LITERATURE CITED

1. M. J. Rice, S. R. Phillpot, A. R. Bishop, and D. K. Campbell, Phys. Rev. B, 34, No. 6, 4139-4148 (1986).
2. S. Suhai, Phys. Rev. B, 29, No. 8, 4570-4581 (1984).
3. E. M. Shustorovich, Zh. Strukt. Khim., 5, No. 2, 325-329 (1964).
4. L. Salem and H. Longuet-Higgins, Proc. Roy. Soc. A, 255, No. 1283, 434-443 (1960).
5. I. A. Misurkin and A. A. Ovchinnikov, Zh. Strukt. Khim., 5, No. 6, 888-891 (1964).
6. J. H. Callomon and B. P. Stoicheff, Canad. J. Chem., 35, No. 4, 372-382 (1957).
7. C. Coulson and A. Golebiewski, Proc. Phys. Soc., 78, No. 506, 1310-1320 (1961).
8. A. A. Ovchinnikov, A. E. Belinskii, I. A. Misurkin, and I. I. Ukrainskii, Int. J. Quant. Chem., 22, No. 5, 761-774 (1982).
9. T. S. Zhuravleva and I. A. Misurkin, Zh. Strukt. Khim., 5, No. 4, 656-658 (1964).

A PROGRAM SET FOR FINDING STRUCTURE-PROPERTY
CORRELATIONS USING TOPOLOGICAL INDICES

I. I. Baskin, M. I. Stankevich,
R. O. Devdariani, and N. S. Zefirov

UDC 547.8:513.83:519.55

The relationship of the structure and properties of chemical compounds has long attracted the attention of chemists. However, major advances in this area began to be achieved only about 20 years ago with the introduction of computers, permitting the treatment of extensive data. In this regard, special importance is found for the development of various mathematical models relating structure and property and their programming for various computer classes, including, in particular, personal computers.

In this work, we describe a program set designed for the analysis of the relationship of structure and property in a class of organic compounds using topological indices. Ruvra [1, 2] has reviewed many examples of topological indices and discussed their various applications for the establishing structure-property correlations. An analysis was also given for the abovementioned method. In this approach, the molecular structure is described using a finite set of topological indices $x_1,...., x_k$ (invariants of the "weighted" molecular graphs). Then, a function $f(x_1,...., x_k)$ is selected, such that $y \approx f(x_1,..., x_k)$ for all the compounds of the set studied, where y is the numerical value of the property. The equation thereby obtained, which relates the structural parameters $x_1,..., x_k$ and a property, is then extrapolated to other compounds of this class.

The following operations are provided in our program set:

1) the creation of files for chemical structure and property data;

2) the calculation of various topological indices, including indices determined for "weighted" graphs;

3) the classification of atoms of various chemical types in the structures of the data file relative to a given classification scheme;

4) selection of "weights" for the atoms of different classes using optimization methods with the subsequent distribution of these "weights" over the data file;

and 5) derivation of correlation equations.

The set has a modular structure, consisting of a lead program and set of slave programs. The lead program permits the selection of the direction of operations 1-5, carries out the loading and filing of other programs, operating in a concrete direction, enters information into the data file or computes it, and organizes the derivation of correlations and the printout of results onto a screen or on a printer. The slave programs include those for computer graphics, calculation of indices, classification, optimization, and distribution of atomic "weights." These programs will be examined in greater detail below.

A special program for graphic input and output of chemical structural information is included for the establishment of a data file in the program set [3]. This program is used for the introduction of structural formulas in the conventional chemical form into the computer by simply drawing them on the display screen. A table of bonds of the corresponding structural formula is automatically constructed using this figure. All the structural and quantitative

property data are transferred to a disk, on which the file data is thereby created. When necessary, the data may be transferred to the screen and edited. For convenience, structural formulas may be introduced into the file without hydrogen atoms and their corresponding chemical bonds.

The program set provides for the calculation of the topological indices of several categories. The first category of indices includes those, whose designation does not require the assignment of specially selected "weights" to the apices of the molecular graphs. For example, these include bondedness indices of the order $h \geq 1 \chi^h$, Hosoya index Z, and indices, which are functions of the distance matrix D such as the Wiener index W, graph diameter, and information indices $I_D{}^E$, $I_D{}^W$, and $I_D{}^M$ [2]. The information indices based on division of the graph apices into equivalence classes according to first-proximity patterns also belong to this category of indices [2].

Another class of indices includes those, which permit us to differentiate atoms of different chemical nature. An example of such indices is the weighted bondedness indices, h-th-order $^v\chi^h$, which are analogs of simple bondedness indices. Thus, the establishment of $^v\chi^1$ for each edge (i,j) of the graph is calculated by multiplication of the "weights" of the corresponding apices $\delta_i \cdot \delta_j$, which are then summed over all the edges.

The following algorithm was used for construction of the distance matrix $D = (d_{ij})$. The i-th apex of the graph is fixed and labelled "0" and their neighbors are labelled "1." The neighbors of the apices labelled "1" are labelled "2" if they are not already labelled and so on until all the apices are labelled. The value of the label of the j-th apex is equal to the d_{ij} element of matrix D. An algorithm to search the isomorphous subgraphs in the graph was used to calculate the χ^h indices ($h > 1$) and Z. Another method is also provided for calculation of the Z index based on the search-and-return procedure (the algorithm and program were developed by S. S. Trach).

The classification of atoms of different chemical types is carried out to establish "weighted" indices. Using suitable programs, atoms may be classified according to their chemical symbols. Furthermore, the additional subdivision of atoms with the same symbol may be carried out according to the degree of the corresponding graph indices. In a more detailed classification, carbon atoms included in five- or six-membered aromatic rings with and without substituents are assigned to different classes. Finally, atoms may be classified according to first-proximity patterns encountered in the compound set studied. All the classes are numbered and the number of the corresponding class is assigned to each atom in all the structures of the data file.

Two variants are provided for the distribution of the atomic "weights." In one variant, the "weights" indicated in the program text are used such as $\delta_i = (Z_i - h_i)^{-\frac{1}{2}}$, where Z_i and h_i are the nuclear charge and number of adjacent hydrogen atoms of the i-th atom [2]. In the other variant, the class number and "weights" selected by some procedure is flashed onto the display screen, introduced from the keyboard, and automatically distributed over the entire data file.

The selection of the "weights" using optimization methods appears most practical. For this purpose, the corresponding program is used, realizing the Hooke—Jeeves method [5]. This program is used for the selections of "weights" δ_i assuming that $y = \Sigma\delta_i \cdot \delta_j + \text{const}$, where y is the numerical value of the property and summation proceeds over all the bonds (i,j) (modifications of this equation are possible).

We note that there are infinitely many topological indices. All these indices cannot be programmed. Thus, we have developed clear instructions, which permit the user to expand the program set to new topological indices, including "weighted" indices without modification of the other programs. Analogously, other classification and optimization programs may be included into the set.

Several variants are provided for the establishment of index-property correlation equations. In the first variant, numerical values of the property and index are transferred to the display screen during the calculation of the topological indices for data file structures. The coordinate system, in which the points with corresponding coordinates are flashed, is also indicated. Then, linear regression analysis is also automatically carried out and the parameters of the approximation line and correlation coefficient are transferred to the screen. In the other variant, the values of the index and properties are introduced from the keyboard. In the case of a one-variable function, graphs are constructed, which may be printed. The type of the approximating function, which may be polynomial, exponential, etc., is given by the user himself. Standard library programs are used for multidimensional regression analysis.

The possibility exists for analysis of the structure-property relationship using a part of the data file. For this purpose, the user should first indicate the number of the necessary structures. The programs were compiled in BASIC for the Iskra-226 computer. The computation time for the W, Z, χ^1, and χ^3 indices (for a subgraph chain) with 15 apices on the Iskra-226 computer is 15, 4, Z, and 73 sec, respectively. The program set described above was used to search for quantitative relationships between the structure and melting point of selenophene and indanedione compounds, the entropy and viscosity of hydrocarbons, and hydrophobicity and solubility of alcohols. Information on the programs for calculation of topological indices for various types of computers is given by Peredunova [6, 8] and Shtain [7].

LITERATURE CITED

1. D. Ruvra, Chemical Applications of Topology and Graph Theory [Russian translation], Mir, Moscow (1987).

2. M. I. Stankevich, I. V. Stankevich, and N. S. Zefirov, Usp. Khim., **57**, No. 3, 337-365 (1988).
3. N. S. Zefirov, I. I. Baskin, and S. S. Trach, Zh. Vses. Khim. Obshch. im. D. I. Mendeleeva, **32**, No. 1, 112-113 (1987).
4. M. I. Stankevich, I. I. Baskin, and N. S. Zefirov, Zh. Strukt. Khim., **28**, No. 6, 136-137 (1987).
5. B. Bundy, Optimization Methods [Russian translation], Radio i Svyaz', Moscow (1988).
6. I. V. Peredunova, V. E. Kuz'min, and Yu. P. Konovorotskii, Zh. Strukt. Khim., **24**, No. 4, 145-147 (1983).
7. L. M. Shtain, A. S. Kabankin, and M. A. Landau, Zh. Strukt. Khim., **26**, No. 6, 129-131 (1985).
8. I. V. Peredunova and Yu. A. Kruglyak, Zh. Strukt. Khim., **29**, No. 1, 174-176 (1988).

PAKE NUCLEAR RESONANCE DOUBLET IN STRUCTURAL
INVESTIGATIONS UNDER MAGNETIC FIELD
ORIENTATIONAL MODULATION

B. F. Alekseev, M. B. Gaifullin,
and E. A. Sizova

UDC 539.143.43

In nuclear spin systems containing almost isolated pairs or triplets of interacting nuclei, complex multiplet structure is observed for the NMR structure [1-3]. In the simple case of a pair of nuclei coupled to each other by a dipole—dipole interaction (as the protons in the water molecule in crystal hydrates), the spectrum shows two lines of the Pake nuclear resonance doublet [3]. Analysis of such multiplet spectra, especially their angular relationships, gives important chemical structural information on the sample studied [1-3], including internuclear distances, orientation of the internuclear vectors, and the nature of nuclear motion.

The "differential course" of the spectrum upon adiabatic orientational modulation (OM) of the magnetic field has been discussed repeatedly in the literature relative to structural studies of crystals with magnetically-anisotropical interactions. This still uncommon procedure for spectral detection has been used in ESR [2, 4-8] and NMR spectroscopy [9]. Features of the Pake nuclear magnetic resonance doublet observed under conditions of adiabatic magnetic field OM are discussed for the first time relative to structural investigation of crystals.

Let us single out the only part significant for the subsequent secular part in the Hamiltonian of a system of two nuclei with spin $I = \frac{1}{2}$ and identical magnetomechanical ratio γ related by dipole—dipole interaction (generally accepted designations):

$$\mathcal{H} = - \gamma \hbar H_0 (I_{1z} + I_{2z}) + \frac{\gamma^2 \hbar^2}{r^3} \left(1 - 3\cos^2 \theta_0\right) \left[I_{1z}I_{2z} - \frac{1}{4} \left(I_1^+ I_2^- + I_1^- I_2^+ \right) \right]. \tag{1}$$

Four eigenvectors $|\psi_{1,2,3,4}\rangle$ of the stationary Schroedinger equation are known. Here, the vector $|\psi_1\rangle \equiv |\beta_1\beta_2\rangle$. The energy of this vector gradually decreases with increasing state number.

Let us assume that the modulating magnetic field $H_m(t) = H_m\cos\Omega t$ is linearly polarized and directed along the x-axis of the laboratory coordinate system (LCS), while the static magnetic field H_0 (the LCS z-axis) and vector r (z'-axis), which connects the coupled nuclei, are located in the same plane $x0z$. The effect of perpendicular modulation corresponds to rocking of the flux vector of the resultant field $H(t) = H_0 + H_m(t)$ in the $x0z$ plane relative to its position at an angle θ_0 to the z'-axis in the absence of modulation.

At low modulation amplitude when $H_m/H_0 = k \ll 1$, we retain only first-order terms relative to k in the expansions for $H(t)$ and $\cos^2\theta(t)$:

$$H(t) \simeq H_0; \quad \cos^2\theta(t) \simeq \cos^2\theta_0 - k\sin 2\theta_0\cos \Omega t.$$

Then, we obtain a Hamiltonian with a weakly oscillating dipole component instead of stationary Hamiltonian (1),

$$\mathcal{H}(t) = - \gamma \hbar H_0 (I_{1z} + I_{2z}) + \frac{\gamma^2 \hbar^2}{r^3} \left(1 - 3\cos^2 \theta_0 + 3k \sin 2\theta_0 \cos \Omega t\right) \times$$
$$\times \left[I_{1z}I_{2z} - \frac{1}{4} \left(I_1^+ I_2^- + I_1^- I_2^+ \right) \right]. \tag{2}$$