Supporting Information

MRlogP: transfer learning enables accurate logP prediction using small experimental training datasets

Yan-Kai Chen[1], S. Shave[1*], and Manfred Auer[1*]

[1] School of Biological Sciences, University of Edinburgh, The King's Buildings, Edinburgh, Scotland, UK.

*Correspondence: s.shave@ed.ac.uk, manfred.auer@ed.ac.uk

MRlogP source code available at https://github.com/JustinYKC/MRlogP

500k training set composition

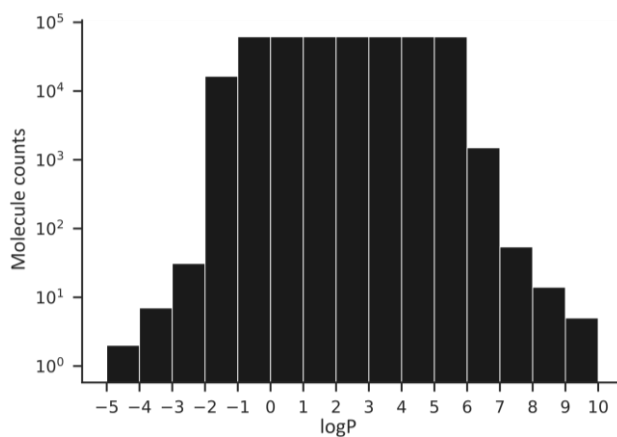**Figure S1.** Binned distribution of logP values in the 500K training set.

Hyperparameter scanning

Table S1. Variables explored in hyperparameter scan (grid search). * - Number of hidden layers are multiples of the number of input descriptors representing a molecule (316) with the multipliers 0.5, 1, 1.5, 2 and 4 applied.

Parameter name	Values
Training epochs	1, 2, 3, 5, 10, 15, 20, 25, 30
Batch size	32, 64, 128
Dropout rate	0.1, 0.2, 0.4
Learning rate	1×10^{-3}, 1×10^{-4}
Hidden layers	3, 4, 5
Nodes per layer	316, 474, 632, 1264 *

Table S2. The hyperparameter scan result of the top 20 models. * - Hyperparameters: (No. of hidden layers, Nodes for the first hidden layer, Batch size, Training epochs, Dropout rate, Learning rate).

Model	Hyperparameter *	Validation RMSE
7-6	$(3, 1264, 32, 30, 0.1, 1\times10^{-4})$	0.38377
1519-5	$(3, 1264, 32, 25, 0.1, 1\times10^{-4})$	0.38662
1303-5	$(3, 1264, 32, 20, 0.1, 1\times10^{-4})$	0.38690
1535-5	$(5, 1264, 32, 25, 0.1, 1\times10^{-4})$	0.38848
30-6	$(3, 1264, 64, 30, 0.1, 1\times10^{-3})$	0.38889
54-6	$(3, 1264, 128, 30, 0.1, 1\times10^{-3})$	0.38932
1103-5	$(5, 1264, 32, 15, 0.1, 1\times10^{-4})$	0.38934
14-6	$(4, 1264, 32, 30, 0.1, 1\times10^{-3})$	0.39074
31-6	$(3, 1264, 64, 30, 0.1, 1\times10^{-4})$	0.39078
15-6	$(4, 1264, 32, 30, 0.1, 1\times10^{-4})$	0.39085
1527-5	$(4, 1264, 32, 25, 0.1, 1\times10^{-4})$	0.39087
1607-5	$(5, 1264, 32, 25, 0.2, 1\times10^{-4})$	0.39093
79-6	$(3, 1264, 32, 30, 0.2, 1\times10^{-4})$	0.39118
70-6	$(5, 1264, 128, 30, 0.1, 1\times10^{-3})$	0.39133
87-6	$(4, 1264, 32, 30, 0.2, 1\times10^{-4})$	0.39139
47-6	$(5, 1264, 64, 30, 0.1, 1\times10^{-4})$	0.39144
1614-5	$(3, 1264, 64, 25, 0.2, 1\times10^{-3})$	0.39165
1599-5	$(4, 1264, 32, 25, 0.2, 1\times10^{-4})$	0.39207
1383-5	$(4, 1264, 32, 20, 0.2, 1\times10^{-4})$	0.39298
1518-5	$(3, 1264, 32, 25, 0.1, 1\times10^{-3})$	0.39302

10-fold cross validation

Table S3. The result of 10-fold cross validation within the top 20 models. * Average values and standard deviation are given.

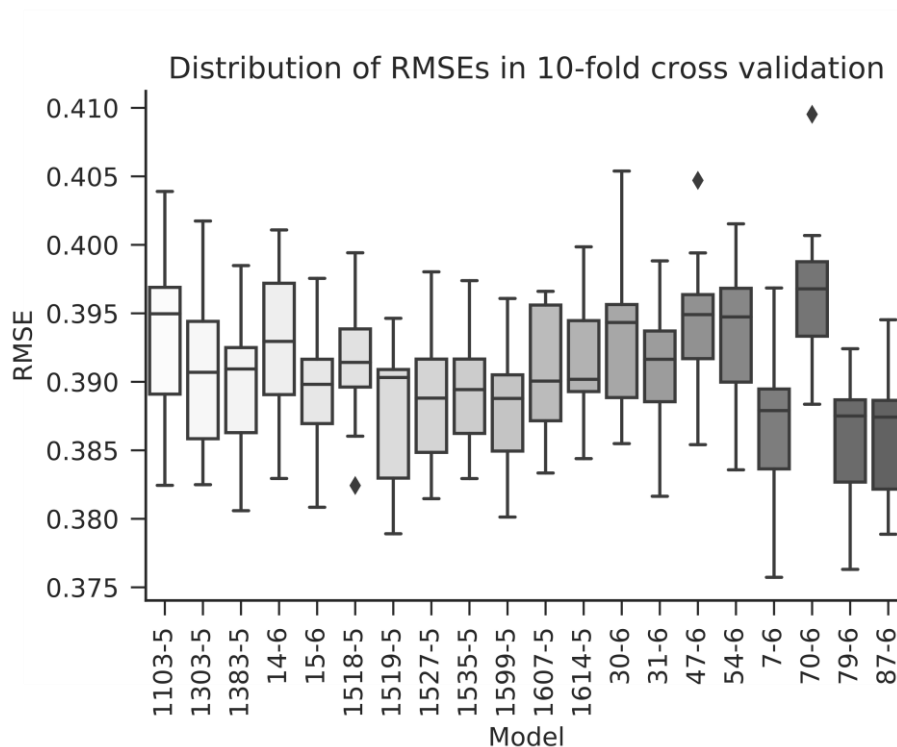| Model	Cross validation RMSE *
79-6	0.38584 ± 0.00534
87-6	0.38650 ± 0.00541
7-6	0.38709 ± 0.00596
1519-5	0.38782 ± 0.00572
1599-5	0.38856 ± 0.00497
1527-5	0.38910 ± 0.00554
15-6	0.38934 ± 0.00544
1535-5	0.38981 ± 0.00479
1383-5	0.39000 ± 0.00543
1303-5	0.39071 ± 0.00585
1607-5	0.39075 ± 0.00484
31-6	0.39107 ± 0.00522
1614-5	0.39158 ± 0.00512
1518-5	0.39175 ± 0.00528
14-6	0.39232 ± 0.00618
54-6	0.39355 ± 0.00552
30-6	0.39368 ± 0.00622
1103-5	0.39418 ± 0.00678
47-6	0.39451 ± 0.00554
70-6	0.39691 ± 0.00577

Figure S2. Boxplots showing RMSEs achieved by top 20 models in 10-Fold Cross Validation.

Transfer learning

Table S4. Hyperparameter scan used for transfer learning. Best scoring hyperparameter combination found to be: Training epochs for new output layer = 1, Number of training epochs for unfrozen hidden layers = 30, Number of unfrozen hidden layers = 2, resulting in RMSEs of 0.988, and 0.715 for Physprop_DL and Reaxys_DL respectively.

Parameter name	Values
Training epochs for new output layer	1, 2, 3, 4, 5
Number of training epochs for unfrozen hidden layers	1, 2, 3, 4, 5, 10, 15, 20, 30, 40
Number of unfrozen hidden layers	1, 2, 3

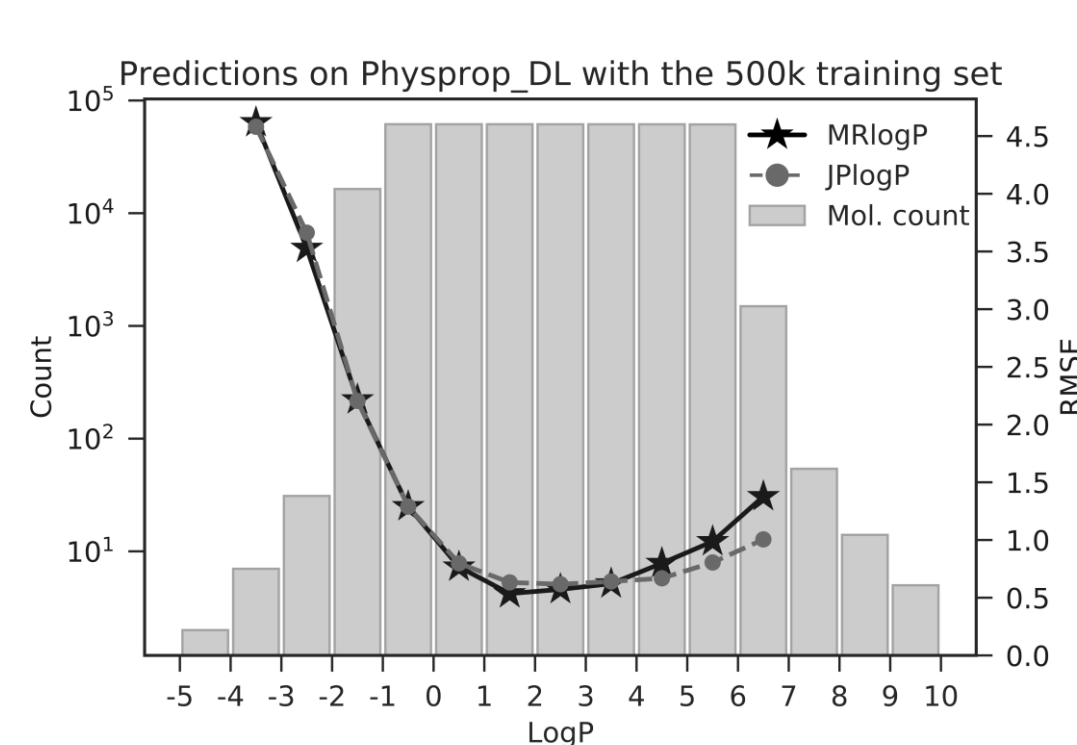**MRlogP and JPlogP performance against the Physprop_DL test set**

Figure S3. Histogram of logP bin occupancy in the 500k training set, along with MRlogP performance on the Physprop_DL test set a perbin basis. JPlogP performance shown for comparison. Strikingly, and as expected, performance of MRlogP is highly dependent on the number of example molecules present within a logP bin.

Table S5. Druglike molecules within the Martel dataset used for transfer learning, SMILES representation given, followed by experimentally determined logP values and calculated QED scores.

SMILES	Measured logP	QED score
COc1ccccc1[C@@H]1CC(=O)Nc2cc(C)c(C)cc21	4.17	0.909
COc1ccc2c(c1)O[C@](O)(C(F)(F)F)CC2=O	2.79	0.839
C=C(c1ccc(N(C)C)cc1)c1ccc(N(C)C)cc1	5.3	0.828
Cc1ccc([C@@H](O)C#CC(O)(c2ccccc2)c2ccccc2)o1	3.53	0.726
CC(C)(C)c1cc(Cc2ccccc2)n[nH]1	4.35	0.816
C[C@H]1C[C@H](NC(=S)NCc2csc3ccccc23)CC(C)(C)C1	6.21	0.757
CCNc1nc(NC(C)C)nc(Oc2ccc(OCC)nn2)n1	3.14	0.755
N#CC1=C(N)Oc2c(c(=O)oc3ccccc23)[C@H]1c1cccs1	2.34	0.695
CCOC(=O)CC1=Nn2c(nnc2-c2cccs2)SC1	2.98	0.810
O=C(Nc1ccc(C(F)(F)F)cc1)[C@H]1CCc2ccccc2O1	4.5	0.906
O=C1C[C@@H](c2c(F)cccc2Cl)c2cc3c(cc2N1)OCCO3	3.96	0.864
CSc1scc(-c2cccs2)c1C#N	3.69	0.735
O=C(Nc1ccccc1C(F)(F)F)NC1CCCCC1	4	0.841
CCn1c(NC[C@@H](O)c2ccccc2)nc2c1c(=O)n(C)c(=O)n2C	3.24	0.707
N#Cc1nc(N2CCCC2)nc(N(c2ccccc2)c2ccccc2)n1	5.38	0.717
N#CCc1nc(-c2ccc3c(c2)OCCCO3)cs1	2.68	0.842
COC(=O)[C@H]1O[C@@H](c2ccccc2)O[C@H]1C(=O)OC	2.49	0.754
CC(C)(C)S(=O)(=O)/C(C#N)=C/Nc1ccc(OC(F)(F)F)cc1	4.28	0.841
CCC[C@@H]1C(=O)C[C@H](C(C)C)C(C#N)(C#N)C1(C#N)C#N	3.3	0.789
Cc1c(C)c(C)c(CSc2nnc(CC#N)n2C)c(C)c1C	4.14	0.807
CCOc1nc(C#N)nc(N2CCOCC2)n1	2.46	0.727
Cc1cc(C)c([C@H]2[C@@H]3CCCC=C3C(C#N)=C(N)C2(C#N)C#N)c(C)c1	5.48	0.826
C[C@@]1(c2ccc(Cl)cc2Cl)OC[C@@H](Cn2cncn2)O1	3.68	0.873
CC1(C)[C@H](C(=O)N2CCOCC2)CC[C@]1(C)C(=O)N1CCOCC1	2.46	0.757
Cc1nc(-c2ccc(Cl)cc2)cs1	3.65	0.696
CCCn1c(=O)c(Br)cn(-c2ccccc2C)c1=O	2.82	0.870
COC(=O)COc1cc2occ(-c3csc(C)n3)c(=O)c2cc1C	3.9	0.676
Cc1ccc2c(C)nc(Nc3nc4c(c(=O)[nH]3)CCCC4)nc2c1	3.92	0.757
CCN(CC)S(=O)(=O)N1CCC[C@H](C(=O)NC2CCCCC2)C1	3.58	0.796
C[C@@H](NS(C)(=O)=O)c1ccccc1	1.8	0.797
CC[C@]1(C)NC(=S)N2N1C(=S)N[C@@]2(C)CC	2.44	0.730
CCCc1[nH]nc2c1[C@H](c1ccsc1C)C(C#N)=C(N)O2	3.83	0.911
CC(C)(C)C(=O)Nc1ccc(NC(=O)C(C)(C)C)cc1	2.79	0.865
COc1ccc(-c2nnc(N)s2)cc1OC	2.34	0.881
N#CCc1nc(-c2ccc3c(c2)OCCO3)cs1	2.92	0.830
Nc1ccccc1Nc1ccccc1	2.67	0.703
O=C1c2ccccc2C(=O)N1c1ccccc1[S+]([O-])C(F)(F)F	3.3	0.790
CCN(CC)C(=O)[C@@]12CC[C@@](C)(C(=O)O1)C2(C)C	2.66	0.722
COc1ccc2c(C)c(C)c(=O)oc2c1C	3.45	0.690
CC(C)OC(=O)c1ccc2[nH]c3c(c2c1)CCCC3	4.97	0.834

COc1ccc([C@@H]2CCCC[C@H]2O)cc1	2.85	0.805
Cc1cccc(NC(=O)NC(C)C)c1C	2.7	0.767
Cc1cc(=O)oc2cc(OC(C)C)ccc12	3.33	0.727
Cc1cc(Cl)ccc1O[C@@H](C)C(=O)NC1CCCCC1	5.45	0.844
CC(C)(C)NC(=S)Nc1ccc(F)cc1	2.71	0.719
CC1(C)CC(=O)C2=C(C1)OC(N)=C(C#N)C21CCC2(CC1)OCCO2	3.24	0.725
Cc1cc(C)n(CCOC(=O)[C@H]2[C@H](C=C(Cl)Cl)C2(C)C)n1	4.76	0.774
N#CC1=C(N)OC2=NN=C(c3cccs3)[C@@H]2[C@@H]1C1CCCCC1	4.64	0.903
C[C@H]1CN(c2nc3ccccn3c(=O)c2C=C(C#N)C#N)C[C@@H](C)O1	3.24	0.775
CCn1c(SCC(=O)NC(C)(C)C)nnc1C1CCCCC1	3.2	0.843
COc1ccc(CNC(=O)C2CCCC2)cc1	2.6	0.867
COc1cc(C(=O)Nc2ccccc2C(C)(C)C)cc(OC)c1OC	3.61	0.881
O=C(CS(=O)(=O)Cc1ccccc1)Nc1ccc2c(c1)OCCO2	3.19	0.894
CCCC(=O)N1C[C@]2(C)CN(C(=O)CCC)C[C@](C)(C1)C2	2.69	0.799
CCOC(=O)N1CCN(c2nc(C(F)(F)F)nc3c(C)cc(C)cc23)CC1	5.92	0.795
Cc1ccc(S(=O)(=O)N2CCCN(S(=O)(=O)c3ccc(C)cc3C)CC2)c(c)c1	5.01	0.738
CCOC(=O)c1c(NC(=O)[C@@H]2C[C@H]3CC[C@@H]2C3)sc(C(=O)N(CC)CC)c1C	5.05	0.692
N#CC1=C(N)OC2=C(C(=O)C[C@H](c3ccccc3)C2)[C@H]1c1ccccc1F	3.07	0.877
COc1ccc(OCC(=O)N2CCN(C(=O)c3ccco3)CC2)cc1	2.76	0.824
CCOc1ccc(C2=NN=C3OC(N)=C(C#N)[C@@H](c4ccco4)[C@@H]32)c(OCC)c1	4.48	0.806
CC1=CC(C(F)(F)F)(C(F)(F)F)C(C#N)=C(N)N1c1ccc(C(N)=O)cc1	4.41	0.757
COc1ccccc1Nc1nc(OCC(F)(F)F)nc(OCC(F)(F)F)n1	5.27	0.715
CC[C@H](C(=O)N1CCCN(C(=O)[C@@H](CC)c2ccccc2)CC1)c1ccccc1	4.93	0.729
Cc1ccc(S(=O)(=O)N(CCO)Cc2ccccc2)cc1	3.69	0.889
Cc1ccc(-c2cc(C(F)F)n3ncc(C(=O)N4CCSCC4)c3n2)cc1	4.55	0.684
O=C1c2ccccc2C(=O)N1c1ccccc1C(F)(F)F	3.55	0.753
O=C(NC1CCCC1)C1c2ccccc2Oc2ccccc21	3.95	0.909
CCc1ccc(OCC(=O)NCCc2ccc(OC)c(OC)c2)cc1	4.22	0.760
O=C(/C=C/c1ccc2c(c1)OCO2)Nc1ccc2c(c1)OCO2	3.72	0.883
COc1cccc(Cn2cc(N)cn2)c1	1.55	0.822
Cc1ccccc(C(=O)Nc2ccc(NC(=O)c3ccccc3)c(C)c2)c1	4.14	0.713
Cn1ncc(Cl)c1-c1nc(-c2nonc2N)no1	2.77	0.721
O=C([C@H]1CCCO1)N1CCN(C(=S)Nc2ccc(Oc3ccccc3)cc2)CC1	4.18	0.777
COc1ccc([C@@H]2[C@@H]3CCCC=C3C(C#N)=C(N)C2(C#N)C#N)cc1COCC(F)(F)F	5.5	0.723
O=C(c1nn2c(c1Br)N[C@H](c1cccs1)C[C@H]2C(F)(F)F)N1CCCC1	4.47	0.729
CCc1nc2sc(C#N)c(N)n2c(=O)c1C#N	2.09	0.792
O=C(c1c(F)cccc1F)N1CCCC1	1.88	0.696
C[C@@H](OC(=O)c1cccc(C(F)(F)F)c1)C(N)=O	2.75	0.842
Cc1cccc(NC(=O)CSc2ccccn2)c1C	3.19	0.867
COc1cc2c(cc1OC)[C@@]1(C)N(CC2)C(=O)OC12CCCC2	3.47	0.830
O=C(c1cccnc1)N1CCN(S(=O)(=O)c2ccccc2)CC1	2.3	0.848
Cc1cc(C)n(C2=NCCN2S(=O)(=O)c2cc(C(F)(F)F)cc(C(F)(F)F)c2)n1	4.7	0.672
COc1cccc(NC(=O)CSc2nnc3ccccn23)c1	2.45	0.732

SMILES		
N#CC1=C(N)C(C#N)(C#N)[C@@H](c2ccccc2Cl)[C@@H]2CSCC=C12	4.5	0.833
CC[C@@H]1CCCCN1C(=O)[C@H]1CC(c2cc(Cl)c(OC)cc2OC)=NO1	4.55	0.781
Cc1nc(N(C)C)c2sc(=S)n(-c3ccccc3)c2n1	3.62	0.677
CCOC(=O)C1CCN(S(=O)(=O)c2nc3nc(C)cc(C)n3n2)CC1	2.71	0.732
O=C(Cn1c(=O)n(-c2ccccc2)c(=O)n1CC(=O)N1CCCCC1)N1CCCCC1	2.78	0.706
COc1cc2c3c(c1)/C(=C1/SC(NC(C)=O)=NC1=O)C(=O)N3C(C)(C)C=C2C	4.62	0.736
CC1=C(C(=O)N2CCCCC2)[C@H](c2cccs2)C2=C(O)CCCC2=N1	3.37	0.843
CCCCC(=O)N[C@H]1[C@H]2C[C@@H]3C[C@H](C2)C[C@H]1C3	3.8	0.796
Cc1nccn1S(=O)(=O)c1ccc(Cl)cc1	2.93	0.826
COc1cccc([C@@H]2C3=C(CCCC3=O)N=C(C)[C@H]2C(=O)OC[C@@H]2CCCO2)c1	3.62	0.709
O=C1c2ccc(-c3ccc4c(c3)C(=O)N(CCO)C4=O)cc2C(=O)N1CCO	3.19	0.731
Cc1ccc(S(=O)(=O)/N=C2\SC(=CN(C)C)C(=O)N2C2CCCCC2)cc1	4.75	0.716
CC1(C)OC[C@H](COc2c(F)c(N3CCOCC3)nc(F)c2Cl)O1	4.07	0.764
COc1cccc(C2=NN=C3OC(N)=C(C#N)[C@H](c4cc(OC)c(OC)cc4Br)[C@@H]32)c1	4.52	0.694
COC(=O)c1ccc(N(C(=O)C(F)(F)F)[C@@H]2C=CS(=O)(=O)C2)cc1	2.05	0.763
CNc1nc2c(C)cccn2c(=O)c1C=O	2.68	0.757
CC(=O)Nc1nc2cc3sc(NC(C)=O)nc3cc2s1	3.13	0.762
CCOC(=O)c1[nH]c2c(c1C)C(Cl)=C(C=O)CC2	3.92	0.676
C[C@@H](NC(=O)C1CCN(C(=O)c2ccc(F)cc2)CC1)c1ccccc1	3.48	0.912
C[C@@]1(c2ccc(Cl)cc2)NC(=S)N2[C@@H]1NC(=S)N2c1ccc(Cl)cc1	4.94	0.728
CN(C)C(/C=C/Nc1cc(Cl)cc(Cl)c1)=C(C#N)C#N	1.59	0.678
N#C[C@@H](Cc1cccs1)c1ncc(C(F)(F)F)cc1Cl	4.09	0.825
CC1(C)CC(=O)C2=C(C1)OC(N)=C(C#N)[C@H]2c1ccc(C(F)(F)F)cc1	4.55	0.813
COC(=O)c1c(C)sc2c1OC(N)=C(C#N)[C@@H]2c1cc(Br)cs1	5.06	0.758
CC(C)(C)C(=O)[C@H]1[C@H](c2ccsc2)C(C#N)(C#N)[C@H]2C=Cc3ccccc3N12	3.37	0.744
Cc1ccc(-n2c(C)cc(=O)cc2C)cc1C	2.95	0.733
COc1cc([C@@H]2C(C#N)=C(N)Oc3cc(C)[nH]c(=O)c32)cc(OC)c1OC	2.55	0.844
COc1c(Cl)ccc(Cl)c1C(=O)Nc1ccc(C)cc1C	4	0.874
CCOC(=O)c1nc(N)sc1C	1.74	0.704
CC(C)CCn1c(NCCO)nc2c1c(=O)n(C)c(=O)n2C	2.93	0.775
Cc1ccc(NC(=O)[C@@H]2CCCN2C(=O)OC(C)(C)C)cc1C	3.9	0.905
CCCC[C@@]1(c2ccccc2)NC(=S)N(c2ccc(C)cc2)N1	5.1	0.804
CCCC(=O)OCC(=O)Nc1ccc(F)cc1Cl	3.02	0.839
Cc1ccccc1-c1noc(CCl)n1	3.3	0.712
O=C1[C@@H](CSCc2ccccc2)NC(=S)N1C1CCCCC1	4.58	0.837
CCOC(=O)N1CCN(C(=O)c2ccc(N3C(=O)[C@H]4CCCC[C@@H]4C3=O)cc2)CC1	3.74	0.709
CC(C)CCn1c(NCCO)nc2c1c(=O)n(C)c(=O)n2C	5.36	0.775
CN1C(=CC=CC(=O)C(F)(F)F)C(C)(C)c2ccccc21	4.41	0.769
C[C@H]1CCc2nc(NC(=O)c3ccc(F)c(F)c3)sc2C1	4.79	0.916
CCO/C(O)=C/c1nc2nonc2nc1O	1.61	0.728
Cc1ccccc1OCC(=O)NCCNC(=O)COc1ccccc1C	4.25	0.673
COc1nc(Cl)nc(Nc2ccccc2)n1	2.89	0.886
CCN(C(=O)CN1C(=O)[C@H]2CCCC[C@H]2C1=O)c1nnc(-c2ccc(Cl)cc2)s1	4.91	0.676

COc1ccc([C@@H](C)n2c(C)c3c(C)nnc(C)c3c2C)cc1	3.43	0.724
CCc1cc2c3c(c1)[C@]1(C(=O)N3C(C)(C)C=C2C)C(C#N)=C(N)OC2=C1C(=O)CC(C)(C)C2	5.39	0.694
CCCCN1C(=O)N(C)[C@@H]2[C@H]1N(C)C(=O)N2CCCC	2.69	0.747
CCc1ccc([C@@H]2C(C(=O)OC(C)C)=C(C)N=C3SCCC(=O)N32)cc1	4.75	0.753
Cc1cc2c(c(=O)o1)[C@@H](c1cccc(F)c1F)C(C#N)=C(N)O2	3.1	0.872
O=C(CSc1nnc(-c2cccs2)o1)NCCN1C(=O)CSC1=O	2.31	0.719
O=C(NCCc1ccccc1)c1ccc(Cl)cc1	3.71	0.896
O=C(c1cccc(S(=O)(=O)N2CCOCC2)c1)N1CCN(c2cccc(Cl)c2)CC1	4.34	0.716
Cc1cccc(C)c1NC(=O)Cn1cc(C(=O)C2CC2)c2ccccc21	3.47	0.692
C[C@@H]1CC[C@@]2(CC1)N=C(c1ccccc1)C(=O)N2CC(=O)Nc1ccc2c(c1)OCCO2	5.07	0.798
C[C@]12CO[C@]3(C(=O)NCC(F)(F)F)C[C@H]1CC[C@]23C	3.11	0.840
COCCCNC(=O)c1cc(Br)c(C)c(C)c1C	2.93	0.848
Cc1nn(-c2ccccc2)c(N)c1-c1ccccc1	3.27	0.755
Cn1ncc(Br)c1C(=O)N1CCCCCC1	2.63	0.793
COc1cc([C@H]2NC(=O)NC(C)=C2C(=O)OC(C)C)ccc1OC(F)F	3.75	0.752
CC1=NN(C(=O)c2ccccc2Cl)[C@](O)(c2ccc(Cl)cc2)C1	3.73	0.888
COc1ccccc1/C=N/n1c(C)cc(C)c(C#N)c1=O	3.08	0.810
O=C(NC[C@H]1CCCO1)c1cc2nc(C3CC3)cc(C(F)(F)F)n2n1	3.33	0.916
C[C@@H]1CCCCN1C(=O)CCCCC(=O)N1CCCC[C@H]1C	3.42	0.730
CCCC1=C(C(=O)OCC)[C@H](c2ccc3c(c2)OCO3)NC(=O)N1	3.59	0.808
FC(F)Oc1ccc(NC(=S)NC[C@@H]2CCCO2)cc1Cl	3.74	0.806
Cc1c(CN(C)C(=O)[C@@H](C)n2nc(C(F)F)c(Br)c2C)cnn1C	3.12	0.769
O=C(c1sc2cc(Cl)ccc2c1Cl)N1CCN(C(=S)NC[C@H]2CCCO2)CC1	4.83	0.704
Cc1c(Cl)c(C(F)(F)F)nn1[C@@H](C)C(=O)N1CCOCC1	3.08	0.837
O=C(CN1N=N[C@@H]2C(=O)N(c3ccc(Cl)cc3Cl)C(=O)[C@@H]21)Nc1ccccc1	4.3	0.773
O=C(NC1CC1)c1ccc(COc2c(F)c(F)cc(F)c2F)o1	3.54	0.676
CCOC(=O)N1CCN(C(=O)c2noc3c2C[C@@H](C(C)(C)C)CC3)CC1	4.52	0.807
CC(C)COc1ccc(C(=O)N2N=C(C(F)F)C[C@@]2(O)C(F)F)cc1	5.16	0.791
CCCNC(=O)N1CCN(C(=O)NCCC)CC1	1.04	0.782
CC(=O)N[C@H]1[C@@H](Oc2cccc3ccccc23)O[C@@H]2COC(C)(C)O[C@H]2[C@@H]1O	3.8	0.837
CCOCc1c(C)cc(C)c([C@@H]2C(C#N)=C(N)Oc3[nH]nc(C)c32)c1C	4.09	0.878
O=S(=O)(/N=C\c1ccc(Br)cc1)c1cccs1	2.14	0.810
O=C(CN1C(=O)c2ccccc2S1(=O)=O)Nc1ccc(S(=O)(=O)N2CCCCCC2)cc1	4.73	0.703
COc1cc(C(=O)Nc2cccc(C#N)c2)cc(OC)c1OC	3.44	0.917
CCOc1ccc(Br)cc1[C@H]1C2=C(CC(C)(C)CC2=O)Nc2ncnn21	3.85	0.814
CS(=O)(=O)N(CC(=O)Nc1c(F)cccc1F)c1cc(Cl)c(Cl)cc1Cl	4.34	0.699
C[C@H](C(=O)NC1CCCC1)N(c1ccc(Cl)cc1)S(C)(=O)=O	3.48	0.892
CC(=O)c1ccc(C(C)(C)C)cc1NC(=O)c1ccccc1	5.13	0.850
Cc1ccc([C@@H](O)P(=O)(OCC(C)C)OCC(C)C)cc1	4.44	0.716
C/C(=C\c1ccco1)[C@H]1C(C#N)=C(N)Oc2n[nH]c(-c3ccccc3)c21	4.82	0.749
COC(=O)c1cccc(NC(=O)c2ccc(C)cc2)c1	3.72	0.871
COc1cc(OC)c([C@H]2C3=C(CCCC3=O)Nc3ncnn32)cc1OC	2.19	0.899

O=P(Oc1ccccc1)(Oc1ccccc1)O[C@@H]1C=CS(=O)(=O)C1	2.48	0.727
COc1ccc(N2C(=O)Cc3ccccc3C2=O)cc1	2.56	0.784
C[C@H](NC(=O)C1CC1)c1ccccc1	1.85	0.776
Cc1cccc(NC(=O)Cc2ccc(Br)cc2)c1C	4.05	0.900
CC[C@@H](SC1=Nc2ccccc2C2=N[C@@H](Cc3ccccc3)C(=O)N12)C(=O)N1CCOCC1	3.95	0.683
Cc1cc(SCC(=O)NCc2ccc3c(c2)OCO3)nc2ccccc12	4.15	0.698
CC[C@H](C)NC(=O)NC(C(F)(F)F)C(F)(F)F	2.84	0.757
CCc1ccccc1NC(=O)[C@]12CC[C@](C)(C(=O)O1)C2(C)C	3.31	0.871
C[C@@H](CCc1ccccc1)NC(=O)C(=O)NC[C@@H]1COc2ccccc2O1	4.8	0.765
C[C@@H](NC(=O)CN(C1CCCCC1)S(=O)(=O)c1ccccc1)c1ccccc1	4.9	0.766
Cc1sc2ncn(CC(=O)N[C@@H]3CCC[C@@H](C)[C@H]3C)c(=O)c2c1C	4.34	0.928
O=C(NC[C@@H]1CCCO1)c1sc2ccccc2c1Cl	3.56	0.941
O=C(c1nn(-c2ccccc2)c(=O)c2ccccc12)N1CCOCC1	2.47	0.716
CC1CCN(S(=O)(=O)N2CCC(C(=O)N[C@@H](C)c3ccccc3)CC2)CC1	4.92	0.835
O=C(Nc1ccccc1C(F)(F)F)C1CCC1	2.55	0.846
COCON1C(=O)CN(C(C)=O)[C@H]1c1ccc(Cl)cc1Cl	1.98	0.793
Cc1ccc(N2C(=O)N(c3cccc(Cl)c3)[C@@H]3CS(=O)(=O)C[C@@H]32)cc1	3.56	0.755
Cc1cc(C)nc(-n2nc(C)c3c2NC2=CC(C)(C)CC(=O)[C@H]2[C@@H]3c2ccco2)n1	4.34	0.688
O=C(OCC(=O)N1CCC(Cc2ccccc2)CC1)C1CCCCC1	4.98	0.767
COC(=O)CN(c1ccccc1)S(=O)(=O)c1ccc(OC)cc1	3.43	0.755
Cc1[nH]c2ccccc2c1C1=CN2CCS(=O)(=O)N=C2S1	1.56	0.876
Cc1onc(-c2c(F)cccc2Cl)c1C(=O)OCC(=O)N1[C@H](C)CCC[C@H]1C	4.38	0.701
CC[C@H]1CS/C(=N\c2cccc(C)c2)N1C(=O)CCN1C(=O)c2ccccc2C1=O	4.96	0.683
CCOC(=O)c1[nH]c(C)c(C(=O)OCC(=O)N2CCc3ccccc32)c1C	4.47	0.817
OC[C@H]1CCCN(c2ncnc3scc(-c4ccccc4)c23)C1	3.9	0.799
CC(=O)N1Cc2ccccc2C[C@H]1C(=O)OCc1ccc(C#N)cc1	2.41	0.809
CC(C)NC(=S)N1CCN(S(=O)(=O)c2cccc(C(F)(F)F)c2)CC1	3.9	0.796
O=C(Cn1cnnn1)N(Cc1ccco1)Cc1cccs1	4.33	0.690
CCOC(=O)c1[nH]c(C)c(C(=O)OCC(=O)N(C)[C@@H]2CCS(=O)(=O)C2)c1C	3.35	0.694
CC(=O)NCCc1ccc(S(=O)(=O)N2CCOCC2)cc1	2	0.853
COC(=O)CCNc1nc(-c2cccnc2)nc2sc3c(c12)CCCC3	4.69	0.693
CCOC(=O)C1=C(C)NC(C)=C(C(=O)OC)[C@H]1c1ccco1	3.67	0.859
O=C(Nc1ccc2c(c1)OCO2)[C@@H]1COc2ccccc2O1	3.28	0.920
CCOC(=O)/C=c1/s/c(=C/c2c(F)cccc2F)c(=O)n1CC(=O)N(C)C	3.17	0.688
COc1cc(/C=C/C(=O)N[C@H]2CCCc3ccccc32)cc(OC)c1OC	4.27	0.786
C[C@H]1CN(C(=O)c2csc3c2CCCC3)/C(=N/c2ccc(Cl)cc2)S1	5.53	0.673
COc1cc2c(cc1OC)CN(C(=O)c1cccc(S(=O)(=O)N3CCCC3)c1)CC2	4.2	0.708
COC(=O)C1=C(C)N=C(C)[C@@H](C(=O)OC)[C@@H]1c1cccc(Cl)c1	4.22	0.795
CC(C)[C@@H](Sc1nnc(C(F)(F)F)n1C)C(=O)NC(N)=O	2.12	0.808
O=C(Cc1coc2cc3c(cc12)CCC3)O[C@H]1CCCCC1=O	4.43	0.813
CC(C)NC(=O)COC(=O)c1ccc(S(=O)(=O)N2CCCCCC2)cc1	3.78	0.759
COc1cccc(C(=O)N2CCC[C@H](c3ccccc3)C2)c1	4.04	0.863
O=C(Cn1cc(C(F)(F)F)cc(Cl)c1=O)N1CCN(C(=O)c2ccco2)CC1	2.27	0.766
COc1cccc(NC(=O)[C@H](OC(=O)C2=COCCO2)c2ccccc2)c1	3.81	0.788

SMILES		
CCC(CC)NC(=O)c1ccc(Cl)c(S(=O)(=O)N2CCCC2)c1	3.6	0.849
CCOC(=O)C1=C(C)NC(=O)C[C@@H]1c1cc(OC)c(OC)cc1Br	3.93	0.772
COC(=O)c1ccc(C)c(NC(=O)Cc2coc3cc4c(cc23)CCC4)c1	4.78	0.705
CCN(CC)S(=O)(=O)c1ccc(/C=C/C(=O)NC2CCCCC2)cc1	4.34	0.755
O=C(COC(=O)c1ccccc1F)Nc1ccccc1	3.18	0.871
COC(=O)c1ccc(Br)c(S(=O)(=O)N2CCCCC2)c1	3.66	0.775
CCOc1ccc(NC(=O)CN(C)C(=O)c2ccc(C)o2)cc1OCC	2.09	0.782
CCOC(=O)N1CCN(C(=O)/C=C/c2ccc(Cl)c(Cl)c2)CC1	4.34	0.780
C[C@H]1C[C@H](C)CN(C(=O)CSc2nnc(-c3ccccc3)n2C)C1	3.68	0.799
COc1ccc(C(=O)Nc2cccc(F)c2)cc1S(=O)(=O)N1CCOCC1	3.64	0.839
Cc1ccc(CNC(=O)C2CCN(S(=O)(=O)c3ccc4c(c3)OCCCO4)CC2)cc1	4.01	0.766
C[C@H](Sc1nnc(-c2ccc(F)cc2)n1-c1ccccc1)C(N)=O	3.38	0.723
C[C@@H](Oc1ccccc1)C(=O)Nc1ccc(OC(F)(F)F)cc1	4.77	0.904
COCCCNC(=O)c1ccccc1C(F)(F)F	2	0.826
Cc1noc(C)c1C(=O)O[C@@H](C)C(=O)Nc1ccc2c(c1)OCCO2	2.78	0.847
C[C@H]1CCCC[C@@H]1NC(=O)COc1ncnc2ccccc12	3.11	0.942
C[C@@H]1CCCC[C@H]1NC(=O)COC(=O)C1=COCCO1	2.48	0.782
CC(C)N(C(=O)CSc1nnc(NCc2ccccc2)s1)C(C)C	4.09	0.720
O=C1Nc2cc(C(=O)NC3CC3)ccc2S/C1=C\c1ccccc1	4.32	0.841
CCOC(=O)c1ccc(NCc2c(C)[nH]c(C(=O)OCC)c2C)cc1	5.09	0.750
c1nc(NC2CCCC2)c2c3c(sc2n1)CCCC3	4.17	0.900
CCC[C@@H](C)NC(=O)COc1ncnc2sc3c(c12)CC[C@@H](C)C3	4.55	0.868
Cc1cc(Br)ccc1NC(=O)CSc1nnc(N)s1	2.69	0.820
Cc1ccc(OCCC(=O)N2CCN(S(=O)(=O)c3cccs3)CC2)cc1	2.98	0.754
CN(C)S(=O)(=O)c1cccc(C(=O)NCCc2ccc(Cl)cc2Cl)c1	4.48	0.808
O=C(OCc1c(F)cccc1Cl)c1cccc(S(=O)(=O)N2CCCC2)c1	4.22	0.722
CN(C)S(=O)(=O)c1ccc(C(=O)OC2CCCCC2)cc1	4.27	0.801
Cc1cccc(NC(=O)C[C@@H]2S/C(=N\c3c(C)cccc3C)N(C)C2=O)c1	4.75	0.863
CCOc1ccccc1OCC(=O)Nc1ccc2ccccc2c1	4.73	0.737
OCC#CC(O)(c1ccc(F)cc1)c1ccc(F)cc1	3.61	0.824
CC(=O)Nc1ccc(/N=C/c2c([O-])oc3ccccc3c2=O)cc1	3.64	0.750
CSc1sc(C(=O)N2CCOCC2)c2c1S(=O)(=O)N(C)c1ccccc1-2	3.11	0.712