RESEARCH PAPER

Accepted 11 June 2001
Revised 27 April 2001
Received 30 October 1999
Indian J. Pharm. Sci., 2001, 63(4) 312-316

Quantitative Relationships Between Topological Indices and Binding Affinities of Antipsychotic Drugs to Alpha-1-Acid Glycoprotein

NEELAM SEEDHER*
Department of Chemistry
Panjab University, Chandigarh-160014

Quantitative structure protein binding relationships (QSPBR) have been derived for the binding of nine antipsychotic phenothiazine derivatives to human alpha-1-acid glycoprotein (AGP). Ninety eight topological indices for the phenothiazine derivatives were calculated using software Polly 2.3. Bivariate correlation matrix showed that most of the statistically significant correlations were obtained for topochemical indices. Four variables, complementary information contents, CIC_0, CIC_1 and CIC_2 and chain connectivity index, $^6\chi_{ch}$, with correlations significant up to 0.01 level were considered to be the predictors of the model. Linear, multiple linear and non-linear regression analysis was performed on the retained variables. The complementary information content for 2nd order neighbourhood of vertices in a hydrogen filled graph, CIC_2, was found to be the best single parameter that can be used in linear or quadratic fit to predict the binding affinities of phenothiazines to AGP.

Quantitative structure-activity relationships (QSARs) have come into widespread use for the prediction of various molecular properties and biological responses[1-4]. The nature and magnitude of drug-protein interaction significantly influences the biological activity of a drug, since most of the receptors for drugs are proteinacious in nature. For example, the administered drugs are extensively and reversibly bound to plasma proteins and drug is transported mainly as a complex with protein. The nature of binding forces and the degree of protein binding has important pharmacokinetic and pharmacodynamic implications, especially when the drugs are highly bound and their apparent volume of distribution is small[5-9]. Since both specific and non-specific drug binding depends on the structure of the drug molecules, quantitative structure-protein binding relationships (QSPBR) can provide a useful method of predicting drug binding affinities and understanding the nature of drug-protein interaction from structural parameters of the drug. Most such studies reported in the literature pertain to correlation of hydrophobicity of drugs to association constants[10-12]. Correlation of topological indices to binding parameters has not received adequate attention so far.

Human serum albumin (HSA), alpha-1-acid glycoprotein (AGP) and lipoproteins (LIPO) are the most important plasma proteins responsible for the binding of drugs in plasma. From the binding studies it follows that HSA accounts mainly for the binding of acidic and neutral drugs, whereas AGP and LIPO associate more readily with basic drugs[13,14]. Due to its greater drug affinity, AGP can be the most important determinant in the plasma binding of antipsychotic phenothiazine derivatives. Present paper reports quantitative relationships between various topological indices and binding affinities of nine antipsychotic phenothiazine derivatives to AGP.

MATERIALS AND METHODS

Database:

Binding parameters (association constants, K and the number of binding sites, n) for the binding of nine

*For correspondence

phenothiazine derivatives to human alfa-1-acid glycoprotein (AGP) were taken from Miyoshi *et al.*[15]. Binding affinities, expressed as log nK, are given in Table 3.

Computation of topological indices:

The ninety eight topological indices (TIs) for nine phenothiazine derivatives used in this study were calculated using Polly 2.3 which uses SMILES notation input of chemical structures[16]. TIs calculated by POLLY 2.3 include the Weiner index, Connectivity indices and information theoretic indices defined on distance matrices of graphs as well as a set of parameters derived on the neighbourhood complexity of vertices in hydrogen-filled

TABLE 1: SYMBOLS AND DEFINITIONS OF TOPOLOGICAL INDICES

Symbol	Definition
I_D^W	Information index for magnitudes of distances between all possible pairs of vertices of a graph
I_D^W	Mean information index for the magnitude of distance
W	Weiner index=half-sum of off-diagonal elements of distance matrix of a graph
I^D	Degree complexity
H^v	Graph vertex complexity
H^D	Graph distance complexity
IC	Information content of distance matrix partitioned by frequency of occurrences of distance h
O	Order of neighbourhood when IC_r reaches its maximum values for the hydrogen-filled graph
I_{ORB}	Information content or complexity of hydrogen-suppressed graph at its maximum neighbourhood of vertices
O_{ORB}	Order of neighbourhood when IC_r reaches its maximum value for the hydrogen-suppressed graph
M_1	A zagreb group parameter=sum of square of degree over all vertices
M_2	A zagreb group parameter=sum of cross-product of degrees over all neighbouring (connected) vertices
IC_r	Mean information index or complexity of a graph based on the r^{th} (r=0-6) order neighbourhood of vertices in a hydrogen-filled graph
SIC_r	Structural information content for r^{th} (r=0-6) order neighbourhood of vertices in a hydrogen-filled graph
CIC_r	Complementary information content for r^{th} (r=0-6) order neighbourhood of vertices in a hydrogen-filled graph
$h\chi$	Path connectivity index of order h=0-6
$h\chi_c$	Cluster connectivity index of order h=3-6
$h\chi_{ch}$	Chain connectivity index of order h=3-6
$h\chi_{pc}$	Path cluster connectivity index of order h=4-6
P_h	Number of paths of length h=0-10
$^h\chi^b$	Bond path connectivity index of order h=0-6
$^h\chi c^b$	Bond cluster connectivity index of order h=3-6
$^h\chi ch^6_{ch}$	Bond chain connectivity index of order h=3-6
$^h\chi^h_{pc}$	Bond path-cluster connectivity index of order h=4-6
$^h\chi^v$	Valence path connectivity index of order h=0-6
$^h\chi c^v$	Valence cluster connectivity index of order h=3-6
$^h\chi^v_{ch}$	Valence chain connectivity index of order h=3-6
$^h\chi^v_{pc}$	Valence path-cluster connectivity index of order h=4-6

molecular graphs. The methods used for the calculation of TIs are described[17]. Table 1 provides a comprehensive list and brief descriptions for these indices.

Data reduction:

Initially all the TIs were transformed by the natural logarithm of the index plus one. This is routinely done to scale the indices since there may be a difference of several orders of magnitude between indices and some may equal zero. From the original set of 98 indices, it was necessary to remove some indices. Nine indices were completely redundant because they had values of zero for all compounds.

On the remaining indices common data reduction techniques such as Variable clustering and Principal component analysis could not be used since the sample size is very small. That is, the number of descriptors is much larger than the number of samples and thus the probability of chance correlations is very high[18]. Bivariate correlation matrix between binding affinities (log nK) and topological indices was obtained using statistical software SPSS for windows (SPSS Inc., Chicago, IL). Only those variables having statistically significant correlations were retained.

Statistical analysis:

Linear regression analysis, multiple linear regression analysis and non-linear regression analysis was performed on the retained variables. The following parameters were determined: the correlation coefficient R, coefficient of determination R^2, the standard error, SE and the significance of the regression model, F.

RESULTS AND DISCUSSION

Topological indices are numerical graph invariants that quantify certain aspects of molecular structure. TIs are sensitive to such structural features as size, shape, bond order, branching and neighbourhood patterns of atoms in molecules.

Bivariate correlation matrix between binding affinities (log nK) and 89 topological indices showed four topological indices, complementary information contents (CIC_0, CIC_1 and CIC_2) and chain connectivity index ($^6\chi_{ch}$) with correlation coefficients 0.866, 0.816 and 0.895 and 0.807, significant up to 0.01 level and five topological indices, bond path connectivity index ($^6\chi^b$), bond chain connectivity index, ($^6\chi_{ch}{}^b$), valence path connectivity index ($^3\chi^v$), valence chain connectivity index ($^6\chi_{ch}{}^v$) and valence path-cluster connectivity index ($^4\chi_{pc}{}^v$) with correlation coefficients 0.683, 0.757, 0.700, 0.770, 0.710 significant up to 0.05 level.

Topological indices used in the present study are usually partitioned into two distinct sets: topostructural indices and topochemical indices[19,20]. Topostructural indices encode information about the adjacency and distances of atoms in molecular structures irrespective of the chemical nature of the atoms involved in the bonding or factors like hybridization states of atoms and number of valence electrons in individual atoms. Topochemical indices are the parameters that quantify information regarding the topology as well as specific chemical properties of the atoms comprising a molecule. It is seen that out of 9 TIs with statistically significant correlations, 8 TIs belong to the set of topochemical indices and only 1 TI, chain connectivity index, belong to the set of topostructural indices. Thus chemical nature of the substituents on the phenothiazine ring plays a major role in determining the binding affinity of these drugs to AGP.

Four TIs with correlation coefficients significant up to 0.01 level (complementary information contents, CIC_0, CIC_1 and CIC_2 and chain connectivity index, $^6\chi_{ch}$) were considered to be predictors of the model. Values of four retained TIs for nine phenothiazine derivatives are given in Table 2. Chain connectivity index ($^6\chi_{ch}$) gives information about phenothiazine ring substitution. Complimentary information content, CIC, obtained from hydrogen filled molecular graph, reflects mainly the steric properties of the drug molecule.

Linear regression analysis performed on the four retained variables gave the following relationships (Eqs. 1-4) between binding affinities (log nK) and the topological indices.

$$\log nk = -9.446\ {}^6\chi_{ch} + 7.109 \quad (1)$$

$R = 0.807$, $R^2 = 0.651$, SE = 0.108, F = 13.054

$$\log nK = -10.308\ CIC_0 + 13.473 \quad (2)$$

$R = 0.866$, $R^2 = 0.750$, SE = 0.092, F = 21.039

$$\log nK = -6.891\ CIC_1 + 10.215 \quad (3)$$

$R = 0.816$, $R^2 = 0.666$, SE = 0.106, F = 13.974

$$\log nK = -5.090\ CIC_2 + 8.251 \quad (4)$$

$R = 0.895$, $R^2 = 0.802$, SE = 0.082, F = 28.279

The coefficient of determination, R^2 gives the fraction of variance explained by the equations. It is seen that the complementary information content, CIC_2 is the only

TABLE 2: VALUES OF FOUR RETAINED TOPOLOGICAL INDICES

Drug	Topological index*			
	$^6\chi_{ch}$	CIC_0	CIC_1	CIC_2
Promazine	0.0835	0.6984	0.5843	0.4064
Chlorpromazine	0.0781	0.6859	0.5556	0.3518
Perazine	0.1123	0.7247	0.6160	0.4476
Prochlorperazine	0.1072	0.7147	0.5941	0.4084
Trifluoperazine	0.1072	0.7116	0.5858	0.4087
Perphenazine	0.1072	0.7169	0.5906	0.4170
Fluphenazine	0.1072	0.7140	0.5830	0.4168
Triflupromazine	0.0781	0.6834	0.5472	0.3562
Promethazine	0.0835	0.6984	0.5843	0.4094

*All the reported values are logarithm of the index plus one. TIs having statistically significant correlations with association constants were retained.

single parameter which can explain more than 80% of the variance in the data. The only topostructural parameter retained in the model, chain connectivity index $^6\chi_{ch}$) could explain only 65.1% of the variance. However, on inclusion of the topochemical parameter, CIC_2 in multiple linear regression analysis (Eq. 5), the explained variance increase to 82.8%.

$$\log nk = -3.122\ ^6\chi_{ch} + 3.895\ CIC_2 + 8.069 \quad (5)$$

$R = 0.910$, $R^2=0.828$, $SE=0.082$, $F=14.489$

Since topostructural parameters do not contain any explicit chemical information about the molecule, it may again be concluded that the chemical nature of the substituents on the phenothiazine ring is important in determining protein binding affinity of these drugs.

Non-liner regression analysis was performed by fitting quadratic equations to the data. Quadratic equations (6-8) gave a better fit of the data for three of the retained indices, complementary information contents (CIC_1 and CIC_2) and chain connectivity index, $^6\chi_{ch}$. Complementary

TABLE 3: COMPARISON OF EXPERIMENTALLY DETERMINED AND CALCULATED BINDING AFFINITIES

Drug	Binding Affinity (log nK)			
	Experimental value*	Calculated from		
		Eq. 4	Eq. 5	Eq. 7
Promazine	6.2788	6.1824	6.2254	6.1490
Chlorpromazine	6.5453	6.4603	6.4549	6.4949
Perazine	6.0682	5.9727	5.9750	6.0566
Prochlorperazine	6.0828	6.1722	6.1436	6.1412
Trifluoperazine	6.1583	6.1707	6.1424	6.1400
Perphenazine	6.1335	6.1285	6.1101	6.1114
Fluphenazine	6.0334	6.1295	6.1109	6.1120
Triflupromazine	6.3979	6.4379	6.4378	6.4576
Promethazine	6.1206	6.1671	6.2137	6.1374

*Values taken from Miyoshi *et al*[15]

information content, CIC_2 could explain 85.7% of the variance in the data.

$$\log nK=0.008\ (CIC_1)^2-101.440\ CIC_1+37.550 \quad (6)$$

$R = 0.850$, $R^2 = 0.722$, SE = 0.105, F = 7.799

$$\log nK=0.004\ (CIC_2)^2-38.709\ (CIC_2)+14.828 \quad (7)$$

$R = 0.926$, $R^2 = 0.857$, SE = 0.075, F = 17.978

$$\log nK=0.008\ (^6\chi_{ch})^2 - 166.920\ (^6\chi_{ch}) +14.381 \quad (8)$$

$R = 0.879$, $R^2 = 0.772$, SE = 0.095, F = 10.184

The complementary information content for 2nd order neighbourhood of vertices in a hydrogen filled graph, CIC_2, is found to be the best single parameter that can be used in linear or quadratic fit to predict the binding affinities of phenothiazines to AGP. Binding affinities can also be predicted by combining chain connectivity index ($^6\chi_{ch}$) with the complementary information content, CIC_2 in multiple linear regression analysis. Experimentally determined binding affinities and those predicted from theoretical molecular descriptors using equations 4, 5 and 7 and given in Table 3. The agreement is fairly good. However, the results can not be generalized since correlation is derived from a small set of compounds.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the contribution of Dr. S.C. Basak, Natural Resources Research Institute, University of Minnesota, 5013 Miller Trunk Highway, Duluth MN55811, USA, for the use of POLLY 2.3 software in the calculation of topological indices.

REFERENCES

1. Basak, S.C. and Gute, B.D., **SAR and QSAR in Environ. Res.**, 1997, 7, 1.
2. Basak, S.C. and Grunwald, G.D., **J. Chem. Inform. Comput. Sci.**, 1995, 35, 366.
3. Kier, L.B. and Hall, L.H., In; Molecular connectivity in structure-activity analysis. Research Studies Press, Letchworth, Herfordshire, England, 1986, 1.
4. Saiakhov, R.D., Stefan, L.R. and Klopman, G., **Perspective Drug Discovery Design**, 2000, 19, 133.
5. Vallner, J.J., **J. Pharm. Sci.**, 1977, 66, 447.
6. Deleve, L.D. and Piafsky, K.M., **Trends Pharmacol. Sci.**, 1981, 2, 283.
7. Lin, J.H., Cocchetto, D.M. and Duggan, D.E., **Clin. Pharmacokinet.**, 1987, 12, 420.
8. Zlotos, G., Bucker, A., Kinzig-Schippers, M., Sorgel, F. and Holzgrabe, U., **J. Pharm. Sci.**, 1998, 87, 215.
9. Seedher, N., Singh, B. and Singh, P., **Indian J. Pharm. Sci.**, 1999, 61, 143.
10. Deschamps-Labat, L., Pehourcq, F., Jagou, M. and Bannwarth, B., **J. Pharm. Biomed. Anal.**, 1997, 16, 223.
11. Laznicek, M. and Laznickova, A. **J. Pharm. Biomed. Anal.**, 1995, 13, 823.
12. Pehourcq, F.D., Radouane, A., Labat, L. and Bannwarth, B., **Pharm. Res.**, 1995, 12, 1535.
13. Morin, D., Zini, R., Ledewyn, S., Colona, J.P., Czajka, M. and Tillement, J.P., **J. Pharm. Sci.**, 1985, 74, 727.
14. Kremer, J.M.H., Wilting, J. and Janssen, L.H.M., **Pharmacol. Rev.**, 1988, 40, 1.
15. Miyoshi, T., Sukimoto, K. and Otagiri, M., **J. Pharm. Pharmacol.**, 1992, 44, 28.
16. Basak, S.C., Harriss, D.K. and Magnuson, V.R., POLLY: copyright University of Minnesota, Duluth, 1988.
17. Balasubramanian, K. and Basak, S.C., **J. Chem. Inform. Comput. Sci.**, 1998, 38, 367.
18. Topliss, J.G. and Edwards, R.P., **J. Med. Chem.**, 1979, 22, 1238.
19. Gute, B.D. and Basak, S.C., **Environ. Res.**, 1997, 7, 117.
20. Basak, S.C., Gute, B.D. and Grunwald, G.D., **J. Chem. Inform. Comput. Sci.**, 1997, 37, 651.