QSAR studies on a number of pyrrolidin-2-one antiarrhythmic arylpiperazinyls

2 AUTHORS:

Alicja Nowaczyk
Nicolaus Copernicus University
28 PUBLICATIONS   **134** CITATIONS

SEE PROFILE

Katarzyna Kulig
Jagiellonian University
75 PUBLICATIONS   **547** CITATIONS

SEE PROFILE

ORIGINAL RESEARCH

QSAR studies on a number of pyrrolidin-2-one antiarrhythmic arylpiperazinyls

Alicja Nowaczyk · Katarzyna Kulig

Abstract The activity of a number of 1-[3-(4-arylpiperazin-1-yl)propyl]pyrrolidin-2-one antiarrhythmic (AA) agents was described using the quantitative structure–activity relationship model by applying it to 33 compounds. The molecular descriptors of the AA activity were obtained by quantum chemical calculations combined with molecular modeling calculations. The resulting model explains up to 91% of the variance and it was successfully validated by four tests (LOO, LMO, external test, and Y-scrambling test). Statistical analysis shows that the AA activity of the studied compounds depends mainly on the PCR and JGI4 descriptors.

Keywords 1-[3-(4-Arylpiperazin-1-yl)propyl]pyrrolidin-2-one derivatives · Antiarrhythmic activity · QSAR analysis

Introduction

α_1-Adrenergic receptors (α_1-AR) are members of the G-protein coupled superfamily of receptors, which modulate intercellular biochemical processes in response to changes in the extracellular concentration of the neurotransmitter norepinephrine and the circulating hormone epinephrine, leading to widespread physiological actions that make them attractive targets for drug discovery (Becker *et al.*, 2004; Golan 2008; He *et al.*, 2008; Zhong and Minneman 1999). They are responsible for a number of physiological functions (Abbas *et al.*, 2006; Graham *et al.*, 1996; Piascik *et al.*, 1999) in:

(a) cardiovascular tissues regarding vascular smooth contraction and blood pressure regulation,
(b) noncardiovascular tissues regarding the human prostate smooth muscle contraction or the regulation of cerebral microcirculation.

Thus, α_1-AR antagonists can be useful in the treatment of hypertension, benign prostatic hyperplasia (BPH), lower urinary track symptoms (LUTS), or cardiac arrhythmia (Carmeliet and Mubagwa, 1998; Chiu *et al.*, 2008; Jain *et al.*, 2008; Koshimizu *et al.*, 2007; Nargund and Grey, 2008; Thiyagarajan, 2002).

Now, in the globalization era, determined by speed, uncertainty and instability people live in increasing stress leading to a rise in the incidence of cardiovascular diseases. Cardiac arrhythmia may be caused by abnormal impulse formation, abnormal impulse propagation, or both (Matyus *et al.*, 1997) it remains a major source of morbidity and mortality in developed countries. For example, between 0.5 and 1 million North Americans and Europeans die each year because of sudden cardiac death, which corresponds to 10–20% of all deaths among adults in the Western world (Goldberger *et al.*, 2008; Huikuri *et al.*, 2001; Kromhout, 2007). In the past decade, the treatment of arrhythmia has been dramatically altered by the development of non-pharmacological therapies, such as targeted ablation of

A. Nowaczyk (✉)
Department of Organic Chemistry, Faculty of Pharmacy, Collegium Medicum in Bydgoszcz, Nicolaus Copernicus University, Sklodowskiej-Curie 9, 85-094 Bydgoszcz, Poland
e-mail: alicja@cm.umk.pl

K. Kulig
Department of Physicochemical Drug Analysis, Faculty of Pharmacy, Medical College Jagiellonian University, 9 Medyczna Str, 30-688 Kraków, Poland

arrhythmogenic tissues and implantable cardioverter defibrillators (ICDs), as well as the limited efficacy and proarrhythmic potential of conventional antiarrhythmic (AA) drugs (Estrada and Darbar, 2008). AA drugs have been classified by Vaughan Williams mainly based on their effects on cardiac action potentials into classes I–IV and later correlated to their effects on Na^+ channel, β-receptors, and K^+ and Ca^{2+} channels (Hashimoto, 2007; Vaughan Williams, 1992).

In the course of our studies directed to search for new α_1-AR antagonists, among which a series of (4-arylpiperazin-1-yl)propylpyrrolidin-2-one or 3-alkyl-3-phenylpyrrolidin-2-one derivatives, it was shown that the compounds obtained also showed marked AA and hypertensive activities. The ED_{50} values determined for a number of them was lower than or comparable with the reference compounds (Kulig et al., 2003, 2004, 2007, 2009; Malawska et al., 2002, 2005). For a large number of chemometric analyses reported in medical research, there are relatively few studies on the application of QSAR analysis to AA species (Debnath et al., 2003; Fumagalli et al., 2005; Pallavicini et al., 2006; Turabekova et al., 2008). In this context, the aim of this study, being a part of our drug design project, is to find a model explaining the AA activity of a series of 1-[3-(4-arylpiperazin-1-yl)propyl]pyrrolidin-2-one derivatives applying the quantitative relationship between structural parameters and AA activity. The quantitative structure–activity relationship (QSAR) equation for our compounds is presented and discussed.

Computational methods

### 1-[3-(4-Arylpiperazin-1-yl)propyl]pyrrolidin-2-one derivatives

Thirty-three analogs of 1-[3-(4-(aryl)piperazin-1-yl)propyl]pyrrolidin-2-one were chosen from the reports published by us between 2002 and 2009 (Kulig et al., 2003, 2004, 2007, 2009; Malawska et al., 2002, 2005). The source publications concern the synthesis of over 70 arylpiperazine derivatives and their pharmacological test results. About 20 of these compounds display a lack of α_1-ARs activity and 40 compounds display a lack of AA activity. These compounds are considered to be irrelevant for the model formulation and they were excluded from the current study. Thus, the set of the remaining 33 compounds displaying both α_1-ARs and AA activity are appropriate for a QSAR analysis and are listed in Table 1. The external set should include about 10–30% of the entire set and should represent activities and structures that cover the whole range of the training set (Gramatica, 2007). Consequently the initial data set was split into two subsets: a training

subset ($N_{TS} = 25$) and a external cross-validation subset included randomly selected compounds number: **1**, **3**, **8**, **17**, **21**, **23**, **25**, **30** ($N_{EXT} = 8$).

Molecular descriptors and methods

In order to identify the effect of the molecular structure on the AA activity a QSAR analysis of the selected compounds was performed.

(1) The AA activity data expressed as ED_{50} (mg/kg) are taken from the source publications and recalculated to ED_{50} (mM/kg). Logarithmic values ($-\log ED_{50}$) are listed in Table 1 as AA observed activity. Each ED_{50} (mg/kg) value was obtained from independent experiments in adrenaline included arrhythmia in anaesthetized rats (Szekeres and Papp, 1975).

(2) For the molecular 3D structure calculations the Gaussian® 03 (version 6.1) package was used (Frisch et al., 2004). The three-dimensional structures of the pyrrolidin-2-one derivatives in their neutral state were obtained through full optimization based on the AM1 quantum chemical procedure. Harmonic vibrational analysis was used to ascertain whether the resulting geometries were the true energy minima structures. All the molecules were minimized until the root mean square (RMS) gradient value was smaller than 10^{-6} a.u. Next, resulting molecular 3D structure was used for the calculation of the descriptors set and to visualize the distribution of charge in a molecule (the map of the electrostatic potential in the form of a 3D plot). In order to obtain reliable energetic and accurate data on electronic properties of molecules the single-point energy calculations were performed at the DFT/B3LYP level of theory using the 6-31G** basis set. Suitable maps of the electrostatic potential were plotted based on the electronic and nuclear charge distribution obtained from the energy calculations results. The Gaussian suite of programs calculates the electrostatic potential maps and surfaces as the distribution of the potential energy of unit positive charge in a given molecular space, with a resolution controlled by the grid density. In Fig. A in the Supplementary file representative plots for extreme difference in the charge distribution pattern are shown (Frisch et al., 1998; Leach, 2001).

(3) For the calculation of the descriptors the Talete srl, DRAGON for Windows Version 5.5-2007 package was used. Dragon descriptors include 22 different logical blocks. The total number of calculated descriptors was 3224. Several criteria were used to reduce this number while optimizing the information content of the descriptors set. First, descriptors for

Table 1 Structures and affinities for AA action of 1-[3-(4-arylpiperazin-1-yl)propyl]pyrrolidin-2-one derivatives used in the current work
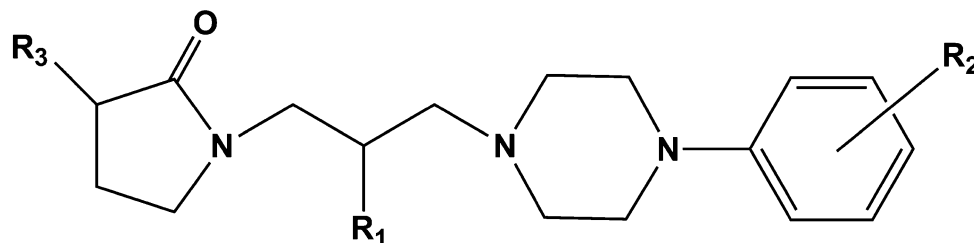
Compounds	AA activity		R_1	R_2	R_3
	Observed	Predicted			
1[a]	2.01	2.09	H	H	H
2	1.79	1.86	H	2-OMe	H
3[a]	1.80	1.79	H	2-Cl	H
4	1.54	1.71	H	2-F	H
5	2.52	2.24	H	2-OEt	H
6	1.45	1.46	H	3-CF$_3$	H
7	1.43	1.43	OH	2-OMe	H
8[a]	1.40	1.44	OH	4-Cl	H
9	1.79	1.58	OH	2-F	H
10	1.64	1.60	OH	3-OMe	H
11	1.97	2.15	OH	2-OEt	H
12	1.55	1.56	OH	2-Me	H
13	2.23	2.21	OH	2-OH	H
14	1.77	1.79	OH	2-OiPr	H
15	1.31	1.31	OH	2-CF$_3$	H
16	1.54	1.53	OH	2,4-diF	H
17[a]	1.48	1.32	OH	2-OMe, 5-Cl	H
18	2.37	2.54	OH	2-OMe	3,3-diPh
19	2.13	2.17	OH	2-CF$_3$	3,3-diPh
20	2.53	2.37	OH	2-Me	3,3-diPh
21[a]	2.66	2.55	OH	2-OEt	3,3-diPh
22	2.38	2.33	OH	H	3,3-diPh
23[a]	1.60	1.88	OH	H	H
24	1.92	1.86	O(CO)NHEt	2-OMe	H
25[a]	2.19	1.99	O(CO)NHiPr	2-OMe	H
26	1.52	1.56	O(CO)NHnPr	2-OMe	H
27	1.77	1.81	O(CO)nPr	2-OiPr	H
28	2.00	2.00	O(CO)NHiPr	2-Cl	H
29	1.66	1.75	O(CO)NHEt	H	H
30[a]	1.88	1.95	O(CO)iPr	H	H
31	1.47	1.51	O(CO)NHnB	H	H
32	1.52	1.42	O(CO)NHnPr	H	H
33	1.36	1.37	H	2-OH	H

The AA expressed as $-\log ED_{50}$ values, in mM/kg

[a] Compounds excluded in the model generation procedures; external data set, AA observed activity by pharmacological tests, AA predicted activity by Eq. 1

which no value was available for all the compounds were disregarded. Second, descriptors of which the value is constant (or near-constant) inside each group of descriptors were excluded. For the remaining descriptors, if two descriptors showed a correlation coefficient greater than 0.9, the one showing of the highest pair correlation with the others descriptors was removed. After these automatic screening procedures, a set of 385 descriptors was obtained for further analysis. To reduce the vast number of descriptors to the 50 that correlated best with the experimental data, the "Feature Selection and Variable Screening" methods available in Statistica® (version 8.0) (2008) software were applied. Then, the chosen descriptors were used as regressors of the model: they are collected in Table A in the Supplementary file and a detailed description of these descriptors can be found in the literature (Todeschini and Consonni, 2002).

Statistical analysis

The Multiple Linear Regression (MLR) (Allison, 1999) and correlation analyses were carried out using the Statistica® (version 8.0) (2008) software. The forward stepwise regression analysis yielded a three-parametric model describing the biological activity as a function of molecular descriptors. The statistical quality of the regression equations was evaluated by parameters such as the correlation coefficient R, the squared correlation coefficient R^2, the adjusted squared correlation coefficient R^2_{adj}, the Root Mean Squared Errors (RMSE) and the variance ratio F. The statistical significance (P level) of a result was determined as $P \leq 0.01$ (Bland, 2000).

The model obtained in this study was validated by calculations of the validated squared correlation coefficient (Q^2) values and prediction error sum of squares (called SPRES) values. The Q^2 values were calculated from the general internal cross-validation procedures "leave-one-out" test (LOO) and "leave-many-out" test (LMO) and external tests (EXT) (Baumann, 2005; Golbraikh and Tropsha, 2002; Hawkins, et al., 2003; Kubinyi 1997a, b). Abbreviations Q^2_{LOO}, Q^2_{LMO}, Q^2_{EXT} (and QS$_{LOO}$, QS$_{LMO}$, QS$_{EXT}$) have been used in their's usual meaning for the tests listed above. In addition, the robustness of the proposed model was checked by permutation testing: parallel models were developed based on a fit to randomly reordered Y-data (Y-scrambling, Y-randomization) (Gramatica, 2007; Tropsha, 2010; Tropsha et al., 2003). According to the basic approach of Wold and Eriksson (1995) all randomization methods consisted of ten randomization runs for any data set size.

All computations were performed on a HP 6200 wx workstation.

Results and discussion

Table 1 reports the observed AA activity, expressed as $-\log$ ED$_{50}$ (mM/kg) values in adrenaline included arrhythmia in anaesthetized rats. All the tested compounds showed AA stimulation as the $-\log$ ED$_{50}$ values are between 1.31 and 2.66.

In this study we have limited the number of presented equations to this of the best regression model of the whole set. The model is given as follows together with the statistical and validation parameters:

$$AA = -60.167(\pm 13.005)\ JGI4\ + 12.334(\pm 3.841)\ PCR + 0.986(\pm 0.213)\ Hy - 20.110(\pm 6.072)$$

(1)

$R = 0.953$, $R^2 = 0.909$, $R^2_{adj} = 0.844$, $F = 14.040$, RMSE $= 0.141$, $N_{TS} = 25$, $N_{EXT} = 8$, $P < 0.01$, $Q^2_{LOO} = 0.744$, QS$_{LOO} = 0.178$, $Q^2_{LMO} = 0.736$, QS$_{LMO} = 0.175$, $Q^2_{EXT} = 0.858$, QS$_{EXT} = 0.168$ $R^2_Y = 0.074$, $Q^2_Y = 0.022$, where N is the number of compounds included in the [training (TS)/external (EXT)] data set, R the correlation coefficient, R^2 the squared correlation coefficient, R^2_{adj} the adjusted squared correlation coefficient, RMSE the root mean squared errors, F the variance ratio, P the significance of the variables in the model, Q^2_{LOO}, Q^2_{LMO}, Q^2_{EXT}, R^2_Y, and Q^2_Y the correlation coefficient of the adequate validation methodologies.

The presented QSAR analysis yields a model incorporating three descriptors. Since the Topliss and Costello rule (1972) allows the use of up to five descriptors for a training set consisting of 25 compounds and the relation $R^2_{adj} < R^2$ is true, the model in not overparametrized. However, for AA action we did not fit any better correlation using more descriptors in multi-parameter correlations. The correlation coefficient R of this relationship is 0.95 and explains up to 91% of all variance data for AA activity. Moreover, the F test value together with RMSE at the P level of 1×10^{-5} suggests that the equation has a good correlation with the data and is statistically significant. Every descriptor in the regression equation must be independent. The correlation between each descriptor was calculated and is presented in form of a Pearson correlation matrix in Table 2. As can be seen from these numbers all predictors have a pair correlation minimal covariance <0.5 which assures that any collinearity of predictors is not present. Table 1 reports the AA activity predicted by Eq. 1. A plot of the predicted activity versus the residual values was prepared to determine the existence of systematic errors in the model

Table 2 Pearson correlation matrix of the parameters used in this study

	JGI4	PCR	Hy
JGI4	1.00		
PCR	0.47	1.00	
Hy	0.39	−0.22	1.00

JGI4 Mean topological charge index of order 4, *PCR* ratio of multiple path count over path count, *Hy* hydrophilic factor

development (see Fig. B in the Supplementary file). The uniform distribution of residues indicates no systematic error (Belsley *et al.*, 2005). The plots of observed AA activities versus those predicted by Eq. 1 together with the corresponding predicted intervals are shown in Fig. C in the Supplementary file. Compound number **5** is out of 91% prediction threshold and exhibits high AA activity in contrast to other compounds of similar structure having low hydrophobic factor i.e., compounds **2**, **4**–**6**. This incidence may be explained by unique structural features. This plot proves that the model as a good descriptive power. Summing up the linear model seems to be adequately fit to the data, all predictors have $P < 0.01$ and one can conclude that all are independently associated with AA activity.

In an attempt to determine the utility of Eq. 1 as model of AA activity four validation analyses were carried out i.e., LOO, LMO, *Y*-scrambling, and external predictivity (Kiralj and Ferreira, 2009). In the field of statistical techniques the LOO and LMO are used for internal validation. From a theoretically acceptable model the R^2 cannot have smaller values than Q^2_{LOO} and Q^2_{LMO} or Q^2_{EXT}. Overall, the best model is achieved when $Q^2_{LOO} \leq R^2 \geq Q^2_{LMO}$ and $Q^2_{LOO} \approx Q^2_{LMO}$. Commonly, $Q^2_{LOO} > 0.5$ is considered as proof of the reasonably predictive capability of the equation. $Q^2_{LOO} > 0.7$ indicates the stable and predictive potential of the equation. Nevertheless a high Q^2_{LOO} value does not indicate a high predictive power of the model. On

the other hand if $R^2 < Q^2_{LOO}$ the model is overfitted. As can be seen from the statistics presented next to Eq. 1 in our case $R^2 > Q^2_{LOO}$, which means that our model is not overfitted. The LMO test is usually used to verify results obtained from the LOO test. In the Q^2_{LMO} procedure ten iterations were performed with five molecules left out in each iteration (e.g., tenfold, 80/20 cross validation) (Kiralj and Ferreira, 2009; Tropsha, 2010). The results of the LMO test are collected in Table 3. On average, the overall test steps $R^2 > Q^2_{LMO}$ and $Q^2_{LOO} \approx Q^2_{LMO}$ which is another proof that the model is not underdetermined. In order to ascertain whether the good results of the model described by Eq. 1 are not due to chance correlation or structural dependency of the training set, the *Y*-scrambling tests were performed. The results of ten runs of *Y*-randomization tests are shown in the Table 4. The average values are smaller than 0.2, which, according to Wold and Eriksson (1995), points to the absence of chance correlation (Kiralj and Ferreira, 2009; Tropsha, 2010). The low R^2_Y and Q^2_Y values prove that our model is valid. To validate the predictive power of the mathematical model more explicitly one needs to conduct validation on the external set of data (Gramatica, 2007; Kiralj and Ferreira, 2009). Therefore, the EXT test was carried out on the groups of compounds including 30% of the data set. As mentioned above, a subset of eight randomly selected compounds was removed from the entire set to be used in the validation procedure. For external compounds (**1**, **3**, **8**, **17**, **21**, **23**, **25**, and **30**) $Q^2_{EXT} = 0.86$ combined with the fact that there are no outliers which exhibit a systematic error, conclusively prove the good predictive potency of the quantitative relationship constructed on the basis of the AA activity. Thus, in our opinion, the derived models can be used for the prediction of the AA commotion for new compounds in a series of analogs. The 3-parametric equation defines the best model for this subset of data. Molecular descriptors incorporated in the equation are: JG4I, PCR, and Hy. All

Table 3 The results of the LMO test

Number of runs	Number of excluded compounds in the LMO test	Q^2_{LMO}	QS_{LMO}
1	26, 22, 33, 11, 20	0.76	0.18
2	13, 9, 33, 29, 22	0.82	0.12
3	20, 7, 32, 14, 24	0.71	0.21
4	24, 20, 9, 19, 16	0.74	0.17
5	29, 28, 32, 20, 33	0.66	0.21
6	24, 6, 18, 14, 19	0.73	0.16
7	24, 9, 13, 20, 16	0.67	0.20
8	16, 27, 20, 22, 13	0.69	0.21
9	22, 19, 14, 27, 9	0.87	0.09
10	14, 5, 32, 2, 13	0.71	0.19
Average values		0.74	0.17

Table 4 R_Y^2 and Q_Y^2 values after ten Y-scrambling tests

Number of runs	Order of compounds in observed y vector in the Y-scrambling test	R_Y^2	Q_Y^2
1	9, 4, 32, 24, 19, 27, 12, 33, 29, 11, 22, 26, 15, 6, 20, 14, 28, 5, 31, 16, 13, 10, 2, 18, 7	0.07	0.01
2	12, 19, 14, 9, 26, 20, 33, 16, 32, 28, 24, 22, 27, 29, 5, 10, 4, 6, 18, 7, 2, 31, 11, 15, 13	0.12	0.05
3	16, 19, 22, 33, 11, 6, 2, 7, 26, 4, 5, 24, 31, 15, 10, 20, 29, 14, 27, 13, 28, 12, 32, 18, 9	0.06	0.02
4	28, 12, 4, 20, 15, 11, 24, 2, 9, 7, 31, 6, 29, 18, 16, 26, 19, 22, 14, 33, 5, 27, 10, 32, 13	0.06	0.01
5	32, 2, 16, 20, 6, 22, 19, 15, 14, 5, 26, 29, 7, 4, 18, 12, 28, 11, 10, 33, 31, 27, 9, 24, 13	0.09	0.01
6	32, 19, 13, 12, 6, 20, 28, 10, 27, 31, 33, 16, 7, 14, 11, 29, 24, 15, 26, 4, 5, 9, 2, 22, 18	0.08	0.05
7	15, 31, 2, 20, 27, 9, 28, 13, 19, 12, 33, 24, 7, 14, 11, 29, 5, 16, 22, 32, 18, 26, 10, 6, 4	0.04	0.00
8	7, 28, 10, 31, 11, 22, 19, 29, 33, 12, 27, 18, 32, 20, 6, 13, 2, 9, 5, 15, 26, 4, 24, 14, 16	0.03	0.00
9	27, 29, 24, 33, 28, 4, 19, 31, 32, 12, 9, 14, 13, 7, 18, 22, 26, 5, 20, 11, 16, 10, 15, 6, 2	0.05	0.00
10	27, 6, 10, 2, 14, 31, 19, 29, 32, 4, 26, 11, 18, 12, 9, 13, 15, 24, 28, 33, 16, 5, 22, 7, 20	0.13	0.07
Average values		0.07	0.02

Table 5 Multiple regression results

	BETA	Standard error	B	Standard error	$t(14)$	P level
Intercept			−20.1101	6.07174	−3.31209	0.005137
JGI4	−0.870898	0.188244	−60.1674	13.00513	−4.62644	0.000392
PCR	1.026828	0.319750	12.3345	3.84092	3.21134	0.006277
Hy	0.604621	0.130843	0.9856	0.21329	4.62095	0.000396

the obtained descriptors belong to different logical blocks of descriptors such as the Topological charge indices (TCI) (JGI4), (Gálvez *et al.*, 1996, 1995, 1994; Rios-Santamarina *et al.*, 1998). The Walk and path counts (PCR) (Diudea *et al.*, 1994; Randic, 1980; Razinger, 1986; Rücker and Rücker, 1993, 2000), and the Molecular properties (Hy) (Todeschini *et al.*, 1997). Brief detailed descriptions of these descriptors can be found in the literature (Todeschini and Consonni, 2002). The obtained model incorporates descriptors of rather structural nature due to the regression coefficient value (see Eq. 1). As can be easily noticed, the descriptors influencing the investigated properties the most are JG4I and PCR. All descriptors related to physico-chemical properties of the molecule (except two) were excluded during the statistical analysis (Table A in the Supplementary file). This means that the structure and geometry of the molecule affect the AA activity, rather than its physico-chemical properties. Looking more closely at the chosen descriptors and their statistics in Table 5 JGI4 and PCR have |BETA| > 1 (Achen, 1982).

The molecular charge distribution plays an important role in many biological and pharmacological activities. Kier and Hall (1999) developed the concept of E-states, an electrotopological-state index for atoms in a molecule. For calculating TCI descriptors, H-depleted molecular structure is represented as a graph G. TCI are calculated using the "inverse square topological distance matrix" where the

charge influence decreases with the square of the distance. Gálvez *et al.* (1996, 1995) introduced the "inverse square topological distance matrix" denoted by D^* in which matrix elements are the inverse square of the corresponding element in the topological distance matrix D. The diagonal entries of the topological distance matrix remain the same, so diagonal entries of D^* are 0. Finally,

$$\mathrm{JGIk} = \frac{\mathrm{GGIk}}{(N-1)} \text{ and } \mathrm{GGIk} = \sum_{i=1, j=i+1}^{i=N-1, j=N} \left|\mathrm{CT}_{i,j}\right| \delta(k, D_{i,j}),$$

(2)

where, δ is where d is Kronecker's delta and $\mathrm{CT}_{ij} = m_{ij} - m_{ji}$ [m stands for the elements of the M matrix $M = A \times D^*$; A is the adjacency ($N \times N$) matrix of the molecular graph G, where N is the number of vertices (atoms different to hydrogen)]. Thus, GGIk represents the sum of all the CT_{ij} terms, with $D_{ij} = k$, being D_{ij} the entries of the topological distance matrix (D) and k ranging from 1 up to 10. These indexes represent a strictly topological quantity plausibly correlating with the charge distribution inside the molecule.

In other words, the TCI estimates the charge transfer between pair of atoms, and hence the global charge transfer in the molecule. The JGI4 parameter varies within the investigated set from 0.040 (compound **1**, unsubstituent) to 0.016 (compound **17**, for which R_1-OH, R_2-2-OMe, 5-Cl,

and R_3-H). In Fig. A in the Supplementary file, the differences in the distribution of the electrostatic charge in compounds **1** and **17** are visualized. Because the sign of the regression coefficient is negative, an increase of this predictor values will result in a decrease in AA activity. This suggests that some unique charge distribution is needed for increase AA activity.

The PCR descriptor is related to the molecular complexity of the graph (Trinajstic, 1992) i.e., to molecular branching and size as derived from the ratio of multiple path count over path count and it is sensitive to the substituent position within the investigated set as it varies from 1.182 (compound **31**, for which O(CO)NHnB substituent R_1 and H substituted R_2 and R_3) to 1.309 (complex derivative 21, for which of R_1-OH, R_2-2-OEt and R_3-3, 3-diPh). Because the sign of the regression coefficient is positive, a decrease of this predictor will result in a decrease in AA stimulation. Our earlier qualitative investigations (SAR) led us to similar conclusions (Kulig *et al.*, 2007; Nowaczyk *et al.*, 2009, 2010). The remaining parameter of the model (Hy) is the hydrophilic factor. It is a simple empirical index related to the hydrophilicity of compounds. In our data set the Hy index varies between -0.8 and 0.4. According to the sign of the BETA coefficient (Table 5), an increase in the hydrophilicity of the compounds will result in an increase in the predicted feature, although the relatively low absolute BETA values indicate that their significance in the model is not crucial.

Conclusions

In this study we have developed a mathematical model for the prediction of the AA activity of a series of 1-[3-(4-arylpiperazin-1-yl)propyl]pyrrolidin-2-ones containing various substituents on the aryl, propyl, and pyrrolidin-2-one moieties. The resulting model displays a good fit with the experimental data, with a correlation coefficient of 0.95 and explains up to 91% of the variance. In addition, the cross-validation coefficients reflecting the predictive power of the regression, Q^2_{LOO} is 0.74, and Q^2_{LMO} is 0.74. The *Y*-scrambling test proved that the good statistics obtained for Eq. 1 are not due to chance correlation or structural dependency of the training set. In addition, the external test showed a Q^2_{EXT} of 0.86 which proves a good predictability of the AA by the model (Eq. 1). The main purpose of this investigation was to determine the parameters which best describe the biological activity of a number of arylpiperazines derivatives. The results obtained here show that the activity of these compounds is mainly determined by the JGI4-, PCR- , and Hy-values. The model provides important information on the structure–activity relationships of these types of compounds at the molecular

level relevant for the design of new AA derivatives. The JGI4 of a potent agent should be as low as possible while PCR- and Hy-values should be high. On the basis of these results in combination with previous evidences we can conclude that the interaction of the 1-[3-(4-arylpiperazin-1-yl)propyl]pyrrolidin-2-one moiety with the arrhythmic species is greatly increased by the structure and the geometry of the molecule rather than its physico-chemical properties. More extensive in silico studies are in progress and will be reported in due course.

References

Abbas SA, Munavvar AS, Abdullah NA, Johns EJ (2006) Involvement of α_1-adrenoceptor subtypes in the cardiac failure in spontaneously hypertensive rats. J Basic Appl Sci 2:59–69

Achen CH (1982) Interpreting and using regression. Sage, London

Allison PD (1999) Multiple regression: a primer. Pine Forge Press, London

Baumann K (2005) Chance correlation in variable subset regression: influence of the objective function, the selection mechanism, and ensemble averaging. QSAR Comb Sci 24:1033–1046

Becker OM, Marantz Y, Shacham S, Inbal B, Heifetz A, Kalid O et al (2004) G protein-coupled receptors: in silico drug discovery in 3D. Proc Natl Acad Sci USA 101:11304–11309

Belsley DA, Kuh E, Welsch RE (2005) Regression diagnostics: identifying influential data and sources of collinearity. John Wiley & Sons, New York

Bland M (2000) Introduction to medical statistics, 3rd edn. Oxford University Press, London

Carmeliet E, Mubagwa K (1998) Antiarrhythmic drugs and cardiac ion channels: mechanisms of action. Prog Biophys Mol Biol 70:1–72

Chiu G, Li S, Connolly PJ, Pulito V, Liu J, Middleton SA (2008) Phenylpiperazinyl) cyclohexylureas: discovery of $\alpha_{1a/1d}$-selective adrenergic receptor antagonists for the treatment of benign prostatic hyperplasia/lower urinary tract symptoms (BPH/LUTS. Bioorg Med Chem Lett 18:640–644

Debnath B, Samanta S, Naskar SK, Roy K, Jha T (2003) QSAR study on the affinity of some arylpiperazines towards the 5-HT1A/α_1-adrenergic receptor using the E-state index. Bioorg Med Chem Lett 13:2837–2842

Diudea MV, Topan M, Graovac A (1994) Layer matrices of walk degrees. J Chem Inf Comput Sci 34:1072–1078

Estrada JC, Darbar D (2008) Clinical use of and future perspectives on antiarrhythmic drugs. Eur J Clin Pharmacol 64:1139–1146

Foresman JB, Frisch A (1998) Exploring chemistry with electronic structure methods. Gaussian, Inc., Pittsburg

Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR et al (2004) Gaussian 03, Revision D.01. Gaussian, Inc., Wallingford

Fumagalli L, Bolchi C, Colleoni S, Gobbi M, Moroni B, Pallavicini M et al (2005) QSAR study for a novel series of ortho monosubstituted phenoxy analogues of α_1-adrenoceptor antagonist WB4101. Bioorg Med Chem 13:2547–2559

Gálvez J, Garcìa R, Salabert MT, Soler R (1994) Charge indexes. New topological descriptors. J Chem Inf Comput Sci 34:520–525

Gálvez J, Garcìa-Domenech R, De Julián-Ortiz V, Soler R (1995) Topological approach to drug design. J Chem Inf Comput Sci 35:272–284

Gálvez J, Garcìa-Domenech R, de Gregorio Alapont C, De Julián-Ortiz V, Popa L (1996) Pharmacological distribution diagrams: a tool for de novo drug design. J Mol Graphics 14:272–276

Golan DE (2008) Principles of pharmacology: the pathophysiologic basis of drug therapy. Lippincott Williams & Wilkins, London

Golbraikh A, Tropsha A (2002) Beware of q2!. J Mol Graphics Mod 20:269–276

Goldberger JJ, Cain ME, Hohnloser SH, Kadish AH, Knight BP, Lauer MS et al (2008) American Heart Association/American College of Cardiology Foundation/Heart Rhythm Society Scientific Statement on Noninvasive Risk Stratification Techniques for identifying patients at risk for sudden cardiac death. A scientific statement from the American Heart Association Council on Clinical Cardiology Committee on Electrocardiography and Arrhythmias and Council on Epidemiology and Prevention. J Am Coll Cardiol 52:1179–1199

Graham RM, Perez DM, Hwa J, Piascik MT (1996) α_1-Adrenergic receptor subtypes molecular structure, function, and signaling. Cir Res 78:737–749

Gramatica P (2007) Principles of QSAR models validation: internal and external. QSAR Comb Sci 26:694

Hashimoto K (2007) Arrhythmia models for drug research: classification of antiarrhythmic drugs. J Pharmacol Sci 103:333–346

Hawkins DM, Basak SC, Mills D (2003) Assessing model fit by cross-validation. J Chem Inf Comput Sci 43:579–586

He Z, Huang L, Wu Y, Wang J, Wang H, Guo L (2008) DDPH: improving cognitive deficits beyond its α_1-adrenoceptor antagonism in chronic cerebral hypoperfused rats. Eur J Pharmacol 588:178–188

Huikuri HV, Castellanos A, Myerburg RJ (2001) Sudden death due to cardiac arrhythmias. N Engl J Med 345:1473–1482

Jain KS, Bariwal JB, Kathiravan MK, Phoujdar MS, Sahne RS, Chauhan BS et al (2008) Recent advances in selective α_1-adrenoreceptor antagonists as antihypertensive agents. Bioorg Med Chem 16:4759–4800

Kier LB, Hall LH (1999) Molecular structure descriptors: the electrotopological state. Academic Press, New York

Kiralj R, Ferreira MMC (2009) Basic validation procedures for regression models in QSAR and QSPR studies: theory and application. J Braz Chem Soc 20:770–787

Koshimizu T, Tanoue A, Tsujimoto G (2007) Clinical implications from studies of α_1 adrenergic receptor knockout mice. Biochem Pharmacol 73:1107–1112

Kromhout D (2007) Epidemiology of cardiovascular diseases in Europe. Public Health Nutr 4:441–457

Kubinyi H (1997a) QSAR and 3D QSAR in drug design Part 1: methodology. Drug Discovery Today 2:457–467

Kubinyi H (1997b) QSAR and 3D QSAR in drug design Part 2: applications and problems. Drug Discovery Today 2:538–546

Kulig K, Malawska B (2003) Estimation of the lipophilicity of antiarrhythmic and antihypertensive active 1-substituted pyrrolidin-2-one and pyrrolidine derivatives. Biomed Chromatogr 17:318–324

Kulig K, Nowicki P, Malawska B (2004) Influence of the absolute configuration on pharmacological activity of antihypertensive and antiarrhythmic drugs. Pol J Pharmacol 56:499–508

Kulig K, Sapa J, Maciag D, Filipek B, Malawska B (2007) Synthesis and pharmacological evaluation of new 1-[3-(4-arylpiperazin-1-yl)-2-hydroxypropyl]-pyrrolidin-2-one derivatives with anti-arrhythmic, hypotensive, and α-adrenolytic activity. Arch Pharm Chem Life Sci 340:466–475

Kulig K, Sapa J, Nowaczyk A, Filipek B, Malawska B (2009) Design, synthesis and pharmacological evaluation of new 1-[3-(4-aryl-piperazin-1-yl)-2-hydroxy-propyl]-3, 3-diphenylpyrrolidin-2-one derivatives with antiarrhythmic, antihypertensive, and α-adrenolytic activity. Eur J Med Chem 44:3994–4003

Leach AR (2001) Molecular modelling: principles and applications. Prentice-Hall, Englewood Cliffs

Malawska B, Kulig K, Filipek B, Sapa J, Maciąg D, Zygmunt M et al (2002) Synthesis, antiarrhythmic, and antihypertensive effects of novel 1-substituted pyrrolidin-2-one and pyrrolidine derivatives with adrenolytic activity. Eur J Med Chem 37:183–195

Malawska B, Kulig K, Gippert A, Filipek B, Sapa J, Maciąg D (2005) Synthesis and development of new 2-substituted 1-[3-(4-arylpiperazin-1-yl)propyl]-pyrrolidin-2-one derivatives with antiarrhythmic, hypertensive, and α-adrenolytic activity. II Farmaco 60:793–803

Matyus P, Varro A, Papp JG, Wamhoff H, Varga I, Virag L (1997) Antiarrhythmic agents: current status and perspectives. Med Res Rev 17:427–451

Nargund VH, Grey ADR (2008) Tamsulosin MR and OCAS (modified release and oral controlled absorption system): current therapeutic uses. Expert Opin Pharmacother 9:813–824

Nowaczyk A, Kulig K, Malawska B (2009) 1-(3-(4-arylpiperazin-1-yl)-propyl)-pyrrolidin-2-one derivatives as α_1-adrenoceptor antagonists: a QSAR studies. QSAR Comb Sci 28:979–988

Nowaczyk A, Przybylski R, Kulig K, Malawska B (2010) Structure–activity relationship studies of a number of α_1-adrenoceptor antagonists and antiarrhythmic agents. Mol Inf 29:343–351

Pallavicini M, Fumagalli L, Gobbi M, Bolchi C, Colleoni S, Moroni B et al (2006) QSAR study for a novel series of ortho disubstituted phenoxy analogues of α_1-adrenoceptor antagonist WB4101. Eur J Med Chem 41:1025–1040

Piascik MT, Soltis EE, Piascik MM, Macmillan LB (1999) α-Adrenoceptors and vascular regulation: molecular, pharmacologic and clinical correlates. Pharmacol Ther 72:215–241

Randic M (1980) Random walks and their diagnostic value for characterization of atomic environment. J Comput Chem 1:386–399

Razinger M (1986) Discrimination and ordering of chemical structures by the number of walks. Theor Chim Acta 70:365–378

Rios-Santamarina I, Garcìa-Domenech R, Gálvez J, Cortijo J, Santamaria P, Marcillo E (1998) New bronchodilators selected by molecular topology. Bioorg Med Chem Lett 8:477–482

Rücker G, Rücker C (1993) Counts of all walks as atomic and molecular descriptors. J Chem Inf Comput Sci 33:683–695

Rücker G, Rücker C (2000) Walk counts, labyrinthicity, and complexity of acyclic and cyclic graphs and molecules. J Chem Inf Comput Sci 40:99–106

STATISTICA StatSoft, Inc. (2008) STATISTICA (data analysis software system), version 8.0. http://www.statsoft.com

Szekeres L, Papp G (1975) Handbook of experimental pharmacology. Springer, Berlin

Talete srl, DRAGON for Windows Version 5.5-2007. http://www.talete.mi.it

Thiyagarajan M (2002) Alpha-adrenoceptor antagonists in the treatment of benign prostate hyperplasia. Pharmacology 65:119–128

Todeschini R, Consonni V (2002) Handbook of molecular descriptors. Wiley-VCH, New York

Todeschini R, Vighi M, Finizio A, Gramatica P (1997) 3D-modelling and prediction by WHIM descriptors. Part 8. Toxicity and

physico-chemical properties of environmental priority chemicals by 2D-TI and 3D-WHIM descriptors. SAR QSAR Environ Res 7:173–193

Topliss JG, Costello RJ (1972) Chance correlations in structure–activity studies using multiple regression analysis. J Med Chem 15:1066–1068

Trinajstic N (1992) Chemical graph theory. CRC Press, Boca

Tropsha A (2010) Best practices for QSAR model development, validation, and exploitation. Mol Inf 29:476–488

Tropsha A, Gramatica P, Gombar VK (2003) The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. QSAR Comb Sci 22:69–77

Turabekova MA, Rasulev BF, Levkovich MG, Abdullaev ND, Leszczynski J (2008) Aconitum and delphinium sp. alkaloids as antagonist modulators of voltage-gated Na+ channels AM1/DFT electronic structure investigations and QSAR studies. Comp Biol Chem 32:88–101

Vaughan Williams EM (1992) Classifying antiarrhythmic actions: by facts or speculation. J Clin Pharmacol 32:964–977

Wold S, Eriksson L (1995) Chemometric methods in molecular design. VCH, Weinheim

Zhong H, Minneman KP (1999) α1-Adrenoceptor subtypes. Eur J Pharmacol 375:261–276