Predicting chemical reactivity of organic materials using machine-learning approach

Byungju Lee[1,2], Jaekyun Yoo[1], and Kisuk Kang[*,1]

1 Department of Materials Science and Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 151-742, Republic of Korea

2 Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

Supporting information

Name	meting point	boiling point	dielectric constant	Electron affinity	Ionization energy
Water	0	100	78.36	0.39	9.27
1,2-Ethanediol	-12.6	197.5	37.7	0.63	7.05
Formamide	2.5	210.5	109.5	0.91	7.89
N-Methylformamide	-3.8	200	182.4	0.86	7.23
Diethylene glycol	-7.8	245.7	31.69	-0.09	6.91
Triethylene glycol	-4.3	288	23.69	0.32	6.84
Tetraethylene glycol	-6.2	327.3	19.7	-0.23	6.87
2-Methoxyethanol	-85.1	124.6	16.93	0.25	6.93
N-Methylacetamide	30.6	206.7	191.3	0.72	7.05
Ethanol	-114.5	78.3	24.55	0.2	7.73
2-Aminoethanol	10.5	170.9	37.72	0.3	6.37
Acetic acid	16.7	117.9	6.17	0.8	8.43
1-Propanol	-126.2	97.2	20.45	0.18	7.45
Benzyl alcohol	-15.3	205.4	12.7	0.97	6.79
1-Butanol	-88.6	117.7	17.51	0.16	7.45
1-Pentanol	-78.2	138	13.9	0.11	7.44
2-Methyl-1-propanol	-108.2	107.9	17.93	0.18	7.22
2-Propanol	-88	82.2	19.92	0.02	7.72
2-Butanol	-114.7	99.5	16.56	0.15	7.56
2-Pentanol	-50	119	13.71	-0.05	7.31
Nitromethane	-28.6	101.2	35.87	2.84	8.68
Propylene carbonate	-54.5	241.7	64.92	1.3	8.39
3-Pentanol	-8	115.3	13.35	-0.05	7.3
Acetonitrile	-43.8	81.6	35.94	1.06	9.46
Dimethyl sulfoxide	18.5	189	46.45	1.49	6.47
Aniline	-6	184.4	6.98	0.49	5.73
Sulfolane	28.4	287.3	43.3	2.35	7.72
Propanenitrile	-92.8	97.3	28.26	1.05	9.17
2-Methyl-2-propanol	25.6	82.3	12.47	-0.01	7.51

N,N-Dimethylformamide	-60.4	153.1	36.71	0.83	6.86
N,N-Dimethylacetamide	-20.1	166.1	37.78	0.78	6.63
1,3-Dimethylimidazolidin-2-one	8.2	225.5	37.6	0.43	6.03
1-Methylpyrrolidin-2-one	-24.4	202	32.2	0.67	6.63
Acetone	-94.7	56.1	20.56	1.28	7.35
1,3-Dimethyl-2-oxohexahydropyrimidine	-23.8	246.5	36.12	0.59	6.11
1,2-Diaminoethane	11.3	116.9	12.9	-0.49	5.7
Cyanobenzene	-12.7	191.1	25.2	1.99	7.42
2-Butanone	-86.7	79.6	18.11	1.26	7.29
Nitrobenzene	5.8	210.9	34.78	3.3	7.57
2-Pentanone	-76.9	102.3	15.38	1.23	7.28
2-Methyl-2-butanol	-8.8	102	5.78	-0.27	7.64
Morpholine	-4.8	129	7.42	-0.39	5.93
Tetramethylurea	-1.2	175.3	23.6	0.66	6.1
Hexamethylphosphoric triamide	7.3	233	29.3	-0.21	5.76
3-Methyl-2-butanone	-92.2	94.9	15.87	1.21	7.14
Dichloromethane	-94.9	39.6	8.93	2.3	8.75
Acetophenone	19.7	202	17.39	2.22	7.13
Pyridine	-41.5	115.3	12.91	1.29	7.47
Cyclohexanone	-32.1	155.7	15.5	1.29	7.02
4-Methyl-2-pentanone	-84.2	117.5	13.11	1.2	7.21
Quinoline	-14.9	237.2	8.95	1.92	6.69
3-Pentanone	-39	102	17	1.25	7.2
3,3-Dimethyl-2-butanone	-52.5	105.9	12.6	1.17	7.04
Methyl acetate	-98.1	56.9	6.68	0.77	8.21

Triethylene glycol dimethyl ether	-45	216	7.6	-0.66	7.33
2,4-Dimethyl-3-pentanone	-69	125.3	17.2	1.16	6.95
1,2-Dimethoxyethane	-69.2	84.6	7.2	-0.9	6.98
Ethyl acetate	-83.6	77.2	6.02	0.71	8.16
2,6-Dimethyl-4-heptanone	-46	168.2	9.91	1.14	7.07
Diethylene glycol diethyl ether	-44.3	188.9	5.7	-0.77	7.04
Tetrahydrofuran	-108.4	66	7.58	-0.56	7.21
Methoxybenzene	-37.5	153.7	4.33	0.36	6.52
Fluorobenzene	-42.2	84.8	5.42	0.71	7.27
1,1-Dichloroethene	-122.6	31.6	4.82	2.04	7.71
Chlorobenzene	-45.6	131.7	5.62	0.78	7.18
Diethyl carbonate	-74.3	126.9	2.82	0.18	9.05
Bromobenzene	-30.9	156	5.4	0.78	7.12
Ethoxybenzene	-29.6	169.9	4.22	0.34	6.49
1,4-Dioxane	11.8	101.3	2.21	-1.57	7.57
Piperidine	-10.5	106.3	5.8	-0.57	5.98
Diethylamine	-49.8	55.5	3.78	-0.79	6.25
Diphenyl ether	26.9	258.1	3.6	0.59	6.54
t-Butyl methyl ether	-108.6	55.2	4.5	-0.71	7.42
Diethyl ether	-116.3	34.5	4.2	-0.89	7.62
Benzene	5.6	80.1	2.27	-0.19	7.9
Di-n-propyl ether	-123.2	90.1	3.39	-0.95	7.57
Toluene	-95	110.6	2.38	-0.15	7.44
1,4-Dimethylbenzene	13.3	138.4	2.27	-0.2	7.13
Di-n-butyl ether	-95.2	140.3	3.08	-0.99	7.56
Carbon disulfide	-111.6	46.3	2.64	1.52	8.49
Tetrachloromethane	-22.9	76.7	2.24	2.32	9.85

cis-Decahydronaphthalene	-43.1	195.8	2.2	-1.32	7.92
n-Pentane	-129.8	36.1	1.84	-1.5	9.02
Ethyl benzene	-94	136	2.41	-0.12	7.41
TetraIin	-31	205	2.77	0	7.03
Propylene glycol methyl ether	-139	121	16.9	0.19	6.81
Ethylene glycol ethyl ether	-70	135	5.3	-0.14	7.19
Ethylene glycol monobutyl ether	-75	171	5.3	-0.14	7.15
Diisopropyl ether	-86	68	3.39	-0.87	7.46
n-Butyl acetate	-73	126	5.6	0.68	8.19
Furfuraldehyde	-37	162	41.9	2.47	6.91
Phenol	41	182	10	0.73	6.41
Tetraethylene glycol dimethyl ether	-30	275	7.9	-1.06	6.92

Table S1. Experimental properties [ref: SMALLWOOD, Ian. Handbook of organic solvent properties. Butterworth-Heinemann, 2012.] (melting point, boiling point and dielectric constant) and calculated properties (adiabatic ionization energy, adiabatic electron affinity) of known solvents

# of electronic features, m	Performance, test set R^2	RMSE
0	0.811	3.48
1	0.897	2.57
5	0.932	2.08
10	0.913	2.36
30	0.919	2.28
50	0.924	2.2

Table S2. Test set R-square and RMSE values of electrophilicity prediction models trained with the different number of electronic features, m. Other hyperparameters are fixed (two hidden layers with an 80×80 neuron size, 1024-bits Morgan fingerprint, learning rate of 10^{-3}, 100,000 learning steps, dropout rate of 0.8, and L2 regularization parameter of 10^{-3})

# of electronic features, m	Performance, test set R^2	RMSE
0	0.646	4.41
1	0.815	3.19
5	0.839	2.97
10	0.831	3.05
30	0.810	3.23
50	0.821	3.13

Table S3. Test set R-square and RMSE values of nucleophilicity prediction models trained with the different number of electronic features, m. Other hyperparameters are fixed (two hidden layers with an 80×80 neuron size, 1024-bits Morgan fingerprint, learning rate of 10^{-3}, 100,000 learning steps, dropout rate of 0.8, and L2 regularization parameter of 10^{-3})

Fingerprint	Performance, test set R^2	RMSE
Morgan (128-bits)	0.917	2.3
Morgan (256-bits)	0.912	2.37
Morgan (512-bits)	0.922	2.23
Morgan (1024-bits)	0.919	2.28
MACCS key (166-bits)	0.911	2.38

Table S4. Test set R-square and RMSE values of electrophilicity prediction models trained with the different types of structural features, i.e. fingerprints. Other hyperparameters are fixed (two hidden layers with an 80×80 neuron size, 30 largest electronic features, learning rate of 10^{-3}, 100,000 learning steps, dropout rate of 0.8, and L2 regularization parameter of 10^{-3}).

Fingerprint	Performance, test set R^2	RMSE
Morgan (128-bits)	0.777	3.5
Morgan (256-bits)	0.794	3.36
Morgan (512-bits)	0.805	3.27
Morgan (1024-bits)	0.810	3.23
MACCS key (166-bits)	0.841	2.95

Table S5. Test set R-square and RMSE values of nucleophilicity prediction models trained with the different types of structural features, i.e. fingerprints. Other hyperparameters are fixed (two hidden layers with an 80×80 neuron size, 30 largest electronic features, learning rate of 10^{-3}, 100,000 learning steps, dropout rate of 0.8, and L2 regularization parameter of 10^{-3}).

ANN structure	Performance, test set R^2	RMSE
40/40	0.911	2.38
60/60	0.920	2.26
80/80	0.919	2.28
100/100	0.930	2.12
200/200	0.926	2.17
300/300	0.937	2.01
400/400	0.942	1.93
500/500	0.937	2.01
600/600	0.938	1.99
700/700	0.939	1.97
1000/1000	0.942	1.93
80/60/40	0.919	2.28
80/60/40/20	0.902	2.5
100/80/60/40/20	0.904	2.48
100/200/100	0.914	2.34
100/50/100	0.931	2.1
100/50/100/50	0.905	2.46
40/40 MACCS keys fingerprint	0.884	2.72
200/200 MACCS keys fingerprint	0.906	2.45
400/400 MACCS keys fingerprint	0.916	2.32

Table S6. Test set R-square and RMSE values of electrophilicity prediction models trained with the different architecture of a neural network. Other hyperparameters are fixed (1024-bits Morgan fingerprint, 30 largest electronic features, a learning rate of 10^{-3}, 100,000 learning steps, dropout rate of 0.8, and L2 regularization parameter of 10^{-3}).

ANN structure	Performance, test set R^2	RMSE
40/40	0.592	4.73
60/60	0.830	3.05
80/80	0.810	3.23
100/100	0.817	3.17
200/200	0.82	3.14
300/300	0.811	3.22
400/400	0.809	3.24
500/500	0.817	3.17
600/600	0.822	3.13
700/700	0.728	3.86
1000/1000	0.062	7.18
80/60/40	0.816	3.17
80/60/40/20	0.830	3.05
100/80/60/40/20	0.818	3.16
100/200/100	0.814	3.2
100/50/100	0.834	3.02
100/50/100/50	0.817	3.17
40/40 MACCS keys fingerprint	0.831	3.05
100/100 MACCS keys fingerprint	0.837	2.99
200/200 MACCS keys fingerprint	0.836	3.00
400/400 **MACCS keys fingerprint**	0.840	2.96
600/600 MACCS keys fingerprint	0.809	3.24
80/60/40/20 MACCS keys fingerprint	0.830	3.05

Table S7. Test set R-square and RMSE values of nucleophilicity prediction models trained with the different architecture of a neural network. Other hyperparameters are fixed (1024-bits Morgan fingerprint, 30 largest electronic features, a learning rate of 10^{-3}, 100,000 learning steps, dropout rate of 0.8, and L2 regularization parameter of 10^{-3}).

Learning rate	Performance, test set R^2	RMSE
$1*10^{-2}$	0.385	5.81
$5*10^{-3}$	0.730	3.85
$1*10^{-3}$	0.810	3.23
$5*10^{-4}$	0.511	5.18
$1*10^{-4}$	0.545	5.00

Table S8. Test set R-square and RMSE values of nucleophilicity prediction models trained with the different learning rates. Other hyperparameters are fixed (two hidden layers with an 80×80 neuron size, 1024-bits Morgan fingerprint, 30 largest electronic features, 100,000 learning steps, dropout rate of 0.8, and L2 regularization parameter of 10^{-3}).

Dropout rate	Performance, test set R^2	RMSE
0.9	0.811	3.22
0.8	0.810	3.23

Table S9. Test set R-square and RMSE values of nucleophilicity prediction models trained with the different dropout rates. (two hidden layers with an 80×80 neuron size, 1024-bits Morgan fingerprint, 30 largest electronic features, learning rate of 10^{-3}, 100,000 learning steps, and L2 regularization parameter of 10^{-3}).

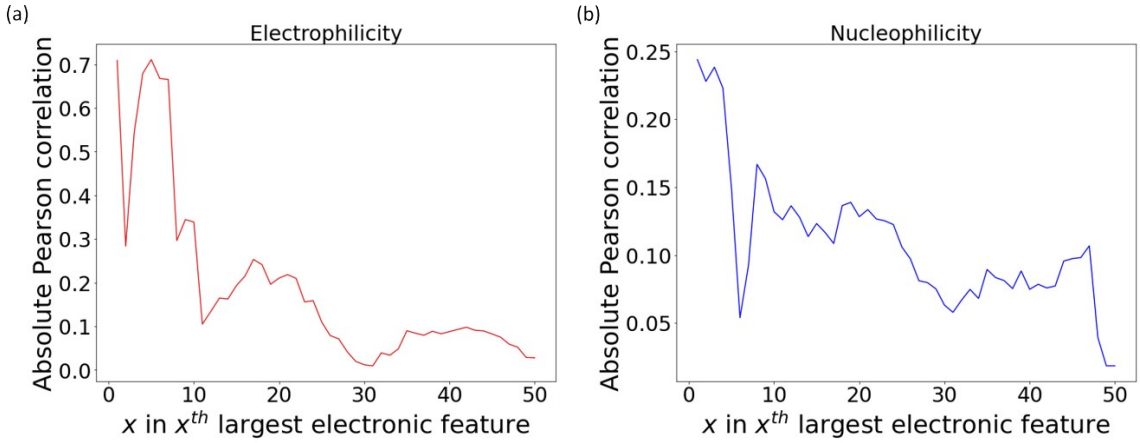


Figure S1. The 50 largest electronic features (*i.e.* local reactivities) vs. absolute Pearson correlation of experimental (a) electrophilicity-local reactivity and (b) nucleophilicity-local reactivity.

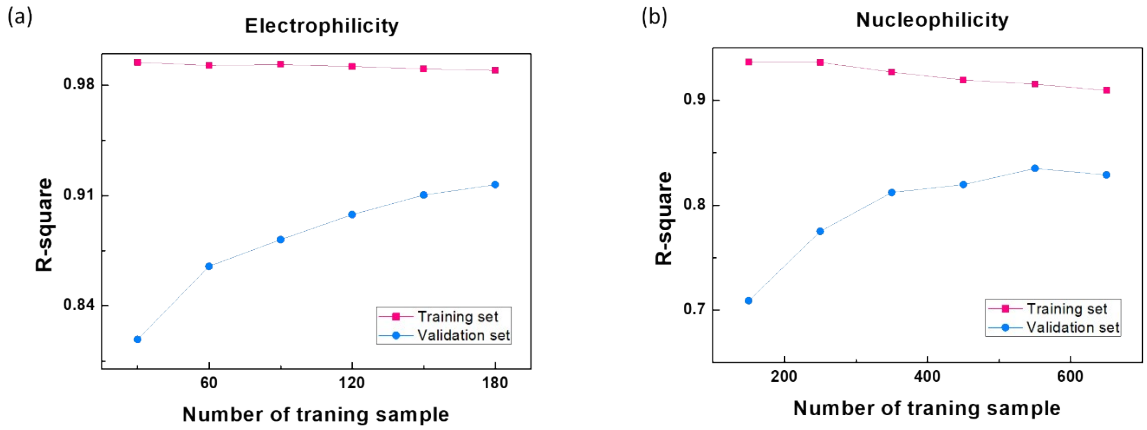


Figure S2. The number of training sample versus R-square value of test and training set, calculated for the final model of (a) electrophilicity and (b) nucleophilicity. It is shown that for both electrophilicity and nucleophilicity cases, the R-square value of validation set improves with increasing number of training sample, which guarantees that the number of training sample is greater than the VC dimension.

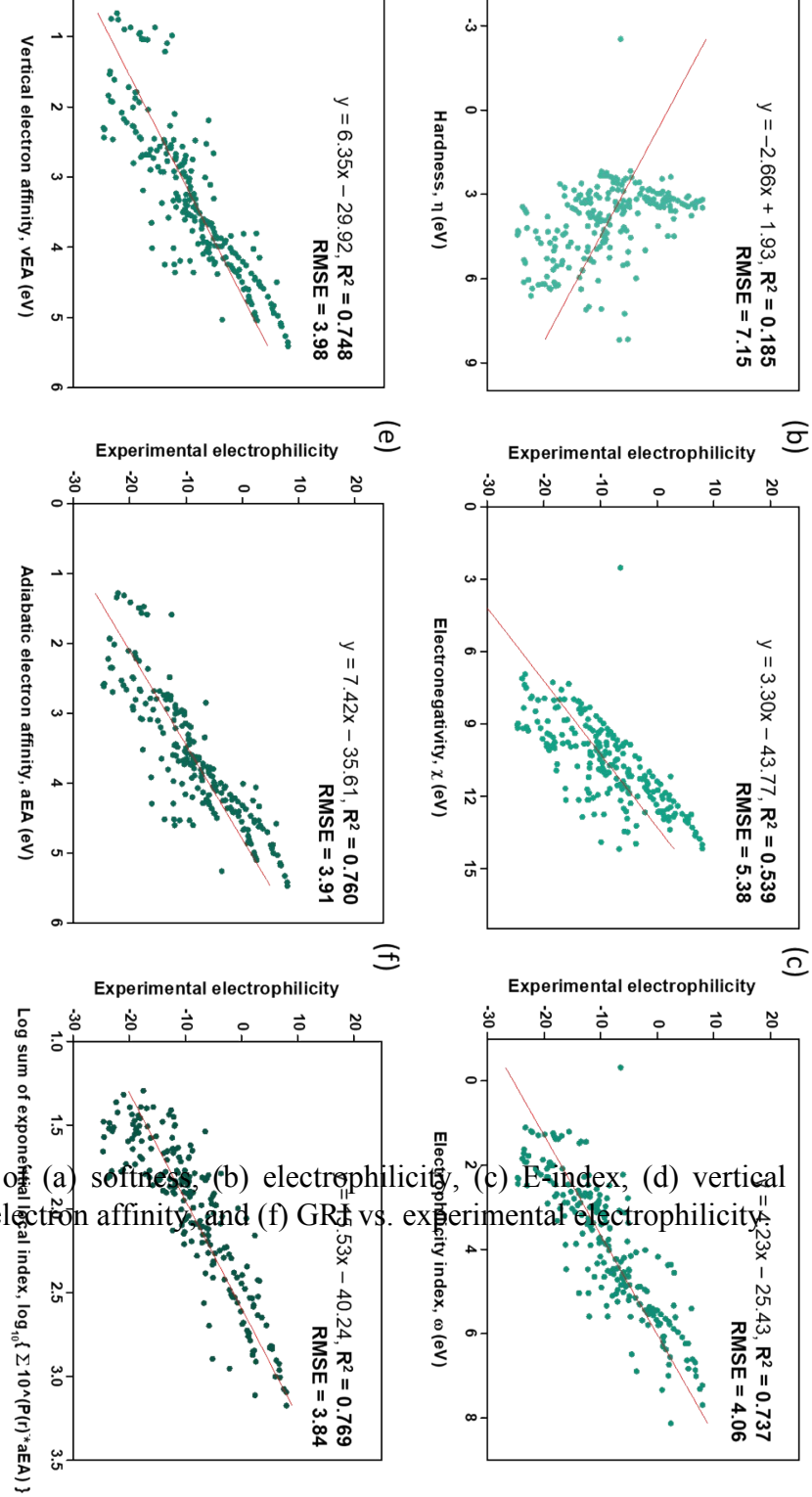


Figure S3. Linear regression of (a) softness, (b) electrophilicity, (c) E-index, (d) vertical electron affinity, (e) adiabatic electron affinity, and (f) GRI vs. experimental electrophilicity.

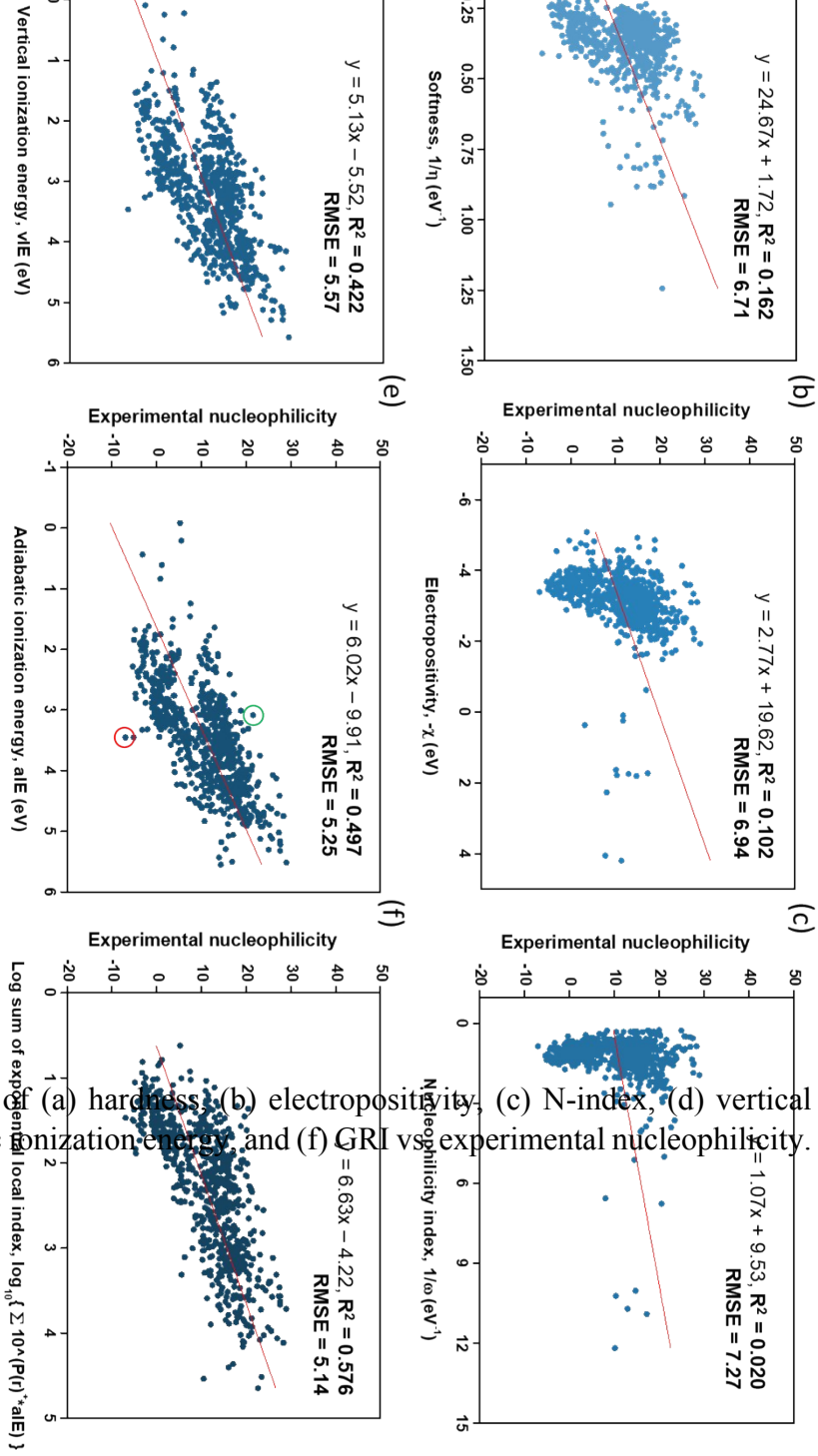


Figure S4. Linear regression of (a) hardness, (b) electropositivity, (c) N-index, (d) vertical ionization energy, (e) adiabatic ionization energy, and (f) GRI vs experimental nucleophilicity.

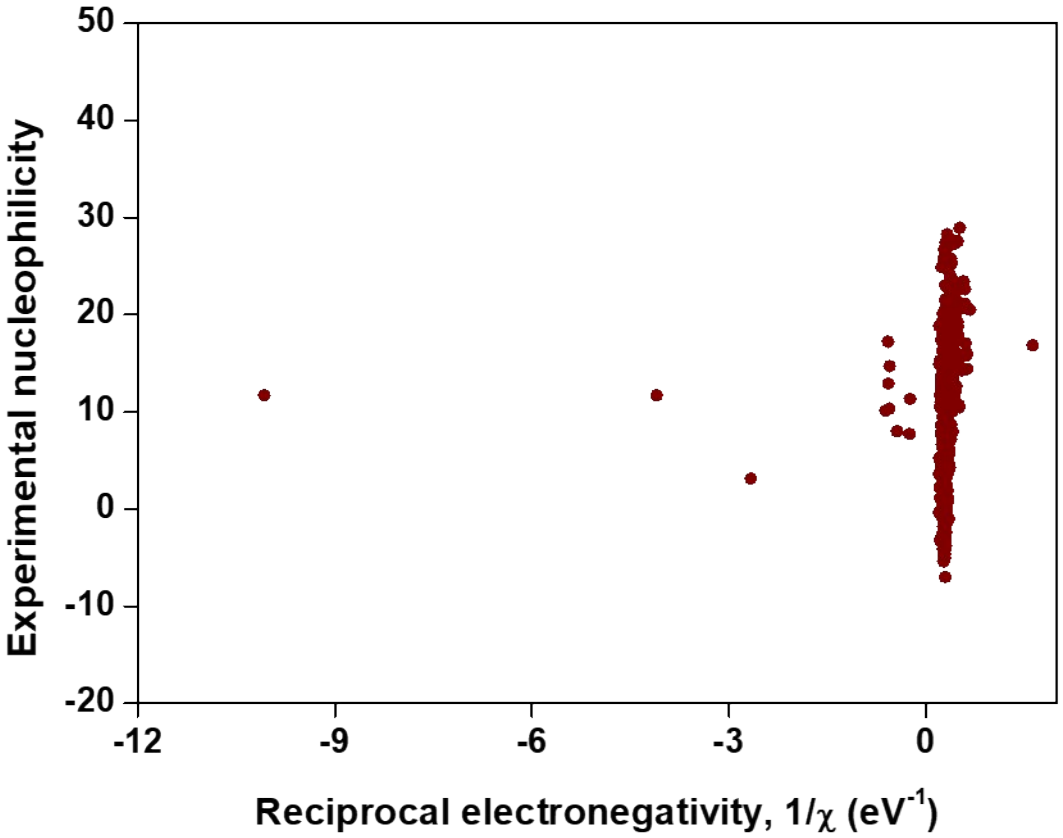


Figure S5. Reciprocal electronegativity vs. nucleophilicity plot.

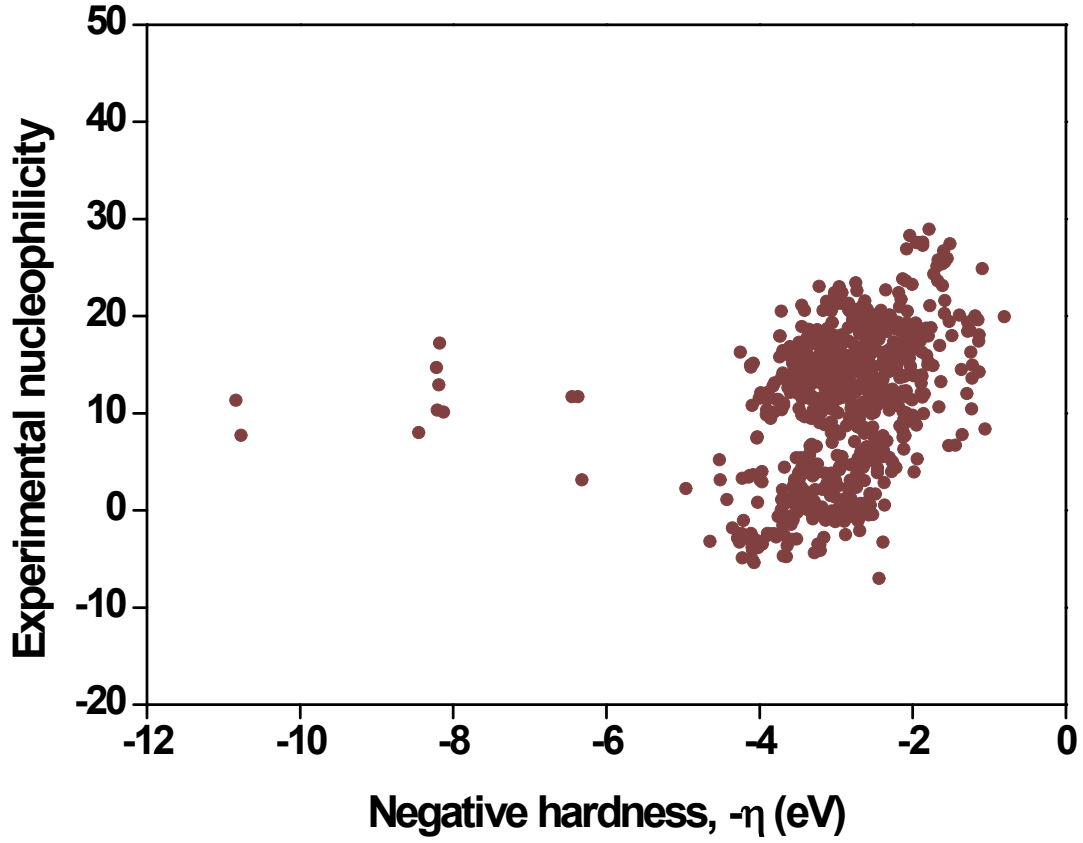
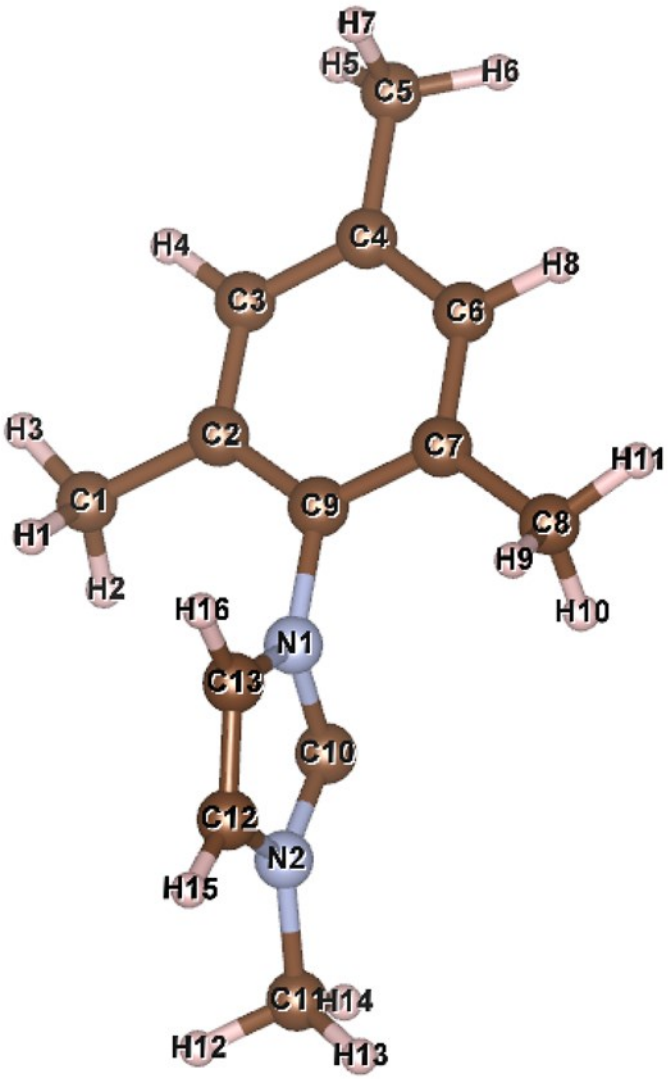
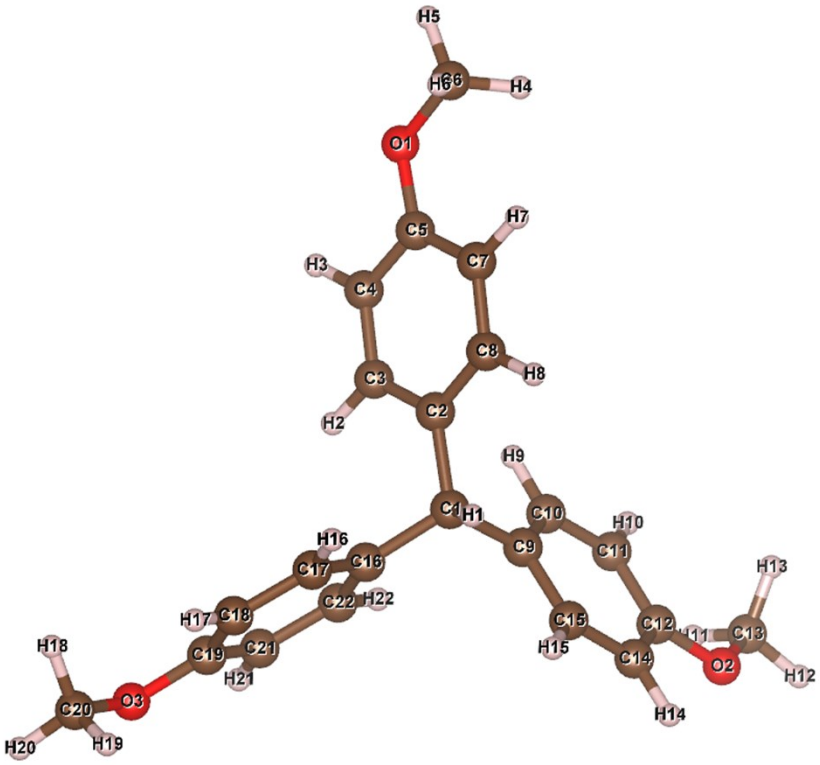


Figure S6. Negative hardness vs. nucleophilicity plot.



Atom name	Spin density
C11	1.00285
C14	0.02211
C15	0.02037
C9	0.00434
C13	0.0035
H31	0.00346
H30	0.00343
C1	0.00131
C6	0.00084
C3	0.00082
C8	0.00077
H18	0.00035
H26	0.00018
C5	9E-05
H21	-2E-05
H20	-2E-05
H22	-4E-05
H19	-5E-05
H23	-6E-05
H24	-9E-05
H16	-0.0001
H29	-0.00019
H25	-0.00028
H17	-0.00048
C4	-0.00093
H28	-0.00148
H27	-0.00155
C2	-0.00169
C7	-0.00177
N10	-0.02667
N12	-0.029

Figure S7. Structure of 3-mesityl-1-methyl-1H-imidazol-3-ium-2-ide (designated as a green circle in Figure S4e), of which nucleophilicity was underestimated in prediction using aIE. Table describes Parr function of each atom.



Atom name	Spin density
C2	0.12567
C10	0.11676
C18	0.10041
C5	0.07757
O6	0.0757
C13	0.07218
O14	0.06895
C21	0.0622
O22	0.06027
C4	0.04693
C24	0.03538
C16	0.02797
H26	0.02761
C12	0.0259
C9	0.02314
C19	0.01875
C11	0.01763
C8	0.01523
C20	0.01208
H29	0.00373
H31	0.00368
H36	0.0035
H38	0.00318
H43	0.003
H44	0.00293
C17	0.00086
H37	0
H30	0
H45	-1E-05
H27	-7E-05
H47	-0.00013
H40	-0.00015
H42	-0.00051
H41	-0.00063
H32	-0.00064
H34	-0.00068
H33	-0.0008
H35	-0.0008
H39	-0.00098
H46	-0.00104
H28	-0.00144
C25	-0.0021
C23	-0.00304
C15	-0.00349
C7	-0.00381
C3	-0.00469
C1	-0.00624

Figure S8. Structure of bis(4-methoxyphenyl)phenylmethane (designated as a red circle in Figure S4e), of which nucleophilicity was overestimated in prediction using aIE. Table describes Parr function of each atom.

LASSO		
Electrophilicity		
Regularization rate (λ)	**Performance, test set R^2**	**RMSE**
10^{-1}	0.827	3.32
10^{-2}	0.814	3.45
10^{-3}	0.608	5.01
10^{-4}	0.484	5.74
Nucleophilicity		
10^{-1}	0.180	6.71
10^{-2}	0.597	4.70
10^{-3}	0.657	4.34
10^{-4}	0.650	4.38

Table S10. Test set R-square and RMSE values of electrophilicity and nucleophilicity trained with LASSO linear regression. Adam optimizer with learning rate 10^{-3}, MACCS keys, 5 largest electronic features, 100,000 learning steps were used.

Ridge		
Electrophilicity		
Regularization rate (λ)	**Performance, test set R^2**	**RMSE**
10^{-1}	0.844	3.16
10^{-2}	0.852	3.08
10^{-3}	0.814	3.45
10^{-4}	0.650	4.73
Nucleophilicity		
10^{-1}	0.448	5.50
10^{-2}	0.526	5.10
10^{-3}	0.473	5.38
10^{-4}	0.482	5.33

Table S11. Test set R-square and RMSE values of electrophilicity and nucleophilicity trained with ridge linear regression. Adam optimizer with learning rate 10^{-3}, MACCS keys, 5 largest electronic features, 100,000 learning steps were used.

Gaussian process		
Electrophilicity		
Kernel function, Learning rate	**Performance, test set R^2**	**RMSE**
Exponential quadratic, 10^{-1}	0.883	2.73
Exponential quadratic, 10^{-2}	0.887	2.69
Exponential quadratic, 10^{-3}	0.678	4.54
Matern3/2, 10^{-1}	0.878	2.79
Matern3/2, 10^{-2}	0.881	2.76
Matern3/2, 10^{-3}	0.612	4.98
Nucleophilicity		
Exponential quadratic, 10^{-1}	0.556	4.94
Exponential quadratic, 10^{-2}	0.253	6.40
Exponential quadratic, 10^{-3}	-1.094	10.72
Matern3/2, 10^{-1}	0.557	4.93
Matern3/2, 10^{-2}	0.304	6.18
Matern3/2, 10^{-3}	-1.260	11.14

Table S12. Test set R-square and RMSE values of electrophilicity and nucleophilicity trained with Gaussian process regression. Adam optimizer, MACCS keys, 5 largest electronic features, 1,000 learning steps were used.

Support vector machine	
Structure, **Regularization rate (λ)**	**Test set accuracy**
Linear, 10^{-1}	0.654
Linear, 10^{-2}	0.643
Linear, 10^{-3}	0.624
Exponential quadratic, 10^{-1}	0.728
Exponential quadratic, 10^{-2}	0.728

Table S13. Test set accuracy of the support vector machine classifier. Adadelta optimizer with learning rate 10^{-3}, MACCS keys, 5 largest electronic features, 10,000 learning steps were used. From 216 electrophilicity and 826 nucleophilicity data, the 178,416 reaction kinetic constants ($k = 10^{N+E}$) was generated by combining N and E values. If N+E > -5, the reaction will occur (label = 1), otherwise the label value would be 0. Simple merge of feature vectors, *i.e.* structural feature vector of electrophilicity, electronic feature vector of electrophilicity, structural feature vector of nucleophilicity and electronic feature vector of nucleophilicity, was used as the input vector.

Random forest	
Number of trees (n), **Maximum depth (d)**	**Test set accuracy**
n = 100, d = 6	0.805
n = 100, d = 10	0.782
n = 300, d = 6	0.801
n = 300, d = 10	0.789
n = 500, d = 6	0.804
n = 500, d = 10	0.780

Table S14. Test set accuracy of the random forest classifier. MACCS keys and 5 largest electronic features were used. From 216 electrophilicity and 826 nucleophilicity data, the 178,416 reaction kinetic constants ($k = 10^{N+E}$) was generated by combining N and E values. If N+E > -5, the reaction will occur (label = 1), otherwise the label value would be 0. Simple merge of feature vectors, *i.e.* structural feature vector of electrophilicity, electronic feature vector of electrophilicity, structural feature vector of nucleophilicity and electronic feature vector of nucleophilicity, was used as the input vector.

Features of local reactivity expression										
	Electronic (30)					Structural (1024)				R^2 value 0.919
	#1	#2	#3	…	#30	#31	#32	…	#1054	
Molecule 1	$10^{P^-(r_{1i})\cdot aEA_1}$	$10^{P^-(r_{1j})\cdot aEA_1}$	$10^{P^-(r_{1k})\cdot aEA_1}$	…	$10^{P^-(r_{1l})\cdot aEA_1}$	Fingerprint for molecule 1				
Molecule 2	$10^{P^-(r_{1a})\cdot aEA_2}$	$10^{P^-(r_{1b})\cdot aEA_2}$	$10^{P^-(r_{2c})\cdot aEA_2}$	…	$10^{P^-(r_{2d})\cdot aIE_2}$	Fingerprint for molecule 2				RMSE 2.28
…						…				
Features of 'only global index' expression										
	Electronic (1)	Structural (1024)								R^2 value 0.839
	#1	#2	#3	…					#1025	
Molecule 1	aEA_1	Fingerprint for molecule 1								
Molecule 2	aEA_2	Fingerprint for molecule 2								RMSE 3.21
…										
Features of 'simple merge of global index and Parr function vector' expression										
	Electronic (31)					Structural (1024)				R^2 value 0.864
	#1	#2	#3	…	#31	#32	#33	…	#1055	
Molecule 1	aEA_1	$10^{P^-(r_{1i})}$	$10^{P^-(r_{1j})}$		$10^{P^-(r_{1l})}$	Fingerprint for molecule 1				
Molecule 2	aEA_2	$10^{P^-(r_{1a})}$	$10^{P^-(r_{1b})}$		$10^{P^-(r_{2d})}$	Fingerprint for molecule 2				RMSE 2.95
…										

Table S15. Detailed feature compositions of the 'local reactivity', 'only global index', and 'simple merge of global index and Parr function vector' expressions which were used for prediction of experimental electrophilicity. $P^-(r_{xy})$ and aEA_x denotes the negative Parr function of *y*th atom in molecule *x* and adiabatic electron affinity of molecule *x*, respectively. For nucleophilicity prediction, $P^+(r_{xy})$ and aIE_x, *i.e.*, positive Parr function and adiabatic ionization energy, were used instead of $P^-(r_{xy})$ and aEA_x.

Name	Predicted E	Predicted N	N(solvent) +E(DMPZ$^+$)	N(DMPZ) +E(solvent)
Water	-18.69	5.28	-4.53	-8.22
1,2-Ethanediol	-15.55	4.48	-5.34	-5.07
Formamide	-17.94	8.77	-1.05	-7.46
N-Methylformamide	-18.04	8.35	-1.47	-7.57
Diethylene glycol	-18.31	-0.2	-10.01	-7.83
Triethylene glycol	-15.83	0.74	-9.07	-5.35
Tetraethylene glycol	-14.87	0.44	-9.38	-4.39
2-Methoxyethanol	-18.97	0.99	-8.83	-8.49
N-Methylacetamide	-20.52	8.73	-1.09	-10.04
Ethanol	-18.46	7.4	-2.42	-7.98
2-Aminoethanol	-17.05	14.15	4.33	-6.57
Acetic acid	-18.5	0.86	-8.96	-8.02
1-Propanol	-19.73	3.57	-6.25	-9.25
Benzyl alcohol	-17.8	2.69	-7.13	-7.32
1-Butanol	-19.87	2.92	-6.89	-9.39
1-Pentanol	-20.13	4.33	-5.49	-9.65
2-Methyl-1-propanol	-18.09	2.76	-7.06	-7.61
2-Propanol	-18.87	0.17	-9.64	-8.39
2-Butanol	-20.38	-0.03	-9.85	-9.91
2-Pentanol	-20.92	0.32	-9.50	-10.44
Nitromethane	-17.28	5.29	-4.53	-6.80
Propylene carbonate	-18.11	-3.82	-13.64	-7.63
3-Pentanol	-20.56	1.69	-8.13	-10.08
Acetonitrile	-17.95	5.62	-4.19	-7.48
Dimethyl sulfoxide	-18.88	6.39	-3.43	-8.41
Aniline	-18.2	12.81	2.99	-7.72
Sulfolane	-11.16	-3.13	-12.95	-0.68
Propanenitrile	-18.02	4.84	-4.98	-7.54

2-Methyl-2-propanol	-18.98	-1.92	-11.73	-8.51
N,N-Dimethylformamide	-20.16	8.19	-1.63	-9.69
N,N-Dimethylacetamide	-22.92	8.75	-1.06	-12.44
1,3-Dimethylimidazolidin-2-one	-19.5	12.26	2.44	-9.02
1-Methylpyrrolidin-2-one	-19.61	10.92	1.11	-9.13
Acetone	-18.09	2.35	-7.47	-7.61
1,3-Dimethyl-2-oxohexahydropyrimidine	-19.99	14.68	4.86	-9.51
1,2-Diaminoethane	-19.7	13.16	3.34	-9.23
Cyanobenzene	-15.95	0.68	-9.14	-5.47
2-Butanone	-20.29	0.48	-9.34	-9.82
Nitrobenzene	-12.84	4.14	-5.67	-2.36
2-Pentanone	-21.99	0	-9.81	-11.51
2-Methyl-2-butanol	-21.34	-0.97	-10.79	-10.86
Morpholine	-18.72	15.2	5.38	-8.24
Tetramethylurea	-23.91	12.33	2.51	-13.43
Hexamethylphosphoric triamide	-18.6	8.84	-0.97	-8.12
3-Methyl-2-butanone	-19.98	-1.29	-11.10	-9.50
Dichloromethane	-17.61	-0.22	-10.03	-7.13
Acetophenone	-16.81	-1.93	-11.75	-6.33
Pyridine	-15.84	12.23	2.41	-5.36
Cyclohexanone	-19.98	1.78	-8.03	-9.51
4-Methyl-2-pentanone	-21	-0.04	-9.86	-10.52
Quinoline	-13.74	10.55	0.74	-3.26
3-Pentanone	-20.5	0.69	-9.12	-10.03

3,3-Dimethyl-2-butanone	-20.8	-2.86	-12.68	-10.32
Methyl acetate	-19.64	-1.21	-11.02	-9.16
Triethylene glycol dimethyl ether	-15.31	-2.37	-12.18	-4.83
2,4-Dimethyl-3-pentanone	-20.26	0.23	-9.58	-9.78
1,2-Dimethoxyethane	-20.84	-2.9	-12.71	-10.36
Ethyl acetate	-22.55	-4.28	-14.09	-12.07
2,6-Dimethyl-4-heptanone	-19.65	-0.71	-10.53	-9.17
Diethylene glycol diethyl ether	-17.74	-4.07	-13.89	-7.26
Tetrahydrofuran	-20.06	-2.6	-12.41	-9.58
Methoxybenzene	-17.07	-1.44	-11.26	-6.60
Fluorobenzene	-15.09	-3.52	-13.33	-4.61
1,1-Dichloroethene	-17.72	-1.42	-11.23	-7.24
Chlorobenzene	-16.71	-1.16	-10.98	-6.23
Diethyl carbonate	-23.82	-7.51	-17.32	-13.35
Bromobenzene	-14.82	-3	-12.82	-4.34
Ethoxybenzene	-20.3	-0.9	-10.71	-9.82
1,4-Dioxane	-19.58	0.35	-9.46	-9.11
Piperidine	-18.71	15.8	5.98	-8.24
Diethylamine	-21.38	15.06	5.24	-10.90
Diphenyl ether	-15.38	-1.13	-10.94	-4.90
t-Butyl methyl ether	-20.1	-1.33	-11.15	-9.62
Diethyl ether	-21.45	-4.55	-14.37	-10.97
Benzene	-18.53	-2.01	-11.83	-8.05
Di-n-propyl ether	-21.82	-2.21	-12.02	-11.34
Toluene	-19.2	-3.05	-12.87	-8.73
1,4-Dimethylbenzene	-19.54	-1.54	-11.36	-9.06
Di-n-butyl ether	-18.06	-2.18	-11.99	-7.58
Carbon disulfide	-18.15	2.32	-7.50	-7.67

Tetrachloromethane	-17	-2.66	-12.47	-6.52
cis-Decahydronaphthalene	-15	2.15	-7.67	-4.52
n-Pentane	-21.2	-2.08	-11.89	-10.72
Ethyl benzene	-20.46	-0.59	-10.40	-9.99
TetraIin	-19.08	-0.97	-10.79	-8.60
Propylene glycol methyl ether	-19.08	2.5	-7.32	-8.61
Ethylene glycol ethyl ether	-20.9	1.28	-8.54	-10.42
Ethylene glycol monobutyl ether	-18.92	2.39	-7.43	-8.44
Diisopropyl ether	-20.16	0.1	-9.71	-9.68
n-Butyl acetate	-23.81	-6.15	-15.97	-13.33
Furfuraldehyde	-17.27	1.79	-8.02	-6.79
Phenol	-16.74	2.36	-7.45	-6.26
Tetraethylene glycol dimethyl ether	-15.4	-1.87	-11.69	-4.92

Table S16. Predicted electrophilicities (E), nucleophilicities (N), and N+E values for reaction against DMPZ species (DMPZ and $DMPZ^{+}$), of 93 solvents