The Hypolipidemic Activity of N-Phenylphthalimide Derivatives: a QSAR Study

M. D. Coutinho Neto , W. C. Pinheiro Filho, B. B. Neto

and M. N. Ramos

Departamento de Química Fundamental, Universidade Federal de Pernambuco, 50739 - Recife (PE) , Brasil

Received: september 9, 1993; december 30, 1993.

Um estudo de relações quantitativas estrutura-atividade (QSAR) foi realizado para um grupo de derivados de N-fenilftalimida usando descritores eletrônicos de orbitais moleculares AM1 juntamente com os parâmetros clássicos , $\Im$, $\Re$ e π . Modelos quantitativos baseados em regressão linear múltipla e regressão de quadrados mínimos parciais foram bem sucedidos na previsão de novos agentes hipolipidêmicos, teoricamente mais eficazes do que aqueles determinados experimentalmente. Dos novos compostos projetados, o derivado *o*-NO foi previsto para ser o mais eficaz.

Electronic descriptors obtained from AM1 molecular orbital calculations are employed, together with the $\Im$, $\Re$ and π parameters , in a statistical analysis of the structure-activity relationships in a group of N-phenylphthalimide derivatives. Quantitative models based on multiple linear regression and on partial least squares regression point to new derivatives expected to be more potent hypolipidemic agents than the N-phenylphthalimides for which experimental activities have been determined . The *o*-NO derivative is predicted to be the most effective of the new compounds.

Key words: *QSAR; phenylphthalimide; hypolipidemic activity.*

Introduction

In the last decade Hall and coworkers , experimenting with CF_1 male mice, observed that several N-substituted phthalimides are potent hypolipidemic agents, even when administered at the relatively low dose of 20 mg/kg/day [1-4]. A sixteen-day treatment with N-phenylphthalimide , for example, reduces the serum cholesterol level in mice to 57% of its initial value . Substitution at the phenyl ring drastically affects the hypolipidemic activity. When an ethyl group is introduced the N-phenylphthalimide becomes almost without effect. On the other hand, if the substituent is an ethoxy group at the *ortho* position, the hypolipidemic activity is increased, and the cholesterol level falls off to 43% after treatment. Such substituent effects have already been analysed, in a qualitative way, by Hall *et al.*[4] They noted that substituents that increase lipophilicity tend to decrease the hypolipidemic activity, whereas electron-withdrawing groups generally lead to more potent drugs. They also observed that N-phenylphthalimide and some of its derivatives are more effective hypolipidemic agents than the commercially available and widely used clofibrate .

In the present work the hypolipidemic activity of a group of N-phenylphthalimides is studied from a quantitative point of view . AM1 molecular orbital calculations [5] are performed to obtain electronic descriptors for all of the molecules studied, and then these parameters are subjected to two statistical treatments, in search of quantitative structure-activity relatioships (QSAR) describing the hypolipidemic action of these compounds .

Methods and Experimental

Experimental. The experimental data were obtained by Hall and co-workers and are reported in ref. 4. N-phenylphthalimide and its derivatives used in this QSAR study were suspended in 1% carboxymethylcellulose in water and administered to CF_1 male mice (25 g) at 20 mg / kg / day intraperitoneally for 16 days. On days 9 and 16 blood was obtained by tail-vein bleeding, and serum was separated by centrifugation for 3 min. The serum cholesterol levels obtained were determined by a modification of the Liebermann-Burchard reaction. Serum was also collected on day 14, and the triglyceride level was determined by a commercial kit (Fisher, Hycel triglyceride test kit). The mice were placed on a commercial diet (U.S. Biochemical Corp. basal atherogenic test diet) for a 2-week period. After the cholesterol and triglyceride levels were assayed and observed to be elevated,

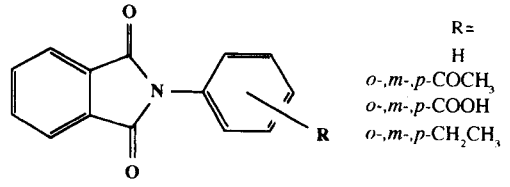

Figure 1. *The N-phenylphthalimides studied in this work. R is the substituent on the phenyl group.*

the mice were administered the test drugs. More biological details and informations with respect to the experimental studies performed by Hall and collaborators, such as enzymatic studies, food intake and cholesterol absorption study, can be found in Ref. 4.

Calculations. As in our recent work on other phthalimides[6], the MOPAC/5 software [7] was employed to perform molecular orbital calculations at the AM1 level. From the wavefunctions thus obtained we extracted the frontier orbital energies, Mulliken net charges, molecular dipole moments and the sums of the absolute values of the net charges over the common skeletal structure of all molecules studied. These variables together with the π, $\Im$ and $\Re$ parameters, which are included to take into account the qualitative observations of Hall *et al.*[4], form a set of 33 descriptors for each N-phenylphthalimide. In geometrical terms, each compound can be thought of as a point in a 33-dimensional space.

The hypolipidemic activities are expressed in terms of pC_{16} and pT_{14}, which are the logarithms of the residual cholesterol and triglyceride levels observed after sixteen- and fourteen-day treatments, respectively. QSAR models relating these activity values to the molecular descriptors were investigated by means of two multivariate techniques: traditional multiple linear regression (MLR), and partial least squares modeling (PLS), which is a kind of regression based on principal component analysis [8]. The PLS calculations were carried out with the SIMCA - 3B package[9].

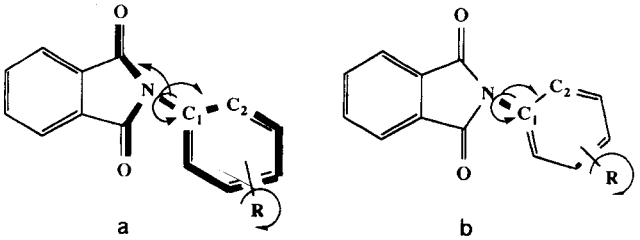

Figure 2. *Regions subjected to geometrical optimization. Optimized bond lengths are shown in bold. Optimized angles are indicated by curved arrows. (a)* Ortho- *substituted derivatives. (b)* Meta- *and para- substituted derivatives.*

Table 1. Experimental, MLR and PLS values of the hypolipidemic activities of N- phenylphthalimide derivatives. pC_{16} and pT_{14} values have been converted back to percent values.[a]

-R	$\%C_{16}$			$\%T_{14}$		
	EXPTL	MLR	PLS	EXPTL	MLR	PLS
1 H	57±3	73	69	61±5	74	75
2 o-$COCH_3$	43±5	51	42	56±9	59	54
3 m-$COCH_3$	84±6	72	71	74±7	69	72
4 p-$COCH_3$	55±4	56	60	66±6	62	65
5 o-C_2H_5	103±11	92	96	92±3	93	95
6 m-C_2H_5	82±10	85	88	95±4	91	88
7 p-C_2H_5	89±6	91	87	96±4	93	88
8 o-COOH	77±9	58	75	78±3	65	76
9 m-COOH	69±6	75	58	71±8	73	64
10 p-COOH	53±5	52	57	58±6	62	61
average error[b]	6.50	7.3	6.4	5.50	4.9	4.4

[a] *$\%C_{16}$ and $\%T_{14}$ are serum cholesterol and triglyceride levels in CF1 male mice after 16- and 14-day treatments, expressed as percent of control (1% carboxymethylcellulose).*

[b] *For MLR and PLS values the errors are given by* $1/n\ \Sigma|exptl - calcd|$.

The N-phenylphthalimides studied in this work are shown in Figure 1. The molecular orbital calculations were performed at partially optimized geometries. The basic structure was initially given the geometrical parameters resulting from a complete optimization of the parent N-phenylphthalimide molecule and then, for each substituent, the derivative structure was subjected to optimization only on those regions likely to be affected by the substitution. These regions are shown in Figure 2. The calculated rotation barrier for the phenyl ring in phenylphthalimide has a minimum around 23°, with a height of ~ 2.5 kcal. Substitution at the *meta* and *para* positions on the phenyl ring leave the barrier unaffected, but *ortho* substituents lead to a two- to threefold increase in the barrier height, and in some cases to two different minima. All descriptors employed in the statistical models were calculated at the geometry corresponding to the absolute minimum of each molecular structure.

Results and Discussion

(a) Multiple linear regression. The variables to be included in the regression equation were chosen on the basis of the correlations observed between the thirty-three descriptors initially considered and the two hypolipidemic activities. Those descriptors exhibiting the largest correlation coefficients with PC_{16} and pT_{14} were identified and, from these, the ones less correlated with one another were finally selected. Only two descriptors, the charge on the phenyl carbon attached to the nitrogen atom (qC_1) and the lipophilicity constant of the substituent (π), met both criteria. Regression

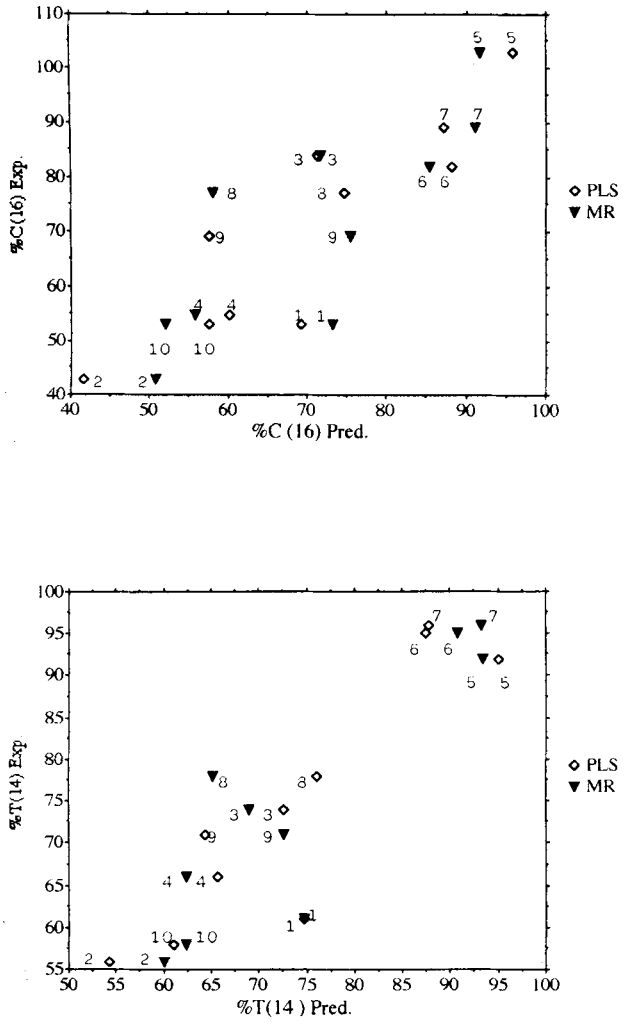

Figure 3. *Experimental and calculated values of the hypolipidemic activities of N-phenylphthalimide derivatives . Triangles stand for MLR values and lozenges represent PLS values. (a) %C_{16} values (b) %T_{14} values . Numbering as in Table 1.*

of pT_{14} and pC_{16} on these parameters led to the following equations:

$$pT_{14} = 1.957 + 0.087\,\pi - 1.207\,qC_1 \qquad (1)$$

with R = 0.87 , F = 10.6 and n = 10 ; and

$$pC_{16} = 2.071 + 0.074\pi - 2.984\,qC_1 \qquad (2)$$

with R = 0.84 , F = 7.50 and n = 10 ;

Smaller values of pT_{14} and pC_{16} correspond to larger activity values. According to these equations, therefore, compounds with a less lipophilic substituent and with a more positive C_1 will tend to be more hypolipidemic. As electronegative substituents at the *ortho* and *para* positions tend to increase qC_1, these results support the conclusions stated in Ref . 4 .

(b) Partial least squares modeling . In this multivariate technique the descriptor matrix is modelled by a set of latent variables, which are closely related to its principal components. In fact, they may be considered principal components slightly rotated in order to maximize the correlation with the variables to be predicted.

Each latent variable is a linear combination of all of the original descriptors, and describes a fraction of the total information contained in the data set. In the present case some of the descriptors were found to contribute little to the first latent variables, and were consequently excluded from the model. The number of descriptors was thus reduced to fifteen: the lipophilicity constant, the dipole moment, the frontier orbital energies, the sum of the absolute atomic charges, and ten individual atomic charges, mostly of the imide ring. The final PLS model contains three uncorrelated latent variables, which together describe 92.5 % of the total variance. With this model the values of both pC_{16} and pT_{14} can be predicted at the same time. In Table 1 the predictions of both models, MLR and PLS, are compared with the experimental hypolipidemic activities. The average errors of the PLS model are smaller than the errors of the two independent MLR equations, and fall within the corresponding average experimental errors. Both models give more or less the same profile

Table 2 . Predicted hypolipidemic activities , carbon atomic charges and lipophilicity constants for some N-phenylphthalimide derivatives .

-R	(%T_{14})pred MLR	PLS	(%C_{16})pred MLR	PLS	qC_1	π
o-NO	48	40	36	27	0.143	-1.20
p-$SOCH_3$	49	49	43	43	0.107	-1.58
p-NO	52	54	45	45	0.111	-1.20
p-$CONH_2$	52	54	48	45	0.094	-1.49
p-COH	59	58	51	50	0.107-	0.65
o-$COCH_3$[a]	60	54	51	42	0.109	0.55

a o- (N- phthalimido) acetophenone is the most active compound used in model building (see Table 1).

for the distribution of hypolipidemic activities, as can be seen in Figure 3. The *o*- $COCH_3$ and *p*- COOH derivatives are predicted to be more potent hypolipidemic agents than the parent molecule, in agreement with experiment. The MLR and PLS values correlate well with each other (correlation coefficients of 0.84 and 0.91 for %C_{16} and %T_{14}, respectively).

(c) New hypolipidemic agents. Although the PLS results are in better agreement with the experimental findings, they are harder to interpret, due to the relatively large number of descriptor variables involved (fifteen, in this case). The MLR

values, on the other hand, are determined by only two variables, the lipophilicity constant and the charge on the C_1 atom. According to equations 1 and 2 substituents that lower the first variable and raise the second should lead to drugs with larger hypolipidemic activities. Several such substitutents were therefore selected based on these criteria, and AM1 molecular orbital calculations were performed on the corresponding derivative structures, to yield values of qC_1 needed for both statistical models, whereas the corresponding values of π were obtained from Ref. 10. Feeding these values into the model equations resulted in the predicted activities shown on Table 2, where they are compared with the activities of the most effective phthalimide used in the model building phase , *o*- (N - phthalimido) acetophenone. All the new compounds are predicted to be stronger hypolipidemic agents, although by a relatively small margin, and the predictions of both models are in very good agreement.

Conclusion

The statistical analysis reported in this work indicates that the hypolipidemic activity of N- phenylphthalimide derivatives may be related to the charge on the carbon atom to which the phenyl group is attached and to the lipophilicity constant of the substituent. Compounds selected on the basis of these considerations are calculated to be more active than any of the molecules employed for model building. They may be viewed therefore as potential candidates for new experimental screening studies of the same kind.

Acknowledgements

We gratefully acknowledge financial support from the Brazilian agencies CNPq , FINEP and CAPES .

References

1. J. M. Chapman Jr., S. D. Wyrick, P. J. Voorstad, J. H. Maguire, G. H. Cocolas, I. H. Hall, *J. Pharm. Sci.* **73** , 1482 (1984).
2. A. R. K. Murthy , J. M. Chapman Jr., S. D. Wyrick, I. H. Hall , *Pharm. Res.* **3** , 286 (1986).
3. A.R.K. Murthy, O.T.Wong, D. J. Reynolds, I. H. Hall, *Pharm. Res.*, **4**, 21 (1987).
4. J. M. Chapman Jr., P. J. Voorstad, G. H. Cocolas, I. H. Hall, *J. Med. Chem.* **26**, 237 (1983).
5. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart, *J. Am. Chem. Soc.* **107**, 3902 (1984).
6. M. N. Ramos, B.B. Neto, *J. Comp. Chem.* **11**, 569 (1990).
7. J. J. P. Stewart, *J. Compt-Aided Mol. Design*, **4**, 1 (1990).
8. (a) H. Wold , in *Systems Under Indirect Observation*, edited by K. G. Joreskog and H. Wold (North-Holland , Amsterdam, 1982), Vol II, Chap. 1; (b) W. G. Glen , W. J. Dunn III, D.R. Scott , *Tetrahedron Comp. Method.* **2**, 349 (1988).
9. Principal Data Components, 2505 Shepard Blvd , Columbia, MO 65201, USA.
10. C. G. Swain, E. C. Lupton , *J. Am. Chem. Soc.* **90**, 4328 (1968).