Self-Organizing Maps for Identification of New Inhibitors of P-Glycoprotein

Dominik Kaiser,[†] Lothar Terfloth,[‡] Stephan Kopp,[§] Jan Schulz,[||] Randolf de Laet,[||] Peter Chiba,[§] Gerhard F. Ecker,*[,†] and Johann Gasteiger[‡]

Emerging Focus Pharmacoinformatics, Department of Medicinal Chemistry, University of Vienna, Althanstrasse 14, 1090 Wien, Austria, Computer-Chemie-Centrum, Institute of Organic Chemistry, University of Erlangen-Nuernberg, Germany, Institute of Medical Chemistry, Medical University of Vienna, Waehringer Strasse 10, 1090 Wien, Austria, and SPECS and BioSPECS B.V., Fleminglaan 16, 2289 CP Rijswijk, The Netherlands

Received May 22, 2006

Self-organizing maps were trained to separate high- and low-active propafenone-type inhibitors of P-glycoprotein. The trained maps were subsequently used to identify highly active compounds in a virtual screen of the SPECS compound library.

Introduction

ATP-dependent transport proteins have been shown to play a major role both for bioavailability of drugs[1] and for development of drug resistance in bacteria and men.[2] In tumor cells, overexpression of these membrane-bound proteins is responsible for decreased intracellular accumulation of therapeutically administered xenotoxins, thus leading to multiple drug resistance. In cancer cells, efflux driven broad spectrum resistance to chemotherapeutic agents is predominantly mediated by the ABCB1 gene product P-glycoprotein (P-gp). Inhibition of P-gp leads to resensitization of multidrug resistant tumor cells in vitro and was thus considered a promising approach for treatment of multidrug resistant tumors.[3] Currently, several compounds are in clinical phase III studies. Although these studies have raised some concerns on the broad clinical use of P-gp inhibitors, there are still ongoing efforts in identifying new P-gp inhibitors.[4] This is mainly due to the involvement of P-gp in bioavailability and brain permeation of drugs.[5] Thus, a phase I clinical study clearly showed that the highly active P-gp inhibitor GF 120918 is able to enhance plasma levels of topotecan by more than 4-fold.[6] There are also several patents claiming P-gp inhibitors as versatile tools for enhancing brain uptake of cns-active drugs.[7]

One of the major characteristics of P-gp is its promiscuity (or multispecificity) in the binding of ligands. This is supported by findings that suggest multiple overlapping binding sites.[8] Use of homologous series of compounds identified both predictive physicochemical parameters and pharmacophoric substructures.[9] CoMFA[a] and CoMSIA analyses lead to distinct 3D-QSAR models for propafenones[10] and phenothiazines.[11] Recently, three-dimensional pharmacophore models have been proposed based on in vitro data for digoxin transport in Caco-2 cells, vinblastine binding in CEM/VLB_{100} cells, and vinblastine and calcein accumulation in LLC-PK1 cells.[12] Additionally, the definition of pharmacophores and alignment was utilized using the genetic algorithm-based similarity program GASP. The proposed general pharmacophore pattern involves two hydrophobic planes, three hydrogen-bond acceptors, and one hydrogen-bond donor.[13] However, applying these models for virtual screening approaches requires multiconformational three-dimensional databases.

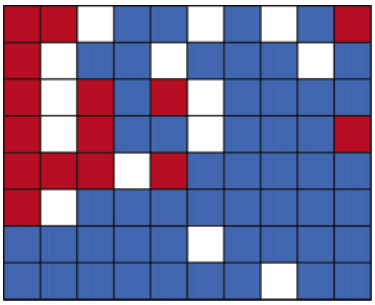


Figure 1. Self-organizing map trained using descriptor set for model 1: red, low active; blue, highly active; white, empty.

In recent years, nonlinear methods were also successfully applied for prediction of polyspecific drug−protein interactions. These include both feedforward backpropagation artificial neural networks and support vector machines.[14] Recently, Wang et al. applied unsupervised and supervised learning approaches for classification of P-glycoprotein substrates and inhibitors.[15] In this paper we describe the use of self-organizing maps (SOMs) for the discovery of new lead compounds in the field of P-gp inhibitors.

Results and Discussion

Identification of new lead compounds via in silico screening of large databases is currently one of the most challenging tasks in the drug development process. Our approach for identification of new P-gp inhibitors is based on the simultaneous presentation of a large compound library (i.e., SPECS) and a training set of 131 P-gp inhibitors covering a broad activity range to a self-organizing map. The concept is based on the expectation that compounds co-localizing with highly active drugs from the training set also show high activity.

The principal ability of SOMs is to obtain a 2D (two-dimensional)-rendering of a multidimensional space, which brings similar compounds in close vicinity on the map (i.e., within the same neuron). SOMs were successfully applied to the design of combinatorial libraries, filtering of HTS libraries, and to distinguish drugs from nondrugs.[16] In a recent paper, Polansky and co-workers described the use of a SOM for screening and development of novel artificial sweetener candidates.[17]

* To whom correspondence should be addressed. Phone: +43-1-4277-55110. Fax: +43-1-4277-9551. E-mail: gerhard.f.ecker@univie.ac.at.
[†] University of Vienna.
[‡] University of Erlangen-Nuernberg.
[§] Medical University of Vienna.
[||] SPECS.

[a] Abbreviations: CoMFA, comparative molecular field analysis; CoMSIA, comparative molecular similarity indices analysis; HTS, high-throughput screening; P-gp, P-glycoprotein; SOM, self-organizing map.

10.1021/jm060604z CCC: $37.00
Published on Web 03/13/2007

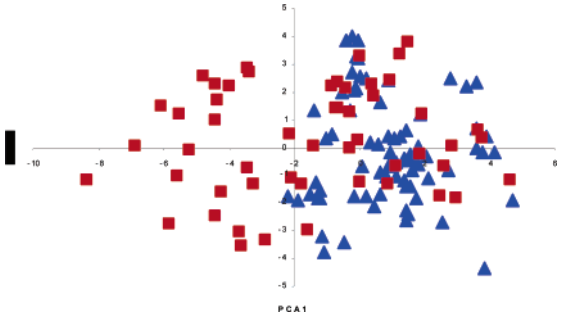


Figure 2. Plot of the first two principal components of descriptors for model 1: red squares, low actives; blue triangles, high actives.

Chart 1. Chemical Structure of Highly Active Propafenone Analogs **1** and **2**

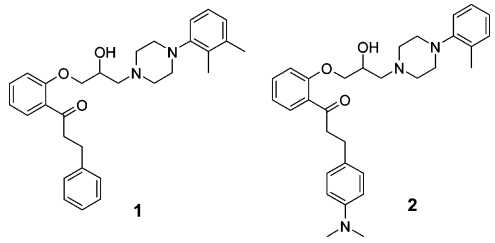
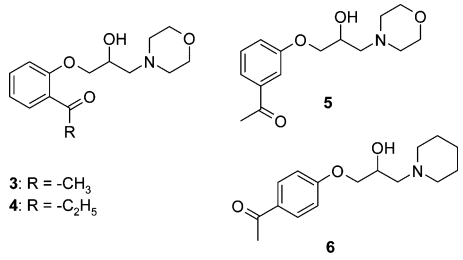


Chart 2. Chemical Structures of Low-Active Propafenone Analogs **3**−**6**



First of all, a training set of 131 propafenone-type inhibitors of P-gp was used to explore the general possibility to distinguish between compounds with high and low activity using self-organizing maps. Using 2D autocorrelation vectors based on seven atom properties, two different models were retrieved that both showed a good discrimination between compounds of high and low activity (Figure 1 and Figure 1, Supporting Information). Comparison with the results obtained from a principal component analysis clearly demonstrated the higher ability of SOMs to separate high- from low-active propafenone-type P-gp inhibitors (Figure 2).

Thus, for identification of new lead compounds, the 131 propafenone analogs were merged with 134 767 compounds from the SPECS database, and the complete data set was used to establish two SOMs under conditions identical to those used in the runs with the propafenones. Obviously, the size of the maps had to be adapted to the higher number of compounds to avoid an unreasonably high number of co-localizations. Table 1 gives the number of SPECS compounds co-localizing with propafenones obtained with the two models and two different network sizes. As expected, the number of co-localizations increases with decreasing network size.

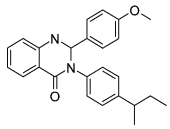
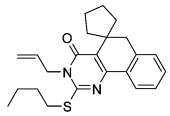
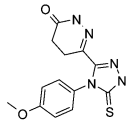
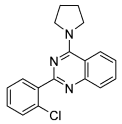
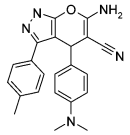
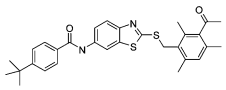
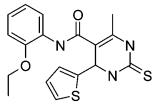
Merging the hit lists from models 1a and 2a, elimination of duplicates and restriction to compounds which co-localize with propafenones with an EC_{50} value < 0.16 μmol/L (which include the top 15% in activity) gave a set of 43 compounds. Twelve out of these 43 compounds were located in the same neuron as **1** and **2** (Chart 1), the hitherto most active propafenone type P-gp inhibitors (EC_{50} = 0.006 and 0.013 μmol/L, respectively). Some of these virtual screening hits showed an identical scaffold, which further reduced the number of structurally diverse hits to seven (Table 2). These compounds were pharmacologically tested in the daunorubicin efflux assay. The results show that two of the compounds were highly active, with EC_{50} values below 1 μmol/L, four compounds had activities between 1 and 10 μmol/L, and only one compound was inactive.

Conversely, compounds co-localizing with low-active propafenone analogs were also tested and used as an additional proof of concept. A total of 22 compounds co-localizing with the least active propafenones **3** (EC_{50} = 207 μmol/L), **4** (EC_{50} = 67 μmol/L), **5** (EC_{50} = 128 μmol/L), and **6** (EC_{50} = 49 μmol/L) were identified. From each subset, the two compounds with the lowest and highest calculated MlogP values were selected and the biological activity of these eight compounds was

Table 1. Dimensions of the SOMs Used and Number of Co-Localizations

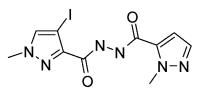
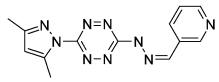
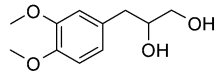
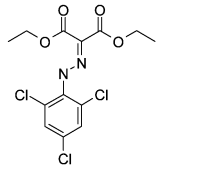
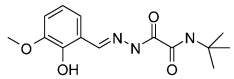
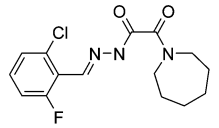
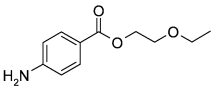
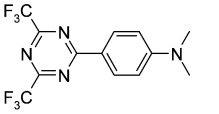
	no. descriptors	dimension	no. co-localizations
model 1a	21	360 × 360	107
model 1b	21	250 × 250	243
model 2a	25	360 × 360	113
model 2b	25	250 × 250	267

Table 2. Chemical Structure and Pharmacological Activity of Compounds Proposed as Highly Active in the In Silico Screen

Structure	Code	EC_{50}[a]
	**7** (AG-690/119727729	0.76
	**8** (AG-690/12887361)	3.50
	**9** (AJ-131/15197008)	> 250[b]
	**10** (AJ-292/13162028)	5.94
	**11** (AJ-292/15089034)	3.08
	**12** (AN-989/14669159)	0.28
	**13** (AO-364/14480185)	8.77

[a] Given in μmol/L. [b] No effect up to this concentration.

Table 3. Chemical Structure and Pharmacological Activity of Compounds Proposed as Inactive in the In Silico Screen

Structure	Code	EC_{50}
	**14** (AK-968/12163579)	> 500
	**15** (AG-227/33912017)	158.3
	**16** (AI-942/13331556)	> 500
	**17** (AF-399/34011064)	17.7
	**18** (AG-205/33114008)	102.2
	**19** (AN-698/40745334)	159.3
	**20** (AF-399/13927090)	311.8
	**21** (AG-690/32530028)	> 100

determined (Chart 2, Table 3). Only one compound showed an EC_{50} value below 100 μmol/L.

To further prove the concept, the SPECS library was screened for compounds structurally similar to the two hits **7** and **12** on the basis of the Tanimoto index values higher than 0.8. Using the Tanimoto index ensures that only structurally very similar compounds are identified. A subset of the compounds was retrieved and biologically tested. In the case of quinazolinones **7**, five out of eight compounds tested showed pharmacological activity in the micromolar and submicromolar range (Table 4). Additionally, there was also a good correlation between calculated MlogP values and $\log(1/EC_{50})$ values found ($r = 0.83$). This is in agreement with previous findings demonstrating that lipophilicity is a general predictive factor for P-gp inhibitory activity.[18] In the case of benzothiazoles **12**, three out of five compounds tested showed pharmacological activities in the submicromolar range (Table 5). These results confirmed that with **7** and **12** indeed two new scaffolds for the design of P-gp inhibitors were identified.

When analyzing the location of compounds **7a**−**7h** and **12a**−**12e** on the 250 × 250 neurons SOM, most of the highly active analogs are indeed located in close vicinity to their parent hits **7** and **12** (Figure 2, Supporting Information), whereas low-active derivatives **7a**, **7b**, **12d**, and **12e** are placed differently. However, it has to be noted that the inactive compound **7f** is located close to **7**, and the active derivatives **7c** and **12c** are not part of the highly active compound cluster around **7** and **12**. This outlines the complexity of the method presented and further strengthens the need for thorough validation procedures.

Table 4. Chemical Structure, Chemical Similarity, Calculated MlogP Values, and Pharmacological Activity of Compounds Structurally Similar to **7** (AG-690/11972772)

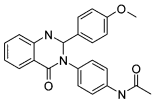
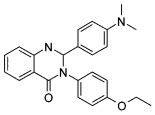
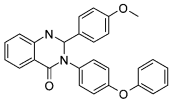
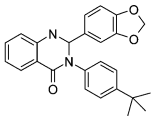
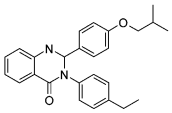
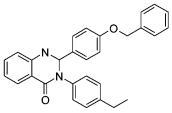
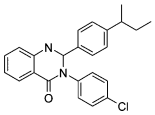
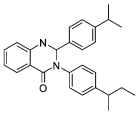
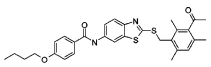
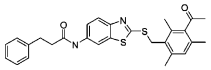
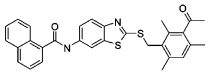
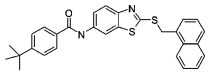
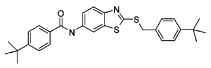
Structure	Code	Simil.	logP	EC_{50}
	**7a** (AG-205/40775655)	0.86	1.86	20.93
	**7b** (AG-205/36953425)	0.85	2.51	25.13
	**7c** (AG-205/40775656)	0.86	2.91	1.00
	**7d** (AO-081/15569243)	0.86	3.12	0.73
	**7e** (AG-205/11867035)	0.90	3.71	1.21
	**7f** (AG-205/37047245)	0.88	4,09	> 100
	**7g** (AG-205/37047268)	0.87	4.31	0.82
	**7h** (AG-690/11972774)	0.91	4.72	0.29

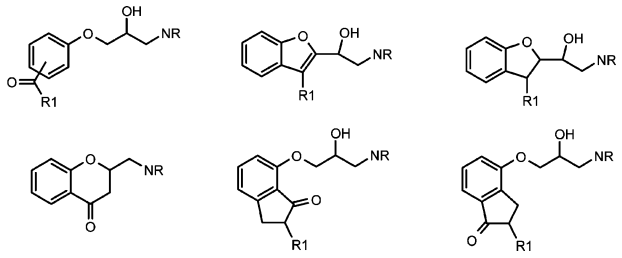
Table 5. Chemical Structure, Chemical Similarity, Calculated LogP Values, and Pharmacological Activity of Compounds Structurally Similar to **12** (AN-989/14669159)

Structure	Code	Simil.	logP	EC_{50}
	**12a** (AN-989/14669164)	0.86	3.16	0.097
	**12b** (AN-989/14669106)	0.88	3.56	0.30
	**12c** (AN-989/14669155)	0.94	3.81	0.11
	**12d** (AN-989/14669060)	0.87	4.49	> 20
	**12e** (AG-205/11373354)	0.90	4.63	> 50

Conclusion

In this paper we demonstrated that self-organizing maps in combination with autocorrelation vectors are a versatile tool for virtual screening of medium-sized compound libraries. Compounds retrieved as potential hits show structural scaffolds differing from those used in the training set. This method is therefore suitable for identification of structurally unrelated, diverse hits and thus represents a versatile tool for scaffold

Chart 3. Scaffolds Used for the Training Set



hopping. Further work will prove whether applicability of this approach extends to other biological targets.

Materials and Methods

Training Set: As a reference data set, our in house library of 131 propafenone-type inhibitors of P-gp was used.[19] The data set comprises phenones, benzofuranes, indanones, and benzopyranes (Chart 3; chemical structures and EC_{50} values are given in Table 1, Supporting Information). Compounds were separated into two activity groups (high activity and low activity) by using a threshold value of 1 μM.

Screening Library: For in silico screening, the SPECS database[20] cleaned from salts, duplicates, and compounds with a molecular weight lower than 100 and higher than 800 was used (number of compounds remaining: 134 767).

Descriptors: For this study, topological autocorrelation vectors for a set of atom properties were used as descriptors. The topological autocorrelation vector **A** for the topological distance d (number of bonds between two atoms) is calculated by

$$\mathbf{A}(d) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} p_i p_j \delta(d,d_{ij}) \qquad \delta(d,d_{ij}) = \begin{cases} 1 \forall d = d_{ij} \\ 0 \forall d \neq d_{ij} \end{cases} \quad (1)$$

with the topological distance of atoms i and j, d_{ij}, and their properties p_i and p_j, respectively.

For both data sets, the following atom properties were calculated using the software package PETRA:[21] atom-polarizability (α_i),[22] σ- (χ_σ), π- (χ_π), and lone pair electronegatives (χ_{lp}),[23,24] and σ- (q_σ),[25] π- (q_π),[26] and total charges ($q_{tot} = q_\sigma + q_\pi$). Subsequently, the topological autocorrelation vectors for distances from 0 to 10 bonds were calculated with AUTOCORR,[27] scaled to unit variance, and used as descriptors for the input vector to the SOM.

Self-Organizing Map: Studies were done with the software package SONNIA.[28] SONNIA is the implementation of a self-organizing network introduced by Teuvo Kohonen.[29] Objects from a multidimensional space are projected into a space of lower dimensionality, usually into a 2D plane. In this projection, the topology of the input space is preserved. Thus, Kohonen neural networks can be applied to cluster objects for similarity perception. The training of the network is unsupervised, that is, the property of interest, here it is the compound's biological activity, is not used during the training process. In the course of training, the objects are randomly presented to the neural network in an iterative manner. For each iteration step, the winning neuron for the input object is identified by determining the neuron having the minimum *Euclidean* distance to the input object. To improve the response of the network, the neurons weights are adapted to become more similar to the input pattern. After termination of training, the response of the network is calculated for each object of the data set. The projection of the data set into the 2D space is then performed by mapping each object into the coordinates of its winning neuron.

Initially, the autocorrelation vectors for each atom property were used separately to select those descriptors that led to a good discrimination between active and inactive propafenones. All maps had a rectangular dimension of 10 × 8 neurons. This network size represents a good compromise between the number of collisions and the number of empty neurons (Figure 2, Supporting Information). Each neuron was color coded according to the activity of the compounds located in the neuron.

As shown earlier, 2D maps allow one to make full use of the human pattern recognizing power. By visual inspection of the color-coded maps (Figure 1), those variables could be selected that led to the best clustering.[30] Best results were obtained with the descriptors α_i, χ_σ, q_σ, q_{tot}, and χ_{lp}. Thus, two models were generated using the following combinations of descriptors: α_i, q_σ, χ_σ (model 1) and α_i, q_{tot}, χ_{lp} (model 2). This procedure was chosen to keep the number of descriptors used in a reasonable relation to the number of training set compounds. Intercorrelated descriptors ($r > 0.95$) were removed from the data matrix to avoid redundancy. Input vectors had 21 dimensions for model 1 and 25 for model 2.

Principal Component Analysis. Principal component analysis was performed using the software package MOE (Chemical Computing Group). Both descriptor combinations (from models 1 and 2) were subject to principal component analysis. A plot of the first two principal components explaining 48.5% of the variance is shown in Figure 2. Compounds with high activity are represented as blue triangles and those with low activity as red squares.

Screening of the SPECS Library. Both input files (propafenones and SPECS-compounds) were merged, and the data were normalized via z-scaling. Two different network sizes were used (Table 1), and the SOMs were trained applying the same conditions as for the small data set of propafenones. In both models, all compounds of the SPECS library co-localizing with a propafenone-type P-gp inhibitor were retrieved. Hit selection was performed as outlined in the Results and Discussion section.

Similarity Search: The SPECS library was converted from the SD file to a UNITY 4.3 database[31] using the "dbimport" command. Subsequently, the UNITY fingerprints were calculated with "db-mkscreen", and the search was done using the SELECTOR similarity search tool[31] with a minimum Tanimoto similarity of 0.8. Alternatively, we also applied autocorrelation vectors and Euclidian distance for similarity searching. However, in this case, also compounds with scaffolds different from the query structure were retrieved. Thus, the standard SELECTOR protocol was used to ensure that predominantly compounds with the same basic structural scaffold were identified.

Calculation of MlogP Values: For calculation of logP values, the method of Moriguchi et al. was used,[32] as implemented in an spl-macro in the SYBYL molecular modeling package. This method has been shown to yield very reliable results, especially when dealing with structurally diverse compounds.

Biological Activity: The pharmacological activity of the compounds was measured in a zero trans efflux protocol using daunorubicin as the fluorochrome.[33] Briefly, multidrug resistant CCRF vcr1000 cells were incubated with daunorubicin, and the time-dependent decrease in mean cellular fluorescence was measured in the absence and presence of various concentrations of the modulator. EC_{50} values were calculated from the concentration−response curve of efflux first-order rate constants (V_{max}/K_m) plotted as a function of the modulator concentration. Thus, the effect of different modulators on the transport rate is measured in a direct functional assay. Values of newly identified hits are given in Tables 2, 4, and 5 and represent the mean of at least three independently performed experiments. Generally, interexperimental variation was below 20%.

Acknowledgment. We gratefully acknowledge financial support by the Austrian Science Fund, Project P-17014-B11. We are grateful to the Vienna University Computer Center for providing us with several software packages and for CPU time at the Schroedinger Cluster.

Supporting Information Available: Table giving the chemical structures and EC_{50} values of 131 propafenone derivatives of the training set and Figure showing the location of these compounds on the self-organizing map trained using the descriptor set for model 2. This material is available free of charge via the Internet at http://pubs.acs.org

References

(1) Chan, L. M. S.; Lowes, S.; Hirst, B. H. The ABCs of drug transport in intestine and liver: Efflux proteins limiting drug absorption and bioavailability. *Eur. J. Pharm. Sci.* **2004**, *21*, 25−51.

(2) Lage, H. ABC-transporters: Implications on drug resistance from microorganisms to human cancers. *Int. J. Antimicrob. Agents* **2003**, *22*, 188−199.

(3) Gottesman, M. M.; Fojo, T.; Bates, S. E. Multidrug resistance in cancer: Role of ATP-dependent transporters. *Nat. Rev. Cancer* **2002**, *2*, 48−58.

(4) Szakacs, G.; Paterson, J. K.; Ludwig, J. A.; Boothe-Genthe, C.; Gottesman, M. M. Targeting multidrug resistance in cancer. *Nat. Rev. Drug Discovery* **2006**, *5*, 219−234.

(5) Breedveld, P.; Beijnen, J. H.; Schellens, J. H. M. Use of P-glycoprotein and BCRP inhibitors to improve oral bioavailability and CNS penetration of anticancer drugs. *Trends Pharmacol. Sci.* **2006**, *27*, 17−24.

(6) Kruijtzer, C. M.; Beijnen, J. H.; Roding, H.; ten Bokkel Huinink, W. W.; Schot, M.; Jewell, R. C.; Paul, E. M.; Schellens, J. H. Increased oral bioavailability of topotecan in combination with the breast cancer resistance protein and P-glycoprotein inhibitor GF120918. *J. Clin. Oncol.* **2002**, *20*, 2943−2950.

(7) Chiba, P.; Ecker, G. F. Inhibitors of ABC-type drug efflux pumps: an overview of the current patent situation. *Expert Opin. Ther. Pat.* **2004**, *14*, 499−508.

(8) Garrigues, A.; Loiseau, N.; Delaforge, M.; Ferte, J.; Garrigos, M.; Andre, F.; Orlowski, S. Characterization of two pharmacophores on the multidrug transporter P-glycoprotein. *Mol. Pharmacol.* **2002**, *62*, 1288−1298.

(9) Avendano, C.; Menendez, J. C. Inhibitors of multidrug resistance to antitumor agents (MDR). *Curr. Med. Chem.* **2002**, *9*, 159−193.

(10) Pajeva, I. K.; Wiese, M. A comparative molecular field analysis of propafenone-type modulators of cancer multidrug resistance. *Quant. Struct.-Act. Relat.* **1998**, *17*, 301−312.

(11) Pajeva, I.; Wiese, M. Molecular modeling of phenothiazines and related drugs as multidrug resistance modifiers: A comparative molecular field analysis study. *J. Med. Chem.* **1998**, *41*, 1815−1826.

(12) Ekins, S.; Kim, R. B.; Leake, B. F.; Dantzig, A. H.; Schuetz, E. G.; Lan, L.; Yasuda, K.; Shepard, R. L.; Winter, M. A.; Schuetz, J. D.; Wikel, J. H.; Wrighton, S. A. Application of three-dimensional quantitative structure−activity relationships of P-glycoprotein inhibitors and substrates. *Mol. Pharmacol.* **2002**, *61*, 974−980.

(13) Pajeva, I. K.; Wiese, M. Pharmacophore model of drugs involved in P-glycoprotein multidrug resistance: Explanation of structural variety (hypothesis). *J. Med. Chem.* **2002**, *45*, 5671−5686.

(14) Xue, Y.; Yap, C. W.; Sun, L. Z.; Cao, Z. W.; Wang, J. F.; Chen, Y. Z. Prediction of P-glycoprotein substrates by a support vector machine approach. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1497−1505.

(15) Wang, Y. H.; Li, Y.; Yang, S. L.; Yang, L. Classification of substrates and inhibitors of P-glycoprotein using unsupervised machine learning approach. *J. Chem. Inf. Model.* **2005**, *45*, 750−757.

(16) Anzali, S.; Gasteiger, J.; Holzgrabe, U.; Polanski, J.; Sadowski, J.; Teckentrup, A.; Wagener, M. The use of self-organizing neural networks in drug design. *Perspect. Drug Discovery Des.* **1998**, 273−299.

(17) Polanski, J.; Gasteiger, J.; Jarzembek, K. Self-organizing neural networks for screening and development of artificial sweetener candidates. *Comb. Chem. High Throughput Screening* **2000**, *3*, 481−495.

(18) Ecker, G.; Chiba, P.; Hitzler, M.; Schmid, D.; Visser, K.; Cordes, H. P.; Csoellei, J.; Seydel, J. K.; Schaper, K. J. Structure−activity relationship studies on benzofurane analogs of propafenone-type modulators of tumor cell multidrug resistance. *J. Med. Chem.* **1996**, *39*, 4767−4774.

(19) Klein, C.; Kaiser, D.; Kopp, S.; Chiba, P.; Ecker, G. F. Similarity based SAR (SIBAR) as tool for early ADME profiling. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 785−793.

(20) SPECS and BioSPECS B.V., Fleminglaan 16, 2289 CP Rijswijk, The Netherlands; http://www.specs.net.

(21) Gasteiger, J. Empirical methods for the calculation of physicochemical data of organic compounds. In *Physical property prediction in organic chemistry*; Jochum, C., Hicks, M. G., Sunkel, J., Eds.; Springer-Verlag: Heidelberg, 1988; pp 119−138. *PETRA*, version 3; Molecular Networks GmbH: Naegelsbachstr. 25, 91052 Erlangen, Germany, http://www.mol-net.de (and can be tested via the internet at http://www2.chemie.uni-erlangen.de/software/petra).

(22) Gasteiger, J.; Hutchings, M. G. Quantification of effective polarisability. Application to studies of X-ray photoelectron spectroscopy and alkylamine protonation. *J. Chem. Soc., Perkin Trans. 2* **1984**, 559−564.

(23) Marsili, M.; Gasteiger, J. π-Charge distributions from molecular topology and π-orbital electronegativity. *Croat. Chem. Acta* **1980**, *53*, 258−259.

(24) Hanebeck, W.; Gasteiger, J. Rapid empirical calculation of the first (n or π) ionization potential of organic molecules. *J. Comput. Chem.* **1993**, *14*, 138−154.

(25) Gasteiger, J.; Marsili, M. Iterative partial equalization of orbital electronegativity—A rapid access to atomic charges. *Tetrahedron* **1980**, *36*, 3219−3228.

(26) Gasteiger, J.; Saller, H. Calculation of the charge distribution in conjugated systems by a quantification of the resonance concept. *Angew. Chem., Int. Ed. Engl.* **1985**, *24*, 687−689.

(27) Bauknecht, H.; Zell, A.; Bayer, H.; Levi, P.; Wagener, M.; Sadowski, J.; Gasteiger, J. Locating biologically active compounds in medium sized heterogeneous datasets by topological autocorrelation vectors. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1205−1213. AUTOCORR is available from Molecular Networks GmbH, Naegelsbachstr. 25, 91052 Erlangen, Germany (http://www.mol-net.de).

(28) *SONNIA*; Molecular Networks GmbH: Naegelsbachstr. 25, 91052 Erlangen, Germany (http://www.mol-net.de).

(29) Kohonen, T. Self-organizing formation of topologically correct feature maps. *Biol. Cyb.* **1982**, *43*, 59−69.

(30) Teckentrup, A.; Briem, H.; Gasteiger, J. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 626−634.

(31) *UNITY* and *SELECTOR*; Tripos GmbH: Martin-Kollar Strasse 13, D-81829 Muenchen, Germany (http://www.tripos.com).

(32) Moriguchi, I.; Hirono, S.; Liu, Q.; Nakagome, I.; Matsushita, Y. Simple method of calculating octanol/water partition coefficient. *Chem. Pharm. Bull.* **1992**, *40*, 127−130.

(33) Ecker, G. F.; Csaszar, E.; Kopp, S.; Plagens, B.; Holzer, W.; Ernst, W.; Chiba, P. Identification of ligand-binding regions of P-glycoprotein by activated-pharmacophore photoaffinity labeling and matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry. *Mol. Pharmacol.* **2002**, *61*, 637−648.

JM060604Z