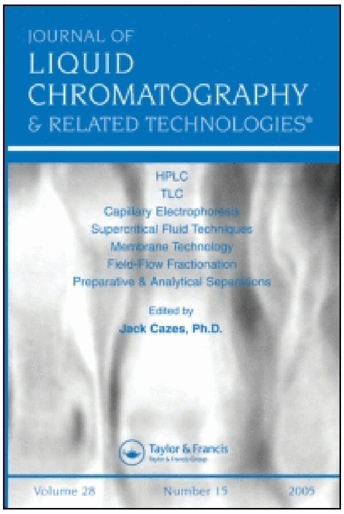

Journal of Liquid Chromatography & Related Technologies
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597273

COMPUTATIONAL CHEMICAL PREDICTION OF THE RETENTION FACTOR OF AROMATIC ACIDS

Toshihiko Hanai[a]; Hiroshi Homma[b]
[a] Health Research Foundation, Institut Pasteur 5F, Kyoto, Japan [b] School of Pharmaceutical Sciences, Kitasato University, Minato-ku, Tokyo, Japan

To cite this Article Hanai, Toshihiko and Homma, Hiroshi(2002) 'COMPUTATIONAL CHEMICAL PREDICTION OF THE RETENTION FACTOR OF AROMATIC ACIDS', Journal of Liquid Chromatography & Related Technologies, 25: 10, 1661 — 1676

To link to this Article: DOI: 10.1081/JLC-120005713

URL: http://dx.doi.org/10.1081/JLC-120005713

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016

J. LIQ. CHROM. & REL. TECHNOL., 25(10&11), 1661–1676 (2002)

COMPUTATIONAL CHEMICAL PREDICTION OF THE RETENTION FACTOR OF AROMATIC ACIDS

Toshihiko Hanai[1],* and Hiroshi Homma[2]

[1]Health Research Foundation, Institut Pasteur 5F,
Sakyo-ku, Kyoto, 606-8225, Japan
[2]School of Pharmaceutical Sciences, Kitasato University,
Minato-ku, Tokyo, 108-0072, Japan

ABSTRACT

The retention factors of aromatic acids were predicted from their log *P* and p*K*a values in reversed-phase liquid chromatography in eluent at a given pH. The log *P* value, a modified CAChe-log P^{TM} and *V*log P^{TM}, was calculated by computational chemical methods. The p*K*a values in this study were derived from partial charges of atoms by computational chemical methods instead of Hammett's equations. The precision of predicted retention factors was dependent on the accuracy of p*K*a values.

INTRODUCTION

Several approaches have been used to develop systems for the determination of quantitative structure-retention relationships (QSRR) in liquid

---

*Corresponding author. E-mail: thanai@attglobal.net

www.dekker.com

chromatography. These have included log *P*, van der Waals volume, ionization potential, delocalization energy, connectivity index, and free energy. The maximum retention factor (*k*m) was related to their octanol-water partition coefficient (log *P*) or van der Waals volume (VWV). Methods for prediction of retention factors using log *P* and VWV have advantages and disadvantages. The log *P* method worked well for groups of compounds such as non-ionizable compounds, phenolic compounds, aromatic acids, and nitrogen containing compounds.[1] However, one equation derived from Rekker's log *P* or Boder's log P^{TM} could not be applied for mixtures of different groups for prediction of retention factors with high precision. This disadvantage was less critical for a new log *P* calculation system suitable for reversed-phase liquid chromatography. The VWV method shows poor handling of isomers, but addition of ionization potential, for the calculation improved the precision.

Furthermore, the retention factors of ionized compounds in reversed-phase liquid chromatography were predicted from the following Eq. (1)

$$k = \frac{(k\mathrm{m} + k\mathrm{i}[\mathrm{H}^+]/[K\mathrm{a}])}{(1 + [\mathrm{H}^+]/[K\mathrm{a}])} \tag{1}$$

where *k*m is the maximum retention factor of the molecular form related to log *P* values or van der Waals volumes, *k*i is the minimum retention factor of 100% ionized form, and $[\mathrm{H}^+]$ is the hydrogen ion concentration. The dissociation constant, p*K*a can be calculated from Hammett's equation or measured by titration or liquid chromatography.

In this report, the prediction of the retention factor of aromatic acids in eluent at a given pH is described using Vlog P^{TM}, a modified CAChe-log P^{TM} (Nlog *P*), and p*K*a derived from atom partial charge calculated by the CAChe$^{\mathrm{TM}}$ program.

EXPERIMENTAL

The computers used for the calculations were an upgraded Macintosh 8100/250 and a Toshiba Dynabook 4000X. The software used was CAChe$^{\mathrm{TM}}$ and *V*log P^{TM} from Fujitsu (Tokyo). The chemical calculations were performed without modification of the program.[2] A modified CAChe-log P^{TM} calculation method was also used to study the performance in reversed-phase liquid chromatography. The Cricket-Graph$^{\mathrm{TM}}$ program from Computer Associates (San Diego, CA, USA) was used for data handling. The retention factors of aromatic acids were from literatures measured previously by reversed-phase liquid chromatography using an octadecyl-bonded silica gel,[3,4] polystyrene gel,[5] and polyvinylalcohol gel,[6] and were used to study the evaluation of computationally calculated values.

RESULTS AND DISCUSSION

Prediction of Maximum Retention Factor

The modified CAChe-log P^{TM} equation was:

$$
\begin{aligned}
N\log P &= -1.451 + 0.06687 \times (\mathrm{SAS}) + 0.004180 \times (\mathrm{HOF}) - 0.008687 \times \\
&(\mathrm{HOF}/\mathrm{H_2O}) + (-1.8193 \times (\mathrm{N1}) - 1.344 \times (\mathrm{OH}))/\surd((\mathrm{N1}) + (\mathrm{OH})) \\
&- 1.8193\surd(\mathrm{N_2}) - 0.937\surd(\mathrm{O}) - 1.981\surd(\mathrm{CO}) - 0.68 \times (\mathrm{NO_2}) - 2.444 \times \\
&(\mathrm{COOH})\mathrm{al} - 1.844 \times (\mathrm{COOH})\mathrm{ar},
\end{aligned}
$$

where SAS, HOF, and HOF/water are solvent accessible surface area, heat of formation in vacuum, and heat of formation in water, respectively, as calculated by ProjectLeader$^{\mathrm{TM}}$ (part of CAChe$^{\mathrm{TM}}$ program) using MOPAC,[2] and N, O, OH, CO, NO_2, and COOH are numbers of nitrogen, oxygen, hydroxyl, carbonyl, nitro, and carboxyl groups (allyphatic, aromatic), respectively.

The correlations between *N*log *P* and *V*log P^{TM} values were as follows:

$$N\log P = 1.103 \times V\log P + 0.091,\ r^2 = 0.847,\ (n = 62)$$

for all acids, and

$$N\log P = 1.142 \times V\log P + 0.091,\ r^2 = 0.948,\ (n = 29)$$

for benzoic acid derivatives as shown in Table 1.

The *N*log *P* calculation method was derived from *C*log P^{TM} and *M*log *P* data for phenolic and nitrogen-containing compounds, where *M*log *P* was the measured octanol-water partition coefficient.[7] The results indicated that *V*log P^{TM} may have selectivity for certain compounds compared to *M*log *P*.

The reliability of *N*log *P* and *V*log P^{TM} was analyzed with log *k* values of aromatic acids, measured on five different packing materials in reversed-phase liquid chromatography. These packing materials were two octadecyl-bonded silica gels (410ODS from Tosoh, Tokyo and ERC 1000 from ERMA, Tokyo), one octadecyl-bonded vinylalcohol copolymer (602ODP from Asahikasei, Tokyo), and two polystyrene gels (Hitachi Gel 3011 and 3013 from Hitachi, Tokyo). The retention factors of aromatic acids were measured at low pH, as their molecular

Table 1. Properties and Retention Factors of Aromatic Acids

						log *k*				
Aromatic Acid	*V*log*P*	*N*log*P*	p*K*a[1]	p*K*a[2]	p*K*a[3]	ODS[4]	ODP[5]	ODS[6]	3013[7]	3011[8]
Benzoic acid	1.485	1.952	4.200	4.19	4.20	0.574	0.5955	0.765	0.631	0.800
2-Methylbenzoic acid	1.757	2.272			3.91	0.824	0.8673			
3-Methylbenzoic acid	1.757	2.439			4.26	0.867	0.8833			
4-Methylbenzoic acid	2.020	2.452	4.379		4.34	0.847	0.8676	1.128		
2,4-Dimethylbenzoic acid	2.268	2.765			4.05	1.117	1.1329			
2,5-Dimethylbenzoic acid	2.197	2.754			3.97	1.117	1.1177			
2,6-Dimethylbenzoic acid	1.948	2.663			3.62	0.815	0.8805			
3,4-Dimethylbenzoic acid	2.312	2.830			4.40	1.082	1.1087			
3,5-Dimethylbenzoic acid	2.471	2.918			4.32	1.160	1.1714			
2,4,6-Trimethylbenzoic acid	2.532	3.157			3.76	1.116	0.9444			
4-Ethylbenzoic acid	2.461	2.883	4.388		4.35	1.157	1.1638	1.515		
2-Chlorobenzoic acid	2.158	2.571			2.92	0.710	0.7707			
3-Chlorobenzoic acid	2.334	2.699			3.83	0.993	1.1691			
4-Chlorobenzoic acid	2.354	2.700	3.894		3.96	1.010	1.1887	1.325		
2,4-Dichlorobenzoic acid	2.976	3.285			2.68	1.178	1.6724			
2,5-Dichlorobenzoic acid	2.971	3.284			2.55	1.110	0.6004			
2,6-Dichlorobenzoic acid	2.723	3.274			1.64	0.789	0.6523			
3,4-Dichlorobenzoic acid	3.012	3.315			3.50	1.371	1.6730			
3,5-Dichlorobenzoic acid	3.093	3.403			2.46	1.442	1.3143			
2-Bromobenzoic acid	2.227	2.916			2.85	0.774	0.8755			
3-Bromobenzoic acid	2.462	2.952			3.86	1.073	1.3073			
4-Bromobenzoic acid	2.510	2.953	3.864		3.98	1.096	1.1850	1.429		
2-Hydroxybenzoic acid	1.060	1.023	2.807	3.06	3.07			0.912	0.819	0.940

3-Hydroxybenzoic acid	1.075	1.113	4.093	4.16	4.07			0.246	0.180	0.370
4-Hydroxybenzoic acid	1.002	1.114	4.682	4.53	4.58			0.262	0.104	0.310
3,4-Dihydroxybenzoic acid	0.726	1.012	4.456	4.49	4.45			−0.181	−0.123	0.120
3,5-Dihydroxybenzoic acid	0.790	0.715	4.068	4.12	3.94			−0.242	−0.161	0.084
3,4,5-Trihydroxybenzoic acid	0.584	0.978	4.364	4.40	4.32			−0.495	−0.332	−0.022
3-Methoxybenzoic acid	1.708	1.971		4.09	4.09			0.883	0.669	0.910
4-Hydroxy-3-methoxybenzoic acid	1.238	1.100	4.474	4.48	4.47			0.148	0.133	0.390
Phenylacetic acid	1.367	1.929	4.334	4.29	4.30	0.597	0.3618	0.787	0.629	0.810
2-Methylphenylacetic acid	1.670	2.193			4.20	0.809	0.5714			
3-Methylphenylacetic acid	1.890	2.365	4.067		4.32	0.877	0.6312			
2-Chlorophenylacetic acid	2.055	2.519			3.75	0.841	0.6994			
4-Chlorophenylacetic acid	2.210	2.633	4.185		4.22	0.979	0.9028	1.270		
2-Hydroxyphenylacetic acid	1.060	0.922	4.225	4.14				0.382	0.254	0.440
4-Hydroxyphenylacetic acid	0.790	0.973	4.404	4.40	4.43			0.112	0.088	0.290
2,5-Dihydroxyphenylacetic acid	0.311	0.846	3.964	4.06				−0.254	−0.223	0.064
3,4-Dihydroxyphenylacetic acid	0.402	0.872	4.356	4.39	4.39			−0.156	−0.131	0.130
3-Methoxyphenylacetic acid	1.496	1.851	4.319	4.27	4.26			0.843	0.667	0.860
4-Methoxyphenylacetic acid	1.474	1.844	4.399	4.35	4.40			0.808	0.639	0.840
4-Hydroxy-3-methoxyphenyl-acetic acid	0.952	0.968	4.425	4.37	4.39			0.178	0.026	0.370
L(+)Mandelic acid	0.404	0.986		3.42	3.38	0.076	−0.2345	0.232		0.350
3,4-Dihydroxymandelic acid	−0.461	0.830			3.49					
3-Methoxymandelic acid	0.620	0.977	3.334	3.47	3.33			0.316	0.148	0.420
4-Hydroxy-3-methoxymandelic acid	0.110	0.883		3.55	3.50				−0.243	0.065
trans-Cinnamic acid	2.200	2.291	4.376	4.37	4.38	0.887	0.9169	1.165	0.929	1.100
4-Hydroxycinnamic acid	1.619	1.451	4.595	4.57	4.63			0.354	0.252	0.500
3,4-Dihydroxycinnamic acid	1.222	1.319	4.554	4.53	4.55			0.066	0.076	0.280

(*continued*)

Table 1. Continued

						log *k*				
Aromatic Acid	*V*log*P*	*N*log*P*	p*K*a[1]	p*K*a[2]	p*K*a[3]	ODS[4]	ODP[5]	ODS[6]	3013[7]	3011[8]
3-Methoxycinnamic acid	2.722	2.885			4.31	1.180	1.2169		0.970	1.190
4-Hydroxy-3-methoxycinnamic acid	1.772	1.437	4.570	4.56	4.56			0.444	0.290	0.610
3-Phenylpropionic acid	1.619	2.411	4.691		4.59	0.849	0.6320	1.108		
4-Phenylbutyric acid	2.196	2.976			4.72	1.093	0.8650	1.418		
Hyppuric acid	0.711	0.358		3.58	4.30	0.031	−0.3839		0.083	0.320
2-Hydroxy-hyppuric acid	0.318	−0.571	3.506	3.58				0.434	0.318	0.520
Indole-3-acetic acid	1.732	0.974	4.590	4.65	3.80	0.580	0.6524	0.785	0.613	0.800
5-Hydroxy-indole-3-acetic acid	1.091	0.846		4.57					0.022	0.260
Indolepropionic acid	2.092	1.631		4.81	5.01	0.827	0.9407		0.836	1.040
Indolebutylic acid	2.522	2.110	4.781							
2-Naphtoic acid	2.699	3.148		4.25	4.24				1.336	1.490
3-Hydroxy-2-naphtoic acid	2.161	2.184		2.89					1.501	1.590
Uric acid	−0.875	−1.990								−0.250

[1] p*K*a by liquid chromatography,[11] [2] p*K*a by titration,[5] [3] p*K*a by Hammett's equation,[10] [4] ERC1000,[4] [5] 602ODP,[6] [6] 410ODS,[3] [7] Hitachi 3013,[5] [8] Hitachi 3011.[5]

forms were correlated with *V*log P^{TM} and *N*log *P* values. The relationships between *N*log *P* or *V*log P^{TM} and log *k* values of different groups of acids varied on these packing materials.

The retention times of aromatic acids collected from literatures were measured under different conditions at different locations, but the selectivity was the same under these conditions. The selectivity of packing materials was analyzed from the chromatographic behavior of aromatic acids, based on their dissociation constants collected in Table 1. 2-Hydroxybenzoic acid was eliminated from the calculation due to low p*K*a value. The correlations between their log *P* and retention factors were analyzed from the slope and interception of the following equations.

The relationships on ERC1000 were:

$\log k$ (benzoic acids) $= 0.482\ (V\log P) - 0.063,\ r^2 = 0.891,\ n = 10$
$\log k$ (phenylacetic acids) $= 0.468\ (V\log P) - 0.022,\ r^2 = 0.846,\ n = 6,$
$\log k$ (benzoic acids) $= 0.502\ (N\log P) - 0.353,\ r^2 = 0.904,\ n = 10,$
$\log k$ (phenylacetic acids) $= 0.409\ (N\log P) - 0.121,\ r^2 = 0.945,\ n = 6.$

Those on 602ODP were:

$\log k$ (benzoic acids) $= 0.459\ (V\log P) - 0.012,\ r^2 = 0.750,\ n = 10,$
$\log k$ (phenylacetic acids) $= 0.591\ (V\log P) - 0.449,\ r^2 = 0.954,\ n = 6,$
$\log k$ (benzoic acids) $= 0.436\ (N\log P) - 0.152,\ r^2 = 0.634,\ n = 10,$
$\log k$ (phenylacetic acids) $= 0.511\ (N\log P) - 0.574,\ r^2 = 0.864,\ n = 6.$

The relationships on 410ODS were:

$\log k$ (benzoic acids) $= 1.198\ (V\log P) - 1.116,\ r^2 = 0.929,\ n = 8,$
$\log k$ (phenylacetic acids) $= 0.942\ (V\log P) - 0.587,\ r^2 = 0.976,\ n = 8,$
$\log k$ (cinnamic acids) $= 1.117\ (V\log P) - 1.396,\ r^2 = 0.934,\ n = 4,$
$\log k$ (benzoic acids) $= 0.939\ (N\log P) - 0.995,\ r^2 = 0.813,\ n = 8,$
$\log k$ (phenylacetic acids) $= 0.809\ (N\log P) - 0.694,\ r^2 = 0.842,\ n = 8,$
$\log k$ (cinnamic acids) $= 1.015\ (N\log P) - 1.141,\ r^2 = 0.948,\ n = 4.$

Those on Hitachi 3013 were:

$\log k$ (benzoic acids) $= 0.906\ (V\log P) - 0.838,\ r^2 = 0.947,\ n = 8,$
$\log k$ (phenylacetic acids) $= 0.759\ (V\log P) - 0.501,\ r^2 = 0.938,\ n = 8,$
$\log k$ (cinnamic acids) $= 0.864\ (V\log P) - 1.086,\ r^2 = 0.876,\ n = 4,$
$\log k$ (benzoic acids) $= 0.724\ (N\log P) - 0.764,\ r^2 = 0.862,\ n = 8,$
$\log k$ (phenylacetic acids) $= 0.679\ (N\log P) - 0.623,\ r^2 = 0.879,\ n = 8,$
$\log k$ (cinnamic acids) $= 0.826\ (N\log P) - 0.955,\ r^2 = 0.983,\ n = 4.$

Those on Hitachi 3011 were:

$$\log k \text{ (benzoic acids)} = 0.849\ (V\log P) - 0.543,\ r^2 = 0.967,\ n = 8,$$
$$\log k \text{ (phenylacetic acids)} = 0.686\ (V\log P) - 0.198,\ r^2 = 0.959,\ n = 8,$$
$$\log k \text{ (cinnamic acids)} = 0.839\ (V\log P) - 0.807,\ r^2 = 0.958,\ n = 4,$$
$$\log k \text{ (benzoic acids)} = 0.682\ (N\log P) - 0.478,\ r^2 = 0.890,\ n = 8,$$
$$\log k \text{ (phenylacetic acids)} = 0.612\ (N\log P) - 0.305,\ r^2 = 0.894,\ n = 8,$$
$$\log k \text{ (cinnamic acids)} = 0.741\ (N\log P) - 0.581,\ r^2 = 0.917,\ n = 4.$$

Octadecyl-bonded polyvinylalcohol gel, 602ODP, is stable in buffer solution, and its capability as a standard column for QSRR was examined. *V*log *P* and *N*log *P* demonstrated different behavior, and the log *k* value of phenylacetic acid was smaller than that of benzoic acid. Generally, the log *k* value of phenylacetic acid was larger than that of benzoic acid in reversed-phase liquid chromatography.[8] The comparison of log *k* values measured on ERC1000 and 602ODP indicated that the chromatographic behavior on 602ODP was exceptional, and this octadecyl-bonded polyvinylalcohol gel may not be suitable for QSRR in reversed-phase liquid chromatography despite its stability in buffer solution.

Polystyrene gel is also stable in buffer solution. The first study of QSRR was carried out using 10 μm μStyragel from Waters (Milford, MA), and the possibility of a prediction of retention factor was demonstrated.[7] No significant difference was found between an octadecyl-bonded silica gel (410ODS) and polystyrene gels (Hitachi 3011 and 3013) based on log *k* values of aromatic acids. Both *N*log *P* and *V*log *P* can be used for QSRR, but the error of predicted log *k* values of aromatic acids may be larger than that of phenolic compounds.[9]

Prediction of Dissociation Constant

Dissociation constants can be predicted using Hammett's equations $pKa = A + B\Sigma\sigma$ where A and B are constant and σ is Hammett's σ constant.[10] A is the p*K*a value of non-substituted compounds and B has to be selected for each series of compounds. There were no standard equations for A and B values. The reference values of dissociation constants were used to predict dissociation constants from atom partial charges calculated by the MM/AM1 program (part of CAChe™).

The calculated partial charges of the carboxyl group hydrogens are listed in Table 2. The correlation of partial charges between groups of acids was calculated to study the usefulness of partial charges instead of Hammett's σ constant for prediction of p*K*a values.

Table 2. Partial Charge and Predicted p*K*a

	Partial Charge				p*K*a[1]				p*K*a[2]			
Substitutes	BA	PhA	MA	CA	BA	PhA	MA	CA	BA	PhA	MA	CA
H	0.2455	0.2447	0.2476	0.2436	4.293	4.291	3.681	4.523	4.197	4.291	3.681	4.523
2-Methyl-	0.2448	0.2427	0.2468	0.2433	4.389	4.569	3.786	4.566	4.305	4.601	3.800	4.570
3-Methyl-	0.2451	0.2426	0.2472	0.2434	4.348	4.583	3.734	4.551	4.259	4.617	3.741	4.554
4-Methyl-	0.2450	0.2425	0.2472	0.2432	4.361	4.596	3.734	4.580	4.274	4.632	3.741	4.586
2,4-Dimethyl-	0.2442	0.2422	0.2471	0.2436	4.471	4.638	3.747	4.523	4.398	4.679	3.756	4.523
2,5-Dimethyl-	0.2444	0.2423	0.2478	0.2436	4.443	4.624	3.655	4.523	4.367	4.663	3.651	4.523
2,6-Dimethyl-	0.2416	0.2424	0.2446	0.2435	4.826	4.610	4.075	4.537	4.799	4.648	4.129	4.539
3,4-Dimethyl-	0.2446	0.2427	0.2467	0.2430	4.416	4.569	3.799	4.608	4.336	4.601	3.815	4.617
3,5-Dimethyl-	0.2447	0.2422	0.2468	0.2432	4.402	4.638	3.786	4.580	4.320	4.679	3.800	4.586
2,4,6-Trimethyl-	0.2411	0.2424	0.2443	0.2431	4.895	4.610	4.114	4.594	4.876	4.648	4.174	4.602
2-Ethyl-	0.2448	0.2428	0.2470	0.2433	4.389	4.555	3.760	4.566	4.305	4.586	3.771	4.570
3-Ethyl-	0.2451	0.2428	0.2471	0.2434	4.348	4.555	3.747	4.551	4.259	4.586	3.756	4.554
4-Ethyl-	0.2449	0.2425	0.2471	0.2432	4.375	4.596	3.747	4.580	4.290	4.632	3.756	4.586
2-Chloro-	0.2453	0.2438	0.2474	0.2444	4.320	4.416	3.707	4.410	4.228	4.431	3.711	4.397
3-Chloro-	0.2480	0.2447	0.2498	0.2447	3.951	4.291	3.392	4.367	3.811	4.291	3.353	4.350
4-Chloro-	0.2476	0.2448	0.2493	0.2448	4.006	4.277	3.458	4.353	3.873	4.275	3.427	4.334
2,4-Dichloro-	0.2472	0.2456	0.2490	0.2455	4.061	4.166	3.497	4.254	3.935	4.151	3.472	4.224
2,5-Dichloro-	0.2466	0.2459	0.2489	0.2453	4.143	4.124	3.510	4.282	4.027	4.105	3.487	4.255
2,6-Dichloro-	0.2459	0.2438	0.2478	0.2449	4.238	4.416	3.655	4.339	4.135	4.431	3.651	4.318
3,4-Dichloro-	0.2498	0.2455	0.2511	0.2458	3.705	4.180	3.222	4.211	3.534	4.167	3.159	4.177
3,5-Dichloro-	0.2502	0.2459	0.2511	0.2461	3.650	4.124	3.222	4.169	3.472	4.105	3.159	4.130
2-Bromo-	0.2446	0.2435	0.2470	0.2446	4.416	4.458	3.760	4.381	4.336	4.477	3.771	4.366
3-Bromo-	0.2484	0.2450	0.2502	0.2461	3.896	4.249	3.340	4.169	3.750	4.244	3.293	4.130

(*continued*)

Table 2. Continued

	Partial Charge				p*K*a[1]				p*K*a[2]			
Substitutes	BA	PhA	MA	CA	BA	PhA	MA	CA	BA	PhA	MA	CA
4-Bromo-	0.2481	0.2447	0.2497	0.2451	3.937	4.291	3.405	4.310	3.796	4.291	3.368	4.287
2-Hydroxy-	0.2569	0.2471	0.2518	0.2421	2.734	3.958	3.130	4.736	2.438	3.919	3.054	4.759
3-Hydroxy-	0.2467	0.2443	0.2486	0.2446	4.129	4.347	3.550	4.381	4.012	4.353	3.532	4.366
4-Hydroxy-	0.2456	0.2429	0.2477	0.2432	4.279	4.541	3.668	4.580	4.182	4.570	3.666	4.586
2,4-Dihydroxy-	0.2438	0.2417	0.2515	0.2416	4.525	4.707	3.169	4.806	4.459	4.756	3.099	4.838
2,5-Dihydroxy-	0.2443	0.2424	0.2522	0.2428	4.457	4.610	3.077	4.636	4.382	4.648	2.994	4.649
2,6-Dihydroxy-	0.2508	0.2447	0.2547	0.2439	3.568	4.291	2.749	4.480	3.379	4.291	2.621	4.476
3,4-Dihydroxy-	0.2466	0.2443	0.2486	0.2441	4.143	4.347	3.550	4.452	4.027	4.353	3.532	4.444
3,5-Dihydroxy-	0.2478	0.2445	0.2495	0.2445	3.978	4.319	3.432	4.395	3.842	4.322	3.397	4.381
3,4,5-Trihydroxy-	0.2477	0.2446	0.2496	0.2445	3.992	4.305	3.418	4.395	3.858	4.307	3.383	4.381
2-Methoxy-	0.2430	0.2413	0.2456	0.2418	4.635	4.763	3.944	4.778	4.583	4.818	3.979	4.806
3-Methoxy-	0.2458	0.2437	0.2483	0.2441	4.252	4.430	3.589	4.452	4.151	4.446	3.577	4.444
4-Methoxy-	0.2449	0.2423	0.2468	0.2429	4.375	4.624	3.786	4.622	4.290	4.663	3.800	4.633
4-Hydroxy-3-methoxy-	0.2457	0.2436	0.2480	0.2436	4.266	4.444	3.628	4.523	4.166	4.462	3.621	4.523

[1] Based on p*K*a by liquid chromatography.[11]

[2] Based on p*K*a by Hammett's equation.[10]

$$\begin{aligned}
&\mathrm{pc(PhA)} = 0.419 \times \mathrm{pc(BA)} + 0.141,\ r^2 = 0.711,\ n = 37,\\
&\mathrm{pc(MA)} = 0.569 \times \mathrm{pc(BA)} + 0.108,\ r^2 = 0.546,\ n = 37,\\
&\mathrm{pc(CA)} = 0.131 \times \mathrm{pc(BA)} + 0.212,\ r^2 = 0.119,\ n = 37,\\
&\mathrm{pc(MA)} = 0.908 \times \mathrm{pc(PhA)} + 0.027,\ r^2 = 0.342,\ n = 37,\\
&\mathrm{pc(CA)} = 0.509 \times \mathrm{pc(PhA)} + 0.120,\ r^2 = 0.444,\ n = 37,\\
&\mathrm{pc(CA)} = 0.128 \times \mathrm{pc(MA)} + 0.212,\ r^2 = 0.068,\ n = 37,
\end{aligned}$$

where, pc, PhA, BA, MA, and CA are partial charges of phenylacetic, benzoic, mandelic, and cinnamic acid derivatives, respectively. The correlations were poor, and, therefore, further calculations were performed without ortho derivatives. Finally, excellent correlation coefficients were obtained as follows:

$$\begin{aligned}
&\mathrm{pc(PhA)} = 0.621 \times \mathrm{pc(BA)} + 0.091,\ r^2 = 0.846,\ n = 21,\\
&\mathrm{pc(MA)} = 0.828 \times \mathrm{pc(BA)} + 0.044,\ r^2 = 0.983,\ n = 21,\\
&\mathrm{pc(CA)} = 0.535 \times \mathrm{pc(BA)} + 0.112,\ r^2 = 0.950,\ n = 21,\\
&\mathrm{pc(PhA)} = 0.651 \times \mathrm{pc(BA)} + 0.083,\ r^2 = 0.937,\ n = 20^*,\\
&\mathrm{pc(MA)} = 1.215 \times \mathrm{pc(PhA)} + 0.048,\ r^2 = 0.956,\ n = 20^*,\\
&\mathrm{pc(CA)} = 0.789 \times \mathrm{pc(PhA)} + 0.052,\ r^2 = 0.928,\ n = 20^*.
\end{aligned}$$

The partial charge of phenylacetic acid was exceptionally high. Therefore, some calculations were performed without phenylacetic acid (*). These results indicated that partial charge could be used for prediction of p*K*a values instead of Hammett's σ constant.

The correlation between p*K*a values measured by liquid chromatography and partial charges of non-substituted aromatic acids including phenol was:

$$\mathrm{p}K\mathrm{a} = -210.485 \times (\text{partial charge}) + 55.797,\ r^2 = 0.994,\ (n = 8) \qquad (2)$$

where the reference compounds used for this calculation were benzoic acid, phenylacetic acid, 3-phenylpropionic acid, mandelic acid, *trans*-cinnamic acid, indole-3-acetic acid, indole-3-butyric acid, and phenol in Table 1. The constant A of Hammett's equation can be calculated from Eq. (2).

The correlation between reference Δp*K*a values measured by liquid chromatography in Table 1 and Δpartial charge of substituted benzoic acids was:

$$\Delta\mathrm{p}K\mathrm{a} = -136.749 \times (\Delta\text{partial charge}) + 0.170,\ r^2 = 0.831,\ (n = 13) \qquad (3)$$

where Δpartial charge = partial charge (benzoic acid derivatives) − partial charge (benzoic acid).

The correlation between reference ΔpKa values and absolute values of partial charge of substituted phenols[7] was:

$$\Delta pKa = -210.429 \times (\Delta\text{partial charge}) + 0.228, r^2 = 0.894, (n = 42). \tag{4}$$

There are insufficient numbers of dissociation constants available to allow derivation of equations such as Eqs 3 and 4 for other aromatic acids. Therefore, $B\Sigma\sigma$ of Hammett's equation was obtained from Eqs 3 and 4:

$$B\Sigma\sigma = (-12.412 \times pKa - 85.573) \times \Delta(\text{parital charge of substituent}) \tag{5}$$

where B is $(-12.412 \times \text{pKa} - 85.573)$ and $\Sigma\sigma$ is Δ(partial charge of substituent). The predicted pKa values (pKa^1) from Eqs 2 and 5 are given in Table 2.

The correlation between predicted and reference pKa values of benzoic acid derivatives was:

$$pKa(\text{predicted}) = 0.937 \times pKa(\text{LC}) - 0.057, r^2 = 0.831, (n = 13).$$

The correlation between ΔpKa values from Hammett's equations and δpartial charge of substituted benzoic acids, except ortho-substituted compounds and 4-hydroxy and 4-methoxybenzoic acids, was:

$$\Delta pKa = -154.890 \times (\Delta\text{partial charge}) + 0.029, r^2 = 0.971, (n = 15) \tag{6}$$

where Δpartial charge = partial charge (benzoic acid derivatives) − partial charge (benzoic acid).

There are insufficient numbers of dissociation constants available to allow derivation of equations such as Eqs 3 and 4 for other aromatic acids. Therefore, $B\Sigma\sigma$ of Hammett's equation was obtained from Eqs 3 and 4:

$$B\Sigma\sigma = (-9.356 \times pKa - 116.314) \times \Delta(\text{parital charge of substitute}) \tag{7}$$

where B is $(-9.356 \times pKa - 116.314)$ and $\Sigma\sigma$ is Δ(partial charge of substituent). The predicted pKa values (pKa^2) from Eqs 3 and 7 are given in Table 2 with partial charges.

The ortho-effect of pKa values of ortho-substituted benzoic acids was 0.6177 and 1.3558 unit for a methyl group and a halogen atom, respectively. Therefore, pKa of benzoic acid derivatives can be predicted from Eqs 2 and 3 and ortho-effects. The correlation between predicted and reference pKa values for benzoic acid derivatives was as follows:

$$pKa(\text{pred}) = 0.991 \times pKa\text{ref} - 0.017, r^2 = 0.982, (n = 22).$$

Prediction of Retention Factor

The retention factors of aromatic acid derivatives listed in Table 3 were predicted from $N\log P$ and atom partial charge using the equations. The measured data are from original data for Ref. 11. First, $\log k$ of these acids at low pH, molecular from, were predicted from their $N\log P$ values listed in Table 1. The standard acids used are 3,5-dihydroxy-, 4-hydroxy, 4-methyl, 4-ethyl benzoic acids, and benzoic acid. Then the retention factors in a given pH eluent were predicted using Eq. (1), with the dissociation constants listed in Table 2 that were derived from partial charge and pKa measured by liquid chromatography. The precision of predicting retention factors of 100% ionized form aromatic acids was poor, due to the very short retention time influenced by the void volume of system. The real void volume is also affected by molecular shape and size of the analyte. The precision of void volume measurement can be improved in a preparative scale liquid chromatography, but the measurement is not practical in analytical scale liquid chromatography.[1] Therefore, the measured retention factors at a high pH were also used to study the practical use of this new QSRR system.

As listed in Table 3, the error between predicted and measured retention factors was large for 2-hydroxy, 4-bromo, and 4-methoxybenzoic acid and 3-methoxy cinnamic acid. The correlation coefficients were calculated between predicted and measured retention factors, excluding these four acids.

pH 2.54 : predicted $k = 0.935(\text{measured } k) - 0.315$, $r^2 = 0.977$ ($n = 27$),
pH 3.55 : predicted $k = 0.885(\text{measured } k) - 0.207$, $r^2 = 0.980$ ($n = 27$),
pH 4.05 : predicted $k = 0.815(\text{measured } k) - 0.127$, $r^2 = 0.984$ ($n = 27$),
pH 4.56 : predicted $k = 0.752(\text{measured } k) - 0.038$, $r^2 = 0.988$ ($n = 27$),
pH 5.05 : predicted $k = 0.656(\text{measured } k) + 0.029$, $r^2 = 0.985$ ($n = 27$),
pH 5.55 : predicted $k = 0.768(\text{measured } k) + 0.024$, $r^2 = 0.987$ ($n = 27$).

The retention factors were predicted, as ki was zero due to difficulty of ki prediction, the correlation coefficients were:

pH 3.55 : predicted $k = 0.876(\text{measured } k) - 0.225$, $r^2 = 0.979$ ($n = 27$),
pH 4.05 : predicted $k = 0.791(\text{measured } k) - 0.164$, $r^2 = 0.982$ ($n = 27$),
pH 4.56 : predicted $k = 0.688(\text{measured } k) - 0.101$, $r^2 = 0.984$ ($n = 27$),
pH 5.05 : predicted $k = 0.513(\text{measured } k) - 0.048$, $r^2 = 0.981$ ($n = 27$),
pH 5.55 : predicted $k = 0.400(\text{measured } k) - 0.032$, $r^2 = 0.961$ ($n = 27$).

These results suggest, that this new prediction method of retention factor is an advanced, simple method, compared with the old method using the empirical calculation methods of Rekker's $\log P$ and Hammett's pKa. The concentration effect of organic modifier on pKa[12–14] may improve the precision of predicted retention factor, however, pH measurement methods should be standardized.

Table 3. Measured and Predicted

		Measured k pH				
Aromatic Acid	k_i	3.55	4.05	4.56	5.05	5.55
Mandelic acid	0.2062	1.026	0.634	0.381	0.278	0.213
3-Methoxymandelic acid	0.2938	1.232	0.768	0.490	0.388	0.327
3,4,5-Trihydroxybenzoic acid	0.0103	0.300	0.247	0.170	0.113	0.046
3,4-Dihydroxybenzoic acid	0.0309	0.593	0.536	0.407	0.253	0.113
3,5-Dihydroxybenzoic acid	0.0516	0.496	0.397	0.237	0.144	0.077
2,4-Dihydtoxybenzoic acid	0.1392	0.850	0.459	0.263	0.186	0.155
4-Hydroxybenzoic acid	0.0722	1.124	1.052	0.814	0.521	0.217
3-Hydroxybenzoic acid	0.1237	1.541	1.237	0.758	0.423	0.201
4-Hydroxy-3-methoxybenzoic acid	0.0825	1.325	1.191	0.866	0.536	0.242
Benzoic acid	0.3041	5.216	4.304	2.691	1.552	0.675
3-Methoxybenzoic acid	0.4278	6.655	5.222	3.031	1.639	0.789
4-Methoxybenzoic acid	0.6495	2.469	1.402	0.912	0.758	0.687
2-Hydroxybenzoic acid	1.1600	17.220	12.560	6.794	3.691	1.923
4-Chlorobenzoic acid	1.4540	21.570	15.550	8.330	4.546	2.366
4-Bromobenzoic acid	0.4227	6.515	5.825	4.201	2.598	1.180
4-Methylbenzoic acid	0.6804	12.550	10.970	7.479	4.351	1.974
4-Etthylbenzoic acid	1.6290	30.510	26.730	18.240	10.950	4.814
3,4-Dihydroxyphenylacetic acid	0.0516	0.634	0.562	0.402	0.227	0.113
2,5-Dihydroxyphenylacetic acid	0.1237	0.423	0.376	0.263	0.186	0.134
4-Hydroxyphenylacetic	0.0928	1.196	1.062	0.758	0.464	0.191
4-Hydroxy-3-Methoxyphenyl-acetic acid	0.1031	1.407	1.247	0.890	0.552	0.247
Phenylacetic acid	0.3866	5.686	4.902	3.320	1.933	0.907
2-Hydroxyphenylacetic acid	0.5617	2.211	1.918	1.419	0.969	0.691
3-Methoxyphenylacetic acid	0.4845	6.438	5.541	3.684	2.227	1.057
4-Methoxyphenylacetic acid	0.4536	6.046	5.309	3.708	2.242	1.057
4-Chlorophenylacetic acid	1.2940	16.770	13.850	8.629	4.876	2.469
3,4-Dihydroxycinnamic acid	0.0894	1.098	1.010	0.768	0.510	0.242
4-Hydroxycinnamic acid	0.1546	2.165	1.995	1.521	0.995	0.485
Cinnamic acid	0.9742	13.730	11.970	8.258	4.892	2.356
4-Hydroxy-3-methoxycinnamic acid	0.1959	2.660	2.443	1.838	1.186	0.577

Retention Factors of Aromatic Acids

	Predicted k Using k_i pH					Predicted k Using $k_i = 0$ pH				
pH 2.54	3.55	4.05	4.56	5.05	5.55	3.55	4.05	4.56	5.05	5.55
0.948	0.633	0.428	0.293	0.237	0.216	0.545	0.284	0.111	0.039	0.013
0.932	0.627	0.458	0.355	0.315	0.301	0.487	0.240	0.090	0.031	0.010
0.934	0.689	0.441	0.207	0.085	0.035	0.686	0.436	0.199	0.075	0.025
0.995	0.799	0.564	0.298	0.137	0.067	0.792	0.550	0.275	0.110	0.037
0.574	0.432	0.291	0.160	0.092	0.065	0.418	0.263	0.119	0.045	0.015
0.732	0.675	0.583	0.424	0.275	0.190	0.662	0.548	0.351	0.168	0.063
1.201	1.023	0.782	0.460	0.236	0.130	1.012	0.755	0.413	0.174	0.061
1.199	0.974	0.710	0.414	0.239	0.163	0.949	0.654	0.324	0.128	0.044
1.170	0.995	0.759	0.449	0.236	0.136	0.982	0.728	0.394	0.165	0.058
5.655	4.836	3.709	2.182	1.101	0.585	4.790	3.599	1.985	0.842	0.297
5.858	4.958	3.763	2.218	1.174	0.688	4.887	3.598	1.932	0.805	0.281
5.793	5.123	4.141	2.682	1.547	0.972	5.039	3.932	2.289	1.011	0.363
1.014	1.141	1.153	1.158	1.159	1.160	0.135	0.047	0.015	0.005	0.002
22.548	17.080	11.468	6.059	3.203	2.041	16.703	10.703	4.922	1.869	0.626
35.975	25.649	15.909	7.267	2.969	1.269	25.526	15.670	6.926	2.576	0.857
14.255	12.438	9.799	5.939	2.986	1.505	12.347	9.576	5.522	2.422	0.866
31.626	27.722	21.992	13.481	6.863	3.508	27.510	21.469	12.495	5.518	1.981
0.768	0.669	0.528	0.324	0.170	0.094	0.662	0.510	0.292	0.127	0.045
0.732	0.683	0.600	0.445	0.286	0.186	0.673	0.574	0.387	0.195	0.075
0.925	0.848	0.722	0.500	0.290	0.167	0.840	0.700	0.453	0.219	0.083
0.917	0.825	0.683	0.456	0.265	0.162	0.813	0.653	0.398	0.182	0.067
5.420	4.646	3.584	2.148	1.133	0.649	4.587	3.443	1.897	0.804	0.283
0.842	0.763	0.687	0.618	0.583	0.569	0.605	0.377	0.168	0.063	0.021
4.692	4.202	3.454	2.276	1.299	0.781	4.145	3.312	1.997	0.908	0.331
4.632	4.307	3.752	2.696	1.593	0.896	4.271	3.656	2.486	1.263	0.491
19.920	16.980	12.987	7.676	3.982	2.237	16.775	12.506	6.825	2.875	1.009
1.755	1.569	1.282	0.819	0.425	0.212	1.559	1.257	0.769	0.354	0.130
2.240	2.062	1.764	1.221	0.682	0.356	2.048	1.729	1.146	0.567	0.217
10.585	9.660	8.165	5.575	3.176	1.800	9.567	7.920	5.067	2.425	0.909
2.182	1.991	1.682	1.147	0.651	0.367	1.972	1.633	1.045	0.500	0.187

Furthermore, artificial neural networks[15] combined with computational chemical calculation, will improve the precision of predicted molecular properties. *V*log *P*^{TM} may be calculated by neural network calculation.

CONCLUSION

The CAChe-log *P*^{TM} calculation method was modified by the computational chemical methods. The p*K*a values were derived from partial charge of atom by computational chemical methods instead of Hammett's equations. The retention factors of aromatic acids were predicted from these new log *P* and p*K*a values in reversed-phase liquid chromatography in a given pH eluent. The precision of predicted retention factors depended on the accuracy of p*K*a values. This prediction method of retention factor seems an advanced method, however more experimental data are required to study for other compounds.

REFERENCES

1. Hanai, T. Ed.; *HPLC, A Practical Guide*; The Royal Society of Chemistry: Cambridge, 1999.
2. *Fujitsu CAChe^{TM} Manual*, 1998.
3. Hanai, T.; Hubert, J. J. Chromatogr. **1984**, *316*, 261–265.
4. Arai, Y.; Yamaguchi, J.; Hanai, T. J. Chromatogr. **1987**, *400*, 21–26.
5. Hanai, T.; Tran, K.C.; Hubert, J. J. Chromatogr. **1982**, *239*, 385–395.
6. Arai, Y.; Hirukawa, Y.M.; Hanai, T. Nippon Kagaku Kaishi **1986**, *1986*, 969–975.
7. Hanai, T.; Koizumi, T.K.; Kinoshita, T. J. Liq. Chrom. & Rel. Technol. **2000**, *23*, 363–385.
8. Hanai, T.; Tran, K.C.; Hubert, J. J. Chromatogr. **1982**, *239*, 385–395.
9. Hanai, T. Chromatographia **1979**, *12*, 77–82.
10. Perrin, D.D., Dempsey, B., Serjeant, E.P., Eds.; *pKa Prediction for Organic Acids and Bases*; Chapman and Hall: London, 1981.
11. Hanai, T.; Hubert, J.; J., HERC & CC **1984**, *7*, 524–528.
12. Bosch, E.; Bou, P.; Allemann, H.; Roses, M. Anal. Chem. **1996**, *68*, 3651–3657.
13. Bosch, E.; Espinosa, S.; Roses, M. J. Chromatogr. A **1998**, *824*, 137–146.
14. Roses, M.; Rived, F.; Bosch, E. J. Chromatogr. A **2000**, *867*, 45–56.
15. Loukas, Y.L. J. Chromatogr. A **2000**, *904*, 119–129.

Received January 18, 2002
Accepted February 15, 2002
Manuscript 5727