Supporting information

An Automated force field Topology Builder (ATB) and repository: version 1.0

Alpeshkumar K. Malde,[1] Le Zuo,[1] Matthew Breeze,[1] Martin Stroet,[1] David Poger,[1] Pramod C. Nair,[1] Chris Oostenbrink[2,3] and Alan E. Mark[,1,4]*

[1]School of Chemistry and Molecular Biosciences,

[2]Leiden/Amsterdam Center for Drug Research, Division of Molecular Toxicology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands,

[3]Institute of Molecular Modeling and Simulation, University of Natural Resources and Applied Life Sciences, Vienna, Austria, and

[4]Institute for Molecular Bioscience, University of Queensland, St. Lucia, QLD 4072, Australia.

* Corresponding author

AUTHOR EMAIL ADDRESSES: **a.e.mark@uq.edu.au**, **a.malde@uq.edu.au**

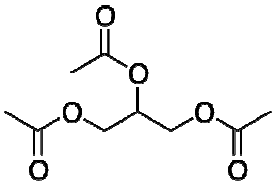
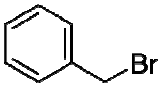
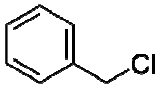
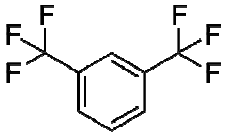
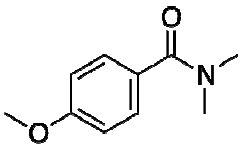
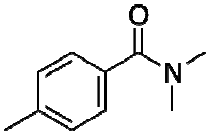
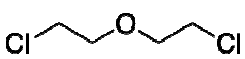
Table S1. A comparison between the experimental and calculated free energies of hydration (ΔG_{hyd}) for biologically relevant small organic molecules using parameters assigned by the Automated Topology Builder (ATB). The RMSD between the QM optimized structure as obtained from the ATB and the one obtained after MD simulation of 1 ns in SPC water using the GROMOS force field assigned by the ATB is also shown. Values are in $kJ.mol^{-1}$. No errors have been reported for the experimental data on these molecules.

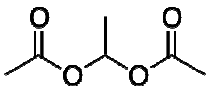
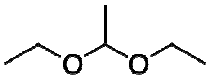
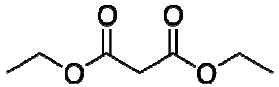
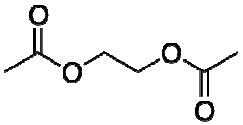
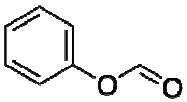
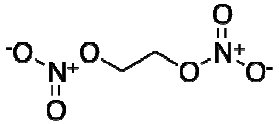
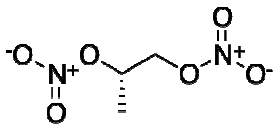
RNME on the ATB	**Name**	**$\Delta G_{hyd;\ exp}$**	**$\Delta G_{hyd;\ cal}$**	**$\lvert\Delta G_{hyd;cal} - \Delta G_{hyd;\ exp}\rvert$**	**All-atom RMSD (nm)**
Alcohols					
ISOP	*Isopropanol*	-19.9	-23.0±0.8	3.1	0.008
G090	*2-methyl-1-propanol*	-18.9	-22.3±1.2	3.4	0.034
G091	*2-butanol*	-19.1	-19.1±1.1	0.0	0.085
G092	*2-methyl-2-propanol*	-18.9	-19.7±0.9	0.8	0.057
G094	*3-methyl-1-butanol*	-18.5	-21.3±1.5	2.8	0.067
G095	*2-pentanol*	-18.7	-19.9±1.5	1.2	0.054
G096	*3-pentanol*	-18.2	-15.5±1.6	2.7	0.045
G097	*2-methyl-2-butanol*	-18.5	-23.9±1.1	5.4	0.064
G099	*2,3-dimethyl-2-butanol*	-16.4	-19.7±1.1	3.3	0.130
G100	*3-hexanol*	-17.0	-17.6±1.6	0.6	0.024
G101	*4-methyl-2-pentanol*	-15.6	-19.3±1.8	3.7	0.122
G102	*2-methyl-3-pentanol*	-16.3	-13.1±1.6	3.2	0.047
G103	*2-methyl-2-pentanol*	-16.4	-22.3±1.4	5.9	0.053
G105	*4-heptanol*	-16.8	-18.4±1.9	1.6	0.032
G108	*Cyclopentanol*	-23.0	-21.4±1.2	1.6	0.053
G109	*Cyclohexanol*	-22.9	-20.6±1.2	2.3	0.052
G110	*Cycloheptanol*	-22.9	-19.0±1.3	3.9	0.077
G111	*Phenol*	-27.7	-28.9±1.2	1.2	0.044
F035	*2-methylphenol*	-24.6	-24.4±1.5	0.2	0.049
_I0H	*4-methylphenol*	-26.6	-25.8±1.1	0.8	0.045
Average				**2.4**	
Aldehydes					
G146	*Ethanal*	-14.7	-15.3±0.6	0.6	0.006
G147	*Propanal*	-14.4	-14.4±-10.4	0.0	0.014
G148	*Butanal*	-13.3	-9.7±1.0	3.6	0.028
G149	*Pentanal*	-12.7	-10.2±1.3	2.5	0.096
G158	*Benzaldehyde*	-16.8	-13.4±0.8	3.4	0.076
Average				**2.0**	
Alkanes					
_ICJ	*Ethane*	7.7	7.8±0.5	0.1	0.000
_I0I	*Propane*	8.2	8.4±0.6	0.2	0.000
G006	*Pentane*	9.8	10.0±1.1	0.2	0.073
G034	*Methyl butane*	10.0	10.7±1.0	0.7	0.071
G008	*Dimethyl propane*	10.5	10.0±0.7	0.5	0.006
_ICZ	*Hexane*	10.4	10.1±0.9	0.3	0.031
G013	*Heptane*	11.0	11.5±1.4	0.5	0.079
G015	*Octane*	12.1	11.4±1.9	0.7	0.064
Average				**0.4**	
Alkenes					
G026	*Ethene*	5.3	-1.3±0.5	6.6	0.004
G027	*Propene*	5.3	8.4±0.6	3.1	0.105
G028	*1-Butene*	5.8	8.3±0.8	2.5	0.026
G031	*trans-2-pentene*	5.6	8.4±0.8	2.8	0.133
G007	*1-hexene*	7.0	10.4±1.3	3.4	0.078
G039	*Cyclopentene*	2.3	6.6±0.7	4.3	0.015
G040	*Cyclohexene*	1.5	6.0±0.8	4.5	0.020
Average				**3.9**	
Alkyl benzenes					
G058	*Benzene*	-3.6	-3.8±0.8	0.2	0.007

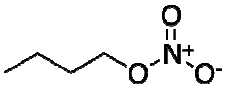
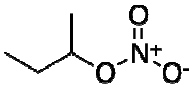
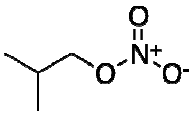
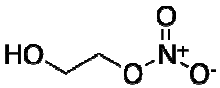
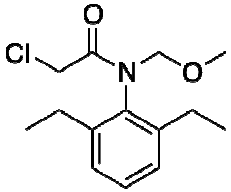
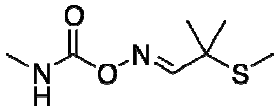
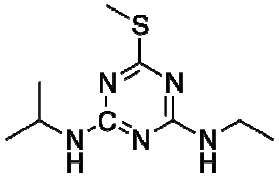
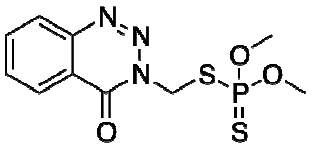
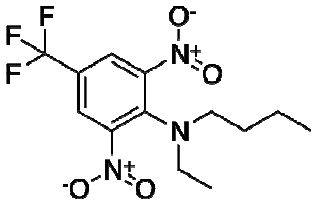
_I0L	*Methyl benzene (toluene)*	-3.7	2.8±0.8	6.5	0.011
G060	*Ethyl benzene*	-3.3	-0.1±1.0	3.2	0.066
G064	*Propyl benzene*	-2.2	0.5±1.4	2.7	0.082
G067	*n-Butyl benzene*	-1.7	0.3±1.6	2.0	0.061
G065	*Isopropyl benzene*	-1.3	3.2±1.2	4.5	0.102
G068	*sec-Butyl benzene*	-1.9	0.2±1.7	2.1	0.021
Average				**3.0**	
Alkynes					
G049	*Ethine*	0.0	1.4±0.5	1.4	0.000
G050	*Propine*	-1.3	5.2±0.6	6.5	0.012
G051	*1-Butine*	-0.7	4.1±0.6	4.8	0.017
G052	*1-Pentine*	0.0	5.5±1.0	5.5	0.081
G053	*1-Hexine*	1.2	8.0±1.1	6.8	0.030
Average				**5.0**	
Amides					
_I0X	*Acetamide*	-40.6	-39.0±1.0	1.6	0.005
_I0J	*Propyl amide*	-39.4	-42.5±0.9	3.1	0.009
G235	*N-methyl acetamide*	-42.2	-28.3±0.9	3.9	0.008
G236	*N,N-dimethyl acetamide*	-28.3	-34.4±1.1	6.1	0.015
Average				**3.7**	
Amines (Primary)					
G191	*Methyl amine*	-19.1	-11.3±0.6	7.8	0.035
G192	*Ethyl amine*	-18.8	-9.8±0.7	9.0	0.060
G193	*Propyl amine*	-18.4	-10.5±0.9	7.9	0.034
_I0F	*Butyl amine*	-18.3	-9.0±1.0	9.3	0.041
Average				**8.5**	
Carboxylic acids					
_I0M	*Acetic acid*	-28.0	-29.8±1.1	1.8	0.009
_I0K	*Propanoic acid*	-27.0	-31.6±1.0	4.6	0.014
G161	*Butanoic acid*	-26.6	-30.6±1.6	4.0	0.027
Average				**3.5**	
Cycloalkanes					
G018	*Cyclopropane*	3.1	9.5±0.6	6.4	0.000
G019	*Cyclopentane*	5.0	5.4±0.7	0.4	0.003
G021	*Methyl cyclopentane*	6.7	10.5±0.8	3.8	0.008
CHE	*Cyclohexane*	5.1	2.9±0.8	1.2	0.012
G022	*Cycloheptane*	3.3	3.6±0.8	0.3	0.047
G023	*Methyl cyclohexane*	7.1	10.0±1.0	2.9	0.006
Average				**2.5**	
Esters					
G162	*Methyl methanoate*	-11.6	-19.7±1.5	8.1	0.023
G163	*Ethyl methanoate*	-11.1	-21.8±1.8	10.7	0.073
G164	*Methyl ethanoate*	-13.9	-17.1±1.5	3.2	0.010
G165	*Propyl methanoate*	-10.4	-14.4±1.9	4.0	0.077
G167	*Ethyl ethanoate*	-12.9	-19.1±1.6	6.2	0.011
G168	*Methyl propanoate*	-12.3	-15.0±1.9	2.7	0.045
G190	*Methyl benzoate*	-17.9	-17.0±1.8	0.9	0.095
Average				**5.1**	
Ketones					
G130	*Dimethylketone*	-16.1	-15.9±0.7	0.2	0.005
G131	*Methyl ethyl ketone*	-15.2	-12.5±0.8	2.7	0.023
G132	*Methyl-1-propyl ketone*	-14.8	-12.8±1.2	2.0	0.078
G133	*Diethyl ketone*	-14.3	-10.7±1.0	3.6	0.087
Average				**2.1**	
Thiols and sulfides					
_I0A	*Methyl thiol*	-5.2	-8.4±0.5	3.2	0.000
G288	*Ethyl thiol*	-5.4	-6.8±0.7	3.4	0.067
G289	*Benzene thiol*	-10.7	-14.2±0.9	3.5	0.057
G290	*Dimethyl sulfide*	-6.5	1.2±0.6	7.7	0.000

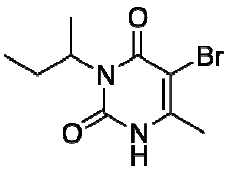
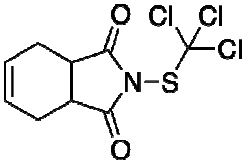
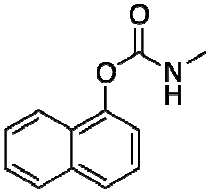
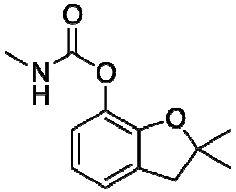
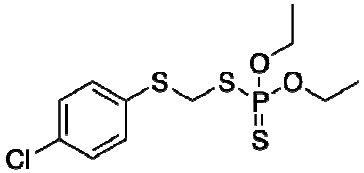
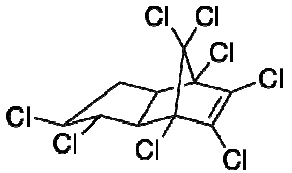
G291	*Diethyl sulfide*	-6.0	-0.4±0.9	5.6	0.080
G292	*Methyl thiobenzene*	-11.4	-4.6±1.2	6.8	0.066
Average				**5.0**	
Overall average				**3.3**	

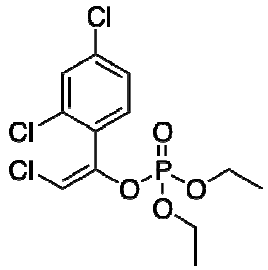
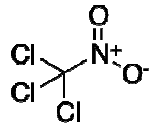
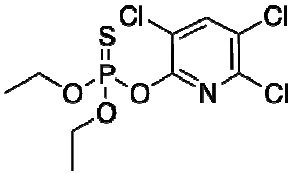
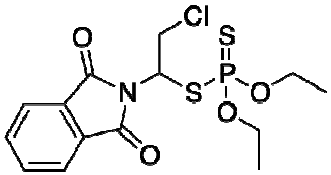
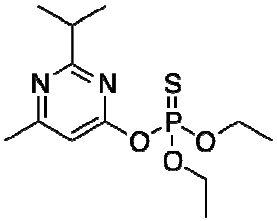
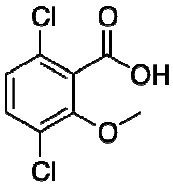
Table S2. A comparison between the experimental and calculated free energies of hydration (ΔG_{hyd}) for the drug-like and drug molecules from SAMPL0, 1 and 2 challenges using parameters assigned by the Automated Topology Builder (ATB). The RMSD between the QM optimized structure as obtained from the ATB and the one obtained after MD simulation of 1 ns in SPC water using the GROMOS force field assigned by the ATB is also shown. Only molecules with ≤ 40 atoms are considered. Values are in $kJ.mol^{-1}$.

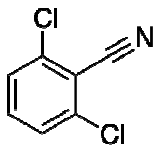
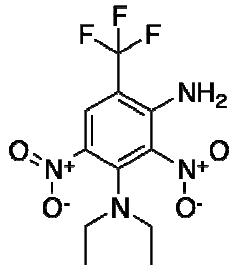
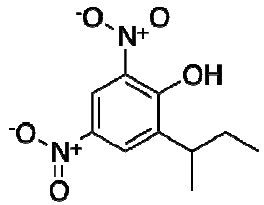
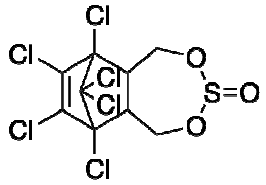
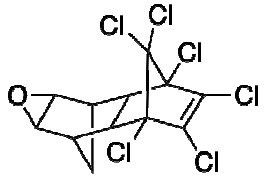
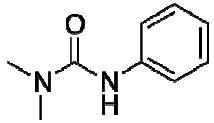
RNME on the ATB	**Name**	**Chemical structure**	**$\Delta G_{hyd;\ exp}$**	**$\Delta G_{hyd;\ cal}$**	**$\lvert\Delta G_{hyd;cal} - \Delta G_{hyd;\ exp}\rvert$ ± error**	**All-atom RMSD (nm)**
SAMPL0/CUP8 dataset						
M494	*Glycerol triacetate*		-36.7±0.8	-45.9±4.1	9.2±4.9	0.106
S002	*Benzyl bromide*		-9.9±0.8	-4.1±1.1	5.8±1.9	0.068
S003	*Benzyl chloride*		-8.1±0.8	-0.3±1.0	7.8±1.8	0.079
S004	*m-bis-trifluoromethylbenzene*		4.5±0.8	26.7±1.1	22.2±1.9	0.103
S005	*N,N-dimethyl-p-methoxy benzamide*		-46.0±0.8	-43.5±1.9	2.5±2.7	0.068
S006	*N,N,4-trimethylbenzamide*		-40.8±0.8	-38.8±1.4	2.0±2.2	0.048
S007	*Bis-2-chloroethyl ether*		-17.7±0.8	-3.1±1.8	14.6±2.6	0.094

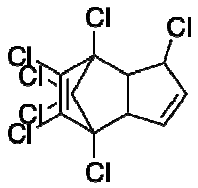
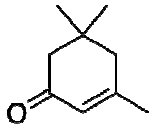
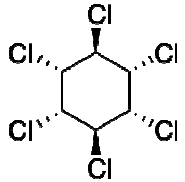
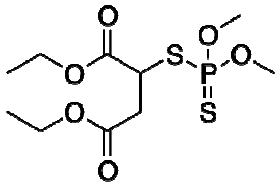
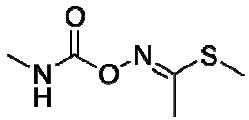
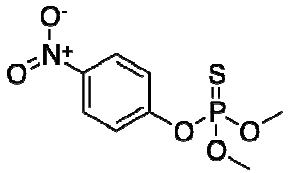
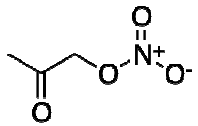
S008	*1,1-diacetoxyethane*		-20.8±0.8	-26.5±3.7	5.7±4.5	0.026
G253	*1,1-diethoxyethane*		-13.7±0.8	-9.3±2.4	4.4±3.2	0.087
G256	*1,4-dioxane*		-21.1±0.8	-6.3±1.3	14.8±2.1	0.006
M291	*Diethyl propanedioate*		-25.1±0.8	-29.9±2.7	4.8±3.5	0.179
G251	*Dimethoxymethane*		-12.2±0.8	-5.7±1.0	6.5±1.8	0.074
S013	*Ethylene glycol diacetate*		-26.5±0.8	-29.7±2.5	3.2±3.3	0.190
G254	*1,2-diethoxyethane*		-14.8±0.8	-3.6±1.8	11.2±2.6	0.092
G291	*Diethyl sulfide*		-6.0±0.8	0.2±0.9	6.2±1.7	0.092
S016	*Phenyl formate*		-16.0±0.8	-18.8±1.6	2.8±2.4	0.012
F135	*Imidazole*		-41.0±0.8	-37.6±0.8	3.4±1.6	0.016
SAMPL1 dataset						
_JCE	*Nitroglycol*		-24.0±0.4	-11.4±2.2	12.6±2.6	0.052
8002	*1,2-dinitroxypropane*		-20.7±0.4	-11.5±2.2	9.2±2.6	0.119

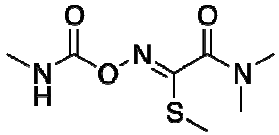
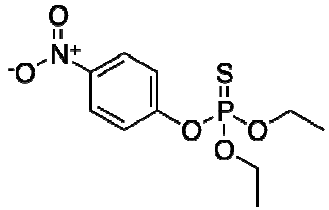
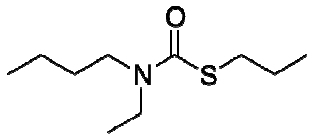
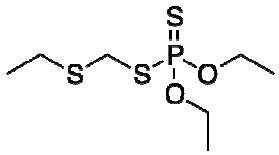
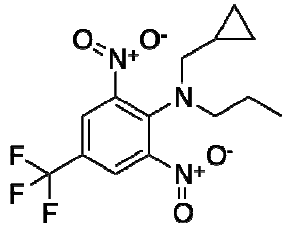
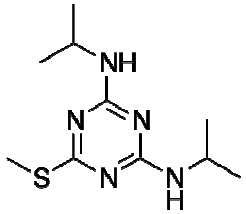
8003	*Butyl nitrate*		-8.7±0.4	-6.3±1.4	2.4±1.8	0.105
8004	*2-butyl nitrate*		-7.6±0.4	-0.3±1.7	7.3±2.1	0.117
8005	*Isobutyl nitrate*		-7.9±0.4	-4.7±1.4	3.2±1.8	0.062
8006	*Ethyleneglycol mononitrate*		-34.2±0.4	-25.5±1.3	8.7±1.7	0.112
8007	*Alachlor*		-34.4±1.2	-23.3±2.9	11.1±4.1	0.113
8008	*Aldicarb*		-41.2±0.4	-41.9±2.7	0.7±3.1	0.071
8009	*Ametryn*		-32.0±1.4	-21.5±2.2	10.5±3.6	0.083
8010	*Azinphosmethyl*		-42.0±5.7	-47.7±2.9	5.7±8.6	0.244
8011	*Benefin*		-14.7±8.1	-11.7±4.9	3.0±13	0.156

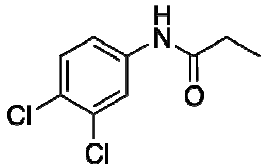
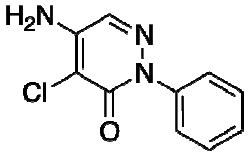
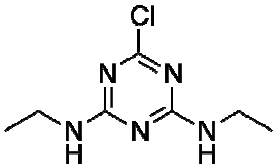
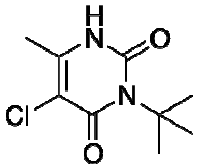
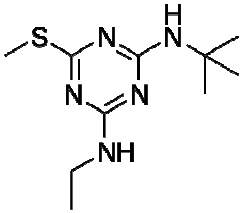
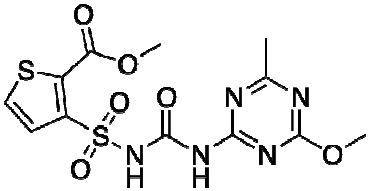
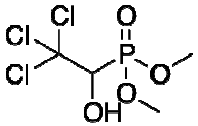
8013	*Bromacil*		-40.7±8.1	-68.0±2.3	27.3±10.4	0.023
8014	*Captan*		-37.7±8.1	-23.9±2.2	13.8±10.3	0.118
8015	*Carbaryl*		-39.5±0.4	-40.9±1.9	1.4±2.3	0.101
8016	*Carbofuran*		-40.2±1.3	-36.6±1.9	3.6±3.2	0.021
8017	*Carbophenothion*		-27.2±3.5	-32.8±3.3	5.6±6.8	0.255
8018	*Chlordane*		-14.4±0.4	1.2±1.5	15.6±1.9	0.018

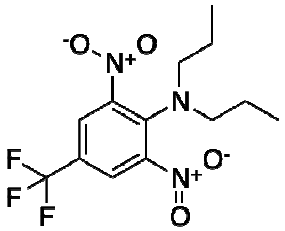
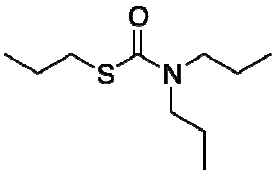
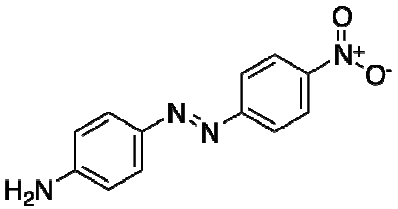
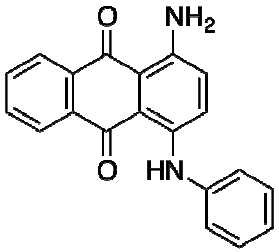
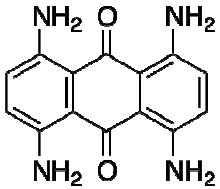
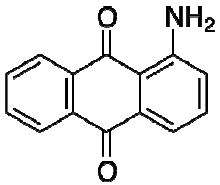
8019	*Chlorfenvinphos*		-29.6±5.7	-27.1±3.7	2.5±9.4	0.234
8021	*Chloropicrin*		-6.1±0.4	7.0±.0.9	13.1±1.3	0.115
8022	*Chlorpyrifos*		-21.1±0.9	-22.8±2.7	1.7±3.6	0.263
8023	*Dialifor*		-24.0±8.1	-41.0±4.4	17.0±12.5	0.146
8024	*Diazinon*		-27.1±0.5	-27.3±3.3	0.2±3.8	0.249
8025	*Dicamba*		-41.3±8.1	-39.5±2.3	1.8±10.4	0.035

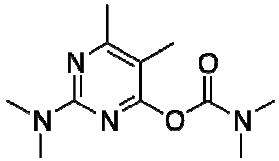
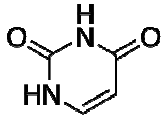
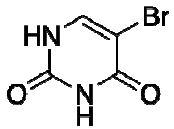
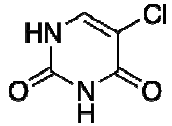
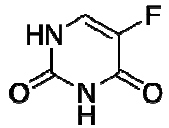
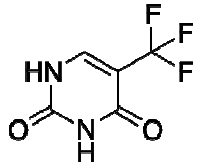
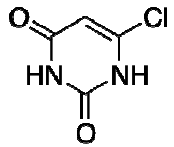
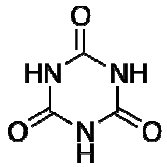
8026	*Dichlobenil*		-19.7±8.1	-24.1±1.0	4.4±9.1	0.024
8027	*Dinitramine*		-23.7±8.1	-30.4±3.7	6.7±11.8	0.092
8028	*Dinoseb*		-26.1±8.1	-30.3±3.2	4.2±11.3	0.067
8029	*Endosulfan alpha*		-17.7±1.1	-8.0±1.4	9.7±2.5	0.015
8030	*Endrin*		-20.2±0.4	-1.0±1.4	19.2±1.8	0.015
8032	*Fenuron*		-38.2±8.1	-41.5±1.5	3.3±9.6	0.047

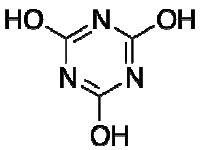
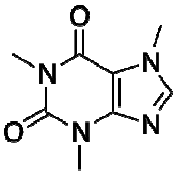
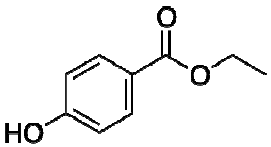
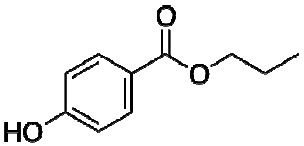
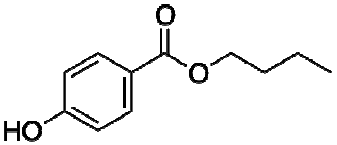
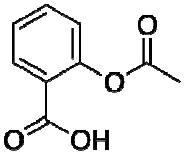
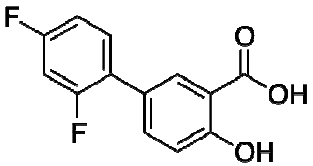
8033	*Heptachlor*		-10.7±0.4	1.6±1.4	12.3±1.8	0.021
8034	*Isophorone*		-21.7±5.7	-24.2±1.4	2.5±7.1	0.013
8035	*Lindane*		-22.8±0.4	9.0±2.4	31.8±2.8	0.097
8036	*Malathion*		-34.1±0.9	-50.7±4.9	16.6±5.8	0.127
8037	*Methomyl*		-44.6±8.1	-43.2±2.6	1.4±10.7	0.062
8038	*Methyparathion*		-30.1±0.4	-32.1±2.5	2.0±2.9	0.161
8041	*Nitroxyacetone*		-25.1±0.4	-15.5±1.4	9.6±1.8	0.069

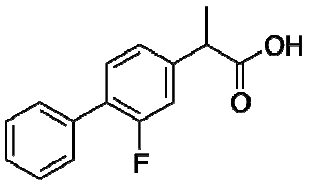
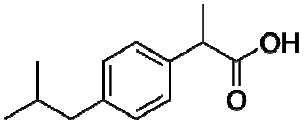
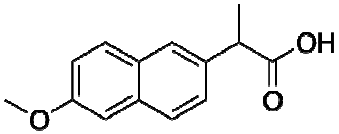
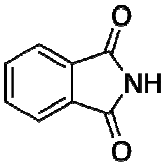
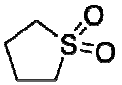
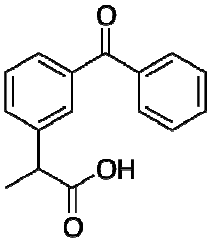
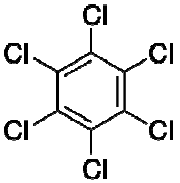
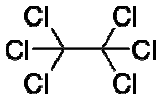
8042	*Oxamyl*		-42.6±8.1	-61.3±3.2	18.7±11.3	0.066
8043	*Parathion*		-28.2±0.4	-37.2±2.6	9.0±3.0	0.142
8044	*Pebulate*		-15.2±8.1	-10.9±2.5	4.3±10.6	0.129
8045	*Phorate*		-18.3±0.4	-24.7±2.7	6.4±3.1	0.167
8046	*Profluralin*		-10.3±5.7	6.4±3.4	16.7±9.1	0.109
8047	*Prometryn*		-35.3±0.4	-20.6±3.0	14.7±3.4	0.057

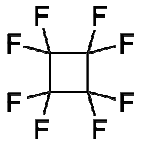
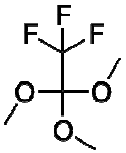
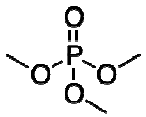
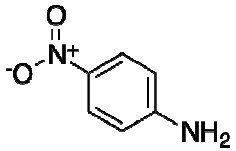
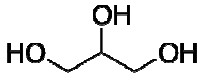
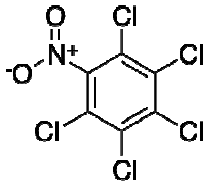
8048	*Propanil*		-32.6±8.1	-34.6±1.5	2.0±9.6	0.045
8049	*Pyrazon*		-68.7±8.1	-60.8±1.7	7.9±9.8	0.047
8050	*Simazine*		-42.8±0.4	-40.5±1.8	2.3±2.2	0.043
8052	*Terbacil*		-46.6±8.1	-53.3±1.9	6.7±10	0.129
8053	*Terbutryn*		-27.9±1.8	-19.6±2.3	8.3±4.1	0.097
8054	*Thifensulfurone*		-67.9±8.1	-99.5±4.1	31.6±12.2	0.262
8055	*Trichlorofon*		-53.3±8.1	-23.5±2.6	29.8±10.7	0.189

8056	*Trifluralin*		-13.6±0.4	-2.4±2.9	11.2±3.3	0.113
8057	*Vernolate*		-17.3±5.7	-12.0±2.6	5.3±8.3	0.085
8058	*4-amino-4'-nitroazobenzene*		-47.0±1.8	-53.8±2.3	6.8±4.1	0.054
8059	*1-amino-4-anilino-anthraquinone*		-31.1±8.1	-43.9±2.4	12.8±10.5	0.074
8060	*1,4,5,8-tetramino-anthraquinone*		-37.4±5.7	-74.6±1.9	37.2±7.6	0.052
8061	*1-amino-anthraquinone*		-33.3±5.7	-37.0±1.5	3.7±7.2	0.018

8063	*Pirimor (pirimacarb)*		-39.4±8.1	-31.1±2.8	8.3±10.9	0.039
SAMPL2 dataset						
2S01	*Uracil*		-69.4±1.2	-72.1±1.1	2.7±2.3	0.013
F008	*5-bromouracil*		-76.0±2.3	-64.5±2.2	11.5±4.5	0.066
2S03	*5-chlorouracil*		-74.2±3.3	-76.4±1.1	2.2±4.4	0.010
2S04	*5-fluorouracil*		-70.8±3.7	-74.7±1.0	3.9±4.7	0.013
2S06	*5-trilfuoromethyluracil*		-64.7±0.7	-62.1±1.2	2.6±1.9	0.090
2S07	*6-chlorouracil*		-66.2±5.1	-61.6±1.1	4.6±6.2	0.013
2S08	*Cyanuric acid - keto tautomer*		-76.4±1.1	-82.4±1.1	6.0±2.2	0.019

	- Enol tautomer			-75.0±3.5		0.093
2S09	*Caffeine*		-52.9±3.1	-65.0±1.3	12.1±4.4	0.018
2S11	*Ethyl paraben*		-38.5±1.3	-54.6±1.8	16.1±3.1	0.129
2S12	*Propyl paraben*		-39.2±0.9	-54.6±2.4	15.4±3.3	0.144
2S13	*Butyl paraben*		-36.5±1.1	-52.0±2.6	15.5±3.7	0.085
2S14	*Acetylsalicylic acid*		-41.6±0.8	-46.5±2.5	4.9±3.3	0.087
2S15	*Diflunisal*		-39.3±0.8	-44.9±3.5	5.6±4.3	0.110

2S16	*Flurbiprofen*		-35.2±0.7	-34.5±2.7	0.7±3.4	0.077
2S17	*Ibuprofen*		-29.3±2.7	-31.9±2.6	2.6±5.3	0.219
2S18	*Naproxen*		-42.7±0.8	-45.0±2.6	2.3±3.4	0.034
2S19	*Phthalimide*		-40.2±2.1	-47.1±1.1	6.9±3.2	0.017
2S22	*Sulfolane*		-36.0±1.3	-32.9±1.1	3.1±2.4	0.037
2S23	*Ketoprofen*		-45.1±0.8	-53.1±2.8	8.0±3.6	0.274
2S24	*Hexachlorobenzene*		-9.6±4.9	4.3±1.1	13.9±6.0	0.016
2S25	*Hexachloroethane*		-5.9±0.4	4.7±1.2	10.6±1.6	0.009

2S26	*Octofluorocyclobutane*		-12.6±0.1	32.6±0.9	45.2±1.0	0.007
2S27	*Trimethyl orthotrifluoroacetate*		-3.3±0.8	12.1±1.7	15.4±2.5	0.062
2S28	*Trimethyl phosphate*		-36.4±0.4	-26.3±1.7	10.1±2.1	0.068
2S29	*4-nitroaniline*		-41.9±0.5	-55.7±1.2	13.8±1.7	0.021
G250	*Glycerol*		-56.1±4.2	-58.3±1.6	2.2±5.8	0.103
2S31	*pentachloronitrobenzene*		-21.8±1.0	-2.6±1.5	19.2±2.5	0.016