ISSN 1070-3632, Russian Journal of General Chemistry, 2013, Vol. 83, No. 1, pp. 154–156. © Pleiades Publishing, Ltd., 2013.
Original Russian Text © I.G. Zenkevich, 2013, published in Zhurnal Obshchei Khimii, 2013, Vol. 83, No. 1, pp. 164–166.

LETTERS TO THE EDITOR

On the New Proposals of Modification of the Nomenclature of Chemical Compounds

I. G. Zenkevich

St. Petersburg State University, Universitetskii pr. 26, St. Petersburg, 198504 Russia
e-mail: izenkevich@mail15.com

Received July 26, 2012

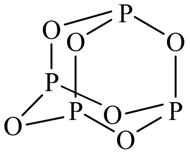
DOI: 10.1134/S1070363213010349

The University of Kragujevac (Serbia Republic) publishes a series of monographs (the series editor, Professor Ivan Gutman), joined under the theme of Mathematical Chemistry Monographs (MCM). The eleventh edition of this series consists of the monograph by Professor Seymour B. Elk (Elk Technical Associates, New Milford, New Jersey, USA) entitled "The Structure-Nomenclature Cycle of Chemistry" [1]. The author of the monograph is a renowned expert in matters of molecular topology, geometry, and the grounds of graph theory in chemistry, chemometrics, and orismology (the science of defining the technical and scientific terms). Key papers of the author of the monograph (over 50 publications) correspond to the period of 1990–2000. The publication [1] is hardly accessible to specialists because of the limited edition (100 copies). It is expedient to review the monograph [1] from the standpoint of analysis of current trends in the fields of chemistry like mathematical chemistry and chemometrics.

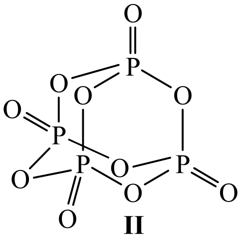
Monograph [1] continues a systematic exposition of the author's suggestions concerning modification of current nomenclature of chemical compounds, which has begun previously (Elsevier, 2004) in the book "A New Unifying Biparametric Nomenclature that Spans all of Chemistry" [2]. The main achievement of the author seems not just the refinement of the nomenclature but the development of encoding the structural features of both organic and inorganic compounds insufficiently accounted for by the modern IUPAC rules. The text is prefaced by a detailed dictionary that explains more than 200 chemical and chemometric terms, including those newly introduced by the author. The monograph is ended by the section considering the topology-nomenclature features of the polymer structures. The main part of the monograph concerns the problems of compounds with low molecular weight (in the author's terminology, "limited geometry").

The priority in the nomenclature of organic compounds has always been an adequate one-to-one verbal representation of the relative positions of atoms in molecules [3]. Among the key provisions of the monograph [1] is a concept of necessary detailed coding not only of the mutual arrangement of atoms, but also of the type of chemical bonds. Thus, in addition to conventional in chemistry σ- and π-bonds it is proposed to use the concept of α-, β-, γ-, and ζ-bonds, the covalent bonds with the order of approximately 0.5, 1.5, 2.5 and less than 1/3, respectively. This consideration is aimed at objects complex enough for the "classical" nomenclature like boron hydrides, phosphorus sulfides, polycyclic structures (cubane and some of its analogues, like cuneane and isocuneane), the crown ether complexes with metal cations, metal cation complexes with cyclopentadiene ("cenes"), metal carbonyls, some tautomeric structures, aromatic compounds, ets. It is hardly possible to present the basic rules of the symbolism proposed in [1] in this brief report, therefore it is advisable to show some of the most illustrative examples of its application.

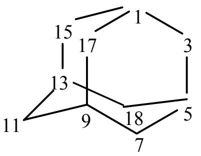
Thus, the phosphorus(III) dimeric oxide $(P_2O_3)_2 = P_4O_6$, which is suggested to have an adamantane-like structure **I**, in the proposed system [1,2] is "named" $(P1O1)_4{:}^{(1\text{-}9,5\text{-}13)}(1O1)$; dimeric phosphorrus(V) oxide (**V**) $(P_2O_5)_2 = P_4O_{10}$ (**II**) is "named" $(P1O1)_4{:}^{(1\text{-}9,5\text{-}13)}(1O1);^{(1,5,9,13)}(2O)$, and the corresponding hydrocarbon structural analog, the adamantane itself (tricyclo[3.3.1.13,7]decane, the



I



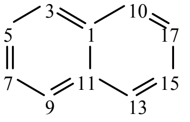
II



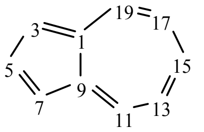
III

numbering of carbon atoms in structure **III** corresponds to the proposed in [2]) is "named" $(\underline{C}1\mathbf{C}1)_4{:}^{(1\text{-}9,5\text{-}13)}(1\mathbf{C}1)$.

Already this example shows clearly enough that these descriptors are not the names of the compounds, but their structural codes. The numbers in the rows correspond to the bond orders, subscripts outside the parentheses show the multiplicity of the structural elements in parentheses, the numbers in superscripts indicate the numbers of atoms bound by the bridging fragments. A single underline of the character of element ($\underline{C}$) corresponds to the presence at the atom of a single hydrogen atom (tertiary carbon). Therewith, two **C** characters shown in the adamantane "name" in bold, in accordance with the original symbolism [1, 2] should be underlined with two lines (secondary carbon atoms with two hydrogen atoms), but to find and use the corresponding symbol in the Microsoft Word software is a very non-standard operation, so it preferrably is not used. Localization of hydrogen atoms can be written in another way, as is evident from the examples of naphthalene (**IV**): $[(C\beta)_{10}{:}^{(1\text{–}11)}(\beta);^{(3,5,7,9,13,15,17,19)}(1H)]$ and azulene (**V**) $[(C\beta)10{:}^{(1\text{-}9)}(1)^{(3,5,7,11,13,15,17,19)}(1H)]$. This possibility illustrates some "redundancy" and the variability of the proposed coding system and can undoubtedly be regarded as its advantage.



IV



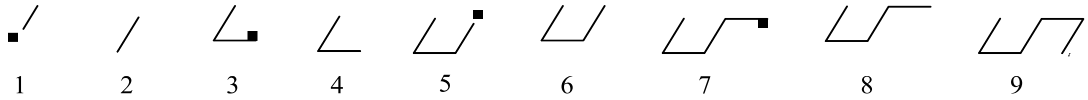
V

Two last examples include notification of β-bonds with the order of about 1.5.

Ferrocene $Fe(C_5H_5)_2$, a molecule which can be described using ζ-links, is coded by a set of characters: $Fe[\zeta(C\beta)_4C\zeta;^{(3\text{-}11)}\beta;^{(1\text{-}5,1\text{-}7,1\text{-}9)}(\zeta)]_2$. When the bond orders are not obvious the question marks can be used, as, for example, for iron pentacarbonyl, which is described by the code $Fe(?C3O{:})_5$. Examples of α-bonds are given by the boron hydrides. For example, the code of compound B_5H_9 (a trigonal bipyramid of the B atoms) has the form: $(B\alpha H\alpha)_4B{:}^{(1\text{-}5,1\text{-}9,1\text{-}13,5\text{-}17,9\text{-}17,13\text{-}17)}.$ $\alpha;^{(1\text{-}9,1\text{-}13,5\text{-}13,5\text{-}17,9\text{-}17)}(\alpha H\alpha)$. The concept of the α-bonds makes it easy to encode the tautomeric equilibrium [1, pp. 179–181], like keto-enol, imine-enamine, oxime-nitroso, and others.

However, the complexity should be noted of the interpretation of the notation system and the method of coding chemical structures proposed in [1], especially in the early stages of study. By this criterion, it can be considered intermediate between the previously proposed systems of line coding the structures of organic compounds like the WLN (Wiswesser Line Notation) [4] and SMILES (Simplified Molecular Input Line Entry Specification) [5]. It should be borne in mind that at present (at the avalanche-like increase in the amount of scientific information), any attempts to revise the long-established systems of notation are unlikely to be effective. The author of [1] not explains the monograph purpose and, therefore, the target remains unclear and, above all, the circle of readers to whom it is addressed. It is unlikely that it can be recommended for students whose primary goal is to master the "traditional" nomenclature [3]. As for the relatively narrow circle of specialists, for them the value of the new coding system at a wide use of the IUPAC nomenclature is also not clear. The concept of monographs [1,2] seems to be more in line not with a new system of nomenclature, but with another form of line coding structures of chemical compounds.

In connection with the above, it is interesting to draw an analogy between the recommendations offered in [1] and the following example. In the early 1970s, (that is, before the formation of chemometrics as an independent scientific discipline), Laszlo Molnar (Hungary) had completed the formation of a new system of writing digits and numbers, and performing all arithmetic operations based on the principles of shorthand, known as "stenometry" [6]. The first nine digits (see below the symbols in the firs row of and related Arabic numerals in the second row) is actually quite rational (in this case, unusually, a special symbol for zero can be omitted):



However, to explain arithmetic operations using graphic ones with these symbols [6] already took a few dozen pages (about 70). As a result, there was no examples of serious application of the stenometry to date, and widespread use of personal computers since the 1980s has made ineffective further development of this analogous approach. Consequently, such a revision of the grounds of arithmetic resulted in, though not uninteresting, but in fact unhelpful recommendations. The same fate may expect too radical proposals to revise the guidelines in the chemistry and, in particular, in chemical nomenclature.

Meanwhile, in the modern nomenclature of organic compounds do exist, if not a serious problem, but, at least, unresolved issues. Thus, by definition, any nomenclature is focused on specific chemical compounds [3]. However, many analytical methods for the determination of the structure of organic compounds (including chromatomass-spectrometry) may not result in an unambiguous answer on the structural level, but can reveal unequivocally the assignment of a compound a certain class (as a result of reference to the corresponding homologous series). In solving such problems, paradoxically, there are problems with the correct names of homologous series. It is not very easy, for example, to name correctly such a series as the aromatic unsaturated carbonyl compounds, especially taking into account that, depending on the type of conjugation of three structural fragments of molecules containing a π-system it should be split into several subgroups characterized by different UV- and mass-spectral parameters.

To solve such problems, in [7,8] a coding system was proposed for the homologous series, suggesting the use of only well-known and readily interpretable chemical symbols. So, the aromatic system is designated by the symbol Ar, carbonyl group by CO, and double bond by "=." The presence of conjugation is encoded by the symbol "**.**," its absence by "**,**." Then, for aromatic unsaturated carbonyl compounds, we obtain seven options of conjugation of the structural fragments of molecule, all are distinguishable by the spectral characteristics of homologs:

Ar,CO,= Ar**.**CO,= Ar,CO**.**= Ar**.**=,CO Ar**.**CO**.**=
CO**.**Ar**.**= CO**.**=**.**Ar

The proposed method has been used to interpret the results of group identification on the chromatographic retention parameters on the polar and nonpolar stationary phases [7], and to verify the correctness of "assembly" of a molecular structure by the evaluation of chromatographic retention indices using the new versions of additive schemes [8]. Thus we can assume that any system of coding the names of both specific chemical compounds and homologous series cannot be considered as a target, but seem to be rather a tool for solving various chemical problems.

Nevertheless, the considered monograph [1] certainly is interesting as an example of a logically consistent system of coding complex structures of chemical compounds. We can expect its further development in several ways.

REFERENCES

1. Elk, S.E., *The Structure-Nomenclature Cycle of Chemistry*, Kraguevac, Serbia: Univ. Kraguevac, 2011.
2. Elk, S.B., *A New Unifying Biparametric Nomenclature That Spans all of Chemistry*, Amsterdam: Elsevier, 2004.
3. Khlebnikov, A.F. and Novikov, M.S., *Sovremennaya nomenklatura organicheskikh soedinenii: uchebno-spravochnoe posobie* (The Modern Nomenclature of Organic Compounds: Training and Reference Manual), St. Petersburg: NGO "Professional," 2004.
4. Smith, E.G., *The Wiswesser Line Formula Chemical Notation*, New York: McGray Hill Co., 1968.
5. *http://en.wikipedia.org/wiki/Simplified_molecular_input_line_entry_specification* (July 2012).
6. Molnar, L., Stenometry. *A New Method for Rapid Calculations*, Budapest: Terra, 1972.
7. Zenkevich, I.G. and Chupalov, A.A., *Vestn. S.-Peterburg. Univ., Ser. Fiz. i Khim.*, 1994, no. 4, p. 45.
8. Zenkevich, I.G. and Ukolov, A.I., *Zh. Strukt. Khim.*, 2010, vol. 51, no. 4, p. 671.