October-December 2005 Volume 1 Issue 1-2

Trade Science Inc.

Medicinal CHEMISTRY
An Indian Journal

Full Paper

MCAIJ, 1(1-2), 2005 [1-6]

Relations Between Molecular Structure Of Terpenoids And Its Antimycotic Activity Against *Candida Albicans*

Eduardo Castro
INIFTA, Departamento de Química,
Facultad de Ciencias Exactas, UNLP,
Suc.4, C.C. 16, La Plata 1900 (ARGENTINA)
Tel.: +54 221 424 7430; Fax: +54 221 4244642
E-mail: castro@quimica.unlp.edu.ar
jubert@arnet.com.ar

Fabiana Maguna, María Nuñez, Nora Okulik
Facultad de Agroindustrias, UNNE,
Sáenz Peña 3700, Chaco (ARGENTINA)

Received: 25th June, 2005
Accepted: 4th August, 2005
Web Publication Date: 31st August, 2005

ABSTRACT

To analyze a series of terpenoids in order to select a set of molecular descriptors which are capable to predict suitably the antimycotic activity against *Candida albicans*. The multiple linear regressions were carried out and the best model was selected according to square correlation coefficient (r^2) and the standard deviation(s). We found a predictive model statistically significant which is built with five descriptors and with $r^2 = 0.97$ and $s = 0.40$. The cross-validation method and the Jacknife test were used for validation of the model. The model presented in this work seems to be statistically significant, with a rather good capacity to predict the antimycotic activity against *Candida albicans*.

KEYWORDS

Terpenoids;
QSAR;
Antimycotic activity;
Candida albicans.

INTRODUCTION

Classically, the design of new drugs has been based on the search of some first compound opposing leaders fortuitously or by means of a random screening of natural and synthetic products. At the present time the rational design of new drugs is the method most frequently used to look for such compounds. A first approach toward a rational design comes from the ideas of the Quantitative Structure/Activity and Structure/Property Relationship (QSAR /QSPR) theory[1], which are unquestionably of great importance in modern chemistry and biochemistry. The concept of QSAR/QSPR is to transform searches for compounds with desired properties using chemical intuition and experience into a mathematically quantified and computerized procedure. Once a correlation between structure and activity/property is found, any number of compounds, including those not yet synthesized, can be readily

Full Paper

screened on the computer in order to select structures with desired properties. Thus, the QSAR/QSPR approach conserves resources and accelerates the process of development of new molecules for use as drugs, materials, additives, or for any other purpose[2].

A variety of QSPR/QSAR models have been studied using various parameters including several well known physicochemical properties and other molecular descriptors such as geometric, electronic or electrostatic, polar, steric, and graph-theoretical topological indices. Among these descriptors, the recently developed three-dimensional (3D) descriptors are particularly interesting because they take into account the geometric conformation and the nature of the bonding of groups within a molecule and they compare quite well with the two-dimensional (2D) molecular descriptors.

In Chemical Graph Theory, molecular structures are normally represented as hydrogen-suppressed graphs, whose vertex and edges act as atoms and covalent bonds, respectively. Graph-theoretical indices, also known as topological indices, are descriptors that characterize a molecular graph and they are capable to give account of their structural properties in order to obtain the relations used in discrimination and prediction studies. They have shown their usefulness in classification analysis, and in general, in the modeling of biological activities.

The conventional 2D and 3D topological indices characterize a molecule as a whole, that is, molecular size or shape, such as molecular connectivity index (χ), Hosoya′s index (Z), Balaban′s index(J), Schulz′s index (MTI), etc.[1].

One of the yeast of more clinical importance as model microorganism for the study of the mycotic infections is *Candida albicans*. This microorganism presents like one of the essential and dynamic structures the cellular wall that is also responsible for the form and of the cellular integrity. As this cellular wall is absent in the animal cell, this organ can be a good leader for the development of new antifungal agents[3].

It is also known that the organism defends in natural form of the infections caused by yeast, but this aggression type can originate a serious problem when associating to other situations, as the immunodeficiency, for example. Although there is a wide collection of antimicrobial agents for the treatment of bacterial infections, this is reduced and it is of limited effectiveness when they are antifungal agents. This situation has stimulated the search of antimicrobial agents in natural sources different from the habitual ones going the attention to superior organisms as the vegetables[4]. This situation has stimulated the search of antimicrobial agents in natural sources different from the habitual (just as microorganism) doing focuses upon the vegetables.

At present it is well known the antimicrobial activity of essential oils, such as Menta piperita, Orégano sp., Salvia fructicosa, etc.[5-7]. The type of antimicrobial activity shown by essential oils varies from partial or complete inhibition of growth to bactericidal or fungicidal activity.

According to several studies[8-10] the antimicrobial activity presented by the essential oils is due, in great measure, to the presence of a type of denominated compounds 'terpenoids'. Recent investigations have shown that the site of action of terpenoides is at the cell membrane and the proposed action mechanisms change in agreement with the functional group of the terpenoid molecules. Nevertheless, the effectiveness of the compounds depends strongly of their water solubility and of their ability to penetrate that cellular membrane[11].

Our research team is working in the study of some series of compounds to define groups of molecular/topological indices that allow us to predict antimycotic activities against *Candida* albicans in a rather acceptable form. Particularly, the interest has been centered in indices that are capable to describe physico-chemical properties such as the octanol/water partition coefficient (Log P) and hydrophilic factor, molecular properties as molecular size and total absolute charge, and the 3D Balaban´s index. This work deals with the correlation between the previously mentioned descriptors and the biological activity of interest, centered in the antimycotic activity of a group of terpenoids molecules.

MATERIALS AND METHODS

The antimicrobial activity against *Candida albicans* of a group of 32 terpenoids group was taken from the bibliography[11]. The values reported as minimum inhibitory concentrations, MIC, in ppm were expressed as mg/ml, in TABLE 1 (available in SUPPLEMENTARY INFORMATION) are presented the studied compounds and the corresponding MIC.

All structures were constructed using the HyperChem 6.03[12]. structure format and were saved as .hin files. Initially, physiochemical, topological, constitutionals and based charge descriptors were calculated using Dragon software free version 5.0[13].

Using the set of 32 terpenoids, multiple linear regression models were developed with the Statgraphics Plus package 4.0[14]. The quality of the model was considered as statistically satisfactory on the basis of squared correlation coefficient (r^2), standard deviation (s) and F-statistics (F). All the parameters in the models were significant at a 90% confidence level.

Once selected the descriptors that presented the highest predictive power, the prediction model was built with the 32 compounds. The molecular descriptors selected were the octanol/water partition coefficient (Log P), information index on molecular size (ISIZ), hydrophilic factor (Hy), total absolute charge index (Qtot), and the 3D Balaban´s index (J3D). These descriptors were analyzed by a correlation matrix to verify that they did not present collinearity among them.

The predictive ability in the adjusted model was carried out cross validation using the method of LOO (leave-one-out), through the defined values for r^2_{cv} and s_{cv}, according to the equations 1 and 2.

Eq. 1

$$r^2_{cv} = 1.0 - \frac{\sum\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}$$

Eq. 2

$$s_{cv} = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{N - M - 1}}$$

Where

y_i = experimental value
$\hat{y}_i$ = predicted value
$\bar{y}$ = mean value Yi
N = number of samples used for model
M = number of describers

The predictive capability in the training set was carried out using the jackknife r^2 (r^2_j) values. For any given compound, C_j, its corresponding r^2_j values can be determined by deleting this compound from the regression analysis and computing the resulting squared correlation coefficient, r^2, from the original model using (n-1) data points. The unduely high r^2_j values might indicate outliers and/or biases, and those with low r^2_j values might be considered the influential points in the data set, respectively (1).

RESULTS AND DISCUSSION

Initially we have calculated 44 indices. Then, according to the methodology described before we selected the octanol/water partition coefficient (Log P), information index on molecular size (ISIZ), hydrophilic factor (Hy), total absolute charge index (Qtot) and the 3D Balaban's index (J3D) because they had the highest power in the prediction of the antimycotic activity against *Candida albicans*. Their values are presented in TABLE 1 (available in SUPPLEMENTARY INFORMATION).

The predictive model with their respective coefficients and statistical parameters are presented in the equation (3) for the 32 terpenoids:

Eq. 3

MIC = -6.3632 - 0.0023* Log P + 0.0598* ISIZ- 1.6832*Hy + 2.0543*Qtot - 0.4912*J3D

n= 32; r^2 = 0.9096; s= 0.6676; F= 52.33

The collinearity analysis among the descriptors was carried out by the correlation matrix, where the pairwise correlations were examined using the correlation coefficients, r^2, and we excluded those descriptors having values greater than 0.55. Data are presented in TABLE 2 (available in SUPPLEMENTARY INFORMATION).

Then, the jackknife test was applied to the data group with the purpose of verifying the presence of biases. The jackknife results are presented in TABLE 3 (available in SUPPLEMENTARY INFORMATION), indicating eight data points higher than the limits of the mean r^2 values (0.909 $\pm$ 0.0056). After excluding these data points **(compounds 20, 24, 25,**

Full Paper

Eq. 4

$$MIC = -7.1053 - 0.0021*Log\ P + 0.0618*ISIZ - 1.5792*Hy + 2.1766*Qtot - 0.5343*J3D$$

$$n = 27;\ r^2 = 0.9703;\ s = 0.4092,\ F = 137.45$$

29, and 32), the following equation(4) was obtained:

The values of calculated MIC in equation 4 with the adjusted model and residuals ones are presented in TABLE 3 (available in SUPPLEMENTARY INFORMATION).

One of the descriptors selected for the model is the log P because offers information on the balance of the solubility of a compound between the water and the oil. This property is directly related with the solubility of the terpenoides, which is a fundamental characteristic to cross in the lipidic bilayer of cellular wall of the microorganism. However the effectiveness also depends on its ability to penetrate the cellular wall. In the present model the describer Log P shows a negative coefficient indicating an inverse relationship with the activity in question.

The information index on molecular size (ISIZ) is other descriptor used in the model. This descriptor is associated with the capability of the compound to cross the cellular wall of the microorganisms[11]. In the adjusted model it presents a directly proportional relationship with the activity against *Candida albicans* and it is demonstrated by its positive coefficient in the model.

We have also used the index Balaban's (J), one of the most representative topological indices, which has been interpreted as an index of the degree of folding of a molecule[15]. This index is also known as distance index described by a topological distance matrix and it is used to characterize the constitution and the configuration of a molecule by a single number[16]. In our study we have incorporated this descriptor because the terpenoides group present isomer compounds and was necessary to incorporate a 3D describer being capable to discriminate among these compounds. In the correlation it is observed an inverse relationship with the biological activity.

Electrical charges in the molecule are obviously the driving force of electrostatic interactions and they are important in many chemical reactions and physico-chemical properties of the compounds[2]. Thus, charge based descriptors have been widely employed as chemical reactivity indices or as measures of weak intermolecular interactions. We have used the total absolute charge index, Qtot (electronic charge index) and we find that is related in a direct way with the activity under consideration.

Next, we have carried out the validation of the adjusted model (equation 4) by the method of LOO and the obtained values are the following:

$r^2_{cv} = 0.9821$ $\qquad$ $s_{cv} = 0.4176$

Figure 1 shows the experimental MIC values versus the calculated MIC. It can be observed the predictive ability of the adjusted model. Our results show that the model predicts the antimycotic activity of the studied group of 27 terpenoids at a 97% level.

In Figure 2 it can be observed the dispersion of

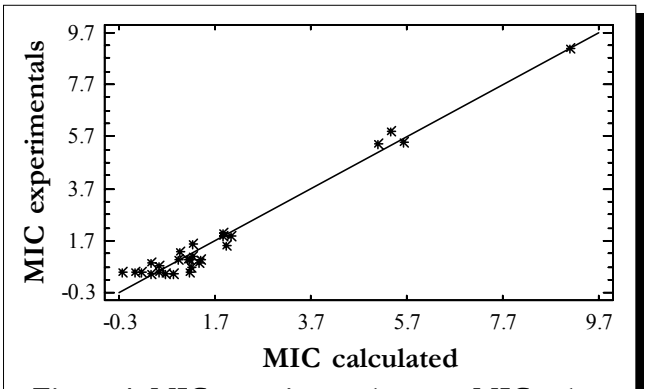

Figure 1: MIC experimental versus MIC calculated for the 27 terpenoids

the residual regarding the values of MIC calculated for the 27 terpenoids. When the correlation is high the residuals should form a "cloud" homogeneous

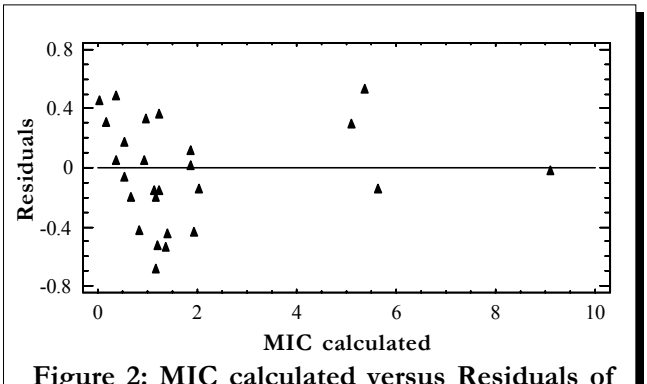

Figure 2: MIC calculated versus Residuals of 27 terpenoids studied

around the central line. In this case there are only two points located quite far from the central line, in agreement with the correlation statistically significant opposing.

CONCLUSIONS

The QSAR equations is a linear model which relates variations in biological activity with regard to values of computed (or measured) properties for a series of molecules. To obtain a significant correlation, it is crucial that appropriate descriptors should be employed, and they may be theoretical, empirical, or derived from readily available experimental characteristics of the structures.

The selection of descriptors is an important first step in any QSAR study because only if the relationship between the selected descriptors and activity is significative the activity predictions will be possible.

One of the properties that can provide a valuable insight in the antimycotic activity is the grade of solubility of the terpenoids in the lipidic bilayer of the cellular wall because many membranes must be crossed for compounds to get to the target site. Those compounds with greatest hydrophobicity will become localized in the membranes they encounter initially. For this reason we have used as descriptor the octanol/water partition coefficient (log P), which shows an inverse relationship with the antimycotic activity.

Another index which has demonstrated to be valuable is the information index on molecular size (ISIZ), which shows a directly proportional relationship with the activity against *Candida albicans*.

Two representative real number topological indices are the connectivity index and Balaban's *J* index, interpreted as an index of the degree of folding of a molecule. In our work they can be considered to be very important because terpenoid groups present isomer compounds, so that it was necessary to incorporate a 3D descriptor being capable to discriminate among these compounds.

It has been proven that local electron densities or charges are important in many chemical reactions and physicochemical properties of compounds. From this point of view, we have found significant to employ a total absolute charge descriptor, such as Qtot.

The model presented in this work, built with those descriptors discussed before, demonstrates to be statistically significant and it has a capacity to predict the antimycotic activity against *Candida albicans* in a relatively high degree.

The results of the present study gives new evidence of the importance of selecting descriptors that are capable to give a suitable basic explanation of the phenomenon under study.

SUPPLEMENTARY INFORMATION AVAILABLE STATEMENT

Following information is available as SUPPLEMENTARY INFORMATION.

- TABLE 1: MIC values and molecular descriptors used
- TABLE 2: Correlation matrix
- TABLE 3: MIC and calculated residuals with two models

ACKNOWLEDGMENTS

The authors thanks SECYT-UNNE for financial support (PI 671). F.M. thanks SECYT-UNNE for a scholarship.

REFERENCES

[1] M.Núñez, F.Maguna, N.Okulik, E.Castro; Bioorganic and Medicinal Chemistry Letters, **14**, 5611–5617 **(2004)**.

[2] M.Karelson, V.Lobanov, A.Katritzky; Chem.Rev., **96**, 1027-1043 **(1996)**.

[3] M.Gil García; Estudio del proteoma de la pared celular de *Candida albicans*. Aplicaciones al estudio de dianas de antifúngicos y al desarrollo de estrategias de diagnóstico e inmunoprotección; published in http://www.ucm.es/info/gyp/proteomica/nov01juandecierva.htm; 5/10/04.

[4] E.Ferro, N.Canela de Alvarenga; Revista de Ciencia y Tecnología Dirección de Investigaciones – UNA; 1 N° 2 **(2000)**.

[5] G.Iscan, N.Kirimer, M.Kurkcuoglu, K.Husnu Can Baser, F.Demirci; J.Agric.Food Chem., **50**, 3943-3946 **(2002)**.

[6] A.Sivropoulou, E.Papanikolaou, C.Nikolaou, S. Kokkini, T.Lanaras, M.Arsenakis; J.Agric.Food Chem., **44**, 1202-1205 **(1996)**.

Full Paper

[7] A.Sivropoulou, C.Nikolaou, E.Papanikolaou, S. Kokkini, T.Lanaras, M.Arsenakis; J.Agric.Food Chem., **45**, 3197-3201 **(1997)**.
[8] P.Suppakul, J.Miltz, K.Sonneveld, S.Bigger; J.Agric. Food Chem., **51**, 3197-3207 **(2003)**.
[9] J.Sikkema, J.DeBont, B.Poolman; Microbiological Rev., **59**, 201-222 **(1995)**.
[10] I.Helander, H.Alakomi, K.Lavta-Kala, T.Mattila-Sandholm, I.Pol, E.Smid, L.Gorris, A.Von Wright; J. Agric.Food Chem., **46**, 3590-3595 **(1998)**.
[11] S.Griffin; Physic Bulletin, **30**, 262 **(1979)**.
[12] HyperChem.Version 6.03, Hypercube, Inc., http://www.hyper.com **(2001)**.
[13] Dragon Evaluation version, Software version 5.0, by R.Todeschini, V.Consonni, A.Mauri, M.Pavan; Milano Chemometrics and QSAR Research Group. Website: http://www.talete.mi.it **(2004)**.
[14] StatGraphics Plus. Versión 4.0, Statistical Graphics Corp., **(1999)**.
[15] M.Randic, M.Pompe; J.Chem.Inf.Comput.Sci., **41**, 575-581 **(2001)**.
[16] C.Mihali, S.Nikoli, N.Trinajsti; J.Chem.Inf.Comput. Sci., **32**, 28-37 **(1992)**.