

Molecular simulation of the (GPx)-like antioxidant activity of ebselen derivatives through machine learning techniques

Luis Calle^{a,b}, Yovani Marrero-Ponce^{c,d}, José R. Mora^{c,*}

^a Instituto de Investigación e Innovación en Salud Integral (ISAIN), Facultad de Ciencias Médicas, Universidad Católica Santiago de Guayaquil, Guayaquil 09013493, Ecuador.

^b Faculty of Pharmacy, University of Granada, Granada 18011, Spain.

^c Universidad San Francisco de Quito, Grupo de Química Computacional y Teórica (QCT-USFQ), Departamento de Ingeniería Química, Diego de Robles y Vía Interoceánica, 17-1200-841 Quito, Ecuador.

^d Universidad San Francisco de Quito, Grupo de Medicina Molecular y Traslacional (MeM&T), Colegio de Ciencias de la Salud (COCSA), Escuela de Medicina, Diego de Robles y Vía Interoceánica, 17-1200-841 Quito, Ecuador.

* Corresponding Author: José R. Mora, email: jrmora@usfq.edu.ec

SUPPLEMENTARY MATERIAL

	Page
Table S1. Dataset used for the modelling with the experimental values of $\log(v_0)$	S2
Table S2. Topographical descriptors used for the construction of the QSAR models.	S3
Table S3. Models obtained for the different machine learning techniques	S8
Table S4. Attributes of each model M1-M21	S9
Table S5. Evaluation of the applicability domain of the external set (from reference 41) in model M2 and prediction of $\log(v_0)$	S10
Table S6. Evaluation of the applicability domain of the external set (from reference 41) in model M15 and prediction of $\log(v_0)$	S12
Table S7. Evaluation of the applicability domain of the external set (from reference 41) in model M19 and prediction of $\log(v_0)$	S14

Table S1. Dataset used for the modelling with the experimental values of $\log(v_o)$

Molecule	Smiles	$\log(v_o)$		
		H ₂ O ₂	<i>t</i> -BuOOH	Cum-OOH
2a (=A)	<chem>c1cc2c(cc1)c(=O)n([se]2)c1ccccc1</chem>	2.094	1.439	1.796
2b	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H](Cc1ccc(cc1)O)C(=O)N[C@@H](C(C)C)C(=O)OC</chem>	1.904	1.29	2.072
2c	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H](Cc1ccccc1)C(=O)N[C@@H](C(C)C)C(=O)OC</chem>	1.986	0.924	1.45
2d	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H](Cc1ccccc1)C(=O)N[C@@H]([C@H](C)CC)C(=O)OC</chem>	1.75	0.748	1.461
2e	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H](Cc1ccccc1)C(=O)N[C@@H](C)C(=O)OC</chem>	2.428	0.748	1.736
2f	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H](Cc1cc2(cccc2)[nH]1)C(=O)N[C@@H](C)C(=O)OC</chem>	2.328	1.35	1.977
2g	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H](Cc1cc2(cccc2)[nH]1)C(=O)NCC(=O)OC</chem>	2.167	1.27	2.338
2h	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H](C(C)C)C(=O)N[C@@H](C)C(=O)OC</chem>	2.339	1.74	1.941
2i	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H](C(C)C)C(=O)N[C@@H](C)C(=O)N[C@@H](Cc1ccccc1)C(=O)OC</chem>	2.328	1.241	1.646
3a	<chem>c1cc2c(cc1)c(=O)n([se]2)c1ccc(cc1)CN1CCCC1</chem>	1.497	---	0.23
3b	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H]1c2c(CCC1)cccc2</chem>	1.169	---	0.762
3c	<chem>c1cc2c(cc1)c(=O)n([se]2)c1ccc(cc1)CCN1CCCC1</chem>	1.77	---	1.313
3d	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H](C)c1c2ccccc2ccc1</chem>	1.588	---	1.204
3e	<chem>c1cc2c(cc1)c(=O)n([se]2)c1ccc(cc1)OCCN1CCCC1</chem>	1.636	---	1.057
3f	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@@H](c1ccc(cc1)N(O)O)C</chem>	1.578	---	1.367
3g	<chem>c1(cc2c(cc1OC)c(=O)n([se]2)c1ccc(cc1)CN1CCCC1)OC</chem>	2.15	---	1.684
3h	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@@H](c1ccc(cc1)OC)C</chem>	2.263	---	2.082
3i	<chem>c1(cc2c(cc1OC)c(=O)n([se]2)c1ccc(cc1)CN1CCCC1)OC</chem>	1.491	---	0.762
4a	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H](Cc1ccccc1)C(=O)OCC</chem>	2.242	---	---

4b	<chem>c1(cc2c(cc1OC)c(=O)n([se]2)c1ccc(cc1)CN(C)Cc1cccc1)OC</chem>	2.082	---	---
4c	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H](C)C1CCCC1</chem>	2.022	---	---
4d	<chem>c1(cc2c(cc1OC)c(=O)n([se]2)c1ccc(cc1)CCN1CCCC1)OC</chem>	2.082	---	---
4e	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H](C)C(C)C</chem>	2.194	---	---
4f	<chem>c1(cc2c(cc1OC)c(=O)n([se]2)c1ccc(cc1)CCN1CCCC1)OC</chem>	1.798	---	---
4g	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H](CO)C(C)C</chem>	2.073	---	---
4h	<chem>c1(cc2c(cc1OC)c(=O)n([se]2)c1ccc(cc1)OCCN1CCCC1)OC</chem>	2.105	---	---
4i	<chem>c1cc2c(cc1)c(=O)n([se]2)[C@H]1[C@@H](CCCC1)n1c(=O)c2c(cccc2)s1</chem>	2.11	---	---
4j	<chem>c1(cc2c(cc1OC)c(=O)n([se]2)c1ccc(cc1)OCCCN1CCCC1)OC</chem>	2.077	---	---
4k	<chem>c1(cc2c(cc1OC)c(=O)n([se]2)c1ccc(cc1)OCCCN1CCCC1)OC</chem>	1.955	---	---
5a	<chem>c1cc2c(cc1)c(=O)n([se]2)c1c2cccnc2c(cc1)O</chem>	2.201	1.449	1.925
5b	<chem>c1(cc2c(cc1)c(=O)n([se]2)c1c2cccnc2c(cc1)O)Cl</chem>	2.171	1.461	1.972
5c	<chem>c1cc2c(cc1Cl)c(=O)n([se]2)c1c2cccnc2c(cc1)O</chem>	2.187	1.473	2.078
5d	<chem>c1(cc2c(cc1)c(=O)n([se]2)c1c2cccnc2c(cc1)O)F</chem>	2.123	1.417	1.98
5e	<chem>c1cc2c(cc1F)c(=O)n([se]2)c1c2cccnc2c(cc1)O</chem>	2.21	1.47	2.073
5f	<chem>c1cc2c(c(c1)F)c(=O)n([se]2)c1c2cccnc2c(cc1)O</chem>	2.397	1.435	2.257
5g	<chem>c1(cc2c(cc1)c(=O)n([se]2)c1c2cccnc2c(cc1)O)C(F)(F)F</chem>	2.005	1.403	1.935
5h	<chem>c1(cc2c(cc1)c(=O)n([se]2)c1c2cccnc2c(cc1)O)OC</chem>	1.947	1.238	1.667
5i	<chem>c1cc2c(cc1OC)c(=O)n([se]2)c1c2cccnc2c(cc1)O</chem>	1.986	1.328	1.822
5j	<chem>c1(cc2c(cc1OC)c(=O)n([se]2)c1c2cccnc2c(cc1)O)OC</chem>	1.872	1.21	1.691
5k	<chem>c1cc2c(cc1O)c(=O)n([se]2)c1c2cccnc2c(cc1)O</chem>	2.262	1.651	2.184
5l	<chem>c1cc2c(cc1O)c(=O)n([se]2)c1c2cccnc2c(c(c1)Cl)O</chem>	2.373	1.845	1.886
5m	<chem>c1cc2c(cc1O)c(=O)n([se]2)c1c2cccnc2c(c(c1)I)O</chem>	2.008	1.356	2.297

Table S2. Topographical descriptors used for the construction of the QSAR models.

P2_B_BB_nCi_2_MP0_H_A_KA_p-s_MAS	GM_Q_BB_Ci(2.0;-2.0)_2_SS3_T_SRW_h_MAS	RA_B_AB_nCi_2_MP5_n_C_LGP[4-6]_e-p_MAS
HM_B_AB_Ci(2.0;-2.0)_2_NS4_H_A_KA_v-e_MAS	N3_B_AB_nCi_2_SS7_X_KA_r-c_MAS	Q2_Q_AB_Ci(2.0;-2.0)_2_MP6_n_X_LGP[1-3]_h_MAS
N3_F_AB_Ci(2.0;-2.0)_2_SS7_T_KA_m_MAS	AM_Q_AB_Ci(2.0;-2.0)_2_MP1_A_KA_a_MAS	P2_Q_AB_Ci(2.0;-2.0)_2_NS5_n_T_LGP[1;2;6]_a_MAS
K_B_AB_nCi_2_SS8_X_NSRW_v-s_MAS	N3_B_AB_nCi_2_SS7_H_T_KA_e-c_MAS	RA_B_AB_Ci(2.0;-2.0)_2_SS5_n_X_LGP[1;2;6]_r-h_MAS
P2_F_BB_nCi_2_SS3_H_A_KA_p_MAS	S_B_AB_nCi_2_NS2_T_SRW_c-s_MAS	K_B_AB_nCi_2_SS6_C_LGP[2-4]_r-c_MAS

S_B_AB_nCi_2_SS7_A_SRW_v-e_MAS	N1_B_AB_nCi_2_MP5_n_X_KA_m-h_MAS	P2_Q_AB_nCi_2_NS6_X_LGP[1;2;6]_a_MAS
AM_B_BB_nCi_2_SS7_H_T_NSRW_c-m_MAS	N3_Q_AB_nCi_2_MP8_T_KA_v_MAS	RA_F_AB_nCi_2_NS8_X_LGP[4-6]_a_MAS
HM_F_AB_nCi_2_SS3_H_T_NSRW_v_MAS	AM_B_BB_Ci(2.0;-2.0)_2_SS7_H_A_LGP[1-3]_a-r_MAS	I50_F_AB_nCi_2_SS4_H_n_P_LGP[4-6]_a_MAS
V_Q_AB_nCi_2_NS4_P_NSRW_m_MAS	K_B_AB_Ci(2.0;-2.0)_2_SS4_H_n_X_LGP[2-4]_r-h_MAS	Q1_B_AB_Ci(2.0;-2.0)_2_MP2_n_X_LGP[1;2;6]_a-p_MAS
VC_B_AB_Ci(2.0;-2.0)_2_NS7_n_X_NSRW_v-s_MAS	I50_B_AB_nCi_2_NS4_n_X_LGP[1-3]_r-s_MAS	K_B_AB_nCi_2_NS4_C_LGP[1;2;6]_m-s_MAS
S_B_AB_nCi_2_SS1_X_NSRW_v-h_MAS	S_B_AB_nCi_2_NS3_X_LGP[1;2;6]_v-s_MAS	Q3_B_AB_nCi_2_MP8_T_LGP[1;2;6]_e-c_MAS
RA_B_AB_nCi_2_NS3_P_SRW_v-h_MAS	HM_F_AB_nCi_2_MP5_A_LGP[1-3]_m_MAS	P2_B_BB_Ci(2.0;-2.0)_2_MP8_H_A_LGP[2-4]_r-h_MAS
HM_B_AB_nCi_2_MP3_A_KA_p-s_MAS	N3_Q_AB_nCi_2_SS5_C_LGP[4-6]_p_MAS	I50_B_AB_nCi_2_SS7_T_LGP[1;2;6]_a-s_MAS
SD_B_AB_nCi_2_NS2_T_NSRW_a-m_MAS	P2_F_BB_Ci(2.0;-2.0)_2_SS5_H_A_LGP[2-4]_p_MAS	Q2_Q_AB_nCi_2_MP6_X_LGP[4-6]_e_MAS
MX_Q_AB_nCi_2_SS8_T_KA_h_MAS	K_B_AB_nCi_2_SS2_P_LGP[2-4]_a-e_MAS	SD_B_AB_Ci(2.0;-2.0)_2_MP2_H_n_X_LGP[2-4]_r-h_MAS
I50_F_AB_nCi_2_MP5_X_KA_r_MAS	VC_B_AB_nCi_2_NS4_X_LGP[4-6]_m-e_MAS	P2_B_AB_Ci(2.0;-2.0)_2_NS2_n_T_LGP[1;2;6]_a-r_MAS
S_B_AB_nCi_2_SS3_T_NSRW_r-s_MAS	VC_B_BB_Ci(2.0;-2.0)_2_SS3_X_LGP[1;2;6]_r-h_MAS	Q3_B_AB_nCi_2_SS8_T_LGP[1]_m-h_MAS
AM_B_BB_nCi_2_NS5_H_A_KA_p-h_MAS	S_B_AB_nCi_2_SS1_X_LGP[1;2;6]_v-h_MAS	I50_Q_AB_nCi_2_MP6_T_LGP[1;2;6]_a_MAS
S_F_AB_Ci(2.0;-2.0)_2_SS4_H_n_T_KA_m_MAS	VC_B_AB_nCi_2_NS6_P_LGP[4-6]_m-s_MAS	K_Q_AB_nCi_2_SS5_C_LGP[2-4]_c_MAS
S_B_AB_nCi_2_SS2_T_NSRW_p-s_MAS	V_B_AB_Ci(2.0;-2.0)_2_SS4_H_n_X_LGP[2-4]_m-e_MAS	VC_B_AB_nCi_2_SS6_X_LGP[4-6]_c-p_MAS
K_B_BB_Ci(2.0;-2.0)_2_SS3_H_P_SRW_e-s_MAS	K_B_AB_nCi_2_SS5_X_LGP[2-4]_a-s_MAS	I50_B_AB_nCi_2_NS4_H_n_P_LGP[4-6]_h-s_MAS
I50_B_AB_nCi_2_SS3_T_KA_e-s_MAS	HM_Q_BB_Ci(2.0;-2.0)_2_NS4_H_A_LGP[2-4]_v_MAS	SD_B_AB_Ci(2.0;-2.0)_2_MP8_H_n_X_LGP[4-6]_p-s_MAS
I50_B_AB_nCi_2_MP3_T_NSRW_p-h_MAS	S_B_AB_nCi_2_SS6_T_LGP[2-4]_v-p_MAS	I50_Q_AB_nCi_2_SS4_H_n_P_LGP[4-6]_h_MAS
S_Q_AB_nCi_2_SS2_T_NSRW_e_MAS	HM_B_BB_Ci(2.0;-2.0)_2_MP5_H_A_LGP[1;2;6]_e-p_MAS	K_Q_AB_nCi_2_NS5_T_LGP[2-4]_h_MAS
S_B_AB_nCi_2_NS4_X_KA_r-h_MAS	P2_F_AB_Ci(2.0;-2.0)_2_SS8_n_T_LGP[1;2;6]_v_MAS	GM_Q_AB_nCi_2_SS3_X_LGP[1;2;6]_h_MAS
HM_Q_BB_Ci(2.0;-2.0)_2_SS6_T_NSRW_r_MAS	K_B_AB_nCi_2_SS7_T_LGP[2-4]_e-p_MAS	N2_B_AB_nCi_2_MP8_A_LGP[4-6]_a-e_MAS
K_B_AB_nCi_2_SS2_A_KA_a-r_MAS	SD_B_AB_nCi_2_NS5_P_LGP[1;2;6]_m-h_MAS	RA_B_AB_Ci(2.0;-2.0)_2_SS7_n_P_LGP[4-6]_r-m_MAS
SD_F_AB_Ci(2.0;-2.0)_2_NS4_H_n_T_KA_a_MAS	GM_Q_AB_Ci(2.0;-2.0)_2_MP1_n_T_LGP[1;2;6]_p_MAS	HM_B_BB_Ci(2.0;-2.0)_2_NS3_H_A_LGP[1;2;6]_v-m_MAS
I50_B_AB_nCi_2_NS4_X_NSRW_e-s_MAS	N1_B_AB_nCi_2_NS7_X_LGP[4-6]_c-p_MAS	HM_B_AB_nCi_2_NS1_X_LGP[1;2;6]_r-h_MAS
K_B_AB_nCi_2_SS4_X_SRW_a-m_MAS	K_Q_AB_nCi_2_SS3_P_LGP[2-4]_m_MAS	SD_F_AB_Ci(2.0;-2.0)_2_NS5_H_n_X_LGP[1;2;6]_a_MAS
N1_B_AB_nCi_2_NS7_T_KA_a-c_MAS	K_F_AB_nCi_2_SS8_n_A_LGP[2-4]_a_MAS	HM_B_AB_nCi_2_SS8_X_LGP[1;2;6]_psa-v_MAS
P3_Q_BB_Ci(2.0;-2.0)_2_MP1_A_KA_r_MAS	HM_F_AB_nCi_2_SS3_X_LGP[1-3]_v_MAS	SD_B_AB_Ci(2.0;-2.0)_2_MP2_H_n_X_LGP[2-4]_e-h_MAS
I50_B_AB_nCi_2_SS6_T_NSRW_v-h_MAS	HM_B_BB_Ci(2.0;-2.0)_2_SS5_H_A_LGP[1;2;6]_m-p_MAS	K_B_AB_nCi_2_NS2_T_LGP[2-4]_a-v_MAS
P2_Q_AB_nCi_2_SS7_H_A_SRW_s_MAS	K_B_BB_Ci(2.0;-2.0)_2_NS7_H_X_LGP[4-6]_c-m_MAS	Q3_B_AB_nCi_2_SS7_T_LGP[1;2;6]_v-e_MAS
K_F_AB_nCi_2_NS2_T_KA_a_MAS	N1_Q_AB_nCi_2_SS8_H_n_T_LGP[4-6]_r_MAS	P2_Q_BB_Ci(2.0;-2.0)_2_NS2_H_A_LGP[2-4]_h_MAS

GM_B_AB_nCi_2_SS8_T_KA_h-s_MAS	I50_Q_AB_nCi_2_SS4_T_LGP[1;2;6]_v_MAS	Q2_B_AB_nCi_2_MP7_A_LGP[4-6]_a-r_MAS
I50_B_AB_nCi_2_NS6_n_X_NSRW_p-h_MAS	K_B_BB_Ci(2.0;-2.0)_2_SS4_H_X_LGP[4-6]_e-s_MAS	K_Q_AB_Ci(2.0;-2.0)_2_SS5_H_n_P_LGP[4-6]_s_MAS
I50_F_AB_nCi_2_NS8_n_X_NSRW_m_MAS	S_F_AB_nCi_2_NS4_T_LGP[2-4]_p_MAS	AM_B_BB_Ci(2.0;-2.0)_2_SS8_H_A_LGP[2-4]_r-c_MAS
GM_B_AB_nCi_2_NS2_A_NSRW_r-s_MAS	GM_B_AB_nCi_2_SS6_X_LGP[1;2;6]_e-p_MAS	SD_B_AB_Ci(2.0;-2.0)_2_MP1_H_n_X_LGP[1;2;6]_a-h_MAS
HM_B_AB_nCi_2_MP5_H_A_SRW_m-p_MAS	V_B_AB_nCi_2_SS7_T_LGP[2-4]_e-p_MAS	AM_B_AB_nCi_2_SS4_C_LGP[1-3]_c-h_MAS
S_Q_AB_nCi_2_SS6_T_NSRW_h_MAS	VC_B_AB_nCi_2_SS4_T_LGP[1;2;6]_e-h_MAS	Q2_B_AB_nCi_2_MP6_X_LGP[4-6]_p-s_MAS
P2_F_AB_nCi_2_NS5_T_KA_a_MAS	GM_Q_AB_nCi_2_SS7_H_T_LGP[2-4]_s_MAS	HM_B_AB_nCi_2_SS1_X_LGP[1;2;6]_r-p_MAS
P2_F_BB_nCi_2_MP7_H_A_KA_a_MAS	N2_B_BB_Ci(2.0;-2.0)_2_SS4_H_T_LGP[1-3]_m-h_MAS	SD_Q_AB_nCi_2_SS2_P_LGP[2-4]_h_MAS
VC_B_BB_Ci(2.0;-2.0)_2_SS4_X_NSRW_r-h_MAS	K_F_AB_nCi_2_SS7_A_LGP[1;2;6]_a_MAS	S_F_BB_Ci(2.0;-2.0)_2_SS5_H_X_LGP[4-6]_a_MAS
HM_B_AB_Ci(2.0;-2.0)_2_SS1_T_SRW_p-s_MAS	VC_F_AB_nCi_2_NS4_A_LGP[4-6]_m_MAS	P2_B_AB_nCi_2_SS3_X_LGP[1;2;6]_r-h_MAS
I50_B_AB_nCi_2_NS5_X_NSRW_r-h_MAS	S_B_AB_nCi_2_NS4_T_LGP[2-4]_v-h_MAS	P3_F_AB_nCi_2_SS1_H_T_LGP[1]_r_MAS
RA_B_AB_nCi_2_MP0_T_SRW_a-s_MAS	K_B_AB_nCi_2_NS4_T_LGP[2-4]_e-p_MAS	K_B_AB_nCi_2_SS2_T_LGP[2-4]_e-s_MAS
S_B_AB_nCi_2_SS7_T_NSRW_p-h_MAS	SD_B_AB_nCi_2_SS2_X_LGP[2-4]_v-h_MAS	RA_B_AB_nCi_2_MP5_C_LGP[4-6]_v-m_MAS
P2_F_AB_Ci(2.0;-2.0)_2_SS6_T_NSRW_s_MAS	AM_B_AB_nCi_2_SS6_X_LGP[1-3]_v-c_MAS	HM_Q_AB_nCi_2_SS2_X_LGP[1-3]_r_MAS
K_B_AB_nCi_2_NS5_X_SRW_r-e_MAS	AM_Q_AB_Ci(2.0;-2.0)_2_SS8_n_T_LGP[1;2;6]_h_MAS	N2_B_AB_nCi_2_MP5_H_n_T_LGP[4-6]_a-m_MAS
I50_F_AB_nCi_2_NS3_n_X_NSRW_h_MAS	N3_Q_AB_nCi_2_MP7_H_n_T_LGP[4-6]_s_MAS	Q2_Q_AB_Ci(2.0;-2.0)_2_SS4_n_P_LGP[4-6]_a_MAS
SD_Q_AB_nCi_2_SS7_n_A_NSRW_v_MAS	GM_B_AB_nCi_2_NS2_A_LGP[1;2;6]_r-s_MAS	S_B_AB_nCi_2_NS7_T_LGP[2-4]_m-p_MAS
K_B_AB_Ci(2.0;-2.0)_2_SS1_H_n_X_NSRW_a-s_MAS	K_F_AB_Ci(2.0;-2.0)_2_SS8_H_n_X_LGP[4-6]_a_MAS	I50_B_AB_nCi_2_MP2_T_LGP[1;2;6]_a-r_MAS
GM_B_AB_nCi_2_NS0_H_A_KA_m-s_MAS	S_Q_AB_Ci(2.0;-2.0)_2_NS8_H_n_X_LGP[4-6]_r_MAS	HM_Q_BB_Ci(2.0;-2.0)_2_NS4_H_A_LGP[1;2;6]_r_MAS
N2_B_AB_nCi_2_NS2_H_T_KA_p-s_MAS	S_B_AB_nCi_2_SS3_T_LGP[2-4]_v-p_MAS	GM_F_BB_Ci(2.0;-2.0)_2_NS7_H_A_LGP[1;2;6]_psa_MAS
P2_B_AB_nCi_2_SS2_X_KA_a-v_MAS	SD_B_AB_nCi_2_SS6_X_LGP[2-4]_a-r_MAS	P2_B_BB_Ci(2.0;-2.0)_2_SS7_H_T_LGP[4-6]_a-c_MAS
SD_B_AB_nCi_2_SS7_T_SRW_a-p_MAS	I50_B_AB_nCi_2_NS3_n_X_LGP[2-4]_v-h_MAS	MX_B_AB_nCi_2_SS5_H_n_P_LGP[1-3]_v-s_MAS
AM_B_BB_Ci(2.0;-2.0)_2_MP4_A_SRW_r-h_MAS	I50_B_AB_nCi_2_NS3_T_LGP[1;2;6]_a-p_MAS	Q1_B_AB_Ci(2.0;-2.0)_2_SS8_n_P_LGP[4-6]_a-psa_MAS
VC_B_AB_nCi_2_SS5_T_KA_r-h_MAS	V_F_AB_Ci(2.0;-2.0)_2_MP6_H_n_P_LGP[4-6]_p_MAS	I50_B_AB_nCi_2_MP1_n_P_LGP[1-3]_m-e_MAS
I50_B_AB_nCi_2_SS8_T_KA_a-r_MAS	P2_B_AB_nCi_2_SS3_H_T_LGP[2-4]_e-h_MAS	HM_Q_AB_nCi_2_NS1_X_LGP[1;2;6]_r_MAS
I50_F_AB_nCi_2_NS0_T_KA_s_MAS	K_Q_AB_nCi_2_SS4_T_LGP[2-4]_a_MAS	RA_Q_AB_nCi_2_SS2_T_LGP[1;2;6]_h_MAS
K_B_AB_nCi_2_NS3_T_NSRW_p-s_MAS	N3_B_AB_nCi_2_NS5_C_LGP[4-6]_e-h_MAS	P2_B_AB_Ci(2.0;-2.0)_2_SS8_H_A_LGP[1]_m-s_MAS
VC_B_AB_nCi_2_SS2_T_KA_c-s_MAS	V_B_AB_Ci(2.0;-2.0)_2_MP8_H_n_X_LGP[4-6]_c-s_MAS	HM_B_AB_Ci(2.0;-2.0)_2_MP4_H_A_LGP[1]_v-m_MAS
HM_B_BB_nCi_2_SS7_H_T_SRW_v-h_MAS	HM_Q_BB_Ci(2.0;-2.0)_2_NS3_H_A_LGP[2-4]_p_MAS	AM_B_BB_Ci(2.0;-2.0)_2_MP8_H_A_LGP[1;2;6]_h-s_MAS
N3_Q_AB_nCi_2_MP2_A_SRW_r_MAS	AM_F_AB_Ci(2.0;-2.0)_2_SS3_n_T_LGP[1;2;6]_e_MAS	N2_B_AB_nCi_2_MP4_A_LGP[4-6]_r-c_MAS
GM_Q_AB_Ci(2.0;-2.0)_2_MP3_H_T_SRW_s_MAS	N2_Q_AB_nCi_2_NS6_H_n_T_LGP[1]_a_MAS	I50_Q_AB_nCi_2_SS2_H_n_P_LGP[1-3]_h_MAS

GM_Q_AB_nCi_2_NS4_X_SRW_h_MAS	N2_B_AB_nCi_2_MP6_T_LGP[4-6]_a-r_MAS	AM_B_AB_Ci(2.0;-2.0)_2_MP8_n_T_LGP[1;2;6]_a-psa_MAS
K_F_AB_nCi_2_SS5_T_SRW_r_MAS	RA_B_BB_Ci(2.0;-2.0)_2_SS7_X_LGP[1;2;6]_p-h_MAS	S_B_AB_nCi_2_SS7_P_LGP[1;2;6]_a-psa_MAS
RA_B_AB_nCi_2_NS4_T_NSRW_v-h_MAS	VC_Q_AB_nCi_2_SS4_P_LGP[4-6]_a_MAS	HM_Q_AB_Ci(2.0;-2.0)_2_SS6_T_LGP[2-4]_e_MAS
RA_B_AB_nCi_2_MP2_T_KA_a-e_MAS	S_B_AB_nCi_2_SS7_T_LGP[2-4]_e-p_MAS	N3_B_AB_nCi_2_NS4_P_LGP[4-6]_v-c_MAS
K_F_AB_nCi_2_SS6_T_KA_e_MAS	P3_B_AB_nCi_2_SS7_A_LGP[1;2;6]_e-p_MAS	SD_Q_AB_nCi_2_NS6_P_LGP[2-4]_h_MAS
I50_Q_AB_nCi_2_NS0_P_KA_h_MAS	HM_B_AB_nCi_2_SS2_X_LGP[1;2;6]_v-h_MAS	Q1_B_AB_Ci(2.0;-2.0)_2_MP8_n_X_LGP[1-3]_a-psa_MAS
RA_B_BB_Ci(2.0;-2.0)_2_MP4_P_SRW_a-e_MAS	SD_B_AB_nCi_2_NS2_A_LGP[1;2;6]_a-h_MAS	SD_B_AB_nCi_2_NS3_X_LGP[1;2;6]_e-h_MAS
N3_F_AB_nCi_2_SS1_P_NSRW_h_MAS	I50_Q_AB_nCi_2_SS5_H_n_P_LGP[4-6]_s_MAS	K_Q_AB_Ci(2.0;-2.0)_2_SS4_H_n_P_LGP[4-6]_a_MAS
HM_B_AB_nCi_2_MP7_X_SRW_e-s_MAS	Q1_F_AB_nCi_2_MP4_T_LGP[4-6]_m_MAS	SD_Q_AB_nCi_2_NS8_T_LGP[2-4]_a_MAS
N3_Q_AB_nCi_2_MP4_X_NSRW_p_MAS	VC_B_AB_nCi_2_SS5_T_LGP[4-6]_v-m_MAS	I50_Q_AB_nCi_2_MP8_n_X_LGP[2-4]_s_MAS
AM_B_AB_Ci(2.0;-2.0)_2_MP2_T_NSRW_a-psa_MAS	Q3_Q_AB_nCi_2_MP8_H_n_P_LGP[4-6]_a_MAS	N3_B_AB_nCi_2_SS6_C_LGP[4-6]_c-h_MAS
N2_Q_AB_nCi_2_MP5_T_NSRW_s_MAS	V_B_AB_nCi_2_SS2_P_LGP[1;2;6]_a-v_MAS	Q3_B_AB_nCi_2_SS3_n_X_LGP[2-4]_a-s_MAS
Q1_B_AB_nCi_2_SS7_T_NSRW_v-m_MAS	VC_B_AB_nCi_2_SS3_T_LGP[1;2;6]_r-s_MAS	N2_Q_AB_nCi_2_MP4_H_n_T_LGP[1]_a_MAS
K_B_AB_nCi_2_NS7_X_SRW_e-s_MAS	HM_Q_AB_Ci(2.0;-2.0)_2_SS4_H_A_LGP[1]_s_MAS	K_Q_AB_nCi_2_SS5_T_LGP[2-4]_psa_MAS
N3_B_AB_nCi_2_SS5_T_KA_r-h_MAS	I50_Q_AB_nCi_2_NS8_n_P_LGP[1;2;6]_h_MAS	P2_Q_AB_Ci(2.0;-2.0)_2_SS8_H_A_LGP[1]_s_MAS
P2_F_AB_nCi_2_NS5_X_NSRW_a_MAS	VC_F_AB_nCi_2_SS7_T_LGP[4-6]_r_MAS	S_B_AB_Ci(2.0;-2.0)_2_SS5_H_n_X_LGP[4-6]_a-v_MAS
N3_B_AB_nCi_2_NS1_T_KA_v-h_MAS	MX_Q_AB_nCi_2_NS1_n_P_LGP[1-3]_h_MAS	SD_B_AB_nCi_2_NS1_P_LGP[1;2;6]_v-h_MAS
GM_Q_AB_Ci(2.0;-2.0)_2_SS5_T_SRW_a_MAS	Q2_B_AB_Ci(2.0;-2.0)_2_MP8_n_P_LGP[4-6]_r-c_MAS	RA_Q_AB_Ci(2.0;-2.0)_2_MP6_n_P_LGP[2-4]_s_MAS
N2_F_AB_nCi_2_MP4_n_X_KA_a_MAS	I50_B_AB_nCi_2_NS3_T_LGP[1]_e-s_MAS	VC_B_AB_nCi_2_SS8_C_LGP[4-6]_c-h_MAS
S_F_AB_nCi_2_NS2_T_NSRW_h_MAS	AM_B_BB_Ci(2.0;-2.0)_2_NS2_H_A_LGP[2-4]_m-e_MAS	Q2_F_AB_Ci(2.0;-2.0)_2_MP8_n_X_LGP[1;2;6]_a_MAS
HM_B_BB_Ci(2.0;-2.0)_2_MP2_A_KA_r-e_MAS	I50_B_AB_nCi_2_NS4_T_LGP[1]_m-p_MAS	N2_B_AB_nCi_2_SS6_A_LGP[1]_psa-c_MAS
AM_F_AB_nCi_2_MP4_C_KA_c_MAS	P2_B_BB_Ci(2.0;-2.0)_2_SS7_H_T_LGP[4-6]_v-s_MAS	HM_B_AB_nCi_2_SS3_C_LGP[1-3]_psa-h_MAS
N3_B_AB_nCi_2_NS4_X_SRW_m-h_MAS	RA_B_AB_nCi_2_MP8_C_LGP[4-6]_e-p_MAS	HM_B_AB_nCi_2_SS7_C_LGP[1;2;6]_psa-h_MAS
RA_B_AB_nCi_2_SS8_X_NSRW_r-c_MAS	P2_B_BB_Ci(2.0;-2.0)_2_SS2_H_T_LGP[2-4]_p-h_MAS	HM_B_AB_Ci(2.0;-2.0)_2_SS8_H_A_LGP[1]_r-v_MAS
V_Q_AB_nCi_2_NS8_P_NSRW_a_MAS	V_B_AB_nCi_2_SS5_P_LGP[1;2;6]_a-e_MAS	I50_F_AB_nCi_2_NS5_n_X_LGP[2-4]_m_MAS
SD_B_AB_nCi_2_MP1_X_NSRW_a-e_MAS	I50_F_AB_nCi_2_NS8_n_X_LGP[1-3]_s_MAS	RA_F_AB_Ci(2.0;-2.0)_2_MP2_n_X_LGP[1-3]_a_MAS
I50_B_AB_nCi_2_MP7_X_KA_e-s_MAS	I50_F_AB_nCi_2_SS7_T_LGP[1;2;6]_s_MAS	N3_B_AB_nCi_2_MP4_A_LGP[4-6]_a-c_MAS
RA_F_AB_nCi_2_MP1_T_NSRW_p_MAS	N1_B_AB_nCi_2_NS5_A_LGP[4-6]_a-c_MAS	P2_B_AB_Ci(2.0;-2.0)_2_SS3_T_LGP[2-4]_h-s_MAS
SD_F_AB_Ci(2.0;-2.0)_2_SS6_H_n_X_NSRW_a_MAS	HM_B_AB_nCi_2_NS7_C_LGP[1-3]_p-h_MAS	RA_B_AB_Ci(2.0;-2.0)_2_SS1_n_X_LGP[1-3]_a-s_MAS
S_B_AB_nCi_2_NS8_T_NSRW_m-s_MAS	V_B_AB_nCi_2_SS5_A_LGP[1;2;6]_psa-c_MAS	GM_Q_AB_nCi_2_NS6_C_LGP[1;2;6]_s_MAS
RA_B_AB_nCi_2_SS8_P_KA_r-h_MAS	I50_B_AB_nCi_2_NS1_T_LGP[1]_m-s_MAS	K_B_AB_nCi_2_NS7_P_LGP[1;2;6]_a-p_MAS

MX_B_AB_nCi_2_SS2_X_NSRW_v-h_MAS	MX_F_AB_nCi_2_SS6_T_LGP[1]_h_MAS	P3_F_BB_Ci(2.0;-2.0)_2_SS6_H_A_LGP[1;2;6]_r_MAS
Q2_B_AB_nCi_2_MP8_A_NSRW_a-m_MAS	VC_B_AB_nCi_2_NS4_T_LGP[4-6]_r-v_MAS	N1_Q_AB_nCi_2_NS3_P_LGP[1]_a_MAS
SD_B_AB_nCi_2_SS3_X_KA_v-h_MAS	N3_B_AB_nCi_2_SS8_A_LGP[4-6]_r-e_MAS	Q1_B_BB_Ci(2.0;-2.0)_2_MP2_X_LGP[1;2;6]_a-psa_MAS
K_B_AB_nCi_2_NS2_T_SRW_a-psa_MAS	N2_B_AB_nCi_2_MP8_H_n_T_LGP[1]_a-p_MAS	Q2_F_AB_nCi_2_MP7_A_LGP[4-6]_c_MAS
Q2_Q_AB_nCi_2_SS2_T_KA_m_MAS	S_F_AB_nCi_2_SS3_T_LGP[2-4]_a_MAS	VC_F_AB_nCi_2_NS4_A_LGP[4-6]_a_MAS
I50_B_AB_nCi_2_MP7_X_KA_h-s_MAS	K_B_AB_nCi_2_SS8_T_LGP[2-4]_r-e_MAS	MN_F_AB_nCi_2_MP1_H_n_P_LGP[1-3]_a_MAS
N1_B_AB_nCi_2_SS6_H_T_KA_c-h_MAS	HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS	I50_Q_AB_nCi_2_SS1_T_LGP[1;2;6]_c_MAS
I50_Q_AB_nCi_2_NS6_T_NSRW_a_MAS	P3_F_BB_Ci(2.0;-2.0)_2_MP8_H_A_LGP[1;2;6]_p_MAS	P3_Q_AB_nCi_2_NS2_C_LGP[1-3]_r_MAS
K_Q_AB_nCi_2_NS3_A_KA_psa_MAS	HM_Q_AB_nCi_2_MP2_X_LGP[1;2;6]_p_MAS	Q1_B_BB_Ci(2.0;-2.0)_2_MP1_X_LGP[1;2;6]_e-c_MAS
Q3_B_AB_nCi_2_SS5_T_SRW_r-m_MAS	I50_B_AB_nCi_2_SS1_T_LGP[1]_p-h_MAS	Q3_B_AB_nCi_2_SS8_H_n_X_LGP[2-4]_a-h_MAS
I50_F_AB_nCi_2_MP5_X_KA_psa_MAS	HM_B_AB_nCi_2_SS5_X_LGP[1;2;6]_r-p_MAS	VC_B_AB_nCi_2_NS5_X_LGP[4-6]_a-psa_MAS
N2_B_AB_nCi_2_SS1_H_T_KA_v-c_MAS	HM_B_AB_nCi_2_NS2_X_LGP[1;2;6]_r-p_MAS	Q1_B_AB_nCi_2_SS7_T_LGP[1;2;6]_a-c_MAS
I50_B_AB_nCi_2_SS6_T_SRW_r-p_MAS	HM_B_AB_Ci(2.0;-2.0)_2_SS8_H_C_LGP[2-4]_r-p_MAS	Q1_B_AB_nCi_2_SS8_X_LGP[4-6]_e-c_MAS
I50_B_AB_nCi_2_SS8_T_SRW_r-c_MAS	I50_B_AB_nCi_2_SS4_T_LGP[1]_a-s_MAS	Q1_B_BB_Ci(2.0;-2.0)_2_SS2_X_LGP[1;2;6]_r-c_MAS
SD_B_AB_nCi_2_NS1_P_NSRW_v-h_MAS	N1_B_AB_nCi_2_MP7_C_LGP[4-6]_psa-v_MAS	HM_F_AB_Ci(2.0;-2.0)_2_SS4_T_LGP[2-4]_p_MAS
AM_Q_AB_Ci(2.0;-2.0)_2_NS2_A_SRW_r_MAS	K_B_BB_Ci(2.0;-2.0)_2_NS8_H_X_LGP[4-6]_a-psa_MAS	HM_Q_AB_Ci(2.0;-2.0)_2_SS2_H_C_LGP[2-4]_v_MAS
S_B_AB_nCi_2_SS7_T_SRW_m-h_MAS	VC_B_AB_Ci(2.0;-2.0)_2_SS8_n_P_LGP[4-6]_psa-p_MAS	RA_B_AB_Ci(2.0;-2.0)_2_SS1_n_X_LGP[1;2;6]_p-h_MAS
HM_Q_AB_nCi_2_MP5_C_SRW_c_MAS	I50_B_AB_nCi_2_NS7_H_n_P_LGP[4-6]_a-c_MAS	S_B_AB_nCi_2_NS8_X_LGP[2-4]_r-v_MAS
N1_B_AB_nCi_2_SS4_H_T_KA_c-s_MAS	N2_Q_AB_nCi_2_NS2_C_LGP[1]_s_MAS	Q2_Q_AB_nCi_2_MP4_A_LGP[4-6]_p_MAS
Q1_B_AB_nCi_2_MP4_A_KA_a-r_MAS	HM_B_AB_nCi_2_NS2_X_LGP[1;2;6]_h-s_MAS	Q1_B_AB_Ci(2.0;-2.0)_2_SS8_n_X_LGP[1;2;6]_c-h_MAS
RA_B_AB_nCi_2_SS6_T_NSRW_v-h_MAS	I50_B_AB_nCi_2_NS5_H_n_P_LGP[1-3]_h-s_MAS	Q1_B_AB_Ci(2.0;-2.0)_2_MP8_n_X_LGP[1;2;6]_c-s_MAS
K_B_AB_nCi_2_NS6_T_SRW_a-h_MAS	AM_B_AB_Ci(2.0;-2.0)_2_SS4_T_LGP[2-4]_r-v_MAS	MN_Q_AB_nCi_2_MP8_H_n_P_LGP[4-6]_a_MAS
N1_B_AB_nCi_2_SS1_T_NSRW_c-h_MAS	Q3_B_AB_nCi_2_SS2_n_X_LGP[1-3]_p-s_MAS	Q1_F_AB_nCi_2_MP8_A_LGP[4-6]_a_MAS
K_Q_AB_nCi_2_SS1_P_NSRW_a_MAS	RA_Q_AB_nCi_2_NS6_C_LGP[4-6]_h_MAS	HM_Q_AB_Ci(2.0;-2.0)_2_NS6_H_A_LGP[1]_p_MAS
RA_B_AB_nCi_2_NS4_C_NSRW_h-s_MAS	VC_Q_AB_nCi_2_SS4_n_C_LGP[4-6]_a_MAS	Q3_B_AB_nCi_2_SS6_n_P_LGP[1;2;6]_r-m_MAS
Q1_F_AB_Ci(2.0;-2.0)_2_MP4_n_X_NSRW_a_MAS	K_F_AB_nCi_2_NS3_P_LGP[1;2;6]_a_MAS	N3_B_AB_nCi_2_MP8_A_LGP[4-6]_a-e_MAS
HM_B_AB_Ci(2.0;-2.0)_2_SS1_H_C_NSRW_v-s_MAS	VC_B_AB_nCi_2_SS4_A_LGP[4-6]_r-s_MAS	Q1_F_AB_nCi_2_MP7_A_LGP[4-6]_a_MAS
SD_B_AB_nCi_2_NS7_A_NSRW_e-c_MAS	MX_Q_AB_nCi_2_MP5_H_n_P_LGP[4-6]_c_MAS	I50_B_AB_nCi_2_SS2_T_LGP[1]_a-s_MAS
Q1_F_AB_Ci(2.0;-2.0)_2_MP7_n_X_LGP[1;2;6]_c_MAS	N2_B_AB_nCi_2_SS4_H_n_T_LGP[1]_r-c_MAS	N1_B_AB_nCi_2_MP8_H_n_T_LGP[4-6]_r-c_MAS
AM_B_AB_nCi_2_MP2_X_LGP[1;2;6]_a-c_MAS	P2_Q_AB_nCi_2_SS1_X_LGP[1;2;6]_a_MAS	HM_Q_AB_nCi_2_MP8_X_LGP[1;2;6]_v_MAS

Table S3. Models obtained for the different machine learning techniques

Reduction reaction of H₂O₂						
Technique	Model	Size	R ²	MAE	Q _{cv} ² (5-fold)	MAE
MLR	M1	6.000	0.911	0.070	0.848	0.089
	M2	6.000	0.907	0.070	0.858	0.083
	M3	5.000	0.886	0.078	0.835	0.092
	M4	5.000	0.885	0.075	0.843	0.091
	M5	4.000	0.823	0.087	0.771	0.103
	M6	6.000	0.886	0.075	0.794	0.105
SMOreg	M7	5.000	0.875	0.082	0.783	0.111
	M8	6.000	0.889	0.069	0.862	0.081
	M9	6.000	0.798	0.092	0.729	0.115
IBK	M10 (knn=4)	5.000	0.866	0.089	0.799	0.115
	M11 (knn=3)	4.000	0.837	0.088	0.729	0.121
RF	M12	6.000	0.980	0.043	0.735	0.120
	M13	6.000	0.973	0.044	0.630	0.131
	M14	6.000	0.979	0.047	0.683	0.137
Reduction reaction of Cum-OOH						
MLR	M15	5.000	0.925	0.110	0.891	0.140
	M16	4.000	0.900	0.127	0.814	0.177
	M17	4.000	0.892	0.124	0.856	0.146
Reduction reaction of t-BuOOH						
MLR	M18	4.000	0.937	0.052	0.856	0.077
	M19	4.000	0.938	0.051	0.874	0.072
	M20	3.000	0.904	0.060	0.863	0.072
	M21	3.000	0.888	0.068	0.860	0.078

Table S4. Attributes of each model M1-M21

Model	Attributes		
M1	K_B_AB_nCi_2_NS3_T_NSRW_p-s_MAS	Q2_B_AB_nCi_2_MP8_A_NSRW_a-m_MAS	Q1_F_AB_Ci(2.0;-2.0)_2_MP4_n_X_NSRW_a_MAS
	VC_B_BB_Ci(2.0;-2.0)_2_SS3_X_LGP[1;2;6]_r-h_MAS	HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS	GM_F_BB_Ci(2.0;-2.0)_2_NS7_H_A_LGP[1;2;6]_psa_MAS
M2	Q2_B_AB_nCi_2_MP8_A_NSRW_a-m_MAS x 10 ⁻³ (H_2O_2)	Q1_F_AB_Ci(2.0;-2.0)_2_MP4_n_X_NSRW_a_MAS x 10 ⁻³ (H_2O_2)	VC_B_BB_Ci(2.0;-2.0)_2_SS3_X_LGP[1;2;6]_r-h_MAS (H_2O_2)
	HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS ($D^{H_2O_2}$)	VC_Q_AB_nCi_2_SS4_n_C_LGP[4-6]_a_MAS ($E^{H_2O_2}$)	GM_F_BB_Ci(2.0;-2.0)_2_NS7_H_A_LGP[1;2;6]_psa_MAS ($B^{H_2O_2}$)
M3	Q1_F_AB_Ci(2.0;-2.0)_2_MP4_n_X_NSRW_a_MAS	I50_B_AB_nCi_2_NS3_T_LGP[1]_e-s_MAS	HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS
	GM_F_BB_Ci(2.0;-2.0)_2_NS7_H_A_LGP[1;2;6]_psa_MAS	SD_Q_AB_nCi_2_NS8_T_LGP[2-4]_a_MAS	
M4	Q2_B_AB_nCi_2_MP8_A_NSRW_a-m_MAS	Q1_F_AB_Ci(2.0;-2.0)_2_MP4_n_X_NSRW_a_MAS	I50_B_AB_nCi_2_NS3_T_LGP[1]_e-s_MAS
	GM_F_BB_Ci(2.0;-2.0)_2_NS7_H_A_LGP[1;2;6]_psa_MAS	HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS	
M5	K_B_AB_nCi_2_NS4_T_LGP[2-4]_e-p_MAS	HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS	P3_F_BB_Ci(2.0;-2.0)_2_MP8_H_A_LGP[1;2;6]_p_MAS
	GM_F_BB_Ci(2.0;-2.0)_2_NS7_H_A_LGP[1;2;6]_psa_MAS		
M6	Q1_F_AB_Ci(2.0;-2.0)_2_MP4_n_X_NSRW_a_MAS	HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS	K_Q_AB_Ci(2.0;-2.0)_2_SS4_H_n_P_LGP[4-6]_a_MAS
	Q3_B_AB_nCi_2_SS3_n_X_LGP[2-4]_a-s_MAS	K_Q_AB_nCi_2_SS5_T_LGP[2-4]_psa_MAS	P2_B_AB_Ci(2.0;-2.0)_2_SS3_T_LGP[2-4]_h-s_MAS
M7	S_B_AB_nCi_2_NS4_X_KA_r-h_MAS	Q1_F_AB_Ci(2.0;-2.0)_2_MP4_n_X_NSRW_a_MAS	HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS
	K_Q_AB_Ci(2.0;-2.0)_2_SS4_H_n_P_LGP[4-6]_a_MAS	P2_B_AB_Ci(2.0;-2.0)_2_SS3_T_LGP[2-4]_h-s_MAS	
M8	Q1_F_AB_Ci(2.0;-2.0)_2_MP4_n_X_NSRW_a_MAS	I50_B_AB_nCi_2_NS3_T_LGP[1]_e-s_MAS	HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS
	GM_F_BB_Ci(2.0;-2.0)_2_NS7_H_A_LGP[1;2;6]_psa_MAS	SD_Q_AB_nCi_2_NS8_T_LGP[2-4]_a_MAS	P2_B_AB_Ci(2.0;-2.0)_2_SS3_T_LGP[2-4]_h-s_MAS
M9	HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS	P3_F_BB_Ci(2.0;-2.0)_2_MP8_H_A_LGP[1;2;6]_p_MAS	VC_B_AB_nCi_2_SS4_A_LGP[4-6]_r-s_MAS
	I50_B_AB_nCi_2_NS4_H_n_P_LGP[4-6]_h-s_MAS	GM_F_BB_Ci(2.0;-2.0)_2_NS7_H_A_LGP[1;2;6]_psa_MAS	Q3_B_AB_nCi_2_SS3_n_X_LGP[2-4]_a-s_MAS
M10	V_Q_AB_nCi_2_NS4_P_NSRW_m_MAS	HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS	K_Q_AB_nCi_2_NS5_T_LGP[2-4]_h_MAS
	GM_F_BB_Ci(2.0;-2.0)_2_NS7_H_A_LGP[1;2;6]_psa_MAS	VC_B_AB_nCi_2_SS8_C_LGP[4-6]_c-h_MAS	
M11	HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS	K_Q_AB_nCi_2_NS5_T_LGP[2-4]_h_MAS	GM_F_BB_Ci(2.0;-2.0)_2_NS7_H_A_LGP[1;2;6]_psa_MAS
	I50_B_AB_nCi_2_SS8_T_SRW_r-c_MAS		
M12	S_B_AB_nCi_2_SS1_X_NSRW_v-h_MAS	N3_Q_AB_nCi_2_MP2_A_SRW_r_MAS	K_B_AB_Ci(2.0;-2.0)_2_SS4_H_n_X_LGP[2-4]_r-h_MAS
	AM_B_AB_nCi_2_SS6_X_LGP[1-3]_v-c_MAS	SD_B_AB_nCi_2_NS2_A_LGP[1;2;6]_a-h_MAS	P2_B_BB_Ci(2.0;-2.0)_2_SS2_H_T_LGP[2-4]_p-h_MAS
M13	GM_Q_AB_Ci(2.0;-2.0)_2_MP1_n_T_LGP[1;2;6]_p_MAS	S_B_AB_nCi_2_NS4_X_KA_r-h_MAS	P2_F_AB_nCi_2_NS5_X_NSRW_a_MAS
	Q2_Q_AB_Ci(2.0;-2.0)_2_SS4_n_P_LGP[4-6]_a_MAS	P2_B_AB_Ci(2.0;-2.0)_2_SS3_T_LGP[2-4]_h-s_MAS	Q1_B_AB_nCi_2_MP4_A_KA_a-r_MAS

M14	K_B_AB_Ci(2.0;-2.0)_2_SS4_H_n_X_LGP[2-4]_r-h_MAS P2_B_BB_Ci(2.0;-2.0)_2_SS2_H_T_LGP[2-4]_p-h_MAS	N3_B_AB_nCi_2_NS5_C_LGP[4-6]_e-h_MAS I50_B_AB_nCi_2_NS1_T_LGP[1]_m-s_MAS	SD_B_AB_nCi_2_NS2_A_LGP[1;2;6]_a-h_MAS HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS
M15	HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS ($A^{Cum-OOH}$) S_B_AB_nCi_2_NS7_T_LGP[2-4]_m-p_MAS ($D^{Cum-OOH}$)	Q2_Q_AB_Ci(2.0;-2.0)_2_MP6_n_X_LGP[1-3]_h_MAS x 10 ⁻³ ($B^{Cum-OOH}$) S_B_AB_nCi_2_SS7_P_LGP[1;2;6]_a-psa_MAS ($E^{Cum-OOH}$)	K_B_AB_nCi_2_NS4_C_LGP[1;2;6]_m-s_MAS ($C^{Cum-OOH}$)
M16	Q2_Q_AB_Ci(2.0;-2.0)_2_MP6_n_X_LGP[1-3]_h_MAS P2_B_AB_Ci(2.0;-2.0)_2_SS3_T_LGP[2-4]_h-s_MAS	K_B_AB_nCi_2_NS4_C_LGP[1;2;6]_m-s_MAS	S_B_AB_nCi_2_SS7_P_LGP[1;2;6]_a-psa_MAS
M17	HM_B_BB_Ci(2.0;-2.0)_2_NS4_H_T_LGP[4-6]_p-h_MAS S_B_AB_nCi_2_NS7_T_LGP[2-4]_m-p_MAS	Q2_Q_AB_Ci(2.0;-2.0)_2_MP6_n_X_LGP[1-3]_h_MAS	K_B_AB_nCi_2_NS4_C_LGP[1;2;6]_m-s_MAS
M18	S_B_AB_nCi_2_SS7_T_SRW_m-h_MAS N2_B_AB_nCi_2_SS6_A_LGP[1]_psa-c_MAS	RA_B_AB_nCi_2_SS6_T_NSRW_v-h_MAS	AM_Q_AB_Ci(2.0;-2.0)_2_MP1_A_KA_a_MAS
M19	S_B_AB_nCi_2_SS7_T_SRW_m-h_MAS ($A^{t-BuOOH}$) N2_B_AB_nCi_2_SS6_A_LGP[1]_psa-c_MAS ($D^{t-BuOOH}$)	N1_B_AB_nCi_2_SS4_H_T_KA_c-s_MAS ($B^{t-BuOOH}$)	AM_Q_AB_Ci(2.0;-2.0)_2_MP1_A_KA_a_MAS x 10 ⁻³ ($C^{t-BuOOH}$)
M20	VC_B_AB_Ci(2.0;-2.0)_2_NS7_n_X_NSRW_v-s_MAS	HM_B_AB_Ci(2.0;-2.0)_2_SS1_T_SRW_p-s_MAS	N2_B_AB_nCi_2_SS6_A_LGP[1]_psa-c_MAS
M21	N1_B_AB_nCi_2_SS4_H_T_KA_c-s_MAS	AM_Q_AB_Ci(2.0;-2.0)_2_MP1_A_KA_a_MAS	N2_B_AB_nCi_2_SS6_A_LGP[1]_psa-c_MAS

Table S5. Evaluation of the applicability domain of the external set (from reference 41) in model M2 and prediction of log(ν_0)

			Applicability Domain				
Name	SMILES	Predicted ν_0 by M2	PCA-range	Euclidean distance	City-block distance	Probability density	Consensus domain
6	c1(ccc(cc1)S(=O)(=O)N)[Se]CCO	3.604	false	false	false	false	false
7	c1(ccc(cc1)S(=O)(=O)N)[Se]CCc1ccccc1	5.267	false	false	false	false	false
8	c1(ccc(cc1)S(=O)(=O)N)[Se]C	5.673	false	false	false	false	false
9	c1(ccc(cc1)S(=O)(=O)N)[Se]C(C)C	6.103	false	false	false	false	false
10	c1(ccc(cc1)S(=O)(=O)N)[Se]CC#C	5.737	false	false	false	false	false
11	c1(ccc(cc1)S(=O)(=O)N)[Se]CCOC	3.201	false	false	false	false	false
12	c1(ccc(cc1)S(=O)(=O)N)[Se]CC=C	5.844	false	false	false	false	false
13	c1(ccc(cc1)S(=O)(=O)N)[Se]C1CCCCC1	4.928	false	false	false	false	false
14	c1(ccc(cc1)S(=O)(=O)N)C[Se]c1ccccc1	3.828	false	false	false	false	false
15	c1(ccc(cc1)S(=O)(=O)N)C(=[Se])N	5.087	false	false	false	false	false
16	c1(ccc(cc1)S(=O)(=O)N)c1cnc([se]1)c1ccc(cc1)F	2.563	true	true	true	true	true

17	<chem>c1(ccc(cc1)S(=O)(=O)N)c1cnc([se]1)C[Se]c1cccc1</chem>	2.21	false	false	false	false	false
18	<chem>c1(ccc(cc1)S(=O)(=O)N)c1cnc([se]1)C[Te]c1ccc(cc1)C</chem>	2.26	false	true	false	false	false
19	<chem>c1(cc(c(cc1)N)S(=O)(=O)N)[Se]C#N</chem>	2.865	false	false	false	false	false
20	<chem>c1(cc(c(cc1)N)S(=O)(=O)N)[Se][Se]c1cc(c(cc1)N)S(=O)(=O)N</chem>	3.426	false	false	false	false	false
21	<chem>c1(cc(ccc1)S(=O)(=O)N)[Se][Se]c1cc(ccc1)S(=O)(=O)N</chem>	4.275	false	false	false	false	false
22	<chem>c1(c(cccc1)S(=O)(=O)N)[Se][Se]c1ccccc1S(=O)(=O)N</chem>	3.718	false	false	false	false	false
23	<chem>c1cc(ccc1)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	1.736	false	false	false	false	false
24	<chem>c1cc(ccc1F)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	1.908	false	false	false	false	false
25	<chem>c1cc(c(cc1)F)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	1.895	false	false	false	false	false
26	<chem>c1cc(cc(c1)C(F)(F)F)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	1.492	false	false	false	false	false
27	<chem>c1cc(c(cc1)Cl)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	1.608	false	false	false	false	false
28	<chem>c1cc(ccc1)NC(=[Se])Nc1cccc(c1)S(=O)(=O)N</chem>	1.697	false	false	false	false	false
29	<chem>c1cc(ccc1F)NC(=[Se])Nc1cccc(c1)S(=O)(=O)N</chem>	1.85	false	false	false	false	false
30	<chem>c1cc(ccc1)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	1.906	true	true	true	true	true
31	<chem>c1cc(c(cc1)OC)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	2.018	true	true	true	true	true
32	<chem>c1cc(c(cc1)F)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	2.087	false	true	true	true	true
33	<chem>c1cc(cc(c1)C(F)(F)F)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	1.808	false	true	true	true	true
34	<chem>c1cc(c(cc1)Cl)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	1.996	false	true	false	false	false
35	<chem>c1cc(ccc1Br)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	2.081	false	true	true	true	true
36	<chem>c1cc(ccc1I)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	2.046	false	true	true	false	true
37	<chem>N(C(=[Se])NCCc1ccc(cc1)S(=O)(=O)N)c1cc2cccc2cc1</chem>	1.655	true	true	true	true	true
38	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1cccc1</chem>	1.693	false	false	false	false	false
39	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1cc2c(cc1)OCO2</chem>	-1.329	false	false	false	false	false
40	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1cc(cc(c1)Br)OC(F)(F)F</chem>	1.808	false	false	false	false	false
41	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1cccc1OC(=O)C</chem>	2.922	false	false	false	false	false
42	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1ccc(cc1)CCCCCCC</chem>	1.643	true	true	true	true	true
43	<chem>N(C(=[Se])Nc1cccc(c1)S(=O)(=O)N)C(=O)c1cccc1</chem>	1.579	false	false	false	false	false
44	<chem>N(C(=[Se])Nc1cccc(c1)S(=O)(=O)N)C(=O)c1cc(ccc1)C</chem>	2.087	false	true	false	false	false
45	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cccc1</chem>	1.738	false	false	false	false	false
46	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cccc(c1)C</chem>	2.167	false	true	true	true	true
47	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1ccc(c(c1)C)Br</chem>	2.238	false	true	false	false	false

48	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cccc1OC(=O)C</chem>	3.009	false	false	false	false	false
49	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cc(cc(c1)OC(F)(F)F)Br</chem>	1.995	false	false	false	false	false
50	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cc2c(cc1)OCO2</chem>	13.514	false	false	false	false	false
51	<chem>N(C(=[Se])NCCc1ccc(cc1)S(=O)(=O)N)C(=O)c1cccc1</chem>	1.88	true	true	true	true	true
52	<chem>N(C(=[Se])NCCc1ccc(cc1)S(=O)(=O)N)C(=O)c1cccc(c1)C</chem>	1.713	true	true	true	true	true
53	<chem>N(C(=[Se])NCCc1ccc(cc1)S(=O)(=O)N)C(=O)c1cc2c(cc1)OCO2</chem>	2.426	false	true	false	true	true
54	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]C#N</chem>	3.04	false	false	false	false	false
55	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se][Se]c1ccc(cc1)S(=O)(=O)N</chem>	5.925	false	false	false	false	false

Table S6. Evaluation of the applicability domain of the external set (from reference 41) in model M15 and prediction of $\log(u_0)$

Name	SMILES	Predicted u_0 by M15	Applicability Domain				
			PCA-range	Euclidean distance	City-block distance	Probability density	Consensus domain
6	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]CCO</chem>	4.443	false	false	false	false	false
7	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]CCc1cccc1</chem>	1.555	false	false	false	false	false
8	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]C</chem>	4.761	false	false	false	false	false
9	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]C(C)C</chem>	4.895	false	false	false	false	false
10	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]CC#C</chem>	2.463	false	false	false	false	false
11	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]CCOC</chem>	4.32	false	false	false	false	false
12	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]CC=C</chem>	3.493	false	false	false	false	false
13	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]C1CCCCC1</chem>	2.79	false	false	false	false	false
14	<chem>c1(ccc(cc1)S(=O)(=O)N)C[Se]c1cccc1</chem>	3.932	false	false	false	false	false
15	<chem>c1(ccc(cc1)S(=O)(=O)N)C(=[Se])N</chem>	4.797	false	false	false	false	false
16	<chem>c1(ccc(cc1)S(=O)(=O)N)c1cnc([se]1)c1ccc(cc1)F</chem>	1.485	false	true	true	false	true
17	<chem>c1(ccc(cc1)S(=O)(=O)N)c1cnc([se]1)C[Se]c1cccc1</chem>	2.707	false	false	false	false	false
18	<chem>c1(ccc(cc1)S(=O)(=O)N)c1cnc([se]1)C[Te]c1ccc(cc1)C</chem>	2.123	false	false	true	false	false
19	<chem>c1(cc(c(cc1)N)S(=O)(=O)N)[Se]C#N</chem>	5.191	false	false	false	false	false
20	<chem>c1(cc(c(cc1)N)S(=O)(=O)N)[Se][Se]c1cc(c(cc1)N)S(=O)(=O)N</chem>	1.371	false	false	false	false	false
21	<chem>c1(cc(ccc1)S(=O)(=O)N)[Se][Se]c1cc(ccc1)S(=O)(=O)N</chem>	1.761	false	false	false	false	false
22	<chem>c1(c(cccc1)S(=O)(=O)N)[Se][Se]c1cccc1S(=O)(=O)N</chem>	1.374	false	false	false	false	false
23	<chem>c1cc(ccc1)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	3.426	false	false	false	false	false

24	<chem>c1cc(ccc1F)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	3.468	false	false	false	false	false
25	<chem>c1cc(c(cc1)F)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	3.406	false	false	false	false	false
26	<chem>c1cc(cc(c1)C(F)(F)F)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	2.686	false	false	false	false	false
27	<chem>c1cc(c(cc1)Cl)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	3.33	false	false	false	false	false
28	<chem>c1cc(ccc1)NC(=[Se])Nc1cccc(c1)S(=O)(=O)N</chem>	3.272	false	false	false	false	false
29	<chem>c1cc(ccc1F)NC(=[Se])Nc1cccc(c1)S(=O)(=O)N</chem>	3.364	false	false	false	false	false
30	<chem>c1cc(ccc1)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	1.803	false	false	false	false	false
31	<chem>c1cc(c(cc1)OC)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	1.652	false	false	false	false	false
32	<chem>c1cc(c(cc1)F)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	1.745	false	false	false	false	false
33	<chem>c1cc(cc(c1)C(F)(F)F)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	1.451	false	false	false	false	false
34	<chem>c1cc(c(cc1)Cl)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	1.746	false	false	false	false	false
35	<chem>c1cc(ccc1Br)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	1.724	false	false	false	false	false
36	<chem>c1cc(ccc1I)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	1.343	false	false	false	false	false
37	<chem>N(C(=[Se])NCCc1ccc(cc1)S(=O)(=O)N)c1cc2cccc2cc1</chem>	1.204	false	false	false	false	false
38	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1cccc1</chem>	3.395	false	false	false	false	false
39	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1cc2c(cc1)OCO2</chem>	3.18	false	false	false	false	false
40	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1cc(cc(c1)Br)OC(F)(F)F</chem>	4.022	false	false	false	false	false
41	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1cccc1OC(=O)C</chem>	3.494	false	false	false	false	false
42	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1ccc(cc1)CCCCCCC</chem>	2.458	false	false	false	false	false
43	<chem>N(C(=[Se])Nc1cccc(c1)S(=O)(=O)N)C(=O)c1cccc1</chem>	3.333	false	false	false	false	false
44	<chem>N(C(=[Se])Nc1cccc(c1)S(=O)(=O)N)C(=O)c1cc(ccc1)C</chem>	2.955	false	false	false	false	false
45	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cccc1</chem>	3.299	false	false	false	false	false
46	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cccc(c1)C</chem>	3.018	false	false	false	false	false
47	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1ccc(c(c1)C)Br</chem>	2.894	false	false	false	false	false
48	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cccc1OC(=O)C</chem>	3.431	false	false	false	false	false
49	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cc(cc(c1)OC(F)(F)F)Br</chem>	3.84	false	false	false	false	false
50	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cc2c(cc1)OCO2</chem>	3.09	false	false	false	false	false
51	<chem>N(C(=[Se])NCCc1ccc(cc1)S(=O)(=O)N)C(=O)c1cccc1</chem>	2.644	false	false	false	false	false
52	<chem>N(C(=[Se])NCCc1ccc(cc1)S(=O)(=O)N)C(=O)c1cccc(c1)C</chem>	2.499	false	false	false	false	false
53	<chem>N(C(=[Se])NCCc1ccc(cc1)S(=O)(=O)N)C(=O)c1cc2c(cc1)OCO2</chem>	2.565	false	false	false	false	false
54	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]C#N</chem>	5.959	false	false	false	false	false

55	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se][Se]c1ccc(cc1)S(=O)(=O)N</chem>	1.777	false	false	false	false	false
----	---	-------	-------	-------	-------	-------	-------

Table S7. Evaluation of the applicability domain of the external set (from reference 41) in model M19 and prediction of $\log(u_0)$

Name	SMILES	Predicted u_0 by M19	Applicability Domain				
			PCA-range	Euclidean distance	City-block distance	Probability density	Consensus domain
6	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]CCO</chem>	-2.818	false	false	false	false	false
7	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]CCc1ccccc1</chem>	-1.692	false	false	false	false	false
8	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]C</chem>	-3.418	false	false	false	false	false
9	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]C(C)C</chem>	-0.571	false	false	false	false	false
10	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]CC#C</chem>	-2.331	false	false	false	false	false
11	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]CCOC</chem>	-2.536	false	false	false	false	false
12	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]CC=C</chem>	-2.477	false	false	false	false	false
13	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]C1CCCCC1</chem>	-1.297	false	false	false	false	false
14	<chem>c1(ccc(cc1)S(=O)(=O)N)C[Se]c1ccccc1</chem>	-2.293	false	false	false	false	false
15	<chem>c1(ccc(cc1)S(=O)(=O)N)C(=[Se])N</chem>	15.092	false	false	false	false	false
16	<chem>c1(ccc(cc1)S(=O)(=O)N)c1cnc([se]1)c1ccc(cc1)F</chem>	6.411	false	false	false	false	false
17	<chem>c1(ccc(cc1)S(=O)(=O)N)c1cnc([se]1)C[Se]c1ccccc1</chem>	3.431	false	false	false	false	false
18	<chem>c1(ccc(cc1)S(=O)(=O)N)c1cnc([se]1)C[Te]c1ccc(cc1)C</chem>	3.417	false	false	false	false	false
19	<chem>c1(cc(c(cc1)N)S(=O)(=O)N)[Se]C#N</chem>	-1.986	false	false	false	false	false
20	<chem>c1(cc(c(cc1)N)S(=O)(=O)N)[Se][Se]c1cc(c(cc1)N)S(=O)(=O)N</chem>	-0.742	false	false	false	false	false
21	<chem>c1(cc(ccc1)S(=O)(=O)N)[Se][Se]c1cc(ccc1)S(=O)(=O)N</chem>	-1.003	false	false	false	false	false
22	<chem>c1(c(cccc1)S(=O)(=O)N)[Se][Se]c1ccccc1S(=O)(=O)N</chem>	-1.025	false	false	false	false	false
23	<chem>c1cc(ccc1)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	4.194	false	false	false	false	false
24	<chem>c1cc(ccc1F)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	4.243	false	false	false	false	false
25	<chem>c1cc(c(cc1)F)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	3.972	false	false	false	false	false
26	<chem>c1cc(cc(c1)C(F)(F)F)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	4.03	false	false	false	false	false
27	<chem>c1cc(c(cc1)Cl)NC(=[Se])Nc1ccc(cc1)S(=O)(=O)N</chem>	4.008	false	false	false	false	false
28	<chem>c1cc(ccc1)NC(=[Se])Nc1cccc(c1)S(=O)(=O)N</chem>	4.168	false	false	false	false	false
29	<chem>c1cc(ccc1F)NC(=[Se])Nc1cccc(c1)S(=O)(=O)N</chem>	4.216	false	false	false	false	false
30	<chem>c1cc(ccc1)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	1.391	false	true	true	false	true

31	<chem>c1cc(c(cc1)OC)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	0.969	false	false	true	false	false
32	<chem>c1cc(c(cc1)F)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	1.155	false	true	true	false	true
33	<chem>c1cc(cc(c1)C(F)(F)F)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	1.319	false	false	false	false	false
34	<chem>c1cc(c(cc1)Cl)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	1.198	false	true	true	true	true
35	<chem>c1cc(ccc1Br)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	0.92	false	false	true	false	false
36	<chem>c1cc(ccc1I)NC(=[Se])NCCc1ccc(cc1)S(=O)(=O)N</chem>	0.849	false	false	false	false	false
37	<chem>N(C(=[Se])NCCc1ccc(cc1)S(=O)(=O)N)c1cc2cccc2cc1</chem>	1.428	false	false	true	false	false
38	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1cccc1</chem>	1.675	false	true	true	false	true
39	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1cc2c(cc1)OCO2</chem>	1.794	false	true	true	false	true
40	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1cc(cc(c1)Br)OC(F)(F)F</chem>	1.127	false	false	false	false	false
41	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1cccc1OC(=O)C</chem>	1.766	false	false	false	false	false
42	<chem>N(C(=[Se])Nc1ccc(cc1)S(=O)(=O)N)C(=O)c1ccc(cc1)CCCCCCC</chem>	1.717	false	true	false	false	false
43	<chem>N(C(=[Se])Nc1cccc(c1)S(=O)(=O)N)C(=O)c1cccc1</chem>	1.644	false	false	false	false	false
44	<chem>N(C(=[Se])Nc1cccc(c1)S(=O)(=O)N)C(=O)c1cc(ccc1)C</chem>	1.665	false	false	false	false	false
45	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cccc1</chem>	1.725	false	true	true	false	true
46	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cccc(c1)C</chem>	1.749	false	true	true	false	true
47	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1ccc(c(c1)C)Br</chem>	1.285	false	false	true	false	false
48	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cccc1OC(=O)C</chem>	1.826	false	true	true	false	true
49	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cc(cc(c1)OC(F)(F)F)Br</chem>	1.202	false	false	false	false	false
50	<chem>N(C(=[Se])Nc1c(ccc(c1)S(=O)(=O)N)O)C(=O)c1cc2c(cc1)OCO2</chem>	1.861	false	true	true	false	true
51	<chem>N(C(=[Se])NCCc1ccc(cc1)S(=O)(=O)N)C(=O)c1cccc1</chem>	-1.087	false	false	false	false	false
52	<chem>N(C(=[Se])NCCc1ccc(cc1)S(=O)(=O)N)C(=O)c1cccc(c1)C</chem>	-1.046	false	false	false	false	false
53	<chem>N(C(=[Se])NCCc1ccc(cc1)S(=O)(=O)N)C(=O)c1cc2c(cc1)OCO2</chem>	-0.585	false	false	false	false	false
54	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se]C#N</chem>	-2.219	false	false	false	false	false
55	<chem>c1(ccc(cc1)S(=O)(=O)N)[Se][Se]c1ccc(cc1)S(=O)(=O)N</chem>	-1.216	false	false	false	false	false

