



The influence of R and S configurations of a series of amphetamine derivatives on quantitative structure–activity relationship models[☆]



Maíra A.C. Fresqui^{a,*}, Márcia M.C. Ferreira^{b,2}, Milan Trsic^{a,1}

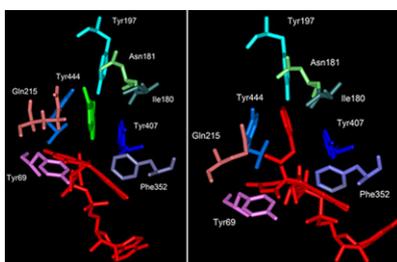
^a Institute of Chemistry of São Carlos, University of São Paulo, Av. Trabalhador São-carlense, 400, POB 780, 13560-970 São Carlos, SP, Brazil

^b Institute of Chemistry, University of Campinas – UNICAMP, POB 6154, 13083-970 Campinas, SP, Brazil

HIGHLIGHTS

- ▶ The QSAR model is not dependent of ligand conformation.
- ▶ Amphetamines were analyzed by quantum chemical, steric and hydrophobic descriptors.
- ▶ CHELPG atomic charges on the benzene ring are one of the most important descriptors.
- ▶ The PLS models built were extensively validated.
- ▶ Manual docking supports the QSAR results by pi–pi stacking interactions.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 10 August 2012

Received in revised form 25 October 2012

Accepted 2 November 2012

Available online 15 November 2012

Keywords:

Quantitative structure–activity relationship

Amphetamine derivatives

Electronic

Hydrophobic and steric descriptors

R and S configurations of amphetamines

ABSTRACT

Chiral molecules need special attention in drug design. In this sense, the R and S configurations of a series of thirty-four amphetamines were evaluated by quantitative structure–activity relationship (QSAR). This class of compounds has antidepressant, anti-Parkinson and anti-Alzheimer effects against the enzyme monoamine oxidase A (MAO A). A set of thirty-eight descriptors, including electronic, steric and hydrophobic ones, were calculated. Variable selection was performed through the correlation coefficients followed by the ordered predictor selection (OPS) algorithm. Six descriptors (CHELPG atomic charges C3, C4 and C5, electrophilicity, molecular surface area and log *P*) were selected for both configurations and a satisfactory model was obtained by PLS regression with three latent variables with $R^2 = 0.73$ and $Q^2 = 0.60$, with external predictability $Q^2 = 0.68$, and $R^2 = 0.76$ and $Q^2 = 0.67$ with external predictability $Q^2 = 0.50$, for R and S configurations, respectively. To confirm the robustness of each model, leave-*N*-out cross validation (LNO) was carried out and the *y*-randomization test was used to check if these models present chance correlation. Moreover, both automated or a manual molecular docking indicate that the reaction of ligands with the enzyme occurs via pi–pi stacking interaction with Tyr407, inclined face-to-face interaction with Tyr444, while aromatic hydrogen–hydrogen interactions with Tyr197 are preferable for R instead of S configurations.

© 2012 Elsevier B.V. All rights reserved.

[☆] Paper presented at the XIII Conference on Chemometrics in Analytical Chemistry (CAC 2012), Budapest, Hungary, 25–29 June 2012.

* Corresponding author. Present address: Theoretical and Applied Chemometrics Laboratory, Institute of Chemistry, University of Campinas – UNICAMP, Zip Box: 6154, 13083-970 Campinas, SP, Brazil. Tel.: +55 19 35213102; fax: +55 19 35213023.

E-mail addresses: maira@iqsc.usp.br, mairacarvalho@gmail.com (M.A.C. Fresqui), marcia@iqm.unicamp.br (M.M.C. Ferreira), cra612@gmail.com (M. Trsic).

¹ Tel.: +55 16 3373 9946; fax: +55 16 3373 9903.

² Tel.: +55 19 35213102; fax: +55 19 35213023.

1. Introduction

The amphetamine family is the most common group of central nervous system stimulant drugs. They inhibit the monoamine oxidase enzyme (MAO, EC 1.4.3.4, isoforms A and B), a flavoenzyme that catalyzes the oxidation of biogenic amines, increasing levels of neurotransmitters like norepinephrine, serotonin (5-HT) and dopamine in the brain [1,2]. The inhibition of this enzyme has been widely studied for application as a clinical treatment of neuropathological disorders such as Parkinson's and Alzheimer's diseases and depression [3,4].

For many years, efforts have been directed toward understanding features of the enzyme, such as its substrate binding site and chemical mechanism of action [4]. Thus, these features may be useful for the discovery of new potent and selective MAO inhibitors. Different classes of drugs have been studied by several authors worldwide to design new MAO ligands, like coumarin [5], indol [6], pyrrole [7], and amphetamine [8,9] derivatives. Such ligands differ in their reversibility and selectivity with respect to the substrate, and potency levels. Whereas MAO is the target for a number of clinically used drug inhibitors [10], both experimental and computational efforts have been undertaken with the aim to identify a new class of potent and selective compounds. Several theoretical works, including quantitative structure–activity relationship (QSAR) studies of MAO inhibitors, have been performed for different classes of compounds [5–9]. However, there are only few studies relating the electronic structure of amphetamine derivatives and their configuration to their antidepressant effects or for Parkinson's or Alzheimer's diseases.

Theoretical studies of active compounds in terms of frontier orbital energies, the energy of the highest occupied molecular orbital (HOMO) and the energy of the lowest unoccupied molecular orbital (LUMO), have been employed for many years in QSAR [11]. However, the limitations of the HOMO–LUMO approach for electron transfer was previously pointed out by Fukui [12,13]. The new strategy of effective reactive-orbital (effective-for-reaction molecular orbital, so-called, FERMO) energies developed by da Silva et al. [14] is pointing to better results than the HOMO energy approach when applied to biomolecules [15].

The goal of QSAR methodology is to build a regression model for a training set using, for example, structural, steric and electronic parameters, commonly known as descriptors. This mathematical relationship, after being validated by statistical methods and chemical intuition, is used to predict the biological activity of new compounds. It is useful in understanding and explaining the mechanism of drug action at the molecular level, providing some insights for the design of new compounds with desirable biological properties [16].

In the present work, a QSAR study of a set of thirty-four amphetamine derivatives (Fig. 1) is presented. The model is based on electronic quantum chemical, hydrophobic and steric descriptors. The dependence on molecular configurations in the PLS

models, i.e., the R and S amphetamine derivatives, are investigated and discussed. The choice of these parameters as descriptors was based on the electronic characteristics of interaction between this class of ligand and the enzyme, as described elsewhere [9]. The biological activity data were from the literature [8,17,18].

Protein binding sites exhibit highly selective recognition of other molecules, such as new ligands. This has been used in the design of new selective molecules to modify the target property. This is possible when the X-ray crystallography receptor is available and by applying, for example, docking techniques. Thus, two different docking approaches were performed between the amphetamine molecule and the MAO A enzyme (2Z5Y PDB code) [19]. These approaches could indicate whether the R or S configuration difference is relevant for its activity, according to QSAR models.

2. Materials and methods

2.1. Chemical structure database and biological activity

The biological activities of isoform A of MAO (MAO A) inhibitors extracted from the literature [8,17,18], were studied using a crude rat brain mitochondrial suspension. IC_{50} values were obtained from plots of inhibition percentages using the methodology described by Scorza et al. [8], Hurtado-Guzman et al. [17] and Sterling et al. [18]. The reported experimental values of biological activity do not bring any information about the stereochemistry of these compounds, so the same value was used for both the R and S configurations, of the compound. Although the measurements were performed using similar techniques, some caution is appropriate to guarantee that the data are comparable. In this sense, it is relevant that the biological activity of one compound, the selegiline molecule, has the same reported IC_{50} value.

The experimental IC_{50} values in $\mu\text{M L}^{-1}$ concentrations were converted into M L^{-1} concentrations and later into their corresponding pIC_{50} ($-\log IC_{50}$) values which are listed in Tables 1 and 2, this unit transformation guarantees that all biological activities have a positive value. Compounds that were described as inactive (IC_{50} higher than $100 \mu\text{M L}^{-1}$ or with values not provided) are not appropriate for a quantitative study and were not included in this work. However, those molecules that have IC_{50} above $100 \mu\text{M L}^{-1}$ and had their biological activity described quantitatively, despite being considered inactive, were included in the model. The biological activities are well distributed within the considered range of pIC_{50} from 3.62 to 7.00 log units, as shown in the histogram of Fig. 2.

Since most of the compounds have one chiral carbon atom, both conformers, R and S, were analyzed. The S isomer is similar to the serotonin molecule, the natural substrate of MAO A. However, according to the results published by Vallejos et al. [9], the R form was also considered. For compounds that had their crystallographic structure determined, such as MDMA [20], or had

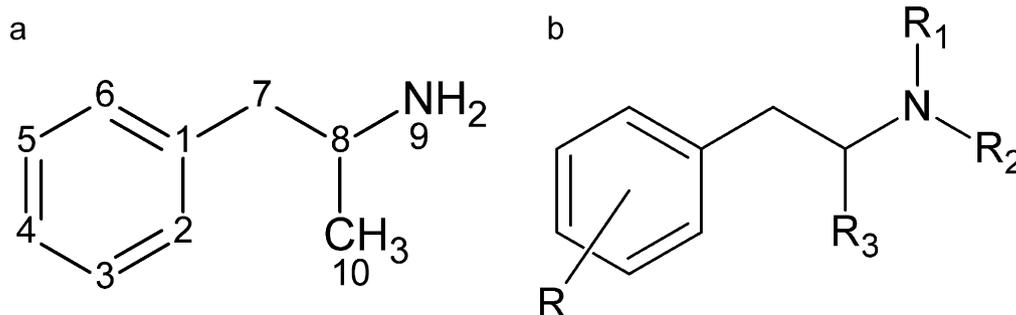
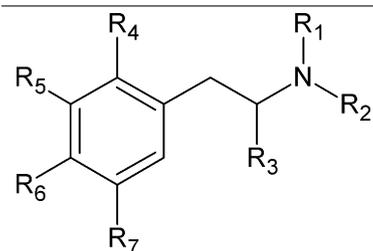


Fig. 1. (a) Structure of the amphetamine molecule, (b) general structure of amphetamine derivatives. The list of substituents is shown in Tables 1 and 2.

Table 1

Biological activities (MAO A) for compounds synthesized by Scorza [8] (Amphetamine – MDMA) and Hurtado-Guzmán [17] (NMMTA – MTAB).



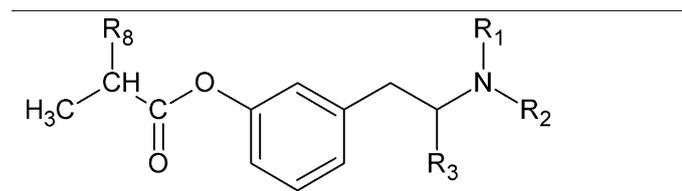
Compound	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	R ₇	pIC ₅₀ MAO A
Amphetamine	H	H	CH ₃	H	H	H	H	4.96
PCA	H	H	CH ₃	H	H	Cl	H	5.40
(+)-MTA	H	H	CH ₃	H	H	SCH ₃	H	6.96
(-)-MTA	H	H	CH ₃	H	H	SCH ₃	H	5.69
ETA	H	H	CH ₃	H	H	SCH ₂ CH ₃	H	7.00
ITA	H	H	CH ₃	H	H	SCH(CH ₃) ₂	H	6.40
4-EtOA	H	H	CH ₃	H	H	OCH ₂ CH ₃	H	6.70
4-MetOA	H	H	CH ₃	H	H	OCH ₃	H	6.52
3,4-DMA	H	H	CH ₃	H	OCH ₃	OCH ₃	H	4.70
Amiflamine	H	H	CH ₃	CH ₃	H	N-(CH ₃) ₂	H	5.70
5Br-2,4-DMA	H	H	CH ₃	OCH ₃	H	OCH ₃	Br	4.89
2,4-DMA	H	H	CH ₃	OCH ₃	H	OCH ₃	H	6.22
ALEPH-2	H	H	CH ₃	OCH ₃	H	SCH ₂ CH ₃	OCH ₃	5.49
ALEPH-1	H	H	CH ₃	OCH ₃	H	SCH ₃	OCH ₃	5.29
DOB	H	H	CH ₃	OCH ₃	H	Br	OCH ₃	4.00
DOM	H	H	CH ₃	OCH ₃	H	CH ₃	OCH ₃	4.62
2Br-4,5-MDA	H	H	CH ₃	Br	H	-OCH ₂ O-		4.89
2Br-4,5-DMA	H	H	CH ₃	Br	H	OCH ₃	OCH ₃	5.03
2Cl-4,5-MDA	H	H	CH ₃	Cl	H	-OCH ₂ O-		5.2
MDA	H	H	CH ₃	H		-OCH ₂ O-	H	5.03
MDMA	H	CH ₃	CH ₃	H		-OCH ₂ O-	H	4.52
NMMTA	H	CH ₃	CH ₃	H	H	SCH ₃	H	6.05
DMMTA	CH ₃	CH ₃	CH ₃	H	H	SCH ₃	H	5.68
NEMTA	H	CH ₂ CH ₃	CH ₃	H	H	SCH ₃	H	5.74
DEMTA	CH ₂ CH ₃	CH ₂ CH ₃	CH ₃	H	H	SCH ₃	H	5.19
NPMTA	H	CH ₂ CH ₂ CH ₃	CH ₃	H	H	SCH ₃	H	5.62
NAMTA	H	CH ₂ CHCH ₂	CH ₃	H	H	SCH ₃	H	5.46
MTAB	H	H	CH ₂ CH ₃	H	H	SCH ₃	H	6.08

the biological activity determined for one specific configuration ((+/-)-4-Methylthioamphetamine (MTA)), the experimental configuration and its biological activity was used to build the PLS model.

The mechanism of MAO inhibition considered in this work for the interaction between the ligands and the enzyme active site was first proposed by Salach [21] and later revised by other research groups [4,10,22,23]. The flavin adenine dinucleotide (FAD) cofactor could react at positions N5 or C4 (Fig. 3). This reaction position will be defined by the ligand characteristics, as reported [10].

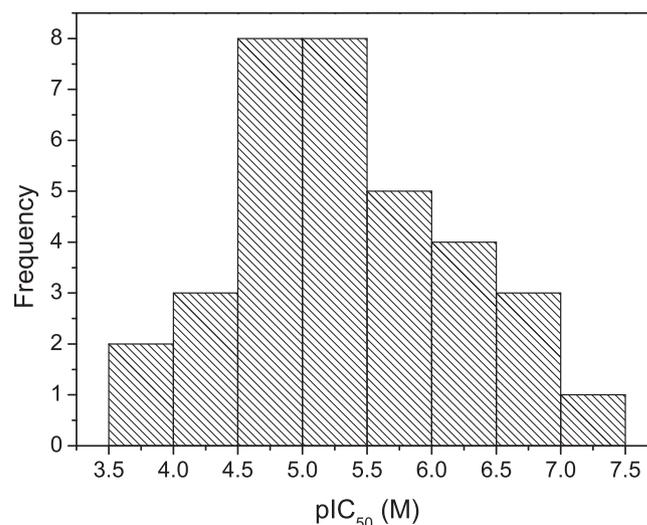
Table 2

Biological activities (MAO A) for compounds synthesized by Sterling [18].



Compound	R ₁	R ₂	R ₃	R ₈	pIC ₅₀ MAO A
45a	H	H	H	CH ₃	4.02
45b	H	H	H	CH ₂ CH ₃	3.74
45c	H	H	H	CH ₂ CH ₂ CH ₃	3.62
46a	H	CH ₃	H	CH ₃	4.66
46b	H	CH ₃	H	CH ₂ CH ₃	4.70
48a	CH ₃	CH ₃	H	CH ₃	4.29

The pharmacophore mode of reaction between the FAD and the amphetamines divides this class of molecules into two different groups following the approach of the MAO adduct, according to the reaction position in the FAD molecule. Thus, normally, the molecules which have a propargyl group bonded at N9 interact with N5 from FAD (see Figs. 1 and 3). Other substances studied in this

**Fig. 2.** Histogram showing the distribution of the biological activity IC₅₀ in log units for the studied compounds.

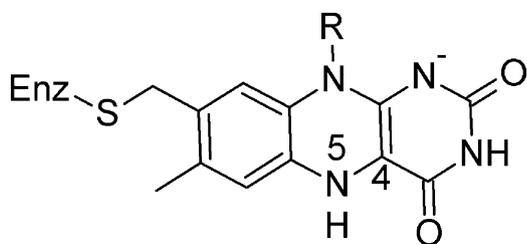


Fig. 3. Structure of the flavin adduct formations at the MAO enzyme proposed by Binda et al. [10].

work interact with C4 from FAD (Fig. 3) and their chemical formulas are presented in Tables 1 and 2.

2.2. Computational methods and chemometrics

All quantum chemical calculations were carried out using Gaussian 03 [24] software. The partition coefficient was calculated by ALOGPS [25,26] and other geometric descriptors were calculated by using the Marvin program [27]. The descriptors used in the QSAR study were analyzed by the chemometric methods of hierarchical cluster analysis, HCA [28], principal component analysis (PCA) [16,28] and partial least squares (PLS) regression [16,28] were implemented in Pirouette program version 3 [29].

The quantum chemical methods Austin model 1 (AM1) [30], the Hartree–Fock (HF) [31,32], and Density-Function-Theory (DFT) [33,34] (B3LYP functional [35,36]), both with 6-31G, 6-31G(d), 6-31G(d,p), 6-31+G(d,p) and 6-31++G(d,p) wavefunctions, and Møller–Plesset (MP2) [37] with 6-31G basis set were applied to analyze which theoretical method would be more appropriate for the molecular geometry optimization of 3,4-methylenedioxymethamphetamine (MDMA, ecstasy). After the

geometry optimization of MDMA at all the mentioned theory levels, a comparison between crystal and theoretical molecular geometries was performed. In this work, two different approaches were applied. In one, the root mean square (RMS) analysis was carried out by using HyperChem software, version 7.1 [38]. In the other, the multivariate method of PCA was applied to autoscaled bond distances and bond and dihedral angles, which formed a **X** matrix (39, 13). According to the two analyses, the HF/6-31G(d,p) method seems to be the most appropriate; thus it was employed for the complete set of molecules under investigation.

It is well known that atomic charges based on electrostatic potential are more realistic with electronegative atoms. Thus, CHELPG [39] instead of Mulliken atomic charges, as used in Vallejos et al. [9], were used in the models presented here. For the charges on carbon atoms C1–C6 (qC1–qC6) from the benzene ring, carbon C8 (qC8) and nitrogen N9 (qN9) atoms (Fig. 1) the total energy (ϵ_{TOTAL}), the HOMO, LUMO and FERMO energies (ϵ_{HOMO} , ϵ_{FERMO} and ϵ_{LUMO} , respectively), and finally, the dipole moment were calculated. HOMO and LUMO energies were also used to calculate the five descriptors: absolute softness (S), absolute hardness (η), electronic chemical potential (μ), absolute electronegativity (χ), and electrophilicity (ω) indexes described by Pearson [40] and Parr and Von Szentpaly [41]. The above mentioned indexes were recalculated using the FERMO energies instead of the HOMO energies, generating new reactivity indexes labeled as S' , η' , μ' , χ' , ω' .

One can select the FERMO orbitals through two consecutive approaches: by visual examination of the molecular orbital (MO) shapes and, further, by calculating the MO contribution on the reactive ligand atom that is supposed to drive the reaction, as described by Solomon and co-workers [42].

The steric descriptors calculated with the Marvin program were molar refractive index and six different molecular surface areas while molecular volume was calculated with the Gaussian program.

Table 3
Molecular descriptors selected for R configurations.

	qC3	qC4	qC5	Electrophilicity	XlogP3	ASA.P
2,4-DMA	-0.43	0.46	-0.37	0.013	1.75	25.66
2Br-4,5-DMA	-0.17	0.19	0.27	0.019	2.46	24.97
2Br-4,5-MDA	-0.13	-0.08	-0.04	0.024	2.18	21.34
3,4-DMA	0.33	0.18	-0.15	0.013	1.2	27.15
4-EtOA	-0.23	0.39	-0.26	0.016	2.14	17.06
4-MetOA	-0.24	0.42	-0.27	0.016	1.77	19.75
5Br-2,4-DMA	-0.46	0.48	-0.14	0.016	2.58	25.34
Amiflamine	-0.38	0.50	-0.42	0.010	2.16	10.94
DOB	-0.21	-0.03	0.36	0.019	2.58	26.03
DOM	-0.42	0.10	0.18	0.012	2.24	23.35
MDA	0.38	0.24	-0.25	0.014	1.64	48.12
PCA	-0.02	0.05	-0.06	0.021	2.43	7.88
Amphetamine	-0.05	-0.14	-0.09	0.016	1.76	7.91
MDMA	-0.23	0.24	0.38	0.013	2.15	69.14
2Cl-4,5MDA	-0.24	0.26	0.33	0.018	2.12	45.95
45a	0.43	-0.28	-0.03	0.017	1.01	44.99
45b	0.42	-0.29	-0.01	0.017	1.37	41.05
45c	0.46	-0.30	-0.02	0.017	1.9	40.33
46a	0.40	-0.24	-0.08	0.017	1.52	42.87
46b	0.42	-0.29	-0.04	0.016	1.88	38.92
48a	0.41	-0.26	-0.06	0.017	1.98	38.81
DEMTA	-0.09	0.22	-0.13	0.020	3.9	12.89
DMMTA	-0.10	0.22	-0.12	0.021	3.16	14.78
MTAB	-0.09	0.22	-0.12	0.022	2.79	16.34
NAMTA	-0.10	0.24	-0.14	0.022	3.34	15.19
NEMTA	-0.10	0.22	-0.12	0.021	3.06	15.82
NMMTA	-0.12	0.23	-0.12	0.021	2.58	15.28
NPMTA	-0.09	0.22	-0.13	0.021	3.59	15.18
(+)MTA	-0.11	0.20	-0.06	0.022	2.31	8.93
(-)MTA	-0.06	0.21	-0.11	0.022	2.31	8.93
ALEPH-1	-0.27	0.20	0.17	0.018	2.17	29.84
ALEPH-2	-0.26	0.19	0.16	0.017	2.54	27.71
ITA	-0.13	0.24	-0.06	0.021	3.11	13.61
ETA	-0.10	0.24	-0.14	0.022	2.68	15.53

Table 4
Molecular descriptors selected for S configurations.

	qC3	qC4	qC5	Eletrophilicity	XlogP3	ASA.P
2,4-DMA	-0.46	0.47	-0.36	0.013	1.75	35.7
2Br-4,5-DMA	-0.16	0.16	0.32	0.020	2.46	36.12
2Br-4,5-MDA	-0.28	0.29	0.34	0.018	2.18	70.46
3,4-DMA	0.28	0.18	-0.10	0.012	1.2	37.66
4-EtOA	-0.26	0.39	-0.23	0.016	2.14	22.98
4-MetOA	-0.27	0.42	-0.24	0.016	1.77	27.15
5Br-2,4-DMA	-0.49	0.46	-0.10	0.017	2.58	35.91
Amiflamine	-0.40	0.50	-0.36	0.010	2.16	14.95
DOB	-0.23	-0.02	0.38	0.019	2.58	36.64
DOM	-0.42	0.06	0.25	0.012	2.24	32.77
MDA	0.33	0.26	-0.21	0.013	1.64	69.89
PCA	-0.06	0.06	-0.02	0.021	2.43	8.94
Amphetamine	-0.09	-0.14	-0.05	0.016	1.76	8.94
MDMA	-0.23	0.24	0.38	0.013	2.15	69.14
2Cl-4,5MDA	-0.27	0.26	0.37	0.018	2.12	70.29
45a	0.43	-0.28	-0.03	0.017	1.01	44.99
45b	0.42	-0.29	-0.01	0.017	1.37	41.05
45c	0.46	-0.30	-0.02	0.017	1.9	40.33
46a	0.40	-0.24	-0.08	0.017	1.52	42.87
46b	0.42	-0.29	-0.04	0.016	1.88	38.92
48a	0.41	-0.26	-0.06	0.017	1.98	38.81
DEMTA	-0.12	0.22	-0.10	0.020	3.9	2.78
DMMTA	-0.12	0.22	-0.10	0.021	3.16	4.85
MTAB	-0.12	0.22	-0.09	0.022	2.79	6.75
NAMTA	-0.123	0.22	-0.09	0.022	3.34	5.96
NEMTA	-0.13	0.22	-0.10	0.021	3.06	5.95
NMMTA	-0.12	0.22	-0.10	0.021	2.58	7.93
NPMTA	-0.11	0.21	-0.09	0.021	3.59	5.26
(+)MTA	-0.11	0.20	-0.06	0.022	2.31	8.93
(-)MTA	-0.06	0.21	-0.11	0.022	2.31	8.93
ALEPH-1	-0.30	0.21	0.201	0.018	2.17	32.44
ALEPH-2	-0.29	0.20	0.18	0.017	2.54	31.86
ITA	-0.18	0.24	-0.03	0.021	3.11	8.94
ETA	-0.14	0.24	-0.10	0.022	2.68	8.93

Among them we can cite van der Waals and solvent accessible surface areas (hydrophobic surface area and polar surface area), indexed by the computer program as van der Waals, ASA, ASA+, ASA-, ASA.H and ASA.P. The partition coefficient $\log P$ was calculated by the ALOGPS program where seven different values (AlogPs, AClogP, AlogP, MlogP, KOWWIN, XlogP2 and XlogP3) were obtained.

To build a reliable QSAR model, a three-step procedure was employed for each configuration, i.e., R and S. Variable selection was first done by excluding those descriptors which showed correlation coefficients lower than 0.3 with MAO A activity. From the remaining descriptors, those highly correlated among themselves, i.e., with a correlation coefficient above 0.90, were also eliminated. In addition, descriptors whose plots versus the dependent variable did not show a uniform distribution or did show pronounced dispersion were also excluded.

Further, the ordered prediction selection method (OPS) [43] was also applied for variable selection. In this step, the regression vector was used as the informative vector and the correlation coefficient of cross-validation, Q^2 , was the criterion used to select the best models. Tables 3 and 4 contain the selected descriptors.

At the third step, the set of nine descriptors selected after variable selection was further refined using the software Pirouette to obtain an optimized model which would fulfill the criteria for being statistically significant, robust and interpretable. The Student t test and Bonferroni test were also performed on the obtained matrices, with 95% of confidence, to check if those matrices are equivalent.

The PLS [16] regression method was employed to model the relationship between the biological activity of the set of compounds and the selected descriptors. In this regression method, the

X matrix of molecular descriptors is linearly related to the **y** vector containing the biological activities (dependent variable). The number of latent variables in the model was defined by leave-one-out (LOO) cross-validation. The final model was validated by leave- N -out (LNO) cross-validations, **y**-randomization [44–47] and sign change [48]. In the LNO cross-validation procedure, N compounds ($N=2, 3, \dots, 17$) were left out from the training set. For a particular N , the data were randomized 10 times, and the average and standard deviation values for Q^2 were used. In the **y**-randomization test, the dependent variable-vector was randomly shuffled 50 times for the investigated sets and new models were built using randomized **y** and the R^2 and Q^2 values were compared with that of the true model.

Exploratory analysis of the 34 amphetamines by hierarchical cluster analysis on autoscaled data was applied to select both the training and test sets, by splitting the complete set of compounds into a training set formed by 26 molecules and a test set with the 8 remaining compounds.

Due to differences in their orders of magnitude, descriptors and biological activities (pIC_{50} in ML^{-1} of MAO A) were autoscaled, i.e., each of them were mean centered and then divided by the respective standard deviation.

Molecular docking is a useful methodology to predict molecular interactions between the ligand and the receptor. Programs are commonly used to position the small molecules into the protein binding site with reliable results. Virtual docking studies were undertaken for the most and the less potent compounds from two different approaches, by automated docking using The AutoDockVina [49] program and by manual docking, where the ligand adduct was drawn into the binding site according to similar X-ray crystallographic structure [50].

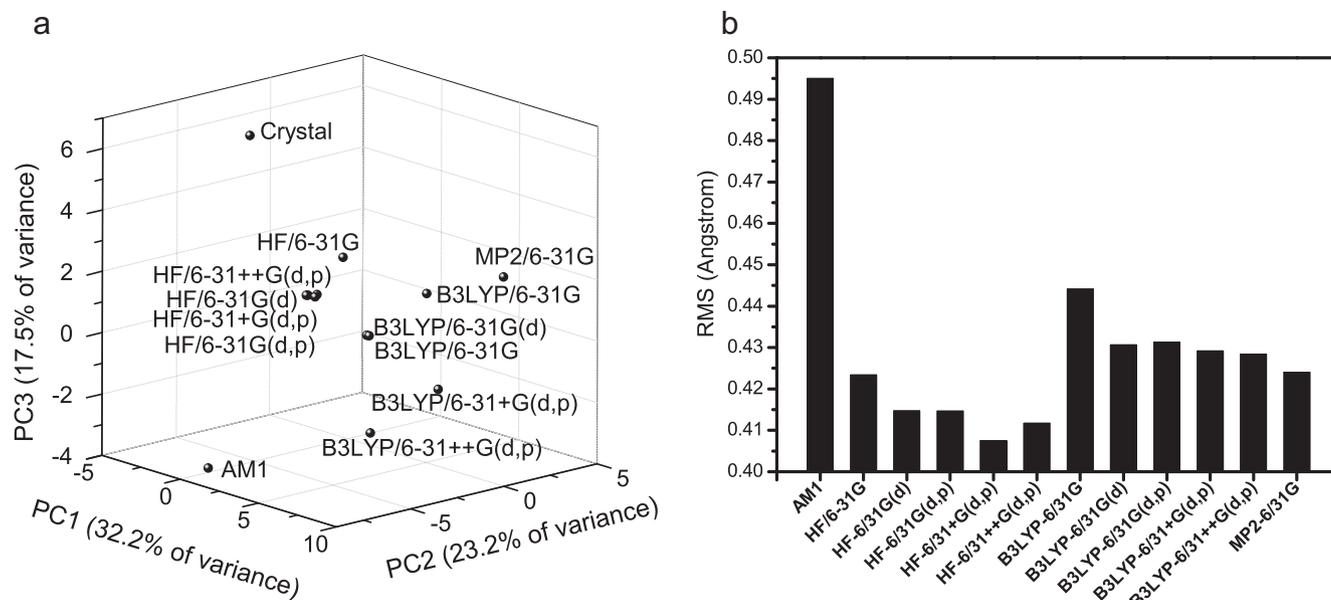


Fig. 4. (a) PC1 versus PC2 versus PC3 scores plot of structural parameters from MDMA for the crystallographic structure, calculated by 12 different quantum chemical methods, (b) RMS of the MDMA structural parameters obtained from 12 different theoretical calculations relative to the experimental crystal structure.

3. Results and discussion

3.1. Chemical structure optimizations and molecular descriptors calculations

Both analyses, PCA and RMS (Fig. 4), comparing the experimental X-ray crystallographic structure [20] reported in the literature and the theoretically obtained geometries of the MDMA molecule showed similar results.

The PC1 × PC2 × PC3 scores plot (Fig. 4a) describing 72.9% of total variance, as well as the RMS plot (Fig. 4b), show that the molecular geometry obtained from the semiempirical method AM1 is the most different from the crystal geometry. One can see in the scores plot that all geometries obtained by applying the HF method have negative PC1 values as well as the crystal structure, however, the DFT and AM1 geometries are at positive PC1 values. Fig. 4b shows that the geometries obtained from HF calculations have smaller RMS and, consequently, they are more similar to the experimental crystal structure, which is in agreement with the chemometric analysis. In the two analyses (PCA and RMS) the selection of an appropriate theoretical method and the wave function used in this work was in agreement with results from the literature [51,52]. This step is quite important to obtain the best 3D geometry of the compounds and so, more reliable descriptor values. Thus, the HF/6-31G(d,p) method was shown to be appropriate for full molecular optimization since its results were the most similar to experimental data.

Although the complete set of molecules has the same mechanism of reaction, the FERMO orbital is not the same for all compounds. The molecules without a sulfur-containing substituent show that FERMO is HOMO-2 while molecules with such a substituent have HOMO-3 as the reactive MO. It is remarkable in which way FERMO is present in the molecules, not necessarily being the same orbital for each molecule. Fig. 5 shows HOMO-3 to HOMO for two compounds: amphetamine and ETA. It is clear that HOMO-2 for amphetamine has similar electron density distributions to that of HOMO-3 for ETA. The FERMO's shapes are in perfect agreement with the shape of the reactive orbital proposed by da Silva [14].

Quantum chemical calculations resulted in the generation of 23 molecular descriptors formed by the total energy, the HOMO,

FERMO and LUMO energies, dipole moment, CHELPG atomic charges on the benzene ring and on the C8 and N9 atoms, the absolute softness, absolute hardness, electronic chemical potential, absolute electronegativity, and electrophilicity indexes calculated from HOMO and LUMO. To complete the quantum chemical descriptors pool, the S , η , μ , χ , ω indexes from FERMO energies instead of those obtained from HOMO energies, were included. Furthermore, seven different partition coefficients, molecular volumes, molecular surface areas and molar refractive indices were also calculated to complete the set of descriptors. The QSAR studies were performed on a data matrix X (34, 38) for both R and S ligand configurations.

Some descriptors that were initially thought to be important, such as dipole moments and FERMO energies, molecular volume, some molecular surface areas and molar refractive indices were eliminated for being poorly correlated to the biological activity (correlation coefficient below 0.3). After performing variable selection by the OPS method the models for both R and S configurations

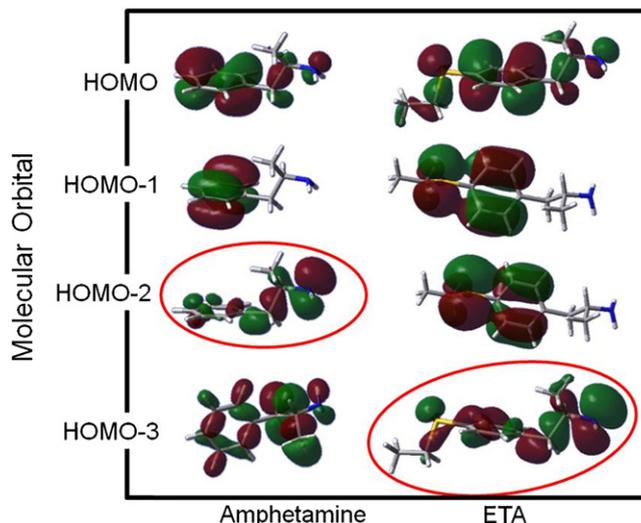


Fig. 5. Scheme of FERMO for two amphetamine derivatives.

Table 5
Statistical parameters of the PLS model for the complete data, training and test sets for R and S molecules.

	R configuration					S configuration				
	R^2	Q^2	SEC	SEV	SEP	R^2	Q^2	SEC	SEV	SEP
Complete data set	0.73	0.60	0.45	0.55		0.76	0.67	0.43	0.50	
Training set	0.74	0.59	0.46	0.53		0.79	0.68	0.41	0.47	
Test set		0.68			0.46		0.50			0.43

R^2 : correlation coefficient of multiple determination, Q^2 : cross-validated correlation coefficient, SEC: root mean square error of calibration, SEV: root mean square error of cross-validation, SEP: root mean square error of prediction.

present the same set of six descriptors, which are the CHELPG atomic charges on the C3, C4 and C5 carbons atoms on the benzene ring and the electrophilicity obtained by the HOMO energy, the polar molecular surface area ASA.P and XlogP3. Tables 3 and 4 contain the values of the selected descriptors for all 34 compounds for molecules in the R and S configurations, respectively.

A deeper interpretation of the selected descriptors shows clearly those descriptors located in the neighborhood of the molecular reaction pocket are not significant and do not bring much information about the ligand–enzyme interaction, and further about the differences in biological activities in the model. The selected descriptors in those models are located in the benzene ring, which is not close to the reactive site or the chiral carbon, suggesting that the biological activity of these compounds is more dependent on the benzene ring interaction with the receptor MAO A than the reaction site.

4. Chemometric analyses

The Student *t* test and Bonferroni test of the data in Tables 3 and 4 showed that they are equivalent with 95% of confidence.

Satisfactory PLS models with three latent variables were obtained for the two data matrices (34, 6). A good agreement can be observed between experimental data and the predicted activity with $R^2 = 0.73$ and $Q^2 = 0.60$ and $R^2 = 0.76$ and $Q^2 = 0.67$ for R and S configurations, respectively (Table 5).

The best PLS model equations for the complete data set are shown in Eqs. (1) and (2) for R and S configurations. These models describe 70.12% and 69.36% of original information for R and S configurations, respectively.

$$y = -0.112qC3 + 0.573qC4 - 0.282qC5 + 0.464\text{Electrophilicity} - 0.136\text{ASA.P} - 0.215\text{XlogP3} \quad (1)$$

$$y = -0.184qC3 + 0.541qC4 - 0.322qC5 + 0.478\text{Electrophilicity} - 0.193\text{ASA.P} - 0.291\text{XlogP3} \quad (2)$$

The final PLS models were then validated by the LNO and *y*-randomization [44–47] tests and were checked for sign change [48]. It is evident from all plots in Fig. 6, for the R configuration, that the model is robust, since it does not suffer from chance correlation [47]. It can be observed that the result obtained for LNO (Fig. 6a) is not greater than 0.1, as recommended [46]. The *y*-randomization results (Fig. 6b–d) indicate that this model does not suffer from chance correlation. All obtained values for Q^2 and R^2 tests are below 0.1 and 0.4, respectively (Fig. 6b) and the intercepts (Fig. 6c and d) are within the acceptable values recommended in the literature, i.e., the intercepts are below the limits of 0.3 and 0.05, respectively [53]. The residues (Fig. 6e) are randomly distributed showing that the model is robust. Configuration S showed a similar graphic pattern and allowed the same observations (data not showed).

After validating the models for the complete data set, the data were split into training and test sets to verify the external

predictability. To identify these two groups, hierarchical cluster analysis, HCA, was carried out on autoscaled data (dendrogram not shown). In this analysis eight compounds were selected for each molecular configuration to form the test set, are them 4-EtOA, 5Br-4,5-DMA, DOB, ALEPH-1, NMMTA – MTAB, 45a and 48a.

Once the training and test sets were defined a new PLS for the training set was built for the six selected descriptors. As expected, these model are similar to the model obtained for the complete data set with 34 compounds, showing a good agreement between experimental data and prediction activity with $R^2 = 0.74$ and $Q^2 = 0.59$ and $R^2 = 0.79$ and $Q^2 = 0.68$ for R and S configurations, respectively (see Table 5).

The PLS model was applied to the test set, and the external validation appears adequate with $Q^2 = 0.68$ and $Q^2 = 0.50$ for R and S molecular configurations, respectively. All these results can be seen in Table 5. Fig. 7a and b shows the pIC_{50} predicted by the model for the training and test sets for R and S configurations, respectively. These models are similar to those obtained for the complete data set (not shown). This analysis, as well as the interpretability of model robustness (Fig. 6) and application of the Student's *t* and Bonferroni tests (with 95% of confidence) to the obtained matrices, indicate that both configurations are suitable to build a reliable PLS model.

5. Molecular interpretation

In QSAR studies it is desirable to obtain a model where the selected molecular properties can be interpreted and can be traced parallel to the mechanism of action [46]. In the pharmacophoric context, the binding pattern of a ligand to its binding site encodes different interactions such as hydrogen-bonds, aromatic pi–pi stacking interactions between the aromatic planar systems, and hydrophobic or electrostatic interactions. In this way, a brief molecular docking with AutoDockVina was performed for amphetamine (the simplest molecular structure) and ETA (the most potent compound). Both ligand configurations were considered. In addition, a manual docking to form the ligand adduct for the amphetamine molecule, was performed. To form the adduct, the ligand was drawn into a MAO A (2Z5Y PDB code) [19] binding site, at the same position as the crystallographic ligand, followed by a fast optimization with the MM+ [54] molecular force field implemented in the molecular modeling package HyperChem. This result is shown in Fig. 8.

To follow the above purpose, the binding modes of the ligands at the substrate/ligand binding site were compared in the crystal structure of the complex [19] (Fig. 8 left) and in the modeled complex (Fig. 8 right). This binding site, as described by Son et al. [19] in the crystallographic structure, consists mainly of pi-electron systems (aromatic, delocalized and other pi-electron-based residues), while other residues are hydrophobic. This is obvious from the pi-stacking (face-to-face) interaction of harmine (HRM) with Glu215, as well as from other interactions characteristic for pi-systems such as inclined face-to-face interaction, perpendicular edge-to-face interactions (so called C–H...pi interactions), and edge-to-edge interactions (interactions between

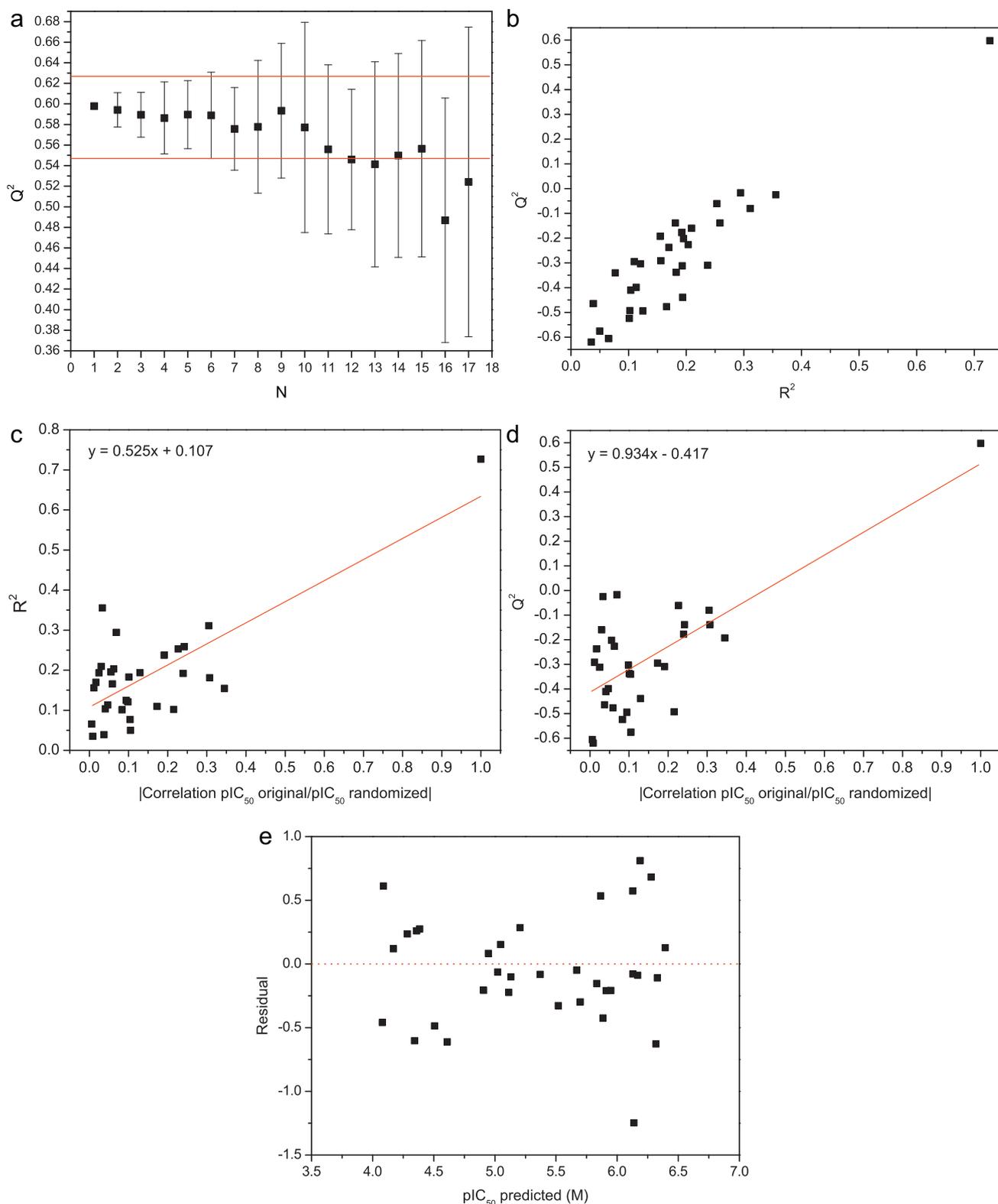


Fig. 6. (a) LNO validation plots, (b–d) y -randomization plots and (e) residuals plot, R configuration.

aromatic hydrogens). This explains the mode of binding of aromatic HRM and the tricyclic pi-system at the final segment of the FAD molecule of MAO (Fig. 8 left). The benzene ring of the ligand is well-accommodated in the pocket involved in the pi–pi stacking interaction with Tyr407, inclined face-to-face interaction with Tyr444, and aromatic hydrogen–hydrogen interactions with Tyr197. Interestingly, the benzene ring interacts additionally with

both the residues and carbonyl oxygens of Ile180 and Asn181. As has already been shown in this work, molecular descriptors accounting for the aromatic character of ligands are important for the regression models. This is consistent with the fact that the aromatic system of the benzene ring of the ligands binds to MAO mainly via interactions with pi-systems of residues and carbonyl groups. Such interactions stabilize the complex significantly,

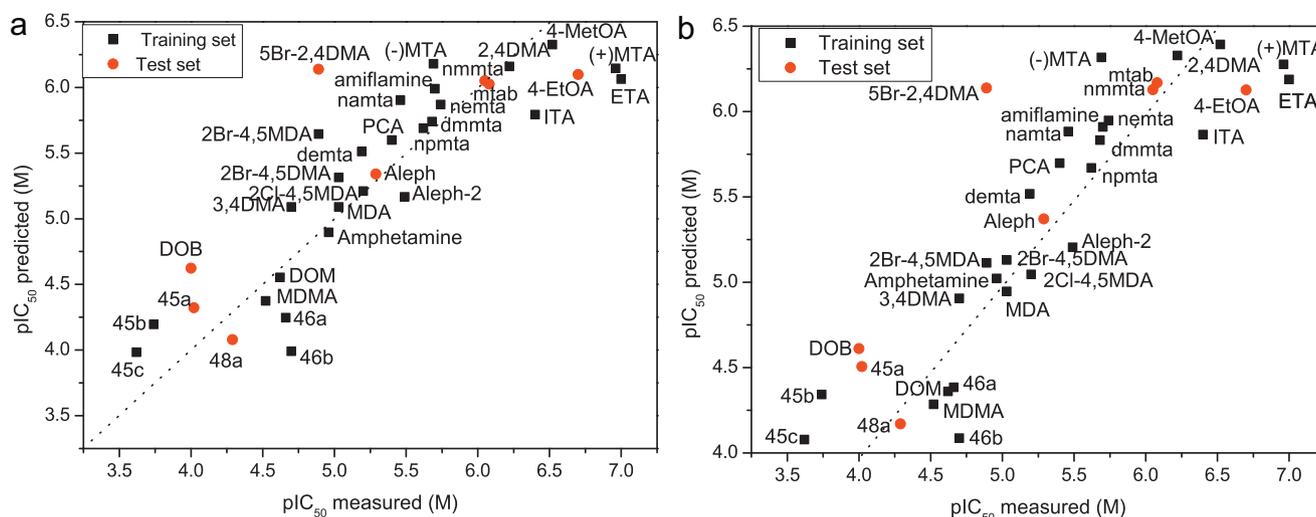


Fig. 7. Measured values for pIC_{50} for MAO A inhibition versus estimated values for regression models for the selected variables (charges $qC3$, $qC4$, $qC5$, electrophilicity, $XlogP3$ and $ASA.P$) after splitting the compounds into training and test sets of (a) R and (b) S molecular configurations. The test set is shown in red and training set is in black. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

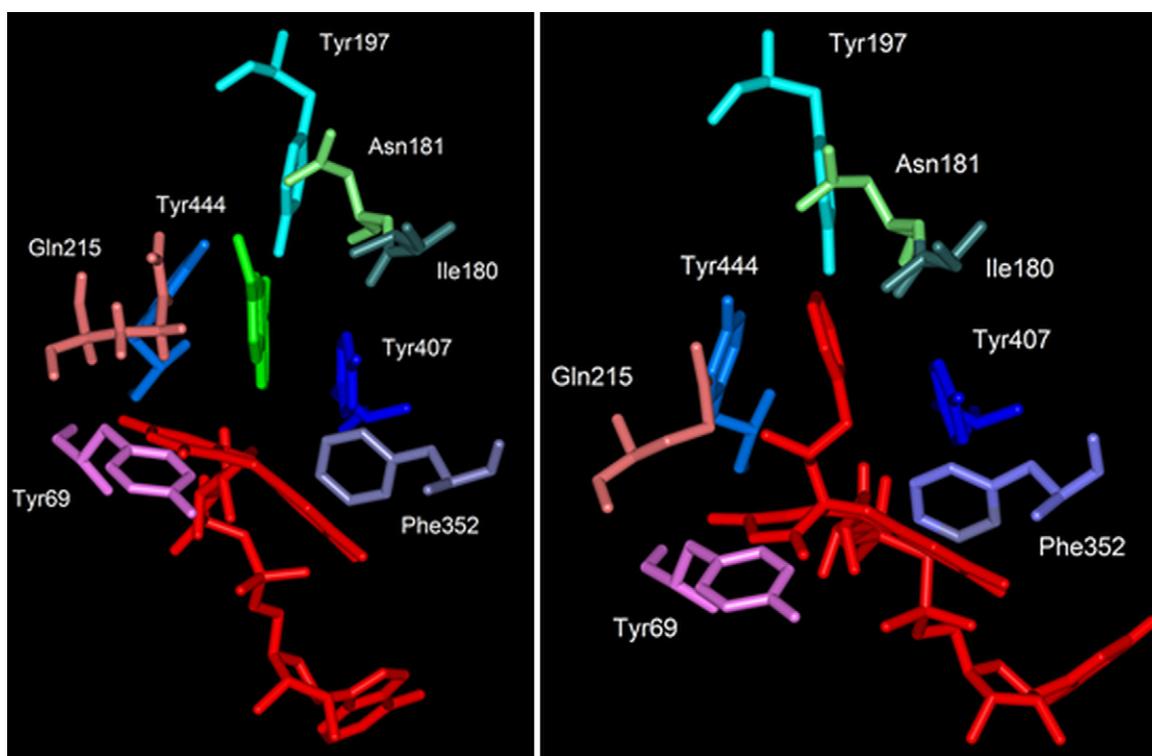


Fig. 8. Comparison of interactions of MAO amino-acids (colored in different tones) and ligands in the crystallographic structure [19] (FAD in red, HRM in green) and in modeled structure (ligand in red) at the substrate/inhibitor binding site. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

meaning that this feature can be used to improve the ligand binding and specificity at the active site in drug design.

The results obtained for $\log P$ and molecular surface area showed a good agreement with the hydrophobic binding site characteristics. The polar molecular surface area ($ASA.P$) is a solvent accessible surface area of all polar atoms. An opposite correlation coefficient between the biological activity and the $ASA.P$ (negative) and $XlogP3$ (positive) were obtained. In general, the compounds with higher $ASA.P$ values have a smaller IC_{50} value and $XlogP3$ results showed that hydrophobic ligands, with high $\log P$ values, are preferentially distributed to hydrophobic compartments. Such behavior is in perfect agreement with ligand–enzyme interaction, which indicates

that more hydrophobic a compound is, the better it binds with its site in MAO A.

6. Conclusion

A multivariate QSAR model for a set of thirty-four amphetamine derivatives was obtained for R and S configurations of amphetamine derivatives. The ligands interact with C4 of FAD (Fig. 3) and are capable of inhibiting the MAO A enzyme.

Prior to building the QSAR model, PCA and RMS analyses supported by relevant literature showed the HF/6-31G(d,p)

quantum-chemical method as appropriate for geometry optimization of these compounds and further electronic descriptor calculations.

A X matrix was formed by 38 descriptor and thirty-four compounds. The OPS algorithm, recently proposed in the literature, was applied for variable selection, and indicated that the CHELPG atomic charges at the benzene ring were the most important parameters to describe the biological activity of the present group of amphetamine derivatives. The PLS models built upon these descriptors were extensively validated indicating that they present satisfactory statistical quality, great prediction power and robustness.

The selected descriptors could explain the interaction of ligands with the enzyme via a π - π stacking interaction and suggest complex formation of the ligand by interacting with the Tyr407 amino acid. Also, ASA_P, XlogP₂ and electrophilicity obtained by the HOMO were selected as important descriptors of these compounds and can be related to the hydrophobic feature of the MAO A binding site.

As the PLS results for both configurations, R and S, of these amphetamines, were similar and confirmed, and by the Student's *t* test, it could be concluded that the models were not dependent of the ligand configuration. In this case, the similarity of both PLS models may be a result of the distance (relatively large) between the selected descriptors and the chiral atom. Therefore, no conclusions could be made about which ligand configuration is the most active, however, one can conclude, in the absence of stereochemical information, that this characteristic is not that important to build into a reliable PLS model and new optimized compounds can be suggested based on one or the other configuration.

Finally, a simple docking study and comparison of the resulting structure with the crystallographic structure of the complex clearly shows the importance of ligand-receptor interactions which are characteristic for π -electron systems; these are: π - π stacking interactions with Tyr407, inclined face-to-face interactions with Tyr444, and aromatic hydrogen-hydrogen interactions with Tyr197. This result is in agreement with the role of selected descriptors that account for the aromatic character of ligands, as can be seen from the regression model.

Acknowledgements

The authors acknowledge the financial support from the Brazilian Agencies Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) of the Ministry for Science and Technology of Brazil, and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Dr. Euzébio Barbosa for his help in the QSAR study, Dr. Ljubica Tasic and Dr. Rudolf Kiralj for their help in the ligand-enzyme interactions study and Prof. Dr. Carol H. Collins for English revision.

References

- [1] S.M. Berman, R. Kuczynski, J.T. McCracken, E.D. London, *Mol. Psychiatry* 14 (2009) 123.
- [2] B.K. Madras, G.M. Miller, A.I. Fischman, *Biol. Psychiatry* 57 (2005) 1397.
- [3] J. Knoll, *Med. Res. Rev.* 12 (1992) 505.
- [4] D.J. Mitchell, D. Nikolic, E. Rivera, S.O. Sablin, S. Choi, R.B. van Breemen, T.P. Singer, R.B. Silverman, *Biochemistry* 40 (2001) 5447.
- [5] F. Chimenti, D. Secci, A. Bolasco, P. Chimenti, B. Bizzarri, A. Granese, S. Carradori, M. Yanez, F. Orallo, F. Ortuso, S. Alcaro, *J. Med. Chem.* 52 (2009) 1935.
- [6] M.O. Ogunrombi, S.F. Malan, G. Terre'Blanche, N. Castagnoli, J.J. Bergh, J.P. Petzer, *Bioorg. Med. Chem.* 16 (2008) 2463.
- [7] G. La Regina, R. Silvestri, V. Gatti, A. Lavecchia, E. Novellino, O. Befani, P. Turini, E. Agostinelli, *Bioorg. Med. Chem.* 16 (2008) 9729.
- [8] M.C. Scorza, C. Carrau, R. Silveira, G. ZapataTorres, B.K. Cassels, M. ReyesParada, *Biochem. Pharmacol.* 54 (1997) 1361.
- [9] G. Vallejos, M.C. Rezende, B.K. Cassels, *J. Comput.-Aided Mol. Des.* 16 (2002) 95.
- [10] C. Binda, M. Li, F. Hubalek, N. Restelli, D.E. Edmondson, A. Mattevi, *Proc. Natl. Acad. Sci. U. S. A.* 100 (2003) 9750.
- [11] S. Subramanian, M.M.C. Ferreira, M. Trsic, *Struct. Chem.* 9 (1998) 47.
- [12] K. Fukui, *Science* 218 (1982) 747.
- [13] K. Fukui, *Angew. Chem. Int. Ed.* 21 (1982) 801.
- [14] R.R. da Silva, T.C. Ramalho, J.M. Santos, J.D. Figueroa-Villar, *J. Phys. Chem. A* 110 (2006) 1031.
- [15] E.B. da Costa, M. Trsic, *J. Mol. Graphics Modell.* 28 (2010) 657.
- [16] M.M.C. Ferreira, *J. Braz. Chem. Soc.* 13 (2002) 742.
- [17] C. Hurtado-Guzman, A. Fierro, P. Iturriaga-Vasquez, S. Sepulveda-Boza, B.K. Cassels, M. Reyes-Parada, *J. Enzyme Inhib. Med. Chem.* 18 (2003) 339.
- [18] J. Sterling, Y. Herzog, T. Goren, N. Finkelstein, D. Lerner, W. Goldenberg, I. Miskolczi, S. Molnar, F. Rantal, T. Tamas, G. Toth, A. Zagyva, A. Zekany, G. Lavian, A. Gross, R. Friedman, M. Razin, W. Huang, B. Kraus, M. Chorev, M.B. Youdim, M. Weinstock, *J. Med. Chem.* 45 (2002) 5260.
- [19] S.Y. Son, A. Ma, Y. Kondou, M. Yoshimura, E. Yamashita, T. Tsukihara, *Proc. Natl. Acad. Sci. U. S. A.* 105 (2008) 5739.
- [20] B.H. Morimoto, S. Lovell, B. Kahr, *Acta Crystallogr. Sect. C: Cryst. Struct. Commun.* 54 (1998) 229.
- [21] J.I. Salach, *Arch. Biochem. Biophys.* 192 (1979) 128.
- [22] L.A.S. Romeiro, C.A.M. Fraga, E.J. Barreiro, *Quim. Nova* 26 (2003) 347.
- [23] J.I. Salach, K. Detmer, M.B.H. Youdim, *Mol. Pharmacol.* 16 (1979) 234.
- [24] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, H. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. Montgomery, J. A. J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N.J. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian, Inc., Wallingford, CT, 2003.
- [25] I.V. Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V. Palyulin, E. Radchenko, N.S. Zefirov, A.S. Makarenko, V.Y. Tanchuk, V.V. Prokopenko, *J. Comput.-Aided Mol. Des.* 19 (2005) 453.
- [26] VCCLAB, 2005. <http://www.vcclab.org>
- [27] MarvinSketch5.9.0, 2012-03-02, trial version.
- [28] R.P.K. Beebe, M.B. Seasholtz, *Chemometrics: A Practical Guide*, John Wiley & Sons, New York, 1998.
- [29] Pirouette Software Version 3, Infometrix Inc., USA, 2007.
- [30] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, *J. Am. Chem. Soc.* 107 (1985) 3902.
- [31] C.C.J. Roothaan, *Rev. Mod. Phys.* 23 (1951) 69.
- [32] C.C.J. Roothaan, *Rev. Mod. Phys.* 32 (1960) 179.
- [33] P. Hohenberg, W. Kohn, *Phys. Rev. B* 136 (1964) B864.
- [34] W. Kohn, L.J. Sham, *Phys. Rev.* 140 (1965) A1133.
- [35] A.D. Becke, *J. Chem. Phys.* 98 (1993) 5648.
- [36] C.T. Lee, W.T. Yang, R.G. Parr, *Phys. Rev. B* 37 (1988) 785.
- [37] C. Moller, M.S. Plesset, *Phys. Rev.* 46 (1934) 0618.
- [38] HyperChem Software Version 7.1, Hyper Co., USA, 2002.
- [39] C.M. Breneman, K.B. Wiberg, *J. Comput. Chem.* 11 (1990) 361.
- [40] R.G. Pearson, *Proc. Natl. Acad. Sci. U. S. A.* 83 (1986) 8440.
- [41] R.G. Parr, L. Von Szentpaly, S.B. Liu, *J. Am. Chem. Soc.* 121 (1999) 1922.
- [42] P. Chen, K. Fujisawa, E.I. Solomon, *J. Am. Chem. Soc.* 122 (2000) 10177.
- [43] R.F. Teofilo, J.P.A. Martins, M.M.C. Ferreira, *J. Chemometr.* 23 (2009) 32.
- [44] N. Bratchell, *J. Chemometr.* 11 (1997) 93.
- [45] A. Golbraikh, A. Tropsha, *J. Mol. Graphics Modell.* 20 (2002) 269.
- [46] R. Kiralj, M.M.C. Ferreira, *J. Braz. Chem. Soc.* 20 (2009) 770.
- [47] A. Tropsha, P. Gramatica, V.K. Gombar, *QSAR Comb. Sci.* 22 (2003) 69.
- [48] R. Kiralj, M.M.C. Ferreira, *J. Chemometr.* 24 (2010) 681.
- [49] O. Trott, A.J. Olson, *J. Comput. Chem.* (2010) 455.
- [50] L. De Colibus, M. Li, C. Binda, A. Lustig, D.E. Edmondson, A. Mattevi, *Proc. Natl. Acad. Sci. U. S. A.* 102 (2005) 12684.
- [51] H. Guo, M. Karplus, *J. Chem. Phys.* 91 (1989) 1719.
- [52] D. Moran, A.C. Simmonett, F.E. Leach, W.D. Allen, P.V. Schleyer, H.F. Schaefer, *J. Am. Chem. Soc.* 128 (2006) 9342.
- [53] L. Eriksson, J. Jaworska, A.P. Worth, M.T.D. Cronin, R.M. McDowell, P. Gramatica, *Environ. Health Perspect.* 111 (2003) 1361.
- [54] N.L. Allinger, X.F. Zhou, J. Bergsma, Theochim – J. Mol. Struct. 118 (1994) 69.