Available online at www.sciencedirect.com
SCIENCE DIRECT®
European Journal of Medicinal Chemistry 39 (2004) 905–916
EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY
www.elsevier.com/locate/ejmech

Original article

A topological sub-structural approach for predicting human intestinal absorption of drugs

Miguel Angel Cabrera Pérez [a,*], Marival Bermejo Sanz [b], Liliana Ramos Torres [a], Ricardo Grau Ávalos [c], Maykel Pérez González [a], Humberto González Díaz [a]

[a] *Department of Drug Design, Center of Chemical Bioactive, Central University of Las Villas, Santa Clara, 54830, Villa Clara, Cuba*
[b] *Department of Pharmacy and Pharmaceutical Technology, University of Valencia, Burjassot, 46100 Valencia, Spain*
[c] *Faculty of Mathematics, Physics and Computation, Central University of Las Villas, Santa Clara, 54830, Villa Clara, Cuba*

Received 12 January 2004; accepted 24 June 2004

Available online 28 September 2004

Abstract

The human intestinal absorption (HIA) of drugs was studied using a topological sub-structural approach (TOPS-MODE). The drugs were divided into three classes according to reported cutoff values for HIA. "Poor" absorption was defined as HIA ≤30%, "high" absorption as HIA ≥80%, whereas "moderate" absorption was defined between these two values (30% < HIA < 79%). Two linear discriminant analyses were carried out on a training set of 82 compounds. The percentages of correct classification, for both models, were 89.02%. The predictive power of the models were validated by three test: a leave-one-out cross validation procedure (88.9% and 87.9%), an external prediction set of 127 drugs (92.9% and 80.31%) and a test set of 109 oral drugs with bioavailability values reported (93.58% and 91.84%). Finally, positive and negative sub-structural contributions to the HIA were identified and their possibilities in the lead generation and optimization process were evaluated.

Keywords: Intestinal absorption; In silico; Sub-structural approach; Discriminant analysis

1. Introduction

The process of discovery and development of drugs is scientifically complex and full of risk, and consequently expensive and time-consuming. The average cost to discover and to develop a new chemical entity (NCE), is typically hundreds of millions of dollars and requires a decade or longer to reach the market-place [1].

Sub-optimal absorption, distribution, metabolism and excretion (ADME) properties are the major reason for the high attrition rates of compounds in development, where more than 90% of all candidates fail [2]. This problem has persisted due to difficulties in obtaining data on ADME properties early in drug discovery. On the other hand, the rate at which NCEs are generated is increasing dramatically using combinatorial chemistry and high-throughput pharmacological screening [3].

Today the traditional in vivo approaches, including the cassette-dosing techniques, have a limited throughput relative to the rate at which NCEs are produced [4]. Other alternative methodology such as in vitro Caco-2 assays offers higher throughput [5] but it is required to synthesize significant quantities of the compound and its correlation with the in vivo pharmacokinetic properties is still under development.

The "in silico" models provide an inexpensive and fast way to assess the ADME properties of a molecule before synthesis and they enable prioritization of molecules for in vitro and in vivo studies [6]. In many cases, these models can predict the properties of molecules based on structural information alone; permitting their use on "virtual" compounds [3].

One area of a visible effort in the application of the "in silico" models has been the prediction of oral absorption. Oral bioavailability is one of the most desirable attributes of a new drug, and the first step in order to obtain a high oral bioavailability is to achieve a good oral absorption. Several

* Corresponding author. Tel.: +53-42-281473, +53-42-281192; fax: +53-42-281130.
E-mail address: migue@gammu.com (M.A.C. Pérez).

0223-5234/$ - see front matter © 2004 Elsevier SAS. All rights reserved.
doi:10.1016/j.ejmech.2004.06.012

molecular descriptors [7–10] and computational methods for the prediction of oral absorption have been reported, being the hydrogen bonding capacity, lipophilicity, molecular size and the molecular surface properties the most important parameters related to this biological process [11–22]. The success of the computational models depends of the molecular descriptors selected to characterize the chemical structure and of the use of appropriate statistical methods. In the last few years the graph–theoretical methods have become a very useful tool for molecular design, especially in quantitative structure–property (QSPR) and quantitative structure–activity (QSAR) relationships [23].

The TOPS-MODE (Topological Sub-structural Molecular Design) approach [24] based on the calculation of the spectral moments of the bond matrix of molecular graph has been used to generate graph–theoretical descriptors with adequate flexibility and interpretability, expressing physical and biological properties in terms of sub-structural features of the molecules. This approach has been successfully applied to different QSPR and QSAR studies [25–33]. The aim of the present work is to find rationality in the prediction of the human intestinal absorption (HIA) using the TOPS-MODE method. This approach will be used not only in the classification of drugs with high, moderate and poor HIA but also in the evaluation and design of new drugs with improved absorption characteristics; using the information generated by the group contributions to the HIA values.

2. Materials and methods

2.1. The TOPS-MODE approach

The TOPS-MODE approach is based on the calculation of the spectral moments of the so-called bond matrix [34], whose theoretical basis has been described in previous reports [25–29]. Nevertheless, an overview of this approach will be given next.

The bond matrix is defined as a square and symmetric matrix whose entries are ones or zeros if the corresponding bonds are adjacent or not. The order of this matrix (m) is the number of bonds in the molecular graph, being two bonds adjacent if they are incident to a common atom. The spectral moments of the edge adjacency matrix are defined as the traces, i.e., the sum of the main diagonal, of the different powers of such matrix.

For the application of the present approach to structure–property relationship, the following steps should be conducted. First, to select an adequate training set with great structural diversity and to draw the hydrogen-depleted molecular graphs for each molecule. Second, to differentiate the molecular bonds with appropriate bond weights. Third, to compute the spectral moments of the bond matrix for each molecule of the data set. Fourth, to find a qualitative structure–pharmacokinetic relationship by using any discrimination statistical technique, such as linear discriminant analysis (LDA):

$$P = a_0\mu_0 + a_1\mu_2 + a_2\mu_2 + \ldots a_k\mu_k + b \quad (1)$$

where P is the studied property (in our case the HIA), μ_k is the kth spectral moment, and the a_k's are the coefficients obtained by LDA. Fifth, to evaluate the predictive power of the discriminant model with different validation procedures. Finally, to compute the contributions of different groups to the HIA values.

2.2. Selection of bond weights and calculation of molecular descriptors

Taking into consideration the QSAR paradigm for structure–permeability correlation used to evaluate the oral absorption [35], several bond weights such as hydrophobicity [36], polar surface area [37], molar refraction [38], Gasteiger–Marsilli atomic charge [39], polarizability [40], van der Waals atomic radii [41] and atomic mass were used for computing the spectral moments of the bond matrix. As most of the approaches to calculate physicochemical properties of fragments are based on atom-additive methods, several transform from atomic to bond contributions were carried out. The way in which these atomic contributions were transformed into bond contributions have been described by Estrada et al. [42]:

$$w(i,j) = \frac{w_i}{\delta_i} + \frac{w_j}{\delta_j} \quad (2)$$

where w_i and δ_i are the atomic weight and vertex degree of the atom i, respectively. The calculation of the TOPS-MODE descriptors was carried out with the computer software MODESLAB 1.0 [43] and using hydrophobicity, polar surface, Gasteiger–Marsilli charges, van der Waals radii, polarizability, molar refraction and atomic mass as bond weights. The input of the software consists of SMILES codes for each compound [44]. The first 10 spectral moments (μ_1–μ_{10}) for each bond weight and the number of bonds in the molecules (μ_0) were calculated. Also, considering the influence of partition coefficient and the polar surface area on intestinal absorption, as well as the non-linear relationship between these properties and the intestinal absorption, the interactions between μ_0 and μ_1 (weighted with hydrophobic and polar surface) were considered as variables.

2.3. Development and assessment of the discrimination functions for HIA

Drug absorption is a complex process which is dependent on several biochemical, physiological, and physicochemical factors. The inter-individual variation in the HIA measurements can be very high, depending on the formulation of the drug, the solubility of the compound in the intestinal fluid,

the dose-dependence, the metabolic transformations and the experimental methodology [45,46]. For these reasons, we selected the qualitative approach for this structure–property relationship study. Taking into consideration the high variability in the reported HIA values, our strategy was to discriminate between HIA ranges, through a LDA method. Consequently, the drugs were divided into three classes according to reported cutoff values for HIA. "Poor" absorption was defined as HIA ≤30%, "high" absorption as HIA ≥80%, whereas "moderate" absorption was defined between these two values (30% < HIA < 79%) [21,22,47].

The data set of 209 compounds was carefully selected from Zhao et al. [45] and Benet et al. [48], and randomly divided in two subsets, one containing 82 drugs (training set) and the other containing 127 compounds (evaluation set). The compounds belonging to the evaluation set were never used in the development of the discriminant functions and they were reserved to assess the final models. All compounds used were predominantly absorbed by passive diffusion on the intestinal wall (see Tables 1 and 2).

The first model was carried out to discriminate the drugs with poor intestinal absorption (HIA < 30%) from those with high-moderate absorption (HIA ≥30%), and the second one to separate compounds with high intestinal absorption (HIA ≥80%) from the rest of the drugs. Another set of 109 compounds, with reported oral bioavailability values (F), was also used to evaluate the quality of the discriminant models [48,49]. Because the value of HIA is always greater than F, the evaluated compounds had values of bioavailability that exceeded the 30%.

2.4. Statistical analysis

The discrimination functions were obtained using the stepwise LDA implemented in the STATISTICA version 5.5 [50]. The default parameters of this program were used for the development of the models. The quality of the models was determined considering the following statistical parameters: the Wilks λ, the Mahalanobis distance, the Fisher ratio and the number of variables in the equation. The λ parameter establishes a perfect discrimination for $\lambda = 0$ and not discrimination when $\lambda = 1$, whereas the Mahalanobis distance indicates the separation of the respective groups. The compounds were considered unclassified (U) by the model when the differences in the percentage of classification between two groups do not differ in more than 5%. However, we have used the posteriori probabilities in order to classify the compounds according to HIA values.

2.5. Computation of group contributions

The computation of group contributions was used to identify the quantitative contribution of a given substructure (negative or positive) to the intestinal absorption process. The general methodology used in this computational approach is as follows. In the first step all the substructure whose contribution we would like to determine were selected. The spectral moments for the substructures were calculated and their contributions to the intestinal absorption were obtained by substitution of their spectral moments in the discriminant models. In this study, 85 real substructures belonging to compounds with high, moderate and poor HIA were selected.

3. Results and discussion

The classification model (for drugs with poor HIA) obtained for the training set is given below together with the statistical parameters of the LDA:

$$\begin{aligned} \mathrm{HIA_L} &= 2.455 + 1.292\mu_1^H - 6.180\times10^{-4}\mu_0\mu_1^{PS} \\ &\quad - 2.489\times10^{-2}\mu_0\mu_1^H \\ N &= 82 \quad \lambda = 0.564 \quad D^2 = 4.91 \quad F_{exp}(3,78) = 20.06 \end{aligned} \tag{3}$$

where λ is the Wilks statistic, D^2 is the squared Mahalanobis distance and F_{exp} is the Fisher ratio. The super-indices H and PS represent the corresponding TOPS-MODE descriptors which were calculated weighting the bond with hydrophobicity and polar surface area, respectively. The statistical parameters show that the Eq. (3) was appropriate to discriminate between compounds with poor and high-moderate HIA.

The three variables in the previous model encoded specific structure information. The first variable ($\mu_1{}^H$) is related to the hydrophobicity or partition coefficient and it has a positive contribution to the HIA values. It is known that absorption drops at both low and high log P and the main reasons of these non-linearties are due to: (1) the inability of weakly lipophilic compounds to penetrate the lipid portion of the membrane and (2) the excessive partitioning of strongly lipophilic compounds into the lipid portion of the membrane and their subsequent inability to pass through the aqueous portion of the membrane [51]. However, until excessive partition values are not reached, increments of log P will increase the HIA values. The other variables ($\mu_0\mu_1{}^{PS}$ and $\mu_0\mu_1{}^H$), with negative contributions to the HIA values, represent the interaction between polar surface area and the hydrophobicity with the number of bonds in the molecules (μ_0). The spectral moment of first order (μ_0) has information about the size of the molecules because it is expressed as linear combination of bond contributions. These results are in agreement with the influence that molecular size has on the intestinal permeability [21]. It is well known that the lipophilicity and the hydrogen bonding capacity, explained by the polar surface area, are two important physicochemical factors related, in a non-linear way, with the passive absorption [52]. For this reason any increment in donors or acceptors hydrogen groups or aliphatic chains will increase the molecular size, reducing the capacity of drug to permeate the biological membranes.

In the Table 1 is illustrated the results in the classification of compounds belonging to the training set.

Table 1
Compounds used and qualitative results obtained in the prediction of the HIA

No	Drug	HIA (%)	Poor HIA model Eq. (3)		High HIA Model Eq. (4)	
			Prob	Class [a]	Prob	Class [a]
Group I (high HIA)						
1	Caffeine	100	93	H-M	93.5	H
2	Desipramine	100	99.5	H-M	98.4	H
3	Diazepam	100	99.1	H-M	99.6	H
4	Ondansetron	100	97.8	H-M	98.6	H
5	Testosterone	100	96.7	H-M	96.5	H
6	Cefadroxil	100	42.6 [b]	P	20.7 [b]	M-P
7	Ofloxacin	100	95.4	H-M	91.6	H
8	Pefloxacin	100	97.1	H-M	93.2	H
9	Naproxen	99	98.7	H-M	93.8	H
10	Prednisolone	99	67.5	H-M	85.4	H
11	Warfarin	98	97.1	H-M	93.5	H
12	Antipyrine	97	97.3	H-M	98.9	H
13	Trimethoprin	97	82.2	H-M	58.2	H
14	Bumetanide	96	90.4	H-M	15.7 [b]	M-P
15	Ibuprofen	95	99.1	H-M	92	H
16	Practolol	95	94.8	H-M	66.4	H
17	Clonidine	95	99.3	H-M	98.7	H
18	Fluconazole	95	47.3 [b]	P	99.7	H
19	Alprenolol	93	98.9	H-M	85.1	H
20	Ketoprofen	92	97.9	H-M	92.9	H
21	Hydrocortisone	91	60.9	H-M	87.4	H
22	Betaxolol	90	98.6	H-M	69.8	H
23	Chloramphenicol	90	86.1	H-M	88.9	H
24	Tenidap	89	89.8	H-M	88.6	H
25	Oxazepam	89	96.8	H-M	97.3	H
26	Sultopride	89	94.4	H-M	86.1	H
27	Felodipine	88	98.2	H-M	98.5	H
28	Acrivastine	88	98.4	H-M	91.8	H
29	Trovafloxacin	88	77	H-M	34.4 [b]	M-P
30	Pindolol	87	97.2	H-M	83.2	H
31	Bupropion	87	99.5	H-M	98.1	H
32	Topiramate	86	70.7	H-M	35.0 [b]	M-P
33	Tolbutamide	85	96.2	H-M	68.2	H
34	Acetylsalicylic acid	84	96.5	H-M	83.5	H
35	Bromazepam	84	95.8	H-M	99.8	H
36	Captopril	84	92.3	H-M	31.7	M-P
37	Sorivudine	82	57.2	H-M	70.9	H
38	Methylprednisolone	82	69.7	H-M	84.3	H
39	Quinidine	81	96.5	H-M	97.8	H
40	Dexamethasone	80	71	H-M	88.2	H
41	Acebutolol	80	92.8	H-M	59.8	H
42	Acetaminophen	80	95.3	H-M	83.8	H
43	Guanabenz	80	99.5	H-M	63.1	H
Group 2 (moderate HIA)						
44	Mercaptoethanesulfonicacid	77	86.3	H-M	96.6	M-P
45	Famciclovir	77	70	H-M	39.1 [b]	H
46	Ceftizoxime	72	44.5 [b]	P	97.4	M-P
47	Cimetidine	64	86.5	H-M	61.3	M-P

(continued on next page)

Table 1
(continued)

No	Drug	HIA (%)	Poor HIA model Eq. (3)		High HIA Model Eq. (4)	
			Prob	Class [a]	Prob	Class [a]
48	Metolazone	64	95	H-M	183 [b]	H
49	Ampicillin	62	55.9	H-M	58.6	M-P
50	Furosemide	61	87.5	H-M	74.7	M-P
51	Fenoterol	60	94.3	H-M	84.8	M-P
52	Nadolol	57	90.2	H-M	60	M-P
53	Tranexamic acid	55	93.1	H-M	60.7	M-P
54	Eflornithine	55	82.8	H-M	87.3	M-P
55	Metformin	53	93.2	H-M	96.4	M-P
56	Atenolol	50	90.8	H-M	58.8	M-P
57	Rimiterol	48	94.3	H-M	64.7	M-P
58	Cymarin	47	28.6 [b]	P	52	U
59	Sulpiride	44	87.7	H-M	51	U
60	Metaproterenol	44	95	H-M	70.8	M-P
61	Methyldopa	41	82.8	H-M	90.9	M-P
62	Famotidine	38	9.7 [b]	P	99.9	M-P
63	Ascorbic acid	35	61.7	H-M	95.2	M-P
64	AAFC	32	70.7	H-M	46.5 [b]	H
65	Fosfomycin	31	80.4	H-M	80.6	M-P
66	Fosmidomycin	30	61.5	H-M	96.9	M-P
Group 3 (poor HIA)						
67	Lincomycin	28	36.6	P	92.5	M-P
68	Netivudine	28	47.4	P	64	M-P
69	Foscarnet sodium	17	63.3 [b]	H-M	99.9	M-P
70	Adefovir	16	36.7	P	93.6	M-P
71	Tacrolimus [c]	15	1.1	P	72.1	M-P
72	Enalaprilate	10	85.7 [b]	H-M	62.1	M-P
73	Cefuroxime	5	73.3 [b]	H-M	55.8	M-P
74	Albendazole [c]	5	96.5 [b]	H-M	29.0 [b]	H
75	Cidofovir	3	32.1	P	99	M-P
76	Ganciclovir	3	43.1	P	95	M-P
77	Acarbose	2	0	P	96.8	M-P
78	Ouabain	1.4	2.3	P	90.6	M-P
79	Kanamycin	1	0.1	P	99.9	M-P
80	Neomycin	1	0	P	99.6	M-P
81	Lactulose	0.6	1.9	P	99.7	M-P
82	Raffinose	0.3	0.2	P	99.6	M-P

[a] Classification of HIA range H: high; P: poor and U: unclassified; H-M: high-moderate, M-P: moderate-poor.
[b] Wrong classified compounds.
[c] Data selected from Goodman and Gilman's [47].

The model classified correctly the 92.42% of drugs with high-moderate HIA values in the training set while 75.00% of the compounds with poor HIA were also well classified. The global correct classification for the Eq. (3) was 89.02%. The percentages of false positives and false negatives in the training set were 4.9% (4/82) and 6.1% (5/82), respectively. False positives are those compounds with poor HIA that were classified, by the model, as high-moderate HIA compounds, whilst the false negatives are those compounds with high-moderate HIA that were classified as having poor HIA (see Table 1). From a practical point of view it is considered more important avoiding false negatives because those are the compounds that will be rejected due to their poor predicted intestinal absorption and therefore they will never be evaluated experimentally, and their true absorption would never be discovered. On the contrary, the false positive compounds eventually will be detected.

The most important criterion for the quality of the discriminant model is based on the statistics of the external evaluation set. The Eq. (3) classified correctly the 95.45% and 76.47% of compounds with high-moderate and poor HIA, respectively. The global classification was 92.91%. The percentages of false negative and false positive compounds were 3.9% (5/127) and 3.1% (4/127), respectively. In Table 2 the classification of compounds in the external evaluation set is presented. Also, a cross-validation leave-one-out procedure

Table 2
Data used and results of the external prediction set

Compounds	HIA (%)	Class [a]	Prob	Class [b]	Prob	Compounds	HIA (%)	Class [a]	Prob	Class [b]	Prob
Valproic acid	100	H-M	98.6	H	79	Oxprenolol	95	H-M	98.5	H	79.2
Corticosterone	100	H-M	80.3	H	93.5	Timolol	95	H-M	94.8	M-P [c]	20.8
Loracarbef	100	H-M	75.2	H	81	Codeine	95	H-M	93.4	H	99.3
Lormetazepam	100	H-M	98.7	H	99.1	Flumazenil	95	H-M	95.6	H	98.2
Salicylic acid	100	H-M	96.7	H	65.2	Amrinone	93	H-M	86.1	H	86.8
Cephalexin	100	H-M	59.5	M-P [c]	46.5	Naloxone	91	H-M	85.2	H	97.1
Imipramine	100	H-M	99.6	H	99.3	Phenytoin	90	H-M	96.8	H	94.6
Progesterone	100	H-M	97.1	H	98.8	Terazocin	90	H-M	81.8	H	84.4
Theophylline	100	H-M	90.6	H	87.2	Sulindac	90	H-M	94	H	98.6
Verapamil	100	H-M	97.1	H	71.9	Ketorolac	90	H-M	96.5	H	91.2
Cisapride	100	H-M	93.1	H	72.3	Meloxicam	90	H-M	79.9	U	52.5
Sudoxicam	100	H-M	78.3	H	54.2	Amphetamine	90	H-M	98.4	H	91.8
Glyburide	100	H-M	82.4	M-P [b]	39.8	Felbamate	90	H-M	88.6	M-P [c]	20.5
Gallopamil	100	H-M	94.9	H	64.1	Nizatidine	90	H-M	80.6	H	53.6
Mexiletine	100	H-M	98.8	H	90.7	Alprazolam	90	H-M	98.3	H	99.7
Nefazodone	100	H-M	96.9	H	92	Tramadol	90	H-M	98.3	H	95.7
Camazepam	100	H-M	97.4	H	98.9	Nisoldipine	90	H-M	91.9	H	88.4
Indomethacin	100	H-M	98.4	H	91.6	Tiagabine	90	H-M	96.7	M-P [c]	46.6
Levonorgestrel	100	H-M	96.6	H	97.8	Dihydrocodeine	89	H-M	94.4	H	98.7
Tenoxicam	100	H-M	81	M-P [c]	44.5	Nitrendipine	88	H-M	91.1	H	92.3
Oxatomide	100	H-M	98.2	H	86.6[b]	Saccharin	88	H-M	91	H	89.7
Desipramine	100	H-M	99.5	H	98.4	Moxonidine	88	H-M	92.4	H	93.9
Fenclofenac	100	H-M	99.4	H	98.6	Nicotinic acid	88	H-M	93	H	83.3
Diclofenac	100	H-M	99.2	H	93.5	Lamivudine	87	H-M	77.4	M-P [c]	18.7
Granisetron	100	H-M	93.9	H	98.9	Levodopa	86	H-M	82.2	M-P [c]	7.6
Ethinylestradi	100	H-M	98.2	H	91.6	Morphine	85	H-M	91.1	H	97.7
Isoxicam	100	H-M	87	H	58.5	Lansoprazole	85	H-M	87	H	99.3
Lornoxicam	100	H-M	85.9	H	59.6	Oxyfedrine	85	H-M	97.6	H	86.7
Nicotine	100	H-M	97.6	H	98.8	Viomycin	85	P [c]	0	M-P [c]	0.1
Piroxicam	100	H-M	90.7	H	71.8	Propiverine	84	H-M	99.2	H	91.2
Stavudine	100	H-M	85.8	H	76.3	Mifobate	82	H-M	94.5	H	72.3
Toremifene	100	H-M	99.7	H	97.6	Digoxin	81	P [c]	0.4	M-P [c]	2.9
Cyproterone	100	H-M	91.9	H	99.4	Flecainide	81	H-M	98.7	H	89.5
Praziquantel	100	H-M	96.7	H	99	Piroximone	81	H-M	87	H	92.7
Cicaprost	100	H-M	93.3	H	90	Ethambutol	80	H-M	93.5	M-P [c]	21
Aminopyrine	100	H-M	97.2	H	99.3	Isoniazid	80	H-M	76.4	H	61.2
Propranolol	99	H-M	99.2	H	86.4	Omeprazole	80	H-M	73.3	H	98.9
Nordiazepam	99	H-M	98.3	H	99.4	Methadone	80	H-M	99.5	H	98.2
Carfecillin	99	H-M	52.7	M-P [b]	34.4	Urapidil	78	H-M	97.6	H [c]	60.1
Lamotrigine	98	H-M	94.9	H	89.1	Propylthiouracil	76	H-M	92.1	H [c]	74.4
Tolmesoxide	98	H-M	92.3	H	97	Cycloserine	73	H-M	72.7	M-P	34.8
Viloxazine	98	H-M	98.9	H	91.1	Spironolactone	73	H-M	88.2	H [c]	92.6
Atropine	98	H-M	96.3	H	94.3	Recainam	71	H-M	99.6	H [c]	81
Minoxidil	98	H-M	95	M-P [c]	44.3	Hydrochlorothizide	69	H-M	76.7	M-P	37.8
Ximoprofen	98	H-M	95.9	H	71.2.	Ziprasidone	60	H-M	96.9	H [c]	92.1
Disulfiram	97	H-M	99.5	M-P [c]	28.4	Reproterol	60	H-M	67.5	M-P	16.7
Clofibrate	97	H-M	99.5	H	98.2	Pirbuterol	60	H-M	85.1	M-P	26.9
Venlafaxine	97	H-M	98.6	H	94.6	Phenoxymethylpenic	*59*	H-M	75.2	H [c]	65
Torasemide	96	H-M	85.3	H	71.5	Vigabatrin	58	H-M	90	M-P	34.2
Trapidil	96	H-M	97.2	H	96.9	Sumatriptan	57	H-M	90	H [c]	95.9
Labetalol	95	H-M	92.9	M-P [c]	25.8	Amiloride	50	H-M	57.9	M-P	1
Metoprolol	95	H-M	97.8	H	76.9	Guanoxon	50	H-M	96.4	M-P	29
Sotalol	95	H-M	91.6	H	5G.6	Capreomycin	50	P [b]	0	M-P	0.2
Cefetamet pivoxil	47	P [c]	22.9	M-P	7.3	Iothalamatesodium	1.9	H-M [c]	98.3	H [c]	100
Azithromycin	37	P [c]	2.6	M-P	1.6	Streptomycin	1	P	0	M-P	0
Fosinopril	36	H-M	68.4	M-P	4.2	Aztreonam	1	P	9.1	M-P	0.7

(continued on next page)

Table 2
(continued)

Compounds	HIA (%)	Class [a]	Prob	Class [b]	Prob	Compounds	HIA (%)	Class [a]	Prob	Class [b]	Prob
Pravastatin	34	H-M	64.4	M-P	5.3	Cromolyn	0.5	P	39.3	M-P	3.8
Cyclosporin	28	P	0	M-P	0.1	Gentamicin	0	P	3.2	M-P	2.2
Bromocriptine	28	P	30.8	H [b]	98.4	Cefotaxime [a]	0	P	8.7	M-P	5.9
Mannitol	16	P	14.4	M-P	0.3	Imipenem [a]	0	H-M [b]	71.3	M-P	7.2
k-strophanthoside	16	P	0	H [c]	53.8	Vancomycin [a]	0	P	0	M-P	9.2
Distigmine bromide	8	H-M [c]	97.5	H [c]	99.8	Paromomycin [a]	0	P	0	M-P	0.6
Doxorubicin	5	P	3.7	M-P	25.9	Amphotericin B	0	P	0	M-P	0.6
Olsalazine	2.3	H-M [c]	91	M-P	3.4						

Classification of HIA range H: high; P: poor and U: unclassified; H-M: high-moderate; M-P: moderate-poor.
[a] Classification of HIA by Eq. (3).
[b] Classification of HIA by Eq. (4).
[c] Wrong classified compounds.

(LOO) was carried out to confirm the reliability of the model, being the percentage of correct classification 87.8%. Only one compound was unclassified, for a global classification of 88.9%. Both validation procedures evidenced the predictive power of the Eq. (3).

The classification model to discriminate the drugs with high HIA from those with moderate-poor HIA in the training set is given below, along with the statistical parameters of the LDA. The classification results are illustrated in Table 1.

$$\mathrm{HIA_H} = 2.923 - 0.080\mu_1^{\mathrm{PS}} + 0.069\mu_1^{\mathrm{MR}} + 3.22\mu_1^{\mathrm{AC}} - 0.022\mu_1^{\mathrm{AC}} - 0.037\mu_0\mu_1^{H} \quad (4)$$

$$N = 82 \quad \lambda = 0.473 \quad D^2 = 4.46 \quad F(5, 76) = 16.92$$

As can be observed in Eq. (4) some descriptors related to polar surface (μ_1^{PS}), molar refraction (μ_1^{MR}), atomic charge (μ_1^{AC}), atomic mass (μ_1^{MA}) and the interaction between the partition coefficient (hydrophobicity) and the number of bonds in the molecules ($\mu_0\mu_1$) have a direct effect on the high HIA values. In this equation the variables, weighted with the polar surface and the atomic mass, have a negative contribution to the HIA. It is a logical result because both properties are not independent variables. The size of a compound could be increased for addition of carbon or halogens atoms, leading to a higher partition coefficient or for addition of heteroatoms (N, O), leading to a higher hydrogen bonding capacity. High values of lipophilicity could cause poor solubility while a high H-bonding may result in unfavorable membrane permeability. Both properties tend to affect the intestinal absorption process [13,53]. The μ_1 spectral moment weighted with Gasteiger–Marsilli atomic charge has a positive regression coefficient. Nevertheless, this descriptor has a negative value in all the evaluated drugs, therefore when its magnitude is increased (molecules with heteroatoms, mainly N and O) the intestinal absorption is decreased. Finally, the spectral moment weighted with the molar refraction has a positive contribution to the HIA. This property is related with the molecular volume and molecular weight, for this reason although it has a positive contribution to the HIA, extreme values of this property will decrease the absorption of drugs. Even though each variable was analyzed separately, the interrelation among them will explain the general effect on the HIA.

The model classified correctly the 88.37% of drugs with high HIA values and 89.74% the compounds with moderate-poor HIA in the training set, for a global correct classification of 89.02%. The overall accuracy of the model was 91.25% (73/80). The percentages of false positive and false negative compounds in the training set were 4.89% (4/82) and 6.10% (5/82), respectively (see Table 1).

The Eq. (4) classified correctly the 83.51% and 72.22% of compounds with high and moderate-poor HIA in the evaluation set, respectively, for a global classification of 80.31%. The percentage of false negative and positive compounds were 11.81% (15/127) and 7.87% (10/127), respectively. The overall prediction was 80.95%. The cross validation leave-one-out procedure showed an 87.88% of correct classification. In Table 2 the classification of compounds in the external evaluation set is detailed.

A second assessment of the predictive power of both discriminant models was carried out over 109 compounds with reported oral bioavailability (*F*). The results are depicted in Table 3.

The Eq. (3) (model for poor HIA) classified correctly the 93.58% of drugs as high-moderate HIA. The percentage of compounds with false negatives in the data set was 6.42% (7/109). In order to validate the Eq. (4) (model for high HIA), **51** compounds with $F \geq 80\%$ were considered from the entire data set. The percentage of correct classification was 88.23% (45/51) and the general accuracy of the model was 91.84% (45/49). As can be appreciated, both discriminant models were able to classify compounds inside the selected range of HIA values. One of the main advantages of the present approach in the molecular design field is the possibility to obtain the quantitative contribution of any kind of substructure to the studied property. Several "in silico" methods to predict the oral intestinal absorption have been developed [16–20] but in few of them have been evaluated the influence of structural fragments. By the present approach, we have generated a number of substructures and determined if they contribute positively or negatively to the HIA values. In the

Table 3
Results of the validation of the discriminant models using reported bioavailability (*F*) values

Compounds	*F* (%)	Class [a]	Prob	Class [b]	Prob	Compounds	*F* (%)	Class [a]	Prob	Class [b]	Prob
Chlordiazepoxide	100	H-M	97.1	H	99.8	Trazodone	75	H-M	98.7		
Isosorbide mononitrate	100	H-M	55.4	H	85.7	Amlodipine	74	H-M	89.9		
Phenobarbital	100	H-M	92.8	H	85	Allopurinol	74	H-M	95		
Probenecid	100	H-M	98.7	H	56.1	Cyclophosphamide	74	H-M	99		
Pyrimethamine	100	H-M	97.2	H	79.2	Betamethasone	72	H-M	71		
Sulfadiazine	100	H-M	77.6	H	66	Cinoxacin	72	H-M	83.7		
Sulfamethoxazole	100	H-M	90.2	M-P [c]	46.5	Melphalan	71	H-M	98.5		
Clobazam	100	H-M	98.9	H	99.4	Amantadine	70	H-M	98		
Metronidazole	99	H-M	93.2	H	76.3	Dicloxacillin	70	H-M	74.3		
Clonazepam	98	H-M	95.5	H	99.2	Fluoxetine	70	H-M	99.7		
Lomefloxacin	97	H-M	96.5	H	89.8	Ritonavir	70	P [c]	1.4		
Oxaprozin	97	H-M	98.5	H	83.9	Zolpidem	67	H-M	98.9		
Sulfisoxazole	96	H-M	90.9	U	51.2	Risperidone	66	H-M	93.4		
Chlorpropamide	95	H-M	96.7	H	83.3	Chlorthalidone	64	H-M	85.1		
Glimepiride	95	H-M	64.6	M-P [c]	37.6	Norethindrone	64	H-M	96		
Glipizide	95	H-M	57.8	H	56.6	Doxazosin	63	H-M	73.7		
Hexobarbital	95	H-M	91.7	H	93.2	Finasteride	63	H-M	96.2		
Tinidazole	95	H-M	88.2	H	87.8	Diphenhydramine	61	H-M	99.5		
Tolmetin	95	H-M	98.4	H	86.3	Bepridil	60	H-M	99.7		
Ifosfamide	94	H-M	99.3	H	93.4	Haloperidol	60	H-M	99		
Flurazepam	93	H-M	99.3	H	98.8	Paroxetine	60	H-M	99.1		
Dapsone	93	H-M	92.8	H	56.4	Clofazimine	57	H-M	98.5		
Lorazepam	93	H-M	97.7	H	98.7	Cocaine	57	H-M	96.2		
Flurbiprofen	92	H-M	99.4	H	96.6	Clarithromycin	55	P [c]	2.4		
Primidone	92	H-M	96.3	H	87.8	Clozapine	55	H-M	99.2		
Bisoprolol	91	H-M	98.1	H	58.5	Itraconazole	55	P [c]	16.9		
Diazoxide	91	H-M	95.3	H	96.9	Triamterene	54	H-M	70.7		
Temazepam	91	H-M	98.3	H	98.3	Valaciclovir	53	H-M	58.4		
Diflunisal	90	H-M	99	H	91.2	Clomipramine	52	H-M	99.7		
Mefloquine	90	H-M	98.1	H	99.1	Meperidine	52	H-M	99.3		
Nitrofurantoin	90	H-M	89.5	H	69	Nortriptyline	51	H-M	99.6		
Penbutolol	90	H-M	99.2	H	77.5	Ethinyl estradiol	51	H-M	98.2		
Phenylbutazone	90	H-M	98	H	99.2	Albuterol	50	H-M	94.3		
Chloroquine	89	H-M	99.6	H	98	Nifedipine	50	H-M	88.2		
Tocainide	89	H-M	95.5	H	84.2	Amitriptyline	48	H-M	99.7		
Zalcitabine	88	H-M	82.8	M-P [c]	36.9	Cefixime	47	P [c]	6.9		
Chlorambucil	87	H-M	99.6	H	97.5	Pentazocine	47	H-M	98.9		
Clindamycin	87	H-M	64.2	M-P [c]	30.9	Quinapril	47	H-M	87.8		
Enoxacin	87	H-M	96.4	H	67.5	Amiodarane	46	H-M	99.7		
Carteolol	85	H-M	93.9	H	77.4	Trimetrexate	45	H-M	77.3		
Etodolac	85	H-M	99.2	H	98.1	Midazolam	44	H-M	99.1		
Milrinone	85	H-M	88.9	H	97	Ramipril	44 [c]	H-M	89.3		
Misoprostol	85	H-M	93.2	M-P [c]	44.6	Triazolam	44	H-M	98.7		
Protriptyline	85	H-M	99.6	H	99	Chlorpromazine	40	H-M	99.7		
Flucytosine	84	H-M	85.7	U	51.9	Pimozide	40	H-M	98.2		
Disopyramide	83	H-M	98.1	H	80.8	Phenylephrine	38	H-M	94.6		
Procainamide	83	H-M	98.1	H	74.3	Nafcillin	36	H-M	91.5		
Carbamazepine	80	H-M	98.6	H	96.2	Losartan	35	H-M	96.2		
Pheny1propanolamine	80	H-M	95.9	H	70.6	Didanoside	35	H-M	80.1		
Prednisona	80	H-M	63.5	H	96.7	Nabumetone	35	H-M	99.1		
Rifampicin	80	P [c]	0.1	M-P [c]	0.2	Chlorpheniramine	33	H-M	99.4		
Nitrazepam	78	H-M	93.8			Oxacillin	33	H-M	67.5		
Tetracycline	77	P [c]	9			Indinavir	30	P [c]	29.7		
Metoclopramide	76	H-M	98.1			Ritodrine	30	H-M	97		
Primaquine	75	H-M	96.5								

Classification of compounds H: high; P: poor and U: unclassified; H-M: high-moderate; M-P: moderate-poor.

[a] Classification of HIA by Eq. (3).

[b] Classification of HIA by Eq. (4).

[c] Wrong classified compounds.

Table 4
Contribution of selected substructures to the HIA values

Substructures	Contribution	Substructures	Contribution
F_1	1.38	F_{44}	–0.31
F_2	–0.2	F_{45}	1.16
F_3	1.81	F_{46}	0.08
F_4	0.48	F_{47}	0.74
F_5	–0.71	F_{48}	1.18
F_6	0.45	F_{49}	0.81
F_7	0.12	F_{50}	0.98
F_8	0.62	F_{51}	0.78
F_9	0.16	F_{52}	1.01
F_{10}	0.73	F_{53}	0.77
F_{11}	–1.3	F_{54}	0.61
F_{12}	1.95	F_{55}	0.89
F_{13}	0.88	F_{56}	1.41
F_{14}	0.12	F_{57}	0.44
F_{15}	–1.02	F_{58}	–0.74
F_{16}	0.03	F_{59}	–1.18
F_{17}	1.16	F_{60}	–1
F_{18}	1.15	F_{61}	–0.31
F_{19}	–1.4	F_{62}	1.02
F_{20}	1.69	F_{63}	0.94
F_{21}	0.45	F_{64}	0.28
F_{22}	0.33	F_{65}	–0.55
F_{23}	1.18	F_{66}	–0.49
F_{24}	1.2	F_{67}	1.47
F_{25}	–1.31	F_{68}	–0.56
F_{26}	0.98	F_{69}	–3.85
F_{27}	1.28	F_{70}	0.07
F_{28}	0.06	F_{71}	0.7
F_{29}	0.27	F_{72}	0.3
F_{30}	–0.04	F_{73}	–1.01
F_{31}	–1.46	F_{74}	–1.8
F_{32}	1.27	F_{75}	0.53
F_{33}	–1.41	F_{76}	–0.42
F_{34}	0.6	F_{77}	0.95
F_{35}	–0.47	F_{78}	–1.08
F_{36}	0.68	F_{79}	1.37
F_{37}	0.83	F_{80}	0.42
F_{38}	–3.24	F_{81}	–0.4
F_{39}	1.17	F_{82}	–2.2
F_{40}	–0.34	F_{83}	–2.01
F_{41}	–1.13	F_{84}	1.11
F_{42}	–0.48	F_{85}	1.47
F_{43}	–0.11		

Fig. 1 the 85 selected substructures are shown. The contribution to the HIA values of these substructures were computed by the Eq. (4), considering that this predictive model was established to discriminate the high HIA compounds from those with moderate-poor HIA values. The quantitative values of these contributions are given in Table 4.

A detailed analysis of all selected substructures gives us a better idea in order to explain the drug absorption behavior or to design new candidates with good absorption profiles. Some considerations are the following:

(a) when the aliphatic chain is increased, the contribution to the HIA value is lowered ($F_{61} > F_{58}$; $F_{16} > F_{81}$; $F_{13} > F_{14}$; $F_{20} > F_{79}$). Nevertheless, the substructures with lineal aliphatic chains have a positive contribution to the HIA values and their magnitudes are increased when heteroatomic substitutions are carried out ($F_{84} > F_{37} > F_{51}$).
(b) The number of unsaturations and heteroatomic substitution decrease the contributions to the HIA value ($F_{61} > F_{78} > F_{33}$).
(c) When the polarity and the atomic mass of the substituents are increased, the contribution to the HIA value is decreased ($F_{85} > F_1 > F_{27}$; $F_{12} > F_8 > F_{44}$; $F_{67} > F_9 > F_{66} > F_5$).

Four examples of the importance of the substructure contributions to the HIA value will be discussed next. Sulbactam, a penicillanic acid compound not used in our study, has a HIA value of 5% [53].

In this drug (see Fig. 2) it can be identified some substructures such as: F_{59}, F_{61}, F_{73}, F_{74}, F_{81} and F_{82} which have a negative contribution to the HIA value (see Table 4). In the case of Aztreonam (see Fig. 2), a monocyclic β-lactam compound used in our study, the HIA is 1%. This drug also has several substructures such as: F_{58}, F_{61}, F_{65}, F_{74}, F_{81}, F_{82} and F_{83} with a negative contribution to the HIA value (see Table 4). The last substructure (F_{83}) belongs to the sulphonamide group and it has a highly significant effect in reducing oral bioavailability by *N*-acetylation and *N*-acetyltransferase metabolic reactions [22]. Also two compounds (chlorphenesin and guanfacine), with high HIA values, were used to explain the role of the substructures in the absorption process. As can be seen in Table 4 and Fig. 2 both drugs contain different substructures (F_1, F_{17}, F_{37}, F_{71}, F_{75} and F_{16}, F_{17}, F_{20}, F_{67}, respectively) with a positive impact on the HIA value. It is evident in all the examples that only the pointed out substructures are not the responsible of the poor or high HIA value, because the global effect is the sum of the contributions of all the fragments in the molecules. Moreover, these results evidenced the transferability of found relationships onto compounds that are not included into the training and evaluation sets.

If the previous approach is considered correct, could we design any hypothetical drug with high HIA value? In order to do that, we selected different substructures with positive contribution to the HIA value (F_1, F_{71}, F_{79}) and the hypothetical drug was designed. However, in order to get a general prediction of the absorption the new candidate was evaluated with both discriminant models and the results indicated that the compound has a high intestinal absorption (see Fig. 2).

Finally, the present approach not only permits the correct classification of heterogeneous dataset according to the HIA ranges but also the impact of various structural modifications on compounds' intestinal absorption can be easily visualized. In this sense, the present work represents a step forward in the development of a theory that permits not only the very fast prediction of biopharmaceutical properties but also the easy interpretation of the results in terms of structural features.

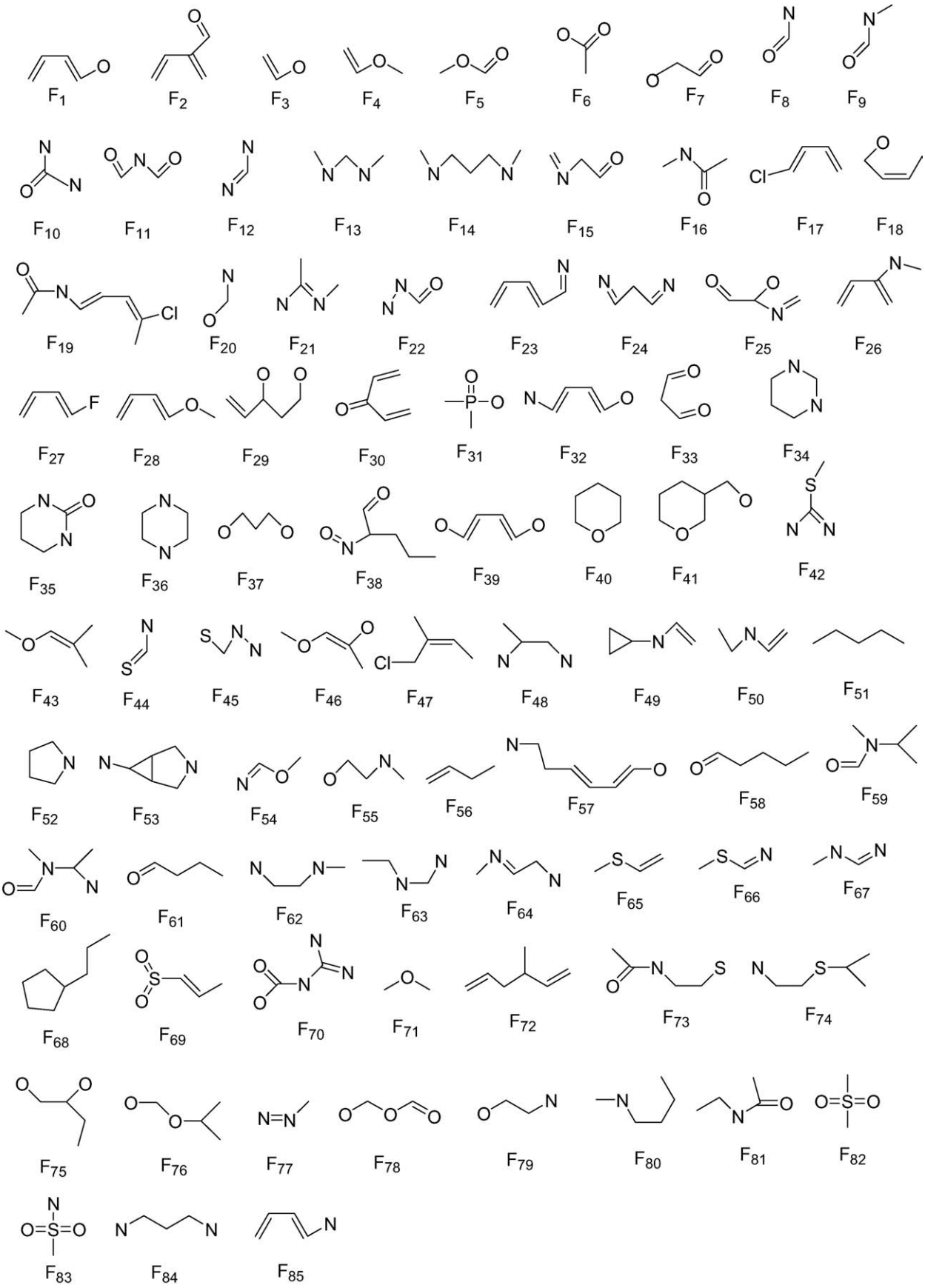
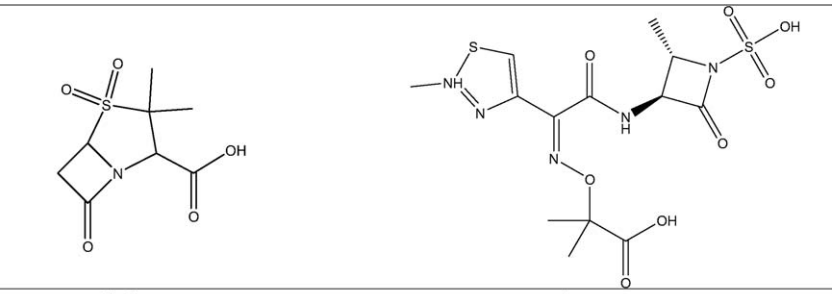
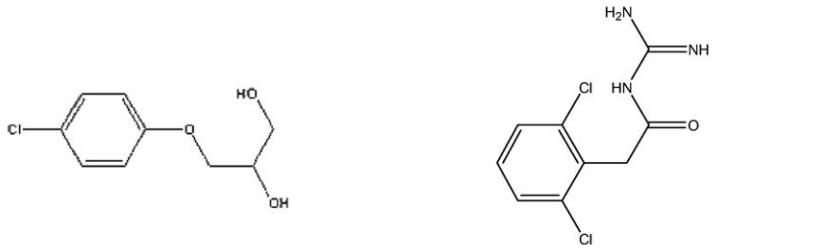


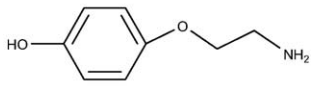
Fig. 1. Structures of selected groups for which their contributions to the HIA are evaluated.



Sulbactam		Aztreonam	
Exp. % HIA = 5	*Poor*	*Exp. % HIA = 1*	*Poor*
Pred. % HIA (Eq.3)	*Poor*	*Pred. % HIA (Eq. 3)*	*Poor*



Chlorphenesin		Guanfacine	
Exp. % HIA = 100	*High*	*Exp. % HIA = 100*	*High*
Pred. % HIA (Eq.4)	*High*	*Pred. % HIA (Eq.4)*	*High*



Hypothetical Drug	
Pred. % HIA (Eq.3)	*High*
Pred. % HIA (Eq.4)	*High*

Fig. 2. Chemical structure of drugs with poor and high HIA.

4. Conclusion

The prediction of HIA has been a goal of pharmaceutical companies due to the importance of this property during the drug development process. For this reason, several in silico methods have been applied in order to predict this property in the early stage of drug development and some of them have become in important tools to select novel chemical compounds for the oral administration. In this study, a topological sub-structural approach has been successfully applied to classify compounds with poor, moderate and high HIA. The procedure has showed that good linear discriminant functions can be obtained using the polar surface, hydrophobicity, molar refraction, atomic charge and atomic mass as weights in the diagonal entries of the bond matrix. The classification of drugs with high, moderate or poor HIA from large dataset can be obtained in a very short time. On the other hand, the present approach permits the identification and quantification of the structural fragments that are responsible for the HIA value. This possibility will play an important role in the generation of new chemical entities with good absorption properties as well as in the optimization process. Refinement of the model is possible with the incorporation of more HIA data, and its power in the design of new drugs with good absorption properties could be improved with the construction and evaluation of large dataset of sub-structural fragments.

Acknowledgements

The authors acknowledge to the owners of the Software MODESLAB 1.0 for their free contribution to the development of this work.

References

[1] J.A. Dimasi, Regul. Toxicol. Pharmacol. 19 (1994) 228–235.
[2] J.A. Dimasi, Clin. Pharmacol. Ther. 69 (2001) 297–307.
[3] P. Stenberg, K. Luthman, P. Artursson, J. Contr. Rel. 65 (2001) 231–243.
[4] M.H. Tarbit, J. Berman, Curr. Opin. Chem. Biol. 2 (1998) 411–416.
[5] P.J. Eddershaw, M. Dickins, Pharm. Sci.Technol. Today 2 (1999) 13–19.
[6] W. Egan, G. Lauri, Adv. Drug Deliv. Rev. 54 (5 54) (2002) 273–289.
[7] L.H. Hall, B. Mohney, L.B. Kier, J. Chem. Inf. Comput. Sci. 31 (1991) 76–82.

[8] B. Lee, F.M. Richards, J. Mol. Biol. 55 (1971) 379–400.
[9] B. Testa, P.A. Carrupt, P. Gaillard, F. Billois, P. Weber, Pharm. Res. 13 (1996) 335–343.
[10] U. Norinder, T. Osterberg, P. Artusson, Eur. J. Pharm. Sci. 8 (1999) 49–56.
[11] U. Norinder, T. Osterberg, Perspect. Delivery Rev 23 (1997) 221–227.
[12] H. Van de Waterbeemd, in: H. Van de Waterbeemd, G. Folkers, R. Guy (Eds.), Pharmacokinetic Optimization in Drug Research; Biological, Physicochemical and Computational Strategies, Verlag HCA, Basel, 2001, pp. 499–511.
[13] C.A. Lipinsky, F. Lombardo, B.W. Dominy, P.J. Feeney, Adv. Drug Deliv. Rev. 23 (1997) 3–25.
[14] A.K. Ghose, V.N. Viswanadhan, J.J. Wendoloski, J. Comb. Chem. 1 (1999) 55–68.
[15] J. Kelder, P.D.J. Grootenhuis, D.M. Bayada, L.P.C. Delbressine, J. Ploemen, Pharm. Res. 16 (1999) 1514–1519.
[16] D.E. Clark, J. Pharm. Sci. 88 (1999) 807–814.
[17] M.D. Wessel, P.C. Jurs, J.M. Tolan, S.M. Muskal, J. Chem. Inf. Comput. Sci. 38 (1998) 726–735.
[18] K. Palm, P. Stenberg, K. Luthman, P. Artusson, Pharm. Res. 14 (1997) 568–571.
[19] M. Kansy, F. Senner, K. Gubernator, J. Med. Chem. 41 (1998) 1007–1010.
[20] T. Oprea, J. Gottfries, J. Mol. Graph. Model. 17 (1999) 261–274.
[21] W.J. Egan, K.M. Merz, J.J. Baldwin, J. Med. Chem. 43 (2000) 3867–3877.
[22] F. Yoshida, J.G. Topliss, J. Med. Chem. 43 (2000) 2575–2585.
[23] D. Ruovray, New Scientist (1993) 35–38.
[24] E. Estrada, SAR QSAR Environ. Res. 11 (2000) 55–73.
[25] E. Estrada, J. Chem. Inf. Comput. Sci. 36 (1996) 844–849.
[26] E. Estrada, J. Chem. Inf. Comput. Sci. 37 (1997) 320–328.
[27] E. Estrada, A. Peña, R. García-Domenech, J. Comput. Aided Mol. Design 12 (1998) 583–595.
[28] E. Estrada, A. Peña, Bioorg. Med. Chem. 8 (2000) 2755–2770.
[29] E. Estrada, E. Uriarte, A. Montero, M. Teijeira, L. Santana, E. De Clercq, J. Med. Chem. 43 (2000) 1975–1985.
[30] E. Estrada, E. Uriarte, SAR QSAR Environ. Res. 12 (2001) 309–324.
[31] E. Estrada, E. Molina, E. Uriarte, SAR QSAR Environ. Res. 12 (2001) 445–459.
[32] M.A. Cabrera, H. González, C. Fernández, J.M. Plá-Delfina, M. Bermejo, Eur. J. Pharm. Biopharm. 53 (2002) 317–325.
[33] M.A. Cabrera, A. Ruiz, C. Fernández, I. González, M. Bermejo, Eur. J. Pharm. Biopharm. 56 (2003) 197–206.
[34] E. Estrada, J. Chem. Inf. Comput. Sci. 35 (1995) 31–33.
[35] H. Van de Waterbeemd, G. Camenisch, G. Folkers, O.A. Raevsky, Quant. Struct.–Act. Relat. 15 (1996) 480–490.
[36] R. Wang, Y. Gao, L. Lai, Perspect. Drug Discov. Des. 19 (2000) 47–66.
[37] P. Ertl, B. Rohde, P. Selzer, J. Med. Chem. 43 (2000) 3714–3717.
[38] A.K. Ghose, G.M. Crippen, J. Chem. Inf. Comput. Sci. 27 (1987) 21–35.
[39] J. Gasteiger, M. Marsilli, Tetrahedron Lett. 34 (1978) 3181–3184.
[40] K.J. Miller, J. Am. Chem. Soc. 112 (1990) 8533–8542.
[41] A. Bondi, J. Phys. Chem. 68 (1964) 441–451.
[42] E. Estrada, E. Uriarte, Y. Gutierrez, H. González, SAR QSAR Environ. Res. 14 (2003) 145–163.
[43] Y. Gutiérrez, E. Estrada, MODESLAB 1.0 (Molecular DEScriptors LABoratory) for Windows, Universidad de Santiago de Compostela, España, 2002 This software can be obtained upon request to Dr. E. Estrada: estrada66@yahoo.com.
[44] D. Weininger, J. Chem. Inf. Comput. Sci. 28 (1998) 31–36.
[45] Y.H. Zhao, J. Le, M.H. Abraham, A. Hersey, P.J. Eddershaw, C.N. Luscombe, D. Boutina, G. Beck, B. Sherborne, I. Cooper, J.A. Platts, J. Pharm. Sci. 90 (2001) 749–784.
[46] L. Barthe, J. Woodley, G. Houin, Fundam. Clin. Pharmacol. 13 (1999) 154–168.
[47] I. Kanfer, J. Pharm. Pharmaceut. Sci. 5 (2002) 1–4.
[48] L.Z. Benet, S. Øie, J.B. Schwartz, in: J.G. Hardman, L.E. Limbird, A.G. Gilman (Eds.), McGraw-Hill, New York, 1996, pp. 1707–1793.
[49] R.J. Bertz, G.R. Granneman, Clin. Pharmacokinet. 32 (1997) 210–258.
[50] StatSoft, STATISTICA version 5.5 for Windows, StatSoft, Inc, 1999.
[51] J. Xu, A. Hagler, Molecules 7 (2002) 566–600.
[52] R.A. Conradi, P.S. Burton, R.T. Borchardt, in: V. Pliška, B. Testa, H. Van de Waterbeemd (Eds.), VCH, Weinheim, 1996, pp. 233–252.
[53] G. Klopman, L.R. Stefan, R.D. Saiakhov, Eur. J. Pharm. Sci. 17 (2002) 253–263.