J. Pharm. Pharmacol. 1996, 48: 240–244
Received November 25, 1994
Accepted September 19, 1995

Correlation of Pharmacological Properties of a Group of Hypolipaemic Drugs by Molecular Topology

R. A. CERCÓS-DEL-POZO, F. PÉREZ-GIMÉNEZ*, J. GÁLVEZ-ALVAREZ, M. T. SALABERT-SALVADOR, F. J. GARCIA-MARCH AND G. M. ANTÓN-FOS

Unit of Drug-design and Molecular Connectivity Investigation, Department of Physical Chemistry, Faculty of Pharmacy, University of Valencia, Valencia, Spain

Abstract

This investigation was undertaken to test the ability of the molecular connectivity model to predict the percentage of plasma protein binding, the percentage of total cholesterol reduction and oral LD50 in rats of a group of hypolipaemic drugs using multi-variable regression equations with multiple correlation coefficients, standard error of estimate, degrees of freedom, F-Snedecor function values, Mallow's CP and Student's *t*-test as criteria of fit.

Regression analyses showed that the molecular connectivity model predicts these properties. Corresponding stability (cross validation) studies were made on the selected prediction models which confirmed their goodness of fit.

The results also demonstrated that the presence of substituents and molecular volume, determine the value of these properties in hypolipaemic drugs.

The quantitative structure-activity relationship (QSAR) method studies the relation between a determined property and the chemical structure of a series of molecules, observing the variation of this property in the analysed series, which could hence be used for the prediction of values of the property in other molecules of the same structural group.

Molecular connectivity is a topological method used to describe the structure of a molecule by means of numbers called indices ($^{m}\chi_t$), calculated from the hydrogen-suppressed graphs of the molecule being studied, which subsequently regress in relation with the experimental values of the physical, chemical or biological properties to connectivity functions (Kier & Hall 1976).

Methods

Molecular structure can be expressed topologically by the hydrogen-suppressed graph. It should be noted that information concerning the contributions of the hydrogen atoms is implicit in this graphical formulation. For alkanes there exists a direct relation between the vertex valence, δ_i, and the number of hydrogen atoms implied at vertex i, H: $\delta i = 4 - H$, where 4 may represent the valence or the number of valence electrons for the carbon atom.

The $^{m}\chi_t$ are terms defined for a subgraph of type t containing m edges, connected in the graph. Disconnected subgraphs are not considered. The order of the subgraph is defined as m. Subgraphs may be conveniently classified into four types: path (p), subgraphs whose subgraph valencies are no greater than 2; cluster (c), subgraphs whose subgraph valencies include at least one of the 3 or 4 but do not include 2; path-cluster (pc), subgraphs whose subgraph valencies must include 2 in addition to 3 or 4; and chain (ch): edge sequences containing at least one cycle.

Correspondence: F. Pérez-Giménez, Departamento de Química-Física, Facultad de Farmacia, Avenida Vicente Andrés Estelles s/n, Burjassot (46100-Valencia), Spain.

The connectivity indices ($^{m}\chi_t$) proposed by Kier & Hall (1976) and based on the index of Randic (Randic 1975) are evaluated as a sum of terms over all the distinct connected subgraphs and are defined by the general equation:

$$^{m}\chi_t = \sum_{j=1}^{n_m} {}^{m}S_j \qquad (1)$$

where m = order of a subgraph, i.e. number of edges of a subgraph; n_m = number of type t subgraphs of order m, and $^{m}S_j$ = quantity calculated for each subgraph and defined by:

$$^{m}S_j = \left[\prod_{i=1}^{m+1}(\delta_i)\right]^{-1/2} \qquad (2)$$

where j denotes the particular set of edges that constitute the subgraph.

All the distinct subgraphs and connectivity indices for isopentane are shown in Fig. 1.

We used $^{8}\chi_t$ indices in our study, ranging from 0 to 4, whose types include path, cluster and path–cluster.

The vertex valences, δ^V, of the unsaturated carbon atoms and the heteroatoms (N, S, O) can be calculated using:

$$\delta^V = Z^V - H \qquad (3)$$

where Z^V is the number of valence electrons of the atom and H is the number of hydrogen atoms attached to it. The empirically derived values for the halogens were also used (Kier & Hall 1986).

A multiple regression analysis was used to find the relationship between the pharmacological properties of hypolipaemic drugs and the connectivity indices (Kier &

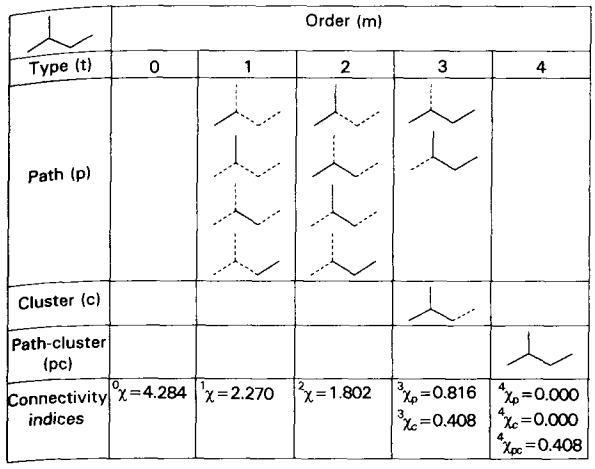


FIG 1. Subgraphs and connectivity indices of isopentane.

Hall 1976):

$$C(\chi) = P = A_0 + \sum_{i=1}^{n} A_i \cdot \chi_i \qquad (4)$$

where P is a property and A_0 and A_i represent the regression coefficients of the obtained equation. Once the connectivity function (eqn 4) is established, its value for a specific molecule may be predicted.

The connectivity indices that were used in this study (Table 1) were calculated using equations 1, 2 and 3 and computer software developed in our department (Ciudad et al 1987). The connectivity function (eqn 4) was obtained by multilinear regression with the BMDP 9R program of the biostatistic package BMDP (Biomedical Computer Programs) (Dixon 1982). To test the quality of the regression equations, the following statistical parameters were used: multiple correlation coefficient (r), standard error of estimate (s.e.), degrees of freedom (d.f.), F-Snedecor function values (F), Mallow's CP and Student's *t*-test (statistical significance), as well as the corresponding cross-validation of the selected functions.

Cross-validation for the selected functions was carried out using the jackknife method (Gray & Shucany 1972), proceeding to eliminate, by means of a random process, n observations and executing a regression program, repeating the process as many times as necessary until all the observations have been eliminated a minimum of once and a maximum of four times, finally comparing the coefficients of the independent variables calculated, the correlation coefficients, standard deviations and the residuals, with those obtained in the selected equation.

The pharmacological properties investigated were protein binding, reduction of total cholesterol and the *LD50* in rat. The experimental values for these properties were obtained from different bibliographic sources (Marshall 1982; Monk & Tood 1987; Vozeh et al 1988; Illingworth & Bacon 1989; Alberts 1990; Grundy et al. 1990; Hunninghake 1990; Pan 1991; Tilly-Kiesi & Tikkanen 1991).

Results and Discussion

The molecular connectivity indices and experimental values for pharmacological properties of nineteen hypolipaemic drugs examined in the present study are shown in Table 1.

The selected equations and statistical parameters for the compounds studied were as follows:

$$\text{Protein binding (\%)} = -(267{\cdot}24 \pm 99{\cdot}60)^0\chi/^0\chi^V - (20{\cdot}41 \pm 2{\cdot}80)^3\chi_c + (420{\cdot}45 \pm 113{\cdot}44) \qquad (5)$$

with n = 11, r = 0·939, s.e. = 11·07, d.f. = 10, F = 29·94, CP = 4 and $P < 0{\cdot}001$; the *t*-value for $^0\chi/^0\chi^V$ was 2·68 ($P < 0{\cdot}05$) and for $^3\chi_c$ was 7·28 ($P < 0{\cdot}001$).

$$\text{Reduction in total cholesterol (\%)} = (0{\cdot}27 \pm 0{\cdot}16)^0\chi^V - (28{\cdot}20 \pm 6{\cdot}80)^2\chi/^2\chi^V + (48{\cdot}11 \pm 9{\cdot}26) \qquad (6)$$

with n = 15, r = 0·829, s.e. = 3·56, d.f. = 14, F = 13·16,

Table 1. Connectivity indices and experimental values for several pharmacological properties of a group of hypolipaemic drugs used in the correlations.

Compound	$^0\chi$	$^0\chi^V$	$^2\chi$	$^2\chi^V$	$^3\chi_c$	$^3\chi_c^V$	Plasma protein binding (%)	Reduction of total cholesterol (%)	LD50 (mmol kg^{-1})
Acifran	10·085	8·636	4·850	3·679	1·031	0·627	88·00	–	13·65
Benfluorex	16·516	13·845	8·475	5·759	1·948	0·664	–	14·00	–
Bezafibrate	16·361	14·909	7·980	6·551	1·744	1·225	95·00	18·00	–
Ciprofibrate	12·508	11·765	7·328	6·718	2·500	2·392	69·50	21·00	–
Clinofibrate	22·987	20·596	11·779	9·675	2·640	1·811	–	11·00	8·54
Clofibrate	11·138	10·445	5·400	4·263	1·367	0·968	96·40	15·00	16·48
Doxazosin	20·922	18·484	10·719	7·810	1·355	0·841	98·50	8·00	
Fenofibrate	16·524	15·533	8·304	6·904	2·030	1·422	99·00	20·00	13·86
Gemfibrozil	12·767	11·617	6·301	5·410	1·517	1·373	96·00	–	19·13
Lovastatin	19·922	18·303	11·162	9·636	1·941	1·454	95·00	27·00	12·36
Niceritrol	24·532	21·621	11·378	8·771	1·540	1·207		16·00	–
Nicotinic acid	5·594	4·512	2·146	1·547	0·260	0·129		10·00	–
Pantethine	26·297	22·658	13·272	11·314	3·032	2·492	–	21·00	18·03
Plafibride	16·165	14·629	8·619	6·392	1·747	1·187	–	–	11·24
Pravastatin	21·215	18·107	10·675	8·460	1·932	1·201	55·00	16·00	–
Probucol	26·224	25·714	16·150	16·941	6·981	7·751	5·00	24·00	–
Simfibrate	20·893	19·597	10·860	8·712	2·733	1·937	–	–	16·19
Simvastatin	20·844	19·096	11·307	9·831	2·508	2·060	95·00	25·00	–
Tiadenol	13·314	12·804	5·950	6·291	0·000	0·000	–	25·10	

Table 2. Cross-validation (jackknife method) for the regression model corresponding to the plasma protein binding values of hypolipaemic drugs.

	Original model (no deletions)		1 deletion per run (11 runs)	
	Regression value	s.d.	Regression value	s.d.
Correlation coefficient	0·939		0·930	0·055
s.d.	11·068		11·586	1·146
Coefficient of $^{0}\chi/^{0}\chi^{v}$	−267·239	99·596	−283·149	61·847
Coefficient of $^{3}\chi_{c}$	−20·414	2·803	−21·219	2·169
Constant	420·448	113·441	439·972	137·222
Average residual	7·317	1·885	7·655	0·571
Residuals <1 s.d.	63·64%		71·90%	
Residuals between 1 and 2 s.d.	36·36%		26·45%	
Residuals >2 s.d.	0%		1·65%	

Table 3. Cross-validation (jackknife method) for the regression model corresponding to the total cholesterol reduction values of hypolipaemic drugs.

	Original model (no deletions)		2 deletions per run (20 runs)	
	Regression value	s.d.	Regression value	s.d.
Correlation coefficient	0·829		0·827	0·022
s.d.	3·555		3·587	0·283
Coefficient of $^{0}\chi^{v}$	0·268	0·156	0·245	0·173
Coefficient of $^{2}\chi/^{2}\chi^{v}$	−28·198	6·799	−30·161	7·617
Constant	48·114	9·255	50·793	10·371
Average residual	2·287	0·590	2·427	0·587
Residuals <1 s.d.	66·67%		68·89%	
Residuals between 1 and 2 s.d.	33·33%		29·33%	
Residuals >2 s.d.	0%		1·78%	

CP = 4 and $P < 0{\cdot}001$; the t-value for $^{0}\chi^{v}$ was 2·71 ($P < 0{\cdot}05$) and for $^{2}\chi/^{2}\chi^{v}$ was 4·15 ($P < 0{\cdot}001$).

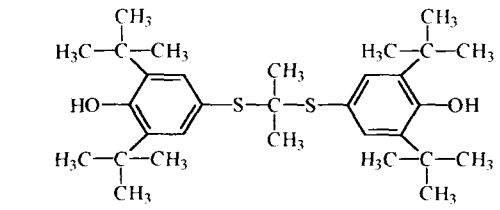
$$\text{LD50 (mmol kg}^{-1}) = -(1{\cdot}49 \pm 0{\cdot}30)^{0}\chi + (14{\cdot}99 \pm 3{\cdot}02)^{3}\chi_{c}^{v} + (18{\cdot}28 \pm 2{\cdot}04) \quad (7)$$

with n = 9, r = 0·900, s.c. = 1·71, d.f. = 8, F = 12·72, CP = 2·55, $P < 0{\cdot}010$; the t-value for $^{0}\chi$ was 4·95 ($P < 0.001$) and for $^{3}\chi_{c}^{v}$ was 4·96 ($P < 0{\cdot}001$).

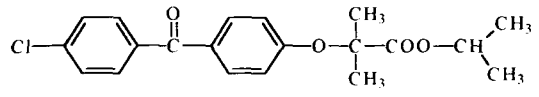
Cross-validation for the selected equations was carried out varying the number of eliminations made and the number of runs for each property in particular, observing that by raising the number of eliminations the model was made more unstable, which on the other hand was to be expected because the degrees of freedom were considerably diminished. Thus, in the case of the percentage of plasma protein binding and LD50 in rat the corresponding stability of 1 elimination was chosen which was repeated a total of 11 and 9 runs, respectively. In the case of percentage of total cholesterol reduction the corresponding stability of 2 elimination was chosen which was repeated a total of 20 runs. This corresponds, in all cases, to approximately 10% of the

Table 4. Cross-validation (jackknife method) for the regression model corresponding to the OLDR values of hypolipaemic drugs.

	Original model (no deletions)		1 deletion per run (9 runs)	
	Regression value	s.d.	Regression value	s.d.
Correlation coefficient	0·900		0·827	0·900
s.d.	1·713		1·686	0·243
Coefficient of $^{0}\chi^{v}$	−1·492	0·301	−1·497	0·322
Coefficient of $^{2}\chi/^{2}\chi^{v}$	14·995	3·023	14·996	3·245
Constant	18·279	2·037	18·343	2·189
Average residual	1·160	0·277	1·177	0·313
Residuals <1 s.d.	77·78%		75·31%	
Residuals between 1 and 2 s.d.	22·22%		22·22%	
Residuals >2 s.d.	0%		2·47%	



Probucol
$^{3}\chi_c = 6.98$
BPP = 5%

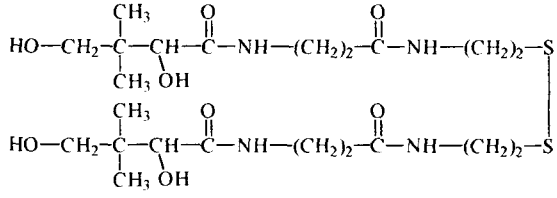


Fenofibrate
$^{3}\chi_c = 2.03$
BPP = 99%

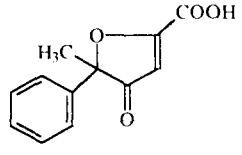
FIG. 2. Percentage of plasmatic protein binding (BPP): chemical structure of probucol and fenofibrate.

eliminated observations, v value recommended by some authors (Kier & Hall 1986) (Tables 2, 3 and 4). The comparison of the results between the obtained values for the selected model and the model of 1 or 2 eliminations shows that the selected equations are very stable as is made patent by the equality of the obtained terms, as well as by the low standard deviations. The analysis of the obtained residuals for the selected model as well as for the 1- or 2-elimination model, manifest minimum discrepancies in the measured value as well as in their standard deviation, an aspect of the study which strengthens the predictive quality of the model.

In equation 5, it is interesting to point out the presence of quotient $^{0}\chi/^{0}\chi^{v}$, a function of the number and nature of the heteroatoms present (Kier & Hall 1986). The $^{3}\chi_c$ index,



Pantethine
$^{0}\chi = 26.30$
OLDR = 13.75 mmol kg^{-1}



Acifran
$^{0}\chi = 10.09$
OLDR = 18.03 mmol kg^{-1}

FIG. 3. LD50 (p.o.) (OLDR) in rats: chemical structure of pantethine and acifran.

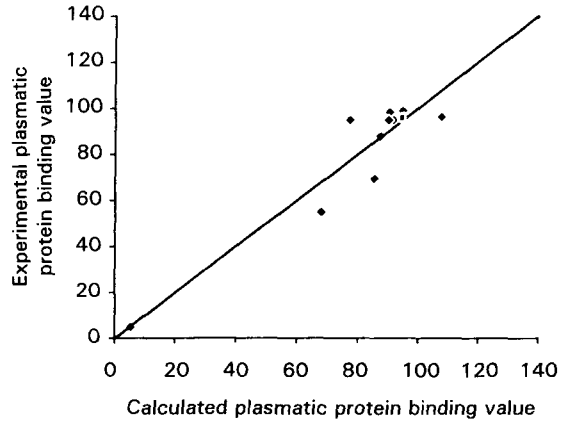


FIG. 4. Correlation between experimental and calculated plasmatic protein binding percentage values of hypolipaemic drugs (eqn 5).

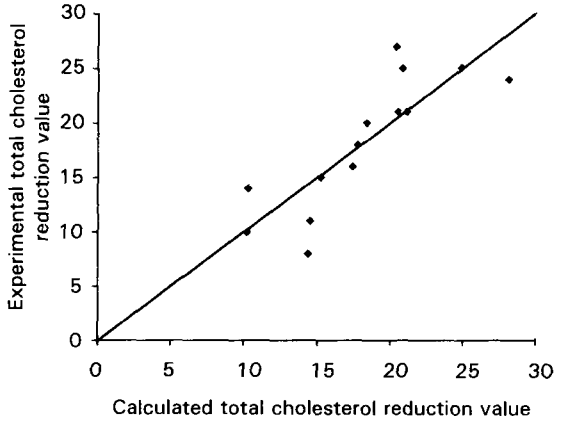


FIG. 5. Correlation between experimental and calculated total cholesterol reduction percentage values of hypolipaemic drugs (eqn 6).

which contributes negatively to the selected equation, is a structural parameter which provides for the number of substituents. Thus, depending on the equation, the number of substituents corresponds a lower degree of plasma protein binding. In the specific case of probucol, a highly-substituted molecule we can see how the value of its percentage of plasma protein binding is the lowest (5%)

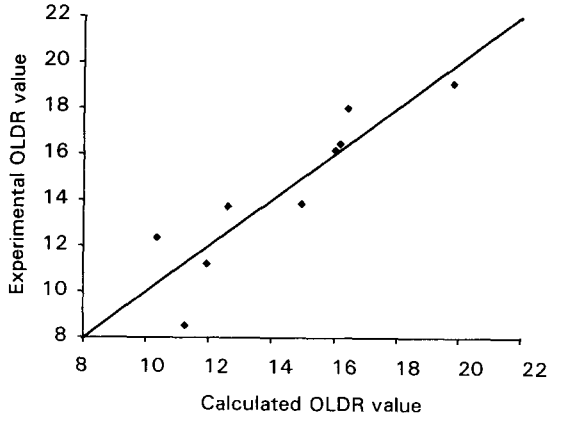


FIG. 6. Correlation between experimental and calculated LD50 (p.o.) values of hypolipaemic drugs (eqn 7).

compared with molecules with few substituents, like fenofibrate (99%) (Fig. 2).

In both equation 5 and equation 6 there is a quotient between non-valence and valence indices of the same order; however, in equation 6 the indices are of second order.

As can be appreciated from equation 7, $^0\chi$ index will contribute to the rise in the value of the property and given that it is related to the molecular volume and the superficial molecular area (Kier & Hall 1986), molecules with a greater volume such as pantethine will possess less toxicity (18·03 mmol kg^{-1}) than those like acifran (13·75 mmol kg^{-1}) whose $^0\chi$ value is considerably less (Fig. 3).

The comparison between the experimental and theoretical values for the studied properties (eqns 5, 6, and 7) is shown in Figs. 4, 5 and 6, respectively.

Conclusion

We conclude that the molecular connectivity method can be applied to the prediction of protein binding, reduction of total cholesterol and LD50 values in rat, for a group of candidate hypolipaemic drugs.

Acknowledgement

The authors wish to thank CICYT, SAF92-0684 (Spanish Ministery of Science and Education) for its economical support of their research effort.

References

Alberts, A. W. (1990) Lovastatin and simvastatin, inhibitors of HMG-CoA reductase and cholesterol biosynthesis. Cardiology 77: 14–21

Ciudad, J., Garcia, R., Galvez, J. (1987) Algoritmo para el cálculo de los indices de conectividad. Su aplicación a la determinación de la refracción molar de un grupo de moléculas. An. Quim. 83: 385–389

Dixon, W. J. (1982) BMD Manual: Biomedical Computer Programs, University of California Press, Los Angeles, CA pp 264–277

Gray H. L., Shucany, W. R. (1972) The Generalized Jackknife Statistic, Marcel Dekker, New York, pp 1–92

Grundy, S. M., Vega, G. L., Garg, A. (1990) Use of 3-hydroxy-3-methylglutaril coenzyme A reductase inhibitors in various forms of dyslipidemia. Am. J. Cardiol. 66: 31B–38B

Hunninghake, M. D. (1990) Drug treatment of dyslipoproteinemia. Endocrinol. Metab. Clin. North. Am. 9: 345–360

Illingworth, D. R., Bacon, S. (1989) Treatment of heterozygous familial hypercholesterolemia with lipid-lowering drugs. Arteriosclerosis 9: 1121–1134

Kier, L. B., Hall, L. H. (1976) In: de Stevens, G. (ed.) Molecular Connectivity in Chemistry and Drug Research, Academic Press, New York, pp 168–184

Kier, L. B., Hall, L. H. (1986) In: Bawden, D. (ed.) Molecular Connectivity in Structure-Activity Analysis, Research Studies Press, Letchworh, pp 1–256

Marshall, F. N. (1982) Pharmacology and toxicology of probucol. Artery 10: 7–21

Monk, J. P., Tood, P. A. (1987) Bezafibrate. A review of its pharmacodynamic and pharmacokinetic properties, and therapeutic use in hyperlipidaemia. Drugs 33: 539–576

Pan, H. Y. (1991) Clinical pharmacology of pravastatin, a selective inhibitor of HMG-CoA reductase. Eur. J. Clin. Pharmacol. 40: 515–518

Randic, M. (1975) On characterization of molecular branching. J. Am. Chem. Soc. 97: 6609–6615

Tilly-Kiesi, M., Tikkanen, M. J. (1991) Low density lipoprotein density and composition in hypercholesterolaemic men treated with HMG-CoA reductase inhibitors and gemfibrozil. J. Intern. Med. 229: 427–434

Vozeh, S., Schmidlin, O., Taeschner, W. (1988) Pharmacokinetic drug data. Clin. Pharmacol. 15: 254–282