

Empirical Procedure for the Calculation of Ionization Constants of Organic Compounds in Water from Their Molecular Volume

Yu. E. Zevatskii and S. S. Lysova

St. Petersburg State Institute of Technology, Moskovskii pr. 26, St. Petersburg, 190013 Russia
e-mail: yuri@newchem.ru

Received August 19, 2008

Abstract—An empirical procedure was used to calculate 278 ionization constants of 271 organic compounds, including phenols, NH acids, benzoic acid derivatives, and mono- and dihydric carboxylic acids, in water. The examined compounds were divided into 11 structural groups. The ionization constants for compounds belonging to a single group were calculated from constant empirical coefficients, molecular volumes, and formulas with an average error of less than 3.5%, the maximal error not exceeding 9.9%.

DOI: 10.1134/S1070428009060049

Unusually rapidly growing interest in calculations of ionization constants (pK_a) of organic compounds has been observed in the past two years. The reason is that pK_a values of potential biologically active compounds are important for biological screening and estimation of pharmacological properties related to pK_a [1–5]. Despite good consistency in experimental pK_a values of organic compounds in water, reported in different publications [6, 7], experimental procedures are continuously improved. This applies to both traditional potentiometric [8], spectrophotometric [9, 10], and capillary electrophoresis methods [11–13] and new procedures, e.g., those utilizing linear pH gradient systems [14]. Extension of the set of available experimental pK_a values of organic compounds in aqueous medium opens wide possibilities in analyzing applicability of theoretical models for the calculation of these constants.

In the recent time, most publications undoubtedly deal with structure–property correlations, in particular QSPR. These methods make use of so-called descriptors, i.e., quantitative parameters that are determined by molecular structure. Depending on a particular procedure, such descriptors may be parameters calculated by semiempirical quantum-chemical methods [2, 15–20], structural topological parameters (CoMFA) [2, 3], fragment (CoMSA) [1, 21] or atomic parameters [1], elements of quantum–topological molecular similarity (QTMS) [5, 22, 23], or several of the above listed (mixed descriptors). Undoubtedly, a strong advantage of QSPR is statistical analysis used to opti-

mize algorithm of selection of significant descriptors. For this purpose, computational neural network-based methodologies are extensively used [15–17, 21, 24]; they ensure less than 10% average relative error in the calculations. Something that is generally not remembered for some time past should be regarded as a disadvantage of these procedures. In most cases descriptors are microcosm elements, i.e. quantities whose values cannot be determined experimentally. For example, in keeping with the uncertainty principle, the relative error in direct measurements (by X-ray diffraction method) of an electron density deficit of 0.2 a.u. at the α -carbon atom with respect to carboxy group should be about 100%. Nevertheless, structure–property correlations extensively develop via search for new descriptors, and there are hopes that new descriptors will include not only “ephemeral” parameters. For instance, Jover et al. [15–17] calculated pK_a values of organic compounds in different solvents using experimentally measurable parameters (Kamlet–Taft parameter and standard heat of vaporization).

Undoubtedly, quantum-chemical methods for the calculation of pK_a values of organic compounds are free from the above disadvantage. The best consistency with the experimental data was achieved with the use of stepwise calculation procedures. In due time, good results were obtained in the calculation of protonation constants (proton affinities) of organic compounds in the gas phase [25]. The error in quantum-chemical calculations of proton affinities (PA), in particular by the complete basis set method (CBS-QB3) [26], was less

than 0.1%. Attempts were made to take into account solvent effect via correlation of PA and pK_a values for a single class of organic compounds [27]; however, these attempts were not continued. Methods have been developed, which make it possible to determine energies of solvation of molecules; this parameter is necessary for the calculation of pK_a . Calculation of solvation energies involves well known difficulties, so that various procedures were proposed to take into account distortion of molecular structure upon dissolution and nonspecific solvation effect on electronic structure of molecules [28–35]. Better results were obtained with the use of models implying formation of complexes with solvent or conjugate base [36–39]. This is quite reasonable, for specific solvation is taken into con-

sideration, which is especially important for aqueous media. In addition, solvate complex and solvent field effect models were proposed [40, 41]. From the practical viewpoint, quantum-chemical methods are not always appropriate since the error in the calculations ranges from 5 to 40% [42]. A probable reason is that each procedure implies a number of approximations and assumptions which are valid only for a specific series of organic compounds but invalid for others. If so, the problem related to improvement of correlations between the results of quantum-chemical calculations and experimental data is quite surmountable.

Correlation analysis methods based on Hammett's fundamental works have not been used alone for the calculation of pK_a values in the recent time. Hammett

Table 1. Groups, numbers of compounds in the groups (N), coefficients a and b , empirical atomic coefficients X_i , and average deviations of the calculated ionization constants pK_a , pK_b , and pK_{BH^+} from the experimental values [6, 7, 16, 27, 49, 50]

Group	N	$a \times 100$	b	Empirical atomic coefficients $X_i, \text{\AA}^3$								$\epsilon_{av}, \%$
				C	H	O	N	F	Cl	Br	I	
<i>meta</i> - and <i>para</i> -Substituted phenols C_6H_6O	28	-2.388	12.95	-5.296	-16.81	4.943	31.44	-14.43	-28.15	-21.68		2.4
<i>ortho</i> -Substituted phenols	36	-1.659	12.31	1.803	-21.69	29.51	80.12	29.44	24.11	16.28		2.9
<i>meta</i> - and <i>para</i> -Substituted benzoic acids $C_7H_6O_2$	26	-2.037	7.431	4.441	-15.78	-1.703	-3.034	-5.014	-18.88	-20.92	-32.48	3.1
<i>ortho</i> -Substituted benzoic acids	16	4.094	-2.694	-1.440	-16.88	-15.88	-39.86	-30.56	-50.63	-53.77	-62.63	2.4
<i>meta</i> - and <i>para</i> -Substituted anilines C_6H_7N	20	-1.254	11.31	22.41	-16.90	-79.08	-137.3	-59.60	-92.84	-132.7		2.2
<i>ortho</i> -Substituted anilines	21	1.872	7.046	116.6	-76.31	-16.23	178.5	-36.56	-27.84	-40.15		1.7
Saturated aliphatic carboxylic acids CHO_2 with functional substituents in the α -position ^a	17	-2.475	4.198	27.68	-30.44	-8.186	-31.81	34.74	-15.56	-35.04	-49.88	2.4
Saturated aliphatic carboxylic acids with no functional substituent in the α -position ^b	24	-1.691	5.997	-40.07	5.471	35.13				5.226		0.9
Unsaturated aliphatic carboxylic acids ^c	11	-7.793	10.22	13.05	-21.39	-0.637						1.4
Aliphatic amines, amino alcohols, and alicyclic amines H_2N^d	34	-2.069	11.04	16.71	-23.65	49.83	33.03					3.0
Pyridine derivatives $C_5H_5N^e$	23	10.67	-8.629	-22.81	-1.755	0.059			-73.04			3.5

^a The group includes α -alkoxy-, α -cyano-, and α -hydroxy derivatives and five derivatives of propanedioic and butanedioic acids.

^b The group also includes alky-substituted carboxylic acids, seven homologous dihydric acids, 4-oxopentanoic acid, and 3-bromopropanoic acid.

^c The group includes seven dienoid acids and their alkyl-substituted derivatives, propynoic acid, but-2-ynoic acid, and two derivatives of but-2-enedioic acid.

^d The group includes 17 alkylamines, 3 allylamines, 6 diamines (including one amino alcohol), diethylenetriamine, 6 amino alcohols, and cyclohexylamine.

^e The group includes pyridine, alkylpyridines (total of 17 compounds), three vinylpyridines, two chloropyridines, and 4-methoxypyridine.

Table 2. Calculated and experimental pK_a values of *meta*- and *para*-substituted phenols in water at 25°C

Compound	ρ , g/cm ³	pK_a (experimental)	pK_a (calculated)	ϵ_{av} , %
Phenol	1.132 [50]	9.99 [6]	9.66	3.3
3-Methylphenol	1.0336 [53]	10.00 [6]	9.73	2.7
4-Methylphenol	1.154 [50]	10.27 [50]	10.17	1.0
3-Ethylphenol	1.0283 [53]	10.07 [6]	10.10	0.3
4-Ethylphenol	1.011 [53]	10.00 [6]	10.02	0.2
3,4-Dimethylphenol	1.138 [50]	10.35 [50]	10.55	2.0
3,5-Dimethylphenol	1.115 [50]	10.2 [50]	10.47	2.6
4-Propylphenol	1.009 [51]	10.34 [7]	10.39	0.5
4-Isopropylphenol	0.99 [51]	10.24 [7]	10.28	0.4
3-Methoxyphenol	1.131 [53]	9.652 [6]	9.41	2.5
3-Ethoxyphenol	1.105 [53]	9.655 [6]	9.73	0.8
1,3-Dihydroxybenzene, pK_{a1}	1.285 [53]	9.44 [6]	9.44	0.0
1,4-Dihydroxybenzene, pK_{a1}	1.358 [53]	9.91 [6]	9.62	2.9
1,3,5-Trihydroxybenzene, pK_{a1}	1.46 [51]	8.45 [6]	9.29	9.9
3-Nitrophenol	1.4854 [54]	8.36 [15]	7.85	6.1
4-Nitrophenol	1.479 [51]	7.18 [15]	7.83	9.1
3-Fluorophenol	1.238 [51]	9.21 [7]	9.30	1.0
3-(Trifluoromethyl)phenol	1.3418 [51]	8.95 [7]	8.92	0.3
3-Chlorophenol	1.268 [53]	9.10 [6]	9.20	1.1
4-Chlorophenol	1.306 [53]	9.43 [6]	9.32	1.2
4-Bromophenol	1.84 [53]	9.34 [6]	9.34	0.0
3-Hydroxybenzoic acid	1.484 [53]	9.85 [6]	9.15	7.1
4-Hydroxybenzoic acid	1.482 [53]	9.23 [6]	9.15	0.9
3,4-Dihydroxybenzoic acid	1.542 [53]	8.67 [6]	8.76	1.1
3,4,5-Trihydroxybenzoic acid	1.694 [53]	8.85 [6]	8.62	2.5
4-(Acetylamino)phenol	1.293 [51]	9.38 [7]	8.90	5.1
Naphthalen-1-ol	1.224 [53]	9.30 [6]	9.59	3.1
Naphthalen-2-ol	1.217 [53]	9.57 [6]	9.56	0.1

constants σ are used as descriptors in some QSPR methods, as well as corrections in quantum-chemical methods. Gross and Seybold [43] compared pK_a values of 36 anilines, determined by correlation analysis and calculated by quantum-chemical methods (Hartree-Fock, GAUSSIAN94, Spartan). The correlation coefficient for the pK_a - σ dependences were poorer than those found by MMSLIE and MEPM, calculated by quantum-chemical methods. Nevertheless, we believe that practical application of empirical methods, including correlation analysis, may occupy leading position if the amounts of time and resources put therein will be the same as those consumed by QSPR or GAUSSIAN.

Our empirical method is based on the assumption that an unambiguous (though nonlinear) dependence exists between ionization constants of organic compounds and their molecular volumes within a definite group. A group consists of organic compounds having a common structural element (elements), e.g., the same protonation or deprotonation center. Theoretical foundations of the proposed procedure were described in general in [44]. This procedure was used by us previously to calculate ionization potentials, polarizabilities [45], and proton affinities [44] of organic compounds from their molecular volumes. Relations were found between thermodynamic parameters of a wide series of

Table 3. Calculated and experimental pK_a values of *ortho*-substituted phenols in water at 25°C

Compound	ρ , g/cm ³	pK_a (experimental)	pK_a (calculated)	ϵ_{av} , %
2-Methylphenol	1.135 [50]	10.33 [50]	10.37	0.4
2,3-Dimethylphenol	1.164 [50]	10.53 [50]	10.79	2.5
2,4-Dimethylphenol	0.965 [53]	10.58 [6]	10.20	3.6
2,5-Dimethylphenol	1.189 [50]	10.4 [50]	10.85	4.4
2,6-Dimethylphenol	1.132 [50]	10.63 [50]	10.71	0.8
2-Ethylphenol	1.0234 [53]	10.2 [6]	10.40	1.9
2-Propylphenol	1.015 [51]	10.47 [7]	10.68	2.0
2-Isopropylphenol	1.012 [51]	10.47 [7]	10.67	1.9
2-Isopropyl-5-methylphenol	0.9257 [53]	10.62 [6]	10.59	0.2
2- <i>tert</i> -Butylphenol	0.9783 [51]	10.28 [7]	10.83	5.4
2,6-Di- <i>tert</i> -butyl-4-methylphenol	1.048 [52]	12.23 [7]	12.72	4.0
2,4,6-Tri- <i>tert</i> -butylphenol	0.864 [51]	12.19 [7]	12.21	0.2
2-Hydroxymethylphenol	1.1613 [53]	9.92 [6]	9.56	3.6
2-Methoxyphenol	1.1288 [53]	9.99 [6]	9.48	5.1
2-Methoxy-4-methylphenol	1.095 [51]	10.28 [7]	9.72	5.5
2-Methoxy-4-(prop-1-en-1-yl)phenol	1.08 [51]	9.88 [7]	9.67	2.2
4-Allyl-2-methoxyphenol	1.0664 [51]	10.19 [7]	9.61	5.7
3-Hydroxy-4-methoxybenzaldehyde	1.196 [53]	8.889 [6]	8.48	4.6
4-Hydroxy-3-methoxybenzaldehyde	1.056 [53]	7.396 [6]	8.02	8.4
2-Hydroxybenzaldehyde	1.1674 [53]	8.34 [6]	8.90	6.8
2-Ethoxyphenol	1.0903 [53]	10.109 [6]	9.70	4.0
2-Nitrophenol	1.657 [53]	7.222 [6]	7.33	1.4
2,4-Dinitrophenol	1.683 [53]	4.08 [6]	3.96	3.0
2-Butyl-4,6-dinitrophenol	1.265 [51]	4.62 [7]	4.50	2.6
2,4,6-Trinitrophenol	1.763 [53]	0.71 [6]	0.72	1.8
4-Chloro-2,6-dinitrophenol	1.74 [53]	2.97 [6]	2.75	7.4
1,2-Dihydroxybenzene, pK_{a1}	1.371 [53]	9.356 [6]	9.60	2.6
1,2,3-Trihydroxybenzene, pK_{a1}	1.453 [53]	9.03 [6]	8.94	1.0
2-Fluorophenol	1.12 [51]	8.7 [7]	8.70	0.0
2-Chlorophenol	1.2634 [53]	8.55 [6]	8.74	2.3
Pentachlorophenol	1.978 [52]	4.7 [7]	4.80	2.1
2-Bromophenol	1.4924 [53]	8.452 [6]	8.48	0.4
2,4-Dibromophenol	2.07 [51]	7.79 [7]	7.69	1.2
2,4,6-Tribromophenol	2.55 [51]	6.8 [7]	6.84	0.6
2-Phenylphenol	1.213 [51]	9.92 [7]	9.70	2.2
2-Aminophenol	1.328 [53]	9.28 [6]	9.07	2.2

organic compounds [46], and pK_a values of carboxylic [47] and NH acids [48] in different solvents were calculated. Nonlinearity of the proposed dependence is determined by atomic composition of organic molecules:

$$pK_a = a\delta V + b, \quad (1)$$

where a and b are empirical coefficients that are constant for a particular group, and

$$\delta V = V + \Sigma(g_i - g'_i)X_i. \quad (2)$$

Here, V is the molecular volume of an organic compound, g_i is the number of i th atoms in its molecule,

g'_i is the number of i th atom in the base structural unit of a group, and X_i is the contribution of i th atom to the molecular volume (its value is constant for compounds of the same group). For example, the base structural unit for the phenol group is C_6H_6O , and for the aniline group, C_6H_7N .

The set of experimental ionization constants of organic compounds in aqueous solutions was collected as follows. Data from studies aimed at determining thermodynamic pK_a values were preferred. Data from publications in which the experimental conditions were not given were excluded; if experimental procedure did not conform to the IUPAC recommendations, the corresponding data were also excluded. In addition, mutual consistency of data given in different publications was taken into account. As a result, the final set

of experimental pK_a values included those reported in [6, 7, 16, 27, 49, 50].

The empirical coefficients were determined by minimization of the sum of squared relative deviations of pK_a values calculated by formula (1) from the corresponding experimental values $pK_a(\text{exp})$ for all compounds within a group:

$$E = \sum \frac{[pK_a(\text{exp}) - pK_a]^2}{[pK_a(\text{exp})]^2} \quad (3)$$

If the quantity E has a global minimum, the following conditions should be met, and their number should be equal to the number of unknown empirical coefficients:

Table 4. Calculated and experimental pK_a values of *meta*- and *para*-substituted benzoic acids in water at 25°C

Compound	ρ , g/cm ³	pK_a (experimental)	pK_a (calculated)	ϵ_{av} , %
Benzoic acid	1.291 [53]	4.204 [6]	4.23	0.6
4-Isopropylbenzoic acid	1.163 [53]	4.36 [6]	4.31	1.1
4-Methoxybenzoic acid	1.385 [53]	4.25 [16]	4.30	1.2
3-Fluorobenzoic acid	1.474 [53]	3.88 [16]	4.00	3.0
4-Fluorobenzoic acid	1.479 [53]	4.14 [6]	4.01	3.2
3-Chlorobenzoic acid	1.496 [53]	3.83 [6]	3.95	3.2
4-Chlorobenzoic acid	1.541 [53]	4.00 [16]	4.06	1.4
3-Bromobenzoic acid	1.845 [53]	3.81 [6]	3.85	1.1
4-Bromobenzoic acid	1.894 [53]	3.99 [6]	3.95	1.1
3-Iodobenzoic acid	2.171 [53]	3.86 [6]	3.91	1.2
4-Iodobenzoic acid	2.184 [53]	3.98 [16]	3.93	1.3
3-Nitrobenzoic acid	1.494 [53]	3.46 [6]	3.46	0.1
4-Nitrobenzoic acid	1.61 [53]	3.44 [6]	3.73	8.4
3,4-Dinitrobenzoic acid	1.674 [53]	2.82 [6]	2.76	2.0
4-Chloro-3-nitrobenzoic acid	1.645 [51]	3.29 [16]	3.16	4.0
Benzene-1,3-dicarboxylic acid, pK_{a1}	1.507 [53]	3.62 [6]	3.68	1.7
Benzene-1,4-dicarboxylic acid, pK_{a1}	1.51 [53]	3.54 [6]	3.69	4.2
3-Hydroxybenzoic acid	1.473 [53]	4.076 [6]	4.29	5.3
4-Hydroxybenzoic acid	1.482 [53]	4.55 [16]	4.31	5.2
3,4-Dihydroxybenzoic acid	1.524 [53]	4.48 [7]	4.08	8.9
3,4,5-Trihydroxybenzoic acid	1.694 [53]	4.21 [6]	4.14	1.7
3-Aminobenzoic acid	1.511 [53]	4.74 [16]	4.74	0.1
4-Aminobenzoic acid	1.374 [53]	4.82 [16]	4.44	7.9
Naphthalene-1-carboxylic acid	1.398 [53]	3.695 [6]	3.55	4.0
<i>N</i> -Benzoylaminoacetic acid	1.371 [53]	3.65 [6]	3.89	6.6
Phenylacetic acid	1.228 [53]	4.312 [6]	4.23	1.8

Table 5. Calculated and experimental pK_a values of *ortho*-substituted benzoic acids in water at 25°C

Compound	ρ , g/cm ³	pK_a (experimental)	pK_a (calculated)	ϵ_{av} , %
2-Iodobenzoic acid	2.249 [51]	2.93 [7]	2.93	0.0
2-Fluorobenzoic acid	1.46 [53]	3.27 [6]	3.27	0.0
2-Bromobenzoic acid	1.929 [51]	2.88 [7]	2.88	0.0
2,5-Dimethylbenzoic acid	1.069 [53]	3.99 [6]	3.97	0.4
Biphenyl-2-carboxylic acid	1.458 [53]	3.46 [6]	3.43	0.8
2,4-Dinitrobenzoic acid	1.672 [53]	1.43 [6]	1.45	1.3
2-Chloro-5-nitrobenzoic acid	1.608 [51]	2.17 [7]	2.21	1.6
2-Methoxybenzoic acid	1.18 [53]	4.09 [6]	3.98	2.7
3,6-Dichloro-2-methoxybenzoic acid	1.57 [51]	1.97 [7]	2.02	2.7
Benzene-1,2-dicarboxylic acid, pK_{a1}	1.593 [53]	2.95 [6]	3.04	2.9
2-Chlorobenzoic acid	1.544 [53]	2.92 [16]	2.82	3.5
2,3-Dihydroxybenzoic acid	1.542 [51]	2.91 [7]	2.80	3.8
2-Nitrobenzoic acid	1.575 [53]	2.19 [16]	2.28	4.0
2-Hydroxy-5-nitrobenzoic acid	1.65 [53]	2.05 [16]	1.96	4.4
2-Chloro-3-nitrobenzoic acid	1.662 [51]	2.02 [16]	1.93	4.5
2-Hydroxybenzoic acid	1.443 [53]	2.98 [6]	3.16	6.1

Table 6. Calculated and experimental pK_b values of *meta*- and *para*-substituted anilines in water at 25°C

Compound	ρ , g/cm ³	pK_b (experimental)	pK_b (calculated)	ϵ_{av} , %
Aniline	1.0217 [53]	9.42 [6]	9.41	0.1
3-Methylaniline	0.98912 [53]	9.29 [6]	9.19	1.0
4-Methylaniline	0.9619 [53]	8.92 [6]	9.13	2.4
3-Ethylaniline	0.9896 [53]	9.30 [6]	9.04	2.8
4-Ethylaniline	0.9679 [53]	8.95 [27]	8.99	0.4
3,5-Dimethylaniline	0.9706 [53]	9.09 [49]	8.99	1.0
3-Methoxyaniline	1.096 [53]	9.80 [6]	10.10	3.1
3-Ethoxyaniline	1.032 [53]	9.83 [6]	9.82	0.1
3-Nitroaniline	0.901 [53]	11.5 [49]	11.61	0.9
4-Nitroaniline	1.424 [53]	12.98 [49]	12.78	1.5
Methyl 3-aminobenzoate	1.232 [53]	10.36 [27]	10.60	2.3
Benzidine	1.25 [53]	9.30 [6]	9.33	0.4
3-Fluoroaniline	1.158 [53]	10.41 [49]	9.84	5.4
4-Fluoroaniline	1.1725 [53]	9.35 [6]	9.87	5.6
3-(Trifluoromethyl)aniline	1.3047 [51]	10.51 [7]	10.48	0.3
3-Chloroaniline	1.2161 [53]	10.48 [6]	10.07	3.9
4-Chloroaniline	1.422 [53]	10.02 [6]	10.39	3.7
3-Bromoaniline	1.5793 [53]	10.49 [6]	10.49	0.0
3-Aminobenzoic acid	1.511 [53]	10.93 [6]	11.12	1.7
4-Aminobenzoic acid	1.374 [53]	11.59 [6]	10.93	5.7

$$\frac{\partial E}{\partial a} = \frac{\partial E}{\partial b} = \frac{\partial E}{\partial X_i} = 0. \quad (4)$$

As molecular volumes we used initially those extrapolated to infinite dilution in aqueous solution. However, better correspondence to the experimental data was achieved with molecular volumes calculated from the molecular weights and relative densities at 20°C or 25°C and atmospheric pressure [50–54]. The intersection of the two sets, experimental pK_a values [6, 7, 16, 27, 49, 50] and relative densities [50–54], included 249 compounds which were divided into 11 groups by statistical combinatorial processing.

The division was initially made with respect to the reaction center. We thus obtained three groups: carboxylic acids, phenols, and NH acids. However, such division did not ensure sufficient agreement with the experimental data. Therefore, each group was additionally divided according to structural features. The calculated empirical coefficients for all isolated groups of organic compounds are collected in Table 1.

The calculated pK_a , pK_b , and pK_{BH^+} values of compounds belonging to the first six groups (Table 1) are listed in Tables 2–7 together with the corresponding experimental values. The remaining 109 calculated ionization constant of compounds belonging to five groups are available from the authors.

Using the obtained coefficients (Table 1) we calculated pK_a , pK_b , and pK_{BH^+} values (water, 25°C) for those compounds for which experimental data are lacking [6, 7, 16, 27, 49, 50] from the known molecular volumes. The results are given in Table 8.

In keeping with the assumed division, some compounds may be classed with different groups. Judging by the accuracy of calculations with empirical coefficients determined for one or another group we can conclude which proton in an organic compound is more labile. This approach was applied in [55], where the mechanism of protonation of difunctional organic compounds was studied. The ionization constants of 2-hydroxy-, 2,3-dihydroxy-, and 2-hydroxy-5-nitro-

Table 7. Calculated and experimental pK_b values of *ortho*-substituted anilines in water at 25°C

Compound	ρ , g/cm ³	pK_b (experimental)	pK_b (calculated)	ϵ_{av} , %
2-Methylaniline	0.99843 [53]	9.61 [27]	9.71	1.0
2-Ethylaniline	0.983 [53]	9.58 [6]	9.53	0.5
2,3-Dimethylaniline	0.9931 [53]	9.3 [6]	9.49	2.0
2,4-Dimethylaniline	0.9723 [53]	9.11 [6]	9.57	5.1
2,5-Dimethylaniline	0.979 [53]	9.47 [6]	9.54	0.8
2,6-Dimethylaniline	0.9842 [53]	10.05 [6]	9.52	5.2
2,4,6-Trimethylaniline	0.9633 [53]	9.62 [6]	9.38	2.5
2-Methoxyaniline	1.0923 [53]	9.51 [27]	9.57	0.7
2-Ethoxyaniline	1.051 [53]	9.53 [6]	9.45	0.8
Methyl 2-aminobenzoate	1.1682 [53]	11.77 [27]	11.97	1.7
Ethyl 2-aminobenzoate	1.1174 [51]	11.82 [7]	11.87	0.4
2-Nitroaniline	1.442 [53]	14.28 [6]	14.19	0.7
2,4-Dinitroaniline	1.615 [51]	18.25 [7]	18.90	3.5
2,4,6-Trinitroaniline	1.762 [53]	24.23 [6]	23.56	2.8
2-Fluoroaniline	1.1478 [53]	10.8 [6]	10.80	0.0
2-Chloroaniline	1.21251 [53]	11.36 [6]	11.22	1.2
2,4-Dichloroaniline	1.567 [51]	12.00 [7]	12.08	0.6
2-Bromoaniline	1.578 [53]	11.4 [27]	11.11	2.5
2,4-Dibromoaniline	2.26 [51]	11.7 [7]	11.85	1.3
2-Aminophenol	1.328 [51]	9.16 [7]	9.30	1.5
2-Aminobenzoic acid	1.412 [53]	11.89 [27]	11.64	2.1

Table 8. Calculated ionization constants pK_a and pK_b of different compounds in water at 25°C

Compound	ρ , g/cm ³	pK_a (calcd.)
<i>meta</i> - and <i>para</i> -Substituted phenols		
3-Butylphenol	0.974 [53]	10.55
4-Butylphenol	0.978 [53]	10.58
4- <i>s</i> -Butylphenol	0.969 [53]	10.52
<i>ortho</i> -Substituted phenols		
4- <i>tert</i> -Butyl-2,6-dimethylphenol	0.959 [53]	11.32
1,3-Dihydroxy-2,4,6-trinitrobenzene	1.829 [53]	0.12
5-Isopropyl-2-methylphenol	0.9772 [53]	10.83
2-Butylphenol	0.975 [53]	10.82
Phenyl 2-hydroxybenzoate	1.2614 [53]	7.88
2-Chloro-3,4-dimethylphenol	1.5538 [53]	10.15
Butyl 2-hydroxybenzoate	1.0728 [53]	9.07
<i>ortho</i> -Substituted benzoic acids		
3-Chloro-2-nitrobenzoic acid	1.566 [53]	2.43
5-Chloro-2-nitrobenzoic acid	1.593 [53]	2.29
Unsaturated aliphatic carboxylic acids		
Hex-2-enoic acid	0.965 [53]	4.84
12-Hydroxyoctadec-9-enoic acid ^a	0.945 [53]	7.14
Octadeca-9,12-dienoic acid ^a	0.9022 [53]	4.40
<i>ortho</i> -Substituted anilines		
2,3,4,6-Tetramethylaniline ^a	0.978 [51]	9.09
2-Bromo-4-methylaniline	1.51 [51]	10.88
2-Bromo-5-methylaniline	1.47 [51]	10.98
2,4,6-Tribromoaniline	2.35 [51]	13.44
2-Methyl-3-nitroaniline	1.378 [51]	13.97
2-Propylaniline	0.9602 [51]	9.40
<i>meta</i> - and <i>para</i> -Substituted anilines		
Ethyl 3-aminobenzoate	1.171 [51]	10.36

^a Compounds for which pK_a (pK_b) values are given arbitrarily, for the concentration of their saturated solutions in water at the given temperature is insufficient to determine these values using traditional experimental techniques.

benzoic acids were calculated only with the use of the coefficients for *ortho*-substituted benzoic acids (Table 5), while the results of calculations with the coefficients determined for phenols were poor; the reason is that the carboxy group is more acidic than phenolic hydroxy group. Likewise, the ionization con-

stant of 2-aminobenzoic acid was calculated with the coefficients for *ortho*-substituted anilines (Table 7). Presumably, intramolecular protonation makes protons at the quaternary nitrogen atom more acidic. In fact, the ionization constants of 3- and 4-aminobenzoic acids, calculated with the coefficients for benzoic acids (Table 4) and for *meta*- and *para*-substituted anilines (Table 6) are approximately similar. The ionization constant of 2-aminophenol, calculated with the coefficients for *ortho*-substituted anilines (Table 7), was better consistent with the experimental value than that calculated using the phenol group coefficients (Table 3). This fact can also be rationalized in terms of higher acidity of protons at a quaternary nitrogen atom, as compared to hydroxy proton involved in five-membered H-chelate ring. The ionization constants of 3-hydroxy-, 4-hydroxy-, 3,4-dihydroxy-, and 3,4,5-trihydroxybenzoic acids were calculated according to the equations found for both phenols (Table 2) and benzoic acids (Table 4). Larger errors were obtained for 4-hydroxy- and 3,4-dihydroxybenzoic acids, and lower errors, for 3-hydroxy- and 3,4,5-trihydroxybenzoic acids. These data somewhat contradict the above explanations, but the difference in the deviations in both directions was lower than the average error of the calculations for all compounds.

The maximal differences between the calculated and experimental ionization constants were observed for the following compounds: 4-hydroxy-3-methoxybenzaldehyde (*ortho*-substituted phenols), 1,3,5-trihydroxybenzene (other phenols), 2-hydroxybenzoic acid (*ortho*-substituted benzoic acids), 3,4-dihydroxybenzoic acid (other benzoic acids), (2*RS*,3*RS*)-2,3-dihydroxybutanedioic acid (aliphatic carboxylic acids), propane-1,2-diamine (amines and amino alcohols), 4-vinylpyridine (pyridine derivatives), 2,4- and 2,6-dimethylaniline (*ortho*-substituted anilines), and 4-aminobenzoic acid (other anilines). The calculations revealed that the empirical coefficients are convergent, i.e., the maximal and average deviations within particular groups decrease as the number of values increases. Therefore, extension of the set of experimental ionization constants of organic compounds should lead to refinement of the empirical coefficients; furthermore, the maximal difference between the calculated and experimental values would be observed for other compounds in a group.

Analysis of our results shows that equation (1) is applicable for the calculation of ionization constants of various organic compounds in water. The empirical

coefficients for three groups of organic compounds, *ortho*-substituted benzoic acids, *ortho*-substituted anilines, and pyridines (Table 1), differ appreciably from those found for the other groups.

REFERENCES

1. Gieleciak, R. and Polanski, J., *J. Chem. Inf. Model.*, 2007, vol. 47, p. 547.
2. Carabias-Martinez, R., Rodrigues-Gonzalo, E., Dominguez-Alvarez, J., and Miranda-Cruz, E., *Anal. Chim. Acta*, 2007, vol. 584, p. 410.
3. Jelfs, S., Ertl, P., and Selzer, P., *J. Chem. Inf. Model.*, 2007, vol. 47, p. 450.
4. Milletti, F., Storchi, L., Sforma, G., and Cruciani, G., *J. Chem. Inf. Model.*, 2007, vol. 47, p. 2172.
5. Popelier, P.L.A. and Smith, P.J., *Eur. J. Med. Chem.*, 2006, vol. 41, p. 862.
6. *Langes Handbook of Chemistry*, Dean, J.A., Ed., New York: McGraw-Hill, 11th ed., 1999, vol. 8, p. 24.
7. Tehan, B.G., Lloyd, E.J., Wong, M.G., Pitt, W.R., Montana, J.G., Manallack, D.T., and Gancia, E., *Quant. Struct.-Act. Relat.*, 2002, vol. 21, p. 457.
8. Kong, X., Zhou, T., Liu, Z., and Hider, R.C., *J. Pharm. Sci.*, 2007, vol. 96, p. 2777.
9. Tam, K.Y. and Takas-Novak, K., *Anal. Chim. Acta*, 2001, vol. 434, p. 157.
10. Tam, K.Y., *Anal. Sci.*, 2001, vol. 17, p. 1203.
11. Caliaro, G.A. and Herbots, A.A., *J. Pharm. Biomed. Anal.*, 2001, vol. 26, p. 427.
12. Jia, Z.J., Ramstad, T., and Zhong, M., *Electrophoresis*, 2001, vol. 22, p. 1112.
13. Kibbey, C.E., Poole, S.K., Robinson, B., Jackson, J.D., and Durham, D., *J. Pharm. Sci.*, 2001, vol. 90, p. 1164.
14. Box, K., Bevan, C., Comer, J., Hill, A., Allen, R., and Reynolds, D., *Anal. Chem.*, 2003, vol. 75, p. 883.
15. Jover, J., Bosque, R., and Sales, J., *QSAR Comb. Sci.*, 2007, vol. 26, p. 385.
16. Jover, J., Bosque, R., and Sales, J., *QSAR Comb. Sci.*, 2008, vol. 27, p. 563.
17. Jover, J., Bosque, R., and Sales, J., *QSAR Comb. Sci.*, 2008, vol. 27, p. 1179.
18. Gross, K.C., Seybold, P., Peralta-Inga, Z., Murray, J.S., and Politzer, P., *J. Org. Chem.*, 2001, vol. 66, p. 6919.
19. Ghasemi, J., Saaidpour, S., and Brown, S.D., *J. Chem. Struct. (Theochem)*, 2007, vol. 805, p. 27.
20. Duchowicz, P.R. and Castro, E.A., *Mendeleev Commun.*, 2002, vol. 12, p. 187.
21. Ivanova, A.A., Baskin, I.I., Palyulin, V.A., and Zefirov, N.S., *Dokl. Akad. Nauk*, 2007, vol. 413, p. 766.
22. Shan, X., Qin, W., Zhou, Z., and Dai, Y., *J. Chem. Eng. Data*, 2008, vol. 53, p. 331.
23. Esteki, M., Hemmateenejad, B., Khayamian, T., and Mohajeri, A., *Chem. Biol. Drug Des.*, 2007, vol. 70, p. 413.
24. Habibi-Yangjeh, A. and Danandeh-Jenagharad, M., *Indian J. Chem., Sect. B*, 2007, vol. 46, p. 478.
25. Epshtein, L.M. and Iogansen, A.V., *Usp. Khim.*, 1990, vol. 59, p. 229.
26. Liptak, M.D., Gloss, K.C., Seybold, P.G., Feloligus, S., and Shields, G.C., *J. Am. Chem. Soc.*, 2002, vol. 124, p. 6421.
27. Pankratov, A.N., Uchaeva, I.M., Doronin, S.Yu., and Chernova, R.K., *Zh. Strukt. Khim.*, 2001, vol. 42, p. 884.
28. Brown, T.N. and Mora-Diez, N., *J. Phys. Chem. B*, 2006, vol. 110, p. 9270.
29. Makowska, J., Makowski, M., and Chmurzynski, L., *J. Phys. Chem. A*, 2004, vol. 108, p. 10354.
30. Brown, T.N. and Mora-Diez, N., *J. Phys. Chem. B*, 2006, vol. 110, p. 20546.
31. Saracino, G.A.A., Improta, R., and Barone, V., *Chem. Phys. Lett.*, 2003, vol. 373, p. 411.
32. Krol, M., Wrona, M., Page, C.S., and Bates, P.A., *J. Chem. Theory Comput.*, 2006, vol. 2, p. 1520.
33. Qin, M., Zhao, G., Duan, C.S., Zhou, Z.Y., and Lu, J.F., *Pol. J. Chem.*, 2006, vol. 80, p. 1865.
34. Lu, H., Chen, X., and Zhan, C.-G., *J. Phys. Chem. B*, 2007, vol. 111, p. 10599.
35. Klicic, J.J., Friesner, R.A., Liu, S.-Y., and Guida, W.C., *J. Phys. Chem. A*, 2002, vol. 106, p. 1327.
36. Han, J. and Tao, F.-M., *J. Phys. Chem. A*, 2006, vol. 110, p. 257.
37. Namazian, M., Zakery, M., Noorbala, M.R., and Coote, M.L., *Chem. Phys. Lett.*, 2008, vol. 451, p. 163.
38. Tao, L., Han, J., and Tao, F.-M., *J. Phys. Chem. A*, 2008, vol. 112, p. 775.
39. Jia, Z.-K., Du, D.-M., Zhou, Z.-Y., Zhang, A.-G., and Hou, R.-Y., *Chem. Phys. Lett.*, 2007, vol. 439, p. 374.
40. Pliego, J.R., Jr. and Riveros, J.M., *J. Phys. Chem. A*, 2002, vol. 106, p. 7434.
41. Kelly, C.P., Cramer, C.J., and Truhlar, D.G., *J. Phys. Chem. A*, 2006, vol. 110, p. 2493.
42. Zevatskii, Yu.E. and Samoilov, D.V., *Russ. J. Org. Chem.*, 2007, vol. 43, p. 483.
43. Gross, K.C. and Seybold, P.G., *J. Org. Chem.*, 2001, vol. 66, p. 6919.
44. Zevatskii, Yu.E. and Vlasov, E.A., *Russ. J. Gen. Chem.*, 2007, vol. 77, p. 234.
45. Zevatskii, Yu.E. and Lysova, S.S., *Russ. J. Appl. Chem.*, 2006, vol. 79, p. 967.
46. Zevatskii, Yu.E. and Samoilov, D.V., *Russ. J. Appl. Chem.*, 2007, vol. 80, p. 230.
47. Zevatskii, Yu.E. and Samoilov, D.V., *Russ. J. Org. Chem.*, 2008, vol. 44, p. 52.
48. Zevatskii, Yu.E. and Samoilov, D.V., *Russ. J. Org. Chem.*, 2008, vol. 44, p. 1737.

49. Chaudry, U.A. and Popelier, L.A., *J. Org. Chem.*, 2004, vol. 69, p. 233.
50. Kharlampovich, G.D. and Churkin, Yu.V., *Fenoly* (Phenols), Moscow: Khimiya, 1974.
51. *CRC Handbook of Chemistry and Physics*, Lide, D.R., Ed., Boca Raton, FL: CRC, 2002.
52. *Khimicheskaya entsiklopediya* (Chemical Encyclopedia), Knunyants, I.L., Ed., Moscow: Sovetskaya Entsiklopediya, 1988, vols. 1, 2.
53. *Novyi spravochnik khimika i tekhnologa* (New Chemist's and Technologist's Handbook), Skvortsov, N.K., Ed., St. Petersburg: Nauch.-Proizv. Ob"edin. "Mir i sem'ya," 2002.
54. Rabinovich, V.A. and Khavin, Z.Ya., *Kratkii khimicheskii spravochnik* (Brief Chemical Reference Book), St. Petersburg: Khimiya, 1994.
55. Deakyne, C.A., Liebman, J.F., Vlasov, E.A., and Zevatskiy, Y.E., *Struct. Chem.*, 2008, vol. 19, p. 609.