CHEMMEDCHEM
Chemistry & Drug Discovery

Supporting Information

A BOILED-Egg To Predict Gastrointestinal Absorption and Brain Penetration of Small Molecules

Antoine Daina and Vincent Zoete*[a]

cmdc_201600182_sm_miscellaneous_information.pdf
cmdc_201600182_sm_Supporting_DataS3.xlsx

Supporting Information

Table of Content

Methods S1. Construction of the HIA-dataset. Data quality is of primary importance to assure significant predictive power of the model. Particular care was given to retrieve the most relevant and up-to-date information on passive gastrointestinal absorption and BBB permeation. We processed and cured the data published by Newby and colleagues,[1] who collected abundant and well-documented information related to human gastrointestinal absorption of small molecules as follow.
For our purpose, measurements performed on prodrugs, salts or molecules without polar atoms were excluded. We performed additional curation by removing entries showing important variability or conflicting FA values. Moreover, compounds for which active transport was seriously suspected were withdrawn. Publicly available databases, both specialized (Metrabase,[2] UCSF-FDA TransPortal,[3] TSdb[4]) and more generalist (DrugBank,[5] ChEMBL[1,6]), were thoroughly queried to detect potential active transporters. Also, literature information on kinetics of absorption was searched to detect slow or saturable rates, which could stem from active transport. The SMILES description[2,7] of all 660 molecules remaining after filtering were visually checked and corrected if necessary. Finally, the chemical structures were neutralized, aromatized and all stereochemistry notations removed by using the Standardizer program (version 14.9.29.0, 2014, http://www.chemaxon.com). The commonly accepted threshold of 30% FA was applied to discriminate 567 well-absorbed (HIA+) from 93 poorly absorbed (HIA-) molecules. The final HIA-dataset is given in Table S1.

Methods S2. Construction of BBB-dataset. We processed and cured the data published by Brito-Sanchez and colleagues,[3,8] who collected abundant and well-documented information related to human brain access for small molecules as follow.
For our purpose, measurements performed on prodrugs, salts or molecules without polar atoms were excluded. We performed additional curation by removing entries showing important variability or conflicting blood-brain partition (log *BB*) values. Moreover, compounds for which active transport or efflux was seriously suspected were withdrawn. Publicly available databases, both specialized (Metrabase,[2,4] UCSF-FDA TransPortal,[3,5] TSdb[4]) and more generalist (DrugBank,[5] ChEMBL[6]), were thoroughly queried to detect potential active transporters. The SMILES description[7] of all 260 molecules remaining after filtering were visually checked and corrected if necessary. Finally, the chemical structures were neutralized, aromatized and all stereochemistry notations removed by using the Standardizer program (version 14.9.29.0, 2014, http://www.chemaxon.com). A threshold of log *BB* = 0 was applied to discriminate 156 permeant (BBB+) from 104 non-permeant (BBB-) molecules. The final BBB-dataset is given in Table S2.

Methods S3. Computation of descriptors. Apart from robustness and speed, the choice of descriptors was based on open access and ease of implementation to allow for general applicability of our method and straightforward reproducibility of the results.

We evaluate the lipophilicity by an atomistic method that breaks up the molecular structure into fragments to which a lipophilic contribution is given according to a fragmental system. The estimated log *P* of a given molecule is the sum of the lipophilic contributions of all its fragments without correction parameter. We choose to rely on the fragmental system proposed by Wildman and Crippen[9] because closely related to that of AlogP98[10] employed by Egan. However and in contrast to AlogP98, the chemical description of the 142 listed fragments is exhaustive and openly available. We implemented the exact SMARTS fragments as described in the original paper[9] in a tailor-made Python script steering the SMARTS recognition pattern engine of OpenBabel.[11] The method and values are referred to as WLOGP.

The polarity of molecules was evaluated by the polar surface area (PSA), *i.e.* the portion of the molecular surface that belongs to polar atoms. Instead of generating tridimensional geometries and actual surfaces, we opted for a well-described fragment-based technique to rapidly and accurately estimate the PSA, called the topological polar surface area (tPSA).[12] Similarly to WLOGP, this method cuts the molecular structure into two-dimensional fragments and a surface contribution is assigned to the 43 atom-types considered as polar. The estimated PSA of a given molecule is the sum of the contributions of all its fragments. The small number of fragments and their precise description makes the method unambiguous and its implementation straightforward. Many packages implementing tPSA exist. For the BOILED-Egg, we computed the tPSA with Openbabel.[11]
All molecules in SMILES description, as well as WLOGP and tPSA values are given in Tables S1 and S2, for HIA- and BBB-datasets, respectively.

Methods S4. Definition of the ellipses. For each dataset, WLOGP and tPSA values were normalized towards standard deviation. In both normalized two-dimensional spaces (HIA and BBB), one can define an ellipse by five parameters: the foci Cartesian coordinates (x_1, y_1) and (x_2, y_2) along with the major axis (or largest diameter, d). By definition, the sum of distances from every point on the ellipse's trace to each focus is equal to d. Note: The final model is denormalized and depicted in the actual WLOGP *versus* tPSA referential that allow for the merging to obtain the BOILED-Egg (Figure 1e).

Methods S5. Classification of molecules. Any point *P* of the normalized two-dimensional plots WLOGP *versus* tPSA can be evaluated as follows: if the sum of distances to foci is smaller than *d*, then *P* is inside the ellipse. If the sum of distances to foci is larger than *d*, then *P* is outside the ellipse. This enabled training the classification models for both datasets. All HIA+ (or BBB+, respectively) molecules were plotted in the normalized physicochemical space and points inside a tentative ellipse were counted as true positives (TP); points outside a tentative ellipse were considered false negatives (FN). Likewise, all HIA- (or BBB-, respectively) molecules were plotted in the normalized physicochemical space and points outside a tentative ellipse were counted as true negatives (TN); points inside a tentative ellipse were considered false positives (FP).

Methods S6. Stochastic optimization of ellipses. To define the starting position of ellipses to be optimized, a 9×9 grid was placed so that it spans the desired physicochemical space of normalized WLOGP *versus* tPSA plots. For HIA: x_1 ranges from 0 to 175; y_1 ranges from -1 to 8; x_2 ranges from x_1 to 175; y_2 ranges from -1 to y_1. For BBB: x_1 ranges from 0 to 120; y_1 ranges from -1 to 8; x_2 ranges from x_1 to 120; y_2 ranges from -1 to y_1. Foci were positioned on these points and the distance between them ($2c$) calculated. For each starting position, four initial largest diameters were set as $d = 1 \times 2c$, $1.5 \times 2c$, $2 \times 2c$ or $3 \times 2c$.

This produced about 10,000 initial vectors of 5 dimensions (x_1, y_1, x_2, y_2, d) per dataset (HIA and BBB) to search for the best classification ellipses. To this end, each initial vector was submitted to 10 independent Monte Carlo (MC) optimizations[13], which consist in i) start from the initial 5-dimension vector and evaluate the ellipse; ii) generate a new tentative ellipse by operating a random change in the five parameters and evaluate the ellipse; iii) determine whether the new ellipse is accepted in accordance with a Metropolis-like criterion; iv) repeat steps ii) and iii) 100,000 times.

The optimization of the five parameters was monitored in steps i) and ii) by the MCC, *i.e.* the classification capacity of the tentative ellipses. To prevent an exaggerated growth of the ellipse at the early steps of MC, we added an empirically determined small constraint on the surface of the ellipse (S) in the normalized space. Finally, the scoring function to evaluate the tentative ellipses was: $Score = MCC - 0.05 \times S$.

The acceptance of a new tentative ellipse followed a Metropolis-like criterion, which implied that an ellipse was retained when its *Score* was higher than the one of the best ellipse so far. In case the change in the parameters led to a lower *Score*, it was kept with a probability defined by the *Score* difference between the tentative ($Score_t$) and the best ellipse so far ($Score_b$), and including a random number (r, between 0 and 1). In a mathematical fashion, the tentative ellipse is accepted if: $Score_t < Score_b$ or $Score_t < Score_b - 0.025 \times r$.

This optimization method was applied on HIA- and BBB-datasets. The best classifying ellipse in terms of MCC for each dataset were denormalized to obtain the final model in the actual WLOGP *versus* tPSA referential (as depicted in Figure 1c and Figure 1d for HIA and BBB, respectively and to Figure 1e for the merged BOILED-Egg model). The normalized and denormalized final ellipses are described in more detail in Data S1 and Figure S1 for HIA, and S Data S2 and Figure S2 for BBB.

Methods S7. Evaluation of classification. In contrast to the Egan egg, all our tentative ellipses were evaluated according to their ability to include as many HIA+ (or BBB+, respectively) and to exclude as many HIA- (or BBB-, respectively) as possible. The Matthews correlation coefficient (MCC) allowed measuring performance of a given ellipse including correctly and incorrectly classified points as expressed by equation (1).

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

MCC varies from -1 for all points misclassified to 1 for perfect classification. A value near 0 reveals random classification.

Methods S8. Cross-validation. In order to assess robustness, the best classification ellipses were submitted to cross-validation. This procedure consists in splitting the entire dataset into a training set and a validation set. The model is re-trained on the new training set and the classification ability evaluated for the molecules belonging to the validation set. In the present case, 10-fold cross-validations were performed for HIA- and BBB-datasets, independently. We attributed every tenth entry of the dataset ranked by tPSA values to the validation set. This was repeated ten times starting from entries number 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 to obtain ten different couples of training/validation sets covering the broadest possible range of polarity. The ellipses were optimized on each training set, and MCC as well as accuracy computed for the molecules belonging to the corresponding validation set. The cross-validated MCC (MCC_{CV}) and cross-validated accuracy were averaged over the ten validation sets (refer to Tables S3 and S4; and to Figures S3 and S4).

Methods S9. Predictions for new molecular structures. Any molecular structure can be plotted in the BOILED-Egg (*e.g.* through the Excel spreadsheet as Data S3). Calculation of both descriptors should be as similar as possible to the protocol described in Method S3. To avoid misclassification by erroneous physicochemical description, the chemical structure should be unsalted and in its neutral form. The tPSA must follow the strict definition of the original paper[12], accounting for sulfur and phosphate polar atoms as well. Most importantly, the log *P* must be calculated with WLOGP and not another method because of possible significant discrepancies in predicted log *P* values.[14]

The best way to assess gastrointestinal absorption and brain penetration is certainly to visually inspect whether a point lies inside or outside of the HIA and BBB ellipses, so of BOILED-Egg's white and yolk, respectively (Figure 1e and Data S3). As an illustration, we plotted in the referential of the BOILED-Egg 46 NCEs accepted by the FDA between January 2014 and September 2015 (http://www.accessdata.fda.gov/scripts/cder/drugsatfda/ accessed October 1st, 2015), excluding prodrugs and compounds of unclear oral bioavailability. As only three of these 46 molecules are part of our HIA-dataset, this test set can be considered as external. Besides, only compounds with enough evidence of brain access are tagged as BBB permeant (refer to FDA-dataset in Table S5). However, one can suspect that for some of these recent molecules, active transporters or efflux pumps remain to be discovered.

Table S1. HIA-dataset. HIA+ and HIA- correspond to compounds considered as well- and poorly absorbed by the gastrointestinal tract, respectively.

Name	SMILES	tPSA [Å^2]	WLOGP	HIA
AAFC_flurocitabine	OCC1OC2C(Oc3nc(=N)c(F)cn23)C1O	100.35	-0.84	+
Acamprosate	CC(=O)NCCCS(O)(=O)=O	91.85	0.48	-
Acarbose	CC1OC(OC2C(CO)OC(OC3C(CO)OC(O)C(O)C3O)C(O)C2O)C(O)C(O)C1NC1C=C(CO)C(O)C(O)C1O	321.17	-8.56	-
Acenocoumarol	CC(=O)CC(c1ccc(cc1)[N+]([O-])=O)c1c(O)c2ccccc2oc1=O	116.25	0.11	+
Acetanilide	CC(=O)Nc1ccccc1	29.10	1.45	+
Acetazolamide	CC(=O)Nc1nnc(s1)S(N)(=O)=O	151.63	4.12	+
Acetohexamide	CC(=O)c1ccc(cc1)S(=O)(=O)N=C(O)NC1CCCCC1	103.88	3.25	+
Acipimox	Cc1cnc(c[n+]1[O-])C(O)=O	75.27	0.42	+
Acitretin	COc1cc(C)c(C=CC(C)=CC=CC(C)=CC(O)=O)c(C)c1C	46.33	1.77	+
Adefovir	Nc1ncnc2n(CCOCP(O)(O)=O)cnc12	146.19	-0.43	-
Adinazolam	CN(C)Cc1nnc2CN=C(c3ccccc3)c3cc(Cl)ccc3-n12	46.25	0.74	+
Albendazole	CCCSc1ccc2[nH]c(NC(=O)OC)nc2c1	92.31	3.05	-
Alcuronium	OCC=C1C[N+]2(CC=C)CCC34C2CC1C1=CN2C5C(=CN(C31)c1ccccc41)C1CC3C5(CC[N+]3(CC=C)CC1=CCO)c1ccccc21	46.94	3.95	-
Alendronic_Acid	NCCCC(O)(P(O)(O)=O)P(O)(O)=O	180.93	-1.27	-
Alfentanil	CCC(=O)N(c1ccccc1)C1(COC)CCN(CCn2nnn(CC)c2=O)CC1	84.51	5.60	+
Alizapride	COc1cc2nn[nH]c2cc1C(=O)NCC1CCCN1CC=C	83.06	3.28	+
Almitrine	Fc1ccc(cc1)C(N1CCN(CC1)c1nc(NCC=C)nc(NCC=C)n1)c1ccc(F)cc1	68.87	-0.42	+
Almotriptan	CN(C)CCc1c[nH]c2ccc(CS(=O)(=O)N3CCCC3)cc12	64.63	3.58	+
Alogliptin	Cn1c(=O)cc(N2CCCC(N)C2)n(Cc2ccccc2C#N)c1=O	97.02	4.66	+
Alprazolam	Cc1nnc2CN=C(c3ccccc3)c3cc(Cl)ccc3-n12	42.85	2.64	+
Alprenolol	CC(C)NCC(O)COc1ccccc1CC=C	41.46	2.18	+
Altretamine	CN(C)c1nc(nc(n1)N(C)C)N(C)C	48.00	2.39	+
Amantadine	NC12CC3CC(CC(C3)C1)C2	26.02	1.91	+
Amdinocillin	CC1(C)SC2C(N=CN3CCCCCC3)C(=O)N2C1C(O)=O	98.51	0.64	-
Amifostine	NCCCNCCSP(O)(O)=O	130.26	4.22	+
Amikacin	NCCC(O)C(=O)NC1CC(N)C(OC2OC(CN)C(O)C(O)C2O)C(O)C1OC1OC(CO)C(O)C(N)C1O	331.94	-8.42	-
Amiloride	NC(N)=NC(=O)c1nc(Cl)c(N)nc1N	158.26	-0.05	+
Aminoglutethimide	CCC1(CCC(=O)NC1=O)c1ccc(N)cc1	71.77	-0.08	+
Aminopyrine	CN(C)c1c(C)n(C)n(-c2ccccc2)c1=N	36.95	1.67	+
Amlodipine	CCOC(=O)C1=C(COCCN)NC(C)=C(C1c1ccccc1Cl)C(=O)OC	99.35	-1.04	+
Amobarbital	CCC1(CCC(C)C)C(O)=NC(O)=NC1=O	81.86	4.71	+
Amosulalol	COc1ccccc1OCCNCC(O)c1ccc(C)cc1S(N)(=O)=O	119.22	1.00	+
Amoxapine	Clc1ccc2Oc3ccccc3N=C(N3CCNCC3)c2c1	36.86	2.29	+
Amphetamine	CC(N)Cc1ccccc1	26.02	1.58	+
Amphotericin_B	CC1OC(OC2(O)CC3OC(O)(CC(O)C3C(O)=O)CC(O)CC(O)C(O)CCC(O)CC(O)CC(=O)OC(C)C(C)C(O)C(C)C=CC=CC=CC=CC=CC=CC=C2)C(O)C(N)C1O	339.84	0.03	-
Amrinone	Nc1cc(c[nH]c1=O)-c1ccncc1	71.74	-0.24	+
Amygdalin	OCC1OC(OCC2OC(OC(C#N)c3ccccc3)C(O)C(O)C2O)C(O)C(O)C1O	202.32	-3.43	-
Anagrelide	Clc1ccc2N=C3NC(=O)CN3Cc2c1Cl	44.27	3.04	+
Anastrozole	CC(C)(C#N)c1cc(Cn2cncn2)cc(c1)C(C)(C)C#N	78.09	5.94	+
Antipyrine	Cc1cc(=O)n(-c2ccccc2)n1C	26.93	1.48	+
Apricitabine	Nc1ccn(C2COC(CO)S2)c(=O)n1	115.64	5.34	+
Aprindine	CCN(CC)CCCN(C1Cc2ccccc2C1)c1ccccc1	6.48	4.39	+
Arbekacin	NCCC(O)C(=O)NC1CC(N)C(OC2OC(CN)CCC2N)C(O)C1OC1OC(CO)C(O)C(N)C1O	297.27	-6.40	-
Aripiprazole	Clc1cccc(N2CCN(CCCCOc3ccc4CCC(=O)Nc4c3)CC2)c1Cl	44.70	0.63	+
Aspirin	CC(=O)Oc1ccccc1C(O)=O	63.57	2.42	+
Atomoxetine	CNCCC(Oc1ccccc1C)c1ccccc1	21.26	3.40	+
Atracurium	COc1ccc(CC2c3cc(OC)c(OC)cc3CC[N+]2(C)CCC(=O)OCCCCCOC(=O)CC[N+]2(C)CCc3cc(OC)c(OC)cc3C2Cc2ccc(OC)c(OC)c2)cc1OC	126.44	6.66	-
Atropine	CN1C2CCC1CC(C2)OC(=O)C(CO)c1ccccc1	49.74	1.91	+
Auranofin	CCP(CC)(CC)=[Au]SC1OC(COC(C)=O)C(OC(C)=O)C(OC(C)=O)C1OC(C)=O	149.54	2.79	-
Azatadine	CN1CCC(CC1)=C1c2ccccc2CCc2cccnc12	16.13	3.33	+
Azathioprine	Cn1cnc(c1Sc1ncnc2[nH]cnc12)[N+]([O-])=O	143.96	-0.18	+
Azelastine	CN1CCCC(CC1)N1NC(Cc2ccc(Cl)cc2)c2ccccc2C1=O	35.58	2.60	+
Azimilide	CN1CCN(CCCCN2C(=O)CN(N=Cc3ccc(o3)-c3ccc(Cl)cc3)C2=O)CC1	72.19	0.42	+
Azithromycin	CCC1OC(=O)C(C)C(OC2CC(C)(OC)C(O)C(C)O2)C(C)C(OC2OC(C)CC(C2O)N(C)C)C(C)(O)CC(C)CN(C)C(C)C(O)C1(C)O	164.63	0.19	+
Azlocillin	CC1(C)SC2C(NC(=O)C(NC(=O)N3CCNC3=O)c3ccccc3)C(=O)N2C1C(O)=O	173.45	-1.37	-
Azosemide	NS(=O)(=O)c1cc(-c2nnn[nH]2)c(NCc2cccs2)cc1Cl	163.27	2.58	-
Aztreonam	CC1C(NC(=O)C(=NOC(C)(C)C(O)=O)c2csc(N)n2)C(=O)N1S(O)(=O)=O	238.20	-0.46	-
Benazepril	CCOC(=O)C(CCc1ccccc1)NC1CCc2ccccc2N(CC(O)=O)C1=O	95.86	1.66	+
Bendroflumethiazide	NS(=O)(=O)c1cc2c(NC(Cc3ccccc3)NS2(=O)=O)cc1C(F)(F)F	134.59	2.13	+
Benorylate	CC(=O)Nc1ccc(OC(=O)c2ccccc2OC(C)=O)cc1	81.65	1.47	+
Benserazide	NC(CO)C(=O)NNCc1ccc(O)c(O)c1O	147.24	1.56	+
Benzbromarone	CCc1oc2ccccc2c1C(=O)c1cc(Br)c(O)c(Br)c1	50.19	0.78	+
Benzydamine	CN(C)CCCOc1nn(Cc2ccccc2)c2ccccc12	30.29	3.42	+
Bephenium	C[N+](C)(CCOc1ccccc1)Cc1ccccc1	9.23	3.19	-
Bepridil	CC(C)COCC(CN1CCCC1)N(Cc1ccccc1)c1ccccc1	15.71	4.30	+
Betahistine	CNCCc1ccccn1	24.92	0.84	+
Betaxolol	CC(C)NCC(O)COc1ccc(CCOCC2CC2)cc1	50.44	4.04	+
Betazole	NCCc1cc[nH]n1	54.37	2.79	+
Bethanechol	CC(C[N+](C)(C)C)OC(N)=O	52.32	0.18	-
Bezafibrate	CC(C)(Oc1ccc(CCNC(=O)c2ccccc2)cc1)C(O)=O	75.35	-0.96	+
Bicalutamide	CC(O)(CS(=O)(=O)c1ccc(F)cc1)C(=O)Nc1ccc(C#N)c(c1)C(F)(F)F	115.06	1.04	+
Bifemelane	CNCCCOc1ccccc1Cc1ccccc1	21.26	3.66	+
Biperiden	OC(CCN1CCCCC1)(C1CC2CC1C=C2)c1ccccc1	23.47	3.47	+
Bisoprolol	CC(C)NCC(O)COc1ccc(COCCOC(C)C)cc1	59.59	5.58	+
Bornaprine	CCN(CC)CCCOC(=O)C1(CC2CCC1C2)c1ccccc1	29.54	4.02	+
Bosentan	COc1ccccc1Oc1c(NS(=O)(=O)c2ccc(cc2)C(C)(C)C)nc(nc1OCCO)-c1ncccn1	151.66	0.03	+
Bromazepam	Brc1ccc2NC(=O)CN=C(c3ccccn3)c2c1	53.33	5.10	+
Bromfenac	Nc1c(CC(O)=O)cccc1C(=O)c1ccc(Br)cc1	80.36	0.31	+
Bromhexine	CN(Cc1cc(Br)cc(Br)c1N)C1CCCCC1	29.26	4.41	+
Budesonide	CCCC1OC2CC3C4CCC5=CC(=O)C=CC5(C)C4C(O)CC3(C)C2(O1)C(=O)CO	93.03	-0.53	+
Buflomedil	COc1cc(OC)c(C(=O)CCCN2CCCC2)c(OC)c1	46.53	3.54	+
Bufuralol	CCc1cccc2cc(oc12)C(O)CNC(C)(C)C	45.33	2.57	+
Bumetanide	CCCCNc1cc(cc(c1Oc1ccccc1)S(N)(=O)=O)C(O)=O	124.57	1.39	+
Bupivacaine	CCCCC1CCCNC1C(=O)Nc1c(C)cccc1C	41.03	5.93	+
Bupropion	CC(NC(C)(C)C)C(=O)c1ccc(Cl)cc1	29.10	3.30	+

Buspirone	O=C1CC2(CCCC2)CC(=O)N1CCCCN1CCN(CC1)c1ncccn1	69.54	1.62	+
Busulfan	CS(=O)(=O)OCCCCOS(C)(=O)=O	103.04	0.30	+
Butylscopolamine	CCCC[N+]1(C)C2CC(CC1C1OC21)OC(=O)C(CO)c1ccccc1	59.06	1.85	-
Caffeine	Cn1cnc2n(C)c(=O)n(C)c(=O)c12	61.77	2.25	+
Calcitriol	CC(CCCC(C)(C)O)C1CCC2C(CCCC12C)=CC=C1CC(O)CC(O)C1=C	60.47	2.83	+
Camazepam	CN(C)C(=O)OC1N=C(c2ccccc2)c2cc(Cl)ccc2N(C)C1=O	61.82	-1.03	+
Candoxatril	COCCOCC(CC1(CCCC1)C(=O)NC1CCC(CC1)C(O)=O)C(=O)Oc1ccc2CCCc2c1	110.00	2.03	+
Capecitabine	CCCCCOC(=O)Nc1nc(=O)n(cc1F)C1OC(C)C(O)C1O	121.70	3.04	+
Capreomycin	CC1NC(=O)C(N)CNC(=O)C(NC(=O)C(NC(=O)C(CNC(=O)CC(N)CCCN)NC1=O)=CNC(N)=O)C1CCN=C(N)N1	358.19	-9.87	-
Captopril	CC(CS)C(=O)N1CCCC1C(O)=O	96.31	3.70	+
Carbamazepine	NC(=O)N1c2ccccc2C=Cc2ccccc12	46.32	1.28	+
Carbamazepine-10,11-epoxide	NC(=O)N1c2ccccc2C2OC2c2ccccc12	58.56	3.24	+
Carbimazole	CCOC(=O)n1ccn(C)c1=S	68.01	-0.31	+
Carfecillin	CC1(C)SC2C(NC(=O)C(C(O)=O)c3ccccc3)C(=O)N2C1C(O)=O	148.82	-1.37	+
Carglumic_Acid	NC(=O)NC(CCC(O)=O)C(O)=O	128.59	-0.34	+
Carmustine	ClCCNC(=O)N(CCCl)N=O	61.69	1.78	+
Carprofen	CC(C(O)=O)c1ccc2c(c1)[nH]c1ccc(Cl)cc21	52.93	1.61	+
Carteolol	CC(C)(C)NCC(O)COc1cccc2NC(=O)CCc12	70.50	0.73	+
Carvedilol	COc1ccccc1OCCNCC(O)OCc1cccc2[nH]c3ccccc3c12	75.63	2.90	+
Cefazolin	Cc1nnc(SCC2=C(N3C(SC2)C(NC(=O)Cn2cnnn2)C3=O)C(O)=O)s1	234.93	-1.02	-
Cefmetazole	COC1(NC(=O)CSCC#N)C2SCC(CSc3nnnn3C)=C(N2C1=O)C(O)=O	239.23	-1.11	-
Cefodizime	CON=C(C(=O)NC1C2SCC(CSc3nc(C)c(CC(O)=O)s3)=C(N2C1=O)C(O)=O)c1csc(N)n1	304.48	0.63	-
Ceforanide	NCc1ccccc1CC(=O)NC1C2SCC(CSc3nnnn3CC(O)=O)=C(N2C1=O)C(O)=O	244.23	-1.24	-
Cefotaxime	CON=C(C(=C)NC1C2SCC(COC(C)=O)=C(N2C1=O)C(O)=O)c1csc(N)n1	209.98	-0.01	-
Cefpirome	CON=C(C(=O)NC1C2SCC(C[n+]3ccccc3)=C(N2C1=O)C(O)=O)c1csc(N)n1	205.68	0.22	-
Cefprozil	CC=CC1=C(N2C(SC1)C(NC(=O)C(N)c1ccc(O)cc1)C2=O)C(O)=O	154.03	5.09	+
Ceftazidime	CC(C)(ON=C(C(=O)NC1C2SCC(C[n+]3ccccc3)=C(N2C1=O)C(O)=O)c1csc(N)n1)C(O)=O	242.98	0.46	-
Celecoxib	Cc1ccc(cc1)-c1cc(nn1-c1ccc(cc1)S(N)(=O)=O)C(F)(F)F	86.30	4.97	+
Cephalothin	CC(=O)OCC1=C(N2C(SC1)C(NC(=O)Cc1cccs1)C2=O)C(O)=O	166.55	0.21	-
Chenodeoxycholic_Acid	CC(CCC(O)=O)C1CCC2C3C(O)CC4CC(O)CCC4(C)C3CCC12C	76.90	3.90	+
Chloral_Hydrate	OC(O)C(Cl)(Cl)Cl	39.82	2.82	+
Chlorambucil	OC(=O)CCCc1ccc(cc1)N(CCCl)CCCl	40.46	4.83	+
Chloramphenicol	OCC(NC(=O)C(Cl)Cl)C(O)c1ccc(cc1)[N+]([O-])=O	118.81	2.65	+
Chlordiazepoxide	CNC1=Nc2ccc(Cl)cc2C(c2ccccc2)=[N+]([O-])C1	56.84	4.35	+
Chlorguanide	CC(C)NC(=N)NC(=N)Nc1ccc(Cl)cc1	83.65	2.82	+
Chlorhexidine	Clc1ccc(NC(=N)NC(=N)NCCCCNC(=N)NC(=N)Nc2ccc(Cl)cc2)cc1	167.58	3.01	-
Chlormezanone	CN1C(c2ccc(Cl)cc2)S(=O)(=O)CCC1=O	62.54	2.52	+
Chlorothiazide	NS(=O)(=O)c1cc2c(cc1Cl)N=CNS2(=O)=O	135.12	2.38	+
Chloroxine	Oc1c(Cl)cc(Cl)c2cccnc12	33.12	3.25	+
Chlorphenesin	OCC(O)COc1ccc(Cl)cc1	49.41	1.71	+
Chlorpheniramine	CN(C)CCN(c1ccc(Cl)cc1)c1ccccn1	19.37	3.43	+
Chlorpropamide	CCCNC(=O)NS(=O)(=O)c1ccc(Cl)cc1	83.55	0.68	+
Chlorthalidone	NS(=O)(=O)c1cc(ccc1Cl)C1(O)NC(=O)c2ccccc12	117.17	3.93	+
Chlorzoxazone	Clc1ccc2oc(=O)[nH]c2c1	45.98	5.06	+
Cibenzoline	C1C(C2=NCCN2)C1(c1ccccc1)c1ccccc1	24.39	2.23	+
Cicaprost	CCC#CCC(C)C(O)C#CC1C(O)CC2CC(CC12)=CCOCC(O)=O	86.89	2.02	+
Cicloprolol	CC(C)NCC(O)COc1ccc(OCCOCC2CC2)cc1	59.95	2.21	+
Cidofovir	Nc1ccn(CC(CO)OCP(O)(O)=O)c(=O)n1	157.71	-1.65	-
Cilazaprilat	OC(=O)C(CCc1ccccc1)NC1CCCN2CCCC(N2C1=O)C(O)=O	110.18	0.36	-
Cilomilast	COc1ccc(cc1OC1CCCC1)C1(CCC(CC1)C(O)=O)C#N	79.37	2.55	+
Cinchonine	OC(C1CC2CCN1CC2C=C)c1ccnc2ccccc12	36.36	2.46	+
Cinolazepam	OC1N=C(c2ccccc2F)c2cc(Cl)ccc2N(CCC#N)C1=O	76.38	3.48	+
Cinoxacin	CCn1nc(C(O)=O)c(=O)c2cc3OCOc3cc12	90.44	0.76	+
Ciprofibrate	CC(C)(Oc1ccc(cc1)C1CC1(Cl)Cl)C(O)=O	46.53	5.06	+
Cisapride	COC1CN(CCCOc2ccc(F)cc2)CCC1NC(=O)c1cc(Cl)c(N)cc1OC	85.61	0.22	+
Citalopram	CN(C)CCCC1(OCc2cc(ccc12)C#N)c1ccc(F)cc1	36.26	3.97	+
Clavulanic_Acid	OCC=C1OC2CC(=O)N2C1C(O)=O	86.99	2.38	+
Clenbuterol	CC(C)(C)NCC(O)c1cc(Cl)c(N)c(Cl)c1	58.22	6.45	+
Clinafloxacin	NC1CCN(C1)c1c(F)cc2c(c1Cl)n(cc(C(O)=O)c2=O)C1CC1	87.71	0.51	+
Clindamycin	CCCC1CC(N(C)C1)C(=O)NC(C(C)Cl)C1OC(SC)C(O)C(O)C1O	127.10	3.93	+
Clobazam	CN1c2ccc(Cl)cc2N(c2ccccc2)C(=O)CC1=O	40.54	5.18	+
Clodronate	OP(O)(=O)C(Cl)(Cl)P(O)(O)=O	134.68	0.43	-
Clofibrate	CCOC(=O)C(C)(C)Oc1ccc(Cl)cc1	35.53	3.06	+
Clomethiazole	Cc1ncsc1CCCl	41.13	3.23	+
Clomipramine	CN(C)CCCN1c2ccccc2CCc2ccc(Cl)cc12	6.48	4.15	+
Clonazepam	[O-][N+](=O)c1ccc2NC(=O)CN=C(c3ccccc3Cl)c2c1	91.02	1.96	+
Clonidine	Clc1cccc(Cl)c1N=C1NCCN1	36.42	1.41	+
Clopidogrel	COC(=O)C(N1CCc2sccc2C1)c1ccccc1Cl	57.61	2.50	+
Clorazepate	OC(=O)C1N=C(c2ccccc2)c2cc(Cl)ccc2NC1=O	78.61	0.89	+
Cloxacillin	Cc1onc(c1C(=O)NC1C2SC(C)(C)C(N2C1=O)C(O)=O)-c1ccccc1Cl	136.22	2.15	+
Codeine	COc1ccc2CC3C4C=CC(O)C5Oc1c2C45CCN3C	41.57	1.02	+
Colestipol	NCCNCCNCCNCCN	105.11	-6.84	-
Conivaptan	Cc1nc2CCN(C(=O)c3ccc(NC(=O)c4ccccc4-c4ccccc4)cc3)c3ccccc3-c2[nH]1	78.02	4.73	+
Cotinine	CN1C(CCC1=O)c1cccnc1	33.20	0.67	+
Coumarin	O=c1ccc2ccccc2o1	30.21	1.79	+
Cromolyn	OC(COc1cccc2oc(cc(=O)c12)C(O)=O)COc1cccc2oc(cc(=O)c12)C(O)=O	173.71	2.11	-
Cyclopenthiazide	NS(=O)(=O)c1cc2c(NC(CN3CCCC3)NS2(=O)=O)cc1Cl	138.04	2.17	+
Cyclophosphamide	ClCCN(CCCl)P1(=O)NCCCO1	51.37	2.46	+
Cymarin	COC1CC(OC2CCC3(C=C)C4CCC5(C)C(CCC5(O)C4CCC3(O)C2)C2=CC(=O)OC2)OC(C)C1O	114.29	3.24	+
Cyproheptadine	CN1CCC(CC1)=C1c2ccccc2C=Cc2ccccc12	3.24	4.10	+
Dalfampridine	Nc1ccncc1	38.77	2.53	+
Dantrolene	[O-][N+](=O)c1ccc(cc1)-c1ccc(C=NN2CC(=O)NC2=O)o1	123.92	0.69	+
Dapiprazole	Cc1ccccc1N1CCN(CCc2nnc3CCCCn23)CC1	37.19	1.53	-
Dapsone	Nc1ccc(cc1)S(=O)(=O)c1ccc(O)cc1	88.56	2.34	+
Deferoxamine	CC(=O)N(O)CCCCCNC(=O)CCC(=O)N(O)CCCCCNC(=O)CCC(=O)N(O)CCCCCN	205.84	0.92	-
Delapril	CCOC(=O)C(CCc1ccccc1)NC(C)C(=O)N(CC(O)=O)C1Cc2ccccc2C1	95.94	2.19	+
Delavirdine	CC(C)Nc1cccnc1N1CCN(CC1)C(=O)c1cc2cc(NS(C)(=O)=O)ccc2[nH]1	118.31	3.12	+
Delmopinol	CCCC(CCC)CCCC1COCCN1CCO	32.70	2.69	+
Desmethyldiazepam	Clc1ccc2NC(=O)CN=C(c3ccccc3)c2c1	41.13	2.23	+
Desogestrel	CCC12CC(=C)C3C(CCC4=CCCCC34)C1CCC2(O)C#C	20.23	4.95	+
Dextromoramide	CC(CN1CCOCC1)C(C(=O)N1CCCC1)(c1ccccc1)c1ccccc1	32.78	2.80	+
Diazepam	CN1c2ccc(Cl)cc2C(=NCC1=O)c1ccccc1	32.67	2.39	+
Diazoxide	CC1=Nc2ccc(Cl)cc2S(=O)(=O)N1	66.48	0.28	+
Diclofenac	OC(=O)Cc1ccccc1Nc1ccc(Cl)cc1Cl	48.95	4.44	+
Dicyclomine	CCN(CC)CCOC(=O)C1(CCCCC1)C1CCCCC1	29.54	4.40	+
Didanosine	OCC1CCC(O1)n1cnc2c1[nH]cnc2=O	92.95	1.89	+
Dienogest	CC12CCC3=C4CCC(=O)C=C4CCC3C1CCC2(O)CC#N	61.02	3.83	+

Diethylcarbamazine	CCN(CC)C(=O)N1CCN(C)CC1	26.79	-0.07	+
Diethylpropion	CCN(CC)C(C)C(=O)c1ccccc1	20.31	2.60	+
Diethylstilbestrol	CCC(=C(CC)c1ccc(O)cc1)c1ccc(O)cc1	40.46	0.67	+
Diflunisal	OC(=O)c1cc(ccc1O)-c1ccc(F)cc1F	57.28	1.91	+
Dihydrocodeine	COc1ccc2CC3C4CCC(O)C5Oc1c2C45CCN3C	41.93	1.12	+
Dihydroergotamine	CN1CC(CC2C1Cc1c[nH]c3cccc2c13)C(=O)NC1(C)OC2(O)C3CCCN3C(=O)C(Cc3ccccc3)N2C1=O	117.87	3.45	+
Diloxanide	CN(C(=O)C(Cl)Cl)c1ccc(O)cc1	40.54	3.38	+
Diphenoxylate	CCOC(=O)C1(CCN(CCC(C#N)(c2ccccc2)c2ccccc2)CC1)c1ccccc1	53.16	1.48	+
Disopyramide	CC(C)N(CCC(C(N)=O)(c1ccccc1)c1ccncn1)C(C)C	59.00	0.67	+
Distigmine	CN(CCCCCCN(C)C(=O)Oc1ccc[n+](C)c1)C(=O)Oc1ccc[n+](C)c1	68.94	4.05	-
Disulfiram	CCN(CC)C(=S)SSC(=S)N(CC)CC	121.24	0.32	+
Dofetilide	CN(CCOc1ccc(NS(C)(=O)=O)cc1)CCc1ccc(NS(C)(=O)=O)cc1	121.26	3.62	+
Dolasetron	O=C(OC1CC2CC3CC(C1)N2CC3=O)c1c[nH]c2ccccc12	62.30	0.54	+
Dothiepin	CN(C)CCC=C1c2ccccc2CSc2ccccc12	28.54	4.52	+
Doxapram	CCN1CC(CCN2CCOCC2)C(C1=O)(c1ccccc1)c1ccccc1	32.78	2.41	+
Doxazosin	COc1cc2nc(nc(N)c2cc1OC)N1CCN(CC1)C(=O)C1COc2ccccc2O1	112.07	4.01	+
Doxepin	CN(C)CCCC1c2ccccc2COc2ccccc12	12.47	3.90	+
Drotaverine	CCOc1ccc(C=C2NCCc3cc(OCC)c(OCC)cc23)cc1OCC	48.55	3.63	+
EDTA	OC(=O)CN(CCN(CC(O)=O)CC(O)=O)CC(O)=O	155.68	-2.07	-
Eflornithine	NCCCC(N)(C(F)F)C(O)=O	89.27	2.88	+
Emtricitabine	Nc1nc(=O)n(cc1F)C1CSC(CO)O1	115.67	-0.91	+
Enalaprilat	CCOC(=O)C(CCc1ccccc1)NC(C)C(=O)N1CCCC1C(O)=O	95.94	1.22	-
Encainide	COc1ccc(cc1)C(=O)Nc1ccccc1CCC1CCCCN1C	41.49	1.73	+
Enoxacin	CCn1cc(C(O)=O)c(=O)c2cc(F)c(nc12)N1CCNCC1	87.13	3.86	+
Enoximone	CSc1ccc(cc1)C(=O)c1[nH]c(=O)[nH]c1C	91.01	0.53	+
Entacapone	CCN(CC)C(=O)C(=Cc1cc(O)c(O)c(c1)[N+]([O-])=O)C#N	133.06	3.77	+
Entecavir	Nc1nc2n(cnc2c(=O)[nH]1)C1CC(O)C(CO)C1=C	129.72	-1.03	+
Enzalutamide	CNC(=O)c1ccc(cc1F)N1C(=S)N(C(=O)C1(C)C)c1ccc(C#N)c(c1)C(F)(F)F	108.26	3.42	+
Ephedrine	CNC(C)C(O)c1ccccc1	32.26	1.00	+
Epristeride	CC(C)(C)NC(=O)C1CCC2C3CC=C4C=C(CCC4(C)C3CCC12C)C(O)=O	65.78	1.27	+
Eprosartan	CCCCc1ncc(C=C(Cc2cccs2)C(O)=O)n1Cc1ccc(cc1)C(O)=O	120.66	4.64	-
Ergotamine	CN1CC(C=C2C1Cc1c[nH]c3cccc2c13)C(=O)NC1(C)OC2(O)C3CCCN3C(=O)C(Cc3cccc3)N2C1=O	118.21	0.94	+
Ertapenem	CC(O)C1C2C(C)C(SC3CNC(C3)C(=O)Nc3cccc(c3)C(O)=O)=C(N2C1=O)C(O)=O	181.57	-0.01	-
Estazolam	Clc1ccc-2c(c1)C(=NCc1nncn-21)c1ccccc1	43.07	3.05	+
Estramustine	CC12CCC3C(CCc4cc(OC(=O)N(CCCl)CCCl)ccc34)C1CCC2O	49.77	1.55	+
Ethacrynic_Acid	CCC(=O)C(=O)c1ccc(OCC(O)=O)c(Cl)c1Cl	80.57	5.18	+
Ethambutol	CCC(CO)NCCNC(CC)CO	64.43	1.66	+
Ethanol	CCO	20.23	0.00	+
Ethionamide	CCc1cc(ccn1)C(N)=S	70.91	1.34	+
Ethosuximide	CCC1(C)CC(=O)NC1=O	46.00	1.77	+
Ethylmorphine	CCN1CCC23C4OC5=C(O)C=CC(CC1C2C=CC4O)C35	52.90	1.71	+
Etidronate	CC(O)(P(O)(O)=O)P(O)(O)=O	154.91	-0.99	-
Etilefrine	CCNCC(O)c1cccc(O)c1	52.01	3.47	+
Etodolac	CCc1cccc2C3CCOC(CC(O)=O)(CC(O)=O)C3Nc12	95.58	1.81	+
Etoricoxib	Cc1ccc(cn1)-c1c(cc(Cl)c[n+]1[O-])-c1ccc(cc1)S(C)(=O)=O	80.39	2.90	+
Etravirine	Cc1cc(cc(C)c1Oc1nc(Nc2ccc(cc2)C#N)nc(N)c1Br)C#N	120.64	4.73	-
Ezetimibe	OC(CCC1C(N(C1=O)c1ccc(F)cc1)c1ccc(O)cc1)c1ccc(F)cc1	60.69	5.56	+
Ezogabine	CCOC(=O)Nc1ccc(NCc2ccc(F)cc2)cc1N	76.36	0.79	+
Febuxostat	CC(C)COc1ccc(cc1C#N)-c1nc(C)c(s1)C(O)=O	111.16	4.07	+
Felbamate	NC(=O)OCC(COC(N)=O)c1ccccc1	104.21	3.50	+
Felodipine	CCOC(=O)C1=C(C)NC(C)=C(C1c1cccc(Cl)c1Cl)C(=O)OC	64.52	-0.29	+
Fenclofenac	OC(=O)Cc1ccccc1Oc1ccc(Cl)cc1Cl	46.53	3.59	+
Fenfluramine	CCNCC(C)c1cccc(c1)C(F)(F)F	12.03	4.57	+
Fenofibrate	CC(C)OC(=O)C(C)(C)Oc1ccc(cc1)C(=O)c1ccc(Cl)cc1	52.49	0.71	+
Fenoprofen	CC(C(O)=O)c1cccc(Oc2ccccc2)c1	46.53	4.41	+
Fenoterol	CC(Cc1ccc(O)cc1)NC(O)Cc1cc(O)cc(O)c1	92.78	3.95	+
Fenspiride	O=C1NCC2(CCN(CCc3ccccc3)CC2)O1	41.57	3.79	+
Feprazone	CC(C)=CCC1C(=O)N(N(C1=O)c1ccccc1)c1ccccc1	40.62	2.61	+
Finasteride	CC(C)(C)NC(=O)C1CCC2C3CCC4NC(=O)C=CC4(C)C3CCC12C	57.78	2.82	+
Flecainide	FC(F)(F)COc1ccc(OCC(F)(F)F)c(c1)C(=O)NCC1CCCCN1	59.30	2.84	+
Fleroxacin	CN1CCN(CC1)c1c(F)cc2C(=O)C(CN(CCF)c2c1F)C(O)=O	63.83	0.72	+
Fluconazole	OC(Cn1cncn1)(Cn1cncn1)c1ccc(F)cc1F	81.47	1.34	+
Flucytosine	Nc1nc(=O)[nH]cc1F	71.77	1.03	+
Flumazenil	CCOC(=O)c1ncn-2c1CN(C)C(=O)c1cc(F)ccc-21	64.09	1.51	+
Flunarizine	Fc1ccc(cc1)C(N1CCN(CCCc2ccccc2)CC1)c1ccc(F)cc1	6.48	5.06	+
Flunisolide	CC1(C)OC2CC3C4CC(F)C5=CC(=O)C=CC5(C)C4C(O)CC3(C)C2(O1)C(=O)CO	93.06	2.72	+
Flunitrazepam	CN1c2ccc(cc2C(=NCC1=O)c1ccccc1F)[N+]([O-])=O	82.25	1.47	+
Fluocortolone	CC1CC2C3CC(F)C4=CC(=O)C=CC4(C)C3C(O)CC2(C)C1C(=O)CO	74.57	1.87	+
Fluoxetine	CNCCC(Oc1ccc(cc1)C(F)(F)F)c1ccccc1	21.26	5.26	+
Fluoxymesterone	CC1(O)CCC2C3CCC4=CC(=O)CCC4(C)C3(F)C(O)CC12C	57.53	3.88	+
Flupenthixol	OCCN1CCN(CCC=C2c3ccccc3Sc3ccc(cc23)C(F)(F)F)CC1	51.38	1.50	+
Flupirtine	CCOC(=O)Nc1ccc(NCc2ccc(F)cc2)nc1N	88.77	2.90	+
Flurazepam	CCN(CC)CCN1c2ccc(Cl)cc2C(=NCC1=O)c1ccccc1F	35.91	3.66	+
Flurbiprofen	CC(C(O)=O)c1ccc(cc1)-c1cccc(F)c1	37.30	4.10	+
Flutamide	CC(C)C(=O)Nc1ccc(c(c1)C(F)(F)F)[N+]([O-])=O	78.76	1.63	+
Fluvastatin	CC(C)n1c(C=CC(O)CC(O)CC(O)=O)c(-c2ccc(F)cc2)c2ccccc12	82.61	3.17	+
Fluvoxamine	COCCCCC(=NOCCN)c1ccc(cc1)C(F)(F)F	55.53	2.79	+
Fomepizole	Cc1cn[nH]c1	28.68	0.72	+
Fosfomycin	CC1OC1P(O)(O)=O	79.55	4.05	+
Fosmidomycin	ON(CCCP(O)(O)=O)C=O	107.51	3.76	+
Frovatriptan	CNC1CCc2[nH]c3ccc(cc3c2C1)C(N)=O	70.59	1.13	+
Galantamine	COc1ccc2CN(C)CCC34C=CC(O)CC3Oc1c24	41.93	1.34	+
Gallopamil	COc1ccc(CCN(C)CCCC(C#N)(C(C)C)c2cc(OC)c(OC)c(OC)c2)cc1OC	72.83	4.20	+
Gatifloxacin	COc1c(N2CCNC(C)C2)c(F)cc2c1n(cc(C(O)=O)c2=O)C1CC1	83.79	2.02	+
Gemfibrozil	Cc1ccc(C)c(OCCCC(C)(C)C(O)=O)c1	46.53	3.67	+
Genaconazole	CC(C(O)(Cn1cncn1)c1ccc(F)cc1F)S(C)(=O)=O	93.06	2.58	+
Gentamicin_C1	CNC(C)C1CCC(N)C(OC2C(N)CC(N)C(OC3OCC(C)(O)C(NC)C3O)C2O)O1	199.73	-3.33	-
Gentamicin_C1a	CCC1CCC(N)C(OC2C(N)CC(N)C(OC3OCC(C)(O)C(NC)C3O)C2O)O1	187.70	-2.53	-
Gestodene	CCC12CCC3C(CCC4=CC(=O)CCC34)C1C=CC2(O)C#C	37.30	3.74	+
Ginkgolide_A	CC1C(=O)OC2CC34C5CC(C(C)(C)C)C33C(O)C(=O)OC3OC4(C(=O)O5)C12O	127.56	0.01	+
Ginkgolide_B	CC1C(=O)OC2C(O)C34C5CC(C(C)(C)C)C33C(O)C(=O)OC3OC4(C(=O)O5)C12O	148.07	-1.91	+
Gitoxin	CC1OC(CC(O)C1O)OC1C(O)CC(OC2C(O)CC(OC3CCC4(C)C(CCC5C4CCC4(C)C(C(O)CC54O)C4=CC(=O)OC4)C3)OC2C)OC1C	180.08	1.52	+
Glibornuride	Cc1ccc(cc1)S(=O)(=O)NC(=O)NC1C(O)C2(C)CCC1C2(C)C	103.55	2.20	+
Gliclazide	Cc1ccc(cc1)S(=O)(=O)NC(=O)NN1CC2CCCC2C1	86.36	5.75	+
Glimepiride	CCC1=C(C)CN(C(=O)NCCc2ccc(cc2)S(=O)(=O)NC(=O)NC2CCC(C)CC2)C1=O	130.69	-0.25	+
Glipizide	Cc1cnc(cn1)C(=O)NCCc1ccc(cc1)S(=O)(=O)NC(=O)NC1CCCCC1	138.36	0.94	+
Gliquidone	COc1ccc2c(c1)C(=O)N(CCc1ccc(cc1)S(=O)(=O)NC(=O)NC1CCCCC1)C(=O)C2(C)C	130.05	-0.82	+
Glyceryl1Nitrate	OCC(O)CO[N+]([O-])=O	99.07	3.25	+

Glycopyrrolate	C[N+]1(C)CCC(C1)OC(=O)C(O)(C1CCCC1)c1ccccc1	46.53	1.97	-
Glymidine	COCCOc1cnc(NS(=O)(=O)c2ccccc2)nc1	98.24	5.69	+
Granisetron	CN1C2CCCC1CC(C2)NC(=O)c1nn(C)c2ccccc12	49.85	2.41	+
Guanabenz	NC(=N)NN=Cc1c(Cl)cccc1Cl	73.80	1.54	+
Guanadrel	NC(=N)NCC1COC2(CCCCC2)O1	79.87	-0.09	+
Guanfacine	NC(N)=NC(=O)Cc1c(Cl)cccc1Cl	81.26	2.84	+
Guanoxan	NC(=N)NCC1COc2ccccc2O1	80.36	0.55	+
HBED	OC(=O)CN(CCN(CC(O)=O)Cc1ccccc1O)Cc1ccccc1O	121.54	1.27	-
Hexobarbital	CN1C(O)=NC(=O)C(C)(C2=CCCCC2)C1=O	69.89	3.36	+
Hydralazine	NNc1nncc2ccccc12	63.60	2.29	+
Hydrocodone	COc1ccc2CC3C4CCC(=O)C5Oc1c2C45CCN3C	38.49	0.99	+
Hydroflumethiazide	NS(=O)(=O)c1cc2c(NCNS2(=O)=O)cc1C(F)(F)F	135.12	3.99	+
Hydroxychloroquine	CCN(CCO)CCC(C)Nc1ccnc2cc(Cl)ccc12	48.39	0.07	+
Hydroxyurea	NC(=O)NO	75.35	0.41	+
Hydroxyzine	OCCOCCN1CCN(CC1)C(c1ccccc1)c1ccc(Cl)cc1	35.94	1.97	+
Ibutilide	CCCCCCCN(CC)CCCC(O)c1ccc(NS(C)(=O)=O)cc1	77.82	2.54	+
Idazoxan	C1CN=C(N1)C1COc2ccccc2O1	42.59	2.45	+
Ifosfamide	ClCCNP1(=O)OCCCN1CCCl	51.38	1.50	+
Imidapril	CCOC(=O)C(CCc1ccccc1)NC(C)C(=O)N1C(CN(C)C1=O)C(O)=O	115.70	4.54	+
Imipenem	CC(O)C1C2CC(SCCN=CN)=C(N2C1=O)C(O)=O	141.52	-0.77	-
Indapamide	CC1Cc2ccccc2N1NC(=O)c1ccc(Cl)c(c1)S(N)(=O)=O	100.87	4.30	+
Indecainide	CC(C)NCCCC1(C(N)=O)c2ccccc2-c2ccccc12	54.74	2.62	+
Indobufen	CCC(C(O)=O)c1ccc(cc1)N1Cc2ccccc2C1=O	57.53	3.75	+
Indomethacin	COc1ccc2n(cc(CC(O)=O)c2c1)C(=O)c1ccccc1	68.36	1.54	+
Indoprofen	CC(C(O)=O)c1ccc(cc1)N1Cc2ccccc2C1=O	57.61	2.89	+
Iohexol	CC(=O)N(CC(O)CO)c1c(I)c(C(=O)NCC(O)CO)c(I)c(C(=O)NCC(O)CO)c1I	199.89	-1.63	-
Ipratropium	CC(C)[N+]1(C)C2CCC1CC(C2)OC(=O)C(CO)c1ccccc1	46.53	2.47	-
Irbesartan	CCCCC1=NC2(CCCC2)C(=O)N1Cc1ccc(cc1)-c1ccccc1-c1nn[nH]n1	87.07	-1.48	+
Isocarboxazid	Cc1cc(no1)C(=O)NNCc1ccccc1	66.91	2.00	+
Isoniazid	NNC(=O)c1ccncc1	67.59	2.01	+
Isotretinoin	CC(C=CC1=C(C)CCCC1(C)C)=CC=CC(C)=CC(O)=O	37.30	5.60	+
Isoxepac	OC(=O)Cc1ccc2OCc3ccccc3C(=O)c2c1	63.60	1.31	+
Isradipine	COC(=O)C1=C(C)NC(C)=C(C1c1cccc2nonc12)C(=O)OC(C)C	103.50	1.88	+
Itraconazole	CCC(C)n1ncn(-c2ccc(cc2)N2CCN(CC2)c2ccc(OCC3COC(Cn4cncn4)(O3)c3ccc(Cl)cc3Cl)cc2)c1=O	104.64	0.96	+
Kanamycin	CC(=O)NCC1OC(OC2C(N)CC(N)C(OC3OC(CO)C(O)C(N)C3O)C2O)C(O)C(O)C1O	285.69	-7.11	-
Kanamycin_A	COC1CC(OC2CCC3(C=O)C4CCC5(C)C(CCC5(O)C4CCC3(O)C2)C2=CC(=O)OC2)OC(O)C1OC1OC(COC2OC(CO)C(O)C(O)C2O)C(O)C(O)C1O	310.28	-2.99	-
Ketanserin	Fc1ccc(cc1)C(=O)C1CCN(CCn2c(=O)[nH]c3ccccc3c2=O)CC1	75.01	1.20	+
Ketazolam	CN1c2ccc(Cl)cc2C2(OC(C)=CC(=O)N2CC1=O)c1ccccc1	49.81	3.86	+
Ketorolac	OC(=O)C1CCn2c1ccc2C(=O)c1ccccc1	59.22	3.36	+
Ketotifen	CN1CCC(CC1)=C1c2ccsc2C(=O)Cc2ccccc12	48.39	3.20	+
Labetalol	CC(CCc1ccccc1)NCC(O)c1ccc(O)c(c1)C(N)=O	95.03	1.30	+
Lacidipine	CCOC(=O)C1=C(C)NC(C)=C(C1c1ccccc1C=CC(=O)OC(C)(C)C)C(=O)OCC	90.90	1.29	+
Lacosamide	COCC(NC(C)=O)C(=O)NCc1ccccc1	67.31	0.60	+
Leflunomide	Cc1oncc1C(=O)Nc1ccc(cc1)C(F)(F)F	55.12	1.40	+
Letrozole	N#Cc1ccc(cc1)C(c1ccc(cc1)C#N)n1cncn1	78.29	2.93	+
Levamisole	C1CN2CC(N=C2S1)c1ccccc1	40.62	1.45	+
Levetiracetam	CCC(N1CCCC1=O)C(N)=O	63.33	4.03	+
Levomepromazine	COc1ccc2Sc3ccccc3N(CC(C)CN(C)C)c2c1	40.90	1.06	+
Levonorgestrel	CCC12CCC3C(CCC4=CC(=O)CCC34)C1CCC2(O)C#C	37.30	3.96	+
Levoprotiline	CNCC(O)CC12CCC(c3ccccc13)c1ccccc21	32.26	3.18	+
Levosimendan	CC1CC(=O)NN=C1c1ccc(NN=C(C#N)C#N)cc1	112.27	0.96	+
Lidocaine	CCN(CC)CC(=O)Nc1c(C)cccc1C	32.34	2.39	+
Lincomycin	CCCC1CC(N(C)C1)C(=O)NC(C(C)O)C1OC(SC)C(O)C(O)C1O	147.79	-1.24	-
Linezolid	CC(=O)NCC1CN(C(=O)O1)c1ccc(N2CCOCC2)c(F)c1	71.00	1.67	+
Liothyronine	NC(Cc1cc(I)c(Oc2ccc(O)c(I)c2)c(I)c1)C(O)=O	91.76	0.10	+
Lisuride	CCN(CC)C(=O)NC1CN(C)C2Cc3c[nH]c4cccc(C2=C1)c34	50.80	2.30	+
Lofexidine	CC(Oc1c(Cl)cccc1Cl)C1=NCCN1	33.62	2.00	+
Lomefloxacin	CCn1cc(C(O)=O)c(=O)c2cc(F)c(N3CCNC(C)C3)c(F)c12	74.29	2.64	+
Lomustine	ClCCN(N=O)C(=O)NC1CCCCC1	61.77	1.16	+
Lorazepam	OC1N=C(c2ccccc2Cl)c2cc(Cl)ccc2NC1=O	61.44	1.53	+
Lorcainide	CC(C)N1CCC(CC1)N(C(=O)Cc1ccccc1)c1ccc(Cl)cc1	23.55	4.41	+
Lormetazepam	CN1c2ccc(Cl)cc2C(=NC(O)C1=O)c1ccccc1Cl	52.81	1.08	+
Lornoxicam	CN1C(C(=O)Nc2ccccn2)=C(O)c2sc(Cl)cc2S1(=O)=O	135.45	1.34	+
Lovastatin	CCC(C)C(=O)OC1CC(C)C=C2C=CC(C)C(CCC3CC(O)CC(=O)O3)C12	72.72	0.83	+
Loxiglumide	CCCCCN(CCCOC)C(=O)C(CCC(O)=O)NC(=O)c1ccc(Cl)c(Cl)c1	95.94	2.61	+
Loxoprofen	CC(C(O)=O)c1ccc(CC2CCCC2=O)cc1	54.35	1.68	+
Lynestrenol	CC12CCC3C(CCC4=CCCCC34)C1CCC2(O)C#C	20.23	4.39	+
Maprotiline	CNCCCC12CCC(c3ccccc13)c1ccccc21	12.03	4.21	+
Mazindol	OC1(N2CCN=C2c2ccccc12)c1ccc(Cl)cc1	35.83	1.74	+
Mebeverine	CCN(CCCCOC(=O)c1ccc(OC)c(OC)c1)C(C)Cc1ccc(OC)cc1	57.08	5.12	+
Meclofenamic_Acid	Cc1ccc(Cl)c(Nc2ccccc2C(O)=O)c1Cl	49.33	4.36	+
Mefenamic_Acid	Cc1cccc(Nc2ccccc2C(O)=O)c1C	49.33	4.74	+
Melagatran	NC(=N)c1ccc(CNC(=O)C2CCN2C(=O)C(NCC(O)=O)C2CCCCC2)cc1	148.61	0.28	-
Meloxicam	CN1C(C(=O)Nc2ncc(C)s2)=C(O)c2ccccc2S1(=O)=O	136.22	2.81	+
Meperidine	CCOC(=O)C1(CCN(C)CC1)c1ccccc1	29.54	1.83	+
Mepindolol	CC(C)NCC(O)COc1cccc2[nH]c(C)cc12	57.23	4.60	+
Meptazinol	CCC1(CCCCN(C)C1)c1cccc(O)c1	23.47	2.77	+
Mercaptopurine	S=c1nc[nH]c2[nH]cnc12	89.34	0.61	+
Meropenem	CC(O)C1C2C(C)C(SC3CCC(N3)C(=O)N(C)C)=C(N2C1=O)C(O)=O	135.48	-0.72	-
Mesalamine	Nc1ccc(O)c(c1)C(O)=O	83.18	3.47	+
Mesna	OS(=O)(=O)CCS	101.45	2.77	+
Mestranol	COc1ccc2C3CCC4(C)C(CCC4(O)C#C)C3CCc2c1	29.46	4.00	+
Metaproterenol	CC(C)NCC(O)c1cc(O)cc(O)c1	72.60	1.70	+
Methadone	CCC(=O)C(CC(C)N(C)C)(c1ccccc1)c1ccccc1	20.31	4.29	+
Methimazole	Cn1cc[nH]c1=S	52.60	4.68	+
Methocarbamol	COc1ccccc1OCC(O)COC(N)=O	90.93	3.91	+
Methoxyamphetamine	COc1cccc(CC(C)N)c1	35.25	1.58	+
Methsuximide	CN1C(=O)CC(C)(C1=O)c1ccccc1	37.38	0.95	+
Methylergonovine	CCC(CO)NC(=O)C1CN(C)C2Cc3c[nH]c4cccc(C2=C1)c34	68.25	1.56	+
Methylphenidate	COC(=O)C(C1CCCCN1)c1ccccc1	38.33	1.70	+
Methylphenobarbital	CCC1(C(=O)NC(=O)N(C)C1=O)c1ccccc1	66.40	3.26	+
Methylscopolamine	C[N+]1(C)C2CC(CC1C1OC21)OC(=O)C(CO)c1ccccc1	59.06	0.68	-
Metoclopramide	CCN(CC)CCNC(=O)c1cc(Cl)c(N)cc1OC	67.51	3.61	+
Metolazone	CC1Nc2cc(Cl)c(cc2C(=O)N1c1ccccc1C)S(N)(=O)=O	100.88	2.78	+
Metoprolol	COCCc1ccc(OCC(O)CNC(C)C)cc1	50.72	2.33	+
Metronidazole	Cc1ncc(n1CCO)[N+]([O-])=O	87.46	0.32	+
Metyrapone	CC(C)(C(=O)c1cccnc1)c1cccnc1	42.85	0.07	+
Mexiletine	CC(N)COc1c(C)cccc1C	35.25	2.03	+
Mezlocillin	CC1(C)SC2C(NC(=O)C(NC(=O)N3CCN(C3=O)S(C)(=O)=O)c3ccccc3)C(=O)N2C1C(O)=	207.18	-0.62	-

	O			
Mianserin	CN1CCN2C(C1)c1ccccc1Cc1ccccc21	6.48	2.00	+
Mibefradil	CCC(=O)OC1(CCN(C)CCCc2nc3ccccc3[nH]2)CCc2cc(F)ccc2C1C(C)C	58.20	-0.22	+
Miconazole	Clc1ccc(COC(Cn2ccnc2)c2ccc(Cl)cc2Cl)c(Cl)c1	27.05	5.98	+
Mifepristone	CN(C)c1ccc(cc1)C1CC2(C)C(CCC2(O)C#C)C2CCC3CC(=O)CCC3=C12	40.54	2.16	+
Mifobate	COP(=O)(CC(c1ccc(Cl)cc1)P(=O)(OC)OC)OC	90.65	0.84	+
Miglustat	CCCCN1CC(O)C(O)C(O)C1CO	84.09	2.40	+
Milnacipran	CCN(CC)C(=O)C1(CC1CN)c1ccccc1	46.33	2.79	+
Milrinone	Cc1[nH]c(=O)c(cc1-c1ccncc1)C#N	69.26	2.43	+
Minocycline	CN(C)C1C2CC3Cc4c(ccc(O)c4C(=O)C3=C(O)C2(O)C(=O)C(C(N)=O)=C1O)N(C)C	163.74	2.66	+
Minoxidil	Nc1nc(cc[n+]1[O-])N1CCCCC1	67.16	1.27	+
Mirtazapine	CN1CCN2C(C1)c1ccccc1Cc1cccnc21	19.37	1.39	+
Misoprostol	CCCCC(C)(O)CC=CC1C(O)CC(=O)C1CCCCCCC(=O)OC	83.80	1.58	+
Mivacurium	COc1cc2CC[N+](C)(CCCOC(=O)CCC=CCCC(=O)OCCC[N+]3(C)CCc4cc(OC)c(OC)cc4C3Cc3cc(OC)c(OC)c(OC)c3)C(Cc3cc(OC)c(OC)c(OC)c3)c2cc1OC	144.90	7.62	-
Moclobemide	Clc1ccc(cc1)C(=O)NCCN1CCOCC1	41.57	1.04	+
Modafinil	NC(=O)CS(=O)C(c1ccccc1)c1ccccc1	79.18	1.59	+
Molsidomine	CCOC(=O)Nc1c[n+](no1)N1CCOCC1	81.70	2.60	+
Moricizine	CCOC(=O)Nc1ccc2Sc3ccccc3N(C(=O)CCN3CCOCC3)c2c1	96.41	0.25	+
Moxalactam	COC1(NC(=O)C(C(O)=O)c2ccc(O)cc2)C2OCC(CSc3nnnn3C)=C(N2C1=O)C(O)=O	231.60	-1.51	-
Moxifloxacin	COc1c2C(NC3CC3)C=C(C(O)=O)C(=O)c2cc(F)c1N1CC2CCCNC2C1	90.68	4.03	+
Moxisylyte	CC(C)c1cc(OC(C)=O)c(C)cc1OCCN(C)C	38.77	1.55	+
Moxonidine	COc1nc(C)nc(Cl)c1NC1=NCCN1	71.11	0.78	+
Mycophenolic_Acid	COc1c(C)c2COC(=O)c2c(O)c1CC=C(C)CCC(O)=O	93.06	2.69	+
N_Acetylprocainamide	CCN(CC)CCNC(=O)c1ccc(NC(C)=O)cc1	61.36	1.16	+
Nabumetone	COc1ccc2cc(CCC(C)=O)ccc2c1	26.30	3.37	+
Nafarelin	CC(C)CC(NC(=O)C(Cc1ccc2ccccc2c1)NC(=O)C(Cc1ccc(O)cc1)NC(=O)C(CO)NC(=O)C(Cc1c[nH]c2ccccc12)NC(=O)C(Cc1c[nH]cn1)NC(=O)C1CCC(=O)N1)C(=O)NC(CCCNC(N)=N)C(=O)N1CCCC1C(=O)NCC(N)=O	472.13	-2.17	-
Nafronyl	CCN(CC)CCOC(=O)C(CC1CCCO1)Cc1ccc2ccccc2c1	38.77	2.98	+
Nalidixic_Acid	CCN1CC(C(O)=O)C(=O)c2ccc(C)nc12	70.18	2.45	+
Naloxone	Oc1ccc2CC3N(CC=C)CCC45C(Oc1c24)C(=O)CCC35O	69.97	0.64	+
Naltrexone	Oc1ccc2CC3N(CC4CC4)CCC45C(Oc1c24)C(=O)CCC35O	70.00	0.92	+
Naproxen	COc1ccc2cc(ccc2c1)C(C)C(O)=O	46.53	3.57	+
Naratriptan	CNS(=O)(=O)CCc1ccc2[nH]c(cc2c1)C1CCN(C)CC1	73.18	5.10	+
Nateglinide	CC(C)C1CCC(CC1)C(=O)NC(Cc1ccccc1)C(O)=O	66.40	5.10	+
Nedocromil	CCCc1c2oc(cc(=O)c2cc2c1n(CC)c(cc2=O)C(O)=O)C(O)=O	126.81	2.48	-
Nefazodone	CCc1nn(CCCN2CCN(CC2)c2cccc(Cl)c2)c(=O)n1CCOc1ccccc1	55.25	2.80	+
Nefopam	CN1CCOC(c2ccccc2)c2ccccc2C1	12.47	2.38	+
Neomycin	NCC1OC(COC2C(CO)OC(OC3C(O)C(N)CC(N)C3OC3OC(CN)C(O)C(O)C3N)C2O)C(N)C(O)C1O	353.11	-8.85	-
Netilmicin	CCNC1CC(N)C(OC2CC(N)C=C(CN)O2)C(O)C1OC1OCC(C)(O)C(NC)C1O	199.73	-3.20	-
Netivudine	CC#Cc1cn(C2OC(CO)C(O)C2O)c(=O)[nH]c1=O	124.78	-2.73	-
Nevirapine	Cc1ccnc2n(C3CC3)c3ncccc3c(=O)[nH]c12	63.40	-0.51	+
Niacin	OC(=O)c1cccnc1	50.16	1.94	+
Nicorandil	[O-][N+](=O)OCCNC(=O)c1cccnc1	100.88	2.84	+
Nicotine	CN1CCCC1c1cccnc1	16.13	1.14	+
Nilutamide	CC1(C)NC(=O)N(C1=O)c1ccc(c(c1)C(F)(F)F)[N+]([O-])=O	98.79	2.19	+
Nimodipine	COCCOC(=O)C1=C(C)NC(C)=C(C1c1cccc(c1)[N+]([O-])=O)C(=O)OC(C)C	122.91	0.67	+
Nisoldipine	COC(=O)C1=C(C)NC(C)=C(C1c1ccccc1[N+]([O-])=O)C(=O)OCC(C)C	113.43	0.41	+
Nitisinone	[O-][N+](=O)c1cc(ccc1C(=O)C1C(=O)CCCC1=O)C(F)(F)F	100.59	-1.39	+
Nitrazepam	[O-][N+](=O)c1ccc2NC(=O)CN=C(c3ccccc3)c2c1	91.12	2.50	+
Nitroxoline	Oc1ccc([N+]([O-])=O)c2cccnc12	82.69	4.94	+
Nomifensine	CN1CC(c2ccccc2)c2cccc(N)c2C1	29.26	2.32	+
Norethindrone	CC12CCC3C(CCC4=CC(=O)CCC34)C1CCC2(O)C#C	37.30	3.57	+
Norfenfluramine	CC(N)Cc1cccc(c1)C(F)(F)F	26.02	3.75	+
Norgestrel	CCC12C=CC3C(CCC4=CC(=O)CCC34)C1CCC2(O)C#C	37.30	3.74	+
Nortriptyline	CNCCC=C1c2ccccc2CCc2ccccc12	12.03	3.83	+
Octreotide	CC(O)C(CO)NC(=O)C1CSSCC(NC(=O)C(N)Cc2ccccc2)C(=O)NC(Cc2ccccc2)C(=O)NC(Cc2c[nH]c3ccccc23)C(=O)NC(CCCCN)C(=O)NC(C(C)O)C(=O)N1	382.82	-2.71	-
Ofloxacin	CC1COc2c(N3CCN(C)CC3)c(F)cc3c2n1cc(C(O)=O)c3=O	74.60	3.06	+
Olsalazine	OC(=O)c1cc(ccc1O)N=Nc1ccc(O)c(c1)C(O)=O	139.78	2.91	-
Omeprazole	COc1ccc2[nH]c(nc2c1)S(=O)Cc1ncc(C)c(OC)c1C	95.94	4.01	+
Ondansetron	Cn1c2CCC(Cn3ccnc3)C(=O)c2c2ccccc12	39.72	1.07	+
Orlistat	CCCCCCCCCCCC(CC1OC(=O)C1CCCCCC)OC(=O)C(CC(C)C)NC=O	81.70	6.88	-
Orphenadrine	CN(C)CCOC(c1ccccc1)c1ccccc1C	12.47	3.34	+
Oxamniquine	CC(C)NCC1CCc2cc(CO)c(cc2N1)[N+]([O-])=O	93.59	2.78	+
Oxandrolone	CC1(O)CCC2C3CCC4CC(=O)OCC4(C)C3CCC12C	46.53	3.04	+
Oxaprozin	OC(=O)CCc1nc(c(o1)-c1ccccc1)-c1ccccc1	63.32	0.84	+
Oxatomide	O=c1[nH]c2ccccc2n1CCCN1CCN(CC1)C(c1ccccc1)c1ccccc1	43.07	2.74	+
Oxazepam	OC1N=C(c2ccccc2)c2cc(Cl)ccc2NC1=O	61.69	2.15	+
Oxcarbazepine	NC(=O)N1c2ccccc2CC(=N)c2ccccc12	70.00	1.08	+
Oxitropium	CC[N+]1(C)C2CC(CC1C1OC21)OC(=O)C(CO)c1ccccc1	59.06	1.07	-
Oxprenolol	CC(C)NCC(O)COc1ccccc1OCC=C	50.72	1.61	+
Oxybutynin	CCN(CC)CC#CCOC(=O)C(O)(C1CCCCC1)c1ccccc1	49.77	5.18	+
Oxycodone	COc1ccc2CC3N(C)CCC45C(Oc1c24)C(=O)CCC35O	58.86	2.00	+
Oxyfedrine	COc1cccc(c1)C(=O)CCNC(C)C(O)c1ccccc1	58.36	1.35	+
Oxyphenbutazone	CCCCC1C(=O)N(N(C1=O)c1ccc(O)cc1)c1ccccc1	60.77	4.70	+
Palonosetron	O=C1N(CC2CCCc3cccc1c23)C1CN2CCC1CC2	23.55	1.89	+
Pamidronic_Acid	NCCC(O)(P(O)(O)=O)P(O)(O)=O	180.93	-1.66	-
Pancuronium	CC(=O)OC1C(CC2C3CCC4CC(OC(C)=O)C(CC4(C)C3CCC12C)[N+]1(C)CCCCC1)[N+]1(C)CCCCC1	52.60	5.35	-
Pantoprazole	COc1ccnc(CS(=O)c2nc3cc(OC(F)F)ccc3[nH]2)c1OC	104.70	4.71	+
Papaverine	COc1ccc(Cc2nccc3cc(OC)c(OC)cc23)cc1OC	49.77	3.31	+
Paricalcitol	CC(C=CC(C)C(C)(C)O)C1CCC2C(CCCC12C)=CC=C1CC(O)CC(O)C1	60.69	5.70	+
Paromomycin	NCC1OC(OC2C(CO)OC(OC3C(O)C(N)CC(N)C3OC3OC(CO)C(O)C(O)C3N)C2O)C(N)C(O)C1O	347.32	-8.86	-
Paroxetine	Fc1ccc(cc1)C1CCNCC1COc1ccc2OCOc2c1	38.91	2.70	+
Pefloxacin	CCn1cc(C(O)=O)c(=O)c2cc(F)c(cc12)N1CCN(C)CC1	64.79	2.75	+
Pelrinone	Cc1nc(NCc2cccnc2)c(C#N)c(=O)[nH]1	93.95	1.50	+
Penbutolol	CC(C)(C)NCC(O)COc1ccccc1C1CCCC1	41.49	2.15	+
Penciclovir	Nc1nc(=O)c2ncn(CCC(CO)CO)c2[nH]1	130.05	-1.30	-
Penicillin_G	CC1(C)SC2C(CC(=O)Cc3ccccc3)C(=O)N2C1C(O)=O	99.88	1.89	+
Penicillin_V	CC1(C)SC2C(NC(=O)COc3ccccc3)C(=O)N2C1C(O)=O	119.26	2.11	+
Pentamidine	NC(=N)c1ccc(OCCCCCOc2ccc(cc2)C(N)=N)cc1	118.20	2.88	-
Pentazocine	CC1C2Cc3ccc(O)cc3C1(C)CCN2CC=C(C)C	23.47	3.50	+
Pentobarbital	CCCC(C)C1(CC)C(=O)NC(=O)NC1=O	75.17	2.46	+
Pentoxifylline	CC(=O)CCCn1c(=O)n(C)c2ncn(C)c2c1=O	78.87	4.51	+
Perhexiline	C(C(C1CCCCC1)C1CCCCC1)C1CCCCN1	12.03	4.91	+
Perphenazine	OCCN1CCN(CCCN2c3ccccc3Sc3ccc(Cl)cc23)CC1	55.13	4.22	+
Phenacemide	NC(=O)NC(=O)Cc1ccccc1	72.19	0.98	+

Phenazopyridine	Nc1ccc(N=Nc2ccccc2)c(N)n1	89.45	1.02	+
Phencyclidine	C1CCN(CC1)C1(CCCCC1)c1ccccc1	3.24	3.84	+
Phenglutarimide	CCN(CC)CCC1(CCC(=O)NC1=O)c1ccccc1	49.33	3.75	+
Phenindione	O=C1C(C(=O)c2ccccc12)c1ccccc1	34.14	2.85	+
Pheniramine	CN(C)CCC(c1ccccc1)c1ccccn1	16.13	3.17	+
Phenprocoumon	CCC(c1ccccc1)c1c(O)c2ccccc2oc1=O	50.44	5.46	+
Phensuximide	CN1C(=O)CC(C1=O)c1ccccc1	37.38	0.78	+
Phenylbutazone	CCCCC1N(N(C(=O)C1=O)c1ccccc1)c1ccccc1	40.62	3.19	+
Phenylethylmalonamide	CCC(C(N)=O)(C(N)=O)c1ccccc1	86.05	3.41	+
Phenylpropanolamine	CC(N)C(O)c1ccccc1	46.17	0.07	+
Phenytoin	O=C1NNC(=O)C1(c1ccccc1)c1ccccc1	58.20	3.43	+
Phthalylsulfathiazole	OC(=O)c1ccccc1C(=O)Nc1ccc(cc1)S(=O)(=O)Nc1nccs1	162.08	3.59	-
Physostigmine	CNC(=O)Oc1ccc2N(C)C3N(C)CCC3(C)c2c1	44.81	3.53	+
Pimozide	Fc1ccc(cc1)C(CCN1CCC(CC1)n1c2ccccc2[nH]c1=O)c1ccc(F)cc1	41.01	4.11	+
Pindolol	CC(C)NCC(O)COc1cccc2[nH]ccc12	57.28	2.21	+
Pioglitazone	CCc1ccc(CCOc2ccc(CC3SC(=O)NC3=O)cc2)nc1	93.46	2.36	+
Pipemidic_Acid	CCn1cc(C(O)=O)c(=O)c2cnc(nc12)N1CCNCC1	99.98	1.57	+
Piperacillin	CCN1CCN(C(=O)NC(C(=O)NC2C3SC(C)(C)C(N3C2=O)C(O)=O)c2ccccc2)C(=O)C1=O	181.73	-1.71	-
Pirazolac	OC(=O)Cc1nn(cc1-c1ccc(Cl)cc1)-c1ccc(F)cc1	55.12	3.22	+
Pirbuterol	CC(C)(C)NCC(O)c1ccc(O)c(CO)n1	85.49	1.00	+
Piretanide	NS(=O)(=O)c1cc(cc(N2CCCC2)c1Oc1ccccc1)C(O)=O	118.21	0.85	+
Pirmenol	CC1CCCC(C)N1CCCC(O)(c1ccccc1)c1ccccn1	36.36	3.87	+
Piroxicam	CN1C(C(=O)Nc2ccccn2)=C(O)c2ccccc2S1(=O)=O	107.88	-0.60	+
Piroximone	CCc1[nH]c(=O)[nH]c1C(=O)c1ccncc1	78.29	2.66	+
Pizotyline	CN1CCC(CC1)=C1c2ccsc2CCc2ccccc12	31.48	3.99	+
Poldine	C[N+]1(C)CCCC1COC(=O)C(O)(c1ccccc1)c1ccccc1	46.53	2.22	-
Polythiazide	CN1C(CSCC(F)(F)F)Nc2cc(Cl)c(cc2S1(=O)=O)S(N)(=O)=O	149.31	0.11	+
Pomalidomide	Nc1cccc2C(=O)N(C3CCC(=O)NC3=O)C(=O)c12	108.71	2.41	+
Posaconazole	CCC(C(C)O)n1ncn(-c2ccc(cc2)N2CCN(CC2)c2ccc(OCC3COC(Cn4cncn4)(C3)c3ccc(F)cc3F)cc2)c1=O	115.67	-0.35	+
Practolol	CC(C)NCC(O)c1ccc(NC(C)=O)cc1	61.09	3.84	+
Pralidoxime	C[n+]1ccccc1C=NO	37.52	1.11	+
Pramipexole	CCCNC1CCc2nc(N)sc2C1	78.89	0.19	+
Pranlukast	O=C(Nc1cccc2c1oc(cc2=O)-c1nnn[nH]1)c1ccc(OCCCCc2ccccc2)cc1	123.00	4.44	-
Praziquantel	O=C(C1CCCCC1)N1CC2N(CCc3ccccc23)C(=O)C1	40.62	2.82	+
Pregabalin	CC(C)CC(CN)CC(O)=O	62.83	1.99	+
Primidone	CCC1(C(=O)NCNC1=O)c1ccccc1	58.20	0.37	+
Probenecid	CCCN(CCC)S(=O)(=O)c1ccc(cc1)C(O)=O	82.78	2.26	+
Probucol	CC(C)(Sc1cc(c(O)c(c1)C(C)(C)C)C(C)(C)C)Sc1cc(c(O)c(c1)C(C)(C)C)C(C)(C)C	91.06	9.91	-
Procainamide	CCN(CC)CCNC(=O)c1ccc(N)cc1	58.28	2.68	+
Procarbazine	CNNCc1ccc(cc1)C(=O)NC(C)C	53.09	4.16	+
Procyclidine	OC(CCN1CCCC1)(C1CCCCC1)c1ccccc1	23.47	3.45	+
Progesterone	CC(=O)C1CCC2C3CCC4=CC(=O)CCC4(C)C3CCC12C	34.14	4.72	+
Promethazine	CC(CN1c2ccccc2Sc2ccccc12)N(C)C	31.78	3.86	+
Propafenone	CCCNCC(O)COc1ccccc1C(=O)CCc1ccccc1	58.56	2.66	+
Propofol	CC(C)c1cccc(C(C)C)c1O	20.23	3.64	+
Propoxyphene	CCC(=O)OC(Cc1ccccc1)(C(C)CN(C)C)c1ccccc1	29.54	4.17	+
Propylthiouracil	CCCc1cc(=O)[nH]c(=S)[nH]1	80.67	2.62	+
Prothionamide	CCCc1cc(ccn1)C(N)=S	71.00	1.28	+
Protriptyline	CNCCCC1c2ccccc2C=Cc2ccccc12	12.03	4.08	+
Pseudoephedrine	CNC(C)C(O)c1ccccc1	32.26	1.00	+
Pyrazinamide	NC(=O)c1cnccn1	68.82	3.64	+
Pyridostigmine	CN(C)C(=O)Oc1ccc[n+](C)c1	34.47	1.37	-
Pyrimethamine	CCc1nc(N)nc(N)c1-c1ccc(Cl)cc1	77.76	4.48	+
Quinagolide	CCCN1CC(CC2Cc3c(O)cccc3CC12)NS(=O)(=O)N(CC)CC	80.74	1.39	+
Quinalbarbitone	CCCC(C)C1(CC=C)C(=C)NC(O)=NC1=O	61.69	1.50	+
Quinaprilat	CC(NC(CCc1ccccc1)C(O)=O)C(=O)N1Cc2ccccc2CC1C(O)=O	106.60	2.26	+
Quinine	COc1ccc2nccc(C(O)C3CC4CCN3CC4C=C)c2c1	45.40	3.09	+
Rabeprazole	COCCCOc1ccnc(CS(=O)c2nc3ccccc3[nH]2)c1C	96.31	3.61	+
Raloxifene	Oc1ccc(cc1)-c1sc2cc(O)ccc2c1C(=O)c1ccc(OCCN2CCCCC2)cc1	97.49	1.56	+
Ramelteon	CCC(=O)NCCC1CCc2ccc3OCCc3c12	38.33	2.57	+
Reboxetine	CCOc1ccccc1OC(C1CNCCO1)c1ccccc1	39.72	3.37	+
Recainam	CC(C)NCCCNC(=O)Nc1c(C)cccc1C	53.16	0.90	+
Remoxipride	CCN1CCCC1CNC(=O)c1c(OC)ccc(Br)c1OC	50.72	1.99	+
Repaglinide	CCOc1cc(CC(=O)NC(CC(C)C)c2ccccc2N2CCCCC2)ccc1C(O)=O	78.76	4.58	+
Reproterol	Cn1c2ncn(CCCNCC(O)c3cc(O)cc(O)c3)c2c(=O)n(C)c1=O	134.22	2.09	+
Rifabutin	COC1C=COC2(C)Oc3c(C2=O)c2C4=NC5(CCN(CC(C)C)CC5)NC4=C(NC(=O)C(C)=CC=CC(C)C(O)C(C)C(O)C(C)C(OC(C)=O)C1C)C(=O)c2c(O)c3C	203.06	2.22	+
Rilpivirine	Cc1cc(C=CC#N)cc(C)c1Nc1ccnc(Nc2ccc(cc2)C#N)n1	97.05	0.01	+
Rimiterol	OC(C1CCCCN1)c1ccc(O)c(O)c1	72.72	0.80	+
Ritodrine	CC(NCCc1ccc(O)cc1)C(O)c1ccc(O)cc1	72.72	0.57	+
Rivastigmine	CCN(C)C(=O)Oc1cccc(c1)C(C)N(C)C	32.78	2.44	+
Rizatriptan	CN(C)CCc1c[nH]c2ccc(Cn3cncn3)cc12	49.69	1.07	+
Rofecoxib	CS(=O)(=O)c1ccc(cc1)C1=C(C(=O)OC1)c1ccccc1	68.53	2.97	+
Ropinirole	CCCNC(Cc1cccc2NC(=O)Cc12)NCCC	53.16	2.62	+
Roquinimex	CN(C(=O)c1c(O)c2ccccc2n(C)c1=O)c1ccccc1	62.40	2.14	+
Rosiglitazone	CN(CCOc1ccc(CC2SC(=O)NC2=O)cc1)c1ccccn1	96.41	3.15	+
Rufinamide	NC(=O)c1cn(Cc2c(F)cccc2F)nn1	73.58	2.25	+
Ruxolitinib	N#CCC(C1CCCC1)n1cc(cn1)-c1ncnc2[nH]ccc12	83.14	0.97	+
Saccharin	O=C1NS(=O)(=O)c2ccccc12	71.43	-0.13	+
Salbutamol	CC(C)(C)NCC(O)c1ccc(O)c(CO)c1	72.72	2.03	+
Salsalate	OC(=O)c1ccccc1OC(=O)c1ccccc1O	83.83	3.95	+
Saxagliptin	NC(C(=O)N1C2CC2CC1C#N)C12CC3CC(CC(O)(C3)C1)C2	89.65	2.68	+
Scopolamine	CN1C2CC(CC1C1OC21)OC(=O)C(CO)c1ccccc1	62.21	2.42	+
Selegiline	CC(Cc1ccccc1)N(C)CC#C	3.24	2.26	+
Sematilide	CCN(CC)CCNC(=O)c1ccc(NS(C)(=O)=O)cc1	86.89	2.33	+
Sibutramine	CC(C)CC(N(C)C)C1(CCC1)c1ccc(Cl)cc1	3.24	4.74	+
Solifenacin	O=C(OC1CN2CCC1CC2)N1CCc2ccccc2C1c1ccccc1	32.78	2.78	+
Stavudine	Cc1cn(C2OC(CO)C=C2)c(=O)nc1N	90.35	0.78	+
Streptomycin	CNC1C(O)C(O)C(CO)OC1OC1C(OC2C(O)C(O)C(NC(N)=N)C(O)C2NC(N)=N)OC(C)C1(O)C=O	331.43	-7.74	-
Streptozocin	CN(N=O)C(=O)NC1C(O)OC(CO)C(O)C1O	151.92	-2.89	-
Succinylsulfathiazole	OC(=O)CCC(=O)Nc1ccc(cc1)S(=O)(=O)Nc1nccs1	162.08	2.45	-
Sudoxicam	CN1C(C(=O)Nc2nccs2)=C(O)c2ccccc2S1(=O)=O	136.22	2.46	+
Sufentanil	CCC(=O)N(c1ccccc1)C1(COC)CCN(CCc2cccs2)CC1	60.85	2.73	+
Sulbactam	CC1(C)C(N2C(CC2=O)S1(=O)=O)C(O)=O	100.13	-0.10	-
Sulfadiazine	Nc1ccc(cc1)S(=O)(=O)Nc1nccn1	105.54	4.44	+
Sulfamerazine	Cc1ccnc(NS(=O)(=O)c2ccc(N)cc2)n1	106.35	1.76	+
Sulfamethazine	Cc1cc(C)nc(NS(=O)(=O)c2ccc(N)cc2)n1	106.35	2.07	+
Sulfamethizole	Cc1nnc(NS(=O)(=O)c2ccc(N)cc2)s1	134.54	-0.77	+
Sulfamethoxazole	Cc1cc(NS(=O)(=O)c2ccc(N)cc2)no1	106.35	2.37	+

Sulfapyridine	Nc1ccc(cc1)S(=O)(=O)Nc1ccccn1	93.46	2.69	+
Sulfinpyrazone	O=C1C(CCS(=O)c2ccccc2)C(=O)N(N1c1ccccc1)c1ccccc1	76.72	3.33	+
Sulfisomidine	Cc1cc(NS(=O)(=O)c2ccc(N)cc2)nc(C)n1	106.35	2.37	+
Sulfisoxazole	Cc1nsc(NS(=O)(=O)c2ccc(N)cc2)c1C	121.57	3.76	+
Sulindac	CC1=C(CC(O)=O)c2cc(F)ccc2C1=Cc1ccc(cc1)S(C)=O	73.58	2.77	+
Sultopride	CCN1CCCC1CNC(=O)c1cc(ccc1OC)S(=O)(=O)CC	83.83	2.31	+
Sumatriptan	CNS(=O)(=O)Cc1ccc2[nH]cc(CCN(C)C)c2c1	73.58	5.54	+
Suprofen	CC(C(O)=O)c1ccc(cc1)C(=O)c1cccs1	82.33	2.62	+
Tacrine	Nc1c2CCCCc2nc2ccccc12	38.91	0.67	+
Tamsulosin	CCOc1ccccc1OCCNC(C)Cc1ccc(OC)c(c1)S(N)(=O)=O	108.17	-2.08	+
Tapentadol	CCC(C(C)CN(C)C)c1cccc(O)c1	23.47	3.08	+
Tegaserod	CCCCCN=C(N)NNC=C1C=Nc2ccc(CO)cc12	94.80	2.41	+
Teichomycin_A2_factor1	CCCCCC=CCCC(=O)NC1C(O)C(O)C(CO)OC1Oc1c2Oc3ccc(CC4NC(=O)C(N)c5ccc(O)c(Oc6cc(O)cc(c6)C(NC4=O)C(=O)NC4c(c2)cc1Oc1ccc(cc1Cl)C(OC1OC(CO)C(O)C(O)C1NC(C)=O)C1NC(=O)C(NC4=O)c2ccc(O)c(c2)-c2c(OC4OC(CO)C(O)C(O)C4O)cc(O)cc2C(NC1=O)C(O)=O)c5)cc3Cl	662.41	-3.29	-
Temazepam	CN1c2ccc(Cl)cc2C(=NC(O)C1=O)c1ccccc1	52.90	2.36	+
Temozolomide	Cn1nnc2c(ncn2c1=O)C(N)=O	107.98	2.09	+
Tenidap	NC(=O)N1C(=O)C(C(=O)c2cccs2)c2cc(Cl)ccc12	108.57	3.61	+
Terazosin	COc1cc2nc(nc(N)c2cc1OC)N1CCN(CC1)C(=O)C1CCCO1	102.29	4.07	+
Terbinafine	CN(CC=CC#CC(C)(C)C)Cc1cccc2ccccc12	3.24	4.81	+
Terlipressin	NCCCCC(NC(=O)C1CCCN1C(=O)C1CSSCC(NC(=O)CNC(=O)CNC(=O)CN)C(=O)NC(Cc2ccc(O)cc2)C(=O)NC(Cc2ccccc2)C(=O)NC(CCC(N)=O)C(=O)NC(CC(N)=O)C(=O)N1)C(=O)NCC(N)=O	563.45	-9.27	-
Terodiline	CC(CC(c1ccccc1)c1ccccc1)NC(C)(C)C	12.03	4.99	+
Tesaglitazar	CCOC(Cc1ccc(OCCc2ccc(OS(C)(=O)=O)cc2)cc1)C(O)=O	106.94	1.56	+
Testosterone	CC12CCC3C(CCC4=CC(=O)CCC34C)C1CCC2O	37.30	3.88	+
Tetrabenazine	COc1cc2CCN3CC(CC(C)C)C(=O)CC3c2cc1OC	38.77	4.45	+
Thiabendazole	c1nc(cs1)-c1nc2ccccc2[nH]1	69.64	0.95	+
Thioridazine	CSc1ccc2Sc3ccccc3N(CCC3CCCCN3C)c2c1	56.98	2.60	+
Tiacrilast	CSc1ccc2ncn(C=CC(O)=O)c(=O)c2c1	97.42	4.88	+
Tiagabine	Cc1ccsc1C(=CCCN1CCCC(C1)C(O)=O)c1sccc1C	96.83	2.11	+
Tianeptine	CN1c2ccccc2N(CCCCCCC(O)=O)c2ccc(Cl)cc2S1(=O)=O	86.18	0.31	+
Tibolone	CC1CC2=C(CCC(=O)C2)C2CCC3(C)C(CCC3(O)C#C)C12	37.30	3.96	+
Ticagrelor	CCCSc1nc(NC2CC2c2ccc(F)c(F)c2)c2nnn(C3CC(OCCO)C(O)C3O)c2n1	159.29	-1.28	+
Ticarcillin	CC1(C)SC2C(NC(=O)C(C(O)=O)c3ccsc3)C(=O)N2C1C(O)=O	177.55	0.17	-
Ticlopidine	Clc1ccccc1CN1CCc2sccc2C1	31.48	3.28	+
Tilidine	CCCC(=O)C1(CCC=CC1N(C)C)c1ccccc1	20.31	3.57	+
Tiludronic_Acid	OP(O)(=O)C(Cc1ccc(Cl)cc1)P(O)(O)=O	134.68	1.56	-
Tinidazole	CCS(=O)(=O)CCn1c(C)ncc1[N+]([O-])=O	109.57	-1.08	+
Tizanidine	Clc1ccc2nsnc2c1NC1=NCCN1	90.37	-0.74	+
Tobramycin	NCC1OC(OC2C(N)CC(N)C(OC3OC(CO)C(O)C(N)C3O)C2O)C(N)CC1O	268.17	-6.30	-
Tocainide	CC(N)C(=O)Nc1c(C)cccc1C	55.12	4.38	+
Tolazoline	C(C1=NCCN1)c1ccccc1	24.39	0.47	+
Toliprolol	CC(C)NCC(O)COc1cccc(C)c1	41.49	3.47	+
Tolmesoxide	COc1cc(C)c(cc1OC)S(C)=O	54.70	-0.09	+
Tolmetin	Cc1ccc(cc1)C(=O)c1ccc(CC(O)=O)n1C	59.30	2.29	+
Tolrestat	COc1ccc2c(ccc(c2c1)C(F)(F)F)C(=S)N(C)CC(O)=O	81.76	-0.28	+
Tolterodine	CC(C)N(CCC(c1ccccc1)c1cc(C)ccc1O)C(C)C	23.47	5.34	+
Topiramate	CC1(C)OC2COC3(COS(N)(=O)=O)OC(C)(C)OC3C2O1	123.52	3.00	+
Torsemide	CC(C)NC(=O)NS(=O)(=O)c1cnccc1Nc1cccc(C)c1	108.53	4.80	+
Tramadol	COc1cccc(c1)C1(O)CCCCC1CN(C)C	32.70	2.53	+
Tranexamic_Acid	NCC1CCC(CC1)C(O)=O	63.32	1.08	+
Trapidil	CCN(CC)c1cc(C)nc2ncnn12	46.31	2.65	+
Trazodone	Clc1cccc(c1)N1CCN(CCCn2nc3ccccn3c2=O)CC1	45.59	2.47	+
Treosulfan	CS(=O)(=O)OCC(O)C(O)COS(C)(=O)=O	138.53	3.16	+
Triamcinolone	CC12CC(O)C3(F)C(CCC4=CC(=O)C=CC34C)C1CC(O)C2(O)C(=O)CO	114.68	3.42	+
Triazolam	Cc1n[nH]c2cnc(-c3ccccc3Cl)c3cc(Cl)ccc3n12	45.78	1.60	+
Trifluoperazine	CN1CCN(CCCN2c3ccccc3Sc3ccc(cc23)C(F)(F)F)CC1	35.02	4.96	+
Trihexyphenidyl	OC(CCN1CCCCC1)(C1CCCCC1)c1ccccc1	23.47	3.84	+
Trimeprazine	CC(CN(C)C)CN1c2ccccc2Sc2ccccc12	31.78	4.11	+
Trimethobenzamide	COc1cc(cc(OC)c1OC)C(=O)NCc1ccc(OCCN(C)C)cc1	69.21	3.63	+
Trimipramine	CC(CN(C)C)CN1c2ccccc2CCc2ccccc12	6.48	3.74	+
Trofosfamide	ClCCN(CCCl)P1(=O)OCCCN1CCCl	41.93	1.32	+
Tropisetron	CN1C2CCC1CC(C2)OC(=O)c1c[nH]c2ccccc12	44.81	1.01	+
Trovafloxacin	NC1C2CN(CC12)c1nc2n(cc(C(O)=O)c(=O)c2cc1F)-c1ccc(F)cc1F	100.88	0.43	+
Tubocurarine	COc1cc2CC[N+](C)(C)C3Cc4ccc(O)c(Oc5cc6C(Cc7ccc(Oc(c1O)c23)cc7)[N+](C)(C)CCc6cc5OC)c4	77.38	5.44	-
Urapidil	COc1ccccc1N1CCN(CCCNc2cc(=O)n(C)c(=O)n2C)CC1	71.62	0.82	+
Ursodeoxycholic_Acid	CC(CCC(O)=O)C1CCC2C3C(O)CC4CC(O)CCC4(C)C3CCC12C	77.76	4.48	+
Valdecoxib	Cc1onc(c1-c1ccc(cc1)S(N)(=O)=O)-c1ccccc1	94.46	0.61	+
Valsartan	CCCCC(=O)N(Cc1ccc(cc1)-c1ccccc1-c1nnn[nH]1)C(C(C)C)C(O)=O	111.45	3.72	+
Vancomycin	CNC(CC(C)C)C(=O)NC1C(O)c2ccc(Oc3cc4cc(Oc5ccc(cc5Cl)C(O)C5NC(=O)C(NC(=O)C4NC(=O)C(CC(N)=O)NC1=O)c1ccc(O)c(c1)-c1c(O)cc(O)cc1C(NC5=O)C(O)=O)c3OC1OC(CO)C(O)C(O)C1OC1CC(C)(N)C(O)C(C)O1)c(Cl)c2	530.49	-3.42	-
Varenicline_tartrate	C1C2CNCC1c1cc3nccnc3cc21	37.81	1.42	+
Venlafaxine	COc1ccc(cc1)C(CN(C)C)C1(O)CCCCC1	32.70	3.04	+
Vidarabine	Nc1ncnc2n(cnc12)C1OC(CO)C(O)C1O	139.54	-2.30	-
Vilazodone	Cl.NC(=O)c1cc2cc(ccc2o1)N1CCN(CCCCc2c[nH]c3ccc(cc23)C#N)CC1	101.55	0.88	+
Vildagliptin	OC12CC3CC(C1)CC(C3)(C2)NCC(=O)N1CCCC1C#N	75.74	3.68	+
Viloxazine	CCOc1ccccc1OCC1COCCN1	39.72	2.49	+
Vismodegib	CS(=O)(=O)c1ccc(C(=O)Nc2ccc(Cl)c(c2)-c2ccccn2)c(Cl)c1	84.16	-1.84	+
Voriconazole	CC(c1ncncc1F)C(O)(Cn1cncn1)c1ccc(F)cc1F	76.71	2.48	+
Warfarin	CC(=O)CC(c1ccccc1)c1c(O)c2ccccc2oc1=O	67.43	0.30	+
Xamoterol	OC(CNCCNC(=O)N1CCOCC1)COc1ccc(O)cc1	103.29	-0.62	-
Ximoprofen	CC(C(O)=O)c1ccc(cc1)C1CCCC(C1)=NO	69.81	2.69	+
Xipamide	Cc1cccc(C)c1NC(=O)c1cc(c(Cl)cc1O)S(N)(=O)=O	117.87	1.52	+
Zaleplon	CCN(C(C)=O)c1cccc(c1)-c1ccnc2c(cnn12)C#N	74.26	1.81	+
Zanamivir	CC(=O)NC1C(NC(N)=N)C=C(OC1C(O)C(O)CO)C(O)=O	198.22	-3.58	-
Zatebradine	COc1ccc(CCN(C)CCCN2CCc3cc(OC)c(OC)cc3CC2=O)cc1OC	59.95	2.17	+
Zileuton	CC(N(O)C(N)=O)c1cc2ccccc2s1	94.57	1.55	+
Ziprasidone	Clc1cc2NC(=O)Cc2cc1CCN1CCN(CC1)c1nsc2ccccc12	76.69	2.55	+
Zolpidem	CCCN(CCC)C(=O)Cc1c(nc2ccc(C)cn12)-c1ccc(C)cc1	37.61	4.81	+
Zomepirac	Cc1cc(CC(O)=O)n(C)c1C(=O)c1ccc(Cl)cc1	59.30	2.19	+
Zonisamide	NS(=O)(=O)Cc1noc2ccccc12	94.57	4.05	+
Zopiclone	CN1CCN(CC1)C(=O)OC1N(C(=O)c2ncncn12)c1ccc(Cl)cn1	91.12	1.85	+
Zotepine	CN(C)CCOC1=Cc2ccccc2Sc2ccc(Cl)cc12	37.77	4.77	+
Zuclopenthixol	OCCN1CCN(CCC=C2c3ccccc3Sc3ccc(Cl)cc23)CC1	52.01	4.99	+

Table S2. BBB-dataset. BBB+ and BBB- correspond to brain permeant and non-permeant compounds, respectively.

Name	SMILES	tPSA [Å^2]	WLOGP	BBB
1-Butyl-3-phenylthiourea	CCCCNC(=S)Nc1ccccc1	56.15	2.58	+
1-hexyl-4-(4-iodophenyl)piperazine	CCCCCCN1CCN(CC1)c1ccc(I)cc1	6.48	3.23	+
1-hydroxyalprazolam	OCc1nnc2CN=C(c3ccccc3)c3cc(Cl)ccc3-n12	63.30	2.08	-
1-isobutyl-4-(4-iodophenyl)piperazine	CC(C)CN1CCN(CC1)c1ccc(I)cc1	6.48	2.31	+
1-pentanol	CCCCCO	20.23	1.17	+
1,2,3,4-tetrahydroquinoline	C1CNc2ccccc2C1	12.03	1.47	+
1,4-divinyloxybutane	C=COCCCCOC=C	18.46	2.09	+
2-Ethoxy-2-methylpropane (ETBE)	CCOC(C)(C)C	9.23	1.82	+
2-fluoroethyl-carbamimidothioate	NC(=N)SCCF	75.17	1.00	-
2-Methoxy-2-methylbutane	CCC(C)(C)OC	9.23	1.82	+
3-Hydroxy-3-phenylpentanamide (3-HEPP)	CCC(O)(CC(N)=O)c1ccccc1	63.32	1.05	+
4-hydroxyalprazolam	Cc1nnc2C(O)N=C(c3ccccc3)c3cc(Cl)ccc3-n12	63.30	2.37	-
4-phenylbenzoic_acid	OC(=O)c1ccc(cc1)-c1ccccc1	37.30	3.05	-
Abacavir	Nc1nc(NC2CC2)c2ncn(C3CC(CO)C=C3)c2n1	101.88	0.85	-
Acrylonitrile	C=CC#N	23.79	0.7	+
Acyclovir	Nc1nc2n(COCCO)cnc2c(=O)[nH]1	119.05	-1.48	-
Alfentanil	CCC(=O)N(c1ccccc1)C1(COC)CCN(CCn2nnn(CC)c2=O)CC1	85.49	1.00	-
Alniditan	C(CNCC1CCc2ccccc2O1)CNC1=NCCCN1	57.68	0.54	-
Alovudine	Cc1cn(C2CC(F)C(CO)O2)c(=O)[nH]c1=O	84.32	-0.44	-
Alprazolam	Cc1nnc2CN=C(c3ccccc3)c3cc(Cl)ccc3-n12	43.07	3.05	+
Alprenolol	CC(C)NCC(O)COc1ccccc1CC=C	41.49	2.15	-
Aminopyrine	CN(C)c1c(C)n(C)n(-c2ccccc2)c1=O	30.17	1.55	+
Amitriptyline	CN(C)CCC=C1c2ccccc2CCc2ccccc12	3.24	4.17	+
Amoxapine	Clc1ccc2Oc3ccccc3N=C(N3CCNCC3)c2c1	36.86	2.29	+
Amprenavir	CC(C)CN(CC(O)C(Cc1ccccc1)NC(=O)OC1CCOC1)S(=O)(=O)c1ccc(N)cc1	139.57	3.49	-
Antimalaric-ORG34167	NC(CC=C)c1ccccc1-c1noc2ccccc12	52.05	3.75	+
Apaxifylline	CCCn1c2nc([nH]c2c(=O)n(CCC)c1=O)C1CCC(=O)C1	89.75	1.54	-
AR-1D9859	CCCCN1CC2Cc3cc(OC)c(OC)cc3C2C1	21.7	2.69	+
Asenapine	CN1CC2C(C1)c1cc(Cl)ccc1Oc1ccccc21	12.47	3.88	+
Astemizole	COc1ccc(CCN2CCC(CC2)Nc2nc3ccccc3n2Cc2ccc(F)cc2)cc1	42.32	5.20	-
Astmizol-analog-NBI1	Cc1cnn(c1)C1CCN(CC1)c1nc2ccccc2n1Cc1ccc(F)cc1	38.88	4.61	+
Astmizol-analog-NBI10e	Fc1ccc(Cn2c(nc3ccccc23)C2CNCCS2)cc1	55.15	3.32	+
Astmizol-analog-NBI2	CN1CCCC(C1)c1nc2ccccc2n1Cc1ccc(F)cc1	21.06	4.07	+
Astmizol-analog-NBI4g	Fc1ccc(Cn2c(nc3ccccc23)C2CCCNC2)cc1	29.85	3.73	+
Astmizol-analog-NBI9a	Cc1ccc(Cn2c(nc3ccccc23)C2CNCCO2)cc1	39.08	2.35	+
Astmizol-analog-NBI9b	Fc1ccc(Cn2c(nc3ccccc23)C2CNCCO2)cc1	39.08	2.6	+
Atenolol	CC(C)NCC(O)COc1ccc(CC(N)=O)cc1	84.58	0.45	-
Atropine	CN1C2CCC1CC(C2)OC(=O)C(CO)c1ccccc1	49.77	1.55	+
AZ-11713908	CCOc1ccc(Cc2nc3cc(ccc3n2CC2CCCCC2)N(C)S(=O)(=O)c2cccs2)nc1	113.94	6.97	-
AZ-CB-inhibitor-29	CCS(=O)(=O)n1c2CCN(Cc2c2cc(ccc12)C(=O)N1CCC(C)CC1)C1CCOCC1	80.23	3.42	-
AZ-CB-inhibitor-5	CC1CCN(CC1)C(=O)c1ccc2n(CC=C)c3CCN(CC4CC4)Cc3c2c1	28.48	3.49	-
Beloxepin	CN1CCC2(O)C(C1)c1cccc(C)c1Oc1ccccc21	32.7	2.92	+
Beloxepin-analog-ORG32104	Cc1cccc2C3CNCCC3(O)c3ccccc3Oc12	41.49	2.58	+
Benzylidene-tert-butylamine-N-oxide	CC(C)(C)[N+]([O-])=Cc1ccccc1	32.59	2.83	+
Betaxolol	CC(C)NCC(O)COc1ccc(CCOCC2CC2)cc1	50.72	2.33	+
Biperiden	OC(CCN1CCCCC1)(C1CC2CC1C=C2)c1ccccc1	23.47	3.47	+
BMS-CDK2,7,9-inhibitor (SNS-032)	CC(C)(C)c1cnc(CSc2cnc(NC(=O)C3CCNCC3)s2)o1	133.59	2.94	-
BMS-Ibersartan-analog-1	CN1CN(c2ccc(Br)cc2)C2(CCN(CCCC(=O)c3ccc(F)cc3)CC2)C1=O	43.86	3.60	-
BMS-Ibersartan-analog-2	CN1CN(c2ccccc2)C2(CCN(CCCC(=O)c3csc(I)c3)CC2)C1=O	72.10	2.94	-
Bremazocine	CCC12CCN(CC3(O)CC3)C(Cc3ccc(O)cc13)C2(C)C	43.7	2.78	+
Bretazenil	CC(C)(C)OC(=O)c1ncn-2c1C1CCCN1C(=O)c1c(Br)cccc-21	64.43	3.18	+
Bretazenil-derivative-1	CC(C)c1nc(no1)-c1ncn-2c1CN(C)C(=O)c1c(Cl)cccc-21	77.05	2.75	-
Bretazenil-derivative-2 (di-OH)	CN1Cc2c(ncn2-c2cccc(Cl)c2C1=O)-c1noc(n1)C(C)(CO)CO	117.51	0.87	-
Bretazenil-derivative-3 (OH)	CN1Cc2c(ncn2-c2cccc(Cl)c2C1=O)-c1noc(n1)C(C)(C)O	97.28	1.75	-
Bretazenil-derivative-4	CN1Cc2c(ncn2-c2cccc(Cl)c2C1=O)-c1noc(n1)C(C)(O)CO	117.51	0.72	-
Budipine	CC(C)(C)N1CCC(CC1)(c1ccccc1)c1ccccc1	3.24	4.49	+
Bupropion	CC(NC(C)(C)C)C(=O)c1cccc(Cl)c1	29.1	3.3	+
Butamben	CCCCOC(=O)c1ccc(N)cc1	52.32	2.23	+
Butaperazine	CCCC(=O)c1ccc2Sc3ccccc3N(CCCN3CCN(C)CC3)c2c1	52.09	3.77	+
Butyl-acetate	CCCCOC(C)=O	26.3	1.35	+
Campthothecin	CCC1(O)C(=O)OCc2c1cc1-c3nc4ccccc4cc3Cn1c2=O	81.42	1.82	-
Carbon-disulfide	S=C=S	64.18	1.02	+
Carmustine	ClCCNC(=O)N(CCCl)N=O	61.77	1.16	+
Carteolol	CC(C)(C)NCC(O)COc1cccc2NC(=O)CCc12	70.59	1.13	-
Carvedilol-derivate	CC(CF)NCC(O)COc1cccc2[nH]c3ccccc3c12	57.28	3.43	+
Cerivastatin	COCc1c(nc(C(C)C)c(C=CC(O)CC(O)CC(O)=O)c1-c1ccc(F)cc1)C(C)C	99.88	5.04	-
Chlorpromazine	CN(C)CCCN1c2ccccc2Sc2ccc(Cl)cc12	31.78	4.51	+
Clobazam	CN1c2ccc(Cl)cc2N(c2ccccc2)C(=O)CC1=O	40.62	2.61	+
Clonidine-analog-14	Cc1ccc(NC2=NCCN2)c(C)c1	36.42	0.72	-
Clonidine-analog-2	Brc1cccc(Br)c1NC1=NCCN1	36.42	1.63	+
Clonidine-analog-3	Fc1cccc(F)c1NC1=NCCN1	36.42	1.22	-
Clozapine	CN1CCN(CC1)C1=Nc2cc(Cl)ccc2Nc2ccccc12	30.87	2.2	+
Cocaine	COC(=O)C1C2CCC(CC1OC(=O)c1ccccc1)N2C	55.84	1.49	+
Cortisol	CC12CC(O)C3C(CCC4=CC(=O)CCC34C)C1CCC2(O)C(=O)CO	94.83	1.78	-
D-Phenylalanine-L-Proline	NC(Cc1ccccc1)C(=O)ON1CCCC1C(O)=O	92.86	0.18	-
D4r-ligand-1	Clc1ccc(cc1)N1CCN(Cc2cnn3ccccc23)CC1	23.78	2.4	+
Daidzein	Oc1ccc(cc1)-c1coc2cc(O)ccc2c1=O	70.67	2.87	+
Dehydroevodiamine	C[n+]1c2C3Nc4ccccc4C3CCn2c(=O)c2ccccc12	38.96	2.38	+
Delavirdine	CC(C)Nc1cccnc1N1CCN(CC1)C(=O)c1cc2cc(NS(C)(=O)=O)ccc2[nH]1	118.81	2.65	-
Demonomethylchlorpromazine	CNCCCN1c2ccccc2Sc2ccc(Cl)cc12	40.57	4.17	+
Desipramine	CNCCCN1c2ccccc2CCc2ccccc12	15.27	3.15	+
Dextromethorphan	COc1ccc2CC3C4CCCCC4(CCN3C)c2c1	12.47	3	+
Diazepam	CN1c2ccc(Cl)cc2C(=NCC1=O)c1ccccc1	32.67	2.39	+
Diclazuril	Clc1ccc(cc1)C(C#N)c1c(Cl)cc(cc1Cl)-n1ncc(=O)[nH]c1=O	91.54	3.54	-
Didanosine	OCC1CCC(O1)n1cnc2c1[nH]cnc2=O	93.03	-0.53	-
Didemethylchlorpromazine	NCCCN1c2ccccc2Sc2ccc(Cl)cc12	54.56	3.91	+
Didesipramine	NCCCN1c2ccccc2CCc2ccccc12	29.26	2.89	+

Diphenhydramine	CN(C)CCOC(c1ccccc1)c1ccccc1	12.47	3.03	+
Dizocilpine	CC12NC(Cc3ccccc13)c1ccccc21	12.03	2.34	+
Donepezil	COc1cc2CC(CC3CCN(Cc4ccccc4)CC3)C(=O)c2cc1OC	38.77	3.83	+
Doxepin	CN(C)CCC=C1c2ccccc2COc2ccccc12	12.47	3.81	+
Doxylamine	CN(C)CCOC(C)(c1ccccc1)c1ccccn1	25.36	2.81	+
Edelfosine	CCCCCCCCCCCCCCCCCCCCOCC(COP(O)(=O)OCC[N+](C)(C)C)OC	84.03	7.12	-
Enaminone-derivative-DM44	COC(=O)C1C(C)CC(Nc2cccc(c2)[N+]([O-])=O)=CC1=O	105.06	2.51	-
Enaminone-derivative-DM49	COC(=O)C1C(C)CC(Nc2cccc(OC(F)(F)F)c2)=CC1=O	64.63	4.35	-
Enaminone-derivative-DM5 (ADD-196022)	COC(=O)C1C(C)CC(Nc2ccc(Cl)cc2)=CC1=O	55.4	2.84	+
Ergotamine	CN1CC(CC2C1CC1=C3C(NC=C1)C=CC=C23)C(=O)NC1(C)OC2(O)C3CCCN3C(=O)C(Cc3ccccc3)N2C1=O	114.45	0.03	-
Ethanol	CCO	20.23	0.00	+
Ethyl-acetate	CCOC(C)=O	26.3	0.57	+
Etodolac	CCc1cccc2c3CCOC(CC)(CC(O)=O)c3[nH]c12	62.32	3.27	-
Ezlopitant (Pfizer-NK1r-1)	COc1ccc(cc1CNC1C2CCN(CC2)C1C(c1ccccc1)c1ccccc1)C(C)C	24.5	5.67	+
Fentanyl	CCC(=O)N(C1CCN(CCc2ccccc2)CC1)c1ccccc1	23.55	3.76	+
Fluconazole	OC(Cn1cncn1)(Cn1cncn1)c1ccc(F)cc1F	81.65	1.47	-
Fluoromisonidazole	OC(COCF)Cn1ccnc1N(=O)=O	93.10	1.04	-
Fluroxene	FC(F)(F)COC=C	9.23	2.97	+
Fluroxene-analog	FC=COC=CF	9.23	2.72	+
Flutemetamol-iodo-analog	Nc1ccc(cc1I)-c1nc2ccc(O)cc2s1	87.38	3.86	+
Gabapentin	NCC1(CC(O)=O)CCCCC1	63.32	-1.37	-
Galantamine	COc1ccc2CN(C)CCC34C=CC(O)CC3Oc1c24	41.93	1.32	+
Galantamine-analog	COc1ccc2CN(C)CC3=CC(O)CC4Oc1c2C34	41.93	1.14	+
Gepirone-analog-ORG13011	FC(F)(F)c1ccnc(c1)N1CCN(CCCCN2CCCC2=O)CC1	39.68	2.63	+
Granisetron	CN1C2CCCC1CC(C2)NC(=O)c1nn(C)c2ccccc12	50.16	1.94	+
Guanidinothiazole-derivative-SKB-16	CCNC(NC#N)=NCCSCc1ncccc1Br	98.40	1.96	-
Guanidinothiazole-derivative-SKB-19	NC(N)=Nc1nc(cs1)-c1ccccc1	105.53	1.71	-
Guanidinothiazole-derivative-SKB-22	NC(N)=Nc1nc(cs1)-c1cccc(N)c1	131.55	1.31	-
Guanidinothiazole-derivative-SKB-24	CNC(Nc1cccc(c1)-c1csc(N=C(N)N)n1)=NC#N	165.74	0.99	-
Guanidinothiazole-derivative-SKB-26	CC(=O)Nc1cccc(c1)-c1csc(N=C(N)N)n1	134.63	1.48	-
GW-MCHr1-antagonist-1	CCc1ccc(cc1)-n1cnc2c(sc3nccc(N(C)CCc4ccccn4)c23)c1=O	92.15	4.63	-
GW-MCHr1-antagonist-2	CCc1ccc(cc1)-n1cnc2c(sc3nccc(N(C)C)c23)c1=O	79.26	3.62	-
GW-MCHr1-antagonist-7a	O=C(Cc1ccc(OCc2ccccc2)cc1)Nc1ccc2cnn(CCN3CCCC3)c2c1	59.39	4.17	-
GW-MCHr1-antagonist-8	Clc1ccc2oc(=O)cc(NC3CCN(Cc4ccc5OCCc5c4)CC3)c2c1	54.71	3.73	-
Histamine-analog-YG16	NCCc1nccs1	67.15	0.64	-
Histamine-analog-YG19	NCCc1nc(cs1)-c1ccccc1	67.15	2.31	-
Hydroxyzine	OCCOCCN1CCN(CC1)C(c1ccccc1)c1ccc(Cl)cc1	35.94	1.97	+
ICI-Tiotidine-analog-1	CN=C(NC#N)Nc1cccc(c1)-c1csc(N=C(N)N)n1	165.74	0.99	-
ICI-Tiotidine-analog-2	Cc1cnc(N=C(N)N)s1	105.53	0.36	-
Icotidine	COc1cccnc1CCCCNc1nc(=O)c(Cc2ccc(C)nc2)c[nH]1	92.79	2.71	-
Imipramine	CN(C)CCCN1c2ccccc2CCc2ccccc12	6.48	3.49	+
Iodoxylenediamine-1	Ic1ccc(CN2CCCCC2)cc1CN1CCCCC1	6.48	3.2	+
Iodoxylenediamine-2	CCC(C)Nc1ccc(I)c(NC(C)CC)c1	24.06	4.33	+
Isobutyl-acetate	CC(C)COC(C)=O	26.3	1.21	+
Isopentyl-acetate	CC(C)CCOC(C)=O	26.3	1.6	+
Isopropyl-acetate	CC(C)OC(C)=O	26.3	0.96	+
KCL-440	CN(C)Cc1ccc(cc1)-c1c[nH]c(=O)c2cccc(O)c12	56.33	2.81	+
Lapatinib	CS(=O)(=O)CCNCc1ccc(o1)-c1ccc2ncnc(Nc3ccc(OCc4cccc(F)c4)c(Cl)c3)c2c1	114.73	7.34	-
Lidocaine	CCN(CC)CC(=O)Nc1c(C)cccc1C	32.34	2.39	+
Lorazepam	OC1N=C(c2ccccc2Cl)c2cc(Cl)ccc2NC1=O	61.69	2.15	+
Loreclezole	ClC(=Cn1cncn1)c1ccc(Cl)cc1Cl	30.71	3.67	+
Loxapine	CN1CCN(CC1)C1=Nc2ccccc2Oc2ccc(Cl)cc12	28.07	2.63	+
Lupitidine	CN(C)Cc1ccc(CSCCNc2nc(=O)c(Cc3ccc(C)nc3)c[nH]2)o1	112.35	2.57	-
Lupitidine-analog-SKB26	CN(C)Cc1ccc(o1)-c1cccc(Nc2[nH]ccc2[N+]([O-])=O)c1	93.86	4.25	-
Lupitidine-analog-SKF93319	CN(C)Cc1ccc(CSCCNc2nc(=O)c(Cc3ccc4ccccc4c3)c[nH]2)o1	99.46	4.02	-
Lupitidinum	Cc1ccc(Cc2cnc(NCCSCc3ccc(o3)C(C)(C)N)[nH]c2=O)cn1	135.13	2.75	-
Mannitol	OCC(O)C(O)C(O)C(O)CO	121.38	-3.59	-
Maprotiline	CNCCCC12CCC(c3ccccc13)c1ccccc21	12.03	4.21	+
Maropitant (Pfizer-NK1r-antagonist-2)	COc1ccc(cc1CNC1C2CCN(CC2)C1C(c1ccccc1)c1ccccc1)C(C)(C)C	24.5	5.84	+
MCHr-antagonist-JNJ-1	CCN(CC)Cc1ccc2CC(CCc2c1)N1CCN(CCc2ccc(F)cc2)CC1=O	26.79	3.42	+
MCHr-antagonist-JNJ-19	Clc1ccc2[nH]nc(NC3CCN(Cc4ccc5OCOc5c4)CC3)c2c1	62.41	3.3	+
MCHr-antagonist-JNJ-2	Clc1ccc2oc(=O)cc(NC3CCN(CC=Cc4ccccc4)CC3)c2c1	45.48	4.36	+
MCHr-antagonist-JNJ-23c	CCc1ccc(cc1)-c1ccc(cc1)C(=O)N(C)C1CCN(C1)C(=O)N1CCC(C1)N1CCCCC1	47.1	3.6	+
MCHr-antagonist-JNJ-28	COc1ccc2[nH]nc(NC3CCN(Cc4ccc5OCOc5c4)CC3)c2c1	71.64	2.65	+
MCHr-antagonist-JNJ-7c	COc1cc(ccc1OCCN1CCCC1)-n1cnc2cc(sc2c1=O)-c1ccc(Cl)cc1	84.83	4.87	+
Meprobamate	CCCC(C)(COC(N)=O)COC(N)=O	104.64	0.98	-
Mepyramine	COc1ccc(CN(CCN(C)C)c2ccccn2)cc1	28.6	2.51	+
Methadone	CCC(=O)C(CC(C)N(C)C)(c1ccccc1)c1ccccc1	20.31	4.29	+
methoxyflurane	COC(F)(F)C(Cl)Cl	9.23	2.87	+
Methyl_acetate	COC(C)=O	26.30	0.18	-
Methyl-carbamimidothioate	CSC(N)=N	75.17	0.24	-
Methyl-tert-butyl-ether	COC(C)(C)C	9.23	1.43	+
mGluR1-inhibitor-AR1	CN(C)c1ccnc2sc3c(ncn(N4CCCCCC4)c3=O)c12	82.5	2.2	+
Midazolam	Cc1ncc2CN=C(c3ccccc3F)c3cc(Cl)ccc3-n12	30.18	4.21	+
Miloxacin	COn1cc(C(O)=O)c(=O)c2cc3OCOc3cc12	86.99	0.49	-
Mirtazapine	CN1CCN2C(C1)c1ccccc1Cc1cccnc21	19.37	1.39	+
Mitratapid	CCC(C)n1ncn(-c2ccc(cc2)N2CCN(CC2)c2ccc(OCC3COC(CSc4nncn4C)(O3)c3ccc(Cl)cc3)cc2)c1=O	130.00	4.68	-
Monodesmethylpromazine	CNCCCN1c2ccccc2Sc2ccccc12	40.57	3.52	+
MPPF	COc1ccccc1N1CCN(CCN(C(=O)c2ccc(F)cc2)c2ccccn2)CC1	48.91	3.36	+
N-methylthiazolium-DCKA	Cc1c(CCOC(=O)c2cc(=O)c3c(Cl)cc(Cl)cc3[nH]2)sc[n+]1C	92.33	4.22	-
Nalidixic_acid	CCn1cc(C(O)=O)c(=O)c2ccc(C)nc12	72.19	1.42	-
Naproxen	COc1ccc2cc(ccc2c1)C(C)C(O)=O	46.53	3.04	+
Naxagolide	CCCN1CCOC2C1CCc1ccc(O)cc21	32.7	1.79	+
Nebivolol	OC(CNCC(O)C1CCc2cc(F)ccc2O1)C1CCc2cc(F)ccc2O1	70.95	3.2	+
Nitrofurantoin	[O-][N+](=O)c1ccc(C=NN2CC(=O)NC2=O)o1	124.57	-0.27	-
Norclobazam	Clc1ccc2NC(=O)CC(=O)N(c3ccccc3)c2c1	49.41	2.39	+
Norharman	c1ccc2c(c1)[nH]c1cnccc21	28.68	2.72	+
Northioridazine	CSc1ccc2Sc3ccccc3N(CCC3CCCNC3)c2c1	65.87	4.64	+
Olanzapine	CN1CCN(CC1)C1=Nc2ccccc2Nc2sc(C)cc12	59.11	1.92	+
Omeprazole	COc1ccc2nc([nH]c2c1)S(=O)Cc1ncc(C)c(OC)c1C	96.31	3.61	-
Paliperidone	Cc1nc2C(O)CCCn2c(=O)c1CCN1CCC(CC1)c1noc2cc(F)ccc12	84.39	2.80	-
Pentazocine	CC1C2Cc3ccc(O)cc3C1(C)CCN2CC=C(C)C	23.47	3.5	+

Pentyl-acetate	CCCCCOC(C)=O	26.3	1.74	+
Pergolide	CCCN1CC(CSC)CC2C1Cc1c[nH]c3cccc2c13	44.33	3.89	+
Perphenazine	OCCN1CCN(CCCN2c3ccccc3Sc3ccc(Cl)cc23)CC1	55.25	2.8	+
Pfizer-NK1r-antagonist-15	COc1ccc(cc1CNC1CCCNC1c1ccccc1)-c1nccs1	74.42	3.54	+
Pfizer-NK1r-antagonist-17	COc1cc2sc(C)nc2cc1CNC1CCCNC1c1ccccc1	74.42	3.34	+
Pfizer-NK1r-antagonist-19	COc1cc2oc(C)nc2cc1CNC1CCCNC1c1ccccc1	59.32	2.87	+
Pfizer-NK1r-antagonist-20	COc1ccc(cc1CNC1CCCNC1c1ccccc1)-n1nc(C)cc1C	51.11	3.22	+
Pfizer-NK1r-antagonist-21	COc1cc2sc(nc2cc1CNC1CCCNC1c1ccccc1)C(C)(C)C	74.42	4.33	+
Pfizer-NK1r-antagonist-22	CC(C)Oc1cc2sc(C)nc2cc1CNC1CCCNC1c1ccccc1	74.42	4.12	+
Pfizer-NK1r-antagonist-23	COc1cc2sc(nc2cc1CNC1CCCNC1c1ccccc1)-c1ccccc1	74.42	4.7	+
Pfizer-NK1r-antagonist-24	COc1cc2sc(nc2cc1CNC1CCCNC1c1ccccc1)C1CC1	74.42	3.85	+
Pfizer-NK1r-antagonist-4	COc1ccccc1CNC1C2CCN(CC2)C1C(c1ccccc1)c1ccccc1	24.5	4.55	+
Pfizer-NK1r-antagonist-5	COc1ccccc1CNC1CCCNC1c1ccccc1	33.29	2.42	+
Pfizer-NK1r-antagonist-7	COc1ccc(cc1CNC1CCCNC1c1ccccc1)C(C)C	33.29	3.54	+
Pfizer-NK1r-antagonist-8	COc1ccc(cc1CNC1CCCNC1c1ccccc1)C(C)(C)C	33.29	3.72	+
Pfizer12-NK1r-antagonist	COc1ccc(cc1CNC1CCCNC1c1ccccc1)S(C)(=O)=O	75.81	2.91	-
Pfizer13-NK1r-antagonist	COc1cc2CCN(c2cc1CNC1C2CCN(C2)C1C(c1ccccc1)c1ccccc1)S(C)(=O)=O	70.26	4.18	-
Pfizer14-NK1r-antagonist	COc1cc2CC(C)N(c2cc1CNC1CCCNC1c1ccccc1)S(C)(=O)=O	79.05	2.83	-
Pfizer16-NK1r-antagonist	COc1ccc(cc1CNC1CCCNC1c1ccccc1)N1CCCC1=O	53.60	2.17	-
Pfizer18-NK1r-antagonist	COc1ccc(cc1CNC1CCCNC1c1ccccc1)N(c1nc(C)c(C)s1)S(C)(=O)=O	120.18	4.67	-
Phencyclidine	C1CCN(CC1)C1(CCCCC1)c1ccccc1	3.24	3.84	+
Phenylbutazone	CCCCC1C(=O)N(N(C1=O)c1ccccc1)c1ccccc1	40.62	3.03	-
Physostigmine	CNC(=O)Oc1ccc2N(C)C3N(C)CCC3(C)c2c1	44.81	1.01	+
POBN	CC(C)(C)[N+]([O-])=Cc1cc[n+]([O-])cc1	57.75	2.15	-
Practolol	CC(C)NCC(O)COc1ccc(NC(C)=O)cc1	70.59	1.19	-
Prazosin	COc1cc2nc(nc(N)c2cc1OC)N1CCN(CC1)C(=O)c1ccco1	106.95	1.03	-
Prednisolone	CC12CC(O)C3C(CCC4=CC(=O)C=CC34C)C1CCC2(O)C(=O)CO	94.83	1.56	-
Probenecid	CCCN(CCC)S(=O)(=O)c1ccc(cc1)C(O)=O	83.06	3.28	-
Procaine	CCN(CC)CCOC(=O)c1ccc(N)cc1	55.56	1.78	+
Promazine	CN(C)CCCN1c2ccccc2Sc2ccccc12	31.78	3.86	+
Propofol	CC(C)c1cccc(C(C)C)c1O	20.23	3.64	+
Propylhexedrine	CNC(C)CC1CCCCC1	12.03	2.56	+
Quercetin	OC1C(Cc2cc(O)cc(O)c2C1=O)c1ccc(O)c(O)c1	118.22	1.39	-
Quetiapine	OCCOCCN1CCN(CC1)C1=Nc2ccccc2Sc2ccccc12	73.6	1.71	+
Racemorphan	CN1CCC23CCCCC2C1Cc1ccc(O)cc31	23.47	2.7	+
Ribavirin	NC(=O)c1ncn(n1)C1OC(CO)C(O)C1O	143.72	-3.34	-
Ridogrel	OC(=O)CCCCON=C(c1cccnc1)c1cccc(c1)C(F)(F)F	71.78	5.28	-
Risocaine	CCCOC(=O)c1ccc(N)cc1	52.32	1.84	+
Rivastigmine	CCN(C)C(=O)Oc1cccc(c1)C(C)N(C)C	32.78	2.44	+
RO-cyclic-guanidine-analog1	CC1NC(N)=Nc2cccc(Cl)c12	50.41	0.86	-
RO-cyclic-guanidine-analog2	CC1NC(NCC(F)F)=Nc2cccc(Cl)c12	36.42	2.6	+
RO19-4603	CN1Cc2c(ncn2-c2ccsc2C1=O)C(=O)OC(C)(C)C	92.67	1.94	-
Rofecoxib	CS(=O)(=O)c1ccc(cc1)C1=C(C(=O)OC1)c1ccccc1	68.82	3.64	+
Rolipram	COc1ccc(cc1OC1CCCC1)C1CNC(=O)C1	47.56	2.24	+
Sabeluzole	CN(C1CCN(CC(O)COc2ccc(F)cc2)CC1)c1nc2ccccc2s1	77.07	3.82	+
Salicyluric_acid	OC(=O)CNC(=O)c1ccccc1O	86.63	0.21	-
Serotonin-transporter-liagnd-1	CN(C)Cc1ccccc1Sc1ccc(F)cc1N	54.56	3.9	+
Serotonin-transporter-liagnd-2	CN(C)c1ccccc1Sc1ccc(CCF)cc1N	54.56	4.43	+
Sertraline	CNC1CCC(c2ccc(Cl)c(Cl)c2)c2ccccc12	12.03	4.86	+
SKF-H2-antagonist	CN(C)c1cc(ccn1)-c1n[nH]c(N)n1	83.72	0.52	-
Sotalol	CC(C)NCC(O)c1ccc(NS(C)(=O)=O)cc1	86.81	1.66	-
Stavudine	Cc1cn(C2OC(CO)C=C2)c(=O)[nH]c1=O	84.32	-1.03	-
Sulphasalazine	OC(=O)c1cc(ccc1O)N=Nc1ccc(cc1)S(=O)(=O)Nc1ccccn1	149.69	4.59	-
Talnetant-analog	CCC(NC(=O)c1c(C)c(nc2ccccc12)-c1ccccc1)c1ccccc1	41.99	5.77	+
Talsupram	CNCCCC1(SC(C)(C)c2ccccc12)c1ccccc1	37.33	4.69	+
Temelastine	Cc1ccc(Cc2c[nH]c(NCCCCc3ncc(Br)cc3C)nc2=O)cn1	83.56	3.77	-
Tert-butanol	CC(C)(C)O	20.23	0.78	+
Tert-butyl-chlorambucil	CC(C)(C)OC(=O)CCCc1ccc(cc1)N(CCCl)CCCl	29.54	4.64	+
Testosterone	CC12CCC3C(CCC4=CC(=O)CCC34C)C1CCC2O	37.3	3.88	+
Tetrahydropalmatine	COc1cc2CCN3Cc4c(CC3c2cc1OC)ccc(OC)c4OC	40.16	2.52	+
Thiamine-Propyl-Disulfide-DCKA	CC(C)SSC(CCOC(=O)c1cc(=O)c2c(Cl)cc(Cl)cc2[nH]1)=C(C)N(Cc1cnc(C)nc1N)C=O	181.87	5.60	-
Thioperamide	S=C(NC1CCCCC1)N1CCC(CC1)c1c[nH]cn1	76.04	2.42	+
Tibolone	CC1CC2=C(CCC(=O)C2)C2CCC3(C)C(CCC3(O)C#C)C12	37.3	3.96	+
Ticalopride	COC1CNCCC1NC(=O)c1cc(Cl)c(N)cc1OC	85.61	0.66	-
Tiotidine	CNC(NCCSCc1csc(N=C(N)N)n1)=NC#N	191.04	-0.22	-
Tivirapine-derivate	CC1Cn2c3c(CN1CC=C(C)C)cc(Cl)cc3[nH]c2=S	56.05	3.99	+
Tolbutamide	CCCCNC(=O)NS(=O)(=O)c1ccc(C)cc1	83.65	2.86	-
Toliprolol	CC(C)NCC(O)COc1cccc(C)c1	41.49	1.73	+
Topiramate	CC1(C)OC2COC3(COS(N)(=O)=O)OC(C)(C)OC3C2O1	123.92	0.69	-
Tramadol	COc1cccc(c1)C1(O)CCCCC1CN(C)C	32.7	2.53	+
Tranylcypromine-analog-1	NC1CC1c1ccc(F)cc1	26.02	2.06	+
Triazolam	Cc1nnc2CN=C(c3ccccc3Cl)c3cc(Cl)ccc3-n12	43.07	3.7	+
Trimipramine	CC(CN(C)C)CN1c2ccccc2CCc2ccccc12	6.48	3.74	+
urea	NC(N)=O	69.11	-0.98	-
Vinyl-ether	C=COC=C	9.23	1.29	+
Zalcitabine	Nc1ccn(C2CCC(CO)O2)c(=O)n1	90.37	-0.82	-
Zalcitabine-analog	Nc1ccn(C2CC(CO)C(CO)O2)c(=O)n1	110.60	-1.60	-
Zaltidine	Cc1nc(c[nH]1)-c1csc(NC(N)=N)n1	131.71	0.96	-
Zanapezil	O=C(CCC1CCN(Cc2ccccc2)CC1)c1ccc2CCCCNc2c1	32.34	4.21	+
Zidovudine	Cc1cn(C2CC(N=[N+]=N)C(CO)O2)c(=O)[nH]c1=O	120.53	-1.28	-
Zolantidine	C(CNc1nc2ccccc2s1)COc1cccc(CN2CCCCC2)c1	65.63	4.44	+

Table S3. Cross-validation for HIA ellipse.
The 660 entries of the HIA-datasets were sorted according to tPSA and every tenth entry was attributed to the validation set. The remaining entries are considered as the training set. This procedure was repeated 10 times starting from entries 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10. Ten different couples of training/validation sets were so generated.

	number of molecules (training/validation)	MCC_{tr}	ACC_{tr}	MCC_{va}	ACC_{va}
1	595/65	0.6957	0.9265	0.7421	0.9385
2	595/65	0.7064	0.9296	0.6948	0.9254
3	594/66	0.6981	0.9280	0.6948	0.9254
4	593/67	0.6924	0.9263	0.7649	0.9403
5	593/67	0.7064	0.9296	0.6474	0.9104
6	593/67	0.7091	0.9296	0.5298	0.8806
7	594/66	0.7012	0.9281	0.6940	0.9242
8	594/66	0.7233	0.9331	0.4854	0.8788
9	594/66	0.7041	0.9314	0.6140	0.9091
10	594/66	0.6977	0.9298	0.6641	0.9242
	Cross-validation			**MCC_{cv} = 0.6531**	**ACC_{CV} = 0.9156**

MCC_{tr}: Matthews correlation coefficient for training set; ACC_{tr}: accuracy for training set; MCC_{va}: Matthews correlation coefficient for validation set; ACC_{va}: accuracy for validation set; MCC_{cv}: Cross-validated Matthews correlation coefficient; ACC_{CV}: Cross-validated accuracy

Table S4. Cross-validation for BBB ellipse.
The 260 entries of the BBB-datasets were sorted according to tPSA and every tenth entry was attributed to the validation set. The remaining entries are considered as the training set. This procedure was repeated 10 times starting from entries 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10. Ten different couples of training/validation sets were so generated.

	number of molecules (training/validation)	MCC_{tr}	ACC_{tr}	MCC_{va}	ACC_{va}
1	235/25	0.7796	0.8936	0.8402	0.9200
2	233/27	0.7789	0.8927	0.8466	0.9259
3	233/27	0.7949	0.9013	0.7135	0.8519
4	233/27	0.7875	0.8970	0.7689	0.8889
5	233/27	0.7961	0.9013	0.5330	0.7778
6	234/26	0.7990	0.9017	0.6750	0.8462
7	234/26	0.7890	0.8974	0.7542	0.8846
8	235/25	0.7895	0.8979	0.8402	0.9200
9	235/25	0.7980	0.9021	0.6651	0.8400
10	235/25	0.7710	0.8894	0.8333	0.9200
	Cross-validation			**MCC_{cv} = 0.7470**	**ACC_{CV} = 0.8775**

MCC_{tr}: Matthews correlation coefficient for training set; ACC_{tr}: accuracy for training set; MCC_{va}: Matthews correlation coefficient for validation set; ACC_{va}: accuracy for validation set; MCC_{cv}: Cross-validated Matthews correlation coefficient; ACC_{CV}: Cross-validated accuracy

Table S5. FDA-dataset.

46 non-prodrug new chemical entities (NCEs) with unambiguous oral bioavailability accepted by the FDA between January 2014 and September 2015. HIA+ and HIA- entries correspond to orally bioavailable and not bioavailable molecules, respectively. Entries in italic are part of the HIA-dataset used for training.
Correctly classified compounds are in black, misclassified compounds are in red (accuracy 83%). Drugs with enough evidence of brain access are tagged as BBB+.

Name	tPSA [Å^2]	WLOGP	HIA	BBB
Apremilast	127.46	2.62	+	
Avibactam	138.62	-1.33	-	
Belinostat	103.88	2.79	+	
Brexpiprazole	76.81	3.96	+	+
Canagliflozin	118.39	3.06	+	
Cangrelor[a]	335.94	3.6	-	
Ceftazidime[a]	244.76	-1.67	-	
Ceritinib	125.65	6.84	+	
Cholic acid	97.99	3.45	+	
Daclatasvir	174.64	4.81	+	
Dapagliflozin	99.38	1.52	+	
Deoxycholic acid	77.76	4.48	+	
Droxidopa	124.01	-0.78	+	
Edoxaban	164.87	0.57	+	
Eliglustat	74.52	3.57	+	+
Eluxadoline	164.63	3.68	+	
Empagliflozin	108.61	1.29	+	
Finafloxacin	107.59	0.82	+	
Flibanserin	44.27	3.56	+	+
Ibrutinib	99.16	3.84	+	+
Idelalisib	101.38	3.66	+	
Ivabradine	60.47	2.93	+	+
Ivacaftor	82.19	4.89	+	
Ledipasvir	174.64	7.87	+	
Lenvatinib	96.3	4.26	+	
Lumacaftor	97.75	5.33	+	
Metformin	88.99	-1.03	+	
Naloxegol	126.77	1.48	+	
Netupitant	39.68	6.79	+	+
Nintedanib	97.71	2.93	+	
Olaparib	86.63	2.36	+	
Olodaterol	103.54	2.32	-	
Oritavancin[a]	560.98	0.82	-	
Palbociclib	105.04	2.2	+	
Palonosetron	23.55	1.89	+	+
Panobinostat	77.15	3.07	+	+
Peramivir	151.03	-0.35	-	
Pirfenidone	22	2.15	+	+
Rolapitant	50.36	5.39	+	
Sofosbuvir	167.99	1.75	+	
Sonidegib	63.69	4.86	+	+
Suvorexant	80.29	3.35	+	+
Tasimelteon	38.33	2.25	+	+
Tavaborolole	29.46	-0.66	-	
Valsartan	112.07	4.01	+	
Vorapaxar	81.01	5.83	+	

[a]Cangrelor, Ceftazidime and Oritavancin are outside the range of the plot in Figure 2a.

Data S1. Optimization of the HIA ellipse (normalized).

The final ellipse in the referential of tPSA and WLOGP (normalized towards standard deviation) able to discriminate best the 93 poorly from the 567 well-absorbed molecules as defined in the HIA-dataset (Table S1) was determined by Monte-Carlo optimization (MC) of five parameters: the Cartesian coordinates of both foci (x_1, x_2); (y_1, y_2) and the major axis (or largest diameter, d). By geometrical definition, the sum of distances from every point on the ellipse's trace to each focus is equal to d.

About 10,000 5-dimensions starting vectors of (x_1, y_1, x_2, y_2, d) were submitted to 10 independent runs of 100,000-step MC, involving a Metropolis-like criterion (refer to Methods S6).

The best ellipse was optimized (with a small constraint on the surface) until reaching the following parameters (in the normalized referential):

x_1: 0.711 ; y_1: 2.736 (focus 1)

x_2: 1.405 ; y_2: -0.720 (focus 2)

d: 4.052

The classification power is:

TP[a] = 547

TN[b] = 67

FP[c] = 26

FN[d] = 20

MCC[e] = 0.7047

ACC[f] = 0.9303

Refer to Figure S1 for the HIA ellipse further described and depicted in the actual (denormalized) tPSA *versus* WLOGP referential.

---

[a]true positives; [b]true negatives; [c]false positives; [d]false negatives; [e]Matthews correlation coefficient; [f]accuracy

Data S2. Optimization of the BBB ellipse (normalized).

The final ellipse in the referential of tPSA and WLOGP (normalized towards standard deviation) able to discriminate best the 104 brain non-penetrant from the 156 brain penetrant molecules as defined in the BBB-dataset (Table S2) was determined by MC of five parameters: the Cartesian coordinates of both foci (x_1, x_2); (y_1, y_2) and the major axis (or largest diameter, d). By geometrical definition, the sum of distances from every point on the ellipse's trace to each focus is equal to d.

By geometrical definition, the sum of distances from every point on the ellipse's trace to each focus is equal to d.

About 10,000 5-dimensions starting vectors of (x_1, y_1, x_2, y_2, d) were submitted to 10 independent runs of 100,000-step MC, involving a Metropolis-like criterion (refer to Methods S6).

The best ellipse was optimized (with a small constraint on the surface) until reaching the following parameters (in the normalized referential):

x_1: 0.978 ; y_1: 3.261 (focus 1)

x_2: 1.098; y_2: 0.664 (focus 2)

d: 3.428

The classification power is:

TP[a] = 151

TN[b] = 82

FP[c] = 22

FN[d] = 5

MCC[e] = 0.7853

ACC[f] = 0.8962

Refer to Figure S2 for the BBB ellipse further described and depicted in the actual (denormalized) tPSA *versus* WLOGP referential.

---

[a]true positives; [b]true negatives; [c]false positives; [d]false negatives; [e]Matthews correlation coefficient; [f]accuracy

Data S3 (separated Excel file). Cartesian coordinates of the BOILED-Egg's ellipse traces. Hundred points of the HIA and BBB traces are connected to depict both ellipses. The user of the Excel spreadsheet have the possibility to enter WLOGP and tPSA values for up to 100 molecules and the corresponding points will be mapped onto the plot.

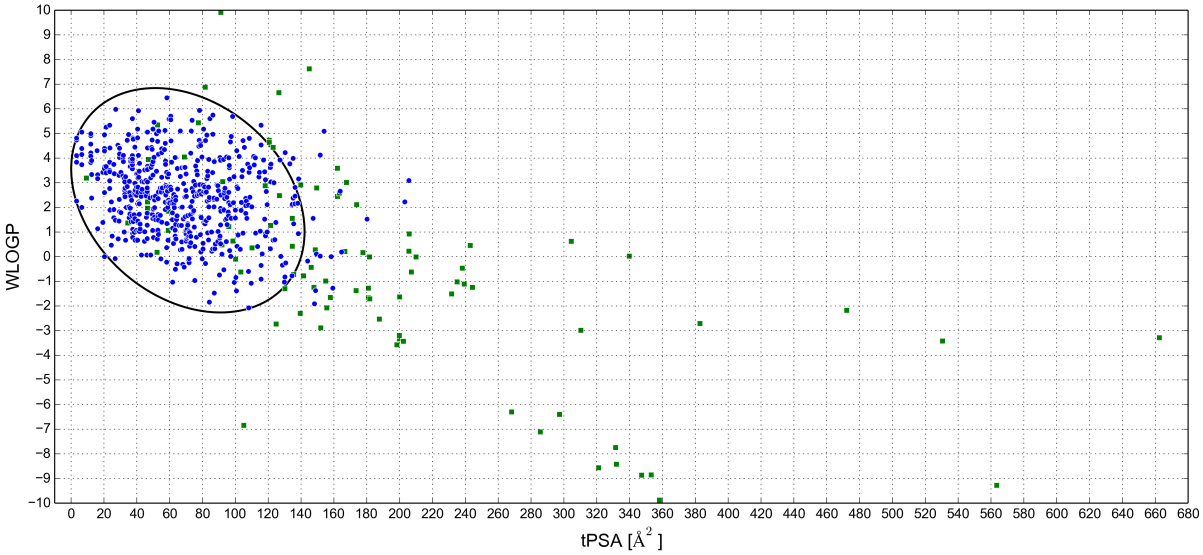

Figure S1. Description of the denormalized HIA ellipse.
The final (denormalized) model allowing for the best classification of 567 well-absorbed (HIA+, blue dots) and 93 poorly absorbed (HIA-, green squares) molecules is depicted as a black ellipse: center at (71.051; 2.292); length of major axis = 142.081; length of minor axis = 8.740 ; angle: -1.031325°. Inside the ellipse, tPSA ranges between 0.0 and 142.1 Å^2 and WLOGP between -2.3 and 6.8.

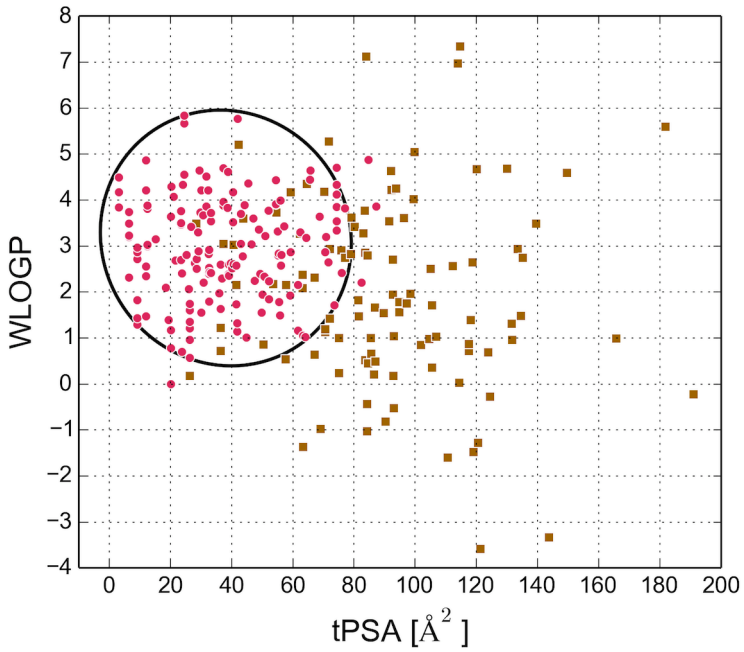

Figure S2. Description of the denormalized BBB ellipse.
The final (denormalized) model allowing for the best classification of 156 permeant (BBB+, pink dots) and 104 non-permeant (BBB-, brown squares) molecules is depicted as a black ellipse: center at (38.117; 3.177); length of major axis = 82.061; length of minor axis = 5.557 ; angle: -0.171887°. Inside the ellipse, tPSA ranges between -2.9 and 79.1 Å^2 and WLOGP between 0.4 and 6.0.

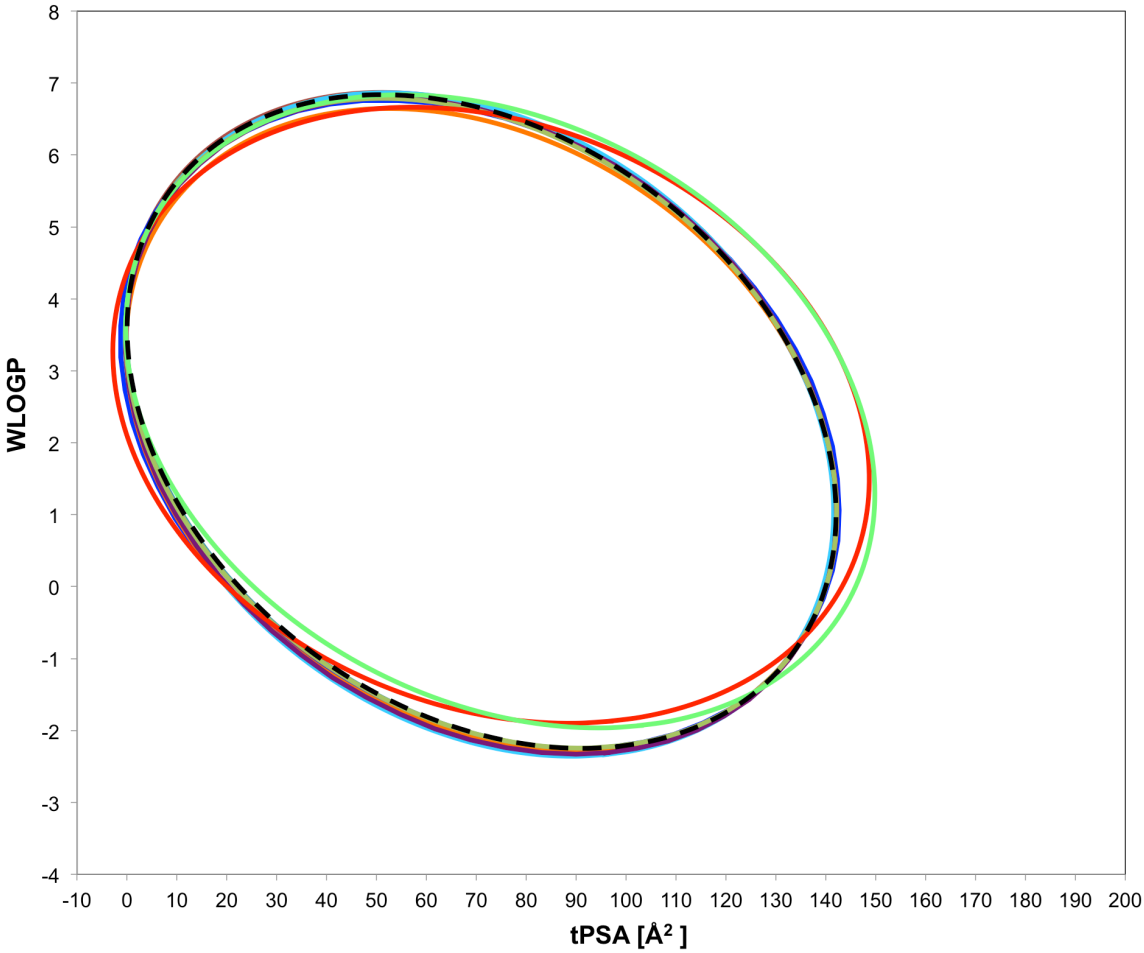

Figure S3. 10-fold cross-validation for the HIA ellipse.
The HIA model is depicted as a dashed black ellipse. The colored lines are the models built from each of the ten cross-validation training sets (refer to Methods S8 and Table S3). The variation is 2.73 $Å^2$ on the lower tPSA; 8.26 $Å^2$ on the upper tPSA and 0.46 for the lower WLOGP; 0.20 for the upper WLOGP.

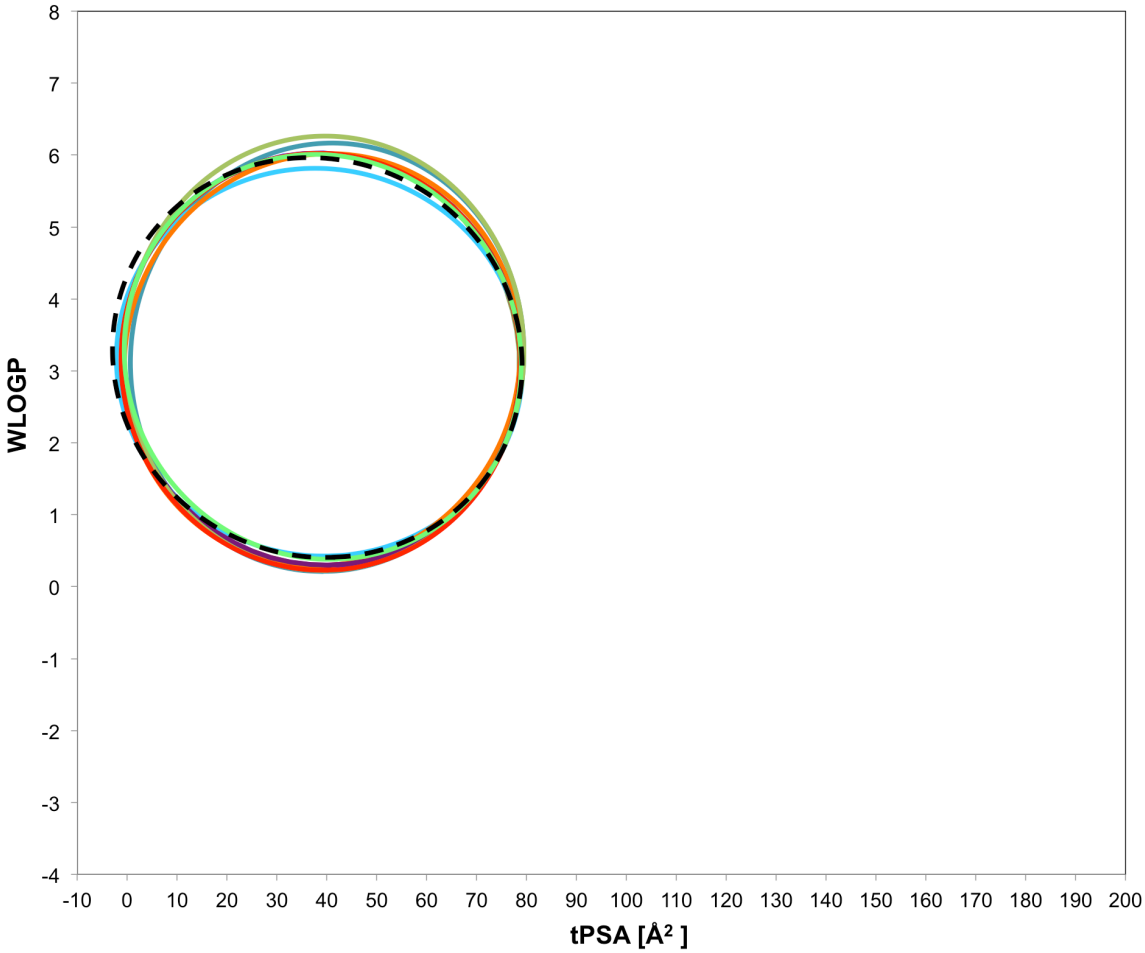

Figure S4. 10-fold cross-validation for the BBB ellipse.
The BBB model is depicted as a dashed black ellipse. The colored lines are the models built from each of the ten cross-validation training sets (refer to Methods S8 and Table S4). The variation is 2.84 $Å^2$ on the lower tPSA; 0.91 $Å^2$ on the upper tPSA and 0.21 for the lower WLOGP; 0.50 for the upper WLOGP.

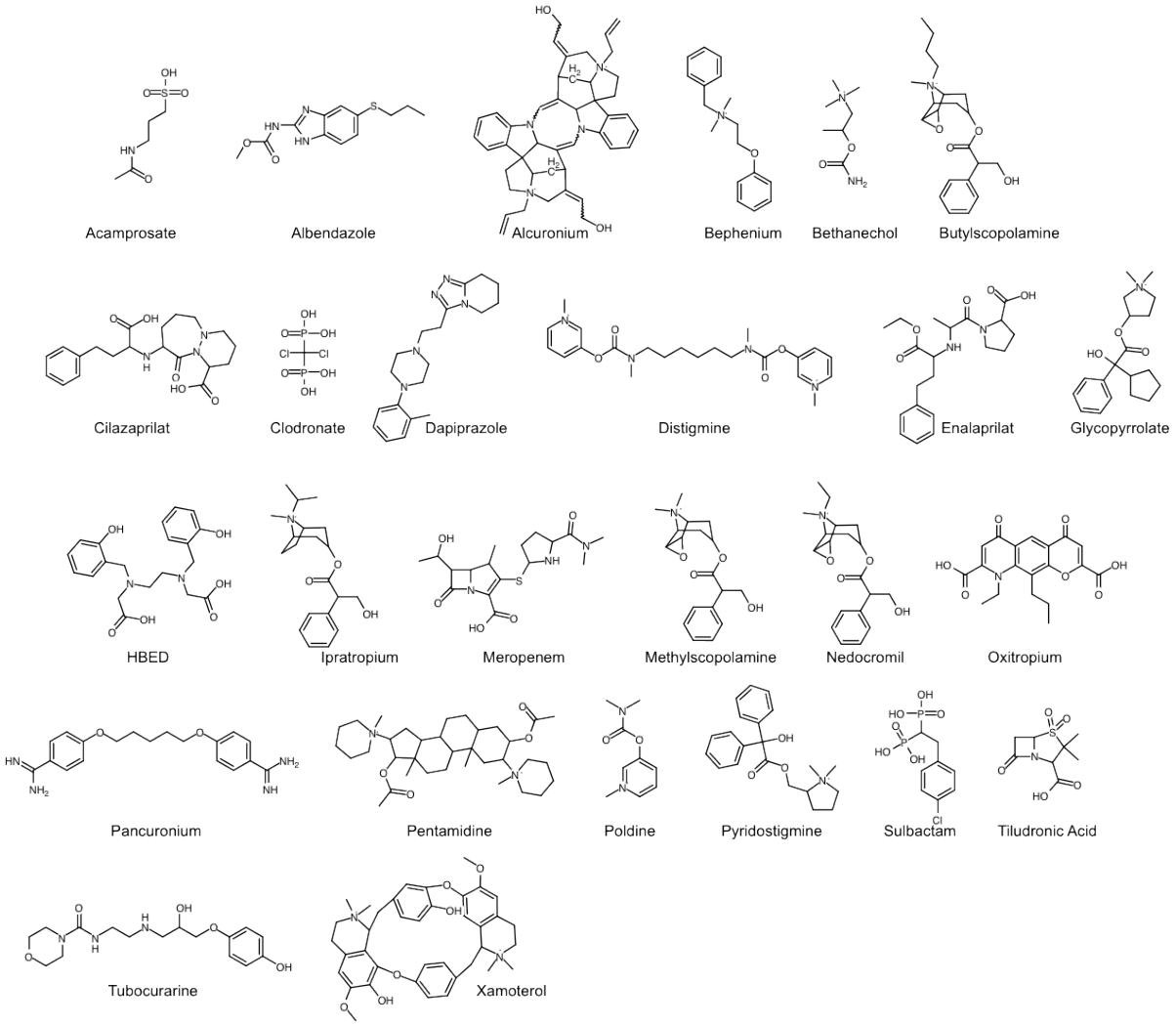

Figure S5. **HIA-dataset false positives** (FP, 26 molecules).

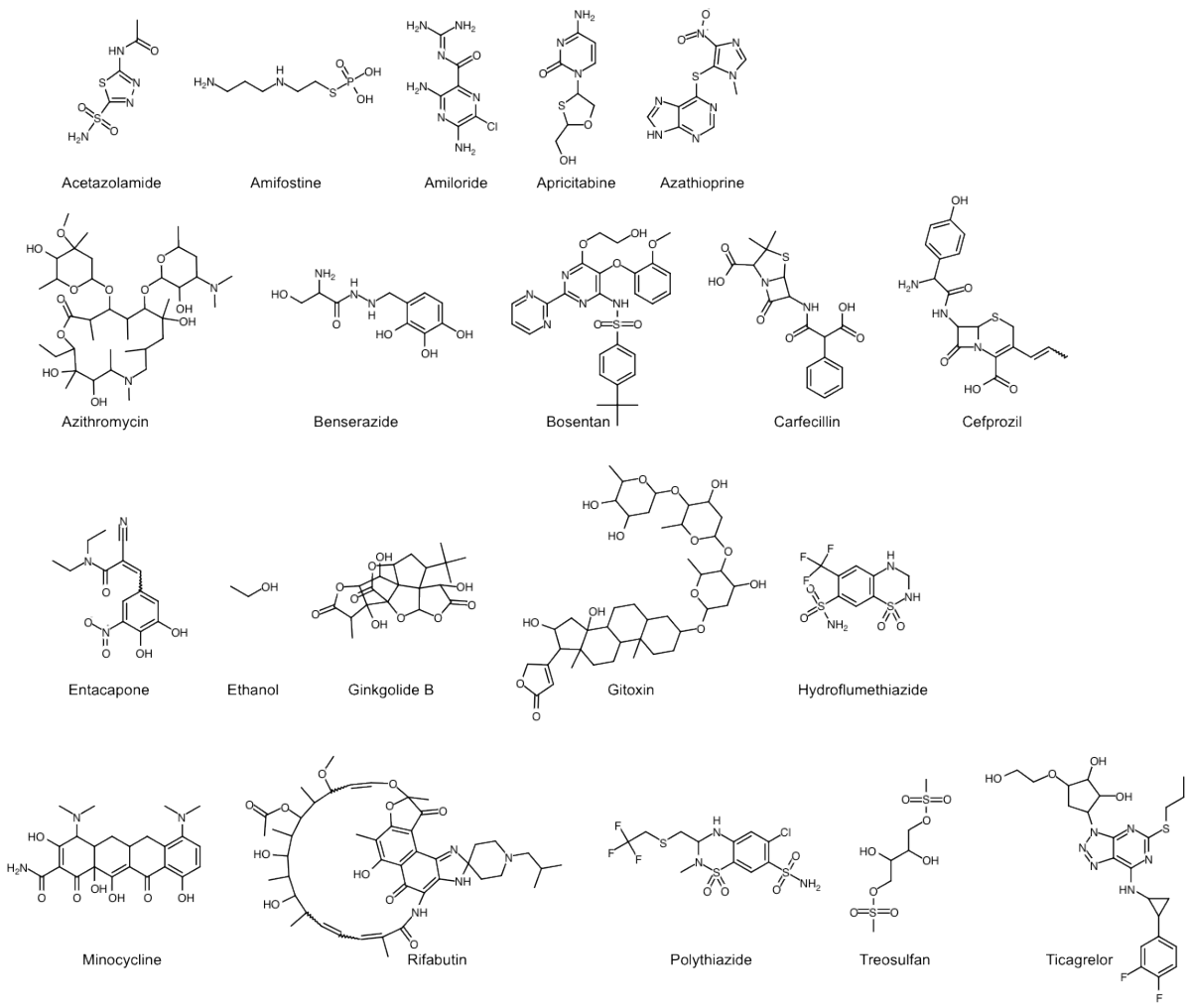

Figure S6. **HIA-dataset false negatives** (FN, 20 molecules).

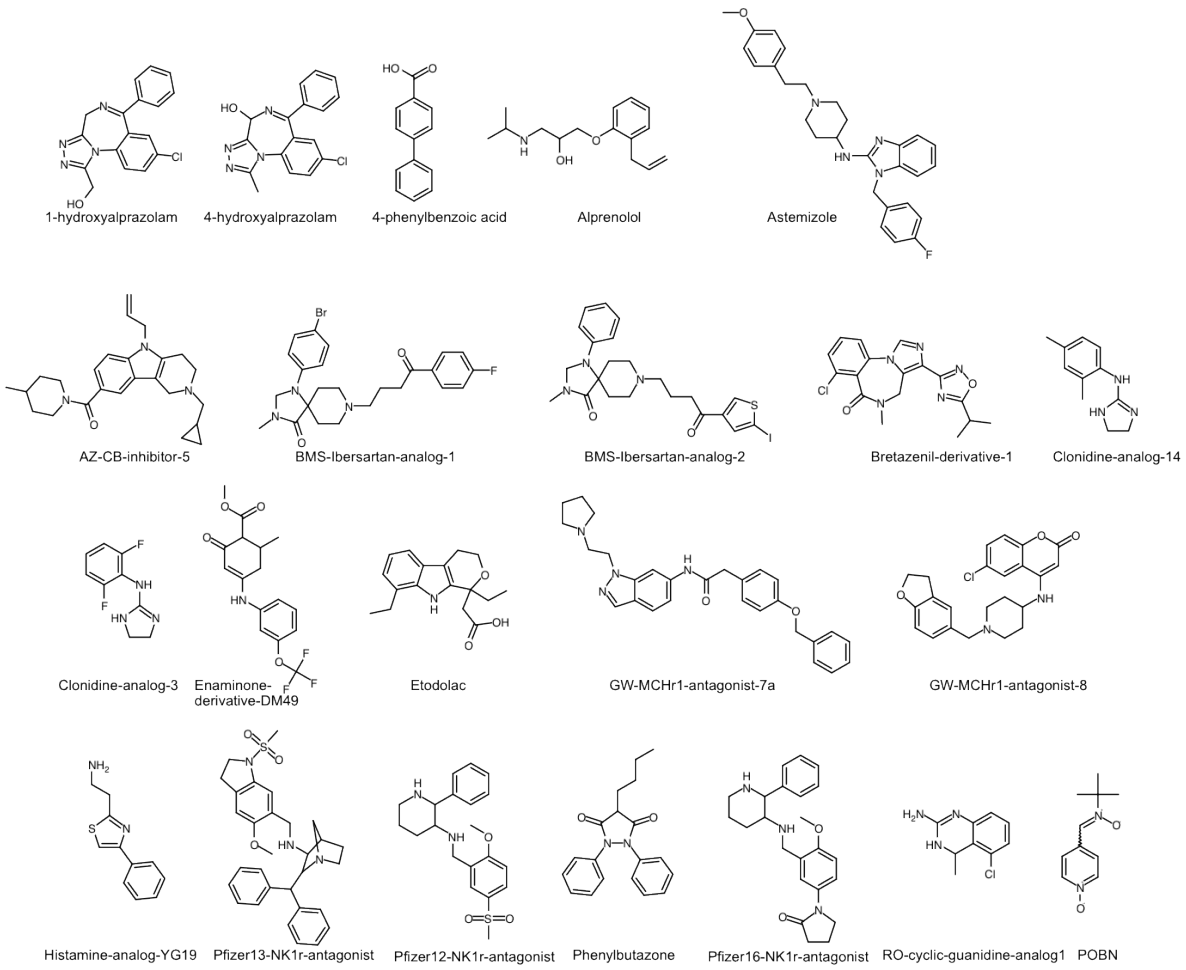

Figure S7. **BBB-dataset false positives** (FP, 22 molecules).

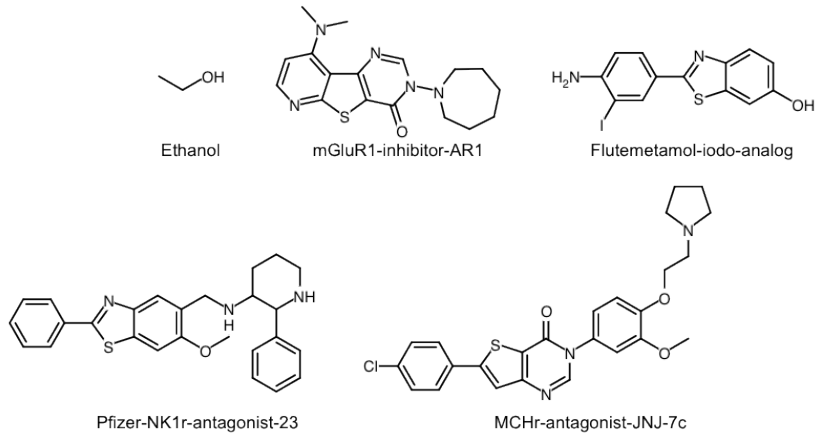

Figure S8. **BBB-dataset false negatives** (FP, 5 molecules).

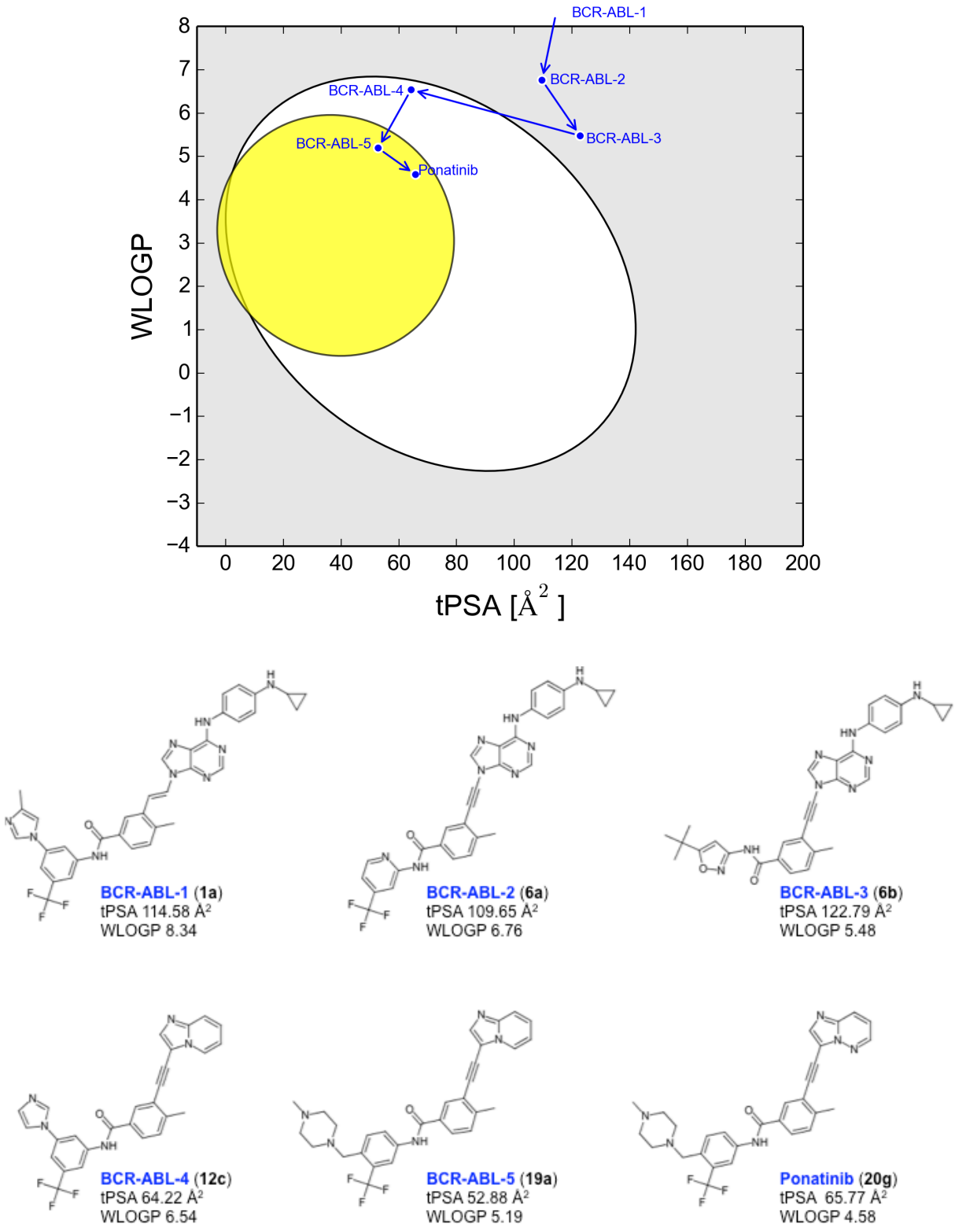

Figure S9. **Optimization of BCR-ABL inhibitors** resulting in the oral anti-cancer drug ponatinib. The medicinal chemistry path shows how to evolve from non-bioavailable leads (BCR-ABL-1, BCR-ABL-2, BCR-ABL-3) to well-absorbed and orally bioavailable drug-candidates (BCR-ABL-1, BCR-ABL-2) and finally develop a potent oral drug (ponatinib, Iclusig®). Molecular structures and identification numbers in brackets are taken from ref[15].

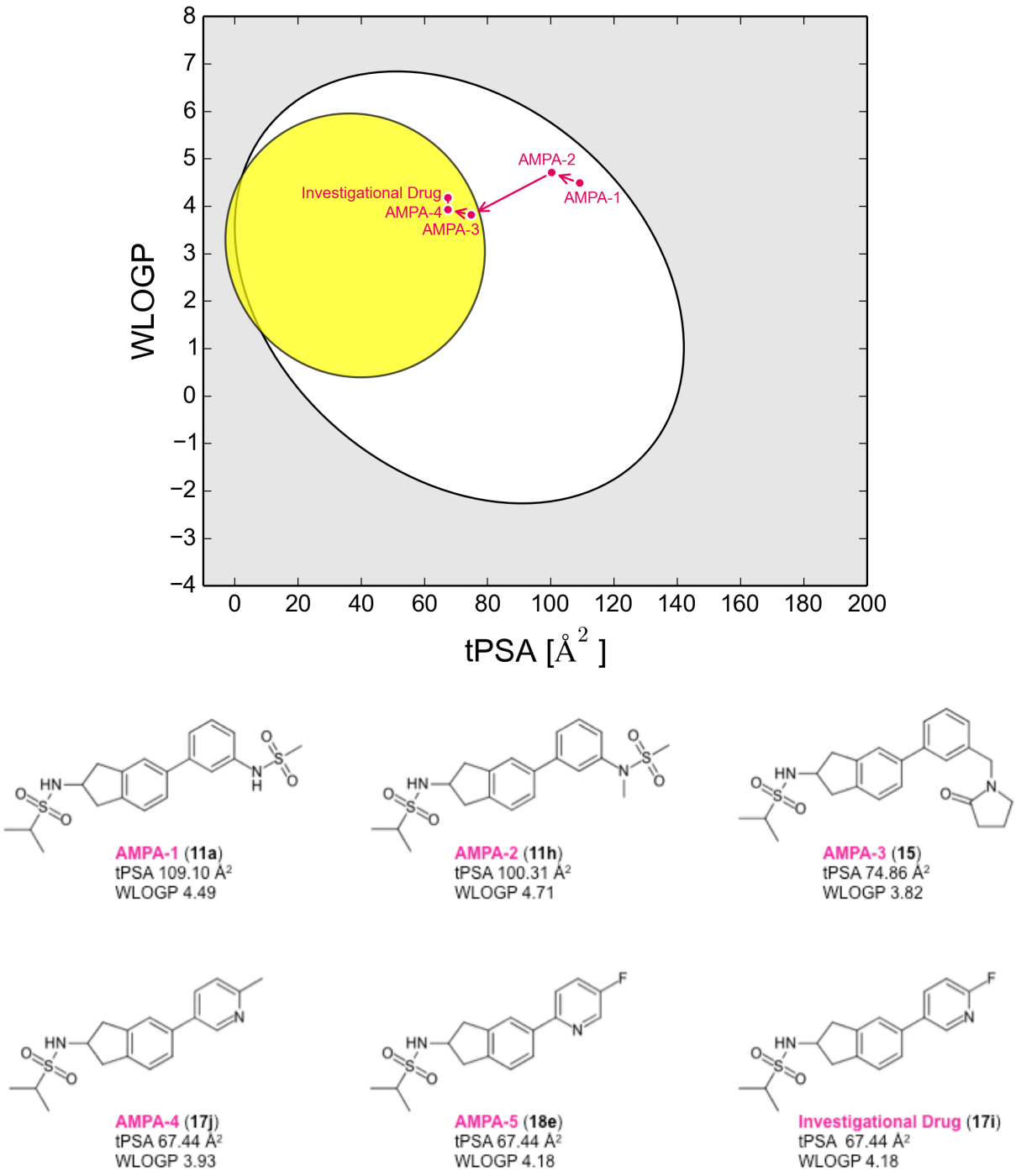

Figure S10. Optimization of AMPA receptor modulators resulting in an investigational drug promoted in clinical tests. The medicinal chemistry path shows how to evolve from BBB non-permeant leads (AMPA-1, AMPA-2, AMPA-3) to optimized compounds with high CNS levels (AMPA-4, AMPA-5) and finally develop a potent investigational drug accessing the brain (Investigational Drug). Molecular structures and identification numbers in brackets are taken from ref[16].

Supporting References

[1] D. Newby, A. A. Freitas, T. Ghafourian, *Eur. J. Med. Chem.* **2015**, *90*, 751–765.

[2] L. Mak, D. Marcus, A. Howlett, G. Yarova, G. Duchateau, W. Klaffke, A. Bender, R. C. Glen, *J. Cheminform.* **2015**, *7*, 31.

[3] K. M. Morrissey, C. C. Wen, S. J. Johns, L. Zhang, S.-M. Huang, K. M. Giacomini, *Clin. Pharmacol. Ther.* **2012**, *92*, 545–546.

[4] M. Zhao, Y. Chen, D. Qu, H. Qu, *Sci. China Life Sci.* **2011**, *54*, 60–64.

[5] V. Law, C. Knox, Y. Djoumbou, T. Jewison, A. C. Guo, Y. Liu, A. Maciejewski, D. Arndt, M. Wilson, V. Neveu, et al., *Nucleic Acids Res.* **2013**, *42*, D1091–D1097.

[6] A. P. Bento, A. Gaulton, A. Hersey, L. J. Bellis, J. Chambers, M. Davies, F. A. Krüger, Y. Light, L. Mak, S. McGlinchey, et al., *Nucleic Acids Res.* **2014**, *42*, D1083–90.

[7] D. Weininger, *J. Chem. Inf. Model.* **1988**, *28*, 31–36.

[8] Y. Brito-Sánchez, Y. Marrero-Ponce, S. J. Barigye, I. Yaber-Goenaga, C. Morell Pérez, H. Le-Thi-Thu, A. Cherkasov, *Mol. Inf.* **2015**, *34*, 308–330.

[9] S. A. Wildman, G. M. Crippen, *J. Chem. Inf. Model.* **1999**, *39*, 868–873.

[10] A. K. Ghose, V. N. Viswanadhan, J. J. Wendoloski, *J Comb Chem* **1999**, *1*, 55–68.

[11] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, G. R. Hutchison, *J. Cheminform.* **2011**, *3*, 33.

[12] P. Ertl, B. Rohde, P. Selzer, *J. Med. Chem.* **2000**, *43*, 3714–3717.

[13] N. Metropolis, S. Ulam, *JASA* **1949**, *44*, 335–341.

[14] A. Daina, O. Michielin, V. Zoete, *J. Chem. Inf. Model.* **2014**, *54*, 3284–3301.

[15] W.-S. Huang, C. A. Metcalf, R. Sundaramoorthi, Y. Wang, D. Zou, R. M. Thomas, X. Zhu, L. Cai, D. Wen, S. Liu, et al., *J. Med. Chem.* **2010**, *53*, 4701–4719.

[16] S. E. Ward, M. Harries, L. Aldegheri, D. Andreotti, S. Ballantine, B. D. Bax, A. J. Harris, A. J. Harker, J. Lund, R. Melarange, et al., *J. Med. Chem.* **2010**, *53*, 5801–5812.