CODEN IEJMAT *Internet Electronic Journal of Molecular Design* **2006**, *5*, 306–319 ISSN 1538–6414
BioChem Press http://www.biochempress.com

Inter*net* Electronic Journal of Molecular Design

June 2006, Volume 5, Number 6, Pages 306–319

Editor: Ovidiu Ivanciuc

Special issue dedicated to Professor Lemont B. Kier on the occasion of the 75th birthday

Topological Virtual Screening and Pharmacological Test of Novel Cytostatic Drugs

María Teresa Llacer,[1] Jorge Gálvez,[2] Ramón García–Domenech,[2] María José Gómez–Lechón,[3] Carmina Más–Arcas,[4] and Jesús Vicente de Julián–Ortiz [5]

[1] Edwards Lifesciences, S.L. Parque Tecnológico de Valencia, 46980 Paterna (Valencia), Spain
[2] Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular, Departamento de Química Física, Universitat de Valencia, Av. Vicent Andrés Estellés s.n. 46100 Burjassot, Spain
[3] Unidad de Hepatologia Experimental, Centro de Investigación, Hospital Universitario La Fe, Valencia, Spain
[4] Servicio de Farmacia, Instituto Valenciano de Oncología, Valencia, Spain
[5] Red de Investigación de Centros de Enfermedades Tropicales, Facultad de Farmacia, Universitat de Valencia, Burjassot, Valencia, Spain

Received: January 24, 2006; Revised: May 8, 2006; Accepted: May 15, 2006; Published: June 30, 2006

Citation of the article:
M. T. Llacer, J. Gálvez, R. García–Domenech, M. J. Gómez–Lechón, C. Más–Arcas, and J. V. de Julián–Ortiz, Topological Virtual Screening and Pharmacological Test of Novel Cytostatic Drugs, *Internet Electron. J. Mol. Des.* **2006**, *5*, 306–319, http://www.biochempress.com.

Inter*net* Electronic Journal of Molecular Design

BioChem Press
http://www.biochempress.com

Topological Virtual Screening and Pharmacological Test of Novel Cytostatic Drugs

María Teresa Llacer,[1] Jorge Gálvez,[2,]* Ramón García–Domenech,[2] María José Gómez–Lechón,[3] Carmina Más–Arcas,[4] and Jesús Vicente de Julián–Ortiz [5]

[1] Edwards Lifesciences, S.L. Parque Tecnológico de Valencia, 46980 Paterna (Valencia), Spain
[2] Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular, Departamento de Química Física, Universitat de Valencia, Av. Vicent Andrés Estellés s.n. 46100 Burjassot, Spain
[3] Unidad de Hepatologia Experimental, Centro de Investigación, Hospital Universitario La Fe, Valencia, Spain
[4] Servicio de Farmacia, Instituto Valenciano de Oncología, Valencia, Spain
[5] Red de Investigación de Centros de Enfermedades Tropicales, Facultad de Farmacia, Universitat de Valencia, Burjassot, Valencia, Spain

Received: January 24, 2006; Revised: May 8, 2006; Accepted: May 15, 2006; Published: June 30, 2006

Internet Electron. J. Mol. Des. **2006**, *5* (6), 306–319

Abstract

Motivation. The main goal of the present work is selecting new cytostatic lead compounds through molecular topology. This is particularly interesting since the finding of new therapeutic alternatives for cancer continues to be a very difficult task as demonstrated by the low number of lead drugs approved by the international agencies in the later years in this field.

Method. Molecular topology, a formalism based on describing the molecules as hydrogen–depleted graphs, as well as linear discriminant analysis, a statistical tool capable to distinguish between two or more categories or objects, have been used to select new cytostatic compounds. All the selected compounds were tested *in vitro* against two human cell cultures: HepG2, hepatocellular carcinoma and HeLa (ATCC CCL2) cell lines, corresponding to cervix epithelioid carcinoma.

Results. A mathematical model comprised of one discriminant function has been developed. The model is able to classify correctly 91.3% of the compounds from the training set. Usnic acid stands among the selected active compounds, showing significant anti–proliferative activity on the two selected lines HepG2 and HeLa, with IC_{50} values of 1.0 and 1.1 μM, respectively. Caffeine showed also significant anti–proliferative activity on HeLa cells. Other compounds such as pyridoxine, atropine and chlortetracycline show moderate inhibitory effect on the HeLa cell line.

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 cytostatic compounds, especially new leads.

Keywords. Linear discriminant analysis; molecular design; cytostatic agents; topological indices; molecular graph; structural descriptors; QSAR; quantitative structure–activity relationships; SAR; structure–activity relationships.

Dedicated to Professor Lemont B. Kier on the occasion of the 75th birthday.

* Correspondence author; phone: 34–963–544–891; fax: 34–963–544–892; E–mail: jorge.galvez@uv.es.

1 INTRODUCTION

The design of new active compounds showing the desired pharmacological properties is a principal objective that focuses the attention of pharmaceutical industry. Molecular topology has widely demonstrated its ability for an easy and efficient characterization of molecular structure through the so–called topological indices, which allow obtaining quantitative relations between molecular structure and activity (QSAR) capable to predict various pharmacological properties [1–3].

In that mathematical formalism a molecule is assimilated to a graph, where each vertex represents one atom and each edge one bond. Starting from the interconnections between the vertices, an adjacency topological matrix is built up in such a way that its elements t_{ij} take the values either one or zero, depending on the vertex *i* is connected or not to the vertex *j*, respectively. The manipulation of this matrix gives origin to a set of topological indices or topological descriptors (TI), which encode information on molecular size, shape and branching, key features of molecular structure. The computation of TIs is very fast and they show the advantage to be truly structural invariants, that is their values are independent of molecular conformations [4,5].

These structural invariants are able to base extra–mechanistical virtual screening methodologies that have demonstrated an actual capability to arise the presence of activity for structurally heterogeneous groups of compounds at various therapeutic areas. In those models, structural similarity is the key. By this means our research group has identified new antiviral [6], antibacterial [7,8], antimalarials [9], bronchodilator [10], antihistaminic [11], antifungal [12], antitoxoplasmatic [13] and also cytostatic [14,15] compounds, most of which can be considered as new leads.

The aim of this work was to search a new topological model to find new cytostatic agents from a widely diverse molecular database. The assessment of the pharmacological activity was carried out by performing the appropriate functional tests.

2 MATERIALS AND METHODS

2.1 Compounds Studied

Two sets of compounds were selected from the Merck Index [16] for linear discriminant analysis: A first group with 137 antineoplastic drugs and a second group of 126 drugs from other therapeutic categories. Tables 2, 3, and 4 illustrate the names of all the drugs included.

2.2 Topological Descriptors

Several types of descriptors have been used in this work: connectivity indices [17,18], topological charge indices, *TCI* [19], connectivity differences and quotients [20], Wiener index [21], *Vn* (number of vertices with topological valence *n*, with *n* = 3 or 4), and other graph–

theoretical descriptors (not outlined here, as they were not selected for the final model). All descriptors were calculated with Desmol1 software [22]. Table 1 shows the symbol, name, definition and references of each descriptor.

Table 1. Descriptors used in this study

Symbol	Name	Definition	Refs.
${}^{k}\chi_{t}$ *k*=0–4 *t*=p,c,pc	Randić–like indices of order *k* and type path (p), cluster (c) and path–cluster (pc)	${}^{k}\chi_{t}=\sum_{j=1}^{{}^{k}n_{t}}\left(\prod_{i\in S_j}\delta_i\right)^{-1/2}$ δ_i, number of bonds, σ or π, of the atom *i* to non–hydrogen atoms. S_j *j*th sub–structure of order *k* and type *t*.	[17]
${}^{k}\chi_{t}^{v}$ *k*=0–4 *t*=p,c,pc	Kier–Hall indices of order *k* and type path (p), cluster (c) and path–cluster (pc)	${}^{k}\chi_{t}^{v}=\sum_{j=1}^{{}^{k}n_{t}}\left(\prod_{i\in S_j}\delta_i^{v}\right)^{-1/2}$ δ_i^{v}, Kier–Hall valence of the atom *i*. S_j *j*th sub–structure of order *k* and type *t*.	[18]
G_k *k*=1–5	Topological charge indices of order *k*	$G_k=\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\lvert\mathbf{M}_{ij}-\mathbf{M}_{ji}\rvert\,\delta(k,\mathbf{D}_{ij})$ **M**=**AQ**, product of the adjacency and inverse squared distance matrices for the hydrogen–depleted molecular graph. **D**, distance matrix. δ, Kronecker delta	[19]
G_k^{v} *k*=1–5	Valence topological charge indices of order *k*	$G_k^{v}=\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\lvert\mathbf{M}_{ij}^{v}-\mathbf{M}_{ji}^{v}\rvert\,\delta(k,\mathbf{D}_{ij})$ $\mathbf{M}^{v}$=$\mathbf{A}^{v}\mathbf{Q}$, product of the electronegativity–modified adjacency and inverse squared distance matrices for the hydrogen–depleted molecular graph. **D**, distance matrix. δ, Kronecker delta	[19]
J_k *k*=1–5	Normalized topological charge indices of order *k*	$J_k=\frac{G_k}{N-1}$	[19]
J_k^{v} *k*=1–5	Normalized valence topological charge indices of order *k*	$J_k^{v}=\frac{G_k^{v}}{N-1}$	[19]
${}^{k}D_{t}$ *k*=0–4 *t*=p,c,pc	Connectivity differences of order *k* and type path (p), cluster (c) and path–cluster (pc)	${}^{k}D_{t}={}^{k}\chi_{t}-{}^{k}\chi_{t}^{v}$	[20]
${}^{k}C_{t}$ *k*=0–4 *t*=p,c,pc	Connectivity quotients of order *k* and type path (p), cluster (c) and path–cluster (pc)	${}^{k}C_{t}=\frac{{}^{k}\chi_{t}}{{}^{k}\chi_{t}^{v}}$	[20]
Vn n=3,4	Number of vertices	Number of vertices with topological valence n	[20]
W	Wiener path number	$W=1/2\sum\delta_{ij}$	[21]

2.3 Linear Discriminant Analysis

Once calculated the topological indices, linear discriminant analysis (LDA) [23] was performed to obtain one discriminant function, DF, capable to select new active compounds. LDA is a useful statistical tool focused to the achievement of equations allowing distinguish between two or more categories of objects. In our case there were two sets of compounds: The first consist of 137 antineoplastic drugs (43 of them with a contrasted cytostatic activity). Molecules were selected including as much structural and functional heterogeneity as possible, so that alkylating agents, antimetabolites, antibiotics, androgens and also other less specific types of antineoplastics were included. The second group, 126 compounds, was comprised of bronchodilator, antihypertensive, antianginal, antiamebic, antiarrhithmic, antituberculosis, anticholinergic, antihistaminic, antihyperlipidemic, antihipertensive, antifungal, analgesic, anticonvulsant, antiemetic, diuretic and antispasmodic drugs, vitamins and some chemical reactives. It is important for that set of inactive compounds to include structural and functional heterogeneity as well as chemical characteristics similar to the active set, particularly a similar molecular size.

The selection of the best discriminant function was carried out using the BMDP 7M package [24] and a training group formed by 43 cytostatic drugs (referenced in Merck Index) and 49 non–antineoplastic compounds selected randomly of the inactive group. The rest of compounds formed the test set.

The method used for the selection of the descriptors was based on the F–Snedecor, and the classification criterion was the shortest Mahalanobis distance (distance of each case to the mean of all cases used in the regression equation). 7M chooses the variables used in computing the linear classification functions in a stepwise manner: at each step the variable that adds the most to the separation of the groups is entered (or the variable that adds the least is removed from) the discriminant function. The quality of the discriminant function is evaluated by the Wilks' λ parameter, which is a multivariate analysis of variance statistic that tests the equality of group means for the variables in the discriminant function and by TP (true positive), TN (true negative), FP (false positive), FN (false negative), sensitivity TPF = (TP/(TP+FN)), specificity TNF = (TN/(TN+FP)), Matthews correlation coefficient, and AUC of the ROC curve. The discriminant ability of the selected function is evaluated through a cross–validation by leave–one–out (jack–knifed classification matrix) [25]: to do this, each compound is classified into a group according to the classification function computed from all the group except the case being classified.

2.4 Pharmacological Distribution Diagrams

These diagrams have been successfully used, for example, to select new antiviral, antimicrobial and antimalarial compounds [6,9,26]. It consists of applying the discriminant function, DF, to both, the group of active compounds and the group of inactive ones. The structures are grouped into the

predicted values of DF intervals, and the frequency of its appearing along each interval of DF is determined for each group. The expectancy "E" to find a molecule with a desired value of DF is so obtained [26]. Thus, for each arbitrary range of DF values, we can define the expectancy of activity E_a as: $E_a = a/(i + 1)$, where a is the number of active compounds in the interval divided by the total number of active compounds and i is the number of inactive compounds in the interval divided by the total number of inactive compounds. The expectancy of inactivity is defined, in a symmetrical way, as $E_i = i/(a + 1)$. In our case, E_a = expectancy of cytostatic activity, and E_i = expectancy of non–cytostatic activity.

When for a given DF function, Ea acquires the form of a distribution and Ei is minimal under the curve, this function is considered as worthy for molecular selection. This allows establishing the DF intervals where the probability of finding new active compounds is maximal with regards to the chance for false active.

2.5 Cytological Tests

Tests for experimental detection of cytostatic activity were developed on two tumor cell lines, namely human hepatocellular carcinoma, HepG2 (ATCC HB 8065) and human cervix epithelioid carcinoma, HeLa (ATCC CCL2) cell lines. Cell culture was Eagle's MEM supplemented with 7 % fetal calf serum, 50 mg/ml streptomycin/ml and 50 mU penicillin/ml. Solutions were filtered through a 0.22 μm porous membrane, and the progressive dilutions of the stock solutions were done in PBS (phosphate buffered saline pH 7.4).

The MTT test, a viability assay, was used as end–point parameter for cytotoxicity and cell proliferation evaluation. This test consists of a reduction of the tetrazolium salt MTT to a blue formazan by mitochondrial succinate dehydrogenase. First, the cytotoxicity of these compounds was studied on confluent monolayers after 72 h exposure of cells to increasing concentrations of the compounds to determine the maximal non–toxic concentration (MNTC). For inhibition of cell proliferation experiments, cells were seeded in 96–well culture plates at a density of 2000 to 4000 cells/well in 100 ml of culture medium. Chemicals were added at sub–cytotoxic concentrations (up to the MNTC) to 24 h cultures, and every 2 days after medium renewal. Control cultures were treated with PBS. Cell proliferation was monitored periodically with the MTT assay. Before the assay, micro titer plates were washed twice with 50 ml PBS at 37ºC, and the assay was performed as described [27,28].

To calculate IC_{50} values (concentrations that produce a 50% of inhibitory effect on cell proliferation), all the results (two–three independent experiments) were transformed to percentage of controls, and the typical sigmoid concentration–effect curves, with all the data were linearized using the LOGIT transformation. The IC values were mathematically interpolated.

Table 2. Results of classification obtained in the LDA study for each compound of the training group

Compound	DF	Prob(+)	Class	Compound	DF	Prob(+)	Class
Cytostatic drugs (Active group, +)							
Amsacrine	–0.94	0.281	–	Lavedustin	0.39	0.597	+
Ancitabine	5.07	0.994	+	Mechloretamine	1.54	0.823	+
Azaserin	3.57	0.973	+	Melphalan	–0.58	0.359	–
Camptothecin	3.62	0.974	+	Methotrexate	5.64	0.996	+
Carmofur	6.07	0.998	+	Mitobronitol	3.55	0.972	+
Carmustine	3.84	0.979	+	Mitoguazone	–2.65	0.066	–
Cyclophosphamide	2.96	0.951	+	Mitomycins	2.18	0.899	+
Cytarabine	4.11	0.984	+	Mitoxantrone	2.08	0.889	+
Dacarbazine	0.09	0.524	+	Nimustine	1.31	0.787	+
Dactinomycin	4.63	0.991	+	Nitracrine	0.48	0.617	+
Damnacanthol	–0.72	0.329	–	Paclitaxol	6.43	0.998	+
Daunorubicin	7.34	0.999	+	Perfosfamide	2.15	0.896	+
Doxifluridine	4.96	0.993	+	Pirarubicin	7.42	0.999	+
Doxorubicin	7.24	0.999	+	Podofilox	2.66	0.935	+
Epirubicin	7.24	0.999	+	Teniposide	8.23	1	+
Estramustine	–3.13	0.042	–	Thiabendazole	1.23	0.775	+
Estreptozocin	3.75	0.977	+	Thiotepa	2.76	0.94	+
Etoposide	8.34	1	+	Verrucarin	1.32	0.791	+
Fluorouracil	3.31	0.965	+	Vinblastine	6.09	0.998	+
Hydroxiurea	5.50	0.996	+	Vincristine	6.26	0.998	+
Idarubicin	6.30	0.998	+	Vindesine	6.43	0.998	+
Ifosfamide	3.06	0.955	+				
Compound	DF	Prob(–)	Class.	Compound	DF	Prob(–)	Class.
Non–cytostatic drugs (Inactive group, –)							
2–amino–4–picoline	–6.34	0.998	–	Guanabenz	–6.82	0.999	–
Acetanilide	–3.85	0.979	–	Homatropine	–1.73	0.85	–
Alprenolol	–3.36	0.966	–	Homonicotinic acid	–2.48	0.923	–
Aminoglutethimide	–5.36	0.995	–	Hydracarbazine	–1.93	0.873	–
Amisometradine	–0.70	0.668	–	Iodochlorhydroxyquin	–5.73	0.997	–
Benzthiazide	–2.25	0.904	–	Lamotrigine	–3.92	0.981	–
Biphenamine	–1.00	0.73	–	Lidocaine	–4.05	0.983	–
Bufexamac	–4.10	0.984	–	Medetomidine	–6.57	0.999	–
Bunitrolol	–4.56	0.99	–	Menadiol	–5.62	0.996	–
Buthiazide	–5.79	0.997	–	Methylhexaneamine	–9.27	1	–
Caramiphen	–3.24	0.962	–	Muzolimine	–5.14	0.994	–
Carbamazepine	–2.24	0.903	–	Naftifine	–4.76	0.991	–
Chloraminophenamide	–8.34	1	–	Narcobarbital	–3.22	0.962	–
Chlorothen	–2.43	0.919	–	Nicotinic acid	1.35	0.205	+
Chlorphenesin	–3.12	0.958	–	P–aminosalicylic acid hydrazide	–5.68	0.997	–
Clofibrate	–2.92	0.949	–	Pindolol	–1.69	0.845	–
Cropropamide	–1.60	0.831	–	Rubijervine	–7.68	1	–
Cycloserine	–1.32	0.789	–	Sulthiame	–2.39	0.916	–
Cycrimine hydrochloride	–3.16	0.959	–	Synephrine	–3.58	0.973	–
Dihydralazine	–4.81	0.992	–	Tenonitrozole	–0.53	0.629	–
Diloxanide	–4.14	0.984	–	Timolol	–1.49	0.815	–
Diphenhydramine	–3.27	0.963	–	Tolpropamine	–5.72	0.997	–
Enfenamic acid	–3.87	0.98	–	Triamterene	2.58	0.07	+
Flucytosine	2.17	0.103	+	Trimethadione	–3.14	0.958	–
Fosfosal	–4.19	0.985	–				

Both, MNTC and IC_{50} were also obtained for two clinically used antineoplastic drugs, 5–fluorouracil and mitomycin C, to get reference values. The cytotoxicity ratios (CTRs) as the quotient between the MNTC values and the IC_{50} ones were also obtained for each cell line. Thus, the larger this ratio the better cytostatic a compound is.

3 RESULTS AND DISCUSSION

The results obtained from LDA, led to the selection of the following discriminant function:

$$DF = -6.56\ {}^4\chi_c - 16.96\ J_1 + 8.56\ J_1^v - 21.34\ J_3 - 7.52\ {}^3C_p + 2.63\ {}^4D_{pc} + 5.28\ {}^4C_{pc} - 0.00048\ W + 0.87\ V_4 - 4.24$$

N = 93 λ = 0.357 F = 16.6 TP = 38 TN = 46 FP = 3 FN = 5

TPF (threshold DF=0) = 0.884 TNF (threshold DF=0) = 0.939 Matthews coeff =0.826 AUC=0.970

This DF function contains connectivity indices that evaluate fundamental topological aspects of each compound ($^4\chi_c$, ^3C_p, $^4D_{pc}$, $^4C_{pc}$) and topological charge indices (J_1, J_1^v, J_3) allowing the evaluation of the distribution of intramolecular charges. The function was capable to classify correctly 88.4 % of actives (38 out of 43) and 93.9 % of the inactive compounds (46 out of 49), which, despite of the large structural heterogeneity of the analyzed compounds, clearly points out its efficacy in discriminating cytostatic activity. Cross–validation (jack–knifed matrix) of the training group showed that 37 (84.1%) of the 43 active compounds and 46 (93.9%) of the 49 inactive compounds were correctly classified.

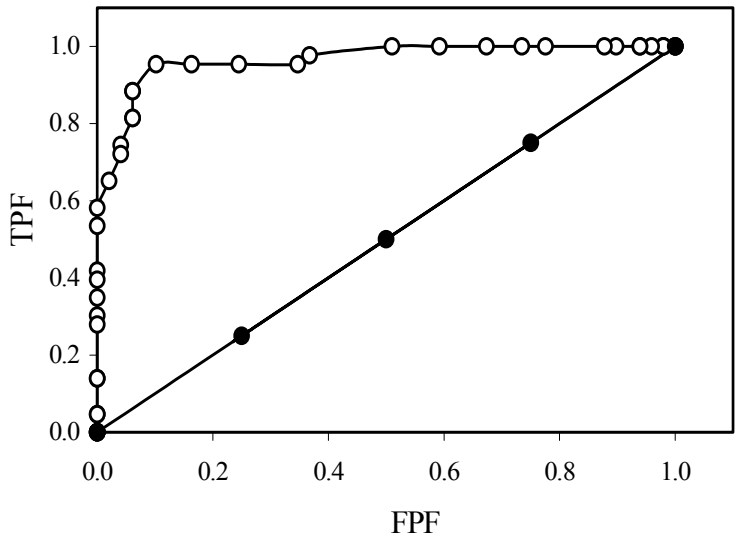


Figure 1. Receiver operating characteristic curve (ROC) for training set (white points) and random classifier (black points). TPF = sensitivity and FPF = 1 – specificity for different thresholds of DF (between –10 and +10).

Table 2 summarizes the results of classification obtained with DF for active and inactive groups, respectively in the training set. A compound will be classified as *active*, +, if DF > 0 (probability assigned, prob(+) > 0.500) and as *inactive*, –, if DF < 0 (prob(–) > 0.500). Receiver operating characteristic curve, ROC, for the training set is shown in the Figure 1. The area under the curve is

0.970. These results confirm that the developed model is not a random classifier on the basis that the area value is significantly higher than 0.5.

Table 3. Results of classification obtained for the test set of actives (+) (antineoplastic drugs not used in the LDA)

Compound	DF	prob(+)	class	Compound	DF	prob(+)	class
Aceglatone	0.56	0.638	+	Mannomustine	1.14	0.759	+
Altretamine	3.80	0.978	+	Medroxyprogesterone	–7.05	0.001	–
Anastrozole	–6.35	0.002	–	Melengestrol	–6.04	0.002	–
Anthramycin	0.77	0.685	+	Menogaril	6.21	0.998	+
Azacitidine	5.75	0.997	+	Mepitiostane	–7.95	0	–
6–azauridine	4.66	0.991	+	6–mercaptopurine	4.55	0.956	+
Batimastat	–4.15	0.016	–	Meturedepa	–5.34	0.005	–
Benzodepa	0.83	0.697	+	Mitolactol	3.55	0.972	+
Bicalutamide	2.81	0.944	+	Mitotane	–8.71	0	–
Bisantrene	1.34	0.793	+	Mopidamol	0.89	0.711	+
Calusterone	–9.28	0	–	Mycophenolic acid	–1.12	0.246	–
Carboquone	0.38	0.595	+	Ninopterin	5.53	0.996	+
Carubicin	7.07	0.999	+	Nordihydroguaiaretic acid	–4.94	0.007	–
Chlorambucil	–3.11	0.043	–	Novembichin	–1.86	0.135	–
Chlormadinone acetate	–5.77	0.003	–	Pentostatin	3.43	0.969	+
Chlornaphazine	–2.91	0.052	–	Phenamet	0.87	0.705	+
Chlorozotocin	3.08	0.99	+	Phenesterine	–9.14	0	–
Colchicine	1.05	0.741	+	Pipobroman	2.58	0.93	+
Defosfamide	2.99	0.952	+	Piposulfan	–0.53	0.371	–
Demecolcine	0.41	0.603	+	Piritrexim	0.67	0.663	+
Denopterin	5.41	0.996	+	Podophyllic acids	2.03	0.884	+
Diaziquone	2.42	0.919	+	Porfiromycin	2.80	0.943	+
Droloxifene	–2.83	0.056	–	Prednimustine	–2.49	0.078	–
Dromostanolone	–9.09	0	–	Procarbazine	–4.42	0.012	–
Edatrexate	3.62	0.974	+	Puromycin	6.32	0.998	+
Eflornithine	4.06	0.983	+	Razoxane	1.50	0.818	+
Elliptinium acetate	–1.21	0.231	–	Retinoic acid	–9.37	0	–
Emitefur	8.51	1	+	Roquinimex	0.46	0.615	+
Epitiostanol	–10.67	0	–	Sobuzoxane	4.63	0.99	+
Etanidazole	3.53	0.972	+	Streptonigrin	8.06	1	+
Etoglucid	6.18	0.998	+	Tegafur	4.34	0.987	+
Fadrozole	–2.14	0.106	–	Temozolomide	2.54	0.927	+
Fenretinide	–7.15	0.001	–	Tenuazonic acid	–4.79	0.008	–
Floxuridine	3.76	0.977	+	Testolactone	–7.44	0.001	–
Fludarabine	6.68	0.999	+	Thiamiprine	2.50	0.924	+
Flutamide	–1.16	0.24	–	Thioguanine	0.66	0.661	+
Formestane	–8.29	0	–	Tomudex (tm)	2.17	0.898	+
Fosfestrol	–7.87	0	–	Topotecan	5.41	0.996	+
Fotemustine	3.35	0.966	+	Toremifene	–2.86	0.054	–
Gemcitabine	9.14	1	+	Triaziquone	0.31	0.577	+
Hexestrol	–7.85	0	–	Triethylenemelamine	6.10	0.998	+
Improsulfan	3.53	0.971	+	Triethylenephosphoramide	3.32	0.965	+
Iobenguane	–7.67	0	–	Triethylenethiophosphoramide	2.76	0.94	+
Irinotecan	6.60	0.999	+	Trimetrexate	1.27	0.782	+
Letrozole	–0.39	0.404	–	Trofosfamide	2.81	0.943	+
Lomustine	0.28	0.57	+	Zorubicin	7.73	1	+
Lonidamine	–3.04	0.046	–				

Table 3 shows the classification results from the application of DF to the antineoplastic drugs not included in the training set (test set). 58 out of the 93 compounds are correctly recognized as cytostatic. A literature search by Medline stands that more than 80% of these compounds had been described as cytostatics, what fits well with the cytostatic related mechanism followed by many of antineoplastics.

Table 4. Results of classification obtained for the test set of inactives (–) (antineoplastic drugs not used in the LDA)

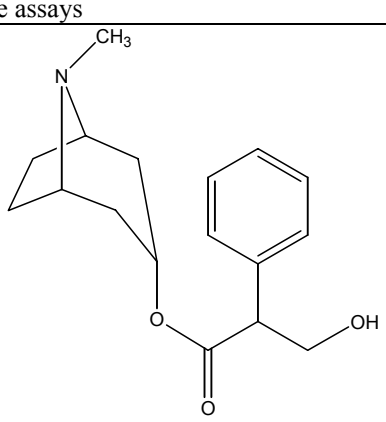
Compound	DF	prob(–)	class	compound	DF	prob(–)	class
α–carotene	–11.37	1	–	Etodolac	–3.21	0.961	–
Acetorphan	–1.70	0.844	–	Felodipine	–2.15	0.896	–
Adenosine	7.08	0.001	+	Fenofibrate	–0.51	0.623	–
Ambuside	–2.51	0.925	–	**Glucosamide**	3.06	0.045	+
Amixetrine	–4.37	0.988	–	Glyconiazide	4.23	0.014	+
Amosulalol	–2.18	0.898	–	G–oryzanol	–8.10	1	–
Antazoline	–2.16	0.897	–	Indapamide	–3.31	0.965	–
Antrafenine	5.04	0.006	+	Isofezolac	–1.86	0.865	–
Apoatropine	–1.75	0.852	–	Itraconazole	1.83	0.134	+
Atenolol	–2.29	0.908	–	Limaprost	–6.27	0.998	–
Atropine	0.31	0.245	+	Mequitazine	–3.25	0.963	–
Beclobrate	–2.19	0.899	–	Methantheline bromide	–1.04	0.739	–
Benazepril	–0.13	0.531	–	Moricizine	0.91	0.286	+
Benzetimide	–3.46	0.969	–	Nabilone	–7.78	1	–
Bermoprofen	–2.38	0.915	–	N–hydroxyethylpromethazine	–2.32	0.91	–
Bisoprolol	–1.95	0.875	–	Nifedipine	–0.29	0.572	–
Bromopride	–2.34	0.912	–	Octopamine	–5.50	0.996	–
Bucumolol	–3.96	0.981	–	Penthienate bromide	–1.01	0.731	–
Caffeine	3.50	0.029	+	Phenobarbital	–2.55	0.928	–
Carazolol	–1.29	0.783	–	Phenyltoloxamine	–3.53	0.971	–
Carbinoxamine	–2.07	0.888	–	Pilsicainide	–3.78	0.978	–
Caroverine	0.37	0.408	+	Piretanide	–0.45	0.609	–
Celiprolol	–1.90	0.869	–	Pridinol	–2.83	0.944	–
Chlorbetamide	–4.82	0.992	–	Propafenone	–1.89	0.868	–
Chlortetracycline	3.63	0.026	+	Propamidine	–0.83	0.696	–
Clentiazem	–0.53	0.629	–	Propyromazine	–1.01	0.731	–
Clobazam	–0.65	0.656	–	Protionamide	–5.65	0.996	–
Clobenzepam	–0.49	0.621	–	Pyridoxal 5–phosphate	–3.38	0.967	–
Cyclonium iodide	–1.12	0.754	–	**Pyridoxine**	0.38	0.294	+
Deoxyepinephrine	–2.75	0.94	–	Sulfoniazide	–0.50	0.621	–
Diamthazole dihydrochloride	–0.66	0.659	–	Sulpiride	–0.57	0.639	–
Dihexyverine	–3.83	0.979	–	Tetracycline	3.60	0.026	+
Dihydrocodeine	–4.65	0.991	–	Thyropropic acid	–7.18	0.999	–
Domperidone	2.49	0.076	+	Timepidium bromide	–3.65	0.975	–
Emedastine	1.34	0.207	+	Trimethobenzamide	1.39	0.198	+
Enalaprilat	0.78	0.314	+	**Usnic acid**	0.44	0.392	+
Encainide	–1.24	0.774	–	Vitamin e	–5.75	0.997	–
Eterobarb	1.84	0.137	+	Yohimbine	–1.64	0.837	–
Ethybenztropine	–3.50	0.971	–				

Table 4 shows the classification obtained for the test set of inactives (inactive compounds group not used in the ALD). Approximately, 80% of the compounds are successfully classified as inactive.

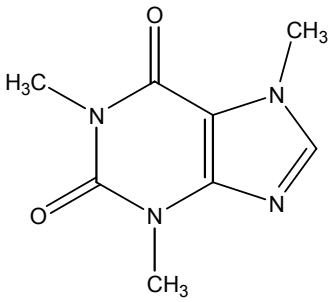
It is noteworthy that, within the inactives group, some elements that were *unexpectedly* classified

as *active* by the function DF, such as etodolac, flucytosine and tetracycline have shown cytostatic activity in recent studies. Thus, the COX–2 inhibitor etodolac was found to inhibit cell invasion of LM–H3 [29]; cytosine deaminase converts the prodrug flucytosine into cytotoxic 5–fluorouracil, which leads to tumor–cell eradication [30] and some tetracyclines with antimicrobial activity were reported to possess cytostatic and cytotoxic activity against mammalian tumor cells, often at high doses [31].

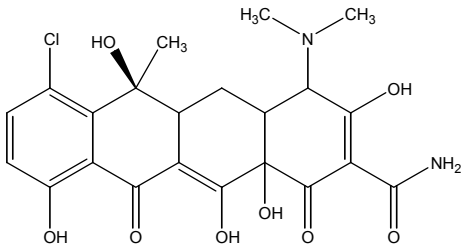
Table 5. Chemical structures of the selected compounds with possible cytostatic activity and drugs used as reference in the assays



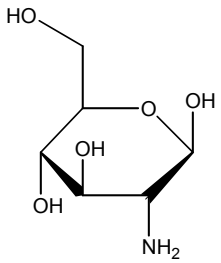
Atropine
DF=0.31



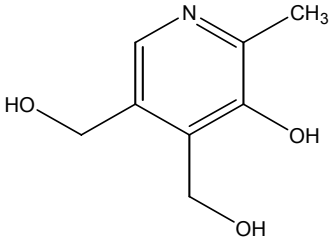
Caffeine
DF=3.50



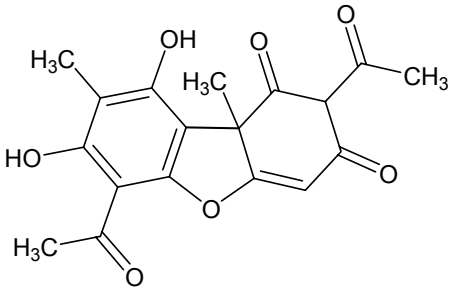
Chlortetracycline
DF=3.63



Glucosamine
DF=3.06

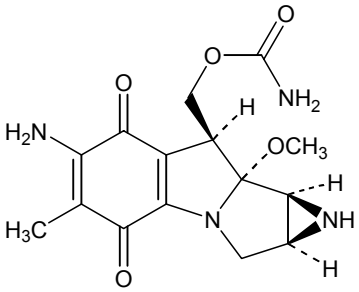


Pyridoxine
DF=0.38



Usnic Acid
DF=0.44

5–Fluorouracil (Ref. drug)



Mitomycin C (Ref. drug)

Figure 2 shows the PDD obtained using DF function. The maximum expectancy zone for new cytostatic compounds is between DF>0 and DF<8.0. In order to check the validity of the proposed topological model, six compounds of the inactive set, which had been classified as active by our model, were selected for testing. Table 5 shows the chemical structures for such six compounds

together with their DF values: pyridoxine, atropine, chlortetracycline, usnic acid, caffeine and glucosamine. No one of these compounds was previously reported to be active in either of the HepG2 or HeLa cell lines.

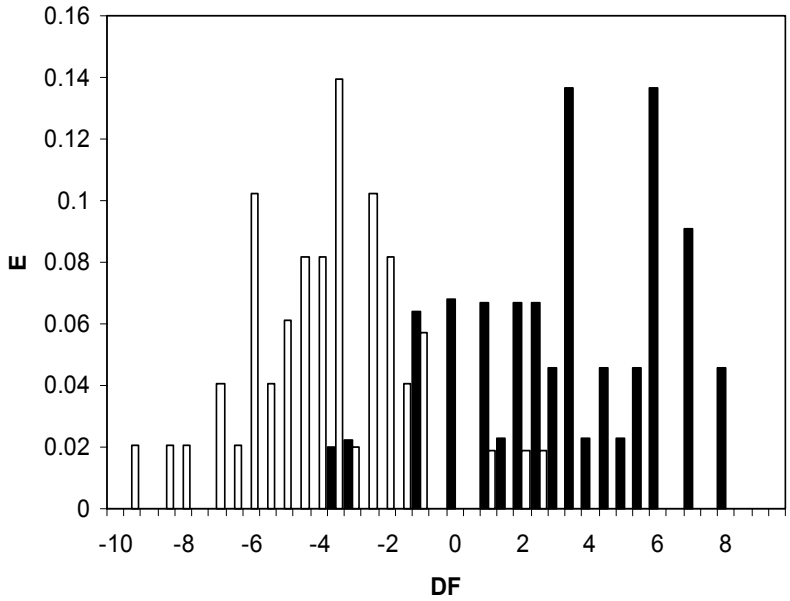


Figure 2. Pharmacological distribution diagram obtained from a set of 93 compounds (44 cytostatic drugs and 49 inactive ones). White line: Ei for non cytostatic drugs. Black line: Ea for cytostatic drugs.

The experimental results obtained for cytotoxicity are illustrated in Table 6, while Table 7 shows the results obtained for the activity test, expressed as IC_{50}. A given compound is considered to be active (to show anti–proliferative effect) if IC_{50} is less than MNTC, which means CTR values above 1.0.

Table 6. Cytotoxicity (expressed as MNTC in μM) of the selected compounds

Compound	MNTC [a] on Cell lines	
	HepG2	HeLa
Pyridoxine	500	1000
Atropine	250	250
Chlortetracycline	21.3	170
Usnic Acid	18	18.1
Caffeine	400	2500
Glucosamine	625	625
5–Fluorouracil (reference drug)	100	–
Mitomycin C (reference drug)	0.7	–

[a] Maximal non toxic concentration (MNTC) calculated after 3 days of treatment of cells with the compounds, evaluated with the MTT test (n = 1–2). Compounds were assayed in the range 3000–0.8 μM.

Table 8 summarizes the CTR values for the selected compounds. Although all of them show anti–proliferative effect at least on one line, the most promising compound is usnic acid (antibiotic extracted of the *Usnea Barbata* lichen), with CTR values of 16 on HeLa and 18 on HepG2 lines, respectively (see Table 8). These values are larger than those of the reference drugs. Other compounds, such as caffeine, show marked inhibitory effect on HeLa cell line.

Table 7. Inhibitory effect of the selected compounds (expressed as IC_{50} in μM) on cell proliferation

Compound	IC_{50} on Cell lines [a]	
	HepG2	HeLa
Pyridoxine	725	200
Atropine	>500	125
Chlortetracycline	100	70
Usnic Acid	1.0	1.1
Caffeine	>6000	125
Glucosamine	2000	625
5–Fluorouracil (reference drug)	85	–
Mitomycin C (reference drug)	0.27	–

[a] IC_{50} calculated after 5 days of treatment of cells with the compounds, valuated with the MTT test (n= 2–3).

Table 8. Cytoxicity ratios, CTRs, of the selected compounds

Compound	CTRs on different Cell lines [a]	
	HepG2	HeLa
Pyridoxine	0.7	5
Atropine	< 0.5	2.0
Chlortetracycline	0.2	1.7
Usnic Acid	18.0	16.0
Caffeine	< 0.06	20.0
Glucosamine	0.3	1.0
5–Fluorouracil (reference drug)	1.17	–
Mitomycin C (reference drug)	2.59	–

[a] CTRs expressed as the MNTC values divided by the IC_{50} ones.

Though, as pointed above, no previous activity report on the tested cell cultures was found for any of our selected compounds, one of them (chlortetracycline) is an antibiotic and, therefore, cannot be considered as new lead. However, the five others may be considered as new cytostatic leads in the HepG2 and HeLa cell lines, given their dissimilar structures as compared to all the remainder of known active compounds. It is also to be emphasized that four of the selected compounds are less harmful for the HeLa and HepG2 cell lines than one of the reference antineoplastic drugs, namely 5–fluorouracil.

The results described herein account for the capability of molecular topology to disclose novel compounds as a possible cytostatic–antineoplastic therapy but also demonstrate the validity of the technique as alternative to other mechanistic approaches in the search of new drugs.

4 CONCLUSIONS

Molecular topology has demonstrated to be a useful methodology for identifying new compounds with cytostatic activity. In this paper, a topological–mathematical model comprising one discriminant function has been developed. The SAR model is able to classify correctly 91.3% of the compounds from the training set. The validation was performed by a cross–validation, leave–one–out. Six compounds were selected as cytostatics and tested on two tumoral cell lines, HepG2 and HeLa.

Acknowledgment

The authors acknowledge financial support from the Spanish *Red de Investigación de Centros de Enfermedades Tropicales*, *RICET*, (*C03/04*) and the SAF2005–PI052128 of the *Fondo de Investigación Sanitaria*, *Ministerio de Sanidad*, Spain, and the grant number BQU2003–07420–C05 of the 'Ministerio de Ciencia y Tecnología' within the Spanish Plan Nacional I+D.

Supplementary Material

Value for topological indices and structures for all compounds in MDL format are deposited as supplementary material.

5 REFERENCES

[1] R. Todeschini and V. Consonni, Handbook of Molecular Descriptors, Vol 11. In *Methods and Principles in Medicinal Chemistry*. R. Mannhold, H. Kubinyi, H. Timmerman, Eds.; Wiley –VCH Verlag GmbH, Weinheim, 2000.

[2] K. Roy and J. T. Leonard, QSAR modeling of HIV–1 reverse transcriptase inhibitor 2–amino–6–arylsulfonylbenzonitriles and congeners using molecular connectivity and E–state parameters. *Bioorg. Med. Chem*. **2004**, *12*, 745–754.

[3] A. Golbraikh, D. Bonchev, and A. Tropsha, Novel *ZE*–Isomerism Descriptors Derived from Molecular Topology and Their Application to QSAR Analysis. *J. Chem. Inf. Comput. Sci*. **2002**, *42*, 769–787.

[4] J. Devillers and A. T. Balaban, *Topological Indices and Related Descriptors in QSAR and QSPAR*; Overseas Publishers Association, Amsterdam 1999.

[5] L. B. Kier and L. H. Hall, *Molecular Connectivity in Structure–Activity Analysis*, Research Studies Press, Letchworth, 1986.

[6] J. V. de Julián–Ortiz, J. Gálvez, C. Muñoz–Collado, R. García–Domenech, and C. Gimeno–Cardona, Virtual combinatorial synthesis and computational screening of new potential anti–herpes compounds. *J. Med. Chem*. **1999**, *42*, 3308–3314.

[7] A. García–García, J. Gálvez, J. V. de Julián–Ortiz, R. García–Domenech, C. Muñoz, R. Guna, and R. Borrás, New agents active against Mycobacterium avium complex selected by molecular topology: a virtual screening method. *J. Antimicrob. Chemother*. **2004**, *53*, 65–73.

[8] R. García–Domenech and J. V. de Julián–Ortiz, Antimicrobial Activity Characterization in a Heterogeneous Group of Compounds. *J. Chem. Inf. Comput. Sci*. **1998**, *38*, 445–449.

[9] N. Mahmoudi, J. V. de Julián–Ortiz, L. Ciceron, J. Gálvez, D. Mazier, M. Danis, F. Derouin, and R. García–Domenech, Identification of new antimalarial drugs by linear discriminant analysis and topological virtual screening, *J. Antimicrob. Chemother*. **2006**, *57*, 489–497.

[10] I. Ríos–Santamarina, R. Garcia–Domenech, J. Gálvez, E. J. Morcillo, P. Santamaría, and J. Cortijo, Getting new bronchodilator compounds from molecular topology. *Eur. J. Pharm. Sci*. **2004**, *22*, 271–277.

[11] M. J. Duart, G: M. Anton–Fos, P. A. Aleman, J. B. Gay–Roig, M. E. Gonzalez–Rosende, J. Galvez, and R. Garcia–Domenech, New Potential Antihistaminic Compounds. Virtual Combinatorial Chemistry, Computational Screening, Real Synthesis, and Pharmacological Evaluation, *J. Med Chem*. **2005**, *48*, 1260–1264.

[12] L. Pastor, R. Garcia–Domenech, S. Wolski, M.D. Garcia, and J. Gálvez, New antifungals selected by Molecular Topology. *Bioorg. Med. Chem. Lett*. **1998**, *8*, 2577–2582.

[13] R. Gozalbes, M. Brun–Pascaud, R. García–Domenech, J. Gálvez, P. M. Girard, J. P. Doucet, and F. Derouin, Anti–Toxoplasma Activities of 24 quinolones and Fluoroquinolones In Vitro: Prediction of Activity by Molecular Topology and Virtual Computational Techniques. *Antimicrob. Agents Chemother.* **2000**, *44*, 2771–2776.

[14] J. Gálvez, M. J. Gomez–Lechón, R. Garcia–Domenech, and J. V. Castell, New Cytostatic Agents Obtained by Molecular Topology. *Bioorg. Med. Chem. Lett*. **1996**, *6*, 2301–2306.

[15] 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.

[16] The Merck Index. An encyclopedia of chemicals, drugs, and biologicals. Ed. S. Budavari, 12th edition. MERCK & CO., INC., Whitehouse Station, NJ, 1996

[17] L. B. Kier, W. J. Murray, M. Randić, and L. H. Hall, Molecular Connectivity V: connectivity series concept applied to density. *J. Pharm. Sci*. **1976**, *65*, 1226–1230.

[18] L. B. Kier and L. H. Hall, General Definition of Valence Delta–Values for Molecular Connectivity. *J. Pharm. Sci*.

1983, *72*, 1170–1173.
[19] J. Gálvez, R. García–Domenech, M. T. Salabert, and R. Soler, Charge indices. New topological descriptors. *J. Chem. Inf. Comp. Sci.* **1994**, *34*, 520–525.
[20] J. Gálvez, R. García–Domenech, J. V. de Julián–Ortiz, and R. Soler, Topological approach to drug design. *J. Chem. Inf. Comp. Sci.* **1995**, *35*, 272–284. Errata: *J. Chem. Inf. Comp. Sci.* **1995**, *35*, 938.
[21] H. Wiener, Structural determination of paraffin boiling points, *J. Am. Chem. Soc.* **1947**, *69*, 17–20.
[22] DESMOL11 software. Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular. Facultad de Farmacia, Universitat de Valencia. Spain
[23] T. Collous, Discriminant Analysis and Applications, Academic Press: New York, 1973.
[24] W. J. Dixon, M. B. Brown, L. Engelmanand, and R. I. Jennrich, *BMDP Statistical Software Manual*, *Vol I.* University of California, Berkeley. Press 1990, 3390–4358.
[25] P. A. Lachenbruch and R. M. Mickey, Estimation of error rates in discriminant analysis. *Technometrics*, **1968**, *10*, 1–11.
[26] J. Gálvez, R. García–Domenech, C. de Gregorio Alapont, J. V. de Julián–Ortiz, and L. Popa, Pharmacological distribution diagrams: a tool for the new drug design. *J. Mol. Graphics*. **1996**, *14*, 272–276.
[27] J. Carmichael, W. G. DeGraff, A. F. Gazdar, J. D. Minna, and J. B. Mitchell, Evaluation of a tetrazolium–based semiautomated colorimetric assay: assessment of chemosensitivity testing. *Cancer Res*. **1987**, *47*, 936–942.
[28] X. Ponsoda, R. Jover, C. Nuñez, M. Royo, J. V. Castelland, and M. J. Gómez–Lechón, Evaluation of the Cytotoxicity of 10 Chemicals in Human and Rat Hepatocytes and in Cell Lines: Correlation Between In Vitro Data and Human Lethal Concentration. *Toxic. In Vitro* **1995**, *9*, 959–966.
[29] N. Matsunaga, N. Yamada, M. Ohira, A. Tachimori, Y. Nishiguchi, H. Nishino, S. Seki, and K. Hirakawa, Combined treatment with selective cyclooxygenase–2 inhibitor and fluorinated pyrimidines for liver metastasis of colon cancer. *Oncology Reports*. **2004**, *11*, 167–71.
[30] S. D. Mahan, G. C. Ireton, B. L. Stoddard, and M. E. Black, Alanine–scanning mutagenesis reveals a cytosine deaminase mutant with altered substrate preference. *Biochem*. **2004**, *43*, 8957–64.
[31] B. L. Lokeshwar, E. Escatel, and B. Zhu, Cytotoxic activity and inhibition of tumor cell invasion by derivatives of a chemically modified tetracycline CMT–3 (COL–3). *Curr. Med. Chem.* **2001**, *8*, 271–279.