Computational Chemistry

Rational Use of Heterogeneous Data in QSAR Modeling of Cyclooxygenase/Lipoxygenase Inhibitors

Alexey A. Lagunin, Athina Geronikaki, Phaedra Eleftheriou, Pavel Pogodin, and Alexey V. Zakharov

J. Chem. Inf. Model., **Just Accepted Manuscript** • Publication Date (Web): 28 Jan 2019

Rational Use of Heterogeneous Data in QSAR Modeling of Cyclooxygenase/Lipoxygenase Inhibitors

Alexey A. Lagunin[*1,2], Athina Geronikaki[3], Phaedra Eleftheriou[4], Pavel V. Pogodin[2], Alexey V. Zakharov[5]*

[1] Pirogov Russian National Research Medical University, Ostrovitianov str. 1, Moscow, 117997, Russia;

[2] Institute of Biomedical Chemistry, Pogodinskaya Str., 10/8, Moscow, 119121, Russia;

[3] School of Pharmacy, Aristotle University, Thessaloniki, 54124, Greece;

[4] School of Health and Medical Care, Alexander Technological Educational Institute of Thessaloniki, Thessaloniki, 57400, Greece;

[5] National Center for Advancing Translational Sciences (NCATS), National Institutes of Health, Rockville, MD 20850, USA

KEYWORDS. Heterogeneous data, QSAR, cyclooxygenase, lipoxygenase, thiazolidinone derivatives, IC_{50}.

ABSTRACT

Numerous studies have been published in recent years with acceptable quantitative QSAR modelling based on heterogeneous data. In many cases, the training sets for QSAR modelling were constructed from compounds tested by different biological assays, contradicting the opinion that QSAR modelling should be based on the data measured by a single protocol. We attempted to develop approaches that help to determine how heterogeneous data should be used for the creation of QSAR models on the basis of different sets of compounds tested by different experimental methods for the same target and the same end-point. To this end, more than one hundred QSAR models for the IC_{50} values of ligands interacting with cyclooxygenase 1,2 (COX) and seed lipoxygenase (LOX), obtained from ChEMBL database, were created using the GUSAR software. The QSAR models were tested on the external set, including 26 new thiazolidinone derivatives, which were experimentally tested for COX-1,2/LOX inhibition. The derivatives' IC_{50} values varied from 89 to 26 μM for LOX, from 200 to 0.018 μM for COX-1 and from 210 to 1 μM for COX-2. This study showed that the accuracy of the models depends on the distribution of IC_{50} values of low activity compounds in the training sets. In the most cases, QSAR models created based on the combined training sets had advantages in comparison with QSAR models based on a single publication. We introduced a new method of combination of quantitative data from different experimental studies based on the data of reference compounds, which was called "scaling".

INTRODUCTION

A significant amount of experimental quantitative data regarding the biological activity of chemical compounds has become available in recent years. These data have been collected using both freely (e.g. PubChem[1], ChEMBL[2], Binding DB[3]) and commercially (e.g. Wombat[4]) available databases. The availability of these data is very attractive and valuable for use in QSAR modelling for different tasks in medicinal chemistry (e.g., virtual screening, lead optimization, ADMET estimation). At the same time, it is known that the quality of the data and its correct use are the most important components of successful QSAR modelling. Several studies have elucidated the problems of data usage in QSAR modelling. Cronin and Schultz wrote that the "data should ideally be measured by a single protocol, ideally even by the same laboratory and by the same workers"[5] and "biological data should ideally be from well-standardized assays, with a clear and unambiguous endpoint"[5]. They consider receptor-binding assays in pharmaceutical research to be good examples of high quality biological data[5]. It is also known that IC_{50} and K_i are the most accurate values derived from experimental pharmacological data[6]. Therefore, the IC_{50} value is a reasonable endpoint for QSAR modelling. However, when we consider the available experimental data presented in scientific publications we can see that in many cases there are different assays for the same target. The represented data differ in the number of studied compounds, the distribution of IC_{50} values, and the maximal and minimal IC_{50} values. In addition, data for the same targets from different organisms are usually used in preclinical studies of drugs. This diversity of data raises questions related to the approaches for data selection for QSAR modelling. How can one use the experimental data from different studies with the same target and the same end-point? Should we select the data provided only by one study or from studies that use the same assay? May the data from several assays be

combined? Which is better: the creation of QSAR models based on a single study and then the use of them to create a consensus prediction or the combination of all the data into one general training set? The problems of data selection and representation for QSAR modelling have been previously analysed (e.g. OECD principles[7], Cronin and Schultz[5], Dearden et al.[8], Fourches et al.[9], Cherkasov et al.[10], Tarasova et al.[11]). However, these studies did not provide answers to many of these questions. We consider the rational selection of heterogeneous experimental data to be essential to QSAR development because many quantitative QSAR modelling studies based on heterogeneous data have been published in the recent years (Table 1).

The aim of this article is to study different approaches for preparation of training sets to answer the abovementioned questions based on the example of the creation of QSAR models for inhibitors of cyclooxygenase (COX) and lipoxygenase (LOX) using the experimental data available in ChEMBL. ChEMBL is one of the most well-known freely available web sources containing data on structures and experimentally determined biological activities (including interactions with targets) for more than 1.7 million compounds. We also introduced a new method for creation of combined training sets from different experimental studies based on the data of reference compounds, which was called “scaling”.

Table 1. Examples of recently published studies with quantitative QSAR modelling based on heterogeneous data of compounds interacting with drug targets.

Target	Unit	Method	N_{tr}	N_{test}	Descriptors	Characteristics	Ref.
HIV-1 reverse transcriptase	IC_{50}	SCR, RBF	22-319	30%	6 Phys-chem. and substructural descriptors	Different assays, median data.	[11]
11 COX from different species	IC_{50}	Gradient-boosting machines, RF, SVM	3228	1/6	797 PaDEL descriptors, Morgan fingerprints calculated by RDkit	Different assays, average data	[12]
Thyroid hormone receptor	IC_{50}	Linear regression	55-139	12.50%	Optimal descriptors on the basis of SMILES attributes	Same assay	[13]
DHFR rat, DHFR homo, F7, IL4, MMP2, CHRM3, NPY1R, NPY2R	IC_{50}	RF, GBM, SVM radial	344-779	1/6	Morgan fingerprints, phys-chem. properties calculated by RDkit	Different assays	[14]
Thyroid hormone receptor	IC_{50}	Random Forest	QSAR: 129-181; Docking: 15-210 positive examples	QSAR: 13-21; Docking: 12-13 positive examples	508 descriptors from Dragon	Different publications with the same assays, averaging data or manual curation [9].	[15]
ACE	IC_{50}	Generalized linear models with a Gamma distribution	245	18	38 phys-chem. descriptors	different assays and species	[16]
HDAC 1, 6	IC_{50}	kNN	95-108	15-18	Dragon	Different assays; normalization of IC_{50} values	[17]
P-gp	Eflux Ratio BA/AB, K_i	SCR, Bayesian approach	94-256	20%	6 phys-chem. and substructural descriptors	Different assays, averaging of IC_{50} values.	[18]

ACE, AChE, BDR; COX-2; DHFR, CB2R	IC_{50}, K_i	ANN	67-1361	10%	FP2 and ECFP6 (1024 bits), MACCS (256 bits) fingerprints	Different assays; averaging of IC_{50} values	[19]
AChE	IC_{50}	MIA-QSAR (Multivariate analysis of images)	26 scaffolds with derivatives	8 scaffolds with derivatives	pixels of 2D images corresponding to chemical structures	Different assays	[20]
PDE 4	IC_{50}	RF, ANN, kNN	812	203	More 1000 descriptors from Dragon	Different assays, exclusion of duplicate structures	[21]
HIV-1 protease	K_i	MLR	64	35	100 phys-chem. descriptors	Different assays	[22]
MHC I molecules	IC_{50}	STR, MLR	126	43	89 phys-chem. properties descriptors	Different assays	[23]
Kinases	IC_{50}	kernel-QSAR based on similarity of sequences and binding sites of kinase, GA, Tanimoto similarity	0-600	25%	FCFP_6 fingerprints	Different assays	[24]
18 antitargets	IC_{50}, K_i	SCR, Bayesian models	60-1366	15-344	6 phys-chem. and substructural descriptors	Different assays, averaging data	[25]
30 antitargets	IC_{50}, K_i	RBF, SCR, Bayesian models	100-4061	20%	6 phys-chem. and substructural descriptors	Different assays, median data	[26]

N_{tr} – number of compounds in the training set(s); N_{test} – number of compounds in the test set(s) ACE – angiotensin-converting enzyme; HDAC – histone deacetylase; AChE – acetylcholinesterase; BDR– benzodiazepine receptor; COX – cyclooxygenase; DHFR – dihydrofolate reductase; CB2R– cannabinoid receptor subtype 2; CHRM3 - human muscarinic acetylcholine receptor M3; NPY1R, NPY2R - human neuropeptide Y receptor type 1,2 PDE 4 – phosphodiesterase 4; MHC I – major histocompatibility class I; P-gp – P glycoprotein; F7 - human factor-7; IL4 - human interleukin-4; MMP2 - human matrix metallopeptidase-2; SVM – Support Vector Machines; RF – Random Forest; kNN – k Nearest Neighbours; ANN – Artificial Neural Network; MLR – Multiple linear regression; STR - stepwise regression; SCR – self-consistence regression; GA – genetic algorithm; RBF - Radial Basis Function; GBM - Gradient Boosting Machines. Dragon - software for the calculation of more 4000 molecular descriptors; FP2, MACCS and FCFP6 – different types of fingerprints.

Cyclooxygenase and lipoxygenase are well known drug targets for non-steroidal anti-inflammatory agents. Considerable experimental data are available about the interactions of drug-like compounds with these targets. The creation of drugs with simultaneous inhibition of cyclooxygenase and lipoxygenase is considered to be a prospective direction in the creation of new, effective and safe non-steroidal anti-inflammatory agents[27]. Compounds with combined COX-1/2 and LOX inhibition have multiple advantages as non-steroidal anti-inflammatory drugs because they act on the two major arachidonic acid metabolic pathways and possess a wide range of anti-inflammatory activities[28]. Since leukotrienes, produced by the action of lipoxygenase, play a role in blood coagulation and gastric tract irritation, LOX inhibition has a positive effect in anti-inflammatory action and prevents the GI tract irritation that results from COX-1 inhibition and the prothrombotic effect resulting from COX-2 inhibition. As a result, recent studies have focused on the development of dual-acting COX/LOX inhibitors[29].

We have experience in QSAR modelling using heterogeneous data[18,25,30,31] and in the creation of new dual cyclooxygenase/lipoxygenase inhibitors based on the analysis of structure-activity relationships[29,32]. This study is based on experimental data obtained during the creation of new inhibitors of cyclooxygenase and lipoxygenase and on IC_{50} values in the literature obtained from ChEMBL for ovine and human COX-1, human COX-2 and soya been lipoxygenase.

MATERIALS AND METHODS

ChEMBL

Data on the structures and IC_{50} values (concentration at half-maximum inhibition) of ligands which were experimentally tested for the inhibition of human and ovine COX-1, human

COX-2 and soya been lipoxygenase (Seed lipoxygenase-1) were retrieved from ChEMBL (version 17). To evaluate experimental errors in the data, we calculated the Standard Deviation (SD) of the IC_{50} values of all the compounds in initial ChEMBL data for which two or more IC_{50} values were given for the same ligand and the same drug target (COX-1, COX-2 and soya LOX) in different experiments. This calculation was made before any transformation of data described below.

Table 2. Standard Deviation (SD) of IC_{50} values of compounds tested on the studied targets that have at least two IC_{50} values.

Target	**Number of compounds**	**Number of IC_{50} values**	**Mean SD of log10(IC_{50}, nM)**
COX-1 human	206	632	0.377
COX-2 human	710	1856	0.403
LOX soya	20	48	0.344
COX-1 ovine	42	201	0.428

The mean SD values from Table 2 show the level of errors which may be achieved by the best QSAR models created based on the data for the studied targets.

The data on compounds tested on human COX-1 were used for creation of training and test sets for the investigation of different approaches for the use of heterogeneous data in QSAR modelling. During the selection of the data, the following limitations related to GUSAR were used:

1. The selected chemical compounds should be single organic electroneutral molecules, with molecular weights varying from 50 Da to 1250 Da;
2. The number of compounds with estimated IC_{50} values discussed in one article should be at least twenty.

Estimation of accuracy of prediction of QSAR models

R^2, Q^2 and RMSE (Root Mean Square Error) values were used for the estimation of the quality and accuracy of the prediction results given by the created QSAR models. R^2 (the coefficient of determination) was calculated for the data used as the training sets (R^2_{tr}) and the external validation test sets (R^2_{test}) using the following equation:

$$R^2 = 1 - \sum (y_{obs} - y_{pred})^2 / \sum (y_{obs} - y_{mean})^2 ,$$

where y_{obs} is an observed dependent variable, y_{pred} is a calculated dependent variable, and y_{mean} is the mean value of the dependent variable (calculated for the values of a training set for R^2_{tr} or for the values of a test set for R^2_{test}). The R^2 values represent the relationship between the predicted and the observed values of the measured biological activity[33] and R^2 values closer to 1 indicate successful predictions. Q^2 is a cross-validated R^2 calculated during the leave-one-out cross-validation procedure using the data from a training set.

RMSE was calculated using the following equation:

$$RMSE = \sqrt{\frac{\sum (y_{obs} - y_{pred})^2}{n}} ,$$

where n is the number of objects.

An RMSE closer to 0 is indicative of a successful prediction.

GUSAR software

All QSAR models were developed using the GUSAR software[25,31,34]. GUSAR uses a combination of three types of descriptors: whole-molecule descriptors, QNA (Quantitative Neighbourhoods of Atoms) descriptors[35], and descriptors based on the PASS (Prediction of

Activity Spectra for Substances) algorithm for predicting the biological activity spectra of compounds based on substructural atom centric MNA (Multilevel Neighbourhoods of Atoms) descriptors[30]. The whole-molecule descriptors used in GUSAR are the topological length, topological volume, lipophilicity, number of positive charges, number of negative charges, number of hydrogen bond acceptors, number of hydrogen bond donors, number of aromatic atoms, molecular weight, and number of halogen atoms. The QNA descriptors are defined by two functions, P and Q. The values for P and Q for each atom *i* are calculated as follows:

$$P_i = B_i \sum_k (Exp(-\frac{1}{2}C))_{ik} B_k, \quad (1)$$

$$Q_i = B_i \sum_k (Exp(-\frac{1}{2}C))_{ik} B_k A_k, \quad (2)$$

where *k* are all the other atoms in the molecule and

$$A_k = \tfrac{1}{2}(IP_k + EA_k), \quad B_k = (IP_k - EA_k)^{-\frac{1}{2}}. \quad (3)$$

In this equation, IP is the first ionization potential, EA is the electron affinity for each atom in electronvolts, and C is the connectivity matrix for the molecule[35]. The standard values of IP and EA of atoms in a molecule were collected from the literature. They have been published previously in the Appendix of one of our earlier publications[35]. Two-dimensional Chebyshev polynomials are used for approximating the functions P and Q over all the atoms in the molecule.

The PASS biological descriptors are calculated using the PASS algorithm[30], which predicts a wide range of biological outcomes including various mechanisms of action describing interaction of ligands with targets, transporter protein binding, gene expression activities, adding up to approximately 6400 "biological activities" at a mean prediction accuracy threshold of at least 95%. The output from PASS is the probability, for each predicted outcome, that the compound will be active (P_a) and the probability that it will be inactive (P_i). The difference

between these two values (P_a-P_i) was used as a molecular descriptor for the regression analysis in GUSAR. The results of PASS prediction for all 6400 "biological activities" are used as descriptors and the ones that show the highest correlation with the dependent variable (IC_{50}) are selected by self-consistent regression.

In GUSAR the scale of QNA and PASS based descriptors ranges from -1 to 1. So, no additional normalization is required for these types of descriptors. Only whole-molecule descriptors are normalized using a standard Z-score normalization procedure.

It has previously been shown that self-consistent regression (SCR) can be successfully applied to different QSAR tasks[25,31,34,36,37]. The basic purpose of the SCR method is to remove the variables that poorly describe the target value. The final number of variables in the QSAR equation selected after the SCR procedure is significantly smaller compared to the initial number of variables. In addition, it has been shown that SCR is robust against noise in the data[37]. The regression coefficients, obtained from SCR, reflect the contribution of each particular descriptor (variable) to the final equation. The higher the absolute value of the coefficient, the greater its contribution. Thus, regression coefficients obtained after SCR can be used for the weighting of descriptors (variables) according to their importance. In this approach, the descriptors are weighted during the calculation of the radial basis functions (RBF) by the coefficients obtained from SCR. RBF-SCR has previously been compared to different modern machine learning approaches using 14 data sets containing nine physicochemical properties and five toxicity endpoints[39]. It was shown that the RBF-SCR method provides more accurate prediction results than other methods, even including consensus predictions of these methods.

Applicability domain estimation

GUSAR simultaneously uses three different approaches for the estimation of the model applicability domains: similarity, leverage, and accuracy assessment:

- *Similarity*. For each compound, the pair-wise distance to each of its three nearest neighbours (3NN) in the training set is calculated using Pearson's correlation coefficient in the space of the independent variables obtained after SCR. The compound is considered to be in the applicability domain of the model if the average of these three distances is less than or equal to 0.7.
- *Leverage*. Leverage calculations are a method for identifying outliers based on the contribution of each molecule to its own predicted value: $Leverage = x^T (X^T X)^{-1} x$, where x is the vector of the descriptors for a test compound, and X is the matrix formed from the rows corresponding to the descriptors of all the molecules in the training set. A compound is considered to be outside the applicability domain of a model if its leverage is higher than the 99th percentile in the distribution of the leverage values calculated for the training set.
- *Accuracy assessment*. In this approach, the applicability domain prediction for each compound is calculated based on the Root Mean Square Error of the prediction for three nearest neighbours in the training set, $RMSE_{3NN}$, (see the Similarity metric above) in relation to the Root Mean Square Error of prediction for the training set as a whole ($RMSE_{train}$): $AD_{value} = RMSE_{3NN} / RMSE_{train}$. In this study, a threshold of 1 was used for the AD_{value}.

Y-scrambling procedure

GUSAR includes a Y-scrambling (Y-randomization) procedure. This procedure allows one to ensure that the developed models do not suffer from overfitting. In this procedure the dependent-variable vector, the Y vector (IC_{50} values), is randomly shuffled and a new QSAR model is developed using the original independent-variable matrix. It is expected that the resulting models should generally have low Q^2 values. This procedure was repeated five times for each model, and then the average Q^2 value calculated.

Consensus modelling

The final predicted values for endpoints were calculated using a weighted average of the predictions from several selected QSAR models. Each model is based on a different set of QNA and PASS "biological" descriptors, and its predictions for each compound are weighted according to the similarity value as calculated during the applicability domain assessment. For each training set, the QSAR models were selected from 1600 QSAR models by internal validation using the 5-fold cross-validation procedure in which 20% of the training set was randomly excluded, and these data were used as an internal test set. Only models which satisfied the criterion of R^2 >0.5, Q^2 >0.5 and R^2 of internal cross-validation > 0.5 were selected for consensus modelling.

Chemistry

General Procedure for the Synthesis of 2-Thiazolylimino-5-arylidene-4-thiazolidinones

The data for compounds 1-12 were reported in our previous paper[40]. The appropriate arylaldehyde (6 mM) was added to a well-stirred solution of 2-(thiazol-2-ylimino)thiazolidin-4-

one (0.8 g, 4 mM) in acetic acid (35 mL) buffered with sodium acetate (8 mM). The solution was refluxed for 4 h and then poured into ice-cold water. The precipitate was filtered and washed with water and the resulting crude product was purified by recrystallization from dioxane.

Biological Assays

In the *in vitro* assays, each experiment was performed at least in triplicate and the standard deviation of the absorbance was less than 10% of the average values.

Soybean Lipoxygenase Inhibition Study in Vitro. Lipoxygenase inhibition was evaluated using soybean LOX type 1b as reported previously[32,41,42]. sLOX-1 is the soybean isoenzyme mostly used in drug screening, although sLOX-3 has also been used[42]. Because of the structural and functional similarities between mammalian LOXs, sLOX is commonly used for both mechanistic and inhibition studies and is widely accepted as a model for LOXs from other sources[44]. It has been shown that the inhibition of plant LOX activity by NSAIDs is qualitatively similar to the inhibition of rat mast cell LOX and can be used as a simple screen for this activity[45]. Similarity in the inhibition behaviour between soybean LOX-1 and human recombinant 5-LOX or soybean LOX-3 and human blood serum LOX has been observed in many cases[43,44]. The tested compounds, dissolved in DMSO, were added to the reaction mixture at a final concentration of 100 μM and were preincubated for 4 min at 28 °C with soybean lipoxygenase at a concentration of $7x10^{-7}$ w/v. The enzymatic reaction was initiated by the addition of sodium linoleate to a final concentration of 0.1 mM. The conversion of sodium linoleate to 13-hydroperoxylinoleic acid was measured at 234 nm. Nordihydroguaretic acid, an appropriate standard inhibitor, was used as the positive control (94.4% inhibition at 0.1 mM, IC_{50} = 31.3 μM).

COX Inhibitor Screening Assay. The COX-1 and COX-2 activities of the compounds were measured using ovine COX-1 and human recombinant COX-2 enzymes included in the "COX

Inhibitor Screening Assay" kit provided by Cayman (Cayman Chemical Co., Ann Arbor, MI). The assay directly measures PGF2a produced by the $SnCl_2$ reduction of COX-derived PGH2[29,32]. The prostanoid production was quantified via enzyme immunoassay using a broadly specific antibody that binds to all the major prostaglandin compounds. The final estimation of % inhibition was performed at a substrate concentration much lower than the saturating concentration. For better visualization of the differences between compounds on a 0–100% inhibition scale, the COX-1 inhibitory activity was tested at an arachidonic acid concentration of 1 μM and the COX-2 inhibitory activity was tested at an arachidonic acid concentration of 0.1 μM. The compounds were added to the reaction mixture at a final concentration of 200 μM. IC_{50} values were calculated for the most active compounds. Naproxen and indomethacin, used as positive controls, were added to the reaction mixture at the same concentration, 200 μM, as the tested compounds.

RESULTS

QSAR modelling for compounds tested as inhibitors of human COX-1

Fourteen sets of compounds tested for the inhibition of human COX-1 were selected from ChEMBL according to the abovementioned rules. The sets were characterized according to the source of COX-1 isoenzyme, the products of the reaction determined to measure the enzyme activity/inhibition and the subsequent calculation of the IC_{50} values of the tested compounds, the number of compounds in the set, and the minimal and maximal determined IC_{50} values (Table 3).

Table 3. Datasets of compounds tested as inhibitors of human COX-1.

No	Determined Product	Enzyme Source	N_{all}	N_{sel}	IC_{50min} (nM)	IC_{50max} (nM)	$\log(IC_{50max}) - \log(IC_{50min})$	Ref.
1	-	U-937 cells	88	21	2	100000	4.7	[46]

2	PGE2	hCOX-1	32	27	2	136000	4.8	[47]
3	PGE2	hCOX-1	50	20	100	265000	3.4	[48]
4*	PGE2	hCOX-1	109	66	5	674000	5.1	[49]
5	PGE2	U937 microsomes	24	20	50	34000	2.8	[50]
6	PGE2	hCOX-1	35	22	5	18300	3.6	[51]
7	PGE2	hCOX-1	96	43	100	723000	3.9	[52]
8	PGH2	human platelet-derived COX-1	25	23	20	30000	3.2	[53]
9	TxB2	blood	34	34	4120	383360	2	[54]
10	TxB2	blood	28	25	1300	800000	2.8	[55]
11	TxB2	blood	31	21	500	15900	1.5	[56]
12	TXB2	blood	39	39	80	60390	2.9	[57]
13	TXB2	blood	36	31	250	49300	2.3	[58]
14	TxB2	blood	29	25	140	25610	2.3	[59]
Total			**690**	**451**				

* - Validation test set; $\mathbf{N_{all}}$ – number of all compounds studied in the publication; $\mathbf{N_{sel}}$ - number of compounds with determined IC_{50} values that were selected for QSAR modelling.

As shown in Table 3, according to the data provided by the 14 selected publications, IC_{50} values were calculated and recorded only for 451 out of the 690 studied compounds that were referred to in the publications. Thus, approximately 35% of the published experimental data could not be used for the creation of quantitative (continuous) QSAR models because of the absence of accurate IC_{50} values. The difference between the log10 values of the minimum and maximum IC_{50} values exceeds 1.5 for all the datasets, which is one of the formal necessary conditions for the creation of reasonable QSAR models[8]. Moreover, the minimum IC_{50} values are less than 10^{-5} M and the maximum IC_{50} values are higher than 10^{-5} M for all the datasets. This is a traditional border between active and inactive compounds used in medicinal chemistry. Therefore, we may initially expect that these data will allow the creation of helpful QSAR models for further studies. Most of the datasets were tested using two types of assays: (1) estimation of the PGE2 production of human recombinant COX-1 and (2) estimation of the TxB2 level using a human whole blood assay. Thus, along with different combination

approaches using the experimental data, we will evaluate the combination of data from sets using different types of assays.

The dataset with the largest number of compounds (dataset #4 in Table 3) was selected as an external validation set for the further estimation of the prediction accuracy of the created QSAR models for the prediction of the IC_{50} values of compounds inhibiting human COX-1[49]. The largest test set provided the possibility to obtain the most statistically significant results of estimation of accuracy for the created QSAR models.

Evaluation of single QSAR models on the basis of individual datasets

We started with the creation of single QSAR models on the basis of the individual datasets mentioned in Table 3. Acceptable QSAR models ($Q^2 > 0.5$ and R^2 of internal cross-validation > 0.5) were obtained for only 5 of the 14 datasets (Table 4). In the second step, a prediction of IC_{50} values was made for the data from the test set and the accuracy of these predictions (R^2_{test} and $RMSE_{test}$) was calculated (Table 4). Table 4 shows that in spite of the acceptable values of R^2_{tr}, Q^2, and $RMSE_{tr}$ (obtained from the internal validation of the models and comparable with the SD of the IC_{50} value (0.377) from Table 2) the values of the prediction accuracy calculated for the test set were unsatisfactory, since low R^2_{test} and high $RMSE_{test}$ values were observed.

Table 4. Characteristics of single QSAR models.

№	Ref	N	R^2_{tr}	Q^2	$RMSE_{tr}$	R^2_{test}	$RMSE_{test}$	AD,%
QSAR models with Q^2 and R^2 of internal cross-validation > 0.5								
10	[55]	25	0.903	0.647	0.459	-0.553	1.318	100
3	[48]	20	0.989	0.721	0.407	-1.958	2.188	3
8	[53]	23	0.967	0.713	0.504	0.097	1.078	89
12	[57]	39	0.796	0.511	0.402	-0.617	1.06	100
14	[59]	25	0.969	0.722	0.339	0.041	1.05	67
QSAR models with Q^2 and R^2 of internal cross-validation < 0.5								

1	[46]	21	1.000	0.465	3.167	0.080	1.014	100
2	[47]	27	0.989	0.498	0.828	-0.595	1.335	100
4	[49] test	66	0.956	0.391	0.826	NA	NA	NA
5	[50]	20	1.000	0.006	1.045	-1.649	1.721	100
6	[51]	22	0.995	0.066	0.754	-0.535	1.310	100
7	[52]	43	0.980	0.001	1.105	-0.161	1.139	100
9	[54]	34	0.874	0.353	1.184	-0.926	1.467	100
11	[56]	21	0.998	0.181	0.476	-0.115	1.116	100
13	[58]	31	0.989	0.007	0.579	-0.142	1.130	100

NA – Not Available; № – indicates the number of a set in Table 3.

Creation of general training sets

In the next step of the study, the creation of the general training sets was performed using different criteria. They were created by combining single datasets and included the following:

1. A general training set, which was created by the combination of data from all thirteen datasets in Table 3 (GeneralTr);
2. A training set that was created by combining the datasets mentioned in Table 4, for which acceptable QSAR models were created (AcceptTr);
3. A training set that was created by combining the data from the datasets in Table 3 with the IC_{50} values calculated from the TxB2 blood level determination (TxB2Tr);
4. A training set that was created by combining the data from the datasets with IC_{50} values calculated from the TxB2 blood level determination, derived from the accepted QSAR models mentioned in Table 4 (AcceptTxB2Tr);
5. A training set that was created by combining the data from the datasets in Table 3 with the IC_{50} values calculated from the determination of the PGE2 levels produced by the action of hCOX-1 *in vitro* (PGE2Tr).

A prediction of IC_{50} values was made for the external test set (Test set #4 from Table 3) by QSAR modelling after the training of GUSAR based on the combined training sets (Table 5). The combined training sets did not overlap with the test set.

Table 5. Characteristics of QSAR models created by combining datasets.

№	Training set	N	R^2_{tr}	Q^2	$RMSE_{tr}$	R^2_{test}	$RMSE_{test}$
1	GeneralTr	348	0.883	0.594	0.698	0.186	1.137
2	AcceptTr	132	0.942	0.683	0.577	0.270	0.952
3	TxB2Tr	175	0.932	0.774	0.446	0.116	1.005
4	AcceptTxB2Tr	89	0.944	0.796	0.473	0.121	0.997
5	PGE2Tr	160	*There was no acceptable QSAR model*				
6	Consensus of prediction results given by the single models from Table 4					0.067	1.021

N – number of structures in a training set.

Table 5 shows that GUSAR could not create any acceptable QSAR model based on the PGE2Tr training set. The best prediction results was derived from a QSAR model created using the AcceptTr training set, which contained the datasets of the QSAR single models mentioned in Table 4. The values of R^2_{test} and $RMSE_{test}$ calculated from the QSAR models created using AcceptTr, TxB2Tr and AcceptTxB2Tr were higher than the values for any single QSAR model in Table 4 or their simple average consensus (the last row in Table 5). All the compounds from the test set were in the applicability domain of the QSAR models based on the AcceptTr, TxB2Tr and AcceptTxB2Tr training sets and 85% of the compounds were for GeneralTr.

The results shown in Table 5 were derived from the combined datasets. This simple combination of datasets leads to the incorporation of duplicates of structures with different IC_{50} values. According to the general rules for QSAR modelling, duplicates in training sets should be avoided[9,10], and manual checking for duplicates and the deletion of error data is recommended[9].

If both experimental properties are highly similar, maintaining the records associated with the structure in the arithmetic average of properties is proposed. If the values are significantly different, the elimination of both records is recommended[9]. This is easily applied when analysing the data from a single laboratory, measured using the same experimental protocol. However, when one attempts to use as much data as possible to create a general training set from all the available studies, it is difficult to recognize which experimental value among dozens is correct, taking into account the special features of the experimental assays and the inevitable errors in publications and databases. When we analysed duplicates in the combined datasets, six compounds with duplicates were found, all of which corresponded to reference compounds. The IC_{50} values of two reference compounds for the studied datasets are presented in Table 6.

Considering Table 6, it seems that all these data cannot be combined in the same training set. The differences are the result of the use of different methods or different conditions when the same method was used. This is one reason why the results obtained for GeneralTr were worst in comparison with the other combined training sets. If one does not aim at analyzing the dependency of the models on a specific parameter and includes limitations based on the value of this parameter (such as the substrate concentration) in the selection of compounds for the training set, this table may be used to choose which data can be combined. No one would feel safe combining data when the values for the reference compound differ by an order of magnitude and taking the average is not appropriate in this case. According to the table, it seems that the data from sets 4, 6 and 2 may be combined and used for the prediction of IC_{50} values because they were experimentally determined by the same method under the same conditions. This is the AcceptTxB2Tr training set, which yields one of the better results in Table 5. Combining the data

from sets 7, 8, 9 and 10 seems to be also acceptable. This is the PGE2Tr training set in Table 5. However, in this case, an acceptable QSAR model could not be created.

Table 6. IC_{50} values of reference compounds in the studied datasets.

	Ref	Celecoxib, nM	Indometacin, nM	Scaling ratio	Product	Source
1	[54]	23470		0.639	TxB2	blood
2	[55]	7000		2.143	TxB2	blood
3	[56]	14000	500	1.071	TxB2	blood
4	[57]	2600		5.769	TxB2	blood
5	[58]	14200	250	1.056	TxB2	blood
6	[59]	2600		5.769	TxB2	blood
7	[51]		100**		PGE2	hCOX-1
8	[48]		100		PGE2	hCOX-1
9	[62]		100		PGE2	hCOX-1
10	[49]*	15000	100**	1	PGE2	hCOX-1
11	[50]	50		300	PGE2	U937 microsomes
12	[46]	5100	2	2.941		U-937 cells
	Average	**9336**	**164**			
	Median	**7000**	**100**			

* - Validation set; ** IC_{50} value using literature data for the same assay; Scaling ratio – a value used for the multiplication of IC_{50} values to scale appropriate data to the IC_{50} value of Celecoxib (15000 nM – IC_{50} value of Celecoxib in the test set); see below in the text for detailed explanation.

Table 6 shows that the experimental IC_{50} values of Celecoxib show a spread from 50 to 23470 nM. The minimum IC_{50} value for Celecoxib was determined by the measurement of the PGE2 level produced by the enzyme derived from U937 microsomes. As mentioned above, the use of the average value of duplicates is one of the approaches used to avoid duplicates. We retrained GUSAR using the average value of duplicates in the combined datasets from Table 5. Moreover, general training sets were also created from all the data (excluding the data from the publication used for the test set) for compounds tested as inhibitors of human COX-1 from ChEMBL (version 17) with average ($CHEMBL_{av}$) or median ($CHEMBL_{med}$) IC_{50} values for

duplicates. The results of the training and the prediction of IC_{50} values for the validation set are shown in Table 7.

Table 7. Characteristics of QSAR models created by combining human COX-1 datasets using average values for reference compounds.

№	Training set	N	R^2_{tr}	Q^2	$RMSE_{tr}$	R^2_{test}	$RMSE_{test}$
1	GeneralTr	333	0.932	0.618	0.600	0.187	1.038
2	AcceptTr	130	0.946	0.714	0.551	0.300	0.884
3	TxB2Tr	169	0.942	0.769	0.456	0.023	1.006
4	AcceptTxB2Tr	87	0.945	0.797	0.477	0.017	1.098
5	PGE2Tr	157	*There was no acceptable QSAR model*				
6	$CHEMBL_{av}$	1137	0.939	0.538	0.653	0.112	0.901 (AD 97%)
7	$CHEMBL_{med}$	1137	0.939	0.542	0.656	0.122	0.895 (AD 97%)
8	Consensus of prediction results given by the single models from Table 3					0.067	1.021

N – number of structures in a training set. $CHEMBL_{av}$ – dataset from all the compounds in ChEMBL tested for human COX-1 inhibition, with addition of average IC_{50} values for the compounds having double/multiple records in ChEMBL; $CHEMBL_{med}$ – dataset from all the compounds in ChEMBL tested for human COX-1 inhibition, with median IC_{50} values for the compounds having double/multiple records in ChEMBL.

Table 7 shows that QSAR models based on the GeneralTr and AcceptTr training sets were improved by the averaging of the data. There were no changes in the accuracy of the prediction for QSAR models based on TxB2Tr and PGE2Tr. The accuracy of prediction for QSAR models based on the AcceptTxB2Tr training set was decreased. Interestingly, the accuracy of prediction based on the $CHEMBL_{av}$ and $CHEMBL_{med}$ training sets was better than that of the other models. The $CHEMBL_{av}$ training set includes average IC_{50} values for compounds having double/multiple records in ChEMBL. The $CHEMBL_{med}$ training set includes median IC_{50} values for the compounds having double/multiple records in ChEMBL. The general model based on the results for all the recorded compounds using the median data for multiply recorded molecules achieved the highest accuracy. The use of median values in training sets 1-5 (Table 7) did not lead to a

change in the prediction results. This may be explained by the small number of duplicates in these sets and the small difference between the given average and median values.

We introduce another approach for the combination of data from different sets – scaling, which is based on analysis of experimental data of reference compounds. If we consider the IC_{50} values for reference compounds (Table 6), we can see that they show considerable differences. Figure 1A shows the distribution of IC_{50} values in the used sets of compounds tested for the inhibition of human COX-1. The tenth set in Figure 1 is the test set. The black squares are the IC_{50} values of Celecoxib, which is a reference compound in most studies, including the test set. The idea of scaling is that the IC_{50} values of compounds in the sets are recalculated so as the IC_{50} of the reference compound becomes equal in all the sets (Figure 1B). The datasets 7, 8, and 9 were not scaled because they didn't contain the reference compound Celecoxib.

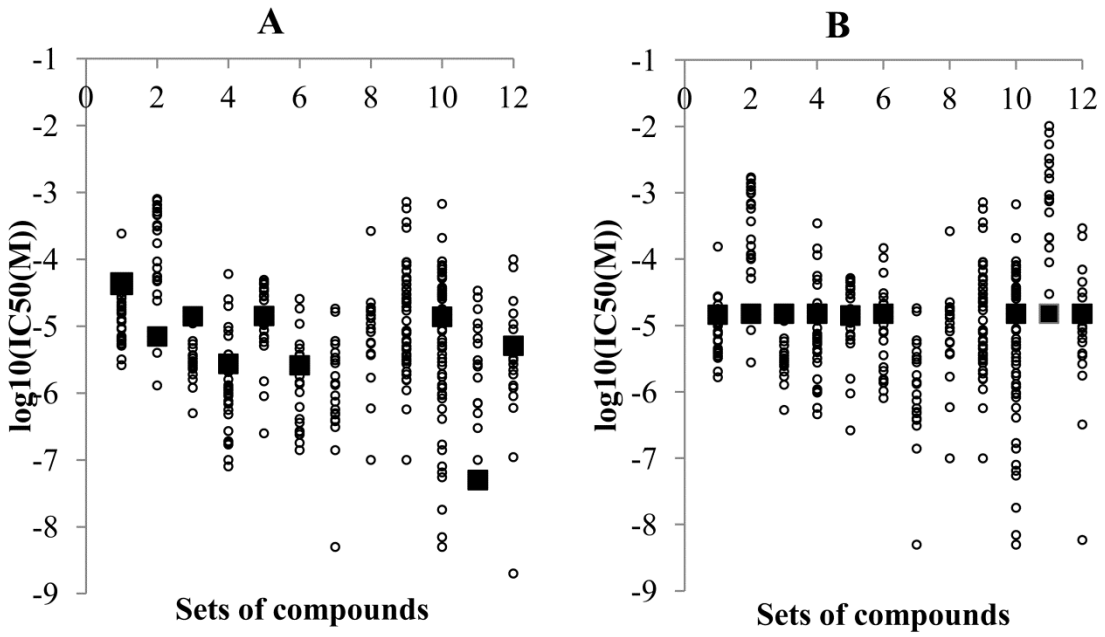


Figure 1. Distribution of IC_{50} values for compounds tested for the inhibition of human COX-1. Black squares – IC_{50} value of Celecoxib. (A) Distribution of IC_{50} values based on the original data. (B) Distribution of IC_{50} values after the scaling procedure.

The scaling was accomplished by the multiplication of the IC_{50} values by an appropriate scaling ratio (see the column "Scaling ratio" in Table 6). The scaling procedure leads to a change in the general distribution of IC_{50} values in a combined training set. An example of this change is shown in Figure 2.

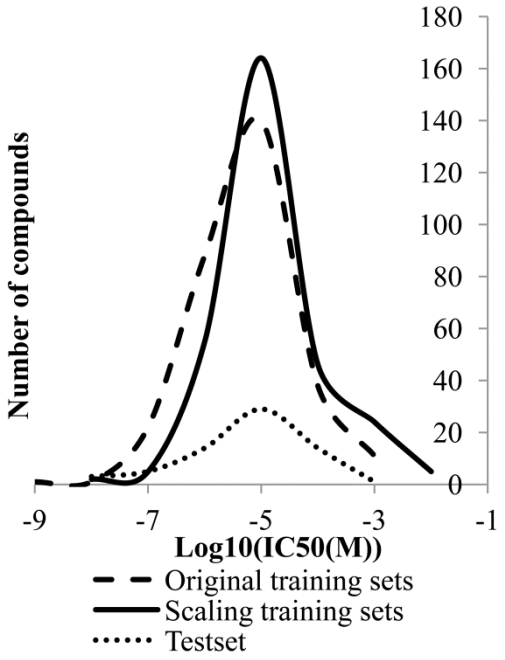


Figure 2. General distribution of IC_{50} values in the sets.

One can see that after the scaling procedure the number of IC_{50} values that are close to the average value increased (Figure 2). At the same time, the distribution of IC_{50} values in the test set (tenth set in Figure 1) was shifted towards the less active compounds. The scaling procedure was applied to the combined training sets mentioned in Table 6. Subsequently, new QSAR models were created and tested on the external validation set (Table 8).

Comparison of the R^2_{test} and RMSE values in Table 7 and Table 8 shows that for some training sets the accuracy of the QSAR models increased after the scaling procedure (GeneralTr,

AcceptTr, PGE2Tr), but it decreased for other training sets (TxB2Tr, AcceptTxB2Tr). In most cases, the accuracy of the models trained on scaled data is higher than accuracy of the prediction results given by the consensus of single models from Table 3.

Table 8. Characteristics of the QSAR models created using the combined training sets after data scaling to the reference compound (Celecoxib, IC_{50} = 15000 nM).

№	Training set	N	R^2_{tr}	Q^2	$Q^2_{Y\text{-}scr}$	$RMSE_{tr}$	R^2_{test}	$RMSE_{test}$	Sim
1	GeneralTr	333	0.680	0.609	-0.257	0.633	0.273	0.890	0.337
2	AcceptTr	130	0.943	0.714	-0.316	0.546	0.308	0.878	0.299
3	TxB2Tr	169	0.913	0.671	0.033	0.448	-0.046	1.081	0.332
4	AcceptTxB2Tr	87	0.928	0.723	0.048	0.482	0.082	1.300	0.511
5	PGE2Tr	157	0.776	0.509	0.028	0.685	0.074	1.017	0.361
6	Consensus of prediction results given by the single models from Table 3						0.067	1.021	

N – number of structures in a training set; $Q^2_{Y\text{-}scr}$ – average Q^2 value calculated based on internal validation. Sim – average value of pair wise similarity (Tanimoto coefficient) between structures from the appropriate training set calculated based on MNA descriptors.

Table 8 shows that $Q^2_{Y\text{-}scr}$ values for all QSAR models on scaling data were less than 0.05. This is significant less in comparison with Q^2 values calculated based on the original data of the training sets and indicates robustness of the models. The $Q^2_{Y\text{-}scr}$ values were better than, or comparable to, R^2_{test} values calculated for the test set by the QSAR models based on individual-dataset in Table 4.

To precisely evaluate the developed models that were based on the different designs (individual, average, scaling) we performed a bootstrap analysis to determine if there were statistically significant differences between the models:

- Each model was used to make a prediction for the same test set (dataset #4 in Table 3). For this comparison to be fair, only results for chemical structures falling in the applicability domain of the predictive model were used.
- The results of each model's prediction were bootstrapped 10 000 times using the R-package called "resample" (https://CRAN.R-project.org/package=resample). The RMSE value was calculated in each bootstrap run. Thus, many instances of RMSE (https://CRAN.R-project.org/package=MLmetrics) were generated for each model.
- RMSE values were used to conduct ANOVA in conjunction with Tukey's HSD (honest significance difference) test.

The results of this procedure indicated that models built using training sets of different design, differed significantly ($P_{value} < 0.05$) from each other in terms of their quality as assessed by RMSE (Figure 3). The illustration was prepared using ggplot2 package (https://cran.r-project.org/web/packages/ggplot2/index.html). QSAR models based on the scaling procedure showed the most accurate results.

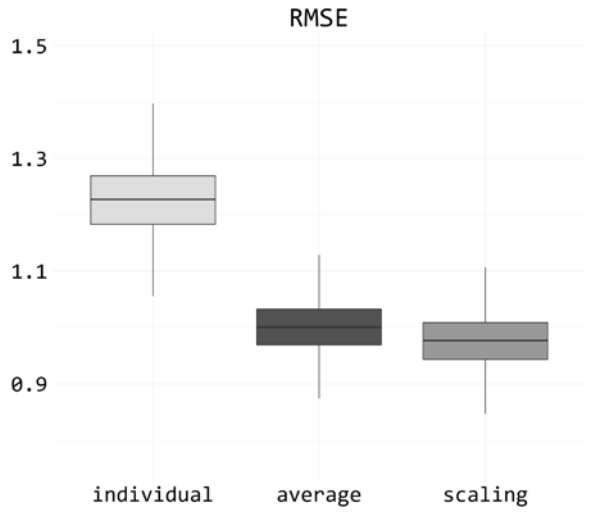


Figure 3. Comparison of RMSE values given by the bootstrap analysis based on the prediction results for the external test set (dataset #4 in Table 3). Individual means RMSE of the

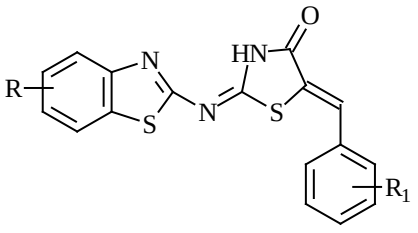
single models from Table 3. Average means RMSE of QSAR models (1-4 models from Table 7). Scaling means RMSE of QSAR models (1-5 models from Table 8).

These approaches to using heterogeneous data for QSAR modelling based on published experimental results were reproduced on ligand interactions with ovine COX-1, human COX-2 and soybean LOX (from ChEMBL using the same limitations as above for the available data on the inhibitors of human COX-1). Novel experimental data concerning new cyclooxygenase/lipoxygenase inhibitors synthesized at the School of Pharmacy of Aristotle University of Thessaloniki and biologically evaluated *in vitro* were used as an external validation set (Table 9) for the created QSAR models.

In Vitro COXs and LOX Inhibition Studies

The influence of the nature and the position of substitution of new thiazolidinone derivatives on the cyclooxygenase and lipoxygenase inhibitory activities were evaluated experimentally (Table 9).

Table 9. Experimental IC_{50} values for COX-1, COX-2 and LOX for the studied thiazolidinone derivatives.



compd	R	R_1	IC_{50}, μM		
			LOX	COX-1,	COX-2,

				(% of inhibition at 10 μM)	(% of inhibition at 10 μM)
1	H	H	89.1	-	- (9.4%)
2	H	2-NO_2	35.5	>200	- (9.4%)
3	H	3-NO_2	42.0	0.3	- (31.7%)
4	H	4-NO_2	50.1	0.5	- (7.9%)
5	H	2-Cl	71.0	0.018	- (58.8%)
6	H	3-Cl	35.5	22.4	- (32.2%)
7	H	4-Cl	50.0	0.3	- (20.3%)
8	H	3-OH	34.7	5.6	- (1.0%)
9	H	4-OH	36.3	7.9	- (20.6%)
10	H	4-OCH_3	28.2	200	- (26.4%)
11	H	3-OMe, 4-OH	35.5	39.8	- (6.6%)
12	H	3,5-OMe, 4-OH	47.9	200	-
13	6- NO_2	H	29.5	100 (61%)	63 (86%)
14	6- NO_2	3- NO_2	81.3	>200	106
15	6- NO_2	4- NO_2	63.2	79	>200
16	6- NO_2	2-Cl	72.4	86	100 (12.7%)
17	6- NO_2	4-Cl	34.7	200 (32%)	106 (61%)
18	6- NO_2	4-OMe	63.1	60 (125%)	200 (53%)
19	6- NO_2	3-OMe, 4-OH	26.3	>200	85 (71%)
20	6- NO_2	3,5-OMe, 4-OH	33.9	74 (95%)	>200 (25%)
21	6-OMe	4- NO_2	33.9	74	>200
22	6-OMe	3-Cl	35.9	200.5	- (26.4%)
23	6-OMe	3-OMe, 4-OH	42.7	>200	- (57%)
24	4-CH_3	H	63.1	4	- (6.7%)
25	4-CH_3	4-NO_2	38.0	190	210
26	4-CH_3	4-Cl	47.9	- (58%)	67 (86%)
Nordihydroguaiaretic acid			31.3	-	-
Celecoxib			-	34.5	0.1

Creation of QSAR models for ligands interacting with ovine COX-1, human COX-2 and soybean LOX and their validation by external test sets with novel cyclooxygenase/lipoxygenase inhibitors

QSAR modelling was performed based on 5 training sets for ovine COX-1 (Table 10), 44 training sets for human COX-2 (Table 11) and 3 training sets for soybean LOX (Table 12). The created QSAR models had not been used to determine which compounds from Table 9 should be

synthesized and experimentally studied as new cyclooxygenase/lipoxygenase inhibitors. The data from Table 9 were used for external validation of the created QSAR models. The borders of the IC_{50} values in Tables 10-12 are given without a scaling correction.

Table 10. Datasets of compounds tested as inhibitors of ovine COX-1.

No	Product	Source	N_{all}	N_{sel}	IC_{50min} (nM)	IC_{50max} (nM)	log10(IC_{50max} - IC_{50min})	Ref.
1	PGF2alpha	Cayman kit	20	20	3200	231200	1.9	[60]
2	PGF2alpha	Cayman kit	28	28	411	366000	2.9	[61]
3	NA	seminal vesicles	49	27	5000	390000	1.9	[62]
4	NA	seminal vesicles	22	22	4620	1150000	2.4	[63]
5	NA	hematin-reconstituted COX-1	25	22	50	17000	2.5	[64]

N_{all} – number of all compounds studied in a publication; **N_{sel}** - number of compounds with IC_{50} values that were selected for QSAR modelling. **NA** – Not Available.

Table 11. Datasets of compounds tested as inhibitors of human COX-2.

No	Product	Source	N_{all}	N_{sel}	IC_{50min} (nM)	IC_{50max} (nM)	log10(IC_{50max} - IC_{50min})	Ref.
1	PGE_2	hCOX-2	32	24	5.92	10000	3.2	[65]
2	PGE_2	blood	90	70	0.3	5720	4.3	[66]
3	PGE2	blood	34	34	300	34610	2.1	[54]
4	PGE2	blood	27	23	190	540000	3.5	[55]
5	PGE2	hCOX-2	25	25	120	19400	2.2	[66]
6	PGE2	hCOX-2	21	21	100	19400	2.3	[67]
7	PGE2	hCOX-2	25	21	3000	200000	1.8	[68]
8	PGE2	hCOX-2	29	26	6	1600	2.4	[69]
9	PGE2	CHO cell	21	21	2	2900	3.2	[70]
10	PGE-2	ECV-304 cells	30	22	10	10000	3	[71]
11	NA	COX-2	25	22	40	66000	3.2	[64]
12	Arachidonic acid	CHO cells	23	20	2	1900	3	[72]
13	NA	COX-2	32	25	20	2240	2	[73]
14	NA	blood	32	24	380	11300	1.5	[73]
15	PGE2	COS cells	24	21	0.44	388	2.9	[74]
16	NA	blood	53	53	60	7048553	5.1	[75]

17	PGE2	COS cells	31	21	0.25	1683	3.8	[76]
18	NA	blood	40	34	5	100000	4.3	[47]
19	PGE2	hCOX-2	39	39	100	23160	2.4	[77]
20	NA	hCOX-2	20	20	3	77900	4.4	[78]
21	PGE2	hCOX-2	25	23	8000	18000000	3.4	[79]
22	PGE2	hCOX-2	35	33	0.5	2900	3.8	[80]
23	PGE2	osteosarcoma cell line 143.98.2	36	20	10	1000	2	[81]
24	PGE2	hCOX-2	35	34	1	15000	4.2	[51]
25	PGE2	hCOX-2	50	45	2	52300	4.4	[48]
26	PGE2	hCOX-2	109	93	1.7	100000	4.7	[49]
27	PGE2	hCOX-2	42	27	14	100000	3.9	[82]
28	PGE2	hCOX-2	96	86	3	100000	4.5	[52]
29	PGF2alpha	hCOX-2	25	22	3.5	30000	3.9	[52]
30	RPP	hCOX-2	49	30	800	510000	2.8	[62]
31	NA	hCOX-2	22	22	190	81000	2.6	[83]
32	NA	hCOX-2	48	29	47	100000	3.3	[84]
33	NA	CHO cells	25	25	2	1530	2.9	[50]
34	NA	blood	25	25	80	9800	2.1	[50]
35	PGE2	blood	36	35	210	59000	2.4	[58]
36	RPP	hCOX-2	74	66	40	25000	2.8	[85]
37	PGE2	hCOX-2	52	49	26	934000	4.6	[86]
38	PGE2	hCOX-2	305	305	1	9500	4	[87]
39	NA	143982 cells	88	59	2	10000	3.7	[46]
40	PGE2	blood	47	43	60	10000	2.2	[88]
41	PGE2	CHO cells	21	20	36	960	1.4	[89]
42	PGE2	blood	29	21	90	47620	2.7	[59]
43	TXB2	blood	23	20	700	26300	1.6	[56]
44	PGE2	blood	39	34	90	21900	2.4	[57]

N_{all} – number of all compounds studied in a publication; **N_{sel}** - number of compounds with IC_{50} values that were selected for QSAR modelling. **NA** – Not Available. RPP - Radiolabeled prostanoid products.

Table 12. Datasets of compounds tested as inhibitors of soybean LOX.

No	Description of assay	Reference compounds (IC_{50}, nM)	Scaling ratio	N_{all}	N_{sel}	IC_{50min} (nM)	IC_{50max} (nM)	log10(IC_{50max} - IC_{50min})	Ref.

1	In vitro inhibitory activity against 15-lipoxygenase obtained from soybean	Nordihydroguaiaretic acid (3500) Luteolol (3200)	9	15	10	100	3500	1.7	[90]
2	Inhibitory concentration against soybean lipoxygenase upon incubation with sodium linoleate (0.1 mM) at RT	NA		13	10	13000	100000	1	[91]
3	Inhibitory activity against soybean 15 LOX	Nordihydroguaiaretic acid (3500) Luteolol (3200)	9	28	16	100	5000	1.7	[92]

N_{all} – number of all compounds studied in a publication; **N_{sel}** - number of compounds with IC_{50} values that were selected for QSAR modelling. **NA** – Not Available. **Scaling ratio** – coefficient of multiplication of IC_{50} values for compounds from the appropriate set. The average value of nordihydroguaiaretic acid (31300 nM) was used for scaling.

A considerable amount of the data from publications concerning COX-1, COX-2 or LOX inhibition could not be used for the creation of QSAR models because of the absence of accurate IC_{50} values (17% for ovine COX-1, 13% for human COX-2 and 36% for soybean LOX data).

The following acceptable QSAR models were created based on the above-mentioned data:

1. Ovine COX-1: 4 QSAR models (Table 13);
2. Human COX-2: 37 QSAR models (Table S1, Supplements and Table 14 (selected 8));
3. Soybean LOX: 1 QSAR model (Table 15).

All these models were tested using the external test sets from compounds tested as novel ovine COX-1, human COX-2 and soybean LOX inhibitors (Table 9). The RMSE values were used for the validation and comparison of the created QSAR models because of the small number of compounds with exact IC_{50} values in this external test set. The results of this evaluation are represented in Tables 13-15 and Table S1.

In spite of the fact that all the selected QSAR models were acceptable based on the internal validation, most of them showed a poor quality of prediction for the external validation sets. No single QSAR model for ovine COX-1 data had an $RMSE_{test}$ value less than 1. Only 6 QSAR models for the human COX-2 data and 1 QSAR model for the soybean LOX data had $RMSE_{test}$ values less than 1 (Tables 13-15). The best QSAR models based on a single publication for each target had the characteristics shown below:

- Ovine COX-1: $RMSE_{test}$ = 1.186;
- Human COX-2: $RMSE_{test}$ = 0.239;
- Soybean LOX: $RMSE_{test}$ = 0.689.

The combined training sets for each target were created using similar procedures to those used for the human COX-1 data (see above). Training sets based on all the data from ChEMBL related to the appropriate targets with average (CHEMBLav) and median (CHEMBLmed) values for duplicates were also created.

Analysis of the published results revealed that most of the datasets of ovine COX-1 data can be combined, leading to the creation of two training sets based on the two types of experimental assays: the Cayman assay kit (human recombinant protein COX-1) and the seminal vesicles-based assay (Table 10). Therefore, two combined training sets were created: CaymanTr, the training set based on the data given by the Cayman kit, and Seminal_vesiclesTr, the training set based on the data given by experiments using seminal vesicles.

The data on human COX-2 were divided into three types (Table 11):

- the data obtained using recombinant human COX-2 (hCOX-2Tr);
- the data obtained using cell cultures (CellsTr);
- the data obtained using blood as a source of COX-2 (BloodTr).

From the diversity of all the possible combined training sets, a GeneralTr was only created for soybean LOX because the data for soybean LOX could not be divided by the type of experiment and only one acceptable single QSAR model was created.

As a result, three types of combined training sets were created: with average data, median data and data scaled to the IC_{50} values of reference compounds (Celecoxib with IC_{50} = 34500 nM for ovine COX-1; Celecoxib, IC_{50} = 100 nM for human COX-2; and nordihydroguaiaretic acid with IC_{50} = 31300 nM for soybean LOX). New QSAR models based on these training sets were created and validated.

Tables 13-15 show a summary of the results of the validation of the QSAR models that were created on the basis of combined training sets for each target, as well as a summary of the results of the validation of the best single QSAR model(s) and the consensus of the prediction results given by single models.

Ovine COX-1

Table 13 shows that there is no single QSAR model based on the combined data with an $RMSE_{test}$ value lower than 1 among the best single QSAR models. However, many of the QSAR models based on the combined data (excluding the models based on Seminal_vesiclesTr) had better $RMSE_{test}$ values in comparison with the consensus (average value) of the prediction results given by all the single QSAR models.

The use of the scaling approach led to an improvement in the $RMSE_{test}$ value for the GeneralTr and Seminal_vesiclesTr data. The use of the CHEMBLmed training set led to the creation of the best QSAR model among those based on combined data ($RMSE_{test}$ = 0.704). The model based on the CHEMBLav training set also has a better $RMSE_{test}$ value (1.091) in

comparison with the best single QSAR model ($RMSE_{test}$ = 1.186) or the best QSAR model based on other combined training sets ($RMSE_{test}$ = 1.141 for Seminal_vesiclesTr with scaled data).

Table 13. Characteristics of the QSAR models created on the basis of the combined ovine COX-1 training sets.

№	Training set	N	R^2_{tr}	Q^2	$RMSE_{tr}$	$RMSE_{test}$
	QSAR models based on a single publication					
1	[60]	20	0.727	0.648	0.310	1.186
2	[61]	28	0.831	0.700	0.362	1.291
3	[62]	27	0.746	0.637	0.307	1.201
4	[63]	22	0.780	0.643	0.371	1.381
5	Consensus of prediction results given by the single QSAR models					1.199
	Training sets with average or median data					
6	GeneralTr	113	0.816	0.758	0.519	1.181
7	AcceptTr	91	0.661	0.559	0.387	1.294
8	CaymanTr	44	NA	NA	NA	NA
9	Seminal_vesiclesTr	49	0.944	0.642	0.404	1.334
	Training sets based on all ChEMBL data for ovine COX-1					
10	CHEMBLav (AD 85%)	700	0.604	0.504	0.766	1.091
11	CHEMBLmed (AD 45%)	700	0.944	0.535	0.728	0.704
	Training sets with scaling data to IC_{50} value of Celecoxib from test set (34500 nM)					
12	GeneralTr	113	0.900	0.864	0.501	1.147
13	AcceptTr	91	0.873	0.840	0.417	1.790
14	CaymanTr	44	NA	NA	NA	NA
15	Seminal_vesiclesTr	49	0.916	0.884	0.408	1.141

N – number of structures in a training set; NA means that the created QSAR model does not meet the minimum requirements (Q^2<0.5); RMSE values are represented on a logarithmic scale ($\log10(IC_{50}(nM))$).

Human COX-2

Table 14 shows a different result for the human COX-2 data. The evaluation set for the human COX-2 inhibition model (Table 9) consisted of only 8 compounds with IC_{50} values varying from IC_{50min} (63000 nM) to IC_{50max} (200000 nM). Since this is a rather narrow interval of

IC_{50} values, analysis of the $RMSE_{test}$ values is more preferable for evaluating the prediction results. Six out of the 37 QSAR models obtained had $RMSE_{test}$ values lower than 1. Most of the training sets of these models were characterized by high IC_{50min} (10^{-7} M) and IC_{50max} (10^{-4}-10^{-5} M) values (Table 11), the last of which is closer to the IC_{50} values for the external evaluation set (Table 9) with IC_{50min} of 10^{-5} M and IC_{50max} of 10^{-4} M. One model (#7) had an R^2_{test} value higher than 0.65, but its $RMSE_{test}$ value was high, indicating poor accuracy (1.923). This may be explained by the fact that the training set for this model has IC_{50max} of 66000 nM and the external evaluation set has IC_{50min} of 63000 nM. Traditionally, the results of QSAR predictions have a tendency to shift to the average value of a training set. Here, we see the same picture when the QSAR model figured out relationships between the structures and the values of activities, but in its own scale. In spite of the poor $RMSE_{test}$ value, this model may be used for the optimization of structures from the test set.

The $RMSE_{test}$ value, which was calculated as an average of the IC_{50} values of the prediction results given by the all single QSAR models (Consensus prediction), was worse than that of the best QSAR models based on single publications, but it was higher than the $RMSE_{test}$ values of the 14 other acceptable QSAR models created on the basis of single publications. All the compounds from the evaluation set were out of applicability domain for 8 of the 37 QSAR models based on single publications (Table S1).

Table 14. Characteristics of the QSAR models created using single and combined human COX-2 datasets.

№	Training set	N	R^2_{tr}	Q^2	$RMSE_{tr}$	$RMSE_{test}$
	Best QSAR models based on single publications					
4	[55]	23	0.859	0.786	0.411	0.239
30	[62]	30	0.957	0.737	0.372	0.645
35	[58]	35	0.969	0.764	0.278	0.710
18	[47]	34	0.754	0.615	0.815	0.742

3	[54]	34	0.943	0.635	0.215	0.776
21	[79]	23	0.987	0.786	0.319	0.801
11	[64]	22	0.970	0.734	0.526	1.923
8	Consensus of prediction results given by the all single QSAR models					1.601
	Training sets with average data					
	GeneralTr	1200	0.793	0.735	0.684	1.452
	AcceptTr	1121	0.828	0.774	0.651	1.327
	hCOX-2Tr	759	0.762	0.703	0.714	1.388
	CellsTr	171	0.711	0.614	0.605	2.703
	BloodTr	241	0.836	0.765	0.568	1.158
	Training sets based on all ChEMBL data for human COX-2					
	CHEMBLav	2253	0.732	0.673	0.723	1.278
	CHEMBLmed	2253	0.734	0.675	0.721	1.334
	Training sets with scaling data to IC_{50} value of Celecoxib from test set (100 nM)					
	GeneralTr	1200	0.776	0.717	0.657	1.700
	AcceptTr	1121	0.795	0.738	0.623	1.467
	hCOX-2Tr	759	0.735	0.673	0.693	1.957
	CellsTr	171	0.794	0.725	0.607	2.369
	BloodTr	241	0.857	0.793	0.577	1.650

№ – indicates the number of a set in Table 11; N – number of structures in a training set; - means that the value of $R^2_{test} \leq 0$; RMSE values are represented on a logarithmic scale ($\log10(IC_{50}(nM))$).

The $RMSE_{test}$ values of the QSAR models created on the basis of the combined training sets were worse than those of the best QSAR models created on the basis of single publications but most of them were better than the consensus of the prediction results based on all the single QSAR models (row 8 in Table 14). The QSAR models based on AcceptTr, hCOX-2Tr and BloodTr achieved higher accuracy in comparison with the QSAR models based on GeneralTr and CellsTr. The use of the scaling procedure only led to an improvement in the QSAR models for the CellsTr training set. The QSAR models based on all the ChEMBL data had better $RMSE_{test}$ values in most cases in comparison with the combined training set, but they were worse than the best QSAR models based on single training sets and had slightly worse $RMSE_{test}$ values compared to the best QSAR models based on the BloodTr training sets. This may be explained by several things, including:

1. The evaluation set contains compounds with weak inhibition activity against human COX-2, whereas the ChEMBL data contains many highly active inhibitors.
2. There are considerable differences in the experimental results derived from different assays. For instance, the IC_{50} values calculated using cells-based assays usually vary from 10^{-6} to 10^{-9} M whereas the IC_{50} values given by the human recombinant COX-2 protein assay vary from 10^{-4} to 10^{-7} M. This also explains the poor $RMSE_{test}$ values for the QSAR models based on the CellsTr training sets and the improvement of the QSAR model based on the CellsTr training set after the scaling procedure. If the same reference compounds were used in most publications, the use of the scaling procedure for all the ChEMBL data would possibly lead to a considerable improvement in the accuracy of the QSAR models.
3. There is a contradictory order of IC_{50} values for the same inhibitors in different assays. For instance, the IC_{50} values of Rofecoxib were half the values of Celecoxib when blood-derived COX-2 (500 nM and 1000 nM, respectively) and COS cells based assays (32 nM and 64 nM, respectively) were used, while the IC_{50} value of Rofecoxib (20 nM) is ten times higher than that of Celecoxib (2 nM) when the CHO cell assay was used. (Table S3).

Soybean LOX

Our study included three training sets with compounds tested for the inhibition of soybean LOX (Table 12). Only one out of the three QSAR models based on single publications was created with reasonable accuracy, estimated by internal validation (Table 15).

Only GeneralTr training sets were created because of the small number of training sets and because it was not possible to group them by different types of assays. Table 15 shows that all the QSAR models created on the basis of the combined training sets showed better accuracy than the QSAR model based on a single publication. Moreover, only 50% of the compounds from the

evaluation set (Table 9) were in the applicability domain of the QSAR model based on a single publication, while the QSAR models based on the combined training sets covered all the compounds from the evaluation set. The QSAR model created using the GeneralTr training set with averaging of the data from duplicates had the highest accuracy. For this target the use of the median value for duplicates from all the ChEMBL data led to a more accurate prediction ($RMSE_{test}$) than the average value for duplicates from all the ChEMBL data.

Table 15. Characteristics of the successful QSAR models for soybean LOX.

№	Ref	N	R^2_{tr}	Q^2	$RMSE_{tr}$	$RMSE_{test}$	AD,%
1	[90]	16	0.959	0.778	0.290	0.689	50
2	GeneralTr average	43	0.913	0.875	0.329	0.197	100
3	GeneralTr scaling	43	0.731	0.651	0.327	0.215	100
4	CHEMBLav	76	0.832	0.755	0.488	0.245	100
5	CHEMBLmed	76	0.844	0.770	0.473	0.240	100

N – number of structures in a training set; AD – part of compounds in Applicability Domain; RMSE values are represented on a logarithmic scale ($\log10(IC_{50}(nM))$).

Discussion and Conclusions

The $RMSE_{test}$ values given by the evaluation of all the created QSAR models are shown in Figure 4. One can see that in most cases the QSAR models created based on combined training sets (3-10 in Figure 4) have better $RMSE_{test}$ values than the QSAR models based on single publications or the consensus of QSAR models based on single publications.

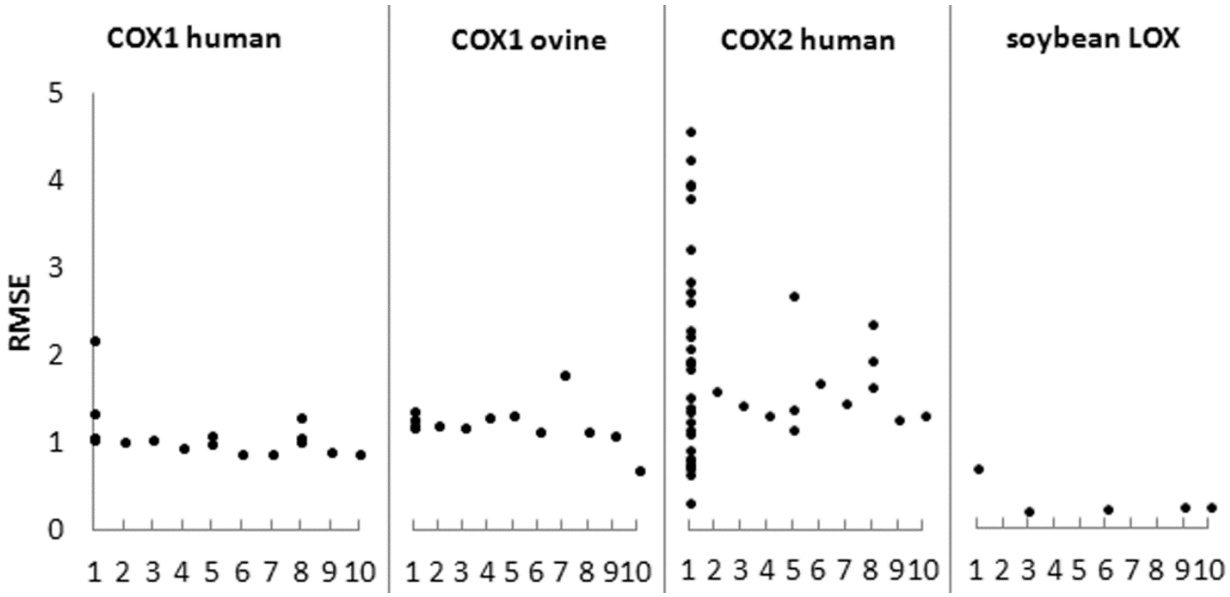


Figure 4. Comparison of RMSE values given for external test sets by different approaches to using training sets in QSAR modelling. 1 – QSAR models based on a single publication; 2 – The consensus of QSAR models based on single publications; 3 - QSAR models based on General training sets with averaged data; 4 - QSAR models based on Accepted training sets with averaged data; 5 - QSAR models based on training sets combined based on the same type of bioassay with averaged data; 6 - QSAR models based on General training sets with averaged data; 7 - QSAR models based on Accepted training sets with averaged data; 8 - QSAR models based on training sets combined based on the same type of bioassay with averaged data; 9 - QSAR models based on training sets with averaged ChEMBL data; 10 - QSAR models based on training sets with median ChEMBL data.

The QSAR models based on a combination of all the available IC_{50} data from ChEMBL for the appropriate target led to the most accurate predictions, in spite of the error in data. This proves the validity of the use of combined data for QSAR modelling, even if the combination is based on data obtained from different assays. Therefore, one of the main results of this study is the validity of the use of ChEMBL for the creation of general training sets and appropriate acceptable QSAR models for the estimation of the IC_{50} values of interaction between drug-like compounds and drug targets.

Based on the results of this study, we can draw the following conclusions:

1. *How can one use the experimental data from different studies with the same target and the same end-point?* – The experimental data from different publications may be combined in several ways:
    a. The creation of a general training set from a specific type of assay (or target source) or from the same assay.
    b. The creation of QSAR models based on single publications or based on the combination of the training sets including the training sets of the best QSAR models into one general training set.
    c. The creation of a general training set from all publications using averaged or median values for duplicates or using a scaling procedure, if reference compounds are available. In some cases, the use of median or scaled values led to a considerable improvement in the quality of the QSAR models in comparison to the use of averaged values.
2. *Should we select the data provided only by one study or from studies that use the same assay?* – The creation of QSAR models based on a single publication should be done with the aim of finding the most accurate model for the optimization of structures. If a publication contains a small number of compounds, its data may be used for QSAR modelling for structure optimization but this kind of model often has a narrow applicability domain and cannot be used for virtual screening. The creation of a general training set from different publications increases the number and diversity of compounds and may lead to a more general QSAR model, which has a wider applicability domain and may be used for virtual screening or the estimation of activity for a broader class of chemicals.

3. *May the data from several assays be combined?* Several assays may be combined into a general training set using averaged or median values for duplicates or using a scaling procedure, if reference compounds are available. However, some data in ChEMBL do not include the detailed description of assays necessary to recognize their type. QSAR models created based on the aggregation of these data and data with known assays show the highest accuracy.
4. *Which is better: the creation of QSAR models based on a single study and then the use of them to create a consensus prediction or the combination of all the data into one general training set?* – Our study showed that the creation of a general training set leads to more accurate QSAR models than the creation of QSAR models based on a single publication and then the use of them to create a consensus prediction.

Our study reveals some imperfections in the data provided in publications by medicinal chemists and we would like to suggest some recommendations to the authors and editors of journals for publishing experimental studies:

1. *Try to obtain and publish exact IC_{50} values for low activity or inactive compounds.* The analysis of the experimental data from publications that evaluate inhibitors of human COX-1, ovine COX-1, human COX-2 and soybean LOX showed that up to 35% of structures cannot be used for QSAR modelling because of the absence of accurate IC_{50} values. Publishing mainly the IC_{50} values for moderate and highly active compounds does not allow the creation of accurate QSAR models that may be used for virtual screening and to distinguish active and inactive compounds.
2. *Provide a clear description of assays.* It allows the aggregation of data based on the type of assay, leading to more accurate QSAR models.

3. *Experimentally validate reference compounds even when using standard kits.* This provides the potential to use a scaling procedure, which demonstrated advantages in comparison with others methods of creation of combined training sets. Moreover, the authors of some publications often referred to the standard values of reference compounds provided by suppliers, while other publications have shown that even with the same kit different authors obtain different activity values for reference compounds. These differences may be related to the use of different substrate concentrations in the activity measurement assays. Depending on the type of inhibition (competitive, non-competitive, or allosteric) changes in the substrate concentration may lead to an increase or a decrease in the IC_{50} values of the compounds. This is why Ki values, when available, may prove to be more helpful in QSAR studies in some cases.

We would like to highlight the importance of the substrate (arachidonic acid) concentration in the IC_{50} values. The Cayman kit suggests an arachidonic acid concentration of 100 micromolar, but it also suggests that researchers may make the appropriate dilutions to achieve the substrate concentration that they consider to be best. The initial concentration of 100 micromole is much higher than the Km value of COX-1 (by more than one order of magnitude) and is not appropriate for the fair estimation of the inhibition of competitive inhibitors. Very little inhibition would be detected if the concentration is too high. As a result, many researchers choose lower substrate concentrations such as 20, 10, 1 or even 0.1 micromolar. These differences in substrate concentrations may lead to IC_{50} values that differ by orders of magnitude. The enzyme concentration may also affect the IC_{50} value. This is why Ki values are preferred when the inhibition activity of compounds has been estimated by different laboratories. Both the substrate and the enzyme concentration are taken in to account in the calculation of Ki.

Unfortunately, the substrate concentration does not affect all types of inhibitors in the same way. The inhibition increases when the substrate concentration decreases for competitive inhibitors, but the opposite occurs for uncompetitive inhibitors. As a result, correction using a scale ratio could improve the results if applied to the same type of inhibitors. However, the mode of inhibition is not mentioned by most studies.

From this point of view, if enough data were available, then better results could be obtained if the results produced by the Cayman method were divided into at least two categories: the results produced using high substrate concentrations (10x Km or higher) and the results produced using substrate concentrations of ~Km or lower. The IC_{50} values predicted using the first training set may agree with the biological results produced using high substrate concentrations and the IC_{50} values predicted using the second training set may agree with the biological results produced using low substrate concentrations. The results that we used as external test sets were obtained using low substrate concentrations.

ASSOCIATED CONTENT

Supplements.PDF includes:

Table S1. Characteristics of accepted QSAR models for human COX-2 datasets.

Table S2. IC50 values of reference compounds in the studied ovine COX-1 datasets.

Table S3. IC50 values of reference compounds in the studied human COX-2 datasets.

Figure S1. Comparison of average RMSE values given for external test sets by different approaches to using training sets in QSAR modeling.

References which were mentioned in Tables of Supplements.

AUTHOR INFORMATION

Corresponding Author

119121, Pogodinskaya Str. 10/8, Moscow, Russia; e-mail: alexey.lagunin@ibmc.msk.ru

Author Contributions

A.A.L. and A.V.Z. designed research and QSAR modeling. P.V.P creation of datasets. A.G. and P.E. synthesis and experimental studies of compounds. The manuscript was written using contributions from all the authors. All authors edited the manuscript and approved the final version.

ACKNOWLEDGMENT

The work performed by Alexey A. Lagunin and Pavel V. Pogodin was supported by the Russian Science Foundation (grant No. 14-15-00449).

ABBREVIATIONS

COX, cyclooxygenase; GUSAR – General Unrestricted Structure-Activity Relationships; IC_{50}, half maximal inhibitory concentration; LOX, lipoxygenase; QNA, Quantitative Neighbourhoods of Atoms descriptors; MNA, Multilevel Neighbourhoods of Atoms descriptors; QSAR, Quantitative structure–activity relationships; PASS, Prediction of the Biological Activity Spectra; RBF, Radial Basis Function; RMSE, Root Mean Square Error; SCR, Self-Consistent Regression.

REFERENCES

(1) Kim, S.; Thiessen, P.A.; Bolton, E.E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B.A.; Wang, J.; Yu, B.; Zhang, J.; Bryant, S.H. PubChem Substance and Compound Databases, *Nucleic Acids Res.* **2016**, *44*, D1202- D1213.

(2) Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A.P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L.J.; Cibrián-Uhalte, E.; Davies, M.; Dedman, N.; Karlsson, A.; Magariños, M.P.; Overington, J.P.; Papadatos, G.; Smit, I.; Leach, A.R. The ChEMBL Database in 2017. *Nucleic Acids Res.* **2017**, *45*, D945-D954.

(3) Liu, T.; Lin, Y.; Wen, X.; Jorrisen, R.N.; Gilson, M.K. BindingDB: a Web-accessible Database of Experimentally Determined Protein-Ligand Binding Affinities. *Nucleic Acids Res.* **2007**, *35*, D198-D201.

(4) Olah, M.; Rad, R.; Ostopovici, L.; Bora, A.; Hadaruga, N.; Hadaruga, D.; Moldovan, R.; Fulias, A.; Mracec, M.; Oprea, T.I. WOMBAT and WOMBAT-PK: Bioactivity Databases for Lead and Drug Discovery, In: S.L. Schreiber, T.M. Kapoor, G. Wess, Eds., *Chemical Biology: From Small Molecules to Systems Biology and Drug Design*; Wiley-VCH, New York, 2007; pp. 760-786.

(5) Cronin, M.T.D.; Schultz, W.T. Pitfalls in QSAR. *J. Mol. Struct. (Theochem)*, **2003**, *622*, 39–51.

(6) van de Waterbeemd, H.; Rose, S. Quantitative Approaches to Structure–Activity Relationships, In: C.G.Wermuth, Eds., *Wermuth's The Practice of Medicinal Chemistry*, 3rd ed. Elsevier Ltd: 2008; pp. 491-512.

(7) The Report from the Expert Group on (Quantitative) Structure-Activity Relationships [(Q)SARs] on the Principles for the Validation of (Q)SARs. OECD Series on Testing and assessment, 2004, 48, pp. 1-206 [Online] http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono(2004)24&doclanguage=en (accessed 17 January 2019).

(8) Dearden, J.C.; Cronin, M.T.; Kaiser, K.L. How not to Develop a Quantitative Structure-Activity or Structure-Property Relationship (QSAR/QSPR). *SAR QSAR Environ Res.* **2009**, *20*, 241-266.

(9) Fourches, D.; Muratov, E.; Tropsha, A. Trust, but Verify: on the Importance of Chemical Structure Curation in Cheminformatics and QSAR Modeling Research. *J. Chem. Inf. Model.* **2010**, *50*, 1189-1204.

(10) Cherkasov, A.; Muratov, E.N.; Fourches, D.; Varnek, A.; Baskin, I.I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y.C.; Todeschini, R.; Consonni, V.; Kuz'min, V.E.; Cramer, R.; Benigni, R.; Yang, C.; Rathman, J.; Terfloth, L.; Gasteiger, J.; Richard, A.; Tropsha, A. QSAR Modeling: Where Have You Been? Where Are You Going To? *J. Med. Chem.* **2014**, *57*, 4977-5010.

(11) Tarasova, O.A.; Urusova, A.F.; Filimonov, D.A.; Nicklaus, M.C.; Zakharov, A.V.; Poroikov, V.V. QSAR Modeling Using Large-Scale Databases: Case Study for HIV-1 Reverse Transcriptase Inhibitors. *J. Chem. Inf. Model.* **2015**, *55*, 1388-1399.

(12) Cortes-Ciriano, I.; Murrell, D.S.;, van Westen, G.J.; Bender, A.; Malliavin, T.E. Prediction of the Potency of Mammalian Cyclooxygenase Inhibitors with Ensemble Proteochemometric Modeling. *J Cheminform.* **2015**, *7*, 1. doi:10.1186/s13321-014-0049-z.

(13) Toropova, A.P.; Toropov, A.A.; Benfenati, E. CORAL: Prediction of Binding Affinity and Efficacy of Thyroid Hormone Receptor Ligands. *Eur. J. Med. Chem.* **2015**, *101*, 452-461.

(14) Cortes-Ciriano, I.; Bender, A. Improved Chemical Structure-Activity Modeling Through Data Augmentation. *J. Chem. Inf. Model.* **2015**, *55*, 2682-2692.

(15) Politi, R.; Rusyn, I.; Tropsha, A. Prediction of Binding Affinity and Efficacy of Thyroid Hormone Receptor Ligands Using QSAR and Structure-Based Modeling Methods. *Toxicol Appl. Pharmacol.* **2014**, *280*, 177-189.

(16) Sagardia, I.; Roa-Ureta, R.H.; Bald, C. A New QSAR Model, for Angiotensin I-Converting Enzyme Inhibitory Oligopeptides. *Food Chem*. **2013**, *136*, 1370-1376.

(17) Zhao, L.; Xiang, Y.; Song, J.; Zhang, Z. A Novel Two-Step QSAR Modeling Work Flow to Predict Selectivity and Activity of HDAC Inhibitors. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 929-933.

(18) Lagunin, A.A., Gloriozova, T.A., Dmitriev, A.V., Volgina, N.E., Poroikov, V.V. Computer Evaluation of Drug Interactions with P-Glycoprotein. *Bull. Exp. Biol. Med.* **2013**, *154*, 521-524.

(19) Myint, K.Z.; Wang, L.; Tong, Q.; Xie, X.Q. Molecular Fingerprint-Based Artificial Neural Networks QSAR for Ligand Biological Activity Predictions. *Mol. Pharm.* **2012**, *9*, 2912-2923.

(20) Bitencourt, M.; Freitas, M.P.; Rittner, R. The MIA-QSAR Method for the Prediction of Bioactivities of Possible Acetylcholinesterase Inhibitors. *Arch. Pharm. (Weinheim)*, **2012**, *345*, 723-728.

(21) Kovalishyn, V.; Tanchuk, V.; Charochkina, L.; Semenuta, I.; Prokopenko, V. Predictive QSAR Modeling of Phosphodiesterase 4 Inhibitors. *J. Mol. Graphics Modell.* **2012**, *32*, 32-38.

(22) Saranya, N.; Selvaraj, S. QSAR Studies on HIV-1 Protease Inhibitors Using Non-Linearly Transformed Descriptors, *Curr. Comput.-Aided Drug Des.* **2012**, *8*, 10-49.

(23) Wang, Y.; Cheng, X.; Lin, Y.; Wen, H.; Wang, L.; Xia, Q.; Lin, Z. TAP-Binding Peptides Prediction by QSAR Modeling Based on Amino Acid Structural Information. *Curr. Comput.-Aided Drug Des.* **2012**, *8*, 50-54.

(24) Martin, E.; Mukherjee, P. Kinase-Kernel Models: Accurate In Silico Screening of 4 Million Compounds Across the Entire Human Kinome. *J. Chem. Inf. Model.* **2012**, *52*, 156-170.

(25) Zakharov, A.V.; Lagunin, A.A.; Filimonov, D.A.; Poroikov, V.V. Quantitative Prediction of Antitarget Interaction Profiles for Chemical Compounds. *Chem Res Toxicol.* **2012**, *25*, 2378-2385.

(26) Lagunin, A.A.; Romanova, M.A.; Zadorozhny, A.D.; Kurilenko, N.S.; Shilov, B.V.; Pogodin P.V.; Ivanov S.M.; Filimonov, D.A.; Poroikov, V.V. Comparison of Quantitative

and Qualitative (Q)SAR Models Created for the Prediction of Ki and IC50 Values of Antitarget Inhibitors. *Front. Pharmacol.* **2018**, *9*, 1136. doi: 10.3389/fphar.2018.01136.

(27) Charlier, C.; Michaux, C. Dual Inhibition of Cyclooxygenase-2 (COX-2) and 5-Lipoxygenase (5-LOX) as a New Strategy to Provide Safer Non-Steroidal Anti-Inflammatory Drugs. *Eur. J. Med. Chem.* **2003**, *38*, 645-659.

(28) Martel-Pelletier, J.; Lajeunesse, D.; Reboul, P.; Pelletier, J.-P. Therapeutic Role of Dual Inhibitors of 5-LOX and COX, Selective and Non-Selective Non-Steroidal Anti-Inflammatory Drugs. *Ann. Rheum. Dis.* **2003**, *62*, 501–509.

(29) Geronikaki, A.A.; Lagunin, A.A.; Hadjipavlou-Litina, D.I.; Eleftheriou, Ph.T.; Filimonov, D.A.; Poroikov, V.V.; Alam, I.; Saxena, A.K. Computer-Aided Discovery of Anti-Inflammatory Thiazolidinones with Dual Cyclooxygenase/Lipoxygenase inhibition. *J. Med. Chem.* **2008**, *51*, 1601-1609.

(30) Lagunin, A.; Zakharov, A.; Filimonov, D.; Poroikov, V. QSAR Modelling of Rat Acute Toxicity on the Basis of PASS Prediction. *Mol. Inform.* **2011**, *30*, 241–250.

(31) Zakharov, A.V.; Peach, M.L.; Sitzmann, M.; Nicklaus, M.C. QSAR Modeling of Imbalanced High-Throughput Screening Data in PubChem. *J. Chem. Inf. Model.* **2014**, *54*, 705–712.

(32) Eleftheriou, P.; Geronikaki, A.; Hadjipavlou-Litina, D.; Vicini, P.; Filz, O.; Filimonov, D.; Poroikov, V.; Chaudhaery, S.S.; Roy, K.K.; Saxena, A. Fragment-Based Design, Docking, Synthesis, Biological Evaluation and Structure–Activity Relationships of 2-Benzo/Benzisothiazolimino–5–Aryliden–4-Thiazolidinones as Cycloxygenase/Lipoxygenase Inhibitors. *Eur. J. Med. Chem.* **2012**, *47*, 111-124.

(33) Massart, D.; Vandeginste, B.; Buydens, L.; De Jong, S.; Lewi, P.; Smeyers-Verbeke, J. Handbook of Chemometrics and Qualimetrics: Part A. // Elsevier Science, Amsterdam, 1997, pp. 867.

(34) Fedorova, E.V.; Buryakina, A.V.; Zakharov, A.V.; Filimonov, D.A.; Lagunin, A.A.; Poroikov, V.V. Design, Synthesis and Pharmacological Evaluation of Novel Vanadium-Containing Complexes as Antidiabetic Agents. *PloS One*, **2014**, *9*, e100386.

(35) Filimonov, D.A.; Zakharov, A.V.; Lagunin, A.A.; Poroikov, V.V. QNA Based "Star Track" QSAR Approach. *SAR QSAR Environ. Res.* **2009**, *20*, 679–709.

(36) Lagunin, A.A.; Zakharov, A.V.; Filimonov, D.A.; Poroikov, V.V. A New Approach to QSAR Modelling of Acute Toxicity. *SAR QSAR Environ. Res.* **2007**, *18*, 285–298.

(37) Zakharov, A.V.; Peach, M.L.; Sitzmann, M.; Filippov, I.V.; McCartney, H.J.; Smith, L.H.; Pugliese, A.; Nicklaus, M.C. Computational Tools and Resources for Metabolism-Related Property Predictions. 2. Application to Prediction of Half-Life Time in Human Liver Microsomes. *Future Med. Chem.* **2012**, *4*, 1933–1944.

(38) Filimonov, D.A.; Akimov, D.V.; Poroikov, V.V. The Method of Self-Consistent Regression for the Quantitative Analysis of Relationships Between Structure and Properties of Chemicals. *Pharm. Chem. J.* **2004**, *38*, 21–24.

(39) Zakharov, A.V.; Peach, M.L.; Sitzmann, M.; Nicklaus, M.C. A New Approach to Radial Basis Function Approximation and Its Application to QSAR. *J. Chem. Inf. Model.* **2014**, *54*, 713–719.

(40) Vicini, P.; Geronikaki, A.; Incerti, M.; Zani, F.; Dearden, J.; Hewitt, M. 2-Heteroarylimino-5-Benzylidene-4-Thiazolidinones Analogues of 2-Thiazolylimino-5-

Benzylidene-4-Thiazolidinones with Antimicrobial Activity: Synthesis and Structure-Activity Relationship. *Bioorg. Med. Chem.* **2008**, *16*, 3714-3724.

(41) Pontiki, E.; Hadjipavlou-Litina, D.; Litinas, K.; Nicolotti, O.; Carotti, A. Design, Synthesis and Pharmacobiological Evaluation of Novel Acrylic Acid Derivatives Acting as Lipoxygenase and Cyclooxygenase-1 Inhibitors with Antioxidant and Anti-Inflammatory Activities. *Eur. J. Med. Chem.* 2011, *46*, 191-200.

(42) Papadopoulou, C.; Geronikaki, A.; Hadjipavlou-Litina, D. Synthesis and Biological Evaluation of New Thiazolyl/Benzothiazolyl-Amides, Derivatives of 4-Phenyl-Piperazine. *Farmaco*, **2005**, *60*, 969-973.

(43) Lapenna, D.; Ciofani, G.; Pierdomenico, S.D.; Neri, M.; Cuccurullo, C.; Giamberardino, M.A.; Cuccurullo, F. Inhibitory Activity of Salicylic Acid on Lipoxygenase-Dependent Lipid Peroxidation. *Biochim. Biophys. Acta*, **2009**, *1790*, 25-30.

(44) Taraporewala, I.B.; Kauffman, J.M. Synthesis and Structure Activity Relationships of Anti-Inflammatory 9,10-Dihydro-9-Oxo-2-Acridinealkanoic Acids and 4-(2-Carboxyphenyl)Aminobenzenealkanoic Acids. *J. Pharm. Sci.* **1990**, *79*, 173-178.

(45) Somvanshi, R.K.; Singh, A.K.; Saxena, M.; Mishra, B.; Dey, S. Development of Novel Peptide Inhibitor of Lipoxygenase Based on Biochemical and BIAcore Evidences. *Biochim. Biophys. Acta*, **2008**, *1784*, 1812-1817.

(46) Almansa, C.; Alfón, J.; de Arriba, A.F.; Cavalcanti, F.L.; Escamilla, I.; Gómez, L.A.; Miralles, A.; Soliva, R.; Bartrolí, J.; Carceller, E.; Merlos, M.; García-Rafanell, J. Synthesis and Structure-Activity Relationship of a New Series of COX-2 Selective Inhibitors: 1,5-Diarylimidazoles. *J. Med. Chem.* **2003**, *46*, 3463-3475.

(47) Wang, J.L.; Limburg, D.; Graneto, M.J.; Springer, J.; Hamper, J.R.; Liao, S.; Pawlitz, J.L.; Kurumbail, R.G.; Maziasz, T.; Talley, J.J.; Kiefer, J.R.; Carter, J. The Novel Benzopyran Class of Selective Cyclooxygenase-2 Inhibitors. Part 2: the Second Clinical Candidate Having a Shorter and Favorable Human Half-Life. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 7159-7163.

(48) Li, J.J.; Norton, M.B.; Reinhard, E.J.; Anderson, G.D.; Gregory, S.A.; Isakson, P.C.; Koboldt, C.M.; Masferrer, J.L.; Perkins, W.E.; Seibert, K.; Zhang, Y.; Zweifel, B.S.; Reitz, D.B. Novel Terphenyls as Selective Cyclooxygenase-2 Inhibitors and Orally Active Anti-Inflammatory Agents. *J. Med. Chem.* **1996**, *39*, 1846-1856.

(49) Penning, T.D.; Talley, J.J.; Bertenshaw, S.R.; Carter, J.S.; Collins, P.W.; Docter, S.; Graneto, M.J.; Lee, L.F.; Malecha, J.W.; Miyashiro, J.M.; Rogers, R.S.; Rogier, D.J.; Yu, S.S.; Anderson, G.D.; Burton, E.G.; Cogburn, J.N.; Gregory, S.A.; Koboldt, C.M.; Perkins, W.E.; Seibert, K.; Veenhuizen, A.W.; Zhang, Y.Y.; Isakson, P.C. Synthesis and Biological Evaluation of the 1,5-Diarylpyrazole Class of Cyclooxygenase-2 Inhibitors: Identification of 4-[5-(4-Methylphenyl)-3-(Trifluoromethyl)-1H-Pyrazol-1-Yl]Benze Nesulfonamide (SC-58635, Celecoxib). *J. Med. Chem.* **1997**, *40*, 1347-1365.

(50) Black, W.C.; Brideau, C.; Chan, C.C.; Charleson, S.; Chauret, N.; Claveau, D.; Ethier, D.; Gordon, R.; Greig, G.; Guay, J.; Hughes, G.; Jolicoeur, P.; Leblanc, Y.; Nicoll-Griffith, D.; Ouimet, N.; Riendeau, D.; Visco, D.; Wang, Z.; Xu, L.; Prasit, P. 2,3-Diarylcyclopentenones as Orally Active, Highly Selective Cyclooxygenase-2 Inhibitors. *J. Med. Chem.* **1999**, *42*, 1274-1281.

(51) Huang, H.C.; Li, J.J.; Garland, D.J.; Chamberlain, T.S.; Reinhard, E.J.; Manning, R.E.; Seibert, K.; Koboldt, C.M.; Gregory, S.A.; Anderson, G.D.; Veenhuizen, A.W.; Zhang, Y.;

Perkins, W.E.; Burton, E.G.; Cogburn, J.N.; Isakson, P.C.; Reitz, D.B. Diarylspiro[2.4]Heptenes as Orally Active, Highly Selective Cyclooxygenase-2 Inhibitors: Synthesis and Structure-Activity Relationships. *J. Med. Chem.* **1996**, *39*, 253-266.

(52) Khanna, I.K.; Weier, R.M.; Yu, Y.; Xu, X.D.; Koszyk, F.J.; Collins, P.W.; Koboldt, C.M.; Veenhuizen, A.W.; Perkins, W.E.; Casler, J.J.; Masferrer, J.L.; Zhang, Y.Y.; Gregory, S.A.; Seibert, K.; Isakson, P.C. 1,2-Diarylimidazoles as Potent, Cyclooxygenase-2 Selective, and Orally Active Antiinflammatory Agents. *J. Med. Chem.* **1997**, *40*, 1634-1647.

(53) Janusz, J.M.; Young, P.A.; Ridgeway, J.M.; Scherz, M.W.; Enzweiler, K.; Wu, L.I.; Gan, L.; Chen, J.; Kellstein, D.E.; Green, S.A.; Tulich, J.L.; Rosario-Jansen, T.; Magrisso, I.J.; Wehmeyer, K.R.; Kuhlenbeck, D.L.; Eichhold, T.H.; Dobson, R.L. New Cyclooxygenase-2/5-Lipoxygenase Inhibitors. 3.7-Tert-Butyl-2, 3-Dihydro-3,3-Dimethylbenzofuran Derivatives as Gastrointestinal Safe Antiinflammatory and Analgesic Agents: Variations at the 5 Position. *J. Med. Chem.* **1998**, *41*, 3515-3529.

(54) Lin, S.J.; Tsai, W.J.; Chiou, W.F.; Yang, T.H.; Yang, L.M. Selective COX-2 Inhibitors. Part 2: Synthesis and Biological Evaluation of 4-Benzylideneamino- and 4-Phenyliminomethyl-Benzenesulfonamides. *Bioorg. Med. Chem.* **2008**, *16*, 2697-2706.

(55) Yamakawa, N.; Suemasu, S.; Matoyama, M.; Tanaka, K.; Katsu, T.; Miyata, K.; Okamoto, Y.; Otsuka, M.; Mizushima, T. Synthesis and Biological Evaluation of Loxoprofen Derivatives. *Bioorg. Med. Chem.* **2011**, *19*, 3299-3311.

(56) Wey, S.J.; Augustyniak, M.E.; Cochran, E.D.; Ellis, J.L.; Fang, X.; Garvey, D.S.; Janero, D.R.; Letts, L.G.; Martino, A.M.;. Melim, T.L; Murty, M.G.; Richardson, S.K.; Schroeder, J.D.; Selig, W.M.; Trocha, A.M.; Wexler, R.S.; Young, D.V.; Zemtseva, I.S.; Zifcak, B.M.

Structure-Based Design, Synthesis, and Biological Evaluation of Indomethacin Derivatives as Cyclooxygenase-2 Inhibiting Nitric Oxide Donors. *J. Med. Chem.* **2007**, *50*, 6367-6382.

(57) Renard, J.F.; Arslan, D.; Garbacki, N.; Pirotte, B.; de Leval, X. Pyridine Analogues of Nimesulide: Design, Synthesis, and In Vitro and In Vivo Pharmacological Evaluation as Promising Cyclooxygenase 1 and 2 Inhibitors. *J. Med. Chem.* **2009**, *52*, 5864-5871.

(58) Puig, C.; Crespo, M.I.; Godessart, N.; Feixas, J.; Ibarzo, J.; Jiménez, J.M.; Soca, L.; Cardelús, I.; Heredia, A.; Miralpeix, M.; Puig, J.; Beleta, J.; Huerta, J.M.; López, M.; Segarra, V.; Ryder, H.; Palacios, J.M. Synthesis and Biological Evaluation of 3,4-Diaryloxazolones: a New Class of Orally Active Cyclooxygenase-2 Inhibitors. *J. Med. Chem.* **2000**, *43*, 214-223.

(59) Julémont, F.; de Leval, X.; Michaux, C.; Renard, J.F.; Winum, J.Y.; Montero, J.L.; Damas, J.; Dogné, J.M.; Pirotte, B. Design, Synthesis, and Pharmacological Evaluation of Pyridinic Analogues of Nimesulide as Cyclooxygenase-2 Selective Inhibitors. *J. Med. Chem.* **2004**, *47*, 6749-6759.

(60) Harrak, Y.; Casula, G.; Basset, J.; Rosell, G.; Plescia, S.; Raffa, D.; Cusimano, M.G.; Pouplana, R.; Pujol, M.D. Synthesis, Anti-Inflammatory Activity, and In Vitro Antitumor Effect of a Novel Class of Cyclooxygenase Inhibitors: 4-(Aryloyl)Phenyl Methyl Sulfones. *J. Med. Chem.* **2010**, *53*, 6560-6571.

(61) Coy, E.D.; Cuca, L.E.; Sefkow, M. COX, LOX and Platelet Aggregation Inhibitory Properties of Lauraceae Neolignans. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6922-6925.

(62) Kalgutkar, A.S.; Kozak, K.R.; Crews, B.C.; Hochgesang, G.P.Jr.; Marnett, L.J. Covalent Modification of Cyclooxygenase-2 (COX-2) by 2-Acetoxyphenyl Alkyl Sulfides, a New Class of Selective COX-2 Inactivators. *J. Med. Chem.* **1998**, *41*, 4800-4818.

(63) Singh, S.K.; Reddy, P.G.; Rao, K.S.; Lohray, B.B.; Misra, P.; Rajjak, S.A.; Rao, Y.K.; Venkateswarlu, A. Polar Substitutions in the Benzenesulfonamide Ring of Celecoxib Afford a Potent 1,5-Diarylpyrazole Class of COX-2 Inhibitors. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 499-504.

(64) Kalgutkar, A.S.; Rowlinson, S.W.; Crews, B.C.; Marnett, L.J. Amide Derivatives of Meclofenamic Acid as Selective Cyclooxygenase-2 Inhibitors. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 521-524.

(65) Orjales, A.; Mosquera, R.; López, B.; Olivera, R.; Labeaga, L.; Núñez, M.T. Novel 2-(4-Methylsulfonylphenyl)Pyrimidine Derivatives as Highly Potent and Specific COX-2 Inhibitors. *Bioorg. Med. Chem.* **2008**, *16*, 2183-2199.

(66) Lu, X.; Zhang, H.; Li, X.; Chen, G.; Li, Q.S.; Luo, Y.; Ruan, B.F.; Chen, X.W.; Zhu, H.L. Design, Synthesis and Biological Evaluation of Pyridine Acyl Sulfonamide Derivatives as Novel COX-2 Inhibitors. *Bioorg. Med. Chem.* **2011**, *19*, 6827-6832.

(67) Qiu, K.M.; Yan, R.; Xing, M.; Wang, H.H.; Cui, H.E.; Gong, H.B.; Zhu, H.L. Synthesis, Biological Evaluation and Molecular Modeling of Dihydro-Pyrazolyl-Thiazolinone Derivatives as Potential COX-2 Inhibitors. *Bioorg. Med. Chem.*, **2012**, *20(22)*, 6648-6654.

(68) Pinto, D.J.P.; Copeland, R.A.; Covington, M.B.; Pitts, W.J.; Batt, D.G.; Orwat, M.J.; Lam, G.N.; Joshi, A.; Chan, Y-C.; Wang, S.; Trzaskos, J.M.; Magolda, R.L.; Kornhauser, D.M. Chemistry and Pharmacokinetics of Diarylthiophenes and Terphenyls as Selective COX-2 Inhibitors. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2907-2912.

(69) Carter, J.S.; Rogier, D.J.; Graneto, M.J.; Seibert, K.; Koboldt, C.M.; Zhang, Y.; Talley, J.J. Design and Synthesis of Sulfonyl-Substituted 4,5-Diarylthiazoles as Selective Cyclooxygenase-2 Inhibitors. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1167-1170.

(70) Lau, C.K.; Brideau, C.; Chan, C.C.; Charleson, S.; Cromlish, W.A.; Ethier, D.; Gauthier, J.Y.; Gordon, R.; Guay, J.; Kargman, S.; Li, C.S.; Prasit, P.; Riendeau, D.; Thérien, M.; Visco, D.M.; Xu, L. Synthesis and Biological Evaluation of 3-Heteroaryloxy-4-Phenyl-2(5H)-Furanones as Selective COX-2 Inhibitors. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 3187-3192.

(71) Sui, Z.; Guan, J.; Ferro, M.P.; McCoy, K.; Wachter, M.P.; Murray, W.V.; Singer, M.; Steber, M.; Ritchie, D.M.; Argentieri, D.C. 1,3-Diarylcycloalkanopyrazoles and Diphenyl Hydrazides as Selective Inhibitors of Cyclooxygenase-2. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 601-604.

(72) Black, W.C.; Brideau, C.; Chan, C.C.; Charleson, S.; Cromlish, W.; Gordon, R.; Grimm, E.L.; Hughes, G.; Leger, S.; Li, C.S.; Riendeau, D.; Thérien, M.; Wang, Z.; Xu, L.J.; Prasit, P. 3,4-Diaryl-5-Hydroxyfuranones: Highly Selective Inhibitors of Cyclooxygenase-2 with Aqueous Solubility. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1195-1198.

(73) Campbell, J.A.; Bordunov, V.; Broka, C.A.; Browner, M.F.; Kress, J.M.; Mirzadegan, T.; Ramesha, C.; Sanpablo, B.F.; Stabler, R.; Takahara, P.; Villasenor, A.; Walker, K.A.; Wang, J.H.; Welch, M.; Weller, P. Rational Design of 6-Methylsulfonylindoles as Selective Cyclooxygenase-2 Inhibitors. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4741-4745.

(74) Beswick, P.; Bingham, S.; Bountra, C.; Brown, T.; Browning, K.; Campbell, I.; Chessell, I.; Clayton, N.; Collins, S.; Corfield, J.; Guntrip, S.; Haslam, C.; Lambeth, P.; Lucas, F.; Mathews, N.; Murkit, G.; Naylor, A.; Pegg, N.; Pickup, E.; Player, H.; Price, H.; Stevens, A.; Stratton, S.; Wiseman, J. Identification of 2,3-Diaryl-Pyrazolo[1,5-B]Pyridazines as Potent and Selective Cyclooxygenase-2 Inhibitors. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5445-5448.

(75) Prasanna, S.; Manivannan, E.; Chaturvedi, S.C. QSAR Studies on Structurally Similar 2-(4-Methanesulfonylphenyl)Pyran-4-Ones as Selective COX-2 Inhibitors: a Hansch Approach. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 313-320.

(76) Swarbrick, M.E.; Beswick, P.J.; Gleave, R.J.; Green, R.H.; Bingham, S.; Bountra, C.; Carter, M.C.; Chambers, L.J.; Chessell, I.P.; Clayton, N.M.; Collins, S.D.; Corfield, J.A.; Hartley, C.D.; Kleanthous, S.; Lambeth, P.F.; Lucas, F.S.; Mathews, N.; Naylor, A.; Page, L.W.; Payne, J.J.; Pegg, N.A.; Price, H.S.; Skidmore, J.; Stevens, A.J.; Stocker, R.; Stratton, S.C.; Stuart, A.J.; Wiseman, J.O. Identification of [4-[4-(Methylsulfonyl)Phenyl]-6-(Trifluoromethyl)-2-Pyrimidinyl] Amines and Ethers as Potent and Selective Cyclooxygenase-2 Inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4504-4508.

(77) Yao, R.S.; Lu, X.Q.; Guan, Q.X.; Zheng, L.; Lu, X.; Ruan, B.F. Synthesis and Biological Evaluation of Some Novel Resveratrol Amide Derivatives as Potential Anti-Tumor Agents. *Eur. J. Med. Chem.* **2013**, *62*, 222-231.

(78) Reitz, D.B.; Li, J.J.; Norton, M.B.; Reinhard, E.J.; Collins, J.T.; Anderson, G.D.; Gregory, S.A.; Koboldt, C.M.; Perkins, W.E.; Seibert, K.; Isakson, P.C. Selective Cyclooxygenase Inhibitors: Novel 1,2-Diarylcyclopentenes are Potent and Orally Active COX-2 Inhibitors. *J. Med. Chem.* **1994**, *37*, 3878-3881.

(79) Wilkerson, W.W.; Copeland, R.A.; Covington, M.; Trzaskos, J.M. Antiinflammatory 4,5-Diarylpyrroles. 2. Activity as a Function of Cyclooxygenase-2 Inhibition. *J. Med. Chem.* **1995**, *38*, 3895-3901.

(80) Li, J.J.; Anderson, G.D.; Burton, E.G.; Cogburn, J.N.; Collins, J.T.; Garland, D.J.; Gregory, S.A.; Huang, H.C.; Isakson, P.C.; Koboldt, C.M.; Logussch, E.W.; Norton, M.B.; Perkins, W.E.; Reinhard, E.J.; Seibert, K.; Veenhuizen, A.W.; Zhang, Y.; Reitz, D.B. 1,2-

Diarylcyclopentenes as Selective Cyclooxygenase-2 Inhibitors and Orally Active Anti-Inflammatory Agents. *J. Med. Chem.* **1995**, *38*, 4570-4578.

(81) Li, C.S.; Black, W.C.; Chan, C.C.; Ford-Hutchinson, A.W.; Gauthier, J.Y.; Gordon, R.; Guay, D.; Kargman, S.; Lau, C.K.; Mancini, J.; Ouimet, N.; Roy, P.; Vickers, P.; Wong, E.; Yuong, R.N.; Zambony, R.; Prasit, P. Cyclooxygenase-2 Inhibitors. Synthesis and Pharmacological Activities of 5-Methanesulfonamido-1-Indanone Derivatives. *J. Med. Chem.* **1995**, *38*, 4897-4905.

(82) Khanna, I.K.; Weier, R.M.; Yu, Y.; Collins, P.W.; Miyashiro, J.M.; Koboldt, C.M.; Veenhuizen, A.W.; Currie, J.L.; Seibert, K.; Isakson, P.C. 1,2-Diarylpyrroles as Potent and Selective Inhibitors of Cyclooxygenase-2. *J. Med. Chem.* **1997**, *40*, 1619-1633.

(83) Song, Y.; Connor, D.T.; Doubleday, R.; Sorenson, R.J.; Sercel, A.D.; Unangst, P.C.; Roth, B.D.; Gilbertsen, R.B.; Chan, K.; Schrier, D.J.; Guglietta, A.; Bornemeier, D.A.; Dyer, R.D. Synthesis, Structure-Activity Relationships, and In Vivo Evaluations of Substituted Di-Tert-Butylphenols as a Novel Class of Potent, Selective, and Orally Active Cyclooxygenase-2 Inhibitors. 1. Thiazolone and Oxazolone Series. *J. Med. Chem.* **1999**, *42*, 1151-1160.

(84) Song, Y.; Connor, D.T.; Sercel, A.D.; Sorenson, R.J.; Doubleday, R.; Unangst, P.C.; Roth, B.D.; Beylin, V.G.; Gilbertsen, R.B.; Chan, K.; Schrier, D.J.; Guglietta, A.; Bornemeier, D.A.; Dyer, R.D. Synthesis, Structure-Activity Relationships, and In Vivo Evaluations of Substituted Di-Tert-Butylphenols as a Novel Class of Potent, Selective, and Orally Active Cyclooxygenase-2 Inhibitors. 2. 1,3,4- and 1,2,4-Thiadiazole Series. *J. Med. Chem.* **1999**, *42*, 1161-1169.

(85) Kalgutkar, A.S.; Marnett, A.B.; Crews, B.C.; Remmel, R.P.; Marnett, L.J. Ester and Amide Derivatives of the Nonsteroidal Antiinflammatory Drug, Indomethacin, as Selective Cyclooxygenase-2 Inhibitors. *J. Med. Chem.* **2000**, *43*, 2860-2870.

(86) Khanna, I.K.; Yu, Y.; Huff, R.M.; Weier, R.M.; Xu, X.; Koszyk, F.J.; Collins, P.W.; Cogburn, J.N.; Isakson, P.C.; Koboldt, C.M.; Masferrer, J.L.; Perkins, W.E.; Seibert, K.; Veenhuizen, A.W.; Yuan, J.; Yang, D.C.; Zhang, Y.Y. Selective Cyclooxygenase-2 Inhibitors: Heteroaryl Modified 1,2-Diarylimidazoles are Potent, Orally Active Antiinflammatory Agents. *J. Med. Chem.* **2000**, *43*, 3168-3185.

(87) Chavatte, P.; Yous, S.; Marot, C.; Baurin, N.; Lesieur, D. Three-Dimensional Quantitative Structure-Activity Relationships of Cyclo-Oxygenase-2 (COX-2) Inhibitors: a Comparative Molecular Field Analysis. *J. Med. Chem.* **2001**, *44*, 3223-3230.

(88) Caturla, F.; Jiménez, J.M.; Godessart, N.; Amat, M.; Cárdenas, A.; Soca, L.; Beleta, J.; Ryder, H.; Crespo, M.I. Synthesis and Biological Evaluation of 2-Phenylpyran-4-Ones: a New Class of Orally Active Cyclooxygenase-2 Inhibitors. *J. Med. Chem.* **2004**, *47*, 3874-3886.

(89) Pommery, N.; Taverne, T.; Telliez, A.; Goossens, L.; Charlier, C.; Pommery, J.; Goossens, J.F.; Houssin, R.; Durant, F.; Hénichart, J.P. New COX-2/5-LOX Inhibitors: Apoptosis-Inducing Agents Potentially Useful in Prostate Cancer Chemotherapy. *J. Med. Chem.* **2004**, *47*, 6195-6206.

(90) Rao, P.N.; Chen, Q.H.; Knaus, E.E. Synthesis and Structure-Activity Relationship Studies of 1,3-Diarylprop-2-Yn-1-Ones: Dual Inhibitors of Cyclooxygenases and Lipoxygenases. *J. Med. Chem.* **2006**, *49*, 1668–1683.

(91) Kontogiorgis, C.A.; Hadjipavlou-Litina, D.J. Synthesis and Antiinflammatory Activity of Coumarin Derivatives. *J. Med. Chem.* **2005**, *48,* 6400-6408.

(92) Rao, P.N.; Chen, Q.H.; Knaus, E.E. Synthesis and Biological Evaluation of 1,3-Diphenylprop-2-Yn-1-Ones as Dual Inhibitors of Cyclooxygenases and Lipoxygenases. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 4842-4845.

Table of Contents use only