J. Pharm. Pharmacol. 1999, 51: 111–117
Received July 24, 1998
Accepted September 22, 1998

© 1999 J. Pharm. Pharmacol.

Discovery of New Antimalarial Compounds by use of Molecular Connectivity Techniques

R. GOZALBES, J. GÁLVEZ, A. MORENO* AND R. GARCÍA-DOMENECH

*Molecular Connectivity & Drug Design Research Unit, Department of Physical Chemistry, Faculty of Pharmacy, University of Valencia, 46100 Burjassot, Valencia, Spain and *Parasitology Service of Charles Nicolle University Hospital, 76031 Rouen, France*

Abstract

Molecular connectivity has been applied to the search for new compounds with antimalarial activity. Linear discriminant analysis and connectivity functions were used to select several potentially suitable drugs which were tested for antimalarial properties by use of an in-vitro micro test which estimates parasite growth by measurement of incorporation of [^{3}H]hypoxanthine.

Hexetidine stands out among the compounds selected. Activity assays were performed with *Plasmodium falciparum* passou and 3CD7 strains, for which the IC50 values (doses resulting in 50% inhibition) were 320 and 400 ng mL^{-1}, respectively. These results are comparable with those obtained for quinine chlorhydrate (IC50 = 60 and 107·8 ng mL^{-1}) and chloroquine sulphate (IC50 = 231 and 415 ng mL^{-1}), the drugs used for reference.

These results demonstrate the usefulness of our topological approach for the selection and design of new lead drugs active against *Plasmodium falciparum.*

More than 15 million chemicals have been discovered or synthesized in laboratories. This extraordinary database is increased annually by nearly six hundred thousand new molecules. Many of these compounds have found no application, although they could be of potential use in the chemical, biological or pharmaceutical industries. Because the tests which must be performed, especially pharmacological and toxicological tests, are usually expensive and time-consuming, during the last few years the pharmaceutical industry has reoriented its research strategy to the development of methods enabling rational selection or design of novel compounds with the desired properties.

Molecular connectivity is a useful tool for describing molecular structure which has been used for efficient analysis of QSAR data. One of the most interesting advantages of molecular topology is the straightforward calculation of simple topological descriptors. With this method each structure is presented as a hydrogen-suppressed graph in which the atoms are represented by vertices and the bonds by edges. The connectivity between each atom to the others is included in a topological matrix, either distance or adjacency, and its mathematical manipulation provides different sets of numbers or topological indexes, TIs, which characterize each molecule at different descriptive structural levels. These descriptors have shown their usefulness in the prediction of diverse physical, chemical and biological properties of various types of compound (Kier & Hall 1986; Julián-Ortiz et al 1996). Recent studies have demonstrated the possibility of finding new antiviral compounds (Gálvez et al 1995), cytostatics (Gálvez et al 1996a), hypoglycaemics (Antón-Fos et al 1994), β-blockers (García-Domenech et al 1997), analgesics (Gálvez et al 1994a; García-Domenech et al 1996; García-March et al 1997) and bronchodilators (Rios-Santamarina et al 1997) which can be regarded as drug leads.

The objective of this work was to use molecular connectivity to generate discriminant functions enabling the selection of molecules with antimalarial activity. Malaria is now the disease with the highest death rate throughout the world,

Correspondence: R. García-Domenech, Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular, Departamento de Química-Física, Facultad de Farmacia, Avda. Vicent Andrés Estellés, s/n. 46100 Burjassot–Valencia, Spain.

especially in tropical countries. It is estimated that 2–3 million people die each year from malaria.

Materials and Methods

The first step is the selection of compounds with antimalarial activity. Next the topological indexes are calculated for each and the statistical techniques of multilinear regression, e.g. linear discriminant analysis, are used to find the activity discriminant functions. These enable recognition of whether or not a compound has the desired pharmacological activity. Next we make use of the discriminant functions to select compounds with the predicted theoretical activity from among the databases that are available. The compounds found are finally submitted to standard pharmacological tests to corroborate their theoretical activity.

Calculation of topological descriptors

Each compound has been characterized by a set of 62 topological indexes, all obtained by use of adjacency and distance matrices. We have used the Kier & Hall (1977) connectivity indexes up to fourth order, and combinations of these (mainly differences and quotients between valence and non-valence indexes, $\Delta^{m}\chi_t = {}^{m}\chi_t - {}^{m}\chi_t^{v}$ and $C^{m}\chi_t = {}^{m}\chi_t / {}^{m}\chi_{t^v}$ respectively), and the more recently introduced charge indexes up to fifth order (Gálvez et al 1994b). The charge indexes evaluate the global charge transferred between pairs of atoms in the molecule.

The topological charge indexes Gk and Jk are defined as:

$$Gk = \sum_{i=1,j=i+1}^{i=N-1,j=N} |CTij|\delta(k, Dij) \qquad (1)$$

and

$$Jk = \frac{Gk}{(N-1)} \qquad (2)$$

where N is the number of vertices (atoms other than hydrogen) and $CT_{ij} = m_{ij} - m_{ji}$, where m stands for the elements of the M matrix:

$$M = A \times D^{*} \qquad (3)$$

where A is the adjacency (N × N) matrix, D* the matrix of inverse square distances, in which their diagonal entries are assigned as 0 and δ (Kronecker's delta).

Hence, Gk represents the sum of all the CT_{ij} terms, with $D_{ij} = k$, D_{ij} being the entries of the topological distance matrix.

Because geometrical factors such as the molecular shape can affect pharmacological activity, a simple set of descriptors named geometrical indexes was also introduced; these evaluate topological or geometrical parameters of the connectivity graph through natural numbers (Gálvez et al 1994a). In this work the PR1, PR2 and S descriptors were particularly useful (PR1 and PR2 are the number of pairs of ramifications separated by one and two edges, respectively, and S is the surface parameter).

All the topological indexes were calculated using computer software developed in our research unit.

Discriminant functions

Stepwise linear discriminant analysis, SLDA, is an useful technique for finding discriminant functions. It is a pattern-recognition method which enables, by combining variables, evaluation of the ability to distinguish among two or more groups of populations. In our work the independent variables were the topological indexes and the discrimination property was the antimalarial activity. The SLDA study is performed with two groups of compounds—the training group (which includes active and inactive compounds), enabling discovery of the discriminant function, and test group (also with active and inactive structures, randomly chosen from the training group), which enables evaluation of the quality of the discriminant function obtained.

Election of connectivity functions was performed with the BMDP 7M package (Dixon 1990). The method used for selection of the descriptors was the *F*-Snedecor parameter, and the classification criterion used was the minimum value of the Mahalanobis distance. The quality of the discriminant function is evaluated through Wilk's *U*-statistical parameter.

Pharmacological distribution diagrams

One way of selecting useful connectivity functions for the molecular selection step is by obtaining pharmacological distribution diagrams; these were developed recently in our research unit and have been used with success in the bronchodilator and antimicrobial therapeutic groups (Gálvez et al 1996b; Rios-Santamarina et al 1997). This consists in applying the discriminant function to the group of active compounds and to a representative group of inactive molecules. The structures are grouped into the predicted value of the property P for each x_i interval, and the frequency with which it appears along each interval of P is determined for each group. The expectancy E of finding a molecule in a given interval x is defined as:

Table 1. Classification of compounds in the training and test series by use of the discriminant function P_1.

Training Set

Active group				Inactive group			
Compound	P_1	*P*	Classn[a]	Compound	P_1	*P*	Classn[a]
Acedapsone	1·24	0·76	+	Androsterone	−1·14	0·88	–
Amodiaquin	0·39	0·60	+	Anisindione	−1·23	0·77	–
Arteether	3·64	0·96	+	Apronalide	0·69	0·67	–
Artemisinin	2·53	0·90	+	Aspirin	−2·13	0·89	–
Cinchonine	0·59	0·64	+	Benomyl	−1·31	0·79	–
Chlorguanide	0·30	0·58	+	Benzphetamine	0·94	0·28	+
Chloroquine	1·21	0·77	+	Bethanidine	−2·43	0·92	–
Chlorproguanil	0·12	0·53	NC	Biguanide	−3·58	0·97	–
Dapsone	2·02	0·78	+	Bixin	−3·92	0·98	–
Halofantrine	5·15	0·99	+	*p*,α-Dibromotoluene	−2·00	0·88	–
Methylarsacetin	1·81	0·85	+	Bucrylate	−2·03	0·88	–
Mefloquine	7·57	1·00	+	Butamisole	−0·27	0·57	–
Pamaquine	1·20	0·77	+	Butoxycaine	0·08	0·48	NC
Pyrimethamine	−1·34	0·21	–	Capsaicin	−0·48	0·62	–
Quinacrine	1·61	0·83	+	Carbimazole	−2·62	0·93	–
Quinine	1·56	0·83	+	Carbuterol	−2·27	0·96	–
Quinoline	−0·61	0·35	–	Caroxazone	−1·55	0·83	–
				Cinoxacin	−2·46	0·92	–
				Cyromacine	−0·58	0·64	–
				Diethylpropion	−0·09	0·52	NC
				Dimecrotic acid	−2·73	0·94	–
				Dithianone	−4·93	0·99	–
				Droxidopa	−1·11	0·75	–
				Fenoterol	1·33	0·21	+

Test Set

Active group				Inactive group			
Compound	P_1	*P*	Classn[a]	Compound	P_1	*P*	Classn[a]
Artemether	3·54	0·96	+	Anthralin	−2·18	0·90	–
Artesunate	3·71	0·97	+	Baclofen	−0·92	0·72	–
Bebeerine	2·32	0·91	+	Barbital	−3·37	0·97	–
Cycloguanil	−2·38	0·06	–	β-Benzalbutyramide	−2·13	0·89	–
Hydroxychloroquine	1·31	0·79	+	Cefuroxime	−1·39	0·80	–
Plasmocid	0·14	0·54	NC	Cotinine	−0·85	0·70	–
Primaquine	0·26	0·57	+	Cropropamide	−0·75	0·68	–
Quinocide	0·65	0·66	+	Dopamine	−2·25	0·90	–
				Emodin	−1·43	0·81	–
				Emylcamate	−1·34	0·85	–

[a]Classification. NC = not classified. The active group comprises antimalarial drugs and the inactive group non-antiprotozoan drugs.

E_a = percentage of active compounds in x/(percentage of inactive compounds in x + 100)

E_i = percentage of inactive compounds in x/(percentage of active compounds in x + 100)

where E_a and E_i are the activity expectancy and inactivity expectancy, respectively.

When for a connectivity function, P_i, E_a takes the form of a distribution curve and E_i tends to 0 along the curve, the overlapping is minimum, so this function can be useful for molecular selection. This enables determination of the intervals of the property where the probability of finding new active compounds is maximum. In this work, we used linear discriminant functions to build anti-Plasmodium activity distribution diagrams.

Pharmacological tests

The compounds selected as having theoretical antimalarial activity were tested by use of the micro-test isotopic method (Brasseur et al 1986). The compounds were tested at different concentrations and the activity was evaluated as a function of the rate of incorporation of a radioactively labelled compound, [^{3}H]hypoxanthine, during 48 h culture. The basis of this method is that

the parasites use the hypoxanthine as a precursor of amino acids, and so the higher the antimalarial activity of the new compounds the lower its uptake. Two strains of *Plasmodium falciparum*, passou and 3CD7, were used in this study.

Results and Discussion

To obtain the discriminant functions, two different studies were performed, SLDA between antimalarial drugs and non-antiprotozoan drugs and SLDA between antimalarial drugs and antiprotozoan drugs without antimalarial activity.

In the first of these studies the discriminant function obtained was:

$$P_1 = 0 \cdot 56 + 3 \cdot 25^3\chi_c - 27 \cdot 88^4\chi_c^v - 8 \cdot 64J_2 \quad (4)$$

$$N = 59;\ U - \text{statistic (Wilk's)} = 0{\cdot}55;\ F = 9{\cdot}83$$

This function can be used to distinguish between potentially active antimalarial compounds and inactive compounds.

Table 1 summarizes the classification results obtained with the P_1 discriminant function. A compound will be selected as an antimalarial if $P_1 > 0$ or as non-antimalarial if $P_1 < 0$. It is

Table 2. Classification of compounds in the training and test series by use of the discriminant function P_2.

Training Set

Active group				Inactive group			
Compound	P_1	*P*	Classn[a]	Compound	P_1	*P*	Classn[a]
Acedapsone	7·89	1·00	+	Acetarsone	−2·04	0·88	–
Amodiaquin	5·50	1·00	+	Azanidazole	−5·40	1·00	–
Arteether	5·02	0·99	+	Bialamicol	−1·45	0·81	–
Artemether	4·63	0·99	+	Bitoscanate	−1·68	0·84	–
Artemisinin	3·95	0·98	+	Carbarsone	−2·52	0·92	–
Artesunate	4·78	0·99	+	Clioquinol	−3·84	0·98	–
Cycloguanil	−0·84	0·31	–	Chlorbetamide	−4·69	0·99	–
Chlorguanide	5·61	1·00	+	Conessine	−1·71	0·84	–
Chloroquine	4·27	0·99	+	Eflornithine	−7·31	1·00	–
Dapsone	0·86	0·71	+	Ethylstilbamidine	−0·45	0·61	–
Halofantrine	4·58	0·99	+	Stibocaptate	−1·47	0·82	–
Mefloquine	0·05	0·53	NC	Phanquinone	−2·05	0·88	–
Pamaquine	4·96	0·99	+	Forminitrazole	−5·87	1·00	–
Plasmocid	2·30	0·91	+	Diloxanide furoate	−4·20	0·99	–
Pyrimethamine	−1·25	0·23	–	Furazolidone	−7·70	1·00	–
Primaquine	3·34	0·97	+	Hydroxystilbamidine	−7·01	1·00	–
Quinacrine	7·83	1·00	+	Metronidazole	−10·6	1·00	–
Quinine	1·30	0·80	+	Monacrin	−1·84	0·86	–
Quinoline	−1·17	0·25	–	Nifuroxime	−8·37	1·00	–
				Paromomycin	−7·15	1·00	–
				Pentamidine	−3·93	0·98	–
				Propamidine	−4·01	0·98	–
				Quinfamide	−2·39	0·91	–
				Sulpharside	−2·41	0·92	–
				Thiocarbamizine	−2·12	0·89	–
				Tryparsamide	−3·12	0·96	–

Test Set

Active group				Inactive group			
Compound	P_1	*P*	Classn[a]	Compound	P_1	*P*	Classn[a]
Bebeerine	5·21	1·00	+	Aminitrozole	−2·95	0·95	–
Cinchonine	−1·73	0·16	–	Dehydroemetine	−1·57	0·92	–
Chlorproguanil	3·82	0·98	+	Emetine	−0·23	0·54	NC
Hydroxychloroquine	4·77	0·99	+	Stilbamidine	−5·48	1·00	–
Methylarsacetin	−0·86	0·30	–	Iodoquinol	−3·59	0·97	–
Quinocide	2·36	0·92	+	Lauroguadine	−0·02	0·50	NC
				Melarsoprol	8·19	0·00	+
				Secnidazole	−6·34	1·00	–
				Tioglycolamide	0·07	0·48	NC

[a]Classification, NC = not classified. The active group comprises antimalarial drugs, the inactive group antiprotozoan non-antimalarial drugs.

apparent from both training and test groups that the average measure of correct prediction is greater than 95%. Those compounds whose probability of classification (Prob. in the table) was between 45% and 55% were not classified (NC).

In the study of SLDA between antimalarial drugs and antiprotozoan drugs without antimalarial activity the function obtained was:

$$P_2 = -2{\cdot}92 + 6{\cdot}45^3\chi_c - 2{\cdot}78^4\chi_{pc} + 0{\cdot}39G_1^v - 1{\cdot}72G_3 - 2{\cdot}24^3\chi_c/^3\chi_c^v - 2{\cdot}62PR1 + 1{\cdot}66PR2 + 0{\cdot}04S \quad (5)$$

$$N = 60;\ U - \text{statistic (Wilk's)} = 0{\cdot}35;\ F = 8{\cdot}88$$

This function, which is more selective, can be used to distinguish between antimalarial compounds and compounds which have antiprotozoan activity but are not antimalarial. Table 2 shows the classification for each compound. In this case, we work with a successful classification probability >95%, which is indicative of the quality of the selected discriminant function.

Figures 1 and 2 show the antimalarial activity distribution diagrams obtained with the functions P_1 and P_2 (white and black bars represent active and inactive sets, respectively). It is apparent that the regions with minimum overlap for the compounds with theoretical antimalarial activity are placed in $P_1 > 1{\cdot}00$ and $P_2 > 1{\cdot}00$, so the highest activity expectancy occurs in these intervals.

After applying the functions P_1 and P_2 to a large data base containing 15 000 (approx.) commercial compounds (Merck 1989; Sigma Aldrich catalogue) we selected as theoretically active those compounds with P_1 and P_2 values within the intervals previously designed (Table 3).

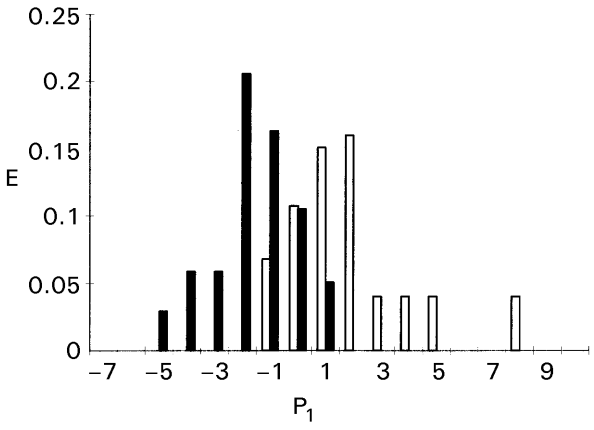

Figure 1. Pharmacological distribution diagram for antimalarial activity obtained using the discriminant function P_1 (black bars indicate inactive group, white bars indicate active group).

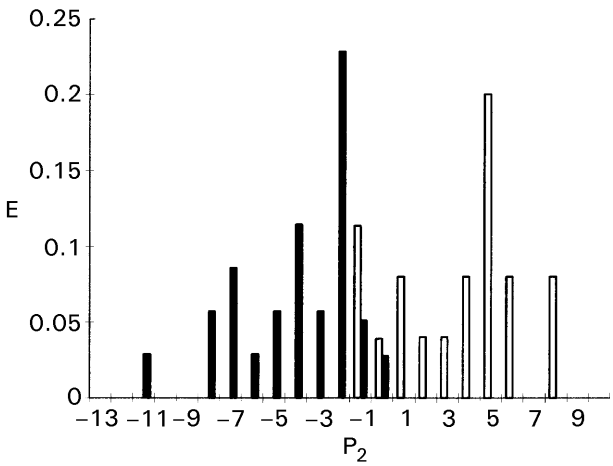

Figure 2. Pharmacological distribution diagram for antimalarial activity obtained using the discriminant function P_2 (black bars indicate inactive group, white bars indicate active group).

The validity of the results is clarified if we perform adequate pharmacological trials to corroborate the expected activity for each compound. In this work we have chosen hexetidine, hydroxyzine and tripelennamine because they are drugs with well-known toxicity (antifungal, tranquilizer minor and antihistaminic, respectively) (Merck 1989). The tests showed that two of the molecules tested had some antimalarial activity. Figures 3 and 4 show the activity results at different concentrations obtained with hexetidine and hydroxyzine, respectively. The study was conducted with two different strains, *Plasmodium falciparum* passou and *Plasmodium falciparum* 3CD7. The drugs used as reference were quinine chlorhydrate and chloroquine sulphate.

Table 3 summarizes the results from the activity test, quantified as IC50 (the dose resulting in 50% inhibition; ng mL^{-1}). Among the compounds selected hexetidine stands out with IC50 values comparable with those of the drugs, e.g. chloroquine sulphate, most commonly used in therapy against malaria. Furthermore, hexetidine is used as an antiseptic mouth wash and is, therefore, a drug with known low toxicity. It could be very interesting to perform research to assess its potential as an antimalarial drug.

These results demonstrate that an appropriate choice of topological indexes is sufficient to enable prediction of antimalarial activity by linear discriminant analysis. Acceptable efficiency is found in the search for new drugs, especially if we take into account the simplicity of the calculations used. Finally, we believe this method could be very

Table 3. Selected compounds, values of discriminant functions P_1 and P_2, and antimalarial activity results for the selected compounds and for the drugs used as referennce.

Compound	P_1	P_2	IC50[a]	
			P. falciparum passou	*P. falciparum* 3CD7
1,5,9-Trimethyl-1,5,9-triaza-cyclononane	2·6	5·4	NT	NT
Alphazurine A	6·8	7·8	NT	NT
3-Piperidine-1,2-propanediol	1·6	1·3	NT	NT
Hexetidine	2·5	10·3	320·0	400·0
Hydroxyzine	1·5	3·9	5300·0	9200·0
Tripelennamine	1·4	3·2	Inactive	Inactive
Reference compounds				
Quinine chlorhydrate	1·6	1·3	60·0	107·8
Chloroquine sulphate	1·2	4·3	231·5	415·0

[a]Dose resulting in 50% inhibition (ng mL^{-1}). NT = not tested.

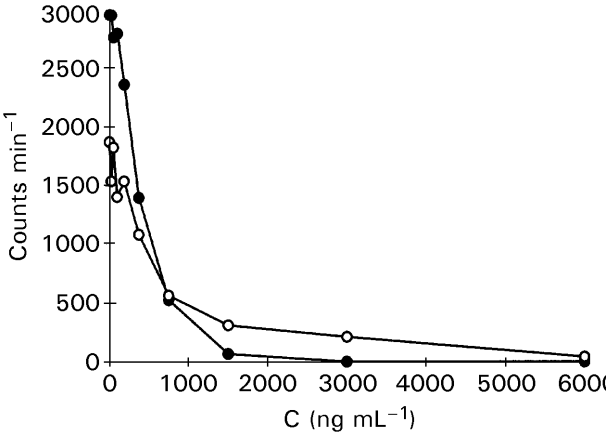

Figure 3. Activity results obtained with hexetidine and *Plasmodium falciparum*: ●, passou strain; ○, 3CD7 strain.

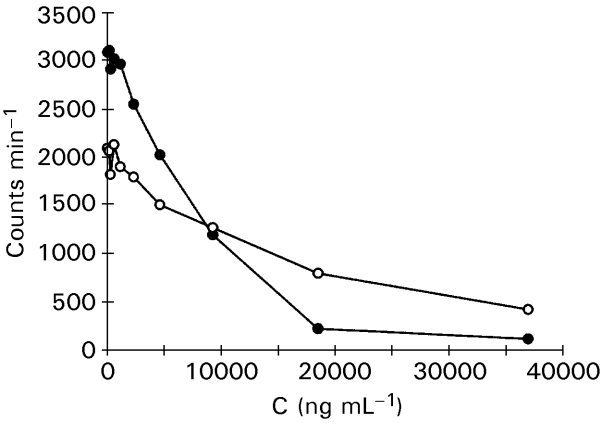

Figure 4. Activity results obtained with hydroxyzine and *Plasmodium falciparum*: ●, passou strain; ○, 3CD7 strain.

useful for the rational design of new antimalarial compounds, because the compounds selected can be considered, because of their chemical structure, as lead drugs and starting points for optimization of antimalarial activity by use of molecular modelling techniques.

Acknowledgements

The authors thank Dr Philippe Brasseur for authorizing the pharmacological tests in the Parasitology Service of the Charles Nicolle University Hospital (Rouen, France), and also Carolina Gregorio and Henry Bainbridge for their help with the editing of this paper. This study has been supported by CICYT, SAF96-0158-C02-02 (Ministerio Español de Educación y Cultura) and GV-3300/95 (Generalitat Valenciana).

References

Antón-Fos, G. M., García-Domenech, R., Perez-Gimenez, F., Peris-Ribera, J. E., García-March, F. J., Salabert-Salvador, M. T. (1994) Pharmacological studies of the new hypoglycaemic compounds 4-(3-methyl-5-oxo-2-pyrazolin-1-yl) benzoic acid and 1-(mesitylen-2-sulphonyl)11*H*-1,2,4-triazole. Arzneim. Forsch. 44: 821–826

Brasseur, P., Druilhe, P., Kouamouo, J., Brandicourt, O., Danis, M., Moyou, S. R. (1986) High level of sensitivity to chloroquine of 72 *Plasmodium falciparum* isolates from southern Cameroon in January 1985. Am. J. Trop. Med. Hyg. 35: 711–716

Dixon, W. J. (1990) BMDP Statistical Software, University of California, Berkeley, USA.

Gálvez, J., Garcia, R., Julian-Ortiz, J. V. de, Soler, R. (1994a) Topological approach to analgesia. J. Chem. Inf. Comput. Sci. 34: 1198–1203

Gálvez, J., García, R., Salabert, M. T., Soler, R. (1994b) Charge indexes. New topological descriptors. J. Chem. Inf. Comput. Sci. 34: 520–525

Gálvez, J., Garcia, R., Julian-Ortiz, J. V. de, Soler, R. (1995) Topological approach to drug design. J. Chem. Inf. Comput. Sci. 35: 272–284

Gálvez, J., Gomez-Lechón, M. J., García-Domenech, R., Castell, J. V. (1996a) New cytostatic agents obtained by molecular topology. Bioorg. Med. Chem. Lett. 6: 2301–2306

Gálvez, J., García-Domenech, R., Gregorio Alapont, C. de, Julián-Ortiz, J. V. de, Salabert-Salvador, M. T., Soler-Roca, R. (1996b) New antibacterial drugs designed by molecular connectivity, advances in molecular similarity. JAI Press Inc. 1: 267–280

García-Domenech, R., García-March, F. J., Soler, R., Galvez, J., Antón-Fos, G. M., Julián Ortiz, J. V. de (1996) New analgesics designed by molecular topology. Quant. Struct.-Act. Relat. 15: 201–207

García-Domenech, R., Gregorio-Alapont, C. de, Julián-Ortiz, J. V. de, Gálvez, J., Popa, L. (1997) Molecular connectivity to find β-blockers with low toxicity. Bioorg. Med. Chem. Lett. 7: 567–572

García-March, F. J., García-Domenech, R., Gálvez, J., Antón-Fos, G. M., Julián-Ortiz, J. V. de, Giner-Pons, R., Recio-Iglesias, M. C. (1997) Pharmacological studies of 1-(*p*-chlorophenyl)propanol and 2-(1-hydroxy-3-butenyl)phenol: two new non-narcotic analgesics designed by molecular connectivity. J. Pharm. Pharmacol. 49: 10–15

Julián-Ortiz, J. V. de, García-Domenech, R., Gálvez, J., Soler, R., García-March, F. J., Antón-Fos, G. M. (1996) Use of topological descriptors in chromatographic chiral separations. J. Chromatogr. A 719: 37–44

Kier, L. B., Hall, L. H. (1977) The nature of structure-activity relationships and their relation to molecular connectivity. Eur. J. Med. Chem. 12: 307–312

Kier, L. B., Hall, L. H. (1986) Molecular Connectivity in Structure-activity Analysis, Research Studies Press, Letchworth, UK, pp 225–246

Merck & Co. Inc. (1989) The Merck Index, 11th edn, Rahway, NJ USA.

Rios-Santamarina, I., García-Domenech, R., Cortijo, J., Santamaría, P., Morcillo, E., Gálvez, J. (1998) New bronchodilators selected by molecular topology. Bioorg. Med. Chem. Lett. 8: 477–482