



Diarylthiazole and diarylimidazole selective COX-1 inhibitor analysis through pharmacophore modeling, virtual screening, and DFT-based approaches

Luminita Crisan¹ · Ana Borota¹ · Alina Bora¹ · Liliana Pacureanu¹

Received: 4 July 2019 / Accepted: 12 August 2019 / Published online: 30 August 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

The current work is focused on in silico modeling of COX-1 inhibitors with enhanced safety gastric profile. A 5-point pharmacophore model, atom-based 3D quantitative structure-activity relationship (3D-QSAR) and electronic properties were computed for a series of COX-1 inhibitors. The best pharmacophore model AAHRR.10 consisting of two hydrogen bond acceptors, one hydrophobic site, and two rings was developed to derive a predictive, statistically significant 3D-QSAR model at three partial least square factors ($R^2 = 0.991$, $SD = 0.059$, $F = 278.5$, $Q^2 = 0.682$, $RMSE = 0.325$, Pearson's $R = 0.903$, Spearman's $\rho = 0.872$). The AAHRR.10 hypothesis was validated by enrichment studies employing a custom-made validation dataset adopting selective COX-1 inhibitors extracted from ChEMBL and decoys generated via DUD methodology. The global reactivity descriptors, such as HOMO and LUMO energies, the HOMO-LUMO gaps, global hardness, softness, Fukui indices, and electrostatic potential, were carried out using density functional theory (DFT) to confirm the key structural features required to achieve COX-1 selectivity. Well-validated AAHRR.10 hypothesis was further used as 3D query in virtual screening of the DrugBank database to detect novel potential COX-1 inhibitors. Docking algorithm was applied to enhance the pharmacophore prediction and to recommend drugs for repositioning, which can interact selectively with COX-1.

Keywords COX-1 · Pharmacophore mapping · DFT · Global reactivity descriptors · Drug repurposing

Introduction

The main cyclooxygenases (COXs) are COX-1, the constitutive enzyme (“housekeeping enzyme”), and the inducible enzyme COX-2 which is responsible for the initiation and maintenance of the inflammation process [1]. COXs are responsible for the biosynthesis of prostaglandins from arachidonic acid (AA) and mediate pathogenic mechanisms [2]. COX-1 stimulates neovascularization and tumor development in the ovary [3]; wherefore, selective inhibition of COX-1 can be

helpful as adjuvant therapy for ovarian cancer. Non-steroidal anti-inflammatory drugs (NSAIDs), the most used classes of analgesics, are prescribed for the treatment of pain, fever, acute and chronic inflammation, and some types of cancers [4]. NSAIDs inhibit both isoforms of COXs, showing various affinities and selectivities. Several side effects including ulcers, gastrointestinal bleeding, deficient coagulation, cerebrovascular, cardiovascular, and renal complications, restraint their use [5, 6]. Selective NSAIDs include COX-1 inhibitors (aspirin, SC-560, mofezolac, FR122047, etc.) [7], COX-2 inhibitors (celecoxib, parecoxib, rofecoxib, parecoxib, valdecoxib, lumiracoxib, etc.) [8, 9], and the competitive non-selective ones (ibuprofen, diclofenac, indomethacin, naproxen, etc.), which inhibit both isoforms showing a preferential trend against COX-1 [10]. Selective COX-2 inhibitors such as celecoxib and parecoxib display analog pharmacological properties but have improved gastrointestinal safety profile compared with traditional NSAIDs [11].

To date, a large number of COX-2 inhibitors (denominated COXIBs) aimed at diminishing gastrointestinal side effects [12] have been developed, but limited attention has been paid

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11224-019-01414-w>) contains supplementary material, which is available to authorized users.

✉ Alina Bora
alina.bora@gmail.com

✉ Liliana Pacureanu
pacureanu@acad-icht.tm.edu.ro

¹ “Coriolan Dragulescu” Institute of Chemistry, Romanian Academy, 24 Mihai Viteazul Ave., 300223 Timisoara, Romania

to design selective COX-1 inhibitors with improved gastrointestinal profile [7, 10, 13, 14]. COX-1 catalytic activity resides in the ability to catch AA among ARG120, TYR355, and the catalytic TYR385. Most of the NSAIDs bearing a carboxylic acid interact with COXs similarly to AA by making a salt bridge with the guanidinium group of the ARG120, situated at the entryway of the hydrophobic binding site channel, which aligns the aromatic moiety of NSAIDs against TYR385, at the top of the active site [15]. The structural factors which determine classical NSAIDs to inhibit preferentially COX-1 are not thoroughly elucidated [16]. Cingolani et al. [17] identified twenty-one out of the twenty-seven co-crystallized NSAIDs-COX-1 complexes where the carboxylate moiety of acidic NSAIDs interacts with the guanidinium group of ARG120. Regardless of their low selectivity, NSAIDs appear to bind more tightly to COX-1 with respect to COX-2, most probably as a result of the energy of the salt bridge interaction between the carboxylate anion of the ligand and the guanidinium cation of ARG120 [18]. The design of selective COX-1 inhibitors relies on the assumption that selective inhibition of COX-1 has to lower pain and inflammation apart from threatening gastric mucosa [19]. In gastric epithelial cells, the dominant isoform is COX-1 [20]. However, gastric tissue exposure is not entirely related to COX-1 inhibition, i.e., 4,5-diarylthiazole and diarylimidazole, which display carboxylic acid group, showed selectivity for COX-1 over COX-2 and low propensity to induce gastric damage [21]. Particularly, NSAIDs with high selectivity for COX-1 contain diarylisoxazole molecular scaffold. The structural evidence for COX-1 binding to diarylisoxazoles has been investigated by X-ray analysis of the COX-1-mofezolac complex that permitted the identification of key binding determinants for structure-activity relationships (SAR) [17]. Hence, the development of new inhibitors which have substantial anti-inflammatory effects but prevent gastrointestinal bleeding remains an open research area [12].

Along these lines, our aim is to establish an *in silico* protocol able to identify selective COX-1 inhibitors with improved gastric profile and to reveal extensive guidelines for the rational design of novel molecules. This goal was accomplished by developing and validating pharmacophore model, custom-made dataset design and screening, and docking, and by density functional theory (DFT) calculations. A pharmacophore model including essential physicochemical features to selectively inhibit COX-1 was obtained. The generated pharmacophore hypothesis was validated by 3D atom-based QSAR aiming at obtaining comprehensive QSAR models that will explain the quantitative differences in the inhibitory activities. The ability of the best pharmacophore hypothesis to identify selective COX-1 inhibitors from a custom-made selective-decoy dataset [22, 23] was verified. Furthermore, we calculated electronic descriptors (“Electronic parameters” section) to confirm the

pharmacophore points required for COX-1 selectivity. The validated pharmacophore hypothesis was used to identify chemical entities from the DrugBank database which were further suggested for repurposing. The probable interactions of the repurposed drugs were identified by docking experiments into COX-1 active site.

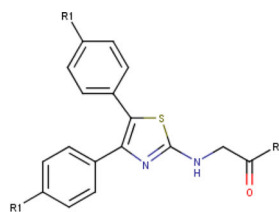
Computational methods

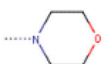
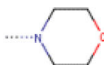
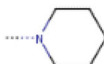
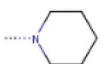
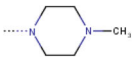
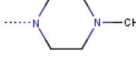
Dataset selection and preparation

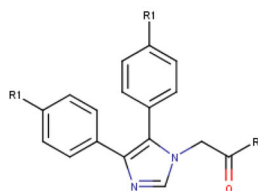
The diarylthiazole and diarylimidazole derivatives as selective COX-1 inhibitors with improved gastric profile were collected from the literature [21]. Most active and selective compounds (7, 8, 15, 16 in Table 1) were used to build the pharmacophore model. The experimental half maximal inhibitory concentration (IC_{50}) values were converted to a logarithmic scale ($pIC_{50} = -\log IC_{50}$). The Marvin tool (ChemAxon) [24] was used for drawing, displaying, and characterizing chemical structures of the investigated ligands. The ionization states and tautomers of each molecule were enumerated using the LigPrep (Schrödinger) in the pH range of 7.2 ± 0.2 [25]. The conformers were generated using default parameters of ConfGen (Schrödinger) [26] by employing the OPLS2005 force field, and only the lowest energy conformer of each ligand was kept.

Validation dataset for virtual screening

The performance of virtual screening (VS) approach is measured by ligand enrichment using benchmarking sets, as shown by numerous publications [27–30]. The datasets of true ligands and their related decoys designated for testing VS methodologies are well-known as validation benchmarking set [31]. The DUD-E database represents an important source of standard and custom-made validation datasets that could be used for the evaluation of VS methods for a wide range of targets [23]. Instead of using DUD-E COX-1 benchmarking set, we take advantage of the ChEMBL25 database [22] to assemble a benchmarking dataset for COX-1 which is focused on selectivity. This custom dataset was built similarly to DUD-E actives sets [23] and further used to test the accuracy of our pharmacophore model. The “duplicate filtering” facility of Instant JChem [32] was applied to a set of 3949 COX-1 inhibitors downloaded from ChEMBL, resulting in 322 unique compounds with an $IC_{50} \leq 1 \mu M$. Out of retrieved 322 unique compounds, 59 compounds satisfy conditions: (i) $IC_{50} \leq 1 \mu M$ related to COX-1 and (ii) the selectivity index related to COX-2 [22] as $SI = IC_{50}COX-2/IC_{50}COX-1 \geq 10$. Afterward, the selectivity custom dataset was subjected to scaffold extraction using Instant JChem [32] resulting in 24 distinct Bemis–Murko chemotypes [33]. Their associated

Table 1 Chemical structures, experimental and predicted pIC_{50} values, selectivity index, and fitness score of diarylthiazole and diarylimidazole analogs, celecoxib, and indomethacin


No.	R	R1	pIC_{50exp}	pIC_{50pred}	SI [†]	Fitness score
1		-H	5.240	5.270	1.440	2.100
2*		-OCH3	5.397	5.690	1.720	2.500
3		-H	5.097	5.170	1.890	2.100
4*		-OCH3	5.221	5.740	2.680	2.700
5*		-H	5.517	5.870	3.630	1.950
6		-OCH3	5.565	5.560	2.920	2.500
7*	-OH	-H	6.377	6.380	25.500	2.260
8	-OH	-OCH3	6.495	6.490	28.840	3.000



decoys including 1300 compounds were generated using the DUD-E free online system (<http://dude.docking.org/>) [23, 34].

The benchmarking dataset was fitted against the AAHRR.10 hypothesis and subsequently ranked by the fitness score values. The fitness score is a measure of how well the selectives and decoys fit with the pharmacophore hypothesis. The compounds showing high values of fitness score were considered true positives if they are COX-1 inhibitors or false positives if they belong to the decoy set. For each compound, the fitness score value is calculated from Eq. 1:

$$\text{Fitness} = W_{\text{site}} \cdot \left(1 - \frac{S_{\text{align}}}{C_{\text{align}}}\right) + W_{\text{vec}} \cdot S_{\text{vec}} + W_{\text{vol}} \cdot S_{\text{vol}} \quad (1)$$

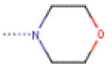
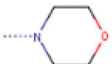
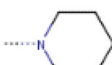
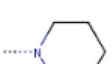
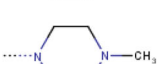
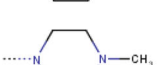
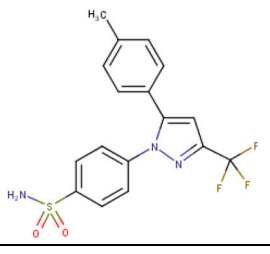
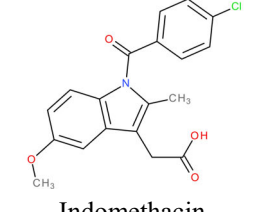
where W_{site} represents the weight of the site score (default is 1.0), S_{align} is the alignment score, C_{align} indicates the alignment cutoff (default is 1.2.), W_{vec} is the weight of the vector

score (default is 1.0), S_{vec} denotes the vector score, W_{vol} shows the weight of volume score (default is 1.0), and S_{vol} is the volume score [35–37].

Pharmacophore modeling and atom-based 3D-QSAR

Phase software (Schrödinger) was engaged to develop the pharmacophore hypothesis [35–37]. Pharmacophore chemical features used contain hydrophobic group (H), aromatic ring (R), hydrogen bond acceptor (A), the hydrogen bond donor (D), negative ionizable group (N), and positive ionizable group (P); no custom features were added. The ligand dataset (Table 1) was divided into actives ($pIC_{50} > 6.1$), inactives ($pIC_{50} < 5.1$), and moderately actives ($6.1 > pIC_{50} > 5.1$). Random selection [35] was employed to split the dataset into a test set of 6 compounds and a training set of 12 (Table 1). All

Table 1 (continued)

No.	R	R1	pIC _{50exp}	pIC _{50pred}	SI [†]	Fitness score
9		-H	5.199	5.270	1.760	0.880
10		-OCH3	5.239	5.250	1.790	2.460
11*		-H	5.070	5.360	1.540	0.980
12		-OCH3	5.134	5.140	1.920	2.520
13		-H	5.380	5.340	2.450	0.710
14		-OCH3	5.514	5.510	2.720	2.520
15	-OH	-H	6.174	6.100	20.070	1.970
16*	-OH	-OCH3	6.268	6.000	23.280	2.640
17			4.790	4.730	0.020	1.760
Celecoxib						
18			6.149	6.200	15.590	1.390
Indomethacin						

*Compounds included in the test set; [†]SI denotes selectivity index

resulted hypotheses were ranked based on survival, energy, volume, vector scores, best active alignment, and the number of site matches [35–37].

Atom-based 3D-QSAR models were generated for the ligand set (Table 1) using partial least squares (PLS) regression based on a grid spacing of 1 Å. To evaluate the quality of all generated hypotheses, the internal and external validation parameters were calculated: (i) training set—squared correlation coefficient (R^2), standard deviation (SD), statistical significance (P), and Fisher's test (F); (ii) test set—squared correlation coefficient (q^2) and Pearson's correlation coefficient (Pearson's R); (iii) complete dataset—Spearman's rank correlation (ρ) The Spearman rank correlation indicates the degree of correlation between experimental and predicted values by

evaluating statistical associations based on the ranks of the data (Eq. 2):

$$\rho = 1 - \frac{6 \times \sum d_i^2}{n(n^2 - 1)} \quad (2)$$

where d_i represents the difference between the ranks of variable Y (activity) and n denotes the number of observations.

Docking

FRED (Fast Rigid Exhaustive Docking; OpenEye) [38–40] docking strategy was used to exhaustively score all possible positions of each ligand in the COX-1 active site. The COX-1

receptor and the ligands were treated as rigid structure during the entire docking process. Ten docking poses were retained for each ligand conformer.

The co-crystal structure of COX-1 complexed with 2-(3,4-bis(4-methoxyphenyl)isoxazol-5-yl)acetic acid (mofezolac) and 3-(5-chlorofuran-2-yl)-5-methyl-4-phenylisoxazole (P6) (PDB ID: 5WBE) [41] was designated as a template for our docking study. This structure was selected based on the similarity between mofezolac and the most active ligand **8** [21]. Generally, the diarylisoxazole moiety is specific for compounds with high selectivity for COX-1 [41]. The 5WBE binding site was used as input and prepared for subsequent grid generation using the Make Receptor (OpenEye) [42]. Active site box (area of the protein where ligands are expected to bind) with the size of 5775 Å³ and two shape contours displaying 80/402 Å³ (inner/outer volumes) were generated, using default parameters without constraints. An RMSD of 0.863 between the atomic coordinates of mofezolac extracted from 5WBE and the docked conformer validated our docking protocol.

Subsequently, the most active and selective COX-1 inhibitors (compounds **7**, **8**, **15**, **16** in Table 1) and the drugs identified by AAHRR.10 hypothesis were docked into the 5WBE receptor. The conformational search ability of OMEGA software (OpenEye) with default parameters was employed to model the ligand conformer ensembles [43–46].

Virtual screening and evaluation techniques

The quality of a pharmacophore is assessed by evaluation of its ability to detect compounds with known biological activity, whereas in terms of predictive accuracy, a valuable pharmacophore model has to identify novel, potential leads that are appropriate for further development. The selected pharmacophore model was used as a 3D query to screen the custom dataset for COX-1. The performance of pharmacophore hypothesis was evaluated using known indicators (Eqs. 3 and 4) [47, 48]: overall prediction accuracy (Acc) and the weighted average precision (W_{PPV}) [49, 50]. The W_{PPV} is highly recommended for imbalanced datasets where the number of decoys is greater than the number of selectives, as the case of our investigation.

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (3)$$

$$W_{PPV} = \frac{\text{PPV}_{\text{CLS } 1} \times \text{N1} + \text{PPV}_{\text{CLS } 2} \times \text{N2}}{\text{N1} + \text{N2}} \quad (4)$$

where TP represents the number of correctly predicted actives/selectives (true positives), TN accounts for the number of correctly predicted decoys (true negatives), FP denotes the

number of mispredicted actives/selectives (false positives), FN stands for the number of mispredicted actives/selectives (false negatives), and N1 is the number of actives (selectives), whereas N2 represents the number of decoys.

Several evaluation metrics (Eqs. 5 and 6) were computed using an *in-house* developed program denoted Evaluation Tool In ChemInformatics (ETICI) [51]. The overall discriminative power for the pharmacophore hypothesis was described by the area under the receiver operating curve (ROC-AUC) [47]. The AUC (Eq. 5) is a measure of the number of actives recovered at any threshold of the hit list, where an AUC of 1 shows the perfect separation between actives (selectives) and decoys. The model performance in the VS test (the early enrichment indicators) was evaluated by calculation of the addition of the early enrichment true positive rates (TPR) at 0.5, 1, 2, 5, and 10% of false positive rates (FPRs) (Eqs. 6–8).

$$\text{AUC} = 1 - \frac{1}{\text{TP} + \text{FN}} \sum_{i=1}^{\text{TP} + \text{FN}} \text{FPR}_i \quad (5)$$

$$\begin{aligned} \text{TPR}_x &= \text{TPR at } x\% \text{FP, where } x \\ &= 0.5\%, 1\%, 2\%, 5\%, 10\% \end{aligned} \quad (6)$$

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (7)$$

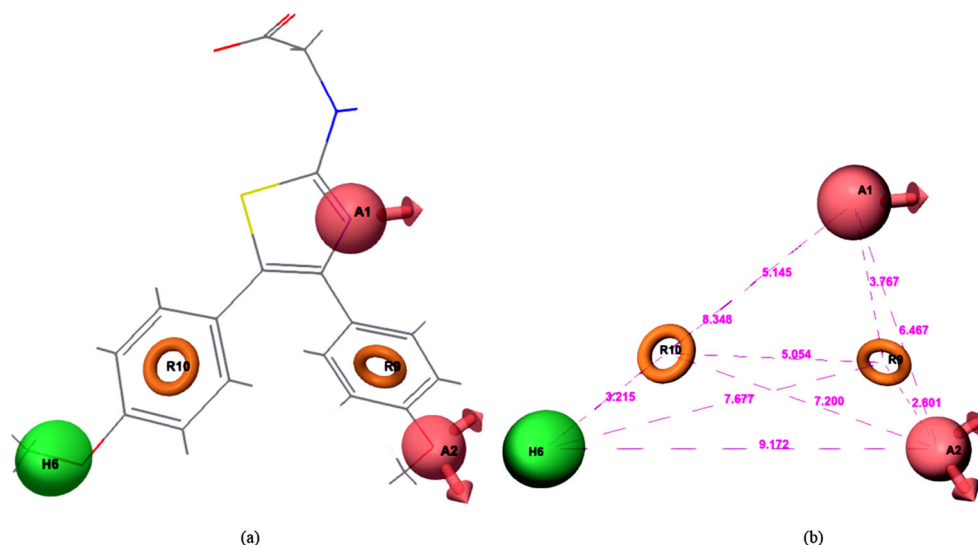
$$\text{FPR} = \frac{\text{FP}}{\text{TN} + \text{FP}} \quad (8)$$

where TPR denotes the fraction of correctly predicted actives/selectives and FPR_i designates the ratio of the number of mispredicted inactives to the total number of inactives when the *i*th active in the ranking list is retrieved.

Computational repurposing

The conventional drug discovery and development are confronted with several challenges such as financial problems, long period of time and multiple steps, and most importantly high attrition rate in preclinical and clinical phases [52]. Detecting new usages for approved drugs and even crowdsourcing of candidate compounds are turning into a promising new endeavor [53]. This scenario seems to be successful and was currently acknowledged by drug agencies, i.e., National Institute of Health (NIH), USA; European Medicine Agency (EMA); and Medical Research Council (MRC), UK, which elaborated policies for research funding to investigate repurposing of drugs [54]. *In silico* approaches such as pharmacophore and molecular docking [30, 55] identified several approved drugs as actives against new targets, which simplify the repositioning for other diseases [56, 57]. The selected drugs that overlapped all

Fig. 1 The pharmacophore hypotheses AAHRR.10 mapped on the most active compound (a); inter-feature distances, acceptor groups A1 and A2 (pink), rings R9 and R10 (orange), and hydrophobic H6 (green) sites (b)



five features of the AAHRR.10 were docked in the 5WBE active site to identify the most suitable orientation and compound binding ability.

Electronic parameters

The conformations of the most active and selective four diarylthiazole and diarylimidazole derivatives (**7**, **8**, **15**, **16**) and celecoxib (**17**) were used as input for DFT studies. The calculated global reactivity descriptors such as the highest occupied molecular orbital (HOMO), the lowest unoccupied molecular orbital (LUMO), the HOMO-LUMO gap energy (ΔE), ionization potential (IP), electron affinity (EA), electronegativity (χ), the global hardness (η), chemical potential (μ), softness (S), electrophilicity (ω), atomic Fukui indices, and atomic electrostatic potential charges (ESP) (Eqs. 9–15) were used (i) to evaluate the reactive sites and the influence of the substituents on the electronic structure of the compounds and (ii) to explain their biological activity/selectivity. All DFT calculations were performed using Jaguar module (Schrödinger) [58, 59]. Complete geometry optimization was assessed using DFT with the

Becke three-parameter exchange potential and Lee-Yang-Parr correlation functional (B3LYP) [60, 61] using 6-31G** basis set [62].

The HOMO orbitals are related to a strong capability to donate electrons (high-energy values) and are straightforwardly correlated with the probability of electrophilic attack. The LUMO orbitals are associated with a strong capability to accept electrons (high-energy values) and are directly associated with the likeliness of nucleophilic attack [63, 64]. The IP and EA descriptors estimate similarly the predisposition of a molecule to accept or donate electrons [63, 65]. The energy difference between the HOMO and LUMO ($\Delta E = E_{\text{HOMO}} - E_{\text{LUMO}}$) gives information regarding chemical reactivity and stability of the molecule [66]. A large ΔE related to global hardness indicates high stability of the molecule, concomitantly with low reactivity, and vice versa [65]. The inverse of the global hardness is defined as softness, S [67]. Chemical potential (μ) describes the escaping tendency of the electron density from the equilibrium state while the negative of the μ , the electronegativity (χ), indicates the ability of an atom in a molecule to attract electrons to itself [68]. The global descriptors were computed according to the following equations (Eqs. 9–15):

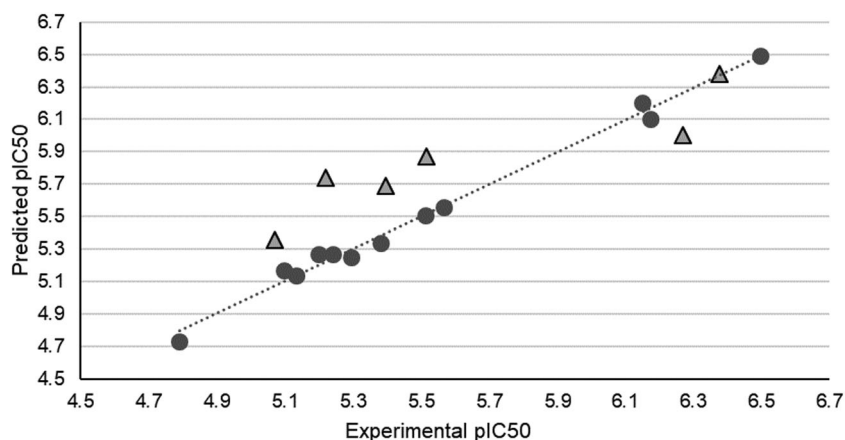
Table 2 The statistical parameters obtained for the atom-based 3D-QSAR model

ID	No.	SD*	R^2 *	F *	p *	RMSE ^{&}	Q^2 ^{&}	Pearson's R ^{&}	Spearman's rho [§]
AAHRR.10	1	0.233	0.812	43.100	6.32E-05	0.373	0.448	0.683	0.878
	2	0.134	0.944	75.600	2.36E-06	0.307	0.627	0.920	0.876
	3	0.059	0.991	278.500	1.98E-08	0.325	0.682	0.903	0.872

*Training set; # test set; § the whole dataset

No., number of PLS factors; SD, standard deviation of the regression; R^2 , the coefficient of determination; F , the ratio of the model variance to the observed activity variance; p , the significance level of variance ratio; RMSE, the root-mean-squared error for the test set predictions; Q^2 , analogs to R^2 based on the test set predictions; Pearson's R , the value for the correlation between the predicted and observed activities for the test set; Spearman's rho is Pearson's correlation coefficient on the ranks of the data

Fig. 2 Plot of experimental versus predicted pIC_{50} values for atom-based 3D-QSAR model AAHRR.10; the black circles illustrates training set compounds, whereas gray triangles depict test set compounds



$$IP = -E_{HOMO} \quad (9)$$

$$EA = -E_{LUMO} \quad (10)$$

$$\mu = -\frac{IP + EA}{2} = \frac{E_{HOMO} + E_{LUMO}}{2} \quad (11)$$

$$\chi = \frac{IP - EA}{2} = -\frac{E_{HOMO} - E_{LUMO}}{2} \quad (12)$$

$$\eta = IP - EA = E_{LUMO} - E_{HOMO} \quad (13)$$

$$S = \frac{1}{\eta} \quad (14)$$

$$\omega = \frac{\mu^2}{2\eta} \quad (15)$$

The Fukui indices are associated with the tendency of atoms to lose or accept electrons [69]. These indices specify which atoms in a molecule are more susceptible to a nucleophilic or an electrophilic attack [70]. The Maestro module (Schrödinger) was used to visualize the influence of the above parameters [71]. These calculations could provide useful

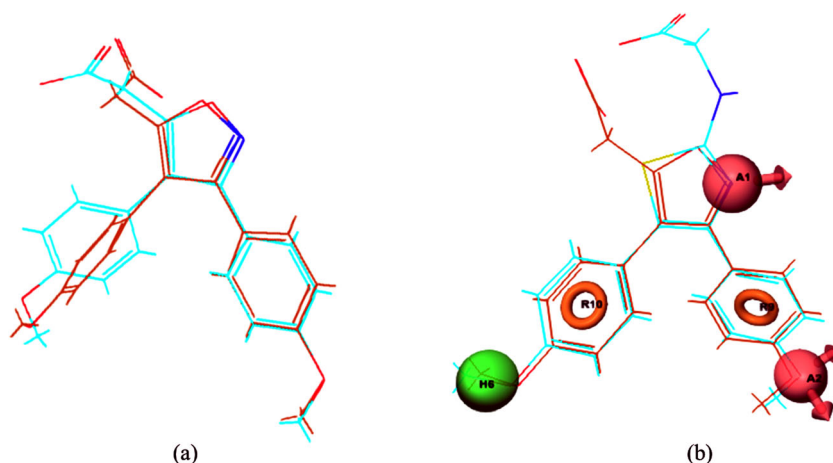
features for designing new COX-1 inhibitors with the enhanced gastric profile.

Results

Pharmacophore modeling

In order to investigate the pharmacophoric space of diarylthiazole and diarylimidazole derivatives, all possible five-featured pharmacophore hypotheses were generated. Among the 29 hypotheses, six displays significant statistics parameters from atom-based 3D-QSAR point of view. The AAHRR.10 hypothesis, including two hydrogen bond acceptors, one hydrophobic site, and two rings (Fig. 1a), mapped on the most active compound **8** displayed the best values for survival score (3.639), site score (0.950), and vector score (0.980). The distances between the pharmacophore sites of the AAHRR.10 reported in Å are shown in Fig. 1b. Compound **8** matches exactly the AAHRR.10 with a perfect fitness score value of 3 (Table 1). This hypothesis was retained for in-depth investigation.

Fig. 3 Overlay of mofezolac bound conformation extracted from co-crystal 5WBE (carbon depicted in cyan) and its pharmacophore fitted conformation (carbon depicted in brown) (a); overlay between pharmacophore fitted conformation of mofezolac (carbon depicted in brown) and the most active compound **8** (carbon depicted in cyan) (b)



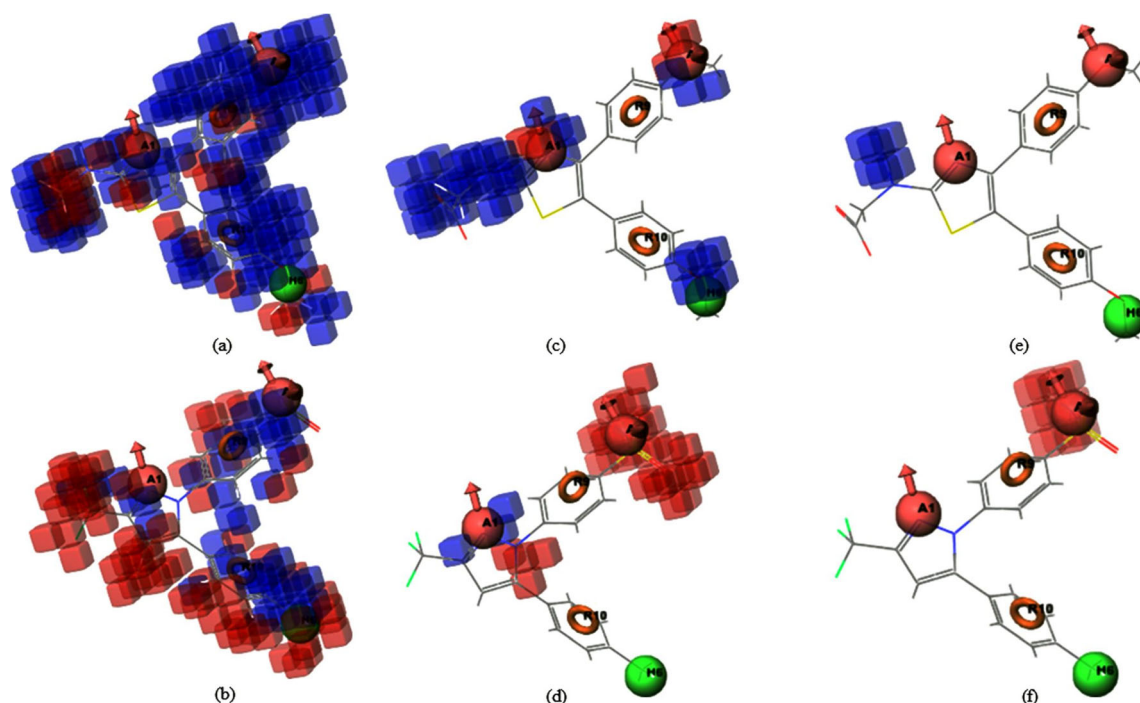


Fig. 4 Hydrophobic non-polar space for the most active compound **8** (a) and the least active compound **17** (b). Electron-withdrawing effect for the most active compound **8** (c) and the least active compound **17** (d).

Hydrogen bond donor field for the most active compound **8** (e) and the least active compound **17** (f)

Atom-based 3D-QSAR

An atom-based 3D-QSAR model containing three PLS factors was developed to validate the AAHRR.10 hypothesis (Table 2 and Fig. 2). The R^2 value of 0.991 for the training set suggests that the matrix of descriptors and the matrix of the experimental inhibitory activities are strongly correlated. The large values of the Fisher test (F) and low values for the significance level of variance ratio (p) and for the standard deviation of the regression (SD) indicate a statistically significant 3D-QSAR model. Also, the good predictive abilities of the model were highlighted by the Q^2 value of 0.682 and the Pearson R value of 0.903, for the test set. The Spearman rank correlation value higher than 0.8 suggests a high positive correlation between the experimental and predicted activities (Table 2).

Moreover, the AAHRR.10 was validated in terms of active conformation prediction by overlaying mofezolac conformation matched on AAHRR.10 with its experimental structure [17, 42] resulting in RMSD of 1.152 Å (Fig. 3a). In a similar manner, the mofezolac and compound **8** conformations matched on the AAHRR.10 showed favorable hydrophobic zones localized on phenyl rings where the presence of methoxy substituent increase inhibitory activity (Fig. 3b).

The AAHRR.10 was employed to generate the contour plot analysis in order to establish the positions of the

substituents, which are responsible for biological activity. The contour maps including hydrophobic, electron-withdrawing, and hydrogen bond donor effects for the most active compound **8** and the least active compound **17** were plotted (Fig. 4a–f). The most favorable regions are depicted in blue, whereas the detrimental features are rendered in red.

For the most active compound **8** (Fig. 4a), the hydrophobic favorable blue region is dispersed over the entire molecule, but the most definite area is delineated by the phenyl ring R9 and its substituent methoxy. On the contrary, the least active compound **17** (Fig. 4b) exhibits a non-favorable hydrophobic area spread on the entire molecule along with a small hydrophobic favorable area localized on R9, thiazole ring and its substituent at position 3. Concerning the electron-withdrawing properties (Fig. 4c, d), the presence of blue cubes reveals that the oxygen atoms belonging to carboxylic and methoxy groups of compound **8** can act as hydrogen bond acceptor. This interaction is supported by X-ray complex of COX-1 with mofezolac [17, 41] where the oxygen atom which belongs to the carboxyl group of the isoxazole substituent interacts by means of hydrogen bond and salt bridge with the nitrogen atom of the guanidine side chain of ARG120 and oxygen by a hydrogen bond with TYR355. In addition to these well-known H bond interactions, two carbon–hydrogen bond interactions were observed: LEU352 with the

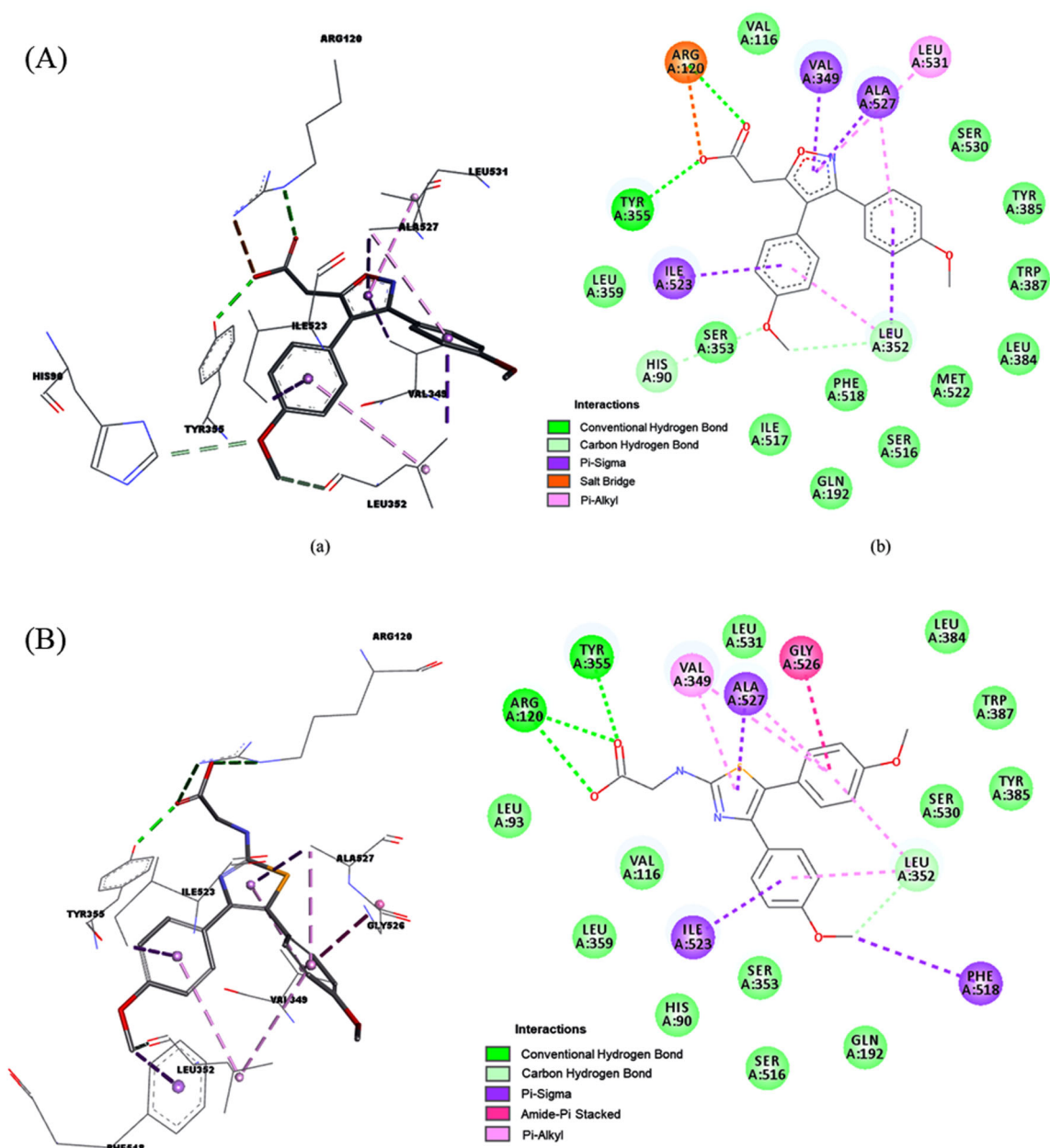


Fig. 5 (A) Interactions of 2-(3,4-bis(4-methoxyphenyl)isoxazol-5-yl)acetic acid (mofezolac) with COX-1 binding site residues (5WBE) [17, 41] illustrated as (a) 3D and (b) 2D versions; (B) the docked complex of 2-[[4,5-bis(4-methoxyphenyl)-1,3-thiazol-2-yl]amino]acetic acid

(compound **8**) with COX-1 binding site residues illustrated as (a) 3D and (b) 2D versions; the (A) and (B) legends describe [71] the conventional hydrogen bond (green), hydrophobic π -sigma, π -alkyl and π -stacked (purple, pink and magenta), and salt bridge (orange) interactions

carbon atom and HIS90 with the oxygen atom of the methoxyphenyl group corresponding to the hydrophobic area H6, (Fig. 5A) [72]. The presence of favorable blue cubes in the region of the $-\text{NH}-$ group of **8** (Fig. 4e) indicates that hydrogen bond donor property of this group will favor the increase of potency.

LEU352, ALA527, and LEU531 form π -alkyl interactions with phenyl ring from position 4, with phenyl ring from position 3, and with the nitrogen atom of isoxazole ring (A1 point on AAHRR.10). ILE523, LEU352, VAL349, and ALA527 display π -sigma interactions with isoxazole and phenyl rings

at positions 3 and 4 of mofezolac. As can be seen in Figs. 3b and 5A [72], the two aromatic rings R9 and R10 (AAHRR.10) correspond to phenyl rings located at positions 3 and 4 of isoxazole ring of mofezolac, forming hydrophobic interactions with the hydrophobic region delineated by ALA527, LEU531, ILE523, LEU352, and VAL349. Hence, the pharmacophore hypothesis developed (A1, H6, R9, and R10) was confirmed by mofezolac interactions with COX-1; moreover, the 3D-QSAR model is statistically reliable and can be used for the design of novel COX-1 inhibitors within the current framework.

Table 3 The performance of the AAHRR.10 hypothesis against the custom-made dataset

Evaluation parameter	Pharmacophore features*	
	A	B
%TPR at 0.5% FPs	12.500	33.333
%TPR at 1% FPs	12.500	33.333
%TPR at 2% FPs	12.500	33.333
%TPR at 5% FPs	18.750	33.333
%TPR at 10% FPs	25.000	33.333
W_{PPV}	0.972	0.967
Acc	0.558	0.833
AUC	0.622	0.706
SD AUC	0.077	0.117
AROCE-AUC	0.378	0.520

*A—Four “essential” pharmacophore features; B—Five “essential” pharmacophore features

Virtual screening

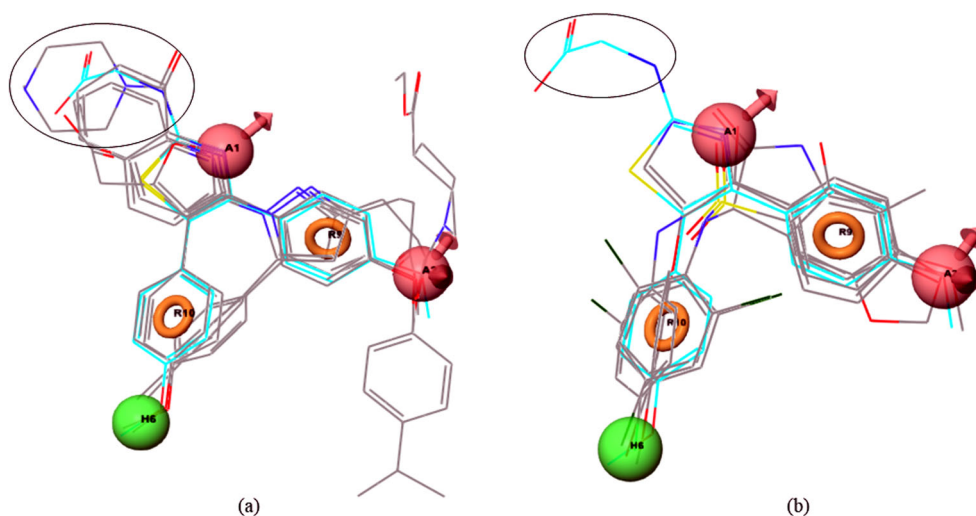
In order to validate the discriminatory ability of AAHRR.10, we performed a VS test against custom-made dataset for COX-1. ROC-AUC; %TPR at 0.5%, 1%, 2%, 5%, and 10% FPs; Acc; and W_{PPV} metrics were calculated to assess the ranking of the selectives in the enrichment list (Table 3). The evaluation parameters expressed by AUC resulted from running AAHRR.10 model yield values greater than 0.7 showing reasonable discriminative performance regarding the retrieval of selective COX-1 inhibitors. The values of early enrichment evaluation parameters indicate a good recognition of selectives at 0.5 to 10% TPR. The Acc values indicate a good accuracy of the model when five “essential” pharmacophore features were considered. The values closest to 1 for W_{PPV} metric reveal a high classification power of the model.

The good values of evaluation parameters associated with the five “essential” pharmacophore features (Table 3) suggest the selective character of the AAHRR.10. Thus, the alignment of the top four selectives and decoys ranked according to fitness score is shown in Fig. 6. As mentioned above, the presence of a hydrogen bond acceptor/donor group in a position corresponding to position 5 of the isoxazole ring of mofezolac, capable to form hydrogen bond or salt bridge interactions with ARG120, is very important for affinity and selectivity towards COX-1. In the case of selective compounds, the substituents matching this position are overlapped (Fig. 6a) while in the case of the decoys, no substituent at this position was present (Fig. 6b).

Molecular docking

Aiming at investigating structural prerequisites which drive affinity and selectivity towards COX-1, a docking experiment of selective compounds (7, 8, 15, 16) was carried out. The docking outcomes were compared with experimental interactions of COX-1 selective inhibitor, mofezolac (Fig. 5A). These selective COX-1 inhibitors showed the same orientation and binding pattern to that of mofezolac (Fig. S1). Both, compound 8 and mofezolac are polar and ionizable molecules. The interactions of the most active compound 8 observed into COX-1 binding site are similar to those which involve the X-ray coordinates of mofezolac (Fig. 5B). Thus, oxygen atoms of the carboxylate group interact via hydrogen bonds with ARG120 and TYR355. Likewise, a carbon–hydrogen bond occurs between 4-OCH₃ group bound to the phenyl ring attached at position 4 of thiazole ring and LEU352. A π - σ interaction was observed between the thiazole ring and ALA527. π -Alkyl contacts were registered with VAL349 (thiazole ring),

Fig. 6 The overlap of the top-ranked four selectives (a) and decoys (b) on AAHRR.10 hypothesis (the carbon atoms of the most active compound 8 are shown in cyan)



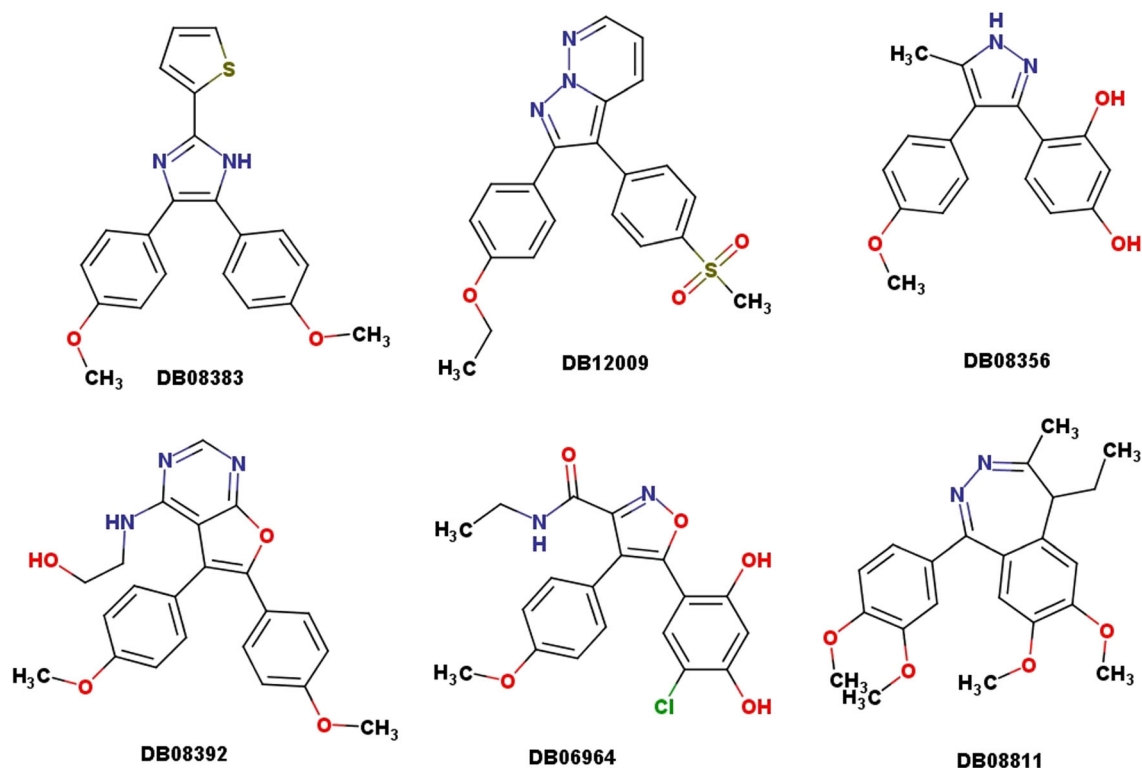


Fig. 7 Chemical structures of the investigational drugs and Tofisopam (**DB08811**) proposed for repurposing using AAHRR.10 hypothesis

phenyl ring of $-C_6H_4-4-OCH_3$ substituent at position 5 of thiazole ring with VAL349, ALA527, and LEU352. Amide- π staking interaction takes place between GLY526 and phenyl substituent at position 5 of the

thiazole ring. Our docking outcomes regarding the most active compound complete the docking simulations of Abdelazeem et al. [21].

Drug bank screening

In order to identify approved and investigational drugs, the AAHRR.10 was used as a template to screen the DrugBank database [73]. The prioritized drugs according to fitness score criterion include drugs which are under testing in various clinical phases: **DB08383** (2.741), **DB12009** (2.575), **DB08356** (2.574), **DB08392** (2.534), **DB06964** (2.492), and approved drug **Tofisopam (DB08811)** (2.100) (Figs. 7 and 8). All six prioritized drugs satisfied the 5-point pharmacophore features. **Tofisopam** is already approved for the treatment of anxiety and alcohol withdrawal [73]. Of the five prioritized drugs, only **DB12009** has been evaluated in trials for the treatment of pain, trauma, neurodynia, dental pain, and hyperalgesia [73].

Further, docking of the six shortlisted drugs within COX-1 active site of 5WBE enabled us to prioritize **DB08392** and **DB06964** due to the hydrogen bond interactions with ARG120. The reproduction of the conventional hydrogen bonding interaction with ARG120 and TYR355, and the hydrophobic interactions with ALA527, with LEU352 and ILE523 residues similar to mofezolac, were noticed (Fig. 9 [72]).

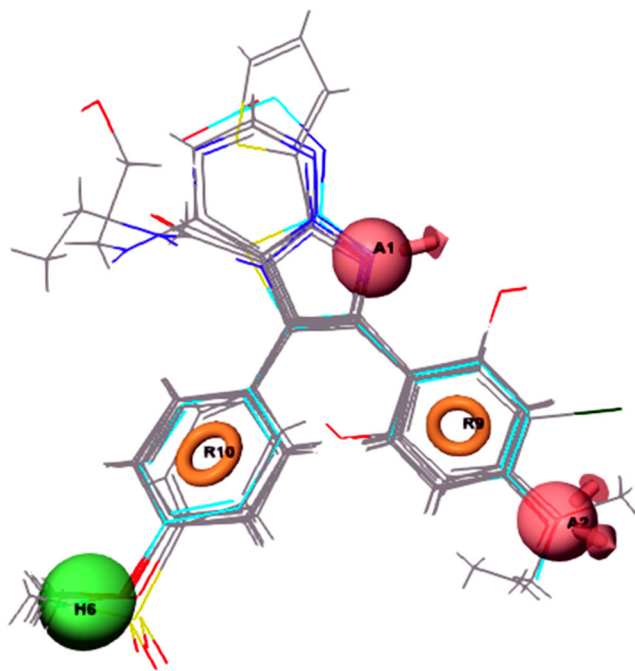
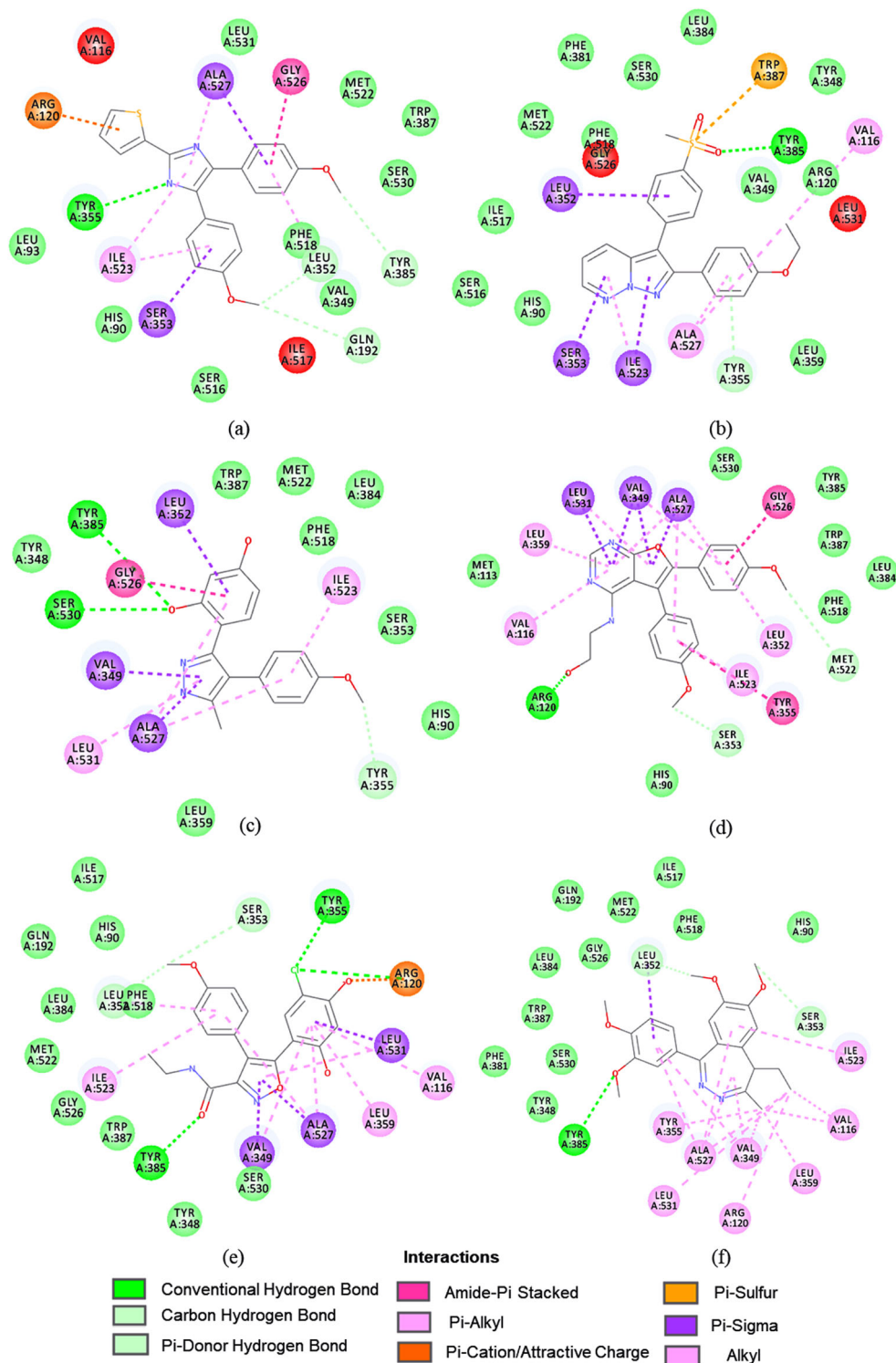


Fig. 8 The investigational drugs and Tofisopam overlaid on the AAHRR.10 hypothesis (the most active compound **8** is shown in cyan)

Fig. 9 The interactions of prioritized drugs with COX-1 residues (PDB ID 5WBE): **DB08383** (a), **DB12009** (b), **DB08356** (c), **DB08392** (d), **DB06964** (e), and **Tofisopam (DB08811)** (f); hydrogen bond (green), hydrophobic (purple), and salt bridge (orange)



All selected drugs experience hydrophobic interactions between aromatic and heteroaromatic rings and binding site residues. π -Alkyl and π -sigma interactions were observed also with VAL116, VAL349, LEU359, PHE518, and LEU531 of COX-1 channel, whereas **DB08811** makes a carbon–hydrogen bond with more polar SER353. For **DB08356**, no

essential H bond interactions and steric bumps were visualized. The analysis of drug-COX-1 interactions displayed several sites of unfavorable steric bumps: (i) **DB08383**—sulfur and carbon atoms of thiophene ring with carbon atom of methyl group of VAL116 (4.2 Å and 1.9 Å), and between carbon atom of methyl group and carbon atom of methyl group of

Table 4 The calculated electronic features: E-HOMO, E-LUMO, HOMO and LUMO gap energies (ΔE), ionization potential (IP), electron affinity (EA), electronegativity (χ), electronic hardness (η),global softness (S), the chemical potential (μ), and electrophilicity index (ω) for the most active compounds (**7**, **8**, **15**, **16**) and celecoxib (**17**)

No.	E-HOMO (eV)	E-LUMO (eV)	ΔE	IP	EA	χ (eV)	η (eV)	S (eV)	μ (eV)	ω (eV)
7	-5.196	-0.964	4.232	5.196	0.964	3.080	4.232	0.236	-3.080	1.211
8	-4.853	-0.694	4.159	4.853	0.694	2.774	4.159	0.240	-2.774	0.925
15	-5.387	-0.708	4.679	5.387	0.708	3.048	4.659	0.214	-3.048	0.993
16	-4.950	-0.446	4.504	4.950	0.446	2.698	4.504	0.222	-2.698	0.808
17	-5.599	-1.723	3.876	5.599	1.723	3.661	3.876	0.258	-3.661	1.729

ILE517 (2.02 Å) (Fig. 9a) [72]; (ii) **DB12009** carbon atom of methyl group and carbon atoms of side chain of LEU531 (2.13 Å and 1.74 Å), carbon atom of methylene group with carbon atom of side chain of LEU531 (1.79 Å), and carbon atom of 4-methanesulfonyl group with oxygen atom of C=O group of GLY526 (2.07 Å). The steric bumps provide information about the rigidity and bulkiness to promote energy favorable positions of the ligand binding to COX-1 active channel. However, the molecular size of ligands represents a crucial element to accommodate into the COX-1 active site, since COX-1 (316 Å³) active site is smaller than in COX-2 (394 Å³). Nevertheless, a number of hydrogen bonds and hydrophobic interactions are registered with essential binding site residues while the steric hindrance restricts favorable interactions. Hence, we can assume that **DB08383** and **DB12009** are not interesting for prioritization as COX-1 ligands.

To sum up, **DB06964** and **DB08392** are the most plausible to inhibit COX-1 in a selective manner due to hydrogen bond interactions with ARG120 and a higher number of π -alkyl and

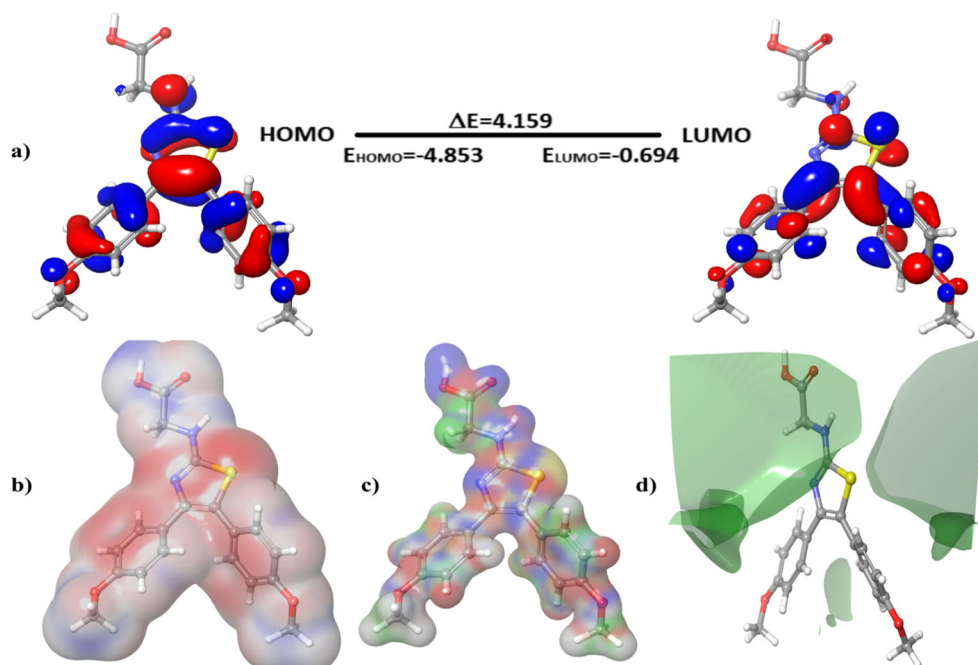
π -sigma contacts, which fixed the compounds in the active site.

Electronic parameters

According to frontier orbital theory, the shapes and symmetries of the HOMO and LUMO orbitals are essential features in predicting the reactivity and stability of a compound in a chemical reaction or ligand-receptor interactions [41]. The calculated electronic features are listed in Table 4. Maps of HOMO and LUMO orbitals, the blue color indicates positive lobes while red color designates negative lobes, are plotted onto the molecular surface of the most active compounds (**7**, **8**, **15**, **16**) and celecoxib (**17**) completed by average local ionization, electrostatic potential and charge distribution profiles (Fig. 10 and Fig. S2).

Analyzing the HOMO energy and the biological activity values of the COX-1 selective inhibitors and celecoxib (Tables 1 and 4), an inverse correlation is clearly observed, suggesting that the HOMO of the selective inhibitors may

Fig. 10 The 3D HOMO and LUMO orbital profiles, the orbital energy values, and the energy gap (a); the electrostatic potential profiles (b); the charge distribution profiles (c); and the average local ionization (d) for the most active compound **8**



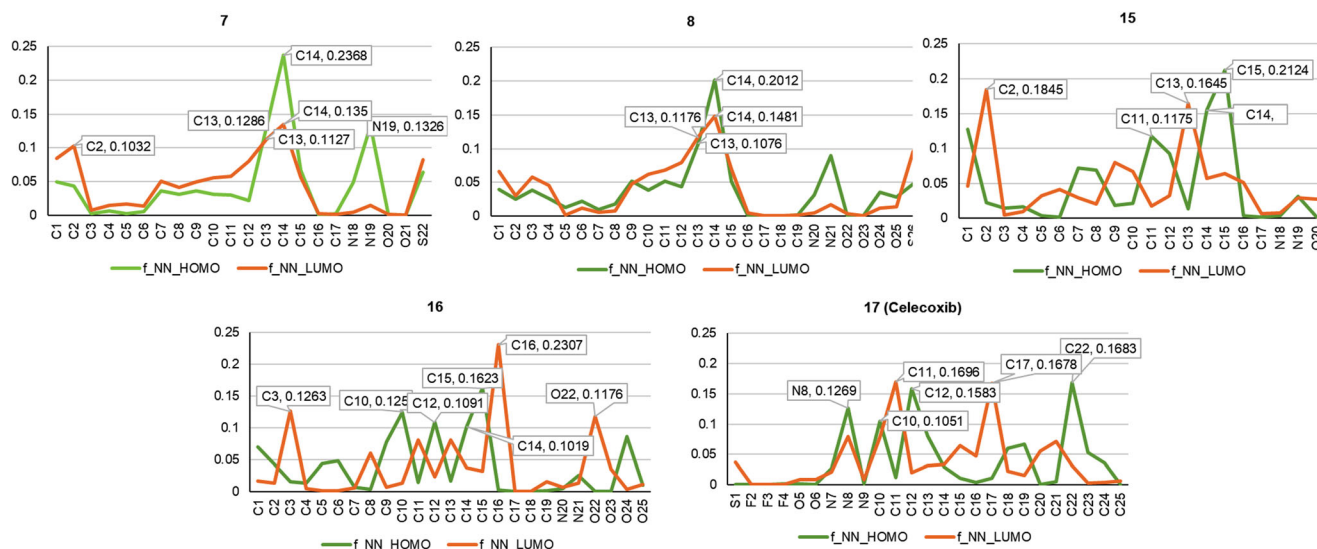


Fig. 11 Plots of the reactivity Fukui indices for compounds **7**, **8**, **15**, **16**, and **17**, where f_{NN_HOMO} (green line) and f_{NN_LUMO} (orange line) indicate $f_{ukui_electron_density_HOMO}$ and $f_{ukui_electron_density_LUMO}$, respectively

transfer easier its electrons to some residues in the active site of COX-1 for binding. The calculations show that compound **8** displays the highest HOMO energy (-4.853 eV) which is in good agreement with its affinity (0.32 μ M). Heteroaromatic rings including nitrogen and sulfur can act as electron donors or acceptors to the residues in the COX-1 binding site. Electron donor thiazole and imidazole rings display the highest electron density of the HOMO (Fig. 10). The phenyl rings corresponding to R9 and R10 in AAHRR.10 exhibit a high electronic density of HOMO that favors hydrophobic interactions. The influence of the HOMO energy on the biological activity can be rationalized in terms of charge transfer, $\pi \cdots \pi$, or $\pi \cdots \sigma$ stacking between phenyl rings of ligands and the binding site residues. A smaller energy gap between HOMO and LUMO orbitals assigned to most active molecule **8** (ΔE of 4.159) illustrates that the molecule is more active, while the larger energy gap of molecule **15** (ΔE of 4.679) indicates high chemical hardness and lower activity, according to the maximum hardness principle (Table 2, Fig. 10) [74]. The large energy gap of a molecule is unfavorable for the electron to be excited from HOMO to LUMO, which leads to a weaker affinity of the COX-1 inhibitors. The larger value of ω indicates also higher reactivity of the chemical system, which can be correlated to biological activity [75]. The average local ionization for the most active compound **8** (Fig. 10) is localized around the carboxyl group, suggesting the favorite site for electrophile attack which provides selectivity towards COX1 by interaction with ARG120 and TYR355. For the most active ligands (Table 4), the inhibitory activity is directly correlated with η ($r^2 = 0.949$) and ΔE ($r^2 = 0.947$) metrics and inversely correlated with the S parameter ($r^2 = 0.952$) (Figs. 10 and S3). These observations are in accord with the previous

investigation which states that the inhibitory properties of drugs on enzymes is governed by an absolute hardness or absolute electronegativity [76].

In accord with AAHRR.10, thiazole and imidazole rings display a hydrogen bond acceptor A1 (Fig. 3b). The most active compounds **7** and **8**, having the LUMO located over the phenyl and methoxyphenyl rings (R9 and R10 on AAHRR.10), suggest the susceptibility of the molecules towards nucleophilic attack. The ESP and the charge distribution profiles highlight the negative potential regions around the oxygen atoms of the carboxyl group extended over the nitrogen atom of the methylcarbamic acid chain and the nitrogen atom of thiazole and imidazole rings. The investigation of ESP is helpful because it influences the initial stages of ligand-receptor interactions and allows prediction of the functional sites [77].

The plots of Fukui indices (Fig. 11) emphasizes the atomic sites susceptible to nucleophilic or/and electrophilic attack. The significant values that correspond to the highest peaks of the Fukui indices were plotted into the callouts. As can be seen from Fig. 11, the atomic Fukui values also correlate with the ESP and charge distribution plots. These outcomes are consistent with the pharmacophore analysis, which illustrates the participation of these moieties in the key ligand-receptor interactions and certify the relationship between electronic properties and potency of diarylthiazole and diarylimidazole analogs.

Conclusion

In the current study, the pharmacophore models, 3D atom-based QSAR, the custom-made dataset screening, docking,

and DFT calculations were achieved based on a series of known diarylthiazole and diarylimidazole COX-1 inhibitors. The best pharmacophore model, AAHRR.10, consisting of two hydrogen bond acceptors, two rings, and one hydrophobic feature, was validated by (i) 3D atom-based QSAR, (ii) resembling the mofezolac interactions with COX-1 (5WBE), and (iii) the custom-made selective/decoy set. The validation results showed a high correlation between experimental and predicted IC_{50} values ($R^2 = 0.991$), satisfactory predictive ability ($Q^2 = 0.682$, and Pearson's $R = 0.903$), and good discriminative performance regarding the retrieval of selective COX-1 inhibitors (AUC = 0.706). These validation outcomes confirmed that AAHRR.10 was the best hypothesis to differentiate the selective COX-1 inhibitors from decoys. Moreover, the AAHRR.10 was used as a 3D query in the DrugBank virtual screening process. Based on the pharmacophore screening, docking analysis, and ligand-protein interactions studies, two investigational drugs, DB06964 and DB08392, have been suggested for repurposing. The impact of electrostatic features of diarylthiazole and diarylimidazole units on their COX-1 inhibitory activity was highlighted using DFT calculations. The global reactivity descriptor outcomes have successfully validated the key pharmacophore features of AAHRR.10. The HOMO-LUMO gap energy was subsequently used to probe the strength and stability of the molecular interactions for the most selective ligands. Accordingly, this theoretical approach can be useful in the development of new selective inhibitors of COX-1 with the improved gastric profile.

Acknowledgments The authors thank Dr. Ramona Curpăn (“Coriolan Dragulescu” Institute of Chemistry, Romanian Academy), for providing access to Schrödinger software acquired through the PN-II-RU-TE-2014-4-422 projects funded by CNCS-UEFISCDI Romania, to OpenEye, BIOVIA Discovery Studio, and Chemaxon for the free academic licenses.

Funding information This project was financially supported by Project 1.2 of the “Coriolan Dragulescu” Institute of Chemistry, Romanian Academy.

Compliance with ethical standards

The authors declare that they have no conflict of interest. The authors alone are responsible for the content and writing of the paper.

References

- Gupta RA, Tejada LV, Tong BJ, Das SK, Morrow JD, Dey SK, DuBois RN (2003). *Cancer Res* 63:906–911
- Litalien C, Beaulieu P (2011) In: Fuhrman BP, Zimmerman JJ (eds) *Pediatric Critical Care*, 4th edn. PA. Elsevier Saunders, Philadelphia, pp 1553–1568
- Marnett LJ (2009). *Annu Rev Pharmacol Toxicol* 49:265–290
- Emery P (1999). *Drugs Today (Barc)* 35:267–274
- Wolfe MM, Lichtenstein DR, Singh G (1999). *N Engl J Med* 340: 1888–1899
- Abdelazeem AH, Khan SI, White SW, Sufka KJ, McCurdy CR (2015). *Bioorg Med Chem* 23:3248–3259
- Perrone MG, Scilimati A, Simone L, Vitale P (2010). *Curr Med Chem* 17:3769–3805
- Howes L (2007). *Ther Clin Risk Man* 3:831–845
- Oniga SD, Pacureanu L, Stoica CI, Palage MD, Crăciun A, Rusu LR, Crisan EL, Aranicu C (2017). *Molecules* 22:1507
- Hawkey CJ (2001). *Best Pract Res Clin Gastroenterol* 15:801–820
- Tacconelli S, Capone ML, Patrignani P (2004). *Curr Pharm Des* 10: 589–601
- Perrone MG, Lofrumento DD, Vitale P, De Nuccio F, La Pesa V, Panella A, Calvello R, Cianciulli A, Panaro MA, Scilimati A (2015). *Pharmacology* 95:22–28
- Di Nunno L, Vitale P, Scilimati A, Tacconelli S, Patrignani P (2004). *J Med Chem* 47:4881–4890
- Liedtke AJ, Crews BC, Daniel CM, Blobaum AL, Kingsley PJ, Ghebreselasie K, Marnett LJ (2012). *J Med Chem* 55:2287–2300
- Smith WL, Garavito RM, DeWitt DL (1996). *J Biol Chem* 271: 33157–33160
- Shamsudin Khan Y, Kazemi M (2015) Gutierrez-de-Tera'n H, Aqvist J. *Biochemistry* 54:7283–7291
- Cingolani G, Panella A, Perrone MG, Vitale P, Di Mauro G, Fortuna CG, Armen RS, Ferorelli S, Smith WL, Scilimati A (2017). *Eur J Med Chem* 138:661–668
- Smith WL, DeWitt DL, Garavito RM (2000). *Annu Rev Biochem* 69:145–182
- Peskar BM (2001). *J Physiol Paris* 95:3–9
- Baumgartner HK, Starodub OT, Joehl JS, Tackett L, Montrose MH (2004). *Gut* 53:1751–1757
- Abdelazeem AH, El-Saadi MT, Safi El-Din AG, Omar HA, El-Moghazy SM (2017). *Bioorg Med Chem* 25:665–676
- Gaulton A, Hersey A, Nowotka M, Bento AP, Chambers J, Mendez D, Mutowo P, Atkinson F, Bellis LJ, Cibrián-Uhalte E, Davies M, Dedmat N, Karlsson A, Magariños MP, Overington JP, Papadatos G, Smit I, Leach AR (2017). *Nucleic Acids Res* 45:D945–D954
- Mysinger MM, Carchia M, Irwin JJ, Shoichet BK (2012). *J Med Chem* 55:6582–6594
- MarvinSketch v17.18.00 (2017) ChemAxon. <https://www.chemaxon.com>. Accessed March 2019
- Schrödinger Release 2016-1: LigPrep v3.1 (2016) Schrödinger LLC, New York, NY. <https://www.schrodinger.com/citations#Jaguar>
- Watts KS, Dalal P, Murphy RB, Sherman W, Friesner RA, Shelley JC (2010). *J Chem Inf Model* 50:534–546
- McGaughey GB, Sheridan RP, Bayly CI, Culberson JC, Kreatsoulas C, Lindsley S, Maiorov V, Truchon JF, Cornell WD (2007). *J Chem Inf Model* 47:1504–1519
- von Korff M, Freyss J, Sander T (2009). *J Chem Inf Model* 49:209–231
- Hu G, Kuang G, Xiao W, Li W, Liu G, Tang Y (2012). *J Chem Inf Model* 52:1103–1113
- Crisan L, Avram S, Pacureanu L (2017). *Mol Divers* 21:385–405
- Bissantz C, Folkers G, Rognan D (2000). *J Med Chem* 43:4759–4767
- Instant JChem v17.17.0 (2017) ChemAxon. <http://www.chemaxon.com>. Accessed March 2019
- Bemis GW, Murcko MA (1996). *J Med Chem* 39:2887–2893
- Nicholls A (2008). *J Comput Aided Mol Des* 22:239–255
- Schrödinger Release 2016-3: Phase v3.8 (2016) Schrödinger, LLC, New York, NY. <https://www.schrodinger.com/citations#Phase>
- Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA (2006). *J Comput Aided Mol Des* 20:647–671. <https://www.schrodinger.com/citations#Phase>

37. Dixon SL, Smondyrev AM, Rao SN (2006). *Chem Biol Drug Des* 67:370–372. <https://www.schrodinger.com/citations#Phase>
38. FRED v.3.2.0.2 OpenEye Scientific Software Inc. Santa Fe NM, USA, <https://www.eyesopen.com>. Accessed March 2019
39. McGann M (2011). *J Chem Inf Model* 51:578–596
40. Kelley BP, Brown SP, Warren GL, Muchmore SW (2015). *J Chem Inf Model* 55:1771–1780
41. RCSB Protein Data Bank, RCSB PDB, <https://www.rcsb.org/structure/5WBE> (accessed in March 2019)
42. Make Receptor v.3.2.0.2 OpenEye Scientific Software Inc., Santa Fe NM, USA <https://www.eyesopen.com>. Accessed March 2019
43. OMEGA v.2.5.1.4 OpenEye Scientific Software Inc. Santa Fe NM, USA <https://www.eyesopen.com>. Accessed March 2019
44. Hawkins PCD, Skillman AG, Warren GL, Ellingson BA, Stahl MT (2010). *J Chem Inf Model* 50:572–584
45. Hawkins PCD, Nicholls A (2012). *J Chem Inf Model* 52:2919–2936
46. Boström J, Greenwood JR, Gottfries J (2003). *J Mol Graph Model* 21:449–446
47. Fawcett T (2006). *Pattern Recognition Letters* 27(8):861–874
48. Matthews BW (1975). *Biochim Biophys Acta (BBA)* 405:442–451
49. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009). *Explor Newsl* 11:10–18
50. Kotsampasakou E, Ecker GF (2017). *J Chem Inf Model* 57:608–615
51. Avram SI, Pacureanu LM, Bora A, Crisan L, Avram S, Kurunczi L (2014). *J Chem Inf Model* 54:2360–2370
52. Ashburn TT, Thor KB (2004). *Nat Rev Drug Discov* 3:673–683
53. Oprea TI, Bauman JE, Bologna CG, Buranda T, Chigaev A, Edwards BS, Jarvik JW, Gresham HD, Haynes MK, Hjelle B, Hromas R, Hudson L, Mackenzie DA, Muller CY, Reed JC, Simons PC, Smagley Y, Strouse J, Surviladze Z, Thompson T, Ursu O, Waller A, Wandinger-Ness A, Winter SS, Wu Y, Young SM, Larson RS, Willman C, Sklar LA (2011). *Drug Discov Today Ther Strateg* 8:61–69
54. Frail DE, Brady M, Escott KJ, Holt A, Sanganee HJ, Pangalos MN, Watkins C, Wegner CD (2015). *Nat Rev Drug Discov* 14:833–841
55. Miller JR, Dunham S, Mochalkin I, Banotai C, Bowman M, Buist S, Dunkle B, Hanna D, Harwood HJ, Huband MD, Karnovsky A, Kuhn M, Limberakis C, Liu JY, Mehrens S, Mueller WT, Narasimhan L, Ogden A, Ohren J, Prasad JV, Shelly JA, Skerlos L, Sulavik M, Thomas VH, VanderRoest S, Wang L, Wang Z, Whitton A, Zhu T, Stover CK (2009). *Proc Natl Acad Sci U S A* 106:1737–1742
56. Swamidass SJ (2011). *Brief Bioinform* 12:327–335
57. Ekins S, Mestres J, Testa B (2007). *Br J Pharmacol* 152:9–20
58. Schrödinger Release 2016-1: Jaguar (2016) version 9.1, Schrödinger, LLC, New York
59. Bochevarov AD, Harder E, Hughes TF, Greenwood JR, Braden DA, Philipp DM, Rinaldo D, Halls MD, Zhang J, Friesner RA (2013). *Int J Quantum Chem* 113:2110–2142
60. Gill PMW, Johnson BG, Pople JA, Frisch MJ (1992). *Chem Phys Lett* 197:499–505
61. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994). *J Phys Chem* 98:11623–11627
62. Lee C, Yang W, Parr RG (1988). *Phys Rev B* 37:785–789
63. Fleming I (2010) In: Fleming I (ed) *Molecular Orbitals and Organic Chemical Reactions* 1st edn. John Wiley & Sons, New York
64. Clare BW (1995). *J Mol Struct Theochem* 337:139–150
65. Zhou Z, Parr RG (1990). *J Am Chem Soc* 112:5720–5724
66. Kosar B, Albayrak C (2011). *Spectrochim Acta A* 78:160–167
67. Ayers PW, Parr RG (2000). *J Am Chem Soc* 122:2010–2018
68. Pearson RG (2005). *J Chem Sci* 117:369–377
69. Fukui K, Yonezawa T, Shingu H (1952). *J Chem Phys* 20:722–725
70. Murray JS, Abu-Awwad F, Politzer P (2000). *J Mol Struct Theochem* 501:241–250
71. Schrödinger Release 2016-1: Maestro v.10.5 (2016) Schrödinger, LLC, New York, NY. <https://www.schrodinger.com/citations#Maestro>
72. Dassault Systèmes BIOVIA (2015) *Discovery Studio Visualizer v4.5* Dassault Systèmes, San Diego, <https://www.3dsbiovia.com>. Accessed April 2019
73. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C, Wilson M (2018). *Nucleic Acids Res* 46:D1074–D1082
74. Parr RG, Chattaraj PK (1991). *J Am Chem Soc* 113:1854–1855
75. Parthasarathi R, Subramanian V, Roy DR, Chattaraj PK (2004). *Bioorg Med Chem* 12:5533–5543
76. Kobayashi S, Hamashima H, Kurihara M, Miyata N, Tanaka A (1998). *Chem Pharm Bull* 46:1108–1115
77. Kenny PW (2009). *J Chem Inf Model* 49:1234–1244

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.