Simplified Molecular Input Line Entry System-based descriptors in QSAR modelling for HIV-protease inhibitors

Md Ataul Islam, Tahir S. Pillay[*, #]

Department of Chemical Pathology, Faculty of Health Sciences, University of Pretoria and National Health Laboratory Service Tshwane Academic Division, Pretoria, South Africa.

[#]*Division of Chemical Pathology, Faculty of Health Sciences, University of Cape Town South Africa*

[*]Correspondence should be addressed to T.S. Pillay, Department of Chemical Pathology, Faculty of Health Sciences, University of Pretoria, Private Bag X323, Arcadia, Pretoria, 0007
Email: tspillay@gmail.com
Phone: +27-123192114
Fax: +27-123283600

Highlights

• QSAR models were developed on HIV-protease inhibitors.
• Descriptors were generated from SMILES-based structures.
• Monte Carlo algorithm was used to correlate the biological activity with descriptors.
• Models were developed with- and without considering influence of inhibitory activity in chemical structures.
• Presence of rings in the molecules was found to be important for inhibition of HIV-protease.

Abstract

Simplified Molecular Input Line Entry System (SMILES) descriptor based quantitative structure-activity relationship (QSAR) study was performed on a set of HIV-protease inhibitors to explore the structural functionalities for inhibition of the HIV-protease. For this purpose a set of HIV-inhibitors was collected from the literature along with their inhibitory constants. Monte Carlo optimization-based CORAL software was used for QSAR model development. Firstly, the dataset was divided into three random splits and secondly each split was divided into training, calibration, test and validation sets. Training set was used for model development whereas the rest of the sets was used to assess the quality of the developed models. QSAR models were developed with and without considering the influence of cyclic rings towards the inhibitory activity. Statistical quality of QSAR models developed from all splits were very good and fulfilled the criteria. The values of R^2, Q^2, s, R^2_{pred} and r^2_m explained that selected models are robust in nature and efficient enough to predict the inhibitory activity of the molecules outside of the training set. Statistical parameters also suggested that the presence of cyclic rings have a crucial impact on inhibitory activity. The

molecular fragments were found to be important for the increase or decrease of the inhibitory activity which explained that models have mechanistic interpretation. This ligand-based QSAR study can provide clear directions to design and modulate potential HIV-protease inhibitors.

Keywords: HIV-protease; Descriptors; QSAR; Monte Carlo method, SMILES; CORAL software

1. Introduction

The human immunodeficiency virus (HIV-1) is a retrovirus which affects the immune system of the body leaving the victim vulnerable to opportunistic infections, malignancies and neurological disorder with progression of the HIV infection. The acquired immunodeficiency syndrome (AIDS) is the most advanced stage of HIV infection, an epidemic disease with significant health challenges worldwide. Since HIV was identified more than three decades ago, it has spread to almost all parts of the world[1, 2]. Owing to its rapid spread it has become a serious global threat and to date, there is no curative treatment for this fatal disease. According to the World Health Organization (WHO), a total of 37.20 million people are living with AIDS and 1.20 million people died in 2014 alone. The HIV-1[1, 3-5] virus shows extensive and dynamic genetic diversity[6] which has implications for the understanding of viral transmission, pathogenesis and diagnosis, and strongly influences strategies for vaccine development. The HIV-1 protease receptor (HIV-1 PR) is a key therapeutic target for the development of anti-HIV inhibitors[7] for the treatment of AIDS as it plays an important role in the maturation and replication of the virus. HIV-1 PR is comprised of a homodimeric C-2 symmetric structure and one catalytic aspartic residue contributed by each monomer along with threonine and glycine residues which are flexible and a flap that favours the binding of substrate and inhibitors. Highly active antiretroviral therapy (HAART), protease inhibitors (PIs) and reverse-transcriptase inhibitors are widely used for the treatment of HIV -infection worldwide and have resulted in the extraordinary success of HIV/AIDS chemotherapy[8-11]. Rapid development of drug-resistant HIV-1 variants and transmission of these resistant viral strains along with the adverse side effects of currently used HIV-1 PIs, remain critical factors which limit the clinical success of HAART[12-14]. Several research groups globally have developed HIV-1 protease inhibitors which show excellent antiviral profiles[15-21]. A few clinically approved HIV-1 protease inhibitors (atazanavir, indivanir, nelfinavir and sequinavir) are available for treatment of

HIV/AIDS but these are very peptide-like and have poor bio-availability. Consequently, there is a demand to identify new potential PIs with enhanced activity against drug resistant variants and exceptional pharmacokinetic and safety profiles.

The primary objective of the present work was to establish the relationships between the chemical structure and corresponding biological activity through statistical methods in terms of QSAR modelling. QSARs are statistically validated and mathematical relationships between experimental or calculated properties obtained from molecular structures with biological activities. The experimental or calculated properties obtained from chemical structures are the descriptors which characterize specific information of the molecule being studied. The QSAR models provide insights into the essential structural information of the ligands or inhibitors which contribute to biological activity[22]. Descriptors are generally categorised as physico-chemical, topological and electronic, geometric and structural, and simple indicatorparameters. Further, descriptors can also be divided based on dimensionality (0D, 1D, 2D or 3D)[23]. Geometry-based descriptors are generally difficult to calculate and require high computational costs and long computational calculation time. Consequently the conformation-independent descriptors such as 0D, 1D and 2D QSAR/QSPR models based on the constitutional and topological molecular features of compounds have been developed as an alternate approach[24, 25]. In general the QSAR models are developed on molecular graph based descriptors[26-28] but the simplified molecular input-line entry system (SMILES) representation can also be considered for the molecular structure[29-31] which can further be used for molecular descriptor calculations followed by development of QSAR models. The descriptors based on SMILES notation depend both on the molecular structure and the property under analysis irrespective of details from the 3D-molecular geometry[22]. The descriptors generated from SMILES-based molecular structure and further used for development of QSAR models is an attractive direction of research work in the field of the QSAR theory and applications[32-34]. Several research groups have already proven the importance of the methodology, which was capable of developing models with a similar or improved quality to the ones built with descriptors containing thousands of 0D–3D descriptors[35-41].

Therefore the current research work was considered to build QSAR models based on the SMILES notation optimal descriptors using the Monte Carlo method for HIV-protease inhibitors followed by analysis and explanation of involvement of molecular fragments responsible for the inhibition of HIV-protease.

2 Materials and methods

2.1 Dataset

In order to develop QSAR models a set of 129 HIV protease inhibitors[42-44] with experimental inhibitory constant (K_i) were collected from literature. The fluorescence resonance energy transfer (FRET) method had been used to determine experimental K_i values by the same group of authors. The molecules of the dataset have a wide range of K_i, from 0.0008 to 237.8 nM. All datasets are provided in the Supplementary file (Tables S1 – S3) with their inhibitory activity. The SMILES format of whole dataset was generated for input purposes. The *Ki* values were converted into pK_i as $pK_i = \log[(1/K_i)\text{x}1000]$ and considered as the endpoint of the QSAR. The whole dataset was randomly divided into three splits (Splits 1 to 3) and further each split was arbitrarily distributed into training, calibration, test and validation sets. Each of the set has specific role in QSAR formulation. The training set was used for model development and calibration and test sets were used to check the predictive ability of developed model. The validation set was used for final estimation of the model using those compounds that were invisible during model formulationi.e. no information from the validation set was involved in the model generation.

2.2 Optimal descriptors

The molecular structures were converted into SMILES format which is one of the convenient representations of chemical compounds. The SMILES format can be used to select optimal molecular descriptors which are mathematical functions of so-called correlation weights (CW) that is "Descriptors of Correlation Weights" (DCW). The Monte Carlo algorithm was used to calculate the DCW. Firstly, the DCW were calculated without considering the influence of the cyclic rings as per following equation (1).

$$DCW_1(SMILES, T, N_{epoch}) = \alpha\sum CW(S_k) + \beta\sum CW(SS_k) + \gamma\sum CW(SSS_k) + \text{x}\cdot CW(NOSP) + \text{y}\cdot CW(HALO) + z\cdot CW(BOND) + t\cdot CW(PAIR) \quad (1)$$

Where, T is the threshold which defines as coefficient for classifying various molecular features extracted from SMILES into two classes: a) active, in which CW is involved in the modelling process and b) rare, in which CW is not involved in the modelling process. The N_{epoch} is the number of epochs used in the Monte Carlo optimization that provides the best statistical results of the calibration set. The S_k represents one or two symbols from SMILES which cannot be inspected separately. On combination of two or three S_k the SS_k and SSS_k are

formed respectively. NOSP, HALO, BOND and PAIR are different descriptors based on presence of absence of different chemical elements and bonds. NOSP refers to nitrogen, oxygen, sulphur and phosphorus; HALO indicates halogen atoms such as fluorine, chlorine and bromine; BOND indicates double (=), triple (#) or stereochemical bonds (@ or @@); and PAIR refers to the probable grouping of pair atoms and/or SMILES attributes (for example double, triple, and stereochemical bonds) that takes place in the structure together. α, β, γ, x, y and t are discrete coefficient with values 0 and 1. Details of the above descriptors with example are explained by Worachartcheewan et al.[39].

In the case of the second version of optimal descriptors calculation, in addition to the equation (1) the influence of cyclic rings on inhibitory activity were also considered and the equation is given below (2).

$$DCW_2(SMILES,T,N_{epoch}) = \alpha\sum CW(S_k) + \beta\sum CW(SS_k) + \gamma\sum CW(SSS_k) + \mathrm{x}\cdot CW(NOSP) + \mathrm{y}\cdot CW(HALO) + z\cdot CW(BOND) + t\cdot CW(PAIR) CW(\mathrm{C3}) + CW(C4) + CW(C5) + CW(C6) + CW(C7) \quad (2)$$

Where, C3, C4, C5, C6 and C7 are three-membered cycles, four-membered cycles, five-membered cycles, six-membered cycles and seven-membered cycles. Details of such descriptors are explained elsewhere [45]. The purpose of optimization is to obtain the maximal correlation coefficient between optimal descriptors and the end point of the training set.

Monte Carlo method was used to calculate the CW which gives the best statistical results for the test set. In order to search the preferable threshold (T*) and preferable number of epochs (N*), range of T and N_{epoch} were selected from 1 to 10 and 1 to 20 respectively. The best (N*, T*) were selected from each split based on best statistical results. The preferable statistics of the calibration set makes it possible to calculate the endpoint value using numerical values of correlation weights from the training set as follows:

$$\text{Endpoint} = C_0 + C_1 \times \text{DCW(SMILES, T, N}_{\text{epoch}}) \quad (3)$$

Where, endpoint is the inhibitory constant and, C_0 and C_1 are constant. The external validation set which was not involved for model building was used to verify the predictive potential of the model.

2.3 Validation

Validation of any statistical model is an essential and crucial step to verify the predictive ability and reliability of the model. Any QSAR model can be validated by means of a) internal validation using training set compounds; b) external validation using test compounds; and c) **Y**-scrambling or randomization of data. Several studies[34, 35, 37, 38, 41] used these validation methodologies on SMILES notation optimal descriptor based QSAR models. In this respect the cross-validated correlation coefficient (Q^2) and error of estimation (s) were calculated based on predicted activity of training compounds. High Q^2(>0.5) and low s explained better predictive ability of the model[46]. Furthermore, in order to confirm the good predictive ability of the training set compounds, the modified r^2 ($r^2_{m(LOO)}$) developed by Roy et al.[47, 48] the $r^2_{m(LOO)}$ was calculated which is the measure of the degree of deviation of the predicted activity from the observed ones. It was reported that model may be considered with $r^2_{m(LOO)} > 0.5$. For calibration, test and external validation set compounds the R^2_{pred} (correlation coefficient) was also calculate to judge the productiveness of developed model. It is reported that models with $R^2_{pred} \geq 0.5$ can be considered for further evaluation. The value of R^2_{pred} concerns the mean experimental activity of the training set compounds. Consequently good statistical value of R^2_{pred} might be found for molecules with a varied range of activity values, but this does not confirm that the predicted activity values are well correlated with observed activity. However, if a high correlation is preserved, there might be a substantial numerical difference between the two values. In order to better indicate the predictive ability of the model, modified r^2 [$r^2_{m(test)}$][49, 50] values were calculated (threshold value = 0.5).In order to verify the chance correlation **Y**-scrambling described by Ojha and Roy[51] was also performed in which ten probes of calculation were carried out. In one probe of calculation, X and Y represent the vectors of experiment and the vector of prediction. First of all exchange of random N1 and random N2 from row X (Y is not modified) were performed thousand times. Further, from above probes the $R^2_{(X,Y)}$ was calculated and represented as R^2_r. The ${}^C R^2_p$ was finally calculated according to the equation (3).

$$ {}^C R^2_p = R \times (R^2 - R^2_r)^{1/2} \qquad (3) $$

Where R^2 and R^2_r were utilized from the non-randomized and randomized model respectively. For acceptance of QSAR model the threshold value of ${}^C R^2_p$ should be greater than 0.5.

3 Results and discussion

In the current research work, HIV-protease inhibition constant activity (K_i) was used as a function of molecular features extracted from SMILES-based attributes. A set of 129 HIV-protease inhibitors were modelled in online freely available CORAL software (http://www.insilico.eu/coral/) which is based on the Monte Carlo optimization technique. In this technique molecular description in the form of SMILES was used to develop well predictive and statistically validated QSAR models and further QSAR models were used to predict the endpoint data that is the pK_i values. Two approaches were considered to develop QSAR models viz. with and without considering influence of cyclic rings of the molecules in inhibitory activity. In order to identify the optimal T and N_{epoch} (T* and N^*_{epoch}) for Monte Carlo optimization the training set molecules were used to develop models using threshold values in the range of 1 to 10 and the number of epochs ranging from 1 to 20 (Figure 1) for each splits. Finally the T* and N* were chosen based on best statistical results and further used for constructing the QSAR models. In case of without considering influence of rings on activity, the best (T*, N*) for splits 1, 2 and 3 were found to be (1,17), (1,12) and (1,19) respectively. On the other hand, the optimal (T*, N*) were (1,19), (1,18) and (1, 12) for splits 1, 2 and 3 respectively by considering the influence of rings on the inhibitory activity. The whole data set was divided into four groups i.e. training, calibration, test and validation sets in each split. The training set was used for model generation while rest of the sets were used for evaluating the predictive performance of generated models.

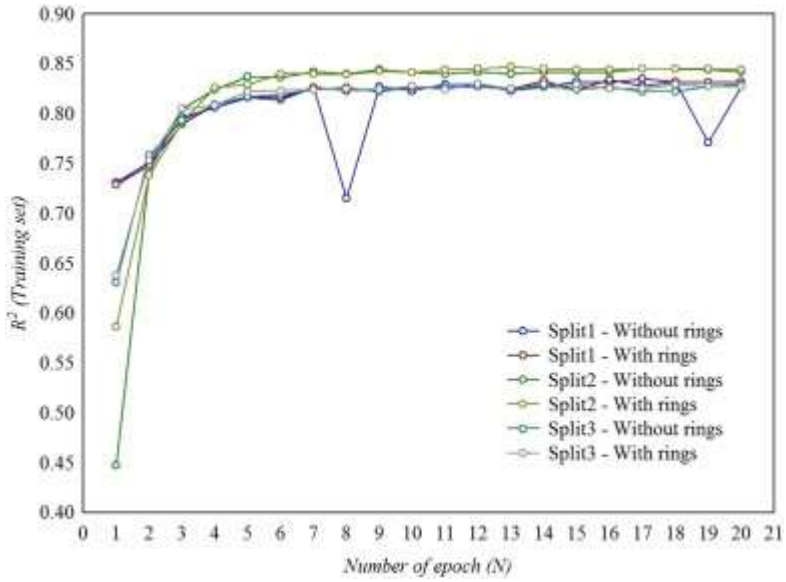


Figure 1: Number of epoch (N) vs. correlation coefficient.

Split 1:

Without considering influence of various cyclic rings:

$$pK_i = 0.0001(\pm\ 0.0255) + 0.0580(\pm\ 0.0005) \times DCW(1,17) \qquad \text{(Model } \mathbf{1}\text{)}$$

Training set: n = 70; R^2 = 0.837; s = 0.448; F = 331; Q^2 = 0.831; s_p = 0.455; r^2_m = 0.830; Δr^2_m = 0.149

Calibration set: n = 20; R^2 = 0.952; s =0.312; F = 332; R^2_{pred} = 0.948; r^2_m = 0.949; Δr^2_m = 0.040

Test set: n = 21; R^2 = 0.799; s = 0.561, F = 72; R^2_{pred} = 0.769; r^2_m = 0.686; Δr^2_m = 0.043

External set: n = 18; R^2 = 0.794; s =0.677, F = 82; R^2_{pred} = 0.781; r^2_m = 0.791; Δr^2_m = 0.043

Considering influence of various cyclic rings:

$$pK_i = -0.0011(\pm\ 0.0267) + 0.0461(\pm 0.0004) \times DCW(1,19) \qquad \text{(Model } \mathbf{2}\text{)}$$

Training set: n = 70; R^2 = 0.811; s = 0.482, F = 346; Q^2 = 0.807, s_p = 0.447; r^2_m = 0.836; Δr^2_m = 0.144

Calibration set: n = 20; R^2 = 0.950; s = 0.315; F = 382; R^2_{pred} = 0.943; r^2_m = 0.953; Δr^2_m = 0.036

Test set: n = 21; R^2 = 0.806; s = 0.553, F = 85; R^2_{pred} = 0.800; r^2_m = 0.690; Δr^2_m = 0.103

External set: n = 18; R^2 = 0.752; s = 0.738, F = 51; R^2_{pred} = 0.729; r^2_m = 0.630; Δr^2_m = 0.080

Split 2:

Without considering influence of various cyclic rings:

$$pK_i = -0.0035(\pm 0.0196) + 0.0485(\pm 0.0003) \times DCW(1,12) \qquad \text{(Model } \mathbf{3}\text{)}$$

Training set: n = 72; R^2 = 0.843; s = 0.454, F = 376; Q^2 = 0.830; s_p = 0.451; r^2_m = 0.843; Δr^2_m = 0.138

Calibration set: n = 20; R^2 = 0.947; s = 0.304; F = 324; R^2_{pred} = 0.943; r^2_m = 0.946; Δr^2_m = 0.046

Test set: n = 19; R^2 =0.781; s = 0.676; F = 61; R^2_{pred} = 0.778; r^2_m = 0.694; Δr^2_m = 0.005

External set: n = 18; R^2 = 0.740; s = 0.644, F = 45; R^2_{pred} = 0.736; r^2_m = 0.531; Δr^2_m = 0.209

Considering influence of various cyclic rings:

$$pK_i = 0.0053(\pm 0.0196) + 0.0652(\pm 0.0004) \times DCW(1,18) \quad \text{(Model } \mathbf{4}\text{)}$$

Training set: n = 72; R^2 = 0.841; s = 0.457; F = 370; Q^2 = 0.828, s_p = 0.454; r^2_m = 0.841; Δr^2_m = 0.130

Calibration set: n = 20; R^2 = 0.941; s = 0.323; F = 286; R^2_{pred} = 0.937; r^2_m = 0.939; Δr^2_m = 0.107

Test set: n = 19; R^2 = 0.784; s = 0.617; F = 89; R^2_{pred} = 0.779; r^2_m = 0.827; Δr^2_m = 0.107

External set: n = 18; R^2 =0.753; s =0.624; F = 45; R^2_{pred} = 0.751; r^2_m = 0.538; Δr^2_m = 0.203

Split 3:

Without considering influence of various cyclic rings:

$$pK_i = -0.00001(\pm 0.0204) + 0.0463(\pm 0.0003) \times DCW(1,19) \quad \text{(Model } \mathbf{5}\text{)}$$

Training set: n = 75; R^2 = 0.829; s = 0.491; F = 355; Q^2 = 0.826; s_p = 0.487; r^2_m = 0.829; Δr^2_m = 0.148

Calibration set: n = 20; R^2 = 0.945; s = 0.335; F = 311; R^2_{pred} = 0.943; r^2_m = 0.945; Δr^2_m = 0.048

Test set: n = 19; R^2 = 0.804; s = 0.592; F = 70; R^2_{pred} = 0.811; r^2_m = 0.614; Δr^2_m = 0.168

External set: n = 15; R^2 =0.881; s =0.381; F = 96; R^2_{pred} = 0.907; r^2_m = 0.791; Δr^2_m = 0.075

Considering influence of various cyclic rings:

$$pK_i = -0.00003(\pm 0.0211) + 0.0393(\pm 0.0002) \times DCW(1,12) \quad \text{(Model } \mathbf{6}\text{)}$$

Training set: n = 75; R^2 = 0.830; s = 0.489; F = 357; Q^2 = 0.827, s_p = 0.486; r^2_m = 0.830; Δr^2_m = 0.148

Calibration set: n = 20; R^2 = 0.945; s = 0.335; F = 311; R^2_{pred} = 0.944; r^2_m = 0.945; Δr^2_m = 0.048

Test set: n = 19; R^2 = 0.773; s = 0.615; F = 58; R^2_{pred} = 0.796; r^2_m = 0.602; Δr^2_m = 0.159

External set: n = 15; R^2 = 0.915; s = 0.311; F = 140; R^2_{pred} = 0.938; r^2_m = 0.883; Δr^2_m = 0.012

Models **1**, **3** and **5** were developed from splits 1, 2 and 3 respectively without considering any influence of rings on the inhibitory activity of the training set molecules. On the other side, Models **2**, **3** and **4** were generated with involvement of the influence of rings on the inhibitory activity of training set molecules of splits 1, 2 and 3 respectively. All models were developed

using best Monte Carlo optimization runs and also according to Organisation for Economic and Co-operation (OECD) principles[52].

Table 1: Statistical parameters of training, calibration, test and validation sets of Models **1** to **6**

Model	set	*n*	R^2	*s*	*F*	Q^2	${}^{C}R^2_p$	R^2_{pred}	$R^2 - Q^2$	r^2_m	Δr^2_m
Model **1**	Training	70	0.837	0.448	331	0.831	0.823	-	0.006	0.830	0.149
	Calibration	20	0.952	0.312	332	-	0.907	0.948	-	0.949	0.040
	Test	21	0.799	0.561	72	-	0.782	0.769	-	0.686	0.043
	Validation	18	0.794	0.677	82	-	-	0.781	-	0.791	0.043
Model **2**	Training	70	0.811	0.482	346	0.807	0.824	-	0.004	0.836	0.144
	Calibration	20	0.950	0.315	382	-	0.937	0.943	-	0.953	0.036
	Test	21	0.806	0.553	85	-	0.796	0.800	-	0.690	0.103
	Validation	18	0.752	0.738	51	-	-	0.729	-	0.630	0.080
Model **3**	Training	72	0.843	0.454	376	0.830	0.830	-	0.013	0.843	0.138
	Calibration	20	0.947	0.304	324	-	0.903	0.943	-	0.946	0.046
	Test	19	0.781	0.676	61	-	0.766	0.778	-	0.694	0.005
	Validation	18	0.740	0.644	45	-	-	0.736	-	0.531	0.209
Model **4**	Training	72	0.841	0.457	370	0.828	0.834	-	0.013	0.841	0.130
	Calibration	20	0.941	0.323	286	-	0.919	0.937	-	0.939	0.107
	Test	19	0.784	0.617	89	-	0.811	0.779	-	0.827	0.107
	Validation	18	0.753	0.624	45	-	-	0.751	-	0.538	0.203
Model **5**	Training	75	0.829	0.491	355	0.826	0.827	-	0.003	0.829	0.148
	Calibration	20	0.945	0.335	311	-	0.923	0.943	-	0.945	0.048
	Test	19	0.804	0.592	70	-	0.787	0.811	-	0.614	0.168
	Validation	15	0.881	0.381	96	-	-	0.907	-	0.791	0.075
Model **6**	Training	75	0.830	0.489	357	0.827	0.825	-	0.003	0.830	0.148
	Calibration	20	0.945	0.335	311	-	0.933	0.944	-	0.945	0.048
	Test	19	0.804	0.592	70	-	0.748	0.811	-	0.614	0.168
	Validation	15	0.881	0.381	96	-	-	0.907	-	0.791	0.075

The statistical results of Models **1** to **6** are depicted in Table 1. Values of Table 1 clearly indicated that all models are statistically reliable and fulfilled the criterion explained by Tropsha et al.[53] and Ojha et al.[48] for a predictive QSAR model such as $R^2 > 0.6$, $Q^2 > 0.5$, ${}^{C}R^2_p > 0.5$, $(R^2 - Q^2) < 0.2$, $R^2_{pred} > 0.5$, $r^2_m > 0.5$ and $\Delta r^2_m < 0.1$. Experimental and predicted activities of all sets are delineated in Figure 2 and supplementary materials (Tables S1 – S3). The high R^2, Q^2 and R^2_{pred} values of all models explained close prediction of inhibitory activity to the experimental activity which supports the robustness of the models. The reliability of the developed models were also assessed by the difference between R^2 and Q^2 (R^2 - Q^2) which is a metric that accounts for the fraction of **Y**-data explained by accumulated chance correlations. It is reported that values of the metric less than 0.2 implies that model

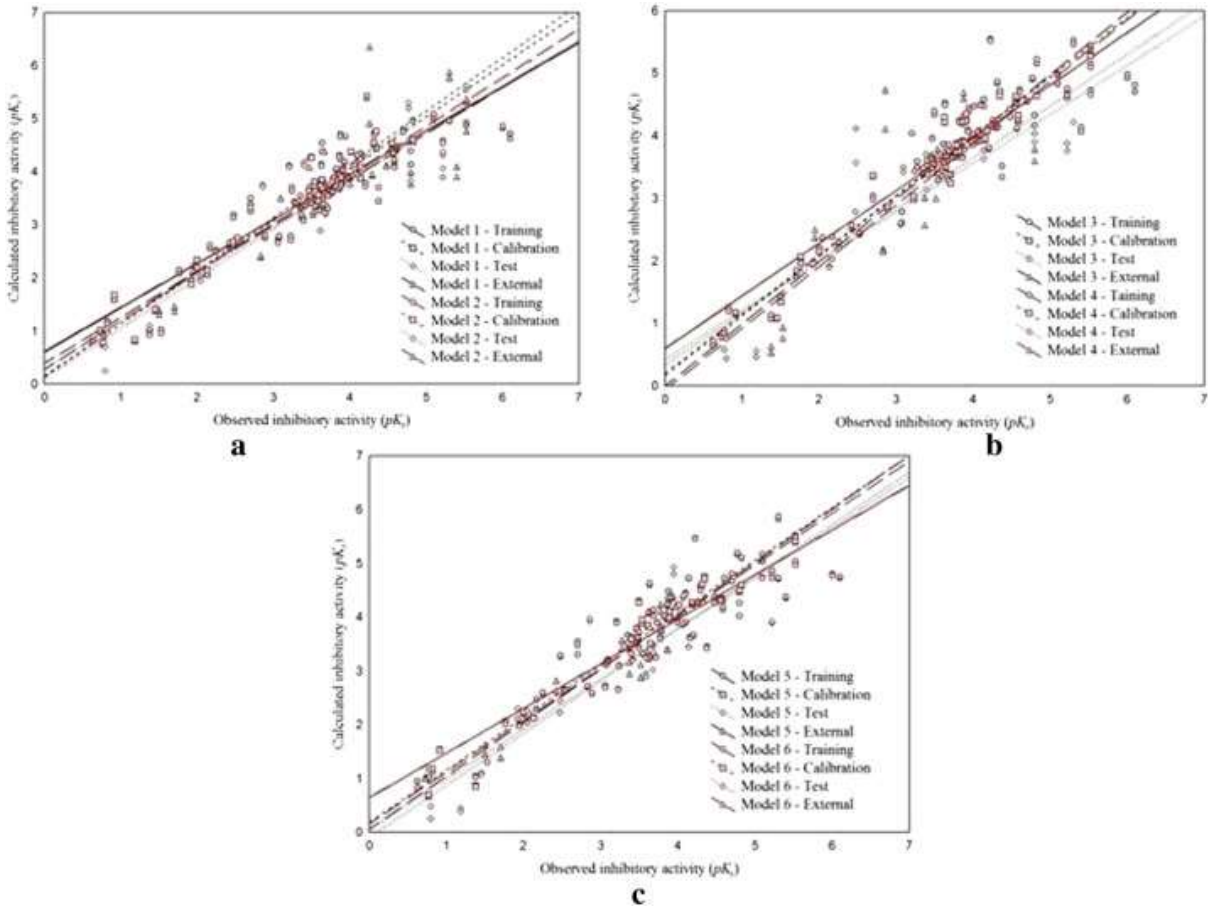


Figure 2: Observed and predicted activities as per Models **1** – **6**. a) Split 1; b) Split 2 and c) Split 3.

developed not by chance[39]. In case of Models **1** to **6** the values of the metric lie between 0.130 to 0.003 which is below the criteria mentioned above, therefore models were developed not by chance. Similarly chance correlations were evaluated by the **Y**-scrambling technique. In the method the Y value (i.e. pK_i) was randomized and reordered with respect to its associated descriptors. Correspondingly the Y-scrambling was performed for 1000 trials in

ten separate runs for all three splits. The average value of each run is given in supplementary file (Tables S4 – S6). If the correlation of randomised runs are better than the original then the models may be developed by chance. In the present study it was observed that no randomised run obtained better correlation than the original run in case of Models **1** to **6**. Therefore Models **1** to **6** were developed not by chance and they are robust in nature. The correlation values of original and randomized runs are given in supplementary file (Tables S4 – S6).

Comparison of results (Table 1) between models developed by considering (Models **2**, **4** and **6**) and without considering (Models **1**, **3** and **5**) influence of the cyclic rings of the molecular scaffolds of training sets on inhibitory activity revealed that the rings have considerable impact on inhibition of HIV-protease as statistical results are better in case of calibration, test and external sets for Models **2**, **4** and **6**. In the current research, descriptors used were based on SMILES notation that means ignoring the three dimensional molecular structures of dataset which reduce the computation time as well as computation resources. All statistical parameters satisfied the criteria except Δr^2_m for external validation set in Models **3** and **4**, where it is more than 0.2. It is noted that if r^2_m is more than 0.5 then Δr^2_m should preferably less than 0.2. But it is also reported that this is not a rigid limitation[54]. Radar plots of training set molecules of Models **1** to **6** were plotted and delineated in Figure 3. The plots explore the fitness of experimental and calculated inhibitory activity. Substantial overlapping of experimental and calculated pK_i values in the radar plots support the idea that the models are efficient enough to predict the biological activity of the molecules of training set molecules.

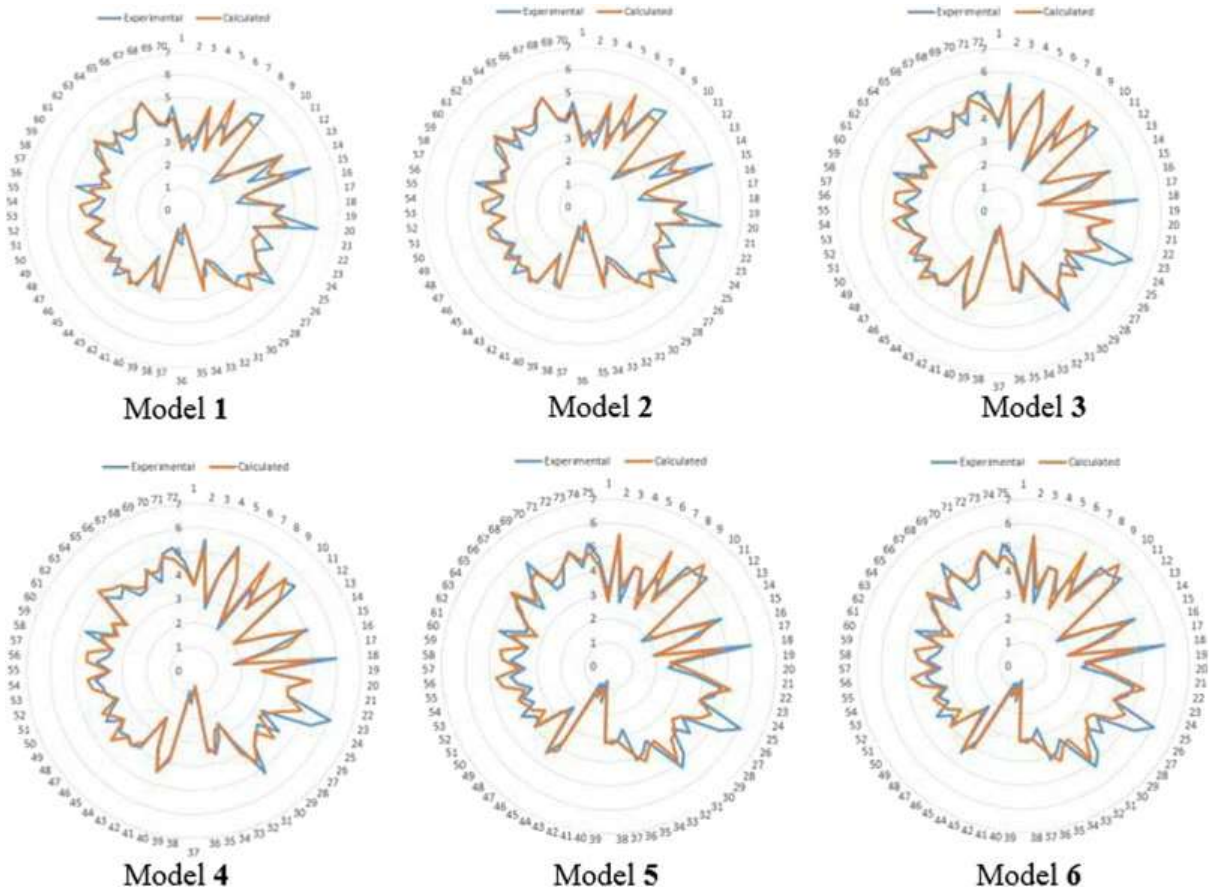


Figure 3: Radar plot showing fitness of predicted and actual activity values of training set of Models **1** to **6**

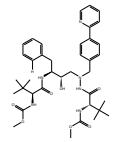
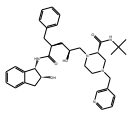
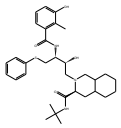
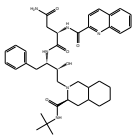
The molecular fragments obtained from SMILES were interpreted to explore chemical information analysed using correlation weights from QSAR models. This analysis was performed on the data samples divided into the following categories: a) molecular features with constant positive values of correlation weights i.e. positive correlation weights in all runs, increased for an endpoint; b) molecular features with negative constant value of correlation weights in all runs decreased for an endpoint; c) unstable molecular features which have both positive and negative correlation weights in several runs. The correlation weights of molecular features greater than '0' implies that molecular features have positive influence on inhibitory activity, if it is less than '0' then molecular features have negative influence on the inhibitory activity. If the correlation weights of molecular features have a mix of greater than and less than '0' values then the impact of molecular features on the inhibitory activity is undefined.

In this manuscript molecule **H1** was considered to explain the impact of molecular features on pK_i. It was observe that smiles attributes '3……' and 'C……1' have positive influence on pK_i in case of Models **3**, **4**, **5** and **6**. The attribute NOSP11100000 (presence of 'N', 'O' and

'S' but absence of 'P') gives negative impact on pK_i in case of Models **2** – **6** but positive for Model **1**. The HALO00000000 (absence of halogen atoms) increases the pK_i in the case of models developed without considering the influence of the rings (Models **1**, **3** and **5**), but decreases the pK_i in the presence of cyclic rings (Models **2**, **4** and **6**). BOND10100000 attributes contribute negative impact on all models except Model **6**. Decrease of pK_i due to presence of oxygen and double bond occurs in all models. Presence of nitrogen atom in case of Models **3** – **6** gives an increased pK_i value while for Models **1** and **2** its involvement decreases the pK_i value. The '++++N---S===' was found to be important for increasing the pK_i in the case of the presence of cyclic rings (Models **2**, **4** and **6**). Another attribute '++++S---B2==' gives negative influence on pK_i in for Models **2** – **5**, but for Models **1** and **6** its impact is positive. Presence of chirality with anti-clockwise ('@') gives negative influence for inhibitory activity for Models **1**, **2**, **4**, and **6**. Influence of chirality with clockwise ('@@') increases the pK_i in all models except Model **2**. From the above discussion it is clear that models developed (Models **1** – **6**) using the HIV-protease inhibitors with considering the inhibitory constant as endpoints have mechanistic interpretations for deducing how the molecular features can be attributed to appropriate atoms and/or molecular fragments.

The pK_i of well-known HIV-protease inhibitors such as atazanavir, indivanir, nelfinavir and sequinavir were calculated as per Models **1** to **6** and given in Table 2. The calculated activity explained that influence of rings have positive impact inhibitory activity as pK_i values were found to be higher in models developed with considering rings (Models **2**, **4** and **6**) in comparison to models developed without considering the influence of rings in inhibitory activity. Therefore the models developed using atomic and molecular fragments using SMILES notation can successfully predict the inhibitory activity of unknown molecules. Therefore the developed models can be used to design new molecular scaffolds of HIV-protease inhibitors with increased or decreased inhibitory activities.

Table 2: Calculated inhibitory activity (*pKi*) of different standard HIV-protease inhibitors based on QSAR models

Molecule	Structure	pK_i					
		Model **1**	Model **2**	Model **3**	Model **4**	Model **5**	Model **6**
Atazanavir		2.324	2.385	2.684	2.873	2.652	2.786
Indivanir		2.358	2.363	2.418	2.571	2.280	2.461
Nelfinavir		2.510	3.171	2.782	2.791	3.195	3.062
Sequinavir		2.524	3.243	3.796	2.650	5.699	3.395

4 Conclusions

The predictive QSAR models for HIV-protease inhibitors were developed with inhibitory constant using online CORAL software. Incorporation of the Monte Carlo optimization technique in CORAL software provides an effective and prominent tool to develop good statistical QSAR models. SMILES-based descriptors were used with and without considering the influence of different rings on the inhibitory activity. The predictive performance of all models were tested. The radar plot also explained the well-fitness of experimental and calculated inhibitory activities of training set molecules. The statistical quality of the models explained that cyclic rings influence the inhibition of HIV-protease. All models have mechanistic interpretations in terms of the importance of molecular features on correlation weights for increases or decreases in the inhibitory activity (pK_i). The developed models can be used to for prediction of inhibitory constants for HIV-protease inhibitors without experiments. The predictive nature of all models was validated according to OECD guidelines. Therefore it can be concluded that the data represented indicates that the suggested approach can be applied to design novel and potential HIV protease inhibitors for the wider community of people with HIV.

Acknowledgement

MA Islam and TS Pillay were funded by the University of Pretoria Vice Chancellor's post-doctoral fellowship and National Research Foundation (NRF), South Africa Innovation post-doctoral fellowship schemes.

5 References

[1] R.C. Gallo, P.S. Sarin, E.P. Gelmann, M. Robert-Guroff, E. Richardson, V.S. Kalyanaraman, D. Mann, G.D. Sidhu, R.E. Stahl, S. Zolla-Pazner, J. Leibowitch, M. Popovic, Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS), Science, 220 (1983) 865-867.
[2] N. Chin'ombe, V. Ruhanya, HIV/AIDS vaccines for Africa: scientific opportunities, challenges and strategies, The Pan African medical journal, 20 (2015) 386.
[3] F. Barre-Sinoussi, J.C. Chermann, F. Rey, M.T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum, L. Montagnier, Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS), Science, 220 (1983) 868-871.
[4] M. Popovic, M.G. Sarngadharan, E. Read, R.C. Gallo, Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS, Science, 224 (1984) 497-500.
[5] M.G. Sarngadharan, M. Popovic, L. Bruch, J. Schupbach, R.C. Gallo, Antibodies reactive with human T-lymphotropic retroviruses (HTLV-III) in the serum of patients with AIDS, Science, 224 (1984) 506-508.
[6] J. Hemelaar, E. Gouws, P.D. Ghys, S. Osmanov, Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004, Aids, 20 (2006) W13-23.
[7] A. Wlodawer, M. Miller, M. Jaskolski, B.K. Sathyanarayana, E. Baldwin, I.T. Weber, L.M. Selk, L. Clawson, J. Schneider, S.B. Kent, Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease, Science, 245 (1989) 616-621.
[8] J.A. Bartlett, R. DeMasi, J. Quinn, C. Moxham, F. Rousseau, Overview of the effectiveness of triple combination therapy in antiretroviral-naive HIV-1 infected adults, Aids, 15 (2001) 1369-1377.
[9] R.M. Gulick, J.W. Mellors, D. Havlir, J.J. Eron, A. Meibohm, J.H. Condra, F.T. Valentine, D. McMahon, C. Gonzalez, L. Jonas, E.A. Emini, J.A. Chodakewitz, R. Isaacs, D.D. Richman, 3-year suppression of HIV viremia with indinavir, zidovudine, and lamivudine, Annals of internal medicine, 133 (2000) 35-39.
[10] F.J. Palella, Jr., K.M. Delaney, A.C. Moorman, M.O. Loveless, J. Fuhrer, G.A. Satten, D.J. Aschman, S.D. Holmberg, Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators, The New England journal of medicine, 338 (1998) 853-860.
[11] R.S. Hogg, K.V. Heath, B. Yip, K.J. Craib, M.V. O'Shaughnessy, M.T. Schechter, J.S. Montaner, Improved survival among HIV-infected individuals following initiation of antiretroviral therapy, JAMA : the journal of the American Medical Association, 279 (1998) 450-454.
[12] L. Waters, M. Nelson, Why do patients fail HIV therapy?, International journal of clinical practice, 61 (2007) 983-990.
[13] J.H. Condra, W.A. Schleif, O.M. Blahy, L.J. Gabryelski, D.J. Graham, J.C. Quintero, A. Rhodes, H.L. Robbins, E. Roth, M. Shivaprakash, et al., In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors, Nature, 374 (1995) 569-571.
[14] F. Clavel, A.J. Hance, HIV drug resistance, The New England journal of medicine, 350 (2004) 1023-1035.
[15] A.G. Tomasselli, R.L. Heinrikson, Targeting the HIV-protease in AIDS therapy: a current clinical perspective, Biochimica et biophysica acta, 1477 (2000) 189-214.
[16] G.S. Reddy, A. Ali, M.N. Nalam, S.G. Anjum, H. Cao, R.S. Nathans, C.A. Schiffer, T.M. Rana, Design and synthesis of HIV-1 protease inhibitors incorporating oxazolidinones as P2/P2' ligands in pseudosymmetric dipeptide isosteres, Journal of medicinal chemistry, 50 (2007) 4316-4328.
[17] A.K. Ghosh, B.D. Chapsal, I.T. Weber, H. Mitsuya, Design of HIV protease inhibitors targeting protein backbone: an effective strategy for combating drug resistance, Accounts of chemical research, 41 (2008) 78-86.
[18] A.M. Wensing, N.M. van Maarseveen, M. Nijhuis, Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance, Antiviral research, 85 (2010) 59-74.
[19] A.K. Ghosh, B.D. Chapsal, A. Baldridge, M.P. Steffey, D.E. Walters, Y. Koh, M. Amano, H. Mitsuya, Design and synthesis of potent HIV-1 protease inhibitors incorporating hexahydrofuropyranol-derived high affinity P(2) ligands: structure-activity studies and biological evaluation, Journal of medicinal chemistry, 54 (2011) 622-634.

[20] A.K. Ganguly, S.S. Alluri, D. Caroccia, D. Biswas, C.H. Wang, E. Kang, Y. Zhang, A.T. McPhail, S.S. Carroll, C. Burlein, V. Munshi, P. Orth, C. Strickland, Design, synthesis, and X-ray crystallographic analysis of a novel class of HIV-1 protease inhibitors, Journal of medicinal chemistry, 54 (2011) 7176-7183.
[21] X. Qiu, G.D. Zhao, L.Q. Tang, Z.P. Liu, Design and synthesis of highly potent HIV-1 protease inhibitors with novel isosorbide-derived P2 ligands, Bioorganic & medicinal chemistry letters, 24 (2014) 2465-2468.
[22] J.V. Zivkovic, N.V. Trutic, J.B. Veselinovic, G.M. Nikolic, A.M. Veselinovic, Monte Carlo method based QSAR modeling of maleimide derivatives as glycogen synthase kinase-3beta inhibitors, Computers in biology and medicine, 64 (2015) 276-282.
[23] K. Roy, S. Kar, R.N. Das, A Primer on QSAR/QSPR Modeling Fundamental Concepts, Springer International Publishing2015.
[24] P.R. Duchowicz, N.C. Comelli, E.V. Ortiz, E.A. Castro, QSAR study for carcinogenicity in a large set of organic compounds, Current drug safety, 7 (2012) 282-288.
[25] A. Talevi, C.L. Bellera, M. Di Ianni, P.R. Duchowicz, L.E. Bruno-Blanch, E.A. Castro, An integrated drug development approach applying topological descriptors, Current computer-aided drug design, 8 (2012) 172-181.
[26] M. Randic, S.C. Basak, A new descriptor for structure-property and structure-activity correlations, Journal of chemical information and computer sciences, 41 (2001) 650-656.
[27] S. Junkes Bda, A.C. Arruda, R.A. Yunes, L.C. Porto, V.E. Heinzen, Semi-empirical topological index: a tool for QSPR/QSAR studies, Journal of molecular modeling, 11 (2005) 128-134.
[28] O. Ivanciuc, Chemical graphs, molecular matrices and topological indices in chemoinformatics and quantitative structure-activity relationships, Current computer-aided drug design, 9 (2013) 153-163.
[29] D. Weininger, SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules, Journal of chemical information and modeling, 28 (1988) 31-36.
[30] D. Weininger, A. Weininger, J.L. Weininger, SMILES. 2. Algorithm for generation of unique SMILES notation Journal of chemical information and modeling, 29 (1989) 97-101.
[31] D. Weininger, SMILES. 3. Depict. Graphical depiction of chemical structures, Journal of chemical information and modeling, 30 (1990) 237-243.
[32] A.A. Toropov, E. Benfenati, SMILES in QSPR/QSAR Modeling: results and perspectives, Current drug discovery technologies, 4 (2007) 77-116.
[33] A.A. Toropov, E. Benfenati, SMILES as an alternative to the graph in QSAR modelling of bee toxicity, Computational biology and chemistry, 31 (2007) 57-60.
[34] A.M. Veselinovic, J.B. Veselinovic, J.V. Zivkovic, G.M. Nikolic, Application of SMILES Notation Based Optimal Descriptors in Drug Discovery and Design, Current topics in medicinal chemistry, 15 (2015) 1768-1779.
[35] A.A. Toropov, A.P. Toropova, R.G. Diaza, E. Benfenati, G. Gini, SMILES-based optimal descriptors: QSAR modeling of estrogen receptor binding affinity by correlation balance, Structural Chemistry, 23 (2012) 529–544.
[36] A.M. Veselinovic, J.B. Milosavljevic, A.A. Toropov, G.M. Nikolic, SMILES-based QSAR models for the calcium channel-antagonistic effect of 1,4-dihydropyridines, Archiv der Pharmazie, 346 (2013) 134-139.
[37] A.M. Veselinovic, J.B. Milosavljevic, A.A. Toropov, G.M. Nikolic, SMILES-based QSAR model for arylpiperazines as high-affinity 5-HT(1A) receptor ligands using CORAL, European journal of pharmaceutical sciences : official journal of the European Federation for Pharmaceutical Sciences, 48 (2013) 532-541.
[38] J.B. Veselinovic, A.A. Toropov, A.P. Toropova, G.M. Nikolic, A.M. Veselinovic, Monte carlo method-based QSAR modeling of penicillins binding to human serum proteins, Archiv der Pharmazie, 348 (2015) 62-67.
[39] A. Worachartcheewan, P. Mandi, V. Prachayasittikul, A.P. Toropova, A.A. Toropov, C. Nantasenamat, Large-scale QSAR study of aromatase inhibitors using SMILES-based descriptors, Chemometrics and Intelligent Laboratory Systems, 138 (2014) 120-126.
[40] Q. Li, X. Ding, H. Si, H. Gao, QSAR model based on SMILES of inhibitory rate of 2, 3-diarylpropenoic acids on AKR1C3, Chemometrics and Intelligent Laboratory Systems, 139 (2014) 120-126.
[41] A.P. Toropova, A.A. Toropov, J.B. Veselinovic, F.N. Miljkovic, A.M. Veselinovic, QSAR models for HEPT derivates as NNRTI inhibitors based on Monte Carlo method, European journal of medicinal chemistry, 77 (2014) 298-305.
[42] A. Ali, G.S. Reddy, H. Cao, S.G. Anjum, M.N. Nalam, C.A. Schiffer, T.M. Rana, Discovery of HIV-1 protease inhibitors with picomolar affinities incorporating N-aryl-oxazolidinone-5-carboxamides as novel P2 ligands, Journal of medicinal chemistry, 49 (2006) 7342-7356.
[43] A. Ali, G.S. Reddy, M.N. Nalam, S.G. Anjum, H. Cao, C.A. Schiffer, T.M. Rana, Structure-based design, synthesis, and structure-activity relationship studies of HIV-1 protease inhibitors incorporating phenyloxazolidinones, Journal of medicinal chemistry, 53 (2010) 7699-7708.

[44] M.K. Parai, D.J. Huggins, H. Cao, M.N. Nalam, A. Ali, C.A. Schiffer, B. Tidor, T.M. Rana, Design, synthesis, and biological and structural evaluations of novel HIV-1 protease inhibitors to combat drug resistance, Journal of medicinal chemistry, 55 (2012) 6328-6341.
[45] M.A. Toropova, A.A. Toropov, I. Raska, Jr., M. Raskova, Searching therapeutic agents for treatment of Alzheimer disease using the Monte Carlo method, Computers in biology and medicine, 64 (2015) 148-154.
[46] H. Kubinyi, F.A. Hamprecht, T. Mietzner, Three-dimensional quantitative similarity-activity relationships (3D QSiAR) from SEAL similarity matrices, Journal of medicinal chemistry, 41 (1998) 2553-2564.
[47] K. Roy, I. Mitra, S. Kar, P.K. Ojha, R.N. Das, H. Kabir, Comparative studies on some metrics for external validation of QSPR models, Journal of chemical information and modeling, 52 (2012) 396-408.
[48] P.K. Ojha, I. Mitra, R.N. Das, K. Roy, Further exploring rm2 metrics for validation of QSPR models, Chemometrics and Intelligent Laboratory Systems, 107 (2011) 194-205.
[49] P. Pratim Roy, S. Paul, I. Mitra, K. Roy, On two novel parameters for validation of predictive QSAR models, Molecules, 14 (2009) 1660-1701.
[50] P.P. Roy, K. Roy, On Some Aspects of Variable Selection for Partial Least Squares Regression Models, Molecular Informatics, 27 (2007) 302-313.
[51] P.K. Ojha, R. Roy, Comparative QSARs for antimalarial endochins: importance of descriptor-thinning and noise reduction prior to feature selection, Chemometrics and Intelligent Laboratory Systems, 109 (2011) 146-161.
[52] OECD, Guidance document on the validation of (quantitative) structure–activity relationships [(Q) SAR] models, ENV/JM/MONO(2007(2), Organisation for Economic Co-operation and Development, Paris, 2007.
[53] A. Tropsha, P. Gramatica, V.K. Gombar, The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models, 22 (2003) 69–77., QSAR & Combinatorial Science, 22 (2003) 69-77.
[54] P.K. Ojha, I. Mitra, R.N. Das, K. Roy, Further exploring r m 2 metrics for validation of QSPR models, Chemometrics and Intelligent Laboratory Systems, 107 (2011) 194-205.

Supplementary materials

Table S1: Observed, predicted activity and DCW, and distribution into the training (+), calibration (-), test (#) and validation (*) sets of split 1 (Models **1** and **2**)

Mol. No.		[1]Exp.	Model **1**		Model **2**		SMILES
			[2]Calc.	DCW	[2]Calc.	DCW	
H1	+	3.059	2.721	46.882	2.656	57.677	O=C(O[C@@](C)([H])CC(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H5	+	3.371	3.126	53.864	3.197	69.425	O=C(O[C@@](C)([H])CCC(C)=O)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H6	+	2.697	3.256	56.104	3.323	72.156	O=C(OCCNC(=O)OC(C)(C)C)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H8	+	4.319	4.736	81.597	4.693	101.884	O=C(OCCN1CCNC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C(C=C3)OC
H12	+	3.229	2.794	48.131	2.714	58.928	O=C(O[C@@](C)([H])CC(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H14	+	4.222	5.388	92.825	5.440	118.118	O=C(OCCNC(=O)C([Cl])([Cl])[H])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H16	+	3.638	3.329	57.353	3.381	73.408	O=C(OCCNC(=O)OC(C)(C)C)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H18	+	5.301	4.954	85.363	4.957	107.624	O=C(OCCN1CCNC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H21	+	5.523	4.897	84.375	4.934	107.127	O=C(OC[C@]1([H])NC(=O)CC1)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H22	+	1.763	2.155	37.133	2.079	45.158	O=C(O[C@@](C)([H])CC(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H24	+	2.167	2.560	44.115	2.620	56.906	O=C(O[C@@](C)([H])CCC(C)=O)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H26	+	3.391	3.114	53.648	3.080	66.881	O=C(OC[C@](C)([H])C(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H28	+	5.097	4.984	85.874	5.036	109.341	O=C(OCCN1CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C(C=C3)OC
H30	+	3.205	4.113	70.860	4.093	88.867	O=C(OC[C@]1([H])NC(=O)CC1)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C(C=C3)OC

H33	+	6.000	4.822	83.077	4.864	105.598	O=C(OCCNC(=O)C([Cl])([Cl])[H])N[C@@](H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H35	+	2.889	2.763	47.604	2.804	60.888	O=C(OCCNC(=O)OC(C)(C)C)N[C@@](H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H36	+	2.476	3.186	54.898	3.138	68.133	O=C(OC[C@](C)([H])C(=O)OC)N[C@@](H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H40	+	4.569	4.331	74.626	4.357	94.607	O=C(OC[C@]1([H])NC(=O)CC1)N[C@@](H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H42	+	4.081	4.054	69.851	3.980	86.416	O=C1O[C@@](H])(CN1C=2C=C/C=C(/F])C=2)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H45	+	6.097	4.721	81.346	4.623	100.384	O=C1O[C@@](H])(CN1C=2C=CC=C(C=2)C(C)=O)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H48	+	3.276	3.432	59.126	3.414	74.135	O=C1O[C@@](H])(CN1C=2C=CC=CC=2)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H49	+	3.770	3.605	62.110	3.624	78.683	O=C1O[C@@](H])(CN1C=2C=C/C=C(/F])C=2)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H51	+	4.377	3.442	59.311	3.704	80.436	O=C1O[C@@](H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H53	+	3.735	3.649	62.862	3.553	77.142	O=C1O[C@@](H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H54	+	3.971	3.924	67.605	3.899	84.662	O=C1O[C@@](H])(CN1C=2C=C/C=C(/F])C=2)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H57	+	5.222	4.591	79.100	4.543	98.630	O=C1O[C@@](H])(CN1C=2C=CC=C(C=2)C(=O)C)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H59	+	4.155	4.040	69.606	4.015	87.187	O=C1O[C@@](H])(CN1C=2C=C/C=C(/F])C=2)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H62	+	3.876	4.707	81.102	4.659	101.155	O=C1O[C@@](H])(CN1C=2C=CC=C(C=2)C(=O)C)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H63	+	4.097	4.084	70.358	3.944	85.645	O=C1O[C@@](H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H65	+	2.699	3.383	58.290	3.508	76.171	O=C1O[C@@](H])(CN1C=2C=CC=C(C=2)C(=O)C)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC([F])([F])[F]
H66	+	2.420	2.543	43.810	2.655	57.654	O=C1O[C@@](H])(CN1C=2C=CC=CC=2)C(=O)N[C@@](H]

							)(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC([F])([F])[F]
H67	+	3.076	2.760	47.546	2.793	60.662	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC([F])([F])[F]
H69	+	3.237	3.705	63.841	3.634	78.905	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CC4)[S](=O)(=O)C5=CC=C(C=C5)OC
H71	+	0.622	0.807	13.895	0.820	17.824	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC=CC(=C5)OC
H72	+	0.724	0.980	16.879	1.030	22.372	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC=CC(=C5)OC
H73	+	1.530	1.023	17.631	0.959	20.831	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC=CC(=C5)OC
H77	+	1.377	1.009	17.391	0.923	20.067	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CCCO4)[S](=O)(=O)C5=CC=CC(=C5)OC
H78	+	0.824	1.183	20.374	1.133	24.616	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CCCO4)[S](=O)(=O)C5=CC=CC(=C5)OC
H79	+	3.686	3.751	64.621	3.690	80.114	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H81	+	3.599	3.300	56.864	3.264	70.868	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H83	+	2.248	2.516	43.353	2.503	54.356	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H85	+	3.866	3.675	63.325	3.692	80.160	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H87	+	3.848	3.994	68.820	3.967	86.139	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H89	+	3.629	3.544	61.063	3.541	76.893	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H91	+	4.201	3.791	65.320	3.726	80.903	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H92	+	3.460	3.661	63.074	3.645	79.150	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H93	+	3.824	4.037	69.562	4.061	88.168	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H

							])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H95	+	3.349	3.329	57.357	3.542	76.916	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H97	+	3.777	3.648	62.851	3.818	82.895	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H98	+	3.876	4.024	69.339	4.233	91.913	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H100	+	4.137	4.344	74.845	4.244	92.139	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H101	+	3.499	3.517	60.600	3.402	73.875	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H102	+	4.097	4.300	74.092	4.315	93.680	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H104	+	3.635	4.546	78.324	4.384	95.194	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H105	+	3.481	3.719	64.080	3.543	76.930	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H106	+	4.796	4.138	71.293	4.395	95.434	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H107	+	3.708	3.311	57.049	3.554	77.170	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H109	+	3.886	3.936	67.820	3.749	81.412	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H110	+	3.622	3.806	65.574	3.669	79.659	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H111	+	4.585	4.182	72.062	4.084	88.677	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H113	+	4.824	4.968	85.588	4.958	107.649	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H114	+	3.627	4.141	71.344	4.117	89.385	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H116	+	4.602	4.387	75.594	4.474	97.147	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[S](C)(=O)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5

H118	+	3.866	4.316	74.364	4.279	92.905	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[N+](=O)[O-])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H119	+	3.932	4.186	72.118	4.198	91.151	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[N+](=O)[O-])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H121	+	4.699	4.796	82.634	4.750	103.139	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/N)C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H123	+	5.097	5.042	86.876	5.085	110.404	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/N)C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H124	+	3.914	4.069	70.104	4.036	87.644	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H127	+	3.788	3.867	66.630	3.898	84.635	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)OC)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H129	+	4.585	4.113	70.872	4.233	91.900	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)OC)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H3	-	1.923	1.947	33.537	1.884	40.924	O=C(OCCNC(=O)C([F])([F])[F])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H7	-	4.553	4.419	76.137	4.544	98.663	O=C(OC[C@@](C)([H])C(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H10	-	1.455	1.372	23.637	1.410	30.632	O=C(O[C@]1([H])CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C(C=C3)OC
H20	-	0.915	1.591	27.403	1.675	36.373	O=C(O[C@]1([H])CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H25	-	2.529	2.690	46.355	2.746	59.637	O=C(OCCNC(=O)OC(C)(C)C)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H29	-	1.181	0.806	13.888	0.833	18.113	O=C(O[C@]1([H])CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C(C=C3)OC
H34	-	2.472	2.633	45.365	2.678	58.157	O=C(O[C@@](C)([H])CCC(C)=O)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H47	-	4.347	4.447	76.622	4.761	103.379	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H50	-	3.638	3.850	66.341	3.693	80.196	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H52	-	3.495	4.272	73.606	4.267	92.651	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4

H61	-	4.143	3.877	66.807	4.096	88.940	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H64	-	2.000	2.336	40.249	2.196	47.697	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC([F])([F])[F]~
H68	-	3.590	3.460	59.610	3.564	77.392	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CC4)[S](=O)(=O)C5=CC=C(C=C5)OC
H74	-	0.769	0.745	12.839	0.789	17.155	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC([F])=C([F])C=C5[F]
H82	-	2.133	2.067	35.613	2.147	46.624	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H83	-	4.013	4.371	75.308	4.383	95.157	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H96	-	3.893	3.778	65.097	3.899	84.648	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H115	-	4.310	4.518	77.840	4.555	98.900	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[S](C)(=O)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H120	-	4.824	4.562	78.605	4.614	100.169	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[N+](=O)[O-])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H128	-	3.539	3.737	64.384	3.817	82.882	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)OC)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H1	#	3.614	2.893	49.841	3.750	81.431	O=C(OC(C)(C)C)N[C@@]([H])(C/C1=C/C=CC=C1)[C@]([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H9	#	5.523	5.550	95.623	5.613	121.860	O=C(OCCN1CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@]([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C(C=C3)OC
H13	#	2.040	2.019	34.786	1.942	42.175	O=C(OCCNC(=O)C([F])([F])[F])N[C@@]([H])(C/C1=C/C=CC=C1)[C@]([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H17	#	2.856	3.752	64.647	3.714	80.652	O=C(OC[C@](C)([H])C(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@]([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H27	#	3.400	4.170	71.848	4.116	89.365	O=C(OCCN1CCNC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@]([H])CN(CCCCC)[S](=O)(=O)C3=CC=C(C=C3)OC
H38	#	4.770	5.203	89.640	5.300	115.081	O=C(OCCN1CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@]([H])CN(CCCCC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H39	#	1.376	1.025	17.654	1.098	23.853	O=C(O[C@]1([H])CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@]([H])CN(CCCCC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4

H43	#	4.180	4.300	74.082	4.050	87.929	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H44	#	5.222	3.892	67.051	4.061	88.169	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H55	#	4.071	4.169	71.836	3.969	86.175	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H58	#	4.796	3.967	68.357	3.828	83.121	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H75	#	0.795	0.701	12.077	0.250	5.444	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC([F])=C([F])C=C5[F]
H86	#	4.066	4.125	71.066	4.048	87.892	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H90	#	3.415	3.342	57.579	3.370	73.171	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H94	#	3.674	3.211	55.318	3.219	69.904	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H99	#	3.684	3.198	55.095	3.392	73.649	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H103	#	3.496	3.474	59.848	3.473	75.417	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H108	#	3.648	3.487	60.079	3.393	73.680	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H112	#	3.544	3.356	57.818	3.243	70.413	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H122	#	3.947	4.666	80.388	4.670	101.386	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/N)C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H126	#	5.222	4.315	74.346	4.371	94.909	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H4	*	5.523	5.315	91.576	5.383	116.866	O=C(OCCNC(=O)C([Cl])([Cl])[H])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H11	*	4.260	4.912	84.626	6.351	137.880	O=C(OC[C@@]1([H])NOCC1)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C(C=C3)OC
H15	*	3.519	3.199	55.113	3.255	70.676	O=C(O[C@@](C)([H])CCC(C)=O)N[C@@]([H])(C/C1=C/C=C

							C=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H19	*	5.301	5.768	99.388	5.877	127.601	O=C(OCCN1CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H23	*	1.503	1.381	23.788	1.307	28.404	O=C(OCCNC(=O)C([F])([F])[F])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H31	*	1.943	2.228	38.382	2.137	46.409	O=C(O[C@@](C)([H])CC(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H32	*	1.706	1.453	25.037	1.365	29.656	O=C(OCCNC(=O)C([F])([F])[F])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H37	*	3.664	4.389	75.614	4.380	95.105	O=C(OCCN1CCNC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H41	*	4.000	3.881	66.867	3.770	81.867	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H46	*	5.398	4.098	70.603	3.909	84.874	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H56	*	4.796	3.761	64.805	3.980	86.416	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H60	*	3.465	4.286	73.837	4.085	88.700	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H70	*	3.097	3.503	60.362	3.493	75.851	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CC4)[S](=O)(=O)C5=CC=C(C=C5)OC
H76	*	0.775	0.962	16.575	0.928	20.162	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC([F])=C([F])C=C5[F]
H80	*	4.481	4.127	71.109	4.105	89.132	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H84	*	2.833	2.386	41.107	2.422	52.603	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H117	*	5.523	4.764	82.082	4.890	106.165	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[S](C)(=O)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H125	*	4.292	3.938	67.858	3.956	85.890	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5

[1]Experimental pK_I; [2]Calculated pK_i

Table S2: Observed, predicted activity and DCW, and distribution into the training (+), calibration (-), test (#) and validation (*) sets of split 2 (Models **3** and **4**)

Mol. No.	Set	[1]Exp.	Model **3**		Model **4**		SMILES
			[2]Calc.	DCW	[2]Calc.	DCW	
H1	+	3.614	3.867	79.727	3.584	54.876	O=C(OC(C)(C)C)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H4	+	5.523	5.144	106.029	5.097	78.081	O=C(OCCNC(=O)C([Cl])([Cl])[H])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H6	+	2.697	3.040	62.686	2.978	45.592	O=C(OCCNC(=O)OC(C)(C)C)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H8	+	4.319	4.148	85.512	4.138	63.375	O=C(OCCN1CCNC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C(C=C3)OC
H9	+	5.523	5.447	112.269	5.377	82.376	O=C(OCCN1CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C(C=C3)OC
H11	+	4.260	4.245	87.510	4.250	65.087	O=C(OC[C@@]1([H])NOCC1)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C(C=C3)OC
H13	+	2.040	2.364	48.765	2.378	36.385	O=C(OCCNC(=O)C([F])([F])[F])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H14	+	4.222	5.545	114.300	5.522	84.606	O=C(OCCNC(=O)C([Cl])([Cl])[H])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H16	+	3.638	3.441	70.957	3.404	52.117	O=C(OCCNC(=O)OC(C)(C)C)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H19	+	5.301	5.519	113.764	5.460	83.650	O=C(OCCN1CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H21	+	5.523	4.728	97.476	4.659	71.365	O=C(OC[C@]1([H])NC(=O)CC1)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H24	+	2.167	2.385	49.206	2.033	31.102	O=C(O[C@@](C)([H])CCC(C)=O)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H25	+	2.529	2.422	49.957	2.443	37.381	O=C(OCCNC(=O)OC(C)(C)C)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H27	+	3.400	3.530	72.783	3.602	55.164	O=C(OCCN1CCNC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C(C=C3)OC
H28	+	5.097	4.829	99.540	4.841	74.165	O=C(OCCN1CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C(C=C3)OC
H30	+	3.205	4.038	83.252	4.040	61.880	O=C(OC[C@]1([H])NC(=O)CC1)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C(C=C3)OC
H32	+	1.706	1.746	36.036	1.842	28.174	O=C(OCCNC(=O)C([F])([F])[F])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3

H33	+	6.000	4.927	101.572	4.987	76.395	O=C(OCCNC(=O)C([Cl])([Cl])[H])N[C@@](H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H35	+	2.889	2.823	58.229	2.868	43.906	O=C(OCCNC(=O)OC(C)(C)C)N[C@@](H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H38	+	4.770	4.901	101.035	4.925	75.439	O=C(OCCN1CCOC1=O)N[C@@](H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H41	+	4.000	3.856	79.502	4.018	61.540	O=C1O[C@@](H])(CN1C=2C=CC=CC=2)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)H])[S](=O)(=O)C4=CC=C(C=C4)OC
H42	+	4.081	4.049	83.474	4.085	62.567	O=C1O[C@@](H])(CN1C=2C=C/C=C(/F])C=2)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)H])[S](=O)(=O)C4=CC=C(C=C4)OC
H45	+	6.097	4.704	96.967	4.815	73.763	O=C1O[C@@](H])(CN1C=2C=CC=C(C=2)C(C)=O)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)H])[S](=O)(=O)C4=CC=C(C=C4)OC
H46	+	5.398	4.133	85.206	4.065	62.264	O=C1O[C@@](H])(CN1C=2C=CC(=CC=2)C(C)=O)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)H])[S](=O)(=O)C4=CC=C(C=C4)OC
H48	+	3.276	3.492	71.998	3.597	55.087	O=C1O[C@@](H])(CN1C=2C=CC=CC=2)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H51	+	4.377	3.515	72.488	3.336	51.071	O=C1O[C@@](H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H52	+	3.495	4.339	89.463	4.395	67.311	O=C1O[C@@](H])(CN1C=2C=CC=C(C=2)C(=O)C)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H53	+	3.735	3.769	77.701	3.645	55.812	O=C1O[C@@](H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H14	+	3.971	3.952	81.485	3.922	60.071	O=C1O[C@@](H])(CN1C=2C=C/C=C(/F])C=2)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H57	+	5.222	4.607	94.978	4.653	71.267	O=C1O[C@@](H])(CN1C=2C=CC=C(C=2)C(=O)C)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H60	+	3.465	3.962	81.693	3.953	60.547	O=C1O[C@@](H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H66	+	2.420	2.508	51.746	2.567	39.277	O=C1O[C@@](H])(CN1C=2C=CC=CC=2)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)H])[S](=O)(=O)C4=CC=C(C=C4)OC([F])([F])[F]
H67	+	3.076	2.785	57.449	2.614	40.002	O=C1O[C@@](H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)H])[S](=O)(=O)C4=CC=C(C=C4)OC([F])([F])[F]
H68	+	3.590	3.407	70.253	3.433	52.567	O=C1O[C@@](H])(CN1C=2C=C/C=C(/F])C=2)C(=O)N[C@@](H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CC4)[S](

							=O)(=O)C5=CC=C(C=C5)OC
H69	+	3.237	3.432	70.764	3.431	52.529	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CC4)[S](=O)(=O)C5=CC=C(C=C5)OC
H71	+	0.622	0.654	13.536	0.714	10.862	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC=CC(=C5)OC
H76	+	0.775	0.857	17.725	0.892	13.592	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC([F])=C([F])C=C5[F]
H77	+	1.377	1.065	22.011	1.133	17.296	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CCCO4)[S](=O)(=O)C5=CC=CC(=C5)OC
H78	+	0.824	1.258	25.982	1.200	18.323	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CCCO4)[S](=O)(=O)C5=CC=CC(=C5)OC
H79	+	3.686	3.759	77.513	3.855	59.044	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H80	+	4.481	4.305	88.744	4.427	67.806	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H81	+	3.599	3.358	69.241	3.497	53.548	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H83	+	2.248	2.278	47.004	2.306	35.288	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H85	+	3.866	3.662	75.498	3.646	55.832	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H86	+	4.066	4.026	83.003	4.067	62.284	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H87	+	3.848	3.929	81.013	3.904	59.789	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H88	+	4.013	4.475	92.244	4.475	68.550	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H89	+	3.629	3.528	72.741	3.546	54.293	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H91	+	4.201	3.794	78.235	3.806	58.280	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC

H93	+	3.824	4.243	87.477	4.214	64.545	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H95	+	3.349	3.406	70.227	3.483	53.332	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H97	+	3.777	3.673	75.743	3.741	57.289	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H98	+	3.876	4.219	86.974	4.312	66.050	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H100	+	4.137	4.581	94.448	4.474	68.530	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H101	+	3.499	3.635	74.944	3.544	54.273	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H102	+	4.097	4.497	92.716	4.494	68.833	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H104	+	3.635	4.522	93.226	4.491	68.795	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H105	+	3.481	3.575	73.723	3.562	54.538	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H106	+	4.796	4.328	89.235	4.165	63.789	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H108	+	3.648	3.418	70.485	3.498	53.567	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H109	+	3.886	3.783	77.989	3.919	60.019	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H112	+	3.544	3.284	67.728	3.398	52.028	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H113	+	4.824	5.152	106.209	5.224	80.029	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H115	+	4.310	4.837	99.707	4.863	74.499	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[S](C)(=O)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H116	+	4.602	4.740	97.718	4.701	72.003	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[S](C)(=O)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S]

							(=O)(=O)C5=CC=C4OCOC4=C5
H118	+	3.866	4.195	86.487	4.227	64.735	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[N+](=O)[O-])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H119	+	3.932	4.098	84.498	4.064	62.240	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[N+](=O)[O-])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H121	+	4.699	4.547	93.747	4.559	69.828	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/N)C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H122	+	3.947	4.451	91.758	4.396	67.333	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/N)C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H123	+	5.097	4.996	102.989	4.967	76.094	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/N)C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H126	+	5.222	4.679	96.460	4.751	72.783	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H129	+	4.585	4.327	89.210	4.262	65.280	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)OC)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H3	-	1.923	1.962	40.494	1.952	29.860	O=C(OCCNC(=O)C([F])([F])[F])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H7	-	4.553	4.521	93.208	4.623	70.810	O=C(OC[C@@](C)([H])C(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H10	-	1.455	1.077	22.249	1.087	16.594	O=C(O[C@]1([H])CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C(C=C3)OC
H12	-	3.229	2.980	61.465	3.027	46.344	O=C(O[C@@](C)([H])CC(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H20	-	0.915	1.149	23.744	1.170	17.868	O=C(O[C@]1([H])CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H22	-	1.763	1.961	40.465	2.066	31.608	O=C(O[C@@](C)([H])CC(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H23	-	1.503	1.344	27.765	1.417	21.649	O=C(OCCNC(=O)C([F])([F])[F])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H43	-	4.180	4.074	83.984	4.083	62.529	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H47	-	4.347	4.704	96.981	4.637	71.029	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C

							4=CC=C(C=C4)OC
H55	-	4.071	3.977	81.995	3.920	60.034	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H59	-	4.155	3.938	81.183	3.956	60.585	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H64	-	2.000	2.193	45.249	2.142	32.769	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC([F])([F])[F]~
H65	-	2.699	3.356	69.211	3.364	51.501	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC([F])([F])[F]
H107	-	3.708	3.382	69.731	3.235	49.532	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H111	-	4.585	4.231	87.231	4.328	66.285	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H114	-	3.627	4.206	86.706	4.294	65.772	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H117	-	5.523	5.285	108.949	5.272	80.765	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[S](C)(=O)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H120	-	4.824	4.644	95.729	4.635	71.001	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[N+](=O)[O-])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H124	-	3.914	4.231	87.219	4.343	66.517	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H127	-	3.788	3.879	79.968	3.853	59.014	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)OC)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H2	#	3.059	2.579	53.194	2.602	39.819	O=C(O[C@@](C)([H])CC(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H18	#	5.301	4.220	87.007	4.221	64.649	O=C(OCCN1CCNC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H26	#	3.391	3.709	76.484	3.143	48.111	O=C(OC[C@](C)([H])C(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H29	#	1.181	0.459	9.520	0.552	8.383	O=C(O[C@]1([H])CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C(C=C3)OC
H34	#	2.472	2.787	57.477	2.459	37.627	O=C(O[C@@](C)([H])CCC(C)=O)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3

H36	#	2.476	4.111	84.755	3.568	54.636	O=C(OC[C@](C)([H])C(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H40	#	4.569	4.111	84.747	4.123	63.154	O=C(OC[C@]1([H])NC(=O)CC1)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H44	#	5.222	3.880	79.993	3.756	57.523	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H50	#	3.638	3.709	76.480	3.662	56.077	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H58	#	4.796	4.036	83.217	3.903	59.768	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H61	#	4.143	3.769	77.702	3.627	55.541	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H70	#	3.097	3.491	71.985	3.413	52.264	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CC4)[S](=O)(=O)C5=CC=C(C=C5)OC
H74	#	0.769	0.580	12.021	0.844	12.868	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC([F])=C([F])C=C5[F]
H75	#	0.795	0.434	9.006	0.749	11.403	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC([F])=C([F])C=C5[F]
H82	#	2.133	1.914	39.500	1.886	28.835	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H90	#	3.415	3.430	70.730	3.385	51.827	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H94	#	3.674	3.296	67.973	3.285	50.288	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H96	#	3.893	3.770	77.732	3.904	59.784	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H99	#	3.684	3.272	67.470	3.383	51.793	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H5	*	3.371	3.003	61.935	2.569	39.313	O=C(O[C@@](C)([H])CCC(C)=O)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H15	*	3.519	3.405	70.206	2.994	45.838	O=C(O[C@@](C)([H])CCC(C)=O)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3

H17	*	2.856	4.729	97.484	4.103	62.847	O=C(OC[C@](C)([H])C(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H31	*	1.943	2.362	48.736	2.492	38.133	O=C(O[C@@](C)([H])CC(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H37	*	3.664	3.602	74.278	3.686	56.438	O=C(OCCN1CCNC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H39	*	1.376	0.531	11.015	0.635	9.656	O=C(O[C@]1([H])CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H49	*	3.770	3.684	75.969	3.664	56.115	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H56	*	4.796	3.783	78.004	3.594	55.028	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H62	*	3.876	4.593	94.676	4.686	71.781	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H63	*	4.097	4.022	82.915	3.936	60.282	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H72	*	0.724	0.846	17.507	0.781	11.889	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC=CC(=C5)OC
H73	*	1.530	0.930	19.239	0.761	11.586	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC=CC(=C5)OC
H84	*	2.833	2.182	45.015	2.144	32.792	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H92	*	3.460	3.698	76.246	3.643	55.784	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H103	*	3.496	3.551	73.212	3.564	54.575	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H110	*	3.622	3.686	76.000	3.756	57.524	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H125	*	4.292	4.134	85.230	4.180	64.022	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H128	*	3.539	3.782	77.979	3.691	56.518	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)OC)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5

[1]Experimental pK_I; [2]Calculated pK_i

Table S3: Observed, predicted activity and DCW, and distribution into the training (+), calibration (-), test (#) and validation (*) sets of split 3 (Models **5** and **6**)

Mol. No.	Set		Model 5		Model 6		SMILES
		[1]Exp	[2]Calc.	DCW	[2]Calc.	DCW	
H2	+	3.059	2.732	59.023	2.692	68.549	O=C(O[C@@](C)([H])CC(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H4	+	5.523	5.532	119.535	5.492	139.826	O=C(OCCNC(=O)C([Cl])([Cl])[H])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H6	+	2.697	3.311	71.535	3.310	84.267	O=C(OCCNC(=O)OC(C)(C)C)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H8	+	4.319	4.284	92.563	4.273	108.781	O=C(OCCN1CCNC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C(C=C3)OC
H11	+	4.260	4.262	92.091	4.264	108.555	O=C(OC[C@@]1([H])NOCC1)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C(C=C3)OC
H12	+	3.229	2.663	57.549	2.654	67.578	O=C(O[C@@](C)([H])CC(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H14	+	4.222	5.464	118.062	5.454	138.855	O=C(OCCNC(=O)C([Cl])([Cl])[H])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H16	+	3.638	3.243	70.061	3.272	83.296	O=C(OCCNC(=O)OC(C)(C)C)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H18	+	5.301	4.712	101.812	4.735	120.566	O=C(OCCN1CCNC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H19	+	5.301	5.823	125.807	5.875	149.576	O=C(OCCN1CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H21	+	5.523	5.014	108.326	5.040	128.318	O=C(OC[C@]1([H])NC(=O)CC1)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H22	+	1.763	2.073	44.799	2.014	51.272	O=C(O[C@@](C)([H])CC(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H24	+	2.167	2.293	49.545	2.475	63.017	O=C(O[C@@](C)([H])CCC(C)=O)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H26	+	3.391	3.336	72.071	3.330	84.780	O=C(OC[C@](C)([H])C(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H28	+	5.097	4.736	102.335	4.733	120.515	O=C(OCCN1CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C(C=C3)OC

H30	+	3.205	3.927	84.854	3.898	99.256	O=C(OC[C@]1([H])NC(=O)CC1)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C(C=C3)OC
H31	+	1.943	2.005	43.325	1.976	50.302	O=C(O[C@@](C)([H])CC(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H33	+	6.000	4.806	103.837	4.775	121.579	O=C(OCCNC(=O)C([Cl])([Cl])[H])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H35	+	2.889	2.584	55.837	2.593	66.020	O=C(OCCNC(=O)OC(C)(C)C)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H36	+	2.476	3.267	70.598	3.292	83.810	O=C(OC[C@](C)([H])C(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H37	+	3.664	4.054	87.588	4.057	103.290	O=C(OCCN1CCNC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H38	+	4.770	5.164	111.583	5.196	132.300	O=C(OCCN1CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H40	+	4.569	4.355	94.102	4.361	111.041	O=C(OC[C@]1([H])NC(=O)CC1)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H42	+	4.081	4.082	88.190	4.076	103.776	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H45	+	6.097	4.729	102.178	4.753	121.010	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H46	+	5.398	4.348	93.946	4.380	111.522	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H49	+	3.770	3.642	78.681	3.594	91.513	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H51	+	4.377	3.455	74.651	3.428	87.280	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H53	+	3.735	3.908	84.438	3.899	99.259	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H56	+	4.796	4.023	86.924	4.028	102.542	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H57	+	5.222	4.857	104.942	4.871	124.009	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H59	+	4.155	3.629	78.419	3.604	91.756	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[

							C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H50	+	3.465	3.779	81.660	3.801	96.780	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H62	+	3.876	4.277	92.407	4.281	108.990	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H63	+	4.097	3.896	84.176	3.908	99.502	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H65	+	2.699	3.548	76.654	3.477	88.536	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC([F])([F])[F]
H67	+	3.076	3.167	68.422	3.105	79.048	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC([F])([F])[F]
H69	+	3.237	3.100	66.979	3.087	78.587	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CC4)[S](=O)(=O)C5=CC=C(C=C5)OC
H70	+	3.097	3.216	69.495	3.194	81.309	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CC4)[S](=O)(=O)C5=CC=C(C=C5)OC
H71	+	0.622	0.950	20.533	0.827	21.052	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC=CC(=C5)OC
H73	+	1.530	1.296	28.009	1.298	33.044	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC=CC(=C5)OC
H76	+	0.775	1.017	21.975	1.171	29.819	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC([F])=C([F])C=C5[F]
H77	+	1.377	1.043	22.541	1.024	26.075	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CCCO4)[S](=O)(=O)C5=CC=CC(=C5)OC
H78	+	0.824	1.123	24.260	1.191	30.322	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CCCO4)[S](=O)(=O)C5=CC=CC(=C5)OC
H79	+	3.686	4.130	89.235	4.027	102.529	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H80	+	4.481	4.373	94.485	4.264	108.569	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H83	+	2.248	2.568	55.485	2.604	66.287	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H84	+	2.833	2.696	58.250	2.721	69.286	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[

							H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H87	+	3.848	3.978	85.958	3.989	101.556	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H89	+	3.629	3.214	69.446	3.213	81.797	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H91	+	4.201	3.655	78.968	3.675	93.571	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H92	+	3.460	3.783	81.732	3.793	96.570	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H93	+	3.824	4.026	86.983	4.030	102.610	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H95	+	3.349	3.666	79.205	3.517	89.534	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H98	+	3.876	4.477	96.729	4.353	110.837	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H100	+	4.137	4.719	101.961	4.735	120.561	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H101	+	3.499	3.712	80.198	3.722	94.762	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H102	+	4.097	4.453	96.205	4.431	112.816	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H104	+	3.635	4.603	99.445	4.628	117.839	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H105	+	3.481	3.595	77.682	3.615	92.040	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H106	+	4.796	4.266	92.174	4.265	108.583	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H107	+	3.708	3.259	70.411	3.252	82.784	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H109	+	3.886	3.942	85.172	3.989	101.556	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H111	+	4.585	4.313	93.187	4.344	110.595	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5

H113	+	4.824	5.100	110.193	5.108	130.049	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H114	+	3.627	4.093	88.430	4.095	104.250	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C(C)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H116	+	4.602	4.729	102.170	4.734	120.519	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[S](C)(=O)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H117	+	5.523	4.972	107.421	4.971	126.560	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[S](C)(=O)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H118	+	3.866	4.219	91.168	4.244	108.048	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[N+](=O)[O-])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H119	+	3.932	4.347	93.933	4.362	111.047	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[N+](=O)[O-])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H121	+	4.699	4.672	100.937	4.812	122.517	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/N)C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H123	+	5.097	5.043	108.952	5.167	131.556	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/N)C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H125	+	4.292	4.568	98.705	4.497	114.505	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H126	+	5.222	4.811	103.956	4.735	120.546	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H129	+	4.585	4.140	89.461	4.195	106.805	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)OC)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H1	-	3.614	3.821	82.562	3.851	98.039	O=C(OC(C)(C)C)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H7	-	4.553	4.352	94.042	4.313	109.817	O=C(OC[C@@](C)([H])C(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H9	-	5.523	5.395	116.559	5.412	137.791	O=C(OCCN1CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C(C=C3)OC
H20	-	0.915	1.530	33.062	1.525	38.832	O=C(O[C@]1([H])CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H25	-	2.529	2.652	57.311	2.631	66.990	O=C(OCCNC(=O)OC(C)(C)C)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H27	-	3.400	3.626	78.339	3.594	91.505	O=C(OCCN1CCNC1=O)N[C@@]([H])(C/C2=C/C=CC=C2

							)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C(C=C3)OC
H39	-	1.376	0.872	18.838	0.847	21.555	O=C(O[C@]1([H])CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C4=CC=C3N=C[S]C3=C4
H43	-	4.180	4.232	91.430	4.273	108.800	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H47	-	4.347	4.753	102.699	4.714	120.024	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H52	-	3.495	4.289	92.669	4.271	108.747	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H58	-	4.796	4.476	96.711	4.498	114.521	O=C1O[C@@]([H])(CN1C=2C=CC(=CC=2)C(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H64	-	2.000	2.216	47.878	2.241	57.063	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC([F])([F])[F]~
H74	-	0.769	0.671	14.499	0.700	17.827	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC([F])=C([F])C=C5[F]
H82	-	2.133	2.128	45.977	2.122	54.024	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H88	-	4.013	4.221	91.209	4.226	107.596	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H96	-	3.893	4.106	88.714	3.998	101.798	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H108	-	3.648	3.502	75.663	3.507	89.293	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H115	-	4.310	4.601	99.406	4.616	117.520	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[S](C)(=O)=O)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H120	-	4.824	4.590	99.183	4.599	117.087	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)[N+](=O)[O-])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4N=C[S]C4=C5
H128	-	3.539	3.897	84.210	3.958	100.764	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)OC)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H3	#	1.923	2.108	45.543	2.300	58.563	O=C(OCCNC(=O)C([F])([F])[F])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H10	#	1.455	1.102	23.814	1.062	27.046	O=C(O[C@]1([H])CCOC1=O)N[C@@]([H])(C/C2=C/C=C

							C=C2)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C(C=C3)OC
H13	#	2.040	2.040	44.069	2.262	57.592	O=C(OCCNC(=O)C([F])([F])[F])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H17	#	2.856	3.926	84.822	3.970	101.086	O=C(OC[C@](C)([H])C(=O)OC)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H29	#	1.181	0.444	9.589	0.384	9.770	O=C(O[C@]1([H])CCOC1=O)N[C@@]([H])(C/C2=C/C=CC=C2)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C(C=C3)OC
H34	#	2.472	2.225	48.071	2.437	62.046	O=C(O[C@@](C)([H])CCC(C)=O)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H44	#	5.222	3.895	84.159	3.910	99.544	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H50	#	3.638	3.791	81.922	3.792	96.537	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H55	#	4.071	4.360	94.195	4.391	111.799	O=C1O[C@@]([H])(CN1C=2C=CC([F])=C([F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H61	#	4.143	3.443	74.389	3.438	87.523	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/OC)C(C)=C4
H68	#	3.590	2.950	63.738	2.889	73.563	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C[C@@]4([H])CC4)[S](=O)(=O)C5=CC=C(C=C5)OC
H75	#	0.795	0.253	5.461	0.474	12.067	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC([F])=C([F])C=C5[F]
H86	#	4.066	3.850	83.194	3.871	98.557	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H90	#	3.415	3.215	69.459	3.194	81.308	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H94	#	3.674	3.018	65.220	3.017	76.811	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H99	#	3.684	3.470	74.965	3.340	85.038	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H103	#	3.496	3.445	74.442	3.418	87.016	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H112	#	3.544	3.306	71.424	3.331	84.796	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])

							CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H122	#	3.947	4.800	103.702	4.930	125.516	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/N)C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H5	*	3.371	2.951	63.769	3.154	80.293	O=C(O[C@@](C)([H])CCC(C)=O)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C2=CC=C(C=C2)OC
H15	*	3.519	2.883	62.295	3.116	79.322	O=C(O[C@@](C)([H])CCC(C)=O)N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(C[C@@](C)([H])CC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H23	*	1.503	1.449	31.319	1.622	41.287	O=C(OCCNC(=O)C([F])([F])[F])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C2=CC=C(C=C2)OC
H32	*	1.706	1.381	29.845	1.584	40.316	O=C(OCCNC(=O)C([F])([F])[F])N[C@@]([H])(C/C1=C/C=CC=C1)[C@](O)([H])CN(CCCCC)[S](=O)(=O)C3=CC=C2N=C[S]C2=C3
H41	*	4.000	4.002	86.470	3.909	99.530	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H48	*	3.276	3.562	76.962	3.428	87.266	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4
H54	*	3.971	4.210	90.954	4.194	106.775	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H66	*	2.420	2.821	60.947	2.634	67.056	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC([F])([F])[F]
H72	*	0.724	1.030	22.253	0.994	25.298	O=C1O[C@@]([H])(CN1C=2C=C/C=C(/[F])C=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(C/C4=C/C=C[S]4)[S](=O)(=O)C5=CC=CC(=C5)OC
H81	*	3.599	3.366	72.722	3.251	82.770	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)CO
H85	*	3.777	4.234	91.478	3.389	86.293	O=C1O[C@@]([H])(CN1C=2C=CC([F])=CC=2)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H97	*	3.622	4.070	87.937	4.116	104.797	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)OC([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C5=CC=C4OCOC4=C5
H110	*	3.914	4.440	95.940	4.107	104.555	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)C)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H124	*	3.788	3.769	81.446	4.380	111.506	O=C1O[C@@]([H])(CN1C=2C=CC=C(C=2)NC(=O)OC)C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=CC=C(C=C4)OC
H127	*	3.866	3.410	73.685	3.840	97.765	O=C1O[C@@]([H])(CN1C=2C=CC=CC=2C([F])([F])[F])C(=O)N[C@@]([H])(C/C3=C/C=CC=C3)[C@@](O)([H])CN(CC(C)(C)[H])[S](=O)(=O)C4=C/C=C(/N)C=C4

[1]Experimental pK_I; [2]Calculated pK_i

Table S4: Y-randomization test of Split 1 (Models **1** and **2**)

	Model **1**			Model **2**		
	Training	Calibration	Test	Training	Calibration	Test
n	70	20	21	70	20	21
Original	0.837	0.942	0.799	0.811	0.950	0.806
1	0.006	0.001	0.004	0.062	0.000	0.088
2	0.029	0.001	0.007	0.028	0.076	0.003
3	0.003	0.000	0.002	0.000	0.021	0.024
4	0.049	0.001	0.003	0.050	0.025	0.121
5	0.012	0.271	0.005	0.002	0.097	0.000
6	0.006	0.060	0.006	0.055	0.009	0.026
7	0.011	0.023	0.009	0.003	0.098	0.007
8	0.008	0.362	0.021	0.014	0.003	0.001
9	0.011	0.073	0.016	0.004	0.025	0.063
10	0.000	0.017	0.103	0.024	0.001	0.106
R^2_r	0.013	0.081	0.018	0.024	0.036	0.044
$^{C}\boldsymbol{R^2_p}$	**0.823**	**0.907**	**0.782**	**0.824**	**0.937**	**0.796**

Table S5: Y-randomization test of Split 2 (Models **3** and **4**)

	Model **3**			Model **4**		
	Training	Calibration	Test	Training	Calibration	Test
n	72	20	19	72	20	19
Original	0.843	0.947	0.781	0.841	0.941	0.784
1	0.024	0.071	0.001	0.000	0.006	0.001
2	0.066	0.130	0.192	0.021	0.051	0.107
3	0.000	0.006	0.007	0.031	0.100	0.012
4	0.000	0.071	0.002	0.002	0.006	0.201
5	0.059	0.005	0.011	0.002	0.001	0.022
6	0.019	0.029	0.002	0.006	0.005	0.059
7	0.006	0.216	0.010	0.015	0.232	0.021
8	0.048	0.147	0.068	0.012	0.008	0.127
9	0.034	0.036	0.004	0.030	0.023	0.006
10	0.006	0.161	0.000	0.028	0.000	0.010
R^2_r	0.026	0.087	0.030	0.015	0.043	0.057
$\mathbf{{}^{C}R^2_p}$	**0.830**	**0.903**	**0.766**	**0.834**	**0.919**	**0.811**

Table S6: Y-randomization test of Split 3 (Models **5** and **6**)

	Model **5**			Model **6**		
	Training	Calibration	Test	Training	Calibration	Test
n	75	20	19	75	20	19
Original	0.829	0.945	0.804	0.830	0.945	0.773
1	0.001	0.008	0.033	0.019	0.001	0.016
2	0.001	0.017	0.024	0.001	0.006	0.091
3	0.000	0.003	0.029	0.004	0.037	0.001
4	0.001	0.099	0.074	0.001	0.000	0.013
5	0.005	0.007	0.005	0.003	0.048	0.129
6	0.018	0.203	0.026	0.030	0.032	0.000
7	0.004	0.040	0.076	0.000	0.062	0.083
8	0.001	0.056	0.001	0.000	0.037	0.154
9	0.002	0.002	0.075	0.011	0.005	0.006
10	0.010	0.012	0.008	0.030	0.025	0.006
R^2_r	0.004	0.045	0.035	0.010	0.025	0.050
$^{C}\boldsymbol{R^2_p}$	**0.827**	**0.923**	**0.787**	**0.825**	**0.933**	**0.748**