European Journal of Medicinal Chemistry 45 (2010) 1724–1730

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

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

Original article

QSAR study of flavonoids and biflavonoids as influenza H1N1 virus neuraminidase inhibitors

Andrew G. Mercader*, Alicia B. Pomilio

PRALIB (UBA-CONICET), Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Junín 956, C1113AAD Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 16 October 2009
Received in revised form
30 December 2009
Accepted 5 January 2010
Available online 14 January 2010

Keywords:
QSAR
Flavonoids
Biflavonoids
Influenza H1N1 virus neuraminidase inhibition
Enhanced replacement method

ABSTRACT

We performed a predictive analysis based on Quantitative Structure–Activity Relationships (QSAR) of a very important property of flavonoids which is the inhibition (IC_{50}) of influenza H1N1 virus neuraminidase. The best linear model constructed from 20 molecular structures incorporated four molecular descriptors, selected from more than a thousand geometrical, topological, quantum-mechanical and electronic types of descriptors.

The obtained model suggests that the activity depends on the electric charges, masses and polarizabilities of the atoms present in the molecule as well as its conformation. The model showed good predictive ability established by the theoretical and external test set validations.

1. Introduction

Influenza virus expresses two envelope glycoproteins: hemagglutinin and neuraminidase (NA) [EC 3.2.1.18]. The hemagglutinin is known to mediate the binding of viruses to target cells *via* sialic acid residue in glycoconjugates. This binding is a key step of the viral infection [1]. The NA facilitates the movement of the virus to and from sites of infection in the respiratory tract by taking charge of catalyzing the cleavage of neuraminic acid residues [1–3]. Therefore, the inhibition of influenza virus neuraminidase has the possibility of blocking an influenza virus infection. Because of the importance of this enzyme in the pathogenesis of influenza virus infection and the close correspondence of the conserved residues of the active sites from NAs of all influenza A viruses, the enzyme has been regarded as a drug target for the treatment of influenza [4].

In recent years, several flavonoids and biflavonoids have been reported as showing anti-influenza virus activity by inhibiting NAs [3,5–8].

Flavonoids (substituted phenyl–benzopyranes) are low molecular weight compounds that are widespread in the plant kingdom, are relatively easy to synthesize and show several interesting biological activities in enzymatic systems.

Clearly it is of great interest to be able to predict the NA inhibition by compounds that have no experimental values yet, as well as attempting to determine the structural parameters that the NA inhibition depends on. A generally accepted approach for overcoming the lack of experimental data in complex chemical phenomena is the analysis based on Quantitative Structure–Activity Relationships (QSAR) [9].

In the present study, we investigated a QSAR model for the inhibition of NA enzyme by flavonoids that could serve as a guide for the rational design of further potent and selective inhibitors of this family of compounds; no such a study has been found in the literature.

A great number of structural molecular descriptors including definitions of all classes was explored using the recently proposed Enhanced Replacement Method (ERM) [10] to select the best subset of variables.

2. Methods

2.1. Data set

In the present study we used a training set of 20 and a test set of 5 flavonoids and biflavonoid derivatives of different classes for which their activities were reported in the literature [3,5–8].

The experimental influenza virus (H1N1) NA inhibitory activity of the data set was measured using a standard fluorimetric assay [11]. The 50% inhibitory concentration (IC_{50}) is defined as the

* Corresponding author. Tel.: +54 11 6091 3759; fax: +54 11 6091 2100x3759.
E-mail addresses: andrewgmercader@gmail.com (A.G. Mercader), pomilio@ffyb.uba.ar (A.B. Pomilio).

0223-5234/$ – see front matter © 2010 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmech.2010.01.005

concentration (μg/mL) of NA inhibitor necessary to reduce NA activity by 50% relative to a reaction mixture containing virus but no inhibitor. Table 1 and Fig. 1 summarize the molecular structures, name of the substance, experimental IC_{50} and log IC_{50} of the above mentioned flavonoid derivatives. The IC_{50} of Oseltamivir and Zanamivir were included at the end of Table 1 as reference substances, both of them exhibit a very high NA inhibitory activity, nevertheless they are synthetic compounds that may present many side effects in contrast to flavonoids which as known are natural and innocuous.

2.2. Molecular descriptors

The structures of the compounds were firstly pre-optimized with the Molecular Mechanics Force Field (MM+) procedure included in the Hyperchem 6.03 package [12], and the resulting geometries were further refined by means of the semiempirical method PM3 (Parametric Method-3) using the Polak–Ribière algorithm and a gradient norm limit of 0.01 kcal $Å^{-1}$. The molecular descriptors were computed using the software Dragon 5.0 [13], parameters of all types were calculated such as Constitutional, Topological, Geometrical, Charge, GETAWAY (Geometry, Topology and Atoms-Weighted AssemblY), WHIM (Weighted Holistic Invariant Molecular descriptors), 3D-MoRSE (3D-Molecular Representation of Structure based on Electron diffraction), Molecular Walk Counts, BCUT descriptors, 2D-Autocorrelations, Aromaticity Indices, Randic Molecular Profiles, Radial Distribution Functions, Functional Groups and Atom-Centred Fragments [14]. In addition 4 quantum-chemical descriptors (molecular dipole moments, total energies, homo–lumo energies) not provided by the program Dragon were added to the descriptors pool. The resulting total pool thus consisted of $D = 1186$ descriptors.

2.3. Model search

It is our purpose to search the set **D**, containing D descriptors, for an optimal subset **d** of $d \ll D$ ones with minimum standard deviation S,

$$S = \frac{1}{(N-d-1)}\sum_{i=1}^{N} \mathrm{res}_i^2 \quad (1)$$

by means of the Multivariable Linear Regression (MLR) technique. In this equation N is the number of molecules in the training set, and res_i the residual for molecule i, the difference between the experimental property (**p**) and predicted property ($\mathbf{p_{pred}}$). More precisely, we want to obtain the global minimum of $S(\mathbf{d})$ where **d** is a point in a space of $D!/[d!(D-d)!]$ ones. A full search (FS) of optimal variables is impractical because it requires $D!/[d!(D-d)!]$ linear regressions. Hence an alternative method is necessary, we have used to select the optimum set of descriptors the Enhanced Replacement Method (ERM) [10] as a search algorithm that produces linear regression QSPR–QSAR models that are quite close the FS ones with much less computational work. This technique approaches the minimum of S by judiciously taking into account the relative errors of the coefficients of the least-squares model given by a set of d descriptors $\mathbf{d} = \{X_1, X_2, \ldots, X_d\}$. The ERM gives models with better statistical parameters than the Forward Stepwise Regression procedure [15] and the more elaborated Genetic Algorithms [16].

The Kubinyi function (FIT) is a statistical parameter that closely relates to the Fisher ratio (F), but avoids the main disadvantage of the latter that is too sensitive to changes in small d values, and poorly sensitive to changes in large d values. The FIT(**d**) criterion has a low sensitivity to changes in small d values and a substantially increasing sensitivity for large d values. The greater the FIT value the better the linear equation; it is given by

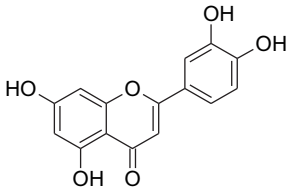
$$\mathrm{FIT} = \frac{R(d)^2(N-d-1)}{\left(N+d^2\right)\left(1-R(d)^2\right)} \quad (2)$$

where $R(d)$ is the correlation coefficient for a model with d descriptors.

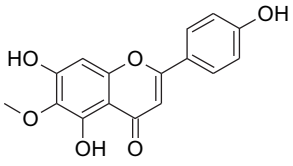
Table 1
Experimental IC_{50} (μg/mL), experimental log IC_{50}, predicted (Eq. (3)) log IC_{50} and residual.

Number	Class	Compound name or chemical name	IC_{50} Exp. (μg/mL)	log IC_{50} Exp.	log IC_{50} Pred.	Residual
Training set						
1	Flavones	Luteolin	9.65	0.984	0.991	−0.007
2		Hispidulin (Dinatin)	13.90	1.143	1.189	−0.046
3		Scutellarein	14.48	1.161	1.275	−0.114
4		Galuteolin (luteolin 7-*O*-β-D-glucopyranoside)	21.25	1.328	1.277	0.051
5		Kaempferol 3-O-β-xylopyranosyl-(1 → 2)-β-D-glucopyranoside	24.50	1.389	1.427	−0.038
6		5,7,4′-Trihydroxy-8-methoxyflavone	13.15	1.119	1.199	−0.080
7	Flavonols	Kaempferol	16.77	1.225	1.174	0.050
8		Quercetin	17.65	1.247	1.250	−0.004
9		Myricetin	26.29	1.420	1.327	0.093
10		Rutin	31.14	1.493	1.462	0.031
11	Isoflavones	Daidzein	9.43	0.975	0.966	0.009
12		Genistein	20.84	1.319	1.230	0.089
13	Aurones	Sulfuretin [2-(*Z*-3′,4′-Dihydroxyphenylidene)-6-hydroxy-2, 3-dihydrobenzofuran-3-one]	8.00	0.903	0.884	0.019
14		2-(*E*-4′-Hydroxyphenylidene)-6-hydroxy-2,3-dihydrobenzofuran-3-one	5.59	0.748	0.837	−0.089
15		2-(*E*-Benzylidene)-6-hydroxy-2,3-dihydrobenzofuran-3-one	17.15	1.234	1.168	0.067
16		2-(*E*-4′-Hydroxyphenylidene)-4,6-dihydroxy-2,3-dihydrobenzo- furan-3-one	6.92	0.840	0.844	−0.004
17	Flavans	7-*O*-Galloyltricetinflavan	15.70	1.196	1.184	0.012
18		7,4′-Di-*O*-galloyltricetinflavan	30.00	1.477	1.530	−0.053
19	Biflavonoids	Ginkgetin (biflavone)	55.00	1.740	1.770	−0.030
20		Hinokiflavone-sialic acid	19.10	1.281	1.237	0.044
Test set						
21	Flavones	Apigenin	8.54	0.932	0.926	0.006
22		Vitexin	20.85	1.319	1.333	−0.014
23		Chrysin	11.62	1.065	1.203	−0.138
24	Flavonol	Rhamnocitrin (3,4′,5-trihydroxy-7-methoxyflavone)	15.46	1.189	1.162	0.028
25	Biflavonoid	Hinokiflavone	41.80	1.621	1.840	−0.218

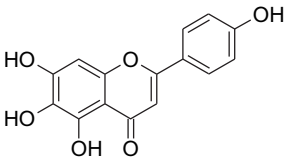
Note: as reference IC_{50} values of synthetic compounds Oseltamivir and Zanamivir are 1.66 nM (6.81×10^{-4} μg/mL) and 1.17 nM (3.89×10^{-4} μg/mL), respectively [28].



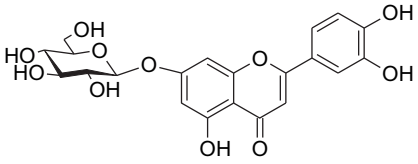
(1) Luteolin



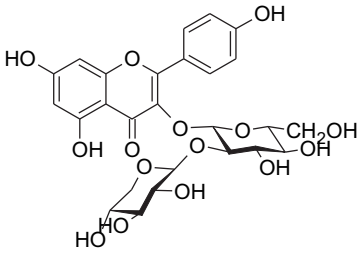
(2) Hispidulin (Dinatin)



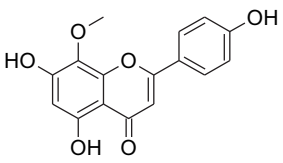
(3) Scutellarein



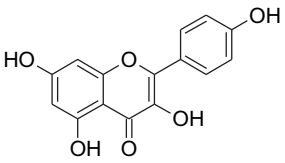
(4) Galuteolin (luteolin 7-*O*-β-*D*-glucopyranoside)



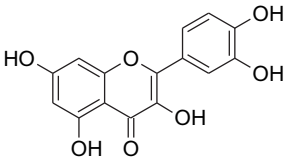
(5) Kaempferol 3-O-β-xylopyranosyl-(1→2)-β-D-glucopyranoside



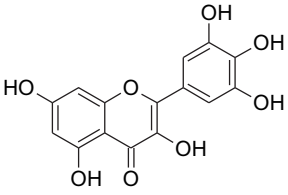
(6) 5,7,4'-Trihydroxy-8-methoxyflavone



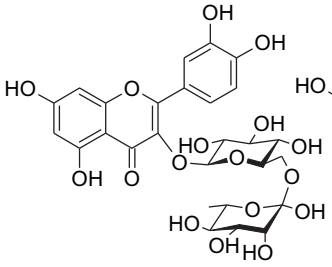
(7) Kaempferol



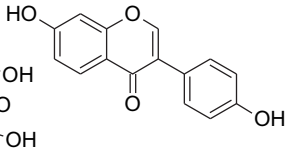
(8) Quercetin



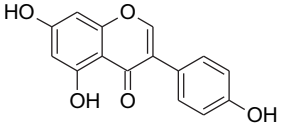
(9) Myricetin



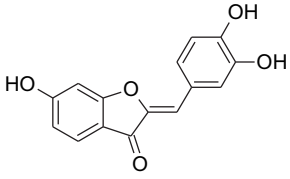
(10) Rutin



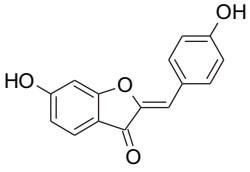
(11) Daidzein



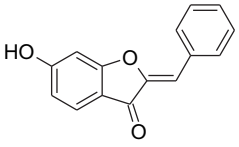
(12) Genistein



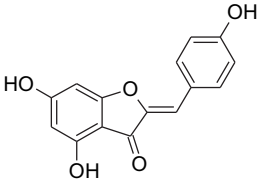
(13) Sulfuretin [2-(*Z*-3',4'-Dihydroxyphenylidene)-6-hydroxy-2,3-dihydrobenzofuran-3-one]



(14) 2-(*E*-4'-Hydroxyphenylidene)-6-hydroxy-2,3-dihydrobenzofuran-3-one

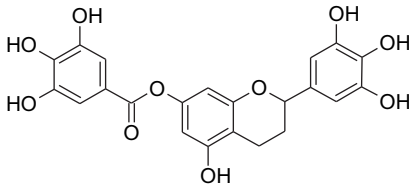


(15) 2-(*E*-Benzylidene)-6-hydroxy-2,3-dihydrobenzofuran-3-one

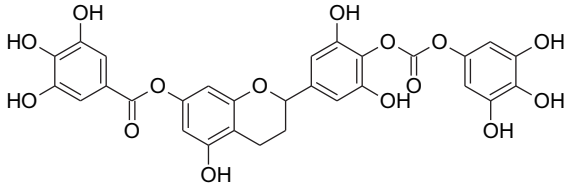


(16) 2-(*E*-4'-Hydroxyphenylidene)-4,6-dihydroxy-2,3-dihydrobenzo- furan-3-one

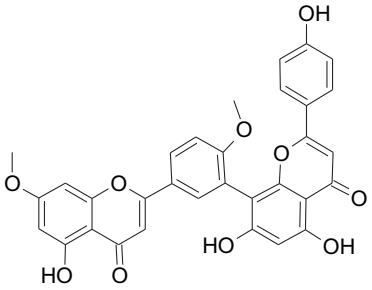
Fig. 1. Molecular structures of the 25 flavonoid derivatives studied.



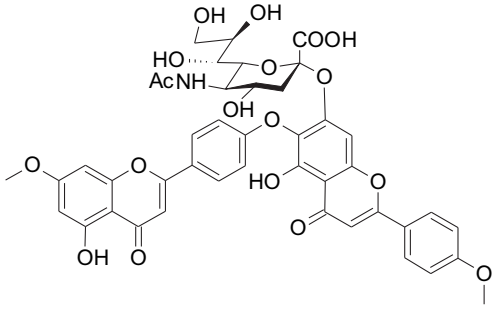
(17) 7-O-Galloyltricetinflavan



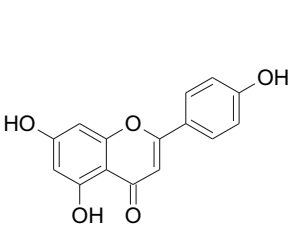
(18) 7,4'-Di-*O*-galloyltricetinflavan



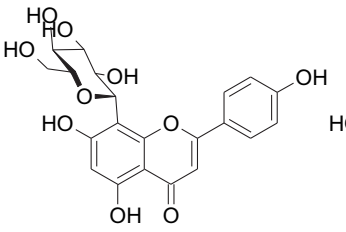
(19) Ginkgetin (biflavone)



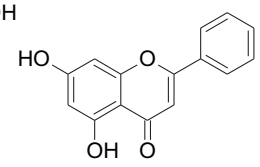
(20) Hinokiflavone-sialic acid



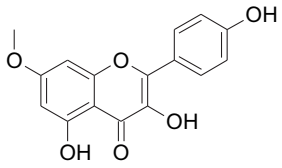
(21) Apigenin



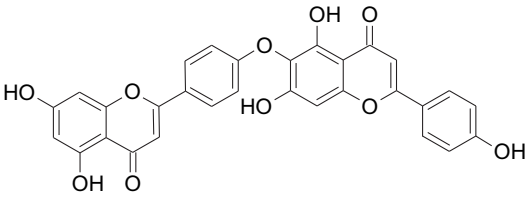
(22) Vitexin



(23) Chrysin



(24) Rhamnocitrin (3,4',5-trihydroxy-7-methoxyflavone)



(25) Hinokiflavone

Fig. 1. (*continued*).

In this work we have used the FIT function in conjunction with the rule of thumb that at least 5 data points should be present for each fitting parameter [17] to set the optimal number of molecular descriptors (d_{op}) in the linear regression equation.

As a theoretical validation of all the models we choose the well-known Leave-One-Out (loo) and the Leave-More-Out Cross-Validation procedures (l-*n*%-o) [18], where *n*% accounts for the number of molecules removed from the training set. We generated 1,000,000 cases of random da removal for l-*n*%-o, where *n*% = 25% (5 molecules).

3. Results and discussion

By means of the ERM we searched the total pool of $D = 1186$ descriptors and obtained optimal models with $d = 1,2,3,4$ parameters linking the molecular structure of the flavonoids compounds

with their inhibitory activity. The statistical parameters arriving from the models with different number of descriptors showed that the best option that complied with the above mentioned rule [17] was $d_{opt} = 4$. Thus finding that the optimal QSAR model according to ERM was:

$$\log IC_{50} = 0.6762(\pm 0.4) - 11.9587(\pm 1)JGI1 + 0.6004(\pm 0.1)\mathrm{AMW} + 0.7815(\pm 0.1)Mor30m - 13.2189(\pm 2)G1p \quad (3)$$

$$N = 20, R = 0.9706, S = 0.0654, \mathrm{FIT} = 6.7718, p < 10^{-5}$$

$$R_{loo} = 0.9501, S_{loo} = 0.0854, R_{l-25\%-o} = 0.7795, S_{l-25\%-o} = 0.1918\ \mathrm{RMSE_{TS}} = 0.1163$$

Here, the absolute errors of the regression coefficients are given in parentheses, p is the significance of the model, and $\mathrm{RMSE_{TS}}$ stands for root mean squared errors of the test set. In our calculations we used the computer system Matlab 5.0 [19].

A summary of the linear models with 1 to d_{opt} parameters calculated by ERM is shown in Table 2. The details of the molecular descriptors of Table 2 are displayed in Table 3.

With the purpose of demonstrating that Eq. (3) does not result from happenstance, we resort to a widely used approach to establish the model robustness: the so-called *y*-randomization [20]. It consists of scrambling the experimental property **p** in such a way that activities do not correspond to the respective compounds. After analyzing 1,000,000 cases of *y*-randomization, the smallest S value obtained in this way $S = 0.0986$ was larger than the one coming from the true calibration ($S = 0.0654$). This result suggests that the model is robust, that the calibration is not a fortuitous correlation, and that we have derived a reliable structure–activity relationship. The prediction from any QSAR model cannot be intrinsically better than the experimental data employed to develop the model [21], because of the shortage of experimental measurements, in the future when more experimental information of these compounds is available another study should be performed to support the present results.

The plot of predicted *vs.* experimental log IC_{50} shown in Fig. 2 suggests that the 20 flavonoid derivatives from the training set and 5 from the test set follow a straight line. The predicted inhibitory potencies given by Eq. (3) for the training and test sets are shown in Table 1. The behavior of the residuals in terms of the predictions illustrated in Fig. 3 shows normal distributions for both sets. The only molecule that presents a residual slightly larger than 2.5S is molecule number **25** (Hinokiflavone, residual 3.3S). It is not possible to determine if such a deviation is either a physical (meaningful) result or a statistical consequence of present selection of descriptors in Eq. (3); however it is probable that the deviation was caused by the low number of biflavonoids available in the training set which led to a shortage of structural information on this class of flavonids during the formation of the model.

The correlation matrix shown in Table 4 reveals that the descriptors of the linear model are not seriously inter-correlated ($R_{ij} < 0.8005$), which justifies the appearance of all parameters in the equation. The predictive power of the linear model is satisfactory as revealed by its stability upon the inclusion and/or exclusion of compounds, measured by the statistical parameters $R_{loo} = 0.9501$ and $l - n\% - o$ $R_{l-25\%-o} = 0.7795$. According to the literature, $R_{l-n\%-o}$ must be higher than 0.71 in order to have a validated model [22].

Table 2
Linear QSAR models for the training set of log IC_{50} ($N = 20$). The best relationship appears in bold.

Model	Descriptors used	R	S	FIT	$\mathrm{RMSE_{TS}}$
M1	*Mor27u*	0.7507	0.1639	1.1070	0.2083
M2	*BIC*1, *RCI*	0.8801	0.1212	2.4350	0.1909
M3	*IC*1, *D/Dr*05, *G1m*	0.9269	0.0988	3.3661	0.2204
M4	AMW, *JGI*1, *Mor*30*m*, *G1p* (Eq. (3))	0.9706	0.0654	6.7718	0.1163

Table 3
Symbols for molecular descriptors involved in different models.

Molecular descriptor	Type	Description
Mor27u	3D-MoRSE	3D-MoRSE - signal 27/unweighted.
*BIC*1	Topological	Bond information content (neighborhood symmetry of 1-order).
RCI	Aromaticity Indices	Jug RC index.
*IC*1	Topological	Information content index (neighborhood symmetry of 1-order).
*D/Dr*05	Topological	Distance/detour ring index of order 5.
G1m	WHIM	1st component symmetry directional WHIM index/weighted by atomic masses.
*JGI*1	Topological	Mean topological charge index of order1.
AMW	Constitutional	Average molecular weight.
*Mor*30*m*	3D-MoRSE	3D-MoRSE - signal 30/weighted by atomic masses.
G1p	WHIM	1st component symmetry directional WHIM index/weighted by atomic polarizabilities.

The molecular descriptors appearing in the linear Eq. (3) merge two- and three-dimensional aspects of the molecular structure, and can be classified as follows: (i) a Topological descriptor: *JGI*1, mean topological charge index of first order; (ii) a Constitutional descriptor: AMW, average molecular weight; (iii) a 3D-MoRSE descriptor: *Mor*30*m*, signal 30 weighted by atomic masses; and (iv) a WHIM descriptor: *G1p*, 1st component symmetry directional WHIM index weighted by atomic polarizabilities.

Topological charge indices were proposed to evaluate the charge transfer between pairs of atoms, and therefore the global charge transfer in the molecule [23,24]. In order to obtain the indices definitions first we need to define the matrix **M** as:

$$\mathbf{M} = \mathbf{A} \cdot \mathbf{D}^{-2} \quad (4)$$

where **A** is the adjacency matrix and $\mathbf{D}^{-2}$ the reciprocal square distance matrix, the diagonal elements of the distance matrix stay unchanged. **M** is the Galvez matrix and is a squared simetric matrix $a \times a$ were a is the number of atoms in the molecule.

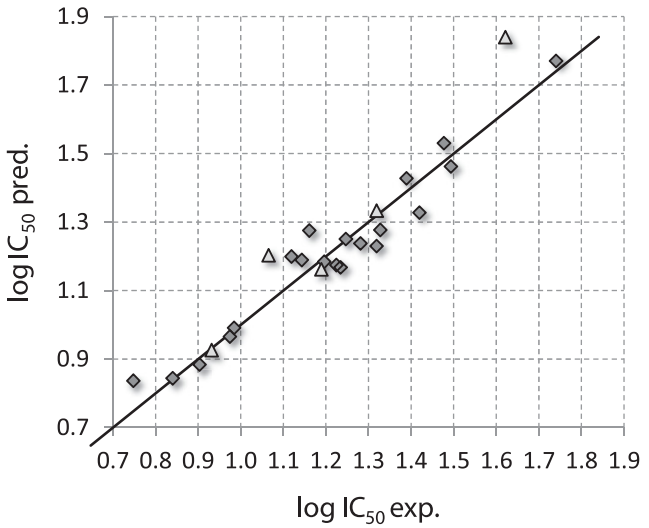


Fig. 2. Predicted (Eq. (3)) *vs.* experimental log IC_{50} for the training (rhombus) and test (triangles) sets.

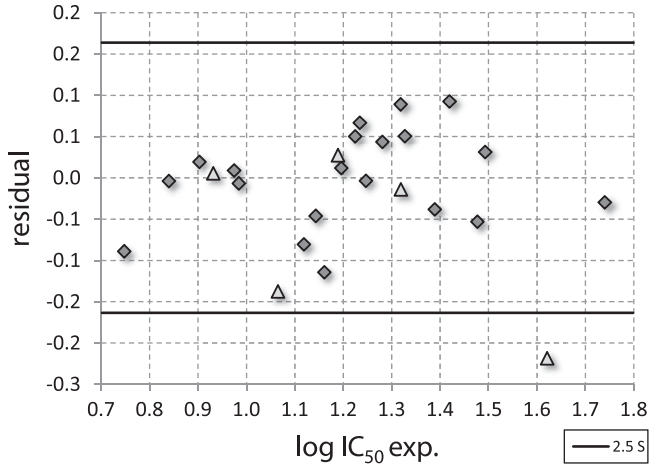


Fig. 3. Dispersion plot of the residuals for the training and test sets according to Eq. (3).

An unsymmetric charge term matrix **CT** is derived from the matrix **M**:

$$CT_{ij} = \begin{cases} \delta_i & si\ i = j \\ m_{ij} - m_{ji} & si\ i \neq j \end{cases} \quad (5)$$

where m_{ij} are the elements of **M** and δ_i is the degree of the vertex of the i^{th} atom. The diagonal entries of the **CT** matrix represent the topological valence of the atoms; the off-diagonal entries CT_{ij} represent a measure of the net charge transferred from the atom i to the atom j.

For each path of length L, a topological charge index **GGIL** is defined as:

$$\mathbf{GGIL} = \sum_{i=1}^{a} \sum_{j=1}^{a} \left| CT_{ij} \right| \cdot \delta(L, d_{ij}) \quad (6)$$

The mean topological charge index of order L (in *JGI*1, $L = 1$), is defined as:

$$\mathbf{JGIL} = \frac{\mathbf{GGIL}}{a - 1} \quad (7)$$

Constitutional descriptors are the most simple and commonly used descriptors, reflecting the molecular composition of a compound independently from molecular connectivity and conformations

The 3D-MoRSE (3D Molecule Representation of Structure based on Electron diffraction) descriptors provide 3D information from the three-dimensional structure of a molecule using a molecular transform derived from an equation used in electron diffraction studies. Several atomic properties can be taken into account, thus giving high flexibility to this representation of a molecule. The simplified form of the transform is:

$$I(s) = \sum_{i=2}^{N} \sum_{j=1}^{i-1} A_i A_j \frac{\sin sr_{ij}}{sr_{ij}} \quad s = 0, \ldots, 31.0\ \text{Å}^{-1} \quad (8)$$

Table 4
Correlation matrix for descriptors of Eq. (3) ($N = 20$).

	AMW	*JGI*1	*Mor*30*m*	*G*1*p*
AMW	1	0.5107	0.2385	0.3926
*JGI*1		1	0.3407	0.2924
*Mor*30*m*			1	0.8005
*G*1*p*				1

where N is the number of atoms; r_{ij} is the distance between atoms i and j; A_i can be any atomic property of atom i such as atomic number, mass, partial atomic charge, or atomic polarizability; s is a reciprocal distance. The value of s was considered only at discrete positions within a certain range. Normally 32 equidistant values between 0 and 31 $Å^{-1}$ were chosen. The choice of the range of s and the number of values to be considered determined the resolution of the code for representing the 3D structure [25,26]. For the case of *Mor*30*u*, an atomic mass weighted scheme was used and s was equal to 29 $Å^{-1}$.

WHIM (Weighted Holistic Invariant Molecular Descriptors) descriptors are based on statistical indices calculated on the projections of atoms along principal axes [27]. The aim is to capture 3D information regarding size, shape, symmetry and atom distributions with respect to invariant reference frames. To calculate them a weighted covariance matrix is obtained from different weighting schemes for the atoms: the unweighted case, atomic mass, van der Waals volume, Sanderson atomic electronegativity, atomic polarizability and electrotopological state indices. Depending on the weighting scheme different covariances matrices and hence different principal axes are obtained. Essentially the WHIM descriptors provide a variety of principal axes with respect to a defined atomic property. For each weighting scheme, a set of statistical indices is calculated on the atoms projected onto the principal axes (ie principal components). Descriptor *G*1*p* is a first component symmetry directional WHIM descriptors that involves the atomic polarizabilities as weighting scheme. These types of descriptors are univariate statistical indices calculated on the scores of the individual principal components.

The standardization of the regression coefficients of Eq. (3) allows assigning greater importance to the molecular descriptors that exhibit the largest absolute standardized coefficients [15]. In our case we have

$$JGI1(0.9444) > \text{AMW}(0.8664) > G1p(0.8102) > Mor30m(0.5811) \quad (9)$$

where the standardized coefficients are shown in parentheses. The ranking of contributions given by Eq. (9) suggests that topological charge descriptor *JGI*1 is the most relevant variable for the present set of flavonoids, thus indicating a significant dependence of the inhibitory activity on the charge of the atoms that form the molecule. The second descriptor in Eq. (9) is the average molecular weight implying that the activity has a significant dependence on the size of the molecules. This is further supported by the presence on Eq. (9) of a 3D-MoRSE descriptor (*Mor*30*m*) that is also weighted by atomic masses. The third WHIM descriptor (*G*1*p*) suggests a dependence on the atomic polarizabilities, although the contribution of this descriptor is lower than the first two of Eq. (9) it is still significant.

Additionally, since *G*1*p* and *Mor*30*m* encode tri-dimensional information that depends on the conformation of the molecule, it is possible to argue that the inhibitory activity of the present set of flavonoid derivatives has a considerable dependence on conformational changes

4. Conclusion

In this paper we constructed a predictive QSAR model of inhibitory activity against influenza H1N1 virus neuraminidase (NA) for 20 flavonoids of different classes using four molecular descriptors that take into account 2D- and 3D-aspects of the molecular structure. The model exhibited good predictive ability established by the theoretical and test set validations. The analysis of the model suggests that

the activity depends on the electric charges, masses and polarizabilities of the atoms present in the molecules as well as the conformation of the molecule. We expect the proposed model to be a useful tool in the prediction of the NA inhibitory activity, in a fast and costless manner, for any future studies that may require an estimation of this important activity of flavonoids, such as determination of candidates for synthesis. It is advisable that when more experimental information is available a second study to support the present results is performed.

Acknowledgements

The authors want to thank the National Research Council of Argentina (CONICET) and Universidad de Buenos Aires (UBA) for financial support; MINCYT for the electronic library facilities. ABP is a Senior Research Member of CONICET; AGM thanks CONICET for a Post-Doctoral Research Fellowship in PRALIB (CONICET, UBA), on leave from INIFTA (CONICET, UNLP).

References

[1] D.C. Wiley, J.J. Skehel, Annu. Rev. Biochem. 56 (1987) 365–394.
[2] K.G. Murti, R.G. Webster, Virology 149 (1986) 36–43.
[3] A.L. Liu, H.D. Wang, S.M. Lee, Y.T. Wang, G.H. Du, Bioorg. Med. Chem. 16 (2008) 7141–7147.
[4] A.L. Liu, Y.T. Wang, G.H. Du, Drugs Future 30 (2005) 799–806.
[5] F. Wei, S.-C. Ma, L.-Y. Ma, P.P.-H. But, R.-C. Lin, I.A. Khan, J. Nat. Prod. 67 (2004) 650–653.
[6] Y. Li, K.-T. Leung, F. Yao, L.S.M. Ooi, V.E.C. Ooi, J. Nat. Prod. 69 (2006) 833–835.
[7] K. Miki, T. Nagai, T. Nakamura, M. Tuji, K. Koyama, K. Kinoshita, K. Furuhata, H. Yamada, K. Takahashi, Heterocycles 75 (2008) 879–885.
[8] K. Miki, T. Nagai, K. Suzuki, R. Tsujimura, K. Koyama, K. Kinoshita, K. Furuhata, H. Yamada, K. Takahashi, Bioorg. Med. Chem. Lett. 17 (2007) 772–775.
[9] C. Hansch, A. Leo, Exploring QSAR: Fundamentals and Applications in Chemistry and Biology. Am. Chem. Soc., Washington, D.C, 1995.
[10] A.G. Mercader, P.R. Duchowicz, F.M. Fernandez, E.A. Castro, Chemom. Intell. Lab. Syst. 92 (2008) 138–144.
[11] S. Bantia, C.D. Parker, S.L. Ananth, L.L. Horn, K. Andries, P. Chand, P.L. Kotian, A. Dehghani, Y. El-Kattan, T. Lin, T.L. Hutchison, J.A. Montgomery, D.L. Kellog, Y.S. Babu, Antimicrob. Agents Chemother. 45 (2001) 1162–1167.
[12] HYPERCHEM 6.03 (Hypercube). http://www.hyper.com.
[13] DRAGON Release 5.0 Evaluation Version. http://www.disat.unimib.it/chm.
[14] R. Todeschini, V. Consonni, Handbook of Molecular Descriptors. Wiley VCH, Weinheim, Germany, 2000.
[15] N.R. Draper, H. Smith, Applied Regression Analysis. John Wiley&Sons, New York, 1981.
[16] S.-S. So, M. Karplus, J. Med. Chem. 39 (1996) 1521–1530.
[17] C. Hansch, Comprehensive Drug Design. Pergamon Press, New York, 1990.
[18] D.M. Hawkins, S.C. Basak, D. Mills, J. Chem. Inf. Model. 43 (2003) 579–586.
[19] Matlab 5.0 The MathWorks Inc. http://www.mathworks.com/
[20] S. Wold, L. Eriksson, in: H.V.D. Waterbeemd (Ed.), Statistical Validation of QSAR Results, VCH, Weinheim, 1995, pp. 309–318.
[21] H. Liu, P. Gramatica, Bioorg. Med. Chem. 15 (2007) 5251–5261.
[22] A. Golbraikh, A. Tropsha, J. Mol. Graph Model. 20 (2002) 269–276.
[23] J. Galvez, R. Garcia-Domenech, M.T. Salabert, R. Soler, J. Chem. Inf. Comput. Sci. 34 (1994) 520–525.
[24] J. Galvez, R. Garcia-Domenech, J.V. de Julian-Ortiz, R. Soler, J. Chem. Inf. Comput. Sci. 35 (1995) 272–284.
[25] J. Gasteiger, J. Sadowski, J. Schuur, P. Selzer, L. Steinhauer, V. Steinhauer, J. Chem. Inf. Comput. Sci. 36 (1996) 1030–1037.
[26] J.H. Schuur, P. Selzer, J. Gasteiger, J. Chem. Inf. Comput. Sci. 36 (1996) 334–344.
[27] R. Todeschini, P. Gramatica, Quantitative Structure-Activity Relationships 16 (1997) 113–119.
[28] V. Escuret, E. Frobert, M. Bouscambert-Duchamp, M. Sabatier, I. Grogc, M. Valette, B. Lina, F. Morfina, O. Ferraris, J. Clin. Virol. 41 (2008) 25–28.