DOI: 10.17344/acsi.2023.8081 *Acta Chim. Slov.* **2023,** *70,* 318–326

Scientific paper

Monte Carlo Optimization Based QSAR Modeling of Angiotensin II Receptor Antagonists

Nemanja Nikolić,[1] Tomislav Kostić,[2,3] Mlađan Golubović,[3,4] Tamara Nikolić,[1] Marija Marinković,[5] Velimir Perić,[4] Sara Mladenović[3] and Aleksandar M. Veselinović[6,*]

[1] *Policlinic "Dr Nikolić", Niš, Serbia.*

[2] *Clinic for Cardiology, Clinical Center Niš, Niš, Serbia.*

[3] *Faculty of Medicine, University of Niš, Niš, Serbia.*

[4] *Department for Cardiac Surgery, Clinic for Anaesthesiology and Intensive Therapy, Clinical Center Niš, Niš, Serbia.*

[5] *Clinic for Pulmology, Clinical Center Niš, Niš, Serbia*

[6] *Departmant of Chemistry, Faculty of Medicine, University of Nis, Nis, Serbia*

* *Corresponding author: E-mail: aveselinovic@medfak.ni.ac.rs*
Fax: +381 18 4238770; Phone: +381 18 4570029

Received: 02-16-2023

Abstract

The pathogenesis of essential hypertension, congestive heart failure, and reno-vascular hypertension is related to angiotensin II. This study presents QSAR modeling for a set of compounds acting as angiotensin II receptor antagonists based on the Monte Carlo optimization with molecular graph-based and SMILES notation based descriptors. Conformation independent QSAR models were developed for three random splits. Various statistical approaches were used to assess the statistical quality of the developed models, and the obtained results were very good. This study used a novel statistical metric known as the index of ideality of correlation for the final assessment of the model, and the results that were obtained suggested that the model was good. Also, molecular fragments which account for the increases and/or decreases of a studied activity were defined and then used for the computer-aided design of new compounds as potential angiotensin II receptor antagonists. The final assessment of the designed inhibitors, was performed with the use of molecular docking studies, highlighting exceptional correlation with the QSAR modeling results. The methodology which is presented in this research can be applied for seeking new agents for cardiovascular disorders treatment by angiotensin II receptor antagonism.

Keywords: Angiotensin II receptor antagonists; hypertension; QSAR; Molecular modeling; Drug design

1. Introduction

The pathogenesis of essential hypertension, congestive heart failure, and reno-vascular hypertension is related to angiotensin II associated with renin-angiotensin system (RAS), because (RAS) has important role in the regulation of cardiovascular homeostasis and electrolyte/fluid balance in both normotensive and hypertensive subjects.[1–3] This effect of angiotensin II could be associated with the mediation through selective membrane bound angiotensin II receptors type 1 (AT1) and type 2 (AT2) and this feature can be used for the treatment of above stated conditions with the application of of angiotensin-converting enzyme (ACE) inhibitors. RAS a major target for drug discovery programs in the pharmaceutical industry was established after clinical success of ACE inhibitors as therapeutics used for the treatment of hypertension and congestive heart failure.[4–6] Unfortunately the application of ACE inhibitors leads to occasional side effects, like angioneurotic edema and dry cough.[7–9] These side effects are related to the increase of bradykinin and substance P concentration, caused by the inhibition of

these peptides degradation.[10,11] To overcome this issue the alternative route has been suggested that will have the direct mode of intervening in the RAS with minimal potential side effect based on inhibition of the interactions of the primary effector hormone angiotensin II at the receptor level.[12–14] In light of the given facts, there is still a need to develop a reliable QSAR model for angiotensin II receptor antagonism that can be used to develop therapeutics for the treatment of hypertension and congestive heart failure.

A Monte Carlo optimization method in which the studied activity is treated as a random event has emerged as a promising approach in QSAR modeling in recent years. This method is based on the conformation-independent approach with optimal descriptors based on topological molecular features and molecules in the Simplified Molecular Input Line Entry System (SMILES) notation.[15–17] One of the primary advantages of the described method over more commonly used ones is its simplicity and efficiency. Also, this method can determine molecular fragments (calculated as SMILES notation descriptors) that have an influence on studied activity and that can be associated with the chemical structures of studied compounds. The main aim of this research is the development of a conformation-independent QSAR model based on the Monte Carlo optimization method for the angiotensin II receptor antagonism. Further, one of the main aims of this research was to define SMILES notation descriptors associated with molecular fragments that have both positive and negative influences on angiotensin II receptor antagonism. Molecular docking studies were used as the "final validator" of the established QSAR models and designed molecules antagonism potential

2. Materials and Methods

2. 1. Development and Validation of QSAR Models

As the first step in developing appropriate QSAR models, molecules obtained from literaturewere drawn using ACD/ChemSketch software v.11.0 and converted into the SMILES notation using the same software.[18,19] Chemical structures of compounds used for QSAR modeling with their SMILES notation are presented in Supporting Information and their general structures in Figure 1.

As the dependent variable for the development of the QSAR model, we used the inhibitor activities rabbit uterine membrane AT1 (IC_{50}) converted to -log10(IC_{50}) and given as pIC_{50} and this numerical values are presented in Table S1, Supplementary material. After we finished constructing the appropriate database, we made three different random splits of the main molecule database into two sets-the training set, which included 56 compounds (75%), and the test set with 19 compounds (25%), and we checked the normality of the activity distribution according to published method.[20] To establish conformation-independent QSAR models we applied software called CORAL (CORrelation and Logic, http://www.insilico.eu/coral) based on the Monte Carlo method and its algorithm, which treats the pertinent activity as a random event. We took into consideration two types of molecular descriptors based on the molecular graph and SMILES notation. Based on molecular graphs, invariants were defined as local graph invariants: Morgan extended connectivity index of increasing order (EC0), path numbers of length 2 and 3 (p2, p3), valence shells of range 2 and 3 (s2, s3), and the Code of Nearest Neighbors (NNCk). In recent years, Simplified Molecular Input-Line Entry System (SMILES) notation has become one of the most convenient representations, especially used in chemoinformatics because SMILES notation is considered as a very convenient alternative to the molecular graph. This fact is very appealing for medicinal chemistry since correlating molecular fragments with molecular graph-based descriptors can be quite challenging. For QSAR modeling, SMILES notation can be used to define molecular optimal descriptors (DCW), where DCW can be calculated as a function of SMILES notation according to Equation 1.

$$DCW(T,N_{epoch}) = \Sigma CW(S_k) + \Sigma CW(SS_k) + \Sigma CW(SSS_k) + \Sigma CW(EC0_k) + \Sigma CW(PT2_k) + \Sigma CW(PT3_k) + \Sigma CW(VS2_k) + \Sigma CW(VS3_k) + \Sigma CW(NNC_k) \quad (1)$$

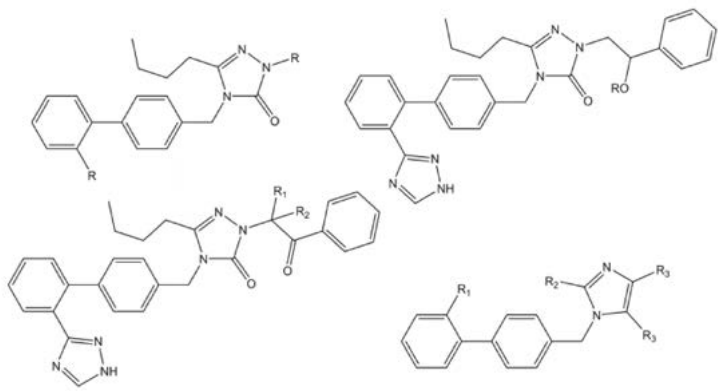


Figure 1. General chemical structures of molecules used for QSAR models development.

In this research, we used all SMILES notation based descriptors: local, global, and HARD-index. One of the main features of the developed QSAR model with the application of the Monte Carlo method is that we calculate correlation weight (CW), a numerical value for each used optimal descriptor.[17] The manner in which this process is achieved is based on generating suitable random numbers and observing how that fraction of numbers obeys some property or properties, in which CW values are randomly assigned to all used optimal descriptors, both molecular graph and SMILES notation based descriptors, in each independent Monte Carlo run. The Monte Carlo optimization process is applied further to calculate the numerical data for the correlation weights, which give the maximal value of the correlation coefficient between studied activity and used optimal descriptors. For this purpose, the Monte Carlo method uses two parameters: threshold (T) and the number of epochs (N_{epoch}). For the development of QSAR models, we used values of 0–10 for T and 0–70 for Nepoch, from which the search for the most predictive combination of T and N_{epoch} was concluded according to published methodology.[15–27] The development of a robust model capable of predicting the properties of new molecules in an objective, reliable, and precise manner is the main goal of any QSAR modeling process. We used the following methods to determine the goodness of the established QSAR models: internal validation using the training set, external validation using the validation set, and data randomization (Y-scrambling test). This was done by using statistical parameters such as the correlation coefficient (r^2), cross-validated correlation coefficient (q^2), standard error of estimation (s), mean absolute error (MAE), Fischer ratio (F), root-mean-square error (RMSE), R_m^2, and MAE-based metrics.[20–25] Recently, the so-called Index of Ideality of Correlation (IIC) has been suggested as a novel criterion for the estimation of the predictive potential of QSAR models, considering not only the correlation coefficient but also the arrangement of the cluster of dots-images relative to the diagonal, in coordinates observed-calculated values of the studied endpoint, and we calculated IIC according to Equations 2–5 as the QSAR model final estimator.[25–27]

$$\Delta_k = observed_k - calculated_k \quad (2)$$

Having data on all Δ_k for the test set, one can calculate sum of negative and positive values of Δ_k similar to mean absolute error (MAE):

$$^{-}MAE_{test} = \frac{1}{^{-}N}\sum_{k=1}^{^{-}N}|\Delta_k| \quad \Delta_k < 0,\ ^{-}N \text{ is the number of } \Delta_k < 0 \quad (3)$$

$$^{+}MAE_{test} = \frac{1}{^{+}N}\sum_{k=1}^{^{+}N}|\Delta_k| \quad \Delta_k \geq 0,\ ^{+}N \text{ is the number of } \Delta_k \geq 0 \quad (4)$$

$$IIC_{test} = r_{test} \times \frac{min(^{-}MAE_{test}, ^{+}MAE_{test})}{max(^{-}MAE_{test}, ^{+}MAE_{test})} \quad (5)$$

2. 2. Molecular Docking Studies

Molegro Virtual Docker (MVD) software was used for docking studies with geometrically optimized ligands using MMFF94 force field implemented in Marvin sketch (Marvin 6.1.0, 2013, ChemAxon) software. As the target for docking studies crystal structure of the angiotensin II type 2 receptoror (AT2R) (PDB: 7jni) was used. For, MVD rigid receptor structure and flexible structure for ligands was used for prerfoming docking studies. MVD yields both hydrophobic (mostly related to steric and Van der Waals interactions) and hydrophilic interactions, including identification of hydrogen bonds between amino acids from the active site and studied ligands. These interactions can be quantified through "scoring" functions, calculated numerical values related to relevant binding energies.[28] For most enzymes there is rule of thumb, the higher the interaction between receptor and ligand is the higher inhibition is achieved, so for this reason obtained numerical values for "scoring" functions could be used to assess the potential inhibition effect of studied ligands[29]. In this research following "scoring" functions were calculated and used further for inhibitory potential estimation: VdW, Steric, Hbond, NoHbond, Pose energy, Electro, ElectroLong, MolDock, and Rerank Score, and complete molecular docking protocol was validated according to published methodology.[31,32] Maestro Version 11.1.012, release 2017-1 was used for showing two-dimensional representations of the interactions between the studied molecules and the amino acids angiotensin II type 2 receptoror active site.

3. Results and Discussion

The applicability domain (AD) is a fundamental characteristic based on which the selection of molecules is done.[32–34] For defining AD we apply published methodology and we determined that all molecules in this study were within the range of AD defined and we did not identify any outliers[17]. Using the Least Squares method, the best developed QSAR models for the studied activity, regarding T and N_{epoch} values, are presented in the form of Eq. 6–8.

$$\text{Split 1: } pIC_{50} = 1.9668(\pm 0.0410) + 0.0580(\pm 0.0004)\times DCW(4,11) \quad (6)$$

$$\text{Split 2: } pIC_{50} = 2.0290(\pm 0.0447) + 0.1134(\pm 0.0009)\times DCW(4,12) \quad (7)$$

Table 1. The statistical quality of the developed QSAR models for angiotensin II receptor antagonism

		Training set						Test set					
		r^2	q^2	CCC	IIC	s	MAE	F	r^2	q^2	CCC	IIC	s	MAE	F
Split 1	1 run	0.8046	0.8917	0.7774	0.7902	0.388	0.294	222	0.8421	0.9104	0.9175	0.7842	0.402	0.315	91
	2 run	0.8145	0.8978	0.7278	0.8012	0.378	0.284	237	0.8755	0.9176	0.9347	0.8292	0.375	0.281	119
	3 run	0.8234	0.9032	0.8449	0.8119	0.368	0.288	252	0.8909	0.9217	0.9438	0.8551	0.362	0.268	139
	Av	0.8142	0.8976	0.7834	0.8011	0.378	0.289	237	0.8695	0.9166	0.9320	0.8228	0.380	0.288	116
Split 2	1 run	0.8878	0.9406	0.8773	0.8775	0.292	0.230	427	0.8029	0.8940	0.8960	0.7551	0.450	0.352	69
	2 run	0.8688	0.9298	0.8678	0.8593	0.315	0.239	358	0.8200	0.9036	0.9055	0.7752	0.450	0.362	77
	3 run	0.8951	0.9446	0.9461	0.8852	0.282	0.216	461	0.8231	0.9020	0.9072	0.7742	0.424	0.359	79
	Av	0.8839	0.9383	0.8971	0.8740	0.296	0.228	415	0.8153	0.9000	0.9029	0.7682	0.441	0.358	75
Split 3	1 run	0.8485	0.9180	0.8576	0.8353	0.360	0.297	302	0.8813	0.9046	0.9382	0.8558	0.335	0.283	126
	2 run	0.8814	0.9369	0.8136	0.8712	0.319	0.259	401	0.8465	0.8784	0.9189	0.8158	0.375	0.290	94
	3 run	0.8417	0.9140	0.9174	0.8290	0.368	0.300	287	0.8578	0.8896	0.9254	0.8297	0.360	0.288	103
	Av	0.8572	0.9230	0.8629	0.8452	0.349	0.285	330	0.8620	0.8909	0.9275	0.8338	0.357	0.287	108

r^2 – Correlation coefficient q^2 – Cross-validated correlation coefficient CCC – Concordance correlation coefficient IIC – Index of ideality of correlation s – Standard error of estimation
MAE – Mean absolute error F – Fischer ratio Av – Average value for statistical parameters obtained from three independent Monte Carlo optimization runs

Split 3: $pIC_{50} = 2.1274(\pm 0.0486) + 0.0659(\pm 0.0006) \times DCW(1,7)$ (8)

The values of the statistical metrics that helped us to determine the quality of the developed QSAR models for angiotensin II antagonism are available in Table 1. They indicate that the applied method was capable of establishing a QSAR model with good reproducibility, which we tested using the concordance correlation coefficient. We evaluated the predictability of the developed QSAR model using values presented in Table 1, and the developed model was proved valid. In addition, the model was classified as valid using MAE-based metrics. We performed the final evaluation of the developed QSAR models both for the training and the test set using the index of ideality of correlation and obtained values that suggest that developed QSAR models have a high predictive potential. Further, we applied Y-randomization, which implied scrambling of Y values in 1000 trials in ten separate runs, to assess the sturdiness of the developed QSAR models.[20] The obtained values presented in Table 2 indicate that there was no correlation by chance among the developed models. In regards to obtained values for statistical methods, we obtained the best QSAR model from the first split.

Table 2. Y-randomization of the best QSAR models (best optimization run) for three independent splits

	Split 1		Split 2		Split 3	
	Training	Test	Training	Test	Training	Test
0	0.8234	0.8909	0.8951	0.8231	0.8485	0.8813
1	0.0054	0.0403	0.0025	0.06	0.0822	0.0027
2	0.0849	0	0.0002	0.0066	0.0027	0.0496
3	0.0005	0.0001	0.0027	0.0009	0.0054	0
4	0.0004	0.0031	0.0681	0.0193	0.0054	0.0725
5	0.0059	0.0299	0.0146	0.081	0.0424	0.0866
6	0.0347	0.0086	0.0013	0.099	0.0027	0.0041
7	0.0068	0.013	0.0065	0.1115	0.0033	0.0056
8	0.0145	0.0756	0.0076	0.0792	0.0596	0.0048
9	0.0002	0.0128	0.0048	0.1635	0.0213	0.0021
10	0.0003	0.0141	0.0182	0.0039	0.0004	0.0255
R_r^2	0.0154	0.0198	0.0127	0.0625	0.0225	0.0254
$^{C}R_p^2$	0.8157	0.881	0.8887	0.7912	0.8371	0.8685

$^{C}R_p^2 = R \times (R^2 - R_r^2)^{1/2}$ should be > 0.5

Also, we observed that the best model was obtained with a T value of 4, whereas the best N_{epoch} value amounted to 11. The best Monte Carlo optimization runs (the highest value for r^2) for the developed QSAR models for all splits are shown in Figure 2 in the form of graphical representations.

Determining molecular fragments, defined as the SMILES notation optimal descriptors having a positive and negative influence on the examined activity, was

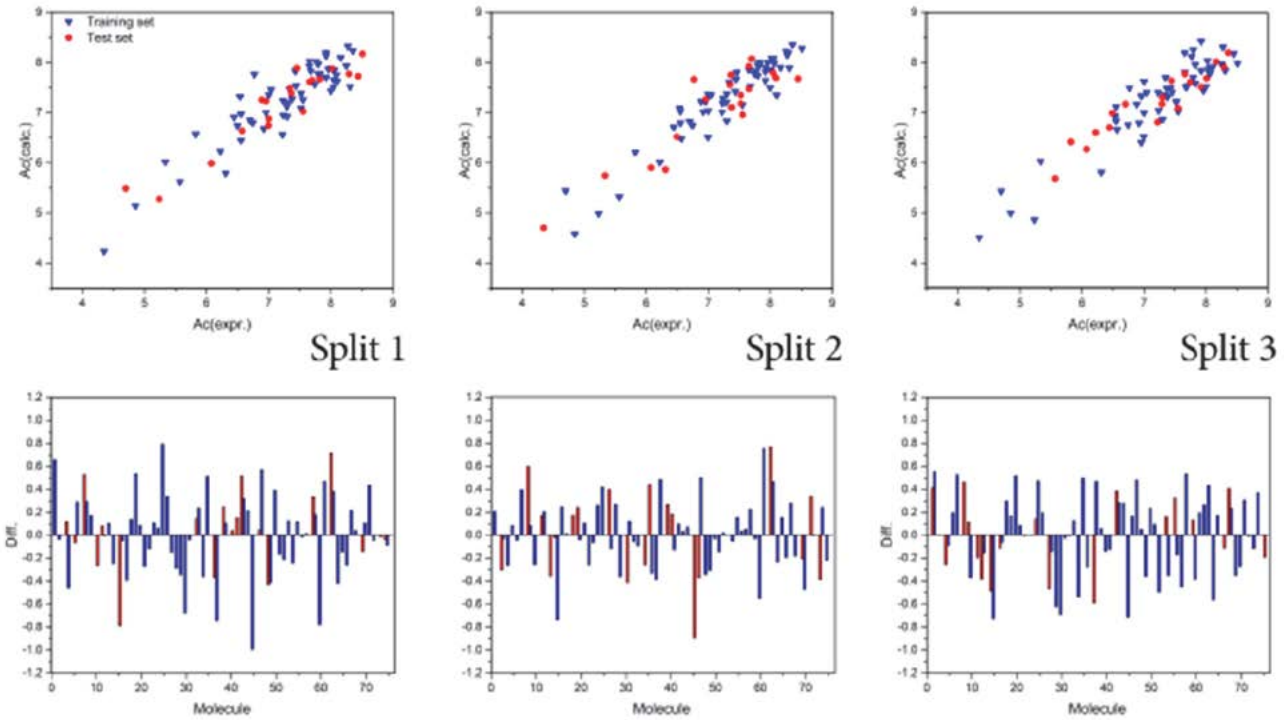


Figure 2. Graphical presentation of the best Monte Carlo optimization runs (the highest value for r^2) for the developed QSAR models.

among the main goals of this research.[17,29,35–38] Table S2 (Supplementary Material) contains the full list of molecular descriptors, which are based both on the SMILES notation and the molecular graph. The calculation example of molecule's both summarized correlation weight (DCW) and studied activity (pED_{50}) is presented in Table 3, where molecular graph-based descriptors were omitted with the aim of achieving an easier interpretation.

According to the published methodology we classified obtained SA_Ks as promoters of angiotensin II receptor antagonism.[17,31,32–35] In Table 4 we enlisted selected SA_Ks with their mechanistic interpretation while the complete list is given in Table S2 (Supporting Information). We presented the analysis of molecular fragments' contribution to angiotensin II receptor antagonism in Figure 3. In presented Figure, green color indicate groups that have positive, while red color indicate groups that have negative influence on corneal permeability. As already stated each SA_K contributes with its CW value.

The computer-aided design of five new potential antagonists whose structures presented in Figure 4 was generated from the conformational-independent results obtained from developed QSAR models. The template molecule was molecule **A**, a molecule taken from initial data base, since it is one of the least chemically exploited molecules. Table 4 contains the list of all the designed molecules, as well as their calculated values for the pIC_{50}.

Based on the obtained results from QSAR modelling, the SMILES notation descriptors associated with molecular fragments with a positive impact on pIC_{50} for angiotensin II receptor antagonism activity and that yield increase in its activity are: "C....." – carbon atom or a methyl group, and "O....." – oxygen atom or a hydroxyl group, where both fragments with positive impact on pIC_{50} numerical value and whose addition lead to the increase of calculated pIC_{50} values for molecule **A1** in comparison to calculated pEC_{50} values for template molecule **A**; also molecule **A1** had additional fragments related to molecular branching – "(....." and "(...C....." both promoters of pIC_{50} increase; further molecular branching was obtained with molecule **A2** with further addition of "C.....", "(....." and "(...C....." fragments that lead to further increase of pIC_{50} numerical value. In molecules **A3** and **A4** oxygen atom was changed with nitrogen atom

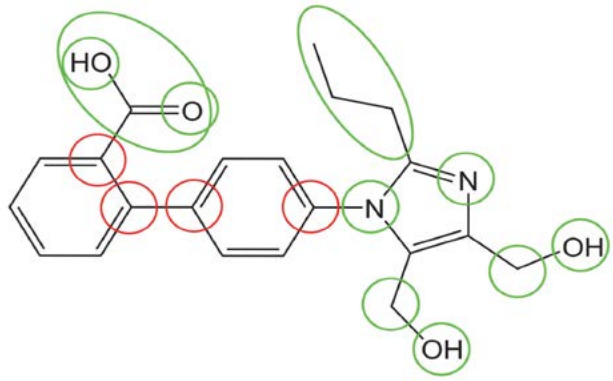
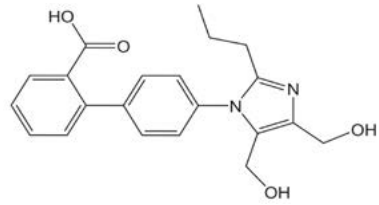


Figure 3. Molecular fragments contribution to angiotensin II receptor antagonism (green – increase, red – decrease).

Table 3. Example of DCW calculation



SMILES notation: O=C(O)c1ccccc1c1ccc(cc1)n1c(CO)c(nc1CCC)CO
DCW: 63.10893
Ac(calc.): 5.6237

O.....	0.3453	c.....	−0.0995	c...(.....	0.3361	c...c...(...	−0.8827
=.....	0.1241	1.....	0.2002	C...(.....	2.0522	c...c...1...	0.1897
C.....	0.905	C.....	0.905	O...C.....	−0.6716	c...1...(...	0.1978
(.....	−0.9914	C.....	0.905	O...(.....	1.0042	n...(...1...	0.4829
O.....	0.3453	C.....	0.905	c...(.....	0.3361	1...n...(...	0.9117
(.....	−0.9914	(.....	−0.9914	c...(.....	0.3361	n...1...c...	0.2376
c.....	−0.0995	C.....	0.905	n...(.....	0.1064	1...c...(...	−1.0596
1.....	0.2002	O.....	0.3453	n...c.....	0.2222	c...(...C...	0.6814
c.....	−0.0995	O...=.....	−0.2385	c...1.....	−0.927	O...C...(...	0.3324
c.....	−0.0995	C...=.....	−1.4962	C...1.....	0.4536	C...O...(...	−0.5291
c.....	−0.0995	C...(.....	2.0522	C...C.....	0.2526	c...(...O...	−0.2309
c.....	−0.0995	O...(.....	1.0042	C...C.....	0.2526	(...c...(...	0.3821
c.....	−0.0995	O...(.....	1.0042	C...(.....	2.0522	n...(...c...	−6.9698
1.....	0.2002	c...(.....	0.3361	C...(.....	2.0522	c...n...(...	1.5402
c.....	−0.0995	c...1.....	−0.927	O...C.....	−0.6716	n...c...1...	−0.0889
1.....	0.2002	c...1.....	−0.927	O...=...C...	0.4779	c...1...C...	0.1189
c.....	−0.0995	c...c.....	−0.2115	=...C...(...	−0.5437	C...C...1...	−0.6848
c.....	−0.0995	c...c.....	−0.2115	O...(...C...	0.4236	C...C...C...	−0.5269
c.....	−0.0995	c...c.....	−0.2115	(...O...(...	0.1847	C...C...(...	0.3714
(.....	−0.9914	c...c.....	−0.2115	c...(...O...	−0.2309	C...(...C...	0.3742
c.....	−0.0995	c...1.....	−0.927	1...c...(...	−1.0596	O...C...(...	0.3324
c.....	−0.0995	c...1.....	−0.927	c...1...c...	0.1224	Cmax.1.....	−0.4456
1.....	0.2002	c...1.....	−0.927	c...c...1...	0.1897	Nmax.0.....	0.3887
(.....	−0.9914	c...1.....	−0.927	c...c...c...	−0.5425	Omax.4.....	−0.2055
n.....	0.171	c...c.....	−0.2115	c...c...c...	−0.5425	Smax.0.....	2.5295
1.....	0.2002	c...c.....	−0.2115	c...c...c...	−0.5425	NOSP01000000	0.3921
c.....	−0.0995	c...(.....	0.3361	c...c...1...	0.1897	HALO00000000	−5.5836
(.....	−0.9914	c...(.....	0.3361	c...1...c...	0.1224	BOND10000000	1.7716
C.....	0.905	c...c.....	−0.2115	1...c...1...	0.0916	++++N---O===	9.883
O.....	0.3453	c...1.....	−0.927	c...1...c...	0.1224	++++O---B2==	1.4457
(.....	−0.9914	1...(.....	−1.147	c...c...1...	0.1897	++++N---B2==	1.2326
c.....	−0.0995	n...(.....	0.1064	c...c...c...	−0.5425	10001000000	0.2082
(.....	−0.9914	n...1.....	−0.0595	c...c...(...	−0.8827		
n.....	0.171	c...1.....	−0.927	c...(...c...	1.4185		

leading to switching of "O....." and "(...O....." fragments with "N....." and "(...N....." both with pIC_{50} numerical value increase feature. Both molecules **A3** and **A4** have higher value for pIC_{50} in comparison to molecule A pIC_{50} numerical value. Since fragments "O....." and "(...O....." have higher numerical values for CW in comparison to "N....." and "(...N....." CW numerical values, calculated values for molecules **A1** and **A2** pIC_{50} were higher in comparison to molecules **A3** and **A4** pIC50 values. Molecule **A5** has additional "O...C...(...", "O...=...C...", "O...C....." fragments, all promoters of pIC_{50} increase, in comparison to molecule **A**, leading to pIC_{50} numerical value increase.

To assess the developed QSAR models' predictability and to validate them further, all designed molecules and template molecule A were subjected to molecular docking studies with angiotensin II type 2 receptoror. Numerical values for all calculated "scoring" functions are presented in Table 6. When the assessment of the inhibitory potency is made different scoring functions should be taken into consideration, since they are related to different ligand-amino acids interactions. According to obtained re-

Table 4. Mechanistic interpretation of selected SA_Ks

SA_k	Increase
C.....	Carbon atom
N.....	Nitrogen atom
O.....	Oxygen atom
(...(..... (..... (...C...(... (...N...(... N...(..... (...O...(... C...(...C...	Branching in molecule as such, branching in molecule on either carbon, nitrogen or oxygen atom
O...C...(... O...=...C... O...C.....	Fragments associated with carboxyl group
	Decrease
1..... 2..... C...(...1...	Presence of one or two rings in molecule
O...=.....	Oxygen atom with double bond
c...(.....	Branching on benzyl group

sults for MolDock and ReRank "score" functions molecule with the potentially highest inhibitory activity is molecule A2 and this result is in correlation with the results from QSAR modeling. The molecule with the lowest values for MolDock and ReRank "score" functions was template molecule A, which is also in good correlation with the results obtained from QSAR modeling. The detailed definitions of other "scoring" functions and their potential impact on inhibitory activity can be found in the literature[27].

Highest energy related to close electrostatic interactions, calculated with Electro "scoring" function, is indentified for molecules M2 and lowest for molecule M5. Further, both highest and lowest energies related to long electrostatic interactions, calculated with ElectroLong "scoring" function, were identified for same molecules as for Electro "scoring" function. Highest energy related to hydrogen bonds is indentified for molecules A4 and lowest for molecule A5. Also same molecules had the highest and lowest energy related to the hydrogen bonding energy (protein-ligand) as calculated if the directionality of the hydrogen bond was not taken into account calculated with NoHBond90 "scoring" function. Highest energy related to steric interactions, calculated with Steric "scoring" function, is indentified for molecules A1 and lowest for molecule A. For Van der Walls energies the highest values was calculated, with application of VdW "scoring" function, for molecule A4, and lowest for molecule A4. Highest energy from overall interactions between ligand and receptor, calculated with Energy "scoring" function, was obtained for molecule A2 and lowest for molecule A.

All interactions between the selected molecules and amino acids from angiotensin II type 2 receptoror active site are identified and 2D representation of hydrogen bonds, hydrophobic, and hydrophilic interactions inside the binding pocket are presented in Figures in the Supplementary Information section, while the best-calculated poses for all designed molecules inside the active site of angiotensin II type 2 receptoror are presented in Figure 5. According to obtained results there are two clusters of molecules inside angiotensin II type 2 receptoror active site. Molecules A, A2, A4 and A5 (cluster 1) were docked in one part of active site, while molecules A1 and A3 (cluster 2) in other. Molecules from cluster 1 had hydrogen bonds with amino acids ARG182, LYS215

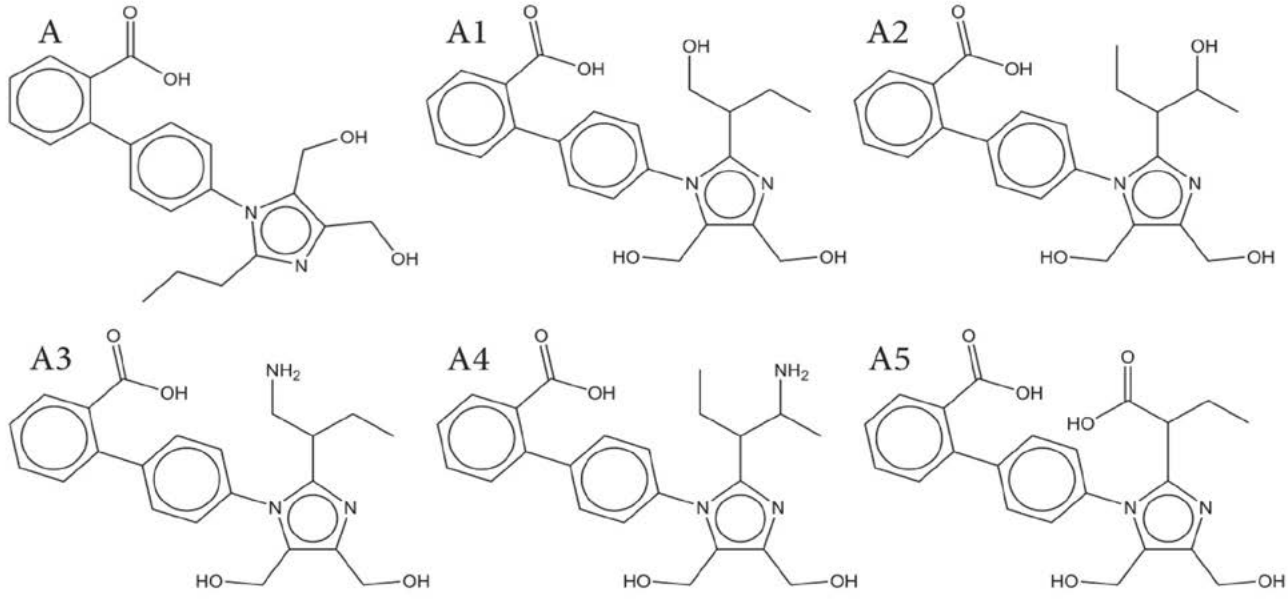


Figure 4. Chemical structures of designed molecules.

Table 5. The list of all designed molecules with their SMILES notation and calculated activities

Molecule	SMILES Notation	Ac(calc.)
A	O=C(O)c1ccccc1c1ccc(cc1)n1c(CO)c(nc1CCC)CO	5.6237
A1	O=C(O)c1ccccc1c1ccc(cc1)n1c(CO)c(nc1C(CC)CO)CO	6.2100
A2	O=C(O)c1ccccc1c1ccc(cc1)n1c(CO)c(nc1C(CC)C(C)O)CO	6.7472
A3	O=C(O)c1ccccc1c1ccc(cc1)n1c(CO)c(nc1C(CC)CN)CO	6.1397
A4	O=C(O)c1ccccc1c1ccc(cc1)n1c(CO)c(nc1C(CC)C(C)N)CO	6.5279
A5	O=C(O)c1ccccc1c1ccc(cc1)n1c(CO)c(nc1C(CC)C(=O)O)CO	6.4697

Table 6. Score values (kcal/mol) for all computer-aided designed compounds

Molecule	Electro	ElectroLong	Steric	VdW	HBond	NoHBond90	Energy	PoseEnergy	RerankScore
A	−8.03766	−2.77855	−139.592	−45.6996	−9.69917	−9.71146	−151.25	−145.534	−112.478
A1	−1.31333	−1.95883	−154.897	−46.1529	−5	−5.38596	−158.14	−150.184	−123.34
A2	−10.8113	−5.73813	−148.83	−42.3753	−9.10444	−10.3858	−167.669	−157.889	−125.207
A3	−8.75907	−2.86037	−152.509	−41.8093	−7.5	−8.39838	−156.857	−150.872	−121.193
A4	−8.3428	−2.78078	−153.414	−26.587	−11.4747	−11.5507	−157.525	−155.959	−121.392
A5	−1.29674	−1.9214	−152.925	−44.0886	−2.5	−2.5	−156.511	−147.798	−120.706

and ILE304, while molecules from cluster 1 had hydrogen bonds with amino acid PRO301. Also molecules from cluster 2 shown π-π interactions with amino acid TRP100.

4. Conclusion

Developing robust QSAR models for angiotensin II receptor antagonism that possess good predictability, which is determined by utilizing various statistical parameters, represents the main aim of this research. Calculations of the conformation independent models, which were developed in accordance with the optimal descriptors and derived from a local graph and the SMILES notation invariants, were performed by employing the Monte Carlo optimization method. Applying a range of statistical techniques yielded the evaluation of the developed QSAR models' predictive potential and robustness. The high applicability of the developed QSAR models is displayed by the realized numerical values applied to validate the mentioned. The Monte Carlo optimization method successfully determined molecular fragments, used in QSAR modeling as the SMILES notation fragments with a positive and negative effect on angiotensin II receptor antagonism and the mentioned were used for the computer-aided design of novel compounds with higher pIC_{50} values. The final validator of the developed QSAR model and the designed molecules' potential inhibitory effect were the molecular docking studies, and the obtained results show good inter-correlation. In summary, new therapeutics for the treatment of hypertension and congestive heart failure can be sought by applying the methodology presented in this research.

We have no conflict of interest to disclose.

Acknowledgments

This work is supported by the Ministry of Education and Science, the Republic of Serbia and the Faculty of Medicine, University of Niš, Republic of Serbia (Internal project No. 70). The authors would like to thank the Ministry of Education, Science and Technological Development of Republic of Serbia (Grant No: 451-03-68/2022-14/200113) for financial support.

5. References

1. C. M. Ferrario, *J. Cardiovasc. Pharmacol.* **1990**, *15*, 51–55 **DOI:**10.1097/00005344-199000153-00001
2. M. Burnier, *Circulation* **2001**, *103*, 904–912 **DOI:**10.1161/01.CIR.103.6.904
3. M. Burnier, H.R. Brunner, *Lancet* **2003**, *355*, 637–645 **DOI:**10.1016/S0140-6736(99)10365-9
4. V. Agrawal, J. K. Gupta, S.S. Qureshi, V. K. Vishwakarma, *Indian Heart J.* **2016**, *68*, 856–861. **DOI:**10.1016/j.ihj.2016.06.010
5. B. Raymer, D. Ebner, *Expert Opin Ther Pat.* **2015**, *25*, 1175–1190. **DOI:**10.1517/13543776.2015.1061997
6. M. A. Zaman, S. Oparil, D. A. Calhoun, *Nat. Rev. Drug Discov.* **2002**,*1*, 621–636. **DOI:**10.1038/nrd873
7. C. M. Ferrario, A. E. Mullick, *Pharmacol. Res.* **2017,** *125*, 57–71. **DOI:**10.1016/j.phrs.2017.05.020
8. A. Nehme, K. Zibara, *Hypertens. Res.* **2017**, *40*, 903–909. **DOI:**10.1038/hr.2017.65
9. H. Cheng, R.C. Harris, *Expert Opin. Drug Saf.* **2006**, *5*, 631–641. **DOI:**10.1517/14740338.5.5.631

10. R. Igic, R. Behnia, *Curr. Pharm. Des.* **2007**, *13*, 1199–1214. **DOI:**10.2174/138161207780618876
11. D. Regoli, G. E. Plante, F. Gobeil Jr, *Pharmacol. Therapeut.* **2012**, *135*, 94–111. **DOI:**10.1016/j.pharmthera.2012.04.002
12. C. Guang, R. D. Phillips, B. Jiang, F. Milani, *Arch. Cardiovasc. Dis.* **2012**, *105*, 373–385. **DOI:**10.1016/j.acvd.2012.02.010
13. K. Hanif, H. K. Bid, R. Konwar, *Hypertens. Res.* **2010**, *33*, 11–21. **DOI:**10.1038/hr.2009.184
14. J. Kućmierz, W. Frąk, E. Młynarska, B. Franczyk, J. Rysz, *Int. J. Mol. Sci.* **2021**, *22*, 9669. **DOI:**10.3390/ijms22189669
15. M. A. Toropova, I. Jr. Raška, A. A. Toropov, M. Raškova, *Comb. Chem. High T. Scr.* **2016**, *19*, 676–687. **DOI:**10.2174/1386207319666160725145852
16. A. P. Toropova, A. A. Toropov, *Mini-Rev. Med. Chem.* **2018**, *18*, 382–391. **DOI:**10.2174/1389557517666170927154931
17. A. M. Veselinović, J. B. Veselinović, J. V. Živković, G. M. Nikolić, *Curr. Top. Med. Chem.* **2015**, *15*, 1768–1779.
18. M. C. Sharma, *Interdiscip. Sci.* **2016**, *8*, 1–0. **DOI:**10.1007/s12539-016-0176-5
19. A. Parate, S. C. Chaturvedi, *Med. Chem. Res.* **2012**, *21*, 1166–1178. **DOI:**10.1007/s00044-011-9622-4
20. P. K. Ojha, K. Roy, *Chemometr. Intell. Lab.* **2011**, *109*, 146–161. **DOI:**10.1016/j.chemolab.2011.08.007
21. P. K. Ojha, I. Mitra, R. N. Das, K. Roy, *Chemometr. Intell. Lab.* **2011**, *107*, 194–205. **DOI:**10.1016/j.chemolab.2011.03.011
22. P. P. Roy, J. T. Leonard, K. Roy, *Chemometr. Intell. Lab.* **2008**, *90*, 31–42. **DOI:**10.1016/j.chemolab.2007.07.004
23. K. Roy, R. N. Das, P. Ambure, R. B. Aher, *Chemometr. Intell. Lab.* **2016**, *152*, 18–33. **DOI:**10.1016/j.chemolab.2016.01.008
24. L. I. Lin, *Biometrics*, **1989**, *45*, 255–268. **DOI:**10.2307/2532051
25. V. Stoičkov, D. Stojanović, I. Tasić, S. Šarić, D. Radenković, P. Babović, D. Sokolović, A. M. Veselinović, *Struct. Chem.* **2018**, *29*, 441–449. **DOI:**10.1007/s11224-017-1041-9
26. A. A. Toropov, A. P. Toropova, *Mutat. Res.- Gen. Tox. En.* **2017**, *819*, 31–37. **DOI:**10.1016/j.mrgentox.2017.05.008
27. A. M. Veselinović, A. Toropov, A. Toropova, D. Stanković-Đorđević, J. B. Veselinović, *New J. Chem.* **2018**, *42*, 10976–10982. **DOI:**10.1039/C8NJ01034J
28. R. Thomsen, M. H. Christensen, *J. Med. Chem.* 2006, 499, 3315–3321. **DOI:**10.1021/jm051197e
29. M. Zivkovic, M. Zlatanovic, N. Zlatanovic, M. Golubović, A. M. Veselinović, Mini-Rev. *Med. Chem.* 2020, 20, 1389–1402. **DOI:**10.2174/1389557520666200212111428
30. S. A. Amin, N. Adhikari, S. Gayen, T. Jha, *J. Biomol. Struct. Dyn.* 2017, 36, 590–608. **DOI:**10.1080/07391102.2017.1288659
31. S. A. Amin, N. Adhikari, S. Gayen, T. Jha, *J. Biomol. Struct. Dyn.* 2019, 37, 4528–4541. **DOI:**10.1080/07391102.2018.1552895
32. D. Gadaleta, G.F. Mangiatordi, M. Catto, A. Carotti, O. Nicolotti, *IJQSPR*, **2016**, *1*, 45–63. **DOI:**10.4018/IJQSPR.2016010102
33. P. Gramatica, *QSAR Comb. Sci.* **2007**, *26*, 694–701. **DOI:**10.1002/qsar.200610151
34. P. Gramatica, A. Sangion, *J. Chem. Inf. Model.* **2016**, *56*, 1127–1131. **DOI:**10.1021/acs.jcim.6b00088
35. P. Kumar, A. Kumar, J. Sindhu *SAR QSAR Environ. Res.* **2019**, *30*, 525–541. **DOI:**10.1080/1062936X.2019.1629998
36. S. Ahmadi, S. Lotfi, S. Afshari, P. Kumar, E. Ghasemi, *SAR QSAR Environ. Res.* **2021**, *32*, 1013–1031. **DOI:**10.1080/1062936X.2021.2003429
37. S. Ahmadi, S. Lotfi, P. Kumar, *Toxicol. Mech. Methods* **2022**, *32*, 302–312. **DOI:**10.1080/15376516.2021.2000686
38. S. Lotfi, S. Ahmadi, P. Kumar, *RSC Adv.* **2021**, *11*, 33849–33857. **DOI:**10.1039/D1RA06861J

Povzetek

Patogeneza esencialne hipertenzije, kongestivnega srčnega popuščanja in renovaskularne hipertenzije je povezana z angiotenzinom II. Ta študija predstavlja modeliranje QSAR za nabor spojin, ki delujejo kot antagonisti receptorjev angiotenzina II, na podlagi optimizacije Monte Carlo z deskriptorji na osnovi molekularnih grafov in zapisov SMILES. Konformacijsko neodvisni modeli QSAR so bili razviti za tri naključne razdelitve. Za oceno statistične kakovosti razvitih modelov smo uporabili različne statistične pristope, dobljeni rezultati pa so bili zelo dobri. Za končno oceno modela smo uporabili novo statistično metriko, znano kot indeks idealnosti korelacije, in dobljeni rezultati kažejo, da je bil model dober. Prav tako so bili definirani molekularni fragmenti, ki so odgovorni za povečanja in/ali zmanjšanja proučevane aktivnosti, in nato uporabljeni za računalniško podprto načrtovanje novih spojin kot potencialnih antagonistov receptorjev angiotenzina II. Končna ocena načrtovanih zaviralcev je bila izvedena z uporabo študij molekularnega sidranja, ki poudarjajo izjemno visoko stopnjo korelacije z rezultati modeliranja QSAR. Metodologijo, ki je predstavljena v tej raziskavi, je mogoče uporabiti pri iskanju novih učinkovin za zdravljenje srčno-žilnih obolenj z antagonizmom receptorjev angiotenzina II.