Pergamon

Bioorganic & Medicinal Chemistry Letters 10 (2000) 2033–2036

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

Molecular Topology: A Useful Tool for the Search of New Antibacterials

C. de Gregorio Alapont,[a] R. García-Domenech,[a] J. Gálvez,[a,*] M. J. Ros,[b] S. Wolski[b] and M. D. García[b]

[a]*Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular Departamento de Química Física, Facultad de Farmacia, Universitat de Valencia, 46100 Burjassot, Valencia, Spain*

[b]*Departamento de Microbiología, Facultad de Ciencias Biológicas, Universitat de Valencia, 46100 Burjassot, Valencia, Spain*

Received 20 October 1999; revised 5 July 2000; accepted 6 July 2000

Abstract—Molecular topology has been applied to find new lead antibacterial compounds. Among the selected compounds, hesperidin, neohesperidin and Mordant Brown 24 stand out, with minimum inhibitory concentrations 90, $MIC_{90} < 0.3$ mg / mL.

The difficulty in finding new antibacterial drugs with a specific spectrum of action, appearance of acquired resistance due to their indiscriminate use and the necessity of antimicrobial drugs with pharmacokinetical characteristics, such as making them useful to fight particular diseases, have made the research of new anti-infective agents to be one of the prime objectives of medicinal chemistry.

Finding a new therapeutic activity for a compound which provides pharmacological and toxicological information means an important saving of money and time which improves its pharmaceutical development as an alternative medicine.

At present, different methods are used for that purpose; one of them is Molecular Topology, particularly Molecular Connectivity, which is a useful instrument to find quantitative structure–activity relationships (QSAR).

In this method, each structure is assimilated as a hydrogen suppressed graph where the atoms are represented by vertices and the bonds by edges, the connectivity between each atom to the others is included in the topological matrix, either distance or adjacency.

Mathematical use of it provides different sets of numbers (topological descriptors), which conjointly can be used to characterise each molecule at different descriptive structural levels.

Topological indices have demonstrated their utility in the prediction of the diverse physical, chemical and biological properties for different types of compounds,[1–3] recently their utility has been demonstrated on the design of new antivirals,[4] cytostatics,[5] sedative/hypnotics,[6] analgesics,[7] antifungals[8] and bronchodilators,[9] many of which can be considered as lead drugs.

In this formalism, the first step is the selection of the compounds with antibacterial activity, then the topological descriptors are calculated for each one. Next, different statistical techniques of multilinear regression are used (such as stepwise linear discriminant analysis, SLDA),[10] in order to find the discriminant functions able to recognise whether a given compound has the desired pharmacological activity or not.

Next, we make use of the discriminant functions for the selection of the compounds showing the predicted theoretical activity, amongst the actual database. The compounds found should finally be submitted to standard pharmacological tests in order to corroborate their theoretical activity.

In this work we have used Kier and Hall's connectivity indices,[11] as well as the more recently introduced indices of differences of path lengths, DPs.[12] These encode important information about electronic interferences in aromatic molecules. Let us suppose that a

*Corresponding author. Tel.: +34-6-3864891; fax: +34-6-386-48-92; e-mail: jorge.galvez@uv.es

0960-894X/00/$ - see front matter
PII: S0960-894X(00)00406-6

wave originated in the vertex i is moving through two different paths p_1 and p_2 to a neighbouring vertex j. In a quantical context, we can suppose that the considered wave can be the one associated to an electron and, therefore, it can produce a constructive interference in j. The kinetic energy of the electron results to be proportional to $(1/\Delta p_{12})_2$. In this way, we introduce the difference of path lengths indices DP as a parameter useful in QSAR studies. A DP index of order k is defined as:

$$DP_k = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\delta(k, D_{ij})}{(\Delta p_{ij})^2} \quad \text{and}$$

$$DP_k^v = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\delta(k, D_{ij})}{(\Delta p_{ij})^2}(1 + \Delta(EN_{ij}))$$

where $\Delta(EN_{ij})$ are the differences of electronegativities between the i and j vertices placed at a topological distance k.

Furthermore, SLDA is a useful technique to find discriminant functions with the ability to distinguish between two groups or populations. The SLDA competes nowadays with success in front of other techniques more complex as they are the neural networks.[13] The method used for descriptors selection with SLDA was based on the F-Snedecor parameter. The classification criterion used was the minimum value of Mahalanobis distance and the quality of the discriminant function is evaluated through the Wilk's U-statistical parameter. A large set of structurally heterogeneous compounds, with both antibacterial and non-antibacterial activity, has been analysed by SLDA. Each group was separated into training and test groups. By this method the discriminant function obtained can be validated. The function chosen was:

$$DF = 0.934\Delta^0\chi + 5.993DP_1^v - 3.635$$

$$N = 355;\ F = 98;\ U\,(\text{Wilk's}) = 0.56$$

where DP_1^v is the valence index of differences of path at distance 1, and $\Delta^0\chi$ is the difference between Kier's connectivity $^0\chi$ and the valence connectivity index $^0\chi^v$. $^0\chi$ takes into account the σ electrons, whereas $^0\chi^v$ the σ, π and lone pairs, therefore, $\Delta^0\chi$ would take into account the presence of π and lone pairs of electrons in the molecule.

A compound will be selected as antibacterial if DF >0 or as non-antibacterial if DF < 0. In the training group, 132 compounds out of 161 were classified correctly as actives (82% prediction) and 82 of 91 as inactives (90.1%). In the test group, 60 of 68 were classified as actives (88.2%) and 32 of 35 as inactives (91.4%). As may be seen, in both training and test groups, the overall accuracy is around 82% for active compounds and around 90% for inactive compounds. On the other hand, in the majority of cases, we work within a success probability higher than 80% (see Prob. (act.) and Prob. (inact.) in Table 1).

Table 1 shows the classification results obtained from the applying the DF to a representative set of compounds (an electronic copy can be supplied with the descriptors of all the compounds used in the SLDA study on request).

When the DF function is applied to the previous set of compounds, a pharmacological distribution diagram (PDD) can be constructed (see Fig. 1.) by representing the expectancy for each classification group into every interval of DF. In general, the expectancy[14] for a group A into a given interval x, is defined as:

$$E_A = \frac{\text{Percentage of A in x}}{(\text{Percentage of non-A in x} + 100)} \tag{4}$$

In our case, E_A = activity antibacterial expectancy; and E_I = inactivity antibacterial expectancy.

This diagram is another way describing pharmacological similarity, so we defined them as Topological Similarity Patterns. In spite of having used a large group of molecules, the profiles of PDD for both training and test groups, are very similar. The highest activity expectancy takes place for the DF >1.5. Furthermore, Figure 1 shows only a very small overlapping region, which is indicative of the discriminant power of DF function. In order to increase the discriminant power, the PDD was used, so the threshold for activity is determined by the overlapping interval.

After applying the DF function to different structures contained in our databases, we focused on the search and selection of natural compounds whose pharmaceutical development can be interesting due to their expected low toxicity. We have selected as theoretical active compounds those showing a value of DF >1.0. The viability of the method was confirmed by the adequate experimental antibacterial tests.

The method used for the assessment of in vitro antibacterial activity of the selected molecules is the Agar Diffusion Susceptibility Test,[15] which is based on the diffusion of a drug solution deposited on a filter-paper disc. The discs are transferred aseptically to Petri dishes containing agar media (Mueller–Hinton agar) uniformly seeded with a suspension in LB solutions of *Staphylococcus aureus, Escherichia coli, Pseudomonas aeruginosa,* or *Proteus mirabilis*. The imbibed antibacterial solution is left to diffuse through the agar during 2 h at 4 °C and the plate is incubated at 37 °C during 24–48 h. The antibacterial activity is determined by measuring the diameter of the growth inhibition zones around the discs after the incubation.

Measurements of minimum inhibitory concentrations (MIC) have been obtained by the Broth Dilution Method,[15] which is based on the growth inhibition of the mould in a uniform solution of the antibacterial agent in a fluid medium (Mueller–Hinton broth). For this, a serial dilution of antibacterial is inoculated with the microorganisms and incubated during 18–24 h at

Table 1. Results obtained applying the linear discriminant analysis (function DF) at a representative group of antibacterial and no antibacterial compounds

Training group									
Active compounds					Inactive compounds				
Compound	$\Delta^0\chi$	DP_I^v	Prob (act.)	Class	Compound	$\Delta^0\chi$	DP_I^v	Prob (inact.)	Class
Acetosulfone	3.338	0.109	0.534	+	Acifran	1.449	0.244	0.694	–
Apramycin	6.709	0.388	0.993	+	Acipimox	1.626	0.178	0.741	–
Bencylpenicillin	1.489	0.547	0.738	+	Aldicarb	0.637	0.000	0.954	–
Cefazolin	1.674	0.744	0.916	+	Allicin	0.220	0.000	0.969	–
Ceforanide	2.986	0.549	0.920	+	Altretamine	0.781	0.089	0.915	–
Cefotiam	2.096	0.744	0.942	+	Amefenac	1.059	0.059	0.908	–
Cefoxitin	2.510	0.566	0.891	+	Amitraz	0.390	0.059	0.949	–
Cephaloridine	1.282	0.785	0.906	+	Antipyrine	0.559	0.241	0.841	–
Dapsone	1.839	0.109	0.220	–	Antrafenine	3.061	0.366	0.195	+
Difloxacin	2.785	0.454	0.844	+	Azapicyl	1.142	0.119	0.865	–
Doxycycline	4.213	0.189	0.807	+	Azaserine	2.094	0.000	0.843	–
Fleroxacin	3.407	0.365	0.850	+	Benzene	0.000	0.000	0.974	–
Flumequine	1.903	0.266	0.435	–	Carpofen	0.855	0.272	0.770	–
Lenampicillin	2.554	0.791	0.970	+	Clofibrate	0.693	0.066	0.930	–
Mafenide	1.839	0.025	0.146	–	Clofibrate acid	0.947	0.066	0.913	–
Miloxacin	2.177	0.508	0.809	+	Difunisal	2.649	0.112	0.620	–
Moxalactam	4.011	0.479	0.952	+	Doxorubicin	4.980	0.378	0.036	+
Nifuradene	2.340	0.447	0.774	+	Benzoic acid	0.852	0.000	0.945	–
Norfloxacin	2.240	0.385	0.682	+	Buspirone	1.248	0.445	0.451	+
Ofloxacin	2.462	0.510	0.848	+	Canthaxanthin	0.598	0.069	0.935	–
Pipacycline	4.811	0.367	0.955	+	Etifoxine	0.432	0.209	0.879	–
Pipemidic acid	1.878	0.450	0.693	+	Fluoresone	1.616	0.064	0.851	–
p-Sulfanilylbencylamine	1.839	0.080	0.192	–	Glutamic acid	2.126	0.000	0.839	–
Salazosufafadimidine	3.126	0.267	0.708	+	Hydroxylphenamate	1.573	0.000	0.897	–
Sulbenicillin	3.035	0.547	0.923	+	Hydroxylzine	0.908	0.209	0.823	–
Sulfamoxole	2.052	0.291	0.507	+	Lovastatin	1.748	0.206	0.683	–
Talampicillin	2.255	0.791	0.961	+	Malaoxon	1.877	0.000	0.868	–
Tetroxoprim	2.301	0.183	0.404	–	Malathion	1.712	0.000	0.884	–
Ticarcillin	2.042	0.724	0.932	+	Nimustine	1.544	0.178	0.755	–
Tobramycin	3.524	0.172	0.666	+	Paracetamol	1.059	0.064	0.906	–
Test group									
Amoxicillin	2.464	0.581	0.896	+	AAS	1.449	0.034	0.889	–
Azlocillin	2.631	0.749	0.965	+	Apazone	1.118	0.360	0.607	–
Carumonam	5.037	0.551	0.988	+	Azacosterol	0.813	0.087	0.913	–
Cefoperazone	3.667	0.728	0.985	+	Carmustine	0.657	0.000	0.954	–
Cefpimizole	4.822	0.871	0.998	+	Flurbiprofen	1.474	0.039	0.883	–
Cinoxacin	2.008	0.478	0.751	+	Clorophene	0.349	0.066	0.949	–
Furaltadone	2.861	0.669	0.955	+	Emylcamate	1.020	0.000	0.936	–
Gentamicin	4.536	0.280	0.907	+	Meprobamate	2.041	0.000	0.849	–
Methacycline	4.213	0.189	0.807	+	Beclobrate	0.693	0.066	0.930	–
Netilmicin	4.536	0.280	0.907	+	Nicomol	3.464	0.509	0.066	+
Ribostamycin	6.203	0.326	0.984	+	Nicotine	0.982	0.119	0.881	–
Rosoxacin	1.411	0.287	0.355	–	Omethoate	1.187	0.000	0.926	–
Sulfoxone	3.550	0.109	0.583	+					
Sulguano	2.467	0.291	0.602	+					

37 °C. MIC is the minor drug concentration without visible growth of the bacterial.

The results are presented in Table 2. It is remarkable that hesperidin (predominant flavonoid in lemons and sweet oranges[16]) with MIC_{90} for *S. aureus* and *P. aeurginosa* 12 times lower than chloramphenicol (which is used as the reference drug) and the neohesperidin (sweetening agent, especially used in chewing gum and dentifrices[16]) with MIC_{90} for *P. mirabilis, S. aureus* and *P. aeurginosa* similar to chloramphenicol. Other active compounds were silymarine (antihepatotoxic principle isolated from the seeds of the milk thistle, *Slybum marianum*[16]), and the colorants Mordant Brown 24 and morine, with MIC_{90} within the same order of magnitude as chloramphenicol. Moreover, compounds such as amaranth (additive food) and fraxine (obtained from the horse chestnut tree[16]), showed an acceptable antibacterial activity for the four strains tested.

It is to be emphasised that no previous report on antibacterial activity was found in the literature for these compounds.

These results demonstrate that by an adequate choice of topological descriptors, it is possible to discriminate the antibacterial activity of a compound and, therefore, the usefulness of molecular topology in the search of these type of drugs is demonstrated as well. It is still necessary to carry out toxicological tests of the selected compounds

Table 2. Classification results for each compound selected and values of the MIC_{90} (mg/mL) obtained for the strains *P. mirabilis, E. coli, S. aureus* and *P. aeurginosa*[a]

Compound	Prob. (act.)	*P. mirabilis,* (CECT170)	*E. coli,* (CECT405)	*S. aureus* (CECT240)	*P. aeurginosa* (CECT108)
Chloramphenicol (reference drug)		0.10	0.01	0.10	0.10
Amaranto	0.670	1.80	1.80	1.80	1.80
Fraxine	0.869	1.45	1.45	1.45	1.45
Hesperidin	0.623	—	—	0.008	0.008
Mordant Brown 24	0.747	0.13	0.13	0.13	0.13
Morine	0.778	0.12	1.38	0.12	0.12
Niflumic acid	0.526	1.80	1.80	0.16	—
Neohesperidin	0.989	0.30	0.30	0.30	0.30
Silymarine	0.965	0.7	—	0.08	0.70

[a]CECT = Colección española de Cultivos Tipo (Dpto Microbiología. Univ. Valencia, España).

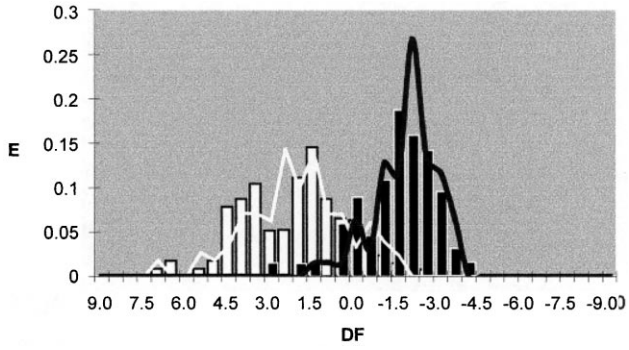

Figure 1. Pharmacological distribution diagram for antibacterial activity obtained by using the discriminant function DF (the bars represent the training group, and the lines, the test group; the white colour represents the compounds with antibacterial activity and the black colour, the compounds without it).

to permit their applicability to the antibacterial treatment in both animals and humans, and that will be developed in later works.

Acknowledgements

This study has been support by CICYT, SAF96-0158-C02-02 (Spanish Ministry of Culture and Science), GV99-91-1-12 (Generalitat Valenciana). The authors acknowledge Prof. Mishra (Sambalpur University, India) for careful reading of the manuscript and useful suggestions.

References and Notes

1. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure–Activity Analysis*; Research Studies Press; Letchworth, England, 1986; pp 225–246.
2. García-Domenech, R.; Villanueva, A.; Gálvez, J.; Gozalbes, R. *J. Chim. Phys.* **1999**, *96*, 1172.
3. De Julián-Ortiz, J. V.; García-Domenech, R.; Galvez, J. *J. Chromatogr.* **1996**, *719*, 37.
4. De Julián-Ortiz, J. V.; Gálvez, J.; Muñoz-Collado, C.; García-Domenech, R.; Gimeno-Cardona, C. *J. Med. Chem.* **1999**, *42* (no. 17), 3308.
5. Gálvez, J.; Gomez-Lechón, M. J.; García-Domenech, R.; Castell, J. V. *Bioorg. Med. Chem. Lett.* **1996**, *19*, 2301.
6. Estrada, E.; Peña, A.; García-Domenech, R. *J. Comput.-Aid. Mol. Design* **1998**, *12*, 583.
7. Gálvez, J.; Garcia-Domenech, R.; De Julian-Ortiz, J. V.; Soler, R. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1198.
8. Pastor, L.; García-Domenech, R.; Wolski, S.; García, M. D.; Gálvez, J. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2577.
9. Ríos-Santamarina, I.; García-Domenech, R.; Gálvez, J. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 477.
10. Dixon, W. J.; Brown, M. B.; Engelman, L.; Jennrich, R. I. BMDP Statistical Software Manual, Vol. 1, University of California Press, 1990; pp 339–358.
11. Kier, L. B.; Hall, L. H. *Eur. J. Med. Chem.* **1977**, *12*, 313.
12. Gálvez, J.; García-Domenech, R.; De Gregorio Alapont, C. *J. Comput.-Aid. Mol. Design.* **2000** (in press).
13. Aoyama, T.; Suzuku, Y.; Ichikawa, H. *J. Med. Chem.* **1990**, *33*, 2583.
14. Gálvez, J.; García-Domenech, R.; De Gregorio-Alapont, C.; De Julian-Ortiz, J. V.; Popa, L. *J. Mol. Graphics* **1996**, *14*, 272.
15. Lennette, E. H.; Balows, A.; Hausler, W. J.; Shadomy, H. J. *Manual de Microbiología Médica*, Panamericana, 4th ed., 1987; pp 1191–1223.
16. *The Merck Index*, 12th ed, Merck & Co. Inc., Whitehouse Station, 1996.