Supplementary Information for ***Acta Phys. -Chim. Sin.*** **2015**, *31* (5), 893-898
doi: 10.3866/PKU.WHXB201503193

基于范数指数定量构效关系预测 β-环糊精络合常数

钱海城 [1] KANWAL Shahid[1] 贾青竹 [2] 王 强 [1,*] 计惠芬 [1]
朱志臣 [3] 夏淑倩 [4] 马沛生 [4]

([1] 天津科技大学材料科学与化学工程学院, 天津 300457; [2] 天津科技大学海洋科学与工程学院, 天津 300457; [3] 天津城市建设学院理学院, 天津 300384; [4] 天津大学化工学院, 天津 300072)

Norm Index-Based Quantitative Structure-Activity Relationship to Predict β-Cyclodextrin Complex Binding Constants

QIAN Hai-Cheng[1] KANWAL Shahid[1] JIA Qing-Zhu[2] WANG Qiang[1,*]
JI Hui-Fen[1] ZHU Zhi-Chen[3] XIA Shu-Qian[4] MA Pei-Sheng[4]

([1]*School of Material Science and Chemical Engineering, Tianjin University of Science and Technology, Tianjin 300457, P. R. China*; [2]*School of Marine Science and Engineering, Tianjin University of Science and Technology, Tianjin 300457, P. R. China*; [3]*School of Science, Tianjin University of Urban Construction, Tianjin 300384, P. R. China*; [4]*School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, P. R. China*))

[*]Corresponding author. Email: wang_q@tust.edu.cn; Tel: +86-22-58088631, +86-13652089933.

Table S1 Observed and predicted β-cyclodextrin complex binding constant (log*K*) calculated by using various methods for 233 compounds

NO.	Name	CAS	Expt.	Pérez-Garrido work	This work			
					[1]PLS	[2]GRNN	[3]MLS	[4]LS-SVM
1	*carbon tetrachloride	56-23-5	2.2	2.25	2.3	2.2	2.3	2.21
2	chloroform	67-66-3	1.43	1.39	1.19	1.43	1.19	1.51
3	methanol	67-56-1	-0.49	-0.74	-0.44	-0.48	-0.439	-0.32
4	acetonitrile	1975/5/8	-0.27	-0.38	0.24	-0.244	0.237	-0.41
5	acetaldehyde	75-07-0	-0.64	-0.03	0.025	-0.0546	0.0251	-0.43
6	*ethanol	64-17-5	-0.03	-0.08	0.63	-0.0683	0.628	-0.04
7	1,2-ethanediol	107-21-1	-0.19	0.17	0.88	0.951	0.884	0.10
8	acetone	67-64-1	0.42	0.4	1.19	0.948	1.19	0.08
9	*1-propanol	71-23-8	0.57	0.54	0.91	0.385	0.906	0.59
10	2-propanol	67-63-0	0.63	0.77	1.96	0.381	1.96	0.64
11	1,3-propanediol	504-63-2	0.67	0.7	1.15	1.52	1.15	0.71
12	*tetrahydrofuran	109-99-9	1.47	0.9	1.69	1.28	1.69	1.56
13	cyclobutanol	2919-23-5	1.18	1.25	1.36	1.29	1.36	1.44
14	1-butanol	71-36-3	1.22	1.07	1.25	1.33	1.25	1.26
15	2-butanol	78-92-2	1.19	1.34	1.89	1.32	1.89	1.42
16	*2-methyl-1-propanol	78-83-1	1.62	1.35	1.74	1.32	1.74	1.53
17	2-methyl-2-propanol	75-65-0	1.68	1.33	3.46	1.32	3.46	1.65
18	1,4-butanediol	110-63-4	0.64	1.12	1.39	1.96	1.39	1.14
19	diethylamine	109-89-7	1.36	1.22	1.34	0.956	1.34	1.22
20	cyclopentanol	96-41-3	2.08	1.87	2.39	1.87	2.39	2.38
21	1-pentanol	96-41-3	1.8	1.49	1.53	1.91	1.53	1.59
22	2-pentanol	6032-29-7	1.49	1.87	2.16	1.9	2.16	1.71
23	3-pentanol	584-02-1	1.35	1.7	1.96	1.9	1.96	1.75
24	2-methyl-1-butanol	137-32-6	2.08	1.71	1.8	1.9	1.8	1.90
25	2-methyl-2-butanol	75-85-4	1.91	1.93	3.08	1.9	3.08	2.04
26	*3-methyl-1-butanol	123-51-3	2.25	1.88	1.98	1.9	1.98	1.87
27	3-methyl-2-butanol	598-75-4	1.92	1.82	2.32	1.9	2.32	2.00
28	2,2-dimethyl-1-propanol	75-84-3	2.71	1.94	2.85	1.9	2.85	2.15
29	1,5-pentanediol	111-29-5	1.22	1.43	1.6	2.25	1.6	1.59
30	*1,4-dibromobenzene	106-37-6	2.97	3.14	2.11	2.97	2.11	2.88
31	1,4-diiodobenzene	624-38-4	3.17	3.38	2.15	2.36	2.15	2.70
32	3,5-dibromophenol	626-41-5	2.56	3.2	2.37	2.56	2.37	2.64
33	3,5-dichlorophenol	591-35-5	2.07	2.99	2.12	2.07	2.12	2.09
34	*1-chloro-4-nitrobenzene	100-00-5	2.15	2.39	1.87	2.44	1.87	2.23
35	fluorobenzene	462-06-6	1.96	1.37	2.11	2.22	2.11	1.97
36	bromobenzene	108-86-1	2.5	2.31	2.31	2.5	2.31	2.52
37	iodobenzene	591-50-4	2.93	2.44	2.21	2.01	2.21	2.35

38	3-fluorophenol	372-20-3	1.7	1.6	2.23	2.07	2.23	2.01
39	*4-fluorophenol	371-41-5	1.73	2.25	2.3	2.2	2.3	2.21
40	3-chlorophenol	108-43-0	2.28	2.41	2.28	2.46	2.28	2.35
41	4-chloropheno	106-48-9	2.61	2.47	2.27	2.46	2.27	2.35
42	3-bromophenol	591-20-8	2.51	2.52	2.53	2.58	2.53	2.57
43	4-bromophenol	106-41-2	2.65	2.58	2.52	2.58	2.52	2.57
44	3-iodophenol	626-02-8	2.93	2.65	2.34	2.35	2.34	2.34
45	4-iodophenol	540-38-5	2.98	2.71	2.33	2.35	2.33	2.33
46	*nitrobenzene	98-95-3	2.04	1.75	1.99	2.65	1.99	2.15
47	4-nitrophenol	100-02-7	2.39	1.8	2.03	2.5	2.03	2.26
48	benzene	54682-86-9	2.23	1.4	2.21	1.47	2.21	1.89
49	phenol	108-95-2	1.98	1.71	2.39	2.19	2.39	2.02
50	hydroquinone	123-31-9	2.05	1.93	2.5	2.21	2.5	1.96
51	*4-nitroaniline	100-01-6	2.48	2.14	2.14	2.51	2.14	2.29
52	aniline	62-53-3	1.6	1.92	2.48	1.8	2.48	2.07
53	sulfanilamide	63-74-1	2.76	2.72	2.36	2.77	2.36	2.49
54	cyclohexanol	108-93-0	2.67	2.55	2.55	2.24	2.55	2.58
55	1-hexanol	111-27-3	2.33	1.79	1.8	2.26	1.8	1.94
56	*2-hexanol	626-93-7	1.98	2.28	2.4	2.26	2.4	2.02
57	2-methyl-2-pentanol	590-36-3	1.99	2.55	3.17	2.26	3.17	2.27
58	3-methyl-3-pentanol	77-74-7	2.15	2.27	2.9	2.26	2.9	2.33
59	4-methyl-2-pentanol	108-11-2	2.04	2.4	2.7	2.26	2.7	2.24
60	3,3-dimethyl-2-butanol	464-07-3	2.75	2.28	3.13	2.26	3.13	2.48
61	*1,6-hexanediol	629-11-8	1.69	1.65	1.81	2.38	1.81	2.17
62	benzonitrile	100-47-0	2.23	1.84	2.11	2.24	2.11	2.43
63	benzothiazole	95-16-9	2.38	2.59	2.4	2.64	2.4	2.63
64	4-nitrobenzoic acid	62-23-7	2.34	1.71	1.75	2.5	1.75	2.42
65	*benzaldehyde	100-52-7	1.78	1.89	1.84	2.21	1.84	2.00
66	benzoic acid	65-85-0	2.12	2.05	2.19	2.27	2.19	1.99
67	4-hydroxybenzaldehyde	123-08-0	1.75	2.04	1.97	2.27	1.97	2.02
68	4-hydroxybenzoic acid	99-96-7	2.2	2.09	2.24	2.7	2.24	2.20
69	benzyl chloride	100-44-7	2.45	2.7	1.85	2.53	1.85	2.55
70	*toluene	108-88-3	2.09	2.03	2.44	2.01	2.44	2.30
71	benzyl alcohol	100-51-6	1.71	2.25	2.27	2.36	2.27	2.17
72	anisole	100-66-3	2.32	2.11	2.1	2.36	2.1	1.98
73	m-cresol	108-39-4	1.98	2.24	2.59	2.36	2.59	2.18
74	p-cresol	106-44-5	2.4	2.3	2.57	2.36	2.57	2.15
75	4-methoxyphenol	150-76-5	2.21	2.27	2.22	2.4	2.22	2.07
76	*3-methoxyphenol	150-19-6	2.11	2.19	2.23	2.4	2.23	2.10
77	4-hydroxybenzyl alcohol	623-05-2	2.16	2.39	2.39	2.4	2.39	2.22
78	hydrochlorothiazide	58-93-5	1.76	1.94	2.1	1.76	2.1	1.82
79	n-methylaniline	100-61-8	2.12	2.38	2.35	2.21	2.35	2.38

80	1-butylimidazole	4316-42-1	2.19	2.96	2.39	2.47	2.39	2.96
81	1-heptanol	111-70-6	2.85	2	2.04	2.49	2.04	2
82	*phenylacetylene	536-74-3	2.36	2.07	2.55	2.35	2.55	2.07
83	*thianaphthene	95-15-8	3.23	2.83	2.6	2.62	2.6	2.83
84	4-fluorophenyl acetate	405-51-6	2.11	2.24	2.13	2.44	2.13	2.24
85	3-fluorophenyl acetate	701-83-7	1.91	2.13	2.22	2.43	2.22	2.13
86	4-chlorophenyl acetate	876-27-7	2.5	2.93	2.28	2.52	2.28	2.93
87	3-chlorophenyl acetate	13031-39-5	2.44	2.84	2.41	2.52	2.41	2.84
88	4-bromophenyl acetate	1927-95-3	2.68	3.02	2.95	2.68	2.95	3.02
89	3-bromophenyl acetate	35065-86-2	2.67	2.94	2.94	2.67	2.94	2.94
90	4-iodophenyl acetate	33527-94-5	3	3.15	2.25	2.73	2.25	3.15
91	3-iodophenyl acetate	42861-71-2	3.07	3.06	2.3	2.73	2.3	3.06
92	*4-nitrophenyl acetate	830-03-5	2.13	1.91	1.9	2.46	1.9	1.91
93	acetophenone	98-86-2	2.27	2.34	2.22	2.39	2.22	2.34
94	phenyl acetate	122-79-2	2.1	2.39	2.4	2.49	2.4	2.39
95	methyl benzoate	93-58-3	2.5	2.24	2.1	2.49	2.1	2.24
96	3-hydroxyacetophenone	121-71-1	2.06	2.35	2.34	2.48	2.34	2.35
97	4-hydroxyacetophenone	99-93-4	2.18	2.44	2.28	2.48	2.28	2.44
98	*acetoanilide	103-84-4	2.2	2.65	2.52	3.44	2.52	2.65
99	p-xylene	106-42-3	2.38	2.61	2.7	2.34	2.7	2.61
100	ethylbenzene	100-41-4	2.59	2.55	2.41	2.34	2.41	2.55
101	phenetole	103-73-1	2.49	2.57	2.34	2.46	2.34	2.57
102	2-phenylethanol	1960/12/8	2.15	2.72	2.31	2.44	2.31	2.72
103	3-ethylphenol	620-17-7	2.6	2.66	2.51	2.44	2.51	2.66
104	*4-ethylphenol	123-07-9	2.69	2.75	2.53	2.44	2.53	2.75
105	4-ethoxyphenol	622-62-8	2.33	2.66	2.28	3.15	2.28	2.66
106	3-ethoxyphenol	621-34-1	2.35	2.58	2.34	3.13	2.34	2.58
107	3,5-dimethoxyphenol	500-99-2	2.34	2.25	2.18	2.74	2.18	2.25
108	n-ethylaniline	103-69-5	2.34	2.83	2.53	2.25	2.53	2.83
109	n,n-dimethylaniline	121-69-7	2.36	2.8	2.66	2.25	2.66	2.8
110	*barbita	57-44-3	1.78	2.39	2.14	2.07	2.14	2.39
111	*cyclooctanol	696-71-9	3.3	3.31	2.77	2.55	2.77	3.31
112	1-octanol	111-87-5	3.17	2.13	2.27	2.75	2.27	2.13
113	2-octanol	123-96-6	3.13	2.74	2.73	2.73	2.73	2.74
114	quinoline	91-22-5	2.12	2.47	2.64	2.2	2.64	2.47
115	3-cyanophenyl acetate	55682-11-6	1.49	2.24	2.24	2.42	2.24	2.24
116	4-hydroxycinnamic acid	4501-31-9	2.83	2.61	2.28	2.69	2.28	2.61
117	ethyl benzoate	93-89-0	2.73	2.53	2.26	2.75	2.26	2.53
118	4'-hydroxypropiophenone	70-70-2	2.63	2.7	2.43	2.74	2.43	2.57
119	3'-hydroxypropiophenone	13103-80-5	2.61	2.61	2.49	2.74	2.49	2.57
120	*p-tolyl acetate	140-39-6	2.49	2.78	2.64	2.75	2.64	2.56
121	3-methylphenyl acetate	122-46-3	2.21	2.69	2.62	2.75	2.62	2.55

122	4-methoxyphenyl acetate	1200-06-2	2.45	2.45	2.31	2.72	2.31	2.50
123	4-propylphenol	645-56-7	3.55	3.14	2.66	2.99	2.66	3.18
124	*3-propylphenol	621-27-2	3.28	3.05	2.64	2.98	2.64	3.18
125	4-isopropylphenol	99-89-8	3.58	3.18	2.86	2.96	2.86	3.18
126	3-isopropylphenol	618-45-1	3.44	3.08	2.91	2.95	2.91	3.18
127	4-isopropoxyphenol	7495-77-4	2.86	3.08	2.78	2.8	2.78	3.25
128	*2-norbornaneacetate	—	3.59	3.42	2.94	2.76	2.94	3.23
129	1-benzylimidazole	4238-71-5	2.61	3.12	2.75	3.04	2.75	3.23
130	1-benzylimidazole	4238-71-5	2.93	2.95	2.48	2.58	2.48	2.70
131	4-ethylphenyl acetate	3245-23-6	2.83	2.97	2.69	2.6	2.69	2.78
132	3-ethylphenyl acetate	3056-60-8	2.68	2.82	2.67	2.59	2.67	2.81
133	*4-ethoxyphenyl acetate	—	2.54	2.63	2.16	2.48	2.16	2.71
134	3-ethoxyphenyl acetate	—	2.49	2.47	2.09	2.47	2.09	2.70
135	allobarbital	52-43-7	1.98	2.28	2.56	2.58	2.56	2.10
136	4-n-butylphenol	1638-22-8	3.97	3.44	2.75	3.79	2.75	3.85
137	*3-n-butylphenol	—	3.76	3.35	2.8	3.78	2.8	3.86
138	3-isobutylphenol	30749-25-8	4.21	3.45	2.99	3.75	2.99	4.00
139	4-sec-butylphenol	99-71-8	4.18	3.41	2.88	3.75	2.88	3.97
140	3-sec-butylphenol	3522-86-9	4.06	3.31	2.87	3.74	2.87	3.98
141	4-tert-butylphenol	30704-63-3	4.56	3.69	3.45	3.73	3.45	4.00
142	*3-tert-Butylphenol	585-34-2	4.41	3.58	3.51	3.72	3.51	3.99
143	menadion	58-27-5	2.27	2.42	2.33	2.62	2.33	2.41
144	sulfapyridine	144-83-2	2.7	2.68	2.8	2.74	2.8	2.67
145	sulfamonomethoxine	1220-83-3	2.48	1.87	1.93	2.49	1.93	2.32
146	sulfisoxazole	127-69-5	2.32	2.58	2	2.31	2	2.39
147	4-n-propylphenyl acetate	—	3.15	3.13	2.87	2.89	2.87	3.07
148	3-n-propylphenyl acetate	—	3.28	2.96	2.84	2.88	2.84	3.00
149	4-isopropylphenyl acetate	2664-32-6	2.88	3.26	2.99	2.88	2.99	2.98
150	3-isopropylphenyl acetate	—	3.36	3.09	3.01	2.87	3.01	2.90
151	*4-n-amylphenol	14938-35-3	4.19	3.65	2.86	3.64	2.86	4.21
152	4-tert-amylphenol	80-46-6	4.7	3.84	3.38	3.47	3.38	4.28
153	*carbutamide	339-43-5	2.29	2.82	3.47	2.24	3.47	2.28
154	pentobarbital	76-74-4	3.01	2.79	2.77	3.03	2.77	3.10
155	amobarbital	57-43-2	3.07	3.01	2.82	3.02	2.82	2.95
156	thiopental	76-75-5	3.28	3.4	3.24	2.92	3.24	2.91
157	dibenzofuran	132-64-9	2.97	2.77	2.88	2.77	2.88	2.88
158	dibenzothiophene	132-65-0	3.48	3.39	3.27	3.14	3.27	2.85
159	phenazine	92-82-0	2.41	2.69	2.8	2.51	2.8	2.65
160	thianthrene	92-85-3	3.57	3.82	3.45	3.15	3.45	3.54
161	*carbazole	86-74-8	2.44	3.01	3.09	2.64	3.09	2.83
162	*phenoxazine	135-67-1	2.69	2.85	2.94	2.75	2.94	2.51
163	phenothiazine	92-84-2	2.73	3.2	3.35	2.53	3.35	2.66

164	furosemide	54-31-9	1.78	2.47	2.26	1.78	2.26	1.89
165	phenobarbital	1950/6/6	3.22	2.5	2.83	2.86	2.83	3.22
166	*sulfisomidine	515-64-0	2.1	2.32	2.82	2.17	2.82	2.24
167	sulfamethomidine	3772-76-7	2.33	1.94	2.31	2.38	2.31	2.23
168	sulfadimethoxine	122-11-2	2.26	1.5	2.36	2.32	2.36	2.28
169	4-n-butylphenyl acetate	—	3.62	3.26	2.96	3.77	2.96	3.56
170	3-n-butylphenyl acetate	—	3.66	3.08	2.8	3.76	2.8	3.63
171	3-isobutylphenyl acetate	—	3.83	3.24	2.96	3.76	2.96	3.81
172	*4-tert-butylphenyl acetate	—	3.85	3.72	3.03	3.77	3.03	3.64
173	cyclobarbital	52-31-3	2.71	2.9	2.98	2.87	2.98	3.08
174	hexobarbital	56-29-1	3.08	3.02	2.94	2.88	2.94	3.05
175	1-adamantaneacetate	—	4.32	4.04	4.08	3.74	4.08	4.15
176	acridine	260-94-6	2.33	2.91	3.01	2.58	3.01	2.87
177	phenanthridine	229-87-8	2.57	2.82	2.94	2.58	2.94	2.83
178	xanthene	92-83-1	2.71	2.99	3.03	2.57	3.03	2.80
179	n-phenylanthranilic acid	91-40-7	2.89	2.85	2.76	2.81	2.76	2.93
180	*mephobarbital	115-38-8	3.16	2.53	2.95	3.12	2.95	3.02
181	4-n-amylphenyl acetate	—	3.8	3.35	3.06	3.67	3.06	3.67
182	flufenamic acid	530-78-9	3.1	2.75	2.51	3.04	2.51	2.91
183	meclofenamic acid	644-62-2	2.67	3.38	3.61	2.66	3.61	2.59
184	nitrazepam	146-22-5	1.97	1.97	2.22	1.97	2.22	2.38
185	flurbiprofen	51543-40-9	3.69	3.02	2.93	3.14	2.93	3.24
186	sulfaphenazole	526-08-9	2.35	2.17	2.63	2.34	2.63	2.43
187	*bendroflumethiazide	73-48-3	2.49	2.4	2.25	2.49	2.25	2.49
188	*mefenamic acid	61-68-7	2.49	2.4	2.96	2.8	2.96	2.77
189	acetohexamide	968-81-0	2.94	3.18	3.26	2.89	3.26	2.96
190	fludiazepam	3900-31-0	2.33	2.45	3.02	2.31	3.02	2.38
191	nimetazepam	2011-67-8	1.73	1.99	2.36	2.06	2.36	1.88
192	fenbufen	36330-85-5	2.63	3.19	2.34	2.92	2.34	2.72
193	ketoprofen	22071-15-4	2.85	2.77	2.75	2.99	2.75	2.58
194	*medazepam	2898/12/6	2.4	3.09	3.62	2.31	3.62	2.10
195	progabide	62666-20-0	2.53	2.98	2.62	2.59	2.62	2.41
196	griseofulvin	78739-00-1	1.47	1.56	2.69	1.5	2.69	1.68
197	tolnaftate	2398-96-1	3.83	3.38	3.99	3.77	3.99	3.70
198	prostacyclin	35121-78-9	2.94	3.7	2.47	2.97	2.47	2.99
199	triamcinolone	124-94-7	3.37	3.32	3.67	3.31	3.67	3.40
200	cortisone	1953/6/5	3.35	3.49	3.7	3.56	3.7	3.58
201	prednisolone	50-24-8	3.56	3.65	3.84	3.56	3.84	3.52
202	hydrocortisone	50-23-7	3.6	3.77	3.98	3.56	3.98	3.55
203	*corticosterone	50-22-6	3.85	3.89	3.86	3.61	3.86	3.63
204	dexamethasone	1950/2/2	3.65	3.63	3.87	3.59	3.87	3.58
205	betamethasone	378-44-9	3.73	3.82	3.79	3.59	3.79	3.57

206	paramethasone	53-33-8	3.4	3.59	3.55	3.59	3.55	3.58
207	cortisone-21-acetate	1950/4/4	3.62	3.45	3.27	3.62	3.27	3.64
208	prednisolone-21-acetate	52-21-1	3.76	3.63	3.45	3.56	3.45	3.74
209	*hydrocortisone-21-acetate	1950/3/3	3.51	3.69	3.66	3.55	3.66	3.52
210	fluocinoloneacetonide	67-73-2	3.48	2.97	3.39	3.49	3.39	3.55
211	triamcinolone acetonide	76-25-5	3.51	3.39	3.78	3.51	3.78	3.59
212	spironolactone	1952/1/7	4.44	3.79	4.72	4.43	4.72	4.20
213	dehydrocholic acid	81-23-2	3.38	3.39	3.6	3.54	3.6	3.36
214	chenodeoxycholic acid	474-25-9	4.36	4.74	4.72	4.43	4.72	4.29
215	ursodeoxycholic acid	128-13-2	4.51	–	4.76	4.43	4.76	4.32
216	*cholic acid	81-25-4	3.5	4.38	4.59	3.5	4.59	3.66
217	hydrocortisone-17-butyrate	13609-67-1	3.23	3.25	3.69	3.23	3.69	3.21
218	cinnarizine	298-57-7	3.64	3.71	2.81	3.64	2.81	3.57
219	*cycloheptanol	502-41-0	3.23	2.94	2.65	2.41	2.65	2.90
220	*2-methoxyethanol	109-86-4	0.22	0.58	0.805	1.52	0.805	0.46
221	3-hydroxycinnamic acid	14755-02-3	2.56	2.54	2.33	2.68	2.33	2.36
222	ethyl 4-hydroxybenzoate	120-47-8	3.01	2.49	2.28	2.72	2.28	2.50
223	ethyl 4-aminobenzoate	1994/9/7	2.69	2.81	2.42	2.52	2.42	2.66
224	4-methylcinnamic acid	1866-39-3	2.65	3.04	2.49	2.58	2.49	2.72
225	sulfadiazine	68-35-9	2.52	2.25	2.56	3	2.56	2.53
226	l-a-o-benzylglycerol	56552-80-8	2.11	3.22	2.22	2.88	2.22	2.82
227	sulfamerazine	127-79-7	1.97	2.37	2.48	2.12	2.48	2.46
228	butyl 4-hydroxybenzoate	94-26-8	3.39	2.86	2.57	3.23	2.57	3.18
229	butyl 4-aminobenzoate	94-25-7	3.19	3.14	2.76	3.53	2.76	3.39
230	*benzidine	92-87-5	3.35	3.54	3.1	2.64	3.1	3.61
231	*triflumizole	99387-89-0	2.66	2.6	2.66	2.65	2.66	2.60
232	diazepam	439-14-5	2.33	2.75	3.15	2.34	3.15	2.39
233	prostaglandine e2	363-24-6	3.09	2.91	2.81	3.09	2.81	3.02

[*]used for external test set; [1]PLS: the partial least square based method; [2]GRNN: the general neural network based method; [3]MLS: the multiple linear regressions based method; [4]LS-SVM: the least-squares support vector machine based method

Example of log*K* prediction

Prediction of log*K* of Methanol (CH_3OH) is exampled as follows. For this compound, the total number of group *N* is 6, and molecular weight $\boldsymbol{M}$ is 32.0419.

Firstly, the structure for the Methanol is drawn and optimized by using the software of heperchem.7.0.

Secondly, the distant matrix $\boldsymbol{M}_0$, $\boldsymbol{M}_1$, $\boldsymbol{M}_2$, $\boldsymbol{M}_3$, $\boldsymbol{M}_4$ and the extended matrix $\boldsymbol{M}_E$ is described as below.

$$M_0 = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 2 \\ 1 & 0 & 2 & 2 & 2 & 1 \\ 1 & 2 & 0 & 1 & 1 & 3 \\ 1 & 2 & 2 & 0 & 2 & 3 \\ 1 & 2 & 2 & 2 & 0 & 3 \\ 2 & 1 & 3 & 3 & 3 & 3 \end{bmatrix} \quad M_1 = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$M_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 2 & 2 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 0 & 2 & 0 \\ 0 & 2 & 2 & 2 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad M_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 3 & 3 & 3 & 3 \end{bmatrix}$$

$$M_4 = \begin{bmatrix} 0 & 1.433 & 1.0915 & 1.0915 & 1.0915 & 1.9277 \\ 1.433 & 0 & 2.0477 & 2.1089 & 2.1089 & 0.99115 \\ 1.0915 & 2.0477 & 0 & 1.7701 & 1.7701 & 2.831 \\ 1.0915 & 2.1089 & 1.7701 & 0 & 1.773 & 2.3163 \\ 1.0915 & 2.1089 & 1.7701 & 1.773 & 0 & 2.3163 \\ 1.9277 & 0.99115 & 2.831 & 2.3163 & 2.3163 & 0 \end{bmatrix}$$

$$M_E = \begin{bmatrix} 1.8 & 1.4 & 1.2 & 1.2 & 1.2 & 1.2 \\ 12 & 16 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$M_{D,0} = \begin{bmatrix} 3.2400 & 3.5200 & 3.1600 & 3.1600 & 3.1600 & 4.1600 \\ 3.5200 & 1.9600 & 3.6800 & 3.6800 & 3.6800 & 2.6800 \\ 3.1600 & 3.6800 & 1.4400 & 3.4400 & 3.4400 & 4.4400 \\ 3.1600 & 3.6800 & 3.4400 & 1.4400 & 3.4400 & 4.4400 \\ 3.1600 & 3.6800 & 3.4400 & 3.4400 & 1.4400 & 4.4400 \\ 4.1600 & 2.6800 & 4.4400 & 4.4400 & 4.4400 & 1.4400 \end{bmatrix}$$

$$M_{D,1} = \begin{bmatrix} 3.2400 & 3.5200 & 3.1600 & 3.1600 & 3.1600 & 2.1600 \\ 3.5200 & 1.9600 & 1.6800 & 1.6800 & 1.6800 & 2.6800 \\ 3.1600 & 1.6800 & 1.4400 & 1.4400 & 1.4400 & 1.4400 \\ 3.1600 & 1.6800 & 1.4400 & 1.4400 & 1.4400 & 1.4400 \\ 3.1600 & 1.6800 & 1.4400 & 1.4400 & 1.4400 & 1.4400 \\ 2.1600 & 2.6800 & 1.4400 & 1.4400 & 1.4400 & 1.4400 \end{bmatrix}$$

$$\boldsymbol{M}_{\mathrm{D},2}=\begin{bmatrix} 3.2400 & 2.5200 & 2.1600 & 2.1600 & 2.1600 & 4.1600 \\ 2.5200 & 1.9600 & 3.6800 & 3.6800 & 3.6800 & 1.6800 \\ 2.1600 & 3.6800 & 1.4400 & 3.4400 & 3.4400 & 1.4400 \\ 2.1600 & 3.6800 & 3.4400 & 1.4400 & 3.4400 & 1.4400 \\ 2.1600 & 3.6800 & 3.4400 & 3.4400 & 1.4400 & 1.4400 \\ 4.1600 & 1.6800 & 1.4400 & 1.4400 & 1.4400 & 1.4400 \end{bmatrix}$$

$$\boldsymbol{M}_{\mathrm{D},3}=\begin{bmatrix} 3.2400 & 2.5200 & 2.1600 & 2.1600 & 2.1600 & 2.1600 \\ 2.5200 & 1.9600 & 1.6800 & 1.6800 & 1.6800 & 1.6800 \\ 2.1600 & 1.6800 & 1.4400 & 1.4400 & 1.4400 & 4.4400 \\ 2.1600 & 1.6800 & 1.4400 & 1.4400 & 1.4400 & 4.4400 \\ 2.1600 & 1.6800 & 1.4400 & 1.4400 & 1.4400 & 4.4400 \\ 2.1600 & 1.6800 & 4.4400 & 4.4400 & 4.4400 & 1.4400 \end{bmatrix}$$

$$\boldsymbol{M}_{\mathrm{D},4}=\begin{bmatrix} 3.2400 & 3.9530 & 3.2515 & 3.2551 & 3.2551 & 4.0877 \\ 3.9530 & 1.9600 & 3.7277 & 3.7889 & 3.7889 & 2.6711 \\ 3.2515 & 3.7277 & 1.4400 & 3.2101 & 3.2101 & 4.2710 \\ 3.2551 & 3.7889 & 3.2101 & 1.4400 & 3.2130 & 3.7563 \\ 3.2551 & 3.7889 & 3.2101 & 3.2130 & 1.4400 & 3.7563 \\ 4.0877 & 2.6711 & 4.2710 & 3.7563 & 3.7563 & 1.4400 \end{bmatrix}$$

$$\boldsymbol{M}_{\mathrm{D},5}=\begin{bmatrix} 144 & 145 & 193 & 13 & 13 & 14 \\ 145 & 144 & 194 & 14 & 14 & 13 \\ 193 & 194 & 256 & 18 & 18 & 19 \\ 13 & 14 & 18 & 1 & 3 & 4 \\ 13 & 14 & 18 & 3 & 1 & 4 \\ 14 & 13 & 19 & 4 & 4 & 1 \end{bmatrix}$$

$$\boldsymbol{M}_{\mathrm{D},6}=\begin{bmatrix} 144 & 145 & 193 & 13 & 13 & 12 \\ 145 & 144 & 192 & 12 & 12 & 13 \\ 193 & 192 & 256 & 16 & 16 & 16 \\ 13 & 12 & 16 & 1 & 1 & 1 \\ 13 & 12 & 16 & 1 & 1 & 1 \\ 12 & 13 & 16 & 1 & 1 & 1 \end{bmatrix}$$

$$\boldsymbol{M}_{\mathrm{D},7}=\begin{bmatrix} 144 & 144 & 192 & 12 & 12 & 14 \\ 144 & 144 & 194 & 14 & 14 & 12 \\ 192 & 194 & 256 & 18 & 18 & 16 \\ 12 & 14 & 18 & 1 & 3 & 1 \\ 12 & 14 & 18 & 3 & 1 & 1 \\ 14 & 12 & 16 & 1 & 1 & 1 \end{bmatrix}$$

$$\boldsymbol{M}_{D,8}=\begin{bmatrix} 144 & 144 & 192 & 12 & 12 & 12 \\ 144 & 144 & 192 & 12 & 12 & 12 \\ 192 & 192 & 256 & 16 & 16 & 19 \\ 12 & 12 & 16 & 1 & 1 & 4 \\ 12 & 12 & 16 & 1 & 1 & 4 \\ 12 & 12 & 19 & 4 & 4 & 1 \end{bmatrix}$$

$$\boldsymbol{M}_{D,9}=\begin{bmatrix} 144.0000 & 145.4330 & 193.0915 & 13.0951 & 13.0951 & 13.9277 \\ 145.4330 & 144.0000 & 194.0477 & 14.1089 & 14.1089 & 12.9911 \\ 193.0915 & 194.0477 & 256.0000 & 17.7701 & 17.7701 & 18.8310 \\ 13.0951 & 14.1089 & 17.7701 & 1.0000 & 2.7730 & 3.3163 \\ 13.0951 & 14.1089 & 17.7701 & 2.7730 & 1.0000 & 3.3163 \\ 13.9277 & 12.9911 & 18.8310 & 3.3163 & 3.3163 & 1.0000 \end{bmatrix}$$

The norm values of the extended matrix are calculated and shown as in Table S2.

Table S2 Norm values of extended matrix

extended matrix	norm(M,fro)	extended matrix	norm(M,fro)
$\boldsymbol{M}_{D,0}$	20.66014	$\boldsymbol{M}_{D,5}$	406.32864
$\boldsymbol{M}_{D,1}$	13.14692	$\boldsymbol{M}_{D,6}$	405.52204
$\boldsymbol{M}_{D,2}$	13.42391	$\boldsymbol{M}_{D,7}$	404.63295
$\boldsymbol{M}_{D,3}$	11.97504	$\boldsymbol{M}_{D,8}$	404.32398
$\boldsymbol{M}_{D,4}$	20.12715	$\boldsymbol{M}_{D,9}$	406.73901

Lastly, from parameters in Table S2, the calculated log*K* value of Methanol is -0.439 while the experimental log*K* is -0.49.