

ENUMERATION OF CATAFUSENES, DIAMONDOID HYDROCARBONS,
AND STAGGERED ALKANE C-ROTAMERS

Alexandru T. Balaban

Institute of Chemistry, The Polytechnic, Bucharest, Roumania

(received: March 1976)

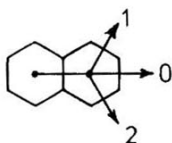
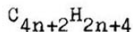
Counting formulas for catafusenes (earlier work based on dualist graphs) are briefly reviewed. It is now reported that the number of normal alkane staggered C-rotamers C_kH_{2k+2} is the same as that of non-branched catafusenes with $k - 1$ benzenoid rings (if enantiomers of C-rotamers are counted together). For branched systems, the situation is somewhat more complex. Diamondoid hydrocarbons have dualist graphs which are staggered alkane or cycloalkane C-rotamers so that this approach may also count (with some restrictions) diamondoid hydrocarbons. Spin-offs of this investigation are more rigorous definitions and systematizations, coding and nomenclature systems, and correlations with chemical properties (resonance energy, chirality, optical rotatory power).

Catafusenes

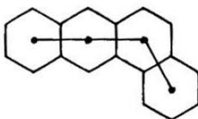
Several years ago, the enumeration of all possible cata-condensed polycyclic benzenoid hydrocarbons (catafusenes) was made possible by using "the dualist graph approach" : the centre of each benzenoid ring becomes a vertex of a new graph, and two vertices are joined by an edge whenever the rings are condensed (ortho-fused). Unlike "dual graphs", the new graph thus obtained has certain restrictions, namely that all edges have a constant length, and that their angles may only be 120° or 180° ; also, there is no point corresponding to the external region, as dual graphs have. Therefore this new graph, formerly called ¹ characteristic graph, is better called

dualist graph.

Polycyclic benzenoid hydrocarbons (polyhexes) are classified ¹ into two main groups, catafusenes, whose dualist graphs are acyclic, and perifusenes, whose dualist graphs have circuits. All catafusenes with the same number n of condensed benzenoid rings are isomeric and have molecular formula

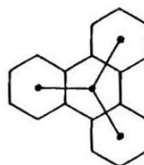


I



01

II



1(.)

III

The enumeration of non-branched catafusenes was made by assigning three digits : 0, 1, and 2, to the three directions in which a catafusene can be annelated to a higher homologue (formula I). It was shown ¹ that the number F_n of non-branched catafusenes with n benzenoid rings is

$$F_n = (3^{n-2} + 4 \times 3^{(n-3)/2} + 1)/4 \quad \text{for } n \text{ odd}$$

$$F_n = (3^{n-2} + 2 \times 3^{(n-2)/2} + 1)/4 \quad \text{for } n \text{ even.}$$

Each catafusene can be coded by a string of digits indicating the angles between the edges of the dualist graph (digit 0 stands for 180° , digits 1 or 2 for 120° or 240°), starting from one end, with the convention that the smallest number formed by the string of digits is adopted as the code (in most cases, e. g. II, different numbers result when one starts from the other end or when one reverts the convention concerning digits 1 and 2). This code can be used in a new nomenclature for catafusenes, e. g. II can be named [01]tetra-catafusene.

A complicated formula results for the enumeration of branched catafusenes. ² Their code can be formulated by indi-

ating the branch(es) in round brackets after the digit where branching occurs ; a one-vertex branch is indicated by a dot, e. g. III is [1(.)]tetracatafusene.^{3,*} The conventions for finding the code as the minimal number ignore the brackets, but have additional rules.³

Perifusenes have so far not been amenable to a mathematical enumeration, but the numbers of the first terms in the series were calculated by a computer program.⁴ Numerical data for catafusenes and perifusenes with $n = 1 - 7$ are presented in table 1.

On the graphite lattice, there are three orientations of edges, hence a bidimensional dualist graph of a catafusene superimposable with this lattice needs three digits : they can be 0, 1, and 2 as above (or in other applications 1, 2, and 3, e. g. for coding configurations of annulenes).^{5,6}

Topological correlations between codes of catafusenes and chemical properties such as resonance energy E , or polarographic reduction potential, were found,⁷ depending linearly on two parameters : the number n of condensed benzenoid rings, and the number a of zero digits in the code :

$$E = 1.75 n - 0.15 a + 0.25 \quad (\text{in beta-units}).$$

Staggered normal-alkane C-rotamers⁸

Diamondoid hydrocarbons, e. g. adamantane, diamantane, etc., can be enumerated by a similar procedure, constructing their dualist graphs ; in these cases, these graphs are no longer bidimensional, but tridimensional dualist graphs : their vertices are the centres of adamantane units.

These dualist graphs are superimposable on the diamond lattice, hence they have fixed tetrahedral angles (109.5°) and are identical with the carbon skeletons of staggered alkane or cycloalkane conformers (C-rotamers). Therefore we started to enumerate staggered alkane C-rotamers, and found that for n -alkanes^{**} $C_k H_{2k+2}$ their numbers R_k are the same as those of

* The codes and dualist graphs may be seen on formulas II-III.

** Here, n stands for normal, i. e. non-branched, therefore we use henceforth the index k for the number of vertices in the dualist graph.

Table 1. Numbers of catafusenes, staggered C-rotamers, and diamondoid hydrocarbons ^a

n or k	Catafusenes and polyhexes				Staggered C-rotamers of ^c alkanes and cycloalkanes				Diamondoid hydrocarbons ^c							
	lin. CF	br. CF	total CF	PF ^e PH ^e	lin. CR	br. CR	br. CR	total CR	lin. DHR	br. DHR	br. DHR	total DHR				
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
1	1	0	1	0	1	1	0	0	0	1	1	0	0	0	0	1
2	1	0	1	0	1	1	0	0	0	1	1	0	0	0	0	1
3	2	0	2	1	3	1	0	0	0	1	1	0	0	0	0	1
4	4	1	5	2	7	2	1	0	0	3	2	0	1	0	0	3
5	10	2	12	10	22	4	2	1	0	7	3	1	3	0	0	7
6	25	12	37	45	82	10	13	1	1	25	7	2	10	4	1	24
7	70	53	123	210	333	25	61	7	2	95	13	9	33	31	2	88

^a Abbreviations : CF = catafusenes ; CR = staggered C-rotamers ; PF = perifusenes ; PH = polyhexes ; wo. = without ; w. = with ; cy. = cyclic ; lin. = linear (non-branched) ; br. = branched ; DHR = regular diamondoid hydrocarbon ; DHR = irregular DH.

^b The number of vertices in the dualist graph of the CF or DH, and of carbons in the CR.

^c Enantiomers are not counted separately, but in pairs.

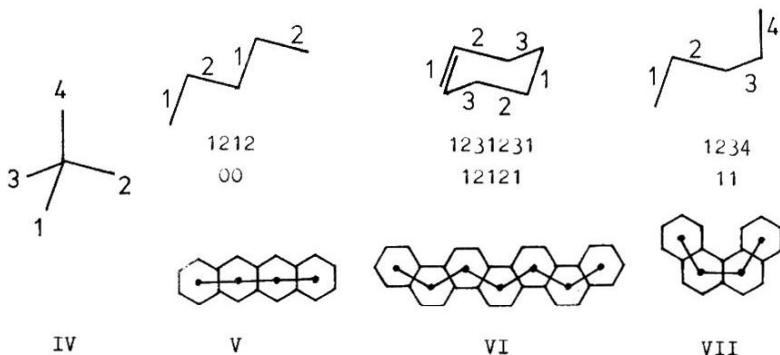
^d Calculated after ref. 2

^e Computer calculations ; ⁴ all other figures were found by the present author.

non-branched catafusenes (if no distinction is made between enantiomeric C-rotamers) :

$$R_k = F_{k-1}$$

This is not unexpected, since also in this case homologation can occur in three directions (cf. formula IV). Since all angles are equal, however, we will use digits 1 - 4 to discriminate among the four orientations of edges on a diamond lattice.⁹

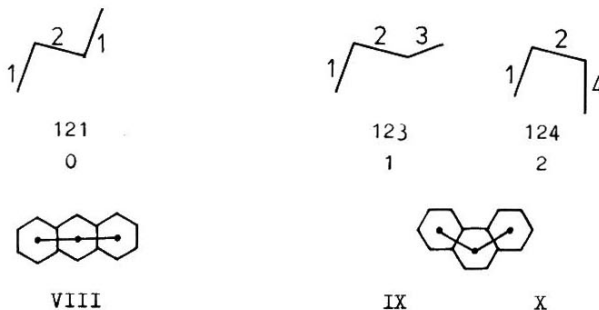


One can obtain a one-to-one correspondence between each n-alkane C-rotamer and each non-branched catafusene. This 1-1 correspondence can be used for converting the 4-digit notation of staggered C-rotamers into a 3-digit notation, as seen from formulas V - VII (in these and subsequent formulas, ordered under one another, are presented : the C-rotamer, the 4-digit code, the 3-digit code, and the corresponding catafusene).

Each triplet in the 4-digit notation corresponds to one of the digits 0, 1, and 2 in the 3-digit notation ; the four-digit code is a degenerate, i. e. an overlapping, code. The triplet can be interpreted as a helicity, e. g. the carbon skeleton of n-butane can exist in three conformations of low energy corresponding to three helicities : trans, VIII, achiral with helicity 0, and two enantiomeric gauche conformations, IX and X, with non-zero helicity symbolized by 1 and 2.

If the convention of the minimal number from the string of digits (ignored in the codes of X) is maintained, then

enantiomeric helicities IX and X cannot be distinguished and then they both correspond to the same catafusene (phenanthrene).



One can devise a "dictionary" between 4-digit triplets and 3-digit codes, as seen in table 2. With the help of this dictionary, any n-alkane staggered C-rotamer can be described either in the 4-digit or in the 3-digit code. For the rules which govern the allowed sequences of digits in the four-digit code, and for the elaboration of the dictionary presented in table 2, the original reference should be consulted.⁸

Table 2. Correspondence between triplets of four digits (1-4) and single digits (0,1,2) of the 3-digit code ^a

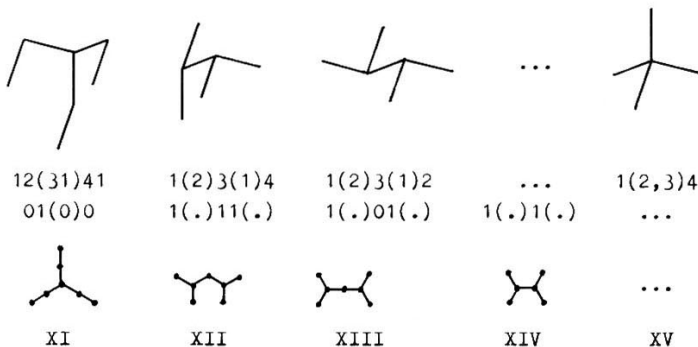
3-Digit code, single digit:	0		1		2	
4-Digit	121	313	123	312	124	314
triplet	131	323	134	324	143	342
of	141	343	142	341	132	321
digits :	212	414	214	413	213	412
	232	424	243	432	234	423
	242	434	231	421	241	431

^a This dictionary is valid when the first digit of the triplet is in an odd position of the code ; if it is in an even position of the 4-digit code, the assignment 1/2 is reversed.

Numerical results for the first terms in the series may be seen in table 1. Formulas were also found for counting separately chiral and achiral staggered C-rotamers, and in the original reference they are presented systematically, individually, with their 3- and 4-digit codes and their point group symmetry symbols.⁸

Staggered branched alkane C-rotamers ⁸

For any branched alkane with no quaternary carbon, the number of its staggered C-rotamers (counting together each pair of enantiomers for the chiral C-rotamers) can be found by a 1-1 correspondence with branched catafusenes. If t is the number of tertiary carbons of the branched alkane C_kH_{2k+2} , then the number of its staggered C-rotamers is equal to the number of branched catafusenes with $n + t - 1$ benzenoid rings and t branching points of its dualist graph (bidimensional) : this means that each branch of the alkane introduces one extra vertex into the bidimensional graph of the corresponding catafusene. As a consequence, with one branching point, the bidimensional dualist graph has as many vertices as the alkane C-rotamer (though their shapes may be dissimilar, as for XI, the positions of the branch coincide) ; no two branching points may be adjacent in the bidimensional dualist graph, though they may in the corresponding alkane C-rotamer, e. g. XII and XIII. Therefore, no alkane corresponds to branched catafusenes like XIV with adjacent branching points, and conversely no bidimensional dualist graph (hence no three-digit code) may correspond to branched alkanes with quaternary carbon atoms like XV.



Numerical formulas were devised for restricted groups of branched alkane staggered C-rotamers. Data may be seen in table 1.

The same approach may be applied to conformers of cycloalkanes superimposable on the diamond lattice, leading to a 3-digit code (provided there are no quaternary carbons) in addition to the 4-digit code developed previously for this case by Saunders⁹).

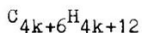
Both for non-branched acyclic and cyclic staggered C-rotamers, there exists a direct correlation between the 3-digit code and the chirality of the rotamer : the sum of the digits in the 3-digit code is zero for all achiral rotamers, while for chiral rotamers, the higher is this sum, the higher is their helicity (and, probably, their optical rotatory power). This zero sum condition is necessary, but insufficient for achirality of C-rotamers : indeed, since there exist chiral rotamers with opposite helicities of triplets which are not symmetrically matched relatively to a plane or centre of symmetry in the molecule, such rotamers will yield a zero sum but will not be achiral ; their optical rotatory power is expected to be extremely small, however, compared with rotamers which yield non-zero sums.

Diamondoid hydrocarbons 10

A diamondoid hydrocarbon has a carbon skeleton which consists of fused adamantane units and is a portion of the diamond lattice.

Though, as mentioned above, all dualist graphs of diamondoid hydrocarbons (DH's for short) are staggered alkane or cycloalkane C-rotamers, the converse is not true.

Whenever in the 4-digit code of the alkane C-rotamer there is no sequence of the form .axya...^{*}, the C-rotamer is a dualist graph of a "regular DH". All regular DH's with the same number k of vertices in their dualist graphs (acyclic branched or non-branched) are isomeric and have formula



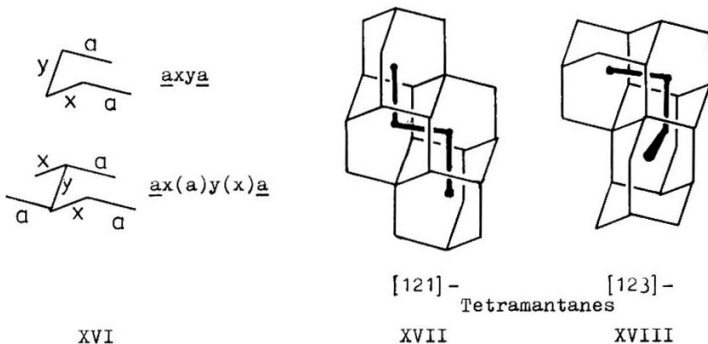
^{*} The letters a, b, x, y stand for one of the digits 1, 2, 3, 4 each ; branching points may exist in the sequence, but in this case the distance between the same digit a which is repeated in the above sequence must be conserved, as in XVI.

When, however, such sequences do exist, the "irregular DH" has fewer carbon and hydrogen atoms than required by the above formula ; both the regular and the irregular DH's with acyclic dualist graphs can be called catamantanes, whereas DH's with cyclic dualist graphs can be named perimantanes ; of course, all perimantanes are irregular.

When sequences ...abxab... exist in the four-digit code of the acyclic C-rotamer, no DH may correspond to it, e. g. there is no hexamantane with acyclic dualist graph 12312 or heptamantane 123123, 121321, 123124, 1231(2)3, 1231(2)4, 12(1)312 and 123(4)12 : in all these cases, the dualist graph is cyclic ; this explains the differences between columns (11) and (17) in table 1.

Because of these restrictions, no general formula for the enumeration of DH's was yet found, but a procedure was developed which served to enumerate the first terms of the series as seen in table 1.

A new nomenclature and coding are proposed on this basis for DH's, which will avoid both the ambiguity of trivial names (there exist two tetramantanes, XVII and XVIII, both regular, whose dualist graphs are the n-butane staggered C-rotamers ; note that the latter DH is chiral ; the third possible tetramantane has a branched dualist graph corresponding to isobutane), and the complexity of systematic names using the von Baeyer system adopted by IUPAC and Chemical Abstracts. As yet, only mono-, di- and triamantane are known, mainly owing to the investigations of Schleyer's school.¹¹



Conclusions

As seen from the above discussion, the method of dualist graphs served to enumerate all possible catafusenes and diamondoid hydrocarbons. The dualist graphs of the latter are themselves part of the diamond lattice and represent carbon skeletons of staggered alkane or cycloalkane C-rotamers.

The numbers of homologous non-branched catafusenes and C-rotamers are identical, and with certain limitations these numbers approximate those of DH's whose dualist graphs are non-branched. Counting formulas for the former numbers were found. Branched acyclic and cyclic systems lead to more complicated counting formulas which were found only for restricted classes, but which can lead to useful algorithms.

The above graph-theoretical approach also served for systematization, definition, coding and nomenclature purposes, as well as for topological correlations with physical and chemical properties of these systems.

Erratum

In a preceding brief exposition of these ideas,¹² in the respective paragraph reference number 52 should have been inserted ; at present, this quotation should be replaced by references 8 and 10 below. Also, in the same paper¹²(p. 43, line 2), the correct notation for triphenylene in the new nomenclature system is that given under formula III of the present paper, namely [1(.)]tetracatafusene.

References

- 1 A. T. Balaban and F. Harary, *Tetrahedron*, 24, 2505 (1968).
- 2 F. Harary and R. C. Read, *Proc. Edinburgh Math. Soc.*, Ser. II, 17, 1 (1970).
- 3 A. T. Balaban, *Tetrahedron*, 25, 2949 (1969).
- 4 D. A. Klarner, *Fibonacci Quart.*, 3, 9 (1965).
- 5 M. Gordon and W. H. T. Davison, *J. Chem. Phys.*, 20, 428 (1952).
- 6 A. T. Balaban, *Tetrahedron*, 27, 6115 (1971).
- 7 A. T. Balaban, *Rev. Roumaine Chim.*, 15, 1243 (1970).
- 8 A. T. Balaban, to appear.
- 9 M. Saunders, *Tetrahedron*, 23, 2105 (1967).
- 10 A. T. Balaban and P. von R. Schleyer, to be published.
- 11 R. C. Bingham and P. von R. Schleyer, *Fortschr. Chem.*

- Forsch., 18, 1 (1971) and further references therein ;
M. A. McKervey, Chem. Soc. Revs., 3, 479 (1974).
12 A. T. Balaban, Match, Nr. 1, 33 (1975), namely p. 44.

Note concerning table 1

Though most balance equations are evident, they are listed below for convenience. All numbers in brackets refer to columns in table 1.

$$(4) = (2) + (3)$$

$$(6) = (4) + (5)$$

$$(11) = (7) + (8) + (9) + (10)$$

$$(17) = (12) + (13) + (14) + (15) + (16)$$

$$(10) \cong (16)$$

$$(11) \cong (17)$$

$$(7) \cong (12) + (13)$$

$$(8) + (9) \cong (14) + (15)$$