SAR and QSAR in Environmental Research

ISSN: 1062-936X (Print) 1029-046X (Online) Journal homepage: http://www.tandfonline.com/loi/gsar20

QuBiLs-MAS method in early drug discovery and rational drug identification of antifungal agents

R. Medina Marrero, Y. Marrero-Ponce, S.J. Barigye, Y. Echeverría Díaz, R. Acevedo-Barrios, G.M. Casañola-Martín, M. García Bernal, F. Torrens & F. Pérez-Giménez

To cite this article: R. Medina Marrero, Y. Marrero-Ponce, S.J. Barigye, Y. Echeverría Díaz, R. Acevedo-Barrios, G.M. Casañola-Martín, M. García Bernal, F. Torrens & F. Pérez-Giménez (2015) QuBiLs-MAS method in early drug discovery and rational drug identification of antifungal agents, SAR and QSAR in Environmental Research, 26:11, 943-958, DOI: 10.1080/1062936X.2015.1104517

To link to this article: http://dx.doi.org/10.1080/1062936X.2015.1104517

SAR and QSAR in Environmental Research, 2015
Vol. 26, No. 11, 943–958, http://dx.doi.org/10.1080/1062936X.2015.1104517

QuBiLs-MAS method in early drug discovery and rational drug identification of antifungal agents

R. Medina Marrero[a,b], Y. Marrero-Ponce[a,c,d,h], S.J. Barigye[a,e]*, Y. Echeverría Díaz[a], R. Acevedo-Barrios[c], G.M. Casañola-Martín[a,d,f], M. García Bernal[b], F. Torrens[g] and F. Pérez-Giménez[d]

[a]*Computer-Aided Molecular 'Biosilico' Discovery and Bioinformatic Research International Network (CAMD-BIR-IN), Cartagena de Indias, Bolivar, Colombia;* [b]*Department of Microbiology, Chemical Bioactive Center, Central University of Las Villas, Villa Clara, Cuba;* [c]*Grupo de Investigación en Estudios Químicos y Biológicos, Facultad de Ciencias Básicas, Universidad Tecnológica de Bolívar, Cartagena de Indias, Bolívar, Colombia;* [d]*Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular, Departamento de Química Física, Facultad de Farmacia, Universitat de València, Valencia, Spain;* [e]*Departamento de Química, Universidade Federal de Lavras, Lavras, MG, Brazil;* [f]*Facultad de Ingeniería Ambiental, Universidad Estatal Amazónica, Puyo, Ecuador;* [g]*Institut Universitari de Ciència Molecular, Universitat de València, Valencia, Spain;* [h]*Grupo de Investigación Microbiología y Ambiente (GIMA). Programa de Bacteriología, Facultad Ciencias de la Salud, Universidad de San Buenaventura, Calle Real de Ternera, 130010, Cartagena (Bolivar), Colombia*

(*Received 20 May 2015; in final form 12 August 2015*)

The QuBiLs-MAS approach is used for the *in silico* modelling of the antifungal activity of organic molecules. To this effect, non-stochastic (NS) and simple-stochastic (SS) atom-based quadratic indices are used to codify chemical information for a comprehensive dataset of 2478 compounds having a great structural variability, with 1087 of them being antifungal agents, covering the broadest antifungal mechanisms of action known so far. The NS and SS index-based antifungal activity classification models obtained using linear discriminant analysis (LDA) yield correct classification percentages of 90.73% and 92.47%, respectively, for the training set. Additionally, these models are able to correctly classify 92.16% and 87.56% of 706 compounds in an external test set. A comparison of the statistical parameters of the QuBiLs-MAS LDA-based models with those for models reported in the literature reveals comparable to superior performance, although the latter were built over much smaller and less diverse datasets, representing fewer mechanisms of action. It may therefore be inferred that the QuBiLs-MAS method constitutes a valuable tool useful in the design and/or selection of new and broad spectrum agents against life-threatening fungal infections.

Keywords: QuBiLs-MAS software; atom-based quadratic indices; linear discriminant analysis; QSAR model; virtual screening, antifungal agent

1. Introduction

In recent times, the incidence of life-threatening fungal infections has increased and is directly related to the increase in the population of patients at risk of developing serious fungal infections, including those placed under major surgery such as solid organ and hematopoietic stem cell transplantation, hemodialysis, or those with HIV infections, chemotherapy-induced

*Corresponding author. Email: sjbarigye@gmail.com

neutropenia, advanced age and premature birth, as well as from the use of broad-spectrum antibiotics and glucocorticosteroids [1–5].

Serious infections are not only produced by the well-known opportunists such as *Candida albicans*, *Cryptococcus neoformans* and *Aspergillus fumigatus,* but also yeast-like fungi such as *Trichosporon* spp., *Rhodotorula* spp. and *Blastoschizomyces capitatus*, hyaline molds such as *Fusarium*, *Acremonium*, *Scedosporium*, *Paecilomyces* and *Trichoderma* species, zygomycetes such as *Rhizopus* spp*., Absidia* spp. and *Rhizomucor* spp., and a wide variety of dematiaceous fungi [3].

The intrinsic resistance, observed in some of these genera, to even new antifungal agents, along with the development of resistance during treatment in others, is becoming a major problem in the management of fungal infections [2]. Furthermore, the clinical utility of the few classes of antifungal drugs on the market is limited by several shortcomings, such as the lack of broad spectrum fungicidal activity, unfavourable routes of administration, severe side effects and undesirable drug–drug interactions [6]. Therefore, novel effective antifungal agents need to be discovered in order to cope with this situation.

Computer-aided drug design has emerged as an important strategy for the 'rational' search of new chemical entities (NCEs) of therapeutic interest. Numerous reports on *in silico* methods for drug design have been published in the literature in recent years [7–13]. Indeed, many pharmaceutical companies and academic institutions have reoriented their research to incorporate computational methods as a means of reducing the hit-to-drug timeline, to increase the number of quality candidate drugs that make the transition from discovery to clinical development and to reduce the attrition rate (currently 90%) of candidate drugs in the clinical stages of the value chain [14]. Several approaches for the computer-aided molecular design and high-throughput *in silico* screening (or virtual high-throughput screening) have been introduced in the literature [15–20].

In recent years, there has been increasing interest in the search of NCEs with fungicidal activity. To this effect, computer-aided drug design techniques based on quantitative structure–activity-relationship (QSAR) studies have played an important role. Unfortunately, almost all the antifungal QSAR studies reported so far are based on rather limited databases, considering only structurally related compounds with *specific* action modes or acting against *a single* fungus species [21–25]. Therefore, most of the previous QSAR studies may be considered as local models based on a small-to-medium chemical space spectrum*,* with limited capacity to predict the activity/inactivity of different ligands on specific molecular targets. As a result, researchers interested in predicting the antifungal activity for a given series of compounds need to use/develop many QSAR equations for the combinations of structurally heterogeneous families of compounds to be studied.

Although some global QSAR equations have been developed in the past few years [26–29], the main drawbacks of these equations are their reduced applicability domain and the rather small diversity of chemical structural patterns considering the actual chemical space. Therefore, the development of a single equation explaining the antifungal activity of structurally heterogeneous series of compounds and covering a broad range of mechanisms of action as possible is of major interest.

In this context, our research group has recently introduced a novel scheme, denominated as TOMOCOMD–CARDD (acronym of Topological MOlecular COMputational Design–Computer Aided 'Rational' Drug Design) able to generate 2D (topologic), 2.5 (3D-chiral) and 3D (topographic and geometric) molecular descriptors, based on application of the discrete mathematics and linear algebra to chemistry. To this effect, atomic, group and atom-type, as well as total linear, quadratic and bilinear molecular indices, have been defined in analogy to

the linear, quadratic and bilinear mathematical maps, and these collectively constitute the QuBiLS-MAS (acronym for Quadratic, Bilinear and Linear MapS based on Graph–Theoretic Electronic-Density Matrices and Atomic weightingS) module of the TOMOCOMD–CARDD suite [30–35]. The QuBiLS-MAS approach has been successfully applied to the prediction of several physical, physicochemical and chemical properties of organic compounds [30, 31] and in the screening of NCEs of therapeutic interests, e.g. tyrosinase inhibitors [36], anthelmintics, antiprotozoals, antibacterial and antimalarial compounds, etc. [8, 9, 37–44]. The satisfactory predictive ability of this ligand-based method is an indication that QuBiLS-MAS descriptors codify important chemical information, and constitute a valuable tool in the drug discovery process [8, 9, 37–42].

The main objectives of the present report are, firstly, to construct a large and structurally diverse antifungal database for modelling the mechanisms of antifungal activity known so far and, secondly, to develop classification models, using the QuBiLS-MAS descriptors and linear discriminant analysis. Finally, the results of the current study are compared with those reported in the literature to gain greater insight on the performance of the obtained antifungal models.

2. Materials and methods

2.1 *Chemical compound dataset*

Though antifungal compounds exhibit enormous structural diversity and action modes, only a small portion of that diversity has been explored for its pharmacological potential so far, and there is thus little reason to believe that this potential has now run dry. For this reason a large database comprising 2142 organic chemicals having a great structural variability, with 1087 of them being antifungal agents [6, 22, 27, 45–103] covering the broadest antifungal mechanisms of actions known so far and the rest inactive ones (1055 compounds having other clinical uses such as antivirals, sedative/hypnotics, diuretics, anticonvulsivants, hemostatics, oral hypoglycemics, antihypertensives, antihelminthics, anticancer compounds and so on) was constructed [104].

The dataset of antifungal agents (active compounds) was chosen considering the largest representation of the action modes so far known, i.e. compounds interfering with cell wall synthesis (chitin synthesis inhibitors such as polyoxins and nikkomycins, and ß-1,3 glucan synthesis inhibitors such as echinocandins), agents interfering with membrane sterols (polyenes, azoles, allylamines and morpholines), protein (sordarins) and DNA synthesis inhibitors (flucytosine and pentamidine analogs), as well as inhibitors of *N*-myristoyltransferase [105]. Several compounds reported as antifungals, but whose mechanisms of action is not known, were also included.

2.2 *Computational methods*

Total and local (atom and atom-type), nonstochastic and stochastic quadratic indices were computed over the k^{th} 'nonstochastic and stochastic graph–theoretical electronic-density matrices' M^k and S^k, respectively, for the molecules in the constructed dataset. In the atom-type quadratic indices formalism, each atom in the molecule is classified into an atom-type (fragment) such as heteroatoms, halogen atoms, aliphatic carbon chains, and aromatic atoms (aromatic rings). The mathematical basis and methodological explanation of this approach have been reported elsewhere [30, 31, 38, 40]. In this study, specifically, we used the *k*th (k = 15)

atom-type (heteroatoms: S, N, O) quadratic fingerprints, not considering and considering H-atoms in the molecular pseudograph, respectively [$q_{kL}(\bar{x}_E)$ and $q_{kL}{}^{H}(\bar{x}_E)$].

In this report, Pauling electronegativity [106] scale was used as the atomic weighting scheme (molecular vector's components). Finally, linear discriminant analysis was performed to find the relationship between the antifungal activity and the quadratic fingerprints generated with the QuBiLS-MAS software.

2.3 *Chemometric method*

Linear discriminant analysis (LDA) was performed using the forward stepwise procedure as a strategy for variable selection [107]. For this experiment, STATISTICA software (version 6.0) was employed. In this way, quantitative models with the following form were obtained:

$$\boldsymbol{P} = a_0\boldsymbol{q}_0(\bar{x}) + a_1\boldsymbol{q}_1(\bar{x}) + \ldots + a_n q_n(\bar{x}) + a_{n+1}\boldsymbol{q}_{0L}(\bar{x}) + a_{n+2}q_{1L}(\bar{x}) + \ldots + a_m q_{mL}(\bar{x}) \quad (1)$$

where $\boldsymbol{P}$ is the biological property (in this study $\boldsymbol{P}$ was designated as ***AFA***, acronym of anti-fungal activity), $\boldsymbol{q_n}(\bar{x})$, the n^{th} total quadratic index, $\boldsymbol{q_{mL}}(\bar{x})$, the m^{th} local quadratic index, and $\boldsymbol{a_n's}$ and $\boldsymbol{a_m's}$, the coefficients obtained by LDA (here, $k^{th} = n^{th}$ or m^{th}). The principle of maximal parsimony (Occam's razor) was taken into consideration as a strategy for model selection. Accordingly, models having the highest statistical significance, but keeping as few parameters as possible were selected.

The quality of the models was assessed by examining Wilks' λ parameter (U statistic), which takes values ranging from 0 (perfect discrimination) to 1 (no discrimination), the square Mahalanobis distance (D^2), which indicates the separation between active and inactive groups, the Fisher ratio (F) and its corresponding p level [p(F)]. Finally, the calculation of percentages of global good classification (accuracy), sensibility, specificity (also known as 'hit rate'), false positive rate (also known as 'false alarm rate) and Matthews correlation coefficient (**C**) in the training and test sets were also used to evaluate the models [108].

The Randić method for orthogonalisation of descriptors was followed to avoid the exclusion of descriptors, on the basis of their colinearity with other variables included in the model [109–113]. As a first step, an appropriate order of orthogonalisation was considered following the order in which the variables were selected from the forward stepwise search procedure of the statistical analysis. The first variable (V_1) is taken as the first orthogonal descriptor ^{1}O (V_1), and the second one (V_2) is orthogonalised with respect to it [$^{2}O(V_2)$]. The residual of its correlation with $^{1}O(V_1)$ is that part of the descriptor V_2 not reproduced by $^{1}O(V_1)$. Similarly, from the regression of V_3 versus $^{1}O(V_1)$, the residual is the part of V_3 that is not reproduced by $^{1}O(V_1)$, and it is labelled $^{1}O(V_3)$. The orthogonal descriptor $^{3}O(V_3)$ is obtained by repeating this process in order to also make it orthogonal to $^{2}O(V_2)$. The process is continually repeated until all variables are completely orthogonalised and the orthogonal variables are then used to obtain the new model. For this orthogonalisation procedure, the entire chemical compound dataset was considered.

3. Results and discussion

3.1 *Development of linear discriminant functions*

The dataset was randomly split into multiple (10-fold) training set and test sets containing 1436 compounds (717 antifungal and 719 inactive) and 706 compounds (370 antifungal and 336 inactive), respectively. Iterative external validation offsets any possible variation in the results due to a biased splitting of the chemical compound dataset. The best discrimination

functions, obtained with non-stochastic and stochastic quadratic indices, respectively, for the training set are given below:

$$\text{AFA} = -5.01 + 5.94 \times 10^{-4}\boldsymbol{q}_5(\bar{x}) - 1.24 \times 10^{-4}\boldsymbol{q}_6(\bar{x}) - 0.02\boldsymbol{q}_{1\text{L}}{}^{\text{H}}(\bar{x}_{\text{E}}) - 1.59 \times 10^{-7}\boldsymbol{q}_{12\text{L}}(\bar{x}_{\text{E-H}}) \quad (2)$$

n (training) = 1436, λ = 0.41, D^2 = 5.69, F(4,1431) = 509.72, Rcan = 0.766, χ^2 = 1268.38, $p < 0.0001$

$$\text{AFA} = -4.44 + 0.20^{\text{s}}\boldsymbol{q}_{11}(\bar{x}) + 0.50^{\text{s}}\boldsymbol{q}_{8\text{L}}{}^{\text{H}}(\bar{x}_{\text{E}}) - 0.08^{\text{s}}\boldsymbol{q}_0{}^{\text{H}}(\bar{x}) - 0.62^{\text{s}}\boldsymbol{q}_{6\text{L}}{}^{\text{H}}(\bar{x}_{\text{E}}) \quad (3)$$

n (training) = 1436, λ = 0.39, D^2 = 6.07, F(4,1431) = 544.51, Rcan = 0.777, χ^2 = 1324.69, $p < 0.0001$

where n is the number of compounds, λ is Wilks' lambda, D^2 is the square Mahalanobis distance, F is the Fisher ratio, p value is the significance level, and Rcan and χ^2 are the correlation coefficient and chi-square parameter of the canonical LDA analysis, respectively.

While Equation (2) classified correctly 90.73% of the compounds in the training set, misclassifying only 133 out of 1436 chemicals, Equation (3) classified correctly 92.47% of compounds, misclassifying only 108 chemicals. As it can be appreciated from Table 1, stochastic quadratic indices were better in predicting the antifungal activity than non-stochastic quadratic indices in the training set, not only because of their better accuracy and Matthew's correlation coefficient, but also due to their higher sensitivity, specificity and lower false positive rate. In general terms, however, both models were good in describing the antifungal activity of chemical compounds. The classification for all active and inactive training compounds according to Equations (2) and (3) is available as supplementary information SM1 and SM2, respectively, in the supplementary material which is available via the multimedia link on the online article webpage.

The results obtained from the training set provide information on the predictive power of the developed models. However, the earnest predictive power is assessed on a set of compounds not employed in the model building, i.e. the test set [114, 115]. For this purpose, the performance of the obtained discriminant functions (Equations (2) and (3)) was assessed on the external test set. In this case, Equation (2) correctly classified 92.16% (274/282) of the active compounds and 91.96% (141/160) of the inactives, whereas Equation (3) correctly classified 87.56% (270/282) of the actives and 91.96% (143/160) of the inactives, giving

Table 1. Results of the training and prediction series performances using the atom-based quadratic indices.

	Matthew's correlation coefficient	*Accuracy 'Q_{Total}'(%)*	*Sensitivity (%)*	*Specificity (%)*	*False positive rate"false alarm rate' (%)*
Non-stochastic descriptors (Equation (2))					
Training set	0.81	90.73	91.07	90.44	9.59
Test set	0.84	92.06	92.16	92.66	8.03
Stochastic descriptors (Equation (3))					
Training set	0.85	92.47	91.21	93.56	6.25
Test set	0.79	89.66	87.56	92.30	8.03

overall accuracies of 92.06% (27/442) and 89.66% (29/442), respectively (for more details, see Table 1). Contrary to what was observed in the training set, non-stochastic quadratic indices were slightly superior in predicting the antifungal activity in the test set. Nonetheless, the predictive power of both models may be considered satisfactory considering the examined statistical parameters. The classification for all active and inactive test compounds according to Equations (2) and (3) is available as supplementary information SM3 and SM4, respectively, in the supplementary material which is available via the multimedia link on the online article webpage.

3.2 *Orthogonalisation of descriptors*

In the orthogonalisation process, molecular descriptors are transformed in such a way that they do not mutually correlate with each other. With this procedure, the exclusion of descriptors, based on their colinearity with other variables previously included in the model, is avoided as a way to improve the statistical interpretation of the model by using interrelated indices [109–113]. Both, the non-orthogonal descriptors and the derived orthogonal descriptors contain the same information. Therefore, the same statistical parameters of the QSAR models are obtained [109–113]. It is known that the interrelation among different descriptors can result in highly unstable regression coefficients, which makes it almost impossible to determine the relative importance of an index included in a model. In other cases, however, strongly interrelated descriptors can enhance the quality of a model because the small fraction of a descriptor that is not reproduced by its strongly interrelated pair can provide positive contributions to the model. Furthermore, the coefficients of the QSAR model, based on orthogonal descriptors, are stable against the inclusion of new descriptors, facilitating the interpretation of the regression coefficients and the evaluation of the role of individual fingerprints in the QSAR model.

The results of the orthogonalisation of molecular descriptors included in both models are shown in Table 2. Equations (2a) and (3a) represent the final models with the orthogonalised molecular indices, whereas in the symbolism ${}^{m}\mathrm{O}[\boldsymbol{q}_k(\bar{x})]$, the superscript m expresses the order of importance of the variable $[\boldsymbol{q}_k(\bar{x})]$ after a preliminary forward stepwise analysis, and O means orthogonal (see Table 2). As can be observed there is total correspondence for all statistical parameters for the orthogonal descriptor-based and linear descriptor-based models.

This feature facilitates the interpretation of the coefficients in the LDA-QSAR equations. Therefore, ${}^{m}\mathrm{O}(\boldsymbol{q}_k(\bar{x}))$ may be classified according to the distance k into short- (0–5), mid- (6–10) and long-range non-stochastic and stochastic quadratic indices. As may be observed in Table 2, short-, middle- and long-range total and atom-type (heteroatoms and H-atom bonding to heteroatoms) quadratic indices all contribute to the models, with the local quadratic indices offering greater contribution to the models' performance. Nevertheless, total variables such as the zero order indices were included in the models, indicating that the size and atomic composition of the molecules are important for their activity. The high contribution of local variables may be explained by the fact that the mechanisms of action of antifungal drugs are direct and local in nature. Therefore weak non-covalent interactions propitiated by the heteroatoms' electronic distribution are very important for their interaction with receptors. However, the inclusion of local variables of superior order, in both models, demonstrates that an adequate molecular environment is also required, for the interaction of antifungal drugs with their pharmacological target.

3.3 *Comparison with other approaches for antifungal activity*

In the past few years, various *in silico* methods have been used to develop ligand-based classification models of antifungal activity [27–29]. However, an exhaustive comparison between these models and the ones developed herein is not possible because of the differences in the used experimental data. Therefore, the comparison is based on the number and diversity of chemical structural patterns contained by the data, as well as on some classification and statistical parameters. Table 3 shows the comparison between antifungal models developed using the QuBiLS-MAS method and other approaches reported in the literature.

Firstly, the dataset used to develop QuBiLS-MAS based models is a lot larger (more than 20 times) than those employed for the rest of the models (see models 4–9 in Table 3), in addition to presenting broader structural diversity. However, models 4–9 were built from datasets essentially comprising azoles, allylamines and thiocarbamates, and thus representative of a reduced chemical structural space. Such models possess a much smaller applicability domain, which in turn compromises their applicability in virtual screening tasks. In addition, these models cover a short range of mechanisms of action, dealing almost exclusively with the inhibition of the biosynthesis of ergosterol, as compared with the broad range of mechanisms covered by our models.

Even then, with the exception of model 6 with a global percentage of correct classification of 96.92% for the training set, the LDA-based QuBiLS-MAS models yielded higher accuracy values than the rest of all the models reported in the literature (see Table 3), although some of these were built using non-linear techniques such as support vector machines (SVM) and decision trees (C4.5), which are generally known to yield better fitted and more robust models.

Table 2. Results of Randić's orthogonalisation analysis.

Non-orthogonal quadratic indices							
$q_5(\bar{x})$	$q_6(\bar{x})$	${}^{1}q_{L}^{H}(\bar{x}_E)$	$q_{12L}(\bar{x}_{E\text{-}H})$	${}^{s}q_{11}(\bar{x})$	${}^{s}q_{8L}^{H}(\bar{x}_E)$	${}^{s}q_{0}^{H}(\bar{x})$	${}^{s}q_{6L}^{H}(\bar{x}_E)$
1.00	0.99	0.61	0.54	1.00	0.90	0.99	0.90
	1.00	0.55	0.54		1.00	0.87	0.99
		1.00	0.56			1.00	0.87
			1.00				1.00
Orthogonal quadratic indices							
$O(\mathbf{q}_5(x))$	$O(\mathbf{q}_6(x))$	$O(\mathbf{q}_{1L}{}^{H}(x_E))$	$O(\mathbf{q}_{12L}(x_{E\text{-}H}))$	${}^{1}O({}^{s}\mathbf{q}_{11}(x))$	${}^{2}O({}^{s}\mathbf{q}_{8L}{}^{H}(x_E))$	${}^{3}O({}^{s}\mathbf{q}_{0}{}^{H}(x))$	${}^{4}O({}^{s}\mathbf{q}_{6L}{}^{H}(x_E))$
1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
	1.00	0.00	0.00		1.00	0.00	0.00
		1.00	0.00			1.00	0.00
			1.00				1.00

LDA based QSAR models derived from the orthogonal non-stochastic and stochastic quadratic indices

$AFA = -0.004 + 3.24{}^{1}O(q_5(\bar{x})) - 1.47{}^{2}O(q_{12L}(\bar{x}_{E\text{-}H}))\ 10.81{}^{3}O(q_{1L}{}^{H}(\bar{x}_E)) - 1.001{}^{4}O(q_6(\bar{x}x))$ (Equation 2a)
$n = 1436$; $\lambda = 0.41$; $D^2 = 5.69$; $F(4,1431) = 509.72$; Rcan = 0.766; $\chi^2 = 1268.38$; $C = 0.81$; $Q_{Total} = 90.73$

$AFA = -0.007 + 2.38{}^{1}O({}^{s}q_{11}(\bar{x})) - 4.67{}^{2}O({}^{s}q_{8L}{}^{H}(\bar{x}_E)) - 17.56{}^{3}O({}^{s}q_{0}{}^{H}(\bar{x}))\ -37.0{}^{4}O({}^{s}q_{6L}{}^{H}(\bar{x}_E))$ (Equation 3a)
$n = 1436$; $\lambda = 0.39$; $D^2 = 6.07$; $F(4,1431) = 544.51$; Rcan = 0.777; $\chi^2 = 1324.69$; $C = 0.85$; $Q_{Total} = 92.47$

Table 3. Comparison between the models developed in this study with other cheminformatic approaches.

	Classification models of the antifungal activity							
Models' features to be compared[a]	*Eq. 2 (LDA)*	*Eq. 3 (LDA)*	*Eq. 4 (LDA)*	*Eq. 5 (LDA)*	*Eq. 6 (LDA)*	*Eq. 7 (SVM)*	*Eq. 8 (k-NN)*	*Eq. 9 (C4.5)*
n total	2142	2142	94	94	90	94	94	94
n antifungals	1087	1087	42	42	49	42	42	42
Wilks' λ	0.41	0.39	0.392	0.387	0.32	–	–	–
F	509.72	544.51	14.5	19.5	17.2	–	–	–
D^2	5.69	6.07	–	–	–	–	–	–
p level	<0.0001	<0.0001	–	–	–	–	–	–
χ^2	1268.38	1324.69	–	–	–	–	–	–
Rcan	0.766	0.777	–	–	–	–	–	–
Variables in the model	4	4	9	9	8	60(30)*	30*	30*
Learning set								
n total	1436	1436	94	94	65	94**	94**	94**
n antifungals	717	717	42	42	36	42**	42**	42**
Matthews correlation coefficient (*C*)	0.81	0.85	0.74	0.77	0.94			
Accuracy 'Q_{Total}'	90.73	92.47	87.23	88.29	96.92	84.0(89.4)**	76.5**	75.6**
Specificity (%)	90.44	93.56	87.50	91.89	97.22	–	–	–
Sensitivity (%)	91.07	91.21	83.33	80.95	97.22	91.0(97.1) **	71.7	73.5
False + Rate (%)	9.56	6.25	9.62	5.76	3.44	–	–	–
Families of drugs[b]	Broad range	Broad range	Short range	Short range	Short range	Short range	Short range	Short range
Validation methods								
Validation method[c]	i	i	ii	ii	i	iii	iii	iii
n total	706	706	–	–	25	–	–	–
n antifungals	370	370	–	–	13	–	–	–
Matthews correlation coefficient (*C*)	0.84	0.79	0.71	0.75	0.60	–	–	–
Accuracy 'Q_{Total}'	92.06	89.66	85.56	87.80	80.00	77.8	75.0	85.7
Specificity (%)	92.66	92.30	85.71	88.57	83.33	–	–	–
Sensitivity (%)	92.16	87.56	85.71	83.78	76.92	–	–	–
False + Rate (%)	8.03	8.03	14.58	8.88	16.66	-	-	-
Families of drugs[b]	Broad range	Broad range	Short range	Short range	Short range	Short range	Short range	Short range

*Equation fitted employed the whole set of 62 variables, between bracket number of predictors selected by using genetic algorithm based SVM [29].
**Result obtained from cross-validation experiment by using five-fold out.
[a]Equations (2) and (3) are reported in this work; the models 4 and 5 were reported by Pastor et al. [27]; Equation (6) was reported by García-Domenech et al. [28] and models 7–9 were reported by Li *et al*. [29].
[b]Only compound families with a wide representativity were taken into account.
[c]Validation methods are: (i) external prediction series; (ii) leave-20%-out; and (iii) an independent test set.

Additionally, it is important to note that model 6 possesses a much higher degree of freedom (eight variables) relative to the QuBiLS-MAS model (four variables) and therefore its favourable result may probably be due to overfitting. It is thus not surprising that the test set accuracy for model 6 falls significantly to 80% (from 96.92% for the training set), while the models 2 and 3 yield accuracy values of 92.06% and 89.66%, respectively. It may therefore be concluded that the LDA-based QuBiLS-MAS models obtained in the present study possess high robustness and predictive power, in addition to a wide applicability domain, and therefore constitute valuable tools for use in the virtual screening of compounds with antifungal activity.

4. Concluding remarks and future outlooks

In the past two decades, the number of patients with severe fungal infections has increased dramatically and concern over the rapid development of resistance to the few antifungal drugs available has risen [116]. Despite aggressive management, the prognosis of invasive fungal diseases, in particular those caused by filamentous fungi, continues to be dismal, with mortality rates exceeding 80% in selected categories of patients [117].

Although the need for new drugs is clear, progress in this area is slow and unpredictable. It is suggested that the ideal antifungal agent of the future should have a broad spectrum of fungicidal activity and be without mechanism-based host toxicity. The antifungal models developed in the present work cover a wide chemical structural domain, a broad range of mechanisms of action, coupled with their high accuracy, sensitivity and specificity to predict the antifungal activity, and thus may be considered as valuable tools for use in the screening of novel and broad spectrum antifungal agents.

Program availability

The QuBiLs-MAS software (portable standalone) and the respective user manual are freely available online at www.tomocomd.com.

Supplementary material

The complete list of compounds used in training and prediction sets, as well as their structures, posterior classification and scores according to LDA-based QSAR models is available free of charge in the supplementary material which is available via the multimedia link on the online article webpage.

Acknowledgements

M.-P.Y. acknowledges Valencia University, Spain, for partial financial support as well as the programme 'Estades Temporals per an Investigadors Convidats' for a fellowship to work at Valencia University (2015). S.J.B. acknowledges financial support from CNPq.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] H.M. Blumberg, W.R. Jarvis, J.M. Soucie, J.E. Edwards, J.E. Patterson, M.A. Pfaller, M.S. Rangel-Frausto, M.G. Rinaldi, L. Saiman, R.T. Wiblin, and R.P. Wenzel, *Risk factors for candidal bloodstream infections in surgical intensive care unit patients: The NEMIS prospective multicenter study. The National Epidemiology of Mycosis Survey*, Clin. Infect. Dis. 33 (2001), pp. 177–186.

[2] G. Maschmeyer, *The changing epidemiology of invasive fungal infections: New threats*, Int. J. Antimicrob. Agents 27(Suppl. 1) (2006), pp. 3–6.

[3] M. Nucci and K.A. Marr, *Emerging fungal diseases*, Clin. Infect. Dis. 41 (2005), pp. 521–526.

[4] J.R. Rees, R.W. Pinner, R.A. Hajjeh, M.E. Brandt, and A.L. Reingold, *The epidemiological features of invasive mycotic infections in the San Francisco Bay area, 1992–1993: Results of population-based laboratory active surveillance*, Clin. Infect. Dis. 27 (1998), pp. 1138–1147.

[5] T.J. Walsh, A. Groll, J. Hiemenz, R. Fleming, E. Roilides, and E. Anaissie, *Infections due to emerging and uncommon medically important fungal pathogens*, Clin. Microbiol. Infect. 10(Suppl 1) (2004), pp. 48–66.

[6] D.A. Dickman, H. Ding, Q. Li, A.M. Nilius, D.J. Balli, S.J. Ballaron, J.M. Trevillyan, M.L. Smith, L.S. Seif, K. Kim, A. Sarthy, R.C. Goldman, J.J. Plattner, and Y.L. Bennani, *Antifungal rapamycin analogues with reduced immunosuppressive activity*, Bioorg. Med. Chem. Lett. 10 (2000), pp. 1405–1408.

[7] R. Gozalbes, J. Galvez, A. Moreno, and R. Garcia-Domenech, *Discovery of new antimalarial compounds by use of molecular connectivity techniques*, J. Pharm. Pharmacol. 51 (1999), pp. 111–117.

[8] Y. Marrero-Ponce, A. Montero-Torres, C.R. Zaldivar, M.I. Veitia, M.M. Perez, and R.N. Sanchez, *Non-stochastic and stochastic linear indices of the 'molecular pseudograph's atom adjacency matrix': Application to 'in silico' studies for the rational discovery of new antimalarial compounds*, Bioorg. Med. Chem. 13 (2005), pp. 1293–1304.

[9] Y. Marrero-Ponce, M. Iyarreta-Veitia, A. Montero-Torres, C. Romero-Zaldivar, C.A. Brandt, P.E. Avila, K. Kirchgatter, and Y. Machado, *Ligand-based virtual screening and in silico design of new antimalarial compounds using nonstochastic and stochastic total and atom-type quadratic maps*, J. Chem. Inf. Model. 45 (2005), pp. 1082–1100.

[10] J.H. McKie, K.T. Douglas, C. Chan, S.A. Roser, R. Yates, M. Read, J.E. Hyde, M.J. Dascombe, Y. Yuthavong, and W. Sirawaraporn, *Rational drug design approach for overcoming drug resistance: Application to pyrimethamine resistance in malaria*, J. Med. Chem. 41 (1998), pp. 1367–1370.

[11] E. Estrada, A. Pena, and R. Garcia-Domenech, *Designing sedative/hypnotic compounds from a novel substructural graph-theoretical approach*, J. Comput. Aided Mol. Des. 12 (1998), pp. 583–595.

[12] E. Estrada, E. Uriarte, A. Montero, M. Teijeira, L. Santana, and E. De Clercq, *A novel approach for the virtual screening and rational design of anticancer compounds*, J. Med. Chem. 43 (2000), pp. 1975–1985.

[13] H. Gonzalez-Diaz, E. Tenorio, N. Castanedo, L. Santana, and E. Uriarte, *3D QSAR Markov model for drug-induced eosinophilia – –theoretical prediction and preliminary experimental assay of the antimicrobial drug G1*, Bioorg. Med. Chem. 13 (2005), pp. 1523–1530.

[14] P.D. Lyne, *Structure-based virtual screening: An overview*, Drug. Discov. Today 7 (2002), pp. 1047–1055.

[15] R. Todeschini and V. Consonni, *Molecular Descriptors for Chemoinformatics*, Wiley-VCH, Weinheim, 2009.

[16] M.H.J. Seifert, K. Wolf, and D. Vitt, *Virtual high-throughput in silico screening*, Biosilico (2003), pp. 143–149.

[17] S.J. Barigye, Y. Marrero-Ponce, F. Pérez-Giménez, and D. Bonchev, *Trends in information theory-based chemical structure codification*, Mol. Divers. 19 (2014), pp. 305–319.

[18] S.J. Barigye and M.P. Freitas, *2D-Discrete Fourier transform: Generalization of the MIA-QSAR strategy in molecular modeling*, Chemom. Intell. Lab. Syst. 143 (2015), pp. 79–84.

[19] S. Barigye, Y. Marrero-Ponce, Y.M. López, O.M. Santiago, F. Torrens, R.G. Domenech, and J. Galvez, *Event-based criteria in GT-STAF information indices: Theory, exploratory diversity analysis and QSPR applications*, SAR QSAR Environ Res. 24 (2013), pp. 3–34.
[20] S.J. Barigye and M.P. Freitas, *Is molecular alignment an indispensable requirement in the MIA-QSAR method?*, J. Comput. Chem. 36 (2015), pp. 1748–1755.
[21] R. Gollapudy, S. Ajmani, and S.A. Kulkarni, *Modeling and interactions of* Aspergillus fumigatus *lanosterol 14-alpha demethylase 'A' with azole antifungals*, Bioorg. Med. Chem. 12 (2004), pp. 2937–2950.
[22] V.M. Gokhale and V.M. Kulkarni, *Understanding the antifungal activity of terbinafine analogues using quantitative structure–activity relationship (QSAR) models*, Bioorg. Med. Chem. 8 (2000), pp. 2487–2499.
[23] P. Purushottamachar and V.M. Kulkarni, *3D-QSAR of N-myristoyltransferase inhibiting antifungal agents by CoMFA and CoMSIA methods*, Bioorg. Med. Chem. 11 (2003), pp. 3487–3497.
[24] S.K. Pranav, Kumar and V.M. Kulkarni, *Insights into the selective inhibition of* Candida albicans *secreted aspartyl protease: A docking analysis study*, Bioorg. Med. Chem. 10 (2002), pp. 1153–1170.
[25] F. Fratev and E. Benfenati, *3D-QSAR and molecular mechanics study for the differences in the azole activity against yeastlike and filamentous fungi and their relation to P450DM inhibition. 1. 3-Substituted-4(3H)-quinazolinones*, J. Chem. Inf. Model. 45 (2005), pp. 634–644.
[26] H. Gonzalez-Diaz, F.J. Prado-Prado, L. Santana, and E. Uriarte, *Unify QSAR approach to antimicrobials. Part 1: Predicting antifungal activity against different species*, Bioorg. Med. Chem. 14 (2006), pp. 5973–5980.
[27] L. Pastor, R. Garcia-Domenech, J. Galvez, S. Wolski, and M.D. Garcia, *New antifungals selected by molecular topology*, Bioorg. Med. Chem. Lett. 8 (1998), pp. 2577–2582.
[28] R. García-Domenech, A. Catalá-Gregori, C. Calabuig, G. Antón-Fos, L. del Castillo, and J. Gálvez, *Predicting antifungal activity: A computational screening using topological descriptors*, Internet Electron. J. Mol. Des. 7 (2002), pp. 339–350.
[29] S.-W. Chen, Z.-R. Li, and X.-Y. Li, *Prediction of antifungal activity by support vector machine approach*, J. Mol. Struct. (THEOCHEM) 731 (2005), pp. 73–81.
[30] Y. Marrero-Ponce, *Total and local quadratic indices of the molecular pseudograph's atom adjacency matrix: Applications to the prediction of physical properties of organic compounds*, Molecules 8 (2003), pp. 687–726.
[31] Y. Marrero-Ponce, *Linear indices of the 'molecular pseudograph's atom adjacency matrix': Definition, significance-interpretation, and application to QSAR analysis of flavone derivatives as HIV-1 integrase inhibitors*, J. Chem. Inf. Comput. Sci. 44 (2004), pp. 2010–2026.
[32] Y. Marrero-Ponce, A. Meneses-Marcel, J.A. Castillo-Garit, Y. Machado-Tugores, J.A. Escario, A.G. Barrio, D.M. Pereira, J.J. Nogal-Ruiz, V.J. Aran, A.R. Martinez-Fernandez, F. Torrens, R. Rotondo, F. Ibarra-Velarde, and Y.J. Alvarado, *Predicting antitrichomonal activity: A computational screening using atom-based bilinear indices and experimental proofs*, Bioorg. Med. Chem. 14 (2006), pp. 6502–6524.
[33] C.R. García-Jacas, Y. Marrero-Ponce, L. Acevedo-Martínez, S.J. Barigye, J.R. Valdés-Martiní, and E. Contreras-Torres, *QuBiLS-MIDAS: A parallel free-software for molecular descriptors computation based on multilinear algebraic maps*, J. Comp. Chem. 35 (2014), pp. 1395–1409.
[34] C.R. Garcia-Jacas, Y. Marrero-Ponce, S.J. Barigye, J.R. Valdes-Martini, O.M. Rivera-Borroto, and J. Olivero-Verbel, *N-linear algebraic maps for chemical structure codification: A suitable generalization for atom-pair approaches?*, Curr. Drug Metab. 15 (2014), pp. 441–469.
[35] C.R. García-Jacas, L. Aguilera-Mendoza, R. González-Pérez, Y. Marrero-Ponce, L. Acevedo-Martínez, S.J. Barigye, and T. Avdeenko, *Multi-server approach for high-throughput molecular descriptors calculation based on multi-linear algebraic maps*, Mol. Info. 34 (2015), pp. 60–69.
[36] G.M. Casanola-Martin, M.T. Khan, Y. Marrero-Ponce, A. Ather, M.N. Sultankhodzhaev, and F. Torrens, *New tyrosinase inhibitors selected by atomic linear indices-based classification models*, Bioorg. Med. Chem. Lett. 16 (2006), pp. 324–330.

[37] Y. Marrero-Ponce, J.A. Castillo-Garit, E. Olazabal, H.S. Serrano, A. Morales, N. Castanedo, F. Ibarra-Velarde, A. Huesca-Guillen, E. Jorge, A. del Valle, F. Torrens, and E.A. Castro, *TOMO-COMD-CARDD, a novel approach for computer-aided 'rational' drug design: I. Theoretical and experimental assessment of a promising method for computational screening and in silico design of new anthelmintic compounds*, J. Comput. Aided Mol. Des. 18 (2004), pp. 615–634.

[38] Y. Marrero-Ponce, J.A. Castillo-Garit, E. Olazabal, H.S. Serrano, A. Morales, N. Castanedo, F. Ibarra-Velarde, A. Huesca-Guillen, A.M. Sanchez, F. Torrens, and E.A. Castro, *Atom, atom-type and total molecular linear indices as a promising approach for bioorganic and medicinal chemistry: Theoretical and experimental assessment of a novel method for virtual screening and rational design of new lead anthelmintic*, Bioorg. Med. Chem. 13 (2005), pp. 1005–1020.

[39] Y. Marrero-Ponce, Y. Machado-Tugores, D.M. Pereira, J.A. Escario, A.G. Barrio, J.J. Nogal-Ruiz, C. Ochoa, V.J. Aran, A.R. Martinez-Fernandez, R.N. Sanchez, A. Montero-Torres, F. Torrens, and A. Meneses-Marcel, *A computer-based approach to the rational discovery of new trichomonacidal drugs by atom-type linear indices*, Curr. Drug Discov. Technol. 2 (2005), pp. 245–265.

[40] Y. Marrero-Ponce, R.M. Marrero, F. Torrens, Y. Martinez, M.G. Bernal, V.R. Zaldivar, E.A. Castro, and R.G. Abalo, *Non-stochastic and stochastic linear indices of the molecular pseudograph's atom-adjacency matrix: A novel approach for computational in silico screening and 'rational' selection of new lead antibacterial agents*, J. Mol. Model. (Online) 12 (2006), pp. 255–271.

[41] Y. Marrero-Ponce, R. Medina-Marrero, F. Torrens, Y. Martinez, V. Romero-Zaldivar, and E.A. Castro, *Atom, atom-type, and total nonstochastic and stochastic quadratic fingerprints: A promising approach for modeling of antibacterial activity*, Bioorg. Med. Chem. 13 (2005), pp. 2881–2899.

[42] M.C. Vega, A. Montero-Torres, Y. Marrero-Ponce, M. Rolon, A. Gomez-Barrio, J.A. Escario, V.J. Aran, J.J. Nogal, A. Meneses-Marcel, and F. Torrens, *New ligand-based approach for the discovery of antitrypanosomal compounds*, Bioorg. Med. Chem. Lett. 16 (2006), pp. 1898–1904.

[43] M.A.M. Alho, Y. Marrero-Ponce, S.J. Barigye, A. Meneses-Marcel, Y.M. Tugores, A. Montero-Torres, A. Gómez-Barrio, J.J. Nogal, R.N. García-Sánchez, and M.C. Vega, *Antiprotozoan lead discovery by aligning dry and wet screening: Prediction, synthesis, and biological assay of novel quinoxalinones*, Bioorg. Med. Chem. 22 (2014), pp. 1568–1585.

[44] J.A. Castillo-Garit, Y. Marrero-Ponce, S.J. Barigye, R. Medina-Marrero, M.G. Bernal, J.M. de la Vega, F. Torrens, V.J. Arán, F. Pérez-Giménez, and R. García-Domenechd, *In silico antibacterial activity modeling based on the TOMOCOMD-CARDD approach*, J. Braz. Chem. Soc. 26 (2015), pp. 1218–1226.

[45] P.M. Bedi, M.P. Mahajan, and V.K. Kapoor, *Amidine derived 1,3-diazabuta-1,3-dienes as potential antibacterial and antifungal agents*, Bioorg. Med. Chem. Lett. 14 (2004), pp. 3821–3824.

[46] T.G. Brayman and J.W. Wilks, *Sensitive assay for antifungal activity of glucan synthase inhibitors that uses germ tube formation in* Candida albicans *as an end point*, Antimicrob. Agents Chemother. 47 (2003), pp. 3305–3310.

[47] A. Bryskier, *Novelties in the field of anti-infective compounds in 1999*, Clin. Infect. Dis. 31 (2000), pp. 1423–1466.

[48] V. Buchta, M. Pour, P. Kubanova, L. Silva, I. Votruba, M. Voprsalova, R. Schiller, H. Fakova, and M. Spulak, *In vitro activities of 3-(halogenated phenyl)-5-acyloxymethyl- 2,5-dihydrofuran-2-ones against common and emerging yeasts and molds*, Antimicrob. Agents Chemother. 48 (2004), pp. 873–878.

[49] J.M. Bueno, J.C. Cuevas, J.M. Fiandor, S. Garcia-Ochoa, and F. Gomez de las Heras, *Antifungal sordarins. Synthesis and structure-activity relationships of 3',4'-fused dioxolane and dioxane derivatives*, Bioorg. Med. Chem. Lett. 12 (2002), pp. 121–124.

[50] A. Carta, G. Paglietti, and M.E. Rahbar Nikookar, P. Sanna, L. Sechi, and S. Zanetti, *Novel substituted quinoxaline 1,4-dioxides with in vitro antimycobacterial and anticandida activity*, Eur. J. Med. Chem. 37 (2002), pp. 355–366.

[51] K.V. Clemons and D.A. Stevens, *Efficacies of two novel azole derivatives each containing a morpholine ring, UR-9746 and UR-9751, against systemic murine coccidioidomycosis*, Antimicrob. Agents Chemother. 41 (1997), pp. 200–203.

[52] M. Debono and R.S. Gordee, *Antibiotics that inhibit fungal cell wall development*, Annu. Rev. Microbiol. 48 (1994), pp. 471–497.
[53] M. Del Poeta, W.A. Schell, C.C. Dykstra, S. Jones, R.R. Tidwell, A. Czarny, M. Bajic, A. Kumar, D. Boykin, and J.R. Perfect, *Structure-in vitro activity relationships of pentamidine analogues and dication-substituted bis-benzimidazoles as new antifungal agents*, Antimicrob. Agents Chemother. 42 (1998), pp. 2495–2502.
[54] M. Del Poeta, W.A. Schell, C.C. Dykstra, S.K. Jones, R.R. Tidwell, A. Kumar, D.W. Boykin, and J.R. Perfect, *In vitro antifungal activities of a series of dication-substituted carbazoles, furans, and benzimidazoles*, Antimicrob. Agents Chemother. 42 (1998), pp. 2503–2510.
[55] D.W. Denning, *Echinocandins and pneumocandins – –A new antifungal class with a novel mode of action*, J. Antimicrob. Chemother. 40 (1997), pp. 611–614.
[56] J.M. Dominguez, V.A. Kelly, O.S. Kinsman, and M.S. Marriott, *F. Gomez de las Heras, and J.J. Martin, Sordarins: A new class of antifungals with selective inhibition of the protein synthesis elongation cycle in yeasts*, Antimicrob. Agents Chemother. 42 (1998), pp. 2274–2278.
[57] H. Ebiike, M. Masubuchi, P. Liu, K. Kawasaki, K. Morikami, S. Sogabe, M. Hayase, T. Fujii, K. Sakata, H. Shindoh, Y. Shiratori, Y. Aoki, T. Ohtsuka, and N. Shimma, *Design and synthesis of novel benzofurans as a new class of antifungal agents targeting fungal N-myristoyltransferase. Part 2*, Bioorg. Med. Chem. Lett. 12 (2002), pp. 607–610.
[58] S. Emami, M. Falahati, A. Banifatemi, and A. Shafiee, *Stereoselective synthesis and antifungal activity of (Z)-trans-3-azolyl-2-methylchromanone oxime ethers*, Bioorg. Med. Chem. 12 (2004), pp. 5881–5889.
[59] B. Favre and N.S. Ryder, *Characterization of squalene epoxidase activity from the dermatophyte* Trichophyton rubrum *and its inhibition by terbinafine and other antimycotic agents*, Antimicrob. Agents Chemother. 40 (1996), pp. 443–447.
[60] J.M. Fostel and P.A. Lartey, *Emerging novel antifungal agents*, Drug. Discov. Today 5 (2000), pp. 25–32.
[61] H. Goker, D.W. Boykin, and S. Yildiz, *Synthesis and potent antimicrobial activity of some novel 2-phenyl or methyl-4H-1-benzopyran-4-ones carrying amidinobenzimidazoles*, Bioorg. Med. Chem. 13 (2005), pp. 1707–1714.
[62] V.M. Gokhale and V.M. Kulkarni, *Selectivity analysis of 5-(arylthio)-2,4-diaminoquinazolines as inhibitors of* Candida albicans *dihydrofolate reductase by molecular dynamics simulations*, J. Comput. Aided Mol. Des. 14 (2000), pp. 495–506.
[63] K. Hata, J. Kimura, H. Miki, T. Toyosawa, M. Moriyama, and K. Katsu, *Efficacy of ER-30346, a novel oral triazole antifungal agent, in experimental models of aspergillosis, candidiasis, and cryptococcosis*, Antimicrob. Agents Chemother. 40 (1996), pp. 2243–2247.
[64] E. Herreros, M.J. Almela, S. Lozano, F. Gomez de las Heras, and D. Gargallo-Viola, *Antifungal activities and cytotoxicity studies of six new azasordarins*, Antimicrob. Agents Chemother. 45 (2001), pp. 3132–3139.
[65] E. Herreros, C.M. Martinez, M.J. Almela, M.S. Marriott, F.G. De Las Heras, and D. Gargallo-Viola, *Sordarins: In vitro activities of new antifungal derivatives against pathogenic yeasts*, Pneumocystis carinii, *and filamentous fungi*, Antimicrob. Agents Chemother. 42 (1998), pp. 2863–2869.
[66] Y. Kamai, T. Harasaki, T. Fukuoka, S. Ohya, K. Uchida, H. Yamaguchi, and S. Kuwahara, *In vitro and in vivo activities of CS-758 (R-120758), a new triazole antifungal agent*, Antimicrob. Agents Chemother. 46 (2002), pp. 367–370.
[67] J.K. Khan, H. Montaseri, M. Poglod, H.Z. Bu, Z. Zuo, S.M. Salama, M. Daneshtalab, and R.G. Micetich, *Interspecies comparison of pharmacokinetics of the novel triazole antifungal agent SYN-2869 and its derivatives*, Antimicrob. Agents Chemother. 44 (2000), pp. 910–915.
[68] I. Kubo, P. Xiao, and K. Fujita, *Antifungal activity of octyl gallate: Structural criteria and mode of action*, Bioorg. Med. Chem. Lett. 11 (2001), pp. 347–350.

[69] B. Lal, V.G. Gund, N.B. Bhise, and A.K. Gangopadhyay, *Mannich reaction: An approach for the synthesis of water soluble mulundocandin analogues*, Bioorg. Med. Chem. 12 (2004), pp. 1751–1768.

[70] B. Lal, V.G. Gund, A.K. Gangopadhyay, S.R. Nadkarni, V. Dikshit, D.K. Chatterjee, and R. Shirvaikar, *Semisynthetic modifications of hemiaminal function at ornithine unit of mulundocandin, towards chemical stability and antifungal activity*, Bioorg. Med. Chem. 11 (2003), pp. 5189–5198.

[71] S. Li, Z. Zhang, A. Cain, B. Wang, M. Long, and J. Taylor, *Antifungal activity of camptothecin, trifolin, and hyperoside isolated from* Camptotheca acuminata, J. Agric. Food Chem. 53 (2005), pp. 32–37.

[72] S.N. Lopez, M.V. Castelli, S.A. Zacchino, J.N. Dominguez, G. Lobo, J. Charris-Charris, J.C. Cortes, J.C. Ribas, C. Devia, A.M. Rodriguez, and R.D. Enriz, *In vitro antifungal evaluation and structure-activity relationships of a new series of chalcone derivatives and synthetic analogues, with inhibitory properties against polymers of the fungal cell wall*, Bioorg Med Chem 9 (2001), pp. 1999–2013.

[73] S.M. Mandala, R.A. Thornton, J. Milligan, M. Rosenbach, M. Garcia-Calvo, H.G. Bull, G. Harris, G.K. Abruzzo, A.M. Flattery, C.J. Gill, K. Bartizal, S. Dreikorn, and M.B. Kurtz, *Rustmicin, a potent antifungal agent, inhibits sphingolipid synthesis at inositol phosphoceramide synthase*, J. Biol. Chem. 273 (1998), pp. 14942–14949.

[74] S.M. Mandala, R.A. Thornton, M. Rosenbach, J. Milligan, M. Garcia-Calvo, H.G. Bull, and M.B. Kurtz, *Khafrefungin, a novel inhibitor of sphingolipid synthesis*, J. Biol. Chem. 272 (1997), pp. 32709–32714.

[75] K. Masubuchi, T. Okada, M. Kohchi, T. Murata, M. Tsukazaki, O. Kondoh, T. Yamazaki, Y. Satoh, Y. Ono, T. Tsukaguchi, K. Kobayashi, N. Ono, T. Inoue, I. Horii, and N. Shimma, *Synthesis and antifungal activities of novel 1,3-beta-d-glucan synthase inhibitors. Part 2*, Bioorg. Med. Chem. Lett. 11 (2001), pp. 1273–1276.

[76] M. Masubuchi, H. Ebiike, K. Kawasaki, S. Sogabe, K. Morikami, Y. Shiratori, S. Tsujii, T. Fujii, K. Sakata, M. Hayase, H. Shindoh, Y. Aoki, T. Ohtsuka, and N. Shimma, *Synthesis and biological activities of benzofuran antifungal agents targeting fungal N-myristoyltransferase*, Bioorg. Med. Chem. 11 (2003), pp. 4463–4478.

[77] M. Masubuchi, K. Kawasaki, H. Ebiike, Y. Ikeda, S. Tsujii, S. Sogabe, T. Fujii, K. Sakata, Y. Shiratori, Y. Aoki, T. Ohtsuka, and N. Shimma, *Design and synthesis of novel benzofurans as a new class of antifungal agents targeting fungal N-myristoyltransferase. Part 1*, Bioorg. Med. Chem. Lett. 11 (2001), pp. 1833–1837.

[78] M. Matsumoto, K. Ishida, A. Konagai, K. Maebashi, and T. Asaoka, *Strong antifungal activity of SS750, a new triazole derivative, is based on its selective binding affinity to cytochrome P450 of fungi*, Antimicrob. Agents Chemother. 46 (2002), pp. 308–314.

[79] G. Menozzi, L. Merello, P. Fossa, S. Schenone, A. Ranise, L. Mosti, F. Bondavalli, R. Loddo, C. Murgioni, V. Mascia, P. La Colla, and E. Tamburini, *Synthesis, antimicrobial activity and molecular modeling studies of halogenated 4-[1H-imidazol-1-yl(phenyl)methyl]-1,5-diphenyl-1H-pyrazoles*, Bioorg. Med. Chem. 12 (2004), pp. 5465–5483.

[80] S. Moreau, M. Varache-Lembege, S. Larrouture, D. Fall, A. Neveu, G. Deffieux, J. Vercauteren, and A. Nuhrich, *(2-Arylhydrazonomethyl)-substituted xanthones as antimycotics*: Synthesis *and fungistatic activity against Candida species*, Eur. J. Med. Chem. 37 (2002), pp. 237–253.

[81] N.H. Nam, S. Sardari, M. Selecky, and K. Parang, *Carboxylic acid and phosphate ester derivatives of fluconazole*: *Synthesis and antifungal activities*, Bioorg. Med. Chem. 12 (2004), pp. 6255–6269.

[82] K. Nimura, Y. Niwano, S. Ishiduka, and R. Fukumoto, *Comparison of in vitro antifungal activities of topical antimycotics launched in 1990s in Japan*, Int. J. Antimicrob. Agents 18 (2001), pp. 173–178.

[83] Y. Niwano, N. Kuzuhara, Y. Goto, Y. Munechika, H. Kodama, K. Kanai, M. Yoshida, T. Miyazaki, and H. Yamaguchi, *Efficacy of NND-502, a novel imidazole antimycotic agent, in experimental*

models of Candida albicans *and* Aspergillus fumigatus *infections*, Int. J. Antimicrob. Agents 12 (1999), pp. 221–228.

[84] P. Panneerselvam, R.R. Nair, G. Vijayalakshmi, E.H. Subramanian, and S.K. Sridhar, *Synthesis of Schiff bases of 4-(4-aminophenyl)-morpholine as potential antimicrobial agents*, Eur. J. Med. Chem. 40 (2005), pp. 225–229.

[85] C.A. Quesnelle, P. Gill, M. Dodier, and D. St Laurent, M. Serrano-Wu, A. Marinier, A. Martel, C.E. Mazzucco, T.M. Stickle, J.F. Barrett, D.M. Vyas, and B.N. Balasubramanian, *Sordaricin antifungal agents*, Bioorg. Med. Chem. Lett. 13 (2003), pp. 519–524.

[86] C.K. Ryu, H.Y. Kang, Y.J. Yi, K.H. Shin, and B.H. Lee, *Synthesis and antifungal activities of 5/6-arylamino-4,7-dioxobenzothiazoles*, Bioorg. Med. Chem. Lett. 10 (2000), pp. 1589–1591.

[87] D.J. Sheehan, C.A. Hitchcock, and C.M. Sibley, *Current and emerging azole antifungal agents*, Clin. Microbiol. Rev. 12 (1999), pp. 40–79.

[88] C.E. Stephens, F. Tanious, S. Kim, W.D. Wilson, W.A. Schell, J.R. Perfect, S.G. Franzblau, and D.W. Boykin, *Diguanidino and 'reversed' diamidino 2,5-diarylfurans as antimicrobial agents*, J. Med. Chem. 44 (2001), pp. 1741–1748.

[89] J.S. Tkacz and B. DiDomenico, *Antifungals: What's in the pipeline*, Curr. Opin. Microbiol. 4 (2001), pp. 540–545.

[90] N. Tsuchimori, R. Hayashi, N. Kitamoto, K. Asai, T. Kitazaki, Y. Iizawa, K. Itoh, and K. Okonogi, *In vitro and in vivo antifungal activities of TAK-456, a novel oral triazole with a broad antifungal spectrum*, Antimicrob. Agents Chemother. 46 (2002), pp. 1388–1393.

[91] J.M. Urbina, J.C. Cortes, A. Palma, S.N. Lopez, S.A. Zacchino, R.D. Enriz, J.C. Ribas, and V.V. Kouznetzov, *Inhibitors of the fungal cell wall. Synthesis of 4-aryl-4-N-arylamine-1-butenes and related compounds with inhibitory activities on beta(1–3) glucan and chitin synthases*, Bioorg. Med. Chem. 8 (2000), pp. 691–698.

[92] M.L. Vargas, M.V. Castelli, V.V. Kouznetsov, G.J. Urbina, S.N. Lopez, M. Sortino, R.D. Enriz, J.C. Ribas, and S. Zacchino, *In vitro antifungal activity of new series of homoallylamines and related compounds with inhibitory properties of the synthesis of fungal cell wall polymers*, Bioorg. Med. Chem. 11 (2003), pp. 1531–1550.

[93] I. Yalcin, I. Oren, O. Temiz, and E.A. Sener, *QSARs of some novel isosteric heterocyclics with antifungal activity*, Acta Biochim. Pol. 47 (2000), pp. 481–486.

[94] A. Yotsuji, K. Shimizu, H. Araki, K. Fujimaki, N. Nishida, R. Hori, N. Annen, S. Yamamoto, H. Hayakawa, H. Imaizumi, Y. Watanbe, and H. Narita, *T-8581, a new orally and parenterally active triazole antifungal agent: In vitro and in vivo evaluations*, Antimicrob. Agents Chemother. 41 (1997), pp. 30–34.

[95] Y.Z. Zhang, X. Sun, D.J. Zeckner, R.K. Sachs, W.L. Current, and S.H. Chen, *8-Amido-Bearing pseudomycin B (PSB) analogue: Novel antifungal agents*, Bioorg. Med. Chem. Lett. 11 (2001), pp. 123–126.

[96] S.Y. Ablordeppey, P. Fan, S. Li, A.M. Clark, and C.D. Hufford, *Substituted indoloquinolines as new antifungal agents*, Bioorg. Med. Chem. 10 (2002), pp. 1337–1346.

[97] A.K. Saha, L. Liu, R.L. Simoneaux, M.J. Kukla, P. Marichal, and F. Odds, *Novel antifungals based on 4-substituted imidazole: a combinatorial chemistry approach to lead discovery and optimization*, Bioorg. Med. Chem. Lett. 10 (2000), pp. 2175–2178.

[98] D. Barrett, A. Tanaka, K. Harada, H. Ohki, E. Watabe, K. Maki, and F. Ikeda, *Synthesis and biological activity of novel macrocyclic antifungals: Acylated conjugates of the ornithine moiety of the lipopeptidolactone FR901469*, Bioorg. Med. Chem. Lett. 11 (2001), pp. 479–482.

[99] A. Bryskier, *Novelties in the field of anti-infectives in 1997*, Clin. Infect. Dis. 27(1998), pp. 865–883.

[100] A. Martinez, S. Ferrer, I. Santos, E. Jimenez, J. Sparrowe, J. Regadera, F.G. De Las Heras, and D. Gargallo-Viola, *Antifungal activities of two new azasordarins, GW471552 and GW471558, in experimental models of oral and vulvovaginal candidiasis in immunosuppressed rats*, Antimicrob. Agents Chemother. 45 (2001), pp. 3304–3309.

[101] M.H. Serrano-Wu and D.R. St Laurent, C.E. Mazzucco, T.M. Stickle, J.F. Barrett, D.M. Vyas, and B.N. Balasubramanian, *Oxime derivatives of sordaricin as potent antifungal agents*, Bioorg. Med. Chem. Lett. 12 (2002), pp. 943–946.
[102] G. Turan-Zitouni, Z.A. Kaplancikli, M.T. Yildiz, P. Chevallet, and D. Kaya, *Synthesis and antimicrobial activity of 4-phenyl/cyclohexyl-5-(1-phenoxyethyl)-3-[N-(2-thiazolyl)acetamido]thio-4H -1,2,4-triazole derivatives*, Eur. J. Med. Chem. 40 (2005), pp. 607–613.
[103] Y.M. Na, M. Le Borgne, F. Pagniez, G. Le Baut, and P. Le Pape, *Synthesis and antifungal activity of new 1-halogenobenzyl-3-imidazolylmethylindole derivatives*, Eur. J. Med. Chem. 38 (2003), pp. 75–87.
[104] *The Merck Index*, Chapman & Hall, New York, 1996.
[105] J. Ruiz-Herrera and G. San-Blas, *Chitin synthesis as target for antifungal drugs*, Curr. Drug Targets Infect. Disord. 3 (2003), pp. 77–91.
[106] L. Pauling, *The Nature of Chemical Bond*, Cornell University Press, Ithaca, NY, 1939.
[107] STATISTICA, version. 6.0, Statsoft.
[108] P. Baldi, S. Brunak, Y. Chauvin, C.A. Andersen, and H. Nielsen, *Assessing the accuracy of prediction algorithms for classification: An overview*, Bioinformatics 16 (2000), pp. 412–424.
[109] M. Randić, *Resolution of ambiguities in structure-property studies by use of orthogonal descriptors*, J. Chem. Inf. Comput. Sci. 31 (1991), pp. 311–320.
[110] M. Randić, *Orthogonal molecular descriptors*, New J Chem 15 (1991), pp. 517–525.
[111] M. Randić, *Correlation of enthalpy of octanes with orthogonal connectivity indices*, J. Mol. Struct. (Theochem) 233 (1991), pp. 45–59.
[112] E. Estrada and E. Uriarte, *Recent advances on the role of topological indices in drug discovery research*, Curr. Med. Chem. 8 (2001), pp. 1573–1588.
[113] E. Estrada, S. Vilar, E. Uriarte, and Y. Gutierrez, *In silico studies toward the discovery of new anti-HIV nucleoside compounds with the use of TOPS-MODE and 2D/3D connectivity indices. 1. Pyrimidyl derivatives*, J. Chem. Inf. Comput. Sci. 42 (2002), pp. 1194–1203.
[114] S. Wold and L. Erikson, *Chemometric methods in molecular design*, in *Chemometric Methods in Molecular Design*, H. van de Waterbeemd, ed., VCH Publishers, Weinheim, 1995, pp. 309–318.
[115] A. Golbraikh and A. Tropsha, *Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection*, Mol. Divers. 5 (2002), pp. 231–243.
[116] J. Loeffler and D.A. Stevens, *Antifungal drug resistance*, Clin. Infect. Dis. 36 (2003), pp. S31–S41.
[117] J.D. Nosanchuk, *Current status and future of antifungal therapy for systemic mycoses*, Recent Patents Anti-Infective Drug Discov. 1 (2006), pp. 75–84.