European Journal of Medicinal Chemistry 44 (2009) 4051–4056

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Original article

Multi-target spectral moment: QSAR for antifungal drugs vs. different fungi species

Francisco J. Prado-Prado [a,b,c,*], Fernanda Borges [c], Lazaro G. Perez-Montoto [a,b], Humberto González-Díaz [a,b,c,**]

[a] Department of Microbiology and Parasitology, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain
[b] Department of Organic Chemistry, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain
[c] Physic-Chemical Molecular Research Units, Department of Organic Chemistry, Faculty of Pharmacy, University of Porto, 4150-047 Porto, Portugal

ARTICLE INFO

Article history:
Received 25 March 2009
Received in revised form
22 April 2009
Accepted 28 April 2009
Available online 5 May 2009

Keywords:
Multi-target QSAR
Markov model
Antifungal drugs
Spectral moments
Linear discriminant analysis

ABSTRACT

The most important limitation of antifungal QSAR models is that they predict the biological activity of drugs against only one fungal species. This is determined due the fact that most of the up-to-date reported molecular descriptors encode only information about the molecular structure. Consequently, predicting the probability with which a drug is active against different fungal species with a single unifying model is a goal of major importance. Herein, we use the Markov Chain theory to calculate new multi-target spectral moments to fit a QSAR model that predicts the antifungal activity of more than 280 drugs against 90 fungi species. Linear discriminant analysis (LDA) was used to classify drugs into two classes as active or non-active against the different tested fungal species whose data we processed. The model correctly classifies 12 434 out of 12 566 non-active compounds (98.95%) and 421 out of 468 active compounds (89.96%). Overall training predictability was 98.63%. Validation of the model was carried out by means of external predicting series, the model classifying, thus, 6216 out of 6277 non-active compounds and 215 out of 239 active compounds. Overall training predictability was 98.7%. The present is the first attempt to calculate, within a unifying framework, the probabilities of antifungal action of drugs against many different species based on spectral moment's analysis.

1. Introduction

Infections caused by fungi have increased dramatically during the past decades. Systemic mycoses mainly appear concomitant with other diseases or are caused by treatment with chemotherapeutics, for instance with cytostatics. In any case, the most dangerous fungal diseases are those related to opportunistic infections of immunocompromised patients because they have an immune system disorder, such as HIV infection. Endocarditis caused by fungi has been observed in patients with drug dependence. So far, many of the new compounds could not be introduced into therapy, mainly for two reasons: (i) poor selectivity and (ii) difficulties in getting into the cytoplasm of fungal cells and/or in gastrointestinal absorption [1–4].

In this sense, a very important role may be played by computer-aided drug design techniques based on multi-target quantitative structure–activity relationships (mt-QSAR) studies. It means that they are models connecting the structure of drugs with the biological activity against different targets (microbial species in the case of antimicrobial drugs) [5,6]. This kind of study may also help in a multi objective optimization (MOOP) of desired properties or activity of drugs against different targets; see for instance the recent works carried out by Cruz-Monteagudo in the topic [7,8]. In principle, up to date there are over 1500 molecular descriptors that may be generalized and used to solve the former problem [9–12]. Many of these indices are known as topological indices (TIs) or simply invariants of a molecular graph, whose vertices are atoms weighed with physicochemical properties (mass, polarity, electro negativity, or charge) [13]. Unfortunately, almost antifungal QSAR studies reported up-to-date are based on molecular descriptors and databases of structurally parent compounds applicable to only one single fungi species. As a result, the researcher interested on predicting, for instance, the antifungal activity for a given series of compounds needs to use/develop as many QSAR equations as combinations of families of compounds vs. fungi species are necessary to be predicted.

* Corresponding author. Physic-Chemical Molecular Research Units, Department of Organic Chemistry, Faculty of Pharmacy, University of Porto, 4150-047 Porto, Portugal.
** Corresponding author.
E-mail addresses: fenol1@hotmail.com (F.J. Prado-Prado), gonzalezdiazh@yahoo.es, humberto.gonzalez@usc.es (H. González-Díaz).

0223-5234/$ – see front matter © 2009 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmech.2009.04.040

Therefore, it is of major interest the development of one single unified equation explaining the antifungal activity of structurally heterogeneous series of compounds against as many fungus species as possible [14,15]. In fact, other mt-QSAR approaches, with demonstrated usefulness, have been introduced recently in medicinal chemistry [16–18]. We introduced a Markov Model encoding molecular backbones information. The method was named the MARCH-INSIDE, MARkovian CHemicals IN SIlico Design [19]. It allowed us introducing matrix invariants such as stochastic entropies, potentials, and spectral moments for the study of molecular properties [20,21]. Specifically, the stochastic spectral moments introduced by our group have been largely used for small molecule mt-QSAR problems including the design of fluckicidal, anticancer and antihypertensive drugs [22]. Applications to macromolecules have been restricted to the field of RNA without applications to proteins [23–26]. The QSAR models based on different MARCH-INSIDE indices may be very useful to optimize important aspects such as activity, toxicity or pharmacokinetics using one single model in many bioorganic and medicinal chemistry problems such as: estimation of anticoccidial activity, modelling the interaction between drugs and HIV-packaging-region RNA, and predicting proteins and virus activity [27–29]. In recent studies, the MARCH-INSIDE method has been extended to encompass molecular environment, interesting information in addition to molecular structure [30,31]. In three recent reviews, we discussed the multiple applications of MARCH-INISDE to classic QSAR, macromolecular QSAR, and specially mt-QSAR [32–34]. However, we have never used before stochastic spectral moments to develop a mt-QSAR for antifungal drugs. In this work, we develop, for the first time, a single linear equation based on these previous ideas to predict the antifungal activity of drugs against different species.

2. Methods

2.1. Markov thermodynamics for drug-target step-by-step interaction

Let us consider a hypothetical situation in which a drug molecule is free in the space at an arbitrary initial time (t_0). It is then interesting to develop a simple stochastic model for a step-by-step interaction between the atoms of a drug molecule and a molecular receptor at the time of triggering the pharmacological effect. For the sake of simplicity, from now on, we are going to consider a general structure-less molecular receptor or drug-target, understanding by structure-less receptor a receptor whose chemical structure is not taken into consideration. In our model, we approach this problem considering the free energy ${}^{k}g_{ij}(s)$ of interaction between an atom in the drug and the drug receptor after k-steps or previous interactions. We state that ${}^{k}g_{ij}(s)$ is also a state function and the symbol g points precisely to Gibbs energy. s indicates that this energy depends on the specific drug target in different microbial species. Afterwards, the interaction has to define the free energy of interaction ${}^{k}g_{ij}(s)$ between the j-th atom and the receptor for a specific microbial species (s) given that i-th atom has been interacted at a previous time t_k. So, one can suppose that, atoms begin binding to this receptor in discrete intervals of time t_k. However, there are several alternative ways in which such step-by-step binding process may occur. In this picture, the free energy ${}^{1}g_{ij}(s)$ can be defined by analogy as dependent on a constant for the atom-target interaction $\Gamma_{ij}(s)$ [14, 22]:

$$ {}^{1}g_{ij}(s) = -RT\log {}^{1}\Gamma_{ij}(s). \tag{1} $$

The present approach to antimicrobial–receptor interaction has two main drawbacks. The first is the difficulty of defining the constants. In this work, we solve the first question by estimating the use of the ratio of occurrence $n_j(s)$ of the j-th atom on active molecules against a given species (frequency of effective interactions) with respect to the number of atoms of the j-th class in the molecules tested against the same species $n_T(s)$. Consequently, one of the most important steps is the change on the value of the atomic weights used ${}^{k}g_{ij}(s)$ for different pathogen species. Regarding ${}^{1}\Gamma_{ij}(s)$, we must take into account that once the j-th atoms have interacted, the preferred candidates for the next interaction are those i-th atoms bound to j by a chemical bond [22]:

$$ {}^{1}\Gamma_{ij}(s) = \left(\alpha_{ij}\frac{n_j(s)}{n_T(s)} + 1\right) = e^{{}^{1}g_{ij}(s)/RT} \tag{2} $$

where α_{ij} are the elements of the atom adjacency matrix, $n_j(s)$, $n_T(s)$, and ${}^{1}g_{ij}(s)$ have been defined in the paragraph above, R is the universal gases constant, and T the absolute temperature. The number 1 is added to avoid forbidden negative values as inputs for the logarithmic function. The second problem relates to the description of the interaction process at higher times $t_k > t_1$. Therefore, Markov Chain theory enables a simple calculation of the probabilities with which the drug–receptor interaction takes place in the time until the antimicrobial effect is achieved. As depicted in Fig. 1, this model deals with the calculation of the probabilities (${}^{k}p_{ij}$) with which any arbitrary molecular atom j-th binds to the structure-less molecular receptor given that other atom i-th has been bound before; along discrete time periods t_k ($k = 1, 2, 3, \ldots$); ($k = 1$ in grey), ($k = 2$ in blue) and ($k = 3$ in red) throughout the chemical bonding system. The procedure described here considers the atoms of the molecule as states of the Markov Chain. The method arranges all the ${}^{1}g_{ij}(s)$ free energies of interaction as a squared table of $n \times n$ dimension. After normalization of the matrix we can built up the corresponding stochastic matrix $\mathbf{{}^{1}\Pi(s)}$, which has the elements ${}^{1}\pi_{ij}(s)$. The matrix is called the 1-step drug–target interaction stochastic matrix. $\mathbf{{}^{1}\Pi(s)}$ is too built as a squared table of order n, where n represents the number of atoms in the molecule. The elements ${}^{1}\pi_{ij}(s)$ of the 1-step drug–target interaction stochastic matrix are the binding probabilities with which a j-th atom binds to a structure-less molecular receptor given that other i-th atoms have been interacted before at a time $t_1 = 1$[22]:

$$ {}^{1}\pi_{ij}(s) = \frac{{}^{1}g_{ij}(s)}{\sum_{a=1}^{n} {}^{1}g_{ia}(s)} = \frac{\alpha_{ij}(-RT)\log\left(\frac{n_j(s)}{n(s)} + 1\right)}{\sum_{a=1}^{n} \alpha_{ia}(-RT)\log\left(\frac{n_a(s)}{n_T(s)} + 1\right)} = \frac{\alpha_{ij}\log\left(\frac{n_j(s)}{n_T} + 1\right)}{\sum_{a=1}^{n} \alpha_{ia}\log\left(\frac{n_j(s)}{n_T(s)} + 1\right)}. \tag{3} $$

Such a model is stochastic *per se* (probabilistic step-by-step atom–receptor interaction in time) but also considers molecular connectivity (the step-by-step atom union in space throughout the chemical bonding system). One can calculate the atomic spectral moments ${}^{k}\mu_s(j) = {}^{k}\pi_{jj}(s)$, values on the main diagonal ($i = j$) of the matrix, in order to numerically characterize the propensity with which a specific atom interacts several times with a drug receptor. In addition, the ${}^{k}\mu_s(j)$ can be summed for specific atom sets (AS) to create local molecular descriptors ${}^{k}\mu_s(AS)$ for the drug–target interaction. Herein the AS used were as follows: halogens (X), unsaturated carbons (C_{ins}), saturated carbons (C_{sat}), heteroatom (Het), and hydrogen bound to heteroatom (H-Het). The corresponding symbols of the local spectral moments for these AS are: $\mu_k(X,s)$, $\mu_k(C_{ins},s)$, $\mu_k(C_{sat},s)$, $\mu_k(Het,s)$, $\mu_k(H\text{-}Het,s)$. Finally, the sum of all atoms (it means that AS = Total contains all atoms) is useful as a total molecular descriptor.

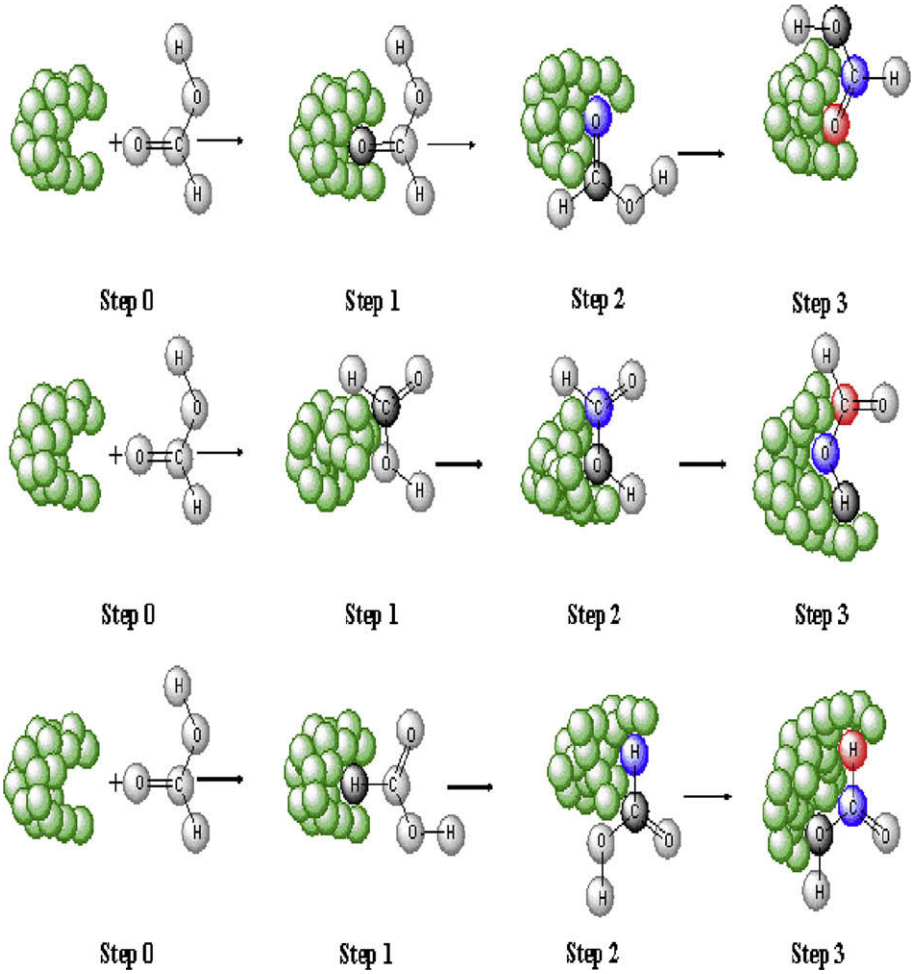

Fig. 1.

$$ {}^{k}\mu_s(\mathrm{AS}) = Tr\left(\left[{}^{1}\mathbf{\Pi}(\mathbf{s})\right]^{k}\right)_{j\in \mathrm{AS}} = Tr\left(\begin{bmatrix} {}^{1}\pi_{11}(\mathrm{s}) & {}^{1}\pi_{12}(\mathrm{s}) & . & . & {}^{1}\pi_{1n}(\mathrm{s}) \\ {}^{1}\pi_{21}(\mathrm{s}) & {}^{1}\pi_{22}(\mathrm{s}) & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ {}^{1}\pi_{n1}(\mathrm{s}) & . & . & . & {}^{1}\pi_{nn}(\mathrm{s}) \end{bmatrix}^{k}\right)_{j\in \mathrm{AS}} = \sum_{i=j\in \mathrm{AS}}^{n} \pi_{ij}(\mathrm{s}) \quad (4) $$

2.2. Statistical analysis

As a continuation of the previous sections, we can attempt to develop a simple linear QSAR with the general formula:

$$ \mathrm{Actv} = a_0{}^{0}\mu_s(\mathrm{AS}) + a_1{}^{1}\mu_s(\mathrm{AS}) + a_2{}^{2}\mu_s(\mathrm{AS}) + a_3{}^{3}\mu_s(\mathrm{AS})_s\cdots + a_k{}^{k}\mu_s(\mathrm{AS}) + b_0. \quad (5) $$

Here, ${}^{k}\mu_s(\mathrm{AS})$ are the spectral moments described above that act as molecule–target interaction descriptors specific for each drug–microbial specie pair. We selected Linear Discriminant Analysis (LDA) to fit the classification functions. The model deals with the classification of a set of compounds as active or not-active against different microbial species [35]. A dummy variable (Actv) was used to codify the antimicrobial activity. This variable indicates either the presence (Actv = 1) or absence (Actv = −1) of antimicrobial activity of the drug against specific species. In Eq. (5), a_k represents the coefficients of the classification function, determined by the least square method as implemented in the LDA module of the STATISTICA 6.0 software package [36]. Forward stepwise strategy was set as the one used for variable selection [35]. The quality of LDA models was determined by examining Wilk's U statistic, Fisher ratio (F), and the p-level (p). We also inspected the accuracy, sensitivity, and specificity of the method. The validation of the model was corroborated by external validation series [35].

2.3. Data set

The data set was formed by a set of marketed and/or very recently reported antifungal drugs which low reported low $\mathrm{MIC}_{50} < 10\ \mu\mathrm{M}$ against different fungus. The data set was

conformed to more than 280 different drugs experimentally tested against some species of a list of 90. Not all drugs were tested in the literature against all listed species so we were able to collect 19 550 cases (drug/species pairs) instead of 280 × 90 cases. The names or codes and activity for all compounds as well as the references used to collect them are depicted in Table 1SM of the supplementary material file.

3. Results and discussion

One of the main advantages of the present stochastic approach is the possibility of deriving average thermodynamic parameters depending on the probability of the states of the MM. The generalized parameters fit on a more clearly physicochemical sense with respect to our previous ones [20,37,38]. More specifically, this work introduces for the first time a single linear QSAR equation model to predict the antifungal activity of drugs against different fungal species. The best model found was:

$$\begin{aligned} \text{actv} = & -3.44 \cdot {}^{5}\mu_s(\text{Het}) - 3.18 \cdot {}^{2}\mu_s(\text{H-Het}) \\ & - 3.85 \cdot {}^{3}\mu_s(C_{sat}) 4.76 \cdot {}^{4}\mu_s(C_{sat}) - 4.61 \cdot {}^{5}\mu_s(C_{sat}) \\ & + 28.26 \cdot {}^{0}\mu_s(T) - 29.26 \\ \lambda = 0.33;\ \chi^2 = 14367.94;\ p < 0.001. & \end{aligned} \quad (6)$$

In the model, the coefficient λ is the Wilk's statistic; statistic for the overall discrimination, χ^2 is the Chi-square, and p the error level. In this equation, ${}^{k}\mu_s$ were calculated for the total (T) of atoms in the molecule or for specific collections of atoms. These collections are atoms with a common characteristic for instance: heteroatom (Het) and hydrogen bound to heteroatom (H-Het) and saturated Carbon atoms (C_{sat}). The model correctly classifies 12 434 out of 12 566 non-active compounds (98.95%) and 421 out of 468 active compounds (89.96%). Overall training predictability was 98.63%. The validation of the model was carried out by means of external predicting series, the model classifying, thus, 6216 out of 6277 non-active compounds and 215 out of 239 active compounds, see Table 1. These results were received very well by the authors who developed LDA and also non-linear QSAR classification models [39–72].

The most interesting characteristic of the present model is that the ${}^{k}\mu_s$ used as molecular descriptors depend both on the molecular structure of the drug and the fungus species against which the drug must act. The codification of the molecular structure is basically due to the use of the adjacent factor α_{ij} to encode atom–atom bonding, molecular connectivity. The other aspect that allows encoding molecular structural changes is that the spectral moment ${}^{k}\mu_s$ are atom-class specific. This property is related to the definition of ${}^{k}\mu_s$. The values of these species, specific atomic standard free energies reported herein for the first time, are given in Table 2 for some atoms and more than 90 species. For example, one change in

Table 1
Results of the model, analysis, validation.

Parameter	%	Classes	Non-active	Antifungal
Analysis				
Sensitivity	98.95	Non-active	**12 434**	132
Specificity	89.96	Antifungal	47	**421**
Accuracy	98.63			
Validation				
Sensitivity	99.03	Non-active	**6216**	61
Specificty	89.96	Antifungal	24	**215**
Accuracy	98.70			

The positive cases are in bold.

Table 2
Standard atomic free energy values for atom–receptor interactions.

Fungi	C	H	N	O	S	Cl	F
Absidia coryambifera	0.26	0.27	0.21	0.28	0.00	0.30	0.11
Achaetomium strumarium	0.55	0.55	0.57	0.44	0.00	0.85	0.54
Acremonium spp.	0.94	0.95	0.96	0.94	0.30	1.10	0.87
Apophysomyces elegans	0.22	0.23	0.18	0.26	0.00	0.30	0.15
Aspergillus candidus	0.66	0.65	0.69	0.63	0.00	0.60	0.73
Aspergillus flavipes	0.72	0.69	0.77	0.65	0.00	0.60	0.82
Aspergillus flavus	0.21	0.19	0.19	0.20	0.00	0.29	0.23
Aspergillus fumigatus	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Aspergillus glaucus	0.60	0.60	0.60	0.60	0.00	0.60	0.60
Aspergillus nidulans	0.02	0.01	0.02	0.00	0.18	0.00	0.05
Aspergillus niger	0.03	0.03	0.03	0.03	0.00	0.03	0.04
Aspergillus ochraceus	0.42	0.40	0.44	0.48	0.00	0.18	0.52
Aspergillus spp.	0.11	0.11	0.10	0.12	0.00	0.12	0.09
Aspergillus sydowii	1.75	1.73	1.81	1.64	0.00	1.94	1.74
Aspergillus terreus	0.64	0.61	0.66	0.58	0.00	0.75	0.78
Aspergillus ustus	0.24	0.21	0.33	0.13	0.00	0.30	0.22
Aspergillus versicolor	0.93	0.95	0.99	0.95	0.95	0.88	0.89
Bipolaris spp.	0.99	1.03	0.94	1.05	0.90	1.00	0.93
Bjerkandera adusta	0.60	0.59	0.57	0.59	0.00	0.60	0.69
Blastomyces dermatitidis	1.34	1.34	1.38	1.33	0.00	1.23	1.40
Candida albicans	0.03	0.03	0.02	0.05	0.01	0.04	0.01
Candida dubliniensis	0.01	0.01	0.01	0.01	0.00	0.01	0.00
Candida famata	0.33	0.31	0.27	0.33	0.00	0.30	0.44
Candida glabrata	0.25	0.25	0.20	0.27	0.25	0.26	0.19
Candida guilliermondii	0.32	0.31	0.33	0.34	0.27	0.37	0.30
Candida kefyr	0.08	0.07	0.06	0.01	0.40	0.12	0.03
Candida krusei	0.02	0.02	0.02	0.02	0.00	0.01	0.03
Candida lusitaniae	0.03	0.02	0.03	0.02	0.00	0.02	0.07
Candida neoformans	0.29	0.27	0.28	0.26	0.00	0.18	0.65
Candida parapsilosis	0.01	0.00	0.01	0.00	0.05	0.00	0.02
Candida tropicalis	0.02	0.02	0.02	0.02	0.00	0.03	0.02
Candida utilis	0.22	0.22	0.24	0.41	0.00	0.30	0.30
Chaetomium atrobrunneum	0.26	0.27	0.27	0.20	0.00	0.48	0.23
Chaetomium globosum	0.22	0.24	0.19	0.19	0.00	0.30	0.09
Chaetomium nigrocolor	0.27	0.26	0.29	0.20	0.00	0.60	0.30
Chrysosporium spp.	1.23	1.24	1.37	1.12	0.90	1.23	1.33
Coccidioides immitis	0.24	0.23	0.19	0.22	0.00	0.30	0.15
Cokeromyces recurvatus	0.49	0.47	0.42	0.36	0.00	0.81	0.36
Coprinus species	0.12	0.15	0.15	0.25	0.00	0.00	0.00
Cryptococcus laurentii	0.06	0.05	0.03	0.08	0.00	0.12	0.09
Cryptococcus neoformans	0.02	0.02	0.01	0.02	0.05	0.03	0.01
Cunninghamella spp.	1.32	1.37	1.11	1.46	0.00	1.57	0.15
Epidermophyton floccosum	0.21	0.20	0.25	0.13	0.30	0.21	0.26
Fusarium oxysporum	0.52	0.42	0.44	0.17	0.60	0.90	0.00
Fusarium solani	0.13	0.10	0.18	0.06	0.00	0.22	0.40
Fusarium spp.	0.23	0.20	0.31	0.16	0.00	0.54	0.30
Madurella mycetomatis	0.05	0.03	0.10	0.02	0.00	0.08	0.90
Malassezia furfur	0.26	0.25	0.22	0.33	0.00	0.39	0.23
Malassezia pachydermatis	0.50	0.51	0.45	0.60	0.30	0.41	0.61
Malassezia slooffiae	0.29	0.29	0.28	0.28	0.00	0.30	0.27
Malassezia sympodialis	0.22	0.23	0.14	0.27	0.00	0.22	0.11
Microsporum audouinii	0.25	0.26	0.22	0.25	0.00	0.30	0.19
Microsporum canis	0.16	0.16	0.12	0.20	0.00	0.28	0.06
Microsporum cookei	0.35	0.36	0.32	0.29	0.00	0.32	0.30
Microsporum ferrugineur	0.30	0.30	0.29	0.28	0.00	0.30	0.30
Microsporum fulvum	0.16	0.18	0.15	0.22	0.00	0.08	0.15
Microsporum gallinae	0.28	0.29	0.26	0.27	0.00	0.30	0.23
Microsporum gypseum	0.18	0.19	0.22	0.26	0.00	0.16	0.25
Microsporum nanum	0.23	0.25	0.25	0.26	0.00	0.08	0.30
Microsporum praecox	0.37	0.36	0.41	0.29	0.00	0.15	0.63
Microsporum racemosum	0.12	0.14	0.07	0.19	0.00	0.00	0.11
Mucor spp.	0.73	0.74	0.61	0.77	0.00	0.93	0.54
Paecilomyces lilacinus	0.07	0.08	0.07	0.09	0.00	0.08	0.04
Paecilomyces spp.	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Paecilomyces variotii	0.06	0.06	0.06	0.06	0.00	0.03	0.08
Rhizopus oryzae	0.72	0.73	0.58	0.73	0.00	0.85	0.78
Rhizopus spp.	0.96	0.99	0.90	1.02	0.48	1.10	0.78
Rhodotorula glutinis	0.16	0.15	0.16	0.15	0.00	0.22	0.19
Saccharomyces cerevisiae	0.13	0.12	0.17	0.12	0.00	0.30	0.25
Saksenaea vasiformis	0.35	0.29	0.42	0.21	0.00	0.40	0.57
Scedosporium apiospermum	0.10	0.10	0.15	0.10	0.00	0.06	0.15
Scedosporium prolificans	0.09	0.08	0.10	0.07	0.00	0.06	0.12
Schizophyllum commune	0.55	0.51	0.47	0.39	0.00	0.90	0.38
Sporobolomyces salmonicolor	0.29	0.28	0.27	0.21	0.00	0.64	0.25

Table 2 (*continued*)

Fungi	C	H	N	O	S	Cl	F
Trichophyton ajelloi	0.20	0.20	0.19	0.25	0.00	0.23	0.21
Trichophyton balcaneum	0.28	0.29	0.26	0.27	0.00	0.30	0.23
Trichophyton concentricum	0.28	0.29	0.26	0.27	0.00	0.30	0.23
Trichophyton erinacei	1.64	1.73	1.57	1.83	0.00	1.13	1.65
Trichophyton interdigitale	0.47	0.47	0.51	0.37	0.00	0.32	0.73
Trichophyton mentagrophytes	0.16	0.16	0.12	0.20	0.00	0.28	0.06
Trichophyton phaseoliforme	0.47	0.47	0.51	0.37	0.00	0.32	0.73
Trichophyton rubrum	0.16	0.17	0.13	0.22	0.00	0.28	0.06
Trichophyton schoenleinii	0.28	0.29	0.26	0.27	0.00	0.30	0.23
Trichophyton simii	0.33	0.33	0.34	0.31	0.00	0.30	0.36
Trichophyton tonsurans	0.30	0.30	0.30	0.30	0.00	0.30	0.30
Trichophyton verrucosum	0.34	0.35	0.32	0.29	0.00	0.32	0.30
Trichophyton violaceum	0.21	0.23	0.21	0.28	0.12	0.10	0.28
Trichosporon asahii	0.08	0.07	0.12	0.05	0.00	0.22	0.15
Trichosporon spp.	0.21	0.23	0.18	0.26	0.00	0.30	0.15
Unidentified basidiomycetes	0.28	0.28	0.23	0.29	0.00	0.30	0.21
Wangiella dermatitidis	0.25	0.23	0.27	0.30	0.30	0.00	0.00

the molecular structure, e.g. F by O, necessarily implies a change in the moments of interaction. Moreover, the most interesting fact is that ${}^{k}\mu_{s}$ are the molecular descriptors reported for antimicrobial mt-QSAR studies able to distinguish among a large number of fungal species.

As an end result of the above mentioned flexible definition of the present approach, it was possible to model for the first time some very heterogeneous and diverse data with more than 19 000 cases. The posterior validation probabilities predicted for every drug–species pair are depicted in Table 1SM of the supplementary material file. The present work is the first reported unifying model, using moments ${}^{k}\mu_{s}$ as a molecular descriptor that allows a predicting antifungal activity of any organic compound against a very large diversity of fungal pathogens.

4. Conclusion

The present mt-QSAR methodology with a very large data set improves significantly the previous QSAR models and may help to perform MOOP of drug activity against different species. It determines that the mt-QSAR methodology may be able to predict the biological activity of drugs in more general situations than the traditional QSAR models, whose greatest limitation is predicting the biological activity of drugs against only one fungi species. This mt-QSAR methodology improves models using spectral moments as a molecular descriptor that allow predicting antifungal activity of any organic compound against a very large diversity of fungal pathogens.

Acknowledgments

The authors sincerely thank the kind attention as well as useful comments from the editor Prof. A. Monge-Vega and two unknown referees. Prado-Prado, F. acknowledges the financial support from the Xunta de Galicia for a one-year post-doctoral position (Research Project IN89A 2008/75-0). González-Díaz, H. acknowledges the financial support of the Programme Isidro Parga Pondal and one-year post-doctoral position (Research Project IN89A 2008/117-0), both funded by the Xunta de Galicia and the European Research funds from European Social Fund (ESF).

Appendix. Supplementary material

Supplementary material associated with this article can be found in the online version, at doi:10.1016/j.ejmech.2009.04.040.

References

[1] A.M. Fedder, B. Morn, J.K. Moller, Ugeskr. Laeger 168 (2006) 363–366.
[2] A. Chopra, G.K. Khuller, Crit. Rev. Microbiol. 11 (1984) 209–271.
[3] S.M. Grant, S.P. Clissold, Drugs 39 (1990) 877–916.
[4] C.V. Vartian, D.M. Shlaes, A.A. Padhye, L. Ajello, Am. J. Med. 78 (1985) 703–707.
[5] F.J. Prado-Prado, H. Gonzalez-Diaz, O.M. de la Vega, F.M. Ubeira, K.C. Chou, Bioorg. Med. Chem. 16 (2008) 5871–5880.
[6] F.J. Prado-Prado, O.M. de la Vega, E. Uriarte, F.M. Ubeira, K.C. Chou, H. Gonzalez-Diaz, Bioorg. Med. Chem. 17 (2009) 569–575.
[7] M. Cruz-Monteagudo, F. Borges, M.N. Cordeiro, J.L. Cagide Fajin, C. Morell, R.M. Ruiz, et al., J. Comb. Chem. 10 (2008) 897–913.
[8] M. Cruz-Monteagudo, F. Borges, M.N. Cordeiro, J. Comput. Chem. 29 (2008) 2445–2459.
[9] R. Todeschini, V. Consonni, Handbook of Molecular Descriptors, Wiley-VCH, Bicocca, Italy, 2002.
[10] H. Kubinyi, J. Cancer. Res. Clin. Oncol. 116 (1990) 529–537.
[11] M.P. González, A.H. Morales, H. González-Díaz, Polymer 45 (2004) 2073–2079.
[12] Y. Marrero-Ponce, R.M. Marrero, F. Torrens, Y. Martinez, M.G. Bernal, V.R. Zaldivar, et al., J. Mol. Model. (Online) (2005) 1–17.
[13] E. Estrada, E. Molina, J. Chem. Inf. Comput. Sci. 41 (2001) 791–797.
[14] H. González-Díaz, F.J. Prado-Prado, L. Santana, E. Uriarte, Bioorg. Med. Chem. 14 (2006) 5973–5980.
[15] H. González-Díaz, F. Prado-Prado, J. Comput. Chem. 29 (2008) 656–657.
[16] Y. Marrero-Ponce, H.G. Diaz, V.R. Zaldivar, F. Torrens, E.A. Castro, Bioorg. Med. Chem. 12 (2004) 5331–5342.
[17] E. Molina, H. González-Díaz, M.P. Gonzalez, E. Rodriguez, E. Uriarte, J. Chem. Inf. Comput. Sci. 44 (2004) 515–521.
[18] H. González-Díaz, A. Perez-Bello, E. Uriarte, Y. Gonzalez-Diaz, Bioorg. Med. Chem. Lett. 16 (2006) 547–553.
[19] H. González-Díaz, R. Molina-Ruiz, I. Hernandez, MARCH-INSIDE v3.0 (MARkov CHains INvariants for SImulation & DEsign), Cuba; Windows supported version under request to the main author contact email: <gonzalezdiazh@yahoo.es>. version 3.0 2007.
[20] H. González-Díaz, G. Aguero, M.A. Cabrera, R. Molina, L. Santana, E. Uriarte, et al., Bioorg. Med. Chem. Lett. 15 (2005) 551–557.
[21] R. Ramos de Armas, H. González-Díaz, R. Molina, M. Perez Gonzalez, E. Uriarte, Bioorg. Med. Chem. 12 (2004) 4815–4822.
[22] F. Prado-Prado, H. González-Díaz, L. Santana, E. Uriarte, Bioorg. Med. Chem. 15 (2007) 897–902.
[23] H. González-Díaz, E. Uriarte, Bioorg. Med. Chem. Lett. 15 (2005) 5088–5094.
[24] H. González-Díaz, O. Gia, E. Uriarte, I. Hernadez, R. Ramos, M. Chaviano, et al., J. Mol. Model. 9 (2003) 395–407.
[25] H. González-Díaz, L. Saiz-Urra, R. Molina, L. Santana, E. Uriarte, J. Proteome Res. 6 (2007) 904–908.
[26] L. Saiz-Urra, H. González-Díaz, E. Uriarte, Bioorg. Med. Chem. 13 (2005) 3641–3647.
[27] H. González-Díaz, R. Molina, E. Uriarte, Bioorg. Med. Chem. Lett. 14 (2004) 4691–4695.
[28] H. González-Díaz, G. Ferino, G. Podda, E. Uriarte, ECSOC 12 (G1) (2008) 1–10.
[29] H. González-Díaz, A. Sanchez-Gonzalez, Y. Gonzalez-Diaz, J. Inorg. Biochem. 100 (2006) 1290–1297.
[30] H. González-Díaz, M.A. Cabrera-Pérez, G. Agüero-Chapín, M. Cruz-Monteagudo, N. Castañedo-Cancio, M.A. del Río, et al., Chemometr. Intell. Lab. Syst. 94 (2008) 160–165.
[31] M. Cruz-Monteagudo, H. González-Díaz, G. Agüero-Chapin, L. Santana, F. Borges, R.E. Domínguez, et al., J. Comput. Chem. 28 (2007) 1909–1922.
[32] H. González-Díaz, Y. González-Díaz, L. Santana, F.M. Ubeira, E. Uriarte, Proteomics 8 (2008) 750–778.
[33] H. Gonzalez-Diaz, F. Prado-Prado, F.M. Ubeira, Curr. Top. Med. Chem. 8 (2008) 1676–1690.
[34] H. González-Díaz, S. Vilar, L. Santana, E. Uriarte, Curr. Top. Med. Chem. 7 (2007) 1025–1039.
[35] H. Van Waterbeemd, Discriminant analysis for activity prediction, in: H. Van Waterbeemd (Ed.), Chemometric Methods in Molecular Design, Wiley-VCH, New York, 1995, pp. 265–282.
[36] STATISTICA (Data Analysis Software System), Version 6.0, StatSoft. Inc., Tulsa, USA, 2002.<www.statsoft.com.Statsoft>.
[37] H. González-Díaz, M. Cruz-Monteagudo, R. Molina, E. Tenorio, E. Uriarte, Bioorg. Med. Chem. 13 (2005) 1119–1129.
[38] M. Cruz-Monteagudo, H. González-Díaz, Eur. J. Med. Chem. 40 (2005) 1030–1041.
[39] E. Estrada, S. Vilar, E. Uriarte, Y. Gutierrez, J. Chem. Inf. Comput. Sci. 42 (2002) 1194–1203.
[40] S. Vilar, G. Cozza, S. Moro, Curr. Top. Med. Chem. 8 (2008) 1555–1572.
[41] S. Vilar, E. Estrada, E. Uriarte, L. Santana, Y. Gutierrez, J. Chem. Inform. Model. 45 (2005) 502–514.
[42] S. Vilar, H. Gonzalez-Diaz, L. Santana, E. Uriarte, J. Comput. Chem. 29 (2008) 2613–2622.
[43] S. Vilar, E. Quezada, L. Santana, E. Uriarte, M. Yanez, N. Fraiz, et al., Bioorg. Med. Chem. Lett. 16 (2006) 257–261.
[44] S. Vilar, L. Santana, E. Uriarte, J. Med. Chem. 49 (2006) 1118–1124.
[45] J.A. Castillo-Garit, Y. Marrero-Ponce, J. Escobar, F. Torrens, R. Rotondo, Chemosphere 73 (2008) 415–427.

[46] J.A. Castillo-Garit, Y. Marrero-Ponce, F. Torrens, Bioorg. Med. Chem. 14 (2006) 2398–2408.
[47] J.A. Castillo-Garit, Y. Marrero-Ponce, F. Torrens, R. Garcia-Domenech, J. Pharm. Sci. 97 (2008) 1946–1976.
[48] J.A. Castillo-Garit, Y. Marrero-Ponce, F. Torrens, R. Garcia-Domenech, V. Romero-Zaldivar, J. Comput. Chem. 29 (2008) 2500–2512.
[49] J.A. Castillo-Garit, Y. Marrero-Ponce, F. Torrens, R. Rotondo, J. Mol. Graph. Model. 26 (2007) 32–47.
[50] Y. Marrero Ponce, J.A. Castillo Garit, D. Nodarse, Bioorg. Med. Chem. 13 (2005) 3397–3404.
[51] Y. Marrero-Ponce, J.A. Castillo-Garit, J. Comput. Aided Mol. Des. 19 (2005) 369–383.
[52] Y. Marrero-Ponce, J.A. Castillo-Garit, E. Olazabal, H.S. Serrano, A. Morales, N. Castanedo, et al., J. Comput. Aided Mol. Des. 18 (2004) 615–634.
[53] Y. Marrero-Ponce, J.A. Castillo-Garit, E. Olazabal, H.S. Serrano, A. Morales, N. Castanedo, et al., Bioorg. Med. Chem. 13 (2005) 1005–1020.
[54] Y. Marrero-Ponce, R. Medina-Marrero, J.A. Castillo-Garit, V. Romero-Zaldivar, F. Torrens, E.A. Castro, Bioorg. Med. Chem. 13 (2005) 3003–3015.
[55] Y. Marrero-Ponce, A. Meneses-Marcel, J.A. Castillo-Garit, Y. Machado-Tugores, J.A. Escario, A.G. Barrio, et al., Bioorg. Med. Chem. 14 (2006) 6502–6524.
[56] K.C. Chou, H.B. Shen, Biochem. Biophys. Res. Commun. (2006).
[57] H. Shen, K.C. Chou, Biochem. Biophys. Res. Commun. 334 (2005) 288–292.
[58] M. Wang, J. Yang, Z.J. Xu, K.C. Chou, J. Theor. Biol. 232 (2005) 7–15.
[59] K.C. Chou, H.B. Shen, J. Cell. Biochem. (2006).
[60] S.Q. Wang, J. Yang, K.C. Chou, J. Theor. Biol. (2006).
[61] H.B. Shen, K.C. Chou, Biochem. Biophys. Res. Commun. 337 (2005) 752–756.
[62] K.Y. Feng, Y.D. Cai, K.C. Chou, Biochem. Biophys. Res. Commun. 334 (2005) 213–217.
[63] Y.D. Cai, K.Y. Feng, W.C. Lu, K.C. Chou, J. Theor. Biol. (2005).
[64] Y.D. Cai, K.Y. Feng, W.C. Lu, K.C. Chou, J. Theor. Biol. 238 (2006) 172–176.
[65] H.B. Shen, K.C. Chou, Bioinformatics 22 (2006) 1717–1722.
[66] H.B. Shen, J. Yang, K.C. Chou, Amino Acids 33 (2007) 57–67.
[67] J. Chen, H. Liu, J. Yang, K.C. Chou, Amino Acids 33 (2007) 423–428.
[68] H.B. Shen, J. Yang, K.C. Chou, Expert Rev. Proteomics 4 (2007) 453–463.
[69] H.B. Shen, K.C. Chou, Biochem. Biophys. Res. Commun. 363 (2007) 297–303.
[70] T.L. Zhang, Y.S. Ding, K.C. Chou, J. Theor. Biol. 250 (2008) 186–193.
[71] Y.S. Ding, T.L. Zhang, K.C. Chou, Protein Pept. Lett. 14 (2007) 811–815.
[72] H.B. Shen, K.C. Chou, Protein Eng. Des. Sel. 20 (2007) 561–567.