Monatsh Chem 139, 401–405 (2008)
DOI 10.1007/s00706-007-0819-7
Printed in The Netherlands

Monatshefte für Chemie
Chemical Monthly

Hydrophobic moments as physicochemical descriptors in structure-activity relationship studies of P-glycoprotein inhibitors

Gerhard König[1], Peter Chiba[2], Gerhard F. Ecker[1]

[1] Emerging Field Pharmacoinformatics, Department of Medicinal Chemistry, University of Vienna, Wien, Austria
[2] Institute of Medical Chemistry, Medical University of Vienna, Wien, Austria

Received 5 July 2007; Accepted 30 September 2007; Published online 14 March 2008

Abstract Lipophilicity is one of the major determining physicochemical descriptors for P-glycoprotein (*P-gp*) inhibitory activity. In order to consider lipophilicity as a space directed property, we apply the concept of hydrophobic moments on a set of propafenone-type inhibitors of P-glycoprotein and use them as descriptors in QSAR analyses. While the 0^{th} moment is the sum of the atomic hydrophobicity coefficients, which is a measure for the total hydrophobicity of the molecule, the 1^{st} moment (or hydrophobic dipole) is a measure for the asymmetry of the distribution of hydrophobicities and therefore is analogous to the electrostatic dipole. The use of these hydrophobic dipole moments as independent variables remarkably improved the predictive power of QSAR models obtained.

Keywords Computer chemistry; Propafenone; Multidrug resistance; Molecular modeling; Antitumor agents.

Introduction

The exact characterization of hydrophobicity and hydrophobic distribution patterns is still one of the major challenges in modeling and computational drug design. Though a plethora of descriptors are available for the treatment of solvation free energies and distribution coefficients between polar and apolar phases, most of these descriptors scarcely integrate data about the spatial arrangement of hydrophobic groups. CoMFA [1] and CoMSIA [2] very well depict hydrophobic fields, but both methods require a thorough conformational analysis and alignment of the ligands.

In protein modeling, hydrophobic moments have been employed as a measure for amphiphilicity of helices and the hydrophobicity distribution within proteins [3]. Since hydrophobic interactions (and amphiphilicity) are driving forces in ligand receptor interaction, it seems more than justified to extend the application of hydrophobic moments to QSAR studies. Recently, *Pajeva* and *Wiese* have shown that for a series of inhibitors of the multidrug efflux pump P-glycoprotein hydrophobicity represents a space directed molecular property rather than a simple overall descriptor [4]. Based on these findings they also introduced the concept of hydrophobic dipoles and concluded that such descriptors are able to characterize the space directionality of hydrophobicity in the case of multidrug-resistance related drugs and thus provide means to design hydrophobic complementarity [5]. Within this paper we aim at extending this concept of hydrophobic moments to a quantitative level and to explore their potential as descriptors in QSAR studies of propafenone-type inhibitors of P-glycoprotein.

Correspondence: Gerhard F. Ecker, Emerging Field Pharmacoinformatics, Department of Medicinal Chemistry, University of Vienna, Althanstrasse 14, 1090 Wien, Austria. E-mail: gerhard.f.ecker@univie.ac.at

Results and discussion

Recently, *Pajeva* and *Wiese* have shown that in case of *P-gp* interaction lipophilicity may be regarded as space directed property [4]. For a series of propafenone-type P-glycoprotein inhibitors we could demonstrate that with increasing lipophilicity of the substituents on the amine moiety also the statistical significance of the indicator variables denoting the substitution pattern on the central aromatic ring system increases [18]. This indicated that the distribution of hydrophobicity within the molecules influences the mode of interaction with *P-gp*. To further explore this hypothesis we explored the applicability of hydrophobic moments for use as descriptors in multiple linear regression analysis.

Hydrophobic moments

Every distribution of charges can be represented by a multipole expansion. The zeroth and first order terms in the expansion are called the monopole and dipole moments, while the higher terms are denoted as quadrupole, octupole, *etc*. In analogy to electronic charges, dipoles can also be applied to hydrophobicity (however, not in the same classical context). This approach has already been utilized for the characterization of helical amphiphilicity [6–8].

For this purpose, the total hydration has to be divided into contributions from individual atoms. However, since no generally accepted measure for hydrophobicity is available, the choice of an adequate numerical representation of hydrophobicity still poses a problem. Usually, such values are derived from distribution coefficients and/or the free energy of transfer between water and an apolar phase [9]. This choice of the apolar reference is of significance for the later use, since the apolar phase should approximately correspond to the environment of the binding pocket of the ligands. Additionally, the application of hydrophobicity measures to medicinal chemistry demands the treatment of an extensive chemical space (*e.g*., parameters for halogens), which holds true for only a small minority of hydrophobicity scales.

Once this conundrum is solved, it is possible to apply the equations for the calculation of electrostatic multipoles to hydrophobicity. According to Eq. (1), the monopole m_0 can be calculated as a sum over all atoms i of the respective atomic contributions to the hydration free energy (a_i):

$$m_0 = \sum_i a_i \tag{1}$$

The same holds true for the dipole, which characterizes the amphiphilicity and separation of the different contributions to the hydration free energy (Eq. (2)).

$$\gamma_H = \sum_i \vec{r}_i a_i \tag{2}$$

However, since the monopole m_0 in practically all (bio-) molecular systems is not equal to zero, the first-order moment m_1 will depend on the origin of the coordinate system. Therefore an adaptation of the equation is necessary, as it has been proposed by *Eisenberg et al*. [6] (Eq. (3)):

$$\gamma_H = \sum_i \vec{r}_i a_i - \bar{a} \sum_i \vec{r}_i \quad \bar{a} = \left(\frac{\sum_i a_i}{n} \right) \tag{3}$$

In this equation $\bar{a}$ is the mean hydrophobicity of the molecule and n denotes the number of atoms in the molecule. With the use of hydrophobicity differences about the mean hydrophobicity of the molecule, the origin of the system can be assigned arbitrarily. However, this necessary adaptation may lead to a ''loss'' of information, since the net hydrophobicity and the position of the centroid are subtracted from the absolute of the hydrophobic moment (Eq. (4)):

$$\bar{a} \sum_i \vec{r}_i = \sum_i \bar{a} \vec{r}_i = \sum_i \frac{\bar{a} n \vec{r}_i}{n} = m_0 \left(\frac{\sum_i \vec{r}}{n} \right) \tag{4}$$

Therefore it may be necessary to compensate this loss of information by the additional use of the monopole (*i.e*., the solvation free energy or $\log P$). For the comparison of molecules with very different numbers of atoms it may also become practical to use the intensive hydrophobic moment, as outlined in Eq. (5). Here, the hydrophobic moment is divided by the number of atoms (n_i) and thus it becomes independent of the size of the molecule.

$$\gamma_H^{\text{int.}} = \frac{\sum_i \vec{r}_i [a_i - \bar{a}]}{n_i} \tag{5}$$

Quantitative structure activity relationship studies

In order to assess the performance of hydrophobic moments as descriptors in quantitative structure

Table 1 Chemical structures, biological activity, and hydrophobic moments for the data set used. The field R marks the hydrophobic region in the vicinity of the nitrogen atom: A denotes morpholine, B piperidine, C *p-F*-phenylpiperazine and D *o*-tolylpiperazine. R^1 gives the substituent according to the general formula given in Fig. 1

Code	*R*	R^1	$-\log EC_{50}$	$\log P$	$\lvert\gamma\rvert$	$\lvert\gamma_{\text{int.}}\rvert$
GPV0005	B	*o*-CO$(CH_2)_2$*Ph*	0.2227	3.83	246.3525	4.3992
GPV0012	B	*o*-COCH_2CH_3	−1.1555	2.47	66.2065	1.4393
GPV0017	B	*o*-COCH_3	−1.5046	1.99	52.4596	1.2200
GPV0027	D	*o*-CO$(CH_2)_2$*Ph*	1.5735	4.33	144.8716	2.1305
GPV0031	C	*o*-CO$(CH_2)_2$*Ph*	1.1549	4.19	108.9920	1.6768
GPV0045	C	*o*-COCH_3	−0.3195	2.35	70.8079	1.3617
GPV0046	A	*o*-COCH_2CH_3	−2.3164	1.06	7.5651	0.1719
GPV0049	A	*o*-COCH_3	−1.8281	0.58	19.8480	0.4841
GPV0057	A	*o*-CO$(CH_2)_2$*Ph*	−0.5617	2.42	83.3588	1.5437
GPV0073	B	*p*-CO$(CH_2)_2$*Ph*	0.0123	3.83	477.6084	8.5287
GPV0134	C	*p*-CO$(CH_2)_2$*Ph*	−0.4040	4.19	289.6450	4.4561
GPV0135	B	*m*-CO$(CH_2)_2$*Ph*	0.3726	3.87	422.0438	7.5365
GPV0149	A	*p*-CO$(CH_2)_2$*Ph*	−0.8373	2.42	158.9834	2.9441
GPV0157	C	*m*-CO$(CH_2)_2$*Ph*	−0.0490	4.23	252.4391	3.8837
GPV0180	B	*o*-CO$(CH_2)_2$Napht.	0.7645	5.05	459.7770	7.4158
GPV0319	C	*o*-CO*Ph*	0.8327	4.01	63.2013	1.0712
GPV0374	A	*o*-CO$(CH_2)_2$Napht.	0.1367	3.64	217.5738	3.6262
GPV0384	A	*m*-COCH_3	−2.1087	0.62	21.0075	0.5124
GPV0385	B	*m*-COCH_3	−0.9578	2.03	52.5307	1.2216
GPV0386	C	*m*-COCH_3	−1.0031	2.39	65.9897	1.2690
GPV0389	B	*p*-COCH_3	−1.6900	1.99	79.1752	1.8413
GPV0390	C	*p*-COCH_3	−1.0750	2.35	57.9224	1.1139
GPV0391	A	*p*-COCH_3	−2.4801	0.59	32.7717	0.7993
GPV0577	D	*p*-COCH_3	−0.1903	2.49	82.8968	1.5072
GPV0596	D	*m*-COCH_3	0.2700	2.53	90.0581	1.6374
GPV0626	D	*o*-COCH_3	0.6038	2.49	65.1960	1.1854
GPV0647	D	*m*-CO$(CH_2)_2$*Ph*	1.1308	4.37	305.4490	4.4919
GPV0653	A	*m*-CO$(CH_2)_2$*Ph*	−0.5410	2.46	153.4707	2.8421
GPV0794	C	*o*-COCH_2CH_3	0.0778	2.82	56.7319	1.0315
GPV0797	D	*o*-COCH_2CH_3	0.7747	2.97	115.6657	1.9942
GPV0863	D	*p*-COCH_2CH_3	0.0640	2.97	141.7554	2.4441
GPV0897	B	*o*-CO*Ph*	−1.1262	2.25	12.7978	0.2666

activity studies, we performed multiple linear regression analyze using different combinations of hydrophobicity related descriptors. As shown in Table 1, the driving force of the ligand-receptor interaction is the hydrophobicity ($\log P(o/w)$) itself. This outcome is consistent with previous studies and the promiscuity of the target. In addition to the hydrophobic interactions, the proposed descriptors were able to significantly improve the predictive power of the QSAR models, which indicates certain specificity in the binding mode. While the value of the hydrophobic moment ($|\gamma|$) alone was sufficient to improve the $\log P$-based model, the addition of the spatial components leads to a finer definition of the hydrophobic asymmetry and thus seemingly to a better quantitative description of ligand protein interaction. Here, the contribution to the hydrophobic moment on the x axis is most significant for the model. The other spatial dimensions play only a minor role. This indicates the impact of the hydrophobicity difference between residues R and R^1. Generally, a higher hydrophobicity on residue R will lead to a higher activity.

Furthermore, the use of intensive hydrophobic moments ($|\gamma$ int.$|$) leads to a slight improvement of

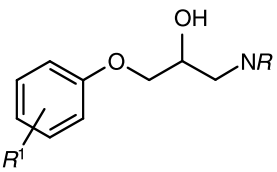

Fig. 1 General formula for P-glycoprotein inhibitors

the model, as compared with the normal hydrophobic moment. Since the differences in atom numbers are negligible, this result can probably be attributed to the fact that the x component contains most of the information about the length of the hydrophobic moment and thus the intensive hydrophobic moment provides a more orthogonal dimension.

In order to compare these results to with those obtained by using standard ADME-type descriptors, a set of 11 descriptors (Number of atoms, $\log P(o/w)$, molecular refractivity, number of rotatable single bonds, number of hydrogen bond acceptor atoms, number of hydrogen bond donor atoms, sum of atomic polarizabilities, sum of *VDW* surface areas of pure hydrogen bond acceptors, sum of *VDW* surface areas of pure hydrogen bond donors, total hydrophobic *van der Waals* surface area, polar surface area) were applied on the same dataset. The final model showed an R^2 value of 0.86 and a predictive power Q^2 of 0.74, which is slightly lower than the values obtained with the hydrophobic moments. However, it shall be noted that though both results rely on the conformation, only the analysis of the hydrophobic moments is dependent on a correct structural superposition and thus requires an additional effort.

To emphasize the dependency on the superposition, five QSAR analyses with randomly oriented ligands were conducted, using the $\log P(o/w)$, $|\gamma$ int.$|$, x, y, and z component descriptors. This approach yielded a R^2 of 0.78 ± 0.01 and a Q^2 of 0.67 ± 0.05, which is hardly an improvement over the result of the $\log P + |\gamma|$ model alone. Since $|\gamma|$ is independent of the orientation, this further strengthens the importance of a proper alignment when using hydrophobic moments as independent variables in QSAR studies.

In conclusion, we could demonstrate that for a set of P-glycoprotein inhibitors with systematically varied hydrophobicity distribution pattern use of the hydrophobic moment as an independent variable yields to models with higher predictive power. This strengthens recent findings that for ligands of P-glycoprotein hydrophobicity has to be considered as a space directed property rather than an overall physicochemical property. The use of hydrophobic moments seems especially useful where amphiphilicity and differences in the distribution of hydrophobic and hydrophilic groups are involved in ligand binding and recognition. This should be especially interesting in the case of interactions which take place at polar/apolar interfaces. Furthermore, hydrophobic dipoles may also serve as a first rough alignment rule for 3D-shape similarity calculations pursued under the framework of our recently introduced SIBAR approach [10].

Materials and methods

Compounds

In total thirty-two compounds with varying hydrophobicity distributions were used in the present study (see Table 2). Compounds were prepared as described previously and follow the general structure given in Fig. 1. The compound design was based on the definition of two hydrophobic areas, one in the vicinity of the nitrogen atom (R) and one on the central aromatic ring (R^1). Variation was achieved by using morpholine, piperidine, *p*-F-phenylpiperazine and *o*-tolylpiperazine on the amine side (R) and miscellaneous acyl groups on the central aromatic core (see Table 2). The corresponding 3D structures were generated using CORINA [11] and aligned based on the aryloxypropanol-substructure.

Calculation of hydrophobic moments

Atomic solvation contributions can be calculated with the parameters of *Eisenberg et al.* [12], *Wang et al.* [13], *Wildman* and *Crippen* [14], and MOE $\log P(o/w)$ [15]. Also potential surface input data are in principle supported. However, for the presented application only the MOE $\log P(o/w)$ fragment constants were selected to calculate the hydrophobic moments. While contributions below the mean hydrophobicity were interpreted as hydrophile, positive values were marked as hydrophobic. The descriptor $\log P$ is given by the sum of all atomic contributions. $|\gamma|$ gives the Euclidian norm of the first hydrophobic moment, as it has been defined in theory. x, y, and z denote the separate contributions to the hydrophobic moment in the directions of the x, y, and z axis. Results for $\log P$, $|\gamma|$ and $|\gamma$ int.$|$ see Table 2.

Table 2 Statistic results of models with different hydrophobic moment descriptors

	R^2	Q^2
$\log P$	0.74	0.70
$\log P + \|\gamma\|$	0.78	0.74
$\log P + \|\gamma$ int.$\|$	0.77	0.74
$\log P + x$	0.82	0.78
$\log P + y$	0.74	0.67
$\log P + z$	0.74	0.68
$\log P + \|\gamma\| + x$	0.83	0.78
$\log P + \|\gamma\| + x + y + z$	0.87	0.79
$\log P + \|\gamma$ int.$\| + x + y + z$	0.89	0.82

Hydrophobic moment descriptors have been implemented in an SVL-extension to the MOE chemical computing environment, called HYDRAA (HYDRation Analysis and Alignment). The code for this program will be provided on the SVL-exchange homepage [16].

QSAR studies

Linear structure-activity relationships were derived by regression analyses (partial least squares analysis), which were carried out within the QSAR-suite of MOE and in the R software environment for statistical computing. MOE was also used for leave-one-out cross-validations.

Pharmacological testing

The pharmacological activity of the compounds was measured in a zero trans efflux protocol using daunorubicine as the fluorochrome [17]. Briefly, multidrug resistant CCRF-CEM vcr 1000 cells were incubated with daunorubicine and the decrease in mean cellular fluorescence in dependence of time was measured in presence of various concentrations of the modulator. EC_{50} values were calculated from the concentration-response curve of efflux V_{max}/K_m *vs* concentration of the modulator. Thus, the effect of different modulators on the transport rate is measured in a direct functional assay. Values are given in Table 1 and are the mean of at least three independently performed experiments. Generally, interexperimental variation was below 20%.

Acknowledgements

We are grateful to the Austrian Science Fund for financial support (grant no. L344-N17).

References

1. Cramer RD, Patterson DE, Bunce JD (1988) J Am Chem Soc 110:5959
2. Klebe G, Abraham U, Mietzner T (1994) J Med Chem 37:4130
3. Silverman BD (2003) Protein-Struct Funct Genet 53:880
4. Pajeva I, Wiese MJ (1998) Med Chem 41:1815
5. Pajeva I, Wiese M (2001) Compt Rend Acad Bulg Sci 54:81
6. Eisenberg D, Weiss RM, Terwilliger TC (1982) Nature 299:371
7. Eisenberg D, Weiss RM, Terwilliger TC, Wilcox W (1982) Faraday Symp Chem S 17:109
8. Eisenberg D, Wesson M, Yamashita M (1989) Chem Scripta 29A:217
9. Biswas KM, Devido DR, Dorsey JG (2003) J Chromatogr A 1000:637
10. Zdrazil B, Kaiser D, Kopp S, Chiba P, Ecker GF (2007) QSAR Combinat Chem 26:669
11. Gasteiger J, Hiller C, Rudolph C, Sadowski J (1991) Abstr Pap Am Chem Soc 202:36
12. Eisenberg D, Yamashita M, Wilcox W, Talafous J, Wesson M (1987) Biophys J 51:a22
13. Wang JM, Wang W, Huo SH, Lee M, Kollman PA (2001) J Phys Chem B 105:5055
14. Wildman SA, Crippen GM (1999) J Chem Inf Comput Sci 39:868
15. Labute P, MOE LogP[Octanol/Water], User Manual
16. http://svl.chemcomp.com/
17. Chiba P, Ecker GF, Schmid D, Drach P, Tell B, Goldenberg S, Gekeler V (1996) Mol Pharmacol 49:1122
18. Pleban K, Hoffer C, Kopp S, Peer M, Chiba P, Ecker GF (2004) Arch Pharm 337:328