Inter*net* Electronic Journal of Molecular Design

July 2002, Volume 1, Number 7, Pages 339–350

Editor: Ovidiu Ivanciuc

Special issue dedicated to Professor Milan Randić on the occasion of the 70th birthday
Part 3

Guest Editor: Mircea V. Diudea

Predicting Antifungal Activity: A Computational Screening Using Topological Descriptors

Ramón García–Domenech,[1] Ana Catalá–Gregori,[2] Carmen Calabuig,[3]
Gerardo Antón–Fos,[2] Lucas del Castillo,[4] and Jorge Gálvez[1]

[1] Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular. Dept. Quimica–Física, Facultad de Farmacia, Universitat de Valencia, 46100 Burjassot, Valencia, Spain
[2] Dept. Química, Bioquímica y Biología Molecular, Universidad Cardenal Herrera–CEU, 46113 Moncada, Valencia, Spain
[3] Dept. Bioquímica y Biología Molecular, Facultad de Farmacia Universitat de Valencia, Spain
[4] Dept. Microbiología y Ecología, Facultad de Farmacia Universitat de Valencia, Spain

Received: March 27, 2002; Revised: June 21, 2002; Accepted: July 3, 2002; Published: July 31, 2002

Citation of the article:
R. García–Domenech, A. Catalá–Gregori, C. Calabuig, G. Antón–Fos, L. del Castillo, and J. Gálvez, Predicting Antifungal Activity: A Computational Screening Using Topological Descriptors, *Internet Electron. J. Mol. Des.* **2002**, *1*, 339–350, http://www.biochempress.com.

Inter*net* Electronic Journal of Molecular Design

BioChem Press
http://www.biochempress.com

Predicting Antifungal Activity: A Computational Screening Using Topological Descriptors[#]

Ramón García–Domenech,[1,*] Ana Catalá–Gregori,[2] Carmen Calabuig,[3] Gerardo Antón–Fos,[2] Lucas del Castillo,[4] and Jorge Gálvez[1]

[1] Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular. Dept. Quimica–Física, Facultad de Farmacia, Universitat de Valencia, 46100 Burjassot, Valencia, Spain
[2] Dept. Química, Bioquímica y Biología Molecular, Universidad Cardenal Herrera–CEU, 46113 Moncada, Valencia, Spain
[3] Dept. Bioquímica y Biología Molecular, Facultad de Farmacia Universitat de Valencia, Spain
[4] Dept. Microbiología y Ecología, Facultad de Farmacia Universitat de Valencia, Spain

Received: March 27, 2002; Revised: June 21, 2002; Accepted: July 3, 2002; Published: July 31, 2002

Internet Electron. J. Mol. Des. **2002**, *1* (7), 339–350

Abstract

Motivation. Recently, the opportunistic fungal infections have risen dramatically. One of the biggest problems facing nowadays the antifungal therapy is the arising of drugs resistance strains for most of the drugs currently used in clinic practice. Therefore, it is important to find new antifungal candidate compounds, particularly new leads, able to become the basis for developing new drugs.

Method. Molecular topology, a formalism based on describing the molecules as hydrogen–depleted graphs, as well as linear discriminant analysis, a statistical tool able to distinguish between two or more categories or objects, have been used to design new antifungal compounds.

Results. A topological–mathematical model comprised of two discriminant functions have been developed. The model is able to classify correctly 98.5% of the inactive compounds from the training set. The model validation was performed in two ways. The first one was to check the literature sources available to confirm the predicted antifungal activity. In the second test, for those compounds not found in the literature, experimental tests were performed to check the antifungal activity. A set of four compounds was selected to be tested in the laboratory against *C. albicans*, *C. glabrata* and *S. cerevisiae*. All compounds, namely anethole, 2–methyl–4,5–diphenyloxazole, 2–mercaptobenzoxazole and β–naphtyl caproate, showed activity against the three species with MIC_{50} ranging between 25 and 100 μg/mL.

Conclusions. The results confirm other previous results from our group, regarding the usefulness of molecular graphs and topological indices as effective tools to discover new antifungal compounds, especially new leads.

Keywords. Linear discriminant analysis; antifungal activity; topological indices; molecular graph.

[#] Dedicated to Professor Milan Randić on the occasion of the 70th birthday.

* Correspondence author; phone: 34–963–864–891; fax: 34–963–864–892; E–mail: ramon.garcia@uv.es.

1 INTRODUCTION

Infections due to *Candida* species are the most common of the fungal infections [1]. *Candida* species produce a broad range of infections, ranging from non life threatening muco–cutaneous illnesses to invasive process that may involve virtually any organ [2]. The frequency of opportunistic infections caused by the fungus *Candida albicans* is very high and is expected to continue to increase as the number of immuno–compromised patients rises [3]. Additionally, until recently, *Candida glabrata* was considered a relatively nonpathogenic fungal organism of the human mucosal tissues.

However, with the increased use of immunosuppressive agents, mucosal and systemic infections caused by *C. glabrata* have increased significantly, especially in the human immunodeficiency virus–infected population [4]. Antifungal drugs currently being used in clinic include polyene antibiotics, azole derivatives and 5–fluorocytosine. However, one significant obstacle preventing successful antifungal therapy is the dramatic increase in drug resistance, especially against azole antimycotics [1].

It is urgent, therefore, to find new chemical structures with a wide antifungal spectrum that can be taken as a starting point for the development and optimization of new antifungal agents. The most successful set of molecular descriptors for drug design, QSAR, and classification of chemical libraries are the molecular connectivity indices [5–7]. These indices are based on a graph–theoretical invariant introduced by Randić 25 years ago in order to compute a branching index for alkanes [8].

Topological indices have demonstrated their utility in the prediction of diverse physical, chemical and biological properties for different types of compounds [9–12]. Their utility was recently demonstrated in the design of new antivirals [13], cytostatics [14], sedative/hypnotics [15], analgesics [16], antihistaminics [17], anticancer [18], antimicotics [19], antimalarials [20], bronchodilators [21].

In a recent paper, we have applied molecular topology to identify a QSAR model able to predict the antifungal activity [22]. Using this QSAR model, several compounds were selected as potentially active. Some of them, such as benztropine mesylate and diclopentamethylenthiuram disulfide, showed activity against *C. albicans* similar to miconazole, the reference drug. The overall accuracy in the selection of new candidates was about 60%.

In this paper we aim to improve the former results, using new topological indices as well as applying the discriminant analysis only to synthetic compounds showing antifungal activity (the antibiotics are taken out from the study).

2 MATERIALS AND METHODS

2.1 Chemical Data

The set of antifungal compounds included into the training set for the linear discriminant analysis was collected from the Merck Index [23], and contains allylamines, imidazoles, thiocarbamates, triazoles and other derivatives. The inactive group was also gathered from the Merck Index from among those drugs not referenced to show antifungal activity. The set of compounds is presented in Table 4.

2.2 Structural Descriptors

In order to characterize the structure of each compound, a set of graph–theoretical descriptors was calculated for each compound by using Hall's MOLCONN–Z [24], and Desmol13 [25] programs. The indices calculated belong to the following families (Table 1): Randić–Kier–Hall subgraph connectivity indices $^{m}\chi t$ [26] up to order four, and their corresponding valence indices [27], topological charge indices [28], topological constitutional indices [29], kappa indices [30], atom type E–state indices [31] and Wiener index [32].

Table 1. Symbols for Topological Indices and Their Definitions

Symbol	Definition	Reference
$^{m}\chi_p$	Path connectivity index of order m = 0 to 4	[26]
$^{m}\chi^{v}_p$	Path valence connectivity index of order m = 0 to 4	[27]
$^{m}\chi_c$	Cluster connectivity index of order m = 0 to 4	[26]
$^{m}\chi^{v}_c$	Cluster valenceconnectivity index of order m = 0 to 4	[27]
$^{m}\chi_{pc}$	Path–cluster connectivity index of order m = 0 to 4	[26]
$^{m}\chi^{v}_{pc}$	Path–cluster valence connectivity index of order m = 0 to 4	[27]
$\Delta^{m}\chi_i$	$^{m}\chi_i - {}^{m}\chi_i^{v}$ of order m = 0 to 4	[27]
$C^{m}\chi_i$	$^{m}\chi_i / {}^{m}\chi_i^{v}$ of order m = 0 to 4	[27]
G_m	Topological charge index of order m = 1 to 5	[28]
J_m	Bond topological charge index of order m = 1 to 5	[28]
G^{v}_m	Valence topological charge index or order m = 1 to 5	[28]
J^{v}_m	Valence bond topological charge index of order m = 1 to 5	[28]
κ_i	Kappa indices	[30]
$S^{T}i$	Atom type E–state indices	[31]
V_3	Number of vertices of degree 3	[29]
W	Wiener path number	[32]
Numhbd	Number of hydrogen bond donors	[24]
Numhba	Number of hydrogen bond acceptors	[24]
Pri	Number of vertices pair of degree 3 at *i*-th topological distance	[29]

2.3 Linear Discriminant Analysis

The objective of the linear discriminant analysis, LDA, which is considered as a heuristic algorithm able to distinguish between two or more categories or objects, is to use the structural descriptors and find linear functions able to discriminate between the active and inactive compounds. One set was made by the antifungal drugs, while the second set contained the inactive

compounds. The discriminant ability is evaluated by the percentage of correct classifications into each group.

LDA was performed by using the BMDP 7M package [33]. The selection of the descriptors was based on the F Snedecor parameter, and the classification criterion was the shortest Mahalanobis distance (distance of each compound to the mean of all compounds used in the regression equation). The software 7M chooses the variables used in computing the linear classification functions in a stepwise manner, *i.e*. at each step the variable that adds the most to the separation of the groups is entered into (or the variable that adds the least is removed from) the discriminant function. The quality of the discriminant function is evaluated by the Wilk's lambda parameter λ, which is a multivariate analysis of variance statistic that tests the equality of group means for the variable(s) in the discriminant function.

2.4 Pharmacological Tests of Antifungal Activity

The viability of the topological–mathematical models used to search and select new compounds with antifungal activity was confirmed by the adequate experimental microbiological tests. The *in vitro* sensibility tests were performed with three species: *Candida albicans* SC 5314, *Candida glabrata* and *Sacharomyces cerevisiae* X21801A. The drug concentration able to inhibit the cellular growth on at least 50% (MIC_{50}) was calculated by a liquid medium dilution method [35] YNB (glucose 20 g, $YNB_{W/O\ SO4(NH4)2}$ 1.67 g, $SO_4(NH_4)_2$ 5 g in a liter of distilled water).

2.4.1 Inoculate Preparation

The strains were prepared with an YPD solid medium and incubated at 28 ºC during 24 hours. From each culture, a 1 mm diameter colony was taken and suspended into tubes containing 5 mL YPD medium (Yeast extract 10 g, peptone 20 g, glucose in water 20 g, per liter), up to stationary phase.

2.4.2 Antifungal dilutions

The testing compounds were provided by Sigma. A 8:2 V/V mixture ethanol/DMSO was used as solvent because of the water insolubility of the tested products. A set of sterile test tubes containing 5 mL each of culture medium was prepared. From a standard 10 mg/mL solution, the corresponding dilutions were prepared to reach the concentrations of 0.5, 5, 10, 25, 50, 75, 100 and 125 μg/mL on the test tubes. As a control, a test tube containing just medium and solvent was used.

2.4.3 Inoculation

5 μL of the cells suspension was added into each tube. After stirring they were left for 24 hours at 28 ºC. In order to evaluate either the growth or the inhibition of cell cultures, cell concentration was determined by counting the cells on a Neubauer chamber. Miconazole was used as reference.

3 RESULTS AND DISCUSSION

In this work, an SAR model based on molecular topology was developed to predict new molecules with antifungal activity. The set of topological indices was calculated for every molecule into the training set (*i.e.* the set of molecules used to get the predictive model) and linear discriminant analysis (LDA) was applied to find the best discriminant functions. LDA was performed on two set of compounds: The first was formed only by antifungal drugs, and the second one of molecules not described as showing antifungal activity and therefore expected to be inactive. Furthermore, two different sets of topological indices were used separately in the discriminant analysis. One of them, DF_1, included connectivity–like indices, topological charge indices and other in house molecular descriptors accounting for molecular shape constitutional indices, while the second, namely DF_2, contained other indices, mainly the electrotopological indices. Tables 2 and 3 show the coefficients for each function as well as the overall classification within each group. As may be realized, both models are rather similar (about 80% of correct classification within the active and over 90% within the inactive set).

Table 2. Discriminant Function and Classification Matrix Obtained by Linear Discriminant Analysis on Antifungal Activity Study using Connectivity and Charge Descriptors

Discriminant function DF_1			Classification matrix		
Variable	Coefficient	Group	% Correct	Active	Inactive
Constant	10.631				
$^4\chi_{pc}$	–1.455	Active	83.3	35	7
G_2	1.000				
G_2^v	–0.620	Inactive	90.4	5	47
$\Delta^1\chi$	–6.912				
$C^2\chi$	–18.134	Total	87.2	40	54
$\Delta^4\chi_p$	7.385				
$C^4\chi_{pc}$	7.573		λ (Wilks' Lambda) = 0.392		
PR1	–0.864		Approximate F–Statistic = 14.5		
V_3	0.269				

Table 3. Discriminant Function and Classification Matrix Obtained by Linear Discriminant Analysis on Antifungal Activity Study using Electrotopological Descriptors

Discriminant function DF_2			Classification matrix		
Variable	Coefficient	Group	% Correct	Active	Inactive
Constant	1.511				
S^T(aaaC)	1.673	Active	81.0	34	8
S^T(aNHa)	–1.835				
S^T(=N–)	0.747	Inactive	94.2	3	49
S^T(=S)	1.152				
S^T(Cl)	0.195	Total	88.3	37	57
W	–0.076		λ (Wilks' Lambda) = 0.387		
Numhbd	–1.045		Approximate F–Statistic = 19.5		

Table 4 illustrates the probability and *DF* values obtained for each of the discriminant functions and compounds studied. For most of the cases, the classification probability for every drug is over 90%, a fact that clearly supports the quality of the discrimination achieved.

Table 4. Classification Results Obtained from the Linear Discriminant Analysis

Active	Prob.	DF_1	Prob.	DF_2	Class	Inactive	Prob.	DF_1	Prob.	DF_2	Class
Bifonazole	0.96	3.22	0.99	5.06	+	Alclofenac	0.85	–1.69	0.50	0.01	–
Biphenamine	0.04	–3.11	0.48	–0.09	–	Ambroxol	0.47	0.12	0.99	–5.03	–
Buclosamide	0.87	1.86	0.24	–1.15	–	Amiodarone	0.82	–1.52	0.59	–0.35	–
Butenafine	0.45	–0.21	0.93	2.55	–	Amrinone	0.54	–0.15	0.97	–3.53	–
Butoconazole	0.99	4.20	0.99	4.88	+	Apomorphine	1.00	–5.86	0.80	–1.38	–
Chlordantoin	0.62	0.48	0.80	1.38	+	Aprobarbital	0.87	–1.87	0.95	–2.87	–
Chlormidazole	0.99	4.43	0.96	3.25	+	Atenolol	0.93	–2.65	0.99	–4.50	–
Chlorphenesin	0.55	0.20	0.32	–0.77	–	Azathioprine	0.99	–4.62	0.96	–3.26	–
Ciclopirox	0.58	0.31	0.16	–1.67	–	Bambuterol	1.00	–6.54	0.97	–3.48	–
Cloconazole	0.99	4.90	0.93	2.58	+	Captopril	0.64	–0.58	0.81	–1.44	–
Clotrimazole	1.00	6.03	0.83	1.61	+	Carbamazepine	0.32	0.76	0.97	–3.55	–
Cloxyquine	0.37	–0.52	0.88	2.00	–	Carbidopa	0.99	–5.08	1.00	–6.82	–
Diamthazole	0.05	–2.86	0.93	2.54	–	Carteolol	0.94	–2.68	0.99	–4.22	–
Econazole	1.00	6.27	0.99	4.84	+	Carvacrol	0.64	–0.56	0.70	–0.83	–
Enilconazole	0.94	2.76	0.99	4.44	+	Cetirizine	0.98	–3.68	0.87	–1.93	–
Exalamide	0.81	1.43	0.10	–2.26	–	Citalopram	0.73	–1.02	0.91	–2.30	–
Fenticonazole	1.00	7.80	0.94	2.83	+	Clonidine	0.18	1.54	0.05	3.04	+
Fluconazole	0.91	2.26	1.00	8.68	+	Codeine	0.99	–4.64	0.99	–4.72	–
Flucytosine	0.97	3.40	0.42	–0.33	–	Cotinine	0.96	–3.09	0.60	–0.40	–
Flutrimazole	0.99	4.19	0.51	0.06	+	Deflazacort	1.00	–6.42	0.94	–2.76	–
Halethazole	0.86	1.82	1.00	6.27	+	Dopamine	0.53	–0.11	0.98	–3.89	–
Hexetidine	0.69	0.78	0.04	–3.17	–	Doxepine	0.83	–1.56	0.83	–1.62	–
Isoconazole	0.99	5.15	1.00	5.80	+	Fenproporex	0.39	0.46	0.70	–0.83	–
Itraconazole	1.00	6.35	0.99	5.15	+	Flumetramide	1.00	–6.93	0.86	–1.82	–
Ketoconazole	1.00	5.34	0.79	1.34	+	Guanabenz	0.95	–2.87	0.30	0.83	–
Lanoconazole	0.45	–0.21	0.93	2.58	–	Indapamide	0.97	–3.44	0.99	–4.30	–
Loflucarban	1.00	6.33	1.00	5.34	+	Ketoprofen	0.95	–3.01	0.89	–2.13	–
Miconazole	1.00	6.04	1.00	5.82	+	Labetalol	0.92	–2.43	1.00	–6.69	–
Naftifine	0.84	1.68	0.95	2.99	+	Lacidipine	1.00	–6.51	0.98	–4.03	–
Nifuratel	1.00	5.31	0.87	1.90	+	Lisinopril	0.99	–4.43	1.00	–7.37	–
Omoconazole	1.00	5.52	0.99	4.39	+	Methoxamine	1.00	–5.59	0.97	–3.53	–
Oxiconazole	1.00	5.65	1.00	8.88	+	Metoclopramide	0.99	–4.92	0.95	–2.90	–
Saperconazole	1.00	5.22	0.92	2.49	+	Mianserin	0.79	–1.31	0.89	–2.07	–
Sertaconazole	0.99	4.24	1.00	7.65	+	Nandrolone	1.00	–6.47	0.97	–3.58	–
Sulconazole	1.00	5.21	0.99	4.90	+	Norepinephrine	0.95	–2.89	0.99	–5.16	–
Tenonitrozole	0.90	2.19	0.16	–1.67	–	Norgestrel	1.00	–5.60	0.99	–4.34	–
Terconazole	0.98	4.03	0.99	4.42	+	Omeprazole	0.76	–1.16	1.00	–6.44	–
Tioconazole	1.00	6.63	0.99	4.87	+	Oxymetazoline	0.96	–3.05	0.52	–0.09	–
Tolciclate	0.47	–0.11	0.98	3.86	–	Pentazocine	0.96	–3.11	0.96	–3.19	–
Tolindate	0.90	2.23	0.99	4.62	+	Praziquantel	0.94	–2.81	0.92	–2.46	–
Tolnaftate	0.95	2.90	1.00	8.43	+	Salmeterol	0.84	–1.65	1.00	–5.95	–
Ujothion	0.36	–0.57	0.99	4.37	–	Simvastatin	0.94	–2.74	0.98	–4.03	–
						Sotalol	0.99	–5.08	0.97	–3.45	–
						Sulpiride	1.00	–7.79	0.99	–4.83	–
						Tamoxifen	0.68	–0.73	0.91	–2.30	–
						Terazosin	1.00	–5.42	0.95	–2.87	–
						Terbutaline	0.94	–2.68	0.99	–4.35	–
						Terfenadine	0.40	0.41	1.00	–5.92	–
						Timolol	0.98	–4.07	0.96	–3.10	–
						Tramadol	0.92	–2.48	0.91	–2.28	–
						Verapamil	1.00	–5.59	0.96	–3.14	–
						Xipamide	0.97	–3.59	0.99	–4.90	–

Moreover, although the quality of both discriminant functions is similar, their joint use leads to a very significant improvement for the inactive set. As long as the goal is to minimize the risk of

selecting false active candidates, this is the best strategy. Indeed, while the percentage of accuracy gets down up to 60% within the actives, that percentage increases up to 98% for inactive. In theory, this would mean that the probability to select a false active is just 1/100. Anyway, it should work as a very good filter for the selection/design of new drug candidates. Table 4 (the columns six and twelve) illustrates the results obtained for each one of the analyzed compounds. The only one false active selected by the model is clonidine, which is supposed to be inactive just because no literature references were found on its activity. Though experimental tests should be carried out to confirm its inactivity, still the overall accuracy of the inactive set classification is 51/52, namely 98%.

The stability of the discriminant functions DF_1 and DF_2 was tested taking random test sets having 20% of the compounds, using the same variables and performing new classification analyses. As outlined in Table 5, the models were stable on random tests as well as fully reproducible.

Table 5. Classification Matrix Obtained in the Random Test using DF_1 and DF_2

		DF1			DF2	
	Correct %	Actives	Inactives	Correct %	Actives	Inactives
Non–random						
Actives	83.3	35	7	81.0	34	8
Inactives	90.4	5	47	94.2	3	49
Test actives	–	–	–	–	–	–
Test inactives	–	–	–	–	–	–
Run 1	λ=0.381	F=11.7		λ=0.39	F=15	
Actives	91.2	31	3	81.8	27	6
Inactives	92.7	3	38	93.2	3	41
Test actives	87.5	7	1	88.9	8	1
Test inactives	72.7	3	8	87.5	1	7
Run 2	λ=0.38	F=12		λ=0.41	F=14	
Actives	87.9	29	4	77.1	27	8
Inactives	95.1	2	39	95.6	2	43
Test actives	77.8	7	2	85.7	6	1
Test inactives	81.8	2	9	100.0	0	7
Run 3	λ=0.40	F=11		λ=0.39	F=15	
Actives	85.3	29	5	85.3	29	5
Inactives	91.1	4	41	91.1	4	41
Test actives	100	8	0	87.5	7	1
Test inactives	85.7	1	6	71.4	2	5
Run 4	λ=0.39	F=11		λ=0.40	F=14	
Actives	83.9	26	5	81.6	31	7
Inactives	93.3	3	42	95.0	2	38
Test actives	81.8	9	2	75.0	3	1
Test inactives	85.7	1	6	91.7	1	11
Run 5	λ=0.39	F=10		λ=0.40	F=14	
Actives	79.3	23	6	84.8	28	5
Inactives	92.5	3	37	92.7	3	38
Test actives	84.6	11	2	77.8	7	2
Test inactives	100	0	12	100.0	0	11
Average						
Actives	85.7	138	23	82.1	142	31
Inactives	92.9	15	197	93.5	14	201
Test actives	85.7	42	7	83.8	31	6
Test inactives	85.4	7	41	91.1	4	41

Figure 1 illustrates the pharmacological distribution diagrams with the expectancy profile for both active and inactive compounds for each interval of DF_1 and DF_2. In general, expectancy for a set *A* along each interval *x* is defined as: *E* = Percentage of *A* in *x* / (Percentage of non–*A* in *x* + 100) [34]. Despite the presence of an overlapping region, it is remarkable that for the values $7 > DF_1 > 1$ and $9 > DF_2 > 0$, most of the active compounds are located within while virtually no inactive compound is found. Thus, within these intervals, the classification was considered optimal.

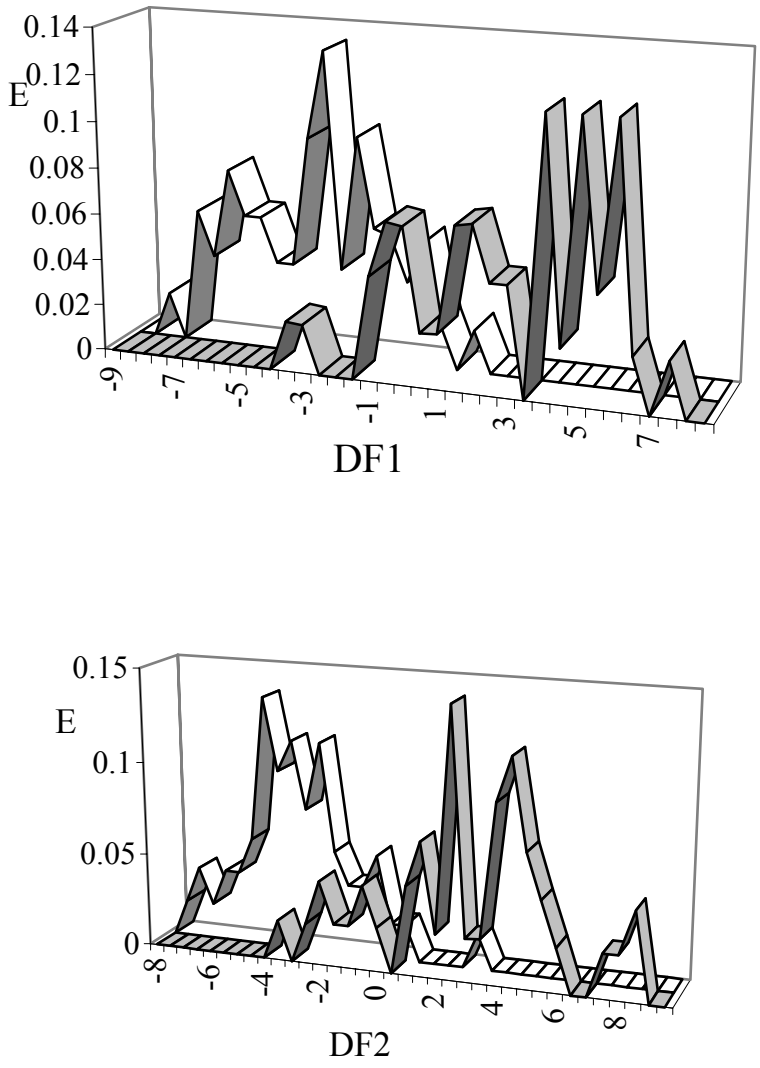


Figure 1. Pharmacological distribution diagrams for antifungal activity from the DF_1 and DF_2 discriminant functions (white lines: inactive drugs; black lines: active drugs).

The selected discriminant function, DF_1 includes the topological descriptors PR_1 (pairs of ramifications located at a topological distance 1) and V_3 (number of vertices with topological valence 3). Both terms encode pure structural information. The topological assembly is defined basically by combinations of connectivity indices such as $\Delta^1\chi$, $\Delta^4\chi_p$, $C^2\chi$ and $C^4\chi_p$. Finally, the G_2

and G_2^v indices are the topological charge indices, which evaluate intramolecular charge transfers between atoms situated at a topological distance $i = 2$. They take into account molecular properties such as dipole moments and electronic polarizability [28]. All these topological indices can be considered as global molecular descriptors as they describe the whole structure of molecules instead of a substructure or region inside it.

The DF_2 discriminant function is essentially comprised of descriptors that can be considered as local molecular descriptors. The E–state index for an atom in a molecule represents the accessibility of that atom. It is a combination of electron richness or deficiency together with topological accessibility [36]. In this work we have used the atom type E–state indices, S^T_i.

Table 6. Test Compounds (Compounds Not Included in the Training of the Discriminant Functions)

Compound	DF_1	DF_2	Activity	References
Aliconazole	7.0	4.9	Antifungal	Clin Microbiol Rev 1988 Apr;1(2):187–217
Allicin	5.4	1.0	Antifungal	Microbes Infect 1999 Feb;1(2):125–9
Azaconazole	3.0	6.8	Antifungal	Toxicology 1985 Jan;34(1):1–11
Benzoxazole	2.8	6.1	Antifungal	Acta Biochim Pol 2000;47(2):481–6
Butyrolactone	5.4	1.0	Antifungal	J Antibiot (Tokyo) 1997 Sep;50(9):742–9
Broxaldine	1.0	4.0	Antifungal	J Assoc Physicians India 1973 Mar;21(3):295–8
Climbazole	4.8	2.9	Antifungal	Mycoses 1996 Jul–Aug;39(7–8):309–12
Chloropicrin	4.7	3.8	Fungicide	Appl Environ Microbiol 2001 Jul;67(7):3245–57
Chlorpyrifos	1.3	8.7	Pesticide	Fresenius J Anal Chem 2001 Dec;371(8):1134–8
Clofentezine	0.2	4.4	Pesticide	J Chromatogr A 1998 Oct 9;823(1–2):11–6
Croconazole	4.9	2.6	Antifungal	Clin Microbiol Rev 1988 Apr;1(2):187–217
Crotonic acid	10.4	0.2	Antifungal	J Med Chem 1976 Aug;19(8):1069–72
Dichlofenthion	3.7	7.8	Herbicide	Environ Sci Technol 2001 Jan 15;35(2):398–405
Furaspor	1.4	0.3	Antifungal	J Med Assoc Ga 1967 Jul;56(7):286–8
Halothane	0.0	0.9	Anest/Antifun	Minerva Med 1986 Nov 10;77(42–43):2007–10
Lombazole	3.1	6.0	Antifungal	Antimicrob Agents Chemother 1986 Aug;30(2):238–44
Oxadiazon	2.2	4.5	Herbicide	Water Res 2002 Jan;36(1):315–29
Surecide	5.1	5.3	Insecticide	Environ Sci Health B 1987 Apr;22(2):149–70
1;1;2;2–tetrachloroethane	2.4	3.1	Antifungal	Teratog Carcinog Mutagen 1989;9(6):349–57
Valconazole	4.7	3.9	Antifungal	Am J Obstet Gynecol 1991 Oct;165(4Pt2):1200–6
Vinclozolin	2.9	0.9	Fungicide	J Agric Food Chem 1999 Aug;47(8):3372–80
Anethole	1.6	0.4	Anestesic	Planta Med 2001 Aug;67(6):564–6
Brompheniramine	1.1	2.5	Antihistaminic	J Ocul Pharmacol 1994 Winter;10(4):665–75
Clemizole	4.6	2.7	Antihistaminic	J Auton Pharmacol 1999 Oct;19(5):281–9
β–naphtyl caproate	2.3	3.1	reactive	Aldrich Catalogue
2–mercaptobenzoxazole	4.4	4.5	reactive	Aldrich Catalogue
Zindotrine	2.1	9.6	Bronchodilator	Drug Intell Clin Pharm 1988 Oct;22(10):760–3
Lotifazole	6.5	7.4	AINE	Agents Actions 1984 Jan;14(1):93–101
2–methyl–4,5–diphenoxazole	3.6	8.6	reactive	Aldrich Catalogue

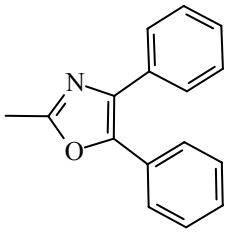
Although all these results are interesting it still remains the test on the applicability of the model to the search of new antifungals, particularly new leads. The only way to check this applicability is through an external validation by applying the discriminant functions for compounds not included in the training set. The predicted activity of these compounds is checked either by experimental

tests or by literature sources. In our case, we followed a two steps process. In the first step, a computational screening (using the antifungal topological–mathematical model, DF_1 and DF_2 functions) was applied to our databases. A set of compounds classified as highly probable antifungals was selected. The results are outlined in Table 6. Many of them, such as allicin, benzoxazole, climbazole, were described as antifungals in the literature, which is a very encouraging result. However, other compounds showed pharmacological activities, for instance anethole (anestesic and carminative) [23]. It would be of great interest to test it as antifungal.

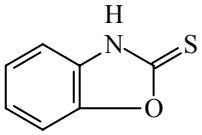
Table 7. Minimal Inhibitory Concentration MIC_{50} for Each Compound and Strain Assayed

Compound	*C. albicans*	*C. glabrata*	*S. cerevisiae*
		MIC_{50} (μg/mL)	
2–methyl–4,5–diphenyloxazole	25	50	125
2–mercaptobenzoxazole	25	25–50	25
Beta–naphtyl caproate	50	75–100	>100
Anethole	75	25–50	25
Miconazole (reference drug)	<0.5	0.5	0.5

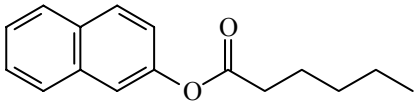
2–methyl–4,5–diphenyloxazole



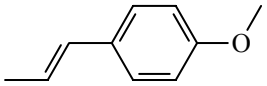
2–mercaptobenzoxazole



β–naphtyl caproate



Anethole



In the second step, four compounds were selected for experimental tests on three highly representative fungal species: *C. albicans*, *C. glabrata* and *S. cerevisiae*. The compounds were: 2–methyl–4,5–diphenyloxazole, 2–mercaptobenzoxazole, β–naphtyl caproate and anethole. After the experimental test, all of them exhibited antifungal activity albeit far away from the reference drug (miconazol). In fact, while miconazol has MIC_{50} values below 0.5 μg/mL, the selected compounds have MIC_{50} values between 25 and 125 μg/mL (see Table 7). Although the activity level of these four compounds is not close to that of the miconazol, a clear antifungal activity was found for every candidate. We are looking forward at improving these results in the future by finding more potent candidates. However, our current results are significant because they demonstrate the straightforward way in which molecular topology can identify new drug candidates, particularly new antifungals.

4 CONCLUSIONS

Molecular topology has demonstrated to be a useful methodology for identifying new compounds with antifungal activity. In this paper, a topological–mathematical model comprising two discriminant functions have been developed. The SAR model is able to classify correctly 98.5% of the inactive compounds from the training set. The SAR model validation was performed in two ways. The first one was performed by checking the literature sources available to confirm the predicted antifungal activity. In the second test, those compounds not found in the literature sources were experimentally tested for the antifungal activity. Four compounds were selected to be tested in the laboratory against *C. albicans*, *C. glabrata* and *S. cerevisiae*. All compounds, namely anethole, 2–methyl–4,5–diphenyloxazole, 2–mercaptobenzoxazole and β–naphtyl caproate, showed activity against the three species with MIC_{50} ranging between 25 and 100 μg/mL. The results confirm other previous results from our group, regarding the usefulness of molecular graphs and topological indices as effective tools in descovering new antifungal compounds.

Acknowledgment

The authors acknowledge the financial support of this research by Generalitat Valenciana (GV2001–047) and Spanish Ministry of Science and Technology (SAF2000–0223–C03–02).

5 REFERENCES

[1] V. St–Georgiev, Membrane transporters and antifungal drug resistance, *Curr. Drug Targets* **2000**, *1*, 261–84.
[2] J. H. Rex, T. J. Walsh, J. D. Sobel, S. G. Filler, P. G. Pappas, W. E. Dismukes, and J. E. Edwards, Practice guidelines for the treatment of candidiasis. Infectious Diseases Society of America, *Clin. Infect. Dis*. **2000**, *30*, 662–678.
[3] M. D. De–Backer, P. T. Magee, and J. Pla, Recent developments in molecular genetics of *Candida albicans*, *Annu. Rev. Microbiol*. **2000**, *54*, 463–498.
[4] P. L. Fidel, J. A. Vazquez, and J. D. Sobel, *Candida glabrata*: review of epidemiology, pathogenesis, and clinical disease with comparison to *C. albicans*, *Clin. Microbiol. Rev*. **1999**, *12*, 80–96.
[5] L. B. Kier and L. H. Hall, *Molecular Connectivity in Chemistry and Drugs Research*, Academic Press, New York, 1976.
[6] L. B. Kier and L. H. Hall, *Molecular Connectivity in Structure–Activity Analysis*, Research Studies Press, Letchworth, 1986, pp. 20–25.
[7] A. T. Balaban, Chemical Graphs. Part 48. Topological Index J for Heteroatom–Containing Molecules Taking Into Account Periodicities of Element Properties, *MATCH* (*Commun. Math. Chem.*) **1986**, *21*, 115–122.
[8] M. Randić, On Characterization of Molecular Branching, *J. Am. Chem. Soc*. **1975**, *97*, 6609–6615.
[9] R. García–Domenech, A. Villanueva, J. Gálvez, and R. Gozalbes, Aplication de la topologie moléculaire à la prédiction de la viscosité liquide de composés organiques, *J. Chim. Phys*. **1999**, *96*, 1172–1185.
[10] O. Ivanciuc, Building–Block Computation of the Ivanciuc–Balaban Indices for the Virtual Screening of Combinatorial Libraries, *Internet Electron. J. Mol. Des.* **2002**, *1*, 1–9, http://www.biochempress.com.
[11] R. García–Domenech, J. V. de Julián–Ortiz, M. J. Duart, J. M. García–Torrecillas, G. M. Antón–Fos, I. Rios–Santamarina, C. De Gregorio–Alapont, and J. Gálvez, Search of topological pattern to evaluate the toxicity of heterogeneous compounds, *SAR QSAR Environ. Res*. **2001**, *12*, 237–254.
[12] A. A. Toropov and A. P. Toropova, QSAR Modeling of Mutagenicity Based on Graphs of Atomic Orbitals, *Internet Electron. J. Mol. Des.* **2002**, *1*, 108–114, http://www.biochempress.com.
[13] J. V. De Julián Ortiz, J. Gálvez, C. Muñoz Collado, R. García–Domenech, and C. Gimeno–Cardona, Virtual Combinatorial Syntheses and Computational Screening of New Potential Anti–Herpes Compounds, *J. Med. Chem*. **1999**, *42*, 3308–3314.
[14] J. Gálvez, R. García–Domenech, M. J. Gómez–Lechón, and J. V. Castell, Use of Molecular Topology in the

Selection of New Cytostatic Drugs, *J.Mol. Struct.* (*Theochem*) **2000**, *504*, 241–248.
[15] E. Estrada, A. Peña, and R. García–Domenech, Designing Sedative/hypnotic Compounds From a Novel Substructural Graph–theoretical Approach *J. Comput. Aid. Mol. Des.* **1998**, *12*, 583–595.
[16] R. García–Domenech, F. J. García, R. M. Soler, J. Gálvez, G. M. Antón, and J. V. de Julián–Ortiz, New Analgesics Designed by Molecular Topology, *Quant. Struct. –Act. Relat.* **1996**, *15*, 201–207.
[17] E. Casabán, G. M. Antón–Fos, J. Gálvez, M. J. Duart, and R. García–Domenech, Search of New Antihistaminic Compounds by Molecular Connectivity, *Quant. Struct.–Act. Relat.* **1999**, *18*, 35–42.
[18] E. Estrada, E. Uriarte, A. Montero, M. Teijeira, L. Santana, and E. De Clercq, A novel approach for the virtual screening and rational design of anticancer compounds, *J. Med. Chem.* **2000**, *43*, 1975–1985.
[19] R. Gozalbes, M. Brun–Pascaud, R. García–Domenech, J. Gálvez, P. M. Girard, J. P. Doucet, and F. Derouin, Prediction of quinolone activity against *mycobacterium avium* by molecular topology and virtual screening, *Antimicrob. Agents Chemother.* **2000**, *44*, 2764–2770.
[20] R. Gozalbes, J. Gálvez, A. Moreno, and R. García–Domenech, Discovery of new antimalarial compounds by use of molecular connectivity techniques, *J. Pharm. Pharmacol.* **1999**, *51*, 111–117.
[21] I. Rios–Santamarina, R. García–Domenech, J. Cortijo, P. Santamaria, E. J. Morcillo, and J. Gálvez, Natural Compounds with Bronchodilator Activity Selected by Molecular Topology, *Internet Electron. J. Mol. Des.* **2002**, *1*, 70–79, http://www.biochempress.com.
[22] L. Pastor, R. García–Domenech, S. Wolski, M. D. García, and J. Gálvez, New Antifungals Selected by Molecular Topology, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2577–2582.
[23] The Merck Index, Twelfth Edition, **1996**.
[24] L. H. Hall, MOLCONN–Z software, **1995**, Eastern Nazarene College, Quincy, Massachusetts.
[25] Desmol13 software, **2000**, Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular, Facultad de Farmacia, Universitat de València, Spain.
[26] L. B. Kier, W. J. Murray, and M. Randić, L. H. Hall, Molecular Connectivity V: Connectivity Series Concept Applied to Density, *J. Pharm. Sci.* **1976**, *65*, 1226–1230.
[27] L. B. Kier and L. H. Hall, Molecular Connectivity VII: Specific Treatment of Heteroatoms, *J. Pharm. Sci.* **1976**, *65*, 1806–1809.
[28] J. Gálvez, R. García–Doménech, M. T. Salabert, and R. Soler, Charge Indices. New Topological Descriptors, *J. Chem. Inf. Comp. Sci.* **1994**, *34*, 520–525.
[29] R. Moliner, F. Garcia, J. Galvez, R. Garcia–Domenech, and C. Serrano, Nuevos índices topológicos en conectividad molecular. Su aplicación a algunas propieades físico–químicas de un grupo de hidrocarburos alifáticos, *An. Real Acad. Farm.* **1991**, *57*, 287–298.
[30] L. B. Kier, A shape index from molecular graphs, *Quant. Struct–Act. Relat.* **1985**, *4*, 109.
[31] L. B. Kier and L. H. Hall, *Pharm. Res.* An electrotopological–state index for atoms in molecules, **1990**, *7*, 801–807.
[32] H. Wiener, Structural determination of paraffin boiling points, *J. Am. Soc.* **1947**, *69*, 17–20.
[33] W. J. Dixon, BMDP Statistical Software, University of California, Berkeley, **1990**.
[34] J. Gálvez, R. García–Domenech, J. V. de Julián–Ortiz, and L. Popa, Pharmacological distribution diagrams: A tool for de novo drug design. *J. Mol. Graphics*. **1996**, *14*, 272–276.
[35] National Committee for Clinical Laboratory Standards. Reference Method for Broth Dilution Antifungal Susceptibility Testing of Yeast Approved Standard M27–A. Villanova. NCCLS, **1997**.
[36] E. Estrada and E. Uriarte, Recent advances on the role of topological indices in drug discovery research, *Curr. Med. Chem.* **2001**, *8*, 1573–1588.