Table 4. Structures and Data for Data Set[5]

SMILES	name	logPerm [a]	TPSA
O=C([O-])c1ccccc1O	salycilate	-4.924	60.36
CC(=O)Oc1ccccc1C(=O)[O-]	acetylsalicylate	-5.620	66.43
OCC(O)C(O)C(O)C(O)CO	mannitol	-6.745	121.37
CC(C)(C)NCC(O)c1cc(O)cc(O)c1	terbutaline	-6.420	72.71
C=CCc1ccccc1OCC(O)CNC(C)C	alprenolol	-4.393	41.49
CC(C)NCC(O)COc1cccc2ccccc12	propranolol	-4.378	41.49
CC(=O)Nc1ccc(OCC(O)CNC(C)C)cc1	practolol	-6.046	70.58
CC(C)NCC(O)COc1ccc(CC(N)=O)cc1	atenolol	-6.700	84.58
COCCc1ccc(OCC(O)CNC(C)C)cc1	metoprolol	-4.569	50.72
CC12CCC(=O)C=C1CCC3C2CCC4(C)C(O)CCC34	testosterone	-4.286	37.30
O=C(O)c2cc(N=Nc1ccc(O)c(C(=O)O)c1)ccc2O	olsalazine	-6.959	139.78
CC(=O)CC(c1ccccc1)c3c(O)c2ccccc2oc3=O	warfarin	-4.417	67.51
CC14CCC(=O)C=C1CCC3C2CCC(C(=O)CO)C2(C)CC(O)C34	corticosterone	-4.263	74.60
CC14CCC(=O)C=C1CCC3C2CCC(O)(C(=O)CO)C2(C)CC(O)C34	hydrocortisone	-4.668	94.83
CCOC(=O)C1=C(C)NC(C)=C(C(=O)OC)C1c2cccc(Cl)c2Cl	felodipine	-4.644	64.64
CC4CC3C2CCC1=CC(=O)C=CC1(C)C2(F)C(O)CC3(C)C4(O)C(=O)CO	dexomethasone	-4.903	94.83
O=C(O)c3cc(N=Nc2ccc(S(=O)(=O)Nc1ccccn1)cc2)ccc3O	sulphasalazine	-6.886	141.32

[a] permeability [cm/s] through Caco-2 monolayers.

Table 5. Structures and Data for Data Set[16]

SMILES	name	logBB [a]	TPSA
CNC(=NC#N)NCCSCc1nc[nH]c1C	1	-1.42	88.89
Cc1csc(N=C(N)N)n1	2	-0.04	77.30
COc1cccnc1CCCCNc3nc(=O)c(Cc2ccc(C)nc2)c[nH]3	3	-2.00	92.80
CN(C)Cc4ccc(CSCCNc3nc(=O)c(Cc2ccc1ccccc1c2)c[nH]3)o4	4	-1.30	74.16
Cc3ccc(Cc2c[nH]c(NCCSCc1ccc(CN(C)C)o1)nc2=O)cn3	5	-1.06	87.05
Clc1cccc(Cl)c1NC2=NCCN2	6	0.11	36.42
COc2ccc(CN(CCN(C)C)c1ccccn1)cc2	7	0.49	28.60
CN(C)CCCN3c1ccccc1CCc2ccccc23	8	0.83	6.48
CNC(=CN(=O)=O)NCCSCc1ccc(CN(C)C)o1	9	-1.23	86.26
CNC(=CC#N)NCCSCc1csc(N=C(N)N)n1	10	-0.82	125.15
CN(C)c2cc(c1n[nH]c(N)n1)ccn2	11	-1.17	83.73
CCNC(=NC#N)NCCSCc1ncccc1Br	12	-2.15	73.10
O=N(=O)c1cc[nH]c1NCCSCc2ncccc2Br	13	-0.67	86.53
O=N(=O)c1cc[nH]c1NCCSCc2ccccn2	14	-0.66	86.53
O=N(=O)c2c(Cc1ccccc1)c[nH]c2NCCSCc3ccccn3	15	-0.12	86.53
NC(N)=Nc2nc(c1ccccc1)cs2	16	-0.18	77.30
NC(N)=Nc2nc(c1cccc(N)c1)cs2	17	-1.15	103.32
CC(=O)Nc2cccc(c1csc(N=C(N)N)n1)c2	18	-1.57	106.40
CNC(=NC#N)Nc2cccc(c1csc(N=C(N)N)n1)c2	19	-1.54	137.51
CN(C)Cc2ccc(CSCCNc1[nH]ccc1N(=O)=O)o2	20	-1.12	90.02
CN(C)Cc3ccc(CSCCNc2[nH]cc(Cc1ccccc1)c2N(=O)=O)o3	21	-0.73	90.02
CN(C)Cc3ccc(c2cccc(Nc1[nH]ccc1N(=O)=O)c2)o3	22	-0.27	90.02
CN(C)Cc3ccnc(c2cccc(Nc1[nH]ccc1N(=O)=O)c2)c3	23	-0.28	89.77
CC(=O)NCCCOc2cccc(CN1CCCCC1)c2	24	-0.46	41.57
O=C(NCCCOc2cccc(CN1CCCCC1)c2)c3ccccc3	25	-0.24	41.57
OCCCOc2cccc(CN1CCCCC1)c2	26	-0.02	32.70
c3ccc(NCCCOc2cccc(CN1CCCCC1)c2)nc3	27	0.69	37.39
c3cc(CN1CCCCC1)cc(OCCCNc2nccs2)c3	28	0.44	37.39
c4cc(CN1CCCCC1)cc(OCCCNc3nc2ccccc2s3)c4	29	0.14	37.39
c4cc(CN1CCCCC1)cc(OCCCNc3nc2ccccc2o3)c4	30	0.22	50.53
CCC(C)=O	butanone	-0.08	17.07
c1ccccc1	benzene	0.37	0.00
CCC(C)CC	3-methylpentane	1.01	0.00
CCCC(C)CC	3-methylhexane	0.90	0.00
CC(C)O	2-propanol	-0.15	20.23
CC(C)CO	2-methylpropanol	-0.17	20.23
CCCC(C)C	2-methylpentane	0.97	0.00
CCC(C)(C)C	2,2-dimethylbutane	1.04	0.00
FC(F)(F)CCl	1,1,1-trifluoro-2-chloroethane	0.08	0.00
CC(Cl)(Cl)Cl	1,1,1-trichloroethane	0.40	0.00
CCOCC	diethyl_ether	0.00	9.23
FC(F)OC(F)(F)C(F)Cl	enflurane	0.24	9.23
CCO	ethanol	-0.16	20.23
C=COCC(F)(F)F	fluroxene	0.13	9.23
FC(F)(F)C(Cl)Br	halothane	0.35	0.00
CCCCCCC	heptane	0.81	0.00
CCCCCC	hexane	0.80	0.00
FC(F)OC(Cl)C(F)(F)F	isoflurane	0.42	9.23
C	methane	0.04	0.00
CC1CCCC1	methylcyclopentane	0.93	0.00
N#N	nitrogen	0.03	47.58
CCCCC	pentane	0.76	0.00

CCCO	propanol	-0.16	20.23
CC(C)=O	propanone	-0.15	17.07
FC(Br)C(F)(F)F	teflurane	0.27	0.00
Cc1ccccc1	toluene	0.37	0.00
ClC=C(Cl)Cl	trichloroethene	0.34	0.00

[a] $\log(C_{brain} / C_{blood})$.

Table 6. Structures and Data for Data Set[13]

SMILES	name	logPc [a]	TPSA
O=C(NCCNCC(O)COc2ccc(CCOCCOCC1CC1)cc2)N3CCC(O)CC3	H216/44	-0.983	112.52
CC(C)NCC(O)COc1ccc(CC(N)=O)cc1	atenolol	0.009	84.58
CC(C)NCC(O)COc1ccc(NC=O)cc1	H95/71	0.574	70.59
CC(=O)Nc1ccc(OCC(O)CNC(C)C)cc1	practolol	0.539	70.59
CCC(=O)Nc1ccc(OCC(O)CNC(C)C)cc1	H244/45	0.780	70.59
CC(C)NCC(O)COc1cccc2[nH]ccc12	pindolol	1.738	57.28
COCCc1ccc(OCC(O)CNC(C)C)cc1	metoprolol	1.963	50.72
C=CCOc1ccccc1OCC(O)CNC(C)C	oxprenolol	2.078	50.72
C=CCc1ccccc1OCC(O)CNC(C)C	alprenolol	2.384	41.49

[a] permeability [10^{-6} cm/s] through Caco-2 monolayers.

Table 7. Structures and Data for Data Set[17]

SMILES	name	logBB [a]	TPSA
CNCCCN3c1ccccc1CCc2ccccc23	1	1.00	15.26
CN(C)CCCN3c1ccccc1CCc2ccccc23	2	1.05	6.48
CN4CCN3c1ccccc1Cc2ccccc2C3C4	3	0.99	6.48
CN(C)CCC=C2c1ccccc1CCc3ccccc23	4	0.98	3.24
Cc1cccc3c1Oc2ccccc2C4(O)CCN(C)CC34	5	0.82	32.70
CN4CC3c1ccccc1Oc2ccc(Cl)cc2C3C4	6	1.03	12.47
FC(F)(F)c2ccc(N1CCNCC1)nc2Cl	7	1.64	28.16
O=C1CCCN1CCCCN3CCN(c2cc(C(F)(F)F)ccn2)CC3	8	0.16	39.68
Cc1cccc3c1Oc2ccccc2C4(O)CCNCC34	9	0.52	41.49
Clc4ccc3Oc1ccccc1C2CNCC2c3c4	10	0.39	21.26
CN4CCN3c1nccccc1Oc2ccccc2C3C4	11	0.53	28.60
C#CC4(O)CCC3C2C(C)CC1=C(CCC(=O)C1)C2CCC34C	12	0.40	37.30
Clc1cccc(Cl)c1N=C2NCCN2	13	0.11	36.42
NC(=O)N3c1ccccc1CCc2ccccc23	14	0.00	46.33
NC(=O)N4c1ccccc1C2OC2c3ccccc34	15	-0.33	58.86
CNC(=NC#N)NCCSCc1nc[nH]c1C	16	-1.42	88.89
CNC(=NN(=O)=O)NCCSCc1ccc(CN(C)C)o1	17	-1.23	98.62
COc2ccc(CN(CCN(C)C)c1ccccn1)cc2	18	0.49	28.60
COc1ccccc1CCCCNc3nc(=O)c(Cc2ccc(C)nc2)c[nH]3	19	-2.00	79.91
Cc3ccc(Cc2c[nH]c(NCCSCc1ccc(CN(C)C)o1)nc2=O)cn3	20	-1.06	87.05
O=c2[nH]c1ccccc1n2CCCN5CCC(n4c(=O)[nH]c3cc(Cl)ccc34)CC5	21	-0.78	78.83
C=CCC(N)c1ccccc1c2noc3ccccc23	22	0.00	52.05
Cc2nc1CCCCn1c(=O)c2CCN5CCC(c3noc4cc(F)ccc34)CC5	23	-0.02	64.17
Cc2nc1C(O)CCCn1c(=O)c2CCN5CCC(c3noc4cc(F)ccc34)CC5	24	-0.67	84.39
CC(C)c4nc(c2ncn3c1cccc(Cl)c1C(=O)N(C)Cc23)no4	25	-0.30	77.06
CN4Cc2c(c1noc(C(C)(C)O)n1)ncn2c3cccc(Cl)c3C4=O	26	-1.34	97.29
CN4Cc2c(c1noc(C(C)(O)CO)n1)ncn2c3cccc(Cl)c3C4=O	27	-1.82	117.51
Cc3ccc(Cc2c[nH]c(NCCCCc1ncc(Br)cc1C)nc2=O)cn3	28	-1.88	83.56
CN(C)Cc4ccc(CSCCNc3nc(=O)c(Cc2ccc1ccccc1c2)c[nH]3)o4	29	-1.30	74.16
CN(C)c2cc(c1n[nH]c(N)n1)ccn2	30	-1.17	83.73
O=N(=O)c1cc[nH]c1NCCSCc2nccc2Br	31	-0.67	86.53
O=N(=O)c1cc[nH]c1NCCSCc2ccccn2	32	-0.66	86.53
NC(N)=Nc2nc(c1cccc(N)c1)cs2	33	-1.15	103.32
CC(=O)Nc2cccc(c1csc(N=C(N)N)n1)c2	34	-1.57	106.40
CN(C)Cc2ccc(CSCCNc1[nH]ccc1N(=O)=O)o2	35	-1.12	90.02
CN(C)Cc3ccc(CSCCNc2[nH]cc(Cc1ccccc1)c2N(=O)=O)o3	36	-0.73	90.02
CN(C)Cc3ccc(c2cccc(Nc1[nH]ccc1N(=O)=O)c2)o3	37	-0.27	90.02
CN(C)Cc3ccnc(c2cccc(Nc1[nH]ccc1N(=O)=O)c2)c3	38	-0.28	89.77
CC(=O)NCCCOc2cccc(CN1CCCCC1)c2	39	-0.46	41.57
O=C(NCCCOc2cccc(CN1CCCCC1)c2)c3ccccc3	40	-0.24	41.57
OCCCOc2cccc(CN1CCCCC1)c2	41	-0.02	32.70
c3ccc(NCCCOc2cccc(CN1CCCCC1)c2)nc3	42	0.69	37.39
c3cc(CN1CCCCC1)cc(OCCCNc2nccs2)c3	43	0.44	37.39
c4cc(CN1CCCCC1)cc(OCCCNc3nc2ccccc2s3)c4	44	0.14	37.39
c4cc(CN1CCCCC1)cc(OCCCNc3nc2ccccc2o3)c4	45	0.22	50.53

[a] $\log(C_{brain} / C_{blood})$.

Table 8. Structures and Data for Data Set[7]

SMILES	name	logPeff [a]	TPSA[b]
Cc2cc(=O)n(c1ccccc1)n2C	antipirine	-3.35	26.94
NC(=O)n3c1ccccc1ccc2ccccc23	carbamazepine	-3.37	48.03
CNCCCN3c1ccccc1CCc2ccccc23	desipramine	-3.36	15.26
CC(NC(CCc1ccccc1)C(=O)O)C(=O)N2CCCC2C(=O)O	enalaprilat	-4.70	106.93
CC(C)n3c(C=CC(O)CC(O)CC(=O)O)c(c1ccc(F)cc1)c2ccccc23	fluvastatin	-3.62	82.69
NS(=O)(=O)c2cc(C(=O)O)c(NCc1ccco1)cc2Cl	furosemide	-5.30	131.01
NS(=O)(=O)c1cc2c(cc1Cl)NCNS2(=O)=O	hydrochlorothiazide	-5.40	135.11
CC(C(=O)O)c2cccc(C(=O)c1ccccc1)c2	ketoprofen	-3.08	54.37
COCCc1ccc(OCC(O)CNC(C)C)cc1	metoprolol	-3.89	50.72
COc2ccc1cc(C(C)C(=O)O)ccc1c2	naproxen	-3.08	46.53
CC(C)NCC(O)COc1cccc2ccccc12	propranolol	-3.54	41.49
CC(C)(C)NCC(O)c1cc(O)cc(O)c1	terbutaline	-4.52	72.71
COc2ccc(CCN(C)CCCC(C#N)(c1ccc(OC)c(OC)c1)C(C)C)cc2OC	verapamil	-3.17	63.97

[a] effective permeability [10^{-4} cm/s] of human jejunum. [b] sulfur fragments were also included in the calculation of PSA.

Table 9. Structures and Data for Data Set[9]

SMILES	name	FA% [a]	TPSA
COCCc1ccc(OCC(O)CNC(C)C)cc1	metoprolol	102	50.72
O=C3CN=C(c1ccccc1)c2cc(Cl)ccc2N3	nordiazepam	99	41.46
CN3C(=O)CN=C(c1ccccc1)c2cc(Cl)ccc23	diazepam	97	32.67
C=CCOc1ccccc1OCC(O)CNC(C)C	oxprenolol	97	50.72
Cc2cc(=O)n(c1ccccc1)n2C	phenazone	97	26.94
O=C3Nc1ccc(Cl)cc1C(c2ccccc2)=NC3O	oxazepam	97	61.69
C=CCc1ccccc1OCC(O)CNC(C)C	alprenolol	96	41.49
CC(=O)Nc1ccc(OCC(O)CNC(C)C)cc1	practolol	95	70.59
CC(C)NCC(O)COc1cccc2[nH]ccc12	pindolol	92	57.28
O=C(O)c4cn(C1CC1)c3cc(N2CCNCC2)c(F)cc3c4=O	ciprofloxacin	69	74.57
Cc1ccccc1N3C(=O)c2cc(S(N)(=O)=O)c(Cl)cc2NC3C	metolazone	64	92.50
NCC1CCC(C(=O)O)CC1	tranexamic acid	55	63.32
CC(C)NCC(O)COc1ccc(CC(N)=O)cc1	atenolol	54	84.58
CCN1CCCC1CNC(=O)c2cc(S(N)(=O)=O)ccc2OC	sulpiride	36	101.73
OCC(O)C(O)C(O)C(O)CO	mannitol	26	121.37
O=C(O)P(=O)(O)O	foscarnet	17	94.83
O=C(O)c3cc(N=Nc2ccc(S(=O)(=O)Nc1ccccn1)cc2)ccc3O	sulfasalazine	12	141.32
O=C(O)c2cc(N=Nc1ccc(O)c(C(=O)O)c1)ccc2O	olsalazine	2.3	139.78
O=C(CO)C(O)C(OC1OC(CO)C(O)C(O)C1O)C(O)CO	lactulose	0.6	197.36
OCC3OC(OCC2OC(OC1(CO)OC(CO)C(O)C1O)C(O)C(O)C2O)C(O)C(O)C3O	raffinose	0.3	268.68

[a] percentage of drug absorbed after oral administration to humans.