Russian Journal of General Chemistry, Vol. 72, No. 4, 2002, pp. 569–574. Translated from Zhurnal Obshchei Khimii, Vol. 72, No. 4, 2002, pp. 611–616.

Information as a Factor of Self-Organization and Organization of Matter

V. B. Aleskovskii

St. Petersburg State University, St. Petersburg, Russia

Received May 25, 2001

Abstract—Data on biosynthesis and similar artificial processes of chemical-informative synthesis are summarized to establish that information is nothing but the property of matter to adapt to the existence conditions by structuring. In other words, information is a factor of self-organization of matter. The regularities of matter organization on a given level of structural hierarchy, including the prebiotic one, are established. The essence is the following. At the formation of a stationary dissipative structure, the production of information in the process of structuring attains a maximum, so that structuring is the process of producing and materializing information. The absence in inanimate nature of polyatomic chemical individuums, polymers and solids of constant composition, is noted and explained by the negligible probability of formation of polyatomic chemical individuums. However, this probability can be heightened by the afflux of large information produced in the processes of chemical-informative synthesis. The creation of nature-friendly chemical-informative technologies on the basis of biosynthesis is forecasted.

1. PHENOMENON OF NON-STOICHIOMETRY

It is accepted by this time that, in addition to individual compounds of permanent composition, there are individual compounds of variable composition, non-stoichiometric compounds, including all polymers and compounds, i.e. the overwhelming majority of substances [1]. As for solid compounds of permanent composition, they seem to be unreal, since, being crystalline substances, they, at a temperature above 0 K, always contain defects disturbing their stoichiometry. However, the compound variation can only occur at the expense of the medium. According to the mass conservation law, defects arising in the bulk of an isolated substance, cannot change its composition. Therefore, defects are beside the point. Polymers have a statistically average molecular mass and solids have variable composition for the simple reason that they are mixtures of corresponding homologs and analogs, affinitive substances. Unfortunately, neither this fact [2], nor even the synthesis of individual polymers [3] and of solid compounds of permanent composition [4] brought the notion of non-stoichiometry into challenge.

2. CONDITION OF SYNTHESIS OF POLYATOMIC CHEMICAL INDIVIDUUMS

Why is nature inhabited with gaseous and liquid low-molecular individual substances, with their numerous mixtures, with a variety of solid solutions, but with no one solid, generally, polyatomic chemical individuum, except biopolymers? The answer is given by the theory of probability: the formation of polyatomic individuums is an extremely rare, practically unrealizable event. However, the existence of biopolymers, as well as of the above polyatomic products of chemical synthesis, compounds of permanent composition [3, 4], conveys the suggestion that the synthesis of polyatomic individuums is not prohibited by nature.

At executing the synthesis

$$A + B = AB \qquad (1)$$

they always presume that, when process (1) attains equilibrium, a mixture consisting of substance AB and its dissociation products A and B is formed rather than substance AB. If AB is a low-molecular compound, one should take reagents A and B in some excess (and, later on, eliminate their residues from the reaction sphere) to obtain AB in an individual state by shifting the equilibrium toward AB. This technique, however, is suitable for obtaining simplest compounds only. The synthesis of polyatomic, high-molecular, and supramolecular substances under equilibrium conditions gives, instead of a single compound, an unseparable mixture of all countless components consisting of A and B, including isomers and homologs. Individual polyatomic chemical compounds cannot be obtained by equilibrum synthesis.

1070-3632/02/7204-0569 $27.00

However, chemical synthesis of polyatomic chemical individuums is possible in irreversible processes, namely, in such processes where, as in biosynthesis, formation of target products is related to production and materialization of a certain amount of information and its embodiment into the structure of the products [5]. Such processes, as real processes at all, proceed far from equilibrium [6].

3. PRINCIPLES OF CHEMICAL INFORMATICS

In 1948 Wiener published “Cybernetics” including the theory of information. Later on, Shannon showed that the amount of information I_S contained in a text is given by the equation

$$I_S = -\sum_i P_i \ln P_i. \quad (2)$$

Here P_i is the probability of appearing the *i*th symbol in a given text (capital or small letters, punctuation marks, and spacings). Equation (2) accounts for main peculiarities of the text: its distinctness, queerness, and eccentricity, i.e. individuality at all. This can be said about any structure containing whatever units: letters, atoms, atomic groups, clusters, molecular blocks, macromolecules, microscopic bodies, and cells. Multiplying I_S by approximately 1.44, we obtain the amount of information I expressed in bits: $I = I_S/\ln 2 \cong 1.44\ I_S$ bit. The larger the symbol (letter, generally, any structural unit) number n, the smaller quantity P and the larger the amount of information I contained in a given text [7], the information being in it only. Each of approximately 2^n rest possible texts is a random set of same symbols with no information. The information is transmitted only in the case when the symbols (letters, signs) are ordered in a certain sequence given by the program of text transmission, letters being chosen from a two-letter alphabet as a minimum. Just realizing a unique possibility of many others results in reproducing information. The wider the variety of possibilities to make a choice and, therefore, the smaller probability of the random appearance of a given text and the more unusual, unexpected, and original its structure, the larger information is contained in the text, as well as the fuller is the knowledge obtained at reading. Information is the more valuable the more unexpected.

It is of note that, accoording to Shannon’s formula, the amount of information is determined irrespective of its semantic content [7]. In the review [8], two Shannon’s definitions of information are considered, of which “one actually coincides with Boltzmann’s definition of entropy. This information, like Boltzmann’s entropy, is a measure of indeterminacy at a chosen level of statistical description of a system under consideration. For this reason, the term *S* information is used. ... Another definition is more adequate for open systems,”, i.e. the information of open systems [8]. This definition is concretized in the review. The general Shannon’s formula is transformed to find the dependence of information on control parameters. Herewith, it is discovered that this formula “can serve as a measure of information of open systems both in the processes of temporary evolution and at the evolution of control systems in space of control parameters” [8].

Zhdanov [9] took Shannon’s information entropy for a measure of ordering of the chemical structure of “polyatomic aggregates.” He introduced the notion of the active information capacity of a molecule Q which he defined as the sum of information entropy amounts H_i over all discrete states:

$$Q = \sum_i H_i. \quad (3)$$

Degenerate states are not taken into account. A criterion of information capacity, independent of the molecule size, is formulated as the specific information capacity per one atom:

$$I_A = \frac{Q_a}{\sum_i A_i} \text{ bit/atom}, \quad (4)$$

where A_i is the number of atoms on the *i*th level. The ratio of the algebraic sum of bonds with electropositive and electronegative atoms to the total number of carbon atoms in a molecule indicates the average degree of oxidation of a carbon atom in the molecule of a bioorganic compound. “The specific information capacity of a compound I_A is a generalized characteristic of structural-dynamic affluence and, with this, of the affluence of chemical possibilities of a compound. ...The average oxidation degree characterizes the state of atoms in a molecule” [9]. Herewith, junior terms of homologous series much differ from senior ones by smaller values of these parameters. It was noted that “small values of I for nucleic bases are related to that they should possess a rigid structure (i.e. relatively simple structure not containing excess information suitable for executing their function of transfer of biochemical information”[1] [9].

Frenkel [10] used the example of the synthesis of fibrillar proteins, fibrin and collagen, from the globular protein albumin, to establish that the product

[1] Comment of V. Aleskovskii.

of the information amount I, inserted by the initial substances in the process of synthesis, and the energy T consumed in the process is a constant quantity:

$$IT = \text{const.} \quad (5)$$

In other words, the amount of work T consumed in the process of chemical synthesis is inversely proportional to the amount of information contained in the structural units composing the structures of purposive products of the synthesis. It follows from here that using high-molecular and supramolecular compounds (not "bricks" but whole "precast panels") containing a large amount of information, minimizes the energy consumption. The rational joint of large and small structural units is secured by the information contained. The information imparts complementarity to those units which should joint each other to make this at first collision due to thermal motion.

It follows from the aforesaid that the information related to a substance characterizes its nature and processes in which it participates, that is all to be studied by chemistry. According to the dictionary [11], information is an account of the surrounding world and the processes running in it. In our case, it is meant the account included in chemical structure. Therefore, information is knowledge materialized in structures, and any structure is the "text" of a communication on the nature of a corresponding object in its own language. The language of matter is known to a certain extent. Its alphabet and first rules are contained in the Mendeleev table. However, substances themselves use a simpler, configuration language at their interaction. All data on the substance nature and programs of their interaction are written in that language, beginning with recognizing complementary partners (to be bound with a sufficiently large number of bonds) by molecules (macromolecules, supramolecules).

It has been known for a long time that, in parallel with "actual" chemical compounds, there are various molecular compounds: clathrates, molecular crystals, and molecular associates. Considering the appearance of the latter as the process of linking complementary molecules, Lehn [12] showed that the recognition of each other in this process consists in reading out the information contained in these molecules. It is of note that linking of molecules, macromolecules, and supramolecules, as well as of their fragments, occurs when the properties of these particles, colliding at thermal motion, are polar, but, at the same time, they can serve as complementary parts of a corresponding sufficiently stable compound. The complementarity of any structural units of matter is given by their inherent information encoded with a set of neutral and polar attributes (A, A–, A+), such as dimensions, shape, charge, nucleophility, etc. Thus, complementary structural units are programmed to linking each other with the probability close to unity. Therefore, participating in thermal motion, they form a molecular associate or a chemical compound just at the first accidental collision. As a rule, the set of complementarity attributes securing mutual recognition and linking structural units, widens and becomes more complex with increasing dimensionality and number of interacting structural units [13]. For example, the complementarity of globular (of zero dimensionality) structural units, if they are no more than three, is given by the correspondence of these units in shape, the size difference being of no significance. If the number of units is more than three, the possibility of linking of complementary structural units of smaller size in their interspaces appears. There are already two attributes in the set of complementarity attributes: configuration and dimensions. The attributes of complementarity of chain (one-dimensional) structural units are length, diameter, and the sequence of functional groups of various kinds, including hydrogen bonds. The set of complementarity attributes can include many other attributes, among which of special significance are those creating the dimension–structure correspondence of type antigen–antibody. It is understandable that preparation for synthesis of a polyatomic chemical individuum includes elements of architectonics, i.e. aggregating various structural units. Complementarity of interacting structural units is a determining factor in the process of formation of isomorphic mixtures (solid solutions) of polyatomic compounds. Complementarity allows them to joint by as large number of bonds as possible and to substitute each other in the structure of supramolecular associates. If these are completely complementary isostructural compounds of the same elements, they are linked by all their interatomic and intermolecular bonds. In this case, there is perfect isomorphism. Not similar substances can also be sufficiently complementary.

Shannon's equation determines the amount of information contained in a text as dependent on the probability of its formation from the set of symbols typing the text. Shannon's probabilistic-statistical theory of transfer of actual information makes the basement of informatics, a science that studies the laws and ways of collecting, treating, and transferring information which is taken from the knowledge obtained by studying nature. It is seen from the aforesaid that information can be created using structure-formation processes.

4. STRUCTURAL ORGANIZATION OF MATTER

Any material objects (matter, radiation) are discrete. Their structure is organized on one of structural hierarchy levels which is determined by the amount of information contained in a given object. Matter is a hierarchic multilevel dynamic system of interacting structural units constructed from lower level units which, in their turn, consist of units of an even lower level, and those consist of subunits, and so on, up to particles of 10^{-33} cm in size. The substances surrounding us are located on four structural levels, namely molecular, supramolecular, prebiotic, and biological. Herewith, the ubiquitous solid substances are not chemical individuums, but solid solutions whose formation is the most probable process. On the other side, the synthesis of polyatomic chemical individuums has an extremely low probability and is realized only in such processes that produce sufficiently large amounts of information, i.e. in the processes of structural organization of matter.

According to Prigogine's thermodynamic theory, a stationary state corresponds to the minimum entropy production under the external conditions preventing attainment of equilibrium. It is known, however, that entropy production is impossible without formation of a corresponding amount of information. Both in an isolated and in an open system in its stationary state, the sum of the amounts of entropy S and information I is constant [7]. The value of this constant is determined as follows. At an absolute order, $I = 1$ and $S = 0$ and, therefore, $S + I = 1$. Knowing the constant to be unity, we determine: $I = 1 - S$. Putting in this formula S_{min}, the entropy amount produced by a system under consideration in the stationary state, we obtain $I = 1 - S_{min}$. This means that the amount of information produced becomes a maximum in an open dynamic system when it has attained the stationary state.

Entropy is given to a medium. As for information, it is converted into matter as structural forms. Entropy simply cannot be realized beyond matter. By materializing, information imparts to particles of matter both structure and form and, with this, complementarity, i.e. the property to joint only with corresponding partners and only in a certain way, the property to keep and transfer information. If, for example, a dissipative dynamic structure crumbles, the information disappears, whereas the entropy increases up to a maximum: $S \to 1$, $I \to 0$. The information is conserved until a given structure exists. However, since information and structure do not exist without each other, they should be considered as a whole: structure as the body and information as the "soul" of matter, i.e. the property of adapting to existing conditions. Thus, matter structurization is actually the process of n conversion of information into matter.

If n is a big number, the multitude of N structures can be formed from n atoms of m elements. If they are equiprobable, the probability of formation of any one of the structures $P = 1/N$ is so small that the synthesis of a polyatomic individuum can proceed only under special conditions. As follows from the theory of information [14], the probability of a randomized event can be heightened by obtaining a corresponding amount of information, namely

$$I = \log_2 N, \text{ bit.}$$

How to do this? The use of the processes of formation of dissipative structures is not a means to this end, since relatively simple structures arise in these processes and, therefore, a small amount of information is produced. Biosynthesis is quite another matter: this process not only gives most complex structures, but also organisms, and, in principle, can produce unlimited amount of information. Biosynthesis is known to be programmed by genetic information contained in the chain structure of DNA. It generates copies in the course of biosynthesis (transcription of genetic information), and protein molecules are assembled on these copies. With this, proteins receive genetic information that programs self-assembling of these macromolecules. Various biologic structures are built in the course of self-assembling and further structuring.

According to our tentative hypothesis, this complex process is composed of two phase transitions. Crystallization gave permanent supramolecular structures, carriers and translators of a small amount of information. This was the first phase transition. In billions of years, in the prebiotic soup, in the process of sorption–desorption of various monomers on fine particles, polymeric dissipative structures including doubled DNA chains and lipid vesicles appeared. With the passage of time, vesicles captured DNA, separating its chains from the surrounding medium with a semipermeable shall and changing to self-organizing systems. The second phase transition occurred. Penetrating, together with other monomers, inside vesicles, amino acids were collected on DNA as on a matrix, which resulted in filling the vesicles with proteins. With time, the amount of proteins increased so much that vesicles divaricated. The ability to fission made such supramolecular complexes prebiotic individuums. With the appearance of these precursors of organisms in nature, there began the process of transfer of genetic information, increasing the amount of information and heightening creative potential.

Not going into details, we note the following. Supramolecular structures adapted to external conditions, by forming solid solutions and by polymorphous transformations, and remained on the relative low level of structural organization. At the same time, precursors of life were subjected to mutation and, breeding in the process of natural selection, heightened the level of their structural organization, and, with this, the ability of adaptation to external conditions. Further evolution was steeply accelerated due to spontaneous self-assembling and symbiosis appeared since the time of unicellular organisms. The correctness of this, generally not new, hypothesis is proved by experiment. One can mention the matrix synthesis of polypeptides [3] and the chemical assembling of supramolecules [4], as well as the self-assembling of polymolecular assemblies [12], polyatomic chemical individuums. These processes, as established in [15–17], are chemical models of biological processes of matrix synthesis and molecular assembling, i.e. of the processes of transcription and translation of information. Experiment showed that, using such processes, one can obtain any complex polyatomic structures and, therefore, unlimited amount of information.

Chemical assembling is performed by multiply alternating, according to a certain program, irreversible chemisorptions, on a solid surface, of atoms A of reagent 1 and of atoms B of reagent 2, that form a chemical bond with each other. The surface area playing the role of a matrix is constant. The synthesis program gives thermodynamic and hydrodynamic parameters of a given process, the sequence of alteration of monolayers of different composition, the number of cycles of such molecular layering, the character of each reaction, and the reaction duration necessary not for attaining equilibrium, as usual, but for 100% covering the sorbent surface. For example, after n cycles of molecular layering, one can obtain an individual solid compound $[AB]_n$ with observing *de visu* how a programmed substance arises, atom by atom, on the solid surface, and the information introduced by the process program materializes literally.

Molecular self-assembling [11] is still a more effective process of information production, since, proceeding spontaneously, it leads to instantaneous summation of information of many aggregating complementary molecules, macromolecules, and supramolecules, each of the latter introducing a huge amount of information in this process. Combining chemical assembling with self-assembling, one can realize a synthesis, similar to biosynthesis, able to heighten the level of structural organization of matter to a maximum, provided a corresponding program is available. By the way, nothing hinders using, on the first stage, the structure of RNA or a similar synthetic structure.

Thus, by lowering entropy production to a minimum, when a system attains the stationary state, is a consequence of the following. With the passage of time, matter, being under external influence, acquires a structure that minimized both this influence and the own resistance to it. Just this ability of matter to adapt to the existence conditions by self-organization, called information, predetermines the evolution of matter wherever it is. If this is so, information is nothing but as the property of matter to adapt to the existence conditions by structuring.

As noted in [17], the exploration of technology of the above processes of production and materialization of information will transform it to the chemicalinformation technology whose processes, like biological ones, can be harmonically involved in the global ecosystem, i.e. to function with no harm for the habitat. At last, this will allow the mankind to realize the age-long dream “To live in accordance with the nature of things, not to deviate from it, to obey its laws, to take a pattern by it” (Seneca).

ACKNOWLEDGMENTS

I am deeply grateful to all my dear colleagues, the members of the scientific school on highly organized substances, for fruitful collaboration.

This investigation was financially supported by the Russian Foundation for Basic Research (project no. 00-15-97356).

REFERENCES

1. *Khimicheskaya entsiklopediya* (Chemical Encyclopaedia), Knunyants, I.L., Ed., Moscow: Sovetskaya Entsiklopediya, 1995, vol. 4, p. 79.
2. Shchukarev, S.A., *Vest. Leningr. Univ.*, 1947, no. 5, p. 5.
3. Merrifield, R.B., *J. Am. Chem. Soc.*, 1963, vol. 85, p. 2149.
4. Aleskovskii, V.B., *Zh. Prikl. Khim.*, 1974, vol. 47, p. 2145.
5. Aleskovskii, V.B., *Vest. S.-Peterb. Gos. Univ., Ser. 4: Fiz., Khim.*, 1992, no. 25, p. 75.
6. Prigogine, I., *From Being to Becoming*, San Francisco: Freeman, 1980.

7. Kadomtsev, B.B., *Usp. Fiz. Nauk*, 1994, vol. 164, p. 449.
8. Klimantovich, Yu.L., *Usp. Fiz. Nauk*, 1999, vol. 169, p. 443.
9. Zhdanov, Yu.A., *Entropiya informatsii v organicheskoi khimii* (Information Entropy in Organic Chemistry), Rostov-on-Don: Rostov Univ., 1979.
10. Frenkel, S., *Sov. Sci. Rev.*, 1971, p. 110.
11. *Slovar' russkogo yazyka* (Dictionary of Russian), Moscow: Russkii Yazyk, 1981.
12. Lehn, J.M., *Supramolecular Chemistry*, New York: VCH, 1995.
13. Aleskovskii, V.B., *Khimiya nadmolekulyarnykh soedinenii* (Chemistry of Supramolecular Compounds), St. Petersburg: S.-Peterb. Univ., 1996.
14. Kolmogorov, A.N., *Teoriya informatsii* (Theory of Information), Moscow: Nauka, 1987.
15. Aleskovskii, V.B., *Vestn. Akad. Nauk SSSR*, 1975, vol. 6, p. 48.
16. Aleskovskii, V.B., *Zh. Obshch. Khim.*, 1995, vol. 65, no. 11, p. 1928.
17. Aleskovskii, V., *Macromol. Symp.*, 1998, vol. 136, p. 1.
18. Aleskovskii, V.B., *Vestn. S.-Peterb. Gos. Univ., Ser. 4: Fiz., Khim.*, 2001, no. 28.