ARTICLES

Cut-and-Unfold Approach to Fullerene Enumeration

A. M. Livshits* and Yu. E. Lozovik

Institute of Spectroscopy, Russian Academy of Sciences, 142190, Troitsk, Moscow Region, Russia

Received November 10, 2003

A simple and effective approach for the enumeration of fullerene structures is proposed. The method combines the formalism of a fullerene graph cut-and-unfold onto a planar triangular lattice and a topological description of closed quasi-2D clusters. A tabulation of possible fullerene graphs C_n is given for the number of atoms $20 \le n \le 150$.

1. INTRODUCTION

Fullerenes, nanotubes, "onions", and other carbon nanostructures (see, e.g., refs 1,2) are currently of great interest. In particular, they are considered as a possible basis for nanoelectronic technology;[3,4] they are applied in molecular design and in the creation of new ultrastrong or superconductive materials.[5] Fullerenes are demanded in medicine,[6] etc.

Fullerenes C_n are spherical carbon clusters with (almost) sp^2 hybridization, with σ-bonds forming only hexagonal and pentagonal rings. One can represent a fullerene by a cubic graph, or by a convex polyhedron, taking atoms as vertices, and σ-bonds as edges. Taking advantage of the Euler theorem connecting the number of faces (f), the number of vertices (v), and the number of edges (e) of a convex polyhedron

$$f + v - e = 2 \tag{1}$$

one can perceive that an arbitrary fullerene always contains exactly 12 pentagonal rings (faces of the polyhedron). The number of possible fullerene structures C_n quickly grows with n. While there is only one fullerene C_{60} [7] and only one fullerene C_{70} that satisfy the isolated pentagon rule (IPR), at $n = 100$ already 450 different IPR fullerenes C_n can exist.[8] Hundreds of thousands of IPR fullerenes are topologically possible at $n \approx 150$ (see below). Enumeration of possible fullerene structures C_n is an important theoretical and applied problem. The ring-spiral algorithm,[9] top-down approach,[10] and other algorithms of numeric generation of fullerene structures[11−13] are known. Some algorithms are designed for fullerenes of high symmetry only.[14−17] Comparison of the results obtained by different approaches is mandatory, because typically algorithms cannot guarantee that all of the possible fullerene structures have been generated.[18]

Methods in refs 11 and 15−17 use the idea of "cutting and unfolding" a fullerene onto a planar hexagonal or triangular lattice. In the case of a triangular lattice, the centers of pentagonal and hexagonal faces of a fullerene are projected onto the lattice points. According to the method in ref 11, an arbitrary fullerene C_n can be cut and unfolded onto a planar triangular lattice. The resulting net diagram, which is a closed polygon with 22 corners, can be wholly described by 11 vectors of triangular lattice. The area of the net diagram $S(n)$ is equal to $\sqrt{3}/4 \cdot n$, where n is the number of atoms in the fullerene molecule. The reverse procedure allows one to construct a fullerene, starting from its net diagram. To obtain isomers of C_n, one generates all of the net diagrams with the area $S(n)$, and then, on the basis of these net diagrams, one constructs the fullerenes.

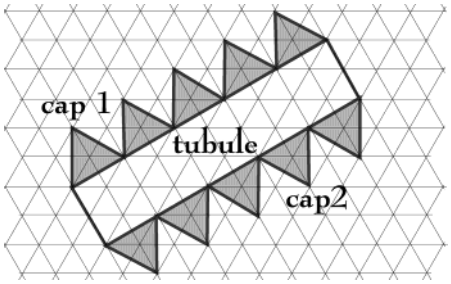

Figure 1. Net diagram of the fullerene (D_{5h}) C_{70}.

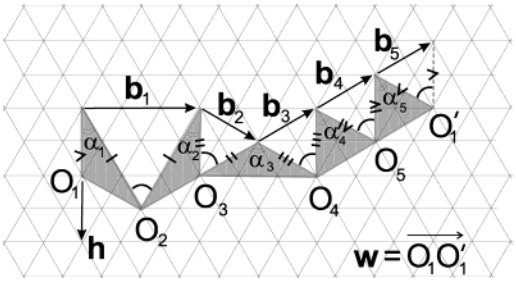

Figure 2. The structure of a cap segment.

The following problem arises. A fullerene can be cut and unfolded onto a planar lattice in many ways. Vice versa, a particular fullerene can be constructed from a number of different net diagrams. So, one needs a method to determine the constructed isomorphic and nonisomorphic fullerene structures. For example, one can generate approximate atomic coordinates[11] of the fullerene, corresponding to a given net diagram. Using the Hückel molecular orbital (HMO) theory, one can then calculate various characteristics of the structure,

* Corresponding author phone: 8(095)339-4834; e-mail: livshits@isan.troitsk.ru.

10.1021/ci0342573 CCC: $27.50
Published on Web 07/09/2004

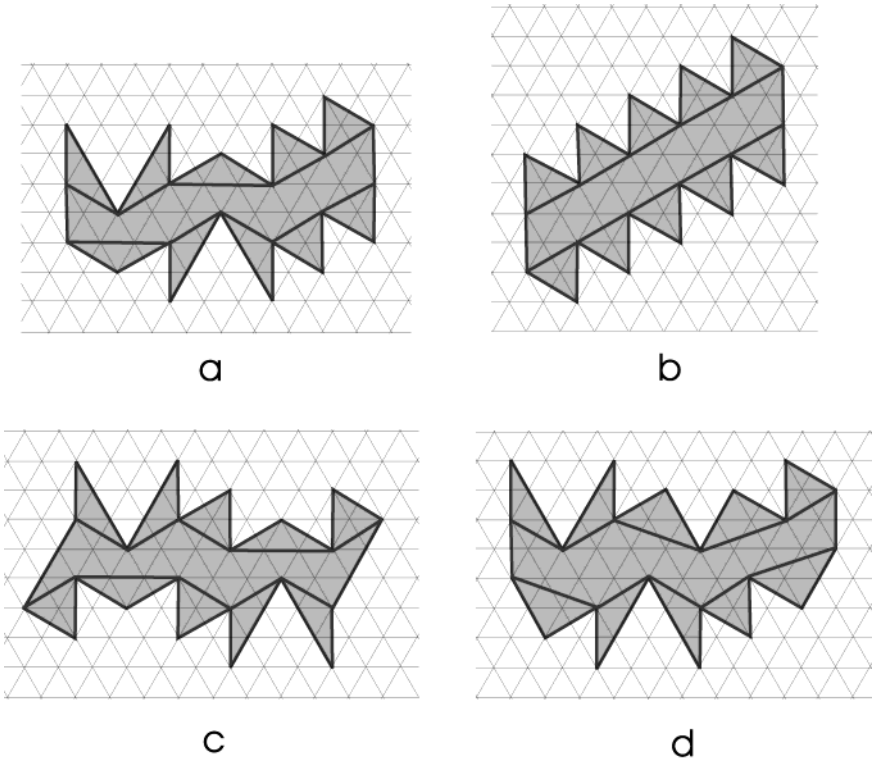

Figure 3. A fullerene can be cut and unfolded onto a flat lattice in many ways. (a−d) Examples of different net diagrams, encoding the same fullerene (I_h) C_{60}.

such as HMO eigenvalues for HOMO and LUMO, or the total HMO energy. These characteristics were used to test the isomorphism of fullerenes for $n \leq 100$.[11] As the number of isomers quickly grows with n, for larger fullerene structures one needs a more effective method of isomorphism testing.

From a methodological point of view, it is desirable to use purely topological characteristics when comparing fullerene graphs. The matrix of distances **D** of a graph unambiguously determines the structure of the graph. However, different matrices, depending on different indexing of the graph vertices, can encode the same graph. Using some special (unique) indexing of the graph vertices, we introduce $\mathbf{D}_0$, which is the matrix of distances in the unique form (see below). On the basis of the net diagram of a fullerene, we determine the matrix $\mathbf{D}_0$ of the graph G*, dual to fullerene graph G. With the use of matrix $\mathbf{D}_0$, we test the isomorphism of generated fullerene structures. We use the cut-and-unfold approach and the topological description of fullerenes to find fullerene structures C_n in the range $20 \leq n \leq 150$. Because most calculations are performed on an integer lattice, the algorithm is highly effective.

2. METHOD

There is a face-dual correspondence between planar triangular and honeycomb lattices: centers of hexagons in the honeycomb lattice are mapped into lattice points of the triangular lattice, and vice versa. One can cut and unfold a fullerene onto a planar honeycomb, as well onto a planar triangular lattice. According to the method in ref 11, fullerenes are cut and unfolded onto a planar triangular lattice. Centers of hexagonal and pentagonal faces of a fullerene are projected onto the lattice points. The resulting net diagram of a fullerene is a planar polygon with 22 corners. It consists of two cap segments and one tubular segment (Figure 1).

A cap segment that is folded forms a cone. A cap segment on a plane is a strip of vertex-sharing triangles (Figure 2).

The following conditions need to be satisfied. The angles between any two adjacent side edges of the adjacent triangles are equal to $\pi/3$; the lengths of these edges are pairwise equal. In other words, a cap segment on a plane is complementary to a sequence of five vertex-sharing equilateral triangles. The angles α_k (where $k = 1, ...5$) are not fixed, but the following equation is satisfied

$$\sum_{k=1}^{5} \alpha_k = 5\pi/3 \qquad (2)$$

A cap segment is wholly determined by the five vectors of a triangular lattice $\mathbf{b}_k$, connecting the tips of cap triangles. The tubule vector **w** and vectors $\mathbf{b}_k$ fulfill the equation $\mathbf{w} = \sum_1^5 \mathbf{b}_k$. The second cap segment in a net diagram is turned through angle $\pi/2$. A tubular segment is placed between the two caps; it is a closed polygon with 12 vertices. Height vector **h** (Figure 2) defines the "sides" of the tubular segment that must be parallel and equal so that the net diagram can be folded. It means that the tubule vectors **w** of both cap segments are also equal. The area of the net diagram is equal to $\sqrt{3}n/4$, where n is the number of carbon atoms in the fullerene.

An outline of the algorithm of fullerene structure generation follows. We make an iteration along the vectors **w** and **h**. For every value of **w**, all of the possible pairs of cap segments with the given tubule vector **w** are taken. Using these pairs of cap segments, we constructed the net diagrams.

Table 1. Number (K) of Fullerene Isomers C_n as a Function of n ($100 \le n \le 110$), Obtained via the Cut-and-Unfold Approach

n	K	n	K
100	285 914	106	497 529
102	341 658	108	604 217
104	419 013	110	713 319

Any two enantiomorphous fullerene structures that are identically cut and unfolded onto a triangular lattice will give mirror image net diagrams (vertices (i,j) of the first diagram are mapped into vertices (j,i) of the second one). We consider here enantiomorphous structures as identical. So, it is enough during the iteration to take only cap segments with vector $\mathbf{w} = (i,j)$: $i \ge j \ge 0$. To find isomers C_n, we choose net diagrams with area $\sqrt{3}n/4$. If necessary, isomers C_n, satisfying the isolated pentagon rule, can be selected.

Because the same fullerene can be constructed from many different net diagrams (see Figure 3), the algorithm above generates numerous duplicates of a given fullerene.

To eliminate duplicates, we proceed in the following way. For every net diagram, we calculate $\mathbf{D}_0$, which is the distance matrix in the unique form (see below). We check to determine if a structure, described by this matrix $\mathbf{D}_0$, was already generated. If it was, we go on to the next step; otherwise the structure and the matrix $\mathbf{D}_0$ are saved.

Let us consider the graph G*, dual to a fullerene graph G: vertices of the graph G* are mapped into faces of the graph G, and vice versa. A fullerene graph can be uniquely embedded on the surface of a sphere. As a result, there is a mutually unique mapping between the graphs G $\Leftrightarrow$ G*. One can define the distance between two vertices of a graph as the number of edges in the shortest path between these vertices. The distance matrix **D** of a graph is a matrix, containing the distances between all of the vertices of that graph. The net diagram of a fullerene C_n contains all of the information about the adjacency of N vertices ($N = n/2 + 2$) of graph G*. So, the distance matrix **D** of graph G* can be calculated. The matrix **D** determines the structure of graph G* and, hence, the structure of the dual graph G.

Different indexing of the graph vertices does not change the topology of the graph. It means that, depending on different indexing of the graph vertices, different distance matrices **D** would encode the same graph. We want to introduce here the special (unique) indexing of the graph vertices, resulting in the special (unique) distance matrix $\mathbf{D}_0$. We use the string representation of a distance matrix, introduced in the following way. All of the matrix elements are written consequently, separated with dots, row by row, from the element $d_{i,i+1}$ until the element $d_{i,N}$:

$$d_{1,2}.d_{1,3}.[..].d_{1,N}.d_{2,3}.d_{2,4}.[..].d_{2,N}.[..].d_{N-1,N} \qquad (3)$$

where i is the row index ($i = 1, ..N - 1$) and N is the rank of the matrix. As the special (unique) indexing of the graph vertices, we choose one that results in the distance matrix with the highest lexicographical value of its string representation. To find this special indexing, all of the possible $N!$ choices must be tested. During the search, the string with the current highest lexicographical value s_{cur}^{max} is kept in the memory. So, during the search, all of the strings can be discarded that start with the substring having less lexicographical value than the corresponding substring of s_{cur}^{max}. Otherwise, the value of s_{cur}^{max} is replaced by the new value. So, in fact, the efficiency of the method is about $O(n^2)$. Other methods of testing fullerene isomorphism exist; see, for example, refs 19 and 20. We use the one described above, as it uses only the distance matrix of the fullerene graph that can be easily calculated in our approach.

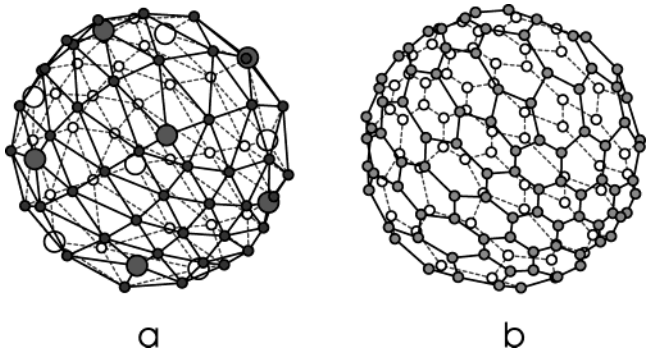

Figure 4. Dual graphs G* and G of a fullerene C_{150}, shown expanded on the surface of a sphere; vertices, placed on the reverse side of the sphere, are blank. (a) Graph G*; vertices-pentamers are marked by big circles, and vertices-hexamers are marked by small circles. (b) Graph G of the same fullerene and in the same space orientation; vertices of graph G can be projected onto faces of G*, and vice versa.

Note that this method of structure topological description can be applied to many mesoscopic systems, for example, a system of Coulomb charges on a sphere, forming on a sphere a "closed triangular lattice with topological defects".[21]

3. RESULTS

We tested the algorithm on the known fullerene structures. The results completely match the standard data for non-IPR fullerenes at $20 \le n \le 100$[8] and for IPR fullerenes at $60 \le n \le 140$.[8]

Using the algorithm suggested, we also determined all fullerene graphs C_n for $100 \le n \le 110$ and fullerene graphs that satisfy the isolated pentagon rule, for $140 < n \le 150$ (see Tables 1 and 2). Note that these data agree with the results of the top-down approach.[10] Graphs G* and G, corresponding to a possible fullerene structure C_{150}, are presented in Figure 4.[22]

The number of possible fullerene isomers C_n increases dramatically as n changes from 100 to 150. Fullerene graphs compatible with the symmetry groups T, T_d, T_h, I, and I_h were found earlier.[14] Comparing our results shows that at $n > 100$, the overwhelming majority of the structures have low symmetry. For example, at $n = 120$, as one can see from the results presented, 10 774 different fullerenes can exist that satisfy the isolated pentagon rule. Yet only one of them has tetrahedral symmetry. All of the other structures have lower types of symmetry. The symmetry noted here is the maximum possible symmetry of a fullerene, matching the symmetry of the corresponding graph. The real symmetry could be lower because of Jahn−Teller distortion.

4. SUMMARY

An algorithm of numeric generation of fullerenes is suggested. It is based on a cut-and-unfold approach to fullerene structures. The net diagram, resulting from cut-

Table 2. Number (K) of Isolated Pentagon C_n Fullerene Isomers as a Function of n ($100 < n \leq 150$), Obtained via the Cut-and-Unfold Approach

n	K	n	K
102	616	128	30 683
104	823	130	39 393
106	1233	132	49 878
108	1799	134	62 372
110	2355	136	79 362
112	3342	138	98 541
114	4468	140	121 354
116	6063	142	151 201
118	8148	144	186 611
120	10 774	146	225 245
122	13 977	148	277 930
124	18 769	150	335 569
126	23 589		

and-unfold operation on a fullerene, is a flat polygon on a triangular lattice. It can be described with 11 vectors of triangular lattice. A net diagram can be easily converted into a fullerene graph. To identify isomers, we use a special representation of the matrix of distances of fullerene graphs. All of the calculations use mostly integer arithmetic and are highly effective.

Net diagrams, encoding fullerenes C_n with isolated pentagons, were found for the number of atoms $n \leq 150$. Net diagrams, encoding non-IPR fullerenes, were found for the number of atoms $n \leq 110$. A net diagram contains all of the information about σ-bonds in the corresponding fullerene, so it can be used as a starting point for further calculations of fullerene quantum-chemical properties.

ACKNOWLEDGMENT

This work was supported by RFBR and Ministry of Science grants.

REFERENCES AND NOTES

(1) Dresselhaus, M. S.; Dresselhaus, G.; Eklund, P. C. *Science of Fullerenes and Carbon Nanotubes: Their Properties and Applications*; Academic Press: San Diego, 1996.

(2) Lozovik, Y. E.; Popov, A. M. Formation and growth of carbon nanostructures: fullerenes, nanoparticles, nanotubes and cones. *Usp. Fiz. Nauk* **1997**, *167*, 751−774.

(3) Georgakilas, V.; Pellarini, F.; Prato, M.; Guldi, D. M.; Melle-Franco, M.; Zerbetto F. Supramolecular self-assembled fullerene nanostructures. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5075−5080.

(4) Sawamura, M.; Kawai, K.; Matsuo, Y.; Kanie, K.; Kato, T.; Nakamura, E. Stacking of conical molecules with a fullerene apex into polar columns in crystals and liquid crystals. *Nature* **2002**, *419*, 702−5.

(5) Capone, M.; Fabrizio, M.; Castellani, C.; Tosatti, E. Strongly correlated superconductivity. *Science* **2002**, *296*, 2364−6.

(6) Lin, A. M.; Fang, S. F.; Lin, S. Z.; Chou, C. K.; Luh, T. Y.; Ho, L. T. Local carboxyfullerene protects cortical infarction in rat brain. *Neurosci. Res.* **2002**, *43*, 317−321.

(7) Kroto, H. W.; Heath, J. R.; O'Brien, S. C.; Curl, R. F.; Smalley, R. E. C_60: Buckminsterfullerene. *Nature* **1985**, *318*, 162−163.

(8) Fowler, P. W.; Manolopoulos, D. E. *An Atlas of Fullerenes*; Oxford University Press: New York, 1995.

(9) Manolopoulos, D. E.; Fowler, P. W. Molecular graphs, point groups, and fullerenes. *J. Chem. Phys.* **1992**, *96*, 7603−7614.

(10) Brinkmann, G.; Dress, A. W. M. A constructive enumeration of fullerenes. *J. Algorithms* **1997**, *23*, 345−358.

(11) Yoshida, M.; Osawa, E. Formalized drawing of fullerene nets. 1. Algorithm and exhaustive generation of isomeric structures. *Bull. Chem. Soc. Jpn.* **1995**, *68*, 2073−2085.

(12) Zhang, B. L.; Wang, C. Z.; Ho, K. M.; Xu, C. H.; Chan, C. T. The geometry of large fullerene cages: C_{72} to C_{102}. *J. Chem. Phys.* **1993**, *98*, 3095−3102.

(13) Liu, X.; Klein, D. J.; Schmaltz, T. G.; Seitz, W. A. Generation of carbon-cage polyhedra. *J. Comput. Chem.* **1991**, *12*, 1252−9.

(14) Fowler, P. W.; Cremona, J. E.; Steer, J. I. Systematic of bonding in non-icosahedral carbon clusters. *Theor. Chim. Acta* **1998**, *73*, 1−26.

(15) Fujita, M.; Saito, R.; Dresselhaus, G.; Dresselhaus, M. S. Formation of general fullerenes by their projection on a honeycomb lattice. *Phys. Rev. B* **1992**, *45*, 13834−6.

(16) Caspar, D. L. D. Deltahedral views of fullerene polymorphism. *Philos. Trans. R. Soc. London, Ser. A* **1993**, *343*, 133−144.

(17) Fowler, P. W.; Cremona, J. E. Fullerenes containing triples of pentagonal rings. *J. Chem. Soc., Faraday Trans.* **1997**, *93*, 2255−2262.

(18) Manolopoulos, D. E.; Fowler, P. W. A fullerene without a spiral. *Chem. Phys. Lett.* **1993**, *204*, 1−7.

(19) Balaban, A. T.; Liu, X.; Klein, D. J.; Babic, D.; Schmalz, T. G.; Seitz, W. A.; Randic, M. Graph invariants for fullerenes. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 396−404.

(20) Babic, D.; Klein, D. J.; Sah, C. H. Symmetry of fullerenes. *Chem. Phys. Lett.* **1993**, *211*, 235−241.

(21) Livshits, A. M.; Lozovik, Yu. E. Coulomb clusters on a sphere: Topological classification. *Chem. Phys. Lett.* **1999**, *314*, 577−583.

(22) Electronic files with net diagrams and distance matrices of other configurations found are available upon request (e-mail: lozovik@isan.troitsk.ru).

CI0342573