Supporting Information

How Does the Methodology of 3D Structure Preparation Influence the Quality of pK_a Prediction?

Stanislav Geidl, Radka Svobodová Vařeková*, Veronika Bendová, Lukáš Petrusek, Crina-Maria Ionescu, Zdeněk Jurka, Ruben Abagyan, and Jaroslav Koča*

E-mail: svobodova@chemi.muni.cz, jkoca@chemi.muni.cz

Contents

1 List of Molecules

Table S1: List of molecules.

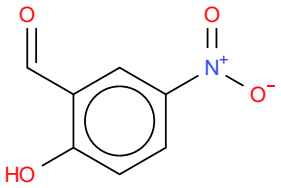
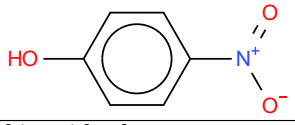
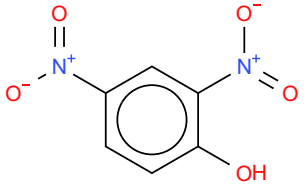
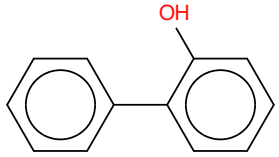
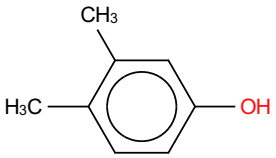
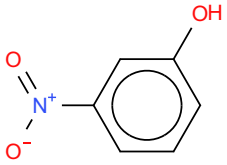
NSC number		**CAS number**	**pK_a**
smiles			
Phenols			
NSC_881		97-51-8	5.51
O=Cc1c(ccc(c1)[N+](=O)[O-])O			
NSC_1317		100-02-7	7.15
c1c(ccc(c1)O)[N+](=O)[O-]			
NSC_1532		51-28-5	4.09
c1(cc(ccc1O)[N+](=O)[O-])[N+](=O)[O-]			
NSC_1548		90-43-7	9.97
c1ccccc1c1ccccc1O			
NSC_1549		95-65-8	10.4
c1c(ccc(c1C)C)O			
NSC_1551		554-84-7	8.36
c1(cc(ccc1)[N+](=O)[O-])O			

Table S1: List of molecules.

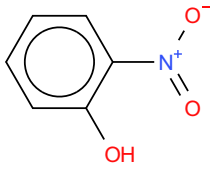
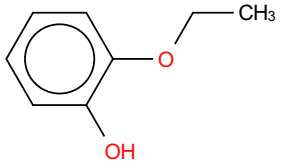
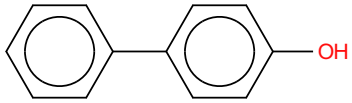
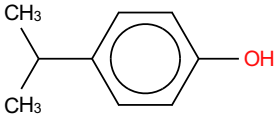
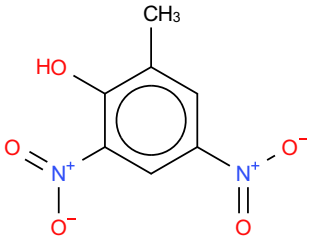
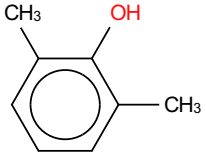
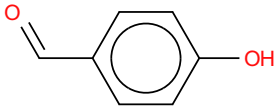
NSC number		**CAS number**	**pK_a**
smiles			
NSC_1552		88-75-5	7.23
c1c(c(ccc1)O)[N+](=O)[O-]			
NSC_1809		94-71-3	10.1
c1c(c(ccc1)O)OCC			
NSC_1858		92-69-3	9.55
c1ccccc1c1ccc(cc1)O			
NSC_1888		99-89-8	10.2
c1(ccc(cc1)C(C)C)O			
NSC_2082		534-52-1	4.31
c1c(cc(c(c1C)O)[N+](=O)[O-])[N+](=O)[O-]			
NSC_2123		576-26-1	10.6
c1(c(cccc1C)C)O			
NSC_2127		123-08-0	7.61
c1cc(ccc1C=O)O			

Table S1: List of molecules.

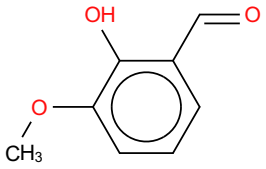
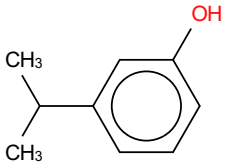
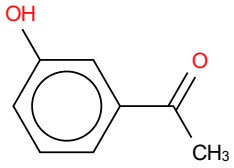
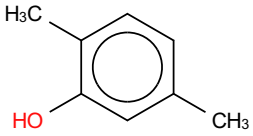
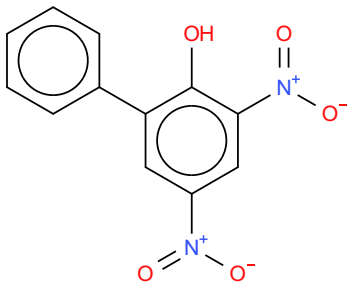
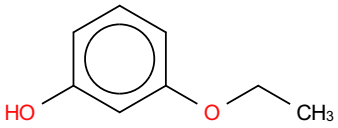
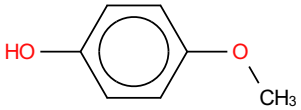
NSC number		**CAS number**	**pK_a**
smiles			
NSC_2150		148-53-8	7.91
c1(c(c(ccc1)OC)O)C=O			
NSC_2209		618-45-1	10.2
c1(cc(ccc1)C(C)C)O			
NSC_2440		121-71-1	9.25
c1c(cccc1O)C(=O)C			
NSC_2599		95-87-4	10.4
c1cc(cc(c1C)O)C			
NSC_2880		731-92-0	3.85
c1ccccc1c1cc(cc(c1O)[N+](=O)[O-])[N+](=O)[O-]			
NSC_4875		621-34-1	9.65
c1c(cc(cc1)O)OCC			
NSC_4960		150-76-5	10.1
c1c(ccc(c1)O)OC			

Table S1: List of molecules.

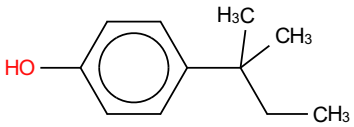
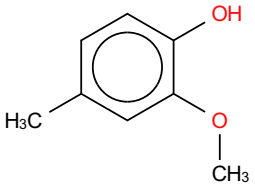
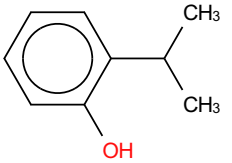
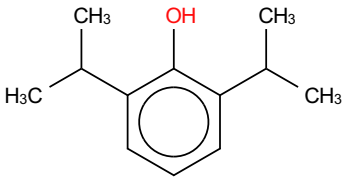
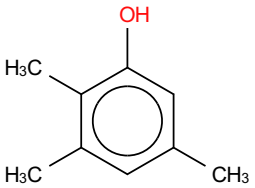
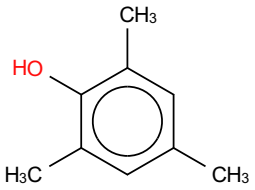
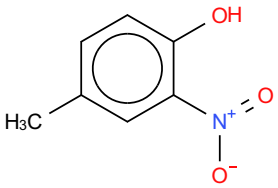
NSC number		**CAS number**	**pK_a**
smiles			
NSC_4965		80-46-6	10.4
c1c(ccc(c1)O)C(CC)(C)C			
NSC_4969		93-51-6	10.3
c1(ccc(cc1OC)C)O			
NSC_5103		88-69-7	10.5
c1c(c(ccc1)O)C(C)C			
NSC_5105		2078-54-8	11.1
c1(c(cccc1C(C)C)C(C)C)O			
NSC_5296		697-82-5	10.7
c1c(cc(c(c1O)C)C)C			
NSC_5353		527-60-6	10.9
c1(cc(cc(c1O)C)C)C			
NSC_5387		119-33-5	7.6

Table S1: List of molecules.

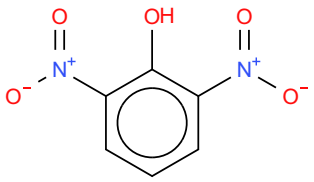
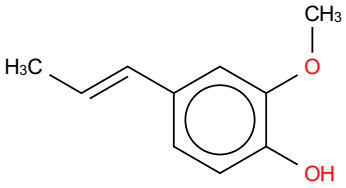
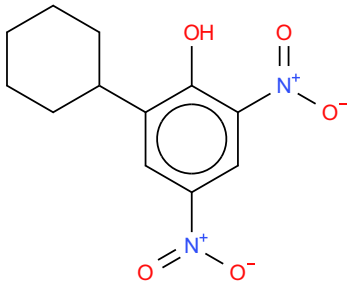
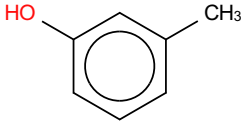
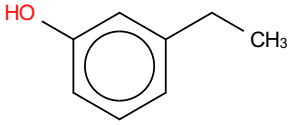
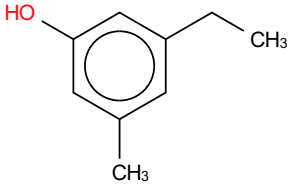
NSC number		**CAS number**	**pK_a**
smiles			
c1(ccc(cc1[N+](=O)[O-])C)O			
NSC_6215		573-56-8	3.97
c1(c(cccc1[N+](=O)[O-])[N+](=O)[O-])O			
NSC_6769		97-54-1	9.88
c1c(c(ccc1/C=C/C)O)OC			
NSC_7739		131-89-5	4.52
C1C(CCCC1)c1cc(cc(c1O)[N+](=O)[O-])[N+](=O)[O-]			
NSC_8768		108-39-4	10.1
c1c(cccc1O)C			
NSC_8873		620-17-7	9.9
c1c(cccc1O)CC			
NSC_8885		698-71-5	10.1
c1c(cc(cc1O)C)CC			

Table S1: List of molecules.

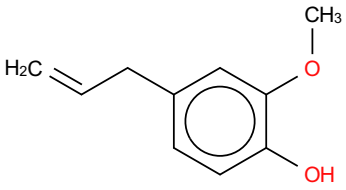
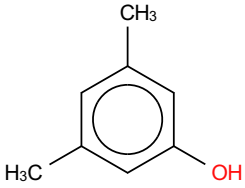
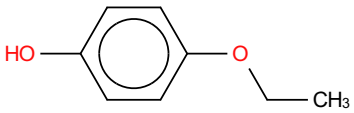
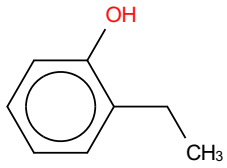
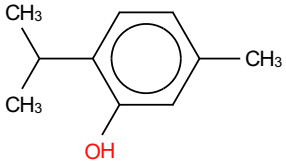
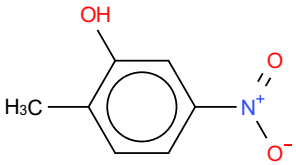
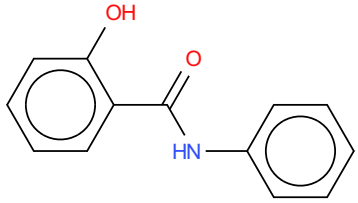
NSC number		**CAS number**	**pK_a**
smiles			
NSC_8895		97-53-0	10.2
c1c(c(ccc1CC=C)O)OC			
NSC_9268		108-68-9	10.2
c1c(cc(cc1C)C)O			
NSC_9885		622-62-8	10.1
c1c(ccc(c1)O)OCC			
NSC_10112		90-00-6	10.2
c1(ccccc1O)CC			
NSC_11215		89-83-8	10.6
c1(cc(c(cc1)C(C)C)O)C			
NSC_12987		5428-54-6	8.59
c1c(ccc(c1O)C)[N+](=O)[O-]			
NSC_14881		87-17-2	7.4
c1(ccccc1C(=O)Nc1ccccc1)O			

Table S1: List of molecules.

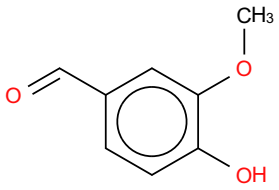
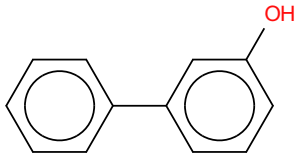
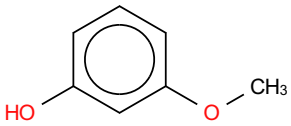
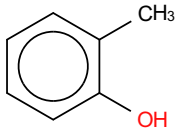
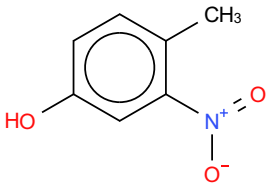
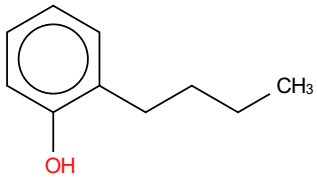
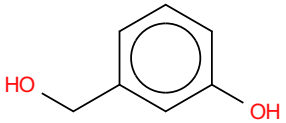
NSC number		**CAS number**	**pK_a**
smiles			
NSC_15351		121-33-5	7.4
O=Cc1cc(c(cc1)O)OC			
NSC_17588		580-51-8	9.64
c1ccccc1c1cccc(c1)O			
NSC_21735		150-19-6	9.65
c1c(cc(cc1)O)OC			
NSC_23076		95-48-7	10.3
c1(ccccc1O)C			
NSC_41205		2042-14-0	8.62
c1(ccc(cc1[N+](=O)[O-])O)C			
NSC_60291		3180-09-4	10.6
c1c(c(ccc1)O)CCCC			
NSC_60735		620-24-6	9.83
c1cc(cc(c1)CO)O			

Table S1: List of molecules.

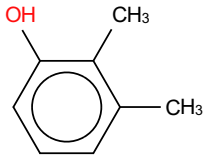
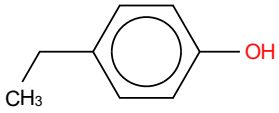
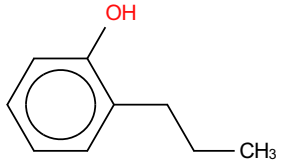
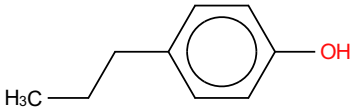
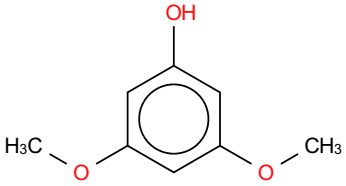
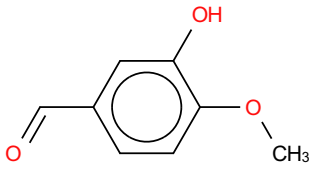
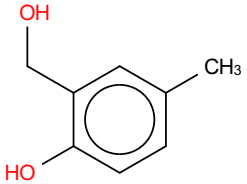
NSC number		**CAS number**	**pK_a**
smiles			
NSC_62011		526-75-0	10.5
c1(c(cccc1C)O)C			
NSC_62012		123-07-9	10
c1c(ccc(c1)CC)O			
NSC_65646		644-35-9	10.5
c1(ccccc1O)CCC			
NSC_65647		645-56-7	10.3
c1c(ccc(c1)CCC)O			
NSC_70955		500-99-2	9.34
c1c(cc(cc1O)OC)OC			
NSC_82996		621-59-0	8.89
O=Cc1ccc(c(c1)O)OC			
NSC_85475		4383-07-7	10.2
c1(cc(c(cc1)O)CO)C			

Table S1: List of molecules.

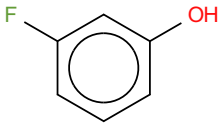
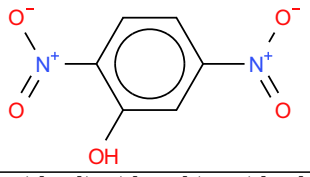
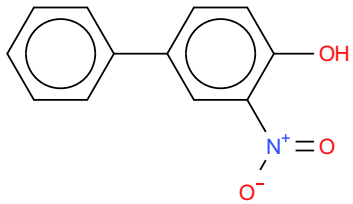
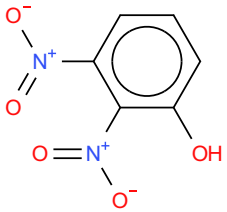
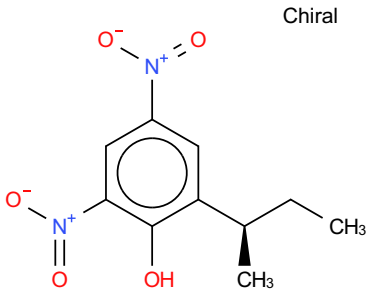
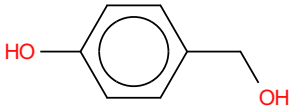
NSC number		CAS number	pK_a
smiles			
NSC_87078		372-20-3	9.21
Fc1cc(ccc1)O			
NSC_90441		329-71-5	5.21
c1cc(cc(c1[N+](=O)[O-])O)[N+](=O)[O-]			
NSC_95810		885-82-5	6.73
c1c(cccc1)c1cc(c(cc1)O)[N+](=O)[O-]			
NSC_156083		66-56-8	4.96
c1ccc(c(c1[N+](=O)[O-])[N+](=O)[O-])O			
NSC_202753		88-85-7	4.62
c1c(c(c(cc1[N+](=O)[O-])[N+](=O)[O-])O)[C@@H](CC)C			
NSC_227926		623-05-2	9.82
c1cc(ccc1O)CO			

Table S1: List of molecules.

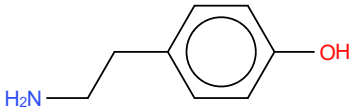
NSC number		**CAS number**	**pK_a**
smiles			
NSC_249188		51-67-2	9.77
c1(ccc(cc1)CCN)O			

Table S1: List of molecules.

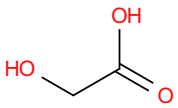
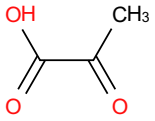
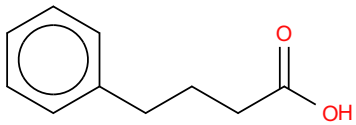
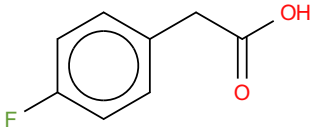
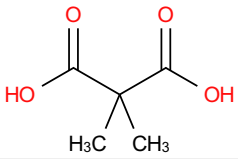
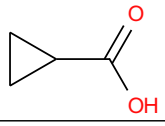
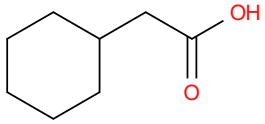
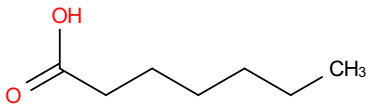
NSC number		CAS number	pK_a
smiles			
Carboxylic acids			
NSC_166		79-14-1	3.83
OCC(=O)O			
NSC_179		127-17-3	2.45
C(=O)(O)C(=O)C			
NSC_295		1821-12-1	4.76
c1ccccc1CCCC(=O)O			
NSC_402		405-50-5	4.24
Fc1ccc(cc1)CC(=O)O			
NSC_836		595-46-0	3.15
C(C(=O)O)(C)(C)C(=O)O			
NSC_1112		1759-53-1	4.83
C1C(C1)C(=O)O			
NSC_2159		5292-21-7	4.8
C1C(CCCC1)CC(=O)O			
NSC_2192		111-14-8	4.8
C(=O)(O)CCCCCC			

Table S1: List of molecules.

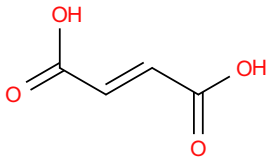
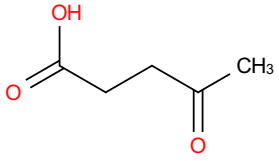
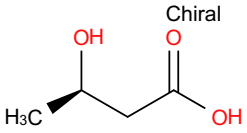
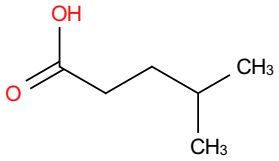
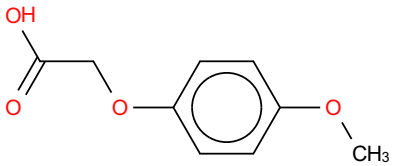
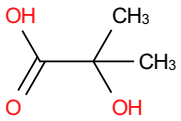
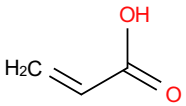
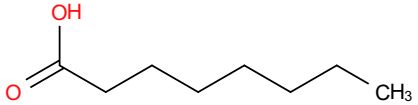
NSC number		**CAS number**	**pK_a**
smiles			
NSC_2752		110-17-8	3.03
C(=C\C(=O)O)/C(=O)O			
NSC_3716		123-76-2	4.64
C(=O)(O)CCC(=O)C			
NSC_3806		300-85-6	4.41
C[C@@H](O)CC(=O)O			
NSC_4126		646-07-1	4.84
C(=O)(O)CCC(C)C			
NSC_4255		1877-75-4	3.21
c1cc(ccc1OCC(=O)O)OC			
NSC_4505		594-61-6	3.61
C(C(=O)O)(O)(C)C			
NSC_4765		79-10-7	4.26
C(=C)C(=O)O			
NSC_5024		124-07-2	4.89
C(=O)(O)CCCCCCC			

Table S1: List of molecules.

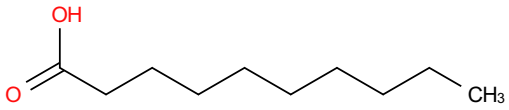
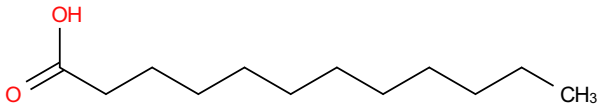
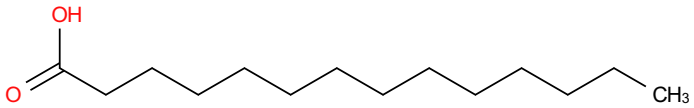
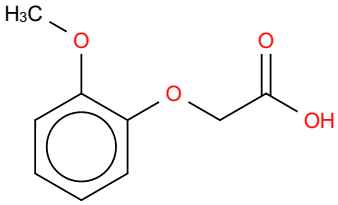
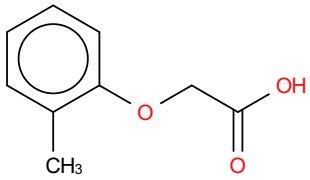
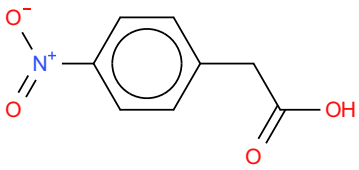
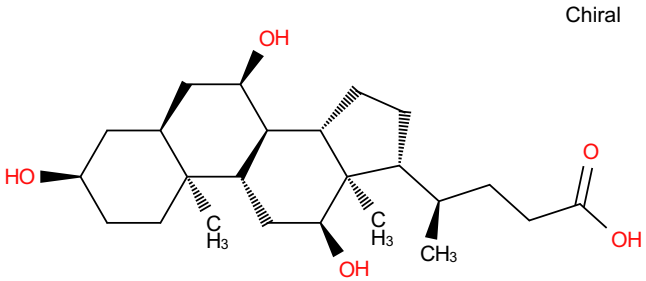
NSC number		**CAS number**	**pK_a**
smiles			
NSC_5025		334-48-5	4.9
C(=O)(O)CCCCCCCCC			
NSC_5026		143-07-7	5.3
C(=O)(O)CCCCCCCCCCC			
NSC_5028		544-63-8	4.9
C(=O)(O)CCCCCCCCCCCCC			
NSC_5165		1878-85-9	3.23
c1(ccccc1OC)OCC(=O)O			
NSC_5293		1878-49-5	3.23
c1c(c(ccc1)C)OCC(=O)O			
NSC_5398		104-03-0	3.85
c1cc(ccc1[N+](=O)[O-])CC(=O)O			
NSC_6135		81-25-4	4.98
[C@@H]([C@@H]1[C@]2([C@H](C[C@H]3[C@H]([C@@H]2CC1)[C@@H](C[C@H]1[C@@]3(CC[C@H](C1)O)C)O)O)C)(C)CCC(=O)O			

Table S1: List of molecules.

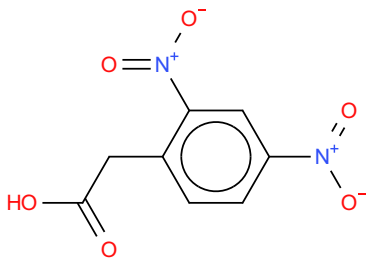
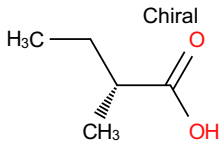
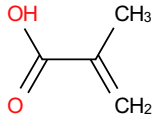
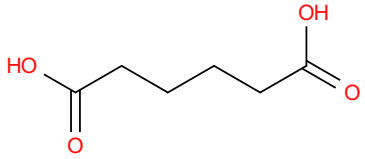
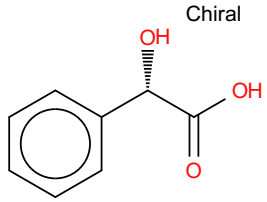
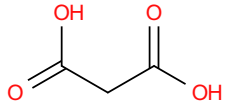
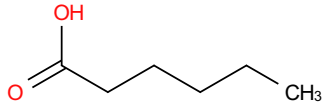
NSC number		**CAS number**	**pK_a**
smiles			
NSC_6225		643-43-6	3.5
c1(cc(ccc1CC(=O)O)[N+](=O)[O-])[N+](=O)[O-]			
NSC_7304		116-53-0	4.81
C(=O)([C@@H](CC)C)O			
NSC_7393		79-41-4	4.65
C(=O)(O)C(=C)C			
NSC_7622		124-04-9	4.44
C(=O)(CCCCC(=O)O)O			
NSC_7925		90-64-2	3.41
c1(ccccc1)[C@@H](C(=O)O)O			
NSC_8124		141-82-2	2.85
C(C(=O)O)C(=O)O			
NSC_8266		142-62-1	4.88
C(=O)(O)CCCCC			

Table S1: List of molecules.

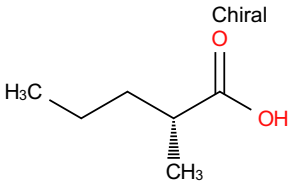
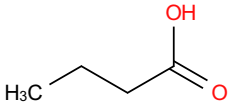
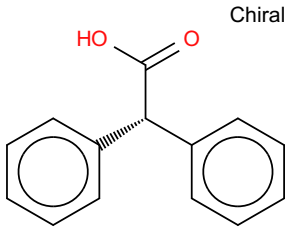
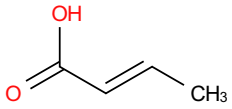
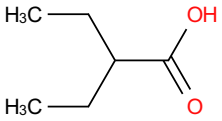
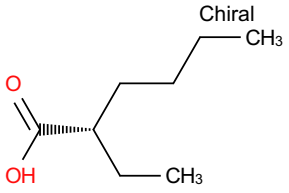
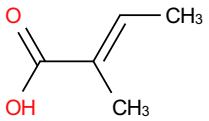
NSC number		**CAS number**	**pK_a**
smiles			
NSC_8406		97-61-0	4.79
C(=O)([C@@H](CCC)C)O			
NSC_8415		107-92-6	4.82
C(CC)C(=O)O			
NSC_8742		117-34-0	3.94
c1ccccc1[C@@H](c1ccccc1)C(=O)O			
NSC_8751		107-93-7	4.17
C(=C\C(=O)O)/C			
NSC_8758		88-09-5	4.71
C(CC)(CC)C(=O)O			
NSC_8881		149-57-5	4.7
C(=O)(O)[C@@H](CC)CCCC			
NSC_8999		80-59-1	4.96
C(=C(\C(=O)O)/C)/C			

Table S1: List of molecules.

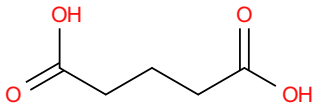
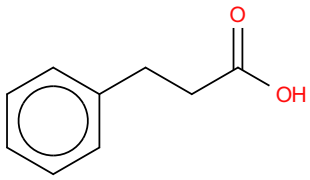
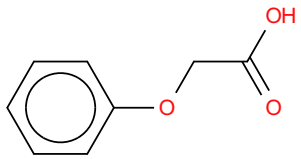
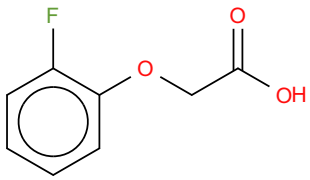
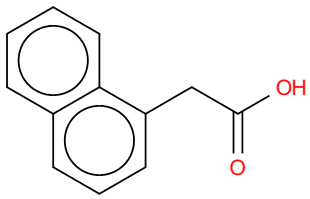
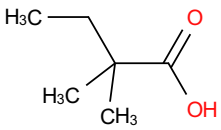
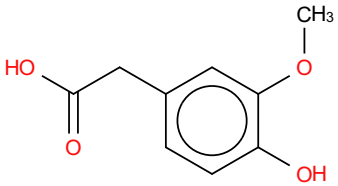
NSC number		**CAS number**	**pK_a**
smiles			
NSC_9238		110-94-1	4.34
C(=O)(O)CCCC(=O)O			
NSC_9272		501-52-0	4.66
c1(ccccc1)CCC(=O)O			
NSC_9810		122-59-8	3.17
c1ccccc1OCC(=O)O			
NSC_10232		348-10-7	3.08
Fc1c(cccc1)OCC(=O)O			
NSC_15772		86-87-3	4.23
c1cccc2c1c(ccc2)CC(=O)O			
NSC_16045		595-37-9	5.03
C(C)C(C)(C)C(=O)O			
NSC_16682		306-08-1	4.41
c1c(c(ccc1CC(=O)O)O)OC			

Table S1: List of molecules.

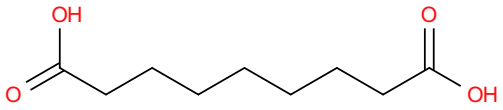
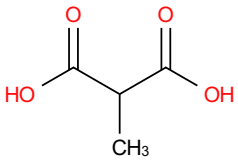
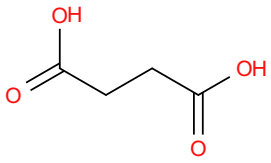
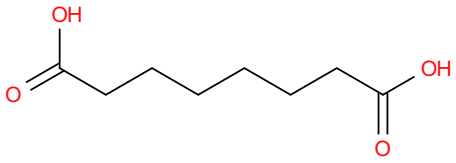
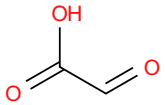
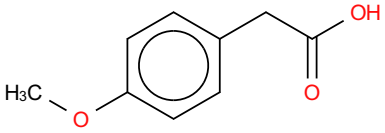
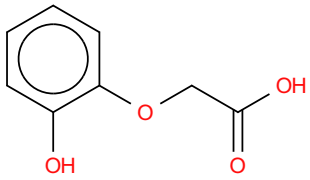
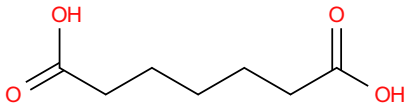
NSC number		**CAS number**	**pK_a**
smiles			
NSC_19493		123-99-9	4.55
O=C(O)CCCCCCCC(=O)O			
NSC_25201		516-05-2	3.12
C(=O)(O)C(C(=O)O)C			
NSC_25949		110-15-6	4.21
C(=O)(O)CCC(=O)O			
NSC_25952		505-48-6	4.52
C(=O)(O)CCCCCCC(=O)O			
NSC_27785		298-12-4	3.3
C(=O)(O)C=O			
NSC_27799		104-01-8	4.36
c1cc(ccc1OC)CC(=O)O			
NSC_30092		6324-11-4	3.02
c1c(c(ccc1)O)OCC(=O)O			
NSC_30112		111-16-0	4.51
C(=O)(O)CCCCCC(=O)O			

Table S1: List of molecules.

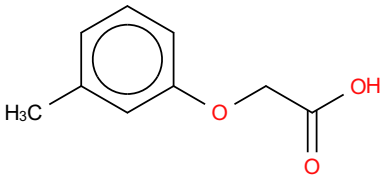
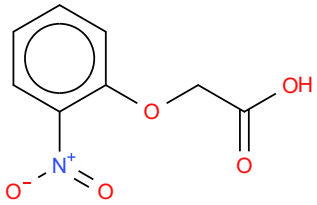
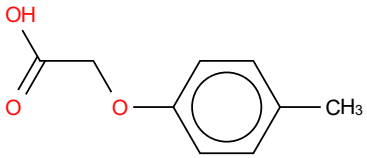
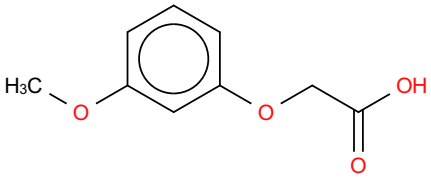
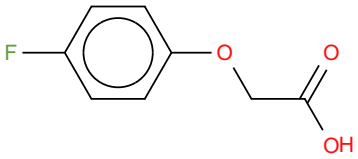
NSC number		**CAS number**	**pK_a**
smiles			
NSC_35913		1643-15-8	3.2
c1c(cc(cc1)C)OCC(=O)O			
NSC_37409		1878-87-1	2.9
c1c(c(ccc1)[N+](=O)[O-])OCC(=O)O			
NSC_38176		940-64-7	3.21
c1cc(ccc1OCC(=O)O)C			
NSC_38177		2088-24-6	3.14
c1c(cc(cc1)OC)OCC(=O)O			
NSC_49589		405-79-8	3.13
c1c(ccc(c1)F)OCC(=O)O			

Table S1: List of molecules.

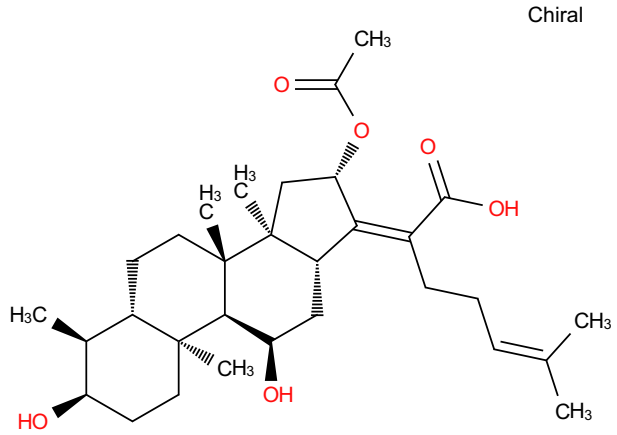
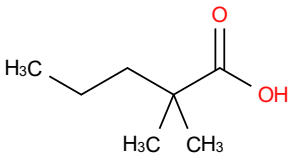
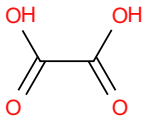
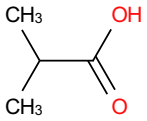
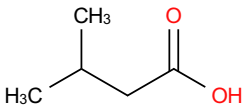
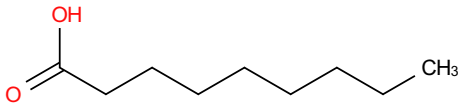
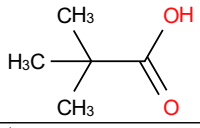
NSC number		**CAS number**	**pK_a**
smiles			
NSC_56192		6990-06-3	5.35
C(=C\1/[C@H](C[C@]2([C@H]1C[C@H]([C@@H]1[C@@]2(CC[C@@H]2[C@@]1(CC[C@H]([C@H]2C)O)C)C)O)C)OC(=O)C)(\CCC=C(C)C)/C(=O)O			
NSC_61983		1185-39-3	5.02
C(CC)C(C)(C)C(=O)O			
NSC_62774		144-62-7	1.25
C(=O)(O)C(=O)O			
NSC_62780		79-31-2	4.84
C(C)(C)C(=O)O			
NSC_62783		503-74-2	4.77
C(C)(C)CC(=O)O			
NSC_62787		112-05-0	4.95
C(=O)(O)CCCCCCCC			
NSC_65449		75-98-9	5.03
C(C)(C)(C)C(=O)O			

Table S1: List of molecules.

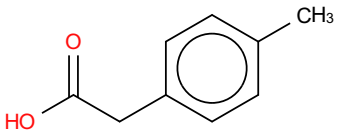
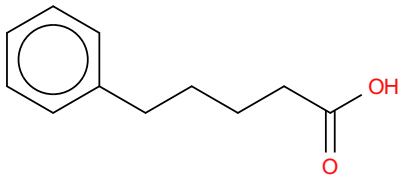
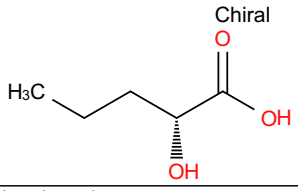
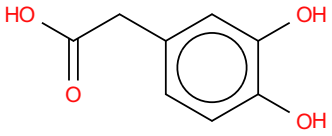
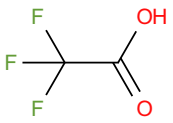
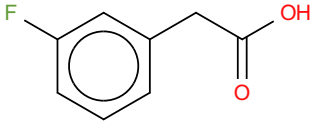
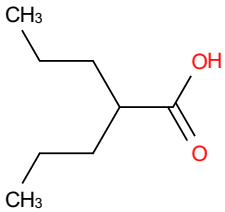
NSC number		**CAS number**	**pK_a**
smiles			
NSC_65595		622-47-9	4.37
c1cc(ccc1CC(=O)O)C			
NSC_65637		2270-20-4	4.88
c1ccccc1CCCCC(=O)O			
NSC_67957		617-31-2	3.59
C(CC)[C@@H](O)C(=O)O			
NSC_73191		102-32-9	4.25
c1c(c(ccc1CC(=O)O)O)O			
NSC_77366		76-05-1	0.52
C(F)(F)(F)C(=O)O			
NSC_88344		331-25-9	4.13
Fc1cc(ccc1)CC(=O)O			
NSC_93819		99-66-1	4.6
C(CCC)(CCC)C(=O)O			

Table S1: List of molecules.

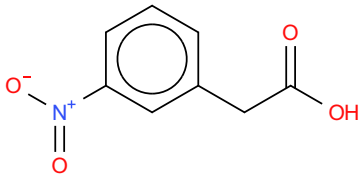
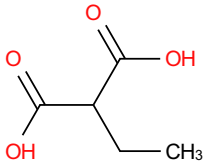
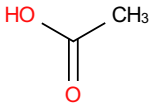
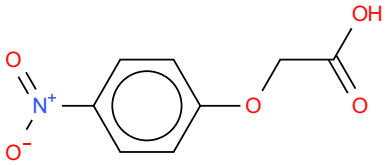
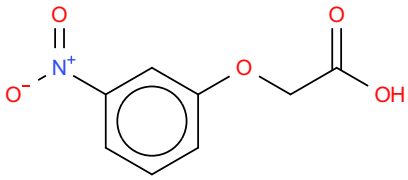
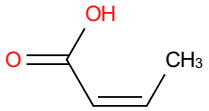
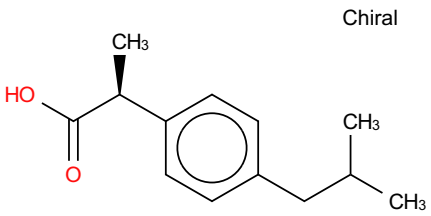
NSC number		**CAS number**	**pK_a**
smiles			
NSC_93911		1877-73-2	3.97
c1c(cc(cc1)[N+](=O)[O-])CC(=O)O			
NSC_96615		601-75-2	2.96
C(C)C(C(=O)O)C(=O)O			
NSC_132953		64-19-7	4.76
C(=O)(O)C			
NSC_166278		1798-11-4	2.89
O(c1ccc(cc1)[N+](=O)[O-])CC(=O)O			
NSC_193418		1878-88-2	2.95
O(c1cccc(c1)[N+](=O)[O-])CC(=O)O			
NSC_206946		3724-65-0	4.17
C(=C\C(=O)O)\C			
NSC_256857		15687-27-1	4.91
c1(ccc(cc1)[C@@H](C(=O)O)C)CC(C)C			

Table S1: List of molecules.

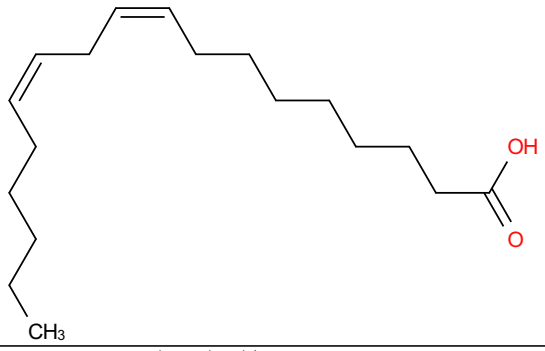
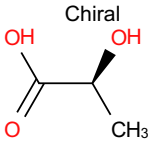
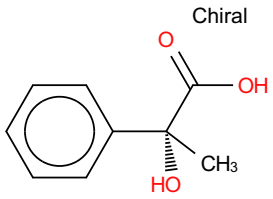
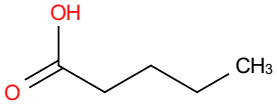
NSC number		**CAS number**	**pK_a**
smiles			
NSC_281243		60-33-3	4.77
C(=C\C/C=C\CCCCCCCC(=O)O)\CCCCC			
NSC_367919		598-82-3	3.86
[C@@H](C(=O)O)(C)O			
NSC_401846		515-30-0	3.53
c1ccccc1[C@@](C)(O)C(=O)O			
NSC_406833		109-52-4	4.84
C(=O)(O)CCCC			

Table S1: List of molecules.

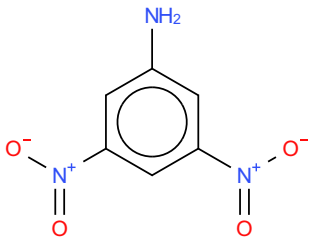
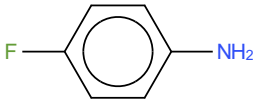
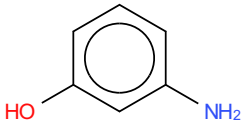
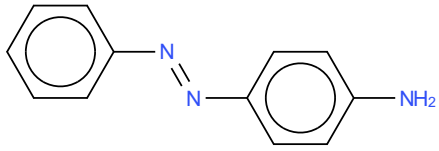
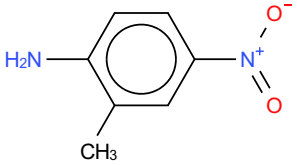
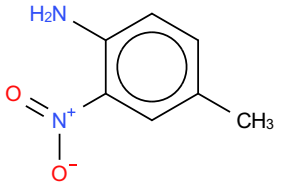
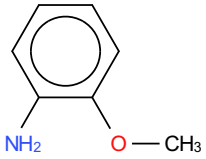
NSC number		CAS number	pK_a
smiles			
Anilines			
NSC_284		618-87-1	0.3
c1c(cc(cc1N)[N+](=O)[O-])[N+](=O)[O-]			
NSC_579		371-40-4	4.65
Fc1ccc(cc1)N			
NSC_1546		591-27-5	4.37
c1ccc(cc1O)N			
NSC_2032		60-09-3	2.82
N(=N\c1ccccc1)/c1ccc(cc1)N			
NSC_2075		99-52-5	1.04
c1cc(cc(c1N)C)[N+](=O)[O-]			
NSC_2759		89-62-3	0.4
c1cc(cc(c1N)[N+](=O)[O-])C			
NSC_3122		90-04-0	4.53
c1c(c(ccc1)N)OC			

Table S1: List of molecules.

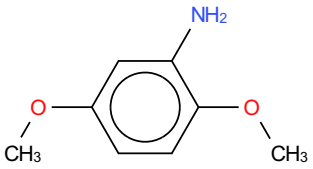
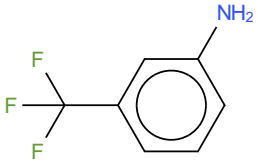
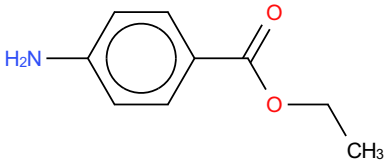
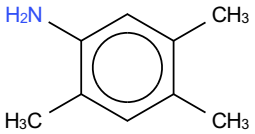
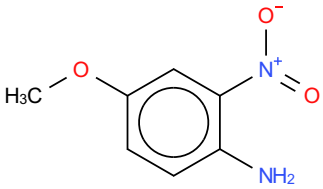
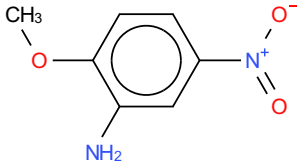
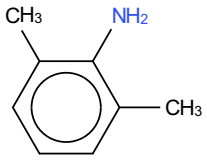
NSC number		**CAS number**	**pK_a**
smiles			
NSC_4138		102-56-7	3.93
c1(ccc(cc1N)OC)OC			
NSC_4540		98-16-8	3.49
c1c(cccc1C(F)(F)F)N			
NSC_4688		94-09-7	2.51
c1c(ccc(c1)N)C(=O)OCC			
NSC_5297		137-17-7	5.09
c1c(c(cc(c1N)C)C)C			
NSC_5509		96-96-8	0.77
c1c(c(ccc1OC)N)[N+](=O)[O-]			
NSC_5510		99-59-2	2.49
c1cc(cc(c1OC)N)[N+](=O)[O-]			
NSC_7098		87-62-7	3.95
c1(cccc(c1N)C)C			

Table S1: List of molecules.

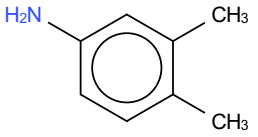
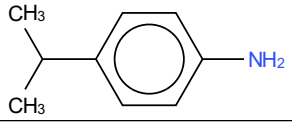
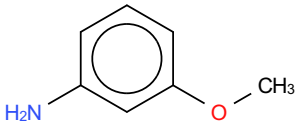
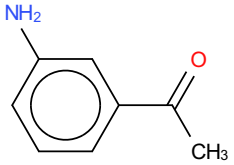
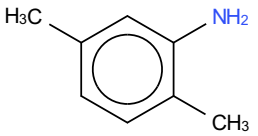
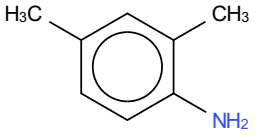
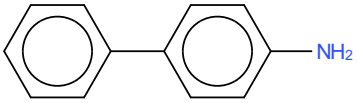
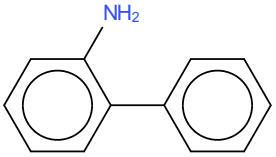
NSC number		**CAS number**	**pK_a**
smiles			
NSC_7099		95-64-7	5.28
c1c(c(ccc1N)C)C			
NSC_7198		99-88-7	4.85
c1cc(ccc1C(C)C)N			
NSC_7631		536-90-3	4.24
c1c(cc(cc1)N)OC			
NSC_7637		99-03-6	3.56
c1c(cccc1N)C(=O)C			
NSC_7639		95-78-3	4.53
c1(ccc(cc1N)C)C			
NSC_7640		95-68-1	4.89
c1(cc(ccc1N)C)C			
NSC_7660		92-67-1	4.35
c1ccccc1c1ccc(cc1)N			
NSC_7661		90-41-5	3.83
c1ccc(c(c1)N)c1ccccc1			

Table S1: List of molecules.

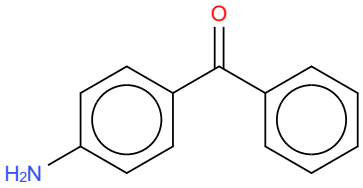
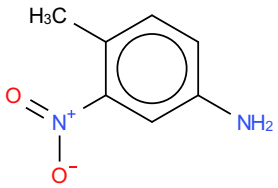
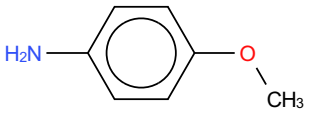
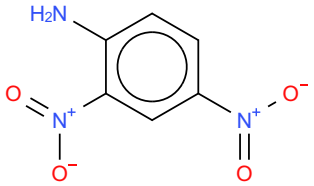
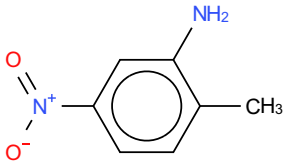
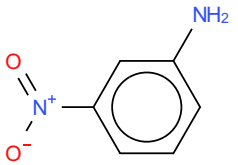
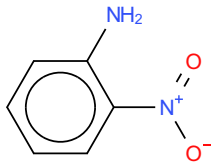
NSC number		**CAS number**	**pK_a**
smiles			
NSC_7665		1137-41-3	2.24
O=C(c1ccc(cc1)N)c1ccccc1			
NSC_7731		119-32-4	3.03
c1cc(cc(c1C)[N+](=O)[O-])N			
NSC_7921		104-94-9	5.34
c1cc(ccc1N)OC			
NSC_8731		97-02-9	-4.25
c1cc(cc(c1N)[N+](=O)[O-])[N+](=O)[O-]			
NSC_8947		99-55-8	2.35
c1(cc(ccc1C)[N+](=O)[O-])N			
NSC_9574		99-09-2	2.47
c1(cc(ccc1)[N+](=O)[O-])N			
NSC_9796		88-74-4	-0.28
c1(ccccc1N)[N+](=O)[O-]			

Table S1: List of molecules.

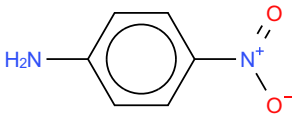
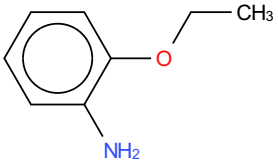
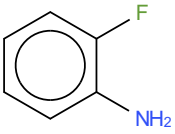
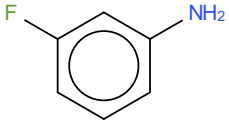
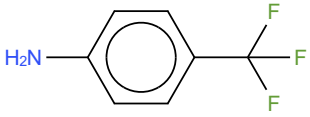
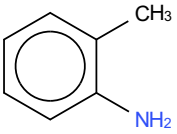
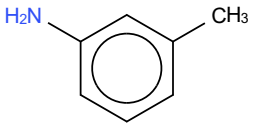
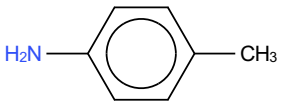
NSC number		CAS number	pK_a
smiles			
NSC_9797		100-01-6	1
c1cc(ccc1N)[N+](=O)[O-]			
NSC_9818		94-70-2	4.43
c1c(c(ccc1)N)OCC			
NSC_10299		348-54-9	3.2
Fc1ccccc1N			
NSC_10300		372-19-0	3.5
Fc1cc(ccc1)N			
NSC_10337		455-14-1	2.45
c1cc(ccc1N)C(F)(F)F			
NSC_15348		95-53-4	4.44
c1(ccccc1N)C			
NSC_15349		108-44-1	4.69
c1c(cccc1N)C			
NSC_15350		106-49-0	5.1
c1cc(ccc1N)C			

Table S1: List of molecules.

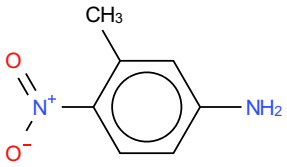
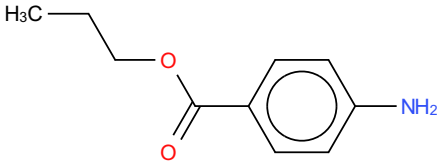
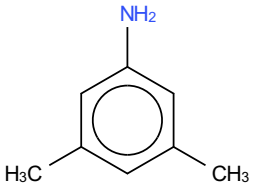
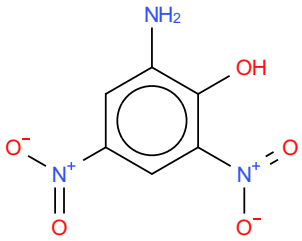
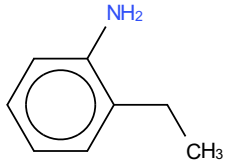
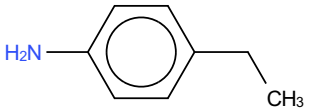
NSC number		**CAS number**	**pK_a**
smiles			
NSC_17041		611-05-2	1.64
c1c(ccc(c1C)[N+](=O)[O-])N			
NSC_23516		94-12-2	2.49
c1cc(ccc1C(=O)OCCC)N			
NSC_26880		108-69-0	4.79
c1c(cc(cc1N)C)C			
NSC_36939		96-91-3	1
c1(c(cc(cc1[N+](=O)[O-])[N+](=O)[O-])N)O			
NSC_62014		578-54-1	4.3
c1(ccccc1N)CC			
NSC_62015		589-16-2	5
c1cc(ccc1N)CC			

Table S1: List of molecules.

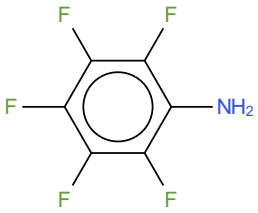
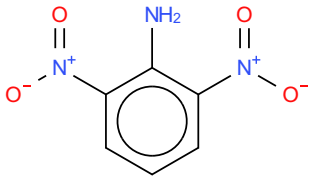
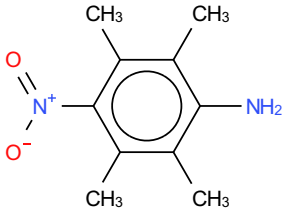
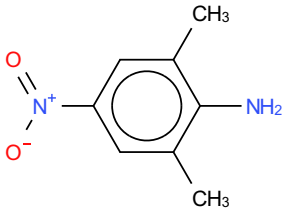
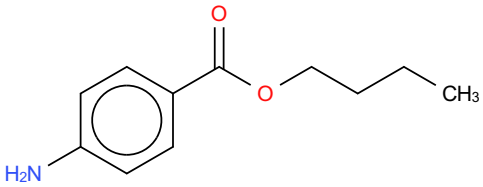
NSC number		**CAS number**	**pK_a**
smiles			
NSC_88320		771-60-8	-0.28
Fc1c(F)c(c(F)c(F)c1F)N			
NSC_93399		606-22-4	-5
c1(c(cccc1[N+](=O)[O-])[N+](=O)[O-])N			
NSC_97134		13171-61-4	2.36
c1(c(c(c(c(c1C)[N+](=O)[O-])C)C)N)C			
NSC_101580		16947-63-0	0.98
c1(c(c(cc(c1)[N+](=O)[O-])C)N)C			
NSC_128464		94-25-7	2.47
O(C(=O)c1ccc(cc1)N)CCCC			

Table S1: List of molecules.

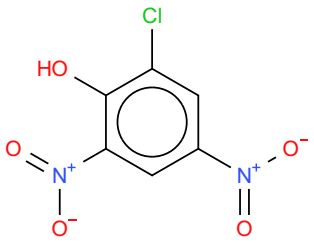
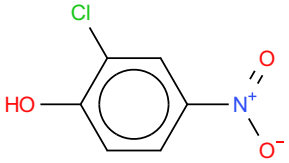
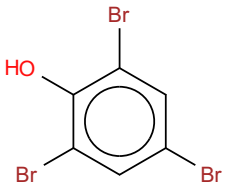
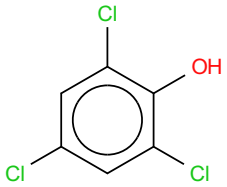
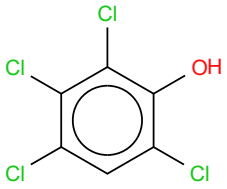
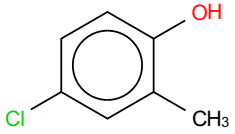
NSC number		**CAS number**	**pK_a**
smiles			
Test set of phenols			
NSC_3		946-31-6	2.1
Clc1cc(cc(c1O)[N+](=O)[O-])[N+](=O)[O-]			
NSC_1316		619-08-9	5.45
Clc1cc(ccc1O)[N+](=O)[O-]			
NSC_2136		118-79-6	6.8
c1(cc(cc(c1O)Br)Br)Br			
NSC_2165		88-06-2	6.23
Clc1cc(Cl)cc(Cl)c1O			
NSC_2428		58-90-2	5.22
Clc1c(c(Cl)cc(Cl)c1Cl)O			
NSC_2851		1570-64-5	9.71
c1(c(cc(cc1)Cl)C)O			

Table S1: List of molecules.

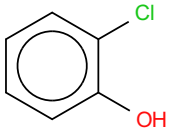
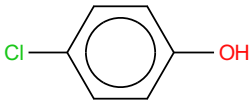
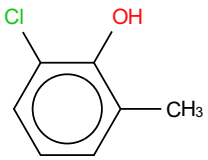
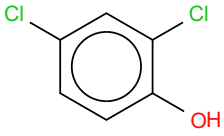
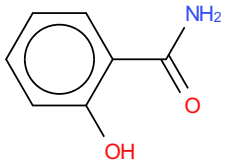
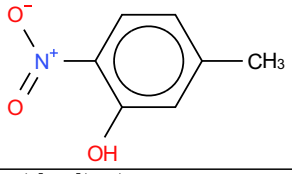
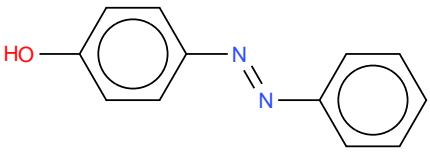
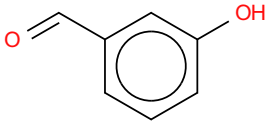
NSC number		**CAS number**	**pK_a**
smiles			
NSC_2870		95-57-8	8.56
Clc1ccccc1O			
NSC_2877		106-48-9	9.41
Clc1ccc(cc1)O			
NSC_2878		87-64-9	8.69
Clc1c(c(ccc1)C)O			
NSC_2879		120-83-2	7.89
Clc1cc(Cl)ccc1O			
NSC_3115		65-45-2	8.37
c1c(c(ccc1)O)C(=O)N			
NSC_3142		700-38-9	7.41
c1cc(cc(c1[N+](=O)[O-])O)C			
NSC_3177		1689-82-3	8.2
c1cc(ccc1O)/N=N/c1ccccc1			
NSC_3504		100-83-4	8.98
O=Cc1cccc(c1)O			

Table S1: List of molecules.

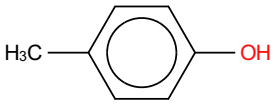
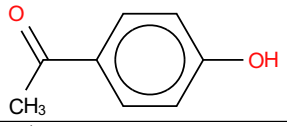
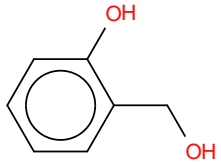
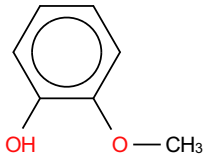
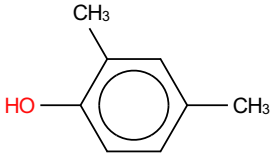
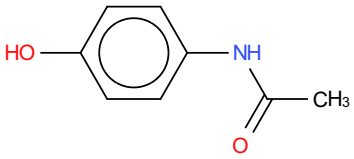
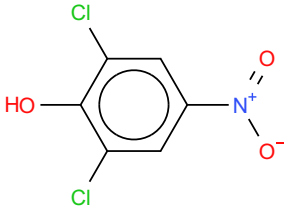
NSC number		**CAS number**	**pK_a**
smiles			
NSC_3696		106-44-5	10.3
c1c(ccc(c1)C)O			
NSC_3698		99-93-4	8.05
c1c(ccc(c1)C(=O)C)O			
NSC_3814		90-01-7	9.84
c1c(c(ccc1)CO)O			
NSC_3815		90-05-1	9.98
c1c(c(ccc1)O)OC			
NSC_3829		105-67-9	10.6
c1c(ccc(c1C)O)C			
NSC_3991		103-90-2	9.38
c1c(ccc(c1)O)NC(=O)C			
NSC_4123		618-80-4	3.55
Clc1cc(cc(Cl)c1O)[N+](=O)[O-]			

Table S1: List of molecules.

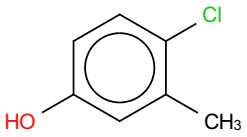
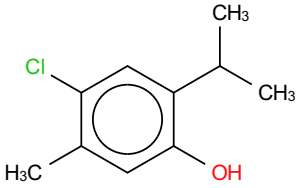
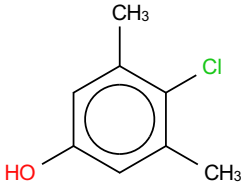
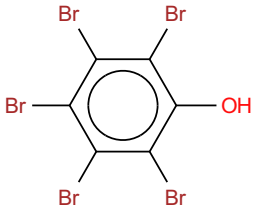
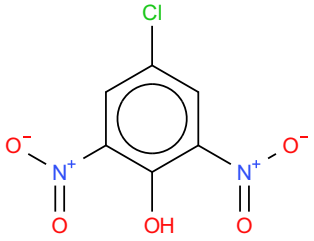
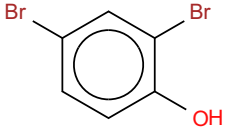
NSC number		**CAS number**	**pK_a**
smiles			
NSC_4166		59-50-7	9.55
Clc1c(cc(cc1)O)C			
NSC_4964		89-68-9	9.98
Clc1cc(c(cc1C)O)C(C)C			
NSC_4970		106-41-2	9.17
Brc1ccc(cc1)O			
NSC_4971		88-04-0	9.7
Clc1c(cc(cc1C)O)C			
NSC_5717		608-71-9	4.62
Brc1c(Br)c(c(Br)c(Br)c1Br)O			
NSC_6212		88-87-9	2.96
Clc1cc(c(c(c1)[N+](=O)[O-])O)[N+](=O)[O-]			
NSC_6213		615-58-7	7.79
Brc1cc(Br)ccc1O			

Table S1: List of molecules.

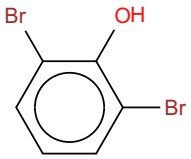
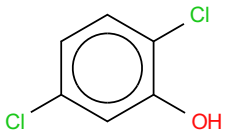
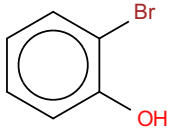
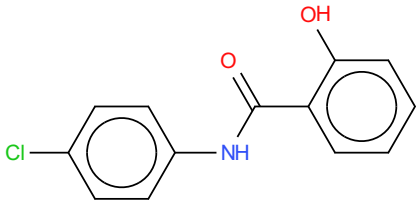
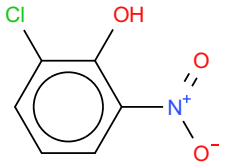
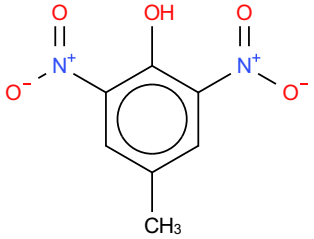
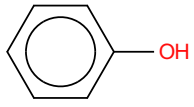
NSC number		**CAS number**	**pK_a**
smiles			
NSC_6214		608-33-3	6.67
Brc1c(c(Br)ccc1)O			
NSC_6296		583-78-8	7.51
Clc1c(cc(Cl)cc1)O			
NSC_6970		95-56-7	8.45
Brc1ccccc1O			
NSC_22902		3679-63-8	7.3
Clc1ccc(cc1)NC(=O)c1ccccc1O			
NSC_28581		603-86-1	5.48
Clc1c(c(ccc1)[N+](=O)[O-])O			
NSC_33870		609-93-8	4.23
c1(c(cc(cc1[N+](=O)[O-])C)[N+](=O)[O-])O			
NSC_36808		108-95-2	9.99
c1ccccc1O			

Table S1: List of molecules.

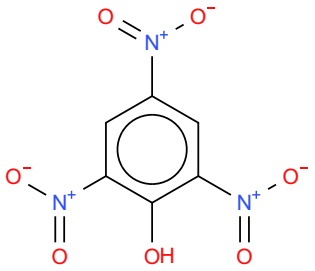
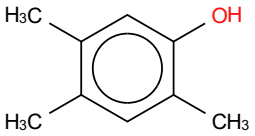
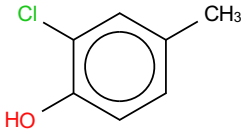
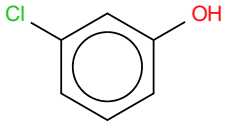
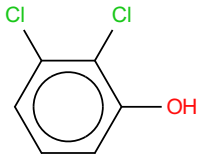
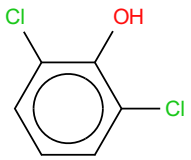
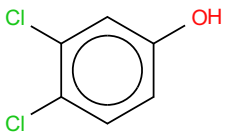
NSC number		**CAS number**	**pK_a**
smiles			
NSC_36947		88-89-1	0.38
c1c(c(c(cc1[N+](=O)[O-])[N+](=O)[O-])O)[N+](=O)[O-]			
NSC_38776		496-78-6	10.6
c1c(c(cc(c1C)C)C)O			
NSC_48848		6640-27-3	8.74
Clc1cc(ccc1O)C			
NSC_59700		108-43-0	9.12
Clc1cc(ccc1)O			
NSC_60646		576-24-9	7.7
Clc1c(Cl)c(ccc1)O			
NSC_60647		87-65-0	6.79
Clc1c(c(Cl)ccc1)O			
NSC_60648		95-77-2	8.63
Clc1cc(ccc1Cl)O			

Table S1: List of molecules.

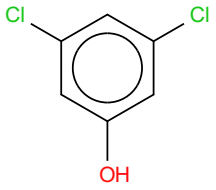
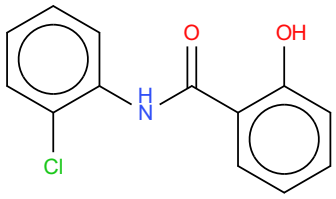
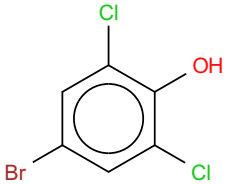
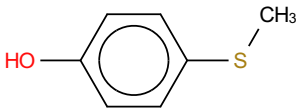
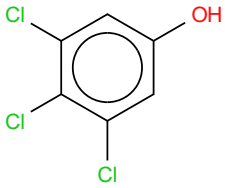
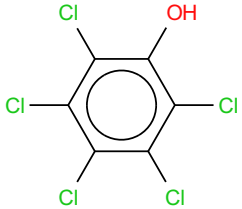
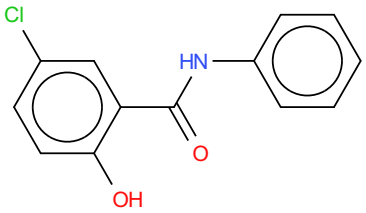
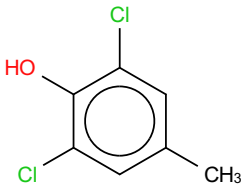
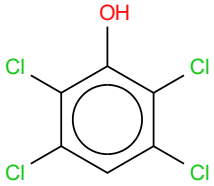
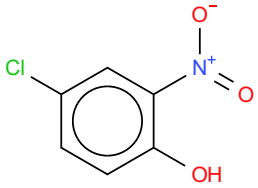
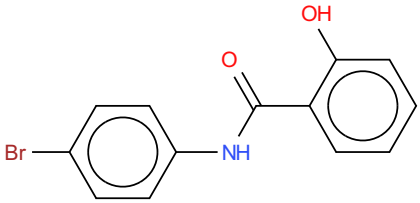
NSC number		**CAS number**	**pK_a**
smiles			
NSC_60649		591-35-5	8.18
Clc1cc(cc(Cl)c1)O			
NSC_63709		1697-18-3	7.31
Clc1c(cccc1)NC(=O)c1ccccc1O			
NSC_74624		3217-15-0	6.21
Clc1cc(Br)cc(Cl)c1O			
NSC_75840		1073-72-9	9.53
S(c1ccc(cc1)O)C			
NSC_243667		609-19-8	7.84
c1(cc(Cl)c(c(c1)Cl)Cl)O			
NSC_263497		87-86-5	4.7
Clc1c(c(c(c(c1Cl)Cl)Cl)Cl)O			

Table S1: List of molecules.

NSC number		**CAS number**	**pK_a**
smiles			
NSC_402600		4638-48-6	6.17
Clc1cc(c(cc1)O)C(=O)Nc1ccccc1			
NSC_407750		2432-12-4	7.19
c1(cc(cc(c1O)Cl)C)Cl			
NSC_407823		935-95-5	5.14
Clc1c(c(Cl)c(Cl)cc1Cl)O			
NSC_520345		89-64-5	6.46
Clc1cc(c(cc1)O)[N+](=O)[O-]			
NSC_526279		2627-77-2	7.31
Brc1ccc(cc1)NC(=O)c1ccccc1O			

2 Summary of Correlations Between Calculated and Experimental pK_a for PM6 Models

Table S2: R^2 describing the correlation between calculated and experimental pK_a for semiempirical QSPR models.

R^2		Class of molecules	Phenols	Carboxylic acids	Anilines	Average
		Charge calculation approach	PM6, MPA	PM6, MPA	PM6, MPA	
Source + Optimization	Balloon	none	0.900	0.854	0.896	0.884
		MM	0.889	0.877	0.935	0.900
		QM	0.886	0.895	0.933	0.904
	Frog2	none	0.939	0.886	0.919	0.914
		MM	0.942	0.931	0.961	0.945
		QM	0.959	0.927	0.890	0.926
	NCI	none	0.950	0.941	0.966	0.952
		MM	0.962	0.936	0.956	0.951
		QM	0.878	0.932	0.940	0.917
	OpenBabel	none	0.949	0.919	0.964	0.944
		MM	0.959	0.920	0.957	0.945
		QM	0.952	0.901	0.887	0.913
	PubChem	none	0.934	0.933	0.952	0.940
		MM	0.915	0.926	0.955	0.932
		QM	0.929	0.919	0.884	0.911
	RDKit	none	0.868	0.839	0.883	0.863
		MM-UFF	0.952	0.895	0.970	0.939
		MM	0.931	0.863	0.957	0.917
		QM	0.940	0.838	0.930	0.903
Average			0.928	0.902	0.933	

Legend	$R^2 \geq 0.95$	$R^2 \geq 0.9$	$R^2 \geq 0.866$	$R^2 \geq 0.833$

3 Quality and Statistical Criteria of All the QSPR Models

Table S3**a**: Quality and statistical criteria of QM and EEM QSPR models.

Charges	Theory level	Basis set	Charge scheme	3D source	Opt	R^2	RMSE	$\overline{\Delta}$	s	F	# molecules
Phenols											
QM	B3LYP	6-31G*	AIM	NCI	NO	0.973	0.341	0.256	0.359	391.9	60
QM	B3LYP	6-31G*	MPA	NCI	NO	0.971	0.353	0.285	0.372	363.2	60
QM	HF	STO-3G	MPA	FROG2	QM	0.969	0.366	0.269	0.386	337.7	60
QM	HF	STO-3G	MPA	FROG2	MM	0.967	0.377	0.294	0.397	317.6	60
QM	B3LYP	6-31G*	MPA	OPENBABEL	MM	0.965	0.389	0.268	0.410	297.3	60
QM	B3LYP	6-31G*	MPA	FROG2	QM	0.963	0.400	0.312	0.422	281.0	60
QM	HF	STO-3G	MPA	PUBCHEM	MM	0.963	0.402	0.307	0.423	278.6	60
QM	B3LYP	6-31G*	MPA	NCI	MM	0.963	0.402	0.287	0.424	277.5	60
QM	B3LYP	6-31G*	AIM	OPENBABEL	NO	0.963	0.403	0.320	0.424	277.3	60
QM	B3LYP	6-31G*	AIM	OPENBABEL	MM	0.961	0.409	0.324	0.431	268.3	60
QM	B3LYP	6-31G*	MPA	RDKIT	MM	0.961	0.411	0.304	0.433	266.2	60
QM	B3LYP	6-31G*	MPA	OPENBABEL	NO	0.961	0.412	0.308	0.434	264.1	60
QM	HF	STO-3G	MPA	OPENBABEL	MM	0.961	0.413	0.310	0.435	263.1	60
QM	B3LYP	6-31G*	NPA	NCI	NO	0.960	0.415	0.333	0.437	260.7	60
QM	HF	STO-3G	MPA	PUBCHEM	NO	0.960	0.417	0.317	0.440	257.3	60
QM	B3LYP	6-31G*	NPA	NCI	MM	0.959	0.419	0.332	0.442	255.0	60
QM	B3LYP	6-31G*	NPA	OPENBABEL	MM	0.959	0.420	0.329	0.443	253.7	60
QM	HF	STO-3G	MPA	NCI	MM	0.958	0.427	0.323	0.450	245.2	60
QM	B3LYP	6-31G*	NPA	OPENBABEL	NO	0.957	0.430	0.335	0.454	241.1	60
QM	B3LYP	6-31G*	MPA	OPENBABEL	QM	0.957	0.432	0.319	0.455	239.7	60
QM	B3LYP	6-31G*	NPA	OPENBABEL	QM	0.956	0.436	0.348	0.459	235.1	60
QM	HF	STO-3G	MPA	OPENBABEL	QM	0.955	0.442	0.317	0.466	228.4	60
QM	HF	STO-3G	MPA	OPENBABEL	NO	0.955	0.443	0.338	0.467	226.9	60
QM	B3LYP	6-31G*	NPA	FROG2	QM	0.953	0.451	0.323	0.476	218.5	60
QM	B3LYP	6-31G*	MPA	PUBCHEM	NO	0.950	0.466	0.320	0.491	204.1	60
QM	B3LYP	6-31G*	AIM	RDKIT	MM	0.950	0.467	0.374	0.492	203.5	60
QM	HF	STO-3G	MPA	RDKIT	MM	0.947	0.478	0.354	0.504	193.5	60
QM	HF	STO-3G	MPA	NCI	NO	0.947	0.480	0.354	0.506	192.2	60
QM	B3LYP	6-31G*	MPA	RDKIT	QM	0.944	0.492	0.369	0.519	181.8	60
QM	HF	STO-3G	MPA	PUBCHEM	QM	0.943	0.494	0.359	0.521	180.2	60
QM	B3LYP	6-31G*	NPA	RDKIT	MM	0.941	0.504	0.412	0.531	173.4	60
QM	B3LYP	6-31G*	AIM	FROG2	QM	0.939	0.514	0.372	0.542	165.9	60
QM	B3LYP	6-31G*	MPA	BALLOON	NO	0.939	0.514	0.367	0.542	165.8	60
QM	B3LYP	6-31G*	AIM	FROG2	MM	0.938	0.518	0.386	0.546	163.3	60
QM	B3LYP	6-31G*	MPA	PUBCHEM	QM	0.936	0.525	0.380	0.553	158.8	60
QM	B3LYP	6-31G*	AIM	OPENBABEL	QM	0.936	0.525	0.400	0.554	158.4	60
QM	B3LYP	6-31G*	AIM	NCI	MM	0.936	0.527	0.389	0.556	157.3	60
QM	HF	STO-3G	MPA	RDKIT	QM	0.935	0.529	0.376	0.557	156.2	60
QM	B3LYP	6-31G*	NPA	PUBCHEM	NO	0.935	0.529	0.401	0.558	156.0	60

Table S3**a**: Quality and statistical criteria of QM and EEM QSPR models.

Charges	**Theory level**	**Basis set**	**Charge scheme**	**3D source**	**Opt**	R^2	**RMSE**	$\overline{\Delta}$	s	F	**# molecules**
QM	B3LYP	6-31G*	MPA	NCI	QM	0.935	0.531	0.390	0.560	154.6	60
QM	B3LYP	6-31G*	NPA	RDKIT	MM	0.934	0.532	0.408	0.561	154.0	60
QM	B3LYP	6-31G*	AIM	RDKIT	MM	0.934	0.534	0.407	0.563	152.8	60
QM	B3LYP	6-31G*	NPA	BALLOON	MM	0.933	0.536	0.422	0.565	151.5	60
QM	B3LYP	6-31G*	NPA	RDKIT	QM	0.933	0.539	0.423	0.569	149.7	60
QM	B3LYP	6-31G*	MPA	FROG2	MM	0.931	0.545	0.408	0.575	146.1	60
QM	HF	STO-3G	MPA	RDKIT	MM	0.931	0.548	0.406	0.577	144.8	60
QM	B3LYP	6-31G*	NPA	PUBCHEM	MM	0.927	0.561	0.437	0.592	137.4	60
Chavez2006	B3LYP	6-31G*	MPA	NCI	MM	0.926	0.565	0.436	0.595	135.6	60
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	MM	0.926	0.566	0.441	0.596	135.2	60
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	QM	0.925	0.570	0.444	0.601	132.9	60
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	QM	0.925	0.571	0.440	0.601	132.7	60
Chavez2006	B3LYP	6-31G*	MPA	NCI	QM	0.924	0.572	0.443	0.603	131.7	60
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	MM	0.923	0.579	0.422	0.610	128.6	60
QM	B3LYP	6-31G*	AIM	RDKIT	QM	0.922	0.579	0.427	0.610	128.4	60
QM	B3LYP	6-31G*	NPA	PUBCHEM	QM	0.922	0.579	0.418	0.611	128.4	60
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	QM	0.921	0.585	0.454	0.616	125.9	60
Svob2007-chal2	HF	STO-3G	MPA	FROG2	QM	0.921	0.585	0.418	0.616	125.8	60
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	NO	0.920	0.588	0.468	0.620	124.1	60
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	MM	0.919	0.591	0.467	0.623	122.9	60
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	QM	0.919	0.593	0.470	0.625	121.9	60
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	NO	0.918	0.594	0.457	0.626	121.6	60
Svob2007-chal2	HF	STO-3G	MPA	FROG2	MM	0.918	0.594	0.437	0.626	121.5	60
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	QM	0.918	0.596	0.472	0.629	120.5	60
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	MM	0.917	0.598	0.476	0.630	119.8	60
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	MM	0.917	0.599	0.467	0.631	119.5	60
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	QM	0.917	0.599	0.476	0.631	119.5	60
QM	HF	STO-3G	MPA	BALLOON	MM	0.917	0.599	0.434	0.631	119.5	60
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	MM	0.917	0.599	0.470	0.631	119.5	60
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	QM	0.916	0.603	0.472	0.635	117.8	60
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	MM	0.916	0.603	0.474	0.636	117.6	60
QM	HF	STO-3G	MPA	BALLOON	QM	0.915	0.607	0.436	0.639	116.1	60
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	MM	0.915	0.607	0.471	0.640	115.9	60
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	NO	0.913	0.615	0.477	0.648	112.7	60
QM	B3LYP	6-31G*	MPA	FROG2	NO	0.912	0.615	0.483	0.648	112.6	60
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	MM	0.912	0.615	0.480	0.649	112.5	60
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	NO	0.912	0.617	0.483	0.650	112.0	60
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	QM	0.912	0.617	0.484	0.651	111.8	60
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	MM	0.911	0.619	0.484	0.653	110.9	60
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	QM	0.911	0.621	0.487	0.654	110.4	60
QM	B3LYP	6-31G*	MPA	PUBCHEM	MM	0.911	0.622	0.437	0.655	110.0	60
QM	B3LYP	6-31G*	MPA	RDKIT	MM	0.909	0.626	0.423	0.660	108.4	60
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	QM	0.909	0.628	0.447	0.662	107.6	60
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	MM	0.908	0.629	0.481	0.663	107.2	60

Table S3**a**: Quality and statistical criteria of QM and EEM QSPR models.

Charges	**Theory level**	**Basis set**	**Charge scheme**	**3D source**	**Opt**	R^2	**RMSE**	$\overline{\Delta}$	s	F	**# molecules**
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	QM	0.908	0.631	0.486	0.665	106.5	60
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	NO	0.908	0.631	0.499	0.665	106.5	60
QM	B3LYP	6-31G*	NPA	BALLOON	NO	0.908	0.631	0.473	0.665	106.5	60
Svob2007-chal2	HF	STO-3G	MPA	FROG2	NO	0.907	0.634	0.470	0.668	105.4	60
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	MM	0.907	0.634	0.498	0.668	105.4	60
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	MM	0.907	0.634	0.498	0.668	105.4	60
QM	B3LYP	6-31G*	NPA	FROG2	MM	0.907	0.635	0.453	0.669	105.2	60
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	NO	0.907	0.635	0.500	0.669	105.0	60
Chavez2006	B3LYP	6-31G*	MPA	FROG2	QM	0.907	0.635	0.500	0.669	105.0	60
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	MM	0.906	0.636	0.492	0.670	104.7	60
Chavez2006	B3LYP	6-31G*	MPA	FROG2	MM	0.906	0.637	0.506	0.671	104.4	60
QM	B3LYP	6-31G*	NPA	FROG2	NO	0.906	0.638	0.479	0.673	103.8	60
Svob2007-chal2	HF	STO-3G	MPA	NCI	NO	0.906	0.639	0.478	0.673	103.6	60
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	QM	0.906	0.639	0.502	0.674	103.5	60
Chavez2006	B3LYP	6-31G*	MPA	NCI	NO	0.906	0.639	0.503	0.674	103.5	60
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	NO	0.904	0.643	0.503	0.678	102.0	60
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	NO	0.904	0.644	0.508	0.679	101.9	60
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	QM	0.904	0.645	0.507	0.680	101.4	60
QM	B3LYP	6-31G*	AIM	BALLOON	NO	0.904	0.645	0.427	0.680	101.3	60
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	NO	0.903	0.649	0.507	0.684	100.1	60
QM	B3LYP	6-31G*	AIM	NCI	QM	0.902	0.651	0.512	0.686	99.5	60
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	MM	0.902	0.652	0.502	0.687	99.1	60
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	NO	0.902	0.653	0.511	0.688	98.8	60
QM	B3LYP	6-31G*	NPA	BALLOON	QM	0.901	0.655	0.474	0.691	98.0	60
QM	B3LYP	6-31G*	AIM	PUBCHEM	NO	0.900	0.657	0.448	0.693	97.4	60
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	MM	0.900	0.658	0.476	0.693	97.2	60
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	NO	0.900	0.658	0.477	0.694	97.0	60
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	NO	0.899	0.661	0.506	0.697	96.1	60
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	MM	0.899	0.661	0.461	0.697	96.1	60
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	QM	0.899	0.661	0.526	0.697	96.1	60
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	MM	0.899	0.661	0.519	0.697	96.0	60
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	QM	0.898	0.664	0.476	0.700	95.1	60
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	NO	0.898	0.665	0.540	0.701	94.9	60
Chavez2006	B3LYP	6-31G*	MPA	FROG2	NO	0.897	0.667	0.542	0.703	94.1	60
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	QM	0.896	0.670	0.481	0.707	93.1	60
QM	HF	STO-3G	MPA	BALLOON	NO	0.896	0.671	0.507	0.707	92.9	60
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	NO	0.896	0.672	0.473	0.708	92.7	60
Svob2007-chal2	HF	STO-3G	MPA	NCI	QM	0.896	0.672	0.468	0.708	92.6	60
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	QM	0.895	0.672	0.527	0.709	92.5	60
QM	B3LYP	6-31G*	MPA	RDKIT	NO	0.895	0.674	0.530	0.710	92.0	60
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	MM	0.895	0.674	0.528	0.711	91.9	60
QM	HF	STO-3G	MPA	FROG2	NO	0.894	0.676	0.487	0.712	91.4	60
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	NO	0.894	0.677	0.504	0.713	91.1	60
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	MM	0.892	0.683	0.524	0.720	89.3	60

Table S3**a**: Quality and statistical criteria of QM and EEM QSPR models.

Charges	**Theory level**	**Basis set**	**Charge scheme**	**3D source**	**Opt**	R^2	**RMSE**	$\overline{\Delta}$	s	F	**# molecules**
QM	B3LYP	6-31G*	AIM	BALLOON	MM	0.891	0.686	0.476	0.723	88.3	60
QM	HF	STO-3G	MPA	NCI	QM	0.891	0.686	0.557	0.724	88.3	60
Svob2007-chal2	HF	STO-3G	MPA	NCI	MM	0.891	0.688	0.473	0.725	87.9	60
QM	B3LYP	6-31G*	AIM	FROG2	NO	0.891	0.688	0.534	0.725	87.8	60
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	QM	0.890	0.691	0.550	0.728	87.1	60
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	NO	0.890	0.691	0.546	0.728	87.0	60
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	NO	0.888	0.695	0.553	0.733	85.8	60
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	MM	0.887	0.698	0.513	0.736	85.0	60
QM	B3LYP	6-31G*	AIM	PUBCHEM	QM	0.886	0.701	0.452	0.739	84.3	60
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	NO	0.885	0.704	0.552	0.742	83.4	60
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	MM	0.885	0.705	0.542	0.743	83.1	60
QM	B3LYP	6-31G*	AIM	RDKIT	NO	0.882	0.716	0.558	0.754	80.4	60
QM	B3LYP	6-31G*	MPA	BALLOON	MM	0.881	0.719	0.466	0.757	79.6	60
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	NO	0.873	0.741	0.553	0.781	74.4	60
QM	B3LYP	6-31G*	MPA	BALLOON	QM	0.871	0.745	0.489	0.786	73.2	60
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	QM	0.869	0.752	0.556	0.792	71.8	60
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	QM	0.869	0.754	0.590	0.794	71.4	60
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	MM	0.868	0.756	0.587	0.797	70.9	60
QM	B3LYP	6-31G*	AIM	PUBCHEM	MM	0.864	0.766	0.526	0.807	68.9	60
QM	B3LYP	6-31G*	NPA	NCI	QM	0.861	0.776	0.651	0.818	66.8	60
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	NO	0.858	0.785	0.625	0.827	65.0	60
QM	B3LYP	6-31G*	AIM	BALLOON	QM	0.856	0.789	0.553	0.831	64.3	60
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	MM	0.852	0.799	0.560	0.843	62.3	60
QM	B3LYP	6-31G*	NPA	RDKIT	NO	0.796	0.940	0.768	0.991	42.1	60
QM	HF	STO-3G	MPA	RDKIT	NO	0.782	0.972	0.801	1.024	38.7	60

Table S3**a**: Quality and statistical criteria of QM and EEM QSPR models.

Charges	Theory level	Basis set	Charge scheme	3D source	Opt	R^2	RMSE	$\overline{\Delta}$	s	F	# molecules
Carboxylic acids											
QM	HF	STO-3G	MPA	NCI	MM	0.938	0.224	0.168	0.236	159.0	82
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	NO	0.932	0.235	0.185	0.247	144.1	82
QM	HF	STO-3G	MPA	NCI	NO	0.931	0.236	0.171	0.249	142.2	82
QM	B3LYP	6-31G*	NPA	NCI	MM	0.929	0.240	0.173	0.252	137.9	82
Chavez2006	B3LYP	6-31G*	MPA	NCI	NO	0.926	0.244	0.189	0.257	132.3	82
QM	HF	STO-3G	MPA	NCI	QM	0.925	0.246	0.181	0.259	130.0	82
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	NO	0.924	0.248	0.199	0.261	128.3	82
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	QM	0.923	0.249	0.182	0.262	127.3	82
Chavez2006	B3LYP	6-31G*	MPA	NCI	QM	0.923	0.249	0.198	0.262	127.1	82
QM	B3LYP	6-31G*	AIM	FROG2	MM	0.922	0.250	0.194	0.263	125.7	82
QM	B3LYP	6-31G*	AIM	NCI	MM	0.922	0.251	0.184	0.264	124.9	82
QM	B3LYP	6-31G*	AIM	NCI	QM	0.921	0.253	0.191	0.267	122.4	82
Chavez2006	B3LYP	6-31G*	MPA	NCI	MM	0.920	0.254	0.197	0.267	121.5	82
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	MM	0.919	0.256	0.198	0.269	120.0	82
QM	B3LYP	6-31G*	AIM	FROG2	QM	0.917	0.258	0.190	0.272	117.2	82
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	MM	0.917	0.259	0.211	0.272	117.0	82
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	QM	0.917	0.259	0.205	0.272	116.8	82
QM	HF	STO-3G	MPA	FROG2	QM	0.917	0.259	0.186	0.272	116.7	82
QM	HF	STO-3G	MPA	PUBCHEM	MM	0.916	0.260	0.186	0.274	115.3	82
QM	B3LYP	6-31G*	AIM	PUBCHEM	MM	0.916	0.261	0.189	0.275	114.8	82
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	QM	0.915	0.263	0.210	0.276	113.1	82
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	QM	0.911	0.268	0.206	0.282	108.3	82
QM	B3LYP	6-31G*	NPA	NCI	NO	0.911	0.268	0.191	0.283	107.8	82
QM	B3LYP	6-31G*	AIM	NCI	NO	0.910	0.270	0.200	0.284	106.5	82
QM	HF	STO-3G	MPA	PUBCHEM	NO	0.909	0.271	0.186	0.285	105.6	82
QM	B3LYP	6-31G*	AIM	PUBCHEM	QM	0.908	0.272	0.203	0.286	104.9	82
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	MM	0.907	0.273	0.210	0.288	103.5	82
QM	HF	STO-3G	MPA	FROG2	MM	0.907	0.273	0.195	0.288	103.4	82
QM	B3LYP	6-31G*	AIM	PUBCHEM	NO	0.907	0.274	0.197	0.288	103.4	82
QM	B3LYP	6-31G*	NPA	FROG2	QM	0.906	0.276	0.202	0.290	101.6	82
QM	B3LYP	6-31G*	NPA	NCI	QM	0.903	0.280	0.201	0.294	98.6	82
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	MM	0.903	0.280	0.217	0.295	98.0	82
QM	B3LYP	6-31G*	NPA	FROG2	MM	0.903	0.280	0.201	0.295	98.0	82
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	MM	0.902	0.281	0.192	0.296	97.4	82
QM	HF	STO-3G	MPA	RDKIT	MM	0.902	0.282	0.216	0.296	97.0	82
QM	HF	STO-3G	MPA	PUBCHEM	QM	0.901	0.282	0.194	0.297	96.6	82
Chavez2006	B3LYP	6-31G*	MPA	FROG2	QM	0.898	0.287	0.199	0.302	93.0	82
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	NO	0.898	0.288	0.225	0.303	92.6	82
QM	B3LYP	6-31G*	NPA	PUBCHEM	QM	0.896	0.290	0.204	0.305	90.9	82
QM	HF	STO-3G	MPA	FROG2	NO	0.896	0.290	0.204	0.305	90.8	82
QM	HF	STO-3G	MPA	RDKIT	MM	0.894	0.292	0.213	0.308	89.4	82
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	QM	0.894	0.293	0.206	0.308	88.8	82

Table S3**a**: Quality and statistical criteria of QM and EEM QSPR models.

Charges	**Theory level**	**Basis set**	**Charge scheme**	**3D source**	**Opt**	R^2	**RMSE**	$\overline{\Delta}$	s	F	**# molecules**
QM	B3LYP	6-31G*	NPA	PUBCHEM	MM	0.892	0.295	0.210	0.310	87.7	82
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	QM	0.892	0.295	0.227	0.311	87.3	82
QM	B3LYP	6-31G*	MPA	NCI	NO	0.891	0.296	0.203	0.311	86.8	82
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	NO	0.891	0.296	0.201	0.312	86.7	82
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	NO	0.891	0.297	0.231	0.313	86.1	82
QM	B3LYP	6-31G*	NPA	PUBCHEM	NO	0.891	0.297	0.207	0.313	86.0	82
QM	HF	STO-3G	MPA	BALLOON	QM	0.890	0.297	0.232	0.313	85.9	82
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	QM	0.889	0.299	0.242	0.314	85.0	82
QM	B3LYP	6-31G*	MPA	NCI	MM	0.889	0.300	0.220	0.315	84.5	82
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	NO	0.889	0.300	0.227	0.315	84.5	82
Chavez2006	B3LYP	6-31G*	MPA	FROG2	MM	0.888	0.300	0.207	0.316	83.9	82
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	NO	0.888	0.300	0.233	0.316	83.9	82
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	MM	0.888	0.301	0.222	0.317	83.6	82
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	NO	0.888	0.301	0.244	0.317	83.6	82
QM	B3LYP	6-31G*	MPA	PUBCHEM	MM	0.885	0.305	0.225	0.321	81.1	82
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	QM	0.884	0.306	0.234	0.322	80.4	82
QM	B3LYP	6-31G*	AIM	FROG2	NO	0.884	0.307	0.220	0.324	79.2	81
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	MM	0.883	0.307	0.248	0.323	80.1	82
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	QM	0.881	0.310	0.241	0.326	78.4	82
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	NO	0.879	0.313	0.259	0.329	76.5	82
QM	B3LYP	6-31G*	MPA	FROG2	NO	0.876	0.316	0.232	0.333	74.8	82
QM	B3LYP	6-31G*	NPA	FROG2	NO	0.876	0.316	0.241	0.333	74.8	82
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	MM	0.876	0.316	0.227	0.333	74.7	82
QM	B3LYP	6-31G*	AIM	OPENBABEL	NO	0.876	0.317	0.249	0.334	74.4	82
Chavez2006	B3LYP	6-31G*	MPA	FROG2	NO	0.875	0.317	0.241	0.334	74.3	82
Svob2007-chal2	HF	STO-3G	MPA	NCI	NO	0.875	0.317	0.244	0.334	74.1	82
QM	B3LYP	6-31G*	AIM	OPENBABEL	MM	0.875	0.318	0.249	0.335	73.8	82
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	QM	0.874	0.319	0.237	0.336	73.3	82
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	MM	0.874	0.319	0.248	0.336	73.3	82
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	NO	0.874	0.319	0.256	0.336	73.2	82
QM	B3LYP	6-31G*	MPA	PUBCHEM	NO	0.873	0.320	0.239	0.337	72.6	82
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	NO	0.873	0.320	0.240	0.337	72.5	82
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	MM	0.871	0.322	0.246	0.339	71.5	82
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	MM	0.871	0.323	0.242	0.340	71.4	82
QM	B3LYP	6-31G*	MPA	PUBCHEM	QM	0.871	0.323	0.235	0.340	71.2	82
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	NO	0.870	0.324	0.231	0.341	70.6	82
QM	HF	STO-3G	MPA	OPENBABEL	NO	0.869	0.325	0.239	0.342	70.3	82
QM	HF	STO-3G	MPA	BALLOON	MM	0.867	0.328	0.228	0.345	68.8	82
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	QM	0.866	0.328	0.221	0.345	68.6	82
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	NO	0.866	0.329	0.232	0.346	68.3	82
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	MM	0.866	0.329	0.260	0.346	68.2	82
QM	HF	STO-3G	MPA	OPENBABEL	MM	0.863	0.332	0.247	0.350	66.8	82
QM	B3LYP	6-31G*	AIM	RDKIT	MM	0.862	0.334	0.254	0.351	66.0	82
QM	HF	STO-3G	MPA	RDKIT	QM	0.861	0.335	0.249	0.353	65.2	82

Table S3**a**: Quality and statistical criteria of QM and EEM QSPR models.

Charges	**Theory level**	**Basis set**	**Charge scheme**	**3D source**	**Opt**	R^2	**RMSE**	$\overline{\Delta}$	s	F	**# molecules**
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	MM	0.860	0.335	0.237	0.353	65.2	82
QM	B3LYP	6-31G*	AIM	RDKIT	MM	0.860	0.336	0.246	0.354	65.0	82
Svob2007-chal2	HF	STO-3G	MPA	NCI	MM	0.860	0.337	0.263	0.354	64.7	82
Svob2007-chal2	HF	STO-3G	MPA	FROG2	MM	0.859	0.337	0.275	0.355	64.3	82
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	QM	0.856	0.341	0.254	0.359	62.7	82
QM	B3LYP	6-31G*	AIM	RDKIT	QM	0.855	0.342	0.269	0.360	62.4	82
QM	B3LYP	6-31G*	MPA	NCI	QM	0.854	0.343	0.244	0.361	61.9	82
QM	B3LYP	6-31G*	MPA	FROG2	QM	0.853	0.344	0.250	0.363	61.3	82
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	NO	0.853	0.345	0.270	0.363	61.3	82
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	MM	0.852	0.345	0.287	0.364	60.9	82
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	QM	0.851	0.346	0.276	0.364	60.6	82
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	QM	0.851	0.346	0.275	0.364	60.6	82
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	QM	0.851	0.346	0.267	0.365	60.5	82
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	MM	0.849	0.350	0.265	0.368	59.2	82
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	MM	0.848	0.350	0.279	0.368	59.2	82
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	NO	0.848	0.350	0.264	0.369	58.9	82
QM	B3LYP	6-31G*	AIM	BALLOON	NO	0.846	0.352	0.268	0.371	58.2	82
QM	B3LYP	6-31G*	NPA	OPENBABEL	NO	0.845	0.353	0.271	0.372	57.8	82
QM	HF	STO-3G	MPA	OPENBABEL	QM	0.845	0.353	0.275	0.372	57.8	82
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	MM	0.844	0.355	0.283	0.374	57.1	82
QM	B3LYP	6-31G*	AIM	BALLOON	MM	0.843	0.356	0.267	0.375	56.7	82
QM	B3LYP	6-31G*	NPA	RDKIT	MM	0.842	0.357	0.259	0.375	56.5	82
QM	B3LYP	6-31G*	NPA	OPENBABEL	MM	0.841	0.358	0.275	0.377	56.0	82
Svob2007-chal2	HF	STO-3G	MPA	FROG2	QM	0.841	0.359	0.282	0.378	55.7	82
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	MM	0.837	0.363	0.290	0.382	54.3	82
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	NO	0.836	0.363	0.285	0.383	54.0	82
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	MM	0.835	0.365	0.287	0.384	53.5	82
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	QM	0.834	0.366	0.291	0.385	53.1	82
Svob2007-chal2	HF	STO-3G	MPA	FROG2	NO	0.832	0.368	0.303	0.387	52.4	82
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	NO	0.832	0.368	0.295	0.388	52.3	82
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	NO	0.831	0.369	0.268	0.388	52.2	82
QM	B3LYP	6-31G*	MPA	FROG2	MM	0.830	0.370	0.253	0.390	51.7	82
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	NO	0.830	0.370	0.317	0.390	51.6	82
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	MM	0.829	0.371	0.284	0.391	51.3	82
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	MM	0.827	0.373	0.319	0.393	50.7	82
QM	B3LYP	6-31G*	AIM	OPENBABEL	QM	0.827	0.374	0.282	0.393	50.5	82
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	NO	0.826	0.374	0.295	0.394	50.3	82
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	QM	0.825	0.376	0.289	0.395	49.9	82
QM	B3LYP	6-31G*	NPA	BALLOON	QM	0.824	0.377	0.280	0.397	49.4	82
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	MM	0.823	0.377	0.303	0.397	49.3	82
QM	HF	STO-3G	MPA	BALLOON	NO	0.823	0.378	0.288	0.398	49.2	82
QM	B3LYP	6-31G*	MPA	RDKIT	MM	0.821	0.380	0.295	0.400	48.6	82
Svob2007-chal2	HF	STO-3G	MPA	NCI	QM	0.821	0.380	0.293	0.400	48.4	82
QM	B3LYP	6-31G*	NPA	BALLOON	NO	0.819	0.383	0.278	0.403	47.7	82

Table S3**a**: Quality and statistical criteria of QM and EEM QSPR models.

Charges	**Theory level**	**Basis set**	**Charge scheme**	**3D source**	**Opt**	R^2	**RMSE**	$\overline{\Delta}$	s	F	**# molecules**
QM	B3LYP	6-31G*	AIM	RDKIT	NO	0.817	0.384	0.295	0.404	47.2	82
QM	B3LYP	6-31G*	NPA	RDKIT	QM	0.814	0.388	0.281	0.408	46.2	82
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	QM	0.807	0.395	0.322	0.415	44.2	82
QM	B3LYP	6-31G*	AIM	BALLOON	QM	0.807	0.395	0.299	0.416	44.1	82
QM	B3LYP	6-31G*	NPA	BALLOON	MM	0.805	0.397	0.298	0.418	43.6	82
QM	B3LYP	6-31G*	NPA	OPENBABEL	QM	0.804	0.398	0.310	0.419	43.4	82
QM	B3LYP	6-31G*	NPA	RDKIT	NO	0.804	0.398	0.296	0.419	43.4	82
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	MM	0.801	0.401	0.328	0.422	42.5	82
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	MM	0.800	0.401	0.336	0.422	42.4	82
QM	B3LYP	6-31G*	NPA	RDKIT	MM	0.797	0.405	0.311	0.426	41.4	82
QM	HF	STO-3G	MPA	RDKIT	NO	0.780	0.421	0.328	0.443	37.6	82
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	QM	0.772	0.429	0.348	0.451	35.8	82
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	QM	0.753	0.446	0.358	0.470	32.2	82
QM	B3LYP	6-31G*	MPA	RDKIT	MM	0.750	0.449	0.353	0.473	31.7	82
QM	B3LYP	6-31G*	MPA	RDKIT	NO	0.723	0.473	0.373	0.497	27.6	82
QM	B3LYP	6-31G*	MPA	BALLOON	NO	0.720	0.475	0.375	0.500	27.2	82
QM	B3LYP	6-31G*	MPA	RDKIT	QM	0.696	0.495	0.356	0.521	24.2	82
QM	B3LYP	6-31G*	MPA	OPENBABEL	QM	0.674	0.513	0.388	0.540	21.9	82
QM	B3LYP	6-31G*	MPA	OPENBABEL	MM	0.665	0.519	0.382	0.547	21.0	82
QM	B3LYP	6-31G*	MPA	OPENBABEL	NO	0.658	0.525	0.383	0.553	20.3	82
QM	B3LYP	6-31G*	MPA	BALLOON	QM	0.618	0.555	0.451	0.585	17.1	82
QM	B3LYP	6-31G*	MPA	BALLOON	MM	0.587	0.577	0.452	0.607	15.0	82

Table S3**a**: Quality and statistical criteria of QM and EEM QSPR models.

Charges	Theory level	Basis set	Charge scheme	3D source	Opt	R^2	RMSE	$\overline{\Delta}$	s	F	# molecules
Anilines											
QM	B3LYP	6-31G*	AIM	RDKIT	MM	0.980	0.312	0.260	0.342	280.6	48
QM	B3LYP	6-31G*	MPA	RDKIT	MM	0.979	0.318	0.251	0.348	271.0	48
QM	B3LYP	6-31G*	MPA	PUBCHEM	MM	0.979	0.320	0.260	0.351	267.0	48
QM	B3LYP	6-31G*	MPA	RDKIT	MM	0.976	0.345	0.261	0.378	228.5	48
QM	B3LYP	6-31G*	MPA	OPENBABEL	MM	0.975	0.346	0.264	0.379	227.1	48
QM	B3LYP	6-31G*	MPA	OPENBABEL	QM	0.974	0.357	0.266	0.391	213.2	48
QM	B3LYP	6-31G*	MPA	PUBCHEM	QM	0.974	0.360	0.281	0.394	210.1	48
QM	B3LYP	6-31G*	MPA	OPENBABEL	NO	0.973	0.361	0.260	0.395	208.8	48
QM	B3LYP	6-31G*	MPA	FROG2	MM	0.973	0.361	0.282	0.395	208.6	48
QM	B3LYP	6-31G*	NPA	RDKIT	MM	0.973	0.363	0.288	0.397	206.6	48
QM	B3LYP	6-31G*	MPA	FROG2	QM	0.973	0.366	0.270	0.401	203.2	48
QM	B3LYP	6-31G*	MPA	NCI	NO	0.970	0.384	0.279	0.421	183.3	48
QM	B3LYP	6-31G*	MPA	BALLOON	QM	0.967	0.400	0.295	0.438	169.1	48
QM	B3LYP	6-31G*	NPA	OPENBABEL	MM	0.967	0.401	0.306	0.439	168.4	48
QM	B3LYP	6-31G*	NPA	NCI	MM	0.967	0.401	0.313	0.439	168.2	48
QM	B3LYP	6-31G*	NPA	RDKIT	MM	0.967	0.402	0.307	0.440	167.2	48
QM	B3LYP	6-31G*	NPA	PUBCHEM	MM	0.966	0.407	0.341	0.445	163.2	48
QM	B3LYP	6-31G*	NPA	NCI	NO	0.966	0.408	0.286	0.447	162.1	48
QM	B3LYP	6-31G*	NPA	OPENBABEL	NO	0.966	0.409	0.294	0.448	161.0	48
QM	B3LYP	6-31G*	NPA	FROG2	MM	0.965	0.413	0.318	0.453	157.8	48
QM	HF	STO-3G	MPA	RDKIT	MM	0.965	0.414	0.327	0.454	157.0	48
QM	B3LYP	6-31G*	MPA	RDKIT	QM	0.964	0.421	0.296	0.462	151.6	48
QM	B3LYP	6-31G*	MPA	NCI	QM	0.959	0.447	0.315	0.490	134.0	48
QM	HF	STO-3G	MPA	RDKIT	MM	0.959	0.450	0.363	0.493	132.1	48
QM	HF	STO-3G	MPA	FROG2	MM	0.958	0.455	0.370	0.498	129.2	48
QM	HF	STO-3G	MPA	OPENBABEL	MM	0.958	0.455	0.367	0.499	129.1	48
QM	B3LYP	6-31G*	MPA	NCI	MM	0.955	0.467	0.311	0.511	122.5	48
QM	HF	STO-3G	MPA	NCI	MM	0.954	0.472	0.363	0.517	119.8	48
QM	B3LYP	6-31G*	MPA	BALLOON	MM	0.953	0.480	0.320	0.526	115.4	48
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	QM	0.953	0.480	0.378	0.526	115.3	48
QM	HF	STO-3G	MPA	OPENBABEL	NO	0.952	0.484	0.367	0.530	113.7	48
QM	HF	STO-3G	MPA	NCI	NO	0.951	0.488	0.367	0.534	111.6	48
QM	HF	STO-3G	MPA	BALLOON	QM	0.948	0.504	0.391	0.552	104.3	48
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	QM	0.946	0.515	0.414	0.564	99.8	48
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	QM	0.945	0.517	0.385	0.566	99.0	48
QM	HF	STO-3G	MPA	NCI	QM	0.942	0.532	0.421	0.582	93.1	48
QM	HF	STO-3G	MPA	PUBCHEM	MM	0.942	0.532	0.402	0.583	92.9	48
QM	HF	STO-3G	MPA	RDKIT	QM	0.940	0.541	0.407	0.592	89.9	48
QM	B3LYP	6-31G*	MPA	PUBCHEM	NO	0.939	0.544	0.412	0.596	88.5	48
Svob2007-chal2	HF	STO-3G	MPA	FROG2	QM	0.939	0.545	0.438	0.597	88.3	48
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	QM	0.939	0.546	0.453	0.598	88.1	48
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	QM	0.939	0.547	0.463	0.599	87.7	48

Table S3**a**: Quality and statistical criteria of QM and EEM QSPR models.

Charges	**Theory level**	**Basis set**	**Charge scheme**	**3D source**	**Opt**	R^2	**RMSE**	$\overline{\Delta}$	s	F	**# molecules**
QM	HF	STO-3G	MPA	PUBCHEM	NO	0.938	0.549	0.405	0.601	87.1	48
QM	B3LYP	6-31G*	AIM	PUBCHEM	NO	0.937	0.554	0.427	0.607	85.2	48
QM	B3LYP	6-31G*	NPA	NCI	QM	0.937	0.555	0.399	0.608	84.9	48
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	QM	0.935	0.565	0.484	0.619	81.8	48
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	QM	0.934	0.566	0.475	0.620	81.3	48
QM	HF	STO-3G	MPA	PUBCHEM	QM	0.934	0.568	0.447	0.622	81.0	48
QM	HF	STO-3G	MPA	FROG2	NO	0.934	0.569	0.459	0.623	80.6	48
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	QM	0.933	0.570	0.461	0.625	80.1	48
QM	B3LYP	6-31G*	NPA	BALLOON	QM	0.933	0.572	0.436	0.627	79.5	48
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	QM	0.930	0.584	0.456	0.639	76.3	48
QM	B3LYP	6-31G*	AIM	OPENBABEL	NO	0.930	0.587	0.444	0.643	75.4	48
QM	B3LYP	6-31G*	NPA	OPENBABEL	QM	0.928	0.592	0.462	0.649	74.0	48
QM	B3LYP	6-31G*	NPA	BALLOON	MM	0.927	0.596	0.459	0.653	73.0	48
QM	B3LYP	6-31G*	AIM	OPENBABEL	MM	0.927	0.597	0.470	0.654	72.6	48
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	QM	0.927	0.598	0.475	0.655	72.3	48
QM	B3LYP	6-31G*	AIM	RDKIT	MM	0.927	0.598	0.459	0.656	72.3	48
QM	B3LYP	6-31G*	NPA	RDKIT	QM	0.927	0.599	0.469	0.656	72.1	48
QM	B3LYP	6-31G*	AIM	FROG2	MM	0.926	0.602	0.481	0.659	71.4	48
Chavez2006	B3LYP	6-31G*	MPA	FROG2	QM	0.926	0.602	0.521	0.660	71.3	48
QM	B3LYP	6-31G*	NPA	FROG2	NO	0.924	0.611	0.492	0.670	69.0	48
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	QM	0.923	0.612	0.485	0.671	68.8	48
QM	B3LYP	6-31G*	AIM	BALLOON	QM	0.921	0.620	0.469	0.679	66.9	48
QM	B3LYP	6-31G*	NPA	PUBCHEM	NO	0.921	0.621	0.440	0.681	66.6	48
QM	B3LYP	6-31G*	AIM	BALLOON	MM	0.921	0.622	0.465	0.681	66.6	48
Svob2007-chal2	HF	STO-3G	MPA	NCI	QM	0.921	0.622	0.490	0.681	66.5	48
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	QM	0.920	0.626	0.491	0.686	65.5	48
QM	B3LYP	6-31G*	AIM	FROG2	NO	0.916	0.642	0.503	0.703	62.0	48
QM	B3LYP	6-31G*	AIM	PUBCHEM	MM	0.916	0.642	0.507	0.704	62.0	48
QM	B3LYP	6-31G*	AIM	NCI	MM	0.914	0.650	0.523	0.712	60.4	48
QM	B3LYP	6-31G*	NPA	BALLOON	NO	0.912	0.656	0.535	0.718	59.3	48
QM	B3LYP	6-31G*	MPA	FROG2	NO	0.911	0.658	0.552	0.721	58.8	48
QM	B3LYP	6-31G*	NPA	FROG2	QM	0.911	0.661	0.524	0.724	58.2	48
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	QM	0.908	0.671	0.525	0.735	56.4	48
QM	B3LYP	6-31G*	AIM	RDKIT	QM	0.908	0.671	0.513	0.735	56.3	48
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	QM	0.907	0.673	0.533	0.737	55.9	48
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	NO	0.907	0.675	0.534	0.739	55.6	48
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	QM	0.905	0.683	0.549	0.748	54.2	48
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	NO	0.904	0.683	0.528	0.749	54.1	48
QM	B3LYP	6-31G*	AIM	NCI	NO	0.903	0.688	0.498	0.753	53.4	48
QM	B3LYP	6-31G*	MPA	BALLOON	NO	0.903	0.689	0.531	0.754	53.2	48
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	MM	0.902	0.693	0.570	0.760	52.4	48
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	MM	0.900	0.700	0.566	0.767	51.3	48
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	NO	0.899	0.702	0.570	0.768	51.0	48
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	MM	0.897	0.709	0.569	0.776	49.9	48

Table S3**a**: Quality and statistical criteria of QM and EEM QSPR models.

Charges	**Theory level**	**Basis set**	**Charge scheme**	**3D source**	**Opt**	R^2	**RMSE**	$\overline{\Delta}$	s	F	**# molecules**
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	MM	0.897	0.711	0.576	0.779	49.5	48
Svob2007-chal2	HF	STO-3G	MPA	FROG2	NO	0.894	0.721	0.569	0.790	48.0	48
QM	B3LYP	6-31G*	AIM	NCI	QM	0.892	0.726	0.591	0.795	47.3	48
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	QM	0.892	0.727	0.581	0.797	47.1	48
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	QM	0.890	0.735	0.602	0.805	46.0	48
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	MM	0.889	0.738	0.551	0.809	45.5	48
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	NO	0.887	0.744	0.540	0.815	44.8	48
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	MM	0.886	0.746	0.577	0.817	44.5	48
QM	B3LYP	6-31G*	NPA	PUBCHEM	QM	0.885	0.749	0.553	0.821	44.1	48
Chavez2006	B3LYP	6-31G*	MPA	NCI	QM	0.884	0.753	0.618	0.825	43.5	48
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	NO	0.882	0.758	0.606	0.831	42.9	48
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	NO	0.881	0.763	0.622	0.836	42.3	48
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	MM	0.880	0.767	0.594	0.840	41.8	48
QM	B3LYP	6-31G*	AIM	OPENBABEL	QM	0.880	0.767	0.615	0.841	41.7	48
Chavez2006	B3LYP	6-31G*	MPA	FROG2	NO	0.879	0.770	0.577	0.844	41.4	48
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	MM	0.875	0.780	0.596	0.855	40.2	48
QM	HF	STO-3G	MPA	FROG2	QM	0.875	0.782	0.668	0.857	39.9	48
QM	HF	STO-3G	MPA	BALLOON	MM	0.874	0.785	0.595	0.860	39.6	48
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	NO	0.874	0.785	0.627	0.860	39.6	48
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	NO	0.874	0.786	0.653	0.861	39.5	48
QM	HF	STO-3G	MPA	OPENBABEL	QM	0.874	0.786	0.674	0.861	39.5	48
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	MM	0.872	0.790	0.624	0.865	39.0	48
Svob2007-chal2	HF	STO-3G	MPA	NCI	NO	0.870	0.796	0.606	0.872	38.4	48
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	NO	0.870	0.797	0.637	0.873	38.3	48
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	QM	0.869	0.801	0.670	0.877	37.8	48
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	QM	0.869	0.801	0.671	0.878	37.8	48
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	QM	0.867	0.805	0.663	0.882	37.4	48
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	MM	0.867	0.806	0.611	0.883	37.3	48
Svob2007-chal2	HF	STO-3G	MPA	FROG2	MM	0.863	0.818	0.647	0.896	36.1	48
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	NO	0.860	0.827	0.652	0.906	35.1	48
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	MM	0.858	0.833	0.664	0.913	34.5	48
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	MM	0.857	0.835	0.646	0.915	34.4	48
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	NO	0.857	0.835	0.623	0.915	34.3	48
Chavez2006	B3LYP	6-31G*	MPA	FROG2	MM	0.857	0.837	0.703	0.917	34.1	48
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	MM	0.855	0.842	0.637	0.923	33.7	48
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	MM	0.854	0.846	0.666	0.926	33.3	48
QM	HF	STO-3G	MPA	RDKIT	NO	0.853	0.847	0.690	0.928	33.2	48
QM	B3LYP	6-31G*	AIM	FROG2	QM	0.853	0.849	0.687	0.930	33.1	48
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	MM	0.852	0.850	0.669	0.931	33.0	48
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	MM	0.852	0.850	0.706	0.931	32.9	48
Chavez2006	B3LYP	6-31G*	MPA	NCI	NO	0.852	0.851	0.683	0.932	32.9	48
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	MM	0.852	0.851	0.683	0.933	32.8	48
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	NO	0.851	0.853	0.688	0.935	32.6	48
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	MM	0.851	0.854	0.701	0.936	32.6	48

Table S3**a**: Quality and statistical criteria of QM and EEM QSPR models.

Charges	**Theory level**	**Basis set**	**Charge scheme**	**3D source**	**Opt**	R^2	**RMSE**	$\overline{\Delta}$	s	F	**# molecules**
QM	B3LYP	6-31G*	NPA	RDKIT	NO	0.851	0.854	0.712	0.936	32.6	48
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	NO	0.849	0.859	0.689	0.941	32.1	48
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	MM	0.848	0.863	0.663	0.945	31.8	48
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	NO	0.847	0.866	0.670	0.948	31.5	48
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	MM	0.846	0.867	0.682	0.950	31.4	48
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	MM	0.845	0.872	0.684	0.955	31.0	48
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	MM	0.845	0.872	0.656	0.955	31.0	48
Svob2007-chal2	HF	STO-3G	MPA	NCI	MM	0.844	0.874	0.663	0.957	30.9	48
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	MM	0.843	0.876	0.689	0.960	30.7	48
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	NO	0.842	0.878	0.664	0.962	30.5	48
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	NO	0.840	0.885	0.692	0.969	30.0	48
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	NO	0.839	0.887	0.695	0.971	29.8	48
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	MM	0.838	0.891	0.721	0.976	29.5	48
QM	HF	STO-3G	MPA	BALLOON	NO	0.836	0.894	0.710	0.979	29.2	48
Chavez2006	B3LYP	6-31G*	MPA	NCI	MM	0.834	0.900	0.702	0.985	28.8	48
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	NO	0.832	0.907	0.732	0.994	28.2	48
QM	B3LYP	6-31G*	AIM	PUBCHEM	QM	0.828	0.918	0.768	1.006	27.4	48
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	NO	0.826	0.922	0.701	1.010	27.2	48
QM	B3LYP	6-31G*	MPA	RDKIT	NO	0.816	0.948	0.814	1.038	25.4	48
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	NO	0.806	0.974	0.806	1.067	23.7	48
QM	B3LYP	6-31G*	AIM	BALLOON	NO	0.805	0.977	0.765	1.070	23.6	48
QM	B3LYP	6-31G*	AIM	RDKIT	NO	0.796	0.998	0.770	1.093	22.3	48

Table S3**b**: Quality and statistical criteria of PM6 and Gasteiger QSPR models.

Charges	**Theory level**	**Basis set**	**Charge scheme**	**3D source**	**Opt**	R^2	**RMSE**	$\overline{\Delta}$	s	F	**# molecules**
Phenols											
QM	PM6	none	MPA	NCI	MM	0.962	0.408	0.319	0.430	269.9	60
QM	PM6	none	MPA	FROG2	QM	0.959	0.419	0.305	0.441	255.7	60
QM	PM6	none	MPA	OPENBABEL	MM	0.959	0.419	0.317	0.441	255.5	60
QM	PM6	none	MPA	RDKIT	MM	0.952	0.455	0.330	0.480	214.5	60
QM	PM6	none	MPA	OPENBABEL	QM	0.952	0.455	0.333	0.480	214.4	60
QM	PM6	none	MPA	NCI	NO	0.950	0.463	0.349	0.488	207.0	60
QM	PM6	none	MPA	OPENBABEL	NO	0.949	0.468	0.352	0.494	202.2	60
QM	PM6	none	MPA	FROG2	MM	0.942	0.501	0.363	0.528	175.0	60
QM	PM6	none	MPA	RDKIT	QM	0.940	0.508	0.371	0.536	170.1	60
QM	PM6	none	MPA	FROG2	NO	0.939	0.514	0.376	0.542	166.1	60
QM	PM6	none	MPA	PUBCHEM	NO	0.934	0.532	0.381	0.561	154.0	60
QM	PM6	none	MPA	RDKIT	MM	0.931	0.547	0.377	0.577	145.2	60
QM	PM6	none	MPA	PUBCHEM	QM	0.929	0.555	0.387	0.585	140.8	60
QM	PM6	none	MPA	PUBCHEM	MM	0.915	0.607	0.443	0.640	116.0	60
QM	PM6	none	MPA	BALLOON	NO	0.900	0.656	0.500	0.692	97.6	60
QM	PM6	none	MPA	BALLOON	MM	0.889	0.692	0.488	0.730	86.7	60
QM	PM6	none	MPA	BALLOON	QM	0.886	0.703	0.495	0.741	83.6	60
QM	PM6	none	MPA	NCI	QM	0.878	0.727	0.586	0.767	77.5	60
QM	PM6	none	MPA	RDKIT	NO	0.868	0.756	0.629	0.797	70.9	60
Gasteiger				NCI	NO	0.747	1.047	0.753	1.103	31.8	60

Table S3**b**: Quality and statistical criteria of PM6 and Gasteiger QSPR models.

Charges	**Theory level**	**Basis set**	**Charge scheme**	**3D source**	**Opt**	R^2	**RMSE**	$\overline{\Delta}$	s	F	**# molecules**
Carboxylic acids											
QM	PM6	none	MPA	NCI	NO	0.941	0.217	0.160	0.229	169.8	82
QM	PM6	none	MPA	NCI	MM	0.936	0.227	0.163	0.239	154.3	82
QM	PM6	none	MPA	PUBCHEM	NO	0.933	0.233	0.172	0.245	146.4	82
QM	PM6	none	MPA	NCI	QM	0.932	0.233	0.171	0.246	146.0	82
QM	PM6	none	MPA	FROG2	MM	0.931	0.235	0.169	0.247	143.7	82
QM	PM6	none	MPA	FROG2	QM	0.927	0.243	0.173	0.256	133.7	82
QM	PM6	none	MPA	PUBCHEM	MM	0.926	0.244	0.174	0.257	133.0	82
QM	PM6	none	MPA	OPENBABEL	MM	0.920	0.254	0.194	0.267	121.5	82
QM	PM6	none	MPA	PUBCHEM	QM	0.919	0.255	0.190	0.269	120.4	82
QM	PM6	none	MPA	OPENBABEL	NO	0.919	0.256	0.196	0.270	119.5	82
QM	PM6	none	MPA	OPENBABEL	QM	0.901	0.282	0.206	0.297	96.6	82
QM	PM6	none	MPA	BALLOON	QM	0.895	0.291	0.224	0.306	90.1	82
QM	PM6	none	MPA	RDKIT	MM	0.895	0.291	0.199	0.307	89.9	82
QM	PM6	none	MPA	FROG2	NO	0.886	0.303	0.209	0.319	82.1	82
QM	PM6	none	MPA	BALLOON	MM	0.877	0.315	0.241	0.332	75.4	82
Gasteiger				NCI	NO	0.870	0.323	0.230	0.340	71.0	82
QM	PM6	none	MPA	RDKIT	MM	0.863	0.332	0.252	0.350	66.7	82
QM	PM6	none	MPA	BALLOON	NO	0.854	0.343	0.257	0.361	62.0	82
QM	PM6	none	MPA	RDKIT	NO	0.839	0.360	0.269	0.379	55.3	82
QM	PM6	none	MPA	RDKIT	QM	0.838	0.361	0.238	0.380	54.8	82

Table S3**b**: Quality and statistical criteria of PM6 and Gasteiger QSPR models.

Charges	**Theory level**	**Basis set**	**Charge scheme**	**3D source**	**Opt**	R^2	**RMSE**	$\overline{\Delta}$	s	F	**# molecules**
Anilines											
QM	PM6	none	MPA	RDKIT	MM	0.970	0.383	0.311	0.420	184.5	48
QM	PM6	none	MPA	NCI	NO	0.966	0.410	0.308	0.449	160.6	48
QM	PM6	none	MPA	OPENBABEL	NO	0.964	0.421	0.315	0.462	151.6	48
QM	PM6	none	MPA	FROG2	MM	0.961	0.436	0.328	0.477	141.5	48
QM	PM6	none	MPA	RDKIT	MM	0.957	0.458	0.338	0.502	127.5	48
QM	PM6	none	MPA	OPENBABEL	MM	0.957	0.459	0.338	0.503	126.9	48
QM	PM6	none	MPA	NCI	MM	0.956	0.466	0.343	0.511	122.8	48
QM	PM6	none	MPA	PUBCHEM	MM	0.955	0.470	0.371	0.515	120.5	48
QM	PM6	none	MPA	PUBCHEM	NO	0.952	0.485	0.365	0.531	113.3	48
QM	PM6	none	MPA	NCI	QM	0.940	0.541	0.434	0.592	89.8	48
QM	PM6	none	MPA	BALLOON	MM	0.935	0.565	0.446	0.619	81.7	48
QM	PM6	none	MPA	BALLOON	QM	0.933	0.573	0.465	0.628	79.3	48
QM	PM6	none	MPA	RDKIT	QM	0.930	0.584	0.459	0.640	76.1	48
QM	PM6	none	MPA	FROG2	NO	0.919	0.631	0.517	0.691	64.5	48
QM	PM6	none	MPA	BALLOON	NO	0.896	0.712	0.553	0.780	49.3	48
QM	PM6	none	MPA	FROG2	QM	0.890	0.732	0.599	0.802	46.4	48
QM	PM6	none	MPA	OPENBABEL	QM	0.887	0.745	0.615	0.816	44.7	48
QM	PM6	none	MPA	PUBCHEM	QM	0.884	0.753	0.643	0.824	43.6	48
QM	PM6	none	MPA	RDKIT	NO	0.883	0.757	0.617	0.830	43.0	48
Gasteiger				NCI	NO	0.737	1.135	0.882	1.243	16.0	48

4 Parameters of All the QSPR Models

Table S4: Parameters of all the QSPR models.

Charges	Theory level	Basis set	Charge scheme	3D source	Opt	$p_{p(H)}$	$p_{p(O)}$	$p_{p(C1)}$	$p_{p(OD)}$	$p_{p(C1D)}$			p_p
Phenols													
QM	B3LYP	6-31G*	AIM	NCI	NO	-148.787	68.843	-10.477	-26.976	-0.754			149.809
QM	B3LYP	6-31G*	MPA	NCI	NO	-212.618	-28.862	-4.259	-1.930	0.065			77.872
QM	HF	STO-3G	MPA	FROG2	QM	-21.823	-62.667	-47.245	-28.557	40.602			-14.151
QM	HF	STO-3G	MPA	FROG2	MM	-30.731	-63.859	16.886	-23.696	-101.353			-8.707
QM	B3LYP	6-31G*	MPA	OPENBABEL	MM	-103.783	-104.319	-14.560	-2.069	7.720			-14.786
QM	B3LYP	6-31G*	MPA	FROG2	QM	-37.001	-66.864	14.857	-15.398	-25.517			-25.108
QM	HF	STO-3G	MPA	PUBCHEM	MM	-38.342	-62.320	57.674	-26.794	-132.738			-11.247
QM	B3LYP	6-31G*	MPA	NCI	MM	-105.820	-96.301	-16.648	1.526	10.636			-6.402
QM	B3LYP	6-31G*	AIM	OPENBABEL	NO	83.264	21.740	-24.031	-60.302	-1.647			-66.798
QM	B3LYP	6-31G*	AIM	OPENBABEL	MM	95.924	35.781	-26.873	-65.038	-1.726			-62.081
QM	B3LYP	6-31G*	MPA	RDKIT	MM	-125.649	-44.555	-4.963	-3.138	-5.772			32.786
QM	B3LYP	6-31G*	MPA	OPENBABEL	NO	-103.097	-106.986	-13.888	-0.754	9.640			-16.777
QM	HF	STO-3G	MPA	OPENBABEL	MM	-87.846	-117.111	-36.778	-5.239	31.317			-6.034
QM	B3LYP	6-31G*	NPA	NCI	NO	-271.174	-10.215	6.658	-2.460	-1.264			132.277
QM	HF	STO-3G	MPA	PUBCHEM	NO	-70.040	-66.021	46.409	-10.727	-124.864			3.225
QM	B3LYP	6-31G*	NPA	NCI	MM	-86.729	-110.154	-15.759	0.344	6.234			-20.270
QM	B3LYP	6-31G*	NPA	OPENBABEL	MM	-89.133	-131.188	22.596	-7.861	-26.821			-41.844
QM	HF	STO-3G	MPA	NCI	MM	-85.423	-102.591	-37.305	0.186	9.856			1.870
QM	B3LYP	6-31G*	NPA	OPENBABEL	NO	-92.101	-123.490	31.884	-9.493	-39.782			-35.224
QM	B3LYP	6-31G*	MPA	OPENBABEL	QM	-67.377	-65.870	11.925	-6.980	-19.699			-6.837
QM	B3LYP	6-31G*	NPA	OPENBABEL	QM	-58.363	-77.501	-2.200	-3.421	-7.216			-14.608
QM	HF	STO-3G	MPA	OPENBABEL	QM	-70.181	-80.662	-25.938	-0.904	2.225			3.797
QM	HF	STO-3G	MPA	OPENBABEL	NO	-89.505	-110.242	-28.521	-4.158	18.624			-3.343
QM	B3LYP	6-31G*	NPA	FROG2	QM	-5.351	-57.571	-28.143	-14.138	26.275			-37.857
QM	B3LYP	6-31G*	MPA	PUBCHEM	NO	-111.673	-59.263	6.787	-7.572	-18.105			16.138
QM	B3LYP	6-31G*	AIM	RDKIT	MM	122.260	66.246	-31.978	-72.701	-6.094			-46.037
QM	HF	STO-3G	MPA	RDKIT	MM	-112.226	-71.568	-1.220	-0.625	-7.487			11.678
QM	HF	STO-3G	MPA	NCI	NO	-223.333	24.787	7.600	-1.588	2.175			62.803
QM	B3LYP	6-31G*	MPA	RDKIT	QM	-84.725	-55.804	-3.945	-1.456	-4.317			10.318
QM	HF	STO-3G	MPA	PUBCHEM	QM	-24.268	-50.702	19.498	-8.200	-102.028			0.966
QM	B3LYP	6-31G*	NPA	RDKIT	MM	-103.106	-72.983	5.749	-2.007	-6.895			7.951
QM	B3LYP	6-31G*	AIM	FROG2	QM	115.517	54.440	-77.818	-41.569	33.841			-23.165
QM	B3LYP	6-31G*	MPA	BALLOON	NO	-171.013	-76.436	4.578	-4.739	-11.853			29.419
QM	B3LYP	6-31G*	AIM	FROG2	MM	139.639	84.256	-63.845	-65.586	11.621			-22.878
QM	B3LYP	6-31G*	MPA	PUBCHEM	QM	-77.851	-71.872	24.297	-5.541	-35.123			-3.570
QM	B3LYP	6-31G*	AIM	OPENBABEL	QM	67.992	49.051	-25.490	-54.499	-0.354			-21.343
QM	B3LYP	6-31G*	AIM	NCI	MM	83.578	29.363	-25.529	-58.992	-4.545			-54.240
QM	HF	STO-3G	MPA	RDKIT	QM	-81.474	-80.312	-5.710	0.338	-10.099			5.089

Table S4: Parameters of all the QSPR models.

QM	B3LYP	6-31G*	NPA	PUBCHEM	NO	-109.690	-95.493	40.990	-5.110	-54.720			-1.689
QM	B3LYP	6-31G*	MPA	NCI	QM	-60.782	-70.142	29.203	-11.332	-36.757			-14.840
QM	B3LYP	6-31G*	NPA	RDKIT	MM	-102.447	-129.536	40.266	-7.931	-45.275			-34.065
QM	B3LYP	6-31G*	AIM	RDKIT	MM	-19.859	19.162	-20.307	-26.553	1.554			22.624
QM	B3LYP	6-31G*	NPA	BALLOON	MM	-115.680	-142.865	85.866	-13.893	-90.101			-41.765
QM	B3LYP	6-31G*	NPA	RDKIT	QM	-82.642	-72.231	3.719	-1.906	-7.988			0.311
QM	B3LYP	6-31G*	MPA	FROG2	MM	-75.478	-77.263	10.052	-20.448	-23.115			-19.327
QM	HF	STO-3G	MPA	RDKIT	MM	-100.858	-135.117	-4.162	-0.416	-7.724			-7.803
QM	B3LYP	6-31G*	NPA	PUBCHEM	MM	-147.206	-153.802	100.280	-6.236	-118.989			-23.000
Chavez2006	B3LYP	6-31G*	MPA	NCI	MM	-461.096	700.773	-134.395	-204.816	66.752			220.716
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	MM	-144.230	846.774	-253.680	-261.328	107.702			350.811
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	QM	124.590	-1400.682	550.853	1313.910	-671.286			-93.804
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	QM	-145.995	519.502	-87.007	-36.928	-12.134			300.063
Chavez2006	B3LYP	6-31G*	MPA	NCI	QM	-447.837	306.349	31.932	174.440	-100.320			213.690
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	MM	-69.845	-32.771	58.986	11.220	-79.442			13.381
QM	B3LYP	6-31G*	AIM	RDKIT	QM	43.042	41.942	-24.286	-39.425	-1.815			1.920
QM	B3LYP	6-31G*	NPA	PUBCHEM	QM	-91.780	-107.495	16.448	11.028	-30.183			-5.981
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	QM	356.912	-479.552	75.427	115.060	-80.851			-152.080
Svob2007-chal2	HF	STO-3G	MPA	FROG2	QM	14.166	-61.594	-18.597	-25.249	-18.474			-18.604
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	NO	-253.000	715.206	-210.017	-714.287	275.413			36.343
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	MM	-259.120	457.715	-96.286	-471.948	168.525			31.707
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	QM	-275.211	99.730	66.422	-80.041	-5.010			43.778
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	NO	339.748	-106.506	-76.055	-441.879	158.548			-209.101
Svob2007-chal2	HF	STO-3G	MPA	FROG2	MM	19.192	-70.789	-17.695	-26.224	-19.078			-22.958
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	QM	56.508	-135.015	38.610	-54.993	-32.155			-111.688
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	MM	57.554	-68.508	4.927	-53.309	-31.371			-78.904
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	MM	354.446	-211.338	-34.468	-341.877	112.440			-212.692
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	QM	-232.034	167.968	18.275	-37.203	-20.224			75.014
QM	HF	STO-3G	MPA	BALLOON	MM	-120.968	-152.436	35.948	1.565	-33.333			-11.185
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	MM	-362.928	76.150	112.340	11.735	-48.146			75.306
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	QM	-82.963	263.193	-5.276	-105.409	-6.583			115.221
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	MM	-269.919	692.866	-282.118	-373.247	207.394			302.644
QM	HF	STO-3G	MPA	BALLOON	QM	-71.164	-87.875	-16.229	0.779	-14.185			2.502
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	MM	107.804	-794.854	292.060	737.901	-403.201			-70.239
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	NO	108.035	-318.110	67.415	128.884	-108.941			-131.684
QM	B3LYP	6-31G*	MPA	FROG2	NO	-78.577	35.244	15.820	-32.257	-41.695			52.731
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	MM	-101.269	302.779	-1.312	-91.909	-17.411			147.971
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	NO	-72.190	550.466	-161.357	-272.956	84.728			170.227
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	QM	-262.634	520.157	-180.308	-238.688	123.835			268.308
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	MM	-73.524	487.708	-129.088	-225.764	60.429			162.914
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	QM	-66.598	380.949	-84.488	-114.313	5.141			161.771
QM	B3LYP	6-31G*	MPA	PUBCHEM	MM	-16.222	10.482	10.517	-55.906	-28.505			-12.354
QM	B3LYP	6-31G*	MPA	RDKIT	MM	-109.982	-109.751	-12.187	-1.154	4.897			-14.805
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	QM	-46.608	-35.167	30.617	1.549	-56.191			7.526
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	MM	133.070	-208.203	49.191	132.252	-98.406			-34.083

Table S4: Parameters of all the QSPR models.

Chavez2006	B3LYP	6-31G*	MPA	BALLOON	QM	-49.378	-192.982	109.881	569.899	-286.723			132.875
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	NO	-144.561	521.731	-221.815	-477.089	250.547			52.384
QM	B3LYP	6-31G*	NPA	BALLOON	NO	-153.575	-74.399	61.802	-20.659	-79.278			22.351
Svob2007-chal2	HF	STO-3G	MPA	FROG2	NO	13.948	-67.394	-5.278	-17.375	-40.767			-21.428
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	MM	-149.261	462.545	-183.959	-427.840	220.372			43.614
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	MM	-40.263	508.819	-172.480	-189.104	43.265			177.326
QM	B3LYP	6-31G*	NPA	FROG2	MM	-79.711	-94.717	20.086	-11.296	-38.277			-19.480
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	NO	-236.467	489.838	-112.179	-302.672	88.025			93.191
Chavez2006	B3LYP	6-31G*	MPA	FROG2	QM	105.948	-62.566	-17.524	-21.981	-21.835			-33.035
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	MM	-89.088	145.304	-21.204	110.703	-87.250			98.862
Chavez2006	B3LYP	6-31G*	MPA	FROG2	MM	95.920	45.640	-60.100	-21.522	-21.626			3.306
QM	B3LYP	6-31G*	NPA	FROG2	NO	-20.406	-47.799	-5.355	-17.272	20.161			-29.421
Svob2007-chal2	HF	STO-3G	MPA	NCI	NO	-131.519	9.741	87.963	13.989	-98.898			38.947
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	QM	-32.549	350.628	-104.467	-17.418	-41.135			179.829
Chavez2006	B3LYP	6-31G*	MPA	NCI	NO	505.873	-483.235	29.086	-19.425	-31.140			-214.367
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	NO	11.492	-40.804	23.403	248.205	-148.852			71.358
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	NO	-48.590	617.610	-206.640	-356.892	114.202			151.787
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	QM	-137.326	460.765	-188.830	-372.399	187.373			91.368
QM	B3LYP	6-31G*	AIM	BALLOON	NO	-15.842	1.709	-4.880	-45.767	-7.163			-23.369
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	NO	-0.017	562.916	-235.927	-267.591	79.222			149.359
QM	B3LYP	6-31G*	AIM	NCI	QM	56.507	31.520	-38.763	-21.031	-1.522			10.579
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	MM	41.117	177.103	-92.390	-105.693	-2.983			22.773
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	NO	213.795	545.438	-407.952	-908.523	517.172			-321.324
QM	B3LYP	6-31G*	NPA	BALLOON	QM	-74.855	-66.199	9.797	-5.085	-24.144			1.419
QM	B3LYP	6-31G*	AIM	PUBCHEM	NO	-14.018	-18.018	30.228	-34.378	-41.552			-28.816
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	MM	-26.281	-30.298	-5.367	-21.256	-15.446			-1.036
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	NO	-26.365	-24.879	-14.064	-25.134	-5.365			-0.258
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	NO	145.328	31.145	-132.190	-197.174	46.107			-134.355
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	MM	-2.995	-85.799	34.834	-0.210	-59.981			-18.297
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	QM	205.170	-86.819	-25.228	-51.657	-1.786			-119.977
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	MM	233.713	312.462	-274.707	-700.103	387.131			-343.943
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	QM	35.743	6.518	-121.772	-81.297	94.113			-14.837
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	NO	-0.756	705.294	-338.256	58.724	-42.436			368.762
Chavez2006	B3LYP	6-31G*	MPA	FROG2	NO	-69.572	2452.333	-1021.007	-1094.902	637.732			612.658
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	QM	-35.513	-67.554	58.448	15.328	-85.335			-1.802
QM	HF	STO-3G	MPA	BALLOON	NO	-146.790	-27.059	1.069	-0.137	-38.249			34.576
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	NO	5.210	-51.959	-12.745	-21.405	-15.646			-13.231
Svob2007-chal2	HF	STO-3G	MPA	NCI	QM	-44.510	-42.602	51.711	13.154	-79.865			7.955
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	QM	-111.463	393.171	-157.145	-324.003	154.908			73.288
QM	B3LYP	6-31G*	MPA	RDKIT	NO	-120.552	-50.198	3.371	-4.380	-15.781			24.970
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	MM	-147.073	444.660	-168.494	-421.586	211.069			32.631
QM	HF	STO-3G	MPA	FROG2	NO	-150.874	-149.488	33.849	4.674	58.823			-7.628
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	NO	-70.040	-29.439	50.517	4.048	-64.316			11.794
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	MM	102.366	329.596	-215.463	-444.923	223.967			-97.502
QM	B3LYP	6-31G*	AIM	BALLOON	MM	56.693	6.631	-2.098	-58.586	-5.880			-75.581

Table S4: Parameters of all the QSPR models.

QM	HF	STO-3G	MPA	NCI	QM	-46.474	-9.843	-41.622	-0.158	3.429			21.667
Svob2007-chal2	HF	STO-3G	MPA	NCI	MM	-46.311	-22.777	17.016	-2.318	-43.028			11.100
QM	B3LYP	6-31G*	AIM	FROG2	NO	-33.636	-57.726	-13.680	-14.946	40.748			-81.607
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	QM	-5.072	736.921	-413.574	-452.815	235.587			262.923
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	NO	411.717	92.415	-222.626	-415.159	211.456			-290.789
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	NO	34.098	711.344	-412.412	-547.202	287.919			154.271
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	MM	28.288	-9.806	-71.546	-64.340	38.195			-14.907
QM	B3LYP	6-31G*	AIM	PUBCHEM	QM	94.460	65.279	-26.084	-48.024	-10.254			-4.224
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	NO	-36.588	467.582	-230.403	-437.911	217.119			35.413
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	MM	-26.241	501.270	-257.676	-414.549	205.931			86.550
QM	B3LYP	6-31G*	AIM	RDKIT	NO	-39.183	-9.365	-15.037	-19.334	-2.480			9.293
QM	B3LYP	6-31G*	MPA	BALLOON	MM	-116.862	-52.018	-35.410	6.166	25.786			30.613
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	NO	-45.644	-71.032	69.436	17.654	-93.891			-0.920
QM	B3LYP	6-31G*	MPA	BALLOON	QM	-81.952	-55.494	5.110	-4.388	-13.926			7.701
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	QM	-35.008	-66.742	73.235	17.670	-109.681			-0.780
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	QM	21.044	71.373	-51.453	-12.907	-11.944			62.298
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	MM	19.946	3.541	-10.366	-13.301	-12.365			1.305
QM	B3LYP	6-31G*	AIM	PUBCHEM	MM	40.555	61.890	75.862	-86.436	-97.222			-15.713
QM	B3LYP	6-31G*	NPA	NCI	QM	-92.378	-10.831	-9.001	-4.633	-9.171			49.772
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	NO	-65.694	254.164	-129.134	462.243	-326.119			457.007
QM	B3LYP	6-31G*	AIM	BALLOON	QM	-6.390	0.618	-5.616	-44.303	-5.116			-29.087
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	MM	-32.314	-52.090	46.248	-0.274	-70.197			-2.202
QM	B3LYP	6-31G*	NPA	RDKIT	NO	-111.075	-23.794	20.978	-15.553	-31.792			38.192
QM	HF	STO-3G	MPA	RDKIT	NO	-84.867	-46.827	-35.345	-1.363	-6.946			15.247
QM	PM6	none	MPA	NCI	MM	-45.470	-74.491	-8.415	3.523	4.298			-6.504
QM	PM6	none	MPA	FROG2	QM	36.116	-4.446	-13.485	-25.682	17.317			-29.776
QM	PM6	none	MPA	OPENBABEL	MM	-46.711	-80.412	-10.101	4.347	9.883			-10.552
QM	PM6	none	MPA	RDKIT	MM	-94.613	-43.921	-1.313	-0.668	-0.620			19.898
QM	PM6	none	MPA	OPENBABEL	QM	-32.757	-51.170	-2.562	0.009	-0.508			-1.997
QM	PM6	none	MPA	NCI	NO	-253.670	-8.884	7.869	-1.377	-1.076			86.978
QM	PM6	none	MPA	OPENBABEL	NO	-49.592	-76.423	-7.619	2.507	6.409			-8.263
QM	PM6	none	MPA	FROG2	MM	45.864	4.257	-8.954	-37.790	9.149			-35.922
QM	PM6	none	MPA	RDKIT	QM	-45.797	-46.260	-1.294	-0.164	-0.973			4.117
QM	PM6	none	MPA	FROG2	NO	0.777	-23.982	-8.733	-8.078	16.825			-12.464
QM	PM6	none	MPA	PUBCHEM	NO	-21.179	-20.006	-6.193	-15.516	4.364			-5.500
QM	PM6	none	MPA	RDKIT	MM	-47.585	-78.692	-2.699	0.058	0.135			-10.154
QM	PM6	none	MPA	PUBCHEM	QM	37.749	5.788	-14.959	-29.451	18.644			-28.956
QM	PM6	none	MPA	PUBCHEM	MM	14.188	12.155	0.974	-41.664	-7.919			-19.488
QM	PM6	none	MPA	BALLOON	NO	-132.186	-42.520	3.406	-6.331	-9.562			31.078
QM	PM6	none	MPA	BALLOON	MM	-82.824	-52.984	1.569	-2.330	-4.766			12.026
QM	PM6	none	MPA	BALLOON	QM	-47.875	-43.797	-3.599	0.655	-0.440			6.877
QM	PM6	none	MPA	NCI	QM	-69.906	-18.753	-2.290	-1.091	-1.924			24.314
QM	PM6	none	MPA	RDKIT	NO	-99.028	-38.700	1.483	-6.131	-6.444			20.736
Gasteiger				NCI	NO	1.899E5	-9.651E3	-68.18	15.03	-22.86			-6.055E4

Table S4: Parameters of all the QSPR models.

Charges	Theory level	Basis set	Charge scheme	3D source	Opt	$p_c(H)$	$p_{c(O1)}$	$p_{c(O2)}$	$p_{c(C1)}$	$p_{c(O1D)}$	$p_{c(O2D)}$	$p_{c(C1D)}$	p_c
Carboxylic acids													
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	MM	5.660	-2.024	157.129	-336.189	196.873	-321.673	465.338	158.572
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	NO	16.756	-15.745	134.099	-296.924	177.660	-306.211	452.941	151.231
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	QM	5.581	-92.125	217.453	-376.584	275.655	-395.768	545.932	172.418
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	MM	-71.256	297.299	-37.309	-68.863	217.923	-156.556	106.367	301.754
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	NO	52.103	145.222	-127.420	14.164	-80.277	47.974	-28.551	7.774
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	QM	-29.615	298.609	2.781	-240.028	54.522	-154.696	265.100	200.538
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	MM	281.290	-572.821	297.688	-652.051	260.868	-337.221	441.573	-364.733
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	NO	384.857	-1045.866	403.343	-634.727	394.138	-373.000	386.034	-591.853
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	QM	271.499	-655.080	346.878	-657.069	320.758	-382.698	479.087	-356.223
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	MM	-19.875	-23.596	43.282	-22.083	214.869	-195.837	179.426	180.025
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	NO	-24.430	-19.257	62.265	-57.086	228.479	-213.396	202.909	185.022
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	QM	-18.604	-11.637	38.813	-27.948	202.049	-208.670	222.239	197.261
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	MM	231.010	-537.009	212.549	-414.190	438.007	-250.992	91.782	-295.340
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	NO	169.464	-593.786	282.774	-380.137	524.463	-336.352	184.233	-195.178
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	QM	70.174	-148.480	-34.996	65.015	373.519	-98.035	-169.031	33.618
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	MM	-11.842	-32.019	164.039	-286.916	197.413	-362.713	554.674	218.172
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	QM	4.783	-137.215	186.826	-266.511	198.312	-386.922	607.307	196.338
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	MM	-13.714	197.151	50.092	-263.108	72.628	-218.285	369.971	192.626
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	NO	14.537	177.713	-25.676	-146.000	73.728	-128.306	182.138	137.728
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	MM	33.450	-15.407	68.048	-117.743	122.454	-256.391	296.634	263.086
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	NO	21.090	-291.803	372.571	-350.008	424.276	-571.082	526.908	284.840
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	QM	42.020	108.251	-132.148	90.288	-46.931	-53.235	125.322	254.586
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	MM	-57.222	-200.218	325.372	-300.009	444.998	-566.812	499.029	407.739
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	NO	5.108	116.333	-28.991	-73.418	-1.466	-124.509	209.589	232.356
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	QM	-48.354	-37.667	306.623	-437.160	259.689	-518.381	604.134	358.054
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	MM	320.851	-386.207	269.070	-314.266	315.021	-368.453	294.710	-82.901
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	NO	56.700	-608.466	259.596	125.477	813.697	-472.136	-29.880	276.404
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	QM	-490.529	509.469	-240.061	287.398	41.819	-77.118	76.070	841.393
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	MM	18.490	-130.668	168.828	-161.957	281.513	-359.678	310.746	275.787
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	NO	20.876	-72.130	88.628	-86.575	212.249	-274.348	236.304	266.827
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	QM	17.271	-9.208	35.065	-56.806	152.793	-233.682	229.928	291.163
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	MM	525.776	-621.922	339.631	-359.527	435.683	-389.468	210.654	-307.706
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	NO	-192.456	-172.679	257.602	-109.889	569.410	-523.634	302.995	542.288
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	QM	99.178	-350.153	264.041	-171.963	501.392	-442.698	236.279	198.389
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	MM	-8.038	-124.530	141.670	-103.005	236.395	-335.815	317.138	299.744
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	QM	26.707	143.207	-154.681	105.585	-101.911	-24.682	132.211	256.795
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	MM	-1.907	-49.644	248.709	-359.985	235.226	-460.297	526.923	324.037
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	NO	40.687	-56.930	166.603	-247.142	217.550	-346.150	350.841	242.796
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	MM	97.802	-73.097	29.697	-88.849	330.501	-330.048	443.662	192.629
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	NO	102.315	-71.675	25.438	-83.452	350.724	-330.692	425.608	195.139
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	QM	105.230	-102.312	4.802	-6.469	319.543	-309.521	408.174	193.992

Table S4: Parameters of all the QSPR models.

Chavez2006	B3LYP	6-31G*	MPA	FROG2	MM	-155.702	359.759	-129.660	76.968	269.156	-176.421	141.964	298.224
Chavez2006	B3LYP	6-31G*	MPA	FROG2	NO	182.399	107.877	-77.854	-88.751	-13.061	-106.647	263.170	71.902
Chavez2006	B3LYP	6-31G*	MPA	FROG2	QM	-10.239	205.002	-59.193	-63.385	222.871	-192.302	227.318	198.522
Chavez2006	B3LYP	6-31G*	MPA	NCI	MM	797.966	-645.970	143.818	-389.905	416.109	-317.077	330.955	-179.437
Chavez2006	B3LYP	6-31G*	MPA	NCI	NO	973.543	-912.196	121.033	-223.582	466.071	-253.415	128.282	-281.567
Chavez2006	B3LYP	6-31G*	MPA	NCI	QM	1046.615	-989.275	251.934	-522.494	465.808	-362.644	391.639	-326.545
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	MM	-11.824	76.532	-27.419	-6.678	297.598	-281.010	358.596	241.940
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	NO	-6.511	55.119	-15.673	-18.786	314.113	-285.270	352.469	233.805
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	QM	11.990	54.410	-38.133	20.770	277.029	-266.802	347.733	230.859
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	MM	431.222	-194.415	-23.571	-128.762	404.974	-158.834	-32.918	-27.905
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	NO	210.716	-115.170	-42.881	25.732	425.662	-186.223	11.559	94.111
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	QM	230.822	-147.474	-70.606	108.663	484.734	-174.832	-74.257	99.564
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	MM	59.405	-87.858	-66.438	189.287	296.181	-245.704	276.501	216.554
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	QM	87.820	-116.010	-15.118	70.037	245.446	-297.739	455.840	207.748
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	MM	29.791	127.863	-26.555	-94.264	241.802	-285.918	426.688	229.687
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	NO	80.694	199.931	-94.738	-46.126	181.088	-192.810	264.542	190.161
QM	B3LYP	6-31G*	AIM	BALLOON	MM	-63.728	-54.960	19.797	-7.834	5.106	-25.333	-15.089	-34.151
QM	B3LYP	6-31G*	AIM	BALLOON	NO	-53.678	-27.855	15.113	-18.649	-7.936	-19.590	-6.648	-28.130
QM	B3LYP	6-31G*	AIM	BALLOON	QM	-50.649	-42.027	-5.870	-8.144	12.450	-9.996	-15.932	-4.158
QM	B3LYP	6-31G*	AIM	FROG2	MM	-127.169	-65.856	11.145	18.069	12.412	-24.925	-23.447	32.238
QM	B3LYP	6-31G*	AIM	FROG2	NO	-82.679	-30.427	7.254	-20.609	-0.495	-11.550	3.925	5.843
QM	B3LYP	6-31G*	AIM	FROG2	QM	-105.857	-86.979	3.307	12.032	17.266	-13.709	-35.490	-27.145
QM	B3LYP	6-31G*	AIM	NCI	MM	-78.593	-36.740	0.764	-13.083	-1.924	-6.575	2.714	3.562
QM	B3LYP	6-31G*	AIM	NCI	NO	-86.975	-43.130	-26.823	1.452	13.184	20.702	-17.666	8.932
QM	B3LYP	6-31G*	AIM	NCI	QM	-129.639	-61.131	-34.376	0.754	9.796	28.046	-12.366	16.035
QM	B3LYP	6-31G*	AIM	OPENBABEL	MM	-61.748	-47.250	28.868	-8.158	3.392	-33.656	-18.309	-34.195
QM	B3LYP	6-31G*	AIM	OPENBABEL	NO	-63.491	-50.684	23.299	-5.149	6.487	-29.610	-21.490	-31.740
QM	B3LYP	6-31G*	AIM	OPENBABEL	QM	-41.586	-25.615	3.196	-14.497	7.922	-15.764	-11.209	-2.858
QM	B3LYP	6-31G*	AIM	PUBCHEM	MM	-97.131	-51.887	-6.701	7.003	8.999	-5.810	-16.618	19.432
QM	B3LYP	6-31G*	AIM	PUBCHEM	NO	-98.414	-55.191	10.810	-9.893	8.537	-11.618	-14.820	-20.183
QM	B3LYP	6-31G*	AIM	PUBCHEM	QM	-119.906	-66.059	-6.349	7.223	11.874	-4.869	-18.457	16.022
QM	B3LYP	6-31G*	AIM	RDKIT	MM	-44.464	-12.475	1.443	-19.805	-5.333	-11.142	1.672	4.794
QM	B3LYP	6-31G*	AIM	RDKIT	QM	-48.841	-31.854	5.823	-20.810	14.153	-19.562	-11.792	-4.775
QM	B3LYP	6-31G*	AIM	RDKIT	MM	-75.485	-55.192	42.750	-17.010	0.875	-38.820	-8.020	-48.825
QM	B3LYP	6-31G*	AIM	RDKIT	NO	-25.977	-13.236	-4.652	-11.056	-7.318	-3.878	-0.739	-3.140
QM	B3LYP	6-31G*	MPA	BALLOON	MM	-46.400	-31.802	-4.438	2.503	2.669	11.541	-25.139	-9.302
QM	B3LYP	6-31G*	MPA	BALLOON	NO	-96.292	-19.883	-0.570	1.697	13.859	5.612	-34.771	20.483
QM	B3LYP	6-31G*	MPA	BALLOON	QM	-43.355	-36.315	5.119	-6.806	1.259	1.164	-6.310	-7.849
QM	B3LYP	6-31G*	MPA	FROG2	MM	-119.729	-19.233	-4.210	4.794	3.423	5.183	-10.825	40.700
QM	B3LYP	6-31G*	MPA	FROG2	NO	-94.218	11.089	-22.854	-24.739	-10.479	20.142	7.313	38.268
QM	B3LYP	6-31G*	MPA	FROG2	QM	-115.943	-5.063	-10.804	-19.526	-6.596	7.852	7.379	41.750
QM	B3LYP	6-31G*	MPA	NCI	MM	-89.187	2.855	-11.694	-19.967	-12.147	12.614	12.493	32.123
QM	B3LYP	6-31G*	MPA	NCI	NO	-52.264	13.735	-27.043	-13.591	-15.165	30.800	-7.089	13.501
QM	B3LYP	6-31G*	MPA	NCI	QM	-124.117	-3.548	-10.004	-14.111	-6.587	9.121	7.124	47.143
QM	B3LYP	6-31G*	MPA	OPENBABEL	MM	-52.514	-23.919	1.670	-0.285	6.558	8.144	-21.637	-0.830

Table S4: Parameters of all the QSPR models.

QM	B3LYP	6-31G*	MPA	OPENBABEL	NO	-52.183	-23.431	5.574	-3.557	4.474	4.263	-14.714	0.053
QM	B3LYP	6-31G*	MPA	OPENBABEL	QM	-46.440	-25.036	12.105	-6.014	4.509	-0.336	-8.612	-2.209
QM	B3LYP	6-31G*	MPA	PUBCHEM	MM	-114.682	13.936	-18.214	-15.628	-11.226	17.255	13.753	53.622
QM	B3LYP	6-31G*	MPA	PUBCHEM	NO	-95.030	8.563	-34.354	-13.939	-5.522	36.531	-7.886	34.667
QM	B3LYP	6-31G*	MPA	PUBCHEM	QM	-129.642	5.427	-22.994	-16.920	-7.800	21.217	9.949	55.354
QM	B3LYP	6-31G*	MPA	RDKIT	MM	-65.422	2.671	-15.796	-12.105	-3.099	18.621	-12.547	18.058
QM	B3LYP	6-31G*	MPA	RDKIT	QM	-46.892	-25.618	5.226	-1.538	-0.777	5.700	-3.935	-0.292
QM	B3LYP	6-31G*	MPA	RDKIT	MM	-86.441	-50.778	-16.525	23.453	17.327	27.165	-39.648	7.329
QM	B3LYP	6-31G*	MPA	RDKIT	NO	-55.620	-28.947	-8.915	-16.143	1.805	6.243	-4.712	3.022
QM	B3LYP	6-31G*	NPA	BALLOON	MM	-50.642	-34.515	-12.235	-8.471	6.929	21.498	-28.462	-20.843
QM	B3LYP	6-31G*	NPA	BALLOON	NO	-50.223	-19.919	-19.509	-11.353	0.247	29.725	-17.479	-10.023
QM	B3LYP	6-31G*	NPA	BALLOON	QM	-52.115	-60.544	-10.926	1.259	11.710	18.279	-39.955	-35.932
QM	B3LYP	6-31G*	NPA	FROG2	MM	-143.422	-58.006	-14.624	15.154	19.366	18.717	-38.769	29.813
QM	B3LYP	6-31G*	NPA	FROG2	NO	-136.507	-38.276	8.353	-39.008	-1.556	-16.136	15.345	37.317
QM	B3LYP	6-31G*	NPA	FROG2	QM	-125.013	-54.043	-4.554	-23.164	9.657	-0.574	-7.486	21.012
QM	B3LYP	6-31G*	NPA	NCI	MM	-97.668	-21.773	-14.865	-15.327	-0.919	18.863	-6.413	21.150
QM	B3LYP	6-31G*	NPA	NCI	NO	-116.015	-76.100	15.062	4.267	15.802	-11.866	-25.240	2.025
QM	B3LYP	6-31G*	NPA	NCI	QM	-125.577	-45.238	-2.376	-16.585	5.398	0.821	-8.356	24.401
QM	B3LYP	6-31G*	NPA	OPENBABEL	MM	-87.166	-29.267	1.375	-0.917	9.752	9.580	-23.175	9.454
QM	B3LYP	6-31G*	NPA	OPENBABEL	NO	-83.446	-26.746	-1.418	-1.476	9.083	12.587	-23.335	8.428
QM	B3LYP	6-31G*	NPA	OPENBABEL	QM	-82.901	-28.474	3.360	-5.253	9.013	4.121	-16.940	11.523
QM	B3LYP	6-31G*	NPA	PUBCHEM	MM	-128.687	-30.331	0.156	-20.921	1.447	-1.825	2.905	38.882
QM	B3LYP	6-31G*	NPA	PUBCHEM	NO	-121.712	-38.537	5.029	-1.663	11.127	0.276	-21.577	25.870
QM	B3LYP	6-31G*	NPA	PUBCHEM	QM	-123.782	-36.049	7.864	-35.225	-2.694	-14.973	14.930	32.457
QM	B3LYP	6-31G*	NPA	RDKIT	MM	-77.875	-49.366	6.124	-12.479	9.695	-4.999	-13.406	-2.589
QM	B3LYP	6-31G*	NPA	RDKIT	QM	-73.611	-28.225	-8.057	-6.807	3.477	14.606	-14.714	4.785
QM	B3LYP	6-31G*	NPA	RDKIT	MM	-66.963	-50.661	-23.414	-5.537	19.143	32.528	-33.628	-15.023
QM	B3LYP	6-31G*	NPA	RDKIT	NO	-51.926	-23.764	-3.489	-17.613	6.071	7.580	-6.672	0.700
QM	HF	STO-3G	MPA	BALLOON	MM	-65.809	-59.120	-9.047	-14.110	6.102	33.624	-32.713	-19.210
QM	HF	STO-3G	MPA	BALLOON	NO	-53.287	-66.783	10.040	-27.559	7.114	19.565	-20.695	-25.212
QM	HF	STO-3G	MPA	BALLOON	QM	-87.214	-100.770	-19.708	-12.198	14.569	28.864	-45.884	-24.294
QM	HF	STO-3G	MPA	FROG2	MM	-113.082	-89.509	-58.668	10.339	8.509	63.325	-43.403	-5.206
QM	HF	STO-3G	MPA	FROG2	NO	-99.620	-54.758	-54.535	-4.554	10.803	72.521	-31.998	2.432
QM	HF	STO-3G	MPA	FROG2	QM	-79.633	-137.152	-29.677	-5.831	21.414	37.528	-53.903	-33.856
QM	HF	STO-3G	MPA	NCI	MM	-80.853	44.528	-73.488	-42.068	-37.722	72.244	4.866	15.008
QM	HF	STO-3G	MPA	NCI	NO	-92.614	-128.833	6.998	5.363	24.070	4.339	-40.049	-24.294
QM	HF	STO-3G	MPA	NCI	QM	-99.369	-87.013	-23.474	-7.445	10.876	38.481	-48.718	-20.239
QM	HF	STO-3G	MPA	OPENBABEL	MM	-72.697	-74.777	15.352	-21.465	14.590	4.056	-20.501	-16.267
QM	HF	STO-3G	MPA	OPENBABEL	NO	-69.181	-75.130	12.716	-21.106	13.264	8.360	-21.900	-18.398
QM	HF	STO-3G	MPA	OPENBABEL	QM	-85.337	-77.759	14.656	-17.669	23.400	4.332	-34.323	-15.232
QM	HF	STO-3G	MPA	PUBCHEM	MM	-94.349	7.240	-107.593	-19.998	-26.149	107.707	-16.836	12.685
QM	HF	STO-3G	MPA	PUBCHEM	NO	-63.295	-164.370	14.203	0.007	37.979	7.432	-61.145	-47.385
QM	HF	STO-3G	MPA	PUBCHEM	QM	-73.185	-11.171	-0.204	-47.164	-16.421	-4.672	3.732	-2.562
QM	HF	STO-3G	MPA	RDKIT	MM	-101.836	-47.152	1.090	-16.130	14.854	7.303	-25.140	1.872
QM	HF	STO-3G	MPA	RDKIT	QM	-90.446	-71.434	-0.866	-12.877	14.246	16.551	-34.556	-13.263

Table S4: Parameters of all the QSPR models.

QM	HF	STO-3G	MPA	RDKIT	MM	-85.596	-65.334	-27.513	-6.789	17.849	48.626	-43.275	-10.595
QM	HF	STO-3G	MPA	RDKIT	NO	-65.317	-27.260	-8.827	-23.010	11.944	17.603	-19.139	1.265
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	MM	-9.498	-16.478	14.793	-28.447	-0.498	1.234	-7.717	-13.919
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	NO	-13.816	-14.415	22.766	-38.179	6.979	-18.180	1.125	-7.849
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	QM	-16.156	-18.713	20.455	-30.348	3.519	-7.391	-6.068	-11.820
Svob2007-chal2	HF	STO-3G	MPA	FROG2	MM	-25.740	26.930	10.502	-51.298	-18.627	11.943	6.427	-4.418
Svob2007-chal2	HF	STO-3G	MPA	FROG2	NO	-27.517	-7.553	20.011	-29.586	-1.371	-15.144	8.166	-0.834
Svob2007-chal2	HF	STO-3G	MPA	FROG2	QM	-11.791	15.914	22.214	-67.482	-12.341	-11.045	18.939	-7.815
Svob2007-chal2	HF	STO-3G	MPA	NCI	MM	73.174	-56.880	24.662	-72.356	13.847	-19.980	-0.330	-48.024
Svob2007-chal2	HF	STO-3G	MPA	NCI	NO	65.320	-61.272	-5.649	-32.029	6.027	19.699	-31.055	-50.327
Svob2007-chal2	HF	STO-3G	MPA	NCI	QM	63.367	-52.811	8.459	-55.013	5.345	-2.222	-10.675	-45.786
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	MM	-24.650	-1.121	-0.736	-14.553	-0.602	18.537	-22.073	-6.270
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	NO	-24.932	0.426	0.121	-16.523	-0.483	17.708	-21.642	-6.096
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	QM	-29.518	-6.614	17.629	-26.718	4.936	-4.340	-9.131	-4.513
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	MM	79.115	-36.773	14.259	-84.073	1.644	-7.268	3.385	-48.340
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	NO	90.471	-56.855	8.327	-73.529	11.237	-6.535	-7.882	-53.555
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	QM	40.836	-47.736	30.916	-63.447	24.066	-21.879	-12.257	-37.779
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	MM	-24.312	-25.243	10.456	-7.445	-0.805	2.533	-10.443	-8.040
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	QM	-21.331	-20.410	11.864	-13.065	-5.847	6.229	-6.633	-10.350
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	MM	-19.080	22.246	-10.798	-28.507	-19.401	35.173	-16.272	-8.197
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	NO	-11.879	12.224	8.576	-45.036	-6.589	0.930	0.324	-5.360
QM	PM6	none	MPA	NCI	NO	-60.643	-54.911	6.001	18.382	13.583	-10.455	-39.876	-6.835
QM	PM6	none	MPA	NCI	MM	-67.460	-30.320	-7.383	-5.584	6.195	7.349	-13.610	4.286
QM	PM6	none	MPA	PUBCHEM	NO	-72.763	-40.699	1.661	1.580	15.317	-2.243	-25.129	5.065
QM	PM6	none	MPA	NCI	QM	-63.098	-32.220	-4.643	-3.409	6.908	3.821	-16.992	1.886
QM	PM6	none	MPA	FROG2	MM	-88.013	-41.695	-8.696	3.686	13.178	8.674	-22.132	10.170
QM	PM6	none	MPA	FROG2	QM	-51.819	-41.760	-7.862	-0.548	13.256	5.722	-23.993	-4.028
QM	PM6	none	MPA	PUBCHEM	MM	-74.905	-25.925	-7.516	-10.085	3.080	7.603	-7.623	8.059
QM	PM6	none	MPA	OPENBABEL	MM	-44.098	-30.929	-8.379	-3.782	10.285	7.451	-21.367	-4.497
QM	PM6	none	MPA	PUBCHEM	QM	-53.215	-23.402	-4.090	-10.177	-0.851	1.954	-8.791	-0.380
QM	PM6	none	MPA	OPENBABEL	NO	-45.106	-30.627	-8.759	-3.581	10.743	8.023	-22.053	-4.192
QM	PM6	none	MPA	OPENBABEL	QM	-42.859	-28.066	-5.447	-0.705	10.083	4.702	-23.535	-3.534
QM	PM6	none	MPA	BALLOON	QM	-15.418	-37.042	-0.694	-4.240	8.088	-2.768	-21.116	-16.294
QM	PM6	none	MPA	RDKIT	MM	-62.551	-25.746	-0.460	-6.769	5.834	-0.081	-12.619	5.247
QM	PM6	none	MPA	FROG2	NO	-32.089	-15.134	2.636	-6.360	-4.574	-2.826	-12.021	-7.567
QM	PM6	none	MPA	BALLOON	MM	-21.356	-34.009	-5.368	-9.780	6.044	2.928	-14.058	-13.642
Gasteiger				NCI	NO	-2.021E4	5.484E3	-7.313E2	-4.562E2	4.481E2	4.740E2	-5.107E2	8.686E3
QM	PM6	none	MPA	RDKIT	MM	-35.805	-38.066	-3.966	-10.032	13.788	2.115	-17.256	-6.939
QM	PM6	none	MPA	BALLOON	NO	-59.647	-19.241	-1.331	-6.397	7.696	1.847	-19.382	3.906
QM	PM6	none	MPA	RDKIT	NO	-57.645	-19.223	-1.042	-15.463	7.863	2.500	-4.999	8.445
QM	PM6	none	MPA	RDKIT	QM	-38.801	-26.462	-0.662	-3.415	7.774	0.037	-19.844	-4.405

Table S4: Parameters of all the QSPR models.

Charges	Theory level	Basis set	Charge scheme	3D source	Opt	$p_{a(H)}$	$p_{a(N)}$	$p_{a(C1)}$	$p_{a(HA)}$	$p_{a(NA)}$	$p_{a(C1A)}$		p_a
Anilines													
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	MM	1794.262	-982.507	-5439.152	2761.729	3620.073	-1913.329		-743.691
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	NO	-408.542	671.403	-274.770	519.656	-636.985	-36.076		-352.942
Bult2002-npa	B3LYP	6-31G*	NPA	BALLOON	QM	-2764.013	-653.293	5671.063	-2171.939	-414.422	1003.241		2718.701
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	MM	2545.247	-318.072	-4179.460	931.836	198.778	250.641		-1920.217
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	NO	-2286.345	-124.454	41.907	2130.640	2563.609	-2234.716		1149.385
Bult2002-npa	B3LYP	6-31G*	NPA	FROG2	QM	-1557.346	-39.313	4652.221	-2511.103	-2214.885	1898.902		1234.443
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	MM	3337.197	-1183.086	-6538.487	2174.638	2627.959	-904.054		-1836.368
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	NO	1213.606	-641.345	-2959.267	1208.884	1481.117	-503.583		-598.450
Bult2002-npa	B3LYP	6-31G*	NPA	NCI	QM	-2691.545	-790.741	2893.650	80.197	1988.065	-779.389		2682.216
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	MM	2240.907	-351.849	-4052.633	1133.185	562.418	-23.357		-1652.428
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	NO	1916.121	-305.800	-3339.137	839.414	273.267	171.533		-1435.857
Bult2002-npa	B3LYP	6-31G*	NPA	OPENBABEL	QM	-2084.349	-58.354	6487.016	-3526.909	-2988.772	2520.467		1716.896
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	MM	3162.844	-1003.675	-5410.372	1372.879	1487.346	-138.345		-1859.313
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	NO	-3897.061	497.380	4484.446	-77.815	292.477	-647.786		2599.303
Bult2002-npa	B3LYP	6-31G*	NPA	PUBCHEM	QM	-697.467	-891.363	1362.351	-591.414	433.737	594.883		1102.582
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	MM	2729.432	-2336.465	-6334.854	2608.803	5022.732	-1762.601		-422.083
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	QM	-2807.527	-677.340	2828.636	241.948	2036.145	-920.170		2679.353
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	MM	2208.244	-262.143	-3696.125	854.580	129.420	249.297		-1701.733
Bult2002-npa	B3LYP	6-31G*	NPA	RDKIT	NO	-546.581	1415.035	-70.866	480.771	-1840.716	232.359		-871.282
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	MM	-1268.655	-104.037	-3382.921	3497.992	3643.282	-3044.376		198.521
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	NO	-1176.272	1246.706	-3110.162	3212.916	2375.322	-2567.901		-464.171
Bult2004-aim	B3LYP	6-31G*	AIM	BALLOON	QM	-5797.975	-1920.996	2628.288	368.855	2949.621	-1654.017		3694.365
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	MM	-5773.511	352.839	450.065	2225.292	2770.384	-2505.841		2151.559
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	NO	-5326.670	-1341.065	1945.079	729.026	1808.769	-1086.823		1856.037
Bult2004-aim	B3LYP	6-31G*	AIM	FROG2	QM	-3626.635	-1476.609	74.034	1574.672	3313.917	-2148.192		2255.368
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	MM	-4435.854	302.813	-1364.432	3198.049	3674.487	-3259.543		1549.470
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	NO	-4528.503	902.382	-152.420	2173.562	2020.362	-2113.168		1261.558
Bult2004-aim	B3LYP	6-31G*	AIM	NCI	QM	-3735.693	-1388.370	-1192.894	2732.023	4259.262	-2997.757		2053.622
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	MM	-4841.096	182.134	-249.535	2406.890	2878.904	-2521.463		1763.018
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	NO	-6166.064	1.922	458.599	2399.178	3245.449	-2754.250		2473.781
Bult2004-aim	B3LYP	6-31G*	AIM	OPENBABEL	QM	-3470.305	-1405.616	161.529	1424.325	3043.956	-1947.745		2135.169
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	MM	-4198.955	120.097	-1101.460	2856.184	3312.780	-2864.833		1486.809
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	NO	-6131.106	492.395	1220.164	1730.513	2386.306	-2249.337		2399.513
Bult2004-aim	B3LYP	6-31G*	AIM	PUBCHEM	QM	-2913.591	-1703.695	-1423.262	2550.101	4232.452	-2828.150		1880.773
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	MM	-8296.387	6654.365	1549.699	2431.967	-1022.048	-2113.250		349.703
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	QM	-3959.791	-461.961	-1499.376	3095.336	4185.176	-3347.141		1794.812
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	MM	-5056.691	198.552	-267.796	2521.615	3060.666	-2683.650		1870.160
Bult2004-aim	B3LYP	6-31G*	AIM	RDKIT	NO	-1659.862	5288.330	-2048.178	2508.321	-1079.062	-1437.820		-2062.447
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	MM	1928.255	-1174.730	-3810.481	2030.357	2271.616	-1173.071		-469.556
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	NO	-976.440	1439.180	-1920.683	1689.016	637.683	-925.504		-381.170
Chavez2006	B3LYP	6-31G*	MPA	BALLOON	QM	-10642.070	-138.437	6328.491	-827.399	1236.013	-837.785		3167.770

Table S4: Parameters of all the QSPR models.

Chavez2006	B3LYP	6-31G*	MPA	FROG2	MM	-6756.797	-909.421	1692.459	1121.726	2103.076	-1152.380		1849.583
Chavez2006	B3LYP	6-31G*	MPA	FROG2	NO	-6749.498	-1044.184	1827.824	1033.696	1890.826	-1350.875		1515.814
Chavez2006	B3LYP	6-31G*	MPA	FROG2	QM	-7800.861	-788.872	4422.792	-480.346	1330.002	-654.759		2455.017
Chavez2006	B3LYP	6-31G*	MPA	NCI	MM	-2162.218	32.604	-742.651	1253.151	1163.530	-807.450		353.000
Chavez2006	B3LYP	6-31G*	MPA	NCI	NO	1107.766	-1418.732	-2375.337	1267.299	1630.086	-625.143		-149.965
Chavez2006	B3LYP	6-31G*	MPA	NCI	QM	-6095.545	-3128.142	1730.309	875.572	3470.686	-1357.678		2396.411
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	MM	-3355.636	217.597	524.337	744.140	432.143	-341.551		600.649
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	NO	-1396.547	-184.740	-1093.158	1229.977	1019.156	-623.479		163.533
Chavez2006	B3LYP	6-31G*	MPA	OPENBABEL	QM	-8301.369	-1910.644	5467.844	-1061.451	1312.640	-250.405		2904.001
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	MM	-1524.014	-1139.739	-1108.238	1287.154	1653.311	-741.366		448.615
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	NO	-7990.144	616.077	3508.006	245.305	581.973	-602.468		1885.218
Chavez2006	B3LYP	6-31G*	MPA	PUBCHEM	QM	-5299.798	-3608.759	1282.328	921.402	3903.578	-1503.356		2359.534
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	MM	1971.318	-631.891	-3115.451	1497.892	1276.744	-646.270		-616.820
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	QM	-6537.361	-3034.567	1568.915	1149.479	3832.311	-1647.814		2488.188
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	MM	-4991.254	249.403	1661.541	498.484	331.360	-281.472		1047.360
Chavez2006	B3LYP	6-31G*	MPA	RDKIT	NO	-1315.693	1765.127	-1371.613	1407.727	-68.094	-530.391		-423.171
QM	B3LYP	6-31G*	AIM	BALLOON	MM	-177.368	362.097	-161.258	-103.749	249.308	46.144		34.947
QM	B3LYP	6-31G*	AIM	BALLOON	NO	-127.676	-111.482	-97.219	-86.901	200.014	89.972		245.858
QM	B3LYP	6-31G*	AIM	BALLOON	QM	126.963	150.348	-62.803	-141.048	108.163	15.720		-8.821
QM	B3LYP	6-31G*	AIM	FROG2	MM	-180.541	14.559	-55.682	-65.885	155.971	63.997		196.866
QM	B3LYP	6-31G*	AIM	FROG2	NO	-76.795	-124.814	-39.281	23.932	97.583	-1.989		158.679
QM	B3LYP	6-31G*	AIM	FROG2	QM	-93.460	-197.384	-45.650	-62.697	54.553	85.335		154.260
QM	B3LYP	6-31G*	AIM	NCI	MM	-234.195	37.979	-38.008	-99.436	84.423	99.845		145.374
QM	B3LYP	6-31G*	AIM	NCI	NO	-334.300	-175.350	84.154	-16.481	-19.817	58.742		300.848
QM	B3LYP	6-31G*	AIM	NCI	QM	75.948	318.883	-82.796	-134.528	95.413	-1.146		-108.945
QM	B3LYP	6-31G*	AIM	OPENBABEL	MM	-189.293	55.274	-75.672	-87.430	155.098	76.901		161.956
QM	B3LYP	6-31G*	AIM	OPENBABEL	NO	-172.279	126.304	-89.862	-68.751	195.869	47.622		152.406
QM	B3LYP	6-31G*	AIM	OPENBABEL	QM	-30.978	-331.936	14.425	-44.212	-4.741	68.224		191.822
QM	B3LYP	6-31G*	AIM	PUBCHEM	MM	-235.584	14.266	-76.203	-138.931	81.318	134.962		118.203
QM	B3LYP	6-31G*	AIM	PUBCHEM	NO	-93.470	44.045	-191.497	75.332	-15.005	10.921		-254.207
QM	B3LYP	6-31G*	AIM	PUBCHEM	QM	-182.328	120.291	-74.960	-62.830	14.915	58.906		-42.202
QM	B3LYP	6-31G*	AIM	RDKIT	MM	-152.627	28.087	-75.692	117.257	54.962	-53.276		-2.519
QM	B3LYP	6-31G*	AIM	RDKIT	QM	79.646	152.079	-77.554	-131.015	125.842	29.380		4.825
QM	B3LYP	6-31G*	AIM	RDKIT	MM	-197.068	66.956	-75.685	-94.932	151.631	81.852		156.804
QM	B3LYP	6-31G*	AIM	RDKIT	NO	-25.592	-321.111	-0.227	20.505	86.955	-12.647		268.108
QM	B3LYP	6-31G*	MPA	BALLOON	MM	-284.150	72.511	-76.683	-4.029	109.029	5.227		76.001
QM	B3LYP	6-31G*	MPA	BALLOON	NO	-338.372	134.298	-48.999	18.604	103.721	-25.844		81.388
QM	B3LYP	6-31G*	MPA	BALLOON	QM	-143.369	-80.018	-50.098	-34.529	48.491	41.794		83.160
QM	B3LYP	6-31G*	MPA	FROG2	MM	-274.832	65.854	-91.737	-38.445	65.432	38.799		43.283
QM	B3LYP	6-31G*	MPA	FROG2	NO	-237.004	62.297	52.337	7.597	-56.961	-7.796		61.393
QM	B3LYP	6-31G*	MPA	FROG2	QM	-182.804	-71.759	-90.362	-32.348	75.732	36.855		77.932
QM	B3LYP	6-31G*	MPA	NCI	MM	-273.501	90.141	-80.660	-40.660	66.508	41.160		42.424
QM	B3LYP	6-31G*	MPA	NCI	NO	-281.085	46.072	-50.228	-4.130	83.572	4.881		91.372
QM	B3LYP	6-31G*	MPA	NCI	QM	-177.742	-58.487	-75.508	-22.424	58.587	29.626		68.715
QM	B3LYP	6-31G*	MPA	OPENBABEL	MM	-271.730	62.001	-93.159	-30.580	65.757	31.075		41.317

Table S4: Parameters of all the QSPR models.

QM	B3LYP	6-31G*	MPA	OPENBABEL	NO	-279.401	73.926	-92.396	-17.400	69.727	20.223		38.974
QM	B3LYP	6-31G*	MPA	OPENBABEL	QM	-171.355	-82.220	-77.807	-32.344	69.924	37.304		84.764
QM	B3LYP	6-31G*	MPA	PUBCHEM	MM	-273.601	49.535	-83.437	-40.419	69.186	41.859		58.949
QM	B3LYP	6-31G*	MPA	PUBCHEM	NO	-196.213	-49.320	-51.393	-51.242	60.885	51.564		101.018
QM	B3LYP	6-31G*	MPA	PUBCHEM	QM	-201.442	-83.056	-62.604	-14.759	100.413	21.324		122.173
QM	B3LYP	6-31G*	MPA	RDKIT	MM	-223.511	122.164	-69.363	64.344	50.042	-52.559		-19.516
QM	B3LYP	6-31G*	MPA	RDKIT	QM	-186.982	-42.610	-79.312	-25.052	70.327	30.983		70.363
QM	B3LYP	6-31G*	MPA	RDKIT	MM	-272.979	61.255	-93.446	-32.577	68.221	32.881		43.869
QM	B3LYP	6-31G*	MPA	RDKIT	NO	-191.947	-10.647	-10.428	11.326	-0.856	-19.297		59.291
QM	B3LYP	6-31G*	NPA	BALLOON	MM	-214.518	-364.932	50.949	-94.537	106.520	80.757		354.305
QM	B3LYP	6-31G*	NPA	BALLOON	NO	-69.249	-792.525	62.328	-166.711	179.728	154.492		533.750
QM	B3LYP	6-31G*	NPA	BALLOON	QM	-1.380	-446.614	-33.527	-43.873	-31.560	37.865		151.143
QM	B3LYP	6-31G*	NPA	FROG2	MM	-307.608	246.860	-111.555	31.691	226.597	-17.942		50.720
QM	B3LYP	6-31G*	NPA	FROG2	NO	-343.611	122.509	10.734	38.431	199.439	-37.112		211.942
QM	B3LYP	6-31G*	NPA	FROG2	QM	-49.108	-522.766	19.929	-52.179	27.820	41.478		280.521
QM	B3LYP	6-31G*	NPA	NCI	MM	-321.572	427.882	-158.216	57.501	85.953	-44.717		-139.074
QM	B3LYP	6-31G*	NPA	NCI	NO	-37.922	-257.644	-108.107	1.518	-78.691	-0.732		-6.309
QM	B3LYP	6-31G*	NPA	NCI	QM	5.960	-233.182	-92.191	9.545	-132.032	-9.946		-52.286
QM	B3LYP	6-31G*	NPA	OPENBABEL	MM	-290.853	191.006	-106.628	22.263	229.805	-8.852		74.421
QM	B3LYP	6-31G*	NPA	OPENBABEL	NO	-308.806	302.484	-126.448	53.203	252.725	-37.767		27.357
QM	B3LYP	6-31G*	NPA	OPENBABEL	QM	-84.789	-445.101	24.759	-54.276	12.717	44.242		258.127
QM	B3LYP	6-31G*	NPA	PUBCHEM	MM	-369.246	423.453	-117.639	27.402	-227.892	-31.330		-254.027
QM	B3LYP	6-31G*	NPA	PUBCHEM	NO	-143.123	-218.467	-72.995	-40.438	-144.648	39.666		16.169
QM	B3LYP	6-31G*	NPA	PUBCHEM	QM	-147.116	213.777	-89.449	94.466	-177.036	-95.663		-211.146
QM	B3LYP	6-31G*	NPA	RDKIT	MM	-243.585	32.452	-65.725	36.440	263.691	-24.438		176.385
QM	B3LYP	6-31G*	NPA	RDKIT	QM	27.985	-371.895	-67.194	-13.770	-39.586	13.694		71.640
QM	B3LYP	6-31G*	NPA	RDKIT	MM	-289.812	181.372	-107.276	15.759	226.961	-2.284		76.724
QM	B3LYP	6-31G*	NPA	RDKIT	NO	-85.641	-556.302	-8.487	20.597	215.790	-21.723		381.757
QM	HF	STO-3G	MPA	BALLOON	MM	-88.478	-204.604	31.469	-74.378	153.730	46.866		203.745
QM	HF	STO-3G	MPA	BALLOON	NO	-221.156	37.382	30.214	103.097	134.753	-151.227		119.184
QM	HF	STO-3G	MPA	BALLOON	QM	-180.902	71.966	-267.500	-67.379	54.185	45.633		-65.105
QM	HF	STO-3G	MPA	FROG2	MM	-216.445	45.565	-158.971	-113.062	488.028	122.949		240.011
QM	HF	STO-3G	MPA	FROG2	NO	-320.013	216.102	-130.545	102.130	326.402	-58.122		104.945
QM	HF	STO-3G	MPA	FROG2	QM	-141.832	-34.716	-91.010	37.718	53.855	-43.666		36.085
QM	HF	STO-3G	MPA	NCI	MM	-200.804	-11.254	-145.735	-158.817	124.905	140.761		62.522
QM	HF	STO-3G	MPA	NCI	NO	-332.358	146.832	-121.009	10.073	-105.340	-32.490		-97.553
QM	HF	STO-3G	MPA	NCI	QM	-193.051	34.612	-180.169	-27.899	-49.407	18.846		-74.096
QM	HF	STO-3G	MPA	OPENBABEL	MM	-225.919	57.918	-161.424	-115.644	463.764	125.403		222.268
QM	HF	STO-3G	MPA	OPENBABEL	NO	-244.728	91.922	-154.939	-95.315	419.033	102.130		189.409
QM	HF	STO-3G	MPA	OPENBABEL	QM	-95.708	-98.137	-57.890	19.735	88.913	-22.295		86.005
QM	HF	STO-3G	MPA	PUBCHEM	MM	-289.913	77.673	-171.420	-95.769	-33.684	92.075		-58.995
QM	HF	STO-3G	MPA	PUBCHEM	NO	-488.226	631.838	-378.554	143.526	-125.621	-115.884		-387.301
QM	HF	STO-3G	MPA	PUBCHEM	QM	-381.322	234.249	-180.957	33.986	-337.670	-58.886		-281.369
QM	HF	STO-3G	MPA	RDKIT	MM	-466.043	346.449	37.015	34.517	394.438	-25.426		191.254
QM	HF	STO-3G	MPA	RDKIT	QM	-114.179	-34.296	-165.441	-47.399	84.332	36.734		20.821

Table S4: Parameters of all the QSPR models.

QM	HF	STO-3G	MPA	RDKIT	MM	-225.731	49.211	-157.754	-126.458	447.332	136.406		218.084
QM	HF	STO-3G	MPA	RDKIT	NO	-72.850	-227.721	4.187	35.351	7.541	-92.646		114.542
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	MM	327.749	186.567	-623.779	-98.407	440.852	63.604		-17.237
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	NO	-157.825	150.383	-81.387	-5.970	374.803	81.535		249.164
Svob2007-chal2	HF	STO-3G	MPA	BALLOON	QM	14.375	-130.925	43.747	-232.293	281.110	321.261		266.463
Svob2007-chal2	HF	STO-3G	MPA	FROG2	MM	417.441	-91.239	-457.434	-376.374	111.960	256.625		-154.711
Svob2007-chal2	HF	STO-3G	MPA	FROG2	NO	-407.498	3.952	306.925	46.293	11.952	75.713		211.088
Svob2007-chal2	HF	STO-3G	MPA	FROG2	QM	-19.347	32.992	-77.404	-112.467	388.681	193.309		267.907
Svob2007-chal2	HF	STO-3G	MPA	NCI	MM	252.376	106.836	-474.769	-155.488	398.949	147.861		42.073
Svob2007-chal2	HF	STO-3G	MPA	NCI	NO	-176.873	140.147	-31.026	9.026	518.423	133.339		391.532
Svob2007-chal2	HF	STO-3G	MPA	NCI	QM	-22.340	104.030	-136.837	-57.816	511.614	172.814		321.088
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	MM	296.063	-29.613	-390.438	-283.669	168.174	207.602		-77.070
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	NO	311.686	-2.431	-429.867	-258.709	191.150	181.875		-84.249
Svob2007-chal2	HF	STO-3G	MPA	OPENBABEL	QM	-1.313	-75.480	4.347	-195.423	274.716	265.292		234.845
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	MM	456.658	333.664	-896.776	2.773	675.562	-10.783		-6.166
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	NO	-45.802	-147.230	111.917	-291.211	174.357	343.945		226.460
Svob2007-chal2	HF	STO-3G	MPA	PUBCHEM	QM	-124.295	126.197	-77.358	28.332	265.777	42.366		164.247
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	MM	-57.781	127.195	-174.632	-58.215	529.934	163.372		320.511
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	QM	-31.522	133.783	-156.134	-17.497	550.086	141.487		338.584
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	MM	168.138	-51.530	-240.586	-274.276	117.461	216.305		-27.038
Svob2007-chal2	HF	STO-3G	MPA	RDKIT	NO	-171.003	142.796	-70.758	32.689	256.971	25.531		163.688
QM	PM6	none	MPA	RDKIT	MM	21.046	-85.215	-94.955	0.361	97.766	14.006		-65.973
QM	PM6	none	MPA	NCI	NO	164.469	-266.893	-87.256	-14.482	45.017	20.552		-39.463
QM	PM6	none	MPA	OPENBABEL	NO	47.051	-456.195	-92.956	-27.674	117.387	45.325		14.950
QM	PM6	none	MPA	FROG2	MM	23.092	-391.282	-94.666	-25.059	118.655	42.385		6.296
QM	PM6	none	MPA	RDKIT	MM	28.463	-406.654	-94.410	-28.330	119.290	45.894		9.431
QM	PM6	none	MPA	OPENBABEL	MM	21.489	-391.180	-93.878	-25.567	119.320	43.100		7.158
QM	PM6	none	MPA	NCI	MM	-20.773	-201.276	-104.026	-27.859	93.751	35.821		-15.794
QM	PM6	none	MPA	PUBCHEM	MM	42.134	-389.544	-112.072	-21.689	134.859	42.220		-12.294
QM	PM6	none	MPA	PUBCHEM	NO	170.131	-127.155	-108.522	1.767	40.565	5.949		-91.091
QM	PM6	none	MPA	NCI	QM	-61.343	67.188	-66.124	-8.223	-4.227	3.305		-21.892
QM	PM6	none	MPA	BALLOON	MM	189.870	-842.507	-44.698	-37.384	74.703	53.888		89.681
QM	PM6	none	MPA	BALLOON	QM	37.547	-241.576	-62.564	-29.843	7.096	26.663		20.396
QM	PM6	none	MPA	RDKIT	QM	-23.466	-131.746	-45.563	-15.707	11.003	15.229		15.521
QM	PM6	none	MPA	FROG2	NO	-24.573	-141.138	-38.603	6.367	13.782	-1.738		7.617
QM	PM6	none	MPA	BALLOON	NO	-49.918	-578.780	2.414	-6.040	102.636	30.921		101.318
QM	PM6	none	MPA	FROG2	QM	-40.710	-233.395	-31.089	-7.310	30.853	15.161		37.168
QM	PM6	none	MPA	OPENBABEL	QM	-52.165	-215.298	-23.176	-5.454	20.761	12.347		42.179
QM	PM6	none	MPA	PUBCHEM	QM	-143.750	159.604	-53.154	10.911	36.228	-5.581		-30.791
QM	PM6	none	MPA	RDKIT	NO	-13.198	-493.481	2.540	-6.639	17.185	13.921		99.396
Gasteiger				NCI	NO	-3.822E5	-8.503E5	-1.416E6	-1.080E4	1.411E6	7.806E3		2.509E5

5 Cross-validation Results

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
Phenols								
QM B3LYP 6-31G*/AIM/NCI NO-OPT R^2=0.973								
1	0.958	0.423	0.315	48	0.992	0.222	0.174	12
2	0.973	0.330	0.234	48	0.977	0.704	0.543	12
3	0.971	0.337	0.253	48	0.921	0.636	0.413	12
4	0.969	0.357	0.283	48	0.971	0.555	0.370	12
5	0.963	0.414	0.304	48	0.971	0.266	0.228	12
QM B3LYP 6-31G*/MPA/NCI NO-OPT R^2=0.971								
1	0.971	0.341	0.282	48	0.974	0.465	0.376	12
2	0.974	0.349	0.271	48	0.959	0.376	0.331	12
3	0.974	0.351	0.273	48	0.952	0.383	0.314	12
4	0.971	0.355	0.285	48	0.976	0.366	0.276	12
5	0.969	0.352	0.289	48	0.980	0.379	0.311	12
QM HF STO-3G/MPA/FROG2 QM-OPT R^2=0.969								
1	0.979	0.290	0.204	48	0.938	0.649	0.489	12
2	0.973	0.356	0.257	48	0.928	0.455	0.361	12
3	0.972	0.361	0.273	48	0.946	0.423	0.290	12
4	0.964	0.394	0.301	48	0.992	0.238	0.195	12
5	0.965	0.371	0.285	48	0.982	0.363	0.274	12
QM HF STO-3G/MPA/FROG2 MM-OPT R^2=0.967								
1	0.970	0.348	0.273	48	0.966	0.505	0.400	12
2	0.976	0.340	0.259	48	0.883	0.570	0.468	12
3	0.971	0.370	0.291	48	0.942	0.436	0.325	12
4	0.963	0.399	0.315	48	0.986	0.289	0.238	12
5	0.961	0.391	0.309	48	0.983	0.333	0.253	12
QM B3LYP 6-31G*/MPA/OPENBABEL MM-OPT R^2=0.965								
1	0.961	0.397	0.283	48	0.976	0.433	0.312	12
2	0.973	0.359	0.261	48	0.956	0.521	0.299	12
3	0.971	0.369	0.261	48	0.950	0.489	0.298	12
4	0.963	0.399	0.264	48	0.982	0.368	0.309	12
5	0.961	0.392	0.274	48	0.974	0.395	0.292	12
QM B3LYP 6-31G*/MPA/FROG2 QM-OPT R^2=0.963								
1	0.969	0.350	0.281	48	0.930	0.644	0.509	12
2	0.965	0.405	0.319	48	0.943	0.419	0.257	12
3	0.967	0.395	0.313	48	0.938	0.435	0.341	12
4	0.961	0.411	0.316	48	0.977	0.369	0.336	12
5	0.961	0.393	0.312	48	0.966	0.450	0.355	12
QM HF STO-3G/MPA/PUBCHEM MM-OPT R^2=0.963								
1	0.957	0.417	0.315	48	0.977	0.380	0.335	12
2	0.970	0.379	0.287	48	0.925	0.490	0.411	12
3	0.969	0.379	0.287	48	0.930	0.538	0.338	12
4	0.958	0.425	0.328	48	0.984	0.301	0.214	12
5	0.964	0.379	0.306	48	0.963	0.524	0.394	12
QM B3LYP 6-31G*/MPA/NCI MM-OPT R^2=0.963								
1	0.957	0.417	0.297	48	0.979	0.411	0.322	12
2	0.968	0.391	0.276	48	0.927	0.468	0.350	12
3	0.967	0.392	0.290	48	0.930	0.476	0.310	12
4	0.959	0.417	0.283	48	0.980	0.362	0.315	12
5	0.966	0.369	0.265	48	0.956	0.527	0.383	12
QM B3LYP 6-31G*/AIM/OPENBABEL NO-OPT R^2=0.963								
1	0.957	0.426	0.343	48	0.974	0.348	0.284	12
2	0.972	0.342	0.269	48	0.952	0.702	0.587	12
3	0.961	0.395	0.321	48	0.964	0.460	0.361	12
4	0.960	0.406	0.313	48	0.961	0.443	0.363	12
5	0.961	0.427	0.342	48	0.945	0.349	0.283	12
QM B3LYP 6-31G*/AIM/OPENBABEL MM-OPT R^2=0.961								

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
1	0.956	0.434	0.350	48	0.976	0.341	0.274	12
2	0.971	0.344	0.270	48	0.947	0.685	0.570	12
3	0.959	0.403	0.323	48	0.963	0.464	0.376	12
4	0.957	0.419	0.324	48	0.965	0.413	0.350	12
5	0.960	0.433	0.341	48	0.940	0.366	0.296	12
QM B3LYP 6-31G*/MPA/RDKIT MM-OPT R^2=0.961								
1	0.966	0.370	0.276	48	0.957	0.557	0.413	12
2	0.962	0.424	0.319	48	0.963	0.394	0.293	12
3	0.964	0.409	0.301	48	0.931	0.452	0.363	12
4	0.957	0.428	0.315	48	0.975	0.360	0.306	12
5	0.963	0.384	0.283	48	0.965	0.573	0.382	12
QM B3LYP 6-31G*/MPA/OPENBABEL NO-OPT R^2=0.961								
1	0.957	0.417	0.316	48	0.976	0.469	0.341	12
2	0.967	0.396	0.293	48	0.949	0.517	0.381	12
3	0.966	0.396	0.307	48	0.933	0.528	0.356	12
4	0.959	0.420	0.303	48	0.981	0.405	0.365	12
5	0.960	0.399	0.300	48	0.964	0.464	0.336	12
QM HF STO-3G/MPA/OPENBABEL MM-OPT R^2=0.961								
1	0.963	0.386	0.283	48	0.962	0.536	0.429	12
2	0.970	0.379	0.283	48	0.931	0.567	0.409	12
3	0.973	0.352	0.266	48	0.507	1.691	0.735	12
4	0.956	0.434	0.324	48	0.989	0.344	0.318	12
5	0.952	0.437	0.331	48	0.984	0.310	0.240	12
QM B3LYP 6-31G*/NPA/NCI NO-OPT R^2=0.960								
1	0.966	0.369	0.303	48	0.943	0.685	0.580	12
2	0.962	0.422	0.336	48	0.940	0.419	0.338	12
3	0.963	0.418	0.331	48	0.941	0.425	0.366	12
4	0.955	0.438	0.351	48	0.985	0.317	0.247	12
5	0.958	0.407	0.328	48	0.965	0.453	0.351	12
QM HF STO-3G/MPA/PUBCHEM NO-OPT R^2=0.960								
1	0.960	0.401	0.309	48	0.958	0.535	0.395	12
2	0.963	0.418	0.323	48	0.954	0.441	0.361	12
3	0.963	0.415	0.303	48	0.943	0.450	0.348	12
4	0.953	0.450	0.348	48	0.988	0.252	0.187	12
5	0.966	0.369	0.298	48	0.943	0.605	0.448	12
QM B3LYP 6-31G*/NPA/NCI MM-OPT R^2=0.959								
1	0.960	0.402	0.312	48	0.866	1.110	0.694	12
2	0.961	0.429	0.337	48	0.954	0.415	0.363	12
3	0.962	0.423	0.344	48	0.942	0.417	0.306	12
4	0.955	0.439	0.347	48	0.981	0.343	0.277	12
5	0.969	0.351	0.266	48	0.944	0.661	0.536	12
QM B3LYP 6-31G*/NPA/OPENBABEL MM-OPT R^2=0.959								
1	0.961	0.396	0.330	48	0.953	0.582	0.407	12
2	0.960	0.436	0.337	48	0.962	0.375	0.305	12
3	0.967	0.394	0.303	48	0.935	0.552	0.417	12
4	0.958	0.423	0.327	48	0.967	0.420	0.341	12
5	0.956	0.419	0.327	48	0.967	0.437	0.328	12
QM HF STO-3G/MPA/NCI MM-OPT R^2=0.958								
1	0.959	0.405	0.303	48	0.950	0.614	0.484	12
2	0.965	0.404	0.299	48	0.905	0.556	0.399	12
3	0.962	0.423	0.322	48	0.935	0.461	0.348	12
4	0.954	0.445	0.334	48	0.980	0.369	0.324	12
5	0.953	0.430	0.322	48	0.974	0.429	0.315	12
QM B3LYP 6-31G*/NPA/OPENBABEL NO-OPT R^2=0.957								
1	0.956	0.423	0.347	48	0.957	0.528	0.374	12
2	0.956	0.454	0.347	48	0.968	0.337	0.291	12
3	0.964	0.408	0.320	48	0.939	0.544	0.416	12
4	0.959	0.418	0.324	48	0.952	0.491	0.394	12
5	0.957	0.415	0.318	48	0.957	0.546	0.413	12

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
QM B3LYP 6-31G*/MPA/OPENBABEL QM-OPT R^2=0.957								
1	0.953	0.436	0.313	48	0.971	0.431	0.326	12
2	0.961	0.428	0.323	48	0.951	0.479	0.367	12
3	0.963	0.417	0.315	48	0.917	0.500	0.367	12
4	0.957	0.431	0.314	48	0.963	0.446	0.380	12
5	0.954	0.428	0.312	48	0.967	0.491	0.297	12
QM B3LYP 6-31G*/NPA/OPENBABEL QM-OPT R^2=0.956								
1	0.954	0.430	0.328	48	0.967	0.484	0.387	12
2	0.957	0.451	0.361	48	0.957	0.392	0.325	12
3	0.960	0.432	0.344	48	0.930	0.461	0.371	12
4	0.954	0.446	0.361	48	0.970	0.401	0.331	12
5	0.960	0.397	0.328	48	0.955	0.605	0.442	12
QM HF STO-3G/MPA/OPENBABEL QM-OPT R^2=0.955								
1	0.955	0.426	0.300	48	0.957	0.517	0.398	12
2	0.964	0.415	0.292	48	0.929	0.573	0.443	12
3	0.964	0.410	0.306	48	0.595	1.365	0.661	12
4	0.950	0.463	0.331	48	0.982	0.363	0.312	12
5	0.950	0.446	0.336	48	0.971	0.460	0.246	12
QM HF STO-3G/MPA/OPENBABEL NO-OPT R^2=0.955								
1	0.954	0.429	0.320	48	0.960	0.525	0.429	12
2	0.965	0.406	0.307	48	0.907	0.615	0.450	12
3	0.972	0.359	0.273	48	0.467	1.816	0.773	12
4	0.949	0.468	0.360	48	0.989	0.355	0.319	12
5	0.947	0.457	0.352	48	0.975	0.416	0.307	12
QM B3LYP 6-31G*/NPA/FROG2 QM-OPT R^2=0.953								
1	0.957	0.414	0.274	48	0.935	0.658	0.483	12
2	0.967	0.394	0.302	48	0.869	0.682	0.398	12
3	0.952	0.474	0.340	48	0.956	0.361	0.289	12
4	0.946	0.481	0.347	48	0.983	0.326	0.298	12
5	0.950	0.445	0.319	48	0.963	0.500	0.404	12
QM B3LYP 6-31G*/MPA/PUBCHEM NO-OPT R^2=0.950								
1	0.955	0.427	0.297	48	0.932	0.640	0.470	12
2	0.947	0.498	0.337	48	0.956	0.374	0.254	12
3	0.950	0.484	0.328	48	0.947	0.404	0.276	12
4	0.944	0.490	0.350	48	0.979	0.421	0.335	12
5	0.970	0.344	0.265	48	0.877	0.901	0.589	12
QM B3LYP 6-31G*/AIM/RDKIT MM-OPT R^2=0.950								
1	0.948	0.469	0.374	48	0.930	0.538	0.362	12
2	0.961	0.398	0.301	48	0.945	0.794	0.671	12
3	0.948	0.452	0.358	48	0.963	0.572	0.444	12
4	0.947	0.466	0.367	48	0.937	0.555	0.446	12
5	0.947	0.494	0.388	48	0.949	0.348	0.300	12
QM HF STO-3G/MPA/RDKIT MM-OPT R^2=0.947								
1	0.952	0.441	0.313	48	0.940	0.650	0.451	12
2	0.955	0.459	0.343	48	0.879	0.610	0.441	12
3	0.962	0.419	0.302	48	0.525	1.542	0.834	12
4	0.939	0.511	0.382	48	0.981	0.350	0.292	12
5	0.939	0.492	0.358	48	0.973	0.460	0.360	12
QM HF STO-3G/MPA/NCI NO-OPT R^2=0.947								
1	0.955	0.425	0.316	48	0.930	0.667	0.504	12
2	0.955	0.463	0.341	48	0.898	0.562	0.410	12
3	0.949	0.486	0.352	48	0.920	0.491	0.400	12
4	0.939	0.512	0.378	48	0.988	0.339	0.311	12
5	0.939	0.492	0.374	48	0.970	0.440	0.346	12
QM B3LYP 6-31G*/MPA/RDKIT QM-OPT R^2=0.944								
1	0.941	0.487	0.357	48	0.952	0.517	0.406	12
2	0.948	0.496	0.363	48	0.909	0.555	0.442	12
3	0.953	0.467	0.344	48	0.858	0.678	0.549	12
4	0.935	0.526	0.401	48	0.982	0.345	0.269	12

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
5	0.955	0.423	0.312	48	0.926	0.812	0.588	12
QM HF STO-3G/MPA/PUBCHEM QM-OPT R^2=0.943								
1	0.935	0.511	0.366	48	0.958	0.508	0.398	12
2	0.947	0.498	0.364	48	0.901	0.531	0.468	12
3	0.947	0.495	0.350	48	0.912	0.541	0.361	12
4	0.939	0.512	0.377	48	0.965	0.435	0.304	12
5	0.958	0.410	0.318	48	0.903	0.772	0.492	12
QM B3LYP 6-31G*/NPA/RDKIT MM-OPT R^2=0.941								
1	0.940	0.490	0.388	48	0.946	0.589	0.518	12
2	0.948	0.496	0.408	48	0.888	0.595	0.473	12
3	0.952	0.472	0.366	48	0.835	0.728	0.626	12
4	0.933	0.535	0.444	48	0.967	0.386	0.263	12
5	0.944	0.469	0.385	48	0.933	0.695	0.565	12
QM B3LYP 6-31G*/AIM/FROG2 QM-OPT R^2=0.939								
1	0.931	0.543	0.376	48	0.937	0.556	0.427	12
2	0.954	0.433	0.320	48	0.953	0.997	0.730	12
3	0.925	0.544	0.401	48	0.947	0.565	0.376	12
4	0.928	0.544	0.386	48	0.950	0.560	0.399	12
5	0.929	0.574	0.406	48	0.975	0.392	0.300	12
QM B3LYP 6-31G*/MPA/BALLOON NO-OPT R^2=0.939								
1	0.941	0.488	0.363	48	0.947	0.634	0.443	12
2	0.937	0.546	0.391	48	0.953	0.389	0.303	12
3	0.950	0.485	0.346	48	0.888	0.704	0.528	12
4	0.943	0.496	0.358	48	0.916	0.618	0.420	12
5	0.933	0.515	0.365	48	0.948	0.554	0.411	12
QM B3LYP 6-31G*/AIM/FROG2 MM-OPT R^2=0.938								
1	0.934	0.528	0.376	48	0.961	0.554	0.467	12
2	0.968	0.363	0.278	48	0.887	1.106	0.744	12
3	0.927	0.536	0.385	48	0.954	0.520	0.438	12
4	0.932	0.530	0.389	48	0.943	0.581	0.389	12
5	0.930	0.571	0.419	48	0.970	0.280	0.236	12
QM B3LYP 6-31G*/MPA/PUBCHEM QM-OPT R^2=0.936								
1	0.929	0.534	0.393	48	0.949	0.554	0.433	12
2	0.935	0.554	0.394	48	0.954	0.424	0.298	12
3	0.939	0.535	0.368	48	0.921	0.501	0.406	12
4	0.929	0.550	0.396	48	0.969	0.439	0.387	12
5	0.960	0.397	0.322	48	0.880	0.900	0.597	12
QM B3LYP 6-31G*/AIM/OPENBABEL QM-OPT R^2=0.936								
1	0.921	0.580	0.422	48	0.973	0.473	0.380	12
2	0.960	0.403	0.311	48	0.924	1.116	0.819	12
3	0.928	0.534	0.409	48	0.908	0.703	0.488	12
4	0.928	0.544	0.385	48	0.946	0.647	0.480	12
5	0.927	0.581	0.418	48	0.955	0.450	0.377	12
QM B3LYP 6-31G*/AIM/NCI MM-OPT R^2=0.936								
1	0.940	0.505	0.387	48	0.923	0.580	0.412	12
2	0.955	0.429	0.300	48	0.918	0.901	0.731	12
3	0.928	0.535	0.401	48	0.967	0.461	0.318	12
4	0.938	0.503	0.356	48	0.929	0.612	0.467	12
5	0.934	0.553	0.418	48	0.966	0.344	0.297	12
QM HF STO-3G/MPA/RDKIT QM-OPT R^2=0.935								
1	0.934	0.514	0.346	48	0.939	0.595	0.447	12
2	0.943	0.520	0.361	48	0.873	0.623	0.479	12
3	0.957	0.446	0.337	48	0.507	1.608	0.972	12
4	0.924	0.572	0.415	48	0.983	0.338	0.287	12
5	0.933	0.516	0.373	48	0.944	0.655	0.461	12
QM B3LYP 6-31G*/NPA/PUBCHEM NO-OPT R^2=0.935								
1	0.931	0.525	0.396	48	0.957	0.638	0.504	12
2	0.940	0.532	0.394	48	0.906	0.542	0.469	12
3	0.941	0.524	0.401	48	0.886	0.621	0.434	12

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
4	0.927	0.558	0.434	48	0.952	0.478	0.304	12
5	0.955	0.424	0.343	48	0.862	0.933	0.632	12
QM B3LYP 6-31G*/MPA/NCI QM-OPT R^2=0.935								
1	0.929	0.532	0.372	48	0.942	0.626	0.470	12
2	0.940	0.534	0.402	48	0.934	0.546	0.427	12
3	0.949	0.487	0.370	48	0.831	0.706	0.533	12
4	0.930	0.547	0.405	48	0.953	0.477	0.373	12
5	0.930	0.527	0.400	48	0.952	0.563	0.370	12
QM B3LYP 6-31G*/NPA/RDKIT MM-OPT R^2=0.934								
1	0.929	0.533	0.396	48	0.946	0.559	0.474	12
2	0.936	0.552	0.428	48	0.919	0.479	0.389	12
3	0.953	0.467	0.366	48	0.802	0.854	0.604	12
4	0.925	0.567	0.438	48	0.970	0.385	0.299	12
5	0.932	0.517	0.392	48	0.940	0.594	0.468	12
QM B3LYP 6-31G*/AIM/RDKIT MM-OPT R^2=0.934								
1	0.923	0.572	0.425	48	0.986	0.539	0.397	12
2	0.946	0.469	0.292	48	0.963	0.996	0.767	12
3	0.924	0.548	0.438	48	0.927	0.618	0.489	12
4	0.935	0.516	0.379	48	0.900	0.764	0.574	12
5	0.930	0.569	0.449	48	0.922	0.532	0.394	12
QM B3LYP 6-31G*/NPA/BALLOON MM-OPT R^2=0.933								
1	0.947	0.463	0.382	48	0.871	0.894	0.574	12
2	0.938	0.541	0.428	48	0.897	0.613	0.529	12
3	0.934	0.556	0.437	48	0.947	0.466	0.348	12
4	0.931	0.543	0.421	48	0.955	0.528	0.423	12
5	0.935	0.507	0.416	48	0.942	0.712	0.565	12
QM B3LYP 6-31G*/NPA/RDKIT QM-OPT R^2=0.933								
1	0.927	0.541	0.403	48	0.945	0.558	0.472	12
2	0.940	0.531	0.411	48	0.816	0.712	0.593	12
3	0.953	0.468	0.375	48	0.794	0.888	0.701	12
4	0.921	0.581	0.466	48	0.974	0.345	0.265	12
5	0.937	0.500	0.386	48	0.935	0.775	0.604	12
QM B3LYP 6-31G*/MPA/FROG2 MM-OPT R^2=0.931								
1	0.924	0.552	0.412	48	0.950	0.640	0.538	12
2	0.942	0.525	0.405	48	0.846	0.647	0.383	12
3	0.940	0.529	0.393	48	0.866	0.627	0.448	12
4	0.932	0.539	0.391	48	0.929	0.588	0.458	12
5	0.925	0.545	0.408	48	0.957	0.595	0.463	12
QM HF STO-3G/MPA/RDKIT MM-OPT R^2=0.931								
1	0.924	0.551	0.393	48	0.950	0.550	0.462	12
2	0.945	0.509	0.361	48	0.792	0.791	0.546	12
3	0.972	0.361	0.286	48	0.310	2.250	1.128	12
4	0.921	0.580	0.434	48	0.974	0.410	0.350	12
5	0.913	0.585	0.444	48	0.978	0.365	0.249	12
QM B3LYP 6-31G*/NPA/PUBCHEM MM-OPT R^2=0.927								
1	0.917	0.579	0.447	48	0.952	0.516	0.387	12
2	0.936	0.550	0.416	48	0.881	0.636	0.549	12
3	0.942	0.519	0.429	48	0.818	0.773	0.531	12
4	0.919	0.589	0.464	48	0.965	0.448	0.347	12
5	0.930	0.526	0.416	48	0.921	0.764	0.550	12
Chavez2006 B3LYP 6-31G*/MPA/NCI MM-OPT R^2=0.926								
1	0.920	0.567	0.436	48	0.981	0.641	0.505	12
2	0.946	0.507	0.412	48	0.780	0.778	0.566	12
3	0.933	0.558	0.410	48	0.886	0.609	0.493	12
4	0.921	0.581	0.460	48	0.937	0.568	0.404	12
5	0.918	0.569	0.437	48	0.950	0.567	0.461	12
Bult2002-npa B3LYP 6-31G*/NPA/NCI MM-OPT R^2=0.926								
1	0.922	0.560	0.447	48	0.974	0.688	0.543	12
2	0.943	0.521	0.419	48	0.789	0.760	0.536	12

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
3	0.935	0.552	0.418	48	0.857	0.666	0.573	12
4	0.925	0.567	0.452	48	0.918	0.642	0.479	12
5	0.917	0.573	0.448	48	0.957	0.558	0.469	12
Bult2002-npa B3LYP 6-31G*/NPA/PUBCHEM QM-OPT R^2=0.925								
1	0.919	0.571	0.435	48	0.957	0.657	0.531	12
2	0.952	0.475	0.387	48	0.743	0.881	0.716	12
3	0.924	0.595	0.457	48	0.929	0.465	0.402	12
4	0.920	0.584	0.464	48	0.930	0.582	0.434	12
5	0.918	0.571	0.446	48	0.949	0.641	0.501	12
Bult2002-npa B3LYP 6-31G*/NPA/NCI QM-OPT R^2=0.925								
1	0.919	0.569	0.453	48	0.976	0.676	0.523	12
2	0.943	0.520	0.425	48	0.783	0.774	0.574	12
3	0.929	0.574	0.429	48	0.897	0.572	0.485	12
4	0.926	0.565	0.444	48	0.911	0.666	0.477	12
5	0.916	0.578	0.445	48	0.957	0.562	0.445	12
Chavez2006 B3LYP 6-31G*/MPA/NCI QM-OPT R^2=0.924								
1	0.918	0.573	0.441	48	0.984	0.673	0.504	12
2	0.945	0.509	0.417	48	0.769	0.799	0.598	12
3	0.928	0.580	0.433	48	0.908	0.550	0.459	12
4	0.921	0.581	0.453	48	0.932	0.594	0.440	12
5	0.915	0.580	0.441	48	0.952	0.557	0.439	12
Svob2007-chal2 HF STO-3G/MPA/RDKIT MM-OPT R^2=0.923								
1	0.911	0.599	0.435	48	0.954	0.518	0.356	12
2	0.937	0.547	0.398	48	0.798	0.741	0.592	12
3	0.925	0.592	0.434	48	0.903	0.578	0.422	12
4	0.919	0.591	0.442	48	0.927	0.571	0.420	12
5	0.932	0.517	0.379	48	0.905	0.862	0.620	12
QM B3LYP 6-31G*/AIM/RDKIT QM-OPT R^2=0.922								
1	0.906	0.632	0.453	48	0.979	0.527	0.373	12
2	0.939	0.501	0.349	48	0.940	1.065	0.805	12
3	0.909	0.599	0.453	48	0.920	0.693	0.491	12
4	0.920	0.575	0.417	48	0.870	0.821	0.568	12
5	0.919	0.613	0.457	48	0.841	0.699	0.440	12
QM B3LYP 6-31G*/NPA/PUBCHEM QM-OPT R^2=0.922								
1	0.906	0.615	0.438	48	0.968	0.485	0.410	12
2	0.931	0.573	0.402	48	0.875	0.642	0.501	12
3	0.925	0.590	0.419	48	0.893	0.574	0.397	12
4	0.910	0.621	0.450	48	0.966	0.414	0.324	12
5	0.951	0.439	0.339	48	0.841	1.007	0.668	12
Chavez2006 B3LYP 6-31G*/MPA/PUBCHEM QM-OPT R^2=0.921								
1	0.911	0.597	0.454	48	0.975	0.629	0.514	12
2	0.949	0.493	0.404	48	0.736	0.889	0.710	12
3	0.920	0.612	0.471	48	0.928	0.460	0.393	12
4	0.916	0.601	0.478	48	0.933	0.560	0.409	12
5	0.917	0.572	0.442	48	0.932	0.697	0.559	12
Svob2007-chal2 HF STO-3G/MPA/FROG2 QM-OPT R^2=0.921								
1	0.941	0.488	0.360	48	0.817	1.034	0.673	12
2	0.947	0.502	0.369	48	0.754	0.892	0.732	12
3	0.924	0.597	0.420	48	0.884	0.587	0.476	12
4	0.910	0.620	0.452	48	0.956	0.456	0.367	12
5	0.903	0.621	0.448	48	0.982	0.458	0.316	12
Chavez2006 B3LYP 6-31G*/MPA/OPENBABEL NO-OPT R^2=0.920								
1	0.904	0.621	0.493	48	0.963	0.485	0.352	12
2	0.947	0.499	0.409	48	0.723	0.886	0.731	12
3	0.923	0.599	0.474	48	0.900	0.562	0.444	12
4	0.919	0.591	0.466	48	0.918	0.632	0.477	12
5	0.910	0.597	0.470	48	0.943	0.578	0.478	12
Chavez2006 B3LYP 6-31G*/MPA/OPENBABEL MM-OPT R^2=0.919								
1	0.902	0.627	0.492	48	0.967	0.475	0.347	12

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
2	0.947	0.499	0.408	48	0.718	0.894	0.734	12
3	0.922	0.604	0.469	48	0.903	0.549	0.442	12
4	0.919	0.591	0.468	48	0.916	0.641	0.485	12
5	0.909	0.600	0.472	48	0.942	0.586	0.477	12
Chavez2006 B3LYP 6-31G*/MPA/RDKIT QM-OPT R^2=0.919								
1	0.909	0.603	0.470	48	0.951	0.610	0.448	12
2	0.939	0.536	0.443	48	0.772	0.815	0.650	12
3	0.916	0.625	0.495	48	0.936	0.458	0.397	12
4	0.919	0.589	0.459	48	0.910	0.646	0.507	12
5	0.919	0.566	0.438	48	0.914	0.802	0.593	12
Chavez2006 B3LYP 6-31G*/MPA/PUBCHEM NO-OPT R^2=0.918								
1	0.903	0.623	0.465	48	0.978	0.527	0.414	12
2	0.945	0.511	0.414	48	0.734	0.888	0.725	12
3	0.922	0.603	0.461	48	0.886	0.585	0.440	12
4	0.914	0.606	0.473	48	0.918	0.618	0.458	12
5	0.914	0.584	0.441	48	0.934	0.671	0.535	12
Svob2007-chal2 HF STO-3G/MPA/FROG2 MM-OPT R^2=0.918								
1	0.930	0.530	0.395	48	0.848	0.941	0.636	12
2	0.948	0.495	0.373	48	0.745	0.947	0.771	12
3	0.919	0.613	0.440	48	0.897	0.554	0.444	12
4	0.907	0.633	0.477	48	0.957	0.441	0.338	12
5	0.903	0.619	0.458	48	0.972	0.518	0.379	12
Bult2002-npa B3LYP 6-31G*/NPA/FROG2 QM-OPT R^2=0.918								
1	0.930	0.529	0.392	48	0.803	1.088	0.663	12
2	0.948	0.496	0.412	48	0.799	0.982	0.873	12
3	0.921	0.608	0.481	48	0.885	0.609	0.510	12
4	0.918	0.594	0.461	48	0.907	0.682	0.519	12
5	0.898	0.635	0.514	48	0.968	0.432	0.343	12
Bult2002-npa B3LYP 6-31G*/NPA/FROG2 MM-OPT R^2=0.917								
1	0.923	0.555	0.415	48	0.868	0.941	0.599	12
2	0.948	0.495	0.413	48	0.801	0.997	0.885	12
3	0.921	0.607	0.483	48	0.882	0.613	0.519	12
4	0.917	0.595	0.461	48	0.906	0.681	0.522	12
5	0.898	0.636	0.518	48	0.968	0.431	0.345	12
Chavez2006 B3LYP 6-31G*/MPA/PUBCHEM MM-OPT R^2=0.917								
1	0.904	0.621	0.463	48	0.973	0.610	0.486	12
2	0.945	0.510	0.418	48	0.725	0.906	0.742	12
3	0.917	0.620	0.480	48	0.912	0.512	0.399	12
4	0.914	0.607	0.473	48	0.918	0.616	0.459	12
5	0.912	0.589	0.452	48	0.935	0.675	0.542	12
Chavez2006 B3LYP 6-31G*/MPA/OPENBABEL QM-OPT R^2=0.917								
1	0.900	0.634	0.495	48	0.970	0.490	0.375	12
2	0.946	0.505	0.415	48	0.709	0.903	0.739	12
3	0.919	0.616	0.479	48	0.907	0.530	0.447	12
4	0.916	0.599	0.480	48	0.916	0.646	0.489	12
5	0.908	0.605	0.476	48	0.937	0.608	0.493	12
QM HF STO-3G/MPA/BALLOON MM-OPT R^2=0.917								
1	0.935	0.511	0.368	48	0.859	0.936	0.643	12
2	0.926	0.593	0.415	48	0.794	0.761	0.503	12
3	0.941	0.522	0.372	48	0.471	1.711	0.845	12
4	0.904	0.641	0.466	48	0.977	0.404	0.363	12
5	0.908	0.603	0.448	48	0.945	0.613	0.415	12
Chavez2006 B3LYP 6-31G*/MPA/RDKIT MM-OPT R^2=0.917								
1	0.908	0.607	0.469	48	0.953	0.626	0.462	12
2	0.938	0.540	0.440	48	0.768	0.825	0.657	12
3	0.914	0.632	0.504	48	0.934	0.468	0.406	12
4	0.920	0.586	0.450	48	0.896	0.691	0.525	12
5	0.913	0.585	0.455	48	0.929	0.724	0.551	12
Bult2002-npa B3LYP 6-31G*/NPA/RDKIT QM-OPT R^2=0.916								

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
1	0.907	0.610	0.469	48	0.944	0.622	0.452	12
2	0.931	0.569	0.465	48	0.798	0.753	0.599	12
3	0.916	0.627	0.491	48	0.924	0.526	0.446	12
4	0.920	0.587	0.447	48	0.889	0.710	0.540	12
5	0.916	0.576	0.450	48	0.924	0.803	0.564	12
Bult2004-aim B3LYP 6-31G*/AIM/NCI MM-OPT R^2=0.916								
1	0.910	0.602	0.476	48	0.980	0.688	0.521	12
2	0.936	0.548	0.445	48	0.770	0.820	0.600	12
3	0.924	0.596	0.452	48	0.864	0.678	0.564	12
4	0.908	0.629	0.502	48	0.938	0.545	0.459	12
5	0.910	0.596	0.463	48	0.936	0.641	0.505	12
QM HF STO-3G/MPA/BALLOON QM-OPT R^2=0.915								
1	0.929	0.534	0.376	48	0.884	0.867	0.619	12
2	0.914	0.636	0.436	48	0.930	0.506	0.391	12
3	0.936	0.545	0.409	48	0.358	2.291	1.009	12
4	0.903	0.645	0.469	48	0.979	0.453	0.409	12
5	0.921	0.561	0.410	48	0.899	0.771	0.511	12
Bult2002-npa B3LYP 6-31G*/NPA/PUBCHEM MM-OPT R^2=0.915								
1	0.907	0.609	0.457	48	0.956	0.723	0.560	12
2	0.941	0.526	0.425	48	0.731	0.894	0.725	12
3	0.917	0.623	0.481	48	0.902	0.551	0.427	12
4	0.912	0.616	0.480	48	0.910	0.648	0.461	12
5	0.908	0.605	0.478	48	0.949	0.679	0.525	12
Bult2002-npa B3LYP 6-31G*/NPA/PUBCHEM NO-OPT R^2=0.913								
1	0.900	0.634	0.475	48	0.977	0.651	0.522	12
2	0.939	0.536	0.436	48	0.728	0.897	0.743	12
3	0.918	0.619	0.480	48	0.874	0.625	0.432	12
4	0.910	0.620	0.485	48	0.899	0.688	0.475	12
5	0.906	0.610	0.473	48	0.947	0.683	0.552	12
QM B3LYP 6-31G*/MPA/FROG2 NO-OPT R^2=0.912								
1	0.909	0.603	0.454	48	0.936	0.694	0.572	12
2	0.944	0.513	0.421	48	0.717	0.975	0.659	12
3	0.915	0.631	0.487	48	0.923	0.584	0.532	12
4	0.901	0.653	0.526	48	0.965	0.466	0.372	12
5	0.900	0.630	0.495	48	0.955	0.591	0.522	12
Bult2002-npa B3LYP 6-31G*/NPA/RDKIT MM-OPT R^2=0.912								
1	0.903	0.624	0.481	48	0.946	0.629	0.465	12
2	0.929	0.579	0.469	48	0.789	0.774	0.619	12
3	0.913	0.637	0.502	48	0.902	0.577	0.481	12
4	0.918	0.591	0.445	48	0.876	0.752	0.565	12
5	0.908	0.602	0.471	48	0.935	0.743	0.526	12
Bult2002-npa B3LYP 6-31G*/NPA/OPENBABEL NO-OPT R^2=0.912								
1	0.893	0.657	0.514	48	0.968	0.476	0.396	12
2	0.940	0.534	0.428	48	0.719	0.920	0.717	12
3	0.914	0.631	0.490	48	0.894	0.561	0.439	12
4	0.914	0.606	0.473	48	0.890	0.728	0.518	12
5	0.905	0.614	0.484	48	0.930	0.661	0.517	12
Bult2004-aim B3LYP 6-31G*/AIM/NCI QM-OPT R^2=0.912								
1	0.905	0.618	0.485	48	0.981	0.695	0.521	12
2	0.935	0.554	0.457	48	0.751	0.853	0.621	12
3	0.918	0.618	0.460	48	0.882	0.660	0.544	12
4	0.904	0.641	0.507	48	0.932	0.571	0.474	12
5	0.905	0.613	0.465	48	0.934	0.648	0.530	12
Bult2002-npa B3LYP 6-31G*/NPA/OPENBABEL MM-OPT R^2=0.911								
1	0.892	0.659	0.513	48	0.967	0.481	0.394	12
2	0.940	0.534	0.429	48	0.711	0.931	0.725	12
3	0.914	0.634	0.489	48	0.891	0.569	0.455	12
4	0.914	0.609	0.472	48	0.890	0.731	0.515	12
5	0.903	0.620	0.489	48	0.932	0.649	0.508	12

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
Bult2002-npa B3LYP 6-31G*/NPA/OPENBABEL QM-OPT R^2=0.911								
1	0.891	0.661	0.517	48	0.969	0.478	0.400	12
2	0.939	0.536	0.430	48	0.710	0.931	0.725	12
3	0.914	0.634	0.483	48	0.851	0.672	0.558	12
4	0.914	0.608	0.473	48	0.888	0.739	0.529	12
5	0.905	0.613	0.480	48	0.923	0.692	0.550	12
QM B3LYP 6-31G*/MPA/PUBCHEM MM-OPT R^2=0.911								
1	0.909	0.604	0.434	48	0.860	0.913	0.527	12
2	0.928	0.584	0.423	48	0.766	0.834	0.498	12
3	0.914	0.633	0.438	48	0.884	0.687	0.469	12
4	0.916	0.599	0.430	48	0.904	0.767	0.517	12
5	0.919	0.565	0.424	48	0.906	0.972	0.740	12
QM B3LYP 6-31G*/MPA/RDKIT MM-OPT R^2=0.909								
1	0.912	0.594	0.399	48	0.916	0.750	0.515	12
2	0.909	0.656	0.439	48	0.927	0.526	0.314	12
3	0.922	0.601	0.420	48	0.813	0.958	0.703	12
4	0.917	0.598	0.405	48	0.883	0.746	0.514	12
5	0.898	0.636	0.446	48	0.945	0.617	0.427	12
Svob2007-chal2 HF STO-3G/MPA/RDKIT QM-OPT R^2=0.909								
1	0.895	0.649	0.454	48	0.944	0.563	0.396	12
2	0.925	0.597	0.432	48	0.760	0.809	0.610	12
3	0.915	0.629	0.438	48	0.870	0.643	0.477	12
4	0.899	0.660	0.482	48	0.940	0.525	0.365	12
5	0.926	0.542	0.391	48	0.875	1.030	0.706	12
Chavez2006 B3LYP 6-31G*/MPA/RDKIT MM-OPT R^2=0.908								
1	0.895	0.651	0.480	48	0.955	0.610	0.487	12
2	0.934	0.557	0.441	48	0.730	0.892	0.724	12
3	0.926	0.586	0.450	48	0.716	1.033	0.503	12
4	0.910	0.621	0.487	48	0.893	0.733	0.531	12
5	0.891	0.655	0.516	48	0.949	0.547	0.415	12
Chavez2006 B3LYP 6-31G*/MPA/BALLOON QM-OPT R^2=0.908								
1	0.911	0.598	0.446	48	0.886	0.972	0.648	12
2	0.933	0.563	0.445	48	0.741	0.870	0.695	12
3	0.911	0.645	0.498	48	0.876	0.609	0.486	12
4	0.912	0.616	0.475	48	0.902	0.722	0.541	12
5	0.890	0.660	0.528	48	0.949	0.546	0.415	12
Bult2004-aim B3LYP 6-31G*/AIM/OPENBABEL NO-OPT R^2=0.908								
1	0.895	0.649	0.506	48	0.945	0.632	0.488	12
2	0.938	0.543	0.435	48	0.698	0.933	0.717	12
3	0.908	0.656	0.520	48	0.892	0.595	0.499	12
4	0.903	0.646	0.504	48	0.922	0.605	0.514	12
5	0.902	0.622	0.483	48	0.920	0.683	0.569	12
QM B3LYP 6-31G*/NPA/BALLOON NO-OPT R^2=0.908								
1	0.911	0.596	0.440	48	0.905	0.809	0.671	12
2	0.912	0.643	0.483	48	0.874	0.606	0.496	12
3	0.911	0.643	0.471	48	0.918	0.612	0.482	12
4	0.919	0.590	0.463	48	0.873	0.831	0.500	12
5	0.899	0.633	0.498	48	0.932	0.719	0.543	12
Svob2007-chal2 HF STO-3G/MPA/FROG2 NO-OPT R^2=0.907								
1	0.915	0.584	0.436	48	0.858	1.004	0.590	12
2	0.941	0.527	0.411	48	0.705	1.023	0.856	12
3	0.911	0.642	0.466	48	0.868	0.648	0.543	12
4	0.901	0.651	0.490	48	0.921	0.599	0.422	12
5	0.885	0.674	0.504	48	0.963	0.467	0.340	12
Bult2004-aim B3LYP 6-31G*/AIM/OPENBABEL MM-OPT R^2=0.907								
1	0.894	0.652	0.503	48	0.943	0.639	0.489	12
2	0.937	0.543	0.434	48	0.691	0.943	0.722	12
3	0.907	0.660	0.516	48	0.891	0.597	0.502	12
4	0.903	0.646	0.503	48	0.918	0.625	0.534	12

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
5	0.900	0.628	0.487	48	0.923	0.673	0.546	12
Bult2002-npa B3LYP 6-31G*/NPA/BALLOON MM-OPT R^2=0.907								
1	0.906	0.615	0.473	48	0.909	0.854	0.552	12
2	0.927	0.589	0.473	48	0.767	0.816	0.655	12
3	0.908	0.655	0.509	48	0.897	0.566	0.471	12
4	0.916	0.602	0.465	48	0.881	0.802	0.608	12
5	0.892	0.653	0.522	48	0.941	0.588	0.475	12
QM B3LYP 6-31G*/NPA/FROG2 MM-OPT R^2=0.907								
1	0.887	0.672	0.481	48	0.963	0.498	0.379	12
2	0.940	0.534	0.424	48	0.722	1.017	0.641	12
3	0.915	0.629	0.453	48	0.839	0.720	0.516	12
4	0.902	0.648	0.470	48	0.929	0.601	0.413	12
5	0.896	0.642	0.459	48	0.941	0.644	0.487	12
Chavez2006 B3LYP 6-31G*/MPA/RDKIT NO-OPT R^2=0.907								
1	0.890	0.665	0.516	48	0.962	0.574	0.448	12
2	0.934	0.556	0.447	48	0.707	0.903	0.760	12
3	0.906	0.664	0.518	48	0.906	0.541	0.441	12
4	0.915	0.604	0.480	48	0.869	0.798	0.605	12
5	0.896	0.642	0.514	48	0.931	0.653	0.481	12
Chavez2006 B3LYP 6-31G*/MPA/FROG2 QM-OPT R^2=0.907								
1	0.923	0.557	0.406	48	0.771	1.193	0.689	12
2	0.942	0.523	0.418	48	0.753	1.063	0.932	12
3	0.908	0.656	0.519	48	0.884	0.588	0.478	12
4	0.906	0.636	0.503	48	0.902	0.711	0.544	12
5	0.886	0.672	0.533	48	0.961	0.478	0.410	12
Chavez2006 B3LYP 6-31G*/MPA/BALLOON MM-OPT R^2=0.906								
1	0.909	0.603	0.442	48	0.883	0.982	0.650	12
2	0.932	0.566	0.451	48	0.731	0.885	0.716	12
3	0.907	0.658	0.508	48	0.900	0.552	0.462	12
4	0.910	0.622	0.482	48	0.900	0.725	0.542	12
5	0.890	0.660	0.526	48	0.945	0.564	0.432	12
Chavez2006 B3LYP 6-31G*/MPA/FROG2 MM-OPT R^2=0.906								
1	0.917	0.577	0.424	48	0.823	1.090	0.672	12
2	0.942	0.523	0.421	48	0.752	1.073	0.940	12
3	0.908	0.654	0.524	48	0.878	0.605	0.489	12
4	0.905	0.637	0.507	48	0.902	0.712	0.543	12
5	0.885	0.675	0.537	48	0.961	0.473	0.412	12
QM B3LYP 6-31G*/NPA/FROG2 NO-OPT R^2=0.906								
1	0.911	0.598	0.440	48	0.885	0.871	0.534	12
2	0.925	0.595	0.435	48	0.763	0.807	0.664	12
3	0.917	0.622	0.463	48	0.812	0.771	0.617	12
4	0.893	0.677	0.513	48	0.955	0.483	0.343	12
5	0.896	0.642	0.484	48	0.945	0.723	0.600	12
Svob2007-chal2 HF STO-3G/MPA/NCI NO-OPT R^2=0.906								
1	0.909	0.603	0.469	48	0.878	0.904	0.510	12
2	0.929	0.581	0.441	48	0.714	0.886	0.731	12
3	0.911	0.644	0.448	48	0.836	0.697	0.566	12
4	0.910	0.621	0.453	48	0.894	0.736	0.596	12
5	0.882	0.683	0.530	48	0.970	0.436	0.334	12
Bult2002-npa B3LYP 6-31G*/NPA/BALLOON QM-OPT R^2=0.906								
1	0.907	0.610	0.467	48	0.900	0.908	0.604	12
2	0.925	0.597	0.481	48	0.771	0.811	0.643	12
3	0.908	0.656	0.504	48	0.829	0.721	0.540	12
4	0.915	0.605	0.463	48	0.875	0.815	0.630	12
5	0.891	0.658	0.526	48	0.938	0.604	0.472	12
Chavez2006 B3LYP 6-31G*/MPA/NCI NO-OPT R^2=0.906								
1	0.890	0.664	0.511	48	0.947	0.638	0.429	12
2	0.932	0.566	0.462	48	0.704	0.908	0.760	12
3	0.913	0.636	0.481	48	0.835	0.700	0.593	12

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
4	0.910	0.620	0.493	48	0.886	0.752	0.544	12
5	0.889	0.664	0.529	48	0.949	0.548	0.437	12
Chavez2006 B3LYP 6-31G*/MPA/BALLOON NO-OPT R^2=0.904								
1	0.898	0.639	0.477	48	0.918	0.867	0.601	12
2	0.931	0.572	0.462	48	0.718	0.895	0.740	12
3	0.908	0.656	0.515	48	0.875	0.637	0.495	12
4	0.909	0.624	0.486	48	0.885	0.759	0.576	12
5	0.889	0.663	0.534	48	0.942	0.601	0.454	12
Bult2002-npa B3LYP 6-31G*/NPA/RDKIT NO-OPT R^2=0.904								
1	0.888	0.669	0.515	48	0.946	0.597	0.490	12
2	0.927	0.587	0.472	48	0.730	0.860	0.715	12
3	0.903	0.673	0.530	48	0.906	0.569	0.418	12
4	0.917	0.595	0.470	48	0.841	0.869	0.657	12
5	0.894	0.649	0.526	48	0.932	0.659	0.518	12
Bult2004-aim B3LYP 6-31G*/AIM/OPENBABEL QM-OPT R^2=0.904								
1	0.890	0.665	0.519	48	0.948	0.644	0.484	12
2	0.935	0.553	0.447	48	0.680	0.956	0.735	12
3	0.902	0.674	0.529	48	0.907	0.555	0.452	12
4	0.899	0.658	0.511	48	0.917	0.628	0.541	12
5	0.898	0.636	0.488	48	0.917	0.699	0.579	12
QM B3LYP 6-31G*/AIM/BALLOON NO-OPT R^2=0.904								
1	0.900	0.633	0.436	48	0.914	0.731	0.464	12
2	0.915	0.632	0.411	48	0.834	0.749	0.554	12
3	0.918	0.618	0.393	48	0.728	1.015	0.712	12
4	0.905	0.639	0.431	48	0.898	0.689	0.435	12
5	0.892	0.653	0.445	48	0.925	0.660	0.347	12
Bult2002-npa B3LYP 6-31G*/NPA/BALLOON NO-OPT R^2=0.903								
1	0.894	0.653	0.504	48	0.936	0.757	0.556	12
2	0.922	0.606	0.483	48	0.757	0.825	0.676	12
3	0.903	0.672	0.522	48	0.888	0.624	0.470	12
4	0.915	0.604	0.469	48	0.855	0.847	0.645	12
5	0.891	0.657	0.530	48	0.932	0.667	0.485	12
QM B3LYP 6-31G*/AIM/NCI QM-OPT R^2=0.902								
1	0.898	0.658	0.505	48	0.928	0.678	0.500	12
2	0.928	0.544	0.411	48	0.910	1.166	0.986	12
3	0.897	0.640	0.507	48	0.897	0.746	0.578	12
4	0.901	0.639	0.471	48	0.912	0.772	0.602	12
5	0.897	0.692	0.552	48	0.911	0.484	0.387	12
Bult2002-npa B3LYP 6-31G*/NPA/RDKIT MM-OPT R^2=0.902								
1	0.889	0.669	0.503	48	0.943	0.644	0.496	12
2	0.923	0.604	0.483	48	0.746	0.850	0.697	12
3	0.925	0.592	0.448	48	0.508	1.236	0.670	12
4	0.908	0.628	0.484	48	0.853	0.847	0.613	12
5	0.887	0.670	0.528	48	0.940	0.608	0.475	12
Bult2004-aim B3LYP 6-31G*/AIM/PUBCHEM NO-OPT R^2=0.902								
1	0.889	0.668	0.510	48	0.951	0.671	0.501	12
2	0.930	0.574	0.472	48	0.711	0.955	0.774	12
3	0.905	0.667	0.521	48	0.865	0.630	0.512	12
4	0.891	0.683	0.536	48	0.932	0.553	0.449	12
5	0.904	0.618	0.479	48	0.897	0.829	0.676	12
QM B3LYP 6-31G*/NPA/BALLOON QM-OPT R^2=0.901								
1	0.919	0.570	0.394	48	0.849	1.069	0.674	12
2	0.900	0.688	0.485	48	0.875	0.588	0.539	12
3	0.903	0.674	0.492	48	0.862	0.718	0.514	12
4	0.888	0.692	0.499	48	0.961	0.504	0.413	12
5	0.920	0.563	0.449	48	0.880	0.975	0.706	12
QM B3LYP 6-31G*/AIM/PUBCHEM NO-OPT R^2=0.900								
1	0.880	0.713	0.503	48	0.968	0.413	0.316	12
2	0.941	0.493	0.312	48	0.892	1.979	1.278	12

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
3	0.910	0.596	0.418	48	0.868	0.923	0.679	12
4	0.883	0.692	0.489	48	0.966	0.572	0.444	12
5	0.907	0.655	0.500	48	0.973	0.754	0.494	12
Svob2007-chal2 HF STO-3G/MPA/OPENBABEL MM-OPT R^2=0.900								
1	0.879	0.698	0.509	48	0.955	0.493	0.341	12
2	0.928	0.583	0.427	48	0.706	0.945	0.685	12
3	0.922	0.605	0.418	48	0.739	0.927	0.751	12
4	0.901	0.652	0.479	48	0.879	0.745	0.502	12
5	0.873	0.708	0.526	48	0.972	0.400	0.290	12
Svob2007-chal2 HF STO-3G/MPA/OPENBABEL NO-OPT R^2=0.900								
1	0.880	0.693	0.507	48	0.949	0.527	0.350	12
2	0.926	0.590	0.433	48	0.723	0.930	0.672	12
3	0.923	0.600	0.416	48	0.733	0.953	0.768	12
4	0.900	0.656	0.485	48	0.885	0.725	0.480	12
5	0.874	0.707	0.523	48	0.971	0.413	0.305	12
Bult2002-npa B3LYP 6-31G*/NPA/NCI NO-OPT R^2=0.899								
1	0.878	0.699	0.534	48	0.962	0.563	0.405	12
2	0.922	0.606	0.474	48	0.702	0.902	0.723	12
3	0.910	0.649	0.490	48	0.813	0.756	0.605	12
4	0.910	0.622	0.475	48	0.851	0.853	0.621	12
5	0.882	0.683	0.543	48	0.944	0.600	0.446	12
Svob2007-chal2 HF STO-3G/MPA/RDKIT MM-OPT R^2=0.899								
1	0.872	0.716	0.497	48	0.968	0.421	0.324	12
2	0.915	0.635	0.445	48	0.774	0.784	0.637	12
3	0.942	0.519	0.386	48	0.615	1.157	0.658	12
4	0.892	0.682	0.480	48	0.932	0.595	0.454	12
5	0.879	0.692	0.510	48	0.950	0.570	0.382	12
Bult2004-aim B3LYP 6-31G*/AIM/PUBCHEM QM-OPT R^2=0.899								
1	0.891	0.662	0.528	48	0.956	0.746	0.592	12
2	0.934	0.560	0.457	48	0.690	0.997	0.823	12
3	0.896	0.696	0.553	48	0.916	0.515	0.431	12
4	0.889	0.690	0.556	48	0.929	0.568	0.473	12
5	0.897	0.638	0.499	48	0.899	0.854	0.697	12
Bult2004-aim B3LYP 6-31G*/AIM/PUBCHEM MM-OPT R^2=0.899								
1	0.889	0.668	0.508	48	0.941	0.768	0.582	12
2	0.930	0.574	0.477	48	0.690	0.986	0.787	12
3	0.898	0.691	0.547	48	0.899	0.545	0.443	12
4	0.891	0.684	0.536	48	0.922	0.592	0.492	12
5	0.900	0.630	0.488	48	0.904	0.829	0.675	12
Svob2007-chal2 HF STO-3G/MPA/PUBCHEM QM-OPT R^2=0.898								
1	0.872	0.717	0.537	48	0.967	0.470	0.339	12
2	0.920	0.615	0.431	48	0.738	0.863	0.681	12
3	0.910	0.648	0.474	48	0.799	0.773	0.612	12
4	0.887	0.695	0.512	48	0.935	0.542	0.395	12
5	0.910	0.598	0.430	48	0.847	0.958	0.578	12
Bult2002-npa B3LYP 6-31G*/NPA/FROG2 NO-OPT R^2=0.898								
1	0.897	0.642	0.508	48	0.877	0.906	0.615	12
2	0.917	0.626	0.525	48	0.788	0.897	0.763	12
3	0.897	0.692	0.564	48	0.896	0.585	0.444	12
4	0.912	0.614	0.490	48	0.843	0.910	0.698	12
5	0.887	0.670	0.558	48	0.917	0.706	0.590	12
Chavez2006 B3LYP 6-31G*/MPA/FROG2 NO-OPT R^2=0.897								
1	0.893	0.657	0.504	48	0.873	0.964	0.705	12
2	0.921	0.610	0.510	48	0.761	0.940	0.763	12
3	0.898	0.690	0.554	48	0.881	0.607	0.503	12
4	0.910	0.620	0.499	48	0.855	0.897	0.713	12
5	0.886	0.673	0.551	48	0.918	0.721	0.578	12
Svob2007-chal2 HF STO-3G/MPA/OPENBABEL QM-OPT R^2=0.896								
1	0.878	0.700	0.522	48	0.942	0.564	0.348	12

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
2	0.925	0.597	0.425	48	0.709	0.956	0.710	12
3	0.915	0.631	0.424	48	0.755	0.870	0.720	12
4	0.896	0.667	0.486	48	0.884	0.727	0.513	12
5	0.869	0.719	0.530	48	0.968	0.431	0.297	12
QM HF STO-3G/MPA/BALLOON NO-OPT R^2=0.896								
1	0.884	0.683	0.509	48	0.929	0.659	0.494	12
2	0.901	0.685	0.505	48	0.828	0.705	0.577	12
3	0.948	0.492	0.356	48	0.332	2.394	0.860	12
4	0.909	0.625	0.508	48	0.832	0.866	0.609	12
5	0.872	0.713	0.548	48	0.966	0.486	0.404	12
Svob2007-chal2 HF STO-3G/MPA/PUBCHEM NO-OPT R^2=0.896								
1	0.869	0.725	0.531	48	0.964	0.448	0.310	12
2	0.916	0.632	0.443	48	0.747	0.842	0.665	12
3	0.909	0.651	0.433	48	0.782	0.810	0.611	12
4	0.888	0.692	0.521	48	0.904	0.664	0.508	12
5	0.903	0.620	0.454	48	0.873	0.871	0.525	12
Svob2007-chal2 HF STO-3G/MPA/NCI QM-OPT R^2=0.896								
1	0.877	0.703	0.515	48	0.946	0.597	0.394	12
2	0.920	0.614	0.439	48	0.712	0.901	0.669	12
3	0.918	0.619	0.397	48	0.738	0.887	0.717	12
4	0.892	0.680	0.479	48	0.904	0.674	0.529	12
5	0.872	0.711	0.514	48	0.963	0.505	0.298	12
Bult2004-aim B3LYP 6-31G*/AIM/RDKIT QM-OPT R^2=0.895								
1	0.890	0.666	0.519	48	0.922	0.760	0.569	12
2	0.918	0.623	0.511	48	0.752	0.874	0.658	12
3	0.890	0.716	0.571	48	0.929	0.486	0.428	12
4	0.897	0.666	0.509	48	0.890	0.707	0.600	12
5	0.892	0.653	0.505	48	0.902	0.827	0.628	12
QM B3LYP 6-31G*/MPA/RDKIT NO-OPT R^2=0.895								
1	0.906	0.615	0.461	48	0.873	0.944	0.785	12
2	0.901	0.684	0.538	48	0.848	0.645	0.542	12
3	0.895	0.700	0.554	48	0.871	0.627	0.506	12
4	0.885	0.703	0.550	48	0.942	0.569	0.492	12
5	0.901	0.627	0.522	48	0.893	0.874	0.590	12
Bult2004-aim B3LYP 6-31G*/AIM/RDKIT MM-OPT R^2=0.895								
1	0.890	0.665	0.512	48	0.920	0.770	0.567	12
2	0.916	0.629	0.511	48	0.759	0.862	0.645	12
3	0.891	0.714	0.565	48	0.874	0.611	0.493	12
4	0.897	0.664	0.502	48	0.883	0.725	0.618	12
5	0.890	0.660	0.518	48	0.913	0.782	0.580	12
QM HF STO-3G/MPA/FROG2 NO-OPT R^2=0.894								
1	0.883	0.686	0.520	48	0.926	0.663	0.404	12
2	0.916	0.632	0.458	48	0.746	0.842	0.640	12
3	0.913	0.638	0.452	48	0.762	0.848	0.697	12
4	0.874	0.734	0.540	48	0.977	0.373	0.297	12
5	0.890	0.661	0.471	48	0.924	0.817	0.589	12
Svob2007-chal2 HF STO-3G/MPA/RDKIT NO-OPT R^2=0.894								
1	0.865	0.737	0.554	48	0.976	0.403	0.301	12
2	0.917	0.627	0.468	48	0.751	0.879	0.714	12
3	0.906	0.661	0.484	48	0.777	0.815	0.636	12
4	0.892	0.679	0.508	48	0.899	0.692	0.494	12
5	0.898	0.637	0.485	48	0.880	0.879	0.550	12
Bult2004-aim B3LYP 6-31G*/AIM/RDKIT MM-OPT R^2=0.892								
1	0.883	0.685	0.515	48	0.927	0.739	0.571	12
2	0.916	0.629	0.503	48	0.738	0.903	0.696	12
3	0.908	0.655	0.503	48	0.688	1.106	0.588	12
4	0.891	0.685	0.518	48	0.895	0.713	0.613	12
5	0.880	0.689	0.545	48	0.920	0.690	0.491	12
QM B3LYP 6-31G*/AIM/BALLOON MM-OPT R^2=0.891								

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
1	0.922	0.559	0.416	48	0.799	1.108	0.577	12
2	0.907	0.664	0.491	48	0.724	0.978	0.585	12
3	0.909	0.652	0.464	48	0.720	1.276	0.838	12
4	0.880	0.716	0.486	48	0.960	0.619	0.450	12
5	0.882	0.683	0.490	48	0.968	0.852	0.623	12
QM HF STO-3G/MPA/NCI QM-OPT R^2=0.891								
1	0.889	0.667	0.520	48	0.907	0.798	0.690	12
2	0.897	0.698	0.556	48	0.865	0.697	0.591	12
3	0.919	0.613	0.510	48	0.700	0.969	0.784	12
4	0.880	0.718	0.579	48	0.935	0.557	0.514	12
5	0.878	0.694	0.581	48	0.953	0.697	0.542	12
Svob2007-chal2 HF STO-3G/MPA/NCI MM-OPT R^2=0.891								
1	0.870	0.723	0.518	48	0.949	0.621	0.427	12
2	0.922	0.608	0.438	48	0.650	0.979	0.754	12
3	0.913	0.635	0.391	48	0.731	0.909	0.717	12
4	0.887	0.698	0.495	48	0.897	0.695	0.479	12
5	0.864	0.733	0.530	48	0.964	0.492	0.287	12
QM B3LYP 6-31G*/AIM/FROG2 NO-OPT R^2=0.891								
1	0.864	0.760	0.518	48	0.978	0.641	0.489	12
2	0.862	0.753	0.536	48	0.960	0.718	0.535	12
3	0.880	0.689	0.495	48	0.829	0.971	0.729	12
4	0.899	0.644	0.453	48	0.796	1.086	0.680	12
5	0.893	0.704	0.499	48	0.852	0.879	0.597	12
Bult2004-aim B3LYP 6-31G*/AIM/BALLOON QM-OPT R^2=0.890								
1	0.898	0.640	0.503	48	0.791	1.205	0.763	12
2	0.919	0.620	0.508	48	0.711	0.941	0.739	12
3	0.886	0.729	0.588	48	0.901	0.588	0.495	12
4	0.894	0.674	0.523	48	0.868	0.793	0.680	12
5	0.876	0.701	0.566	48	0.921	0.676	0.509	12
Bult2004-aim B3LYP 6-31G*/AIM/NCI NO-OPT R^2=0.890								
1	0.881	0.691	0.534	48	0.893	0.840	0.539	12
2	0.916	0.631	0.515	48	0.692	0.934	0.725	12
3	0.901	0.678	0.542	48	0.752	0.870	0.722	12
4	0.894	0.674	0.521	48	0.866	0.794	0.683	12
5	0.873	0.710	0.568	48	0.937	0.622	0.475	12
Bult2004-aim B3LYP 6-31G*/AIM/BALLOON NO-OPT R^2=0.888								
1	0.886	0.677	0.528	48	0.867	1.016	0.674	12
2	0.917	0.627	0.513	48	0.699	0.945	0.755	12
3	0.884	0.736	0.591	48	0.918	0.521	0.430	12
4	0.894	0.676	0.522	48	0.868	0.784	0.687	12
5	0.878	0.696	0.561	48	0.917	0.728	0.562	12
Svob2007-chal2 HF STO-3G/MPA/PUBCHEM MM-OPT R^2=0.887								
1	0.871	0.719	0.520	48	0.889	0.819	0.426	12
2	0.914	0.636	0.475	48	0.692	0.938	0.751	12
3	0.897	0.694	0.510	48	0.797	0.782	0.546	12
4	0.879	0.720	0.539	48	0.918	0.643	0.577	12
5	0.888	0.667	0.495	48	0.881	0.833	0.558	12
QM B3LYP 6-31G*/AIM/PUBCHEM QM-OPT R^2=0.886								
1	0.861	0.768	0.515	48	0.971	0.463	0.366	12
2	0.930	0.537	0.341	48	0.918	1.919	1.343	12
3	0.903	0.622	0.428	48	0.800	1.085	0.719	12
4	0.862	0.754	0.514	48	0.939	0.632	0.473	12
5	0.887	0.725	0.512	48	0.956	0.691	0.395	12
Bult2004-aim B3LYP 6-31G*/AIM/RDKIT NO-OPT R^2=0.885								
1	0.874	0.710	0.557	48	0.930	0.753	0.554	12
2	0.912	0.643	0.515	48	0.701	0.940	0.741	12
3	0.882	0.742	0.583	48	0.841	0.694	0.545	12
4	0.893	0.679	0.511	48	0.847	0.844	0.715	12
5	0.876	0.699	0.558	48	0.909	0.754	0.559	12

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
Bult2004-aim B3LYP 6-31G*/AIM/BALLOON MM-OPT R^2=0.885								
1	0.897	0.643	0.495	48	0.781	1.198	0.733	12
2	0.914	0.638	0.503	48	0.700	0.957	0.747	12
3	0.880	0.748	0.584	48	0.915	0.515	0.409	12
4	0.888	0.693	0.524	48	0.872	0.777	0.671	12
5	0.871	0.713	0.559	48	0.918	0.704	0.507	12
QM B3LYP 6-31G*/AIM/RDKIT NO-OPT R^2=0.882								
1	0.861	0.769	0.608	48	0.896	0.694	0.590	12
2	0.889	0.675	0.531	48	0.792	1.046	0.842	12
3	0.867	0.726	0.585	48	0.842	0.922	0.715	12
4	0.869	0.734	0.583	48	0.908	0.847	0.633	12
5	0.888	0.720	0.577	48	0.767	1.019	0.731	12
QM B3LYP 6-31G*/MPA/BALLOON MM-OPT R^2=0.881								
1	0.948	0.455	0.364	48	0.643	1.538	0.934	12
2	0.875	0.767	0.509	48	0.904	0.521	0.383	12
3	0.881	0.744	0.507	48	0.907	0.652	0.382	12
4	0.868	0.753	0.490	48	0.971	0.647	0.482	12
5	0.870	0.717	0.448	48	0.972	0.863	0.562	12
Svob2007-chal2 HF STO-3G/MPA/BALLOON NO-OPT R^2=0.873								
1	0.857	0.756	0.562	48	0.899	0.768	0.527	12
2	0.894	0.707	0.525	48	0.707	0.904	0.722	12
3	0.886	0.728	0.515	48	0.750	0.862	0.646	12
4	0.874	0.735	0.568	48	0.883	0.846	0.627	12
5	0.871	0.715	0.564	48	0.862	0.892	0.659	12
QM B3LYP 6-31G*/MPA/BALLOON QM-OPT R^2=0.871								
1	0.910	0.602	0.403	48	0.785	1.179	0.686	12
2	0.868	0.789	0.520	48	0.875	0.611	0.447	12
3	0.870	0.779	0.525	48	0.889	0.644	0.420	12
4	0.861	0.773	0.503	48	0.949	0.716	0.559	12
5	0.872	0.711	0.474	48	0.913	0.920	0.652	12
Svob2007-chal2 HF STO-3G/MPA/BALLOON QM-OPT R^2=0.869								
1	0.868	0.728	0.505	48	0.846	0.952	0.609	12
2	0.885	0.736	0.538	48	0.687	0.930	0.716	12
3	0.875	0.764	0.551	48	0.827	0.721	0.577	12
4	0.875	0.731	0.551	48	0.893	0.927	0.754	12
5	0.866	0.728	0.561	48	0.881	0.897	0.572	12
Bult2004-aim B3LYP 6-31G*/AIM/FROG2 QM-OPT R^2=0.869								
1	0.906	0.615	0.419	48	0.550	1.603	0.963	12
2	0.904	0.674	0.547	48	0.710	1.147	0.992	12
3	0.874	0.766	0.598	48	0.813	0.767	0.635	12
4	0.871	0.743	0.598	48	0.847	0.873	0.639	12
5	0.835	0.807	0.633	48	0.953	0.523	0.436	12
Bult2004-aim B3LYP 6-31G*/AIM/FROG2 MM-OPT R^2=0.868								
1	0.904	0.622	0.433	48	0.612	1.504	0.898	12
2	0.903	0.676	0.548	48	0.711	1.159	1.002	12
3	0.872	0.773	0.599	48	0.825	0.738	0.618	12
4	0.870	0.746	0.598	48	0.846	0.877	0.636	12
5	0.837	0.803	0.628	48	0.934	0.627	0.524	12
QM B3LYP 6-31G*/AIM/PUBCHEM MM-OPT R^2=0.864								
1	0.831	0.846	0.576	48	0.985	0.479	0.344	12
2	0.936	0.513	0.359	48	0.808	2.845	1.620	12
3	0.930	0.527	0.388	48	0.585	1.678	0.960	12
4	0.841	0.807	0.538	48	0.897	0.751	0.571	12
5	0.851	0.830	0.588	48	0.894	0.540	0.362	12
QM B3LYP 6-31G*/NPA/NCI QM-OPT R^2=0.861								
1	0.862	0.744	0.621	48	0.815	1.224	0.975	12
2	0.883	0.744	0.615	48	0.623	1.028	0.859	12
3	0.882	0.742	0.625	48	0.680	0.998	0.792	12
4	0.844	0.819	0.696	48	0.922	0.589	0.519	12

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
5	0.861	0.743	0.627	48	0.876	0.954	0.765	12
Bult2004-aim B3LYP 6-31G*/AIM/FROG2 NO-OPT R^2=0.858								
1	0.874	0.712	0.551	48	0.758	1.234	0.873	12
2	0.870	0.784	0.623	48	0.785	0.805	0.666	12
3	0.868	0.785	0.615	48	0.812	0.833	0.708	12
4	0.872	0.742	0.599	48	0.795	1.023	0.771	12
5	0.831	0.818	0.672	48	0.932	0.698	0.527	12
QM B3LYP 6-31G*/AIM/BALLOON QM-OPT R^2=0.856								
1	0.837	0.831	0.594	48	0.972	0.453	0.370	12
2	0.880	0.702	0.460	48	0.927	1.167	0.908	12
3	0.895	0.646	0.481	48	0.721	1.297	0.837	12
4	0.871	0.728	0.535	48	0.718	1.165	0.816	12
5	0.883	0.736	0.568	48	0.591	1.066	0.638	12
Svob2007-chal2 HF STO-3G/MPA/BALLOON MM-OPT R^2=0.852								
1	0.861	0.747	0.502	48	0.808	1.059	0.701	12
2	0.876	0.766	0.546	48	0.634	1.003	0.684	12
3	0.858	0.814	0.554	48	0.809	0.752	0.561	12
4	0.845	0.815	0.585	48	0.897	0.791	0.659	12
5	0.841	0.794	0.584	48	0.888	0.887	0.568	12
QM B3LYP 6-31G*/NPA/RDKIT NO-OPT R^2=0.796								
1	0.813	0.866	0.724	48	0.772	1.217	1.001	12
2	0.812	0.944	0.773	48	0.705	0.989	0.781	12
3	0.815	0.928	0.740	48	0.626	1.255	0.874	12
4	0.766	1.002	0.835	48	0.926	0.666	0.489	12
5	0.799	0.891	0.741	48	0.780	1.149	0.908	12
QM HF STO-3G/MPA/RDKIT NO-OPT R^2=0.782								
1	0.798	0.900	0.719	48	0.733	1.260	1.148	12
2	0.822	0.918	0.760	48	0.498	1.295	0.997	12
3	0.809	0.943	0.754	48	0.391	1.975	1.162	12
4	0.760	1.016	0.868	48	0.872	0.797	0.553	12
5	0.766	0.961	0.779	48	0.810	1.067	0.896	12

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
Carboxylic acids								
Bult2002-npa B3LYP 6-31G*/NPA/BALLOON MM-OPT R^2=0.853								
1	0.871	0.315	0.263	65	0.909	0.290	0.232	17
2	0.878	0.286	0.233	65	0.899	0.410	0.293	17
3	0.894	0.302	0.244	66	0.806	0.340	0.282	16
4	0.901	0.292	0.229	66	0.765	0.377	0.324	16
5	0.879	0.319	0.258	66	0.901	0.256	0.212	16
Bult2002-npa B3LYP 6-31G*/NPA/BALLOON NO-OPT R^2=0.862								
1	0.873	0.313	0.256	65	0.925	0.263	0.216	17
2	0.887	0.276	0.225	65	0.897	0.449	0.339	17
3	0.900	0.293	0.239	66	0.788	0.349	0.291	16
4	0.908	0.282	0.222	66	0.762	0.389	0.334	16
5	0.878	0.321	0.261	66	0.946	0.208	0.167	16
Bult2002-npa B3LYP 6-31G*/NPA/BALLOON QM-OPT R^2=0.862								
1	0.877	0.308	0.253	65	0.915	0.281	0.232	17
2	0.885	0.277	0.228	65	0.878	0.462	0.340	17
3	0.903	0.289	0.232	66	0.780	0.362	0.304	16
4	0.907	0.284	0.228	66	0.772	0.372	0.327	16
5	0.885	0.311	0.254	66	0.903	0.251	0.197	16
Bult2002-npa B3LYP 6-31G*/NPA/FROG2 MM-OPT R^2=0.881								
1	0.828	0.365	0.262	65	0.966	0.197	0.157	17
2	0.910	0.245	0.188	65	0.753	0.609	0.392	17
3	0.873	0.330	0.225	66	0.750	0.406	0.347	16
4	0.863	0.344	0.239	66	0.844	0.340	0.271	16
5	0.864	0.339	0.243	66	0.705	0.462	0.313	16
Bult2002-npa B3LYP 6-31G*/NPA/FROG2 NO-OPT R^2=0.805								
1	0.796	0.397	0.295	65	0.938	0.249	0.186	17
2	0.848	0.320	0.234	65	0.797	0.590	0.404	17
3	0.841	0.370	0.265	66	0.802	0.396	0.338	16
4	0.863	0.344	0.240	66	0.500	0.690	0.461	16
5	0.852	0.354	0.273	66	0.713	0.463	0.290	16
Bult2002-npa B3LYP 6-31G*/NPA/FROG2 QM-OPT R^2=0.891								
1	0.835	0.357	0.247	65	0.969	0.200	0.156	17
2	0.920	0.232	0.174	65	0.761	0.610	0.392	17
3	0.870	0.334	0.228	66	0.841	0.323	0.259	16
4	0.875	0.328	0.218	66	0.569	0.612	0.365	16
5	0.879	0.320	0.221	66	0.558	0.580	0.295	16
Bult2002-npa B3LYP 6-31G*/NPA/NCI MM-OPT R^2=0.863								
1	0.868	0.320	0.243	65	0.955	0.220	0.161	17
2	0.892	0.269	0.195	65	0.868	0.436	0.317	17
3	0.896	0.298	0.215	66	0.856	0.347	0.305	16
4	0.915	0.272	0.207	66	0.749	0.415	0.323	16
5	0.877	0.322	0.245	66	0.947	0.194	0.156	16
Bult2002-npa B3LYP 6-31G*/NPA/NCI NO-OPT R^2=0.862								
1	0.873	0.313	0.247	65	0.951	0.238	0.197	17
2	0.888	0.274	0.205	65	0.897	0.409	0.289	17
3	0.896	0.299	0.234	66	0.871	0.310	0.264	16
4	0.920	0.263	0.204	66	0.700	0.428	0.369	16
5	0.884	0.312	0.250	66	0.933	0.240	0.205	16
Bult2002-npa B3LYP 6-31G*/NPA/NCI QM-OPT R^2=0.868								
1	0.865	0.323	0.259	65	0.950	0.238	0.184	17
2	0.893	0.267	0.201	65	0.835	0.508	0.360	17
3	0.892	0.304	0.225	66	0.877	0.353	0.318	16
4	0.901	0.293	0.224	66	0.801	0.369	0.298	16
5	0.883	0.314	0.237	66	0.752	0.419	0.301	16
Bult2002-npa B3LYP 6-31G*/NPA/OPENBABEL MM-OPT R^2=0.816								
1	0.820	0.373	0.288	65	0.935	0.249	0.195	17
2	0.849	0.318	0.238	65	0.857	0.649	0.411	17
3	0.863	0.344	0.254	66	0.799	0.409	0.322	16

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
4	0.869	0.337	0.255	66	0.713	0.423	0.341	16
5	0.866	0.336	0.256	66	0.700	0.460	0.353	16
Bult2002-npa B3LYP 6-31G*/NPA/OPENBABEL NO-OPT R^2=0.813								
1	0.818	0.374	0.290	65	0.940	0.245	0.189	17
2	0.847	0.320	0.237	65	0.875	0.571	0.368	17
3	0.861	0.346	0.256	66	0.805	0.403	0.322	16
4	0.868	0.337	0.254	66	0.714	0.422	0.336	16
5	0.865	0.338	0.255	66	0.724	0.441	0.342	16
Bult2002-npa B3LYP 6-31G*/NPA/OPENBABEL QM-OPT R^2=0.833								
1	0.828	0.364	0.277	65	0.941	0.240	0.182	17
2	0.864	0.302	0.232	65	0.843	0.697	0.415	17
3	0.869	0.335	0.244	66	0.808	0.415	0.328	16
4	0.878	0.325	0.244	66	0.718	0.423	0.335	16
5	0.868	0.333	0.254	66	0.720	0.444	0.332	16
Bult2002-npa B3LYP 6-31G*/NPA/PUBCHEM MM-OPT R^2=0.896								
1	0.849	0.341	0.258	65	0.960	0.215	0.162	17
2	0.924	0.225	0.175	65	0.752	0.597	0.392	17
3	0.887	0.312	0.222	66	0.741	0.416	0.363	16
4	0.884	0.317	0.225	66	0.835	0.339	0.265	16
5	0.864	0.338	0.249	66	0.931	0.211	0.168	16
Bult2002-npa B3LYP 6-31G*/NPA/PUBCHEM NO-OPT R^2=0.894								
1	0.838	0.354	0.262	65	0.956	0.231	0.208	17
2	0.922	0.228	0.170	65	0.725	0.631	0.419	17
3	0.871	0.333	0.235	66	0.828	0.337	0.293	16
4	0.872	0.332	0.230	66	0.847	0.337	0.261	16
5	0.856	0.349	0.251	66	0.912	0.241	0.195	16
Bult2002-npa B3LYP 6-31G*/NPA/PUBCHEM QM-OPT R^2=0.896								
1	0.868	0.319	0.236	65	0.979	0.177	0.142	17
2	0.920	0.232	0.170	65	0.841	0.485	0.333	17
3	0.904	0.287	0.204	66	0.818	0.371	0.296	16
4	0.903	0.289	0.202	66	0.867	0.336	0.272	16
5	0.890	0.304	0.211	66	0.884	0.273	0.185	16
Bult2002-npa B3LYP 6-31G*/NPA/RDKIT MM-OPT R^2=0.832								
1	0.845	0.346	0.280	65	0.924	0.273	0.206	17
2	0.859	0.308	0.247	65	0.879	0.521	0.345	17
3	0.906	0.285	0.222	66	0.525	0.550	0.461	16
4	0.884	0.317	0.253	66	0.726	0.411	0.350	16
5	0.868	0.334	0.259	66	0.850	0.332	0.266	16
Bult2002-npa B3LYP 6-31G*/NPA/RDKIT QM-OPT R^2=0.841								
1	0.864	0.324	0.253	65	0.942	0.255	0.194	17
2	0.880	0.283	0.235	65	0.794	0.623	0.440	17
3	0.925	0.254	0.195	66	0.584	0.569	0.425	16
4	0.892	0.306	0.241	66	0.780	0.358	0.304	16
5	0.882	0.316	0.249	66	0.856	0.306	0.213	16
Bult2002-npa B3LYP 6-31G*/NPA/RDKIT MM-OPT R^2=0.838								
1	0.851	0.339	0.256	65	0.933	0.256	0.211	17
2	0.870	0.295	0.230	65	0.877	0.461	0.318	17
3	0.887	0.312	0.223	66	0.858	0.392	0.328	16
4	0.896	0.300	0.222	66	0.751	0.420	0.322	16
5	0.863	0.340	0.256	66	0.920	0.240	0.192	16
Bult2002-npa B3LYP 6-31G*/NPA/RDKIT NO-OPT R^2=0.862								
1	0.870	0.316	0.238	65	0.941	0.235	0.191	17
2	0.896	0.264	0.205	65	0.882	0.499	0.385	17
3	0.899	0.295	0.221	66	0.858	0.338	0.269	16
4	0.906	0.285	0.219	66	0.789	0.363	0.273	16
5	0.885	0.311	0.229	66	0.900	0.351	0.252	16
Bult2004-aim B3LYP 6-31G*/AIM/BALLOON MM-OPT R^2=0.843								
1	0.805	0.388	0.316	65	0.941	0.256	0.229	17
2	0.883	0.281	0.230	65	0.745	0.638	0.454	17

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
3	0.841	0.370	0.291	66	0.808	0.399	0.345	16
4	0.855	0.354	0.281	66	0.652	0.411	0.345	16
5	0.832	0.377	0.302	66	0.896	0.306	0.272	16
Bult2004-aim B3LYP 6-31G*/AIM/BALLOON NO-OPT R^2=0.845								
1	0.822	0.370	0.295	65	0.951	0.237	0.199	17
2	0.882	0.281	0.224	65	0.805	0.551	0.421	17
3	0.857	0.351	0.272	66	0.790	0.353	0.292	16
4	0.869	0.336	0.262	66	0.701	0.379	0.293	16
5	0.848	0.359	0.282	66	0.934	0.304	0.279	16
Bult2004-aim B3LYP 6-31G*/AIM/BALLOON QM-OPT R^2=0.860								
1	0.830	0.362	0.300	65	0.919	0.303	0.250	17
2	0.892	0.269	0.207	65	0.738	0.680	0.432	17
3	0.857	0.350	0.284	66	0.786	0.382	0.315	16
4	0.869	0.336	0.267	66	0.691	0.397	0.307	16
5	0.851	0.355	0.287	66	0.909	0.327	0.290	16
Bult2004-aim B3LYP 6-31G*/AIM/FROG2 MM-OPT R^2=0.825								
1	0.821	0.372	0.298	65	0.939	0.271	0.237	17
2	0.871	0.294	0.235	65	0.812	0.533	0.415	17
3	0.860	0.347	0.259	66	0.753	0.403	0.340	16
4	0.859	0.349	0.283	66	0.639	0.421	0.324	16
5	0.854	0.351	0.280	66	0.653	0.511	0.368	16
Bult2004-aim B3LYP 6-31G*/AIM/FROG2 NO-OPT R^2=0.821								
1	0.848	0.342	0.251	65	0.939	0.259	0.188	17
2	0.855	0.311	0.210	65	0.905	0.401	0.263	17
3	0.886	0.313	0.225	66	0.812	0.390	0.312	16
4	0.894	0.303	0.222	66	0.596	0.625	0.421	16
5	0.885	0.312	0.224	66	0.787	0.400	0.235	16
Bult2004-aim B3LYP 6-31G*/AIM/FROG2 QM-OPT R^2=0.870								
1	0.854	0.336	0.258	65	0.932	0.282	0.222	17
2	0.911	0.244	0.180	65	0.801	0.554	0.419	17
3	0.872	0.332	0.253	66	0.906	0.280	0.226	16
4	0.884	0.317	0.240	66	0.745	0.347	0.269	16
5	0.870	0.332	0.251	66	0.907	0.310	0.243	16
Bult2004-aim B3LYP 6-31G*/AIM/NCI MM-OPT R^2=0.833								
1	0.795	0.398	0.308	65	0.944	0.262	0.239	17
2	0.880	0.284	0.218	65	0.736	0.638	0.458	17
3	0.828	0.385	0.296	66	0.858	0.360	0.324	16
4	0.843	0.369	0.279	66	0.712	0.388	0.294	16
5	0.871	0.331	0.274	66	0.016	1.304	0.632	16
Bult2004-aim B3LYP 6-31G*/AIM/NCI NO-OPT R^2=0.843								
1	0.862	0.326	0.270	65	0.942	0.271	0.239	17
2	0.888	0.274	0.219	65	0.863	0.483	0.374	17
3	0.879	0.323	0.270	66	0.899	0.291	0.235	16
4	0.894	0.302	0.245	66	0.742	0.358	0.308	16
5	0.883	0.315	0.259	66	0.806	0.350	0.312	16
Bult2004-aim B3LYP 6-31G*/AIM/NCI QM-OPT R^2=0.838								
1	0.805	0.388	0.310	65	0.932	0.283	0.242	17
2	0.885	0.278	0.210	65	0.647	0.809	0.502	17
3	0.851	0.358	0.289	66	0.733	0.475	0.375	16
4	0.856	0.353	0.281	66	0.661	0.435	0.320	16
5	0.877	0.322	0.269	66	0.076	1.062	0.544	16
Bult2004-aim B3LYP 6-31G*/AIM/OPENBABEL MM-OPT R^2=0.820								
1	0.798	0.395	0.320	65	0.965	0.224	0.191	17
2	0.864	0.302	0.238	65	0.816	0.594	0.422	17
3	0.840	0.371	0.288	66	0.917	0.371	0.310	16
4	0.854	0.356	0.272	66	0.651	0.409	0.342	16
5	0.847	0.359	0.290	66	0.601	0.538	0.361	16
Bult2004-aim B3LYP 6-31G*/AIM/OPENBABEL NO-OPT R^2=0.820								
1	0.786	0.406	0.329	65	0.962	0.222	0.201	17

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
2	0.864	0.301	0.238	65	0.791	0.647	0.453	17
3	0.833	0.379	0.295	66	0.883	0.396	0.332	16
4	0.844	0.367	0.286	66	0.651	0.412	0.340	16
5	0.840	0.368	0.295	66	0.546	0.570	0.389	16
Bult2004-aim B3LYP 6-31G*/AIM/OPENBABEL QM-OPT R^2=0.846								
1	0.817	0.376	0.300	65	0.970	0.206	0.182	17
2	0.887	0.275	0.207	65	0.813	0.623	0.424	17
3	0.855	0.353	0.273	66	0.909	0.366	0.300	16
4	0.868	0.338	0.258	66	0.689	0.392	0.322	16
5	0.873	0.328	0.263	66	0.542	0.588	0.381	16
Bult2004-aim B3LYP 6-31G*/AIM/PUBCHEM MM-OPT R^2=0.859								
1	0.851	0.339	0.263	65	0.948	0.272	0.248	17
2	0.901	0.257	0.191	65	0.807	0.521	0.392	17
3	0.873	0.331	0.258	66	0.698	0.407	0.300	16
4	0.883	0.318	0.238	66	0.761	0.343	0.264	16
5	0.865	0.337	0.268	66	0.899	0.278	0.241	16
Bult2004-aim B3LYP 6-31G*/AIM/PUBCHEM NO-OPT R^2=0.862								
1	0.856	0.333	0.256	65	0.924	0.298	0.232	17
2	0.906	0.251	0.187	65	0.791	0.537	0.394	17
3	0.880	0.322	0.248	66	0.399	0.768	0.382	16
4	0.883	0.319	0.238	66	0.768	0.332	0.256	16
5	0.871	0.331	0.257	66	0.910	0.309	0.241	16
Bult2004-aim B3LYP 6-31G*/AIM/PUBCHEM QM-OPT R^2=0.871								
1	0.879	0.306	0.235	65	0.939	0.276	0.243	17
2	0.911	0.244	0.179	65	0.862	0.457	0.351	17
3	0.891	0.306	0.239	66	0.806	0.327	0.231	16
4	0.908	0.282	0.211	66	0.744	0.353	0.283	16
5	0.884	0.313	0.248	66	0.937	0.224	0.184	16
Bult2004-aim B3LYP 6-31G*/AIM/RDKIT MM-OPT R^2=0.853								
1	0.811	0.382	0.311	65	0.953	0.240	0.215	17
2	0.895	0.265	0.205	65	0.750	0.661	0.437	17
3	0.850	0.359	0.279	66	0.779	0.404	0.348	16
4	0.854	0.356	0.285	66	0.727	0.361	0.295	16
5	0.845	0.361	0.296	66	0.862	0.376	0.319	16
Bult2004-aim B3LYP 6-31G*/AIM/RDKIT QM-OPT R^2=0.847								
1	0.834	0.358	0.286	65	0.906	0.323	0.288	17
2	0.887	0.275	0.213	65	0.744	0.683	0.418	17
3	0.865	0.341	0.277	66	0.822	0.439	0.369	16
4	0.866	0.341	0.270	66	0.720	0.378	0.289	16
5	0.849	0.357	0.292	66	0.865	0.334	0.273	16
Bult2004-aim B3LYP 6-31G*/AIM/RDKIT MM-OPT R^2=0.812								
1	0.833	0.359	0.303	65	0.907	0.304	0.259	17
2	0.858	0.308	0.252	65	0.829	0.520	0.413	17
3	0.867	0.339	0.278	66	0.788	0.397	0.342	16
4	0.871	0.334	0.282	66	0.707	0.394	0.319	16
5	0.846	0.361	0.301	66	0.865	0.296	0.268	16
Bult2004-aim B3LYP 6-31G*/AIM/RDKIT NO-OPT R^2=0.821								
1	0.855	0.334	0.273	65	0.931	0.269	0.209	17
2	0.873	0.291	0.228	65	0.870	0.449	0.382	17
3	0.890	0.307	0.250	66	0.696	0.391	0.306	16
4	0.885	0.316	0.253	66	0.766	0.336	0.270	16
5	0.877	0.323	0.258	66	0.826	0.384	0.315	16
Chavez2006 B3LYP 6-31G*/MPA/BALLOON MM-OPT R^2=0.887								
1	0.911	0.262	0.215	65	0.931	0.264	0.214	17
2	0.913	0.242	0.190	65	0.870	0.416	0.314	17
3	0.918	0.266	0.215	66	0.904	0.234	0.212	16
4	0.934	0.239	0.193	66	0.794	0.339	0.284	16
5	0.918	0.263	0.220	66	0.897	0.255	0.189	16
Chavez2006 B3LYP 6-31G*/MPA/BALLOON NO-OPT R^2=0.891								

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
1	0.918	0.251	0.203	65	0.938	0.257	0.182	17
2	0.918	0.235	0.185	65	0.894	0.375	0.303	17
3	0.931	0.244	0.194	66	0.851	0.275	0.245	16
4	0.940	0.228	0.184	66	0.814	0.327	0.257	16
5	0.919	0.261	0.213	66	0.940	0.201	0.152	16
Chavez2006 B3LYP 6-31G*/MPA/BALLOON QM-OPT R^2=0.889								
1	0.909	0.265	0.211	65	0.937	0.249	0.186	17
2	0.913	0.241	0.189	65	0.863	0.436	0.324	17
3	0.924	0.256	0.202	66	0.850	0.278	0.236	16
4	0.932	0.242	0.190	66	0.807	0.332	0.272	16
5	0.917	0.264	0.213	66	0.903	0.250	0.187	16
Chavez2006 B3LYP 6-31G*/MPA/FROG2 MM-OPT R^2=0.900								
1	0.863	0.325	0.227	65	0.961	0.194	0.152	17
2	0.925	0.224	0.164	65	0.796	0.525	0.318	17
3	0.893	0.303	0.211	66	0.839	0.315	0.258	16
4	0.894	0.303	0.202	66	0.744	0.412	0.290	16
5	0.887	0.309	0.220	66	0.841	0.335	0.232	16
Chavez2006 B3LYP 6-31G*/MPA/FROG2 NO-OPT R^2=0.849								
1	0.846	0.345	0.267	65	0.965	0.181	0.158	17
2	0.886	0.276	0.210	65	0.841	0.472	0.332	17
3	0.879	0.322	0.239	66	0.887	0.321	0.293	16
4	0.916	0.270	0.191	66	0.500	0.735	0.456	16
5	0.883	0.314	0.243	66	0.830	0.359	0.266	16
Chavez2006 B3LYP 6-31G*/MPA/FROG2 QM-OPT R^2=0.910								
1	0.874	0.311	0.223	65	0.965	0.182	0.143	17
2	0.936	0.207	0.157	65	0.792	0.531	0.333	17
3	0.896	0.299	0.208	66	0.892	0.246	0.201	16
4	0.913	0.275	0.191	66	0.426	0.735	0.372	16
5	0.901	0.289	0.207	66	0.855	0.325	0.199	16
Chavez2006 B3LYP 6-31G*/MPA/NCI MM-OPT R^2=0.892								
1	0.908	0.267	0.204	65	0.948	0.226	0.156	17
2	0.920	0.231	0.172	65	0.838	0.465	0.320	17
3	0.921	0.260	0.204	66	0.898	0.251	0.220	16
4	0.944	0.219	0.169	66	0.763	0.383	0.324	16
5	0.918	0.263	0.203	66	0.892	0.278	0.191	16
Chavez2006 B3LYP 6-31G*/MPA/NCI NO-OPT R^2=0.897								
1	0.915	0.257	0.200	65	0.957	0.200	0.140	17
2	0.921	0.229	0.173	65	0.901	0.364	0.275	17
3	0.927	0.251	0.193	66	0.916	0.234	0.203	16
4	0.948	0.212	0.170	66	0.806	0.360	0.284	16
5	0.923	0.255	0.200	66	0.940	0.204	0.173	16
Chavez2006 B3LYP 6-31G*/MPA/NCI QM-OPT R^2=0.896								
1	0.913	0.259	0.209	65	0.955	0.216	0.166	17
2	0.921	0.230	0.178	65	0.829	0.481	0.323	17
3	0.925	0.253	0.202	66	0.897	0.254	0.216	16
4	0.939	0.230	0.185	66	0.823	0.322	0.269	16
5	0.931	0.241	0.191	66	0.826	0.343	0.238	16
Chavez2006 B3LYP 6-31G*/MPA/OPENBABEL MM-OPT R^2=0.857								
1	0.891	0.290	0.229	65	0.934	0.266	0.178	17
2	0.886	0.277	0.208	65	0.926	0.316	0.258	17
3	0.904	0.287	0.222	66	0.888	0.271	0.213	16
4	0.923	0.259	0.200	66	0.750	0.376	0.308	16
5	0.914	0.270	0.211	66	0.818	0.348	0.264	16
Chavez2006 B3LYP 6-31G*/MPA/OPENBABEL NO-OPT R^2=0.851								
1	0.884	0.299	0.238	65	0.932	0.264	0.186	17
2	0.881	0.282	0.214	65	0.913	0.345	0.273	17
3	0.901	0.291	0.230	66	0.868	0.286	0.223	16
4	0.917	0.268	0.210	66	0.747	0.376	0.307	16
5	0.908	0.279	0.223	66	0.820	0.345	0.264	16

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
Chavez2006 B3LYP 6-31G*/MPA/OPENBABEL QM-OPT R^2=0.875								
1	0.899	0.279	0.218	65	0.942	0.239	0.156	17
2	0.899	0.260	0.195	65	0.917	0.336	0.262	17
3	0.913	0.274	0.209	66	0.901	0.263	0.223	16
4	0.932	0.242	0.185	66	0.756	0.370	0.298	16
5	0.919	0.261	0.206	66	0.826	0.341	0.255	16
Chavez2006 B3LYP 6-31G*/MPA/PUBCHEM MM-OPT R^2=0.920								
1	0.876	0.309	0.218	65	0.981	0.136	0.107	17
2	0.944	0.195	0.151	65	0.794	0.525	0.334	17
3	0.905	0.286	0.195	66	0.818	0.340	0.278	16
4	0.914	0.272	0.180	66	0.826	0.336	0.266	16
5	0.892	0.303	0.216	66	0.954	0.177	0.141	16
Chavez2006 B3LYP 6-31G*/MPA/PUBCHEM NO-OPT R^2=0.916								
1	0.866	0.322	0.226	65	0.965	0.181	0.135	17
2	0.940	0.201	0.152	65	0.770	0.554	0.345	17
3	0.888	0.311	0.215	66	0.889	0.266	0.240	16
4	0.900	0.294	0.186	66	0.847	0.321	0.251	16
5	0.884	0.313	0.223	66	0.916	0.243	0.185	16
Chavez2006 B3LYP 6-31G*/MPA/PUBCHEM QM-OPT R^2=0.917								
1	0.905	0.271	0.204	65	0.979	0.139	0.112	17
2	0.939	0.202	0.156	65	0.888	0.386	0.284	17
3	0.928	0.249	0.187	66	0.862	0.298	0.216	16
4	0.933	0.241	0.167	66	0.876	0.298	0.237	16
5	0.920	0.260	0.193	66	0.933	0.212	0.145	16
Chavez2006 B3LYP 6-31G*/MPA/RDKIT MM-OPT R^2=0.886								
1	0.912	0.261	0.202	65	0.933	0.264	0.175	17
2	0.911	0.244	0.191	65	0.924	0.335	0.261	17
3	0.928	0.249	0.198	66	0.802	0.339	0.266	16
4	0.935	0.238	0.181	66	0.819	0.342	0.273	16
5	0.920	0.260	0.205	66	0.913	0.261	0.191	16
Chavez2006 B3LYP 6-31G*/MPA/RDKIT QM-OPT R^2=0.882								
1	0.906	0.269	0.218	65	0.936	0.257	0.191	17
2	0.913	0.242	0.194	65	0.889	0.401	0.305	17
3	0.929	0.248	0.201	66	0.791	0.388	0.286	16
4	0.924	0.256	0.201	66	0.827	0.311	0.248	16
5	0.919	0.262	0.211	66	0.859	0.298	0.219	16
Chavez2006 B3LYP 6-31G*/MPA/RDKIT MM-OPT R^2=0.874								
1	0.897	0.281	0.219	65	0.929	0.272	0.193	17
2	0.903	0.255	0.203	65	0.870	0.422	0.306	17
3	0.912	0.275	0.209	66	0.875	0.287	0.249	16
4	0.928	0.249	0.192	66	0.784	0.372	0.288	16
5	0.906	0.282	0.217	66	0.899	0.265	0.207	16
Chavez2006 B3LYP 6-31G*/MPA/RDKIT NO-OPT R^2=0.900								
1	0.927	0.238	0.185	65	0.950	0.238	0.189	17
2	0.931	0.215	0.174	65	0.924	0.320	0.256	17
3	0.931	0.244	0.192	66	0.933	0.207	0.178	16
4	0.945	0.217	0.170	66	0.853	0.306	0.234	16
5	0.929	0.245	0.190	66	0.942	0.248	0.179	16
QM B3LYP 6-31G*/AIM/BALLOON MM-OPT R^2=0.801								
1	0.828	0.365	0.284	65	0.908	0.356	0.267	17
2	0.813	0.354	0.264	65	0.766	0.588	0.342	17
3	0.881	0.321	0.236	66	0.483	0.514	0.370	16
4	0.874	0.331	0.259	66	0.733	0.477	0.381	16
5	0.842	0.365	0.277	66	0.832	0.326	0.249	16
QM B3LYP 6-31G*/AIM/BALLOON NO-OPT R^2=0.777								
1	0.819	0.374	0.280	65	0.899	0.385	0.299	17
2	0.777	0.386	0.300	65	0.776	0.561	0.306	17
3	0.857	0.350	0.270	66	0.663	0.462	0.374	16
4	0.849	0.361	0.265	66	0.761	0.418	0.355	16

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
5	0.872	0.329	0.263	66	0.186	2.746	1.024	16
QM B3LYP 6-31G*/AIM/BALLOON QM-OPT R^2=0.823								
1	0.786	0.407	0.322	65	0.003	9.592	2.562	17
2	0.750	0.409	0.319	65	0.285	1.128	0.487	17
3	0.765	0.450	0.333	66	0.630	0.516	0.431	16
4	0.756	0.460	0.355	66	0.760	0.477	0.389	16
5	0.801	0.410	0.321	66	0.376	0.676	0.511	16
QM B3LYP 6-31G*/AIM/FROG2 MM-OPT R^2=0.901								
1	0.911	0.262	0.203	65	0.959	0.207	0.174	17
2	0.912	0.242	0.187	65	0.895	0.386	0.254	17
3	0.926	0.252	0.192	66	0.903	0.256	0.204	16
4	0.940	0.227	0.177	66	0.765	0.407	0.293	16
5	0.935	0.234	0.176	66	0.840	0.357	0.267	16
QM B3LYP 6-31G*/AIM/FROG2 NO-OPT R^2=0.880								
1	0.899	0.302	0.209	65	0.701	0.435	0.275	17
2	0.880	0.285	0.204	65	0.865	0.420	0.244	17
3	0.913	0.272	0.207	66	0.682	0.481	0.304	16
4	0.881	0.316	0.228	66	0.794	0.371	0.261	16
5	0.879	0.302	0.213	66	0.920	0.346	0.244	16
QM B3LYP 6-31G*/AIM/FROG2 QM-OPT R^2=0.908								
1	0.904	0.273	0.198	65	0.960	0.199	0.168	17
2	0.923	0.228	0.169	65	0.867	0.455	0.295	17
3	0.919	0.263	0.199	66	0.933	0.267	0.200	16
4	0.944	0.219	0.168	66	0.766	0.421	0.308	16
5	0.914	0.269	0.195	66	0.927	0.233	0.193	16
QM B3LYP 6-31G*/AIM/NCI MM-OPT R^2=0.886								
1	0.902	0.275	0.202	65	0.981	0.180	0.150	17
2	0.906	0.251	0.176	65	0.924	0.409	0.244	17
3	0.939	0.229	0.172	66	0.812	0.369	0.276	16
4	0.939	0.230	0.172	66	0.771	0.396	0.280	16
5	0.918	0.264	0.193	66	0.920	0.226	0.170	16
QM B3LYP 6-31G*/AIM/NCI NO-OPT R^2=0.884								
1	0.922	0.245	0.176	65	0.942	0.232	0.140	17
2	0.910	0.246	0.172	65	0.908	0.357	0.206	17
3	0.948	0.212	0.150	66	0.889	0.355	0.272	16
4	0.945	0.219	0.163	66	0.874	0.355	0.269	16
5	0.925	0.252	0.179	66	0.920	0.227	0.186	16
QM B3LYP 6-31G*/AIM/NCI QM-OPT R^2=0.886								
1	0.901	0.276	0.205	65	0.948	0.224	0.197	17
2	0.897	0.263	0.201	65	0.768	0.575	0.281	17
3	0.916	0.269	0.207	66	0.914	0.297	0.220	16
4	0.942	0.225	0.175	66	0.719	0.456	0.340	16
5	0.923	0.254	0.197	66	0.859	0.320	0.238	16
QM B3LYP 6-31G*/AIM/OPENBABEL MM-OPT R^2=0.844								
1	0.765	0.425	0.329	65	0.767	0.485	0.359	17
2	0.848	0.319	0.246	65	0.104	6.085	1.869	17
3	0.781	0.433	0.319	66	0.727	0.481	0.388	16
4	0.778	0.438	0.338	66	0.711	0.429	0.312	16
5	0.787	0.424	0.319	66	0.758	0.504	0.367	16
QM B3LYP 6-31G*/AIM/OPENBABEL NO-OPT R^2=0.849								
1	0.845	0.346	0.259	65	0.868	0.372	0.269	17
2	0.849	0.318	0.243	65	0.749	0.598	0.385	17
3	0.873	0.331	0.250	66	0.687	0.468	0.344	16
4	0.866	0.340	0.258	66	0.797	0.390	0.300	16
5	0.846	0.361	0.265	66	0.902	0.319	0.256	16
QM B3LYP 6-31G*/AIM/OPENBABEL QM-OPT R^2=0.834								
1	0.768	0.423	0.324	65	0.591	0.623	0.433	17
2	0.768	0.395	0.295	65	0.184	1.200	0.623	17
3	0.849	0.360	0.283	66	0.031	677.501	164.688	16

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
4	0.748	0.467	0.355	66	0.813	0.507	0.435	16
5	0.720	0.487	0.355	66	0.804	0.393	0.351	16
QM B3LYP 6-31G*/AIM/PUBCHEM MM-OPT R^2=0.906								
1	0.899	0.280	0.200	65	0.960	0.196	0.158	17
2	0.898	0.261	0.193	65	0.932	0.301	0.202	17
3	0.931	0.244	0.173	66	0.736	0.404	0.255	16
4	0.942	0.223	0.167	66	0.750	0.428	0.273	16
5	0.917	0.265	0.186	66	0.894	0.280	0.207	16
QM B3LYP 6-31G*/AIM/PUBCHEM NO-OPT R^2=0.914								
1	0.910	0.264	0.196	65	0.945	0.235	0.196	17
2	0.910	0.245	0.187	65	0.453	1.613	0.604	17
3	0.924	0.255	0.183	66	0.891	0.270	0.208	16
4	0.959	0.188	0.144	66	0.366	0.659	0.375	16
5	0.915	0.269	0.202	66	0.964	0.214	0.180	16
QM B3LYP 6-31G*/AIM/PUBCHEM QM-OPT R^2=0.885								
1	0.882	0.302	0.228	65	0.940	0.242	0.198	17
2	0.887	0.275	0.213	65	0.896	0.414	0.296	17
3	0.901	0.292	0.221	66	0.910	0.318	0.253	16
4	0.931	0.245	0.189	66	0.691	0.490	0.336	16
5	0.905	0.283	0.208	66	0.843	0.362	0.267	16
QM B3LYP 6-31G*/AIM/RDKIT MM-OPT R^2=0.798								
1	0.826	0.366	0.278	65	0.864	0.380	0.255	17
2	0.847	0.320	0.241	65	0.194	5.366	1.547	17
3	0.875	0.328	0.233	66	0.627	0.575	0.427	16
4	0.873	0.331	0.265	66	0.576	0.577	0.374	16
5	0.842	0.366	0.284	66	0.727	0.521	0.353	16
QM B3LYP 6-31G*/AIM/RDKIT QM-OPT R^2=0.852								
1	0.728	0.458	0.361	65	0.435	0.726	0.538	17
2	0.741	0.416	0.308	65	0.145	1.250	0.681	17
3	0.741	0.472	0.349	66	0.044	9.966	2.868	16
4	0.677	0.528	0.434	66	0.682	0.387	0.329	16
5	0.720	0.487	0.406	66	0.508	0.700	0.514	16
QM B3LYP 6-31G*/AIM/RDKIT MM-OPT R^2=0.839								
1	0.840	0.351	0.276	65	0.828	0.492	0.316	17
2	0.809	0.358	0.281	65	0.486	0.922	0.432	17
3	0.857	0.350	0.267	66	0.660	0.499	0.382	16
4	0.855	0.354	0.263	66	0.710	0.461	0.389	16
5	0.818	0.393	0.305	66	0.872	0.293	0.247	16
QM B3LYP 6-31G*/AIM/RDKIT NO-OPT R^2=0.778								
1	0.842	0.349	0.280	65	0.877	0.350	0.296	17
2	0.855	0.312	0.247	65	0.665	0.703	0.415	17
3	0.859	0.349	0.282	66	0.756	0.371	0.293	16
4	0.882	0.319	0.254	66	0.587	0.464	0.360	16
5	0.857	0.347	0.273	66	0.816	0.364	0.267	16
QM B3LYP 6-31G*/MPA/BALLOON MM-OPT R^2=0.482								
1	0.543	0.594	0.478	65	0.715	0.551	0.411	17
2	0.486	0.587	0.489	65	0.212	1.092	0.621	17
3	0.725	0.486	0.390	66	0.014	0.930	0.685	16
4	0.620	0.573	0.434	66	0.597	0.663	0.571	16
5	0.620	0.567	0.438	66	0.303	0.664	0.538	16
QM B3LYP 6-31G*/MPA/BALLOON NO-OPT R^2=0.589								
1	0.673	0.502	0.394	65	0.872	0.368	0.314	17
2	0.647	0.486	0.393	65	0.490	0.841	0.507	17
3	0.783	0.432	0.354	66	0.307	0.664	0.515	16
4	0.755	0.460	0.361	66	0.498	0.571	0.450	16
5	0.750	0.459	0.352	66	0.530	0.549	0.481	16
QM B3LYP 6-31G*/MPA/BALLOON QM-OPT R^2=0.521								
1	0.614	0.545	0.438	65	0.612	0.619	0.539	17
2	0.561	0.542	0.444	65	0.255	1.038	0.689	17

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
3	0.671	0.532	0.433	66	0.323	0.684	0.474	16
4	0.659	0.543	0.437	66	0.560	0.697	0.570	16
5	0.645	0.548	0.438	66	0.387	0.623	0.509	16
QM B3LYP 6-31G*/MPA/FROG2 MM-OPT R^2=0.792								
1	0.809	0.383	0.273	65	0.881	0.357	0.243	17
2	0.797	0.369	0.260	65	0.893	0.419	0.317	17
3	0.833	0.379	0.267	66	0.774	0.372	0.262	16
4	0.939	0.230	0.166	66	0.339	0.927	0.494	16
5	0.829	0.380	0.264	66	0.823	0.349	0.271	16
QM B3LYP 6-31G*/MPA/FROG2 NO-OPT R^2=0.814								
1	0.871	0.316	0.232	65	0.878	0.339	0.255	17
2	0.840	0.328	0.245	65	0.946	0.311	0.233	17
3	0.885	0.315	0.235	66	0.812	0.342	0.248	16
4	0.920	0.263	0.203	66	0.609	0.568	0.380	16
5	0.882	0.316	0.239	66	0.821	0.353	0.256	16
QM B3LYP 6-31G*/MPA/FROG2 QM-OPT R^2=0.827								
1	0.829	0.364	0.265	65	0.930	0.276	0.234	17
2	0.841	0.326	0.240	65	0.865	0.492	0.363	17
3	0.847	0.362	0.265	66	0.880	0.285	0.207	16
4	0.901	0.293	0.212	66	0.609	0.614	0.418	16
5	0.875	0.326	0.238	66	0.755	0.438	0.307	16
QM B3LYP 6-31G*/MPA/NCI MM-OPT R^2=0.851								
1	0.872	0.314	0.232	65	0.950	0.257	0.211	17
2	0.886	0.276	0.195	65	0.571	1.377	0.521	17
3	0.913	0.274	0.208	66	0.710	0.414	0.300	16
4	0.909	0.280	0.206	66	0.749	0.408	0.295	16
5	0.889	0.306	0.224	66	0.882	0.288	0.197	16
QM B3LYP 6-31G*/MPA/NCI NO-OPT R^2=0.870								
1	0.877	0.308	0.224	65	0.927	0.279	0.169	17
2	0.912	0.243	0.171	65	0.564	1.125	0.537	17
3	0.900	0.293	0.190	66	0.871	0.332	0.266	16
4	0.919	0.265	0.194	66	0.758	0.444	0.281	16
5	0.881	0.318	0.226	66	0.932	0.219	0.180	16
QM B3LYP 6-31G*/MPA/NCI QM-OPT R^2=0.819								
1	0.829	0.363	0.261	65	0.941	0.280	0.236	17
2	0.822	0.345	0.263	65	0.755	0.623	0.355	17
3	0.889	0.309	0.244	66	0.355	0.692	0.468	16
4	0.891	0.307	0.218	66	0.697	0.510	0.350	16
5	0.877	0.322	0.222	66	0.729	0.445	0.314	16
QM B3LYP 6-31G*/MPA/OPENBABEL MM-OPT R^2=0.638								
1	0.608	0.550	0.418	65	0.876	0.436	0.312	17
2	0.702	0.447	0.350	65	0.013	1.568	0.878	17
3	0.767	0.447	0.349	66	0.225	0.789	0.550	16
4	0.713	0.498	0.362	66	0.697	0.704	0.514	16
5	0.671	0.528	0.390	66	0.571	0.531	0.365	16
QM B3LYP 6-31G*/MPA/OPENBABEL NO-OPT R^2=0.634								
1	0.603	0.553	0.414	65	0.862	0.451	0.334	17
2	0.692	0.454	0.356	65	0.006	1.631	0.910	17
3	0.754	0.460	0.347	66	0.281	0.774	0.537	16
4	0.693	0.515	0.364	66	0.689	0.626	0.463	16
5	0.672	0.527	0.389	66	0.527	0.556	0.368	16
QM B3LYP 6-31G*/MPA/OPENBABEL QM-OPT R^2=0.607								
1	0.622	0.540	0.417	65	0.882	0.439	0.357	17
2	0.685	0.459	0.370	65	0.028	1.444	0.866	17
3	0.749	0.464	0.355	66	0.322	0.712	0.523	16
4	0.718	0.494	0.357	66	0.662	0.687	0.501	16
5	0.693	0.509	0.390	66	0.574	0.559	0.388	16
QM B3LYP 6-31G*/MPA/PUBCHEM MM-OPT R^2=0.846								
1	0.872	0.314	0.236	65	0.924	0.287	0.222	17

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
2	0.864	0.302	0.218	65	0.909	0.351	0.252	17
3	0.882	0.318	0.236	66	0.885	0.252	0.188	16
4	0.925	0.255	0.198	66	0.634	0.506	0.350	16
5	0.890	0.305	0.220	66	0.866	0.328	0.263	16
QM B3LYP 6-31G*/MPA/PUBCHEM NO-OPT R^2=0.831								
1	0.870	0.317	0.241	65	0.868	0.358	0.269	17
2	0.848	0.319	0.249	65	0.884	0.439	0.258	17
3	0.893	0.303	0.221	66	0.751	0.432	0.319	16
4	0.907	0.284	0.218	66	0.699	0.486	0.339	16
5	0.871	0.330	0.232	66	0.877	0.327	0.282	16
QM B3LYP 6-31G*/MPA/PUBCHEM QM-OPT R^2=0.859								
1	0.854	0.335	0.238	65	0.916	0.291	0.251	17
2	0.866	0.299	0.226	65	0.886	0.478	0.332	17
3	0.871	0.333	0.245	66	0.850	0.298	0.199	16
4	0.916	0.269	0.209	66	0.686	0.539	0.337	16
5	0.868	0.334	0.240	66	0.901	0.297	0.234	16
QM B3LYP 6-31G*/MPA/RDKIT MM-OPT R^2=0.756								
1	0.790	0.403	0.309	65	0.927	0.284	0.232	17
2	0.812	0.354	0.285	65	0.742	0.618	0.486	17
3	0.873	0.330	0.237	66	0.499	0.616	0.483	16
4	0.865	0.342	0.271	66	0.570	0.602	0.409	16
5	0.812	0.398	0.308	66	0.868	0.360	0.299	16
QM B3LYP 6-31G*/MPA/RDKIT QM-OPT R^2=0.619								
1	0.650	0.519	0.389	65	0.876	0.424	0.331	17
2	0.648	0.485	0.353	65	0.387	0.926	0.510	17
3	0.825	0.388	0.283	66	0.131	0.877	0.635	16
4	0.698	0.511	0.371	66	0.658	0.449	0.375	16
5	0.706	0.498	0.377	66	0.633	0.523	0.339	16
QM B3LYP 6-31G*/MPA/RDKIT MM-OPT R^2=0.679								
1	0.742	0.446	0.347	65	0.782	0.484	0.386	17
2	0.732	0.424	0.332	65	0.285	1.103	0.586	17
3	0.779	0.436	0.363	66	0.596	0.516	0.332	16
4	0.797	0.419	0.313	66	0.420	0.590	0.526	16
5	0.746	0.463	0.357	66	0.753	0.394	0.327	16
QM B3LYP 6-31G*/MPA/RDKIT NO-OPT R^2=0.624								
1	0.712	0.471	0.369	65	0.851	0.503	0.436	17
2	0.666	0.473	0.377	65	0.694	0.661	0.447	17
3	0.717	0.493	0.385	66	0.849	0.415	0.338	16
4	0.808	0.407	0.341	66	0.342	0.736	0.563	16
5	0.747	0.463	0.362	66	0.490	0.568	0.446	16
QM B3LYP 6-31G*/NPA/BALLOON MM-OPT R^2=0.693								
1	0.776	0.415	0.314	65	0.886	0.336	0.278	17
2	0.734	0.422	0.312	65	0.894	0.386	0.279	17
3	0.851	0.358	0.263	66	0.475	0.552	0.426	16
4	0.847	0.363	0.286	66	0.604	0.533	0.386	16
5	0.810	0.401	0.305	66	0.790	0.400	0.294	16
QM B3LYP 6-31G*/NPA/BALLOON NO-OPT R^2=0.717								
1	0.803	0.390	0.291	65	0.902	0.402	0.307	17
2	0.752	0.407	0.300	65	0.946	0.281	0.224	17
3	0.857	0.351	0.258	66	0.610	0.509	0.377	16
4	0.852	0.358	0.264	66	0.700	0.501	0.327	16
5	0.840	0.368	0.267	66	0.695	0.455	0.337	16
QM B3LYP 6-31G*/NPA/BALLOON QM-OPT R^2=0.733								
1	0.796	0.397	0.296	65	0.913	0.306	0.250	17
2	0.769	0.394	0.294	65	0.714	0.623	0.358	17
3	0.877	0.325	0.246	66	0.456	0.540	0.436	16
4	0.844	0.367	0.273	66	0.732	0.462	0.281	16
5	0.830	0.379	0.284	66	0.779	0.379	0.273	16
QM B3LYP 6-31G*/NPA/FROG2 MM-OPT R^2=0.881								

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
1	0.887	0.296	0.220	65	0.949	0.228	0.151	17
2	0.894	0.266	0.185	65	0.900	0.375	0.237	17
3	0.907	0.283	0.213	66	0.851	0.282	0.214	16
4	0.955	0.198	0.151	66	0.473	0.738	0.439	16
5	0.902	0.288	0.207	66	0.878	0.298	0.213	16
QM B3LYP 6-31G*/NPA/FROG2 NO-OPT R^2=0.819								
1	0.880	0.305	0.233	65	0.851	0.381	0.298	17
2	0.860	0.306	0.231	65	0.910	0.382	0.280	17
3	0.881	0.319	0.243	66	0.853	0.321	0.262	16
4	0.892	0.306	0.234	66	0.739	0.408	0.289	16
5	0.882	0.316	0.243	66	0.817	0.341	0.258	16
QM B3LYP 6-31G*/NPA/FROG2 QM-OPT R^2=0.886								
1	0.888	0.294	0.217	65	0.961	0.200	0.178	17
2	0.905	0.252	0.195	65	0.904	0.426	0.296	17
3	0.902	0.290	0.214	66	0.919	0.228	0.194	16
4	0.949	0.209	0.161	66	0.697	0.507	0.372	16
5	0.900	0.291	0.211	66	0.919	0.236	0.176	16
QM B3LYP 6-31G*/NPA/NCI MM-OPT R^2=0.911								
1	0.916	0.254	0.186	65	0.967	0.180	0.131	17
2	0.934	0.210	0.152	65	0.704	1.146	0.490	17
3	0.940	0.228	0.162	66	0.846	0.318	0.273	16
4	0.950	0.209	0.156	66	0.795	0.364	0.255	16
5	0.925	0.251	0.178	66	0.935	0.209	0.148	16
QM B3LYP 6-31G*/NPA/NCI NO-OPT R^2=0.857								
1	0.895	0.284	0.205	65	0.958	0.199	0.127	17
2	0.890	0.272	0.195	65	0.932	0.302	0.212	17
3	0.930	0.245	0.172	66	0.874	0.379	0.324	16
4	0.941	0.226	0.165	66	0.785	0.455	0.302	16
5	0.904	0.285	0.204	66	0.933	0.206	0.159	16
QM B3LYP 6-31G*/NPA/NCI QM-OPT R^2=0.852								
1	0.888	0.294	0.212	65	0.945	0.228	0.174	17
2	0.879	0.284	0.212	65	0.855	0.488	0.280	17
3	0.909	0.280	0.207	66	0.765	0.352	0.274	16
4	0.947	0.213	0.166	66	0.713	0.502	0.337	16
5	0.900	0.291	0.203	66	0.916	0.236	0.183	16
QM B3LYP 6-31G*/NPA/OPENBABEL MM-OPT R^2=0.771								
1	0.811	0.382	0.296	65	0.939	0.258	0.194	17
2	0.819	0.348	0.273	65	0.783	0.631	0.444	17
3	0.884	0.316	0.250	66	0.609	0.524	0.444	16
4	0.865	0.342	0.262	66	0.826	0.450	0.346	16
5	0.849	0.358	0.278	66	0.820	0.382	0.262	16
QM B3LYP 6-31G*/NPA/OPENBABEL NO-OPT R^2=0.771								
1	0.818	0.375	0.290	65	0.928	0.264	0.197	17
2	0.825	0.343	0.268	65	0.785	0.606	0.426	17
3	0.889	0.309	0.246	66	0.609	0.526	0.450	16
4	0.863	0.344	0.262	66	0.850	0.421	0.312	16
5	0.852	0.354	0.271	66	0.816	0.375	0.267	16
QM B3LYP 6-31G*/NPA/OPENBABEL QM-OPT R^2=0.730								
1	0.765	0.426	0.334	65	0.906	0.311	0.267	17
2	0.782	0.382	0.306	65	0.543	0.811	0.518	17
3	0.843	0.368	0.272	66	0.623	0.532	0.472	16
4	0.849	0.361	0.290	66	0.780	0.590	0.457	16
5	0.821	0.389	0.314	66	0.703	0.491	0.317	16
QM B3LYP 6-31G*/NPA/PUBCHEM MM-OPT R^2=0.849								
1	0.874	0.312	0.225	65	0.951	0.231	0.186	17
2	0.878	0.286	0.212	65	0.907	0.385	0.261	17
3	0.896	0.299	0.210	66	0.856	0.296	0.222	16
4	0.936	0.235	0.176	66	0.643	0.499	0.336	16
5	0.888	0.308	0.220	66	0.898	0.278	0.214	16

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
QM B3LYP 6-31G*/NPA/PUBCHEM NO-OPT R^2=0.872								
1	0.876	0.309	0.220	65	0.928	0.263	0.198	17
2	0.887	0.275	0.197	65	0.863	0.450	0.276	17
3	0.906	0.284	0.206	66	0.654	0.505	0.356	16
4	0.927	0.251	0.174	66	0.745	0.473	0.316	16
5	0.889	0.306	0.220	66	0.887	0.315	0.222	16
QM B3LYP 6-31G*/NPA/PUBCHEM QM-OPT R^2=0.864								
1	0.875	0.311	0.217	65	0.963	0.206	0.188	17
2	0.886	0.276	0.203	65	0.912	0.384	0.272	17
3	0.898	0.297	0.209	66	0.845	0.327	0.232	16
4	0.943	0.221	0.169	66	0.701	0.515	0.328	16
5	0.891	0.304	0.213	66	0.915	0.240	0.181	16
QM B3LYP 6-31G*/NPA/RDKIT MM-OPT R^2=0.768								
1	0.810	0.382	0.277	65	0.940	0.238	0.212	17
2	0.798	0.368	0.279	65	0.902	0.370	0.277	17
3	0.867	0.338	0.248	66	0.655	0.476	0.357	16
4	0.906	0.284	0.213	66	0.589	0.650	0.378	16
5	0.877	0.323	0.240	66	0.352	0.743	0.459	16
QM B3LYP 6-31G*/NPA/RDKIT QM-OPT R^2=0.745								
1	0.772	0.419	0.315	65	0.966	0.253	0.198	17
2	0.783	0.382	0.282	65	0.778	0.580	0.368	17
3	0.873	0.330	0.226	66	0.605	0.605	0.512	16
4	0.843	0.368	0.284	66	0.626	0.522	0.360	16
5	0.829	0.380	0.285	66	0.565	0.565	0.258	16
QM B3LYP 6-31G*/NPA/RDKIT MM-OPT R^2=0.677								
1	0.763	0.428	0.329	65	0.916	0.334	0.301	17
2	0.746	0.412	0.304	65	0.689	0.670	0.414	17
3	0.816	0.398	0.289	66	0.792	0.495	0.411	16
4	0.879	0.324	0.263	66	0.461	0.669	0.522	16
5	0.798	0.413	0.319	66	0.779	0.429	0.340	16
QM B3LYP 6-31G*/NPA/RDKIT NO-OPT R^2=0.728								
1	0.797	0.396	0.283	65	0.857	0.474	0.407	17
2	0.781	0.383	0.285	65	0.551	0.821	0.543	17
3	0.789	0.426	0.317	66	0.880	0.306	0.245	16
4	0.888	0.311	0.258	66	0.421	0.678	0.470	16
5	0.816	0.395	0.288	66	0.750	0.443	0.354	16
QM HF STO-3G/MPA/BALLOON MM-OPT R^2=0.809								
1	0.839	0.353	0.244	65	0.969	0.223	0.181	17
2	0.835	0.332	0.243	65	0.307	1.337	0.481	17
3	0.921	0.261	0.192	66	0.426	0.560	0.392	16
4	0.890	0.308	0.217	66	0.691	0.436	0.346	16
5	0.870	0.331	0.234	66	0.852	0.318	0.205	16
QM HF STO-3G/MPA/BALLOON NO-OPT R^2=0.782								
1	0.805	0.388	0.285	65	0.904	0.376	0.312	17
2	0.813	0.354	0.265	65	0.048	2.500	0.807	17
3	0.872	0.332	0.266	66	0.515	0.557	0.402	16
4	0.848	0.363	0.267	66	0.655	0.469	0.391	16
5	0.839	0.369	0.290	66	0.708	0.431	0.326	16
QM HF STO-3G/MPA/BALLOON QM-OPT R^2=0.842								
1	0.868	0.319	0.252	65	0.962	0.210	0.186	17
2	0.861	0.305	0.239	65	0.562	0.820	0.387	17
3	0.918	0.266	0.202	66	0.702	0.431	0.313	16
4	0.901	0.293	0.229	66	0.853	0.330	0.265	16
5	0.914	0.269	0.206	66	0.777	0.455	0.341	16
QM HF STO-3G/MPA/FROG2 MM-OPT R^2=0.896								
1	0.889	0.292	0.216	65	0.962	0.228	0.144	17
2	0.908	0.248	0.176	65	0.887	0.487	0.330	17
3	0.919	0.264	0.190	66	0.864	0.368	0.284	16
4	0.944	0.221	0.170	66	0.545	0.671	0.378	16

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
5	0.913	0.271	0.198	66	0.871	0.315	0.221	16
QM HF STO-3G/MPA/FROG2 NO-OPT R^2=0.865								
1	0.877	0.308	0.218	65	0.953	0.211	0.158	17
2	0.893	0.268	0.194	65	0.865	0.445	0.271	17
3	0.912	0.274	0.189	66	0.815	0.380	0.301	16
4	0.916	0.270	0.202	66	0.727	0.443	0.275	16
5	0.903	0.286	0.203	66	0.837	0.328	0.227	16
QM HF STO-3G/MPA/FROG2 QM-OPT R^2=0.921								
1	0.897	0.282	0.205	65	0.978	0.145	0.127	17
2	0.948	0.187	0.140	65	0.746	0.597	0.385	17
3	0.919	0.263	0.194	66	0.911	0.267	0.196	16
4	0.948	0.212	0.159	66	0.754	0.454	0.321	16
5	0.910	0.276	0.195	66	0.951	0.186	0.170	16
QM HF STO-3G/MPA/NCI MM-OPT R^2=0.919								
1	0.924	0.242	0.180	65	0.981	0.145	0.126	17
2	0.938	0.203	0.151	65	0.438	2.635	0.865	17
3	0.953	0.202	0.149	66	0.801	0.358	0.288	16
4	0.946	0.217	0.155	66	0.894	0.272	0.212	16
5	0.948	0.209	0.160	66	0.876	0.286	0.199	16
QM HF STO-3G/MPA/NCI NO-OPT R^2=0.895								
1	0.917	0.254	0.184	65	0.979	0.159	0.124	17
2	0.924	0.225	0.158	65	0.767	0.567	0.323	17
3	0.943	0.222	0.158	66	0.913	0.314	0.261	16
4	0.948	0.212	0.155	66	0.876	0.338	0.240	16
5	0.931	0.242	0.174	66	0.916	0.235	0.165	16
QM HF STO-3G/MPA/NCI QM-OPT R^2=0.889								
1	0.917	0.253	0.187	65	0.938	0.242	0.159	17
2	0.902	0.256	0.191	65	0.963	0.251	0.204	17
3	0.932	0.242	0.179	66	0.838	0.306	0.252	16
4	0.947	0.214	0.160	66	0.833	0.376	0.274	16
5	0.933	0.238	0.170	66	0.869	0.299	0.214	16
QM HF STO-3G/MPA/OPENBABEL MM-OPT R^2=0.833								
1	0.832	0.360	0.270	65	0.968	0.199	0.169	17
2	0.865	0.300	0.218	65	0.063	1.879	0.723	17
3	0.903	0.289	0.211	66	0.608	0.479	0.386	16
4	0.881	0.320	0.233	66	0.706	0.401	0.315	16
5	0.870	0.332	0.260	66	0.828	0.348	0.219	16
QM HF STO-3G/MPA/OPENBABEL NO-OPT R^2=0.832								
1	0.839	0.353	0.262	65	0.969	0.190	0.161	17
2	0.867	0.298	0.215	65	0.147	1.579	0.631	17
3	0.911	0.276	0.199	66	0.614	0.487	0.382	16
4	0.883	0.318	0.233	66	0.772	0.372	0.274	16
5	0.875	0.325	0.254	66	0.836	0.336	0.206	16
QM HF STO-3G/MPA/OPENBABEL QM-OPT R^2=0.835								
1	0.815	0.378	0.290	65	0.950	0.255	0.225	17
2	0.865	0.301	0.230	65	0.006	2.495	0.866	17
3	0.866	0.339	0.249	66	0.739	0.440	0.380	16
4	0.860	0.348	0.262	66	0.768	0.388	0.330	16
5	0.880	0.318	0.261	66	0.659	0.481	0.340	16
QM HF STO-3G/MPA/PUBCHEM MM-OPT R^2=0.894								
1	0.904	0.273	0.203	65	0.955	0.238	0.160	17
2	0.911	0.244	0.164	65	0.917	0.434	0.282	17
3	0.923	0.257	0.177	66	0.906	0.314	0.244	16
4	0.943	0.223	0.169	66	0.743	0.426	0.271	16
5	0.920	0.260	0.179	66	0.891	0.296	0.193	16
QM HF STO-3G/MPA/PUBCHEM NO-OPT R^2=0.925								
1	0.896	0.283	0.195	65	0.958	0.233	0.178	17
2	0.932	0.214	0.165	65	0.865	0.503	0.268	17
3	0.919	0.264	0.188	66	0.819	0.392	0.292	16

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
4	0.925	0.254	0.171	66	0.810	0.370	0.270	16
5	0.901	0.289	0.204	66	0.947	0.185	0.139	16
QM HF STO-3G/MPA/PUBCHEM QM-OPT R^2=0.888								
1	0.879	0.306	0.210	65	0.970	0.172	0.151	17
2	0.900	0.259	0.189	65	0.880	0.452	0.299	17
3	0.905	0.286	0.195	66	0.876	0.327	0.226	16
4	0.940	0.228	0.166	66	0.749	0.482	0.303	16
5	0.902	0.287	0.196	66	0.874	0.286	0.216	16
QM HF STO-3G/MPA/RDKIT MM-OPT R^2=0.858								
1	0.877	0.309	0.229	65	0.947	0.225	0.165	17
2	0.883	0.280	0.210	65	0.036	2.296	0.762	17
3	0.918	0.265	0.188	66	0.790	0.441	0.346	16
4	0.929	0.247	0.193	66	0.698	0.470	0.316	16
5	0.902	0.288	0.211	66	0.727	0.424	0.269	16
QM HF STO-3G/MPA/RDKIT QM-OPT R^2=0.845								
1	0.830	0.363	0.275	65	0.971	0.227	0.199	17
2	0.871	0.294	0.209	65	0.025	2.314	0.786	17
3	0.906	0.285	0.200	66	0.716	0.551	0.450	16
4	0.863	0.344	0.258	66	0.816	0.318	0.242	16
5	0.902	0.288	0.212	66	0.704	0.581	0.389	16
QM HF STO-3G/MPA/RDKIT MM-OPT R^2=0.865								
1	0.890	0.291	0.218	65	0.956	0.279	0.226	17
2	0.878	0.286	0.202	65	0.477	0.986	0.415	17
3	0.922	0.260	0.204	66	0.717	0.476	0.342	16
4	0.940	0.227	0.174	66	0.645	0.504	0.397	16
5	0.910	0.276	0.219	66	0.781	0.628	0.354	16
QM HF STO-3G/MPA/RDKIT NO-OPT R^2=0.812								
1	0.774	0.417	0.317	65	0.834	0.478	0.434	17
2	0.863	0.303	0.233	65	0.056	2.050	0.834	17
3	0.780	0.435	0.344	66	0.724	0.394	0.260	16
4	0.807	0.409	0.320	66	0.531	0.488	0.387	16
5	0.801	0.410	0.323	66	0.643	0.476	0.365	16
Svob2007-chal2 HF STO-3G/MPA/BALLOON MM-OPT R^2=0.711								
1	0.784	0.409	0.339	65	0.820	0.408	0.352	17
2	0.736	0.420	0.360	65	0.927	0.333	0.260	17
3	0.846	0.364	0.291	66	0.495	0.553	0.488	16
4	0.807	0.408	0.335	66	0.718	0.396	0.378	16
5	0.831	0.378	0.320	66	0.656	0.503	0.412	16
Svob2007-chal2 HF STO-3G/MPA/BALLOON NO-OPT R^2=0.762								
1	0.819	0.374	0.300	65	0.839	0.387	0.312	17
2	0.795	0.370	0.296	65	0.888	0.390	0.296	17
3	0.856	0.351	0.275	66	0.662	0.454	0.400	16
4	0.830	0.383	0.304	66	0.805	0.310	0.279	16
5	0.867	0.335	0.276	66	0.671	0.520	0.375	16
Svob2007-chal2 HF STO-3G/MPA/BALLOON QM-OPT R^2=0.679								
1	0.738	0.450	0.365	65	0.848	0.375	0.321	17
2	0.729	0.426	0.347	65	0.843	0.479	0.362	17
3	0.825	0.388	0.310	66	0.508	0.601	0.506	16
4	0.771	0.444	0.357	66	0.742	0.379	0.351	16
5	0.809	0.402	0.332	66	0.614	0.561	0.434	16
Svob2007-chal2 HF STO-3G/MPA/FROG2 MM-OPT R^2=0.839								
1	0.837	0.354	0.284	65	0.925	0.278	0.237	17
2	0.855	0.311	0.252	65	0.852	0.444	0.326	17
3	0.881	0.320	0.255	66	0.758	0.420	0.341	16
4	0.884	0.316	0.245	66	0.425	0.702	0.419	16
5	0.861	0.343	0.290	66	0.861	0.352	0.280	16
Svob2007-chal2 HF STO-3G/MPA/FROG2 NO-OPT R^2=0.799								
1	0.803	0.389	0.322	65	0.930	0.290	0.202	17
2	0.820	0.347	0.289	65	0.844	0.462	0.369	17

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
3	0.844	0.367	0.293	66	0.737	0.406	0.355	16
4	0.853	0.356	0.289	66	0.548	0.597	0.455	16
5	0.864	0.339	0.273	66	0.699	0.514	0.422	16
Svob2007-chal2 HF STO-3G/MPA/FROG2 QM-OPT R^2=0.819								
1	0.813	0.380	0.298	65	0.928	0.278	0.219	17
2	0.844	0.323	0.266	65	0.796	0.525	0.374	17
3	0.858	0.350	0.273	66	0.711	0.416	0.352	16
4	0.869	0.336	0.258	66	0.441	0.729	0.433	16
5	0.857	0.347	0.276	66	0.763	0.439	0.373	16
Svob2007-chal2 HF STO-3G/MPA/NCI MM-OPT R^2=0.829								
1	0.839	0.353	0.281	65	0.928	0.288	0.216	17
2	0.855	0.312	0.232	65	0.809	0.537	0.375	17
3	0.893	0.303	0.238	66	0.647	0.481	0.389	16
4	0.881	0.320	0.246	66	0.715	0.415	0.339	16
5	0.853	0.352	0.286	66	0.856	0.302	0.226	16
Svob2007-chal2 HF STO-3G/MPA/NCI NO-OPT R^2=0.846								
1	0.851	0.340	0.265	65	0.950	0.222	0.187	17
2	0.875	0.290	0.222	65	0.645	0.737	0.404	17
3	0.886	0.313	0.244	66	0.797	0.355	0.272	16
4	0.913	0.274	0.208	66	0.663	0.476	0.391	16
5	0.879	0.320	0.250	66	0.847	0.323	0.238	16
Svob2007-chal2 HF STO-3G/MPA/NCI QM-OPT R^2=0.787								
1	0.782	0.410	0.327	65	0.942	0.247	0.185	17
2	0.826	0.341	0.258	65	0.737	0.650	0.419	17
3	0.863	0.343	0.243	66	0.592	0.545	0.410	16
4	0.836	0.377	0.302	66	0.687	0.434	0.358	16
5	0.843	0.364	0.296	66	0.615	0.513	0.344	16
Svob2007-chal2 HF STO-3G/MPA/OPENBABEL MM-OPT R^2=0.765								
1	0.816	0.377	0.317	65	0.848	0.426	0.312	17
2	0.791	0.375	0.322	65	0.899	0.388	0.325	17
3	0.857	0.351	0.295	66	0.674	0.482	0.413	16
4	0.839	0.373	0.320	66	0.761	0.386	0.339	16
5	0.845	0.362	0.306	66	0.738	0.453	0.373	16
Svob2007-chal2 HF STO-3G/MPA/OPENBABEL NO-OPT R^2=0.773								
1	0.819	0.374	0.317	65	0.851	0.421	0.306	17
2	0.793	0.372	0.318	65	0.904	0.383	0.317	17
3	0.860	0.347	0.292	66	0.669	0.477	0.407	16
4	0.840	0.372	0.319	66	0.766	0.374	0.332	16
5	0.847	0.359	0.305	66	0.743	0.449	0.370	16
Svob2007-chal2 HF STO-3G/MPA/OPENBABEL QM-OPT R^2=0.724								
1	0.781	0.411	0.333	65	0.903	0.395	0.277	17
2	0.757	0.404	0.327	65	0.897	0.376	0.311	17
3	0.843	0.367	0.297	66	0.629	0.520	0.450	16
4	0.815	0.400	0.332	66	0.768	0.390	0.328	16
5	0.851	0.355	0.286	66	0.612	0.593	0.425	16
Svob2007-chal2 HF STO-3G/MPA/PUBCHEM MM-OPT R^2=0.838								
1	0.850	0.340	0.272	65	0.959	0.235	0.182	17
2	0.862	0.304	0.236	65	0.895	0.388	0.274	17
3	0.901	0.292	0.228	66	0.707	0.468	0.398	16
4	0.888	0.311	0.235	66	0.775	0.359	0.304	16
5	0.874	0.327	0.261	66	0.871	0.294	0.226	16
Svob2007-chal2 HF STO-3G/MPA/PUBCHEM NO-OPT R^2=0.884								
1	0.866	0.321	0.258	65	0.965	0.221	0.183	17
2	0.899	0.261	0.203	65	0.855	0.443	0.289	17
3	0.904	0.288	0.220	66	0.830	0.372	0.313	16
4	0.902	0.291	0.223	66	0.829	0.352	0.274	16
5	0.878	0.320	0.255	66	0.935	0.209	0.179	16
Svob2007-chal2 HF STO-3G/MPA/PUBCHEM QM-OPT R^2=0.746								
1	0.796	0.397	0.307	65	0.917	0.297	0.238	17

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
2	0.780	0.384	0.292	65	0.907	0.355	0.288	17
3	0.887	0.312	0.249	66	0.554	0.624	0.492	16
4	0.836	0.376	0.286	66	0.775	0.392	0.350	16
5	0.838	0.370	0.295	66	0.760	0.454	0.321	16
Svob2007-chal2 HF STO-3G/MPA/RDKIT MM-OPT R^2=0.700								
1	0.760	0.430	0.359	65	0.922	0.291	0.226	17
2	0.764	0.398	0.321	65	0.875	0.436	0.372	17
3	0.858	0.350	0.285	66	0.329	0.662	0.563	16
4	0.808	0.408	0.333	66	0.683	0.404	0.326	16
5	0.839	0.369	0.305	66	0.622	0.559	0.424	16
Svob2007-chal2 HF STO-3G/MPA/RDKIT QM-OPT R^2=0.641								
1	0.708	0.475	0.385	65	0.888	0.345	0.278	17
2	0.689	0.457	0.367	65	0.886	0.429	0.373	17
3	0.837	0.374	0.287	66	0.202	0.752	0.635	16
4	0.750	0.465	0.382	66	0.665	0.400	0.337	16
5	0.805	0.406	0.327	66	0.534	0.649	0.473	16
Svob2007-chal2 HF STO-3G/MPA/RDKIT MM-OPT R^2=0.744								
1	0.808	0.385	0.312	65	0.873	0.381	0.319	17
2	0.798	0.368	0.295	65	0.875	0.423	0.346	17
3	0.839	0.372	0.301	66	0.706	0.437	0.341	16
4	0.834	0.379	0.305	66	0.778	0.380	0.312	16
5	0.856	0.349	0.283	66	0.580	0.558	0.404	16
Svob2007-chal2 HF STO-3G/MPA/RDKIT NO-OPT R^2=0.760								
1	0.812	0.381	0.301	65	0.918	0.305	0.240	17
2	0.820	0.347	0.273	65	0.861	0.433	0.319	17
3	0.850	0.359	0.277	66	0.739	0.397	0.342	16
4	0.860	0.348	0.277	66	0.606	0.448	0.324	16
5	0.842	0.365	0.292	66	0.807	0.379	0.299	16

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
Anilines								
Bult2002-npa B3LYP 6-31G*/NPA/BALLOON MM-OPT R^2=0.855								
1	0.866	0.850	0.649	38	0.806	0.914	0.744	10
2	0.886	0.767	0.560	38	0.808	1.263	0.960	10
3	0.862	0.879	0.636	38	0.819	0.742	0.593	10
4	0.785	0.897	0.689	39	0.973	0.678	0.583	9
5	0.885	0.717	0.606	39	0.740	1.313	0.927	9
Bult2002-npa B3LYP 6-31G*/NPA/BALLOON NO-OPT R^2=0.826								
1	0.835	0.944	0.696	38	0.863	0.960	0.828	10
2	0.863	0.842	0.661	38	0.607	1.711	1.319	10
3	0.845	0.932	0.686	38	0.742	0.941	0.841	10
4	0.790	0.887	0.685	39	0.920	1.308	0.926	9
5	0.852	0.813	0.674	39	0.695	1.451	1.019	9
Bult2002-npa B3LYP 6-31G*/NPA/BALLOON QM-OPT R^2=0.908								
1	0.931	0.611	0.470	38	0.645	0.966	0.815	10
2	0.910	0.680	0.554	38	0.703	1.138	0.766	10
3	0.937	0.594	0.455	38	0.617	1.006	0.750	10
4	0.868	0.705	0.590	39	0.996	0.724	0.562	9
5	0.911	0.631	0.494	39	0.873	0.942	0.698	9
Bult2002-npa B3LYP 6-31G*/NPA/FROG2 MM-OPT R^2=0.852								
1	0.880	0.806	0.653	38	0.728	1.156	1.001	10
2	0.885	0.770	0.605	38	0.682	1.533	1.145	10
3	0.870	0.853	0.666	38	0.764	0.927	0.780	10
4	0.778	0.912	0.742	39	0.982	0.727	0.590	9
5	0.875	0.746	0.629	39	0.711	1.383	0.983	9
Bult2002-npa B3LYP 6-31G*/NPA/FROG2 NO-OPT R^2=0.904								
1	0.924	0.642	0.494	38	0.714	1.028	0.849	10
2	0.925	0.623	0.470	38	0.674	1.142	0.912	10
3	0.911	0.705	0.548	38	0.884	0.662	0.566	10
4	0.884	0.659	0.547	39	0.986	0.942	0.552	9
5	0.908	0.641	0.486	39	0.865	0.962	0.748	9
Bult2002-npa B3LYP 6-31G*/NPA/FROG2 QM-OPT R^2=0.907								
1	0.920	0.657	0.530	38	0.799	0.784	0.666	10
2	0.907	0.695	0.578	38	0.814	0.934	0.704	10
3	0.929	0.631	0.510	38	0.789	0.889	0.812	10
4	0.902	0.608	0.519	39	0.893	1.050	0.628	9
5	0.914	0.618	0.493	39	0.774	1.409	0.894	9
Bult2002-npa B3LYP 6-31G*/NPA/NCI MM-OPT R^2=0.848								
1	0.869	0.842	0.645	38	0.747	1.060	0.905	10
2	0.882	0.782	0.589	38	0.665	1.671	1.332	10
3	0.859	0.888	0.656	38	0.755	0.813	0.700	10
4	0.768	0.933	0.732	39	0.979	0.680	0.574	9
5	0.879	0.736	0.628	39	0.715	1.345	0.953	9
Bult2002-npa B3LYP 6-31G*/NPA/NCI NO-OPT R^2=0.839								
1	0.847	0.910	0.702	38	0.845	0.982	0.847	10
2	0.871	0.816	0.634	38	0.620	1.638	1.235	10
3	0.857	0.893	0.662	38	0.731	0.986	0.906	10
4	0.773	0.923	0.735	39	0.961	0.864	0.680	9
5	0.883	0.725	0.598	39	0.635	1.497	1.131	9
Bult2002-npa B3LYP 6-31G*/NPA/NCI QM-OPT R^2=0.869								
1	0.912	0.692	0.545	38	0.684	1.359	1.213	10
2	0.886	0.767	0.623	38	0.792	1.130	0.801	10
3	0.891	0.781	0.647	38	0.769	0.927	0.775	10
4	0.812	0.840	0.710	39	0.964	0.785	0.680	9
5	0.866	0.775	0.641	39	0.807	1.073	0.870	9
Bult2002-npa B3LYP 6-31G*/NPA/OPENBABEL MM-OPT R^2=0.857								
1	0.886	0.785	0.621	38	0.784	1.150	0.964	10
2	0.891	0.751	0.581	38	0.664	1.527	1.110	10
3	0.870	0.851	0.662	38	0.774	0.866	0.741	10

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
4	0.783	0.903	0.735	39	0.980	0.637	0.474	9
5	0.884	0.720	0.594	39	0.702	1.399	1.017	9
Bult2002-npa B3LYP 6-31G*/NPA/OPENBABEL NO-OPT R^2=0.851								
1	0.876	0.820	0.666	38	0.793	1.137	0.994	10
2	0.886	0.768	0.617	38	0.626	1.598	1.149	10
3	0.868	0.860	0.677	38	0.750	0.906	0.793	10
4	0.778	0.912	0.754	39	0.980	0.691	0.527	9
5	0.877	0.741	0.628	39	0.694	1.423	1.052	9
Bult2002-npa B3LYP 6-31G*/NPA/OPENBABEL QM-OPT R^2=0.939								
1	0.960	0.463	0.395	38	0.718	0.829	0.737	10
2	0.935	0.580	0.495	38	0.917	0.496	0.363	10
3	0.955	0.501	0.394	38	0.850	0.748	0.684	10
4	0.915	0.564	0.483	39	0.991	0.565	0.404	9
5	0.932	0.550	0.452	39	0.934	0.698	0.543	9
Bult2002-npa B3LYP 6-31G*/NPA/PUBCHEM MM-OPT R^2=0.872								
1	0.891	0.766	0.600	38	0.568	1.036	0.877	10
2	0.895	0.737	0.579	38	0.787	1.236	0.890	10
3	0.890	0.783	0.603	38	0.739	0.907	0.817	10
4	0.805	0.855	0.694	39	0.987	0.584	0.475	9
5	0.896	0.682	0.590	39	0.763	1.253	0.880	9
Bult2002-npa B3LYP 6-31G*/NPA/PUBCHEM NO-OPT R^2=0.874								
1	0.909	0.700	0.562	38	0.759	1.303	1.099	10
2	0.896	0.733	0.573	38	0.723	1.236	0.897	10
3	0.893	0.774	0.616	38	0.770	0.973	0.883	10
4	0.806	0.853	0.721	39	0.987	0.656	0.489	9
5	0.886	0.713	0.597	39	0.789	1.121	0.978	9
Bult2002-npa B3LYP 6-31G*/NPA/PUBCHEM QM-OPT R^2=0.867								
1	0.892	0.764	0.623	38	0.769	1.089	0.994	10
2	0.885	0.770	0.647	38	0.688	1.338	0.976	10
3	0.901	0.746	0.593	38	0.765	1.094	0.969	10
4	0.840	0.774	0.659	39	0.952	1.212	0.715	9
5	0.861	0.788	0.651	39	0.808	1.092	0.824	9
Bult2002-npa B3LYP 6-31G*/NPA/RDKIT MM-OPT R^2=0.843								
1	0.851	0.897	0.688	38	0.866	0.953	0.850	10
2	0.881	0.785	0.654	38	0.430	1.957	1.405	10
3	0.854	0.905	0.688	38	0.786	0.827	0.718	10
4	0.801	0.864	0.679	39	0.940	1.037	0.821	9
5	0.876	0.745	0.607	39	0.720	1.409	1.125	9
Bult2002-npa B3LYP 6-31G*/NPA/RDKIT QM-OPT R^2=0.869								
1	0.912	0.690	0.542	38	0.616	1.336	1.165	10
2	0.890	0.752	0.610	38	0.753	1.180	0.815	10
3	0.893	0.773	0.632	38	0.804	0.994	0.841	10
4	0.805	0.855	0.739	39	0.981	0.653	0.549	9
5	0.864	0.781	0.642	39	0.779	1.173	0.921	9
Bult2002-npa B3LYP 6-31G*/NPA/RDKIT MM-OPT R^2=0.854								
1	0.881	0.803	0.640	38	0.776	1.138	0.988	10
2	0.885	0.772	0.596	38	0.672	1.496	1.091	10
3	0.873	0.843	0.643	38	0.734	0.934	0.815	10
4	0.777	0.915	0.741	39	0.985	0.672	0.484	9
5	0.874	0.750	0.641	39	0.726	1.310	1.000	9
Bult2002-npa B3LYP 6-31G*/NPA/RDKIT NO-OPT R^2=0.840								
1	0.861	0.867	0.657	38	0.893	1.162	0.984	10
2	0.878	0.794	0.645	38	0.610	1.482	1.149	10
3	0.855	0.900	0.691	38	0.797	0.926	0.777	10
4	0.806	0.853	0.715	39	0.968	1.190	0.789	9
5	0.860	0.792	0.624	39	0.748	1.393	1.044	9
Bult2004-aim B3LYP 6-31G*/AIM/BALLOON MM-OPT R^2=0.880								
1	0.909	0.701	0.537	38	0.761	1.219	1.107	10
2	0.895	0.738	0.577	38	0.617	1.242	0.802	10

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
3	0.889	0.786	0.581	38	0.808	0.716	0.641	10
4	0.819	0.825	0.655	39	0.987	0.626	0.543	9
5	0.923	0.588	0.487	39	0.669	1.501	1.186	9
Bult2004-aim B3LYP 6-31G*/AIM/BALLOON NO-OPT R^2=0.870								
1	0.897	0.747	0.578	38	0.676	1.091	0.931	10
2	0.881	0.785	0.629	38	0.689	1.222	0.947	10
3	0.885	0.804	0.619	38	0.783	0.835	0.708	10
4	0.824	0.812	0.667	39	0.947	0.994	0.677	9
5	0.904	0.653	0.532	39	0.709	1.422	1.122	9
Bult2004-aim B3LYP 6-31G*/AIM/BALLOON QM-OPT R^2=0.945								
1	0.958	0.477	0.344	38	0.808	0.689	0.540	10
2	0.951	0.505	0.365	38	0.771	0.805	0.497	10
3	0.953	0.512	0.396	38	0.842	0.603	0.347	10
4	0.929	0.516	0.389	39	0.967	0.590	0.459	9
5	0.948	0.484	0.391	39	0.905	0.769	0.657	9
Bult2004-aim B3LYP 6-31G*/AIM/FROG2 MM-OPT R^2=0.902								
1	0.924	0.640	0.534	38	0.853	0.944	0.746	10
2	0.913	0.671	0.527	38	0.719	1.029	0.780	10
3	0.915	0.690	0.552	38	0.852	0.779	0.710	10
4	0.873	0.690	0.550	39	0.935	0.919	0.733	9
5	0.919	0.603	0.498	39	0.660	1.497	1.239	9
Bult2004-aim B3LYP 6-31G*/AIM/FROG2 NO-OPT R^2=0.887								
1	0.914	0.681	0.489	38	0.699	0.993	0.828	10
2	0.886	0.767	0.562	38	0.868	0.776	0.663	10
3	0.884	0.806	0.564	38	0.953	0.478	0.409	10
4	0.869	0.701	0.531	39	0.909	1.129	0.785	9
5	0.906	0.647	0.463	39	0.786	1.135	0.887	9
Bult2004-aim B3LYP 6-31G*/AIM/FROG2 QM-OPT R^2=0.939								
1	0.953	0.506	0.418	38	0.855	0.730	0.566	10
2	0.935	0.578	0.481	38	0.921	0.486	0.416	10
3	0.949	0.535	0.438	38	0.871	0.637	0.552	10
4	0.915	0.565	0.471	39	0.978	0.546	0.463	9
5	0.949	0.475	0.404	39	0.872	1.010	0.826	9
Bult2004-aim B3LYP 6-31G*/AIM/NCI MM-OPT R^2=0.889								
1	0.906	0.712	0.544	38	0.847	0.969	0.709	10
2	0.902	0.710	0.534	38	0.696	1.197	0.926	10
3	0.895	0.768	0.564	38	0.852	0.658	0.554	10
4	0.854	0.739	0.514	39	0.930	1.063	0.817	9
5	0.935	0.538	0.443	39	0.648	1.558	1.266	9
Bult2004-aim B3LYP 6-31G*/AIM/NCI NO-OPT R^2=0.882								
1	0.909	0.702	0.563	38	0.726	1.075	0.899	10
2	0.898	0.725	0.567	38	0.715	1.153	0.943	10
3	0.890	0.786	0.606	38	0.837	0.700	0.585	10
4	0.888	0.649	0.534	39	0.417	2.491	1.336	9
5	0.911	0.632	0.528	39	0.711	1.324	1.082	9
Bult2004-aim B3LYP 6-31G*/AIM/NCI QM-OPT R^2=0.920								
1	0.935	0.593	0.470	38	0.855	0.779	0.639	10
2	0.926	0.619	0.478	38	0.742	0.930	0.602	10
3	0.921	0.666	0.530	38	0.907	0.498	0.405	10
4	0.890	0.641	0.504	39	0.966	0.640	0.494	9
5	0.943	0.503	0.411	39	0.775	1.167	0.992	9
Bult2004-aim B3LYP 6-31G*/AIM/OPENBABEL MM-OPT R^2=0.897								
1	0.916	0.674	0.553	38	0.855	0.894	0.704	10
2	0.910	0.680	0.522	38	0.679	1.098	0.761	10
3	0.911	0.706	0.561	38	0.850	0.781	0.734	10
4	0.875	0.686	0.541	39	0.903	1.052	0.836	9
5	0.920	0.598	0.492	39	0.637	1.573	1.261	9
Bult2004-aim B3LYP 6-31G*/AIM/OPENBABEL NO-OPT R^2=0.899								
1	0.917	0.669	0.542	38	0.810	0.857	0.719	10

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
2	0.907	0.692	0.537	38	0.718	0.992	0.713	10
3	0.914	0.693	0.553	38	0.834	0.808	0.733	10
4	0.876	0.683	0.538	39	0.907	0.997	0.755	9
5	0.917	0.609	0.526	39	0.662	1.511	1.223	9
Bult2004-aim B3LYP 6-31G*/AIM/OPENBABEL QM-OPT R^2=0.934								
1	0.948	0.529	0.429	38	0.851	0.745	0.595	10
2	0.931	0.599	0.496	38	0.910	0.522	0.413	10
3	0.944	0.559	0.452	38	0.871	0.644	0.584	10
4	0.913	0.573	0.476	39	0.963	0.614	0.479	9
5	0.948	0.480	0.392	39	0.790	1.290	0.892	9
Bult2004-aim B3LYP 6-31G*/AIM/PUBCHEM MM-OPT R^2=0.900								
1	0.929	0.619	0.521	38	0.855	1.102	0.896	10
2	0.918	0.650	0.516	38	0.678	1.189	0.869	10
3	0.910	0.711	0.561	38	0.854	0.703	0.622	10
4	0.865	0.713	0.558	39	0.950	0.873	0.677	9
5	0.928	0.568	0.454	39	0.602	1.822	1.379	9
Bult2004-aim B3LYP 6-31G*/AIM/PUBCHEM NO-OPT R^2=0.907								
1	0.922	0.651	0.515	38	0.784	0.868	0.691	10
2	0.914	0.665	0.516	38	0.759	0.874	0.618	10
3	0.911	0.706	0.536	38	0.887	0.580	0.518	10
4	0.861	0.723	0.588	39	0.992	0.731	0.501	9
5	0.936	0.535	0.455	39	0.786	1.180	0.963	9
Bult2004-aim B3LYP 6-31G*/AIM/PUBCHEM QM-OPT R^2=0.927								
1	0.950	0.518	0.433	38	0.819	0.899	0.673	10
2	0.926	0.618	0.480	38	0.862	0.654	0.466	10
3	0.936	0.598	0.459	38	0.856	0.664	0.537	10
4	0.897	0.623	0.504	39	0.965	0.655	0.521	9
5	0.931	0.554	0.440	39	0.903	0.824	0.712	9
Bult2004-aim B3LYP 6-31G*/AIM/RDKIT MM-OPT R^2=0.875								
1	0.879	0.808	0.594	38	0.875	0.716	0.620	10
2	0.893	0.745	0.580	38	0.640	1.160	0.816	10
3	0.884	0.804	0.611	38	0.797	0.735	0.572	10
4	0.810	0.844	0.649	39	0.984	0.550	0.409	9
5	0.922	0.590	0.464	39	0.707	1.404	1.126	9
Bult2004-aim B3LYP 6-31G*/AIM/RDKIT QM-OPT R^2=0.923								
1	0.941	0.567	0.457	38	0.796	0.834	0.719	10
2	0.928	0.611	0.469	38	0.854	0.808	0.611	10
3	0.928	0.636	0.503	38	0.931	0.568	0.511	10
4	0.901	0.608	0.504	39	0.959	0.732	0.513	9
5	0.938	0.524	0.423	39	0.789	1.126	0.922	9
Bult2004-aim B3LYP 6-31G*/AIM/RDKIT MM-OPT R^2=0.897								
1	0.916	0.675	0.545	38	0.838	0.908	0.722	10
2	0.910	0.682	0.519	38	0.721	1.107	0.852	10
3	0.908	0.717	0.580	38	0.855	0.760	0.689	10
4	0.880	0.671	0.520	39	0.897	1.168	0.923	9
5	0.922	0.590	0.486	39	0.639	1.579	1.233	9
Bult2004-aim B3LYP 6-31G*/AIM/RDKIT NO-OPT R^2=0.857								
1	0.889	0.776	0.568	38	0.823	1.296	0.971	10
2	0.886	0.767	0.591	38	0.644	1.517	1.144	10
3	0.857	0.896	0.674	38	0.880	0.581	0.447	10
4	0.807	0.851	0.669	39	0.984	0.925	0.555	9
5	0.902	0.662	0.505	39	0.741	1.476	1.115	9
Chavez2006 B3LYP 6-31G*/MPA/BALLOON MM-OPT R^2=0.845								
1	0.860	0.871	0.653	38	0.739	1.059	0.763	10
2	0.867	0.830	0.638	38	0.629	1.427	1.002	10
3	0.855	0.901	0.659	38	0.823	0.829	0.774	10
4	0.791	0.886	0.672	39	0.955	0.944	0.743	9
5	0.883	0.722	0.584	39	0.705	1.441	1.047	9
Chavez2006 B3LYP 6-31G*/MPA/BALLOON NO-OPT R^2=0.847								

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
1	0.868	0.846	0.642	38	0.745	1.164	0.810	10
2	0.871	0.816	0.663	38	0.609	1.580	1.278	10
3	0.862	0.879	0.646	38	0.798	0.884	0.814	10
4	0.809	0.847	0.665	39	0.934	1.165	0.850	9
5	0.882	0.726	0.593	39	0.684	1.497	1.096	9
Chavez2006 B3LYP 6-31G*/MPA/BALLOON QM-OPT R^2=0.930								
1	0.941	0.567	0.428	38	0.836	0.704	0.600	10
2	0.938	0.565	0.441	38	0.790	0.905	0.564	10
3	0.948	0.541	0.446	38	0.745	0.783	0.484	10
4	0.903	0.603	0.462	39	0.966	0.543	0.491	9
5	0.931	0.555	0.441	39	0.893	0.827	0.693	9
Chavez2006 B3LYP 6-31G*/MPA/FROG2 MM-OPT R^2=0.857								
1	0.880	0.804	0.660	38	0.790	1.140	0.911	10
2	0.883	0.778	0.618	38	0.606	1.518	1.083	10
3	0.874	0.839	0.693	38	0.794	0.908	0.861	10
4	0.815	0.833	0.680	39	0.914	1.067	0.859	9
5	0.877	0.740	0.616	39	0.743	1.272	1.081	9
Chavez2006 B3LYP 6-31G*/MPA/FROG2 NO-OPT R^2=0.879								
1	0.907	0.711	0.565	38	0.753	1.114	0.959	10
2	0.891	0.751	0.559	38	0.648	1.100	0.805	10
3	0.884	0.807	0.595	38	0.917	0.692	0.563	10
4	0.851	0.747	0.578	39	0.970	1.006	0.637	9
5	0.894	0.689	0.508	39	0.783	1.200	0.938	9
Chavez2006 B3LYP 6-31G*/MPA/FROG2 QM-OPT R^2=0.926								
1	0.942	0.560	0.469	38	0.817	0.824	0.678	10
2	0.928	0.611	0.531	38	0.893	0.707	0.559	10
3	0.936	0.597	0.519	38	0.845	0.684	0.596	10
4	0.915	0.564	0.480	39	0.928	0.834	0.724	9
5	0.921	0.593	0.524	39	0.909	0.773	0.719	9
Chavez2006 B3LYP 6-31G*/MPA/NCI MM-OPT R^2=0.834								
1	0.853	0.891	0.662	38	0.815	1.141	0.899	10
2	0.862	0.845	0.635	38	0.549	1.640	1.151	10
3	0.851	0.913	0.677	38	0.746	0.932	0.882	10
4	0.777	0.914	0.711	39	0.942	1.034	0.777	9
5	0.876	0.744	0.618	39	0.673	1.503	1.187	9
Chavez2006 B3LYP 6-31G*/MPA/NCI NO-OPT R^2=0.852								
1	0.871	0.836	0.664	38	0.803	1.154	0.918	10
2	0.878	0.794	0.630	38	0.633	1.438	1.092	10
3	0.868	0.859	0.659	38	0.778	0.903	0.855	10
4	0.792	0.884	0.717	39	0.975	0.990	0.679	9
5	0.888	0.707	0.597	39	0.684	1.408	1.113	9
Chavez2006 B3LYP 6-31G*/MPA/NCI QM-OPT R^2=0.884								
1	0.909	0.700	0.563	38	0.771	1.133	0.861	10
2	0.900	0.718	0.592	38	0.699	1.281	0.928	10
3	0.894	0.771	0.641	38	0.873	0.765	0.618	10
4	0.845	0.764	0.616	39	0.940	0.790	0.685	9
5	0.906	0.648	0.519	39	0.722	1.306	1.102	9
Chavez2006 B3LYP 6-31G*/MPA/OPENBABEL MM-OPT R^2=0.851								
1	0.869	0.841	0.664	38	0.849	1.056	0.901	10
2	0.880	0.789	0.633	38	0.584	1.546	1.115	10
3	0.871	0.848	0.679	38	0.770	0.951	0.905	10
4	0.812	0.840	0.680	39	0.847	1.278	0.955	9
5	0.876	0.745	0.605	39	0.725	1.339	1.079	9
Chavez2006 B3LYP 6-31G*/MPA/OPENBABEL NO-OPT R^2=0.849								
1	0.870	0.839	0.659	38	0.836	1.094	0.891	10
2	0.882	0.782	0.612	38	0.574	1.586	1.113	10
3	0.868	0.858	0.673	38	0.763	0.967	0.917	10
4	0.813	0.839	0.678	39	0.775	1.523	1.110	9
5	0.878	0.739	0.590	39	0.708	1.395	1.081	9

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
Chavez2006 B3LYP 6-31G*/MPA/OPENBABEL QM-OPT R^2=0.935								
1	0.948	0.530	0.440	38	0.845	0.739	0.646	10
2	0.940	0.556	0.473	38	0.847	0.756	0.571	10
3	0.944	0.561	0.474	38	0.881	0.645	0.541	10
4	0.919	0.550	0.471	39	0.948	0.708	0.619	9
5	0.933	0.547	0.462	39	0.939	0.717	0.626	9
Chavez2006 B3LYP 6-31G*/MPA/PUBCHEM MM-OPT R^2=0.852								
1	0.880	0.804	0.644	38	0.804	1.336	1.063	10
2	0.881	0.784	0.602	38	0.613	1.499	1.064	10
3	0.870	0.852	0.665	38	0.786	0.925	0.879	10
4	0.797	0.872	0.716	39	0.965	0.943	0.692	9
5	0.877	0.740	0.596	39	0.731	1.348	1.004	9
Chavez2006 B3LYP 6-31G*/MPA/PUBCHEM NO-OPT R^2=0.881								
1	0.906	0.714	0.558	38	0.829	1.114	0.977	10
2	0.900	0.717	0.565	38	0.693	1.266	0.954	10
3	0.898	0.755	0.584	38	0.831	0.969	0.864	10
4	0.829	0.800	0.653	39	0.974	0.749	0.598	9
5	0.904	0.654	0.559	39	0.764	1.191	1.003	9
Chavez2006 B3LYP 6-31G*/MPA/PUBCHEM QM-OPT R^2=0.905								
1	0.933	0.603	0.469	38	0.736	1.098	0.850	10
2	0.907	0.692	0.559	38	0.844	0.867	0.662	10
3	0.933	0.614	0.506	38	0.806	1.062	0.742	10
4	0.865	0.712	0.579	39	0.977	0.648	0.487	9
5	0.900	0.669	0.548	39	0.916	0.827	0.681	9
Chavez2006 B3LYP 6-31G*/MPA/RDKIT MM-OPT R^2=0.845								
1	0.854	0.889	0.686	38	0.866	0.921	0.781	10
2	0.885	0.769	0.622	38	0.485	1.828	1.379	10
3	0.859	0.887	0.669	38	0.778	0.900	0.849	10
4	0.799	0.868	0.679	39	0.933	1.050	0.757	9
5	0.877	0.742	0.587	39	0.707	1.405	1.099	9
Chavez2006 B3LYP 6-31G*/MPA/RDKIT QM-OPT R^2=0.892								
1	0.917	0.668	0.529	38	0.710	1.124	0.791	10
2	0.915	0.663	0.524	38	0.655	1.281	0.943	10
3	0.906	0.725	0.586	38	0.877	0.850	0.666	10
4	0.844	0.766	0.630	39	0.972	0.632	0.478	9
5	0.910	0.635	0.504	39	0.721	1.344	1.113	9
Chavez2006 B3LYP 6-31G*/MPA/RDKIT MM-OPT R^2=0.852								
1	0.868	0.843	0.677	38	0.868	0.984	0.871	10
2	0.882	0.781	0.617	38	0.535	1.641	1.161	10
3	0.875	0.837	0.684	38	0.747	0.994	0.930	10
4	0.818	0.826	0.673	39	0.840	1.276	0.918	9
5	0.877	0.742	0.604	39	0.723	1.353	1.099	9
Chavez2006 B3LYP 6-31G*/MPA/RDKIT NO-OPT R^2=0.842								
1	0.866	0.852	0.616	38	0.880	1.194	0.967	10
2	0.857	0.859	0.640	38	0.722	1.424	1.117	10
3	0.848	0.921	0.682	38	0.865	0.761	0.683	10
4	0.832	0.793	0.659	39	0.953	1.450	0.794	9
5	0.877	0.741	0.569	39	0.729	1.421	1.122	9
QM B3LYP 6-31G*/AIM/BALLOON MM-OPT R^2=0.921								
1	0.928	0.626	0.430	38	0.843	0.696	0.597	10
2	0.957	0.472	0.348	38	0.596	1.305	0.783	10
3	0.926	0.643	0.494	38	0.865	0.653	0.465	10
4	0.900	0.612	0.450	39	0.948	0.847	0.606	9
5	0.916	0.611	0.480	39	0.897	0.784	0.584	9
QM B3LYP 6-31G*/AIM/BALLOON NO-OPT R^2=0.805								
1	0.824	0.976	0.757	38	0.822	0.890	0.634	10
2	0.879	0.792	0.623	38	0.537	1.617	1.366	10
3	0.836	0.959	0.757	38	0.632	1.007	0.758	10
4	0.769	0.931	0.768	39	0.902	1.195	0.929	9

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
5	0.813	0.915	0.735	39	0.839	1.108	0.909	9
QM B3LYP 6-31G*/AIM/BALLOON QM-OPT R^2=0.921								
1	0.924	0.643	0.508	38	0.869	0.660	0.552	10
2	0.974	0.366	0.267	38	0.520	1.611	1.227	10
3	0.930	0.625	0.472	38	0.869	0.663	0.480	10
4	0.913	0.572	0.438	39	0.991	1.600	1.054	9
5	0.904	0.655	0.503	39	0.988	0.512	0.460	9
QM B3LYP 6-31G*/AIM/FROG2 MM-OPT R^2=0.926								
1	0.935	0.592	0.469	38	0.861	0.666	0.542	10
2	0.948	0.517	0.416	38	0.716	0.991	0.690	10
3	0.927	0.637	0.500	38	0.891	0.559	0.454	10
4	0.915	0.564	0.437	39	0.969	0.989	0.749	9
5	0.924	0.583	0.461	39	0.972	0.747	0.582	9
QM B3LYP 6-31G*/AIM/FROG2 NO-OPT R^2=0.916								
1	0.929	0.618	0.486	38	0.802	0.812	0.737	10
2	0.948	0.516	0.422	38	0.467	1.408	0.922	10
3	0.925	0.649	0.520	38	0.862	0.755	0.584	10
4	0.877	0.681	0.538	39	0.979	0.526	0.433	9
5	0.921	0.595	0.467	39	0.929	0.903	0.713	9
QM B3LYP 6-31G*/AIM/FROG2 QM-OPT R^2=0.853								
1	0.861	0.868	0.705	38	0.779	0.915	0.845	10
2	0.872	0.812	0.641	38	0.739	1.089	0.818	10
3	0.862	0.879	0.715	38	0.832	0.806	0.695	10
4	0.889	0.645	0.535	39	0.977	2.478	1.400	9
5	0.882	0.725	0.556	39	0.240	3.421	1.592	9
QM B3LYP 6-31G*/AIM/NCI MM-OPT R^2=0.914								
1	0.926	0.634	0.487	38	0.814	0.801	0.674	10
2	0.948	0.520	0.370	38	0.584	1.290	0.765	10
3	0.916	0.687	0.553	38	0.904	0.548	0.515	10
4	0.885	0.656	0.539	39	0.979	0.882	0.695	9
5	0.914	0.620	0.509	39	0.908	0.813	0.607	9
QM B3LYP 6-31G*/AIM/NCI NO-OPT R^2=0.903								
1	0.907	0.710	0.531	38	0.802	0.775	0.620	10
2	0.907	0.693	0.457	38	0.855	0.858	0.678	10
3	0.912	0.701	0.502	38	0.921	0.692	0.536	10
4	0.932	0.503	0.374	39	0.987	1.619	0.864	9
5	0.910	0.636	0.505	39	0.927	0.939	0.588	9
QM B3LYP 6-31G*/AIM/NCI QM-OPT R^2=0.892								
1	0.908	0.707	0.568	38	0.781	0.973	0.822	10
2	0.933	0.589	0.445	38	0.597	1.292	1.032	10
3	0.913	0.696	0.559	38	0.710	0.873	0.737	10
4	0.851	0.747	0.619	39	0.974	0.786	0.689	9
5	0.871	0.759	0.627	39	0.985	0.647	0.559	9
QM B3LYP 6-31G*/AIM/OPENBABEL MM-OPT R^2=0.927								
1	0.938	0.579	0.459	38	0.862	0.699	0.552	10
2	0.947	0.523	0.417	38	0.749	0.948	0.658	10
3	0.930	0.625	0.477	38	0.874	0.596	0.457	10
4	0.916	0.562	0.425	39	0.966	1.007	0.761	9
5	0.926	0.576	0.446	39	0.971	0.748	0.577	9
QM B3LYP 6-31G*/AIM/OPENBABEL NO-OPT R^2=0.930								
1	0.940	0.571	0.431	38	0.883	0.693	0.512	10
2	0.949	0.512	0.388	38	0.713	0.971	0.644	10
3	0.934	0.607	0.460	38	0.885	0.558	0.414	10
4	0.913	0.570	0.443	39	0.964	0.896	0.659	9
5	0.929	0.562	0.415	39	0.963	0.748	0.595	9
QM B3LYP 6-31G*/AIM/OPENBABEL QM-OPT R^2=0.880								
1	0.889	0.776	0.625	38	0.803	0.847	0.711	10
2	0.941	0.551	0.416	38	0.496	1.707	1.336	10
3	0.882	0.814	0.671	38	0.908	0.627	0.539	10

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
4	0.891	0.639	0.520	39	0.978	2.592	1.488	9
5	0.893	0.691	0.527	39	0.319	2.905	1.388	9
QM B3LYP 6-31G*/AIM/PUBCHEM MM-OPT R^2=0.916								
1	0.927	0.630	0.495	38	0.807	0.818	0.677	10
2	0.946	0.528	0.391	38	0.696	1.120	0.833	10
3	0.921	0.665	0.514	38	0.846	0.638	0.525	10
4	0.894	0.630	0.502	39	0.970	0.900	0.706	9
5	0.914	0.621	0.497	39	0.957	0.789	0.590	9
QM B3LYP 6-31G*/AIM/PUBCHEM NO-OPT R^2=0.937								
1	0.950	0.518	0.378	38	0.804	0.702	0.565	10
2	0.938	0.565	0.449	38	0.906	0.570	0.447	10
3	0.944	0.558	0.426	38	0.852	0.642	0.543	10
4	0.916	0.561	0.431	39	0.971	0.655	0.492	9
5	0.940	0.518	0.410	39	0.948	0.755	0.549	9
QM B3LYP 6-31G*/AIM/PUBCHEM QM-OPT R^2=0.828								
1	0.859	0.872	0.709	38	0.724	1.289	1.132	10
2	0.851	0.878	0.746	38	0.539	1.519	1.235	10
3	0.853	0.906	0.743	38	0.681	1.035	0.886	10
4	0.779	0.910	0.752	39	0.949	1.023	0.822	9
5	0.879	0.735	0.626	39	0.218	3.821	1.982	9
QM B3LYP 6-31G*/AIM/RDKIT MM-OPT R^2=0.980								
1	0.982	0.313	0.260	38	0.957	0.326	0.287	10
2	0.985	0.274	0.224	38	0.939	0.478	0.419	10
3	0.980	0.331	0.270	38	0.972	0.260	0.224	10
4	0.972	0.323	0.263	39	0.993	0.303	0.250	9
5	0.984	0.270	0.231	39	0.965	0.518	0.425	9
QM B3LYP 6-31G*/AIM/RDKIT QM-OPT R^2=0.908								
1	0.908	0.704	0.523	38	0.902	0.643	0.531	10
2	0.933	0.586	0.458	38	0.689	0.996	0.755	10
3	0.931	0.623	0.472	38	0.683	0.874	0.626	10
4	0.893	0.633	0.468	39	0.983	1.364	0.928	9
5	0.892	0.695	0.524	39	0.987	0.652	0.528	9
QM B3LYP 6-31G*/AIM/RDKIT MM-OPT R^2=0.927								
1	0.938	0.579	0.443	38	0.860	0.723	0.560	10
2	0.949	0.516	0.402	38	0.724	0.976	0.652	10
3	0.932	0.619	0.461	38	0.867	0.605	0.462	10
4	0.909	0.584	0.449	39	0.977	0.920	0.669	9
5	0.925	0.580	0.443	39	0.974	0.738	0.551	9
QM B3LYP 6-31G*/AIM/RDKIT NO-OPT R^2=0.796								
1	0.818	0.993	0.782	38	0.750	1.182	0.869	10
2	0.840	0.909	0.699	38	0.611	1.429	1.084	10
3	0.798	1.064	0.833	38	0.795	0.706	0.524	10
4	0.783	0.903	0.767	39	0.883	1.459	1.198	9
5	0.781	0.988	0.755	39	0.751	1.227	1.087	9
QM B3LYP 6-31G*/MPA/BALLOON MM-OPT R^2=0.953								
1	0.957	0.480	0.314	38	0.912	0.510	0.371	10
2	0.964	0.429	0.269	38	0.666	1.068	0.557	10
3	0.950	0.530	0.366	38	0.979	0.222	0.193	10
4	0.937	0.485	0.311	39	0.983	0.586	0.482	9
5	0.971	0.358	0.266	39	0.876	0.863	0.577	9
QM B3LYP 6-31G*/MPA/BALLOON NO-OPT R^2=0.903								
1	0.916	0.673	0.518	38	0.808	0.796	0.611	10
2	0.927	0.616	0.463	38	0.591	1.376	1.119	10
3	0.911	0.705	0.550	38	0.857	0.687	0.530	10
4	0.884	0.658	0.491	39	0.887	1.054	0.783	9
5	0.910	0.634	0.498	39	0.856	0.945	0.695	9
QM B3LYP 6-31G*/MPA/BALLOON QM-OPT R^2=0.967								
1	0.970	0.401	0.284	38	0.925	0.415	0.309	10
2	0.980	0.321	0.238	38	0.747	1.015	0.735	10

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
3	0.970	0.413	0.306	38	0.958	0.419	0.350	10
4	0.952	0.423	0.318	39	0.994	0.409	0.348	9
5	0.980	0.302	0.229	39	0.917	0.744	0.554	9
QM B3LYP 6-31G*/MPA/FROG2 MM-OPT R^2=0.973								
1	0.981	0.321	0.265	38	0.919	0.498	0.387	10
2	0.977	0.347	0.265	38	0.890	0.579	0.338	10
3	0.974	0.382	0.293	38	0.988	0.324	0.274	10
4	0.965	0.363	0.261	39	0.987	0.432	0.339	9
5	0.976	0.327	0.256	39	0.943	0.646	0.505	9
QM B3LYP 6-31G*/MPA/FROG2 NO-OPT R^2=0.911								
1	0.929	0.621	0.549	38	0.829	0.900	0.721	10
2	0.918	0.650	0.524	38	0.749	0.922	0.716	10
3	0.923	0.655	0.551	38	0.867	0.743	0.632	10
4	0.878	0.676	0.535	39	0.965	0.632	0.592	9
5	0.925	0.578	0.485	39	0.916	1.079	0.866	9
QM B3LYP 6-31G*/MPA/FROG2 QM-OPT R^2=0.973								
1	0.982	0.312	0.245	38	0.908	0.540	0.423	10
2	0.973	0.372	0.273	38	0.927	0.483	0.342	10
3	0.972	0.397	0.293	38	0.989	0.224	0.198	10
4	0.964	0.368	0.262	39	0.975	0.799	0.559	9
5	0.980	0.297	0.231	39	0.803	1.088	0.737	9
QM B3LYP 6-31G*/MPA/NCI MM-OPT R^2=0.955								
1	0.960	0.466	0.299	38	0.911	0.480	0.378	10
2	0.963	0.435	0.279	38	0.644	1.039	0.494	10
3	0.953	0.513	0.353	38	0.978	0.225	0.188	10
4	0.936	0.491	0.336	39	0.985	0.460	0.379	9
5	0.976	0.326	0.256	39	0.868	0.889	0.606	9
QM B3LYP 6-31G*/MPA/NCI NO-OPT R^2=0.970								
1	0.984	0.291	0.228	38	0.829	0.740	0.465	10
2	0.970	0.392	0.279	38	0.899	0.584	0.419	10
3	0.970	0.408	0.299	38	0.971	0.314	0.267	10
4	0.968	0.345	0.240	39	0.998	1.405	0.556	9
5	0.978	0.310	0.251	39	0.966	0.724	0.474	9
QM B3LYP 6-31G*/MPA/NCI QM-OPT R^2=0.959								
1	0.968	0.415	0.278	38	0.890	0.619	0.495	10
2	0.957	0.470	0.354	38	0.933	0.497	0.432	10
3	0.959	0.479	0.348	38	0.964	0.351	0.299	10
4	0.941	0.470	0.338	39	0.992	0.450	0.394	9
5	0.980	0.302	0.231	39	0.877	0.900	0.632	9
QM B3LYP 6-31G*/MPA/OPENBABEL MM-OPT R^2=0.975								
1	0.984	0.298	0.244	38	0.915	0.507	0.372	10
2	0.978	0.335	0.251	38	0.902	0.548	0.321	10
3	0.976	0.367	0.270	38	0.988	0.295	0.265	10
4	0.965	0.365	0.263	39	0.994	0.311	0.258	9
5	0.978	0.310	0.242	39	0.956	0.605	0.468	9
QM B3LYP 6-31G*/MPA/OPENBABEL NO-OPT R^2=0.973								
1	0.979	0.340	0.256	38	0.936	0.444	0.278	10
2	0.974	0.368	0.264	38	0.937	0.425	0.303	10
3	0.978	0.354	0.254	38	0.954	0.457	0.354	10
4	0.963	0.375	0.257	39	0.992	0.350	0.275	9
5	0.979	0.305	0.232	39	0.918	0.745	0.559	9
QM B3LYP 6-31G*/MPA/OPENBABEL QM-OPT R^2=0.974								
1	0.983	0.302	0.238	38	0.913	0.539	0.442	10
2	0.978	0.338	0.240	38	0.867	0.684	0.500	10
3	0.973	0.389	0.293	38	0.989	0.205	0.174	10
4	0.964	0.367	0.262	39	0.995	0.617	0.404	9
5	0.982	0.285	0.227	39	0.765	1.209	0.787	9
QM B3LYP 6-31G*/MPA/PUBCHEM MM-OPT R^2=0.979								
1	0.982	0.313	0.256	38	0.948	0.360	0.270	10

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
2	0.983	0.300	0.230	38	0.915	0.506	0.354	10
3	0.981	0.324	0.253	38	0.987	0.365	0.314	10
4	0.972	0.322	0.262	39	0.991	0.399	0.332	9
5	0.982	0.287	0.240	39	0.987	0.502	0.375	9
QM B3LYP 6-31G*/MPA/PUBCHEM NO-OPT R^2=0.939								
1	0.954	0.499	0.355	38	0.797	0.718	0.645	10
2	0.940	0.556	0.435	38	0.697	0.974	0.697	10
3	0.940	0.581	0.445	38	0.941	0.401	0.343	10
4	0.911	0.576	0.451	39	0.993	0.553	0.413	9
5	0.963	0.408	0.306	39	0.895	1.045	0.772	9
QM B3LYP 6-31G*/MPA/PUBCHEM QM-OPT R^2=0.974								
1	0.982	0.311	0.249	38	0.896	0.519	0.428	10
2	0.978	0.336	0.258	38	0.901	0.585	0.449	10
3	0.974	0.380	0.290	38	0.963	0.296	0.239	10
4	0.960	0.388	0.305	39	0.996	0.257	0.215	9
5	0.978	0.314	0.249	39	0.964	0.751	0.571	9
QM B3LYP 6-31G*/MPA/RDKIT MM-OPT R^2=0.979								
1	0.980	0.328	0.261	38	0.962	0.300	0.239	10
2	0.983	0.295	0.220	38	0.884	0.584	0.443	10
3	0.981	0.326	0.262	38	0.964	0.327	0.257	10
4	0.973	0.318	0.246	39	0.994	0.382	0.323	9
5	0.988	0.229	0.191	39	0.926	0.674	0.521	9
QM B3LYP 6-31G*/MPA/RDKIT QM-OPT R^2=0.964								
1	0.973	0.381	0.290	38	0.880	0.642	0.467	10
2	0.961	0.447	0.319	38	0.938	0.477	0.323	10
3	0.962	0.463	0.331	38	0.985	0.206	0.189	10
4	0.951	0.429	0.296	39	0.993	0.540	0.408	9
5	0.983	0.279	0.212	39	0.887	0.868	0.634	9
QM B3LYP 6-31G*/MPA/RDKIT MM-OPT R^2=0.976								
1	0.984	0.292	0.236	38	0.913	0.522	0.384	10
2	0.979	0.330	0.242	38	0.896	0.563	0.332	10
3	0.976	0.363	0.260	38	0.986	0.310	0.283	10
4	0.964	0.366	0.265	39	0.995	0.305	0.261	9
5	0.978	0.314	0.249	39	0.975	0.562	0.402	9
QM B3LYP 6-31G*/MPA/RDKIT NO-OPT R^2=0.816								
1	0.822	0.981	0.855	38	0.866	0.954	0.730	10
2	0.867	0.829	0.677	38	0.512	1.531	1.308	10
3	0.834	0.964	0.833	38	0.649	0.923	0.768	10
4	0.770	0.929	0.783	39	0.910	1.260	0.983	9
5	0.825	0.883	0.761	39	0.690	1.386	1.139	9
QM B3LYP 6-31G*/NPA/BALLOON MM-OPT R^2=0.927								
1	0.940	0.570	0.439	38	0.821	0.778	0.603	10
2	0.973	0.374	0.264	38	0.350	2.478	1.088	10
3	0.924	0.651	0.516	38	0.957	0.353	0.281	10
4	0.933	0.500	0.378	39	0.974	1.289	0.908	9
5	0.936	0.536	0.441	39	0.941	0.901	0.593	9
QM B3LYP 6-31G*/NPA/BALLOON NO-OPT R^2=0.912								
1	0.932	0.607	0.480	38	0.828	0.886	0.779	10
2	0.930	0.602	0.438	38	0.658	1.021	0.851	10
3	0.916	0.685	0.560	38	0.896	0.575	0.455	10
4	0.875	0.685	0.545	39	0.992	0.622	0.554	9
5	0.915	0.618	0.526	39	0.917	0.884	0.647	9
QM B3LYP 6-31G*/NPA/BALLOON QM-OPT R^2=0.933								
1	0.934	0.598	0.444	38	0.936	0.513	0.378	10
2	0.974	0.369	0.289	38	0.608	1.245	0.918	10
3	0.935	0.603	0.461	38	0.952	0.489	0.370	10
4	0.938	0.480	0.362	39	0.985	2.459	1.474	9
5	0.926	0.575	0.449	39	0.959	0.739	0.509	9
QM B3LYP 6-31G*/NPA/FROG2 MM-OPT R^2=0.965								

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
1	0.970	0.403	0.303	38	0.915	0.480	0.387	10
2	0.965	0.425	0.323	38	0.933	0.456	0.314	10
3	0.964	0.448	0.347	38	0.971	0.266	0.235	10
4	0.971	0.328	0.250	39	0.981	0.782	0.557	9
5	0.970	0.369	0.286	39	0.927	0.777	0.482	9
QM B3LYP 6-31G*/NPA/FROG2 NO-OPT R^2=0.924								
1	0.934	0.598	0.489	38	0.837	0.728	0.654	10
2	0.934	0.584	0.467	38	0.729	0.881	0.711	10
3	0.939	0.585	0.471	38	0.850	0.800	0.678	10
4	0.884	0.660	0.550	39	0.989	0.427	0.362	9
5	0.937	0.532	0.435	39	0.889	0.952	0.720	9
QM B3LYP 6-31G*/NPA/FROG2 QM-OPT R^2=0.911								
1	0.914	0.682	0.541	38	0.895	0.617	0.485	10
2	0.935	0.577	0.447	38	0.633	1.128	0.712	10
3	0.909	0.714	0.574	38	0.942	0.424	0.365	10
4	0.926	0.527	0.424	39	0.987	2.015	1.274	9
5	0.915	0.617	0.492	39	0.486	2.163	1.130	9
QM B3LYP 6-31G*/NPA/NCI MM-OPT R^2=0.967								
1	0.970	0.400	0.316	38	0.908	0.443	0.356	10
2	0.979	0.326	0.247	38	0.784	0.815	0.479	10
3	0.970	0.411	0.326	38	0.935	0.410	0.357	10
4	0.956	0.406	0.330	39	0.982	0.442	0.394	9
5	0.969	0.373	0.298	39	0.949	0.554	0.408	9
QM B3LYP 6-31G*/NPA/NCI NO-OPT R^2=0.966								
1	0.972	0.390	0.268	38	0.859	0.558	0.471	10
2	0.982	0.303	0.245	38	0.741	0.888	0.538	10
3	0.967	0.427	0.307	38	0.952	0.339	0.196	10
4	0.962	0.378	0.290	39	0.996	1.101	0.574	9
5	0.963	0.409	0.285	39	0.968	0.456	0.359	9
QM B3LYP 6-31G*/NPA/NCI QM-OPT R^2=0.937								
1	0.935	0.591	0.439	38	0.915	0.456	0.400	10
2	0.981	0.313	0.242	38	0.561	1.284	0.828	10
3	0.941	0.573	0.407	38	0.920	0.519	0.447	10
4	0.913	0.571	0.433	39	0.990	0.588	0.510	9
5	0.923	0.587	0.434	39	0.974	0.426	0.303	9
QM B3LYP 6-31G*/NPA/OPENBABEL MM-OPT R^2=0.967								
1	0.973	0.380	0.282	38	0.910	0.500	0.402	10
2	0.967	0.414	0.312	38	0.942	0.423	0.298	10
3	0.967	0.429	0.334	38	0.966	0.283	0.227	10
4	0.969	0.341	0.262	39	0.986	0.710	0.501	9
5	0.971	0.357	0.273	39	0.939	0.729	0.464	9
QM B3LYP 6-31G*/NPA/OPENBABEL NO-OPT R^2=0.966								
1	0.968	0.419	0.307	38	0.937	0.387	0.251	10
2	0.965	0.425	0.305	38	0.925	0.457	0.319	10
3	0.974	0.385	0.287	38	0.910	0.528	0.363	10
4	0.963	0.374	0.263	39	0.988	0.615	0.408	9
5	0.970	0.368	0.266	39	0.932	0.762	0.476	9
QM B3LYP 6-31G*/NPA/OPENBABEL QM-OPT R^2=0.928								
1	0.936	0.587	0.442	38	0.919	0.676	0.468	10
2	0.959	0.458	0.355	38	0.622	1.192	0.771	10
3	0.928	0.636	0.503	38	0.948	0.424	0.373	10
4	0.927	0.524	0.411	39	0.992	2.062	1.255	9
5	0.928	0.569	0.459	39	0.930	0.820	0.592	9
QM B3LYP 6-31G*/NPA/PUBCHEM MM-OPT R^2=0.966								
1	0.977	0.354	0.285	38	0.842	0.630	0.598	10
2	0.970	0.391	0.324	38	0.925	0.499	0.383	10
3	0.968	0.426	0.364	38	0.946	0.361	0.318	10
4	0.955	0.410	0.337	39	0.988	0.470	0.416	9
5	0.969	0.369	0.291	39	0.929	0.753	0.462	9

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
QM B3LYP 6-31G*/NPA/PUBCHEM NO-OPT R^2=0.921								
1	0.925	0.638	0.441	38	0.857	0.587	0.466	10
2	0.921	0.638	0.474	38	0.850	0.669	0.500	10
3	0.923	0.656	0.461	38	0.904	0.532	0.431	10
4	0.917	0.559	0.418	39	0.989	1.401	0.821	9
5	0.956	0.444	0.323	39	0.830	1.296	0.891	9
QM B3LYP 6-31G*/NPA/PUBCHEM QM-OPT R^2=0.885								
1	0.900	0.735	0.555	38	0.677	0.840	0.594	10
2	0.887	0.764	0.562	38	0.759	0.825	0.692	10
3	0.900	0.748	0.557	38	0.734	0.767	0.592	10
4	0.841	0.772	0.593	39	0.987	0.749	0.600	9
5	0.943	0.503	0.398	39	0.258	3.322	1.590	9
QM B3LYP 6-31G*/NPA/RDKIT MM-OPT R^2=0.973								
1	0.980	0.327	0.259	38	0.902	0.543	0.465	10
2	0.982	0.303	0.246	38	0.763	0.815	0.510	10
3	0.974	0.383	0.300	38	0.972	0.290	0.246	10
4	0.967	0.350	0.288	39	0.985	0.488	0.364	9
5	0.976	0.329	0.256	39	0.938	0.625	0.521	9
QM B3LYP 6-31G*/NPA/RDKIT QM-OPT R^2=0.927								
1	0.929	0.618	0.480	38	0.911	0.561	0.435	10
2	0.969	0.402	0.313	38	0.577	1.272	0.906	10
3	0.931	0.623	0.495	38	0.893	0.565	0.485	10
4	0.930	0.511	0.408	39	0.985	2.500	1.435	9
5	0.914	0.621	0.499	39	0.977	0.564	0.434	9
QM B3LYP 6-31G*/NPA/RDKIT MM-OPT R^2=0.967								
1	0.973	0.381	0.282	38	0.912	0.515	0.419	10
2	0.967	0.413	0.311	38	0.940	0.432	0.308	10
3	0.968	0.424	0.326	38	0.955	0.323	0.245	10
4	0.967	0.353	0.271	39	0.987	0.681	0.487	9
5	0.971	0.362	0.281	39	0.944	0.698	0.442	9
QM B3LYP 6-31G*/NPA/RDKIT NO-OPT R^2=0.851								
1	0.868	0.845	0.660	38	0.737	0.946	0.811	10
2	0.872	0.813	0.685	38	0.691	1.111	0.910	10
3	0.867	0.864	0.713	38	0.716	0.954	0.772	10
4	0.815	0.834	0.690	39	0.913	1.052	0.941	9
5	0.850	0.818	0.674	39	0.795	1.137	0.935	9
QM HF STO-3G/MPA/BALLOON MM-OPT R^2=0.874								
1	0.872	0.831	0.649	38	0.858	0.662	0.563	10
2	0.973	0.377	0.283	38	0.442	2.008	0.933	10
3	0.874	0.839	0.638	38	0.864	0.578	0.530	10
4	0.825	0.811	0.627	39	0.993	0.855	0.806	9
5	0.858	0.796	0.628	39	0.942	0.963	0.748	9
QM HF STO-3G/MPA/BALLOON NO-OPT R^2=0.836								
1	0.841	0.927	0.751	38	0.792	0.801	0.623	10
2	0.888	0.759	0.583	38	0.305	1.721	1.370	10
3	0.859	0.888	0.704	38	0.634	0.935	0.753	10
4	0.776	0.917	0.765	39	0.948	0.905	0.737	9
5	0.836	0.856	0.685	39	0.760	1.216	0.969	9
QM HF STO-3G/MPA/BALLOON QM-OPT R^2=0.948								
1	0.952	0.508	0.391	38	0.892	0.545	0.505	10
2	0.972	0.378	0.274	38	0.636	1.152	0.872	10
3	0.954	0.506	0.402	38	0.915	0.520	0.366	10
4	0.939	0.480	0.380	39	0.984	0.909	0.713	9
5	0.948	0.484	0.389	39	0.958	0.860	0.740	9
QM HF STO-3G/MPA/FROG2 MM-OPT R^2=0.958								
1	0.963	0.447	0.361	38	0.905	0.500	0.428	10
2	0.963	0.436	0.355	38	0.873	0.656	0.446	10
3	0.963	0.452	0.356	38	0.942	0.546	0.461	10
4	0.951	0.430	0.340	39	0.990	0.761	0.558	9

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
5	0.960	0.421	0.331	39	0.968	0.676	0.507	9
QM HF STO-3G/MPA/FROG2 NO-OPT R^2=0.934								
1	0.941	0.565	0.445	38	0.863	0.668	0.633	10
2	0.950	0.506	0.395	38	0.672	0.977	0.840	10
3	0.944	0.561	0.450	38	0.870	0.654	0.501	10
4	0.903	0.603	0.505	39	0.988	0.474	0.413	9
5	0.945	0.496	0.397	39	0.891	0.929	0.744	9
QM HF STO-3G/MPA/FROG2 QM-OPT R^2=0.875								
1	0.876	0.818	0.702	38	0.840	0.690	0.644	10
2	0.903	0.708	0.588	38	0.581	1.213	1.054	10
3	0.889	0.787	0.661	38	0.743	0.837	0.721	10
4	0.920	0.548	0.445	39	0.990	3.452	1.678	9
5	0.933	0.549	0.415	39	0.410	2.830	1.285	9
QM HF STO-3G/MPA/NCI MM-OPT R^2=0.954								
1	0.960	0.464	0.349	38	0.907	0.581	0.470	10
2	0.971	0.386	0.288	38	0.751	0.912	0.602	10
3	0.959	0.480	0.357	38	0.930	0.485	0.366	10
4	0.938	0.481	0.357	39	0.988	0.549	0.449	9
5	0.956	0.445	0.356	39	0.940	0.624	0.476	9
QM HF STO-3G/MPA/NCI NO-OPT R^2=0.951								
1	0.961	0.462	0.341	38	0.804	0.717	0.579	10
2	0.969	0.397	0.306	38	0.570	1.136	0.763	10
3	0.957	0.490	0.364	38	0.896	0.504	0.421	10
4	0.930	0.513	0.407	39	0.995	0.602	0.412	9
5	0.962	0.412	0.307	39	0.896	0.835	0.670	9
QM HF STO-3G/MPA/NCI QM-OPT R^2=0.942								
1	0.949	0.523	0.423	38	0.849	0.693	0.603	10
2	0.967	0.412	0.305	38	0.658	1.051	0.786	10
3	0.949	0.532	0.418	38	0.888	0.559	0.495	10
4	0.923	0.538	0.415	39	0.980	0.624	0.501	9
5	0.935	0.539	0.439	39	0.964	0.567	0.469	9
QM HF STO-3G/MPA/OPENBABEL MM-OPT R^2=0.958								
1	0.963	0.448	0.357	38	0.908	0.508	0.429	10
2	0.963	0.437	0.356	38	0.875	0.663	0.448	10
3	0.963	0.457	0.355	38	0.949	0.531	0.451	10
4	0.953	0.421	0.330	39	0.989	0.798	0.588	9
5	0.961	0.416	0.326	39	0.960	0.710	0.518	9
QM HF STO-3G/MPA/OPENBABEL NO-OPT R^2=0.952								
1	0.957	0.483	0.372	38	0.918	0.543	0.467	10
2	0.955	0.484	0.369	38	0.873	0.642	0.412	10
3	0.965	0.439	0.334	38	0.889	0.712	0.522	10
4	0.943	0.462	0.335	39	0.988	0.740	0.589	9
5	0.955	0.448	0.335	39	0.959	0.751	0.570	9
QM HF STO-3G/MPA/OPENBABEL QM-OPT R^2=0.874								
1	0.875	0.823	0.704	38	0.846	0.702	0.665	10
2	0.918	0.650	0.528	38	0.469	1.631	1.327	10
3	0.886	0.799	0.677	38	0.768	0.819	0.694	10
4	0.923	0.539	0.427	39	0.989	3.419	1.751	9
5	0.947	0.485	0.372	39	0.357	3.552	1.738	9
QM HF STO-3G/MPA/PUBCHEM MM-OPT R^2=0.942								
1	0.947	0.537	0.408	38	0.912	0.601	0.470	10
2	0.945	0.533	0.398	38	0.811	0.819	0.661	10
3	0.950	0.529	0.385	38	0.867	0.598	0.423	10
4	0.932	0.507	0.356	39	0.963	0.711	0.571	9
5	0.960	0.424	0.351	39	0.898	1.014	0.675	9
QM HF STO-3G/MPA/PUBCHEM NO-OPT R^2=0.938								
1	0.945	0.546	0.387	38	0.854	0.569	0.463	10
2	0.939	0.560	0.421	38	0.843	0.648	0.475	10
3	0.948	0.540	0.393	38	0.850	0.587	0.433	10

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
4	0.926	0.527	0.386	39	0.972	0.712	0.599	9
5	0.938	0.528	0.394	39	0.949	0.702	0.482	9
QM HF STO-3G/MPA/PUBCHEM QM-OPT R^2=0.934								
1	0.936	0.590	0.474	38	0.891	0.517	0.427	10
2	0.939	0.560	0.439	38	0.756	0.830	0.584	10
3	0.943	0.564	0.437	38	0.844	0.594	0.456	10
4	0.933	0.503	0.409	39	0.979	0.912	0.683	9
5	0.951	0.467	0.346	39	0.724	1.680	1.039	9
QM HF STO-3G/MPA/RDKIT MM-OPT R^2=0.965								
1	0.968	0.414	0.327	38	0.914	0.438	0.380	10
2	0.971	0.389	0.296	38	0.791	0.768	0.509	10
3	0.967	0.433	0.343	38	0.946	0.347	0.280	10
4	0.961	0.381	0.304	39	0.994	0.606	0.443	9
5	0.972	0.352	0.290	39	0.958	0.723	0.537	9
QM HF STO-3G/MPA/RDKIT QM-OPT R^2=0.940								
1	0.941	0.564	0.422	38	0.941	0.527	0.419	10
2	0.966	0.418	0.315	38	0.659	1.063	0.760	10
3	0.949	0.533	0.407	38	0.846	0.590	0.420	10
4	0.931	0.508	0.393	39	0.979	1.591	0.955	9
5	0.932	0.551	0.420	39	0.967	0.552	0.483	9
QM HF STO-3G/MPA/RDKIT MM-OPT R^2=0.959								
1	0.964	0.438	0.353	38	0.903	0.526	0.455	10
2	0.963	0.435	0.356	38	0.878	0.647	0.431	10
3	0.966	0.438	0.336	38	0.937	0.586	0.505	10
4	0.951	0.427	0.337	39	0.990	0.757	0.563	9
5	0.961	0.420	0.334	39	0.971	0.659	0.496	9
QM HF STO-3G/MPA/RDKIT NO-OPT R^2=0.853								
1	0.878	0.812	0.666	38	0.843	1.079	0.914	10
2	0.904	0.704	0.570	38	0.381	1.500	1.172	10
3	0.851	0.912	0.763	38	0.871	0.548	0.429	10
4	0.814	0.836	0.661	39	0.936	1.017	0.901	9
5	0.852	0.814	0.686	39	0.682	1.376	1.085	9
Svob2007-chal2 HF STO-3G/MPA/BALLOON MM-OPT R^2=0.867								
1	0.882	0.800	0.600	38	0.666	0.945	0.740	10
2	0.893	0.744	0.539	38	0.714	1.422	0.974	10
3	0.878	0.827	0.613	38	0.778	0.757	0.600	10
4	0.792	0.883	0.699	39	0.988	0.489	0.380	9
5	0.897	0.677	0.549	39	0.745	1.246	0.947	9
Svob2007-chal2 HF STO-3G/MPA/BALLOON NO-OPT R^2=0.806								
1	0.823	0.979	0.807	38	0.827	1.108	1.040	10
2	0.822	0.959	0.763	38	0.762	1.226	0.980	10
3	0.826	0.986	0.804	38	0.677	0.966	0.840	10
4	0.748	0.972	0.772	39	0.881	1.103	0.809	9
5	0.829	0.873	0.756	39	0.697	1.380	1.079	9
Svob2007-chal2 HF STO-3G/MPA/BALLOON QM-OPT R^2=0.953								
1	0.966	0.429	0.308	38	0.888	0.702	0.637	10
2	0.959	0.458	0.345	38	0.771	0.988	0.783	10
3	0.957	0.493	0.399	38	0.944	0.489	0.363	10
4	0.930	0.512	0.410	39	0.987	0.381	0.297	9
5	0.964	0.399	0.314	39	0.902	0.803	0.665	9
Svob2007-chal2 HF STO-3G/MPA/FROG2 MM-OPT R^2=0.863								
1	0.877	0.814	0.637	38	0.764	0.924	0.804	10
2	0.912	0.676	0.528	38	0.647	1.943	1.202	10
3	0.883	0.808	0.632	38	0.678	0.928	0.761	10
4	0.802	0.862	0.697	39	0.961	0.849	0.559	9
5	0.888	0.708	0.582	39	0.655	1.471	1.108	9
Svob2007-chal2 HF STO-3G/MPA/FROG2 NO-OPT R^2=0.894								
1	0.917	0.670	0.539	38	0.819	1.022	0.918	10
2	0.890	0.753	0.611	38	0.835	0.863	0.752	10

Table S5: Cross-validation results.

	Training set				Test set			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
3	0.921	0.664	0.545	38	0.707	1.013	0.779	10
4	0.838	0.780	0.607	39	0.980	0.471	0.428	9
5	0.910	0.635	0.520	39	0.793	1.118	0.945	9
Svob2007-chal2 HF STO-3G/MPA/FROG2 QM-OPT R^2=0.939								
1	0.954	0.500	0.394	38	0.840	0.744	0.651	10
2	0.941	0.553	0.428	38	0.861	0.801	0.633	10
3	0.942	0.568	0.450	38	0.933	0.479	0.410	10
4	0.927	0.523	0.422	39	0.912	0.941	0.537	9
5	0.951	0.468	0.387	39	0.889	0.853	0.672	9
Svob2007-chal2 HF STO-3G/MPA/NCI MM-OPT R^2=0.844								
1	0.857	0.878	0.628	38	0.799	0.956	0.843	10
2	0.892	0.749	0.595	38	0.489	2.919	1.668	10
3	0.855	0.900	0.668	38	0.741	0.811	0.676	10
4	0.780	0.909	0.734	39	0.965	0.992	0.594	9
5	0.889	0.703	0.594	39	0.656	1.445	1.070	9
Svob2007-chal2 HF STO-3G/MPA/NCI NO-OPT R^2=0.870								
1	0.891	0.768	0.581	38	0.912	1.039	0.888	10
2	0.873	0.810	0.593	38	0.806	1.054	0.929	10
3	0.876	0.832	0.646	38	0.857	0.706	0.547	10
4	0.812	0.839	0.626	39	0.955	0.714	0.614	9
5	0.920	0.598	0.506	39	0.658	1.451	1.033	9
Svob2007-chal2 HF STO-3G/MPA/NCI QM-OPT R^2=0.921								
1	0.937	0.586	0.445	38	0.907	0.796	0.677	10
2	0.940	0.555	0.410	38	0.414	1.760	1.102	10
3	0.922	0.660	0.526	38	0.930	0.510	0.418	10
4	0.887	0.652	0.499	39	0.969	0.586	0.540	9
5	0.949	0.480	0.413	39	0.799	1.132	0.865	9
Svob2007-chal2 HF STO-3G/MPA/OPENBABEL MM-OPT R^2=0.858								
1	0.873	0.830	0.659	38	0.781	0.948	0.857	10
2	0.896	0.731	0.574	38	0.677	1.794	1.103	10
3	0.876	0.832	0.659	38	0.670	0.909	0.772	10
4	0.799	0.868	0.710	39	0.931	0.933	0.611	9
5	0.891	0.698	0.569	39	0.635	1.521	1.147	9
Svob2007-chal2 HF STO-3G/MPA/OPENBABEL NO-OPT R^2=0.860								
1	0.874	0.825	0.645	38	0.776	0.924	0.801	10
2	0.896	0.733	0.572	38	0.685	1.769	1.104	10
3	0.877	0.831	0.647	38	0.671	0.873	0.717	10
4	0.803	0.860	0.706	39	0.918	0.937	0.627	9
5	0.892	0.696	0.568	39	0.647	1.491	1.103	9
Svob2007-chal2 HF STO-3G/MPA/OPENBABEL QM-OPT R^2=0.946								
1	0.961	0.462	0.349	38	0.877	0.722	0.643	10
2	0.953	0.492	0.379	38	0.818	0.913	0.761	10
3	0.947	0.544	0.442	38	0.958	0.414	0.323	10
4	0.921	0.546	0.436	39	0.985	0.413	0.347	9
5	0.957	0.437	0.373	39	0.904	0.821	0.639	9
Svob2007-chal2 HF STO-3G/MPA/PUBCHEM MM-OPT R^2=0.886								
1	0.901	0.733	0.558	38	0.755	0.856	0.685	10
2	0.908	0.689	0.531	38	0.653	1.874	1.134	10
3	0.904	0.733	0.555	38	0.714	0.847	0.716	10
4	0.831	0.796	0.639	39	0.981	0.670	0.439	9
5	0.910	0.633	0.511	39	0.777	1.177	0.924	9
Svob2007-chal2 HF STO-3G/MPA/PUBCHEM NO-OPT R^2=0.874								
1	0.911	0.694	0.535	38	0.832	1.238	1.091	10
2	0.894	0.741	0.590	38	0.786	1.205	0.963	10
3	0.889	0.787	0.633	38	0.810	0.846	0.656	10
4	0.816	0.830	0.679	39	0.964	0.843	0.585	9
5	0.878	0.738	0.610	39	0.877	1.010	0.745	9
Svob2007-chal2 HF STO-3G/MPA/PUBCHEM QM-OPT R^2=0.890								
1	0.909	0.703	0.545	38	0.899	0.949	0.841	10

Table S5: Cross-validation results.

	Training set				**Test set**			
Fold	R^2	**RMSE**	$\overline{\Delta}$	N	Q^2	**RMSE**	$\overline{\Delta}$	N
2	0.892	0.748	0.619	38	0.897	0.827	0.701	10
3	0.904	0.732	0.603	38	0.770	0.816	0.670	10
4	0.844	0.764	0.632	39	0.962	0.686	0.615	9
5	0.940	0.516	0.427	39	0.493	1.971	1.150	9
Svob2007-chal2 HF STO-3G/MPA/RDKIT MM-OPT R^2=0.838								
1	0.858	0.875	0.678	38	0.703	1.088	0.931	10
2	0.859	0.853	0.680	38	0.769	1.349	1.004	10
3	0.848	0.923	0.735	38	0.789	0.792	0.692	10
4	0.778	0.913	0.748	39	0.968	0.903	0.618	9
5	0.866	0.774	0.636	39	0.718	1.329	1.065	9
Svob2007-chal2 HF STO-3G/MPA/RDKIT QM-OPT R^2=0.933								
1	0.952	0.510	0.406	38	0.778	0.840	0.783	10
2	0.935	0.577	0.435	38	0.813	0.923	0.650	10
3	0.934	0.606	0.494	38	0.937	0.445	0.372	10
4	0.918	0.556	0.452	39	0.984	1.002	0.725	9
5	0.956	0.443	0.371	39	0.812	1.062	0.851	9
Svob2007-chal2 HF STO-3G/MPA/RDKIT MM-OPT R^2=0.846								
1	0.859	0.873	0.674	38	0.823	0.980	0.891	10
2	0.889	0.757	0.605	38	0.669	1.911	1.119	10
3	0.867	0.864	0.673	38	0.661	0.948	0.801	10
4	0.792	0.884	0.720	39	0.933	1.055	0.636	9
5	0.881	0.730	0.601	39	0.609	1.571	1.151	9
Svob2007-chal2 HF STO-3G/MPA/RDKIT NO-OPT R^2=0.832								
1	0.839	0.932	0.727	38	0.873	0.930	0.847	10
2	0.873	0.812	0.645	38	0.569	1.557	1.126	10
3	0.843	0.937	0.766	38	0.743	0.828	0.691	10
4	0.812	0.840	0.718	39	0.899	1.326	0.850	9
5	0.838	0.851	0.686	39	0.784	1.162	0.980	9

6 Quality Criteria for Testing of Selected EEM QSPR Models

Table S6: Quality criteria for testing of selected EEM QSPR models.
QSPR model description: phenols, charges: EEM (B3LYP/6-31G*/NPA),
3D structure: NCI with no optimization

Quality criteria	R^2	RMSE	$\overline{\Delta}$
Training set	0.906	0.639	0.503
Test set	0.889	0.771	0.596

QSPR model description: phenols, charges: EEM (HF/STO-3G/MPA),
3D structure: RDKit with no optimization

Quality criteria	R^2	RMSE	$\overline{\Delta}$
Training set	0.894	0.677	0.504
Test set	0.869	0.824	0.631

7 Description of Making Dissociated and Associated Form of Molecules

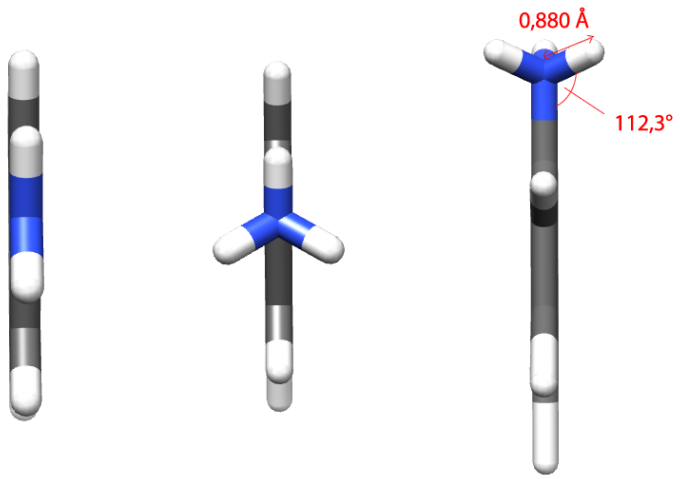


Figure S1: Aniline with its associated form according our simple rules.

Dissociated form of molecule was simply done by removing of hydrogen atom and associated forms was done based simple rules inspired by experimental crystal structure of CCDC-196349[1]. Both hydrogens from amino group were removed and new hydrogen atoms were added with bond length N-H 0.880 Å, bond angle C-N-H 112.3° and dihedral angle respective to the benzene ring 0°, 120° and 240° (see Figure S1).

8 EEM Models Statistics

Table S7: Number and percentage of EEM QSPR models with R^2 higher than a defined limit for individual charge calculation approaches.

EEM charge calculation approach	R^2			R^2_{chrg}
	$\geq$ **0.9**	**(0.9, 0.8>**	$<$ **0.8**	
Svob2007_chal2	67%	30%	4%	0.869
Chaves2006	89%	11%	0%	0.898
Bult2002_npa	86%	14%	0%	0.884
Bult2004_aim	88%	12%	0%	0.883

Note: R^2_{chrg} is the average value of R^2 for all QSPR models, which use charges calculated by a given EEM charge calculation approach.

Table S8: Number and percentage of EEM QSPR models with R^2 higher than a defined limit for individual classes of molecules.

Class of molecules	R^2			R^2_{mol}
	$\geq$ **0.9**	**(0.9, 0.8>**	$<$ **0.8**	
Phenols	62%	38%	0%	0%
Carboxylic acids	18%	50%	29%	3%
Anilines	28%	51%	21%	0%

Note: R^2_{mol} is the average value of R^2 for all QSPR models, which were built for a given class of molecules.

Table S9: Percentage of EEM QSPR models with given R^2 for individual 3D structure sources.

Source	Optimization	R^2			
		$\geq$ 0.9	**(0.9, 0.85>**	**(0.85, 0.8>**	**< 0.8**
Balloon	none	25%	42%	33%	0%
	MM	25%	50%	25%	0%
	QM	**58%**	**33%**	**0%**	**8%**
Frog2	none	17%	67%	17%	0%
	MM	33%	58%	8%	0%
	QM	**58%**	**33%**	**8%**	**0%**
NCI	**none**	**25%**	**67%**	**8%**	**0%**
	mM	33%	33%	33%	0%
	QM	**50%**	**33%**	**17%**	**0%**
OpenBabel	none	25%	42%	33%	0%
	MM	33%	42%	25%	0%
	QM	**67%**	**25%**	**8%**	**0%**
PubChem	**none**	**33%**	**67%**	**0%**	**0%**
	MM	25%	75%	0%	0%
	QM	**42%**	**50%**	**8%**	**0%**
RDKit	none	25%	42%	33%	0%
	UFF	33%	25%	42%	0%
	MM	25%	58%	17%	0%
	QM	**50%**	**42%**	**0%**	**8%**

Note: The optimization procedures which produce the best QSPR models for each source of 3D structures are marked in bold font.

9 Geometrical Properties and Troubles of Selected Structures

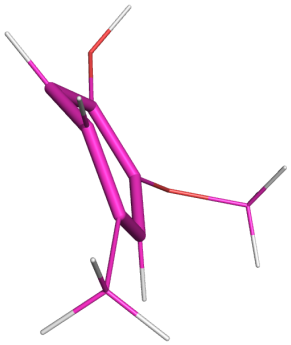
9.1 Class of Molecules: Phenols, Anilines

Troublesome 3D structure source ($R^2 < 0.8$ for at least one QSPR model):

- RDKit, no optimization

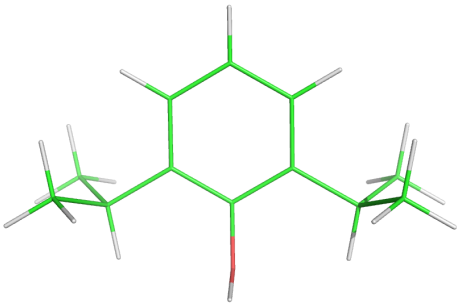
Reasons:

◦ Wrong (non-planar) modeling of benzene cycle
Example: NSC_4969



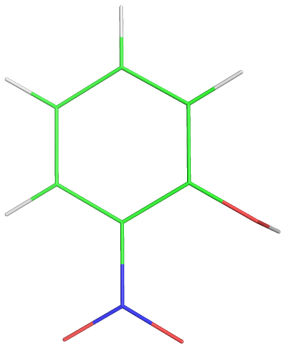
◦ Steric interactions of substituents are not considered
Example: NSC_5105

Correct structure (NCI, no opt.):

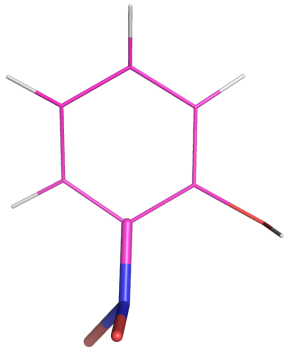


Incorrect structure (RDKit, no opt.):

◦ Non-standard orientation of substituents relative to the ring
Example: NSC_1552
Correct structure (NCI, no opt.):



Incorrect structure (RDKit, no opt.):



9.2 Class of Molecules: Carboxylic Acids

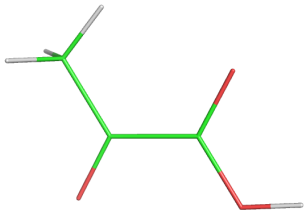
Troublesome 3D structure source ($R^2 < 0.8$ for at least one QSPR model):

- Balloon (all types of optimization)
- OpenBabel (all types of optimization)
- RDKit (all types of optimization expect MM-UFF)

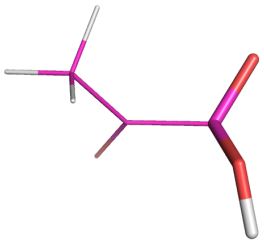
Reasons:

◦ Incorrect modelling of –COOH group:
Example: NSC_179

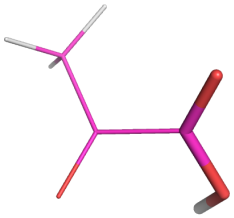
Correct structure (NCI, no opt.):



Incorrect structure (RDKit, no opt.):

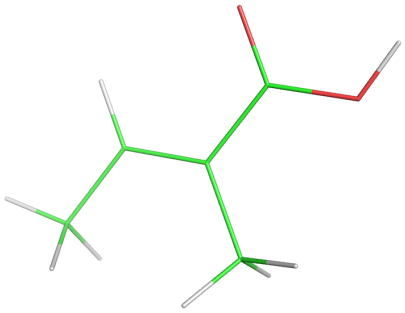


Incorrect structure (Balloon, no opt.):

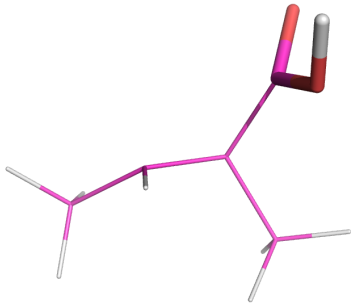


◦ Nonstandard orientation (turn) of –COOH group against the rest of molecule:
Example: NSC_8999

Correct structure (NCI, no opt.):

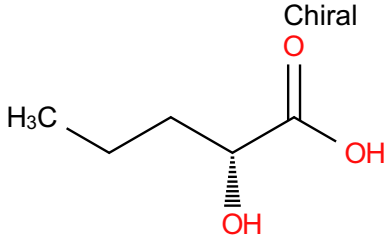


Incorrect structure (RDKit, no opt.):

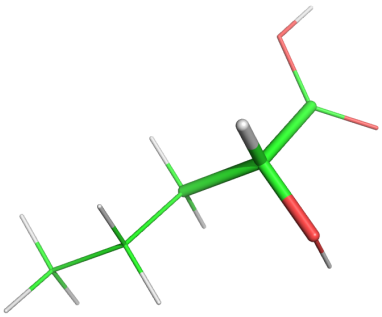


◦ Incorrect modelling of –COOH group:
Example: NSC_67957

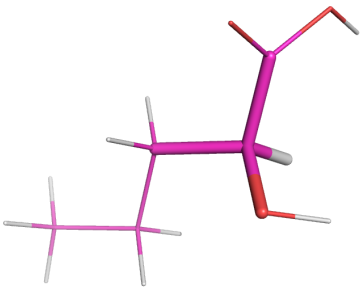
2D structure:



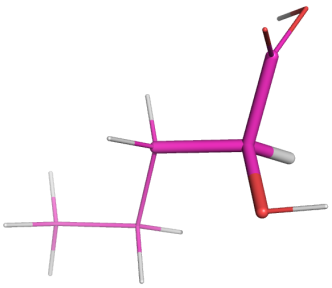
Correct structure (NCI, no opt.):



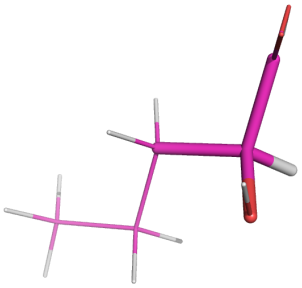
Incorrect structure (RDKit, UFF):



Incorrect structure (Balloon, no opt.):



Incorrect structure (Balloon, MM):



10 Statistical Test of Sensitivity of 3D Structure Source

The sensitivity of a particular QSPR model to a change of molecular class was analyzed via a statistical test, which compared correlation coefficients of three independent populations (i.e. molecular classes). These classes were phenols (with $n_1 = 60$ molecules), carboxylic acids (with $n_2 = 82$ molecules) and anilines (with $n_3 = 48$ molecules). The comparison was done independently for all 4 *ab-initio* QM charge calculation approaches used in our study (i.e., HF, STO-3G, MPA; B3LYP, 6-31G*, MPA; B3LYP, 6-31G*, NPA; B3LYP, 6-31G*, AIM).

Specifically, we tested the null hypothesis: $H_0 : \rho_1 = \rho_2 = \rho_3$, where ρ is the correlation coefficient, against an alternative hypothesis: $H_1 : \exists i, j$ for which $\rho_i \neq \rho_j$.

The test is based on Fisher's z-transformation:

$$z = \frac{1}{2} \ln \frac{(1+r)}{(1-r)} \tag{1}$$

where R is the Pearson correlation coefficient.

After an application of the z-transformation to the hypotheses H_0 and H_1, we obtain these new hypotheses:

$$H'_0 : z_1 = z_2 = z_3 = \bar{z}, \text{ where:}$$

$$\bar{z} = \frac{z_1(n_1 - 3) + z_2(n_2 - 3) + z_3(n_3 - 3)}{n_1 - 3 + n_2 - 3 + n_3 - 3} \tag{2}$$

is a mean of z weighted by $n - 3$.

$H'_1 : \exists i$ for which $z_i \neq \bar{z}$.

Therefore, the tested statistics is described by this equation:

$$K = (n_1 - 3)(z_1 - \bar{z})^2 + (n_2 - 3)(z_2 - \bar{z})^2 + (n_3 - 3)(z_3 - \bar{z})^2 \tag{3}$$

Under the null hypothesis, statistic K follows the χ^2 distribution with $k-1$ degrees of freedom, where k is the number of populations. The test was performed with a statistical level $\alpha = 0.05$.

Table S10: Results of the statistical analysis used for analysis of a 3D structure source sensitivity to a change of a molecular class.

		HF, STO-3G, MPA			B3LYP, 6-31G*, MPA			B3LYP, 6-31G*, NPA			B3LYP, 6-31G*, AIM		
Source	Opt.	K	p	Decision	K	p	Decision	K	p	Decision	K	p	Decision
Balloon	none	2.92	0.23	1	24.45	0	0	6.22	0.04	0	3.98	0.14	1
	MM	2.25	0.32	1	44.28	0	0	13.52	0	0	3.97	0.14	1
	QM	4.36	0.11	1	51.41	0	0	8.13	0.02	0	6.57	0.04	0
Frog2	none	1.98	0.37	1	1.42	0.49	1	2.01	0.37	1	0.86	0.65	1
	MM	10.62	0	0	27.02	0	0	9.04	0.01	0	0.49	0.78	1
	QM	14.68	0	0	28.43	0	0	4.77	0.09	1	5.48	0.06	1
NCI	none	1.09	0.58	1	20.32	0	0	9.16	0.01	0	15.68	0	0
	MM	1.49	0.48	1	12.51	0	0	5.19	0.07	1	0.65	0.72	1
	QM	2.84	0.24	1	13.96	0	0	4.36	0.11	1	0.88	0.64	1
OpenBabel	none	13.14	0	0	71.26	0	0	23.71	0	0	13.1	0	0
	MM	18.14	0	0	75.13	0	0	25.91	0	0	12.16	0	0
	QM	14.78	0	0	66.68	0	0	22.02	0	0	9.22	0.01	0
PubChem	none	5.98	0.05	1	8.94	0.01	0	2.52	0.28	1	1.67	0.43	1
	MM	5.91	0.05	1	23.25	0	0	10.24	0.01	0	2.54	0.28	1
	QM	3.01	0.22	1	19.84	0	0	1.21	0.55	1	3.22	0.2	1
RDKit	none	1.65	0.44	1	9.5	0.01	0	0.88	0.64	1	2.74	0.25	1
	UFF	10.32	0.01	0	41.71	0	0	25.55	0	0	28.99	0	0
	MM	5.84	0.05	1	44.14	0	0	28.66	0	0	9.78	0.01	0
	QM	7.83	0.02	0	47.4	0	0	12.27	0	0	3.94	0.14	1

11 QSPR Models Limitations

The QSPR models provided in the manuscript have the following limitations:

- They are applicable only for the class of molecules, for which they were parameterized (i.e., phenols, carboxylic acids and anilines).
- They require the same type of partial atomic charges, for which they were parameterized (e.g, B3LYP, 6-31G*, MPA; PM6, MPA; Gasteiger-Marsili etc.).
- The quality criteria describing the QSPR models are mentioned with 3 digits precision, but the third digit has lower accuracy. The reason is, the input partial atomic charges have 3 digit precision, but the experimental pK_a values have only 2 digit precision.

References

[1] Lopez-Dupla, E.; Jones, P. G.; Vancea, F. Secondary bonding interactions in some di- and trihaloanilinium halides. *Z. Naturforsch. B Chem. Sci.*, **2003**, *58b*, 191–200.