

Scientific paper

Development of Novel Analgesics Related to TRPV1 Antagonism – *In Silico* Approach

Mladjan Golubovic^{1,2}, Velimir Peric^{1,2}, Marija Stosic^{1,2}, Vladimir Stojiljkovic^{1,2}, Tomislav Kostic^{1,3}, Aleksandar Kamenov^{1,2}, Milan Lazarevic^{1,2}, Dalibor Stojanovic², Aleksandar M. Veselinovic^{4*}

¹ Faculty of Medicine, University of Niš, Niš, Serbia

² University Clinical Center Niš, Clinic for Cardiosurgery, Niš, Serbia

³ University Clinical Center Niš, Clinic for Cardiovascular Disease, Niš, Serbia

⁴ Faculty of Medicine, Department of Chemistry, University of Niš, Niš, Serbia

* Corresponding author: E-mail: aveselinovic@medfak.ni.ac.rs

Fax: +381 18 4238770; Phone: +381 18 4570029

Received: 02-08-2025

Abstract

In the context of pharmacological intervention for pain, Transient Receptor Potential Vanilloid, member 1 (TRPV1), as a non-selective cation channel belonging to the transient receptor potential (TRP) family of ion channels, has emerged as a promising target. However, the availability of selective TRPV1 antagonists and their associated pharmacological properties remains limited. This research paper explores various QSAR modeling techniques applied to a range of piperazinyl-aryl compounds acting as TRPV1 antagonists. The descriptors utilized in the creation of conformation-independent QSAR models included local molecular graph invariants and the SMILES notation, along with the incorporation of the Monte Carlo optimization method as a model development technique. Several statistical methods were employed to evaluate the quality, robustness, and predictive capacity of the developed models, yielding positive results. For the best developed QSAR model following statistical parameters were obtained for training set $R^2 = 0.7155$, CCC = 0.8134, IIC = 0.7430, $Q^2 = 0.6970$, RMSE = 0.645, MAE = 0.489 and F = 157; and for test set $R^2 = 0.9271$, CCC = 0.9469, IIC = 0.9635, $Q^2 = 0.9241$, RMSE = 0.367, MAE = 0.329 and F = 328. Additionally, molecular fragments derived from SMILES notation descriptors, which explain observed changes in the evaluated activity, were identified, leading to the design of four new antagonists. The final validation of the QSAR model and the designed antagonists was conducted through molecular docking, which demonstrated strong correlation with the QSAR modeling results.

Keywords: TRPV1, Pain treatment, QSAR, Molecular modeling, Drug design

1. Introduction

TRPV1 (transient receptor potential cation channel, subfamily V, member 1) is a non-selective cation channel activated by a variety of exogenous and endogenous stimuli. Cloning of TRPV1 and demonstration of its therapeutic value have led to intense research in understanding the molecular mechanisms encompassing the responses of sensory neurons to stimuli such as heat, protons, some endogenous activators such as anandamide and exogenous activators such as capsaicin^{1–5}. TRPV1, also known as VR1 (Vanilloid Receptor 1), is primarily expressed on

unmyelinated pain-sensing nerve fibers (C-fibers) and small A fibers in the dorsal root and trigeminal ganglia^{6,7}. Activation of the channel leads to an influx of calcium and sodium ions into the cell, causing depolarization, and that in turn results in the excitation of primary sensory neurons and ultimately in the perception of pain, though a reduction of the activation thresholds of the channel to other stimuli indicate that agonists of the channel can cause desensitization with therapeutic application in the management of pain. However, such desensitization leads to the side effects of burning sensation, irritation, and neurotoxicity, resulting from continuous influx of calci-

um ions into the cells, thus limiting the clinical use of agonists^{8–10}. Blocking the TRPV1-mediated pain signaling pathways with receptor antagonists is an alternative promising strategy for the development of novel analgesic drugs with potentially fewer side effects^{11–14}. The chemistry and pharmacology of several classes of competitive TRPV1 antagonists, based on chemotypes containing thiourea, urea, and amide groups, have been described and reviewed^{15–18}. They have been classified into four profiles based on their ability to differentially modulate TRPV1 activation by different modes such as capsaicin, pH 5 and heat¹⁹.

Over the years, many antagonists that have entered and progressed into various clinical phases have failed due to the development of hyperthermia as an undesired on-target side-effect. With increased structural, mechanistic and biophysical knowledge on TRPV1, there is an unrecognized need to find more and diverse chemotypes for TRPV1 antagonists. New chemotypes are needed to evolve functionally-selective or modality-selective antagonists that can probe and prove if hyperthermia can be dissociated from the other functions of this ion channel^{21,22}. The need for new chemotypes is encouraged by the fact that there are numerous new indications attributed to TRPV1 for which small molecule chemical probes are required for validation^{23–25}. Apart from validation for newer indications, probes are required to decipher the potential consequences of targeting a single indication. In this regard, a generally applicable TRPV1 antagonist pharmacophore that can enable discovery of novel chemotypes is critical.

Drug discovery and development are inherently time-consuming processes, requiring substantial time, effort, and financial resources. These challenges arise from the need to identify effective and safe therapeutic compounds through extensive experimental screening, optimization, and validation. To mitigate these constraints, chemoinformatics has emerged as an essential discipline, leveraging computational tools and *in silico* methodologies to accelerate various stages of the drug development pipeline. Chemoinformatic approaches offer a wide range of applications, including the identification of novel lead compounds, optimization of pharmacological activity, and improvement of pharmacokinetic and toxicological profiles of compounds with known biological activity^{26–28}. Among the numerous chemoinformatic techniques, quantitative structure-activity relationship (QSAR) modeling is recognized as the most extensively applied and impactful approach. QSAR studies aim to establish predictive relationships between the chemical structure of compounds and their biological activity, providing valuable insights into molecular interactions and activity mechanisms. Contemporary QSAR models are developed using a variety of molecular descriptors, which are computationally derived parameters that capture distinct chemical, physical, and structural features of molecules. These descriptors vary widely, encompassing properties such as electronic, steric, hydrophobic, and topological characteristics. The

construction of QSAR models involves a systematic process. Initially, molecular descriptors are calculated from well-defined molecular structures, each descriptor contributing unique insights into the molecule's behavior. These descriptors are then analyzed to identify their relevance and predictive capacity, enabling the development of robust mathematical equations that quantitatively relate molecular descriptors to observed biological activities. Such models provide a powerful framework for predicting the activity of untested compounds, guiding the rational design of new drugs, and prioritizing experimental validation.

Despite their strengths, QSAR models are not without limitations. The accuracy and reliability of a QSAR model depend heavily on the quality and diversity of the training dataset, the choice of molecular descriptors, and the statistical methods employed. Moreover, the interpretability of these models can sometimes be challenging, especially when complex machine learning algorithms are used. Nevertheless, QSAR remains a cornerstone of modern drug discovery, contributing to significant time and cost savings by enabling the prioritization of promising candidates for further experimental testing^{29–33}. As computational power and chemoinformatics methodologies continue to evolve, the integration of QSAR with other *in silico* approaches, such as molecular docking, pharmacophore modeling, and machine learning, promises to further enhance its predictive capabilities. This integration will likely play a pivotal role in addressing the growing demand for efficient, cost-effective, and innovative drug development strategies.

This research employed various *in silico* methods to identify novel compounds with potential antagonistic effects on the TRPV1 receptor. The study focused on developing QSAR models using conformation-independent molecular descriptors derived from SMILES notation and local graph invariants, integrated with the Monte Carlo optimization method. A key objective was to identify molecular fragments or structural features responsible for TRPV1 antagonism and to explore correlations among the different computational approaches used. The study successfully pinpointed molecular fragments present in small molecules that are critical for ligand-receptor interactions. These findings provide valuable insights into the structural basis of TRPV1 antagonism, offering a foundation for the rational design and development of novel analgesic agents. The identified fragments could be strategically employed to enhance the efficiency and specificity of future drug discovery efforts targeting TRPV1-related pathways.

2. Materials and Method

In this study, a dataset comprising 98 molecules known to TRPV1 antagonism effect was collected from the scientific literature³⁴. The compounds analyzed in

this study were evaluated for their ability to inhibit capsaicin- (CAP, 500 nM) or acid- (pH 5.0) induced uptake of $^{45}\text{Ca}^{2+}$ in Chinese Hamster Ovary (CHO) cells stably expressing rat TRPV1 (rTRPV1), as previously described^{35,36}. Functional activity was reported as $\text{IC}_{50} \pm \text{SEM}$ (nM), based on experimental measurements. These IC_{50} values were collected from [insert database or literature reference], and converted to pIC_{50} values using the standard equation: $\text{pIC}_{50} = -\log_{10}(\text{IC}_{50} \times 10^{-9})$. The dataset was curated to remove duplicates, inconsistent records, and extreme outliers prior to model development. The SMILES notation for all the molecules used in the study, along with their corresponding pIC_{50} values, is provided in Table S1 within the Supplementary Material. To ensure the robustness of the analysis, the dataset was randomly divided into three random splits with two sets: a training set consisting of 73 compounds (75%) and a test set comprising 25 compounds (25%). The normality of the activity distribution for all the dataset splits was assessed following the methodology described in a published reference³⁷.

2. 1. QSAR Modeling Utilizing the Monte Carlo Optimization Method

The Monte Carlo optimization method was employed to construct a conformation-independent QSAR model using a hybrid approach that combined molecular graph-based and SMILES notation-based descriptors. The molecular graph-based descriptors included local graph invariants derived from fundamental graph theory concepts, such as paths and walks, with detailed mathematical definitions available in the literature³⁸. The optimal descriptors identified from the graph-based approach encompassed Morgan extended connectivity indices (EC0), valence shell descriptors for ranges 2 and 3 (s2, s3), path numbers for lengths 2 and 3 (p2, p3), the number of neighboring carbon atoms (Number of Carbon), and the number of neighboring non-carbon atoms (Number of Non Carbon). In parallel, SMILES notation-based descriptors offered mechanistic insights by representing molecular fragments. Each descriptor contributes to the molecule's Descriptor Correlation Weight (DCW), which is computed as the sum of the correlation weights (CW) assigned to all relevant SMILES descriptors. This relationship is mathematically formalized in Equation 1. By combining these two descriptor systems, the hybrid approach provides a robust framework for characterizing molecular features, enabling enhanced predictive accuracy and interpretability in identifying key structural determinants of biological activity.

$$\text{DCW}(\text{T}, \text{N}_{\text{epoch}}) = z\text{CW}(\text{ATOMPAIR}) + x\text{CW}(\text{NOSP}) + y\text{CW}(\text{BOND}) + t\text{CW}(\text{HALO}) + (1) \\ r\text{CW}(\text{HARD}) + \alpha\text{CW}(\text{S}_k) + \beta\text{CW}(\text{SS}_k) + \gamma\text{CW}(\text{SSS}_k)$$

In Equation 1, the variables $z, x, y, t, \alpha, \beta, \gamma$ represent binary values – 1 indicating "yes" and 0 indicating "no" – that determine whether the corresponding SMILES descriptor is included in the QSAR model development. The symbol S_k refers to a SMILES atom represented by a single SMILES notation symbol (or a pair of inseparable symbols) and is associated with local molecular descriptors. Descriptors constructed as linear combinations of two and three SMILES atoms are denoted by SS_k and SSS_k , respectively, to account for interactions between multiple atomic components. The second category of SMILES notation-based descriptors used in the study comprises global descriptors, which capture the overall properties of the studied molecule. These include descriptors such as ATOMPAIR, HALO, BOND, NOSP, and HARD, each defined according to methodologies outlined in references^{39,40}. These global descriptors provide a comprehensive overview of molecular structure and complement the local descriptors in capturing the nuances of molecular behavior. The QSAR model developed in this study integrated both SMILES notation-based descriptors (local and global) and local graph invariants. This hybrid approach facilitated the calculation of the Descriptor Correlation Weight (DCW) for molecules as described in Equation 2, providing a robust and versatile framework for accurately modeling the relationship between molecular features and biological activity.

$$\text{DCW}(\text{T}, \text{N}_{\text{epoch}}) = \text{CW}(\text{S}_k) + \text{CW}(\text{SS}_k) + \text{CW}(\text{SSS}_k) + \text{CW}(\text{EC0}_k) + \text{CW}(\text{PT2}_k) + (2) \\ \text{CW}(\text{PT3}_k) + \text{CW}(\text{VS2}_k) + \text{CW}(\text{VS3}_k) + \text{CW}(\text{NNC}_k)$$

In addition to the previously defined symbols S_k , SS_k and SSS_k , Equation 2 incorporates the following symbols: The Morgan connectivity index of zero order (the hydrogen-suppressed graph was used in this research) – EC0_k , paths of length of 2 and 3 – PT2_k and PT3_k , valence shell 2 and 3 – VS2_k and VS3_k , and Nearest Neighbors – NNC_k ³⁸. The linear regression approach is used to compute the QSAR model (utilizing the training set) as indicated in Equation 3. This is achieved when the numerical data regarding the correlation weights are derived from the model, leading to favorable statistical results for the test set. In this specific study, the search for the optimal combination of T and N_{epoch} was carried out within the ranges of 1–5 for T and 0–50 for N_{epoch} .

$$\text{Ac} = \text{C}_0 + \text{C}_1 \times \text{DCW}(\text{T}, \text{N}_{\text{epoch}}) \quad (3)$$

To thoroughly evaluate the quality, robustness, and predictive reliability of the developed conformation-independent QSAR models, a comprehensive set of validation metrics was employed. These included widely used statistical parameters such as the squared correlation coefficient (R^2), which measures the proportion of variance explained

by the model, and the root-mean-squared error (RMSE), a standard indicator of prediction error magnitude. Cross-validation coefficients (Q^2) were calculated to assess the model's performance in predicting data excluded from the training set, while the F-value was used to determine the statistical significance of the regression. Additionally, the mean absolute error (MAE) was included as a measure of the average deviation between observed and predicted values, providing further insights into model accuracy^{41–44}. To strengthen the reliability of the QSAR models, advanced validation metrics were also applied. These included R_m^2 and MAE-based metrics, which emphasize the model's predictive power for new datasets. The concordance correlation coefficient (CCC) was used to evaluate the agreement between predicted and observed values, while the index of ideality of correlation (IIC) offered insights into the degree to which the correlation between the predicted and observed values aligned with an ideal relationship⁴⁵. The inclusion of these metrics ensured a thorough validation process and a holistic assessment of model performance. A pivotal component of any QSAR model is the establishment of its applicability domain (AD), which defines the chemical space within which the model can make reliable predictions. The AD ensures that predictions are made for compounds structurally and chemically similar to those in the training set, preventing extrapolation into areas of chemical space where the model may be unreliable. In this study, a literature-derived method was employed to determine the AD, as recommended in references^{46–48}. This involved systematic evaluation of the chemical structures and descriptors used in the model, ensuring that predictions adhered to the established AD criteria. A key aspect of this study was the analysis of "statistical defects" in conformation-independent molecular descriptors, particularly $d(A)$, to define the AD. These descriptors, previously utilized in QSAR model construction^{39–40}, were scrutinized to identify potential outliers or anomalies that could affect model reliability. The calculations for AD determination were performed using the CORAL software, which allowed for precise evaluation and correction of these statistical defects. Equation 4 formalized the methodology for this process, ensuring consistency and rigor.

$$d(A) = \frac{|P(A_{train}) - P(A_{test})|}{N(A_{train}) - N(A_{test})} \quad (4)$$

In the equation above, $P(A)_{train}$ and $P(A)_{calib}$ denote the probabilities of a conformation-independent attribute or descriptor (A) in the training and test sets, respectively. Meanwhile, $NA_{(train)}$ and $NA_{(calib)}$ represent the frequency of occurrence of a conformation-independent attribute or descriptor (A) in the training set and the test set, respectively. The statistical SMILES defect (D) is the cumulative sum of the defects, $d(A)$, of all the attributes found in the SMILES notation of the molecules. It is computed according to Equation 5.

$$D = defect(SMILES) = \sum_{k=1}^{NA} d(A) \quad (5)$$

A molecule is labeled as an outlier if it falls outside the defined applicability domain (AD), which happens when its D exceeds 2 times D_{av} , where D_{av} represents the average D calculated for the relevant set (whether it's the training or test set) in which the molecule is located.

2. 2. Molecular Docking

Docking studies were carried out using the Molegro Virtual Docker (MVD) software to evaluate the interactions between potential ligands and the TRPV1 receptor. Ligands were geometrically optimized prior to docking using the MMFF94 force field to ensure accurate structural representations. The three-dimensional structure of the TRPV1 receptor used for docking studies was obtained from the Protein Data Bank (PDB ID: 5IRX), representing the rat TRPV1 channel in complex with the antagonist capsazepine, resolved at 3.27 Å resolution⁴¹. This structure was selected based on its biological relevance and compatibility with the experimental system used in the QSAR dataset. The binding site was defined based on the position of the co-crystallized antagonist and included amino acid residues known to participate in ligand interactions, such as Tyr511, Ser512, Met547, Thr550, and Glu570⁴². These residues form the hydrophobic pocket and polar environment critical for antagonist binding. Protein preparation included removal of crystallographic water molecules, assignment of charges, and optimization of hydrogen bonding network using default MVD settings. The docking grid was centered on the native ligand position, with a radius of 15 Å to ensure full coverage of the binding cavity. Standard MolDock scoring function and search algorithm parameters were applied (maximum iterations = 1500; population size = 50; number of poses = 10 per ligand). MVD employs a hybrid approach in which the receptor structure is treated as rigid, while the ligand structures are allowed flexibility. This approach balances computational efficiency with the ability to account for conformational adaptability of ligands during docking. MVD identifies and quantifies both hydrophobic and hydrophilic interactions between the receptor and the ligands. Hydrophobic interactions primarily encompass Van der Waals forces and steric effects, while hydrophilic interactions involve hydrogen bond formation between ligand atoms and specific amino acid residues in the active site. These interactions are quantified through the use of "scoring" functions, which calculate numerical values corresponding to binding energies^{51,52}. In molecular docking studies, the strength and nature of ligand-receptor interactions are critical indicators of potential inhibitory activity. A general principle applies to enzyme and receptor studies: stronger and more favorable interactions typically correlate with enhanced inhibition potential. For this reason, the "scoring" functions calculated by MVD provide valuable insights into the binding affinity and po-

tential efficacy of the studied ligands⁴⁰. In this research, two primary scoring functions were calculated and utilized for the evaluation of inhibitory potential: MolDock Score and Rerank Score. The MolDockScore represents the primary binding energy calculated during docking, encompassing contributions from steric, electrostatic, and hydrogen bonding interactions. The Rerank Score is a secondary evaluation that re-assesses the binding interactions using additional weighting for certain interaction types, providing a more refined prediction of binding affinity. The docking protocol employed in this study was meticulously validated to ensure its reliability and accuracy. The validation was performed in accordance with established methodologies from the literature⁵³, which involve comparing docking results with experimentally determined binding modes or known inhibitors. This validation step is crucial for confirming that the docking simulations accurately represent the ligand-receptor interactions. Furthermore, the results of these docking studies provide a quantitative basis for ranking ligands based on their predicted binding affinity and inhibitory potential. These findings contribute to the identification of promising compounds for further experimental validation and drug development.

To complement molecular docking and provide an additional layer of binding affinity prediction, we applied KDEEP, a deep learning-based tool for structure-based binding affinity estimation. KDEEP uses 3D convolutional neural networks trained on experimentally validated protein–ligand complexes to predict binding strength and utilizes 3D Convolutional Neural Networks (3DCNN) to enhance the accuracy of binding predictions⁵⁴. It classifies input molecules into eight pharmacophore properties: hydrophobic, aromatic, hydrogen-bond donor and acceptor, positive and negative ionizable, metallic, and total excluded volume. The molecules are then processed using a Deep Convolutional Neural Network (DCNN) model trained on the PDBbind 2016 database (available at <https://playmolecule.com/Kdeep/>)⁵⁵. The docking results for ligand orientation within the TRPV1 receptor active site obtained with MolDock were used to estimate absolute binding affinity with KDeep. For each ligand–TRPV1 complex, KDEEP calculated three key parameters: pKd, binding free energies (ΔG) and ligand efficiencies (LE).

By integrating computational and statistical rigor, the study establishes a robust framework for exploring ligand interactions with the TRPV1 receptor, paving the way for the development of novel therapeutic agents targeting this receptor.

3. Results and Discussion

Table 1 summarizes the numerical values of all the validation metrics used to comprehensively evaluate the quality and performance of the conformation-independent QSAR models developed through the Monte Carlo

Table 1. The statistical quality of the developed conformational-independent QSAR models for TRPV1 antagonism by piperazinyl-aryl compounds

		Training set						Test set									
		R ²	CCC	IIC	CII	Q ²	RMSE	MAE	F	R ²	CCC	IIC	CII	Q ²	RMSE	MAE	F
Split 1	1 run	0.7926	0.8843	0.8200	0.8828	0.7775	0.517	0.410	271	0.8263	0.8851	0.8401	0.9097	0.7799	0.547	0.433	109
	2 run	0.7928	0.8844	0.8201	0.8833	0.7798	0.517	0.398	272	0.8410	0.8923	0.7270	0.9305	0.8077	0.529	0.424	122
	3 run	0.7858	0.8800	0.7743	0.8715	0.7732	0.525	0.399	260	0.8487	0.8859	0.8449	0.9344	0.8124	0.534	0.421	129
	Av	0.7904	0.8829	0.8048	0.8729	0.7768	0.520	0.402	268	0.8387	0.8878	0.8040	0.9249	0.8000	0.537	0.426	120
Split 2	1 run	0.7219	0.8385	0.7826	0.8298	0.7090	0.591	0.413	184	0.8367	0.8657	0.8352	0.8931	0.8087	0.591	0.505	118
	2 run	0.7478	0.8557	0.7539	0.8406	0.7360	0.563	0.398	210	0.8379	0.8485	0.8394	0.8979	0.8100	0.616	0.512	119
	3 run	0.7781	0.8752	0.7690	0.8658	0.7672	0.528	0.373	249	0.8770	0.8894	0.8858	0.9212	0.8535	0.538	0.441	164
	Av	0.7493	0.8565	0.7685	0.8545	0.7374	0.561	0.395	214	0.8505	0.8679	0.8535	0.9041	0.8241	0.582	0.486	134
Split 3	1 run	0.7472	0.8553	0.7757	0.8447	0.7316	0.578	0.439	213	0.9158	0.9535	0.9155	0.9568	0.8977	0.363	0.311	239
	2 run	0.7244	0.8297	0.7895	0.8255	0.7156	0.636	0.481	164	0.9119	0.9448	0.9299	0.9430	0.8932	0.384	0.338	228
	3 run	0.7155	0.8134	0.7430	0.7766	0.6970	0.645	0.489	157	0.9271	0.9469	0.9635	0.9593	0.9241	0.367	0.329	328
	Av	0.7290	0.8328	0.7694	0.8156	0.7147	0.620	0.470	178	0.9183	0.9484	0.9363	0.9530	0.9050	0.371	0.326	265

R² – Correlation coefficient CCC – Concordance correlation coefficient IIC – Index of ideality of correlation Q² – Cross-validated correlation coefficient RMSE – Root mean squared error MAE – Mean absolute error F – Fischer ratio Av – Average value for statistical parameters obtained from three independent Monte Carlo optimization runs

optimization method. These metrics demonstrate the robustness, predictive accuracy, and reproducibility of the models, providing a clear indication of their reliability. The results highlight the strong predictive potential of the QSAR models, with no significant deviations or inconsistencies observed across the dataset. Among the various splits examined during the optimization process, the second split, utilizing a T value of 4 and N_{epoch} of 15, produced the most favorable results. This configuration led to a model with superior performance, as indicated by its validation metrics. Notably, no outliers were identified during the analysis, as the methodology applied for defining the applicability domain (AD) confirmed that all molecules fell within the defined chemical space. This is a crucial finding, as the absence of outliers ensures that the model's predictions are both valid and reliable within the specified AD, thereby enhancing its applicability to unseen compounds. Figure 1 provides a graphical representation of the best-performing QSAR model, which achieved the highest R^2 value across all three splits during the optimal Monte Carlo optimization run. The graph visually illustrates the close agreement between observed and predicted values, highlighting the model's predictive accuracy and its capability to generalize across the dataset. To ensure further validation, the concordance correlation coefficient (CCC) was calculated for all QSAR models. The CCC is a robust metric that evaluates the reproducibility of predictions by measuring the degree of agreement between observed and predicted values. The results demonstrated high reproducibility across all splits, confirming that the

models are consistent and reliable across different configurations. In addition, the mean absolute error (MAE)-based metric was used to evaluate the precision of the models, and the outcomes were rated as "GOOD," further solidifying the models' robustness and reliability. The final layer of validation involved the calculation of the index of ideality of correlation (IIC). This metric assesses how closely the correlation between predicted and observed values aligns with an ideal relationship. The IIC values obtained were highly favorable, suggesting that the developed QSAR models not only perform well but also exhibit a high degree of predictive reliability and alignment with theoretical expectations. The findings of this study strongly indicate that the QSAR models developed using the Monte Carlo optimization method possess exceptional predictive potential, making them valuable tools for future applications in drug discovery and other computational research domains. The comprehensive validation of these models using multiple metrics – including R^2 , RMSE, Q^2 , F-value, MAE, CCC, and IIC – underscores their robustness and generalizability. By incorporating a rigorous methodology for defining and adhering to the applicability domain, the study ensures that these models can be reliably employed for predictions involving structurally similar compounds within the defined chemical space. Furthermore, the absence of outliers and the high concordance between predicted and observed values across all validation steps provide additional confidence in the utility of these models. This thorough validation process paves the way for the practical application of these QSAR models in tasks such

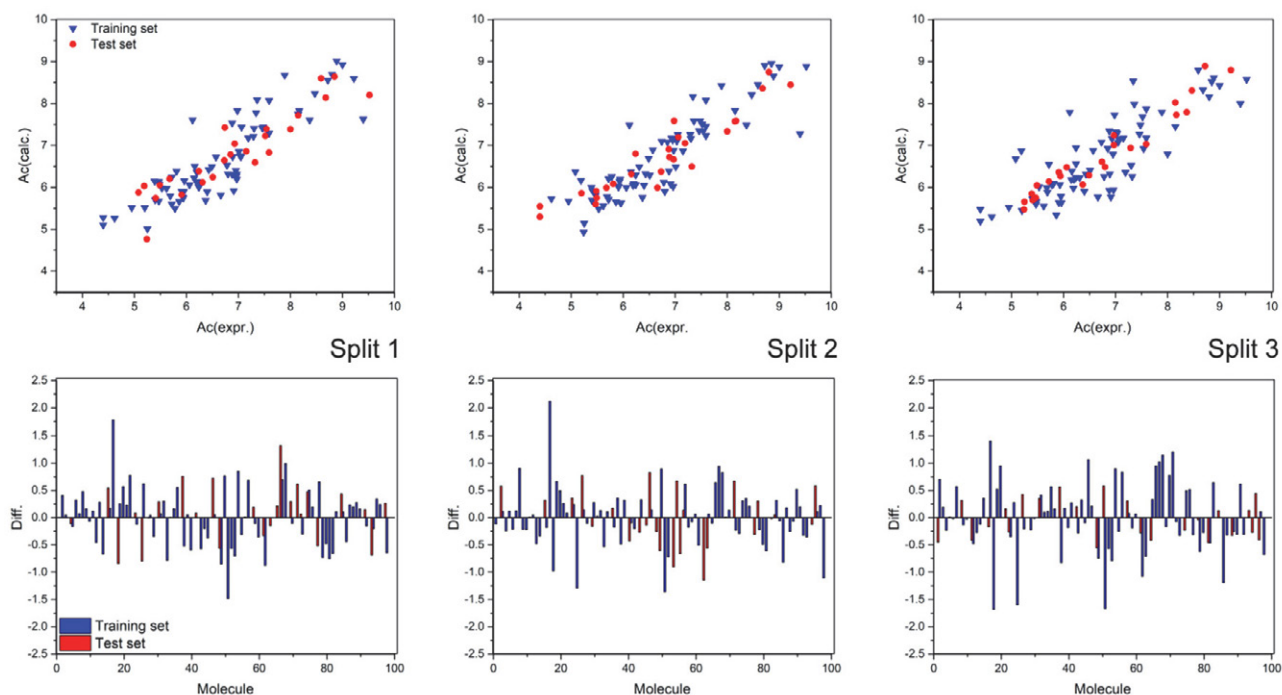


Figure 1. Above) Graphical presentation of the best Monte Carlo optimization runs (the highest value for R^2) for the developed QSAR models; Bellow) Diff. – Difference between experimental and calculated values for pIC_{50} .

as virtual screening, lead optimization, and the prediction of biological activity for new compounds. Ultimately, these models represent a significant advancement in the integration of computational tools into modern drug discovery workflows.

The mathematical formulations for the top-performing QSAR models, as determined by the test set R^2 values for all the splits, are provided in Equations 6–8.

$$\text{Split 1: } \text{pIC}_{50} = -0.4579(\pm 0.0534) + 0.0398(\pm 0.0003) \times \text{DCW}(3,24) \quad (6)$$

$$\text{Split 2: } \text{pIC}_{50} = 0.3013(\pm 0.0420) + 0.0297(\pm 0.0002) \times \text{DCW}(1,10) \quad (7)$$

$$\text{Split 3: } \text{pIC}_{50} = 2.1931(\pm 0.0443) + 0.0339(\pm 0.0003) \times \text{DCW}(3,23) \quad (8)$$

The equations (Eq. 6–8) show that for split 1, the preferred values for T and N_{epoch} are 3 and 24, respectively. For split 2, the preferred values are 1 for T and 10 for N_{epoch} , while for split 3, the preferred values are 3 for T and 23 for N_{epoch} .

There have been significant efforts to apply ligand-based methodologies for developing pharmacophore and QSAR models aimed at TRPV1 antagonists. These studies have utilized various computational approaches, demonstrating the versatility and effectiveness of such methods in exploring the molecular basis of TRPV1 antagonism. Some studies have employed molecular field-based techniques like CoMFA (Comparative Molecular Field Analysis) and CoMSIA (Comparative Molecular Similarity Indices Analysis), as well as Phase QSAR methodologies⁵⁶. These approaches focus on the spatial and electronic properties of molecules to identify patterns correlating with antagonistic activity. Others have used descriptor-based algorithms to determine molecular descriptors most strongly associated with TRPV1 antagonism, enabling the development of predictive QSAR models⁵⁷. In other efforts, 3D alignments of TRPV1 antagonists have been analyzed in the context of homology models. These studies have leveraged the structural insights provided by homology models to explore the spatial arrangements and interactions of antagonists within the receptor binding site^{54,58}. For example, Goldmann et al. utilized publicly available data on TRPV1 antagonists to construct pharmacophore models. Their research involved extensive validation of these models, which were then applied to virtually screen the LifeChem database comprising 305,841 compounds. This exercise identified 12 hits with promising activity and diversity compared to reference antagonists and other active compounds. Goldmann and colleagues further hypothesized that pharmacophore modeling of public data could reveal "pharmacophoric ensembles," helping to differentiate safe compounds from those with undesirable profiles⁵⁹. Similarly, Feng et al. reported constructing hu-

man TRPV1 (hTRPV1) homology models based on the recently released rat TRPV1 (rTRPV1) structure. Their study involved validation using known agonists and antagonists, prediction of binding modes for well-known antagonists, and a virtual screening exercise targeting the putative antagonist binding site. This approach provided valuable insights into the molecular interactions governing TRPV1 antagonism and highlighted potential candidates for further investigation⁶⁰. Kristam et al. developed and validated predictive 3D-QSAR models for a collection of TRPV1 receptor antagonists using CoMFA and Topomer-CoMFA methodologies. These models were applied to screen databases for alternative fragments that could replace key functional groups in known antagonists, such as the disubstituted dimidazolyl moiety (R1 fragment) or the piperazine aryl moiety (R2 fragment)³⁴. Their work underscores the potential of 3D-QSAR techniques in guiding the rational design of structurally optimized TRPV1 antagonists. Together, these studies demonstrate the power of ligand-based methodologies, including pharmacophore modeling, QSAR, and homology modeling, in advancing our understanding of TRPV1 receptor antagonists. They highlight the potential of computational tools to identify novel candidates, optimize known scaffolds, and reveal critical molecular features that differentiate efficacious and safe compounds from those with less desirable profiles. These approaches continue to contribute significantly to the field of TRPV1-targeted drug discovery, offering pathways to innovative therapeutic agents for pain management and other conditions.

A primary objective of this research was to identify molecular fragments, defined as optimal descriptors in SMILES notation, that contribute positively or negatively to the studied activity, as supported by prior studies^{39,40,61,62}. These fragments are essential for understanding the structural features that influence the biological activity of compounds, providing valuable insights for drug design and optimization. The complete set of calculated molecular descriptors, derived from both SMILES notation and molecular graph-based approaches, is provided in Table S2 in the Supplementary Material. These descriptors encompass a wide range of molecular properties and were systematically analyzed to determine their relevance to the studied activity. To aid in understanding, an example calculation of a molecule's summarized Descriptor Correlation Weight (DCW) and its corresponding studied activity (pIC_{50}) is detailed in Table 2. For simplicity, molecular graph-based descriptors are excluded from the example, allowing for a more focused interpretation of the contributions from SMILES-based descriptors. This example highlights the methodology used to link molecular features to biological activity, demonstrating how individual descriptor contributions are aggregated into the DCW. Furthermore, a graphical representation of the molecular fragments associated with the same molecule is provided in Figure 2. This visualization illustrates the structural components of the

Table 2. Example of DCW calculation

SMILES notation:
Nc1ccc(cc1)c1cc(cc2c1NC(N2)N1CCN(CC1)c1ncccc1C(F)(F)F)C(F)(F)F
 DCW = 115.8633
 pIC₅₀(calc.) = 6.1255

SA _k	CW(SA _k)	SA _k	CW(SA _k)	SA _k	CW(SA _k)	SA _k	CW(SA _k)
10001000	-0.0545	C...(...	0.4633	c...c...1...	-0.1625	N...(...C...	1.6575
(...(...	0.2736	c...(...c...	0.2654	C...C...1...	-0.11	N.....	-0.1445
(.....	0.1116	c.....	0.0465	c...c...2...	-2.6721	n.....	0.2927
(...C...(...	0.3527	C.....	-0.3361	c...c...c...	0.1654	N...1.....	-0.7082
(...F...(...	-0.0813	c...1...(...	0.2406	C...N...(...	-0.1388	n...1.....	-0.5053
1...(...	0.8061	C...1...(...	0.1533	c...N.....	0.1804	N...1...C...	-0.0284
1.....	0.2906	c...1.....	0.0952	C...N...1...	0.1723	n...1...c...	0.1866
1...C...(...	-0.3129	C...1.....	-0.5409	c...n...1...	0.0845	N...2.....	0.5679
1...C...(...	0.4098	c...1...c...	0.7524	F...(...	0.3315	N...C...(...	-0.2294
1...N...(...	0.1976	c...1...C...	0.2782	F...(...	0.0918	N...C.....	-0.1215
2...(...	0.3586	c...1...N...	-0.424	F...(...C...	1.0121	n...C.....	0.7165
2.....	0.4303	c...2.....	0.0583	F...(...F...	0.0196	N...c...1...	0.4924
2...C...1...	-0.6572	c...2...c...	0.1318	F.....	0.3709	N...C...C...	-2.8109
BOND00000	0.1922	c...C...(...	0.1144	HALO10000	-0.6764	n...c...c...	2.3328
c...(...	0.1999	C...C...(...	0.0764	N...(...	1.8556	NOSP10000	0.6488
c...c.....	0.1336	C...C.....	-0.4259	N...(...2...	0.6523		

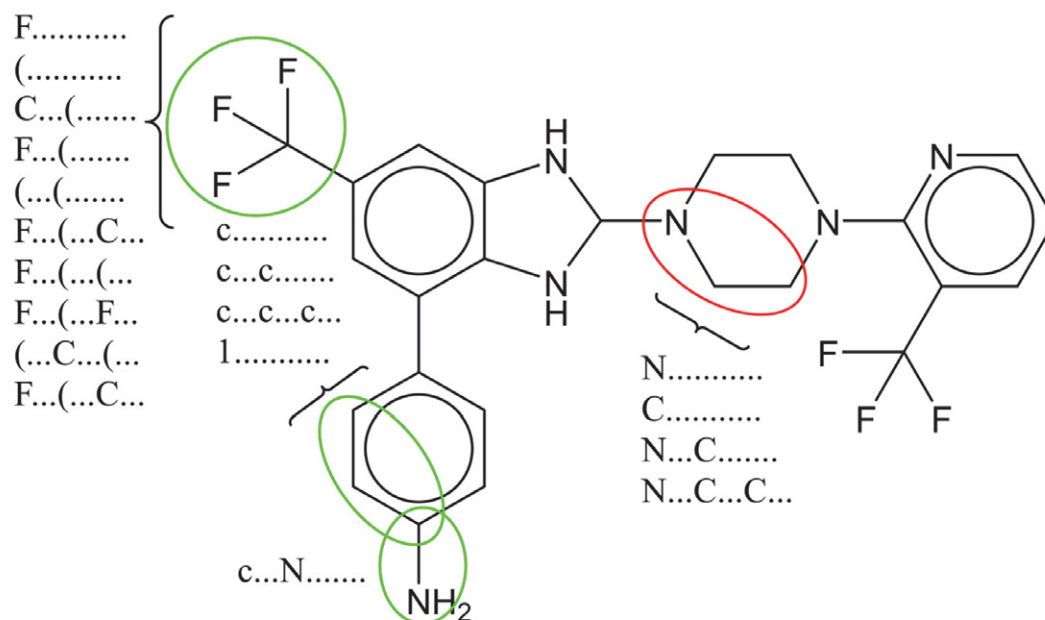


Figure 2. Contribution of molecular fragments to TRPV1 antagonism (Green – Increase, Red – Decrease).

molecule and their respective roles, offering a clear depiction of how specific fragments correlate with the observed activity. Together, these analyses provide a comprehensive framework for identifying and understanding the molecular determinants of TRPV1 antagonism, facilitating the rational design of new compounds with improved efficacy and safety profiles.

The results from QSAR modeling revealed specific molecular fragments, identified through SMILES notation, that significantly influence pIC_{50} activity, either positively or negatively. These findings provide a structural basis for designing new TRPV1 antagonists with enhanced activity. Molecular fragments that positively influenced pIC_{50} activity included those associated with trifluoromethyl groups, such as "F.....", "(.....", "C...(.....", "F...(.....", "(...(.....", "F...(C...", "F...(C...", "F...(F...", "(...C...(C...", and "F...(C...". Fragments associated with aromatic carbon atoms, including "c.....", "c...c.....", and "c...c...c...", as well as the nitrogen atom bonded to aromatic carbon ("c...N....."), also demonstrated positive effects. Fragments that negatively influenced pIC_{50} activity included the nitrogen atom ("N....."), carbon atom ("C....."); and combinations of nitrogen bonded to one or two carbon atoms ("N...C....." and "N...C...C..." respectively). These negatively correlated fragments provided insights into structural features that should be avoided or minimized in antagonist design.

Molecule A1 was created by replacing the amino group with a hydroxyl group in the para position, resulting in the addition of the fragment "O.....". Molecule A2 introduced a chlorine atom at the same position, adding the "Cl....." fragment, while molecule A3 incorporated a methoxy group with the fragment "C...O.....". Molecule A4 introduced an aminomethyl group, adding the "C...N....." fragment, enhancing hydrogen bonding potential. Molecule A5 featured a dimethylamino group, adding the fragments "C...(.....", "(.....", "N...(.....", and "N...(C...", which indicated increased molecular branching around the nitrogen atom. Molecule A6 incorporated an isopropyl group, resulting in fragments such as "C...(.....", "(.....", "C...(.....", and "C...(C..." that also indicated increased branching on carbon atoms. Molecule A7 introduced a carboxyl group, contributing fragments such as "=.....", "=...C...(C...", "C...=.....", "O...(.....", "O...(C...", "O...=.....", "O...C...(C...", and "O...C..." that enhanced electron-withdrawing and polar interaction properties. All introduced fragments in molecules A1–A7 were associated with a positive impact on pIC_{50} activity, leading to higher predicted activity values. The substitution strategy and resulting molecular modifications were validated through increased pIC_{50} values, as summarized in Table 3. These results demonstrate the effectiveness of the CAD approach in integrating QSAR insights for the rational design of novel TRPV1 antagonists with improved biological activity.

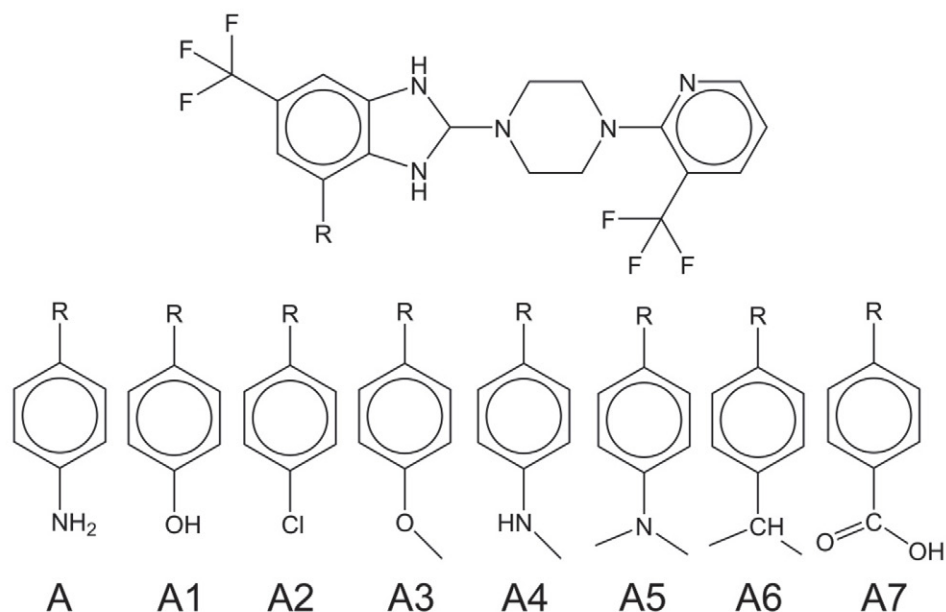


Figure 3. Chemical structures of designed molecules.

The molecular fragments identified as positively influencing activity were applied in a Computer-Aided Design (CAD) process to develop seven novel compounds with enhanced pIC_{50} values. These compounds were designed by introducing functional groups that incorporated new fragments positively associated with activity. Mole-

All the designed molecules, along with the template molecule A, were subjected to molecular docking studies targeting TRPV1 to evaluate and further validate the predictive accuracy of the developed QSAR models. Table 3 provides the numerical values of the calculated "scoring" functions, which reflect the strength of interactions be-

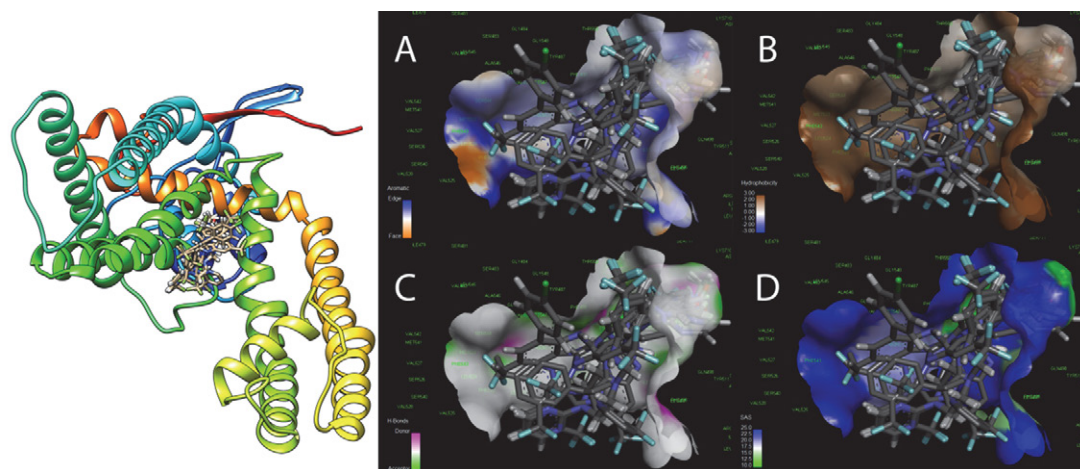
Table 3. The list of all the designed molecules with their calculated activities, score values (kcal/mol), ΔG (kcal/mol), ligand efficiency (LE) (kcal/mol) and pKd

Molecule	pIC ₅₀ (calc.)	Rerank	MolDock	ΔG	LE	pKd
A0	6.1255	-121.35	-118.85	-7.437	-0.1957	5.5089
A1	6.4688	-133.49	-128.49	-8.3795	-0.2328	6.207
A2	6.6225	-137.85	-121.65	-9.3946	-0.2610	6.9590
A3	6.5881	-136.35	-134.32	-8.6117	-0.2328	6.3790
A4	6.4472	-133.22	-130.66	-8.7899	-0.2376	6.5110
A5	7.0963	-139.59	-136.57	-9.2786	-0.2577	6.8730
A6	7.2891	-144.43	-137.47	-9.7236	-0.2559	7.2027
A7	6.7976	-138.48	-130.3	-9.1377	-0.2405	6.7686

tween the ligands and the receptor. Since different scoring functions capture distinct ligand-amino acid interactions, a comprehensive assessment of inhibitory potency requires consideration of all relevant factors. Based on the results obtained for the MolDock and ReRank scoring functions, molecule A6 emerged as the compound with the highest potential inhibitory activity. This finding aligns well with the predictions from QSAR modeling, reinforcing the consistency and reliability of the models. In contrast, the template molecule A exhibited the lowest MolDock and ReRank scores, a result that also corresponds to the predictions made by the QSAR models. These correlations between docking results and QSAR predictions provide strong validation for the approach used in this study. Table 3 presents a list of the designed TRPV1 antagonist candidates, along with their predicted biological activity (pIC₅₀), binding free energy (ΔG), ligand efficiency (LE), and predicted dissociation constant pKd. These values were calculated using a combination of QSAR modeling, molecular docking, and KDEEP-based binding affinity estimation. As expected, compounds with higher predicted pIC₅₀ values—which indicate stronger biological activity and lower effective concentration—generally correspond to more favorable binding energies (lower ΔG values). This relationship is consistent with the principle that stronger

binding (i.e., more negative ΔG) often correlates with increased ligand efficiency and potency. For instance, compound A6 shows the highest predicted pIC₅₀ (7.2891), in line with its strong binding affinity ($\Delta G = -9.72$ kcal/mol) and the highest pKd (7.20), indicating a tighter ligand–receptor complex. Ligand efficiency (LE), defined as the binding energy per heavy atom, also supports this trend and reflects the balance between molecular size and binding strength. Compounds A5 and A6 exhibit both high pIC₅₀ and LE values, suggesting that they are not only potent but also structurally efficient binders. Overall, these results demonstrate a high level of consistency between ligand-based (QSAR) and structure-based (docking and deep learning) predictions, further validating the designed molecules as promising candidates for future experimental testing.

The best-calculated binding poses of all designed molecules within the active site of TRPV1 are illustrated in Figure 4. Further in Figure 4 surfaces of active site (aromatic, hydrophobic, H-Bond and solvent accessibility surface – SAS). These visual representations highlight the spatial orientation and interaction profiles of each molecule, further supporting the docking results and their alignment with QSAR-based predictions. Together, these findings underscore the effectiveness of combining QSAR

**Figure 4.** Left) The best calculated poses for all the designed molecules within the active site of TRPV1; Right) Surfaces inside active site: A) Aromatic, B) Hydrophobic, C) H-Bond and D) SAS

modeling and molecular docking studies in identifying and validating promising TRPV1 antagonists.

This study highlights the potential of the Monte Carlo optimization method for generating diverse and informative QSAR models. However, certain limitations merit further investigation to enhance the method's applicability and reliability. One notable limitation involves the interpretability of certain SMILES-based descriptors. The inclusion of two- and three-atom combinations can result in descriptors such as "...2....." or "#.../.....", which lack clear physical or chemical meaning. This ambiguity makes it challenging to derive mechanistic insights into their influence on biological properties, such as vaginal permeability predictions. Improving the clarity and interpretability of these descriptors would significantly enhance the practical utility of the models. Fragment identification also poses challenges. The method may struggle to detect rare but biologically significant molecular fragments, which could provide valuable insights into permeability or activity. Additionally, the focus on smaller fragments, such as those comprising three atoms, may overlook the contributions of larger structural motifs or long-range interactions that play a critical role in biological processes. The CORAL algorithm, employed within the Monte Carlo optimization framework, introduces another limitation. Its potential prevalence bias may prioritize common fragments within the dataset, potentially underestimating the significance of truly active fragments that are present across all SMILES descriptors. This bias could reduce the algorithm's ability to identify unique features critical for model predictions. While the hybrid approach incorporates descriptors derived from both SMILES notation and molecular graphs, some molecular graph-based descriptors also lack clear mechanistic interpretation. This limitation restricts the ability to connect model predictions with underlying chemical or biological principles, which is essential for advancing scientific understanding and rational compound design. To address these limitations, data preprocessing strategies could be implemented. Preliminary analysis and filtering of molecular fragments and descriptors might help identify those that are most informative and relevant to the studied activity. This approach could reduce noise, improve model interpretability, and enhance the mechanistic insights derived from the models. Additionally, incorporating techniques that balance the emphasis on rare and common fragments may improve the identification of biologically significant features. Future research should focus on addressing these challenges to enhance the capabilities of the Monte Carlo optimization method. Efforts could include refining descriptor definitions, exploring alternative algorithms with reduced bias, and expanding the scope of descriptors to capture larger and more complex molecular interactions. By overcoming these limitations, the Monte Carlo optimization method could become an even more powerful tool in the development of QSAR models and their application in drug discovery and beyond.

4. Conclusion

The primary goal of this study was to develop reliable QSAR models with strong predictive power for TRPV1 antagonism, validated through a comprehensive set of statistical parameters. Conformation-independent QSAR models were constructed using the Monte Carlo optimization method, leveraging optimal descriptors derived from both local graph invariants and SMILES notation. These descriptors provided a robust foundation for the modeling process, enabling the identification of structural features influencing TRPV1 antagonism. The robustness and predictive capabilities of the QSAR models were thoroughly evaluated using various statistical techniques. The validation metrics confirmed the high applicability of the models, demonstrating their effectiveness in predicting biological activity. The Monte Carlo optimization method facilitated the identification of molecular fragments, represented as SMILES notation fragments, with both positive and negative effects on TRPV1 antagonism. These insights offered a deeper understanding of the molecular features contributing to or detracting from activity, supporting the rational design of more effective antagonists. Molecular docking studies served as the final validation step, further confirming the predictive accuracy of the developed QSAR models. The docking results provided an additional layer of evidence for the potential inhibitory effects of the designed molecules. A strong inter-correlation was observed between the calculated pIC_{50} values from the best QSAR model and the interaction energies derived from docking studies with the TRPV1 active site. To further strengthen the predictive workflow, we incorporated KDEEP—a deep learning-based method for protein–ligand binding affinity estimation—which provided additional independent predictions of pK_d and ΔG values. The strong concordance between these results and the QSAR-predicted pIC_{50} values supports the internal validity of the designed compounds and adds confidence to their prioritization. KDEEP thus served as a third, AI-powered validation layer within our CADD pipeline, highlighting its utility in guiding early-phase drug discovery even in the absence of experimental data. This concordance highlights the reliability of the combined QSAR and molecular docking approach in predicting ligand–receptor interactions. The methodology described in this study demonstrates significant potential for broader applications, including the discovery of novel therapeutics for conditions such as atherosclerosis. By targeting the antagonism of Transient Receptor Potential Vanilloid, member 1 (TRPV1), the outlined approach provides a versatile framework for identifying and optimizing drug candidates for a variety of therapeutic areas.

Competing Interests

The authors declare that there are no conflicts of interest in this study.

Acknowledgment

The authors would like to thank the Ministry of Education, Science and Technological Development of Republic of Serbia (Grant No: 451-03-136/2025-03/200113) for financial support.

5. References

1. M. J. Caterina, M. A. Schumacher, M. Tominaga, T. A. Rosen, J. D. Levine, D. Julius, *Nature* **1997**, 389, 816–824. DOI:10.1038/39807
2. M. M. Moran, M. A. McAlexander, T. Bíró, A. Szallasi, *Nat. Rev. Drug Discov.* **2011**, 10, 601–620. DOI:10.2174/138161208783330763
3. R. Vennekens, G. Owsianik, B. Nilius, *Curr. Pharm. Des.* **2008**, 14, 18–31. DOI:10.2174/138161208783330763
4. M. Zhang, Y. Ma, X. Ye, N. Zhang, L. Pan, B. Wang, *Signal Transduct. Target Ther.* **2023**, 8, 261. DOI:10.1038/s41392-023-01464-x
5. L. Basso, C. Altier, *Curr. Opin. Pharmacol.* **2017**, 32, 9–15. DOI:10.1016/j.coph.2016.10.002
6. Y. Kaneko, A. Szallasi, *Br. J. Pharmacol.* **2014**, 171, 2474–2507. DOI:10.1111/bph.12414
7. M. G. Tsagareli, I. Nozadze, *Behav. Pharmacol.* **2020**, 31, 413–434. DOI:10.1097/FBP.0000000000000524
8. A. Szallasi, P. M. Blumberg, *Pharmacol. Rev.* **1999**, 51, 159–212. DOI:10.1016/S0031-6997(24)01403-0
9. O. M. E. Abdel-Salam, G. Mózsik, *Neurochem. Res.* **2023**, 48, 3296–3315. DOI:10.1007/s11064-023-03983-z
10. J. E. Krause, B. L. Chenard, D. N. Cortright, *Curr. Opin. Investig. Drugs* **2005**, 6, 48–57.
11. M. M. Moran, *Annu. Rev. Pharmacol. Toxicol.* **2018**, 58, 309–330. DOI:10.1146/annurev-pharmtox-010617-052832
12. L. A. Roberts, M. Connor, *Recent Pat. CNS Drug Discov.* **2006**, 1, 65–76. DOI:10.2174/157488906775245309
13. B. Nilius, *Curr. Top. Med. Chem.* **2013**, 13, 244–246. DOI:10.2174/1568026611313030002
14. A. Szallasi, D. N. Cortright, C. A. Blum, S. R. Eid, *Nat. Rev. Drug Discov.* **2007**, 6, 357–372. DOI:10.1038/nrd2280
15. D. Y. Okuhara, A. Y. Hsia, M. Xie, *Expert Opin. Ther. Targets* **2007**, 11, 391–401. DOI:10.1517/14728222.11.3.391
16. J. Lázár, L. Gharat, N. K. Joshi, P. M. Blumberg, A. Szallasi, *Expert Opin. Drug Discov.* **2009**, 4, 159–180. DOI:10.1517/17460440802681300
17. D. Souza Monteiro de Araujo, R. Nassini, P. Geppetti, F. De Logu, *Expert Opin. Ther. Targets* **2020**, 24, 997–1008. DOI:10.1080/14728222.2020.1815191
18. C. C. Correll, A. Palani, *Expert Opin. Ther. Pat.* **2006**, 16, 783–795. DOI:10.1517/13543776.16.6.783
19. G. Y. Wong, N. R. Gavva, *Brain Res. Rev.* **2009**, 60, 267–277. DOI:10.1016/j.brainresrev.2008.12.006
20. P. Holzer, *Br. J. Pharmacol.* **2008**, 155, 1145–1162. DOI:10.1038/bjp.2008.351
21. A. Fernández-Carvajal, G. Fernández-Ballester, I. Devesa, J. M. González-Ros, A. Ferrer-Montiel, *Pharmaceuticals* **2012**, 5, 16–48. DOI:10.3390/ph5010016
22. K. Kaszas, J. M. Keller, C. Coddou, S. K. Mishra, M. A. Hoon, S. Stojilkovic, K. A. Jacobson, M. J. Iadarola, *J. Pharmacol. Exp. Ther.* **2012**, 340, 152–160. DOI:10.1124/jpet.111.183053
23. J. Chen, D. H. Hackos, *Naunyn Schmiedeberg's Arch. Pharmacol.* **2015**, 388, 451–463. DOI:10.1007/s00210-015-1088-3
24. B. Nilius, A. Szallasi, *Pharmacol. Rev.* **2014**, 66, 676–814. DOI:10.1124/pr.113.008268
25. M. J. Gunthorpe, A. Szallasi, *Curr. Pharm. Des.* **2008**, 14, 32–41. DOI:10.2174/138161208783330754
26. P. V. Desai, *Future Med. Chem.* **2016**, 8, 1717–1720. DOI:10.4155/fmc-2016-0161
27. S. Ekins, J. Mestres, B. Testa, *Br. J. Pharmacol.* **2007**, 152, 9–20. DOI:10.1038/sj.bjp.0707305
28. S. De Vita, M. G. Chini, G. Bifulco, G. Lauro, *Bioorg. Med. Chem. Lett.* **2023**, 83, 129171. DOI:10.1016/j.bmcl.2023.129171
29. P. G. R. Achary, *Mini Rev. Med. Chem.* **2020**, 20, 1375–1388. DOI:10.2174/1389557520666200429102334
30. S. J. Macalino, V. Gosu, S. Hong, S. Choi, *Arch. Pharm. Res.* **2015**, 38, 1686–1701. DOI:10.1007/s12272-015-0640-5
31. T. Wang, M. B. Wu, J. P. Lin, L. R. Yang, *Expert Opin. Drug Discov.* **2015**, 10, 1283–1300. DOI:10.1517/17460441.2015.1083006
32. M. R. Hasan, A. A. Alsaiari, B. Z. Fakhurji, M. H. R. Molla, A. H. Asseri, M. A. A. Sumon, M. N. Park, F. Ahammad, B. Kim, *Molecules* **2022**, 27, 4169. DOI:10.3390/molecules27134169
33. K. Funatsu, T. Miyao, M. Arakawa, *Curr. Comput. Aided Drug Des.* **2011**, 7, 1–9. DOI:10.2174/157340911793743556
34. R. Kristam, S. N. Rao, A. S. D'Cruz, V. Mahadevan, V. N. Viswanadhan, *J. Mol. Graph. Model.* **2017**, 72, 112–128. DOI:10.1016/j.jmglm.2017.01.010
35. V. I. Ognyanov, C. Balan, A. W. Bannon, Y. Bo, C. Dominguez, C. Fotsch, V. K. Gore, L. Klionsky, V. V. Ma, Y. X. Qian, R. Tamir, X. Wang, N. Xi, S. Xu, D. Zhu, N. R. Gavva, J. J. Treanor, M. H. Norman, *J. Med. Chem.* **2006**, 49, 3719–3742. DOI:10.1021/jm060065y
36. V. K. Gore, V. V. Ma, R. Tamir, N. R. Gavva, J. J. Treanor, M. H. Norman, *Bioorg. Med. Chem. Lett.* **2007**, 17, 5825–5830. DOI:10.1016/j.bmcl.2007.08.044
37. P. K. Ojha, K. Roy, *Chemometr. Intell. Lab.* **2011**, 109, 146–161. DOI:10.1016/j.chemolab.2011.08.007
38. A. A. Toropov, P. Duchowicz, E. A. Castro, *Int. J. Mol. Sci.* **2003**, 4, s272–283. DOI:10.3390/i4050272
39. A. M. Veselinović, J. B. Veselinović, J. V. Živković, G. M. Nikolić, *Curr. Top. Med. Chem.* **2015**, 15, 1768–1779. DOI:10.2174/1568026615666150506151533
40. M. Zivkovic, M. Zlatanovic, N. Zlatanovic, M. Golubović, A. M. Veselinović, *Mini Rev. Med. Chem.* **2020**, 20, 1389–1402. DOI:10.2174/1389557520666200212111428
41. A. Golbraikh, A. Tropsha, *J. Mol. Graph. Model.* **2020**, 20, 269–276. DOI:10.1016/S1093-3263(01)00123-1
42. P. P. Roy, J. T. Leonard, K. Roy, *Chemometr. Intell. Lab.* **2008**, 90, 31–42. DOI:10.1016/j.chemolab.2007.07.004
43. 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

44. 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
45. A. P. Toropova, A. A. Toropov, *Sci. Total Environ.* **2017**, 586, 466–472. DOI:10.1016/j.scitotenv.2017.01.198
46. K. Roy, S. Kar, P. Ambure, *Chemometr. Intell. Lab.* **2015**, 145, 22–29. DOI:10.1016/j.chemolab.2015.04.013
47. P. Gramatica, *QSAR Comb. Sci.* **2007**, 26, 694–701. DOI:10.1002/qsar.200610151
48. A. A. Toropov, A. P. Toropova, A. Lombardo, A. Roncaglioni, E. Benfenati, G. Gini, *Eur. J. Med. Chem.* **2011**, 46, 1400–1403. DOI:10.1016/j.ejmech.2011.01.018
49. Y. Gao, E. Cao, D. Julius, Y. Cheng, *Nature* **2016**, 534, 347–351. DOI:10.1038/nature17964
50. M. Liao, E. Cao, D. Julius, Y. Cheng, *Nature* **2013**, 504, 107–112. DOI:10.1038/nature12822
51. R. Thomsen, M. H. Christensen, *J. Med. Chem.* **2006**, 49, 3315–3321. DOI:10.1021/jm051197e
52. G. Bitencourt-Ferreira, W. F. de Azevedo Jr., *Methods Mol. Biol.* **2019**, 2053, 149–167. DOI:10.1007/978-1-4939-9752-7_10
53. Manisha, S. Chauhan, P. Kumar, A. Kumar, *SAR QSAR Environ. Res.* **2019**, 30, 145–159. DOI:10.1080/1062936X.2019.1568299
54. J. Jiménez, M. Škalič, G. Martínez-Rosell, G. De Fabritiis, *J. Chem. Inf. Model.* **2018**, 58, 287–296. DOI:10.1021/acs.jcim.7b00650
55. M. Torrens-Fontanals, P. Tourlas, S. Doerr, G. De Fabritiis, *J. Chem. Inf. Model.* **2024**, 64, 584–589. DOI:10.1021/acs.jcim.3c01776
56. V. N. Viswanadhan, Y. Sun, M. H. Norman, *J. Med. Chem.* **2007**, 50, 5608–5619. DOI:10.1021/jm070261k
57. W. S. Cheung, R. R. Calvo, B. A. Tounge, S. P. Zhang, D. R. Stone, M. R. Brandt, T. Hutchinson, C. M. Flores, M. R. Player, *Bioorg. Med. Chem. Lett.* **2008**, 18, 4569–4572. DOI:10.1016/j.bmcl.2008.07.035
58. P. R. Kym, M. E. Kort, C. W. Hutchins, *Biochem. Pharmacol.* **2009**, 78, 211–216. DOI:10.1016/j.bmcl.2008.07.035
59. D. Goldmann, P. Pakfeifer, S. Hering, G. F. Ecker, *Future Med. Chem.* **2015**, 7, 24–256. DOI:10.4155/fmc.14.168
60. Z. Feng, L. V. Pearce, X. Xu, X. Yang, P. Yang, P. M. Blumberg, X. Q. Xie, *J. Chem. Inf. Model.* **2015**, 55, 572–588. DOI:10.1021/ci5007189
61. S. A. Amin, N. Adhikari, S. Gayen, T. Jha, *J. Biomol. Struct. Dyn.* **2019**, 37, 4528–4541. DOI:10.1080/07391102.2018.1552895
62. P. Kumar, A. Kumar, J. Sindhu, *SAR QSAR Environ. Res.* **2019**, 30, 525–541. DOI:10.1080/1062936X.2019.1629998

Povzetek

V kontekstu farmakološkega posega pri lajšanju bolečine se je transientni receptorski potencial vaniloid tip 1 (TRPV1), nespecifični kationski kanal iz družine TRP ionskih kanalov, izkazal kot obetavna tarča. Kljub temu pa je razpoložljivost selektivnih antagonistov TRPV1 in njihovih farmakoloških lastnosti še vedno omejena. V tem raziskovalnem članku predstavljamo različne tehnike QSAR modeliranja, uporabljene na nizu piperazinil-arilnih spojin, ki delujejo kot antagonisti TRPV1. Opisi, uporabljeni pri oblikovanju konformacijsko neodvisnih QSAR modelov, vključujejo lokalne molekulske grafe in SMILES notacijo, pri čemer je bila za razvoj modela uporabljena tudi Monte Carlo optimizacija. Za oceno kakovosti, robustnosti in napovedne sposobnosti razvitih modelov smo uporabili več statističnih metod, ki so dale pozitivne rezultate. Za najboljši QSAR model so bili doseženi naslednji statistični parametri: za učno množico $R^2 = 0.7155$, CCC = 0.8134, IIC = 0.7430, $Q^2 = 0.6970$, RMSE = 0.645, MAE = 0.489 in F = 157; za testno množico pa $R^2 = 0.9271$, CCC = 0.9469, IIC = 0.9635, $Q^2 = 0.9241$, RMSE = 0.367, MAE = 0.329 in F = 328. Poleg tega smo identificirali molekulske fragmente, izpeljane iz deskriptorjev SMILES notacije, ki pojasnjujejo opažene spremembe v ocenjeni aktivnosti, kar je vodilo do zasnove štirih novih antagonistov. Končna validacija QSAR modela in zasnovanih antagonistov je bila izvedena z molekulskim sidranjem, ki je pokazalo dobro ujemanje z rezultati QSAR modeliranja.



Except when otherwise noted, articles in this journal are published under the terms and conditions of the Creative Commons Attribution 4.0 International License