How To Enumerate the Connectional Isomers of a Toroidal Polyhex Fullerene

E. C. Kirby*

Resource Use Institute, 14 Lower Oakfield, Pitlochry, Perthshire PH16 5DS, Scotland UK

P. Pollak

The King's School, Canterbury, CT1 2ES, England UK

Received August 30, 1997[⊗]

A simple algorithm is described for enumerating all isomers of a toroidal polyhex fullerene (a boundless network of three-valent carbon atoms embedded in the surface of an imaginary torus where all the rings or faces are hexagons). An implementation for a PC is convenient for structures of at least 10 000 atoms in size. An extension of this program allows fast and automatic compilation of an adjacency matrix and derivation of a closed trigonometric expression yielding eigenvalues for all structures with fewer than 7200 atoms and, apparently, most others, although a few evaluations beyond this size are more complicated. The simplicity of a strictly polyhex toroidal structure means that, in marked contrast to many chemical enumerations, the total number of isomers increases at only a modest rate and (within the range surveyed) does not exceed approximately 30% of the number of atoms.

INTRODUCTION

Various recent events (for example¹) show that there is still interest in chemical, mathematical, and computational features of fullerene structures. One problem of the latter variety arises often in chemistry, and is no less important here. Namely, having defined a class of structures, how many connectional isomers are possible (without at this stage paying any regard to their energetic plausibility or to distinguishing isomers that are isomorphic within graph theory and differ only in such things as their chirality or knottedness)? This problem has recently been solved for "simple" fullerenes, i.e., those that, like buckminsterfullerene (C₆₀) itself, are graph-theoretically planar, and embeddable in the surface of a sphere. Earlier work gave a count of 1812 isomers for C_{60} . This and also other recent results, such as numbers of very small polyhedra,4 have been confirmed and are greatly extended with the aid of the computer program from Bielefeld.5-7 (See also the work by Fowler and Mitchell.8)

Here we elaborate on a method for enumerating isomers within the class of toroidal polyhex fullerenes⁹ and give a few results. This class of structure, consisting of a boundless network of three-valent carbon atoms embedded in the surface of an imaginary torus where all the rings or faces are hexagons, is relatively simple. Besides our own previous work, ^{10,11} where we laid out the mathematical *principles* of a simple encoding and enumeration scheme, other authors ^{12–16} too have discussed (although using slightly different approaches) the mathematics involved. In our earlier work, ^{10,11} however, we did not, except for one or two examples, examine the numbers of isomers. As far as we know no practical algorithm for obtaining an exhaustive list of all distinct connection tables of possible isomers that is simple and convenient to use has yet been made available. It

appears likely (because of mechanical strain) that these pure polyhexes may not in themselves be particularly promising candidates for real chemical construction—the introduction of mixed ring sizes may be necessary. Whatever the final status of these polyhexes, however, it is still of interest to enumerate them in order to enlarge the range of molecular energy calculations and in order to use them as precursors for the design and theoretical construction of more promising mixed ring tori, for we have not addressed directly this wider, and generally more complicated, enumeration problem.

A SUMMARY OF THE MATHEMATICAL BACKGROUND

In a previous paper with Dr. R. B. Mallion¹⁰ we introduced an encoding scheme whereby any toroidal polyhex may be described by a unique string of three integers a,b,d. Such a structure can be represented by a planar lattice of regular hexagons in which each labeled hexagon represents the samelabel hexagon on the torus, resulting in a pattern which repeats itself endlessly in two dimensions. The three integers a,b,d derive from two vectors a,c and b,d, but with nonrectangular axes chosen so that c is zero. These vectors define a parallelogram drawn on the flat lattice which represents the surface of the torus when opposite edges are "identified", i.e., considered identical. (See Figure 1.)

The array a,b,d represents a particular toroidal polyhex but, in general, so do several others, and in order to pick out a unique one we have arbitrarily chosen to minimize first d and then b. The standard code for the example in Figure 1 is 6-0-3, while in Figure 2 we pick out 15-4-1. We can distinguish a reference to the standard canonical code by writing it as TPH(a-b-d), i.e., Toroidal PolyHex (a-b-d), and equivalent forms as simply a'-b'-d'.

Generally speaking, the most common finding is a set of six codes which are all equivalent in the sense that they

[®] Abstract published in *Advance ACS Abstracts*, December 15, 1997.

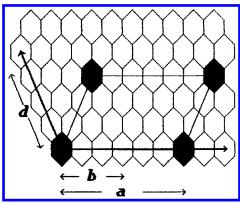


Figure 1. A toroidal polyhex, encoded as a-b-d, represented on a 2D hexagon lattice. One hexagon is arbitrarily selected as the coordinate origin, and shading shows the repeating pattern. In this particular case there are 18 hexagons on the torus, and the code a-b-d has the value 6-3-3, with equivalents 6-0-3 or 3-0-6 obtained by choosing other nonrectangular 120° axes.

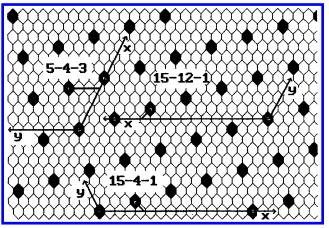


Figure 2. The canonical encoding of a toroidal polyhex. When a number of choices are available for a-b-d, first d and then b is minimized. So in this example the structure is referred to as TPH (15-4-1).

represent the same toroidal-polyhex connectional isomer. However, singlets, doublets, triplets, and, once only, a quartet do also occur (see Table 1). It should be stressed again that we are dealing with distinct connectivities here. Such a set of equivalents might or might not correspond to more than one distinct physical realization of toroidal shape.

If these numbers, representing a toroidal polyhex, are arranged within a matrix as

$$\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$$

then the matrix can be transformed to give all equivalent forms. This amounts to using a set of five rules. (Four were published earlier, 10 but a fifth one, which needed to be made explicit for computer use, is now added.)

Rule 1.

$$\begin{bmatrix} a & b \\ 0 & d \end{bmatrix} = \begin{bmatrix} a & b + g \cdot a \\ 0 & d \end{bmatrix}$$

where g is any integer, positive or negative. This rule is used so that no b outside the range $a > b \ge 0$ need be considered.

Table 1. Canonical Codes for Toroidal Polyhexes Having 6-40 Vertices and Their Equivalent Forms^a

n Code : Equivalent codes							
n Code : 6: 3-2-1 :	Equivate	ent codes					
8: 4-2-1:	(4-3-1)	(2-0-2)	(2-1-2)				
10: 5-2-1:	(5-4-1)	(5-3-1)					
12: 6-2-1:	(6-5-1)	(3-1-2)	(2 0 2)	(2-0-3)	(2-1-3)		
12: 6-3-1: 14: 7-2-1:	(6-4-1) (7-6-1)	(3-2-2) (7- 4 -1)	(3-0-2)	(2-0-3)	(2-1-3)		
14: 7-3-1:	(7-5-1)	(7 4 1)					
16: 8-2-1:	(8-7-1)	(4-1-2)					
16: 8-3-1:	(8-6-1)	(4-3-2)	(4.0.0)	(0 0 1)	(0.1.4)		
16: 8-4-1: 18: 9-2-1:	(8-5-1) (9-8-1)	(4-0-2) (9-5-1)	(4-2-2)	(2-0-4)	(2-1-4)		
18: 9-3-1:	(9-7-1)	(9-4-1)	(9-6-1)	(3-2-3)	(3-1-3)		
18: 3-0-3:	(- , -,	,	(/		,		
20: 10-02-1 :	(10-09-1)				(44 67 4)		
20: 10-03-1 :	(10-08-1)	(05-04-2)	(05-03-2)	(10-04-1)	(10-07-1)		
20: 10-05-1 : 22: 11-02-1 :	(10-06-1) (11-10-1)	(05-02-2) (11-06-1)	(05-00-2)	(02-00-5)	(02-01-5)		
22: 11-02-1 :	(11-10-1)	(11-05-1)	(11-07-1)	(11-08-1)	(11-04-1)		
24: 12-02-1 :	(12-11-1)	(06-01-2)			· ·		
24: 12-03-1 :	(12-10-1)			(04-02-3)	(04-01-3)		
24: 12-04-1 : 24: 12-05-1 :	(12-09-1) (12-08-1)	(04-03-3)	(04-00-3)	(03-00-4)	(03-01-4)		
24: 12-05-1 : 24: 12-06-1 :	(12-06-1)	(06-00-2)	(06-02-2)	(02-00-6)	(02-01-6)		
24: 06-04-2 :	(12 0, 1)	(00 00 2)	(20 01 0)	(32 00 0)	(02 02 0)		
26: 13-02-1 :	(13-12-1)						
26: 13-03-1 :	(13-11-1)	(13-06-1)	(13-08-1)	(13-05-1)	(13-09-1)		
26: 13-04-1 : 28: 14-02-1 :	(13-10-1) (14-13-1)	(07-01-2)					
28: 14-03-1 :	(14-12-1)	(07-06-2)	(07-03-2)	(14-10-1)	(14-05-1)		
28: 14-04-1 ;	(14-11-1)	(14-09-1)	(14-06-1)	(07-05-2)	(07-04-2)		
28: 14-07-1 :	(14-08-1)	(07-02-2)	(07-00-2)	(02-00-7)	(02-01-7)		
30: 15-02-1 : 30: 15-03-1 :	(15-14-1) (15-13-1)	(15-08-1) (15-07-1)	(15-09-1)	(05-02-3)	(05-01-3)		
30: 15-04-1 :	(15-12-1)	(05-04-3)	(13-03-1)	(00-02-0)	(05-01-0)		
30: 15-05-1 :	(15-11-1)	(03-01-5)					
30: 15-06-1 :	(15-10-1)	(03-02-5)	(03-00-5)	(05-00-3)	(05-03-3)		
32: 16-02-1 : 32: 16-03-1 :	(16-15-1) (16-14-1)	(08-01-2) (08-07-2)	(08-03-2)	(16.06.1)	(16-11-1)		
32: 16-03-1 :	(16-14-1)	(16-05-1)		(04-03-4)			
32: 16-07-1 :	(16-10-1)	(08-05-2)	(/	(-, ,,	(11 12 17		
32: 16-08-1 :	(16-09-1)	(08-00-2)	(08-02-2)	(02-00-8)	(02-01-8)		
32: 08-04-2 :	(08-06-2)	(04-02-4)					
32: 04-00-4 : 34: 17-02-1 :	(17-16-1)	(17-09-1)					
34: 17-03-1 :	(17-15-1)	(17-08-1)		(17-12-1)			
34: 17-04-1 :	(17-14-1)	(17-11-1)	(17-07-1)	(17-05-1)	(17-13-1)		
36: 18-02-1 :	(18-17-1)	(09-01-2)	(00 02 2)	(06-02-3)	(06 01 2)		
36: 18-03-1 : 36: 18-04-1 :	(18-16-1) (18-15-1)	(09-08-2) (06-05-3)	(09-03-2) (06-04-3)				
36: 18-05-1 :	(18-14-1)	(09-04-2)	(09-07-2)	(18-08-1)	(18-11-1)		
36: 18-06-1 :	(18-13-1)	(18-07-1)	(18-12-1)	(03-02-6)	(03-01-6)		
36: 18-09-1 :	(18-10-1)	(09-02-2)	(09-00-2)	(02-00-9)	(02-01-9)		
36: 06-00-3 : 38: 19-02-1 :	(06-03-3) (19-18-1)	(03-00-6) (19-10-1)					
38: 19-03-1	(19-10-1)	(19-10-1) (19-09-1)	(19-11-1)	(19-07-1)	(19-13-1)		
38: 19-04-1	(19-16-1)	(19-06-1)	(19-14-1)	(19-15-1)	(19-05-1)		
38: 19-08-1 :	(19-12-1)	/10 C. C.					
40: 20-02-1 : 40: 20-03-1 :	(20-19-1)	(10-01-2) (10-09-2)	(10 02 2)	(20-14-1)	(20.07.1)		
40: 20-03-1 : 40: 20-04-1 :	(20-18-1) (20-17-1)	(20-13-1)	(20-03-2)				
40: 20-05-1	(20-16-1)	(05-04-4)	(05-00-4)	(04-00-5)	(04-01-5)		
40: 20-06-1	(20-15-1)	(04-03-5)					
40: 20-09-1 : 40: 20-10-1 :	(20-12-1) (20-11-1)	(05-02-4) (10-00-2)	(10 02 2)	(02-00-1)	(02 01 1)		
40: 20-10-1 : 40: 10-04-2 :	(10-08-2)		(10-02-2)	(02-00-1)	(UZ-U1-1)		
10. 10 VT L.	(14 00 C)	(10 00 2)					

^a Note that these small objects, and especially the code series xx-2-1 where xx is even, are included for convenient illustration and formal interest only. Mechanical strain makes them highly implausible as realizable toroidal structures, and in addition the latter are topologically planar (but could in principle be fitted round a torus).

Rule 2.

Rule 3.
$$\begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \equiv \begin{bmatrix} a & d-b \\ 0 & d \end{bmatrix}$$
$$\begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \equiv \begin{bmatrix} a \cdot d/h & x \cdot d \\ 0 & h \end{bmatrix}$$

where h is the highest common factor of a,b and x is the multiplicative inverse of $b/h \mod a/h$

Rule 4.

$$\begin{bmatrix} 2 & 0 \\ 0 & d \end{bmatrix} \equiv \begin{bmatrix} 2 & 1 \\ 0 & d \end{bmatrix}$$

This rule (4) is deducible from the others when d is odd but not when d is even.

Rule 5.

$$\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} = \begin{bmatrix} d & 0 \\ 0 & a \end{bmatrix}$$

AN OUTLINE OF THE COMPUTER ALGORITHM

Start with the number of hexagons (which, of course, is equal to half the number of vertices) and store all the pairs of factors a,d ($a \ge d$) of this number.

For each a,d take every b within the range 0 to

$$INT \left[\frac{(a+d)}{2} \right]$$

and apply a filter to the b values to exclude a b value when (i) (b is 0 or 1) and (d = 1); such cases are simple cycles. (ii) The HCF of b with a is less than d; to avoid repetition.

The main routine is described below, but it is convenient at this stage to pick out some special cases where short cuts are available.

- (i) If b/d = 2 and a/d = 3 then the code is a singlet, TPH(a-b-d).
- (ii) If a = 2 then

if d = 1 then there is a triplet TPH(2-0-1), (2-1-1), (1-0-2)

if d = 2 then there is a quartet of equivalent codes; TPH(4,2,1) with (4-3-1), (2-0-2) and (2-1-2) if d > 2 a sextet;

TPH(2d-d-1), (2d-d+1-1), (d-0-2), (d-2-2), (2-0-d), (2-1-d).

(iii) If $a \neq 2$ and b = 0 then

if d = 2 swap a, d and revert to the a = 2 case.

If a = d the result is the singlet TPH(a-0-a), otherwise a triplet Max, 0, Min; Max, Min, Min; and Min, 0, Max where Max and Min refer, respectively to the greater or lesser of a, d.

if $a \neq d$ but either a MOD d = 0 or d MOD a = 0 then

If the lesser of a,d is 2 then swap a,d if $a \ge d$, when the result is the sextet

TPH(2*d*-*d*-1), (2*d*-*d*+1-1), (*d*-0-2), (*d*-2-2), (2-0-*d*), (2-1-*d*).

Finally, an *a,b,d* surviving thus far is passed through the main routine as follows (applying alternately Rule 2 and Rule 3). This yields a sextet or a triple of identical pairs:

Take a copy of a,b,d as L_a,L_b,L_d

Set Array₁ = a,b,d

Evaluate the expression b = (a+d-b) MODa

if b = 0 then revert to test (ii) above.

If $b < L_b$ and $d = L_d$ then $L_b = b$; set the next Array to a.b.d:

find h the HCF of a and b; set p = a/h and q = b/h; find g, such that $g \cdot q$ MOD p = 1, and set $a = d \cdot p$; $b = d \cdot g$ MOD a; d = h, and

if a = 2 then revert to test (ii) above.

If $d < L_d$ then $L_a = a$; $L_b = b$; $L_d = d$;

If $d = L_d$ and $b < L_b$ then $L_b = b$.

Set the next array to a,b,d and repeat for a total of three passes.

For every canonical TPH code selected by this procedure, a simple lexicographic validation check is made before acceptance; one of the statements

[
$$(a = \text{previous } a) \text{ AND } (b > \text{previous } b)$$
]

or

[
$$(a < previous a) AND (b < previous b)$$
]

must be true. Once accepted, the canonical TPH code and its equivalent formulations, if any, can be written to disk file. (A number of options are available as to which information is collected by the program in its current version.) On completion of the isomer enumeration for a given number of vertices n, a further validation check is performed. The number of equivalent codes found must equal (k-3), where k is the sum of factors of n/2, including n/2 and 1.

The translation of a TPH code into an associated connection table, or the calculation of its Hückel spectrum, can usually be done by applying a simple formula and can always be done if there are fewer than 7200 vertices. Although Haigh is developing an analytic treatment for the general case, 15 we find it convenient for practical purposes to classify isomers according to their case number. The logic and the mathematics used is described in our earlier paper, 10 but, in brief, case I arises if and only if b in at least one of the equivalent codes is divisible by the highest common factor of a and d. This is always the case when there are less than 1800 vertices. Case II is a useful supplement or alternative; it can be used if and only if no integer greater than one divides all three of a,b,d. These two cases, which both have simple formulas for their application, cover all the polyhexes up to 7200 vertices. Case III covers all examples which fail attempts to bring them under cases I or II and at present requires more attention to detail. We have not thus far formulated a closed expression for it. Most examples, however, even for large systems, can be dealt with under case I or, failing that, case II. Thus, at the transition numbers, for example, we have for 1800 vertices, 474 case I and 8 case II = 482 isomers, and for 7200 vertices, 2048 case I, 56 case II, and 8 case III = 2112 isomers. The first case II encountered is 450-5-2 and the first case III 900-10-4.

RESULTS: TOTAL NUMBER OF ISOMERS

Table 2 shows isomer counts for a selection of sizes, and in Figure 3 a narrower size range is taken and the isomer count plotted against the number of vertices. For clarity, just a scatter-plot is shown, but if the numbers are examined more closely, periodicity and oscillation within a diverging envelope is apparent. On average the count shows only a modest and near-linear rate of increase with size. This is in marked contrast to many isomer enumerations, which often exhibit an all-too-familiar and seemingly explosive growth of numbers. The objects studied here, which have hexagonal faces only, are comparatively simple. The number of isomers may be expected to reflect the number of distinct parallelo-

Table 2. Total Number of Possible Toroidal Polyhex Isomers for a Selection of Sizes

no. of vertices	no. of isomers	no. of vertices	no. of isomers
60	13	3000	738
100	16	4000	818
120	32	5000	917
180	41	6000	1578
240	68	7000	1466
360	98	8000	1654
1000	186	9000	2382
2000	398	10000	1963

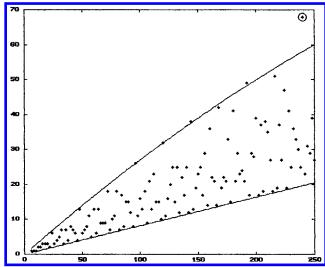


Figure 3. Numbers of isomers for toroidal polyhexes of 6-250 vertices, showing bound lines fitted to enclose most points. Vertical axis, total number of isomers; horizontal axis, number of vertices; upper bound, $y = 0.29x - 2x^2/10^4$; and lower bound, y = 0.082x.

grams of a given size that may be cut from a hexagon lattice, and the number of distinct ways in which these flat polyhexes may be "rolled up" into a torus. The number of such parallelograms depends upon how many pairs of factors the number of hexagons (= half the number of vertices) has and is thus a consequence of the properties of numbers. At first sight it might appear that other shapes besides parallelograms should also be taken into account. In fact this is unnecessary, because once any shape cut from a network is fashioned into a torus (assuming that opposite sides match), it can be "cut and skinned" to produce one of the standard shapes described here, and so will not generate fresh isomers.

A linear regression line for the raw plot of Figure 3 indicates a modest correlation between size and the number of vertices (y = 0.1406x + 0.7075; r = 0.801; s = 7.449), but when the number of factors is taken into account, the correlation, while still not especially high, is considerably improved (y = 0.0191x + 5.3829; r = 0.9743; s = 2.8065).

APPENDIX: PRACTICAL DETAILS

The main program described here is implemented as a pair of DOS EXE files. In the current version, both allow input of a size or a range of sizes. The first then lists and counts standard codes and files them with their equivalent forms. The other gives additional options such as determining case number and when this is I or II evaluating the Hückel spectrum. Conversion of a code to a connection table is delegated to an additional program. It is possible to evaluate selected eigenvectors in closed trigonometric form from a

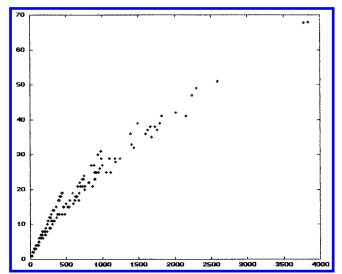


Figure 4. Numbers of isomers related to both the size and the number of size factors, for toroidal polyhexes of 6-250 vertices. Vertical axis, total number of isomers; and horizontal axis, number of vertices × factorability of that number.

code value without necessarily computing them all, but this facility has not so far been included. A recent program written by Mr. Plestenjak²⁰ is slower but more powerful (and almost as compact in its requirement) in that it can take any three-valent connection table (not just polyhex ones) modified to "VGR" format and generate preselected eigenvectors.

As a rough guide to speed, during trials using a 486 CPU running at 100 MHz, it took a little under 1 min to list all of the toroidal polyhexes (1963) having 10 000 vertices (without pursuing case numbers).

REFERENCES AND NOTES

- (1) For example: The award of a Nobel Prize in 1997 to three pioneers of fullerene science; ongoing meetings of the European Human Capital and Mobility Fullerene Network (such as was organized by P. W. Fowler, Exeter, April 1996); the dedication of a recent issue of MATCH (Vol. 33) to "Mathematical aspects of the fullerenes".
- (2) Liu, X.; Schmaltz, T. G.; Klein, D. J. Reply to comment on "Favourable structures for higher fullerenes". Chem. Phys. Lett. 1992, 192,
- (3) Manolopoulos, D. E. Comment on "Favourable structures for higher fullerenes". Chem. Phys. Lett. 1992, 192, 330.
- (4) Jiang, Y. S.; Shao, Y. H.; Kirby, E. C. Topology and stability of trivalent polyhedral clusters. Fullerene Sci. Technol. 1994, 2, 481-
- (5) Brinkmann, G.; McKay, B. plantri.c V1.0. Internet web page: http: //cs.anu.edu.au/~/bdm. (Further details for this, and of refs 5 and 6, can be obtained from Gunnar Brinkmann and Andreas W. M. Dress at Universität Bielefeld, D-33501 Bielefeld, Germany, or Brendan D. McKay, Australian National University, Canberra, ACT 0200 Australia.)
- Brinkmann, G. Generating cubic graphs faster than isomorphism checking. University of Bielfeld Preprint 92-047, 1992. (See also ref
- (7) Brinkmann, G.; Dress, A. W. M. A constructive enumeration of fullerenes, Authors' preprint 1996. (See also ref 5.)
- (8) Fowler, P. W.; Mitchell, D. A sum rule for symmetries and isomer counts of trivalent polyhedra. J. Chem. Soc., Faraday Trans. 1996, 92, 4145-4150.
- (9) Note: Here we adopt a relaxed attitude toward the definition of a fullerene by deliberately excluding any stipulation about ring sizes.
- (10) Kirby, E. C.; Mallion, R. B.; Pollak, P. Toroidal polyhexes. J. Chem. Soc., Faraday Trans. 1993, 89, 1945-1953.
- (11) Kirby, E. C. Cylindrical and toroidal polyhex structures. Croat. Chem. Acta **1993**, 66, 13-26.
- (12) Thomassen, C. Tilings of the Torus and the Klein Bottle and Vertextransitive Graphs on a Fixed Surface. Trans. Am. Math. Soc. 1991, 323, 605-635.

- (13) Klein, D. J. Elemental Benzenoids. J. Chem. Inf. Comput. Sci. 1994, 34, 453–459.
- (14) Klein, D. J.; Liu, X. Elemental carbon isomerism. *Int. J. Quant. Chem.* **1994**, Symposium Suppl. 28, 501–523.
- (15) Haigh, C. W. (of University College of Swansea, University of Wales, Singleton Park, Swansea SA2 8PP, UK.) Personal communication dated 2nd March, 1995.
- (16) Hosoya, H.; Okuma, Y.; Tsukano, Y.; Nakada, K. Multilayered cyclic fence graphs: Novel cubic graphs related to the graphite network. *J. Chem. Inf. Comput. Sci.* 1995, 35, 351–356.
- (17) Note: Carbon tori have recently been discovered among laser-grown Single-Wall Nanotube material (ref 18), but it is uncertain whether
- they are purely hexagonal. Such tori seem to have relatively thin tubes but are probably large enough overall for an all-hexagon structure not to feel unduly strained.
- (18) Liu, J.; Dai, H., Hafner, J. H.; Colbert, D. T.; Smalley, R. E.; Tans, S. J.; Dekker, C. Fullerene "crop circles". *Nature* **1997**, *385*, 780–781.
- (19) Fowler, P. W. et al (of the University of Exeter, UK), 1996; papers in preparation,
- (20) Plestenjak, B.; Pisanski, T. (of the IMFM, Jadranska 19, SI-1000 Ljubljana, Slovenia) personal communication, 1997.

CI970072I