Printed as manuscript

Zevatskiy E. Yuriy

NEW METHOD OF CALCULATION OF PHYSICOCHEMICAL PARAMETERS OF ORGANIC COMPOUNDS

Disciplines 02.00.03 — Organic chemistry,
02.00.04 — Physical chemistry

Thesis abstract
for a degree of the Doctor of Chemistry

Saint-Petersburg
2010

The work is performed in The State Educational Institution of Higher Professional Training
"Saint-Petersburg State Institute of Technology (Technical University)"

Leading organization: The State Educational Institution of Higher Professional Training *"D. Mendeleyev University of Chemical Technology of Russia"* (Moscow)

Thesis abstract is send out 1 June 2010.

GENERAL CHARACTERISTIC OF THE WORK

Quantitative estimation of reaction constants of organic compounds is one of the most important problem, but at the same time it is one of the most difficult task of theoretical organic chemistry. Its solution includes search of quantitative parameters that allow to predict the relative reactivity in the number of substances similar in structure or the preferred direction of the reaction in the presence of several reaction centers, as well as development of calculation methods for noted quantities.

To identify the detail reaction mechanism for determination of the rate-limiting step of the process; to consider the influence of media of the reaction expressed in solvation (specific and nonspecific) effects of the solvent, concentration and ionic effects; to determine the influence of substance structure, appearing in inductive, steric, resonance effect of the substituents; catalytic action – this is not a complete list of challenges facing an organic chemist, who carries out full-rate research of reactivity of certain substances. It is not a surprise that the universal formula that allows overcoming all these difficulties on this way has not been created yet.

Nowadays large-scale investigations in the area of the synthesis of compounds with intended biochemical, pharmacological and physicochemical properties are carried out. New conditions of the synthesis of compounds of diverse structure are examined. Instrumental possibilities of physicochemical methods of analysis are constantly mounting. As a consequence lots and lots of publications reporting newly synthesized compounds and new procedures of organic synthesis appear. It results in a situation when theory doesn't keep up with practice, when the methods of quantitative estimation and calculation of reaction constants of organic compounds are not able to cope with a variety of experimental material annually supplied by the research laboratories. For example, such ordinary, in terms of physics, modification of reaction conditions as the microwave activation results in significant quantitative abnormalities in the reactivity of organic compounds in well-studied processes. Theoretical study of this issue indicates that there is no consensus on the mechanism of microwave impact on the reaction kinetics and thermodynamics. This probably means that it's time to reflect on the experience accumulated in order to systematize

and find the universal factors that determine the behavior of compounds in chemical reactions. One possible solution is an empirical analysis of the research results in the field of organic synthesis with large array of diverse experimental quantitative material.

The purpose and objectives:

To suggest the universal method for calculation of various physicochemical parameters of organic compounds suitable to predict the values of equilibrium constants of protolytic reactions. To achieve this goal it was necessary to solve the following tasks:

- to develop the theoretical basis of the method by means of analysis the existing views on the relationship between physicochemical properties of compounds and molecular structure as well as measurement conditions;
- to develop the mathematical tools for practical application of the method;
- to develop an algorithm of practical application of the method depending on available initial experimental data;
- to determine the scope of the method: list of organic compounds and their physicochemical parameters for that the method can be effectively applied;
- in accordance with the chosen field of application to calculate the values of physicochemical properties of organic compounds for evaluation of reliability of method positions and approximations, as well as the effectiveness of its mathematical tool;
- to carry out experimental measurements of pK_a in different solvents in order to check reliability of suggested method for the values of the acidity constant (pK_a) of organic compounds and to determine the influence of medium.

Scientific novelty:

We offer new element linear empirical method (ELEM) for calculation of the values of physicochemical parameters of organic compounds taking into account its

molecular composition and structure. In the method numerical values of empirical parameters corresponding to atoms of the compound are used.

The conditions of the strict implementation of one fundamental principle of physical organic chemistry – the principle of polylinearity - are discovered.

The principle of "linearity of nonlinear effects" which allows calculating the nonlinear dependencies "structure-property" and "property-property" of organic compounds is well founded. Designed principle of "the element of permanent structure" (EPS) enables to achieve satisfactory accuracy of calculations of acidity constants values (pK_a) of organic compounds by means of ELEM at the reduction of the number of empirical parameters.

New experimental pK_a values of carboxylic acids and amines in three anhydrous organic solvents not published so far in the scientific literature are received. Using the ELEM equations values of standard thermodynamic parameters (enthalpy, entropy, Gibbs energy, isobaric heat capacity), proton affinity, polarizability, ionization potentials of various organic compounds are approximated. The calculations predicted 1010 values of these thermodynamic parameters of 365 organic compounds, the values of polarizabilities of 6 compounds, the values of ionization potentials of 6 compounds; the corresponding experimental values were not found in the literature. The pK_a values of carboxylic acid derivatives, phenols, amines and nitrogen heterocyclic compounds in various media are approximated. Comparison of empirical method values with values calculated by other methods in the literature demonstrated the benefits of ELEM in calculation accuracy in contrast with modern quantum-chemical, empirical and combined methods.

As a consequence 22 pK_a values of organic compounds in water, 17 pK_a values in methanol, 38 pK_a values in formamide, 31 pK_a value in ethanol, 18 pK_a values in nitromethane, 8 pK_a values in DMSO, 10 pK_a values in acetone, 6 pK_a values in acetonitrile, 20 pK_a values in benzonitrile and 10 pK_a values in propylene carbonate are predicted; the experimental data were not in the literature. Based on analysis of empirical parameters values obtained from the ELEM equations, proposed unique concepts of determination of the preferred deprotonation center in bifunctional

compounds and of classification of weak organic acids by type of the deprotonation center.

It is theoretically shown and experimentally confirmed that the molecular volume of organic compounds in the limit of infinite dilution in some solvent is connected to acid-base properties of this compound in the given media.

A new method of spectrophotometric measurement of acidity constants of organic compounds is suggested. New photocell of the unique design allowing registration of current-voltage characteristic of direct photocurrent cutoff without distortion caused by back photocurrent.

Practical importance:

Developed on the basis of ELEM the principles "linearity of nonlinear effects" and "the element of permanent structure" (EPS) allows determination of approximate equations for calculation of any measured physicochemical parameter by the experimental values of the other one in those cases when the explicit dependence between them is not identified or there is low correlation coefficient of values of these quantities.

ELEM equations based on the principles of "linearity of nonlinear effects" and EPS, as well as defined in this study empirical parameters of the atoms can be used for calculations of values of thermodynamic quantities (enthalpy, entropy, Gibbs energy, isobaric heat capacity), proton affinity, polarizability, ionization potentials of various organic compounds; the values of acidity constants of various classes of organic acids in different solvents with a relative error less than 10%. Performed calculations of acid-base properties of compounds in water can be used in order to evaluate biochemical parameters of both known and not yet synthesized compounds.

The experimental pK_a values in non-aqueous solvents (ethanol, methanol, formamide) allow us to set the ratio of prototropic forms of the compounds depending on the medium acidity and, thus, to predict the optimal conditions of the reaction to receive desired products in fine organic synthesis.

The developed photoemission method of spectral registration can be used to determine pK_a in cases when spectra of prototropic forms of compounds are not allowed, i.e. to solve the problems that are not comprehensible by traditional and multiwave methods of spectrophotometric analysis.

Principle defended provisions of the thesis:

- formulation of the purpose and objectives of the work;
- proof of the principle of polylinearity and formulation of two conditions when exact polylinear decomposition of the investigated physicochemical characteristics of compounds as functions of many variables is achieved;
- theoretical positions of the element linear empirical method and its three approximations (single empirical coefficient, corresponding to the particular chemical element, the principle of "linearity of nonlinear effects" and the principle of EPS) making possible to use the method for calculation of acid-base properties of organic compounds;
- calculations: the values of some thermodynamic parameters for 1055 organic compounds, the values of proton affinity of 395 compounds, 363 values of dissociation constants of 33 benzoic acid and acetic acid derivatives in 11 solvents, 192 values of dissociation constants of 24 protonated amines and nitrogen heterocyclic compounds in 8 solvents, the values of polarizabilities and ionization potentials of 78 CHNO-containing organic compounds;
- measurements: 58 values of dissociation constants of carboxylic acids, protonated amines and nitrogen heterocyclic compounds in anhydrous methanol, ethanol and formamide by the potentiometric method and 6 pK_a values of organic acids in water by the spectrophotometric method;
- identification of the preferred center of deprotonation in bifunctional compounds and classification of weak organic acids by type of the deprotonation center on the basis of analysis of empirical parameters in ELEM calculation equations;

- new photoemission method of spectral registration for measurements of equilibrium constants of prototropic forms of organic compounds and determination of the values of ionization constants of three organic pigments by this method.

Approbation of the work: The results were presented at the International conference "Modern Physical Chemistry» (MPC'07) in Kharkiv, 26-30 June 2007, at the XVIII Mendeleev Congress on General and Applied Chemistry (IUPAC) in Moscow, 23-28 September 2007, at the II International conference on computational chemistry (WSEAS COMPUCHEM'08) in Puerto de la Cruz, Tenerife, Spain, 15-17 December 2008, at the I International conference "New Trends in Chemistry of Heterocyclic Compounds" in Kislovodsk, 3-8 May 2009; at the V International Conference on Organic Chemistry «Universities Contribution in the Organic Chemistry Progress» (InterYCOS-2009) in St. Petersburg, 22-25 June 2009; at the XVII International Conference on Chemical Thermodynamics in Russia (RCCT) in Kazan, June 29 - July 3, 2009.

Related publications: 17 articles are published in Russian and international peer-reviewed journals, 2 patents of the Russian Federation are received, 8 abstracts are published in the collected works of international and national scientific conferences.

The structure and the volume of the thesis: The thesis consists of introduction, literature review containing three sections, the chapter on the basics and application of ELEM, discussion of calculations and measurements results, the experimental part, the chapter on development of the new photoemission method of spectral registration to measure the constants of the protolytic equilibria, conclusions, list of references (720 references). The document contains 396 pages of typewritten text, 66 tables, 33 figures, 25 chemical schemes, and 180 mathematical equations.

BODY OF THE WORK

I. Reactivity and acid-base properties of organic compounds. Methods of calculation and experimental determination.

In the first chapter of the analytical review the concepts of reactivity and regioselectivity of organic compounds found in modern scientific literature are examined. The first section outlines methods for determination of reaction rate with the known mechanism based on analysis of the reaction coordinate on the potential energy surface. Methods based on density functional theory are analyzed in details. The methods of quantum molecular dynamics, as well as the group of *Monte-Carlo* statistical methods are reviewed. We have made examples of modifications of the transition state theory used for calculation of the potential energy surface and the rates of chemical reactions. New theory of "no barrier" is described in details. Various prognostic procedures for mechanisms of chemical reactions on the basis of molecular mechanics are investigated.

The second section reviews the current state of correlation analysis.

In the third section the theoretical basis is analyzed and examples of quantum methods of reactivity descriptors are made.

In the fourth section analysis of QSPR methods (quantitative relationships of structure-property), a leader by the number of reports about reactivity evaluation in the scientific literature, is carried out. Examples of methods for computation of the correlation dependences "structure-property" and "property-property" are given. In conclusion the chapter provides a list of actual practical problems that are not solved by modern computational methods.

The second chapter of analytical review outlines the basic modern concepts about acid-base properties of organic compounds and their relation to reactivity. The methods of calculation of the organic compounds acidity in various media are described. Methods that showed the best results of calculations in practice are noted. The conclusion to the second chapter founds the statement that the theoretical

calculations of the protolytic equilibrium constants are far from perfect, despite of elementary quality of proton transfer reactions.

In the third chapter of analytical review the main methods of experimental determination of dissociation constants of organic acids in solutions are analyzed. The classification of solvents, the existing standards and the scale of acidity in them is described. We submit the list of the main problems of experimental measurements of the acidity constants in non-aqueous media: the homo- and hetero-association of ions, methodological errors in calibration of measuring systems, deficiently reliable determination of the acidity scales in solvents, except methanol, ethanol, acetonitrile, DMSO and partly DMF.

In conclusion of analytical review we prove the necessity for research of new methods of calculation and evaluation of reaction constants of organic compounds. Against the background of the existing numerous and varied experimental data, the empirical approach to analysis and prediction of the rate constants and chemical reaction equilibria is of particular importance, since its use requires strict systematization and generalization of experimental data. Based on analyzed theoretical material we suggest that in order to successfully solve the fundamental problem of the quantitative determination of reaction constants of organic compounds there are all prerequisites and necessary components. Based on numerous examples of certainty value of theoretical predictions and procedures of calculation of reaction constants we show that it’s appropriate to carry out the testing of new methods by the example of calculation of the protolytic equilibrium constants. In this area numerous acidity constant values of different organic compounds in a variety of media - protic, aprotic, polar and non-polar solvents - have been accumulated.

II. ELEMENT LINEAR EMPIRICAL METHOD FOR CALCULATION OF PHYSICOCHEMICAL CHARACTERISTICS. BASIC ESTIMATED EQUATION.

The basis for the practical application of this method is the statement that for some pair of physicochemical characteristics X and Y of organic compound in a line of compounds a linear dependence between deviations (δX, δY) of experimental values of these quantities X_{ex} and Y_{ex} and the sums of atomic contributions to these values ($\Sigma g_i X_i$, $\Sigma g_i Y_i$) exists. Mathematically this statement can be expressed by the following equations:

$$\delta X = X_{ex} - \sum_i g_i X_i \qquad (1)$$

$$\delta Y = Y_{ex} - \sum_i g_i Y_i \qquad (2)$$

$$\delta Y = a \cdot \delta X + b\,, \qquad (3)$$

where a and b – constant coefficients, g_i – quantity of i-th atoms in the molecule of the compound (by gross formula), X_i, Y_i – contributions of i-th atom to the observational molecular parameter. The equations (1)-(3) provide a formal mathematical representation of *the principle of linearity of nonlinear effects*, which is one of the approximations of the proposed method (see Chapter IV *"The theoretical positions and mathematical tool of ELEM"*).

The following provision of the method - the atomic contributions (X_i, Y_i) are independent of the state of hybridization of the atom and its belonging to the functional groups, and are determined only by the charge of its nucleus. This approximation has great practical importance, since it allows reducing the number of empirical parameters needed for calculation of the target physicochemical quantity.

It is important that the presence of marked correlation between the experimental values X_{ex} and Y_{ex} in the line of investigated compounds is not necessary for the

application of the method. The equations (1)-(3) are special cases of more general mathematical model of the method, making possible to consider nonlinear effects by means of the principle of polylinearity of V.A. Palm, which is described below, in chapter IV. Based on these equations three main estimated equations of the method (4)-(6) allowed to realize the main purpose of the work - calculation of physicochemical characteristics - were obtained.

The above model was used to approximate and calculate the values of one thermodynamic parameter using the values of other one for 1055 organic compounds with different structures, proton affinity values of 395 compounds, ionization potentials, polarizabilities, and enthalpy of atomization of 78 CHNO-containing organic compounds via molar volume. These calculations were performed by the general equation (4) in minor modifications in each particular case:

$$Y = a \cdot (X_{ex} - A_X \sum_i g_i X_i) + b + B_Y \sum_i g_i Y_i \,, \qquad (4)$$

where a and b – constant coefficients, X_{ex} – experimental physicochemical quantity, related with the target value; Y – the target estimated physicochemical quantity, X_i, Y_i - the atomic contributions to i-th atom in the corresponding molecular value, g_i – quantity of i-th atoms in the molecule by gross formula, A_X, B_Y – calibration coefficients, constant for large groups of considered organic compounds and selected pair of physicochemical parameters.

Balance of the parameters in the equation (4) is very significant in terms of physical sense. The absolute values of all components: aX_{ex}, $aA_X \sum g_i X_i$, $B_Y \sum g_i Y_i$, and b must be a value of the same order of magnitude with the target value Y, and the ratio X_{ex} to $A_X \sum g_i X_i$ and Y to $B_Y \sum g_i Y_i$ - in the limit of 0.1 - 10 in absolute magnitude. If these relations are more than 10, the atomic contributions (X_i, Y_i) are practically insignificant, and if the relation is less than 0.1, it will lead to high calculation error in the target value Y.

Another modification of the method is used for calculation of 363 values of dissociation constants of 33 benzoic acid and acetic acid derivatives in 11 solvents,

192 values of dissociation constants of 24 protonated amines and nitrogen heterocyclic compounds in 8 solvents. In these calculations the target quantity - pK_a – is related with the values of experimental parameters of the solvents by the equation (5):

$$pK_a^{jk} = a_k \left(X_{exj} - c_k \sum_i g_{ij} X_i \right) + Y_j + b_k , \qquad (5)$$

where pK_a^{jk} – the dissociation constant of *k*-th organic compound in *j*-th solvent; X_{exj} – the known molecular parameter of *j*-th solvent: molecular volume (V_M), the enthalpy of protonation in gas phase (PA) or Dimroth-Reichardt solvatochromic parameter (E_T^{30}). g_{ij} – stoichiometric coefficient of *i*-th atom in the gross-formula of *j*-th solvent; a_k, b_k – empirical coefficients related to *k*-th compound, c_k – calibration coefficient, X_i – constant empirical contribution of *i*-th atom to the observational molecular parameter of the solvent; Y_j –constant of *j*-th solvent. During the research of the above empirical coefficients by approximation of the experimental data by the equation (5) we paid attention to the balance of parameters (especially to c_k value) by analogy with the equation (4).

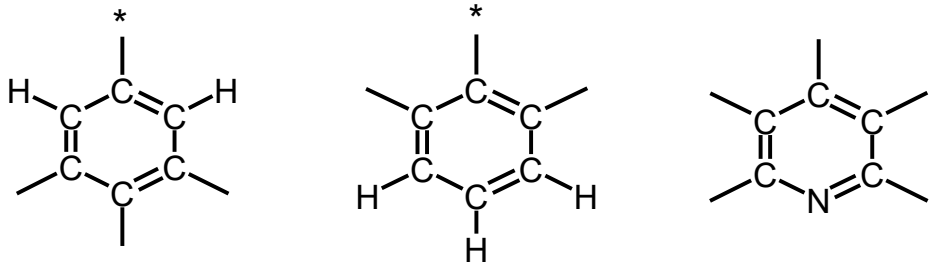
II.1. *The principle of the element of permanent structure*

Studies of organic reactions mechanisms and relations such as "structure-property" indicate that the nearest atoms to the reaction center have more significant effect than the substituting groups at the periphery of the molecule.

In terms of practical application of ELEM this gives grounds to cluster organic compounds on the basis of similarity of the structure – the presence of the common element of the molecule structure; this element has one or several investigated reaction centers. For example, all derivatives of aniline have *the element of permanent structure* (EPS) - benzene ring and amino group. The main purpose of the grouping of compounds is more accurate approximation of values of the target physicochemical parameter within the groups and reducing of the number of

necessary empirical parameters. The more atoms are included in the element of permanent structure, the more aligned influence of all atoms with the same nuclear charge at the periphery of the molecule on the reaction center.

Reaction center * **- OH; -NH_2 (NH_3)$^+$; -COOH**
Реакционный центр * – OH; -NH_2 (NH_3^+); -COOH



ЭПС с *мета-* и *пара-* заместителями | **ЭПС с *орто-* заместителями** | **ЭПС производных пиридина**

ЭПС карбоновых кислот с заместителями в α-положении | **ЭПС карбоновых кислот без заместителей в α-положении** | **ЭПС непредельных карбоновых кислот** | **ЭПС аминов, аминоспиртов**

EPS with *meta*- and *para*-substituents | EPS with orto-substituents | EPS of pyridinium compounds

EPS of carboxylic acids with substituents in the *α*-position | EPS of carboxylic acids without substituents in the *α*-position | EPS of unsaturated carboxylic acids | EPS of amines, amino alcohols

Figure 1. Examples of EPS used in calculation of pK_a values in water.

All above provisions of ELEM were used in calculation of the acidity constants values in water *via* molecular volume of organic compounds. The calculations were performed according to the equation (6), representing a combination of the above principles of linearity of nonlinear effects and EPS without decomposition of the target pK_a value into atomic constituents:

$$\mathrm{p}K_a^j = a \cdot \left(V_j + \sum_i \left(g_{ij} - g_i^* \right) \cdot X_i \right) + b\,, \qquad (6)$$

where pK_a^j – ionization constant of *j*-th organic compound in water, V_j – the value of its molecular volume; g_{ij} – stoichiometric coefficient of *i*-th atom in the gross-formula of this compound, g_i^* – stoichiometric coefficient of *i*-th atom in the gross-formula of the element of permanent structure; X_i – constant empirical coefficient of *i*-th atom, which includes a linear combination of values of atomic volume V_i and atomic contribution to the ionization constant.

In contrast to the equation (5), the parameters a, b are constant within a group of compounds with common EPS. This led to the dramatic reduction in the number of empirical parameters necessary for calculation of the target value with the same amount of initial experimental data compared with the number of empirical parameters in the equation (5). Examples used in this study EPS are shown in Fig. 1.

The unknown empirical parameters in the calculation equations (4)-(6) we found by minimizing of the sum of squares of the relative discrepancy between the experimental Y_{ex} and calculated values of the target magnitude (εY). The explicit form of the equation system that solution is carried out by numerical techniques is given in Chapter IV.

The procedure of calculation of the unknown values of the target physicochemical magnitude by the calculation equations (4)-(6) on the basis of experimental data is described by the following algorithm:

1. Compilation of the representative array of reliable experimental values of the target physicochemical parameter Y_{ex}, that measurements are performed under the same conditions: temperature, pressure, aggregate state.
2. Selection of physicochemical parameter (X_{ex}) with known values and unknown relation with the target physicochemical parameter Y_{ex} in the explicit form.
3. Compilation of the similar representative array of reliable experimental data by the selected quantity *(*X_{ex}*)***.** Grouping of the experimental data of target *(*Y_{ex}) and known *(*Y_{ex}) quantities in the array for approximation by the equations (4) or (5).
4. Selection as the atomic parameters (X_i и Y_i) experimentally determined atomic contributions for the quantities *X* and *Y*, if they exist.

5. If the principle of EPS is not used, the following actions are performed:
   A. To solve the system of linear equations in order to find the empirical parameters *a, b, c,* A_X, B_Y. If the article "4" is not used, then to find the values X_i and Y_i.
   B. To calculate the target values *Y* using the obtained parameters by the relevant equations (4) or (5).

If the EPS principle is used, the sequence of actions is as follows:
   A. To group organic compounds by the principle of structure similarity and to choose EPS for each group.
   B. To solve the system of linear equations in order to find parameters *a, b*, and combined values X_i for each group.
   C. To calculate atomic parameters X_i independent on the type of EPS by the magnitudes of the combined values of parameters X_i.
   D. To calculate target *Y* values by the values *a*, *b,* and X_i by the equation (6).

6. Validation of calculations by the values of the target physicochemical parameter, that did not belong to the original array. It is desirable to obtain values for the test experimentally, not by exception when compiling the original array of experimental data.

7. When using EPS it's necessary to valid the calculations by the equation (6) of the values of target physicochemical parameters *Y* for compounds that were not included in any of the formed groups.

III. EXAMPLES OF APPLICATION OF THE METHOD AND DISCUSSION OF RESULTS

According to the equation (7), which is analogous to the equation (4), without the coefficients A_X and B_Y, the values of enthalpies of atomization, polarizabilities and ionization potentials of 78 CHNO-containing organic compounds by their molecular volume were calculated:

$$Y = a \cdot (V_{ex} - \sum_i g_i V_i) + b + \sum_i g_i Y_i \,. \qquad (7)$$

V_{ex} – value of molecular volume of compound, V_i – atomic volumes calculated by the covalent radii of atoms, g_{ij} – stoichiometric coefficient of *i*-th atom in the gross-formula of this compound.

Table 1. The values of coefficients a and b, as well as correlation coefficients r for calculation by the equation (7) for certain classes of organic compounds.

Value	Enthalpy of atomization			Polarizability			Ionization potential		
Compounds (quantity)	*a*, eV/Å^3	*b*, eV	|*r*|	a_η	*b*, Å^3	|*r*|	*a*, eV/Å^3	*b*, eV	|*r*|
Alkanes (7)	-0.508	23.7	0.9999	-0.053	2.63	0.9987	-1.62	75.2	0.9999
Alkenes (4)	-0.516	26.7	0.9999	-0.050	2.66	0.9984	-1.63	91.5	0.9999
Alkines (4)	-0.498	21.5	0.9999	-0.051	2.73	0.9994	-1.58	83.4	0.9999
n-alcohols (10)	-0.488	10.0	0.9999	-0.043	0.83	0.9996	-1.52	27.6	0.9998
iso-alcohols (3)	-0.506	13.4	0.9999	-0.048	1.17	0.9999	–	–	–
Esters (3)	-0.498	22.6	0.9989	-0.040	0.62	0.9999	-1.58	68.7	0.9989
Ketones (3)	-0.492	14.5	0.9999	-0.054	2.32	0.9995	-1.56	56.1	0.9999
Aldehydes (5)	-0.500	15.1	0.9986	-0.052	2.16	0.9999	-1.67	66.3	0.9994
Carboxylic acids (4)	–	–	–	-0.056	2.22	0.9983	-1.48	35.0	0.9999
Alicycles (3)	-0.649	32.4	0.9999	-0.069	3.71	0.9998	-1.91	89.5	0.9999
Aromatic (10)	-0.536	15.6	0.9940	-0.048	2.14	0.9898	-1.68	87.3	0.9930
(3) (solid)	-0.951	109	0.9963	-0.101	13.3	0.9961	-3.03	392	0.9961
Azines (2) (liquid)	-0.493	7.39	–	-0.087	6.74	–	-0.786	-35.0	–
(3) (solid)	-0.185	-25.5	–	–	–	–	-0.514	-61.3	0.9749
Azoles (3) (solid)	-0.311	-7.45	0.9999	–	–	–	-0.625	-43.7	0.9986
Alicyclic imines (4)	-0.533	11.7	0.9995	–	–	–	-1.66	33.6	0.9998

The parameters *a* and *b*, listed in Table 1, were constant within these groups of compounds. All compounds were divided into groups (Table 1) according to the principles of similarity of functional group or structure. The correlation coefficients of obtained relations were more than 0.997.

The thermodynamic quantities: the standard enthalpy (ΔH_f), entropy (ΔS_f), Gibbs energy (ΔG_f) of formation and isobaric heat capacity (C_p) - 1055 organic compounds were mutually correlated (an array of some experimental values with an array of others) by the equation (8), analogous to the equation (4), but without the coefficient B_Y:

$$Y = a \cdot (X_{ex} - A_X \sum_i g_i X_i) + b + \sum_i g_i Y_i \ , \qquad (8)$$

where X, Y – couple of the above mentioned standard thermodynamic parameters. As the atomic contributions X_i, Y_i the corresponding thermodynamic quantities are used for elements in gas phase.

The coefficient A_X was introduced into the equation (8) to improve the correlation between the values of δX and δY in the initial equations (1)-(3). As a result, 1010 values ΔH_f, ΔS_f, ΔG_f and C_p for 365 organic compounds are calculated; experimental data for these compounds are not in the literature.

In order to apply the method in this case it was important to observe the following conditions:

(1) a number of compounds should not be homologous;

(2) correlation coefficient between the values of atomic amounts $\sum g_i X_i$ and $\sum g_i Y_i$ should not be equal to 1.

For this purpose selected from the NIST database compounds were combined into 9 groups for that these conditions were met: 1 - alkyl derivatives of benzene, 2 - halogen, hydroxy, alkyl derivatives of benzene, 3 - all derivatives of benzene, 4 - carboxylic acids and their derivatives (amides, nitriles, esters), 5 – six-membered heterocycles with 1 and 2 nitrogen atoms; 6 - acyclic monohydric alcohols, thiols, esters, sulfides, 8 - cyclic ethers, sulfides, 9 - all alcohols, ethers, thiols, sulfides, 10 - monohydric alcohols, ethers, aldehydes, ketones of the group C_1 – C_8. It should be noted that some compounds fall into several groups. As a result, for the selected couple of thermodynamic quantities within groups the correlation coefficients between δX and δY by the equation (3) were no less than 0.983. Examples of empirical parameters for the group of alcohols, ethers, thiols and sulfides are given in Table 2.

Table 2. Examples of the values of parameters a, b, A_X and correlation coefficients r in the equation (8) for the pairwise correlations of thermodynamic quantities.

Pair of values $Y-X$	Values of the parameters			Correlation coefficient, r	The number of pairs of known empirical values
	a	b	A_X		
$\Delta C^g - \Delta G^g$	0.070	-8.77	0.54	0.9935	11
$\Delta S^g - \Delta H^l$	0.387	50.3	0.76	0.9836	24
$C_P^g - \Delta H^l$	0.037	-11.9	0.89	0.9901	18
$\Delta S^g - \Delta H^g$	0.383	48.1	0.78	0.9834	27
$C_P^g - \Delta H^g$	0.038	-12.9	0.87	0.9908	21

Calculation of the protonation enthalpies in gas phase (PA) by molecular volume (V_M) of 395 organic compounds was carried out by the formula (4). The coefficients A_X, B_Y, a, b are constant for four groups of organic compounds with maximum error no more than 4% (Table 3).

In the literature generally for the assessment of "accuracy" of the calculation method the correlation coefficient between the calculated and experimental values is used. In this work, as a criterion of accuracy the average (ε_{av}) and the maximum (ε_{max}) relative deviations of calculated values from experimental ones are used. This is done in connection with the fact that when calculation of PA by V_M, as well as in some works of other authors, high correlation coefficient did not give an adequate representation of the estimated error. Thus, at r = 0.95 the relative deviation of calculated values from experimental ones can reach 30%, that makes a correlation dependence unsuitable for practical use.

Table 3. Groups of compounds, the number of compounds in the group N, coefficients a, b, A_X, B_Y, and average relative deviations of calculated protonation energies from experimental data when calculations by the equation (4).

The structure of the group		a	b	A_X	B_Y	ε_{av}, %
Acids without phenyl fragment, alcohols, ketones, aldehydes, esters	37	-0.0108	7.62	-9.65	0.0710	2.0
Alcohols, thiols, ethers, sulfides	50	-0.00750	7.72	-2.12	0.0393	3.9
Pyridine derivatives	22	-0.0165	9.45	-0.287	0.0515	1.0
Substances with the nitrile group	18	-0.00482	8.00	-10.4	0.0344	2.3

III.1. *Calculations of* p*Ka values of organic acids in non-aqueous media*

The problem of calculation of acidity constants in various media includes consideration of the molecular composition of the solvent; it is significantly complicated by the presence of specific solvation. The principle difference of calculations by the equation (5) from calculations by the equation (4) in its various modifications is that the equation (5) relates the molecular parameters referring not to one substance but to different ones.

Despite of the fact that in the literature many experimental pK_a values varied by the structure of compounds in different solvents were found, common combination of data of different works into a single array for processing and analysis is not correct for the following reasons: incoordination of methods and measurement conditions between various sources of experimental data; incorrect calibration of measurement systems and methodic errors associated with misunderstanding of the system of protolytic equilibria and lack of the clearly defined scale of acidity ($pa^*_{H^+}$) especially in low-polar protophobic solvents. This leads to the situation when the pK_a values of one compound in one solvent from different works can vary by 5 log. units.

Therefore, for carboxylic acids in the basic array of experimental data to approximation of ELEM the values of pK_a of substituted benzoic and acetic acids in series of 12 and 5 polar solvents, respectively, were selected. These data were obtained mainly by the same authors and supplemented by recent data obtained by the single method (potentiometric titration) [*Bartnicka H. et al.* // Austr. J. Chem. - 1991. - Vol. 44, N. 8. - P. 1077-1084; 1993. - Vol. 46, N. 1. - P. 31-36]. For amines and nitrogen heterocyclic compounds the considerable array of data obtained by the same authors at the same conditions was absent in the literature, for this reason the selection of data for processing presented an independent analytical task included an assessment of the reliability of available data on the basis of the used experimental technique, and statistical processing with the presence of pK_a values for one compound in several sources.

In the final array for processing of ELEM we included carboxylic acids with 5 or more experimental values of dissociation constants in various solvents, and amines and nitrogen-containing heterocyclic compounds with 4 or more values.

To calculate the pK_a values of organic acids in different media the principle of linearity of nonlinear effects was used. Accordingly, approximation and calculations of the pK_a values of carboxylic acids, amines and nitrogen-containing heterocyclic compounds were carried out by the equation (5). As a result, for all 555 calculated pK_a values of organic acids in different solvents by the equation (5) the relative deviation from the experimental values did not exceed 10%.

For carboxylic acids proton affinity (PA) and Dimroth-Reichardt solvatochromic parameter (E_T^{30}) were used as a molecular parameter of the solvent. For amines and nitrogen heterocyclic compounds molecular volume (V_M) and E_T^{30} were applied.

As a result of calculations the atomic contributions X_i were equal when using both molecular parameters of solvents for one type of organic acids, but they appeared different for various types (carboxylic acids and amines). When use of V_M as a molecular parameter of the solvent it was expected that successful approximations for X_i would be the values of volume of the corresponding atoms, but the results of calculations did not prove it. Therefore, the atomic contributions to the parameters of the solvent were selected by the empiric way. Determined optimal values of the atomic contributions X_i, corresponding to C, N, O, N and S atoms for carboxylic acid, as well as for amines and nitrogen-containing heterocyclic compounds are listed in Table 4.

Table 4. Optimal values of atomic contributions X_i of C, N, O, N and S atoms to the parameters of the solvent (PA, E_T^{30}) for carboxylic acids and to the parameters of the solvent (V_M, E_T^{30}) for amines and nitrogen-containing heterocyclic compounds.

Atom (i)	X_C	X_H	X_O	X_N	X_S
Carboxylic acids ($X_{ex} \equiv$ PA, E_T^{30})					
X_i	1.97	-1.75	2.30	-2.85	0.76
Amines and nitrogen-containing heterocyclic compounds ($X_{ex} \equiv$ PA, V_M)					
X_i	1.88	1.22	-1.75	-6.10	0.50

Optimal values of the constants of solvents Y_j for carboxylic acids are shown in Table 5, and for amines and nitrogen-containing heterocyclic compounds - in Table 6. As can be seen from Tables 5 and 6, the differences between the values of these constants are much smaller for amines and nitrogen-containing heterocyclic compounds in comparison with carboxylic acids.

Values of the constants of solvents Y_j and atomic contribution of oxygen X_O have the greatest influence on the relative error of dissociation constant (εpK_a). The obtained values of coefficients a_k and b_k are not identical for all 33 acids (22 derivatives of benzoic acid and 11 acetic derivatives) as expected, but they are of the same order, furthermore the ratio a_k to b_k is relatively constant. For carboxylic acids turndowns of parameters are: a_k = 0.42 ÷ 0.64, b_k = -49.5 ÷ -62.4, c_k = -7.71 ÷ 4.74 для E_T^{30}. For PA: a_k = 0.0210 ÷ 0.0457, b_k = -15.5 ÷ -30.5, c_k = 8.4 ÷ 34.9. The smallest value c_k = 8.4 corresponds to propionic acid. For the amines and nitrogen-containing heterocyclic compounds the parameters are within: a_k = -0.019 ÷ -0.423, b_k = -2.4 ÷ 25.9 c_k = -1.32 ÷ 22.9 for E_T^{30}. For V_M parameter of the solvent: a_k = 0.109 ÷ 0.178, b_k = -2.4 ÷ -12.2, c_k = 1.57 ÷ 4.08. In this case the anomalous c_k value corresponds to codeine (22.9).

Table 5. Constants Y_j of solvents, used for calculations of pK_a of carboxylic acids by the equation (5), when selection of PA and E_T^{30} as experimental parameters of solvents.

j	Solvent	Y_j	
		$X_{ex} \equiv$ PA	$X_{ex} \equiv E_T^{30}$
1	Acetonitrile	15.41	57.8
2	Acetone	13.62	44.4
3	DMF	5.35	41.4
4	DMSO	5.82	33.2
5	Propylene carbonate	11.90	66.1
6	Water	4.28	23.8
7	Methanol	8.70	28.9
8	Formamide	0.60	37.8
9	Benzonitrile	11.60	77.9
10	Nitromethane	11.80	61.2
11	Ethanol	9.60	27.2

Table 6. Constants of solvents Y_j, used for calculation of pK_a of amines and nitrogen-containing heterocyclic compounds by the formula (5), when selection of V_M and E_T^{30} as experimental parameters of solvents.

j	Solvent	Y_j	
		$X_{ex} \equiv V_M$	$X_{ex} \equiv E_T^{30}$
1	Acetone	8.77	8.21
2	DMSO	5.91	6.01
3	Acetonitrile	12.7	11.8
4	Nitromethane	9.82	9.75
5	Ethanol	8.86	9.63
6	Methanol	10.7	10.5
7	Formamide	6.36	7.73
8	Water	11.1	11.1

The reliability of the ELEM calculations of pK_a values by the formulas (5) and (6) was evaluated in the following ways:

- comparison of the calculated values with the values experimentally obtained in this work;
- comparison with literature values obtained by other calculation methods for the same organic compounds in similar conditions;
- comparison with the calculated values for a series of carboxylic acids (20 compounds) in the same solvents (water, DMSO, methanol); those values were obtained in this work by the DFT method by means of the program Jaguar 7.0;
- comparison with the values calculated in this work by the method of correlation analysis (by known σ-constants) for all *meta*- and *para*-substituted phenols in water investigated by ELEM.

The relative error of the calculation less than 10% of log. units was adopted as the criterion of estimation reliability.

In 25 of 53 studies published in the period from 1998 to 2009 the calculated pK_a values of different organic compounds were obtained by calculation of geometry and thermodynamic parameters of free molecules by quantum-chemical methods *ab initio* within the scope of the theory of the Hartree-Fock self-consistent field or DFT theory with the following consideration of the influence of the solvent by a polarized dielectric continuum in modifications (PCM, CPCM, COSMO, IPCM, SCRF). At the

same time in 5 studies the initial data or the results of calculations were additionally processed by empirical methods (including QSPR), and in 21 work the QSPR method in modern versions (ComSA, ComFA, QTMS, 3D-QSPR) was used for pK_a calculation independently, without *ab initio* calculations of geometry of the studied compounds molecules. The structure of solvate complexes of the studied compounds with 1-2 molecules of the solvent was considered only in 5 of these works. In the above noted literature 4886 values of pK_a calculated for 2382 substances in water and non-aqueous solvents are given; 3004 values are calculated with a relative error of less than 10%.

Comparison of calculations of pK_a values of carboxylic acids, amines and nitrogen heterocyclic compounds obtained by ELEM by the equation (5) with values obtained by another methods published in the literature by the second half of 2009, are listed in Table 7.

Table 7. Comparison of the results of literature pK_a *calculations for carboxylic acids and amines, performed by different methods, with the results of calculations by the ELEM equation (5).*

Designations of columns:

1. Solvents used for pK_a calculations.
2. Calculation model of geometry optimization and consideration of the solvent effects.

If only a model of the influence of the solvent is noted, it means that the quantum-chemical calculation of geometry and energy of the molecule without the solvent was carried out by various methods within the scope of Hartree-Fock SCF theory. In the other cases, the following designations are used: DFT - quantum-chemical calculation carried out by FEC, PCM - the method of polarized dielectric continuum, CPCM – the method of conducting polarized dielectric continuum, IPCM - the method of integral conductive dielectric continuum, COSMO - *conductor-like screening solvation model*, PBSM – the model of solvation of Poisson-Boltzmann, SCRF – the method of self-consistent reaction field. Cluster – quantum-chemical calculation, taking into account geometry of the complex "solvate-solvent", QSPR - the method "structure-property", which was used in processing of results or (and) in calculation of solvation energies or deprotonation of solvated objects, CA - correlation analysis, RI – reactivity indices, SE - descriptors calculated by semi-empirical methods, CNN - used neural networks, ComSA - comparative analysis of molecular surfaces (3D-QSPR), ComFA - comparative analysis of molecular fragments (3D-QSPR).

3. The total number of compounds studied by the method noted in the column 2. In parentheses - the number of pK_a values calculated for them in different ways.

4. The number of calculated pK_a values, meeting the criterion εpK_a <10%. If the number of these values exceeds the half of all calculated ones, it is indicated in bold type.

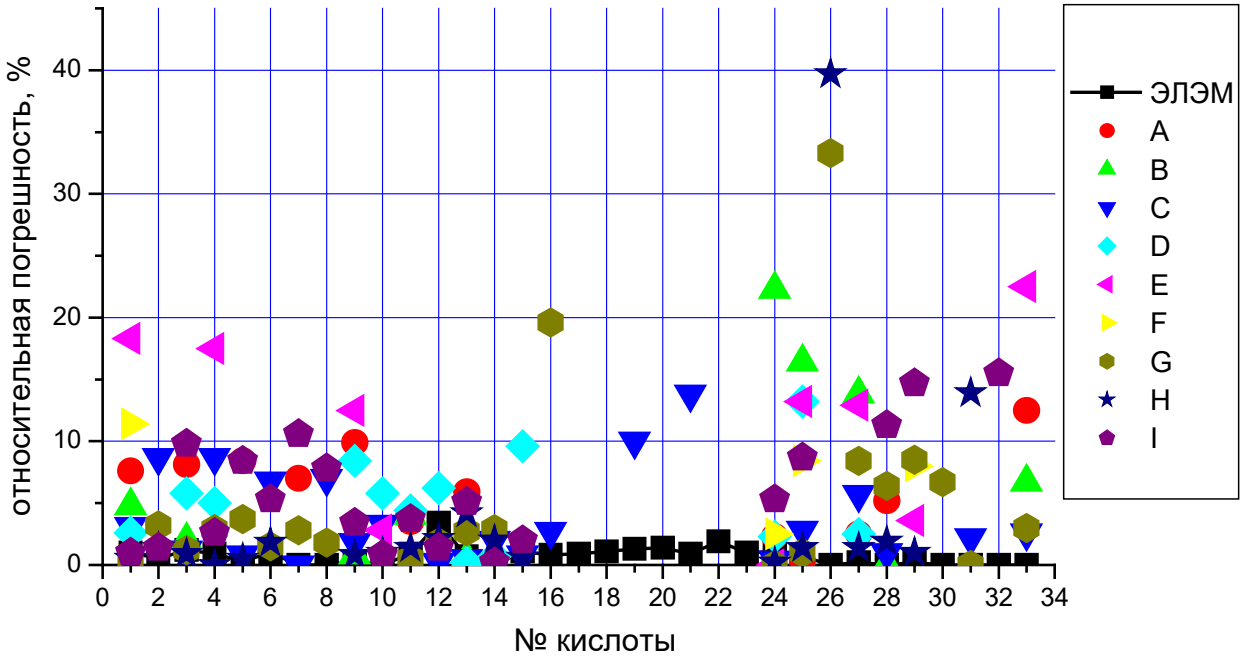
5. The number of compounds examined by the method indicated in the column 2; for these compounds the pK_a calculations using ELEM were carried out. In parentheses - the total number of pK_a values calculated for these acids in different ways.

6. The number of pK_a values from the column 5, for that a relative error of the calculation by the method noted in the column 2 is lower than the error obtained in this work.

1	2	Type of compounds	3	4	5	6
Water	CPCM, QSPR	Substances of different structure	64(64)	**42**	8(8)	0
Water	CPCM	Carboxylic acids	6(183)	**115**	2(33)	0
Water	CPCM	Carboxylic acids	28(56)	10	7(14)	3
Water	PCM	Carboxylic acids	66(132)	60	13(26)	3
Water	PCM	Carboxylic acids	66(132)	20	6(12)	0
Water	CPCM	Carboxylic acids	6(27)	15	3(13)	1
Water	PCM	Carboxylic acids	6	3	3	0
Water	CPCM	Chloroacetic acids, nucleotides	9(60)	0	1(7)	0
Water	CPCM, cluster, QSPR	Carboxylic acids, phenols and pyridinium salts	64(64)	**53**	14(16)	4
Water	Cluster, QSPR	Carboxylic acids, phenols, alcohols	57(57)	11	1(1)	1
Water	PCM, QSPR	Carboxylic acids	16(32)	**17**	4(8)	1
Water	PCM	Different acids	7(7)	**4**	4(4)	0
Water	PCM	Acids, alcohols, thiols	14(14)	5	4(4)	0
Water	PCM	Carboxylic acids	8(16)	3	4(8)	2
Water, DMSO, MeCN	PCM	Carboxylic acids, phenols, amines	8(144)	23	3(54)	2
Water	PCM	Amines, imines, heterocycles	25(72)	18	5(15)	1
Water	PCM, cluster	Substances of different structure	17(17)	5	2(6)	0
DMSO	Cluster	Substances of different structure	106(212)	98	8(16)	0
DMSO	IPCM	Substances of different structure	36(36)	12	1(1)	0
DMSO	PCM	Acids, alcohols and amides, imines	42(42)	**28**	2(2)	1
Different	QSPR, CNN	Carboxylic acids, anilines	44(266)	261	38(167)	69
Different	QSPR, CNN	Benzoic acids	138(519)	501		
Water	QSPR	Carboxylic acids, amines	76(76)	**57**	13(3)	5
Water	QSPR	Anilines	36(36)	**28**	8(8)	3
Water	PCM, SE	Pyridines	15(15)	7	1(1)	1
Water	SE, QSPR	Anilines	63(126)	49	12(24)	0
Water	RI, CA	Substances of different structure	64(64)	44	19(19)	0
Water	Cluster	Carboxylic acids	49	**44**	17(17)	5
Water	DFT/SCRF	Substances of different structure	300	**214**	17(17)	4
Water	SE, QSPR	Substances of different structure	430	**311**	24(26)	6
Water	QSPR	Amines	25	**14**	8	0
Water	QSPR	Substances of different structure	33	**33**	3	0
Water	ComSA, CNN	Carboxylic acids	89(352)	**273**	22(70)	4
Water	QSPR QTMS	Carboxylic acids	5(15)	5	3(9)	0
Water	COSMO	Anilines, bioactive molecules	24(384)	53	4(64)	0

It is seen in Table 7 that only 121 of 689 pK_a values, pointed in the literature, is in better agreement with experimental results than those calculated using ELEM by the equation (5).

A comparison of relative errors of pK_a calculations for carboxylic acids in water and DMSO, obtained by ELEM by the equation (5), with errors of values obtained by other methods published in the literature, is presented graphically in Figures 2 and 3.



"относительная погрешность" – relative error; "№ кислоты" – № of the acid

Figure 2. A comparison of relative errors of pK_a calculations for carboxylic acids in water, obtained by other methods, published in the literature with results obtained by ELEM("ЭЛЭМ") by the equation (5).

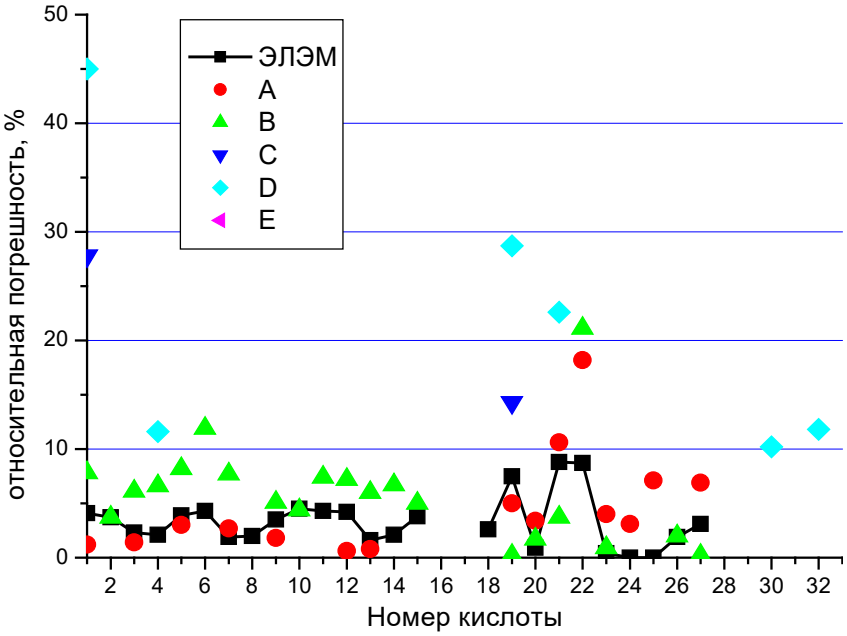
Designations of published data. A. – The method of DFT [*Yu. E. Zevatskii et al.*, Russian Journal of General Chemistry, Vol. 79, No.5, 2009, pp. 772-780.], B. – ["The pK_a prediction module" Jaguar 7.0, Schrodinger L.L.C., Portland O.R., 1991-2003], C. – [*Jover J. et al.,* QSAR Comb. Sci. 2008. Vol. 27. 563.; 1179.], D. – [*Adam K.R.*, J. Phys. Chem,. A. 2002. Vol. 106. P. 11963.], E. – [*Namazian M. et al.*, J. Chem. Thermodyn. 2006. Vol. 38. P. 1495.], F. – [*Klamt A. et al.*, J. Phys. Chem., A. 2003. Vol. 107. P. 9380.], G. – [*Tehan B.G. et al.*, QSAR. 2002. Vol. 21. P. 457.; 473.], H. – [*Tao L. et al.,* J. Phys. Chem. A. 2008. Vol. 112. P. 775.], I. – [*Polanski J. et al.*, J. Chem. Inf. Comp. Sci. 2002. Vol. 42. P. 184].

Designations of carboxylic acids: 1 – PhCOOH, 2 – 2-$NO_2C_6H_4COOH$, 3 – 3-$NO_2C_6H_4COOH$, 4 – 4-$NO_2C_6H_4COOH$, 5 – 3-BrC_6H_4COOH, 6 – 4-BrC_6H_4COOH, 7 – 3-IC_6H_4COOH, 8 – 4-IC_6H_4COOH, 9 – 3-ClC_6H_4COOH, 10 – 4-ClC_6H_4COOH, 11 – 3-MeC_6H_4COOH, 12 – 4-MeC_6H_4COOH, 13 – 3-$MeOC_6H_4COOH$, 14 – 4-$MeOC_6H_4COOH$, 15 – 4-CNC_6H_4COOH, 16 – 2-$MeCOC_6H_4COOH$, 17 – 4-$MeSO_2C_6H_4COOH$, 18 – 4-$NH_2SO_2C_6H_4COOH$, 19 – 2-$HCOC_6H_4COOH$, 20 – 2-$HOOCC_6H_4COOH$, 21 – 2-$MeCOOC_6H_4COOH$, 22 – 2-NOC_6H_4COOH, 23 – 2-$MeSOC_6H_4COOH$, 24 – CH_3COOH, 25 – $ClCH_2COOH$, 26 – $Cl_2CHCOOH$, 27 – EtCOOH, 28 – *i*-PrCOOH, 29 – *t*-BuCOOH, 30 – $CyCH_2COOH$, 31 – $PhOCH_2COOH$, 32 – $Ph_2CHCOOH$, 33 – $PhCH_2COOH$.

Generally the comparison in Fig. 2 shows an advantage of ELEM in accuracy over known calculation methods. However, occasionally the series of organic acids with calculated pK_a reported in the literature, largely aligned with a number of compounds with performed calculations of ELEM (see Table 7). For a more rigorous check of ELEM advantages in calculation accuracy, pK_a values were calculated for a single set of organic compounds in three solvents (water, DMSO, methanol) by the proposed method and another known method. For this purpose in the work we reported pK_a calculations for a number of carboxylic acids different in the structure by the quantum-chemical method of DFT. Acidic dissociation were considered as reactions of proton transfer from the molecule of organic acid HA to the solvent molecule. Calculations of the electronic structure of compounds were carried out by means of software system Jaguar 7.0 by DFT B3LYP method in the basis 6-31G**. Thermodynamic characteristics of equilibria are obtained by calculation of the frequencies of normal modes of compounds on the basis of DFT force fields in the harmonic approximation. The values of the dissociation constants were calculated by the formula $pK_a^{DFT} = \Delta G_{solv}/2.3RT$. Solvation energies of compounds were counted in a polarizable dielectric continuum model. Also, a direct quantum-chemical calculation of pK_a values with the module "*The* pK_a *prediction module*" of the program Jaguar 7.0 was performed. Despite of the over-estimated pK_a values, obtained by the DFT, they generally reflected the trend of the change in the experimental values in a number of the investigated acids. Refined calculation values were obtained by the approximate equation:

$$pK_a^{DFTapp} = a \cdot pK_a^{DFT} + b \,. \qquad (9)$$

a and *b* are: 0.3168 and -1,453 for water, 0.4605 and -0.6746 for methanol, 0.6978 and 3.030 for DMSO, respectively. For a comparison Fig. 2 shows the results of these calculations.



“относительная погрешность” – relative error. “Номер кислоты” – The number of the acid.

Figure 3. Comparison of errors in pK_a calculations for carboxylic acids in DMSO obtained by various methods in the literature, with the results obtained by ELEM(“ЭЛЭМ”) by the equation (5). Designations of published data. A. - The method of DFT [*Yu. E. Zevatskii et al.*, Russian Journal of General Chemistry, Vol. 79, No.5, 2009, pp. 772-780.], B. – [*Jover J. et al.,* QSAR Comb. Sci. 2008. Vol. 27. 563.; *Jover J. et al.*, QSAR Comb. Sci. 2008. Vol. 27. 1179], C. – [*Almerindo G.I. et al.*, J. Phys. Chem., A. 2004. Vol. 108. P. 166-171], D. – [*Fu Y. et al.,* J. Am. Chem. Soc. 2004. Vol. 126, P. 814-822], E. – [*Chipman D.M.*, J. Phys. Chem. A. 2002. Vol. 106. P. 7413-7422].

The numbers of carboxylic acids are as in Fig. 2.

As it is seen in Fig. 3 only some of calculated pK_a values in DMSO, reported in the literature, obtained by various methods, are calculated more accurately than the values estimated by the method ELEM. The accuracy of the calculation of pK_a values in water, performed in the present work by the DFT method in the program Jaguar 7.0, is at the level of other studies used quantum-chemical calculations, and is slightly inferior to the calculations by the QSPR and ELEM methods. Accuracy of calculations of the pK_a values in DMSO implemented in this work by the ELEM methods is comparable with the accuracy of DFT method (Fig. 3).

III.2. *Calculations of pKa values by the EPS principle*

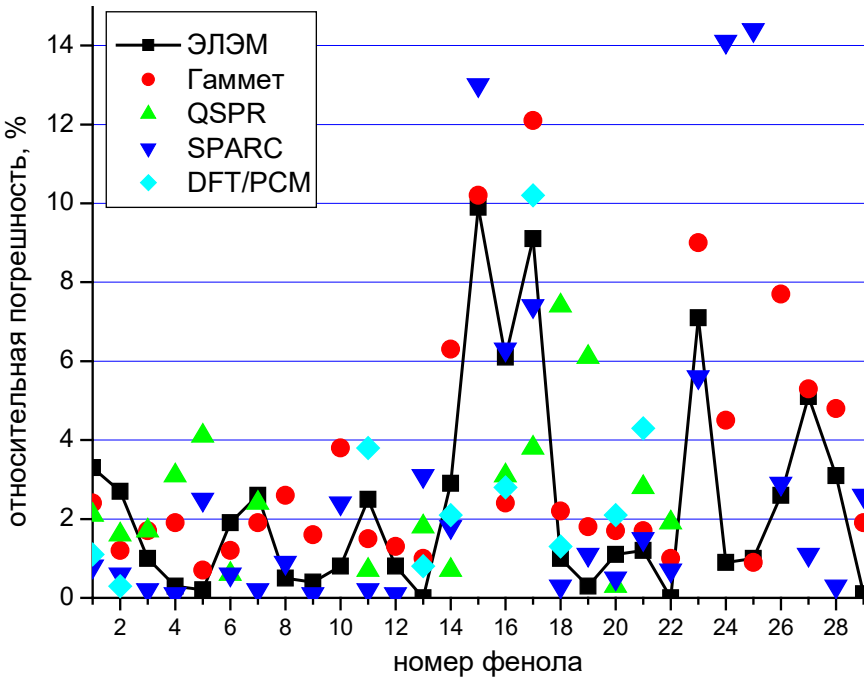
The groups of compounds with the general EPS (see Fig. 1) and empirical coefficients used for calculations of pK_a values in water by the equation (6) are

represented in Table 8.

Table 8. The groups of compounds, the number of compounds in the group N, coefficients a, b, and X_i, as well as the average deviations of calculated values pK_a, pK_b, *and* pK_{BH^+} *from the experimental values.*

Composition of the group	*N*	Linear coefficients		Atomic coefficients X_i, Å^3								ε_{cp}, %
		$a \cdot 10^2$, Å^{-3}	*b*	C	H		N		Cl	Br	I	
Phenols with *meta*- and *para*-substituents	29	-2.388	12.95	-5.30	-16.8	4.94	31.4	-14.4	-28.2	-21.7		2.4
Phenols with *ortho*-substituents	37	-1.670	12.31	1.58	-21.5	29.0	79.9	28.7	23.8	15.9		2.9
Benzoic acids with *meta*- and *para*-substituents	26	-2.037	7.431	4.44	-15.8	-1.70	-3.03	-5.01	-18.9	-20.9	-32.5	3.1
Benzoic acids with *ortho*-substituents	16	4.094	-2.694	-1.44	-16.9	-15.9	-39.9	-30.6	-50.6	-53.8	-62.6	2.4
Anilines and *meta*-*para*-substituents	20	-1.254	11.31	22.4	-16.9	-79.1	-137	-59.6	-92.8	-132		2.2
Ortho-substituted anilines	21	1.872	7.046	116	-76.3	-16.2	178	-36.6	-27.8	-40.1		1.7
Aliphatic carboxylic acids with functional substituents in the α-position	17	-2.475	4.198	27.6	-30.4	-8.19	-31.9	34.7	-15.6	-35.0	-49.9	2.4
Aliphatic carboxylic acids without functional substituents in the α-position	24	-1.691	5.997	-40.0	5.47	35.1				5.23		0.9
Unsaturated aliphatic carboxylic acids	12	-7.975	10.39	13.7	-21.1	-0.66						1.5
Aliphatic amines, amino alcohol and alicyclic amines	34	-2.069	11.00	16.7	-23.7	49.8	33.0					3.0
Pyridine derivatives	23	10.67	-8.629	-22.8	-1.76	0.059			-73.0			3.5

A comparison of relative errors of pK_a values of *meta*- and *para*-substituted phenols in water calculated by the principle of EPS and the equation (6) with the results obtained in this work by the method of correlation analysis (by the known σ-constants), as well as with the results obtained by other methods is shown in Figure 4. The result of comparison shows that for this group of organic acids ELEM exceeds the accuracy of the modified Hammett method, the QSPR method, quantum chemical calculations by DFT method. Although the SPARC program demonstrates the best results (based on the average relative error of calculation), but 3 of 29 pK_a values are calculated with the error of more than 10%.



"относительная погрешность" – the relative error; "номер фенола" – the number of phenols

Figure 4. Comparison of the relative errors of pK_a calculation for different phenols with *meta*- and *para*-substituents in water, published in the literature and obtained in this work by known methods, with the results obtained by ELEM+ EPS ("ЭЛЭМ") equation (6).

Designations of methods:

Гаммет - calculation by the modified Hammett equation, **SPARC –** on-line program for calculations by modern equations of correlation analysis. **QSPR** –work data [*Jover J., et al.*, QSAR Comb. Sci. - 2007. - Vol. 26, N. 3. - P. 385-397], **DFT/PCM –** work data [*Liptak M.D. et al.*, J. Am. Chem. Soc. - 2002. - Vol. 124, N. 22. - P. 6421-6427.]

Designations of phenols:

1 – Phenol, 2 – 3-Methylphenol, 3 – 4-Methylphenol, 4 – 3- Ethylphenol, 5 – 4-Ethylphenol, 6 – 3,4- Dimethylphenol 7 – 3,5-Dimethylphenol, 8 – 4-Propylphenol, 9 – 4-Isopropylphenol, 10 – 4-Butylphenol, 11 – 3- Methoxyphenol, 12 – 3-Ethoxyphenol, 13 – 1,3-Dihydroxybenzene, 14 – 1,4-Dihydroxybenzene, 15 – 1,3,5-Trihydroxybenzene, 16 – 3-nitrophenol, 17 – 4-nitrophenol, 18 – 3-Fluorophenol, 19 – 3-(Trifluoromethyl)phenol, 20 – 3-Chlorophenol, 21 – 4-Chlorophenol, 22 – 4-Bromophenol, 23 – 3-Hydroxybenzoic acid, 24 – 4-Hydroxybenzoic acid, 25 – 3,4-Dihydroxybenzoic acid, 26 – 3,4,5-Trihydroxybenzoic acid, 27 – 4-(Acetylamino)phenol, 28 – 1-Naphthol, 29 – 2-Naphthol.

A comparison of pK_a calculations for phenols, benzoic acid derivatives, anilines, aliphatic carboxylic acids, amines and pyridine derivatives, obtained by ELEM by the equation (6), with the values obtained by another methods, reported in the literature by the second half of 2009, is presented in Table 9.

Summing of numbers in the 6-th column of Table 9 shows, that out of 935 "crossed" pK_a values of different organic compounds in water, only 230 values available in the literature is in better agreement with experimental results than the calculated pK_a values obtained by ELEM using the EPS principle by the equation (6).

Table 9. A comparison of the results of reported pK_a calculations performed by different methods with the results obtained using the EPS principle by the equation (6) of ELEM.

The numbering of the columns, designations of the calculation methods and notes are the same as in Table 7.

1	2	The type of compounds	3	4	5	6
Water	CPCM, QSPR	Substances of different structure	64(64)	**42**	13(13)	3
Water	CPCM	Carboxylic acids	6(183)	**115**	0	0
Water	CPCM	Carboxylic acids	28(56)	10	15(30)	3
Water	CPCM	Carboxylic acids	66(132)	60	30(60)	6
Water	CPCM	Carboxylic acids	66(132)	20	15(30)	5
Water	CPCM	Carboxylic acids	6(27)	15	1(5)	0
Water	PCM	Carboxylic acids	6	3	3	**0**
Water	CPCM	Phenols	20(40)	**36**	14(28)	8
Water	CPCM	Chloroacetic acids, nucleotides	9(60)	0	2(14)	0
Water	CPCM, cluster, QSPR	Carboxylic acids, phenols, and pyridinium salt	64(64)	**53**	(58)	26
Water	Cluster, QSPR	Cl,Br,F-substituted phenols	35(62)	**61**	8(16)	4
Water	Cluster, QSPR	Substances of different structure	57(57)	11	1	0
Water	PCM, QSPR	Carboxylic acids	16(32)	**17**	6(12)	2
Water	PCM	Substances of different structure	14(14)	5	3	0
Water	PCM	Carboxylic acids	7(7)	**4**	3	0
Water	PCM	Carboxylic acids	8(16)	3	2(4)	0
Water, DMSO, MeCN	PCM	Substances with different structure	8(48)	23	3(18)	2
Water	PCM	Amines, imines, heterocycles	25(72)	18	5(15)	1
Water	PCM, cluster	Substances with different structure	17(17)	5	2(2)	0
Water	Cluster	Phenols, imidazole, methanol	5(16)	**11**	3(12)	1
Different	QSPR, CNN	Carboxylic acids, anilines	44(266)	261	67(67)	24
Different	QSPR, CNN	Benzoic acids	138(519)	501		
Different	QSPR, CNN	Phenols	31(276)	**240**	23(23)	10
Water	QSPR	Carboxylic acids, amines	76(76)	**69**	36(36)	14
Water	QSPR	Anilines	36(36)	**28**	11(11)	6
Water	PCM, SE	Pyridines	15(15)	7	6(6)	2
Water	SE,QSPR	Anilines	63(126)	49	28(56)	4
Water	RI, CA	Substances with different structure	64(64)	44	26(26)	1
Water	CA	Phenols	15(15)	12	3(3)	2

III.3. *A comparison of calculated* pK_a *values with the values experimentally obtained in this work*

The results of experimental determination of acidity constants in this study (pK_a^{ex}), the calculated values for the compounds obtained by ELEM by the equation (5) using E_T^{30} as a parameter of the solvent (p$K_a^{(5)}$), as well as reported experimental data (pK_a) are given in Table 10.

*Table 10. Calculated (*p$K_a^{(5)}$*), literature (*pK_a*) and experimental (*pK_a^{ex}*) values of* pK_a *(at 25 ° C) for compounds investigated in this work.*

Carboxylic acids	Methanol				Ethanol				Formamide			
	p$K_a^{(5)}$	pK_a^{ex}	pK_a	ε, %	p$K_a^{(5)}$	pK_a^{ex}	pK_a	ε, %	p$K_a^{(5)}$	pK_a^{ex}	pK_a	ε, %
2-$NO_2C_6H_4COOH$					8.4	8.25	8.27	2.3	4.08	4.46	4.28	8.5
4-IC_6H_4COOH					10.5	9.50	9.70	10.0				
3-$MeOC_6H_4COOH$					10.8	10.25	10.1	5.7				
4-$MeOC_6H_4COOH$					11.4	10.61	10.6	7.8				
3-CNC_6H_4COOH					9.3	9.25	9.08	0.8	5.69	5.32		6.9
2-$MeCOC_6H_4COOH$					9.8	10.10	10.1	3.2	5.68	5.50		3.2
4-$MeSO_2C_6H_4COOH$					9.4	9.45	9.13	0.8	5.34	5.57		4.2
2-$HCOC_6H_4COOH$					10.3	10.00		2.7	6.09	6.07		0.4
2-NOC_6H_4COOH					8.8	9.00		1.7	5.86	6.45		9.1
CH_3COOH									7.49	7.45		6.9
$ClCH_2COOH$									5.01	3.82	4.74	31.2
C_2H_5COOH	11.7	10.03	9.71	16.8	13.6	10.55	10.6	28.5	6.12	7.10	7.18	13.8
$(CH3)_2CHCOOH$	11.0	10.67	9.90	2.7	12.2	10.79		13.1	6.84	7.02		2.5
$(CH_3)_3CCOOH$	11.1	10.67	7.43	4.2	12.4	11.24		10.0	7.00	7.42	5.01	5.6
$PhOCH_2COOH$	8.6	8.35		1.2	9.1	9.17		0.5	5.51	5.05		9.1
$Ph_2CHCOOH$	9.6	9.62		0.2	10.4	10.17		2.6	5.91	5.55		6.4
$PhCH_2COOH$	9.9	10.01	9.51	1.0	10.7	10.19	10.2	5.2	6.81	6.33		7.6
Amines												
$BuNH_2$	12.40	11.86	11.48	4.6	12.2	11.69		4.4	10.50	10.2		2.7
1-Naphthylamine									3.00	3.4		11.7
Et_2NH	11.90	11.92	11.20	0.2	10.7	10.97		8.2	11.50	10.4		10.4
Bu_3N	10.60	10.13		4.6	9.3	10.23		9.1	10.70	8.9		20.8
4-$MeC_6H_4NH_2$									4.30	4.8		9.6
3-$MeC_6H_4NH_2$									4.00	4.3		7.6
4-$BrC_6H_4NH_2$									2.70	3.5		22.4
3-$NO_2C_6H_4NH_2$	3.40	3.78	3.46	10.1	3.1	3.32		6.6	0.95	1.7		42.9
4-$NO_2C_6H_4NH_2$	2.30	2.71	1.55	15.1	2.2	2.84		22.5	1.20	2.0		39.1
PhNHMe									4.10	3.8		7.9
PhNHEt									4.60	4.2		10.3
$PhNMe_2$									4.90	3.8		27.8
$PhNEt_2$									6.50	5.0		30.1
2- Methylpyridine									6.10	5.2		17.9
Quinoline									5.10	4.0		26.9

Empty cells of experimental values in Table 10 indicate that the corresponding pK_a values were found in the literature and were included in the array of input data for approximation and generation of the empirical parameters for pK_a calculation by the equation (5). The error of the calculated values (ε, %) was calculated with respect to the experimental values obtained in this work. Empty cells of published values in Table 10 indicate that after approximation new data were not found in the literature.

As can be seen from Table 10, the most of calculated values obtained empirically (41 of 59) are within certain limits of the relative error - 10% and at the same time they also agree well with new literary experimental data added after approximation.

Significant mismatch between calculated pK_a value of propionic acid in ethanol and obtained experimental value (with the good compliance of the latter to the published value) is due to significant deviation of the values of empirical parameters in the equation (5) obtained for this acid, in comparison with other aliphatic acids.

High values of the relative error of calculation for 4-nitroaniline with small absolute pK_a values and in the absence of published data for this compound in ethanol lead to the following conclusion. To improve the accuracy of calculation prediction it is desirable to include the compounds covering the widest range of pK_a values in the original array of data for approximation and definition of empirical parameters. In the instant case (with low pK_a value) it was impossible in view of the absence of experimental data in the literature.

A marked trend of more significant errors of calculated pK_a values for amines in formamide, in contrast to carboxylic acids, is due to essentially lower number of available experimental values in the literature used for approximation and calculation of empirical parameters. 3-4 values for each amine were found for amines, as compared to carboxylic acids - 13-14 values for each acid.

To check the reliability of estimates of pK_a values in water using the EPS principle and the equation (6), dissociation constants of six compounds were determined spectrophotometrically. Four of these (aniline derivatives) were a part of twenty compounds with predicted values.

The obtained thermodynamic values of pK_a^{ex} of phenols and pK_b^{ex} of anilines in the water, their densities, as well as calculated values of ionization constants and the relative errors ε are given in Table 11.

Table 11. The experimental values pK_a^{ex} *of phenols and* pK_b^{ex} *of anilines in water in comparison with the calculated values, as well as the relative errors ε.*

Compounds	pK_a^{ex}	ρ, g/cm^3	Calculation of pK_a	ε, %
3- *tert*-butylphenol	10.06±0.03	0.974*	10.64	5.8
4- *tert*-butylphenol	10.05±0.02	0.978*	10.40	3.5
	pK_b^{ex}		Calculation of pK_b	
Ethyl-3-aminobenzoate	10.38±0.01	1.171	10.36	0.2
2-bromo-4-methylaniline	10.83±0.03	1.51	10.88	0.5
2-isopropylalanine	9.59±0.03	0.976	9.33	2.7
2-methyl-3-nitroaniline	11.51±0.05	1.378	13.97	21.4

* Determined experimentally in this work

The measured values of ionization constants in aqueous solutions are in good agreement with the results of calculation. The deviations of the calculated values from the experimental ones are within 5.8%. The single exception is pK_b of 2-methyl-3-nitroaniline, for which the error of calculation is 21.4%. It has been suggested that the cause of significant calculation error is an incorrect definition of the empirical atomic coefficient in the equation (6), related to nitrogen. Therefore for anilines with *ortho*-substituents the atomic coefficients were adjusted by the inclusion of three other compounds with pK_b values experimentally found during this work in the original group (Table 11).

As a result of this recalculation, the maximum error of ionization constants of aniline group with *ortho*-substituents decreased, moreover it corresponded to pK_b of 2-methyl-3-nitroaniline (15.8%) again. Thus, the alteration of empirical coefficients associated with an enlargement of processed database of experimental values, reduced the maximum and the average error of calculation in the group.

This fact confirmed the trend detected at the stage of approximation of experimental values of the physicochemical characteristics of compounds. With increasing of the number of experimental data related to compounds, combined into the single array by the principle of structure similarity, the errors of calculated values

obtained by the method equations using empirical coefficients found by approximation (see method algorithm), decrease.

III.4. *The relationship between empirical parameters in used ELEM equations and the type of the reaction center (protonation/deprotonation)*

Test of the predictive capacity of ELEM was conducted as an example of determination of the protonation center of molecules with two or more possible protonation centers. For this purpose for calculation of proton affinity values by molecular volume the equation (4) was used. Based on these data *the principle of the preferred reaction center* was proposed.

For the study we selected amino alcohols that could be protonated in gas phase or by hydroxy- or amino group. All liquid at normal conditions alcohols (19 compounds), amines (62 compounds) and amino alcohols (total 4) were combined in four ways: alcohols, amines, alcohols with amino alcohols and amines with amino alcohols. For each of four combinations the parameters with calculated values of PA were defined by the equation (4) (Table 12).

Table 12. The group of compounds, the number of compounds in the group N, the coefficients a, b, A_X , B_Y in the equation (4) and the average deviations of calculated PA values from the experimental ones.

The structure of the group	N	*a*	*b*	A_X	B_Y	ε_{av}, %
Alcohols	19	-0.0158	7.88	-2.37	0.0564	3.0
Amines	62	0.00707	8.96	19.0	0.0243	2.0
Alcohols + amino alcohols	23	-0.0254	8.05	-0.422	0.0768	4.9
Amines + amino alcohols	66	0.00374	9.12	40.7	0.0361	2.1

As expected, when combining amino alcohols and amines in one group, the accuracy of PA calculation (judging by the average relative deviation ε_{av}) was more than twice higher than when combining amino alcohols and alcohols. On this basis, we made the conclusion about the possibility of application of the method to detect

and predict the preferred center of protonation in those cases when the choice between alternative centers is not so obvious.

To estimate the applicability limits for ELEM in calculating of pK_a values in different solvents, joint pK_a calculation for three different types of organic compounds (carboxylic acids, amines and phenols) was attempted. As initial data we used the same array of experimental pK_a values for carboxylic acids and amines, just as in previous calculations (Table 7), and pK_a values of phenols selected by the same criteria. For calculations the equation (10), resulting from the formal transformation of the equation (5), was used:

$$\mathrm{p}K_a^{jk} = a_k X_j^{ex} - c_k \sum_i X_i + b_k + Y_j \,. \qquad (10)$$

Designations are similar to the equation (5). Approximation of experimental data by the equation (10) enabled us to calculate all empirical parameters.

Calculations were carried out for the arrays of 9-45 compounds in 3-9 solvents. Generally the maximum relative error of calculations did not exceed 14%. So, as a result of processing of pK_a values for 45 compounds (15 compounds of each type) in 5 solvents (water, acetonitrile, DMSO, DMF, methanol) using V_M as a parameter of the solvent, approximation by the equation (10) had the maximum relative error of calculation no more than 11.7% (Table 13, 14).

Table 13. Atomic contributions to molecular volume of the solvent (X_i) and the parameters of the solvents (Y_j) in the equation (10) *for joint* pK_a *calculation for the groups of compounds.*

Solvent	Y_j
Water	-10.83
MeCN	3.03
DMSO	19.67
DMF	26.19
MeOH	-3.51

Atom	X_i
C	-12.06
H	6.04
O	-6.32
N	10.37
S	3.38

Table 14. The parameters a_k, b_k, c_k of the compounds in the equation (10) *and their turndown for 3 groups of 15 compounds.*

The type of the compound	Parameter	a_k	b_k	c_k
Carboxylic acids	average	0.304	26.5	18.0
	turndown	0.186	14.3	10.0
Amines and nitrogen-containing heterocycles	average	0.814	-5.29	6.01
	turndown	0.175	27.4	3.47
Phenols	average	0.423	22.6	16.4
	turndown	0.673	48.1	51.0

As it seen from Table 14, the values of the parameters relating to the objects (a_k, b_k, c_k) are slightly different for the compounds within one group and significantly different for various groups. This means that the average values in the equation (10) are related to the type (chemical nature) of the center of protonation/deprotonation, i.e. to the generic accessory of compounds. The most noticeable relation of the type of reaction center is with the value of the parameter a_k. Although the turndown of a_k values for phenols is quite wide and almost overlaps with the range for carboxylic acids and amines, the differences in mean values mutually between 1 and 2 group and 2 and 3 group are quite obvious.

Similar relationships were observed for other combinations of compounds and solvents. Revealed sensitivity of empirical parameters to the type of the reaction center shows that ELEM identifies the characteristic for each feature type dependence of the investigated property of organic compounds upon the selected experimental parameter, like the parameter of the reaction series (ρ) in the Hammett equations.

III.5. *Calculation of* pK_a *values of nitrogen-containing organic compounds in aqueous-organic mixtures*

ELEM was used for calculation of dissociation constants (pK_a) of nitrogen-containing organic compounds in various aqueous-organic mixtures. It was supposed that the contributions of the solvents to the properties of the mixture were linearly proportional to their mole fractions (χ_{js}). Dimroth-Reichardt parameter (E_T^{30}) was used as a parameter of the mixtures of solvents. It was the most reliable determined

experimental value for the mixtures of solvents. Dissociation constants of *k*-th object in *j*-th mixture of solvents pK_a^{jk} were calculated by the formula:

$$pK_a^{jk} = a_k \left(X_j^{ex} - c_k \sum_s \sum_i \chi_{js} g_{ij} X_i \right) + b_k + \sum_{s=1}^{S} Y_s \cdot \chi_{js}, \qquad (11)$$

where S – the total number of solvents in the studied mixtures, Y_s - coefficient of *s*-th solvent, constant in all mixtures, χ_{js} - its mole fraction in *j*-th mixture. Other designations are similar to the equation (5). Finally for 8 amines in 4 mixtures of 4 solvents and for 7 tetrazoles in 36 mixtures of 5 solvents the maximum relative error of pK_a calculation did not exceed 3.6%.

IV. THEORETICAL POSITIONS AND MATHEMATICAL TOOL OF ELEM

The theory of physically observable quantities of objects belonging to the micro world is described by quantum mechanics. According to modern concepts, the object of the micro world (atom, molecule) is characterized by some vector of state. Physically observable quantity of the object, according to the measurement method, is associated with a linear operator in a linear space, where the vector of state is defined. Omitting the details, the average observable quantity is the trace of a matrix formed by the specified operator when multiplied by the transformed vector of state.

When applied to the practically measured molecular physicochemical characteristics it means the following. Any experimentally measured physicochemical quantity can be represented as a function of the finite number of independent variables. The explicit form of this function is determined by the molecular structure of the compound and, basically, by the measurement process. As a result, a stable and repeatable numerical value of the target value will be received.

According to this concept, there is a certain methodical similarity between experimental data obtained from a variety of measurements of physicochemical characteristics of the compound. Whether it be kinetic, spectrophotometric studies,

measurements of thermodynamic parameters or equilibrium constants, depending on external conditions - all of this can be covered by unified mathematical formalism.

If the function $Y = f(x_1,x_2,\ldots x_i,\ldots)$ is assigned to some measured physicochemical characteristics of organic compound, then atoms forming the molecule of this compound can be used as independent variables. This is the fundamental provision of the element linear empirical method. The choice of organic compounds as objects of the study was dictated by the relatively small number of the structure elements. Initially, during the design of the method CHNO-containing compounds are considered, and then - the compounds of more diverse composition.

Taking into consideration that the explicit form of this function is usually unknown and is entirely determined by the measurement process, for solution of applied chemical tasks and for its approximation the principle of polylinearity (PPL) of V.A. Palm was used. During mathematical investigation of PPL was found inaccuracy in the original conclusion, the removal of which led to the determination of two conditions under that PPL was accurately performed. Polylinear decomposition of the function Y of several independent variables x_i at a point with coordinates $(x_{10},\ldots x_{i0},\ldots, x_{n0})$ by V.A. Palm is the following:

$$Y(x_1,...,x_i,...,x_n) = Y_0 + \sum X_i + \alpha \sum_{i\neq j} \sum_{j} X_i X_j + \alpha^2 \sum_{i\neq j,k} \sum_{j\neq k} \sum_{k} X_i X_j X_k + ... + \alpha^{n-1} \prod_{i} X_i \,, \quad (12)$$

where Y_0 – value of the function Y at the point $(x_{10},\ldots x_{i0},\ldots, x_{n0})$, X_i – function of the corresponding to n variables x_i:

$$X_i = Y(x_{10},...,x_i,...,x_{n0}) - Y_0, \quad (13)$$

α – constant coefficient, which existence is defined at the studied point under two conditions described below

$$\alpha = \left. \frac{\partial^2 Y(x_1,...,x_i,...,x_j,...)}{2\partial X_i \partial X_j} \right|_{(x_{10},...,x_{i0},...,x_{j0},...)} \qquad i \neq j. \quad (14)$$

Basically the function X_i in an explicit form does not depend on "foreign" variable x_j. Formally, it means that

$$\frac{\partial X_i}{\partial x_j} = \delta_{ij}\phi(x_i), \qquad (15)$$

where $\phi(x_i)$ – an arbitrary function of "its" variable x_i, δ_{ij} – the Kronecker symbol.

Variables X_i of atoms x_i are discrete functions of the set of the quantum numbers characterizing the energy state of an atom in the molecule. For calculation of the physicochemical constants of organic compounds, taking into account the quantum numbers related to the valence electrons is sufficient in most cases. From the viewpoint of practical calculations by the formula (12) more attractive solution would be reduction of the number of quantum numbers that affect the atomic function X_i, to one; in this case it would be possible to assign one numerical value for atoms with the same number in the Periodic Table.

Then values of variables X_i in the decomposition (12) would only depend on one quantum number of the atom x_i. For example, the measured physicochemical quantity of formaldehyde molecule can be represented as a function of four variables: X_C, X_O and two X_H.

Studies of the molecular structure show that the charge of the nucleus has the greatest impact on the energy state of the atom in the molecule, if we are not talking about the photochemical and high-energy processes. The dependence of atomic variables only on the nuclear charge is the *first approximation of ELEM*. Conducted calculations of values of molecular volume and dipole moments have confirmed this approach and, moreover, have found the little value of the parameter α in the polylinear form (12).

With the first approximation of ELEM, calculations of the values of molecular parameters of compounds are more accurate than by the additive models. For this purpose it is enough to have corresponding atomic values (ionization potentials, enthalpies of protonation in gas phase, polarizability, etc.) and to calculate α value for compounds of similar structure and molecular composition.

When the atomic coefficients are unknown, the determination of these values is the following procedure. Suppose that for a number of compounds of similar structure there are J experimentally determined values, that are designated as Y_j^{ex}. The magnitude of the sum of the absolute or the relative errors of calculation of polylinear decomposition will have the following expression:

$$\Xi = \sum_{j=1}^{J} \left(Y_j^{ex} - Y_j \right)^2 \quad \text{или} \quad \Xi = \sum_{j=1}^{J} \left(\frac{Y_j^{ex} - Y_j}{Y_j^{ex}} \right)^2, \tag{16}$$

where Y_j – calculated magnitude of the corresponding experimental value, expressed by the equation (12). Suppose that in the structure of compounds there are I various chemical elements. Taking into account the above analysis, for compounds of similar structure one common value α is enough. Then the total number of empirical coefficients, necessary for the calculation of a number of experimental values by the first approximation, is equal to I+2.

At the presence of the global minimum of Ξ value, as a function of I+2 variables, with the need the conditions will be followed:

$$\frac{\partial \Xi}{\partial \alpha} = \frac{\partial \Xi}{\partial X_i} = \frac{\partial \Xi}{\partial Y_0} = 0, \tag{17}$$

Thus, there is a system of I+2 equations, after its solving α, Y_0 and X_i will be available in the case of the nondegenerated system. In the strict sense the obtained values for the empirical coefficients with the solution of system (17) may not fully reflect the position of global minimum the Ξ (if one exists), as this condition is necessary but not sufficient. Several sets of coefficients satisfying to the condition (17) correspond to local minima or saddle points (if there are any). This fact should not be forgotten when numerical solution of the system of the equations (17). Obviously, for calculations of polyatomic molecules the equations (17) represent a system of homogeneous nonlinear equations, and the success of its solutions does not always take place.

In the process of calculations by the first approximation it was found that the polylinear terms in the decomposition (12) for various physicochemical variables

were correlated in a number of certain classes of organic compounds. On this basis, we formulated the principle of linearity of nonlinear effects, reflected by the equations (1)-(3); its use in practice greatly simplified the search of empirical coefficients. The search of empirical coefficients was carried out similarly, but in this case it was necessary to solve the system of linear equations that was easy to implement by numerical methods. For example, when determination of the coefficients in the equation (6) in the system (17), a should be substituted instead of α, and b - instead of Y_0. These equations are linear relating to the target a, b, and X_i.

IV.1. *The conditions of PPL*

The conditions of PPL for exact decomposition of the function of many variables by (12) place the limitation on the scope of its definition in the neighborhood of the "zero point" and require the existence of mixed partial derivatives with respect to all variables at this point. The domain where the PPL is exactly implemented is limited by a curve in the space of dimension equal to the number of independent variables of the function passing through the "zero point".

The first condition in practice means that when fixing one of the independent parameters (variables), the function should not have two or more identical values. For example, if we are talking about the measured physicochemical value of the compounds, then the number of compounds where one of the independent parameters corresponds to the "zero point" should not have equal values of the investigated variable. In general, the coincidence of the values of measured physicochemical parameters of compounds is a rather rare phenomenon, if measured with sufficient accuracy. Furthermore, it should be noted that the first condition applies only to functions of fewer variables. Itself objective function can take several identical values in the studied neighborhood of the "zero point".

The second condition in practical terms means the physical reality of the existence of the point, which neighborhood should be reviewed. For example, a compound with the composition, structure or other parameters, corresponding to the

point with fixed variables, must really exist, the link of its parameters with the variables in the polylinear decomposition must be proved. In extreme case it may be the vacuum state, if the measurements of investigated physicochemical quantity are carried out in discharged gas phase, or the "zero point" should correspond to the composition of the medium where measurements are performed.

In view of the foregoing considerations, it can be concluded that the mathematical restrictions to the use of polylinear decomposition (12) in practice of the study of measured physicochemical quantities will rarely occur. The second condition is easily realizable in practice, in default of the first condition it is usual to introduce an additional independent variable or reduce the field of research.

Since the computation using the calculation equations of ELEM, expressing the basic provisions of the method produced the excellent results, and the equations themselves are a special case of the more fundamental principle (PPL), one can assume that the field of application of the method both in terms of coverage of various classes of organic compounds, and for calculated physicochemical characteristics can be greatly expanded in the future.

V. EXPERIMENTAL PART

pK_a values of compounds with quantities calculated by the ELEM equations and unavailable in the literature experimental data, were determined in water and three organic solvents. To obtain the correct results in the determination of pK_a values in non-aqueous media one should take into account the peculiarities of the protolytic equilibria in these media in order to avoid methodological errors committed in some works:

1. Frequently calibration of the measurement system for the scale pa_{H}^* on the ground of values of the EMF is carried out only by using one buffer system at different buffer ratios in the range of 3-4 log. units, resulting in significant errors of the determination of pa_{H}^* values and, hence, pK_a outside this area.

2. The scale pa_H^* is reliably determined only in a few solvents: methanol, ethanol, DMSO, formamide, and acetonitrile; this is an obstacle to obtain correct data in other media.

3. When determination of pK_a by the titration, frequently the distortion of the titration curve in aprotic solvents due to ion-molecule association and incomplete dissociation of buffer salts (except for the point of semi-neutralization) is not taken into account.

4. In aprotic protophobic solvents with very small water content, pK_a values are sharply reduced, resulting in a spread in pK_a values for one object up to 5 log. units in different authors who did not consider this factor.

In this work at experimental pK_a determination for the exception of these reasons of incorrect values some measures were taken:

1. For calibration we used several buffer systems with the buffer ratio of 1:1 to ensure the correct calibration of pa_H^*values throughout the investigated range.

2. For experimental studies we selected some solvents with reliably determined pa_H^* scale - methanol, ethanol, as representatives of the proton polar solvents; protophobic solvents - DMSO and acetonitrile – were excluded. As the aprotic solvent for the study protophilic formamide was chosen, because impurity of water in it slightly affects the pK_a value.

3. To measure pK_a in formamide we used the potentiometric method of filling, rather than potentiometric titration (with rare exception).

4. To increase the rate of establishment of electrochemical equilibria in organic solvents and to increase the electrical conductivity, tetrabutylammonium chloride (TBACl) as background electrolyte was added in test solutions.

For correct determination of pK_a values, in this work non-standard equipment was developed and used, special experimental techniques were applied, and the components of buffer systems were synthesized.

V.1. *Measurement of* pK_a *values in methanol and ethanol*

For calibration we used four literary pK_a values of acids of the buffer systems in ethanol and three - in methanol; the used buffer systems were picrate, salicylate, benzoate, and veronal ones. Calibration was performed by filling the cells with the relevant buffer solution with a buffer ratio of 1:1 and measurement of EMF of the system after the establishment of its permanent value (3 hours). Calibration in methanol is similar to ethanol. Due to instability of the measured EMF values in ethanol the point corresponding to the picrate buffer system was excluded from the gage graph.

Potentiometric titration using autotitrator "АТП-02" was performed in a specially designed glass measuring cell with a magnetic stirrer, protected from air moisture. Electrodes and hydraulic track of the titrator were connected to a cell by thin tight joint. As a reference electrode silver-chloride electrode "ЭСр-10103" was used; it was filled with 0.1 M solution of TBACl in ethanol for measurements in ethanol and the saturated solution of anhydrous lithium chloride in methanol – for measurements in methanol. It was found that in alcohols in one hour after the start of the experiment the potentials became somewhat permanent and differed from the equilibrium values by no more than 10 %, for this reason portions of titrant were added with an interval of 1 hour. The full curve of the potentiometric titration was not registered. For pK_a calculation the first 3-5 points of the titration curve corresponding to 15-40 % of the final point of titration were used. Processing of the titration curve was performed by the standardized technique, taking into consideration a correction for the activity of ions by the first Debye-Huckel approximation.

V.2. *Measurements of* pK_a *values in formamide*

For pK_a measurement in formamide we mostly used the method of filling. The cell for measurements by filling was two U-tubes with electrodes connected by electrolytic bridge filled with 0.1 M solution of TBACl in formamide. In the test

section with the glass electrode the test solutions with 0.01 M TBACl were placed, but in comparison section - 0.1 M solution of KCl in water.

For calibration and determination of pK_a in formamide we prepared solutions of mixtures of investigated carboxylic acids and their tetraethylammonium salts or amines with their picrates in the ratio 1:1 with the addition of 0.1 M TBACl, and measured EMF of the cell filled with these solutions in the test section. The equilibrium EMF value of the chain was recorded using the "*И*-500" ionomer in 2-3 hours after filling at the average.

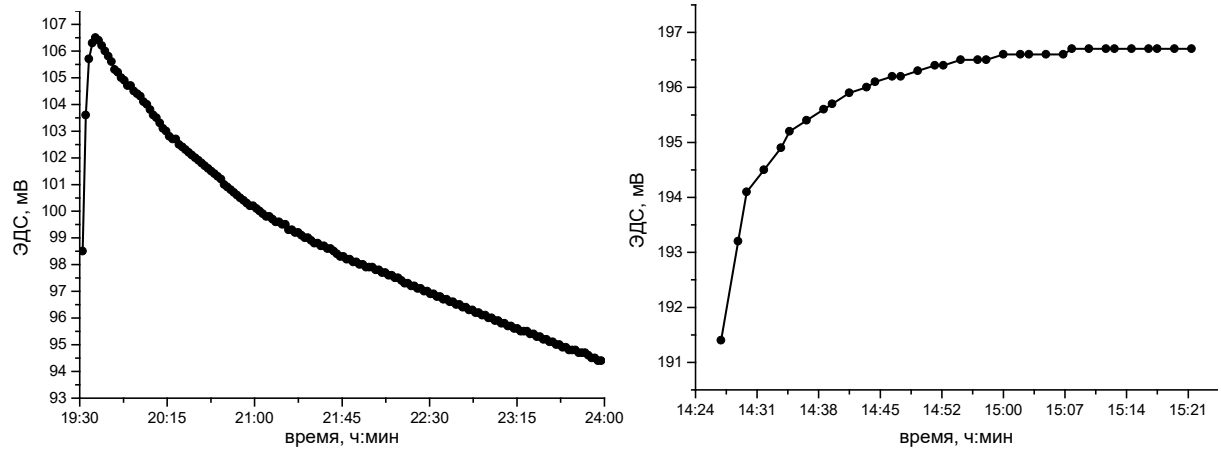
In this study pK_a values of 2,4-dichloroaniline, 3-chloroaniline, pyridine, benzoic acid, and triethylamine were initially selected for calibration.

However when the use of a buffer on the basis of 2,4- dichloroaniline the dependence of EMF on the time cell was expressed not by an exponential curve, but varied monotonically and did not reach the equilibrium value for more than 5 hours yet (Fig. 5a), assuming that the EMF values, measured at interval of a day, were not reproduced. In addition, the gage graph had the slope, far different from the theory (47.81 mV/log. units).

Therefore, we began using 3-chloroaniline instead of 2,4-dichloroaniline. The buffer solution on its basis was stable throughout the experiment series. The glass electrode "ЭС-10-60-3" was replaced by electrode "ЭСЛ-47-07СР", and the concentration of background electrolyte was increased. The new system drift of EMF in time dramatically decreased, the asymptote of the curve appeared, corresponding to the equilibrium EMF value (Fig. *5b*), and the slope of the gage graph was not different from the theoretical value. Thus, final calibration was carried out using four buffer systems: based on 3-chloroaniline, pyridine, benzoic acid and triethylamine.

Picrates of amines and nitrogenous bases were obtained by the standard procedure and qualified by melting point. Tetraethylammonium salts of acetic, *iso*-butyric and propionic acid could not be separated in the form of crystals, so their pK_a values were obtained by the results of the titration of these acids solutions by the solution of triethylamine in formamide with the addition of 0.1 M TBACl. For correlation of these data with the results of the measurements by filling method, test

pK_a measurements of benzoic acid and 2-nitrozobenzoic acid were carried out by different methods; they differed by less than 0.1 log. units.



“ЭДС, мВ“ – EMF, mV
“время, ч:мин” – time, h:min

а *б*

Figure 5. Dependence of EMF of the cell with buffer solutions such as amine / picrate of amine in formamide on time. *a* – 2,4-dichloroaniline, *b* – 3-chloroaniline after modification of the measuring system.

The titration of these acids was carried out on “АТП-02” titrator. Titrant (0.05 M solution of triethylamine in formamide) was added in a sealed glass cell of special design with a measuring electrode and a stirrer; the cell contained the solution of the studied acid with the addition of 0.1 M TBACl. The reference electrode was placed in the U-tube connected to the measuring cell by an electrolytic bridge filled with 0.1 M solution TBACl in formamide. When calibration of the cell for potentiometric titration in formamide we used the same buffer systems as for the filling method. To reduce the total time of titration we previously added 2-4 ml of titrant in the solution (which was equivalent of 15-30% of the final point of titration) and kept the system for 2 hours prior to the establishment of constant EMF.

Using pa_{H}^{*} value at the point of semi-neutralization the pKa value, with allowance for the activity of ions by the second Debye-Hückel approximation, was received.

Tetraethylammonium salts of carboxylic acids were obtained by titration of water-alcohol acids solutions by 20% aqueous solution of tetraethylammonium

hydroxide to the point of equivalence. These solutions were evaporated to dryness in a vacuum. Gel-like substances were usually obtained; they were kept in a desiccator under vacuum over P_2O_5 for several days. The result was very hygroscopic crystals, that deliquesced in the air for several minutes; it made impossible to take a precise test portion. Therefore, they were transferred to a mini-desiccator over P_2O_5 in a pre-dried air-tight box. In the same box the analytic balance and the volumetric flask with the prepared solution of the corresponding acid were placed. A test portion of salt was taken from the box and placed into the flask with the acid solution, covered with lid and took out from the box for mixing and for the further work.

V.3. *Determination of dissociation constants of organic compounds in water by the spectrophotometric method*

Calibration of the pH-metric system was carried out by standard aqueous buffer solutions. Determination of thermodynamic constants of ionization (pK_a^T) was performed in accordance with standard technique.

Two techniques were used for determination of pK_a for 2-methyl-3-nitroaniline. In the first case we used the primary buffer solutions, and in the second - specially designed buffer compositions with constant ionic strength (0.01 M). After comparative evaluation of the results (I technique $pK_a^T = 2.50 \pm 0.06$; II technique $pK_a^T = 2.49 \pm 0.05$), it was concluded that for pK_a determination in aqueous solutions of all test compounds, with not very complicated structure, the application of primary buffer solutions was possible.

V.4. *Experimental determination of molecular volume of compounds in diluted aqueous solutions*

Values of density of highly diluted aqueous solutions with mass fraction of substance w less than 2 % were measured with ultrasonic density detector (Kyoto Electronics DA-500).

For performed calculation of molecular volume of the studied compound in dilute aqueous solutions by the equation (18) the coefficient *d* was determined:

$$\rho(w) = \rho_S + d \cdot w, \qquad (18)$$

where ρ_S – the density of pure solvent. For this purpose on the basis of the experimental data the dependence of the density of the highly diluted aqueous solution ρ on mass fraction of dissolved substance *w* was plotted, and by the angle of slope the coefficient *d* was determined.

Abundant experimental data showed that the dependence of the highly diluted solution on mass fraction of dissolved substance (*w* less than 2 %) is indeed linear with high correlation coefficient.

V.5. *Photoemission method of spectral registration*

Until the present time spectrophotometric method is the most important classical method of experimental pK_a determination. It is extensively used both in the original version, and in various variants.

In addition to the traditional schemes of spectrophotometers with photomultiplier tubes and photodiode matrices with monochromators, for measurement of optical spectra photoemission method can be used. It is based on the dependence of the energy distribution of photoelectrons on incident frequency, i.e. the spectrometry problem is solved purely by means of electronic system. A few works were devoted to the solution of spectroscopy problem by means of electronic system, these works show the possibility of spectral radiation measurements by photoemission technique in terms of continuous and line spectrum [*Kasparov K.N. at al.*, Journal of Applied Spectroscopy. 1975. Vol. 22. N. 3. P. 491-498.; *Hinterregger H.E at al.*, J. Geophys. Res. 1959. Vol. 64. N. 8. P. 961-969.]. The problem of spectrum obtaining requires the use of the serious mathematical tool.

In this investigation the construction of vacuum phototube with a point anode and a hemispherical cathode was first used; it could reduce the contribution of the

photoinduced current from the anode to the total photocurrent. This phototube was used with a metal cathode. In this case, energy spread of photoelectrons was much smaller, and the distortion introduced in the current-voltage characteristic was much less, as compared to semiconductor cathodes.

The total photocurrent contains a function of the integral spectrum of radiation incident on the photocathode from the short waves to the "long-wavelength cutoff", defined by the photoelectric work function and cutoff voltage. In cases where the spectra of protonated and non-protonated forms are well resolved, it would appear reasonable that the integrals of these spectra in the same ranges have different values as well. Thus, the value of the photocurrent depending on pH value of test solution in a cell in optical path incident on the photocathode will be different. The total photocurrent, originated under the influence of light passing through the test solution i_{Σ}, made up the current, resulting in the absorption of the protonated form (BH) – i_{BH}, and non-protonated one (B) – i_{B}, thus:

$$i_{\Sigma} = x i_{\mathrm{BH}} + (1-x) i_{\mathrm{B}} \tag{19}$$

$$x = \frac{K_a}{[\mathrm{H}^+] + K_a}, \tag{20}$$

where x – the degree of ionization. Whence it follows that:

$$i_{\Sigma} = i_B + \frac{K_a}{[\mathrm{H}^+] + K_a}\left(i_{\mathrm{BH}} - i_{\mathrm{B}}\right) = a + \frac{10^{-\mathrm{p}K_a}}{10^{-\mathrm{pH}} + 10^{-\mathrm{p}K_a}} b, \tag{21}$$

where a and b – some constants.

Shown dependence has a jump (the maximum of first-order derivative by pH value) at pH = pK_a.

4-nitrophenol, 2-nitrophenol, and 2,6-dinitrophenol were selected as objects of study. These compounds had pK_a values in water, suitable for measurements (7.15, 7.22, and 3.71, respectively), and well-resolved spectra of the prototropic forms with the absorption maxima position, limited by "long-wavelength cutoff" of photoelectric effect for potassium (500 nm) and transmission cutoff of the used cell (300 nm).

Determination of ionization constants by the photoemission method

Apparatus for current-voltage characteristic (CVC) registration comprised:

- The source of emitting of continuous spectrum in the visible region with a collimator.
- The described above phototube of the unique design with potassium photocathode.
- DC calibrator - voltage calibrator “B1-18”.
- Digital Keithley System Electrometer, 6514 model (measurement limit - 1 fA).
- Electrometer amplifier (EMA) of the unique design (JSC "Spectron-Analyt") for pre-amplification.
- The special construction as protective cover, with the phototube and the electrometer amplifier.
- Personal computer for processing of the measurement results and for control of registration process.

The test solution in the glass cell was placed in the path of the light flux incident on the phototube, located in the protective cover. The values of current at a given voltage were calculated as the average value of 5-20 measurements depending on the strength of photocurrent. Measurement pitch was 0.02 V from 0 to 2 V and 0.5 V – from 2 to 10 V. The absorption spectra of the compounds solutions (indicators) were recorded in the same glass cell on spectrophotometer “СФ-56”.

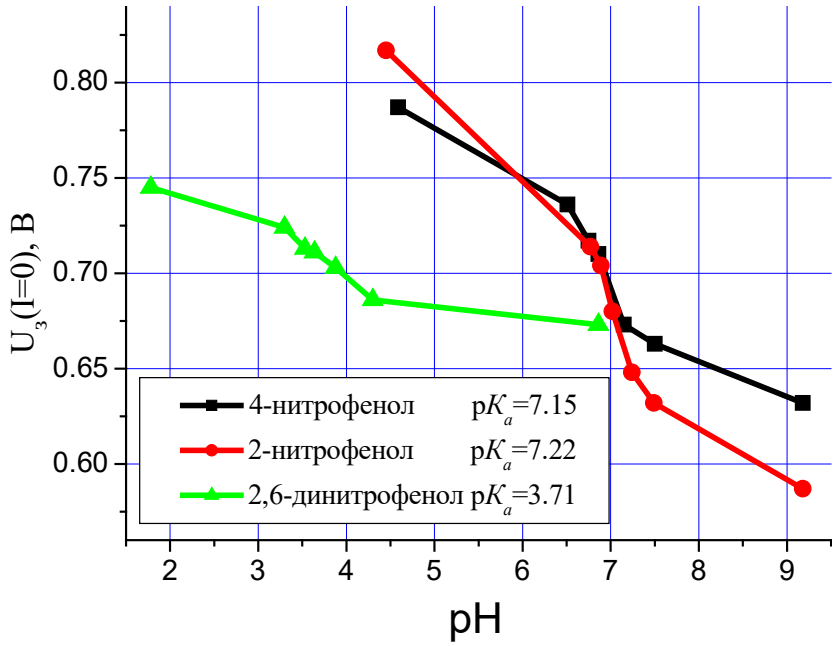
For determination of pK_a values on the basis of CVC of the phototube, the search of features on the original integral CVC, that correlated with pKa values, was taken. As a result, it was found that for all three studied compounds the pH curve of cutoff voltage ($U_3(I=0)$), when total photocurrent I (cathode with anode) reduced to zero, has an indicative jump in the domain corresponding to pK_a value (Fig. 6).

The position of the jump (maximum of first derivative with respect to pH, obtained by interpolation of experimental values with a subsequent numerical

differentiation) on the above curves differs from the literature experimental pK_a values by 0.1-0.2 units of pH, the relative difference is less than 5% (Table 15).

Table 15. pKa values, obtained by photoemission method from the dependence $U_3(I=0)$ on pH (pK_a^{ex}) in comparison with literature values (pK_a), as well as the relative discrepancies (εpK_a).

Compound	pK_a	pK_a^{ex}	εpK_a, %
2- nitrophenol	7.22	7.1	2%
4- nitrophenol	7.15	6.9	4%
2,6-dinitrophenol	3.71	3.8	4%



“4-нитрофенол” – 4-nitrophenol,
“2-нитрофенол” – 2-nitrophenol,
“2,6-динитрофенол” – 2,6-dinitrophenol

Figure 6. Dependencies $U_3(I=0)$ on pH for 4-nitrophenol, 2-nitrophenol, and 2,6-dinitrophenol.

These data give ground to conclude that the proposed method can be used to determine pK_a of organic compounds in the studied region of the spectrum (300 - 500 nm).

CONCLUSIONS

1. A new empirical method for calculation of the physicochemical characteristics of organic compounds was developed. The method took into consideration the composition and the structure of molecules by presentation of physicochemical characteristics as a function of the independent variables; atoms forming the molecule of the compound were selected as these variables.

2. The principles of linearity of nonlinear effects, "the element of permanent structure" and "the preferred reaction center", making possible to apply the proposed method in practice, were proved and confirmed by calculations and experiments.

3. It was shown that the calculations of acidity constants (pK_a) of organic compounds in various media by means of the element linear empirical method (ELEM) had comparable or even better predictive power in comparison with known calculation methods.

4. By means of ELEM values of a series of thermodynamic parameters (enthalpy, entropy, Gibbs energy, the isobaric heat capacity) for 1055 organic compounds, the proton affinity values for 395 compounds, 363 values of dissociation constants for 33 benzoic acid and acetic acid derivatives in 11 solvents, 192 values of dissociation constants for 24 amines in 8 solvents were calculated. The maximum calculation errors did not exceed 10 % in all instances, and the average errors of calculation in the investigated groups of compounds were less than 4%.

5. It was found that the calculations of values of physicochemical variables by ELEM had the convergence property. This means that with increasing in the number of approximated experimental values, obtainable values of empirical coefficients converge to fixed values, moreover the average errors of calculation in the test sampling fall.

6. It is proved that the principle of polylinearity makes possible not approximate, but mathematically accurate calculation of the physicochemical values as a function of a set of independent parameters under two conditions (lack of concurrence of

values of physicochemical variables in a selected series of compounds and the existence of the compound with the parameters, that form a set of values, in that neighborhood the decomposition of the function is performed); the fulfillment of the condition should be easily accomplished in the practice of chemical research.

7. A new photoemission method of spectral registration, protected by two patents of the Russian Federation. A unique procedure of processing the current-voltage characteristics of the phototube, which enables to determine pK_a of organic compounds by the absorption spectra.

8. A new parameter of the solvent was obtained with ELEM. It makes possible to consider quantitatively the influence of medium in pK_a determination of organic compounds. Applicability of this parameter is experimentally confirmed by the example of pK_a calculations in 11 solvents.

9. It is theoretically shown and experimentally confirmed by ELEM that molecular volume of organic compounds in the limit of infinite dilution in some solvent is connected to acid-base properties of this compound.

10. In comparison with existing methods of calculation of the physicochemical characteristics of organic compounds, in most cases the results of ELEM are superior to other methods; it is shown by the example of calculations of acidity constants for the same compounds in different media.

11. The results of the work offer ample opportunities for determination and prediction of various physicochemical properties of organic compounds.

The main content of the work is described in following publications

1. Patent 25598 of the Russian Federation, International Patent Classification U1 7G 01 J 3/00, 3/02. Apparatus for determination of spectrum of electromagnetic radiation, *Yu. E. Zevatskiy, D.V. Samoilov.* - № 2002107548; applied 21.03.2002; published 10.10.2002. Bulletin № 28.
2. Patent 2204811 of the Russian Federation, International Patent Classification C1 7G 01 J 3/12. The method of determination of the electromagnetic radiation spectrum, *Yu. E. Zevatskiy.* - № 2002107932; applied 21.03.2002; published 20.05.2003. Bulletin № 14.
3. *Зевацкий Ю.Э.* Связь диэлектрической проницаемости неполярных газов с потенциалом ионизации молекул // ЖПХ. – 2005. – Т. 78. – № 3. – С. 466-470 (*Yu. E. Zevatskiy,* Relation between dielectric capacity of non-polar gases and ionization potential of molecules, Russian Journal of Applied Chemistry, Vol. 78, No. 3,. 2005, pp. 466-470).
4. *Зевацкий Ю.Э.* Зависимость потенциалов ионизации атомов и многозарядных ионов от заряда ядра // ЖТФ. – 2006. – Т. 76. – № 3. – С. 86-90 (*Yu. E. Zevatskiy,* Dependence of the ionization potential of atoms and multicharged ions on the nuclear charge, Technical Physics Journal, Vol. 76, No. 3,. 2006, pp. 86-90).
5. *Зевацкий Ю.Э., Лысова С.С.* Взаимосвязь молярного объема с энтальпией образования, поляризуемостью и потенциалом ионизации CHNO-содержащих органических веществ // ЖПХ. – 2006. – Т. 79. – № 6. – С. 978-985 (*Yu. E. Zevatskiy, S.S. Lysova,* Relationship between the Molar Volume and Enthalpy of Formation, Polarizability, and Ionization Potential of CHNO-containing Organic compounds, Russian Journal of Applied Chemistry, Vol. 79, No. 6,. 2006, pp. 978-985).
6. *Зевацкий Ю.Э., Самойлов Д.В.* Эмпирический метод расчета термодинамических потенциалов органических веществ // ЖПХ. – 2007. – Т. 80. – № 2. – С. 230-235 (*Yu. E. Zevatskiy, D. V. Samoilov,* Empirical Method for

Calculation of Thermodynamic Potentials of Organic compounds, Russian Journal of Applied Chemistry, Vol. 80., N.2, 2007, pp. 230-235).

7. *Зевацкий Ю.Э., Власов Е.А.* Зависимость констант протонирования органических веществ от молекулярного объёма // ЖОХ. – 2007. – Т. 77. – № 2. – С. 260-267 (*Yu. E. Zevatskiy, E. A. Vlasov,* Dependence of the Protonation Constants of Organic Compounds on Their Molecular Volume, Russian Journal of General Chemistry, Vol. 77, No.2, 2007, pp. 260-267).
8. *Зевацкий Ю.Э., Самойлов Д.В.* Некоторые современные методы оценки реакционной способности органических соединений // ЖОрХ. – 2007. – Т. 43. – № 4. – С. 487-504 (*Yu.E. Zevatskiy, D. V. Samoilov,* Some modern methods for evaluation of the reactivity of organic compounds, Russian Journal of Organic Chemistry, Vol. 43, No. 4, 2007, pp. 487–504).
9. *Yu.E. Zevatskiy, D.V. Samoylov* Empirical method of the account of influence of solvent dissociation constants of some carboxylic acids, International conference MODERN PHYSICAL CHEMISTRY FOR ADVANCED MATERIALS (MPC'07) Kharkiv, 26-30 June, 2007. Book of abstracts: Kharkiv National University, 2007, pp. 332-335.
10. *Зевацкий Ю.Э., Самойлов Д.В.* Эмпирический метод расчета констант диссоциации NH-кислот в смесевых растворителях // XVIII Менделеевский съезд по общей и прикладной химии (IUPAC). Москва, 23-28.09.2007. Тезисы докладов в 5 т., т. 1. М.: Граница, 2007. – С. 225,228 (*Yu. E. Zevatskiy, D.V. Samoilov,* Empirical method for calculation of dissociation constants of NH-acids in mixed solvents, XVIII Mendeleev Congress on General and Applied Chemistry (IUPAC), Moscow, 23-28 September 2007. Book of abstracts in Vol. 5, Vol. 1, Moscow, Granitsa, 2007, pp. 225,228).
11. *Зевацкий Ю.Э., Самойлов Д.В.* Эмпирический метод учета влияния растворителя на константы диссоциации карбоновых кислот // ЖОрХ. – 2008. – Т. 44. – № 1. – С. 59-68 (*Yu. E. Zevatskiy, D. V. Samoilov,* Empirical Method for Consideration of Solvent Effect on the Dissociation Constants of

Carboxylic Acids, Russian Journal of Organic Chemistry, Vol. 44, No. 1, 2008, pp. 59–68).

12. *C.A. Deakyne, J.F. Liebman, E.A. Vlasov, Yu.E. Zevatsky,* Paradigms and paradoxes: analysis of the site of protonation of bifunctional organic compounds with the protonation energy/volume computation method, Structural Chemistry, Vol. 19, № 4, 2008, pp. 609-611.
13. *Yu. E. Zevatskii, D.V. Samoilov,* Empirical method for calculation properties of organic compounds and chemical similarity recognition, Proceedings of the 2nd WSEAS International Conference on computational chemistry (COMPUCHEM'08). Puerto De La Cruz, Tenerife, Spain, December 15-17, 2008, pp. 30-32.
14. *Зевацкий Ю.Э., Самойлов Д.В.* Применение эмпирического метода для учета влияния растворителя на константы ионизации NH-кислот // ЖОрХ. – 2008. – Т. 44. – №12. – С. 1764-1771 (*Yu. E. Zevatskiy, D. V. Samoilov,* Application of Empirical Method for Evaluation of Solvent Effect on the Dissociation Constants of NH-Acids, Russian Journal of Organic Chemistry, Vol. 44, № 12, 2008, pp. 1764-1771).
15. *Зевацкий Ю.Э., Самойлов Д.В., Панина Н.С.* Расчеты констант диссоциации карбоновых кислот эмпирическим и квантовохимическим DFT методами // ЖОХ. – 2009. – Т. 79. – № 5. – С. 772-780 (*Yu. E. Zevatskiy, D.V. Samoilov, N.C. Panina,* Calculations of dissociation constants of carboxylic acids by empirical and quantum chemical DFT methods, Russian Journal of General Chemistry, Vol. 79, No.5, 2009, pp. 772-780).
16. *Лысова С.С., Зевацкий Ю.Э.* Современные проблемы определения значений констант ионизации органических соединений в водных растворах // Известия Санкт-Петербургского государственного технологического института (Технического университета). – 2009. – № 5. – С. 20-25 (*S.S. Lysova, Yu. E. Zevatskiy,* Modern problems of the determination of the values of ionization constants of organic compounds in aqueous solutions, News of Saint-

Petersburg State Institute of Technology (Technical University), No.5, 2009, pp. 20-25).

17. *Колдобский Г.И., Зевацкий Ю.Э.* Микроволновая активация в химии тетразолов // Материалы 1-ой Международной конференции «Новые направления в химии гетероциклических соединений» (Кисловодск, 3-8 мая 2009). – С. 127 (*G.I. Koldobskii, Yu. E. Zevatskiy,* Microwave activation in the chemistry of tetrazoles, Materials of the I International conference "New Trends in Chemistry of Heterocyclic Compounds" (Kislovodsk, 3-8 May 2009), p. 127).
18. *Зевацкий Ю.Э., Лысова С.С.* Эмпирический метод расчета значений констант ионизации в воде по молекулярному объему органических соединений // ЖОрХ. – 2009. – Т. 45. – № 6. – С. 842-850 (*Yu. E. Zevatskiy, S.S. Lysova,* Empirical method of calculation of ionization constants of organic compounds in water solutions on molecular volumes, Russian Journal of Organic Chemistry, Vol. 45, № 6, 2009, pp. 842-850).
19. *S.S. Lysova, Yu.E. Zevatskii,* Element linear empirical method of calculation of ionization constants of organic compounds in water solutions on molecular volumes, Abstr. Fifth International Conference on Organic Chemistry for Young Scientists (InterYCOS-2009) "Universities Contribution in the Organic Chemistry Progress", Saint-Petersburg, June 22-25, 2009, pp. 85-86.
20. *S.S. Lysova, Yu.E. Zevatskii,* Application of the element linear empirical method for calculation of ionization constants organic compounds, Abstracts of the XVII International Conference on Chemical Thermodynamics in Russia (RCCT Kazan, June 29 - July 3, 2009), Vol. 1, p. 53.
21. *Yu.E. Zevatskii,* Development of polylinearity principle in modern physical organic chemistry, Abstracts of the XVII International Conference on Chemical Thermodynamics in Russia (RCCT Kazan, June 29 - July 3, 2009), Vol. 1, p. 318.
22. *Yu.E. Zevatskii, D.V. Samoylov,* Application of metallic photocells for express spectrophotometric determination of equilibrium constants, Abstracts of the XVII

International Conference on Chemical Thermodynamics in Russia (RCCT Kazan, June 29 - July 3, 2009), Vol. 1, p. 319.

23. *Лысова С.С., Зевацкий Ю.Э.* Спектрофотометрическое определение значений констант ионизации некоторых органических соединений в водных растворах // ЖОХ. – 2009. – Т. 79, – № 8. – С. 1311-1313 (*S.S. Lysova, Yu.E. Zevatskiy,* Spectrophotometric determination of values of ionization constants of some organic compounds in aqueous solutions, Russian Journal of General Chemistry, Vol. 79, No.8, 2009, pp. 1311-1313).
24. *Зевацкий Ю.Э., Самойлов Д.В., Мчедлов-Петросян Н.О.* Современные методы экспериментального определения констант диссоциации органических кислот в растворах // ЖОХ. – 2009. – Т. 79. – № 9. – С. 1504-1532 (*Yu.E. Zevatskiy, D.V. Samoylov, N.O. Mchedlov-Petrossyan,* Modern methods of experimental determination of dissociation constants of organic acids in solutions, Russian Journal of General Chemistry, Vol. 79, No.9, 2009, pp. 1504-1532).
25. *Зевацкий Ю.Э., Рузанов Д.О., Самойлов Д.В.* Экспериментальная проверка результатов расчетов констант диссоциации органических соединений в неводных средах // ЖОХ. – 2009. – Т. 79. – № 9. – С. 1533-1537 (*Yu.E. Zevatskiy, D.O. Ruzanov, D.V. Samoylov,* Experimental validation of the calculation results of the dissociation constants of organic compounds in non-aqueous media, Russian Journal of General Chemistry, Vol. 79, No.9, 2009, pp. 1533-1537).
26. *Зевацкий Ю.Э., Самойлов Д.В.* Фотоэмиссионный метод измерения спектров для спектрофотометрического определения констант ионизации // Известия Санкт-Петербургского государственного технологического института (Технического университета). – 2009. – № 6. – С. 44-49 (*Yu.E. Zevatskiy, D.V. Samoylov,* Photoemission method of spectra measurement for the spectrophotometric determination of ionization constants, News of Saint-Petersburg State Institute of Technology (Technical University), No.6, 2009, pp. 44-49).

27. *Зевацкий Ю.Э.* Условия строгого выполнения принципа полилинейности // Известия Санкт-Петербургского государственного технологического института (Технического университета). – 2009. – № 6. – С. 85-87 (*Yu.E. Zevatskiy,* Terms of strict implementation of the principle of polylinearity, News of Saint-Petersburg State Institute of Technology (Technical University), No.6, 2009, pp. 85-87).