# Information Entropy of Fullerenes

Denis Sh. Sabirov[*,†] and Eiji Ōsawa[‡]

[†]Institute of Petrochemistry and Catalysis, Russian Academy of Sciences, 450075 Ufa, Russia

[‡]NanoCarbon Research Institute, AREC, Faculty of Textile Science and Technology, Shinshu University, Ueda, Nagano 386-8567, Japan

Ⓢ Supporting Information

ABSTRACT: The reasons for the formation of the highly symmetric C_{60} molecule under nonequilibrium conditions are widely discussed as it dominates over numerous similar fullerene structures. In such conditions, evolution of structure rather than energy defines the processes. We have first studied the diversity of fullerenes in terms of information entropy. Sorting 2079 structures from *An Atlas of Fullerenes* [Fowler, P. W.; Manolopoulos, D. E. *An Atlas of Fullerenes*; Oxford: Clarendon, 1995.], we have found that the information entropies of only 14 fullerenes (<1% of the studied structures) lie between the values of C_{60} and C_{70}, the two most abundant fullerenes. Interestingly, buckminsterfullerene is the only fullerene with zero information entropy, i.e., an exclusive compound among the other members of the fullerene family. Such an efficient sorting demonstrates possible relevance of information entropy to chemical processes. For this reason, we have introduced an algorithm for calculating changes in information entropy at chemical transformations. The preliminary calculations of changes in information entropy at the selected fullerene reactions show good agreement with thermochemical data.
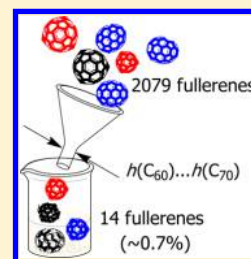
■ INTRODUCTION

The last century is rationalized as a starting point and rapid development of information sciences. Partly, this is due to the working out of perfect computational and communication facilities and originates from the pioneering works that introduced basic definitions of this field. Earliest Hartley's[1] (1928) and Shannon's[2] (1948) works have defined information as a logarithmic measure of uncertainty. Applying this approach to some experience, we expect a finite set of n outcomes A_1, A_2, ..., A_n with the respective probabilities p_1, p_2, ..., p_i. According to Shannon,[2] each outcome is characterized with the uncertainty equal to $-\log p_i$. For the set of possible outcomes, information entropy h (or Shannon entropy) is introduced as a measure of uncertainty:

$$h = -\sum_{i}^{n} p_i \log p_i \tag{1}$$

This concept with more than 50 years of history remains fruitful not only for the theory of information[3] but also for natural sciences. Trucco in 1956 demonstrated that information entropy appropriately describes information capacity of graphs.[4] As molecules in some approaches are approximated as graphs, it was not surprising that the information theory then has come into chemistry. For example, it is used in the original and modified forms to describe structural diversity of organic compounds,[5−17] chemical reactions,[5,18−20] and physicochemical processes in open systems.[21−25] Shannon's approach is an efficient auxiliary tool when diverse nanosystems[24−28] and processes in them are studied (e.g., association of fullerenes[27] or nanotubes[28]).

In the original form, Shannon's approach is applied to atoms and/or bonds,[5,7−9,12−15,18,28] i.e., to the topological description

of molecular structure. Fullerenes are perfect candidates for this approach as a variety of group symmetries of the fullerenes are combined with the strict rules of their structures. Indeed, fullerenes previously became objectives of the topological studies,[29−33] which allow explaining most of their chemical properties and especially their stabilities. Partly due to the topological considerations, the widely used "isolated pentagon rule" has been deduced.[30]

Structural analysis of fullerenes retains high interest due to the following reasons. In spite of the successful rational synthesis of C_{60} from polycyclic hydrocarbons[34] and separation of less available (minor) fullerenes,[35,36] there are blind spots in their formation mechanism under plasma conditions.[37] Promising results in this field have been obtained from Irle and Morokuma's quantum mechanics−molecular dynamics simulations on the formation of fullerene structures from dicarbon molecules.[38] However, those did not reproduce the C_{60} formation in large amounts as these highly symmetric molecules are formed under hard conditions and dominate among the similar and seeming more likely products.[37] We think that some novel ideas concerning the C_{60} formation may come from the studies on structural diversity of the fullerenes family. As C_{60} arises under extremely hard and nonequilibrium conditions, energetic factors become background and evolution of the structure starts defining the process (as it defines numerous previously studied processes[39]).

Shannon's approach has never been applied to quantitative estimation of structural diversity of the fullerenes family. In the present work, we calculated information entropies of fullerenes in terms of the original Shannon's formalism and symmetry

considerations. The data obtained have been compared with the thermodynamic entropies calculated by the DFT method (experimental values obtained only for C_{60} and C_{70}). We have demonstrated that information entropy may be used as a sorting parameter to detect attainable fullerene structures. Hence, its application should not be limited by structural description as it relates to real chemical processes.

■ COMPUTATIONAL DETAILS

Shannon's approach is applied to the analysis of molecular structures in many ways. It depends on what molecular property and/or structural parameter is considered in the aspect of its uncertainty. For example, electron density can be such molecular property for calculation of information entropy.[20,21,24,40,41] Another way (we call it "discrete") is to consider the type of atoms or structural fragments to estimate the uncertainty of the molecular structure.[5,7–9,12–18] In the present work, we apply information entropy to the fullerene molecules in this simplest way. Hence, the proposed outcomes in eq 1 will be unequivalent types of carbon atoms of the fullerene molecules under consideration. The probability of finding atom of i type is given by a simple relation

$$p_i = \frac{N_i}{\sum_i^n N_i} \qquad (2)$$

where N_i is a number of i atoms, and $N = \Sigma N_i$ is a total number of carbon atoms in the molecule. Expression 1 can be rewritten into eq 3 using eq 2 and logarithm to the base of two (the last one allows expressing information entropy in bits):

$$h = -\sum_i^n \frac{N_i}{\sum_i^n N_i} \log_2 \frac{N_i}{\sum_i^n N_i} \qquad (3)$$

This is the main equation we use to analyze fullerene structures in terms of Shannon's approach.

It is interesting to compare information h and thermodynamic entropies $S°$ as the molecular structure defines both of them (though these two quantities are expressed in different units). We have calculated the thermodynamic entropies by the PBE/3ζ density functional theory method,[42] implemented in the Priroda program.[43] Previously, we have efficiently used this method in theoretical studies on diverse fullerene-containing systems and found that it allows reproducing most of the experimentally known properties of fullerenes and their derivatives (see reviews[44,45] and original works where justification of the chosen method is given in detail[46–49]). The thermodynamic entropies of C_{60} (I_h) and C_{70} (D_{5h}) fullerenes, calculated by this method, are 548.47 and 616.17 J mol^{-1} K^{-1} and well agree with experimental gas-phase values (544.0 and 641.0 J mol^{-1} K^{-1}, respectively[50]). Additionally, this method reproduces the known symmetries of fullerenes.[46,47] This is important for the present work requiring accurate symmetry considerations. The optimized geometries of fullerene molecules were used for the determination of their symmetry point groups in the ChemCraft program.[51]

■ RESULTS AND DISCUSSION

Information Entropy of C_{60} and C_{70} Fullerenes. First, we applied eq 3 to C_{60} (I_h) and C_{70} (D_{5h}) (Figure 1). In the case of the highly symmetric C_{60} molecule, all 60 carbon atoms are equivalent. Substituting $n = 1$ and $N = 60$ in eq 3 shows that information entropy of C_{60} equals zero. It reflects the fact that choosing one random atom in the buckminsterfullerene's
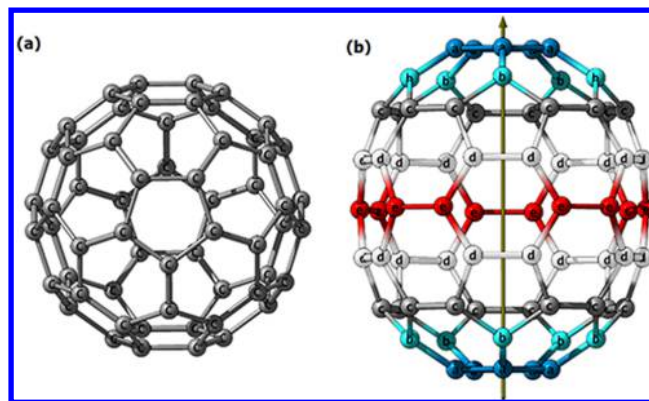
Figure 1. Molecular structures of C_{60} (I_h) (a) and C_{70} (D_{5h}) (b) fullerenes. Five types of atoms of the C_{70} molecule are colored; the C_5 symmetry axis of the C_{70} molecule is shown.

molecule is devoid of uncertainty because all of them are the same. The C_{70} (D_{5h}) molecule consists of the atoms of five types, designated as a–e.[52] These include 10 atoms a, 10 atoms b, 20 atoms c, 20 atoms d, and 10 atoms e, i.e., $n = 5$ and $N = 3 \times 10 + 2 \times 20$. Substituting these in eq 3 leads to $h = 2.236$ for C_{70} (D_{5h}). These simple examples show how information entropy quantitatively describes the difference in the structures of the two most abundant fullerenes.

As shown by calorimetric study,[53] the C_{70} fullerene is thermodynamically more stable in comparison with C_{60}. Hence, in the case of fullerenes, thermodynamic stability does not always correlate with probability of formation. As the buckminsterfullerene, being more abundant, obtains zero value, we propose that lower h values should indicate fullerenes, formed with higher probabilities.

Information Entropies of C_{60} Isomers. According to eq 3, information entropy depends on the number of atom types n and the numbers of atoms of each type N_i. Comparison of C_{60} (I_h) and C_{70} (D_{5h}) fullerenes demonstrates that these two values may be defined by the symmetry of the molecule. Indeed, C_{70} has a nonuniform structure due to its lower symmetry as compared to the truncated icosahedron of C_{60}.

To scrutinize the relation between the symmetry and the information entropy, we consider six isomers of C_{60} with C_1, C_2, C_{2v}, D_{2h}, D_{5d}, and I_h group symmetries (Figures 1a and 2). Previously, these isomeric species became an objective of theoretical[54] and experimental[55] works. As follows from the structural analysis, the number of atom types n in the set above diminishes from 60 to 1 for C_1 and I_h isomers, respectively. Hence, the calculated h values decrease from 5.907 to zero for this series C_1, C_2, C_{2v}, D_{2h}, D_{5d}, and I_h. Quantitatively, symmetry point groups are roughly described by the respective rotational numbers σ. As σ values define fullerene structure, we find that h and σ are correlated (Figure 3a).

Similar trend is found in the case of thermodynamic entropy of C_{60} isomers (Figure 3b) as it is also defined by symmetry. The lowest $S°$ value, as expected, has been found for the I_h isomer (548.47 J mol^{-1} K^{-1}), whereas the highest one corresponds to the C_1 isomer (591.26 J mol^{-1} K^{-1}). As we see, both entropies strongly depend on the fullerene structure, *viz.* its symmetry. Consequently, h and $S°$ values are symbate (Figure 3c). This seems to hold true for other isomeric fullerenes. Note that linear correlations between Shannon entropy (expressed in bit kmol^{-1}) and thermodynamic ones have been previously found for a series of homologues (alkanes,
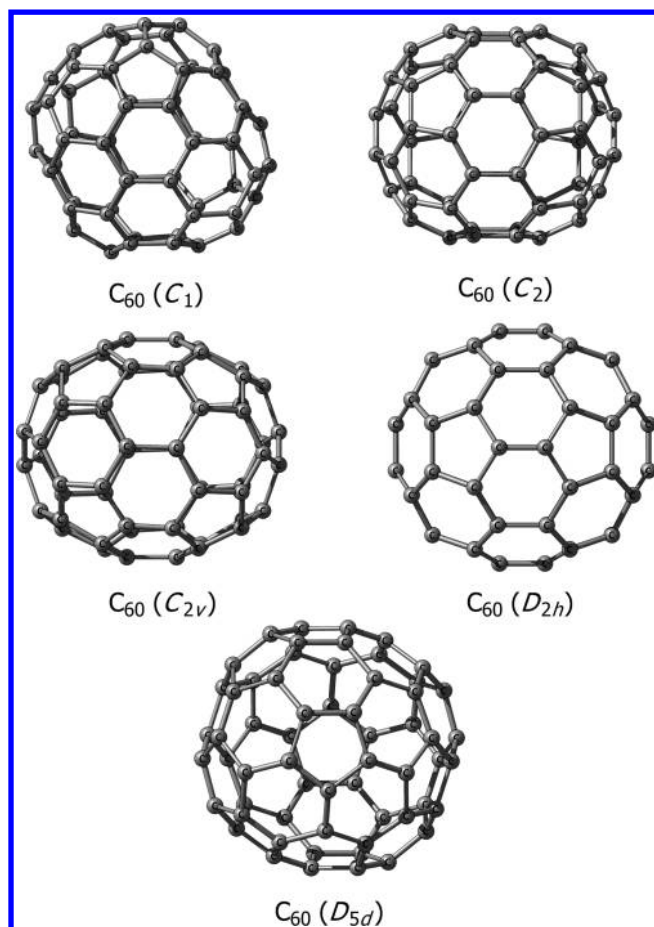
Figure 2. Structures of low-symmetry isomers of buckminsterfullerene.

alkenes, alcohols, *etc.*).[12] Thus, our results on fullerenes are consistent with the previous findings.

Analytical Expressions of Information Entropies for the Selected Classes of Fullerenes. Based on the example above, we conclude that the highest $S°$ and h values should be typical for the least symmetric C_1 fullerene isomers. In this case, the number of atom types n in the molecule equals to N, so each atom type consists of only one atom. These considerations allow assessing the maximal information entropy possible for the set of isomeric fullerenes with the same number of carbon atoms. According to eq 3, it is equal to

$$h_{max} = -\log_2 \frac{1}{N} \tag{4}$$

Similar analytic expressions can be simply derived for other low-symmetry species. Thus, the fullerene molecules with C_i and C_2 symmetries are made up by atoms of $N/2$ types consisting of 2 members. These considerations transforms eq 3 into the following expression:

$$h(C_N(C_i)) = h(C_N(C_2)) = -\log_2 \frac{2}{N} \tag{5}$$

In the case of fullerenes with C_s symmetry, m unequivalent atoms lie in the plane of symmetry, whereas $(N-m)$ others are divided into $(N-m)/2$ types consisting of 2 atoms. This leads to the following expression for h:

$$h(C_N(C_s)) = -\frac{m}{N}\log_2 \frac{1}{N} - \frac{N-m}{N}\log_2 \frac{2}{N} \tag{6}$$

Estimation of the opposite case of h_{min} is more complicated. The lowest h values should be obviously typical for the isomers with icosahedral and tetrahedral symmetries with the highest σ values, 30 and 12, respectively. We do not consider the octahedral symmetry group O_h with $\sigma = 24$ as no conventional fullerenes correspond to it.

Previously, two groups of icosahedral fullerenes representing Goldberg polyhedra with $60a^2$ and $20a^2$ atoms in the molecule (a is integer) were theoretically studied[56] (we should mention the growing theoretical interest in such fullerene structures[57−59]). Their regular structures allow quick counting types of unequivalent atoms in the molecules and the number of atoms of each type (Tables S2 and S3 in the Supporting Information). Application of Shannon's approach to the series of $20a^2$ (C_{20}, C_{80}, C_{180}, C_{320}, C_{500}, C_{720}) and $60a^2$ fullerenes (C_{60}, C_{240}, C_{540}, C_{960}) demonstrate the logarithmic growth of h with an increase in N (Figure 4). This agrees with the previous study on the homological series, in which information entropy increases with the number of atoms in molecule.[12] For example, the molecules of $60a^2$-fullerenes are made up by unequivalent atom groups consisting of 60 and 120 atoms. This allows deriving a formula that describes the relation between h and a values:

$$h(C_{60a^2}) = -\frac{1}{a}\log_2 \frac{1}{a^2} - \frac{a-1}{a}\log_2 \frac{2}{a^2} \tag{7}$$

Fullerenes C_{20} and C_{80}, the first members of the $20a^2$ series, have lower symmetries due to the Jahn−Teller effect, so the h values are higher than the values deduced from their idealized icosahedral symmetries. Precise optimizations of these structures by the PBE/3ζ method indicate reduced symmetries to C_i and C_{3v} for C_{20} and C_{80}, respectively. The propensity of C_{20} and C_{80} for the Jahn−Teller symmetry reduction is explained by the zero band gap (the frontier orbital energy level difference $|\varepsilon_{HOMO} - \varepsilon_{LUMO}|$); in the case of C_{60}, this value equals to 0.7566 $|\beta|$ (where β is the Hückel resonance integral).[29] These cases will be additionally discussed below. Simultaneously, according to our calculations, C_{180}, C_{320} (the next members of the $20a^2$ series), and C_{240} (the second member of the $60a^2$ series) keep the I_h group symmetry.

Information Entropy of the Fullerenes Family. We have briefly studied how both entropies depend on the fullerene size, i.e., total number of carbon atoms in the fullerene molecule. In the case of thermodynamic entropy, we have chosen one isomer for each N from numerous fullerene structures. This simplification is rather rough but suitable for searching a general trend. Indeed, the scattering of the $S°$ values of the mentioned C_{60} isomers is ∼50 J mol^{-1} K^{-1}. This is a small value if we superpose it with an overall $S°$ *versus* N plot (Figure 5). This plot indicates the trend of growing entropy with the increase in the size of the fullerene molecule. We do not discuss here a correlation between $S°$ and N values because it requires considering a larger number of fullerene structures. In the scope of the present work, it is important to recognize that the C_{60} fullerene is not an exceptional compound according to this trend since there are fullerenes with lower and larger thermodynamic entropies.

To study information entropy of the fullerenes family, we have chosen 2079 fullerenes indexed by Fowler and Manolopoulos in their Atlas.[29] No obvious correlation is observed with a total number of carbon atoms in the molecule (Figure 6). To find it out, one should separately study groups of the fullerene isomers. Symmetry considerations alone
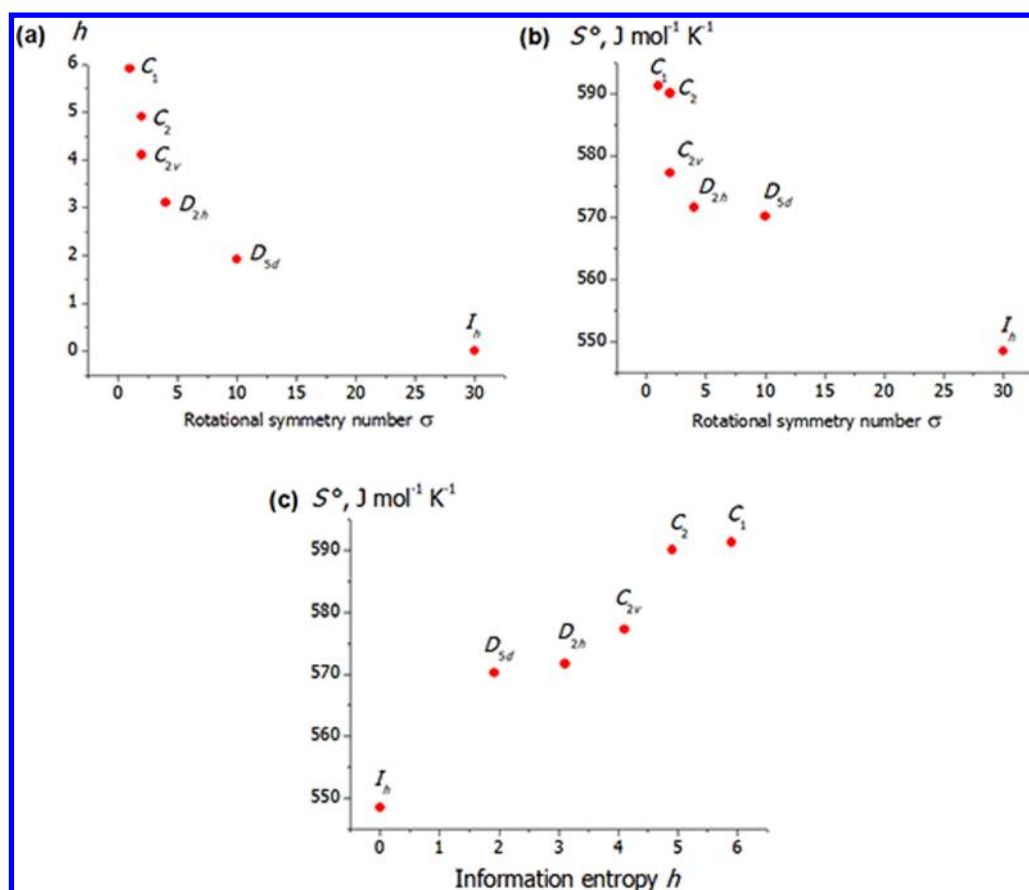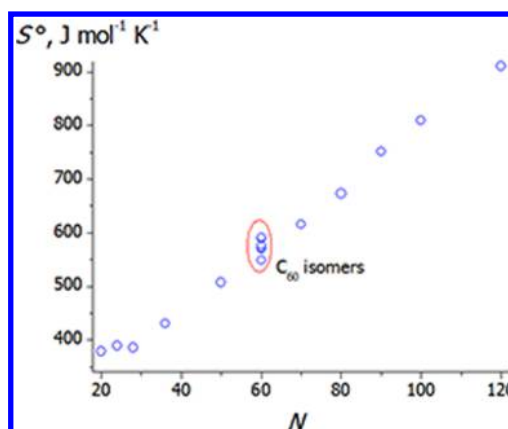
Figure 3. Relations between the calculated information entropy h, thermodynamic entropy $S°$, and rotational symmetry number σ for C_{60} fullerenes. Numerical data associated with the plot are given in the Supporting Information (Table S1).
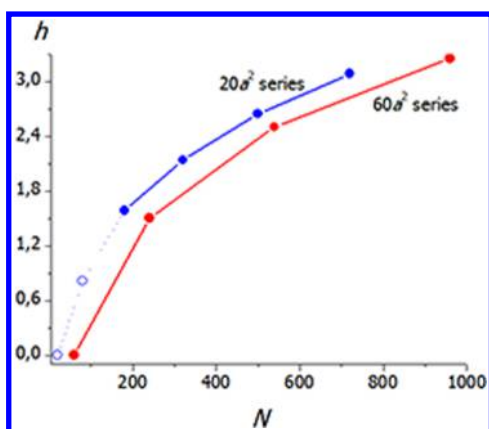
Figure 4. Dependence of information entropy for icosahedral Goldberg fullerenes on the number of carbon atoms in the molecule. Two points within the $20a^2$ series corresponding to C_{20} and C_{80} fullerenes have no color filling to stress that their h values are not fitted into the dependencies presented due to lower symmetries of the respective fullerenes. Numerical data associated with the plot are given in the Supporting Information (Tables S2 and S3).

Figure 5. Dependence of thermodynamic entropy on the number of carbon atoms in the fullerene molecules. Numerical data associated with the plot are given in the Supporting Information (Table S4).

become insufficient for this purpose (except for the isomers with C_1 and C_2 symmetries; for these species analytical expressions eqs 4 and 5 depend on the N number only).

Though there is no correlation between h and N for the whole fullerene family, the h versus N plot provides important information. As follows from Figure 6, most of the fullerenes under study are low symmetric, so their h values swarm in the upper part of the plot. To clearly demonstrate the last sentence, we have separated all the studied fullerenes by the line corresponding to the information entropy of the second abundant fullerene C_{70} (D_{5h}) with $h = 2.236$. Remarkably, only 24 fullerenes (including C_{60} and C_{70}) are characterized with $h \leq h(C_{70}$ $(D_{5h}))$; this is 1.15% of 2079 studied species. We pay particular attention to these selected fullerenes (Figure 7). Their h values and the respective partitioning are listed in Table 1. From the "low-h" members of the fullerenes family, only C_{20} (I_h) and C_{60} (I_h) primarily obtained zero information entropies. However, when calculating h, we operate with the idealized
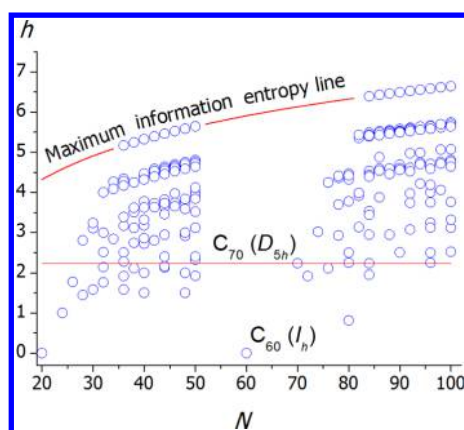
Figure 6. Dependence of information entropy on the number of carbon atoms in the fullerene isomers taken as is from *An Atlas of Fullerenes.*[29] Numerical data associated with the plot are given in the Supporting Information.

group symmetries, *viz.* highest symmetries determined with the topologies of the molecules without consideration of the Jahn–Teller effect reducing the symmetry. The mentioned C_{20} fullerene has a lower symmetry due to this effect.[35] Theoretical studies on the C_{20} fullerene did not succeed in the determination of its group symmetry. Diverse quantum-chemical methods predict structures with C_2, C_{2h}, C_i, D_{3d}, or D_{2h} symmetries for C_{20}.[60] We found C_i symmetry for the C_{20} molecule after the highly accurate optimization by the PBE/3ζ method. Therefore, C_{20} with I_h symmetry and $h = 0$ represents a hypothetical case. According to these considerations, the C_{60} fullerene remains the only fullerene with zero information entropy that makes it different from all other fullerene structures.

Similar to the C_{20} case, we have recalculated h values for the other low-h fullerenes according to the group symmetries obtained after the PBE/3ζ optimization. We use the semi-empirically calculated HOMO–LUMO band gaps[29] as an auxiliary tool for indicating fullerenes, which are prone to undergo the Jahn–Teller symmetry reduction. This way we have ruled out 9 other fullerenes C_{24}-1, C_{26}-1, C_{30}-1, C_{32}-3, C_{38}-2, C_{40}-1, C_{44}-3, C_{76}-2, and C_{80}-7 (in addition to the above-mentioned C_{20} case; Figure 7). The listed fullerenes are mostly hypothetical structures that were not observed in the experiments. Here, only trifluoromethylated derivatives of

Table 1. Fullerene Isomers with Information Entropies (h_{ideal}) Lying in the Range $h(C_{60})...h(C_{70})$ in the Case of Their Idealized Symmetries and after the DFT-Optimization (h)[c]

fullerene	band gap[a]	idealized symmetry	symmetry obtained	h_{ideal}[b]	h
C_{20}-1	0.0000	I_h (1 × 20)	C_i (10 × 2)	0.000	3.322
C_{24}-1	0.0000	D_{6d} (2 × 12)	C_2 (12 × 2)	1.000	3.585
C_{26}-1	0.0669	D_{3h} (1 × 2 + 2 × 6 + 1 × 12)	C_1 (26 × 1)	1.776	4.700
C_{28}-2	**0.0000**	T_d (1 × 4 + 2 × 12)	T_d (1 × 4 + 2 × 12)	**1.449**	**1.449**
C_{30}-1	0.0000	D_{5h} (3 × 10)	C_s (3 × 2 + 6 × 4)	1.585	3.107
C_{32}-3	0.0000	D_{3d} (1 × 2 + 1 × 6 + 2 × 12)	C_i (16 × 2)	1.764	4.000
C_{32}-5	0.3943	D_{3h} (1 × 2 + 3 × 6 + 1 × 12)	D_{3h} (1 × 2 + 3 × 6 + 1 × 12)	2.139	0.250
C_{36}-13	**0.0000**	D_{3h} (2 × 6 + 2 × 12)	D_{3h} (2 × 6 + 2 × 12)	**1.918**	**1.918**
C_{36}-15	**0.0000**	D_{6h} (3 × 12)	D_{6h} (3 × 12)	**1.585**	**1.585**
C_{38}-2	0.0000	D_{3h} (1 × 2 + 2 × 6 + 2 × 12)	C_s (3 × 2 + 8 × 4)	2.115	3.406
C_{40}-1	0.0000	D_{5d} (4 × 10)	C_i (20 × 2)	2.000	4.322
C_{40}-39	0.3731	D_{5d} (2 × 10 + 1 × 20)	D_{5d} (2 × 10 + 1 × 20)	**1.500**	**1.500**
C_{44}-3	0.0000	D_{3d} (1 × 2 + 1 × 6 + 3 × 12)	C_i (22 × 2)	2.128	4.459
C_{44}-73	**0.4059**	T_d (1 × 4 + 1 × 12 + 1 × 24)	T_d (1 × 4 + 1 × 12 + 1 × 24)	**2.163**	**2.163**
C_{48}-186	**0.2135**	D_{6d} (2 × 12 + 1 × 24)	D_{6d} (2 × 12 + 1 × 24)	**1.500**	**1.500**
C_{48}-189	**0.0000**	D_{6d} (4 × 12)	D_{6d} (4 × 12)	**2.000**	**2.000**
C_{50}-271	**0.1031**	D_{5h} (3 × 10 + 1 × 20)	D_{5h} (3 × 10 + 1 × 20)	**1.922**	**1.922**
C_{72}-1	**0.7023**	D_{6d} (2 × 12 + 2 × 24)	D_{6d} (2 × 12 + 2 × 24)	**1.918**	**1.918**
C_{76}-2	0.0000	T_d (1 × 4 + 2 × 12 + 2 × 24)	S_4 (3 × 4 + 4 × 8 + 2 × 16)	2.115	2.985
C_{80}-7	0.0000	I_h (1 × 20 + 1 × 60)	C_{3v} (1 × 2 + 1 × 6 + 6 × 12)	0.811	2.877
C_{84}-20	**0.6962**	T_d (1 × 12 + 3 × 24)	T_d (1 × 12 + 3 × 24)	**1.950**	**1.950**
C_{84}-24	**0.5293**	D_{6h} (3 × 12 + 2 × 24)	D_{6h} (3 × 12 + 2 × 24)	**2.236**	**2.236**

[a]Band gaps in β units (β is the Hückel resonance integral) are taken from ref 29. [b]Calculated for the structures with highest possible (idealized) symmetry. [c]The structures remaining in the range after the DFT-optimization are shown in bold.
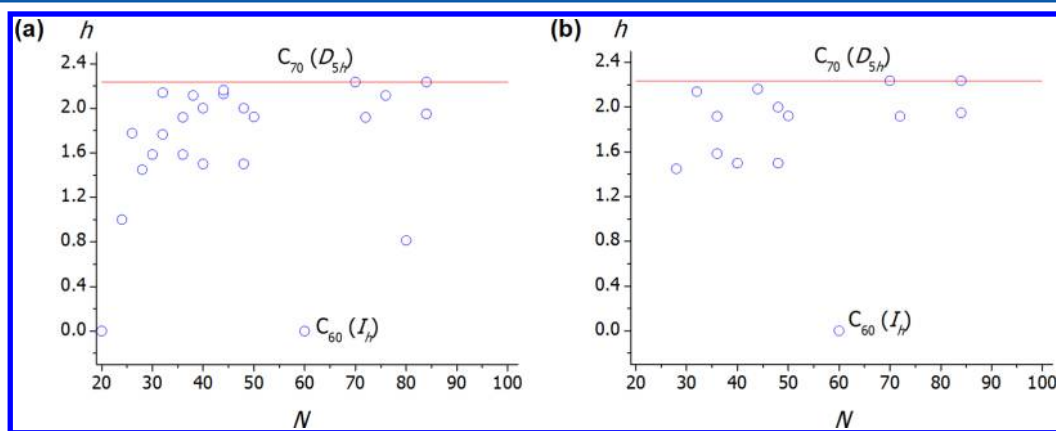
Figure 7. Information entropies of the fullerenes with h from the range $0 \leq h \leq h(C_{70}\ (D_{5h}))$: (a) the initial plot with 24 points; (b) the plot after the additional sorting with consideration of the Jahn–Teller effect (14 points). Numerical data associated with the plots are given in Table 1.

C_{76}-2[61] and endohedral complexes of C_{80}-7 with metals[62] were synthesized.

Most of 14 remaining members (keeping the initial symmetries) are experimentally observed in the original or functionalized forms. For example, the C_{36}-15 fullerene has been obtained as well as its exo- and endohedral compounds.[63] The endohedral complexes of C_{28}-2 with uranium[64] and the exohedral chloro-derivatives of C_{50}-271 were synthesized.[65] Remarkably, only one experimentally known isomer #271 has entered the discussed range of h values, whereas the other 270 isomers of C_{50} have not. Additionally, the highly intensive peaks of C_{32}, C_{44}, C_{36}, and C_{50} were previously detected in the mass spectra of the products of laser ablation of graphite,[66] whereas C_{84}-20 and C_{84}-24 fullerenes are produced in large amounts and separated by high-performance liquid chromatography.[67]

Unfortunately, we must mention some disadvantages of the information entropy approach applied to fullerenes. In addition to the fullerenes produced in larger or smaller amounts, some fullerene isomers with no experimental evidence for their existence (e.g., C_{36}-13, C_{40}-39, C_{48}-186, and C_{48}-189) receive low h values and, hence, enter the range $0 \leq h \leq h(C_{70}\ (D_{5h}))$. Conditions for their stability were theoretically studied, so the ways for their stabilizations were proposed.[35] However, the mentioned fullerenes or their derivatives have not been synthesized.

Additionally, the information entropy values do not provide the ratio of the synthesized isomeric fullerenes. For example, in the case of the C_{84} series, two minor isomers C_{84}-20 (T_d) and C_{84}-24 (D_{6h}) obtain low h values. At the same time, the major ones C_{84}-23 (D_{2d}) and C_{84}-22 (D_2) have higher information entropies (3.440 and 4.392, respectively), so these fullerenes do not match the range $0 \leq h \leq h(C_{70}\ (D_{5h}))$. This is also true for some other rare but experimentally known fullerenes (e.g., C_{76} and C_{78} isomers[36]). The fact that we arbitrarily chose the range $0 \leq h \leq h(C_{70}\ (D_{5h}))$ for sorting the fullerenes partly explains the observed mismatches. We used this range as it lies between the h values of two most abundant fullerenes. More flexible sorting parameters might correct current disadvantages (maybe with joint use of h values with additional structural descriptors). We wish to solve this problem in the continuation of the present work.

Nevertheless, the h-based sorting of fullerenes taking into account the Jahn–Teller effect allows distinguishing 14 members of the fullerenes family (0.67% of the 2079 fullerenes under study); most of them are synthesized or have evidence for existence. According to our approach, the C_{60} fullerene is exceptional as it is the only member of the fullerenes family with zero information entropy. We started the work using this quantity as a structural descriptor. However, it may be relevant to the processes underlying fullerene formation. For example, theory of information, applied to chemistry, interprets Shannon information as a measure of uncertainty as follows: a higher possibility of the formation of some chemical structure corresponds to the lower information in the "message" about its formation (and *vice versa*).[68] Hence, $h = 0$ in the case of C_{60} indicates high possibility of its formation as compared to other fullerene structures.

Changes in Information Entropy during Structure Transformations.

As the formalism depicted above shows efficiency in sorting fullerene structures, we plan to apply information entropy to theoretical studies on the C_{60} formation mechanism in further works. Currently, the mechanism of its formation is one of the unsolved problems of physical chemistry. Many hypotheses have been presented to explain the formation of the highly symmetric C_{60} molecules in the nonequilibrium plasma conditions (some of them reviewed in refs 34, 37, and 38). Apparently, the use of information entropies of reactions Δh_R instead of h values corresponding to isolated compounds is more suitable for mechanistic studies. In this paper, we briefly discuss how the information entropy approach may be applied to the formation mechanism of fullerenes. We should mention that the first steps of application of Shannon's theory to chemical processes were performed years ago by Karreman[18] (1955), Kobozev[22] (1966), Levine[19] (1978), and Zhdanov works[5] (1979) and proposition for information as one of the factors of self-organizing matter has been recently made.[69]

Most of the theories and computational models of the C_{60} formation deal with molecular dynamics simulations and/or calculations of energetic thermodynamic parameters. None of them pays the deserved attention to entropies, either thermodynamic or information. As it is conventional that C_{60} forms from the C_2 plasma,[34,37,38] we define the starting and final points of this process as follows:

$$30C_2 \rightarrow C_{60}(I_h) \tag{8}$$

The calculated $\Delta S_R{}^\circ$ values are -5342.17 and -5438 J K^{-1} mol^{-1}; the first value has been calculated by the PBE/3ζ method, and the second value has been estimated with the measured thermodynamic properties of gaseous C_{60} and C_2 taken from ref 50. As expected, reaction 8 is extremely antientropic from the thermodynamic point of view. Similarly considering changes in the information entropy of reaction 8, we find that that $\Delta h_R = 0$ (as $h(C_{60}) = 0$; and $h(C_2) = 0$ since C_2 consists of two identical carbon atoms). Thus, reaction 8 is isoentropic in terms of the information approach.

The C_2 loss is another important reaction among the processes under fullerene synthesis.[38,70−72] Two processes were theoretically studied:[70,71]

$$C_{60}(I_h) \rightarrow C_{58}(C_s) + C_2 \tag{9}$$

$$C_{70}(D_{5h}) \rightarrow C_{60}(I_h) + 5C_2 \tag{10}$$

The last one is thermodynamically more favorable. As previously calculated, $\Delta E_R(9) = +1080.2$ (ref 70) and $\Delta E_R(10) = +782.9$ kJ mol^{-1} (ref 71), so C_{60} has higher stability against the C_2 loss. These reactions are good examples to demonstrate the formalism for calculation of changes in information entropies Δh_R. It differs from the calculation of thermodynamic parameters of the processes, *viz.* Δh_R is not deduced from the h values of reactants and products considered separately. Indeed, we have two compounds in the right part of eq 10 (with equivalent atoms within each of them, 60 in C_{60} and 10 in five C_2 molecules). Therefore, the probabilities to find atoms of each type equal to 60/70 and 10/70 ($h_{products} = -60/70 \log_2 (60/70) - 10/70 \log_2 (10/70) = 0.592$). Application of Shannon's approach leads to $\Delta h_R(10) = -1.644$. Similarly, for reaction 9, we find $\Delta h_R(9) = +4.974$ ($\Delta h(C_{58}\ (C_s)) = 4.927$; $h_{products} = 4.974$). Thus, Δh_R values for these reactions differ by sign, and the more favorable reaction obtains negative Δh_R that is consistent with the conventional notion that the arising of information is an energetically disfavored process.[22,68,69]

The examples above demonstrate that information parameters of the processes relate to the thermodynamic ones, so in our further works we are going to apply this approach to larger

lists of fullerene reactions. We should note that our algorithm for assessing the information entropy of chemical reactions radically differs from the previous one, proposed by Karreman for calculation of change in topological information.[18] In the mentioned work, values of topological information (I) of reactants and products were calculated as sums of the contributions from the molecules, separately considered. In our approach, molecules of products (or reactants) are jointly considered. We believe that this is more appropriate for real chemical systems. Additionally, calculated *via* our algorithm, Δh values for interdependent reactions obey Hess' law as common thermodynamic reaction parameters (ΔS_R, ΔH_R, or ΔG_R).

Peculiarities and Prospectives of Consideration of Fullerenes in Terms of the Shannon Approach. Before the conclusion, we should mention questions transpired from the present study. Numerical discrimination of chemical structures is not an easy task.[14] As the readers may note, such different molecules as C_2 and C_{60} obtained the same (zero) information entropies. Chemical intuition suggests that despite this fact, these two molecules differ by complexity. We see the reason for this in a disadvantage of the original Shannon approach that estimates the structural diversity of the information carrier but ignores the "sense of the message".[3] To solve this problem (if it will impede studying of chemical processes), additional parameters may be invoked. For example, the joint use of h and N (number of atoms in the molecule) throws light on the question of what molecule, C_2 or C_{60}, is more complex.

Since three parameters are used in the present paper (h, N, and rotational symmetry number σ), we should check if they are related to each other. Previously, it was shown that information entropy is increased with symmetry loss and increase in total number of atoms in the homological series of simple organic compounds.[12] In our case, h depends on σ and N if we deal with the compounds from the same class (e.g., only fullerenes). However, if we extend the consideration to more than one class, h becomes independent of N (the C_2 *versus* C_{60} case). This justifies the joint use of h and N values. The rotational symmetry number cannot be an independent parameter in the case of fullerenes because some minimal number of atoms is needed to build the structure with a certain symmetry (see, e.g., Grünbaum–Motzkin theorem,[73] applicable to the fullerenes).

CONCLUSION

We have studied diversity of fullerene structures in terms of the Shannon approach and shown that information entropy is a good quantity for numerical description of the structural peculiarities of the fullerene isomers. Sorting 2079 fullerenes from Fowler and Manolopoulos' collection[29] (taking into account the Jahn–Teller symmetry reduction), we found that h values of only 14 isomers (less than 1% of all the studied structures) lie in the range between the two most abundant fullerenes C_{60} (I_h) and C_{70} (D_{5h}). Interestingly, C_{60} (I_h) is the only fullerene with zero information entropy that makes it an exclusive structure on the background of the whole fullerene family.

As information entropy sorts fullerene structures, we introduced the algorithm for calculation of its change at chemical transformations of fullerenes. As preliminary comparisons show, the calculated Δh values correlate with the known thermodynamic parameters of fullerene formation reactions.

We will continue the work with application of this approach to mechanistic studies on the C_{60} formation.

ASSOCIATED CONTENT

Ⓢ Supporting Information

Numerical data associated with Figures 3−6; Cartesian coordinates of the optimized fullerene structures from Table 1. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.5b00334.

AUTHOR INFORMATION

Corresponding Author
*E-mail: diozno@mail.ru.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Hartley, R. V. L. Transmission of Information. *Bell Syst. Tech. J.* **1928**, 7, 535−563.
(2) Shannon, C. E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, 27, 379−423.
(3) Nielsen, M. A.; Chuang, I. L. *Quantum Computation and Quantum Information*; Cambridge University Press: 2001; p 822.
(4) Trucco, E. A Note on the Information Content of Graphs. *Bull. Math. Biophys.* **1956**, 18, 129−135.
(5) Zhdanov, Y. A. *Information Entropy in Organic Chemistry*; Rostov University: 1979; p 56.
(6) Lin, S. K. Molecular Diversity Assessment: Logarithmic Relations of Information and Species Diversity and Logarithmic Relations of Entropy and Indistinguishability after Rejection of Gibbs Paradox of Entropy of Mixing. *Molecules* **1996**, 1, 57−67.
(7) Agrafiotis, D. K. On the Use of Information Theory for Assessing Molecular Diversity. *J. Chem. Inf. Model.* **1997**, 37, 576−580.
(8) Godden, J. W.; Stahura, F. L.; Bajorath, J. Variability of Molecular Descriptors in Compound Databases Revealed by Shannon Entropy Calculations. *J. Chem. Inf. Model.* **2000**, 40, 796−800.
(9) Godden, J. W.; Bajorath, J. Differential Shannon Entropy as a Sensitive Measure of Differences in Database Variability of Molecular Descriptors. *J. Chem. Inf. Model.* **2001**, 41, 1060−1066.
(10) Graham, D. J. Information and Organic Molecules: Structure Considerations *via* Integer Statistics. *J. Chem. Inf. Model.* **2002**, 42, 215−221.
(11) Dehmer, M.; Varmuza, K.; Borgert, S.; Emmert-Streib, F. On Entropy-based Molecular Descriptors: Statistical Analysis of Real and Synthetic Chemical Structures. *J. Chem. Inf. Model.* **2009**, 49, 1655−1663.
(12) Bonchev, D.; Kamenski, D.; Kamenska, V. Symmetry and Information Content of Chemical Structures. *Bull. Math. Biol.* **1976**, 38, 119−133.
(13) Bonchev, D.; Trinajstić, N. Information Theory, Distance Matrix, and Molecular Branching. *J. Chem. Phys.* **1977**, 67, 4517−4533.
(14) Bonchev, D.; Mekenyan, O.; Trinajstić, N. Isomer Discrimination by Topological Information Approach. *J. Comput. Chem.* **1981**, 2, 127−148.
(15) Bonchev, D. *Information Theoretic Indices for Characterization of Chemical Structures*; Research Studies Press: Chichester, UK, 1983; 249 pp.
(16) Graham, D. J.; Malarkey, C.; Schulmerich, M. V. Information Content in Organic Molecules: Quantification and Statistical Structure *via* Brownian Processing. *J. Chem. Inf. Model.* **2004**, 44, 1601−1611.
(17) Tarko, L. A New Manner to Use Application of Shannon Entropy in Similarity Computation. *J. Math. Chem.* **2011**, 49, 2330−2344.
(18) Karreman, G. Topological Information Content and Chemical Reactions. *Bull. Math. Biophys.* **1955**, 17, 279−285.
(19) Levine, R. D. Information Theory Approach TO Molecular Reaction Dynamics. *Annu. Rev. Phys. Chem.* **1978**, 29, 59−92.

(20) Nguyen, P. H. Estimating Configurational Entropy of Complex Molecules: A Novel Variable Transformation Approach. *Chem. Phys. Lett.* **2009**, *468*, 90–93.

(21) Ho, M.; Schmider, H. L.; Weaver, D. F.; Smith, V. H.; Sagar, R. P.; Esquivel, R. O. Shannon Entropy of Chemical Changes: S_{N2} Displacement Reactions. *Int. J. Quantum Chem.* **2000**, *77*, 376–382.

(22) Kobozev, N. I. A Physicochemical Modelling of Information and Thinking Processes. I. Thermodynamics of the Information Processes. *Russ. J. Phys. Chem. A* **1966**, *40*, 281–294.

(23) Haken, H. *Information and Self-Organization. A Macroscopic Approach to Complex Systems*; Springer-Verlag: 1988; p 240.

(24) Garcia-Garibay, M. A. The Entropic Enlightenment of Organic Photochemistry: Strategic Modifications. *Photochem. Photobiol. Sci.* **2010**, *9*, 1574–1588.

(25) Nosonovsky, M. Self-Organization at the Frictional Interface for Green Tribology. *Philos. Trans. R. Soc., A* **2010**, *368*, 4755–4774.

(26) Kaatz, F. H.; Bultheel, A. Informational Thermodynamic Model for Nanostructures. *J. Math. Chem.* **2014**, *52*, 1563–1575.

(27) Nemcsics, A.; Nagy, S.; Mojzes, I.; Schwedhelm, R.; Woedtke, S.; Adelung, R.; Kipp, L. Investigation of the Surface Morphology on Epitaxially Grown Fullerene Structures. *Vacuum* **2009**, *84*, 152–154.

(28) Torrens, F.; Castellano, G. Effect of Packing on the Cluster Nature of C Nanotubes: An Information Entropy Analysis. *Microelectron. J.* **2007**, *38*, 1109–1122.

(29) Fowler, P. W.; Manolopoulos, D. E. *An Atlas of Fullerenes*; Clarendon Press: Oxford, 1995; p 392.

(30) Schmalz, T. G.; Seitz, W. A.; Klein, D. J.; Hite, G. E. Elemental Carbon Cages. *J. Am. Chem. Soc.* **1988**, *110*, 1113–1127.

(31) King, R. B. Topological Aspects of Metals in Carbon Cages: Analogies with Organometallic Chemistry. *Russ. Chem. Bull.* **1998**, *47*, 833–840.

(32) Schwerdtfeger, P.; Wirz, L. N.; Avery, J. The Topology of Fullerenes. *WIREs Comput. Mol. Sci.* **2015**, *5*, 96–145.

(33) Randić, M.; Guo, X.; Plavšić, D.; Balaban. A. T. On the Complexity of Fullerenes and Nanotubes. In *Complexity in chemistry, biology, and ecology*; Bonchev, D., Rouvray, D. H., Eds.; Springer: 2005; pp 1–48.

(34) Mojica, M.; Alonso, J. A.; Méndez, F. Synthesis of Fullerenes. *J. Phys. Org. Chem.* **2013**, *26*, 526–539.

(35) Lu, X.; Chen, Z. Curved Pi-Conjugation, Aromaticity, and the Related Chemistry of Small Fullerenes ($<C_{60}$) and Single-Walled Carbon Nanotubes. *Chem. Rev.* **2005**, *105*, 3643–3696.

(36) Kovalenko, V. I.; Khamatgalimov, A. R. Regularities in the Molecular Structures of Stable Fullerenes. *Russ. Chem. Rev.* **2006**, *75*, 981–988.

(37) Ōsawa, E. Formation Mechanism of C_{60} under Nonequilibrium and Irreversible Conditions − An Annotation. *Fullerenes, Nanotubes, Carbon Nanostruct.* **2012**, *20*, 299–309.

(38) Irle, S.; Zheng, G.; Wang, Z.; Morokuma, K. The C_{60} Formation Puzzle "Solved": QM/MD Simulations Reveal the Shrinking Hot Giant Road of the Dynamic Fullerene Self-Assembly Mechanism. *J. Phys. Chem. B* **2006**, *110*, 14531–14545.

(39) Glansdorff, P.; Prigogine, I. *Thermodynamic Theory of Structure, Stability and Fluctuations*; Wiley-Interscience: 1973; p 280.

(40) Sen, K. D.; Proft, F. D.; Borgoo, A.; Geerlings, P. *N*-Derivative of Shannon Entropy of Shape Function for Atoms. *Chem. Phys. Lett.* **2005**, *410*, 70–76.

(41) Noorizadeh, S.; Shakerzadeh, E. Shannon Entropy as a New Measure of Aromaticity, Shannon Aromaticity. *Phys. Chem. Chem. Phys.* **2010**, *12*, 4742–4749.

(42) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(43) Laikov, D. N.; Ustynyuk, Yu. A. PRIRODA-04: A Quantum-Chemical Program Suite. New Possibilities in the Study of Molecular Systems with the Application of Parallel Computing. *Russ. Chem. Bull.* **2005**, *54*, 820–826.

(44) Sabirov, D. Sh.; Bulgakov, R. G.; Khursan, S. L. Indices of the Fullerenes Reactivity. *ARKIVOC* **2011**, *2011*, 200–224.

(45) Sabirov, D. Sh. Polarizability as a Landmark Property for Fullerene Chemistry and Materials Science. *RSC Adv.* **2014**, *4*, 44996–45028.

(46) Sabirov, D. Sh.; Khursan, S. L.; Bulgakov, R. G. 1,3-Dipolar Addition Reactions to Fullerenes: The Role of the Local Curvature of Carbon Surface. *Russ. Chem. Bull.* **2008**, *57*, 2520–2525.

(47) Sabirov, D. Sh.; Bulgakov, R. G. Reactivity of Fullerenes Family towards Radicals in Terms of Local Curvature. *Comput. Theor. Chem.* **2011**, *963*, 185–190.

(48) Sabirov, D. Sh. Anisotropy of Polarizability of Fullerene Higher Adducts for Assessing the Efficiency of Their Use in Organic Solar Cells. *J. Phys. Chem. C* **2013**, *117*, 9148–9153.

(49) Sabirov, D. Sh.; Garipova, R. R.; Bulgakov, R. G. Density Functional Theory Study on the Decay of Fullerenyl Radicals $RC_{60}{}^{\bullet}$, $ROC_{60}{}^{\bullet}$, and $ROOC_{60}{}^{\bullet}$ (R = *tert*-Butyl and Cumyl) and Polarizability of the Formed Fullerene Dimers. *J. Phys. Chem. A* **2013**, *117*, 13176–13183.

(50) *Standard Thermodynamic Properties of Chemical Substances*; CRC Press LLC: 2000.

(51) Chemcraft. http://www.chemcraftprog.com (accessed July 2, 2015).

(52) Zverev, V. V.; Kovalenko, V. I. An Analysis of the Structure of Fullerene C_{70} by Quantum-Chemical Methods. *Russ. J. Phys. Chem. A* **2006**, *80*, 99–105.

(53) Beckhaus, H. D.; Verevkin, S.; Rüchardt, C.; Diederich, F.; Thilgen, C.; ter Meer, H. U.; Mohn, H.; Müller, W. C_{70} is more stable than C_{60}: Experimental Determination of the Heat of Formation of C_{70}. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 996–998.

(54) Ragavachari, K.; Rohlfing, C. M. Imperfect Fullerene Structures: Isomers of C_{60}. *J. Phys. Chem.* **1992**, *96*, 2463–2466.

(55) Tan, Y.-Z.; Liao, Z.-J.; Qian, Z.-Z.; Chen, R.-T.; Wu, X.; Liang, H.; Han, X.; Zhu, F.; Zhou, S.-J.; Zheng, Z.; Lu, X.; Xie, S.-Y.; Huang, R.-B.; Zheng, L.-S. Two I_h-Symmetry-Breaking C_{60} Isomers Stabilized by Chlorination. *Nat. Mater.* **2008**, *7*, 790–794.

(56) Yoshida, M.; Ōsawa, E. Molecular Mechanics Calculations of Giant- and Hyperfullerenes with Eicosahedral Symmetry. *Fullerene Sci. Technol.* **1993**, *1*, 55–74.

(57) Zope, R. R. Electronic Structure and Static Dipole Polarizability of $C_{60}@C_{240}$. *J. Phys. B: At., Mol. Opt. Phys.* **2008**, *41*, 085101.

(58) Calaminici, P.; Carmona-Espindola, J.; Geudtner, G.; Köster, A. M. Static and Dynamic Polarizability of C_{540} Fullerene. *Int. J. Quantum Chem.* **2012**, *112*, 3252–3255.

(59) Wirz, L. N.; Tonner, R.; Hermann, A.; Sure, R.; Schwerdtfeger, P. From Small Fullerenes to the Graphene Limit: A Harmonic Force-Field Method for Fullerenes and a Comparison to Density Functional Calculations for Goldberg−Coxeter Fullerenes up to C_{980}. *J. Comput. Chem.* **2015**, n/a.

(60) Chen, Z.; Heine, T.; Jiao, H.; Hirsch, A.; Thiel, W.; Schleyer, P. v. R. Theoretical Studies on the Smallest Fullerene: From Monomer to Oligomers and Solid States. *Chem. - Eur. J.* **2004**, *10*, 963–970.

(61) Kareev, I. E.; Popov, A. A.; Kuvychko, I. V.; Shustova, N. B.; Lebedkin, S. F.; Bubnov, V. P.; Anderson, O. P.; Seppelt, K.; Strauss, S. H.; Boltalina, O. V. Synthesis and X-ray or NMR/DFT Structure Elucidation of Twenty-One New Trifluoromethyl Derivatives of Soluble Cage Isomers of C_{76}, C_{78}, C_{84}, and C_{90}. *J. Am. Chem. Soc.* **2008**, *130*, 13471–13489.

(62) Guldi, D. M.; Feng, L.; Radhakrishnan, S. G.; Nikawa, H.; Yamada, M.; Mizorogi, N.; Tsuchiya, T.; Akasaka, T.; Nagase, S.; Herranz, M. Á.; Martín, N. A Molecular $Ce_2@I_h$-C_{80} Switch − Unprecedented Oxidative Pathway in Photoinduced Charge Transfer Reactivity. *J. Am. Chem. Soc.* **2010**, *132*, 9078–9086.

(63) Piskoti, C.; Yarger, J.; Zettl, A. C_{36}, a New Carbon Solid. *Nature* **1998**, *393*, 771–774.

(64) Guo, T.; Diener, M. D.; Chai, Y.; Alford, M. J.; Haufler, R. E.; McClure, S. M.; Ohno, T.; Weaver, J. H.; Scuseria, G. E.; Smalley, R. E. Uranium Stabilization of C_{28}: A Tetravalent Fullerene. *Science* **1992**, *257*, 1661–1664.

(65) Xie, S.-Y.; Gao, F.; Lu, X.; Huang, R.-B.; Wang, C.-R.; Zhang, X.; Liu, M.-L.; Deng, S.-L.; Zheng, L.-S. Capturing the Labile Fullerene[50] as $C_{50}Cl_{10}$. *Science* **2004**, *304*, 699.

(66) Kietzmann, H.; Rochow, R.; Ganteför, G.; Eberhardt, W.; Vietze, K.; Seifert, G.; Fowler, P. W. Electronic Structure of Small Fullerenes: Evidence for the High Stability of C_{32}. *Phys. Rev. Lett.* **1998**, *81*, 5378−5381.

(67) Tagmatarchis, N.; Avent, A. G.; Prassides, K.; Dennis, T. J. S.; Shinohara, H. Separation, Isolation and Characterisation of Two Minor Isomers of the [84]Fullerene C_{84}. *Chem. Commun.* **1999**, 1023−1024.

(68) Talanov, V. M.; Ivanov, V. V. Structure as the Source of Information on the Chemical Organization of Substance. *Russ. J. Gen. Chem.* **2013**, *83*, 2225−2336.

(69) Aleskovskii, V. B. Information as a Factor of Self-organization and Organization of Matter. *Russ. J. Gen. Chem.* **2002**, *72*, 569−574.

(70) Boese, A. D.; Scuseria, G. E. C_2 Fragmentation Energy of C_{60} Revisited: Theory Disagrees with Most Experiments. *Chem. Phys. Lett.* **1998**, *294*, 233−236.

(71) Hennrich, F. H.; Eisler, H.-J.; Gilb, S.; Gerhardt, P.; Wellmann, R.; Schulz, R.; Kappes, M. M. Fullerenes: Aspects, Prospects, Limits and Speculations. *Berichte der Bunsengesellschaft für physikalische Chemie* **1997**, *101*, 1605−1612.

(72) O'Brien, S. C.; Heath, J. R.; Curl, R. F.; Smalley, R. E. Photophysics of Buckminsterfullerene and Other Carbon Cluster Ions. *J. Chem. Phys.* **1988**, *88*, 220−230.

(73) Grünbaum, B.; Motzkin, T. S. The Number of Hexagons and the Simplicity of Geodesics on Certain Polyhedra. *Can. J. Math.* **1963**, *15*, 744−751.