrazinyl)	H	H	67 ± 17	7.17	7.84
19	NHCO	Phenyl	2-Pyridyl	H	H	680 ± 110	6.17	7.84
20	NHCO	Phenyl	3-Pyridyl	H	H	880 ± 380	6.06	7.84
21	NHCO	Phenyl	3-(6-Methylpyridyl)	H	H	67 ± 25	7.17	7.84
22	NHCO	Phenyl	2-Quinoxalinyl	H	H	92 ± 13	7.04	7.84
23	NHCO	Phenyl	2-Quinolinyl	H	H	790 ± 15	6.10	7.84
24	NHCO	Phenyl	3-Quinolinyl	H	H	87 ± 13	7.06	7.84
25	NHCO	Phenyl	3-Isoquinolinyl	H	H	4900 (n = 1)	5.31	7.84
26	NHCO	Phenyl	3-Quinolinyl	H	F	60 ± 16	7.22	7.83
27	NHCO	Phenyl	3-Quinolinyl	F	H	38 ± 13	7.42	7.84
28	NHCO	Phenyl	3-Quinolinyl	H	Cl	149 ± 26	6.83	7.83
29	NHCO	Phenyl	3-Quinolinyl	Cl	H	141 ± 19	6.85	7.83
30	NHCO	Phenyl	3-Quinolinyl	H	Me	220 ± 66	6.66	7.84
31	NHCO	Phenyl	3-Quinolinyl	H	NMe_2	293 ± 75	6.53	7.84
32	NHCO	Phenyl	3-Quinolinyl	NO_2	H	45 ± 16	7.35	7.81

[a] IC_{50} values as reported by Roe et al.;[7] n = 1 marks IC_{50} obtained by a single experiment.
[b] $pIC_{50} = -\log(IC_{50})$.
[c] pK_a values were calculated by ACD/Labs,[19] the estimation error was ±0.2 for all compounds, except for compound **3** (±0.24).

activity, were first outlined by the Free–Wilson analysis. A more precise definition of the important functional groups and atoms was performed by pharmacophore identification using Hoechst 33342 as a template. CoMFA and CoMSIA fields were generated from the structures aligned according to the identified pharmacophore groups. A detailed outlier analysis was performed to decide on the relevant selection of compounds for the training and test sets. A number of 3D-QSAR models were derived and tested. Finally, contour

plots were created and interpreted in relation to the mostly contributing 3D-regions. The presented results help for better understanding of structure–activity relationships of this promising class of P-gp inhibitors and thus for directing the rational design of potent and effective MDR modulators.

2. Results and discussion

2.1. Structural and activity data

The main series includes 31 substances based on the anthranilamide nucleus of XR9576 (XR-compounds) taken from the publication of Roe et al.[7] Activity data represent the effect of compounds on the accumulation of [^{3}H] daunorubicin expressed as IC_{50} in the MDR mouse mammary carcinoma cell line EMT6/AR 1.0 calculated relative to a dose of 100 μM verapamil, which restores the level to that of the parental EMT6/P cells. In Table 1, the structures, the pK_a, and the corresponding IC_{50} and pIC_{50} values of the compounds are given. As seen from the pK_a values, at physiological conditions the compounds are partially protonated on the nitrogen atom of the tetrahydroquinoline substituent. Therefore both, protonated and neutral forms, were considered.

2.2. Free-Wilson analysis

The main purpose of this analysis was to identify the significant structural features and to estimate their contributions to MDR reversal by the drugs studied. The evidence of an additive effect of atoms, functional groups, and substituents on activity studied can be directly related to the pharmacophore hypothesis for drug interactions with the putative receptor site and appears as a prerequisite for a good 3D-QSAR.

Compound **3** (Table 1) is the only one that does not possess an anthranilamide-like nucleus and a tetrahydroquinoline substructure, therefore this compound was omitted from the analysis. The remaining 31 compounds were subjected to the analysis using the following structural variables for Y, X, R, R_1, and R_2 in Table 1.

- Y: tetrahydroquinoline substructure bonded directly to the anthranilamide core (Y-db) or through a phenyl (Y-ph);
- X: X–NHCO for –NHC(=O)–, X–CO for –C(=O)–, X–sa for single atoms –O– or –S– and X-db for a direct bond;
- R: R-5, R-6, and R-66 for 5-atom, 6-atom, and two 6-atom aromatic ring system, respectively; R-5-2O and R-5-3O for the furanyl aromatic ring with O in position 2 or 3, respectively; R-6H for the unsubstituted phenyl ring; R-6isp, R-6clx, and R-6NMe for the isopropyl, cyclohexyl, and $–N(CH_3)_2$ para-substituted phenyl ring, respectively; R-5H-3S for the unsubstituted thiophenyl; R-6H-2N and R-6H-3N for the unsubstituted pyridines with N in positions 2 and 3, respectively; R-6Me-2N and R-6Me-3N for the $–CH_3$ substituted pyridines with N in positions 2 and 3, respectively; R-66-2N, R-66-3N for isoquinoline and quinolines with a 2nd or 3rd positioned N, respectively; R-dkp for the diketopiperazine substituent in XR9051 (No. **1**, Table 1);
- R_1: R1-H, R1-(+σ), and R1-(−σ) for substituents with 0 (H-atoms), positive (+σ) or negative (−σ) sigma constants at position 4 of the anthranilamide ring.
- R_2: R2-H, R2-(+σ), and R2-(−σ) for substituents with 0 (H-atoms), positive (+σ) or negative (−σ) sigma constants at position 5 of the anthranilamide ring.

Table 2 represents the structural features used as variables in the Free–Wilson analysis and lists the compounds containing the features. The reference consisted of Y-db, X-sa, R-6H, R1-H, and R2-H. Because of the large number of variables, a stepwise forward selection was applied. The final model was derived using the significant variables only (Table 3A). Eight significant variables were identified, of them R-5H-2O and R1-(−σ) appeared as single points (Table 2). To avoid involvement of variables with one non-zero value and overrepresented (X–NHCO was present in 27 out of 31 compounds), a Free-Wilson analysis was also performed excluding compounds **10**–**16**, **27**, **29**, and **32**, and the corresponding structural features for X (four variables), R-5H (three variables), and R_1 (three vari-

Table 2. Structural features used as variables in the Free-Wilson analysis and the XR-compounds containing them

No.	Structural feature	Compound[a]
	Y	
1	Y-db[b]	**4**–**7**
2	Y-ph	**1**, **2**, **8**–**32**
	X	
3	X-CO	**10**
4	X-sa[b]	**11**, **12**
5	X-db	**13**
6	X-NHCO	**1**, **2**, **4**–**9**, **14**–**32**
	R	
7	R-5H-3S	**14**
8	R-5H-2O	**15**
9	R-5H-3O	**16**
10	R-6H[b]	**7**, **8**, **10**–**13**
11	R-6isp	**4**, **9**
12	R-6clx	**5**
13	R-6NMe	**6**
14	R-6H-2N	**19**
15	R-6H-3N	**17**, **20**
16	R-6Me-3N	**18**, **21**
17	R-66-2N	**23**, **25**
18	R-66-3N	**2**, **22**, **24**, **26**–**32**
19	R-dkp	**1**
	R_1 (4-substituted[7])	
20	R1-H[b]	**1**, **4**–**26**, **28**, **31**
21	R1-(+σ)	**27**, **29**, **32**
22	R1-(−σ)	**2**
	R_2 (5-substituted[7])	
23	R2-H[b]	**1**, **4**–**25**, **27**, **29**
24	R2-(+σ)	**26**, **28**
25	R2-(−σ)	**2**, **30**, **31**

[a] The numbers correspond to those used in Table 1.
[b] These are used in the reference structure.

Table 3. Results of the Free-Wilson analysis of the XR-compounds

A	Significant variables only			
	$n = 31$; $r^2 = 0.874$; $r^2_{cv} = 0.657$ $F(8,22) = 19.12$; $p < 0.000$; SD of estimate: 0.277			
	Coefficient	SD	t	p level
Intercept	5.079	0.219	23.189	0.0000
Y-ph	0.666	0.170	3.927	0.0007
X-NHCO	0.559	0.170	3.294	0.0033
R-5H-2O	−0.824	0.294	−2.803	0.0104
R-6Me-3N	0.866	0.219	3.955	0.0007
R-66-2N	−0.599	0.219	−2.733	0.0121
R-66-3N	0.806	0.143	5.623	0.0000
R1-(−σ)	0.825	0.339	2.431	0.0236
R2-(−σ)	−0.515	0.222	−2.319	0.0301
B	Significant variables only			
	$n = 21$; $r^2 = 0.796$; $r^2_{cv} = 0.619$ $F(4,16) = 15.60$; $p < 0.000$; SD of estimate: 0.327			
	Coefficient	SD	t	p level
Intercept	5.638	0.164	34.450	0.0000
Y-ph	0.738	0.211	3.491	0.0030
R-6Me-3N	0.795	0.267	2.975	0.0089
R-66-2N	−0.670	0.267	−2.507	0.0233
R-66-3N	0.591	0.182	3.244	0.0051

ables). This resulted in a set of 21 compounds described by 15 variables. The reference consisted of Y-db, R-6H, and R2-H. A model with four significant variables was derived (Table 3B).

The QSAR models (Tables 3A and B) have satisfactory fitting ($r^2 = 0.874$ for $n = 31$ and $r^2 = 0.796$ for $n = 21$) and predicting abilities (cross-validated $r^2_{cv} = 0.657$ for $n = 31$ and $r^2_{cv} = 0.619$ for $n = 21$). Among the most significant features with a positive contribution to inhibitory activity are the substituents R-66-3N and R-6Me-3N indicating that the presence of a large aromatic ring system (either substituted pyridine or quinoline) next to the anthranilamide nucleus in the structures may favor inhibitory potency. A heteroatom in position 3 of R appears to be crucial implying a possible involvement of this atom in HB acceptor and/or electrostatic interactions. This observation is also confirmed by the significance of the variable R-66-2N that has a negative contribution to activity. The tetrahydroquinoline substructure with the phenyl ring next to the anthranilamide nucleus (Y-ph) is also among the major features contributing to MDR reversal.

In the model with 31 compounds, the contributing effect of the amide-group (where X=NHCO) can be related to its potential involvement in HB-interactions. The favorable influence of substituents with negative sigma effect in R_1 (R1-(−σ)) could be related to the contributing effect of the methoxy group and can be explained by its potential influence on the HB acceptor abilities of the carbonyl oxygen in the anthranilamide nucleus. In order to get a deeper insight into the role of the substituents in R_1 and R_2, a correlation with the identified structural features was experimented using the lipophilic π-constants. The more hydrophilic substituents in R_1 and R_2 were identified as significant provided the sigma effects were not involved ($r^2 = 0.858$, seven significant variables: Y-ph, X-NHCO, R-5H-2O, R-6Me-3N, R-66-2N, R-66-3N, and the sum of the π-values of R_1 and R_2 with a negative contribution to activity). Obviously, the contributions of the substituents in positions R_1 and R_2 cannot be explained by the π- or σ-constants only but rather by the mixed effect of both, lipophilic and electronic properties.

The predicted versus observed activity values calculated by the Free-Wilson models are shown in Figures 1A and B for the 31 and 21 compound models. Clustering of the compounds is observed indicating that, despite the statistically significant model obtained the significant variables only do not allow us to precisely differentiate between the compounds' activities. The results, however, indicate additive contributions of the identified structural features and can be taken into account to decide on the potential pharmacophore points of the inhibitors.

2.3. Identification of common functional groups and atoms: conformation generation

The XR-compounds are presumed to bind to the same site of P-gp as Hoechst 33342.[8] Thus, identification of

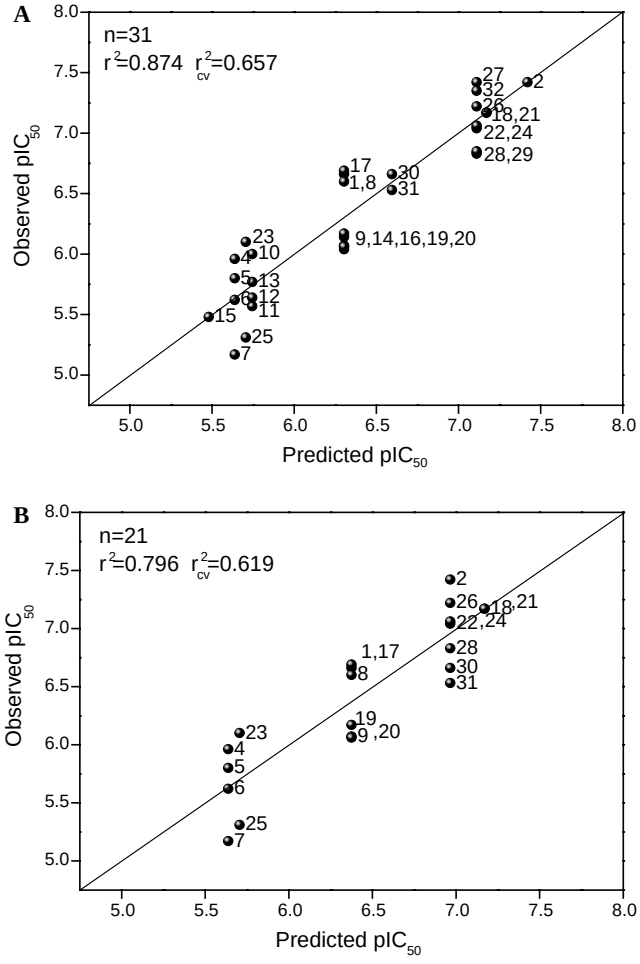


Figure 1. Predicted versus observed activity values of the XR-compounds in the Free-Wilson model (see Table 3). The regression line and the line that corresponds to the perfect prediction overlap each other.

the most appropriate conformations and functional groups and atoms was performed on the basis of Hoechst 33342 using the pharmacophore search[16] and considering in addition the results of the Free-Wilson analysis. All four global minimum conformations of Hoechst 33342 were used (see Section 4 for details on the Hoechst 33342 conformers) to generate a set of structural features that included three aromatic features (two phenyl rings from the benzimidazole rings and the single phenyl group), four acceptors (the benzimidazole N-atoms, the methyl-substituted N-atom, and the ethoxy O-atom), and two donors (the benzimidazole –NH-groups) (Fig. 2A). Some of these features (the three aromatic rings, the acceptor N-atom, and the donor –NH group in the benzimidazole ring next to the phenyl) were previously identified as the main pharmacophore points of the P-gp substrates that bind to the H-site of P-gp.[24] Considering the fact that the XR compounds are inhibitors and could bind to the H-site in different binding modes, none of these nine features was assigned as essential but at least six corresponding features were set as necessary. All four patterns were further tested. For this purpose, the pharmacophore search was applied to those of the conformers of the XR compounds that had the highest number of hits after the stochastic search. The pharmacophore pattern was selected that was found in most XR conformers. The criterion of six consensus features was only reached by compounds with a phenyl ring in the Y substituent. The other compounds achieved only four hits. Therefore, two templates were explored: one for compounds without a phenyl ring in Y, like compounds **4**–**7** and one for compounds with a phenyl ring in Y (Table 1). The most active representative in the series, XR9576, corresponded to the general pattern by six out of nine features (Fig. 2C), while the least active representative **7** (Fig. 2B) missed additionally two of these features.

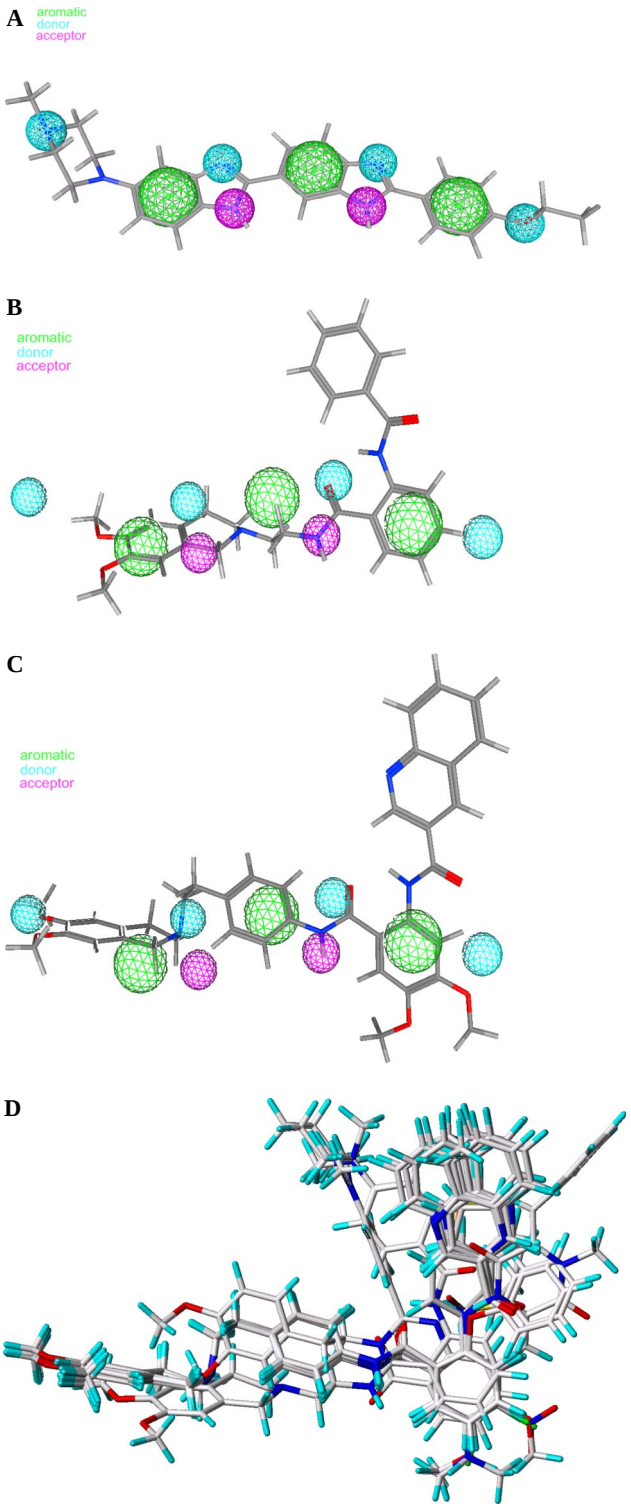


Figure 2. Structures of Hoechst 33342 (A), No. 7 (B), and XR9576 (No. 2) (C) with atoms and groups potentially involved in common interactions as identified on the global minima conformers of Hoechst 33342. (D) Alignment of all compounds as used in the CoMFA and CoMSIA analysis.

Taking into account that in the Free-Wilson analysis the 3rd-positioned heteroatom in R was found to significantly influence the anti-MDR activity, two subgroups were generated: 'left' and 'right' depending on the orientation of the N-atom in R (in Fig. 2C the 'left' orientation is shown). Initially, the tests with both orientations were performed separately using CoMSIA hydrophobic, donor, and acceptor fields to decide on the more relevant conformer for a further use. A model obtained with the 'left' oriented conformers yielded better statistical results, therefore a systematic test procedure was further performed with this model as follows: starting from the compound with the highest prediction error the oppositely oriented conformation was used to generate the next model, and so on, until all opposite conformations were tested. The conformational set with highest q^2 had some 'right' oriented conformers, namely **15**, **20**, **22**, **25**, **28**, and **31**, and was further used to derive the final models. Figure 2D illustrates the superimposed structures as used in the 3D-QSAR analyses.

2.4. 3D-QSAR models of all compounds

The 3D-QSAR analyses were performed in a stepwise manner. First the 31 derivatives possessing the anthranilamide nucleus were investigated as in the Free-Wilson analysis and compound **3** served as an external test compound. LOO cross-validated results of CoMFA and CoMSIA models derived with 31 compounds are summarized in Table 4. The best CoMFA model for

Table 4. PLS statistics of the CoMFA and CoMSIA models

Fields[a]	Neutral										Protonated without no. 3		
	Without no. 3							With no. 3					
	q^2_{cv} [b]	n_{opt} [c]	PRESS[d]	r^2	s	no. 3_{pred}	res.	q^2_{cv} [b]	n_{opt} [b]	PRESS[b]	q^2_{cv} [b]	n_{opt} [b]	PRESS[b]
CoMFA													
both	0.600	5	0.464	0.955	0.155	5.92	0.62	0.610	4	0.458	0.547	3	0.475
s	0.523	3	0.487	0.806	0.311	5.54	0.24	0.554	3	0.481	0.495	3	0.501
e	0.695	6	0.413	0.965	0.140	5.92	0.62	0.695	6	0.421	0.719	5	0.388
s+e	0.631	5	0.445	0.966	0.136	5.93	0.63	0.635	4	0.443	0.609	3	0.441
h-bnd	0.460	6	0.550	0.882	0.257	5.01	−0.29	0.519	6	0.529	0.523	4	0.496
s+e+h-bnd	0.589	5	0.470	0.907	0.224	5.66	0.36	0.626	6	0.467	0.622	3	0.434
both+h-bnd	0.513	4	0.501	0.881	0.248	5.57	0.27	0.589	6	0.489	0.596	3	0.448
CoMSIA													
s	0.484	3	0.506	0.675	0.402	5.47	0.17	0.519	3	0.500	0.493	3	0.502
e	0.673	4	0.411	0.875	0.254	5.74	0.44	0.677	3	0.409	0.703	4	0.392
s+e	0.643	3	0.421	0.823	0.297	5.65	0.35	0.663	3	0.419	0.652	3	0.416
h	0.524	3	0.487	0.730	0.367	5.46	0.16	0.554	3	0.481	0.527	3	0.485
d	0.240	1	0.593	0.875	0.254	5.60	0.30	0.288	1	0.588	0.214	1	0.603
a	0.747	4	0.362	0.342	0.552	5.71	0.41	0.769	5	0.360	0.694	4	0.398
s+h	0.574	3	0.460	0.731	0.365	5.43	0.13	0.599	3	0.456	0.579	3	0.458
s+a	0.757	5	0.361	0.900	0.232	5.66	0.36	0.767	5	0.361	0.696	3	0.389
e+h	0.671	3	0.404	0.817	0.302	5.59	0.29	0.687	3	0.403	0.649	3	0.418
h+a	0.794	6	0.339	0.938	0.186	5.43	0.13	0.805	6	0.337	0.743	5	0.371
s+e+h	0.660	3	0.411	0.815	0.303	5.56	0.26	0.678	3	0.409	0.641	3	0.423
s+h+a	0.778	6	0.352	0.935	0.191	5.54	0.24	0.787	6	0.352	0.726	5	0.384
s+e+h+a	0.755	5	0.363	0.922	0.205	5.65	0.35	0.766	5	0.361	0.746	5	0.369

[a] Fields used: both, s—steric, e—electrostatic, h-bnd—hydrogen bonds, and h—hydrophobicity, a—hydrogen bond acceptor, d—hydrogen bond donor.
[b] q^2_{cv} cross-validated correlation coefficient.
[c] n_{opt} optimal number of components.
[d] PRESS predictive sum of squares.

the neutral form was obtained using the electrostatic field alone ($q^2_{cv} = 0.695$; $n_{opt} = 6$). The CoMFA model based on the standard field (both with internal weighting) led to a decrease in the internal predictivity ($q^2_{cv} = 0.600$) but also to a slightly smaller number of components ($n_{opt} = 5$). For the protonated forms the electrostatic field based model yielded the best result ($q^2_{cv} = 0.719$; $n_{opt} = 5$).

CoMSIA models with the highest q^2 were obtained for the neutral form of the compounds. The acceptor field yielded the best single field model ($q^2_{cv} = 0.747$; $n_{opt} = 4$). In contrast, the donor field was of the lowest internal predictivity ($q^2_{cv} = 0.240$; $n_{opt} = 1$). In the combined models, the highest q^2 were achieved by the combination of hydrophobic and acceptor fields ($q^2_{cv} = 0.794$; $n_{opt} = 6$) and steric with hydrophobic and acceptor fields simultaneously ($q^2_{cv} = 0.778$; $n_{opt} = 6$). Thus, the model with hydrophobic and acceptor fields of the neutral forms appeared to be superior among all the models derived. The best results for the protonated forms were obtained with the combination of all the fields ($q^2_{cv} = 0.746$; $n_{opt} = 5$) and the model based on the hydrophobic and acceptor fields simultaneously had very close statistical parameters ($q^2_{cv} = 0.743$; $n_{opt} = 5$).

The statistics for all 32 compounds are summarized in Table 4. Inclusion of compound **3** in the neutral set led to a slight increase in q^2_{cv} and the ranking of the models remained essentially the same, with the acceptor field alone or in combination with the hydrophobic field leading to the best model.

2.5. 3D-QSAR models of training set compounds

Taking into account that the LOO cross-validation may lead to overoptimistic assessment of predictivity, a precise assessment of the models was performed. For this purpose, the initial set of 31 (protonated form) or 32 compounds (neutral form) was divided into training and test sets.

Nine compounds were randomly selected to form the test set and this procedure was repeated 20 times. The remaining compounds formed the training sets that were used to derive new models and to estimate their external predictivity. Table 5 summarizes the results of the CoMFA and CoMSIA performed with the generated training sets for the neutral forms. Results similar to those given in Table 5 have been obtained for the protonated forms of the XR-compounds (data not reported). For each model, the minimal, maximal, and the average values of the model parameters are shown. In general, compared to the models obtained with all compounds (Table 4), the cross-validated correlation coefficients decreased. This can be related to lowered total variance resulting from the reduced data set used in the training.

As seen from Table 5 the best models are the 's+e+h+a' model, followed by the 's+e' and 's+a' models with average $q^2_{cv} = 0.62$–0.64 and $n_{opt} = 3.3$–3.7, all involving steric and HB-acceptor fields. Thus, once again the role of these fields for MDR reversal was confirmed.

Table 5. CoMFA and CoMSIA models derived with the neutral forms of the compounds: minimal (min), maximal (max), and average (ave) values for 20 training/test set selections are reported

Fields[a]	Neutral: training sets								Neutral: test sets					
	q^2_{cv}[b]			n_{opt}[c]	PRESS[d]	r^2			r^2_{pred}[e]			p^2[f]		
	min	max	ave	range/ave	ave	min	max	ave	min	max	ave	min	max	ave
CoMFA														
both	0.297	0.699	0.497	2–4/2.9	0.505	0.641	0.981	0.859	0.024	0.768	0.473	0.091	0.764	0.574
s	0.219	0.686	0.465	1–4/2.8	0.518	0.577	0.981	0.820	−0.093	0.841	0.453	−0.116	0.839	0.553
e	0.519	0.846	0.679	2–6/4.0	0.414	0.724	0.987	0.917	−0.085	0.818	0.536	−0.144	0.820	0.598
s+e	0.343	0.761	0.582	2–5/3.3	0.464	0.683	0.990	0.890	0.041	0.837	0.523	0.141	0.839	0.609
h-bnd	0.106	0.656	0.330	1–4/2.7	0.580	0.272	0.952	0.789	−0.324	0.759	0.258	−0.075	0.818	0.326
s+e+h-bnd	0.276	0.672	0.478	2–4/3.2	0.519	0.737	0.987	0.889	−0.139	0.796	0.457	0.015	0.790	0.544
both+h-bnd	0.191	0.681	0.415	1–4/2.8	0.543	0.721	0.979	0.871	−0.131	0.784	0.420	0.074	0.783	0.515
CoMSIA														
s	0.233	0.639	0.469	2–4/2.8	0.516	0.451	0.790	0.660	−0.018	0.848	0.465	−0.040	0.850	0.525
e	0.496	0.832	0.658	2–6/3.6	0.422	0.806	0.958	0.884	0.040	0.887	0.588	0.199	0.895	0.651
h	0.182	0.640	0.461	1–6/2.1	0.511	0.596	0.970	0.751	−2.410	0.809	0.224	−1.704	0.813	0.314
d	0.069	0.495	0.211	1–2/1.3	0.607	0.266	0.679	0.403	−1.254	0.641	0.249	−1.303	0.645	0.241
a	0.491	0.771	0.637	3-5/3.8	0.440	0.836	0.953	0.894	0.301	0.941	0.691	0.225	0.958	0.720
s+e	0.463	0.790	0.623	3–5/3.3	0.441	0.790	0.936	0.862	0.122	0.888	0.575	0.350	0.891	0.655
s+h	0.223	0.641	0.482	1–6/2.2	0.503	0.491	0.952	0.744	−1.485	0.886	0.374	−0.944	0.892	0.457
s+a	0.486	0.765	0.619	2–4/3.3	0.446	0.855	0.981	0.911	0.241	0.963	0.679	0.132	0.964	0.710
e+h	0.427	0.760	0.597	2–6/3.1	0.454	0.773	0.969	0.864	0.124	0.880	0.561	0.276	0.883	0.617
h+a	0.384	0.761	0.583	3–4/3.0	0.460	0.764	0.980	0.913	0.047	0.948	0.633	0.023	0.963	0.666
s+e+h	0.420	0.743	0.587	2–6/3.2	0.460	0.759	0.970	0.870	0.116	0.916	0.572	0.246	0.927	0.634
s+h+a	0.441	0.747	0.590	2–5/3.1	0.459	0.813	0.984	0.905	0.164	0.943	0.621	0.143	0.956	0.651
s+e+h+a	0.505	0.778	0.637	2–5/3.7	0.439	0.847	0.988	0.920	0.307	0.935	0.657	0.291	0.937	0.695

[a,b,c,d] See Table 4 for the abbreviations used.
[e] r^2_{pred}—predictive correlation coefficient.[21]
[f] p^2—predictive correlation coefficient.[23]

2.6. Estimation of external predictivity

Table 5 reports the results on the external predictivity of the models by the coefficient r^2_{pred} and p^2 for the test sets with the neutral form for 32 compounds. In general, the r^2_{pred} and p^2 values agreed, the p^2 values being slightly higher than r^2_{pred}.

With CoMFA the best test results were obtained with the electrostatic field in correspondence with the results from the training. In general, the CoMSIA models produced higher r^2_{pred} and p^2 than the CoMFA ones. For both, CoMFA and CoMSIA models there was a correspondence between the best training and test set models. The best external predictivity was achieved by the HB-acceptor ($r^2_{pred} = 0.691$, $p^2 = 0.720$) and the 's+a' model ($r^2_{pred} = 0.679$, $p^2 = 0.710$), but the 's+e+h+a' ($r^2_{pred} = 0.657$, $p^2 = 0.695$) and the 'h+a' models had very close characteristics ($r^2_{pred} = 0.633$, $p^2 = 0.666$).

Comparison of results in Tables 4 and 5 reveals that the hydrophobic field is no longer among the best models. This can be attributed to compound 32 whose predicted activity varied between good and poor with deviations of up to 2 log units depending on the training set. This is also reflected by the high range of r^2_{pred}. We performed a PCA of the fields with the highest contributions to activity of the XR-compounds (steric, electrostatic, hydrophobic, and acceptor) with five components using a column filtering value of 1.0 kcal/mol. The plot of the first versus the second component of the neutral drug forms (Table 4, neutral 's+e+h+a' model) is shown in Figure 3. Cluster analysis of the principal component scores by Ward's method lead, to two main clusters A and B, and compounds **3** and **32** belong to neither of them. A possible reason for the observed deviation of compound **32** could be the CoMSIA parameterization of its nitro-group, whose hydrophobic index was the lowest (−2.76) in the data set. The cluster B can be divided into three subclusters: cluster B1 (compounds

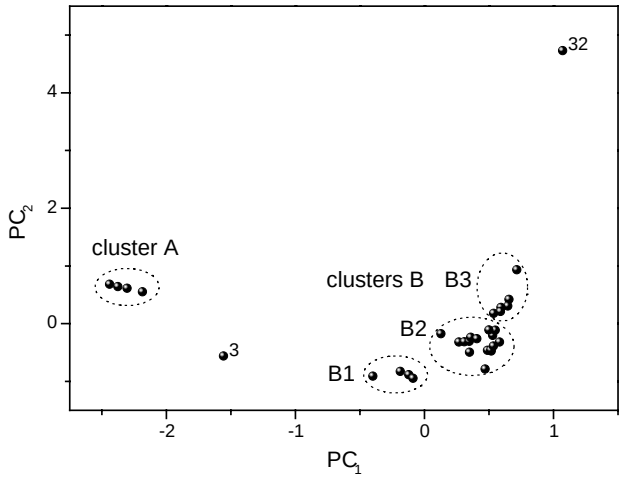


Figure 3. The principal component analysis plot for the steric, electrostatic, hydrophobic, and acceptor fields of the XR-compounds; cluster A: Nos. **4**–**7**; cluster B1: Nos. **10**–**13**; cluster B2: Nos. **1**, **8**, **9**, **14**–**20**, **23**, **25**, and **28**–**31**; cluster B3: Nos. **2**, **21**, **22**, **24**, **26**, and **27**.

10, **11**, **12**, and **13** which do not contain the –NHCO– group in X); cluster B2 (compounds **1**, **8**, **9**, **14**–**20**, **23**, **25**, and **28**–**31**), and cluster B3 (compounds **2**, **21**, **22**, **24**, **26**, and **27**). Notably, the clustering of the compounds in B follows the ranking of the compounds according to their pIC_{50} values that increase from B1 to B3 and resembles the clustering in the Free-Wilson plot (Fig. 1A). The same analysis was also performed with the protonated forms and it yielded similar results (data not shown). This strong deviation of compound **32** explains its variable predicted activity from models with the hydrophobic field.

In general, the best training set models showed very close r^2_{pred} and p^2 values indicating a stable external predictivity of the models obtained.

Figure 4 shows an exemplary plot of predicted versus observed pIC_{50} values of the training and test neutral compounds obtained by a model with r^2_{pred} and p^2 values in the middle range of the intervals obtained ('h+a' model, Table 5). In Figure 4A, the results from the LOO cross-validation are depicted. As seen from the plot, for all except one compound, namely compound **25** (predicted affinity too high), the predicted pIC_{50} values are in good agreement with the observed values ($r^2 = 0.751$). In case of compound **25**, the negative influence of the nitrogen in the unique isoquinolinyl substituent is underestimated. Figure 4B illustrates the very good correspondence between experimental and predicted activities for the external test set. All compounds, except no. **15**, are predicted very well.

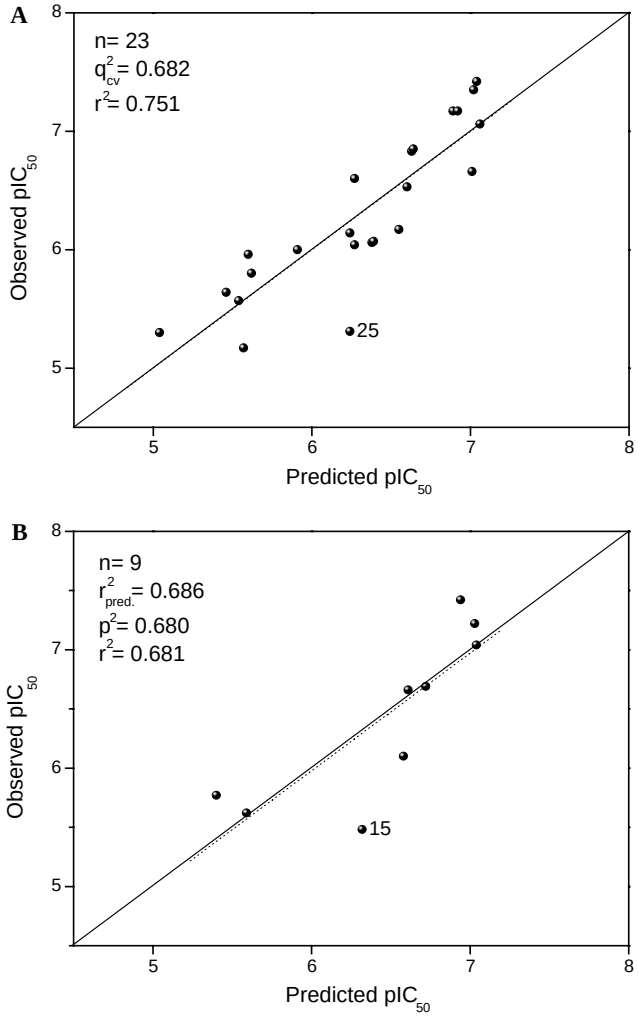


Figure 4. Predicted versus observed activity values obtained with one of the 'h+a' models for the neutral compounds in the training (A) and test (B) sets: (–) perfect prediction; (···) linear fit.

The above results from the model validations illustrate that the models derived are stable: they agree in relation to the most contributing fields and in the q^2_{cv}, r^2_{pred}, and p^2 values.

2.7. Contour plots

Further information from the models obtained can be gained by interpretation of the CoMSIA contour plots. In Figure 5, SD*coefficient plots of the CoMSIA 'h+a' model are shown with the structures of the most active (XR9576, Fig. 5A), the least active (7, Fig. 5B), and the structurally divergent compound **3** (Fig. 5C) in the series. The blue contoured areas mark the favorite influence of an atom with a HB acceptor possibility. One of these positions corresponds to the heteroatom in position 3 as seen from the plot with XR9576 (Fig. 5A), while the other one corresponds to oppositely oriented heteroatoms. The disfavored acceptor regions, where atoms with acceptor capabilities have a negative effect, are contoured in red. The small red region marks the negative contribution of the carbonyl group of compound **3**. The big red area corresponds to the 2nd positioned heteroatoms of the R-substituent. It can be seen on both, left and right sides, in correspondence with the left and right oriented conformations. The big favored hydrophobic area colored in green accounts for the positive effect of the hydrophobic aromatic ring substitution in the R-part of the molecule. The small one can be related again to the unfavorable effect of the R-part heteroatom in ortho-position to the amide bond. Unfavorable hydrophobic interactions are coded in yellow. The big yellow area in Figure 5B points to the negative contribution of the phenyl substituent without any heteroatom with HB-acceptor function. The small yellow contour plot reflects the positive effect of hydrophilic substituents in position 4 in the intermediate phenyl ring, including the nitro group in compound **32**.

3. Conclusion

In the study, 32 anthranilamide derivatives (Table 1) were studied by classical QSAR and 3D-QSAR analyses. These compounds were intentionally synthesized as analogs of the P-gp inhibitor XR9576,[7] one of the most powerful MDR modulators obtained up to now[3–5] and shown to bind to the same site as the P-gp substrate Hoechst 33342.[8] Thus, the profound analysis of the structure–activity relationships of these compounds can help in both, better understanding of their mechanism of P-gp inhibition and rational design of new and more effective MDR modulators.

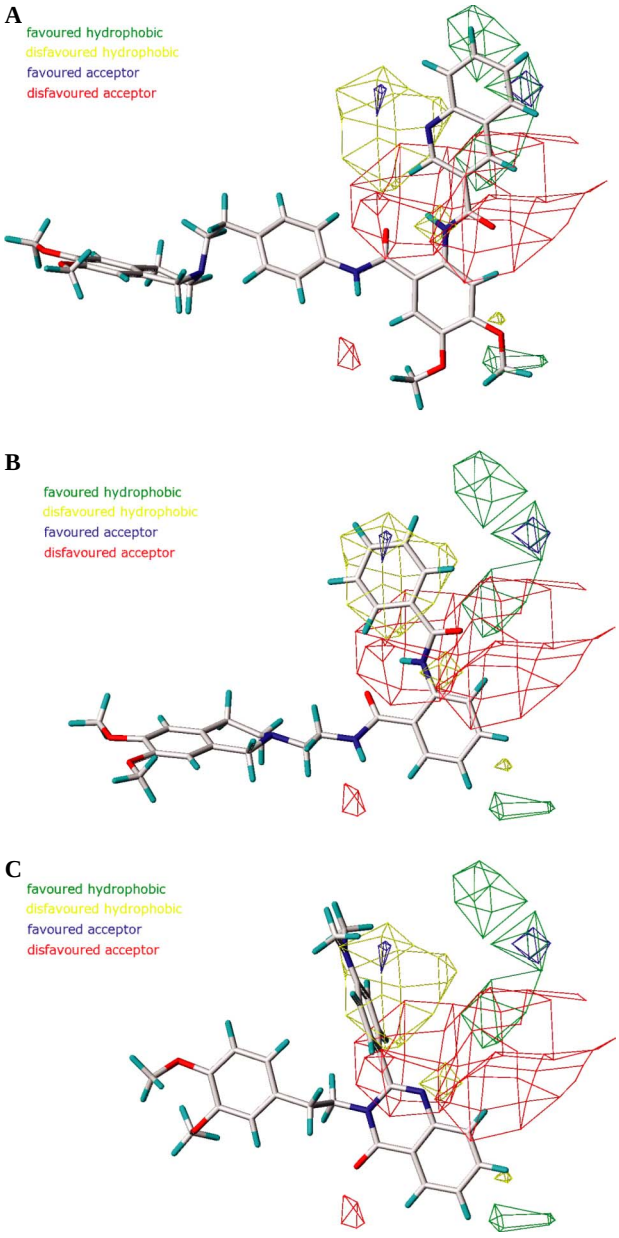


Figure 5. CoMSIA contour plot (SD*Coeffficient; favored level 80%, disfavoured level 20%) of the 'h+a' model (see Table 4) with compound XR9576 (A), 7 (B), and 3 (C).

One of the most important results of this study is identification of the functional groups and atoms possibly related to the binding and inhibitory potency of the compounds studied. The Free-Wilson analysis outlined the presence of a tetrahydroquinoline substructure bonded to the anthranilamide core through a phenyl moiety and of a heteroatom in position 3 in a bulky aromatic ring system in the R-part to have the most significant impact on MDR reversal. The 3D-QSAR contour plots confirmed these results and clearly illustrated that the structural features with the highest contribution to activity variance related mostly to the R-part of the structures (Fig. 5). Considering that the compounds were aligned according to the pharmacophore pattern identified on the structure of Hoechst 33342, it can be suggested that the parts of the XR-structures that have their correspondence in the Hoechst 33342 structure (Fig. 2) can be related to the potency of the compounds to bind to the same site on P-gp as the substrate does. Here, the anthranilamide core and the tetrahydroquinoline substructure bonded to the core through the phenyl moiety appear to be essential. The anthranilamide nucleus provides HB-intercation points; the phenyl moiety is an essential hydrophobic point in correspondence with the pharmacophore pattern of Hoechst 33342 recently obtained.[24] The R-substituents in the molecules that have no corresponding parts in the structure of Hoechst 33342 and that show the most structural variation can presumably be related to their modulating effect. However, the way the XR-compounds may bind to the H-site of P-gp so as to exert an inhibitory effect on its transport activity remains to be established. Efforts can be directed to modifications of this part in order to optimize the modulating effect of the synthesized compounds.

In this study, the 3D-QSARs HB-acceptor, steric, and hydrophobic fields were shown to be the most important 3D properties in agreement with our previous studies about the role of the steric and hydrophobic interactions of the MDR modulators[9,17,18] and the general understanding about the crucial role of the hydrophobic and HB-interactions with P-gp.[10–15] The best models have high internal and external predictivity and can be used for prediction of MDR reversal of new analogs of XR9576.

4. Experimental

4.1. Methods

The pK_a values of the compounds studied were calculated with the program ACD/pK_a[19] employing a fragment-based approach. The so-called 'apparent constants' were calculated: the algorithm mimics the experimental order of protonation and determines the pK_a values, which can be really experimentally measured in water solution. The Free-Wilson analysis was performed with the multiple linear regression module of STATISTICA.[20] The regression models were estimated by the correlation coefficient r^2 and the cross-validated correlation coefficient r^2_{cv}.

Molecular modeling was done with the program Sybyl[21] running under IRIX 6.5 on a Silicon Graphics Octane 2 (R12000) and MOE (Molecular Operating Environment)[16] running under Linux x86 workstation. The conformational analysis was performed by simulated annealing in Sybyl and by the stochastic and pharmacophore search in MOE. In the simulated annealing, repeated heating and cooling of the molecule is simulated to cover the conformational space. In the stochastic search, a scrambling of the absolute coordinates of the atoms is performed followed by minimization of the structure. The pharmacophore search performs a rigid superposition of the computed conformers on a predefined pharmacophore pattern. The force field methods used were Tripos (Sybyl) and MMFF94s (MOE).

The AM1 semiempirical method (MOPAC[22]) was applied for geometrical optimization and calculation of the partial atomic charges. 3D-QSARs were derived with the CoMFA and CoMSIA approaches as implemented in Sybyl. A principal component analysis (PCA), using the factor analysis module of Sybyl, was performed on the CoMSIA fields to select a representative test set of compounds. The external predictive power of the models was estimated by r^2_{pred}[21] and p^2.[23]

4.2. Conformational analysis

As no X-ray data of Hoechst 33342 and the XR9576 analogs were available, the energy minimum conformers of the compounds were generated from the sketched and minimized structure of Hoechst 33342 (Tripos force field, 0.05 kcal mol^{-1} Å^{-1} gradient, no charges). Two techniques were applied: simulated annealing[21] and stochastic search[16] for generation of the Hoechst 33342 structures. Simulated annealing was performed with 100 cycles, 2000 K initial temperature for heating for 2000 fs equilibration, 0 K target temperature for 5000 fs annealing time, and exponential annealing function. The obtained 100 local minima were then optimized by the semiempirical quantum chemistry method AM1 (full optimization, precise convergence, and 'xyz' keyword) as implemented in MOPAC 6.0.[22] In the stochastic search, the MMFF94s force field was used with the default MOE settings and the following parameters were increased to ensure a better covering of the conformational space: energy cutoff 10 kcal/mol, failure limit 200 (contiguous number of attempts to generate a new conformation), RMS tolerance 0.5, iteration limit 20000, and minimization iteration limit 2000. Both techniques produced similar results. Four clusters of conformations were identified and the lowest energy conformer of each cluster was taken for further analysis. The conformers differed in the position of the benzimidazole ring with the attached methylpiperazine. In one pair of conformers, the imidazole nitrogens (–NH and =N–) in both benzimidazole rings had opposite orientations just as shown that in Figure 2A. In another pair, vice versa, the benzimidazole ring with the attached piperazine was rotated by 180°. The four conformers had AM1 heats of formation in the interval from 139.5 to 142.6 kcal/mol. As the heats of formation were very close, all four conformers of Hoechst 33342 were used as templates for identification of its pharmacophore pattern, to decide on the best conformer.

A stochastic search was performed with all XR-compounds. The settings were the same as those used for Hoechst 33342 (see above) but additionally rotation of the amide bonds was allowed. The search for the most appropriate conformations was performed with the pharmacophore search in MOE.[16] The final structures of XR-compounds were built from the templates obtained through the pharmacophore search (see Results). The structures were then minimized again with Tripos force field (0.05 kcal mol^{-1} Å^{-1} gradient, no charges) keeping the pharmacophore template as an aggregate and afterwards optimized with AM1 (full optimization, precise convergence, and 'xyz' and 'mmok' keywords). The protonated forms of the drugs were built from the neutral ones and an additional optimization with AM1 was performed setting a net charge of +1. Compound **3** was aligned on the template of the smaller compounds lacking the phenyl spacer.

4.3. CoMFA and CoMSIA settings

The standard settings were used in the CoMFA calculations:[21] 2 Å regular grid size in all three directions within the automatically created grid box with 4 Å extension beyond the van der Waals volume of the overlaid molecules, sp^3 carbon probe with +1 charge, and a distance dependent ($1/r$) dielectric constant. The following fields were calculated in CoMFA: steric (s), electrostatic (e), both (both), and hydrogen bond (h-bnd). The AM1 point charges were used for calculation of the electrostatic fields. The standard energy cutoff value of 30 kcal/mol with no electrostatic interactions at bad steric contacts was used. The threshold column filtering was set to 1.0 kcal/mol. In CoMSIA, the following similarity indices fields were calculated: steric (s), electrostatic (e), hydrophobic (h), hydrogen bond donor (d), hydrogen bond acceptor (a), hydrogen bond donor (d), and donor-acceptor (da) with the default attenuation factor of 0.3 in the same grid box as used for CoMFA. A common probe atom with 1 Å radius and charge, hydrophobicity, and hydrogen bond property of +1 was used. The indices were evaluated according to the usual CoMSIA protocol with 1.0 kcal/mol column filtering.

The internal predictive power of the models was evaluated first by leave-one-out (LOO) cross-validation basing on the cross-validated coefficient q^2_{cv}, the optimal number of components n_{opt}, and the predictive sum of squares PRESS. The external predictive power of the obtained models was estimated by the predictive correlation coefficient r^2_{pred}.[21] Additionally, a parameter called p^2 was calculated as follows:

$$p^2 = 1 - \frac{\sum_{i,j}[(p_i - p_j) - (m_i - m_j)]^2}{\sum_{i,j}(m_i - m_j)^2},$$

where p_i, p_j are predicted activities and m_i, m_j are measured activities.[23] It describes the predictive power of the model based on the pairwise predicted and observed activity differences and is therefore independent of the additive constants like means of the training and test sets (see Table 5).

Visualization of the best achieved model is presented as SD*Coefficient contour plots (80% favored/20% disfavored level) and illustrates regions where variability in molecular fields explains target property differences.

Acknowledgments

I.P. and M.W. thank the Deutsche Forschungsgemeinschaft for the main financial support (Grant 436 BUL 17/7/04). C.G. and M.W. thank the Deutsche Forschungsgemeinschaft for the financial support to

Graduiertenkolleg 804. I.P. thanks also the National Science Fund of Bulgaria (L-1416).

References and notes

1. Tsuruo, T.; Iida, H.; Tsukagoshi, S.; Sakurai, Y. *Biochem. Pharmacol.* **1982**, *31*, 3138–3140.
2. Thomas, H.; Coley, H. M. *Cancer Control* **2003**, *10*, 159–165.
3. Mistry, P.; Stewart, A. J.; Dangerfield, W.; Okiji, S.; Liddle, C.; Bootle, D.; Plumb, J. A.; Templeton, D.; Charlton, P. *Cancer Res.* **2001**, *61*, 749–758.
4. Walker, J.; Martin, C.; Callaghan, R. *Eur. J. Cancer* **2004**, *40*, 594–605.
5. Kohler, S.; Stein, W. D. *Biotechnol. Bioeng.* **2003**, *81*, 507–517.
6. Martin, C.; Berridge, G.; Mistry, P.; Higgins, C.; Charlton, P.; Callaghan, R. *Br. J. Pharmacol.* **1999**, *128*, 403–411.
7. Roe, M.; Folkers, A.; Ashworth, P.; Brumwell, J.; Chima, L.; Hunjan, S.; Pretswell, I.; Dangerfield, W.; Ryder, H.; Charlton, P. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 595–600.
8. Martin, C.; Berridge, G.; Higgins, C. F.; Mistry, P.; Charlton, P.; Callaghan, R. *Mol. Pharmacol.* **2000**, *58*, 624–632.
9. Wiese, M.; Pajeva, I. *Curr. Med. Chem.* **2001**, *8*, 685–713.
10. Penzotti, J. E.; Lamb, M. L.; Evensen, E.; Grootenhuis, P. D. *J. Med. Chem.* **2002**, *45*, 1737–1740.
11. Ekins, S.; Kim, R. B.; Leake, B. F.; Dantzig, A. H.; Schuetz, E. G.; Lan, L. B.; Yasuda, K.; Shepard, R. L.; Winter, M. A.; Schuetz, J. D.; Wikel, J. H.; Wrighton, S. A. *Mol. Pharmacol.* **2002**, *61*, 974–981.
12. Garrigues, A.; Loiseau, N.; Delaforge, M.; Ferte, J.; Garrigos, M.; Andre, F.; Orlowski, S. *Mol. Pharmacol.* **2002**, *62*, 1288–1298.
13. Pajeva, I. K.; Wiese, M. *J. Med. Chem.* **2002**, *45*, 5671–5686.
14. Klebe, G.; Abraham, U. *J. Comput. Aided Mol. Des.* **1999**, *13*, 1–10.
15. Tsakovska, I. *Bioorg. Med. Chem.* **2003**, *11*, 2889–2899.
16. *MOE 2002.03 (Molecular Operating Enviroment)*, Chemical Computing Group, 1010 Sherbrooke Street West, Suite 910; Montreal, Que., Canada H3A 2R7.
17. Pajeva, I. K.; Wiese, M. *J. Med. Chem.* **1998**, *41*, 1815–1826.
18. Pajeva, I. K.; Wiese, M. *Quant. Struct.-Act. Relat.* **1998**, *17*, 301–312.
19. *ACD/LogD* Suite, version 5.0, Advanced Chemistry Development, Toronto ON, Canada, www.acdlabs.com, 2003.
20. *SPSS*, version 11, SPSS, Chicago IL, USA.
21. *Sybyl*, version 6.9; Tripos, 1699 South Hanley Road, St. Louis, MO 63114-2917; October 2002.
22. *MOPAC 6*, (QCPE No. 445): Department of Chemistry, Indiana University, Bloomington, IN 47405, USA.
23. Silverman, D. *Quant. Struct.-Act. Relat.* **2000**, *19*, 237–246.
24. Pajeva, I. K.; Globish, C.; Wiese, M. *J. Med. Chem.* **2004**, *47*, 2523–2533.