Available online at www.sciencedirect.com
ScienceDirect
European Journal of Medicinal Chemistry 42 (2007) 1370–1381
EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY
http://www.elsevier.com/locate/ejmech

Original article

Dragon method for finding novel tyrosinase inhibitors: *Biosilico* identification and experimental *in vitro* assays

Gerardo M. Casañola-Martín [a,b,c], Yovani Marrero-Ponce [a,b,d,*], Mahmud Tareq Hassan Khan [e,f], Arjumand Ather [g], Khalid M. Khan [h], Francisco Torrens [d], Richard Rotondo [i]

[a] *Unit of Computer-Aided Molecular ''Biosilico'' Discovery and Bioinformatic Research (CAMD-BIR Unit), Department of Pharmacy, Faculty of Chemistry–Pharmacy, Central University of Las Villas, Santa Clara, 54830 Villa Clara, Cuba*
[b] *Department of Drug Design, Chemical Bioactive Center, Central University of Las Villas, Santa Clara, 54830 Villa Clara, Cuba*
[c] *Department of Biological Sciences, Faculty of Agricultural Sciences, University of Ciego de Avila, 69450 Ciego de Avila, Cuba*
[d] *Institut Universitari de Ciència Molecular, Universitat de València, Edifici d'Instituts de Paterna, Poligon la Coma s/n (detras de Canal Nou) P.O. Box 22085, E-46071 Valencia, Spain*
[e] *Pharmacology Research Lab., Faculty of Pharmaceutical Sciences, University of Science and Technology, Chittagong, Bangladesh*
[f] *Department of Pharmacology, Institute of Medical Biology, University of Tromso, Tromso 9037, Norway*
[g] *The Norwegian Structural Biology Centre (NorStruct), University of Tromso, Tromso 9037, Norway*
[h] *HEJ Research Institute of Chemistry, Pakistan*
[i] *Advanced Medisyns, Inc., 601 Carlson Parkway, Suite 1050, Minnetonka, MN 55305, USA*

Received 12 September 2006; received in revised form 18 January 2007; accepted 19 January 2007
Available online 23 February 2007

Abstract

QSAR (quantitative structure–activity relationship) studies of tyrosinase inhibitors employing Dragon descriptors and linear discriminant analysis (LDA) are presented here. A data set of 653 compounds, 245 with tyrosinase inhibitory activity and 408 having other clinical uses were used. The active data set was processed by *k*-means cluster analysis in order to design training and prediction series. Seven LDA-based QSAR models were obtained. The discriminant functions applied showed a globally good classification of 99.79% for the best model Class $= -96.067 + 1.988 \times 10^2$ **X0Av** $+ 91.907$ **BIC3** $+ 6.853$ **CIC1** in the training set. External validation processes to assess the robustness and predictive power of the obtained model were carried out. This external prediction set had an accuracy of 99.44%. After that, the developed models were used in ligand-based virtual screening of tyrosinase inhibitors from the literature and never considered in either training or predicting series. In this case, all screened chemicals were correctly classified by the LDA-based QSAR models. As a final point, these fitted models were used in the screening of new bipiperidine series as new tyrosinase inhibitors. These methods are an adequate alternative to the process of selection/identification of new bioactive compounds. The *biosilico* assays and *in vitro* results of inhibitory activity on mushroom tyrosinase showed good correspondence. It is important to stand out that compound **BP4** ($IC_{50} = 1.72$ μM) showed higher activity in the inhibition against the enzyme than reference compound kojic acid ($IC_{50} = 16.67$ μM) and L-mimosine ($IC_{50} = 3.68$ μM). These results support the role of *biosilico* algorithm for the identification of new tyrosinase inhibitor compounds.

Keywords: Dragon descriptor; LDA-based QSAR model; Tyrosinase inhibitor; Bipiperidine series; Virtual screening

* Corresponding author. Unit of Computer-Aided Molecular "*Biosilico*" Discovery and Bioinformatic Research (CAMD-BIR Unit), Department of Pharmacy, Faculty of Chemistry–Pharmacy, Central University of Las Villas, Santa Clara, 54830 Villa Clara, Cuba. Tel.: +53 42 281192/473 (Cuba); +963543156 (València); fax: +53 42 281130/455 (Cuba); +963543156 (València); cell: 610028990.
E-mail addresses: ymarrero77@yahoo.es, yovani.marrero@uv.es, ymponce@gmail.com, yovanimp@qf.uclv.edu.cu (Y. Marrero-Ponce).
URL: http://www.uv.es/yoma/

0223-5234/$ - see front matter © 2007 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmech.2007.01.026

"*Predictive in silico models could be used for drug target identification, accelerating the selection process of lead compounds…*" Watson, C. Biosilico. **2003**, *1*, 83–84.

1. Introduction

Tyrosinase (EC. 1.14.18.1) is a copper-containing enzyme widely distributed in nature including fungi, higher plants, and animals. This enzyme catalyzes two key reactions in the melanin biosynthesis pathway, the hydroxylation of monophenol to *o*-diphenol (monophenolase activity) and conversion of an *o*-diphenol to the corresponding *o*-quinone (diphenolase activity), involving reactive oxygen species (ROS) [1,2]. Quinones are highly reactive compounds and can polymerize spontaneously to form high-molecular weight compounds or brown pigments (melanins), or react with amino acids and proteins that enhance the brown color produced [3].

Alterations in melanin synthesis occur in many disease states like hyperpigmentation, melasma and age spots [4]. Melanin pigments are also found in mammalian brain and tyrosinase may play a role in neuromelanin formation in the human brain. This mixed-function oxidase could be central to dopamine neurotoxicity and may contribute to the neurodegeneration associated with Parkinson's disease [5]. Melanoma specific anticarcinogenic activity is also known to be linked with tyrosinase activity [6].

The standard topical treatments for hyperpigmentation disorders include tyrosinase inhibitors, some compounds with the inhibitory activity are used in medicine, but the majority of them do not satisfy all requirements of clinical efficacy, or adverse effects are observed [4,7]. As result of these clinical behaviors and other side effects, there has been a constant search for new herbal or synthesized compounds with anti-tyrosinase activity [8–10]. In this sense, one of our group's researches has been focused on finding new potent tyrosinase inhibitors through 'trial-and-error' techniques [11,12].

By other way, the *in silico* techniques have proven their usefulness in the pharmaceutical research for the selection/identification and/or design/optimization of new chemical entities (NCE), to transform early stage drug discovery, particularly in terms of time- and cost-savings [13]. QSAR approaches report a high incidence of the use of different molecular descriptors for the *in silico* drug screening [14–17].

The congeneric data set used in SAR and QSAR studies of tyrosinase inhibitors [11,12,18–20] do not provide enough tools for drug development, this kind of data can only be applied to structural lead optimization. Therefore, database of heterogeneous compounds may be a successful tool in QSAR research of tyrosinase inhibitors and the discovery of novel lead compounds with different structural features and more effective activity [21–23].

In the present paper, we used the Dragon descriptors, extensively applied to describe biological activities [24,25] and linear discriminant analysis (LDA) strategy to find classification functions that allow to discriminate tyrosinase inhibitor compounds from inactive ones. As a final point, the *in silico* selection (identification), isolation, and *in vitro* assays of a new series of compounds were carried out to show the applicability of Dragon descriptors in the *biosilico* drug discovery processes.

2. Materials and methods

2.1. Chemical data set

Selected data set of this study was constructed warranting enough molecular diversity on it. Taking this into account, we selected a data set of 653 organic-chemicals having a great structural variability, 245 of them having tyrosinase inhibitory activity reported and the rest inactive ones [26] (408 compounds having different clinical uses, such as antivirals, sedative/hypnotics, diuretics, anticonvulsivants, hemostatics, oral hypoglycemics, antihypertensives, antihelminthics, anticancer compounds, and so on) were employed.

The database of active compounds was chosen considering a representation of most of the different inhibition modes in the case of the compounds with tyrosinase inhibitory activity. For instance, it includes compounds that belong to different subsystems such as azobenzene derivatives [27], kojic acid tripeptide library [28], disubstituted-oxadiazole analogues [11], longifolene derivatives [29], glycyrrhetinic acid derivatives [30], novel *N*-substituted *N*-nitrosohydroxylamines [31,32], catechins [33], gentisic acid esters [34], hydroxystilbene compounds [35], and benzaldoximes [36]. Fig. 1 shows a representative sample of such inhibitors from these data. In Table 1 of Supporting Information, the names of compounds in the database are given, together with their experimental data taken from the literature. The molecular structures of these 246 tyrosinase inhibitors are given as Supporting Information (see Table 2).

The great structural variability of chemicals in training and prediction series can assure an adequate extrapolation power. In this sense, the selection process is not constrained to compounds with only the structural features included in data and new series of compounds can be discovered. It is an important remark that these data provide a useful tool for scientific research in synthesis, natural-product chemistry, theoretical chemistry and other areas related to the field of tyrosinase inhibitors. A *k*-MCA was carried out to split the active data set into training and prediction tests in a rational way.

2.2. Dragon molecular descriptors

The molecular descriptors were calculated using the Dragon [37] software; these were the Constitutional, Topological, BCUT, Galvez topological charge, 2D autocorrelations, empirical and properties descriptors [38]. Descriptors with constant values inside each group were discarded. For the remaining descriptors, a pairwise correlation analysis for all families of descriptors was carried out. The presented exclusion method was used to reduce, in a first step, the collinearity and correlation between descriptors.

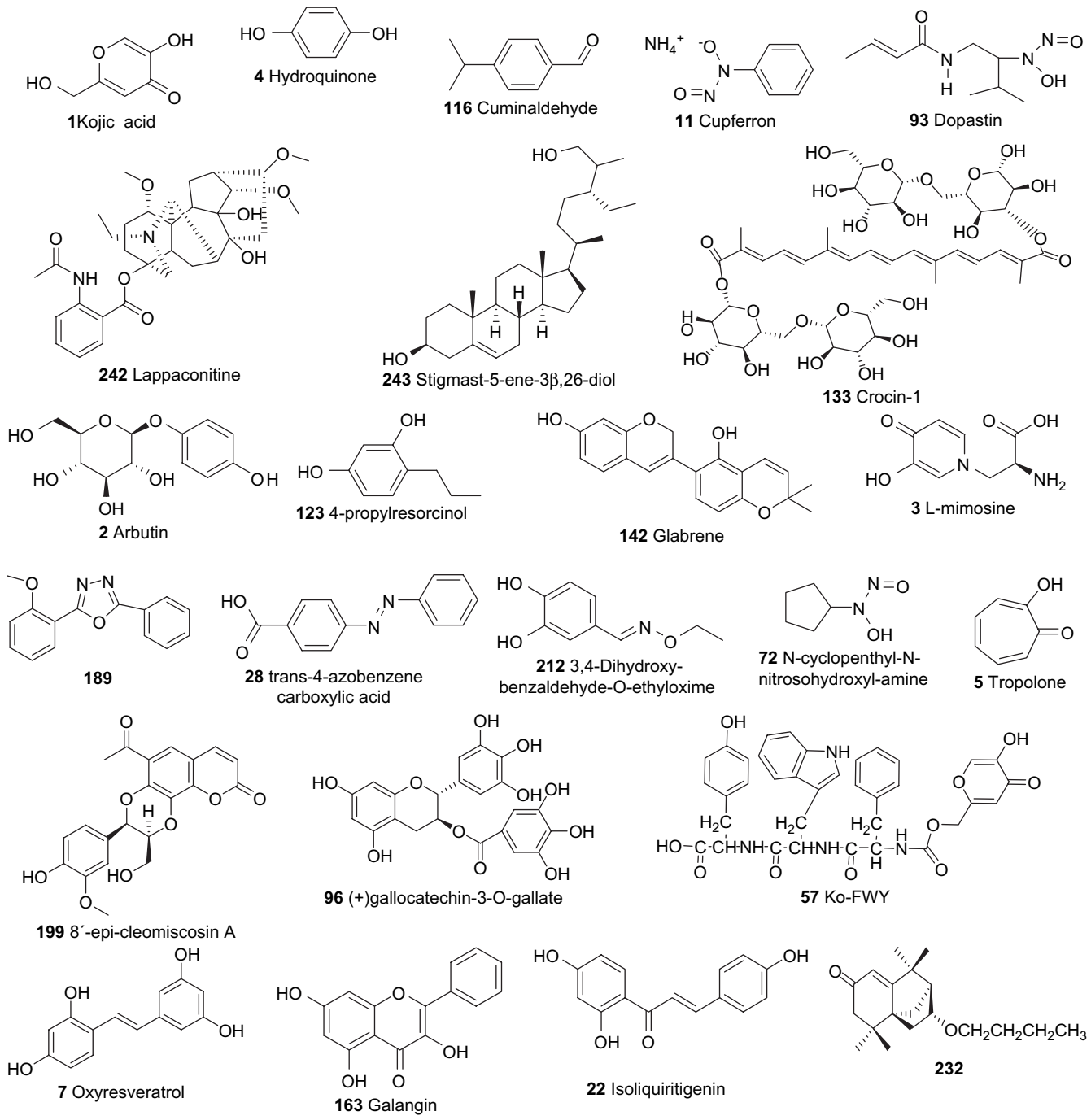

Fig. 1. Random, but not exhaustive, sample of the molecular families of tyrosinase inhibitors studied here.

2.3. Chemometric techniques

2.3.1. k-Means cluster analysis (k-MCA)

The statistical software package STATISTICA was used to develop the *k*-MCA [39]. The number of members in each cluster and the standard deviation of the variables in the cluster (kept as low as possible) were taken into account, to have an acceptable statistical quality of data partitions in the clusters. The values of the standard deviation (SS) between and within clusters, of the respective Fisher's ratio and their *p* level of significance, were also examined [40,41]. Finally, before carrying out the cluster processes, all the variables were standardized. In standardization, all values of selected variables (e.g. the complete molecular descriptor data set after excluding constant values and pair wise correlation analysis) were replaced by standardized values, which are computed as follows: Std. score = (raw score − mean)/Std. deviation.

2.3.2. Linear discriminant analysis (LDA)

LDA was carried out with the STATISTICA software [39]. The considered tolerance parameter (proportion of variance that is unique to the respective variable) was the default value

for minimum acceptable tolerance, which is 0.01. A forward stepwise search procedure was fixed as the strategy for variable selection. The principle of parsimony (Occam's razor) was taken into account as a strategy for model selection. In connection, we selected the model with a high statistical significance but having as few parameters (ak) as possible. The quality of the models was determined by examining Wilks' λ parameter (U statistic), the square Mahalanobis distance (D^2), the Fisher's ratio (F), and the corresponding p level [$p(F)$] as well as the percentage of good classification in the training and test sets. Models with a proportion between the number of cases and variables in the equation lower than 5 were rejected. The biological activity was codified by a dummy variable ''Class''. This variable indicates the presence of either an active compound [(Class) 1] or an inactive compound [Class) − 1]. The classification of cases was performed by means of the posterior classification probabilities. By using the models, one compound can then be classified as active, if $\Delta P\% > 0$, being $\Delta P\%$, $[P(\text{Active}) - P(\text{Inactive})]100$, or as inactive otherwise. P(Active) and P(Inactive) are the probabilities with which the equations classify a compound as active or inactive, respectively.

The statistical robustness and predictive power of the obtained model were assessed using a prediction (test) set. 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 (MCC) in the training and test sets permitted the assessment of the model [42].

2.3.3. Orthogonalization of descriptors

In this study, the Randić method of orthogonalization was used [43]. This orthogonalization process of molecular descriptors was introduced by Randić several years ago as a way to improve the statistical interpretation of the models by using interrelated indices. This method has been described in detail in several publications. Thus, we will give only a general overview here. As a first step, an appropriate order of orthogonalization was considered following the order with which the variables were selected in the forward stepwise search procedure of the statistical analysis. The first variable (V_1) is taken as the first orthogonal descriptor $1O(V_1)$, and the second one (V_2) is orthogonalized with respect to it [$2O(V_2)$]. The residual of its correlation with $1O(V_1)$ is that part of descriptor V_2 not reproduced by $1O(V_1)$. Similarly, from the regression of V_3 versus $1O(V_1)$, the residual is the part of V_3 that is not reproduced by $1O(V_1)$, and it is labeled $1O(V_3)$. The orthogonal descriptor $3O(V_3)$ is obtained by repeating this process in order to also make it orthogonal to $2O(V_2)$. The process is repeated until all variables are completely orthogonalized, and the orthogonal variables are then used to obtain the new model [43–49].

Because the different molecular descriptors included here used entirely ''different types of scales'', the data were standardized so that each variable has a mean 0 and a standard deviation 1. In standardization, all values of selected variables (molecular descriptors) were replaced by standardized values, which were computed as follows: Std. score = (raw score − mean)/Std. deviation.

2.4. Chemical procedures

The synthesis and structural characterization of the bipiperidine series and biological studies and cross references have been reported in some detail elsewhere by other member of our research team [50].

2.5. Experimental corroboration of tyrosinase inhibitory activity

Tyrosinase inhibition assay was performed with kojic acid and L-mimosine as standard inhibitors for the tyrosinase in a 96-well microplate format using a SpectraMax 340 microplate reader (Molecular Devices, CA, USA) according to the method developed by Hearing [51]. Briefly, first the compounds were screened for the *o*-diphenolase inhibitory activity of tyrosinase using L-DOPA as substrate. All the active inhibitors from the preliminary screening were subjected to IC_{50} studies. Compounds were dissolved in methanol to a concentration of 2.5%. Thirty units of mushroom tyrosinase (28 nM from Sigma Chemical Co., USA) was first preincubated with the test compounds in 50 nM Na−phosphate buffer (pH 6.8) for 10 min at 25 °C. Then the L-DOPA (0.5 mM) was added to the reaction mixture and the enzymatic reaction was monitored by measuring the change in absorbance at 475 nm (at 37 °C) due to the formation of the DOPAchrome for 10 min. The percent inhibition of the enzyme was calculated as follows, by using MS Excel®™ 2000 (Microsoft Corp., USA) based program developed for this purpose:

$$\text{Percent inhibition} = [B - S/B]100 \quad (1)$$

Here B and S are the absorbances for the blank and samples, respectively. After screening of the compounds, median inhibitory concentrations (IC_{50}) were also calculated. All the studies have been carried out at least in triplicates and the result represents the mean ± SEM (standard error of the mean). Kojic acid and L-mimosine were used as standard inhibitors for the tyrosinase and both of them were purchased from Sigma Chemical Co., USA.

3. Results and discussion

3.1. Design of training and test set

In the first place, the molecular diversity of active compounds should be assured, and in this sense a hierarchical cluster analysis (CA) is developed with the STATISTICA software [39]. Fig. 2 show a dendrogram, where a large number of different subsets can be observed, proving the structural diversity of the active data set (tyrosinase inhibitors).

Data of 408 drugs having a series of different clinical uses were chosen as inactive set. In this case, these chemicals are untested compounds as tyrosinase inhibitors, and the

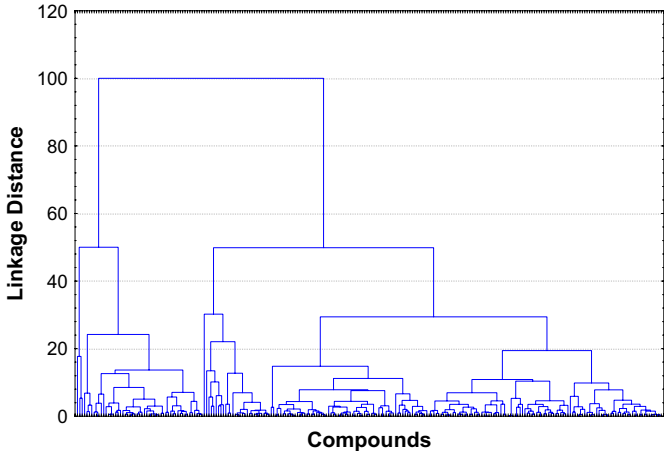

Fig. 2. A dendrogram illustrating the results of the hierarchical *k*-NNCA of the set of tyrosinase inhibitors used in the training and prediction set of the present work.

classifications of these compounds as 'inactive' (non-inhibitors of tyrosinase) do not assure that any inhibitory activity does not exist for those organic-chemicals. This problem can be reflected in the results of classification for the series of inactive chemicals [52].

Second, a *k*-MCA is carried out to ensure that any chemical subsystem selected will be in both learning and external sets, in a representative way. The *k*-MCA was made with active compounds and partitioned the tyrosinase inhibitors into 10 clusters. Topological descriptors were used, with all variables showing p-levels <0.05 for the Fisher's test. The results are shown in Table 1. In the case of inactive (untested) data set, the selection of compounds for every subset (training and test) was made at random.

Afterwards, the selection of the training and prediction sets for the active database was performed by taking, in random way, compounds belonging to each cluster. From these 653 chemicals, 474 were chosen at random to form the training set, being 182 of them active ones and 292 inactive ones. The remaining subseries composed of 63 tyrosinase inhibitors and 116 compounds with different biological properties were prepared as test set for the external validation of the classification models (179 compounds). These chemicals were never used in the development of the classification models. Fig. 3 illustrates graphically the above-described procedure where one independent cluster analysis for active compounds and

Table 1
Main results of the *k*-MCAs, for tyrosinase inhibitors and inactive drug-like compounds

Analysis of variance				
Variables	Between SS[a]	Within SS[b]	Fisher's ratio (F)	p-Level[c]
Tyrosinase inhibitor clusters (*k*-MCA)				
X0Av	0.87	3.57	6.42	0.00
BIC3	2.71	3.46	20.43	0.00
CIC1	185.59	6.03	803.24	0.00

[a] Variability between groups.
[b] Variability within groups.
[c] Level of significance.

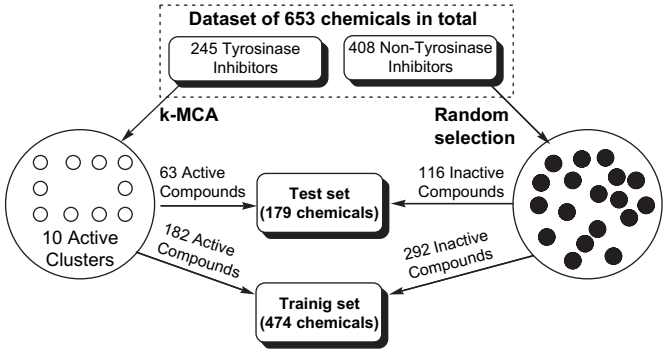

Fig. 3. General algorithm used to design training and test sets.

a random selection for the inactive compounds were performed to select a representative sample for the training and test sets.

3.2. Finding discriminant models

3.2.1. Classification functions

Although many different chemometric techniques could be used to fit discriminant functions, such as SIMCA or neural networks, in our case, we select the linear discriminant analysis (LDA) given the simplicity of the method, in order to derive discriminant functions that permit the classification of compounds as tyrosinase inhibitors or inactive ones. The LDA has become an important tool for the prediction of chemical bioactive properties [47–49,53–57].

In the present study, we developed discriminant functions, using Dragon descriptors as independent variables. Seven LDA-based QSAR models were obtained. The models used the Constitutional, Topological, BCUT, Galvez topological charge, 2D autocorrelations, empirical and properties as molecular descriptors in this order (Eqs. (2)–(8)), respectively. The classification models obtained are given in Table 2, and the meaning of the variables included in the models, are depicted in Table 3.

Table 4 summarizes the prediction performances and the statistical parameters for LDA-based QSAR models with the training set. The equations showed to be statistically significant at p-level <0.0001. The fitted model for Eq. (3) showed the best result in these classification functions, this best model (Eq. (3)) has an appropriate overall accuracy of 99.79% in the training set. The equation showed a high Matthews' correlation coefficient (MCC = 1). MCC quantifies the strength of the linear relation between the molecular descriptors and the classifications, and it may often provide a much more balanced evaluation of the prediction than, for instance, the percentages (accuracy) [42].

Also, we list in Table 4 most of the parameters commonly used in medical statistics [sensitivity, specificity and false positive rate (also known as 'false alarm rate')] for the whole set of developed models. While the sensitivity is the probability of correctly predicting a positive example, the specificity (also known as 'hit rate') is the probability that a positive prediction is correct [42]. These statistical parameters mentioned above,

Table 2
Discriminant models obtained with the 0D–2D Dragon descriptors

$$\text{Class} = -4.438 \times 10^2 + 72.326\,\text{Mv} + 4.061 \times 10^2\,\text{Me} - 2.464\,\text{AMW} + 6.450\,\text{Ms} + 2.036\,\text{nCIC} - 0.690\,\text{nAB} + 2.450\,\text{nR05} \quad (2)$$

$$\text{Class} = -96.067 + 1.988 \times 10^2\,\mathbf{X0Av} + 91.907\,\mathbf{BIC3} + 6.853\,\mathbf{CIC1} \quad (3)$$

$$\text{Class} = -40.001 + 8.931\,\text{BEHv1} + 8.238\,\text{BELm1} - 3.596\,\text{BEHm7} + 4.772\,\text{BELm5} + 6.231\,\text{BELv2} - 2.318\,\text{BEHp4} + 2.795\,\text{BELe4} \quad (4)$$

$$\text{Class} = -5.143 + 32.291\,\text{JGT} - 0.942\,\text{GGI1} + 4.553\,\text{GGI6} - 1.120\,\text{GGI2} + 1.831\,\text{GGI5} \quad (5)$$

$$\text{Class} = -74.331 + 88.885\,\text{ATS1v} - 8.054\,\text{ATS2m} - 2.287\,\text{ATS4m} + 44.373\,\text{ATS3e} \quad (6)$$

$$\text{Class} = -7.020 - 12.331\,\text{Hy} + 1.048\,\text{Ui} \quad (7)$$

$$\text{Class} = -2.854 - 6.898 \times 10^{-5}\,\text{MR} - 7.801 \times 10^{-3}\,\text{PSA} - 3.811 \times 10^{-4}\,\text{MLOGP} \quad (8)$$

Table 3
Symbols of the descriptors used in the QSAR models and their definitions

Symbols	Descriptor definition
Mv	Molecular weight
Me	Mean atomic Sanderson electronegativity (scaled on Carbon atom)
AMW	Average molecular weight
Ms	Mean electrotopological state
nCIC	Number of rings
nAB	Number of aromatic bonds
nR05	Number of fivve-membered rings
X0Av	Average valence connectivity index γ-0
BIC3	Bond information content (neighborhood symmetry of 3-order)
CIC1	Complementary information content (neighborhood symmetry of 1-order)
BEHv1	Highest eigenvalue n.1 of burden matrix/weighted by atomic van der Waals volumes
BELm1	Lowest eigenvalue n.1 of burden matrix/weighted by atomic masses
BEHm7	Highest eigenvalue n.7 of burden matrix/weighted by atomic masses
BELm5	Lowest eigenvalue n.5 of burden matrix/weighted by atomic masses
BELv2	Lowest eigenvalue n.2 of burden matrix/weighted by atomic van der Waals volumes
BEHp4	Highest eigenvalue n.4 of burden matrix/weighted by atomic polarizabilities
BELe4	Lowest eigenvalue n.4 of burden matrix/weighted by atomic Sanderson electronegativities
JGT	Global topological charge index
GGI1	Topological charge index of order 1
GGI2	Topological charge index of order 2
GGI5	Topological charge index of order 5
GGI6	Topological charge index of order 6
ATS1v	Broto–Moreau autocorrelation of a topological structure-lag 1/weighted by atomic van der Waals
ATS2m	Broto–Moreau autocorrelation of a topological structure-lag 2/weighted by atomic masses
ATS4m	Broto–Moreau autocorrelation of a topological structure-lag 4/weighted by atomic masses
ATS3e	Broto–Moreau autocorrelation of a topological structure-lag 3/weighted by atomic Sanderson
Hy	Hydrophilic factor
Ui	Unsaturation index
MR	Ghose–Crippen molar refractivity CCF
PSA	Fragment-based polar surface area
MLOGP	Moriguchi octanol–water partition coefficient (log *P*)

together with the linear discriminant canonical statistics: canonical regression coefficient (R_{can}) as well as Chi-squared (χ^2) and its *p*-level [$p(\chi^2)$] were checked and results are depicted in Table 4.

The canonical transformations of the LDA results with the Topological descriptors (Eq. (3)) give rise to canonical roots with a good canonical correlation coefficient of 0.99. Chi-squared test allows us to assess the statistical signification of this analysis as having a *p*-level <0.0001.

3.2.2. Validation test

The statistical parameters in the complete training data set provide some assessment of the goodness of fit of the models, but it is not enough to assure the predictive power of the models. For that reason, we carried out an external validation process using a test set [58,59].

In this sense, the activity of the compounds in the test set was predicted with the obtained discrimination functions. Eq. (3) shows a 99.44% ($C = 0.99$) in the prediction series. The results of the classifications for all models in the test set are depicted in Table 5. The accuracy and other statistical parameters (sensitivity, specificity and false positive rate) of the test set are depicted in Table 5. These results validate the models for the use in the ligand-based virtual screening taking into consideration that 85.0% is considered as an acceptable threshold limit for this kind of analysis [60].

The results of classification (including the canonical scores) using all developed models for the complete set of organic-chemicals in training data sets are shown in Tables 3 and 4 of Supporting Information. Conversely, the active and inactive compounds in test sets, as well as their classification using all equations, are also given in Supporting Information (see Tables 5 and 6).

3.2.3. Descriptors' orthogonalization process

On the other hand, a good method to eliminate the collinearity is the pairwise correlation analysis, but the correlation between variables can persist, as it was observed after a close inspection of the molecular fingerprints included in the best

Table 4
Prediction performances and statistical parameters for LDA-based QSAR models in the training set

Models[a]	Matthews' correlation coefficient (C)	Accuracy 'Q_{Total}' (%)	Specificity (%)	Sensitivity 'hit rate' (%)	False positive rate (%)	Wilks' λ	D^2	F	Chi-squared (χ^2)	Canonical R (R_{can})[b]
LDA-based QSAR models obtained using the Dragon descriptors										
Eq. (2) (7)	0.99	99.58	98.91	100	0.68	0.04	100.0	1582	1503	0.98
Eq. (3) (3)	**1**	**99.79**	**99.45**	**100**	**0.34**	**0.02**	**180.5**	**6719**	**1779**	**0.99**
Eq. (4) (7)	0.93	96.84	95.13	96.70	3.08	0.21	16.01	253.2	735.3	0.89
Eq. (5) (5)	0.97	98.52	96.79	99.45	2.05	0.18	19.5	433.6	811.5	0.91
Eq. (6) (4)	0.98	98.95	97.33	100	1.71	0.05	80.9	2253	1413	0.97
Eq. (7) (2)	0.93	97.05	100	92.30	0	0.33	8.49	475.2	520.2	0.82
Eq. (8) (3)	0.85	92.40	83.80	99.45	11.99	0.45	5.21	193.8	378.9	0.74

[a] In brackets, the quantity of variables of the models.
[b] Canonical correlation coefficient obtained from the linear discriminant **canonical** analysis.

LDA-based QSAR model. In Table 6, we give the correlation coefficient of the molecular descriptors in Eq. (3).

It is well known that interrelation among the molecular descriptors makes difficult the interpretation of the QSAR model [44–49], and underestimates the utility of the correlation coefficient in a model. To overcome this difficulty, we used the Randić's orthogonalization process of the molecular descriptors. The main philosophy of this approach is to avoid the exclusion of descriptors on the basis of its collinearity with other variables included in the model. However, in some cases strongly interrelated descriptors can enhance the quality of a model because the small fraction of a descriptor which is not reproduced by its strongly interrelated pair can provide positive contributions to the modeling. This process is an approach in which molecular descriptors are transformed in such a way that they do not mutually correlate (see Section 2.3). Both, the non-orthogonal (original) descriptors and the derived orthogonal descriptors contain the same information. Therefore, the same statistical parameters of the QSAR models are obtained [44–49].

In Eq. (9) are shown the results of the orthogonalization of the topological descriptors included in model:

$$\begin{aligned} \text{Class} &= -43.937 + 77.726^{1}O(\mathbf{X0Av}) + 35.142^{2}O(\mathbf{BIC3}) \\ &\quad +4.780^{3}O(\mathbf{CIC1}) \\ N &= 474;\ \lambda = 0.02;\ D^2 = 180.5; \\ F &= 6719;\ R = 0.99;\ \chi^2 = 1779;\ Q = 99.79;\ C = 1 \end{aligned} \tag{9}$$

Table 5
Prediction performances for LDA-based QSAR models in the test set

Models	Matthews' correlation coefficient (C)	Accuracy 'Q_{Total}' (%)	Specificity (%)	Sensitivity 'hit rate' (%)	False positive rate (%)
LDA-based QSAR models obtained using the Dragon descriptors					
Eq. (2)	0.98	98.88	96.9	100	1.7
Eq. (3)	**0.99**	**99.44**	**98.4**	**100**	**0.9**
Eq. (4)	0.89	94.97	92.2	93.7	4.3
Eq. (5)	0.96	98.32	98.4	96.8	0.9
Eq. (6)	0.96	98.32	95.5	100	2.6
Eq. (7)	0.98	98.88	100	96.8	0
Eq. (8)	0.78	88.26	75.0	100	18.1

Here, we used the symbols ${}^{m}O(\mathbf{b})$, where the superscript m expresses the order of importance of the variable (**b**) after a preliminary forward stepwise analysis and O means orthogonal.

It is an important remark here that the orthogonal descriptor-based models coincide with the collinear (i.e. ordinary) topological descriptor-based models in all the statistical parameters. The statistical coefficients of LDA–QSARs λ, F, D^2, C, accuracy, are the same whether we use a set of non-orthogonal descriptors or the corresponding set of orthogonal indices. This is not surprising, because the latter models are derived as a linear combination of the former ones and cannot have more information content than them [44–49].

In the process of orthogonalization, the data were standardized so that each variable has a mean of zero and a standard deviation of 1, because the different molecular descriptors used entirely ''different types of scales''.

3.3. Novel tyrosinase inhibitors through virtual screening identification

One of the most common approaches reported recently in the area of drug discovery is the *in silico* methods, this tool permits the assay of virtual libraries of chemicals, and can predict ahead of time, the likely result of many-year biological-property study. This process is associated to the great costs involved during the discovery of new drug-like compounds by the pharmaceutical industries. Virtual assays can be considered in this case a novel paradigm inside the new automation and information technologies, and can provide to the pharmaceutical industry platforms to translate clinical liabilities into simple, fast and cost-effective *in vitro* screening assays, applicable to the early phases of drug discovery [61].

In this context and with the aim to prove the possibilities of the present approach for the ligand-based virtual screening of

Table 6
Correlation matrix of the variables in Eq. (3)

	Non-orthogonal topological descriptors		
	X0Av	**BIC3**	**CIC1**
X0Av	1.00	0.91	0.45
BIC3		1.00	0.39
CIC1			1.00

Table 7
Results of the virtual screening

Compound[a]	Class[b]	Ref.[c]	Compound[a]	Class[b]	Ref.[c]
Active compounds (tyrosinase inhibitors)					
1 *p*-Nitrophenol	+++++++	A B	**42** Methimazole	+++++++	V
2 3-(3,4-Dihydroxyphenyl)-l-alanine	+++++++	C	**43** BMY-28438	+++++++	V
3 3-Amino-4-hydroxybenzoic acid	+++++++	C	**44** Captopril	+++++++	W
4 4-Amino-3-hydroxybenzoic acid	+++++++	C	**45** Yohimbine	+++++−+	X
5 3,4-Diaminobenzoic acid	+++++++	C	**46** 4-(Phenylazo)phenol	+++++++	Y
6 3-Aminobenzoic acid	+++++++	C	**47** SACat	+++++++	Y
7 4-Aminobenzoic acid	+++++++	C	**48** NPACat	+++++++	Y
8 4,6-*O*-Hexahydroxy-diphenylglucose	+++++++	D	**49** DNPACat	+++++++	Y
9 Tunicamycin	+++++++	E	**50** EDTA	+++++++	Z
10 Methyl-*p*-coumarate	+++++++	F	**51** Dodecyl gallate	+++++++	A1
11 *o*-Phenylphenol	+++++++	F	**52** Gallic acid	+++++++	A1
12 Phenylhydroquinone	+++++++	F	**53** (±)-Flavanone	+++++++	B2
13 Chamaecin	+++++++	F G	**54** (−)-Pinocembrin	+++++++	B2
14 Stearyl glycyrrhetinate	+++++++	H	**55** (±)-Naringenin	+++++++	B2
15 3-Flurotyrosine	+++++++	I	**56** (+)-Dihydromorin	+++++++	B2
16 *N*-Acetyltyrosine	+++++++	I	**57** Flavone	+++++++	B2
17 *N*-Formyltyrosine	+++++++	I	**58** Myricetin	+++++++	B2
18 Gentisic acid	+++++++	J	**59** Artocarpin	+++++++	B2
19 6-BH_4	+++++++	K	**60** Artocarpesin	+++++++	B2
20 7-BH_4	+++++++	K	**61** Isoartocarpesin	+++++++	B2
21 Propylparaben	+++++−+	L	**62** (−)-Angolensin	+++++++	B2
22 Phenylalanine	+++++++	I	**63** Pinosylvin	+++++++	B2
23 Dithiothreitol	++−++−+	M	**64** 4-Prenyloxyresveratrol	+++++++	B2
24 Azelaic acid	++−++++	N	**65** 26	+++++++	B2
25 Undecandioic acid	++−++++	N	**66** 27	+++++++	B2
26 Suberic acid	++−++++	N	**67** 28	+++++−+	B2
27 Sebacic acid	++−++++	N	**68** 29	+++++++	B2
28 Dodecandioic acid	++−++++	N	**69** 30	+++++++	B2
29 Tridecandioic acid	++−++++	N	**70** 31	+++++++	B2
30 Traumatic acid	++−++++	N	**71** 32	+++++++	B2
31 Pantothenic acid	+++++++	I	**72** 34	+++++++	B2
32 5-(Hydroxymethyl)-2-furfural	+++++++	O P	**73** 35	+++++++	B2
33 Hinokitiol	+++++++	Q	**74** 36	+++++++	B2
34 Penicillamine	+++++++	R	**75** 37	+++++++	B2
35 Toluic acid	+++++++	A	**76** 38	+++++++	B2
36	+++++−+	S	**77** 39	+++++++	B2
37	+++++++	S	**78** 40	+++++++	B2
38 3,5-Dihydroxy-4′-*O*-methoxystilbene	+++++++	T	**79** 41	+++++++	B2
39 *p*-Hydroxybenzoic acid	+++++++	U	**80** 2′-*O*-Feruloylaloesin	+++++++	C3
40 *o*-Hydroxybenzoic acid	+++++++	U	**81** Barbaloin	+++++++	C3
41 Cysteine	++−++−+	V			

[a] The molecular structures of these tyrosinase inhibitors are given in Supporting Information (see Table 2).

[b] Results of the classification of compounds in this set: (i) classification of each compound using the obtained models with the Dragon descriptors in the following order: Eqs. (2)−(8).

[c] References taken from the literature: [A]Bubacco, L.; van Gastel, M.; Groenen, E. J. J.; Vijgenboom, E.; Canters, G. W. *J. Biol. Chem.* **2003**, *278*, 7381−7389. [B]van Gastela, M.; Bubaccob, L.; Groenena, E. J. J.; Vijgenboomc, E.; Cantersc, G. W. *FEBS Lett.* **2000**, *474*, 228−232. [C]Gasowskaa, B.; Kafarskia, P.; Wojtasek, H. *Biochim. Biophys. Acta* **2004**, *1673*, 170−177. [D]http://open.cacb.org.tw/index.php (2005-03-03 09:09:51). [E]Takahashi, H.; Parsons, P. G. *J. Invest. Dermatol.* **1992**, *98*, 481−487. [F]Kubo, I.; Niheia, K.; Tsujimoto, K. *Bioorg. Med. Chem.* **2004**, *12*, 5349−5354. [G]Nihei, K.-I.; Yamagiwa, Y.; Kamikawab, T.; Kubo, I. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 681−683. [H]Um, S.-J.; Park, M.-S.; Park, S.-H.; Han, H.-S.; Kwonb, Y.-J.; Sin, H.-S. *Bioorg. Med. Chem.* **2003**, *11*, 5345−5352. [I]http://www.thecosmeticsite.com/formulating/959621.htlm (April 2000). [J]Curto, E. V.; Kwong, C.; Hermersdorfer, H.; Glatt, H.; Santis, C.; Virador, V.; Hearing, V. J.; Dooley, T. P. *Biochem. Pharmacol.* **1999**, *57*, 663−672. [K]Wood, J. M.; Schallreuter-Wood, K. U.; Lindsey, N. J.; Callaghan, S.; Gardner, M. L. G *Biochem. Biophys. Res. Commun.* **1995**, *206*, 480−485. [L]Hori, I.; Nihei, K.-I.; Kubo, I. *Phytother. Res.* **2004**, *18*, 475−479. [M]Naish-Byfield, S.; Cooksey, C. J.; Riley, P. A. *Biochem. J.* **1994**, *304*, 155−162. [N]Nazzaro-Porro, M.; Passi, S. *J. Invest. Dermatol.* **1978**, *71*, 205−208. [O]Sharma, V. K.; Choi, J.; Sharma, N.; Choi, M.; Seo, S.-Y. *Phytother. Res.* **2004**, *18*, 841−844. [P]Kang, H. S.; Choi, J. H.; Cho, W. K.; Park, J. C.; Choi, J. S. *Arch. Pharm. Res.* **2004**, *7*, 742−750. [Q]Sakuma, K.; Ogawa, M.; Sugibayashi, K.; Yamada, K.; Yamamoto, K. *Arch. Pharm. Res.* **1999**, *4*, 335−339. [R]Lovstad, R. A. *Biochem. Pharmacol.* **1976**, *25*, 533−535. [S]Kubo, I.; Kinst-Hori, I.; Yokokawa, Y. *J. Nat. Prod.* **1994**, *57*, 545−551. [T]Regev-Shoshani, G.; Shoseyov, O.; Bilkis, I.; Kerem, Z. *Biochem. J.* **2003**, *374*, 157−163. [U]Bernard, P.; Berthon, J.-Y. *Int. J. Cosmetic Sci.* **2000**, *22*, 219−226. [V]Imada, C.; Sugimoto, Y.; Makimura, T.; Kobayashi, T.; Hamada, N.; Watanabe, E. *Fish. Sci.* **2001**, *67*, 1151−1156. [W]Espín, J. C.; Wichers, H. J. *Biochim. Biophys. Acta* **2001**, *1544*, 289−300. [X]Fuller, B. B.; Drake, M. A.; Spaulding, D. T.; Chaudry, F. *J. Invest. Dermatol.* **2000**, *114*, 268−276. [Y]Borojerdi, S. S.; Haghbeen, K.; Karkhane, A. A.; Fazli, M.; Sabouryc, A. A. *Biochem. Biophys. Res. Commun.* **2004**, *314*, 925−930. [Z]Kong, K.-H.; Hong, M.-P.; Choi, S.-S.; Kim, Y.-T.; Cho, S.-H. *Biotechnol. Appl. Biochem.* **2000**, *31*, 113−118. [A1]Kubo, I.; Chen, Q.-X.; Nihei, K.-I. *Food Chem.* **2003**, *81*, 241−247. [B2]Shimizu, K.; Kondo, R.; Sakai, K. *Planta Med.* **2000**, *66*, 11−15. [C3]Yagi, A.; Kanbara, T.; Morinobu, N. *Planta Med.* **1987**, 515−517.

tyrosinase inhibitors, we chose 81 compounds, whose names are depicted in Table 7. These organic-chemicals were reported to be active compounds from the medicinal chemistry literature (see the last column of Table 7). The molecular structures of these compounds are shown in Table 7 of Supporting Information.

In the first instance, a *k*-NNCA was realized to observe the molecular variability in the data set of the virtual screening. As can be seen in the dendrogram of Fig. 4, there are many different subsystems showing the great molecular diversity of the selected chemicals in this set.

The results of the classification of the compounds in the virtual screening (external set) are summarized in Table 7. At the same time, the values of the $\Delta P\%$ (posterior classification probabilities) and canonical scores of the compounds using all the developed models are given in Table 8 of Supporting Information. All screened chemicals included in this ''simulated'' virtual screening experiment were well classified as active for the best LDA-based QSAR model developed with the topological descriptors (Eq. (3)). The verification of the predictions carried out by all the obtained models comes from the recent reports in the literature from where these compounds were selected (see the last column of Table 7).

After these good results, a next step that should be done is, the inclusion of these 'novel' compounds in the training set, and carrying out the new models to find novel discrimination functions. This new model can be significantly different from the previous one, due to the inclusion of a new structural pattern, but it should be able to recognize a greater number of compounds such as tyrosinase inhibitors. Therefore, this iterative process can improve the quality of the classification models in which a great quantity of compounds with novel structural features is evaluated against the activity of the enzyme.

Several drugs were identified by the discrimination models as possible tyrosinase inhibitors. This result is the most important validation for the models developed here, because we have demonstrated that they are able to detect a series of drugs as active and these chemicals have shown the predicted activity. The drugs with some pharmacological uses selected as

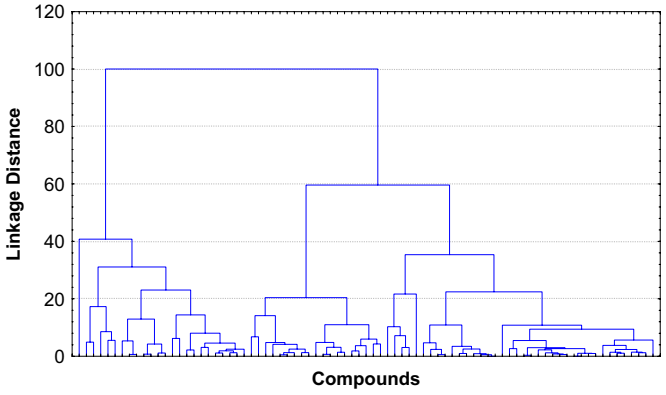

Fig. 4. A dendrogram illustrating the results of the hierarchical *k*-NNCA of the set of active chemicals used for evaluating the predictive ability of the QSAR models for ligand-based virtual screening.

Table 8
Results of ligand-based *in silico* screening and tyrosinase inhibitory activities of new bipiperidine series

Compound[a]	$\Delta P\%$[b]	Scores[b]	$\Delta P\%$[c]	Scores[c]	$\Delta P\%$[d]	Scores[d]	$\Delta P\%$[e]	Scores[e]	$\Delta P\%$[f]	Scores[f]	$\Delta P\%$[g]	Scores[g]	$\Delta P\%$[h]	Scores[h]	$IC_{50} \pm SEM$[i] (μM)
BP1	100	−3.88	100	−8.65	−19.72	0.48	45.76	−0.84	100	−5.08	−96.85	−0.92	87.78	1.67	110.79 ± 0.0583
BP2	100	−4.90	100	−8.48	−87.21	−0.09	99.66	−2.06	100	−5.29	94.98	1.76	87.02	1.64	29.94 ± 0.4289
BP3	100	−4.89	100	−8.47	99.94	2.59	99.57	−2.01	100	−5.88	99.72	2.76	86.98	1.64	6.64 ± 0.3529
BP4	100	−4.45	100	−8.45	95.03	1.50	97.93	−1.65	100	−5.81	99.51	2.56	88.52	1.70	1.72 ± 0.0438
BP5	100	−7.25	100	−7.96	−79.12	0.05	99.99	−2.89	100	−5.22	78.04	1.22	87.01	1.64	18.08 ± 0.3494
BP6	100	−6.08	100	−8.45	99.76	2.26	99.99	−2.93	100	−5.26	97.15	1.95	80.06	1.44	8.76 ± 0.1186
BP7	100	−4.65	100	−8.41	80.78	1.14	64.56	−0.97	100	−5.58	99.24	2.41	87.71	1.67	19.52 ± 0.0003

[a] The molecular structures of these chemicals are shown in Fig. 5.
[b,c,d,e,f,g,h] $\Delta P\% = [P(\text{Active}) - P(\text{Inactive})]100$ as well as canonical scores of each compound in this set: (i) classification of each compound using the obtained models with the Dragon descriptors in the following order: Eqs. (2)–(8).
[i] IC_{50} is the 50% inhibitory concentrations against the enzyme tyrosinase and SEM is the standard error of the mean.

new lead tyrosinase inhibitors have well-established methods of synthesis as well as toxicological, pharmacodynamical and pharmaceutical behaviours are also well known.

3.4. In silico novel tyrosinase inhibitors and experimental results

In the following section, and taking into account all the above steps describe in past sections, we were conducted to explore the ability of our discriminant models to find novel compounds. Besides, the good results in the algorithm presented encouraged us to carry out an *in silico* screening to search novel active compounds not described yet in the literature as tyrosinase inhibitors.

As previously indicated, one of our research teams has focused mainly on trial–error searching for new tyrosinase inhibitors [9,11,12]. At the same time, we are also identifying new drug candidates using computational screening (based on QSAR techniques). For that reason, we perform *in silico* assays for bipiperidine series synthesized and structural characterized, searching novel tyrosinase inhibitors by using the discriminant functions obtained through the Dragon descriptors and LDA technique.

The LDA-based QSAR models were used to evaluate seven compounds and in order to corroborate the predictions were prepared with excellent yields by very economic and simple methods, and evaluated *in vitro* against tyrosinase enzyme. In Table 8, the $\Delta P\%$ values of the compounds in this series, as well as their canonical scores using all the developed models, are given. From these results, we can conclude that the current approach is a suitable alternative for the selection/identification of novel tyrosinase inhibitors which may be used to prevent or treat pigmentation disorders.

A very good coincidence among the theoretical predictions and the observed activity for all the compounds is observed. In the study of the inhibitory activity, all seven compounds showed effectiveness in mushroom tyrosinase inhibition (see Table 8). Compound **BP1** ($IC_{50} = 110.79$ μM) showed mild inhibition against the enzyme, compounds **BP2** ($IC_{50} = 29.94$ μM), **BP5** ($IC_{50} = 18.08$ μM), and **BP7** ($IC_{50} = 19.52$ μM) exhibited pronounced activity when compared with kojic acid, a tyrosinase inhibitor reference. The remaining compounds, **BP3** ($IC_{50} = 6.64$ μM), and **BP6** ($IC_{50} = 8.76$ μM) had more potent activity than kojic acid

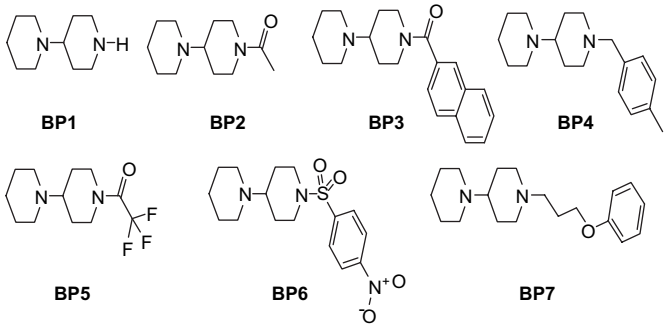

Fig. 5. Molecular structure of the new bipiperidine series.

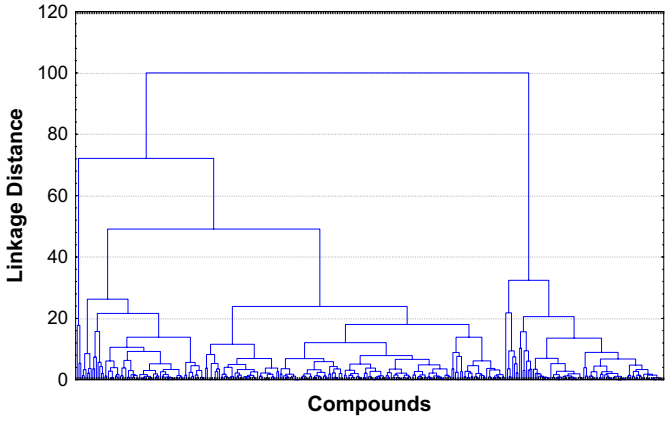

Fig. 6. A dendrogram illustrating the results of the hierarchical *k*-NNCA of the set of all active chemicals (tyrosinase inhibitors) included in training, test, virtual screening and new active bipiperidine series discovery in the present work.

($IC_{50} = 16.67$ μM) but less than L-mimosine ($IC_{50} = 3.68$ μM), another standard tyrosinase inhibitor. Finally, we want to highlight the case of compound **BP4** ($IC_{50} = 1.72$ μM) with a very potent activity against the enzyme, even compared with the reference drugs. In Fig. 5 are shown the structures of the bipiperidine compounds.

A *k*-NNCA for all the active compounds included in the training, test, virtual screening sets and the novel chemicals was carried out. This hierarchical cluster analysis was developed to compare similarities between new discovered active compounds and the complete active data set. The dendrogram illustrates the great diversity of subsystems in the complete data under investigation (see Fig. 6). An exhaustive analysis of each cluster showed that these new compounds were included in many clusters.

The principal impact of these models developed here is their capability to recognize new tyrosinase inhibitors. This is one of the major goals and can be considered as a very promising tool for the future design of new compounds with higher tyrosinase activity. In this sense, compound **BP4** presented more potent effect in the inhibition against the enzyme than L-mimosine (reference drug) and is available consider this organic-chemical as a *hit* for drug-discovery. The identification of novel structural *subsystems* can be made in search of drug-like compounds with such activity, after examining the pharmacological, toxicity, pharmacokinetic properties and good activity in clinical animal assays. Finally, it is important to remark that our aim in this study is to show how the models can be used for potential drug discovery.

4. Conclusions

The melanogenesis disorders, hyperpigmentation and other skin diseases are related to the tyrosinase. This enzyme has become a useful target for the discovery of new tyrosinase inhibitors due to the broad applications in many fields [1–7]. The areas of pharmaceutical, cosmetic, agricultural sciences have

focus in the tyrosinase inhibitors' field research due to the usefulness of this kind of compounds.

However, the cost associated to drug discovery make it slow; for that reason, the implementation of more rational search methodologies is recommended. In this case, the computational tools can aid us to speed up the assaying of drug-like compounds. These more efficient strategies such as vHTS (virtual High-Throughput Screening) can be used in complement with the QSAR models in the virtual assays, and the costs can be reduced in all terms of massive screening [62,63].

In this sense, and knowing that most of the tyrosinase inhibitors described in the literature until today have been discovered through trial–error methods, we have shown the biological *in silico* evaluation with QSAR models of new compounds synthesized and structural characterized.

Besides, we presented the application of the Dragon descriptors to the rational selection of new active compounds against the tyrosinase enzyme. The usefulness to discriminate novel active compounds from inactive ones as tyrosinase inhibitors is depicted. This classification functions obtained were applied to pools of chemicals in simulated virtual screening of compounds with the activity under study exhibiting good results. Active database presented here, can be considered useful for the entire scientist community in the natural-product, theoretical, synthesis chemistry area and others related to the field of tyrosinase inhibitor researches.

The molecular descriptors are becoming an attractive tool for efficient drug design process. Its usefulness is proven here in an experimental screening of novel bipiperidine series using pattern recognition techniques (LDA). The *in vitro* assays of the synthesized and characterized compounds were done to corroborate the *in silico* results. Seven new chemicals exhibited anti-tyrosinase activity, proving that the algorithm presented can constitute a step forward in the search of new structural features with the activity. In this way, and looking for more efficient ways to discover new potent-selective tyrosinase inhibitors which may be used to prevent or treat pigmentation disorders, can be said that, predictive *in silico* models could be used for drug target identification, accelerating the selection process of lead compounds [64].

Acknowledgements

One of the authors (Y.M.-P.) thanks the program '*Estades Temporals per a Investigadors Convidats*' for a fellowship to work at Valencia University (2006–2007). Y.M.-P. also thanks the Generalitat Valenciana, (Spain) for partial financial support as well as support from Spanish MEC (Project Reference: SAF2006-04698). M.T.H.K. is the recipient of a grant from MCBN-UNESCO (grant no. 1056), and fellowships from CIB (Italy) and Associasione Veneta per la Lotta alla Talassemia (AVTL, Italy). F.T. acknowledges financial support from the Spanish MEC DGI (Project No. CTQ2004-07768-C02-01/BQU) and Generalitat Valenciana (DGEUI INF01-051 and INFRA03-047, and OCYT GRUPOS03-173).

Appendix A. Supporting information

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, chemistry and data analysis of the obtained chemicals are available free of charge via Internet at doi:10.1016/j.ejmech.2007.01.026.

References

[1] S. Briganti, E. Camera, M. Picardo, Pigment Cell Res. 16 (2003) 101–110.
[2] A. Sanchez-Ferrer, J.N. Rodriguez-Lopez, F. Garcia-Canovas, F. Garcia-Carmona, Biochim. Biophys. Acta 1247 (1995) 1–11.
[3] L. Vamos-Vigyazo, Crit. Rev. Food Sci. Nutr. 15 (1981) 49–127.
[4] E. Frenk (Ed.), Treatment of Melasma with Depigmenting Agents, Martin Dunitz Ltd., London, 1995.
[5] Y. Xu, A.H. Stokes, W.M. Freeman, S.C. Kumer, B.A. Vogt, K.E. Vrana, Mol. Brain Res. 45 (1997) 159–162.
[6] J.A. Prezioso, M.W. Epperly, N. Wang, W.D. Bloomer, Cancer Lett. 63 (1992) 73–79.
[7] T.P. Dooley, J. Dermatol. Treat. 7 (1997) 188–200.
[8] I. Kubo, I. Kinst-Hori, Y. Kubo, Y. Yamagiwa, T. Kamikawa, H. Haraguchi, J. Agric. Food Chem. 48 (2000) 1393–1399.
[9] T. Şabudak, M.T.H. Khan, M.I. Choudhary, S. Oksuz, Nat. Prod. Res. (2006). doi:10.1080/14786410500196821.
[10] O. Nerya, J. Vaya, R. Musa, S. Izrael, R. Ben-Arie, S. Tamir, J. Agric. Food Chem. 51 (2003) 1201–1207.
[11] M.T. Khan, M.I. Choudhary, K.M. Khan, M. Rani, R. Atta Ur, Bioorg. Med. Chem. 13 (2005) 3385–3395.
[12] V.U. Ahmad, F. Ullah, J. Hussain, U. Farooq, M. Zubair, M.T. Khan, M.I. Choudhary, Chem. Pharm. Bull. (Tokyo) 52 (2004) 1458–1461.
[13] C.J. Manly, S. Louise-May, J.D. Hammer, Drug Discov. Today 6 (2001) 1101–1110.
[14] E. Estrada, E. Uriarte, A. Montero, M. Teijeira, L. Santana, E. De Clercq, J. Med. Chem. 43 (2000) 1975–1985.
[15] H. Gonzales-Diaz, Y. Marrero Ponce, I. Hernadez, I. Bastida, E. Tenorio, O. Nasco, E. Uriarte, N. Castanedo, M.A. Cabrera, E. Aguila, O. Marrero, A. Morales, M. Perez, Chem. Res. Toxicol. 16 (2003) 1318–1327.
[16] J.V. de Julian-Ortiz, C.G. de Alapont, I. Ríos-Santamarina, R. Garcia-Domenech, E. Galvez, J. Mol. Graphics Model 16 (1998) 14–18.
[17] Y. Marrero Ponce, J. Chem. Inf. Comput. Sci. 44 (2004) 2010–2026.
[18] I. Kubo, Q. Chen, K. Nihei, Food Chem. 81 (2003) 241–247.
[19] K. Shimizu, R. Kondo, K. Sakai, Planta Med. 66 (2000) 11–15.
[20] W. Li, I. Kubo, Bioorg. Med. Chem. 12 (2004) 701–713.
[21] G.M. Casanola-Martin, M.T. Khan, Y. Marrero-Ponce, A. Ather, M.N. Sultankhodzhaev, F. Torrens, Bioorg. Med. Chem. Lett. 16 (2006) 324–330.
[22] Y. Marrero-Ponce, M.T.H. Khan, G.M. Casañola-Martín, A. Ather, M.N. Sultankhodzhaev, F. Torrens, R. Rotondo, Y.J. Alvarado, QSAR Comb. Sci. 27 (2007) 469–487.
[23] Y. Marrero-Ponce, M.T.H. Khan, G.M. Casañola-Martín, A. Ather, M.N. Sultankhodzhaev, F. Torrens, R. Rotondo, Chem. Med. Chem. 2 (2007) 449–478.
[24] M.P. Gonzalez, C. Teran, M. Teijeira, P. Besada, M.J. Gonzalez-Moa, Bioorg. Med. Chem. Lett. 15 (2005) 3491–3495.
[25] M. Fernandez, J. Caballero, A.M. Helguera, E.A. Castro, M.P. Gonzalez, Bioorg. Med. Chem. 13 (2005) 3269–3277.
[26] M. Negwer (Ed.), Organic-Chemical Drugs and Their Synonyms, Akademie-Verlag, Berlin, 1987.
[27] K. Komori, K. Yatagai, T. Tatsuma, J. Biotechnol. 108 (2004) 11–16.
[28] H. Kim, J. Choi, J.K. Cho, S.Y. Kim, Y.S. Lee, Bioorg. Med. Chem. Lett. 14 (2004) 2843–2846.

[29] M.I. Choudhary, S.G. Musharraf, M.T.H. Khan, D. Abdelrahman, M. Parvez, F. Shaheen, Atta-ur-Rahman, Helv. Chim. Acta 86 (2003) 3450–3460.
[30] S.J. Um, M.S. Park, S.H. Park, H.S. Han, Y.J. Kwon, H.S. Sin, Bioorg. Med. Chem. 11 (2003) 5345–5352.
[31] M. Shiino, Y. Watanabe, K. Umezawa, Bioorg. Med. Chem. 9 (2001) 1233–1240.
[32] M. Shiino, Y. Watanabe, K. Umezawa, Bioorg. Chem. 31 (2003) 129–135.
[33] J.K. No, D.Y. Soung, Y.J. Kim, K.H. Shim, Y.S. Jun, S.H. Rhee, T. Yokozawa, H.Y. Chung, Life Sci. 65 (1999) PL241–PL246.
[34] E.V. Curto, C. Kwong, H. Hermersdorfer, H. Glatt, C. Santis, V. Virador, V.J. Hearing Jr., T.P. Dooley, Biochem. Pharmacol. 57 (1999) 663–672.
[35] Y.M. Kim, J. Yun, C.K. Lee, H. Lee, K.R. Min, Y. Kim, J. Biol. Chem. 277 (2002) 16340–16344.
[36] J.P. Ley, H.J. Bertram, Bioorg. Med. Chem. 9 (2001) 1879–1885.
[37] R. Todeschini, V. Consonni, M. Pavan, Dragon Software version 2.1. (2002).
[38] R. Todeschini, V. Consonni (Eds.), Handbook of Molecular Descriptors, Wiley-VCH, Germany, 2000.
[39] StatSoft Inc., STATISTICA (Data Analysis Software System).www.statsoft.com (2001).
[40] J.W. Mc Farland, D.J. Gans, in: H. Waterbeemd (Ed.), Chemometric Methods in Molecular Design, VCH Publishers, New York, 1995, pp. 295–307.
[41] R.A. Johnson, D.W. Wichern (Eds.), Applied Multivariate Statistical Analysis, Prentice-Hall, New Jersey, 1988.
[42] A. Baldi, E. Dragonetti, T. Battista, A.M. Groeger, V. Esposito, G. Baldi, D. Santini, Anticancer Res. 20 (2000) 3923–3928.
[43] E. Estrada, S. Vilar, E. Uriarte, Y. Gutierrez, J. Chem. Inf. Comput. Sci. 42 (2002) 1194–1203.
[44] M. Randić, J. Chem. Inf. Comput. Sci. 31 (1991) 311–320.
[45] M. Randić, New J. Chem. 15 (1991) 517–525.
[46] M. Randić, J. Mol. Struct. (Theochem.) 233 (1991) 45–59.
[47] B. Lučić, S. Nikolić, N. Trinajstić, D. Jurić, J. Chem. Inf. Comput. Sci. (1995) 532–538.
[48] D.J. Klein, M. Randić, D. Babić, B. Lučić, S. Nikolić, N. Trinajstić, Int. J. Quantum Chem. 63 (1997) 215–222.
[49] M.-G. Ford, D.-W. Salt, in: H. van de Waterbeemd (Ed.), Chemometric Methods in Molecular Design, VCH Publishers, New York, 1995, pp. 283–292.
[50] K.M. Khan, Z.S. Saify, M.T. Khan, N. Butt, G.M. Maharvi, S. Perveen, N. Ambreen, M.I. Choudhary, R. Atta Ur, C.T. Supuran, J. Enzyme Inhib. Med. Chem. 20 (2005) 401–407.
[51] V.J. Hearing (Ed.), Methods in Enzymology, Academic Press, New York, 1987.
[52] E. Estrada, A. Peña, Bioorg. Med. Chem. 8 (2000) 2755–2770.
[53] Y. Marrero-Ponce, R. Medina-Marrero, Y. Martinez, F. Torrens, V. Romero-Zaldivar, E.A. Castro, J. Mol. Mod. 12 (2006) 255–271.
[54] D.H. Rouvray, in: T. Balaban (Ed.), Chemical Applications of Graph Theory, Academic Press, London, 1976, pp. 180–181.
[55] N. Trinajstić (Ed.), Chemical Graph Theory, CRC Press, Boca Raton, FL, 1992.
[56] E. Estrada, J. Chem. Inf. Comput. Sci. 35 (1995) 31–33.
[57] E. Estrada, A. Ramirez, J. Chem. Inf. Comput. Sci. 36 (1996) 837–843.
[58] S. Wold, L. Erikson, in: H. van de Waterbeemd (Ed.), Chemometric Methods in Molecular Design, VCH Publishers, New York, 1995, pp. 309–318.
[59] A. Golbraikh, A. Tropsha, J. Mol. Graphics Model 20 (2002) 269–276.
[60] J. Gálvez, R. García, M.T. Salabert, R. Soler, J. Chem. Inf. Comput. Sci. 34 (1994) 520–525.
[61] S. Whitebread, J. Hamon, D. Bojanic, L. Urban, Drug Discov. Today 10 (2005) 1421–1433.
[62] J. Xu, A. Hagler, Molecules 7 (2002) 566–700.
[63] H.J.M. Seifert, K. Wolf, D. Vitt, Biosilico. 1 (2003) 143–149.
[64] C. Watson, Biosilico. 1 (2003) 83.