Dualists of 2-dimensional and 3-dimensional classes of molecules: benzenoids and diamondoids

Alexandru T. Balaban

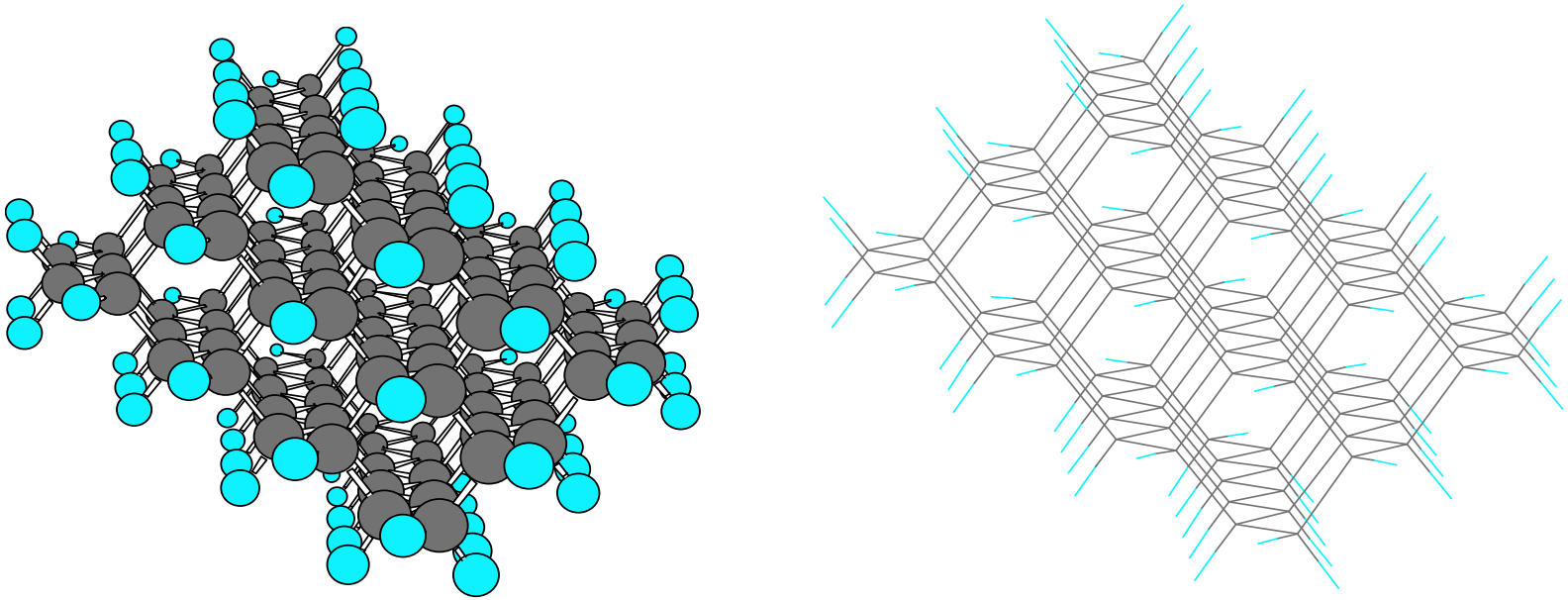
Texas A&M University at Galveston
Galveston, Texas 77551, USA

Portoroz, September 2012

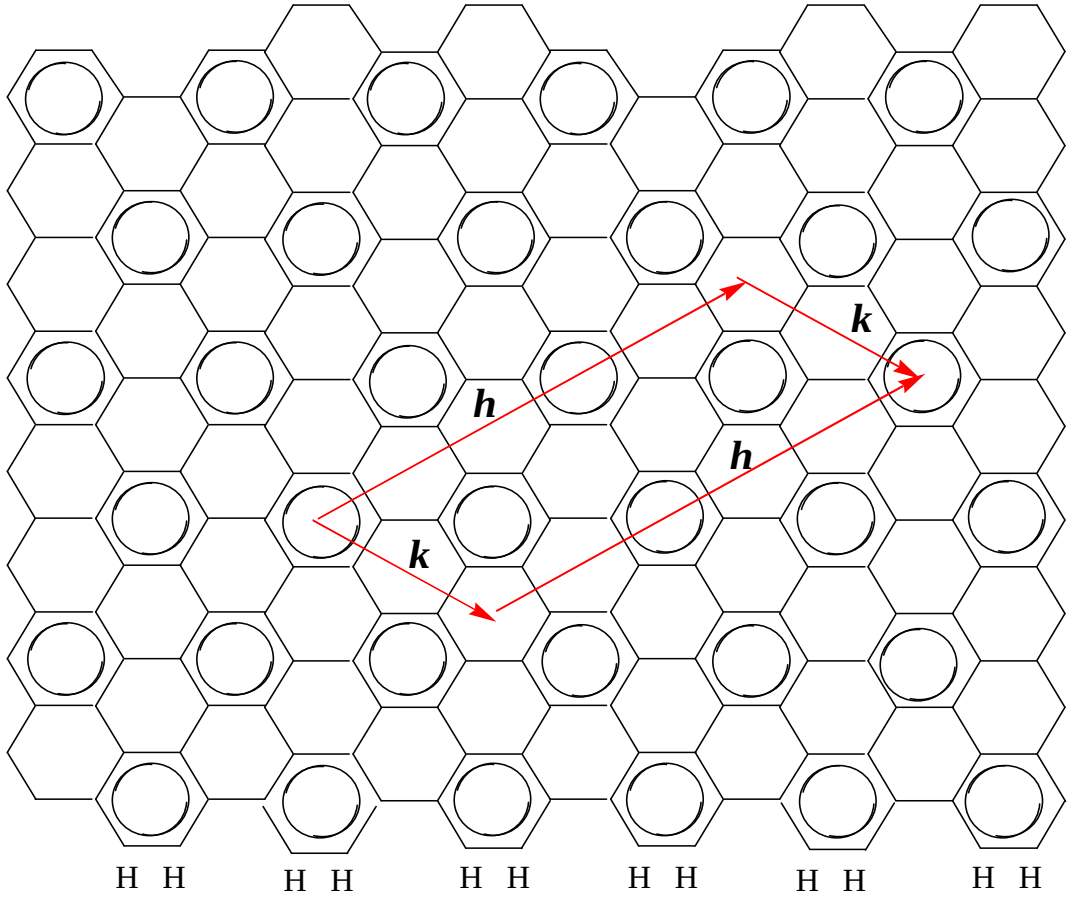
Contrasting properties of diamond and multi-layer graphite

Property	Diamond	Multilayer graphite
Aspect	transparent, highly refractive	opaque, black
Crystalline system	cubic	hexagonal
Density ($g \times cm^{-3}$)	3.514	2.266
Hardness (Mohs scale)	10	between 1 and 2
Electrical conductivity	low	high
Thermal conductivity	high	low
Thermodynamic stability	unstable	stable

Two views of a small portion of diamond with hydrogen atoms connected to dangling bonds



A portion of graphene with Clar circles symbolizing π-electron sextets



Newer carbon allotropes (fullerenes) were discovered by R. F. Curl, H. W. Kroto, and R. E. Smalley:1985 (Nobel Prize in chemistry, 1996.

Shortly afterwards nanotubes, nanocones, and nanotori followed. Single-walled nanotubes (SWNTs) may have many applications, but so far they cannot be obtained as substances with billions of identical molecules, but only (like most Natural or synthetic polymers) as mixtures of similar molecules with wider or narrower dispersion curves.

It was found that electrical and other properties of SWNTs depend on how the graphene sheets are folded: if the Coxeter vectors h and k satisfy the condition

$$h - k \equiv 0 \pmod 3$$

or when h = k, the “armchair” nanotube folds congruently such that the Clar circles overlap, the band gap is zero, and the conductivity is high (metalllic). When k = 0, one has a “zigzag” nanotube, and when h > k >0, one obtains a chiral nanotube, often described as “helical” with non-zero band gap.

All armchair SWNTs and ~1/3 of zigzag and chiral SWNTs are readily conductive, whereas ~2/3 of zigzag SWNTs are semiconductive.

“Honorary” hydrocarbons and chemical elements

The two traditional carbon allotropes (diamond and graphite) are macromolecules with enormous numbers of carbon atoms and a negligible number of hydrogen atoms attached to the peripheral dangling bonds.

One can consider ***graphene*** as an “honorary” polycyclic aromatic hydrocarbon or perifusenic benzenoid, and ***diamond*** as an “honorary” polycyclic saturated hydrocarbon formed from fused adamantane units.

Similarly, atoms of chemical elements have nuclei containing N neutrons and Z protons surrounded by Z electrons. With linearly increasing atomic numbers Z, the number N of neutrons must increase faster for the strong nuclear force to ensure cohesion via π-boson exchange. Therefore ***neutron stars*** can be considered to be “honorary” chemical elements with huge numbers of neutrons and negligible numbers of protons and electrons.

Benzenoids

There are two definitions for **benzenoids (**polycyclic aromatic compounds):

1.Portions of graphene (honeycomb lattice) with hydrogens attached to dangling bonds (planar molecules); and

2.Hydrocarbons consisting of benzenoid rings sharing carbon-carbon bonds (including helicenes)

Hexagonal and trigonal lattices are mutually dual.

Dualists (internal duals) of benzenoids

Dualist vertices are the centers of hexagons, and **edges** connect vertices corresponding to hexagons sharing carbon-carbon bonds.

Balaban, A. T.; Harary, F. *Tetrahedron,* 1968, **24**, 2505-2516: Chemical graphs. V. Enumeration and proposed nomenclature of benzenoid *cata*-condensed polycyclic aromatic hydrocarbons.

Balaban, A. T. *Tetrahedron,* 1969, **25**, 2949-2956: Chemical graphs. VII. Proposed nomenclature of branched *cata*-condensed benzenoid polycyclic hydrocarbons.

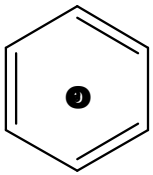
Classification of **benzenoids**

Acyclic dualists: cata-condensed (catafusenes)

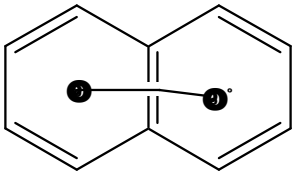
Dualists with triangles: peri-condensed (perifusenes)

Dualists with larger rings: coronoids (coronafusenes)

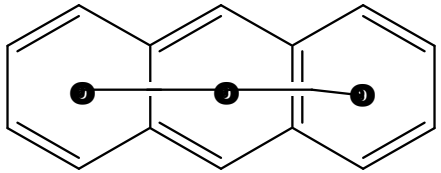
Catafusenes



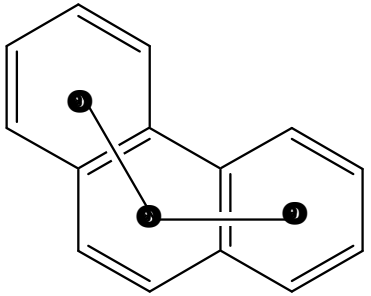
benzene



naphthalene

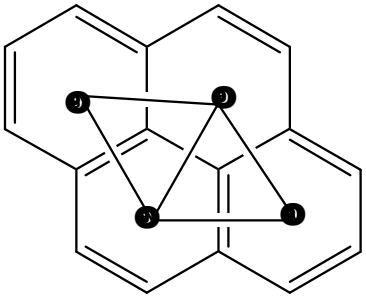


anthracene

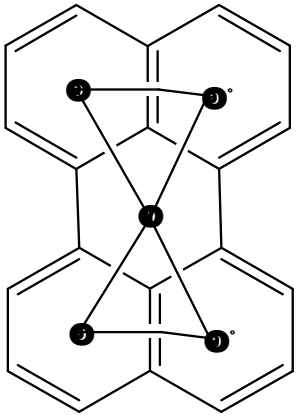


phenanthrene

Perifusenes

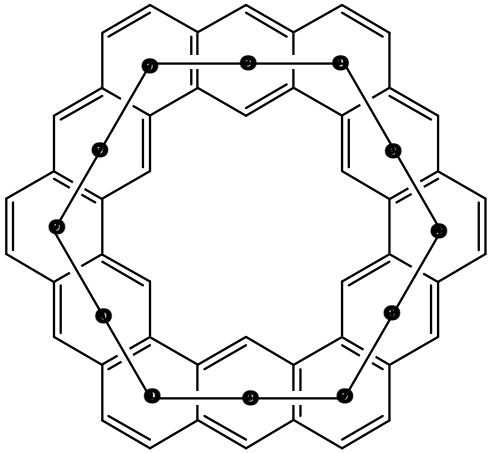


pyrene



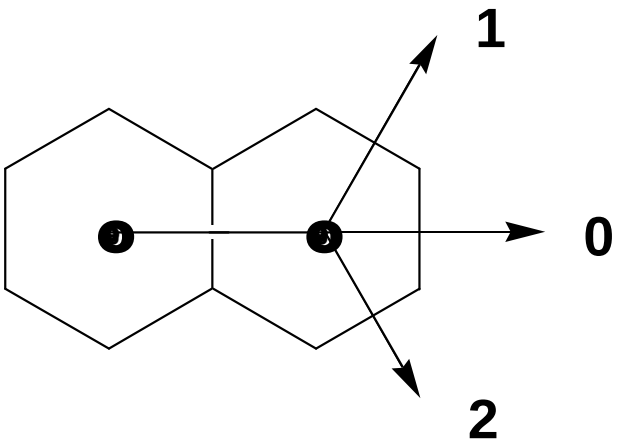
perylene

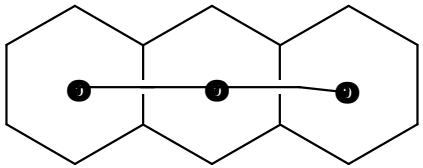
Coronafusene



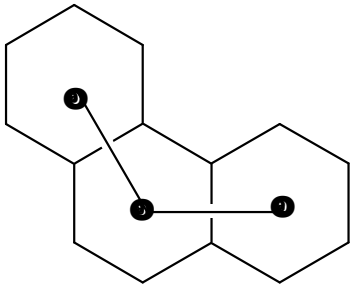
kekulene

Coding of catafusenes

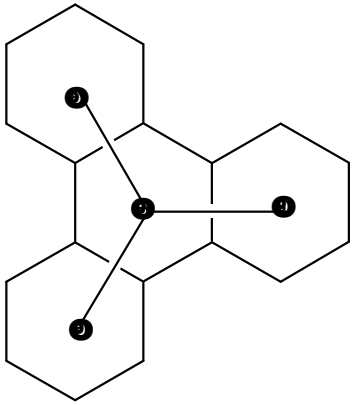




anthracene: 0



phenanthrene: 1



triphenylene: 1(.)

Coding of catafusenes (continued)

1. Select the longest nonbranched sequence of benzenoid rings

2. Assign for each triplet of rings a digit (0, 1,or 2) for the median ring, keeping the 1/2 selection for left/right direction

3. Branches are denoted by digits in brackets; if a branch consists of only one ring, one denotes it by a dot in brackets.

4. Select the starting point so as to minimize the number formed by reading sequentially the digits and ignoring brackets

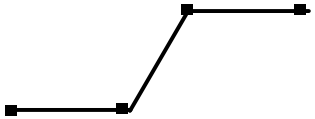
Dualists and codes for all isomeric tetra-catafusenes $C_{18}H_{12}$ with R = 4 benzenoid rings, **5**–**10**. The code has R–2 digits; a pair of brackets is equivalent to one digit.

[00]

5

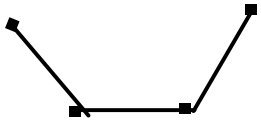
[01]

6



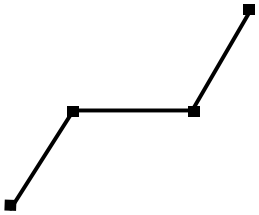
[12]

7



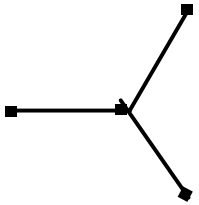
[11]

8



[12]

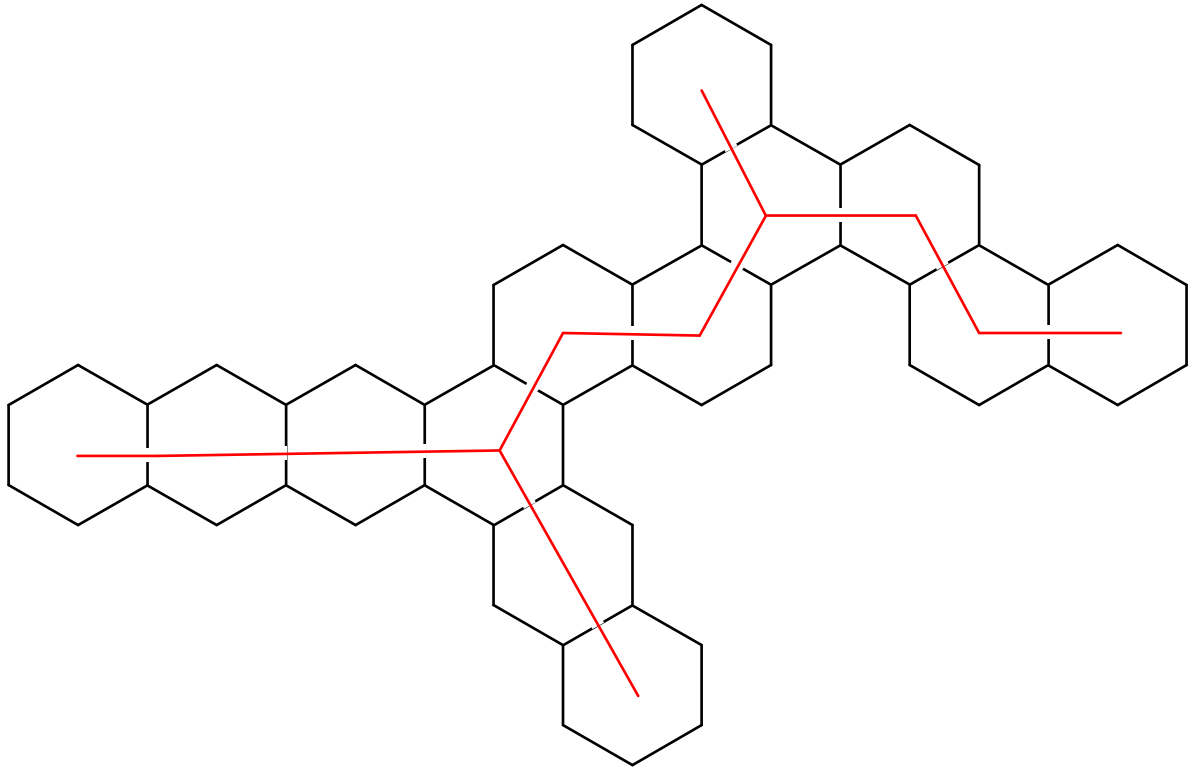
9



[1(.)]

10

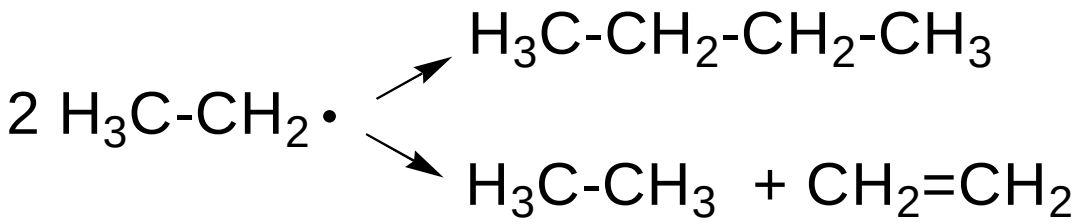
Example of catafusene code



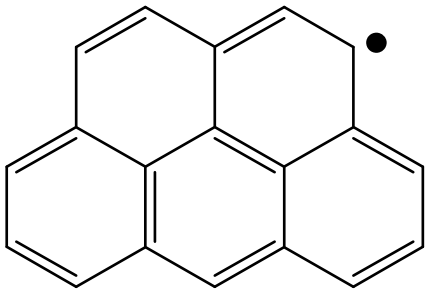
001(0)212(.)21

Special cases of benzenoids

The number of carbon atoms must be even, otherwise the benzenoid is a free radical which is not stable in condensed state because it dimerizes or disproportionates.



Olympiadane is a free radical



Enumerations of benzenoids

A. T. Balaban, J. Brunvoll, J. Cioslowski, B. N. Cyvin, S. J. Cyvin, I. Gutman, He Wenchen, He Wenjie, J. V. Knop, M. Kovacevic, W. R. Müller, K. Szymanski, R. Tosić, and N. Trinajstić, Enumeration of benzenoid and coronoid hydrocarbons. *Zeitschrift für Naturforschung*, **1987**, *42c*, 863-870.

G. Brinkmann, G. Caporossi, and P. Hansen, A survey and new results on computer enumeration of polyhex and fusene hydrocarbons, *Journal of Chemical Information and Computer Sciences*, **2003**, *43,* 842–851.

G. Brinkmann, C. Grothaus, and I. Gutman, Fusenes and benzenoids With perfect matchings, J. Math. Chem. 2007, 42, 909-924.

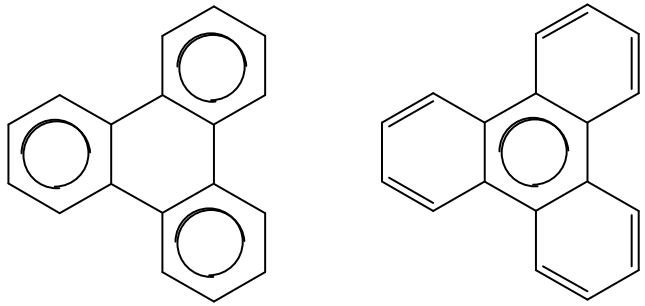
I. Jensen, A parallel algorithm for the enumeration of benzenoid hydrocarbons. *J. Stat. Mech.* (2009) P02065.

I. Gutman, How many benzenoid hydrocarbons? *Bull. Chem.Technol. Macedonia*, 21 (2002) 53-56.

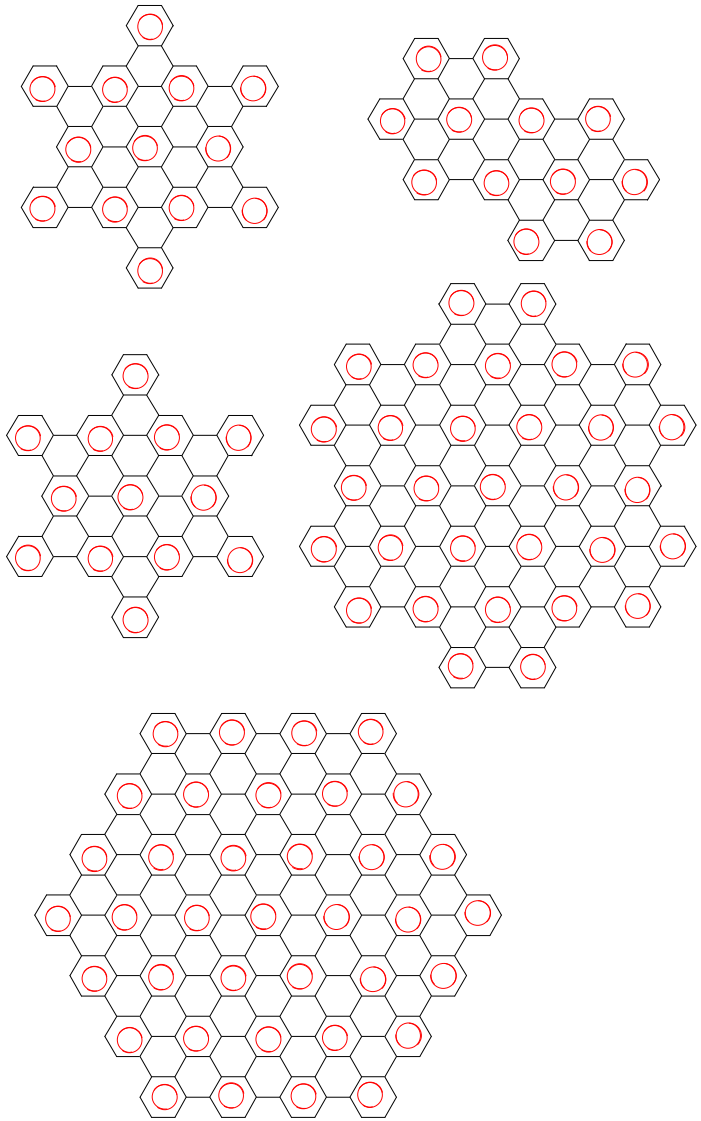
Clar's theory of aromatic sextets

Clar, E. *The Aromatic Sextet*. Wiley, New York, 1972.
Due to competition for the six pi-electrons of a benzenoid ring, in an acene only one ring is to be inscribed with a circle symbolizing the pi-electron sextet.
Migrating sextets are indicated by an arrow for the range.

Triphenylene is a"claromatic" benzenoid, having only Clar sextets and "empty"rings



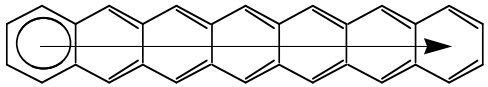
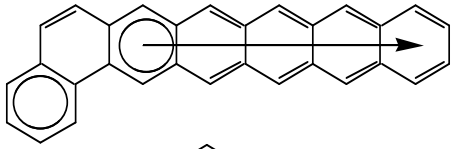
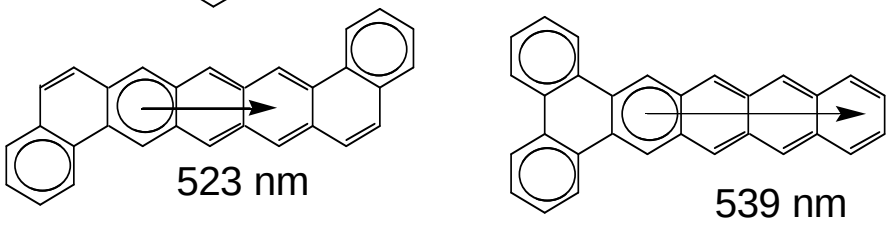
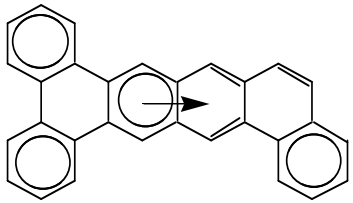
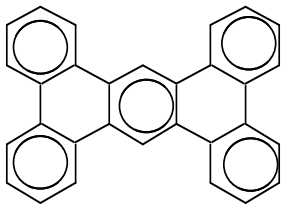
Some claromatic benzenoids synthesized by Klaus Muellen et al.



Isomers are different substances with the same molecular formula: constitutional isomers differ in the way atoms are connected by covalent bonds; stereo-isomers have the same connectivity but differ in the 2-D geometry, or for enantiomers in the 3D-configuration (mirror-images are not superimposable).

Valence isomers are a special class of isomers differing by the redistribution of covalent bonds connecting building blocks, e. g. there are 5 valence isomers of benzene $(CH)_6$ with hydrogen-depleted formulas corresponding to the 5 cubic planar cubic multigraphs with 6 vertices.

Seeing Clar sextets by color changes of isomeric heptacatafusenes

Acenic chain	Structure	λ_{max}	Clar sextets	Color
7		840 nm	1	deep green **deep green**
6		651 nm	2	blue-green **blue-green**
5	 523 nm 539 nm		3	red **red**
4		425 nm	4	yellow **yellow**
3		328 nm	5	colorless

Several physical-chemical properties of benzenoids
(in particular catafusenes) are correlated biparametrically in terms of
(1) the number R of benzenoid rings, and
(2) the number of linearly-fused rings in the longest acenic
portion (or of vertices in the straight portion of the dualist):

Redox potentials
Energy of conjugation
Reaction rate for cycloadditions
Energy of longest wavelength in the UV-Vis absorption spectrum

Formula periodic table for benzenoids (J. R. Dias)
Catafusenes are in the first row

$N_{lc} \setminus d_s$	-2	-1	0	1	2	3
0			$C_{10}H_8$	$C_{14}H_{10}$	$C_{18}H_{12}$	$C_{22}H_{14}$
2			$C_{16}H_{10}$	$C_{20}H_{12}$	$C_{24}H_{14}$	$C_{28}H_{16}$
4			$C_{22}H_{12}$	$C_{26}H_{14}$	$C_{30}H_{16}$	$C_{34}H_{18}$
6		$C_{24}H_{12}$	$C_{28}H_{14}$	$C_{32}H_{16}$	$C_{36}H_{18}$	$C_{40}H_{20}$
8		$C_{30}H_{44}$	$C_{34}H_{16}$	$C_{38}H_{18}$	$C_{42}H_{20}$	$C_{46}H_{22}$
10	$C_{32}H_{14}$	$C_{36}H_{16}$	$C_{40}H_{18}$	$C_{44}H_{20}$	$C_{48}H_{22}$	$C_{52}H_{24}$

J. R. Dias, *Handbook of Polycyclic Hydrocarbons, Part A (Benzenoid Hydrocarbons)*, Elsevier, Amsterdam, 1987, p. 36.

J. R. Dias, A periodic table for polycyclic aromatic hydrocarbons, *Accounts of Chemical Research*, 1985, 18, 241-248.

In Dias's formula periodic table, N_{ic} is the number of internal quaternary carbon atoms, and d_s is the number of disconnections of internal C–C bond edges. For catafusenes, $N_{ic} = 0$.

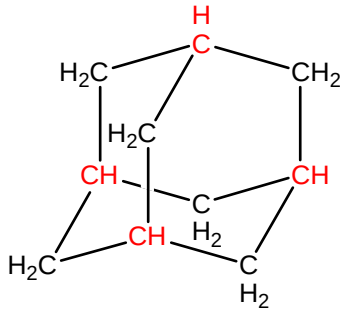
Benzenoids with R rings and molecular formula $C_{q+t}H_t$ have q quaternary carbon atoms and t tertiary carbon atoms (or CH groups). Quaternary carbon atoms are of two kinds: peripheral or internal; their numbers are denoted by p and i (N_{ic} in the Dias notation), respectively.

$$i = N_{ic} = q - t + 6 = R - d_s - 2$$

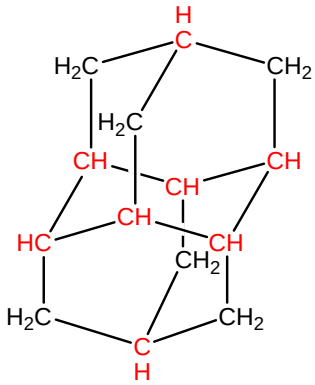
For catafusenes, $i = 0$, $q = p$. their molecular formula is $C_{4R+2}H_{2R+6}$ and the partitioned formula is $C_{2R-4}(CH)_{2R+6}$

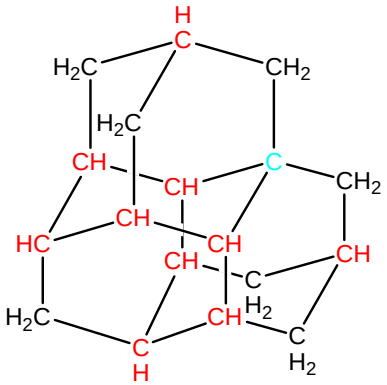
Diamondoids

Adamantane has formula $C_{10}H_{16}$ and the partitioned formula $(CH)_4(CH_2)_6$. It has four tertiary C and H atoms and six secondary C and H atoms.



Diamantane $C_{14}H_{20}$ and **triamantane** $C_{18}H_{24}$ also are single isomers. The latter is the first polymantane with a quaternary carbon atom.





Classification of **diamondoids**

The diamond lattice is self-dual.

Catamantanes have acyclic dualists
Perimantanes have dualists with chair-shaped hexagons
Coronamantanes have dualists with larger rings

Catamantanes with *n* adamantane units can be **regular** when they have molecular formulas $C_{4n+6}H_{4n+12}$. The difference between the numbers of carbon atoms and hydrogen atoms in regular catamantanes is 6.

Irregular catamantanes and **perimantanes** with *n* adamantane units have a lower number of carbon atoms than $4n+6$; the difference between the numbers of carbon atoms and hydrogen atoms in irregular catamantanes or perimantanes is lower than 6.

Coding of catamantanes

Catamantane and catafusene codings are similar.

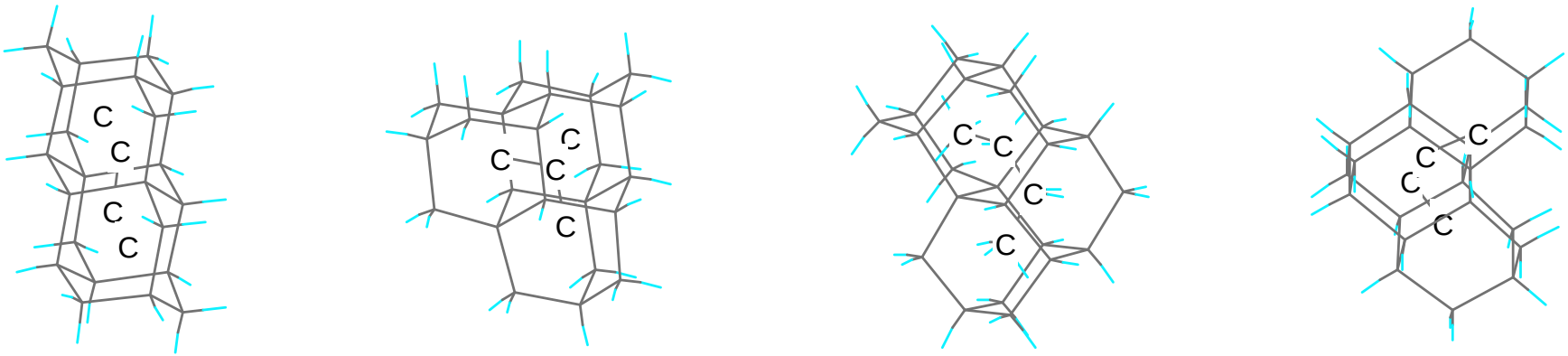
1. Select the longest nonbranched sequence of adamantane units or dualist edges

2. Assign for each edge of the dualist a digit (1, 2, 3,or 4) keeping the selection for each of the four directions

3. Branches are denoted by digits in brackets

4. Select the starting point so as to minimize the number formed by reading sequentially the digits and ignoring brackets

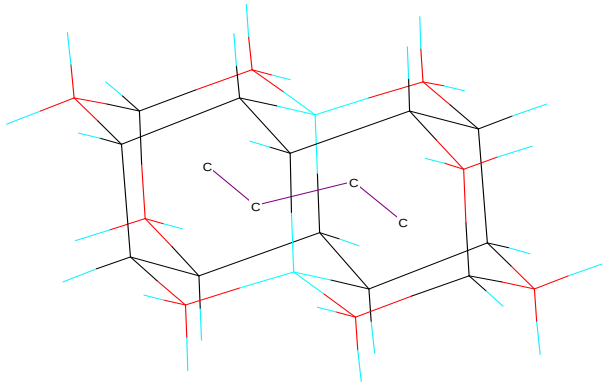
Dualists of diamondoids are isomorphic with carbon skeletons of staggered (cyclo)alkanes

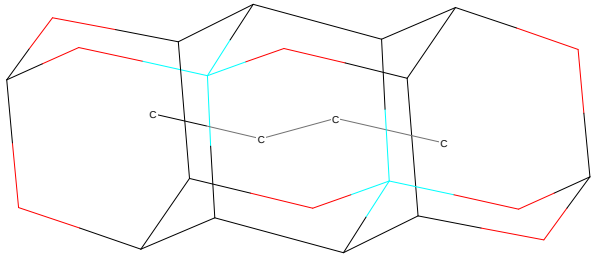


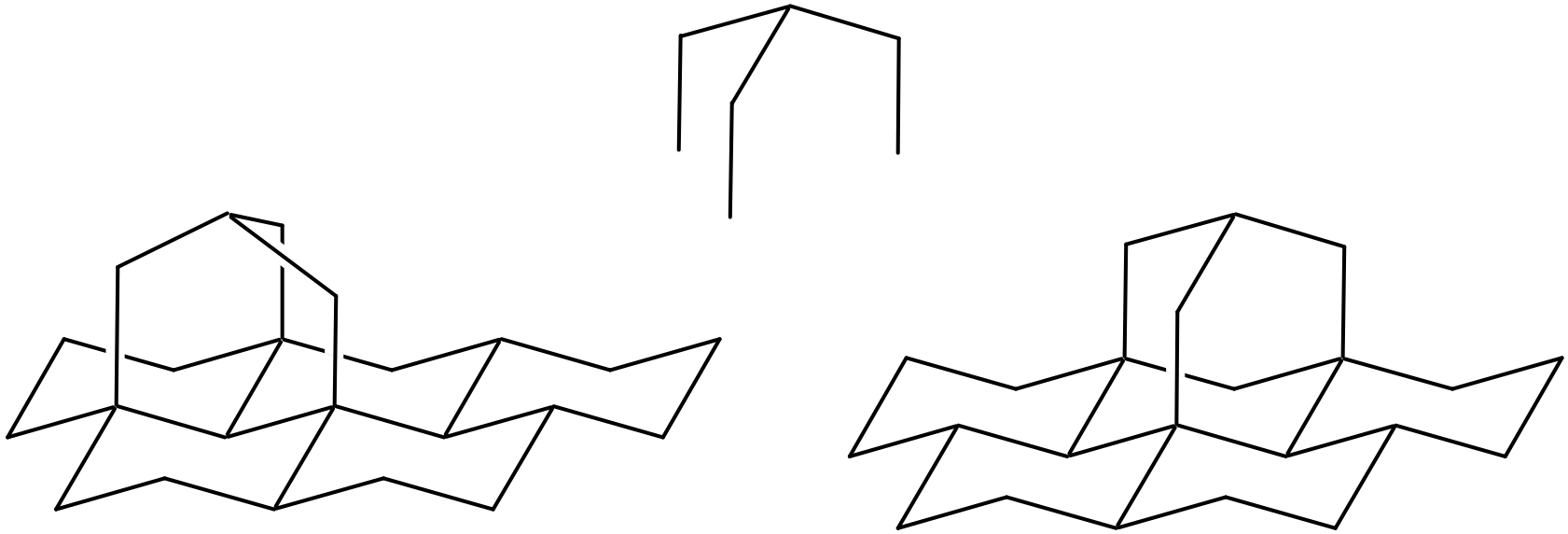
The four possible tetramantanes: the achiral [121]tetramantane, the achiral [1(2)3]tetramantane, and the two chiral enantiomers [123]tetramantane and [124]tetramantane.

Tertiary C–H bonds of adamantane can be more readily substituted (e. g. oxidized or halogenated) than secondary CH_2 bonds. However, higher diamondoids have several non-equivalent such bonds. For synthetic purposes, namely for obtaining new medicinal drugs (e. g. for having rod-like molecules with groups interacting with protein receptors at required distances) one needs accurate knowledge of diamondoid structures in order to obtain the desired selectivity. [121]Tetramantane $C_2(CH)_{12}(CH_2)_8$ serves as an example.

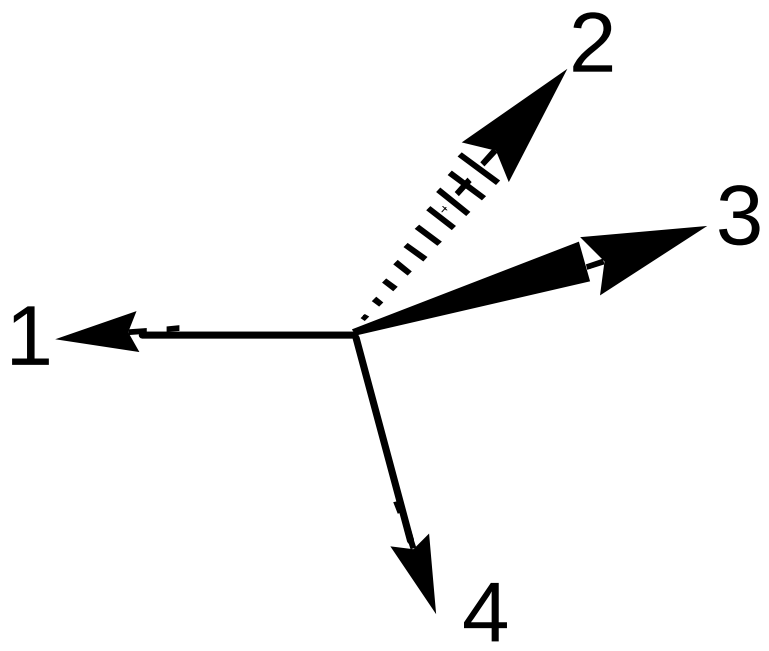
Schwertfeger, H.; Fokin, A. A.; Schreiner, P. R. *Angew. Chem. Int. Ed.* **2008**, *47*, 1022-1036.







The two possibilities of attaching to “perhydro-graphene” an isobutane fragment $(CH)_4$ shown above together with four “dangling bonds”: only the right-hand possibility is allowed, producing an adamantane unit with chair-shaped 6-membered rings in staggered conformations; the left-hand alternative yields tub-shaped 6-membered rings having bonds in eclipsed conformations.



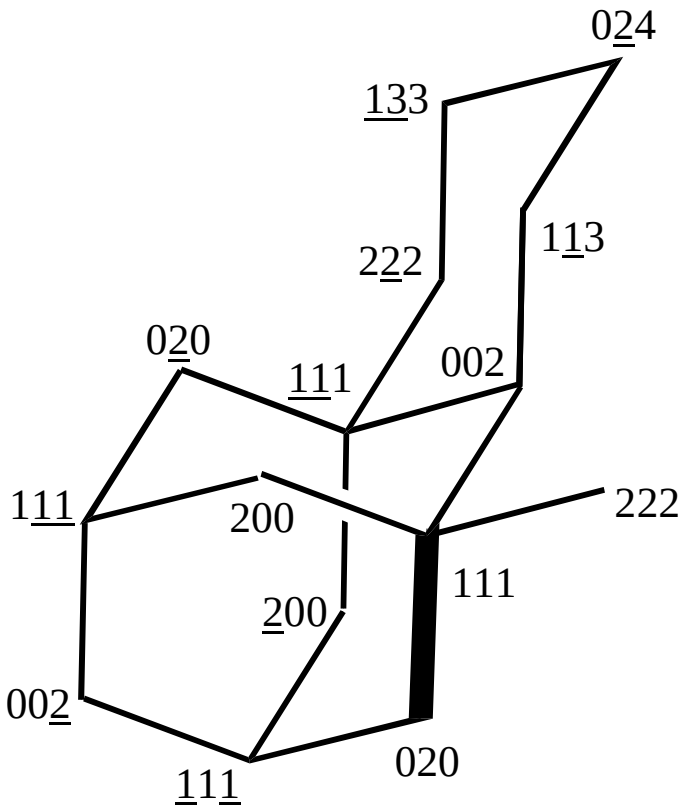
The four tetrahedral directions for coding regular diamondoids

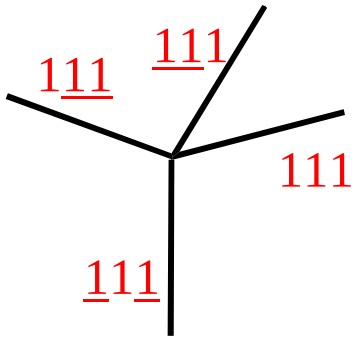
Vertices of the diamond lattice can be given integer coordinates by using a cubical unit cell four units across. With these coordinates, the points of the structure have coordinates (x, y, z) in the Euclidean three-dimensional space satisfying the two equations:

$$x = y = z \pmod 2$$

and

$$x + y + z = 0 \text{ or } 1 \pmod 4$$





Integer coordinates (*xyz*) for carbon atoms in the diamond lattice.

Left: carbon atom coordinates for an adamantane unit with a fused 6-membered ring and one attached carbon atom (minus signs are represented by underlined digits).

Right: on going from the central vertex to its four neighboring vertices, the red numbers indicate changes in the *xyz* coordinates.

There are eight points (modulo 4) that satisfy these conditions:
(0,0,0), (0,2,2), (2,0,2), (2,2,0), (3,3,3), (3,1,1), (1,3,1), (1,1,3)
All of the other points in the integer lattice may be obtained by adding multiples of four to the coordinates *x*, *y*, and *z* of these 8 points. In this integer lattice, adjacent points connected by C–C bonds are at distance $\sqrt{3}$ apart.

A triplet of even numbers is always connected with a triplet of odd numbers.

For economy of space, minus signs are written under each digit, instead of being placed before each digit.

The partition of carbon atoms in diamondoids with n adamantane units according to the vertex degrees of their constitutional hydrogen-depleted carbon scaffolds is indicated by the following relationships for regular catamantanes, where parameter b stands for the number of branches plus the number of degree-4 vertices in the dualist:

Quaternary C groups: $Q = n + b - 2 = S - 6$ (1)
Ternary CH groups: $T = 2(n - b + 2)$ (2)
Secondary CH_2 groups: $S = n + b + 4$ (3)

Thus a diamondoid with molecular formula

$$C_{4n+6}H_{4n+12}$$

can correspond to several valence-isomers with partitioned-formula

$$C_{n+b-2}(CH)_{2n-2b+4}(CH_2)_{n+b+4}$$

according to the value of parameter b

Catamamantanes with n = 1 to 5

Regular											
linear					branched						Irregular linear
n	Code	Symm.	Valence isomer	Formula	Code	Symm.	Valence isomer	Code	Symm.	Formula	Valence isomer
1		T_d	$(CH)_4(CH_2)_6$	$C_{10}H_{16}$							
2	1	D_{3d}	$(CH)_8(CH_2)_6$	$C_{14}H_{20}$							
3	12	C_{2v}	$C(CH)_{10}(CH_2)_7$	$C_{18}H_{24}$							
4	121	C_{2h}	$C_2(CH)_{12}(CH_2)_8$	$C_{22}H_{28}$	1(2)3	C_{3v}	$C_3(CH)_{10}(CH_2)_9$				
	123	C_2	$C_2(CH)_{12}(CH_2)_8$	$C_{22}H_{28}$							
5	1212	C_{2v}	$C_3(CH)_{14}(CH_2)_9$	$C_{26}H_{32}$	12(1)3	C_1	$C_4(CH)_{12}(CH_2)_{10}$	1231	C_s	$C_{25}H_{30}$	$C_2CH)_{16}(CH_2)_7$
	1213	C_1	$C_3(CH)_{14}(CH_2)_9$	$C_{26}H_{32}$	12(3)4	C_s	$C_4(CH)_{12}(CH_2)_{10}$				
	1234	C_2	$C_3(CH)_{14}(CH_2)_9$	$C_{26}H_{32}$	1(2,3)4	T_d	$C_6(CH)_8(CH_2)_{12}$				

For regular catamantanes, partitions of degrees of carbon atoms (upper row: quaternary-tertiary-secondary carbon atoms) and partitions of vertex degrees of dualists (lower row: numbers of vertices with degree 1, 2, 3, 4).

C\ *b* *n* \		0	1	2	3	4	5	6
Units	Formula	Zero branches	One branch	Two branches		Three branches		4 Branches
18 3	$C_{18}H_{24}$	1-10-7 **2**, 1, 0, 0						
22 4	$C_{22}H_{28}$	2-12-8 **2**, 2, 0, 0	3-10-9 **3**, 0, **1**, 0					
26 5	$C_{26}H_{32}$	3-14-9 **2**, 3, 0, 0	4-12-10 **3**, 1, **1**, 0		6-8-12 4, 0, 0, **1**			
30 6	$C_{30}H_{36}$	4-16-10 **2**, 4, 0, 0	5-14-11 **3**, 2, **1**, 0	6-12-12 **4**, 0, **2**, 0	7-10-13 **4**, 1, 0, **1**			
34 7	$C_{34}H_{40}$	5-18-11 **2**, 5, 0, 0	6-16-12 **3**, 3, **1**, 0	7-14-13 **4**, 1, **2**, 0	8-12-14 **4**, 2, 0, **1**	9-10-15 **5**, 0, **1, 1**		
38 8	$C_{38}H_{44}$	6-20-12 **2**, 6, 0, 0	7-18-13 **3**, 4, **1**, 0	8-16-14 **4**, 2, **2**, 0	9-14-15 **4**, 3, 0, **1**	10-12-16 **5**, 1, **1, 1**	9-14-15 5, 0, 3, 0	12-8-18 **6**, 0, 0, **2**

A spectacular symmetry of this periodic table of regular catamantanes $C_{4n+6}H_{4n+12}$ is apparent. Practically all partitions of *C* in can be expressed in terms of parameters *n* and *b* indicating that the numbers of quaternary and secondary carbon atoms increase according to the number *n* of adamantane units and number *b* of branches, whereas the number of tertiary CH groups increases twice as fast. The three components of the triplet *Q*-*T*-*S* indicating the numbers of C, CH, and CH_2 groups, respectively, are:

$$Q = n + b - 2 = S - 6 \quad (1)$$

$$T = 2(n - b + 2) \quad (2)$$

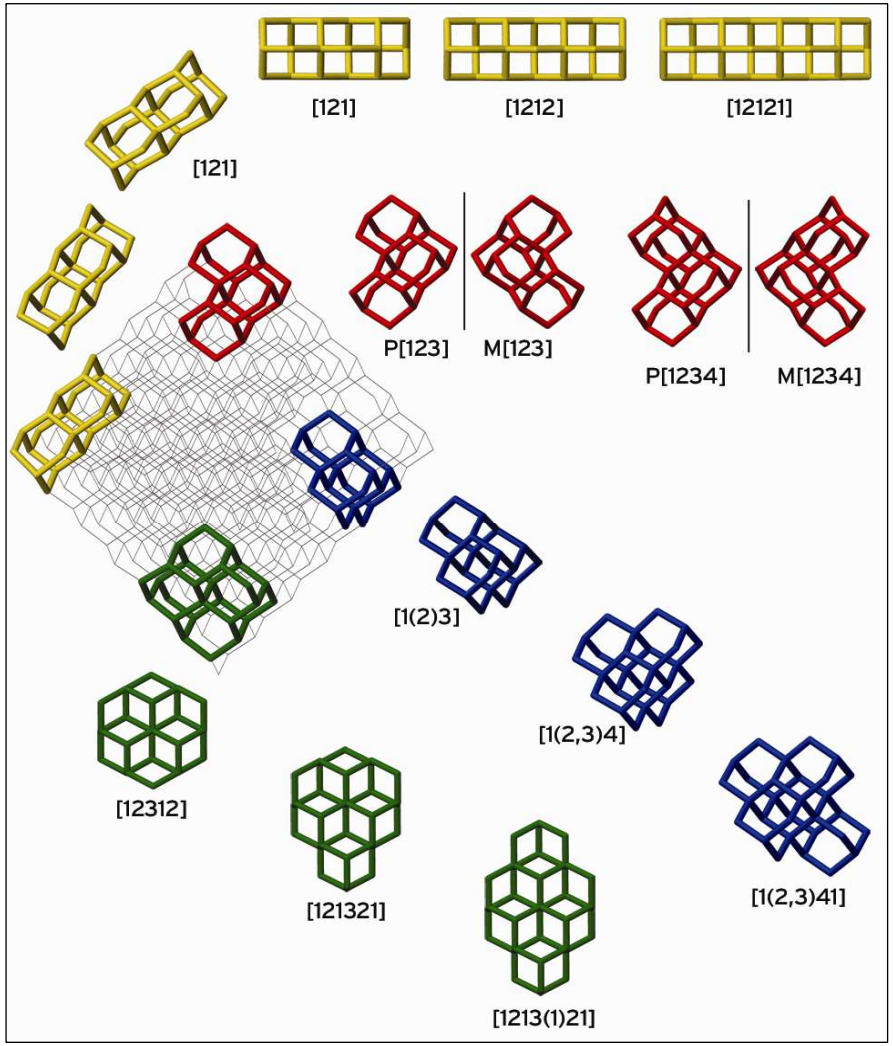
$$S = n + b + 4 \quad (3)$$

In other words, regular catamantanes have the following partitioned formula: $C_{n+b-2}(CH)_{2n-2b+4}(CH_2)_{n+b+4}$ and molecular formula $C_{4n+6}H_{4n+12}$

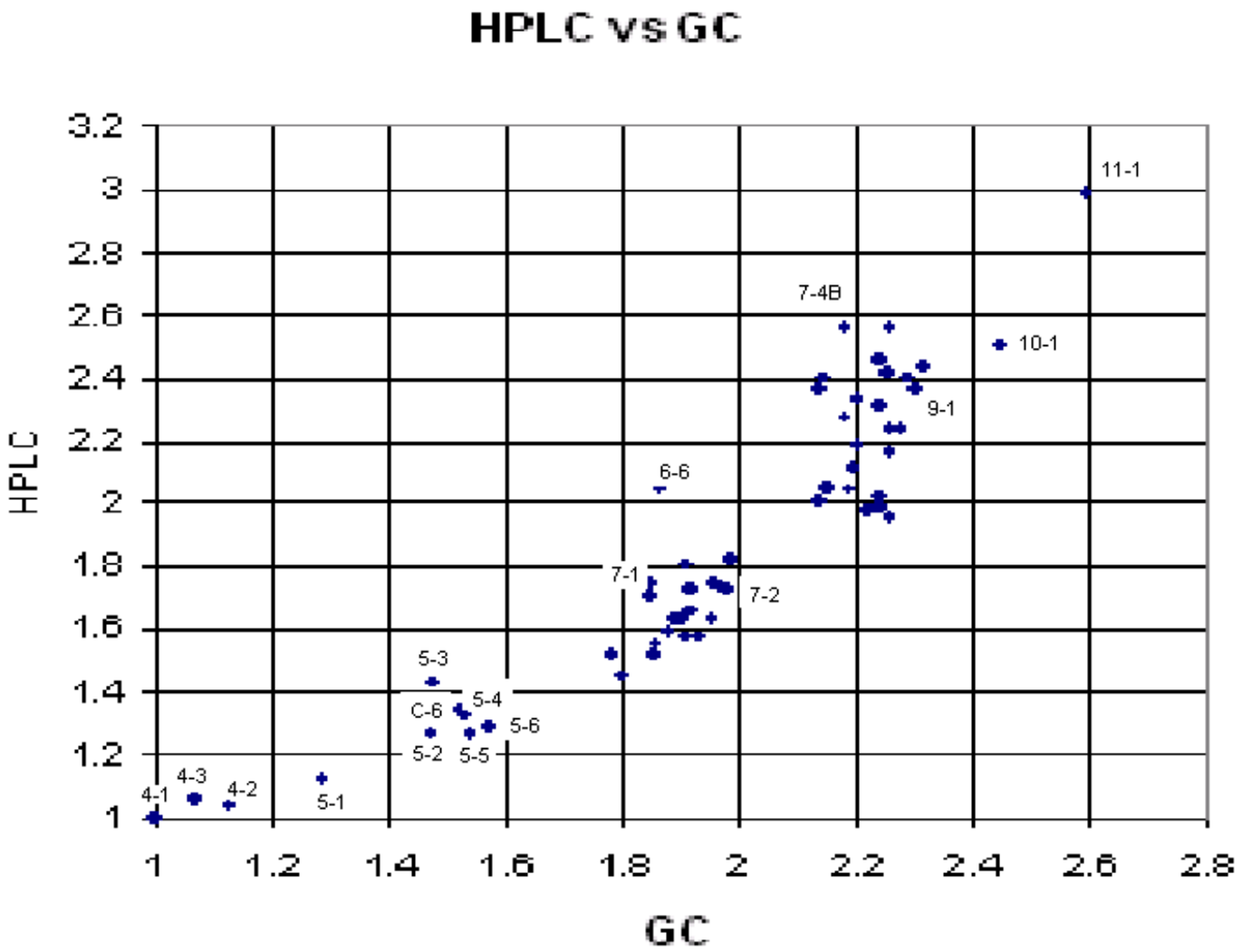
Irregular catamantanes and perimantanes. The same conventions apply as for regular catamantanes, except that diamondoids with equal *n* are not isomeric; molecular formulas are now inscribed in parentheses in each cell of the table.

Units	0	1	2	3	4	5	6
	Irregular catamantanes						Perimantanes
	Zero branches		One branch		Two branches		
n	One proximity	Two proximities	One proximity	Two proximities	One proximity	Two proximities	
5	2-16-7 2, 3, 0, 0 ($C_{25}H_{30}$)						
6	3-18-8 4-16-9 2, 4, 0, 0 ($C_{29}H_{34}$)		4-16-9 3, 2, 1, 0 ($C_{29}H_{34}$)				2-18-6 0, 6, 0, 0 ($C_{26}H_{30}$)
7	4-20-9 5-18-10 2, 5, 0, 0 ($C_{33}H_{38}$)	5-18-9 2, 5, 0, 0 ($C_{32}H_{36}$)	5-18-10 6-16-11 3, 3, 1, 0 ($C_{33}H_{38}$)	5-18-9 3, 3, 1, 0 ($C_{32}H_{36}$)	7-14-12 4, 2, 0, 1 ($C_{33}H_{38}$)	5-18-9 4, 1, 2, 0 ($C_{32}H_{36}$)	4-18-8 1, 5, 1 0 ($C_{30}H_{34}$)

Examples of four structure series of higher diamondoids



HPLC eluent volumes versus the GC retention times



Chromatographic data of identified diamondoids from petroleum

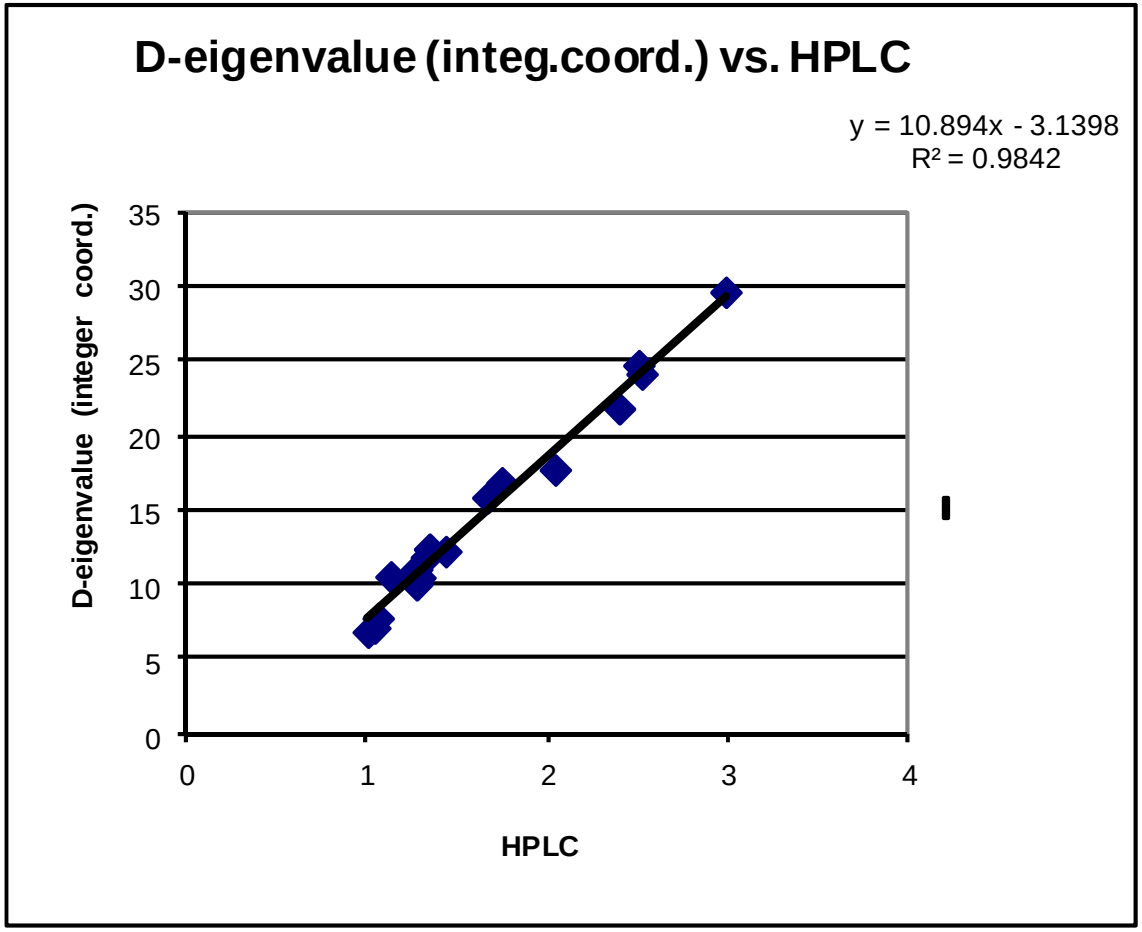
Balaban AT, Klein DJ, Dahl JA, Carlson RMK , Molecular descriptors for natural diamondoid hydrocarbons and quantitative structure-property relationships for their chromatographic data. *The Open Org Chem J* (2007) 1: 13-31

Higher Diamondoid Code	Compound Reference Number	Molecular Formula	M^+ (m/z) Base Peak	GC/MS Relative Retention Times*	ODS HPLC Relative Elution Volumes*
[1(2)3]	4;1	$C_{22}H_{28}$	292	1	1
[121]	4;2	$C_{22}H_{28}$	292	1.069	1.055
[123]	4;3	$C_{22}H_{28}$	292	1.126	1.036
[1(2,3)4]	5;1	$C_{26}H_{32}$	344	1.284	1.126
[12(1)3]	5;2	$C_{26}H_{32}$	344	1.473	1.269
[1212]	5;3	$C_{26}H_{32}$	344	1.479	1.431
[1213]	5;4	$C_{26}H_{32}$	344	1.529	1.323
[12(3)4]	5;5	$C_{26}H_{32}$	344	1.543	1.269
[1234]	5;6	$C_{26}H_{32}$	344	1.57	1.287
[12312]	C-6	$C_{26}H_{30}$	342	1.523	1.341
[12121]	6;6	$C_{30}H_{36}$	396	1.868	2.039
[121321]	7;1	$C_{30}H_{34}$	394	1.847	1.743
[123124]	7;2	$C_{30}H_{34}$	394	1.917	1.664
[121212]	7;4B	$C_{34}H_{40}$	448	2.185	2.521
[12312412]	9;1	$C_{34}H40$	444	2.292	2.396
[1231241(2)3]	10;1	$C_{38}H_{43}$	456	2.451	2.503
[123(1,2)42143]	11;1	$C_{39}H_{38}$	508	2.599	2.987

Molecular invariants (D-eigenvalues) and chromatographic data

Diamondoid	***n***	D-eigenv.	GC/MS	HPLC	Type
[1(2)3]	4	6.87	1.000	1.000	4-cata
[121]	4	7.79	1.069	1.055	4-rod
[123]	4	7.15	1.126	1.036	4-cata
[1(2,3)4]	5	10.58	1.284	1.126	5-bundle
[12(1)3]	5	11.08	1.473	1.269	5-cata
[1212]	5	12.27	1.479	1.431	5-rod
[1213]	5	11.88	1.529	1.323	5-cata
[12(3)4]	5	10.04	1.543	1.269	5-cata
[1234]	5	10.53	1.570	1.287	5-cata
[12121]	6	17.74	1.868	2.039	6-rod
[12312]	6	12.44	1.523	1.341	6-bundle
[121321]	7	16.89	1.847	1.743	7-bundle
[123124]	7	15.88	1.917	1.664	7-bundle
[121212]	7	24.17	2.185	2.521	7-rod
[12312412]	9	21.84	2.292	2.396	9-bundle
[1231241(2)3]	10	24.74	2.451	2.503	10-bundle
[123(1,2)42143]	11	29.65	2.599	2.987	11-bundle

Eigenvalues of the distance matrix for idealized coordinates vs. HPLC relative elution values



Computational challenges

Computational programs for the enumeration of benzenoids including or excluding helicenes.

Computational programs for finding Clar numbers (enumerating Clar structures) of large fullerenes and nanotubes.

Computational programs for obtaining partitioned-formulas for perimantanes, regular and irregular catamantanes, and their point-group symmetries.

Developing coding systems for perifusenes and perimantanes that can assist in nomenclature and enumerations.