REFERENCES AND NOTES

(1) O. Kennard, D. G. Watson, F. H. Allen, W. D. S. Motherwell, W. G. Town, and J. R. Rodgers, "Crystal Clear Data". *Chem. Br.*, **11**, 213-6 (1975).
(2) O. Kennard, D. G. Watson, and W. G. Town, "Cambridge Crystallographic Data Centre I: The Bibliographic File". *J. Chem. Doc.*, **12**, 14-19 (1972).
(3) F. H. Allen, S. Bellard, M. D. Brice, B. A. Cartwright, A. Doubleday, H. Higgs, T. Hummelink, B. G. Hummelink-Peters, O. Kennard, W. D. S. Motherwell, J. R. Rodgers, and D. G. Watson, "The Cambridge Crystallographic Data Centre: Computer-Based Search, Retrieval, Analysis and Display of Information". *Acta Crystallogr., Sect. B*, **B35**, 2331-9 (1979).
(4) F. H. Allen, O. Kennard, W. D. S. Motherwell, W. G. Town, and D. G. Watson, "Cambridge Crystallographic Data Centre II: The Structural Data File". *J. Chem. Doc.*, **13**, 119-23 (1973).
(5) O. Kennard and D. G. Watson, "Molecular Structures and Dimensions", Vol. 1-3, Oosthoek, Utrecht, 1970-1973; O. Kennard, D. G. Watson, and W. G. Town, *ibid.*, Vol. 4, 5, 1974; O. Kennard, D. G. Watson, F. H. Allen, and S. M. Weeds, *ibid.*, Vol 6-11, Bohn, Scheltema, and Holkema, Utrecht, 1975-1980; O. Kennard, D. G. Watson, F. H. Allen, and S. A. Bellard, Vol. 12, D. Reidel, Dordrecht, 1981.
(6) O. Kennard, D. G. Watson, F. H. Allen, N. W. Isaacs, W. D. S. Motherwell, R. C. Pettersen and W. G. Town, "Molecular Structures and Dimensions: Interatomic Distances 1960-65", Vol. A1, Oosthoek, Utrecht, 1973.
(7) F. H. Allen, N. W. Isaacs, O. Kennard, W. D. S. Motherwell, R. C. Pettersen, W. G. Town, and D. G. Watson, "Cambridge Crystallographic Data Centre IV: Preparation of Interatomic Distances 1960-65". *J. Chem. Doc.*, **13**, 211-9 (1973).
(8) O. Kennard, F. H. Allen, and D. G. Watson, "Molecular Structures and Dimensions: Guide to the Literature 1935-76", Bohn, Scheltema, and Holkema, Utrecht, 1977.
(9) F. H. Allen and W. G. Town, "The Automatic Generation of Keywords from Chemical Compound Names: Preparation of a Permuted Name Index with KWIC Layout". *J. Chem. Inf. Comput. Sci.*, **17**, 9-15 (1977).
(10) F. H. Allen, "Cambridge Crystallographic Data Centre V: An Integrated System of Printed Indexes". *J. Chem. Inf. Comput. Sci.*, **20**, 68-75 (1980).
(11) K. M. Crennell, "The Development of a Layout Program for Scientific Texts using an FR80 Microfilm Recorder" Report No. RL-81-038, Rutherford Appleton Laboratory, Chilton, England, 1981. See also K.M. Crennell, "Users Guide to PURL (Preliminary Documentation)", Rutherford Appleton Laboratory, Chilton, England, 1980.
(12) (a) A complete list of journal names is read by BIBTYP at run time. Each name contains preassigned line-break points. (b) The concatenated sequence journal name*volume*page*year normally fits on one line, but for a few long journal names line breaks are provided in the name.
(13) A. V. Hershey, "Calligraphy for Computers" Report No. AD-662-398, U.S. Naval Weapons Laboratory, Dahlgren, VA, 1967. See also N. M. Wolcott and J. Hilsenrath, "A Contribution to Computer Typesetting Techniques: Tables of Coordinates for Hershey's Repertory of Occidental Type Fonts and Graphic Symbols", *NBS Spec. Publ.*, **No. 424** (1976). (We thank Dr. N. M. Wolcott, Mr. J. Hilsenrath, and Mrs. C. Messina of NBS for making the font specifications available to us and Dr. M. F. Chiu of the Rutherford Appleton Laboratory for assistance in its implementation on the FR80).
(14) P. A. Dewar, "The SMOG System Manual", Computing Division, Atlas Centre, Rutherford Appleton Laboratory, Chilton, England, 1976.
(15) B. A. Cartwright, O. Kennard, and D. G. Watson, "Cambridge Crystallographic Data Centre: The Chemical Diagram System", to be submitted for publication.

Periodic Table for Polycyclic Aromatic Hydrocarbons. 2. Polycyclic Aromatic Hydrocarbons Containing Tetragonal, Pentagonal, Heptagonal, and Octagonal Rings

JERRY RAY DIAS

Department of Chemistry, University of Missouri, Kansas City, Missouri 64110

Received December 14, 1981

A formula polycyclic aromatic hydrocarbon (PAH) periodic table previously discovered for exclusively fused hexagonal rings is now extended to include tetragonal, pentagonal, heptagonal, and octagonal rings. The original formula PAH6 periodic table and graph-theoretical equations are equally applicable to polycyclic aromatic hydrocarbons composed of hexagonal rings and/or an equal number of tetragonal and octagonal rings and/or an equal number of pentagonal and heptagonal rings (azulenoids). The maximum number of four-membered rings that a structure corresponding to a PAH6 formula can possess in addition to hexagonal rings is given by $r_{4max} \leq [N_c - 2N_H + 6]/2$, and the maximum number of five-membered rings that a structure corresponding to a PAH6 formula can possess is given by $r_{5max} \leq [N_c - 2N_H + 6]$. N_c and N_H are the number of carbon and hydrogen atoms in the molecular formula, respectively. The formula-structure rule for (noncirculene) totally fused PAH's is $N_{Pc} = N_H - 6 + 2r_4 + r_5 - r_7 - 2r_8$, where N_{Pc} is the number of third-degree peripheral carbon atoms (vertices in a structural graph consisting of only the C-C σ-bond framework) and where r_4, r_5, r_7, and r_8 are the numbers of four-, five-, seven-, and eight-membered rings, respectively. It is believed that this is the first formulation of the above graph-theoretical relationships.

In continuation of previous work[1] to provide a systematic framework for the grouping of polycyclic aromatic hydrocarbons (PAH) according to their variant and invariant graph-theoretical properties, this paper now discusses the formula-structure relationships of PAHs containing tetragonal, pentagonal, heptagonal, and octagonal rings. Eventual goals of this work include computer enumeration of all the PAH structural isomers of a given molecular formula, property correlation studies, and the identification of new directions for possible research pursuance. This approach is based on formula-structure relationships. Another approach based on the hexagonal ring for determining formulas that would predict the number of possible PAH isomers was only partially successful[2] since the hexagonal ring frame-of-reference does not adequately define the invariants. For example, perinaphthalene ($C_{13}H_{10}$) and phenanthrene ($C_{14}H_{10}$) both have three hexagonal rings but different formulas. The restriction to hexagonal rings also precluded consideration of other ring sizes which is one of the thrusts of this current paper.

RESULTS AND DISCUSSION

A review of terminology and graph-theoretical relationships for totally fused PAHs will be presented first.

A glossary of terms used in this text is presented in Table I. The meaning of these terms may be illustrated by the

0095-2338/82/1622-0139$01.25/0 © 1982 American Chemical Society

Table I. Glossary of Terms

c = number of circulene holes
d_i = degree of vertex i of a graph
d_s = net tree disconnections (of internal graph edges)
N_c = total number of carbon atoms in a PAH
N_H = total number of hydrogen atoms in a PAH
N_{Ic} = number of internal carbon atoms in a PAH having a degree of 3
N_{Pc} = number of peripheral carbon atoms in PAH having a degree of 3
PAH6 = polycyclic aromatic hydrocarbon containing exclusively fused hexagonal rings
p_3 = number of graph points (vertices) having a degree of 3
q = number of graph edges (lines or C–C bonds)
q_I = number of internal graph edges
q_P = number of peripheral graph edges
q_b = number of graph edges connecting to phenyl-like ring substituents
r = number of rings
r_{5max} = maximum number of pentagonal rings
r_{4max} = maximum number of tetragonal rings
r_n = number of rings or cycles having n vertices

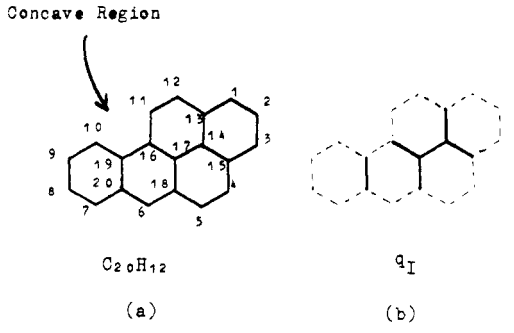


Figure 1. Graph of benzo[*a*]pyrene (a) and its internal edges, q_I (b).

molecular graph (only the σ C–C bonds will be shown throughout this paper) of benzo[*a*]pyrene ($C_{20}H_{12}$), a well-known carcinogen, shown in Figure 1a. Vertices 1–12 have a degree of 2 (i.e., have two incident lines) and are counted by N_H. Similarly, vertices 13–20 have a degree of 3 (p_3), where the vertices 13, 15, 16, and 18–20 are counted by N_{Pc} and vertices 14 and 17 by N_{Ic}. The total number of vertices 1–20 is counted by N_c and the total number of edges (bonds) by q. Figure 1b presents the internal edges, q_I, as bold lines and collectively is called a broken tree (forest) with a single disconnection (d_s) between points 16 and 19.

Table II presents a summary of equations that are applicable to totally fused PAHs and, in general, to graphs composed of a web of hexagons. Equation 1 states that the total number of vertices (N_c) is equal to the sum of the vertices of degree 3 (N_{Pc} and N_{Ic}) and vertices of degree 2 (N_H). Equation 3 relates the total number of edges (q) to the sum of internal (q_I) and peripheral (q_p) edges. Equation 2 can easily be proved by induction, and eq 4 and 5 can be algebraically derived from the other equations. Specialization of well-known theorems by Euler[3] lead to eq 6 and 8. If $G(p,q)$ is an unbroken tree graph of p vertices connected by q lines (e.g., the carbon skeleton of isooctane has $p = 8$ and $q = 7$), then $p = q + 1$. Definition of d_s as being the net number of disconnections (positive integers 1, 2, 3,...) or connections (negative integers $-1, -2, -3,...$) associated with the removal or addition of one or more lines (bonds), respectively, while a constant number of vertices (carbon atoms) are retained results in eq 7 for unbroken and broken trees (forests) which have been specialized for the internal edges of fused PAHs; Figure 1b is a broken tree (forest) with a single disconnection.

A comprehensive summary of all possible PAH6 formulas is provided by Table III. In Table III every compound in a particular column has the same d_s value, and every compound

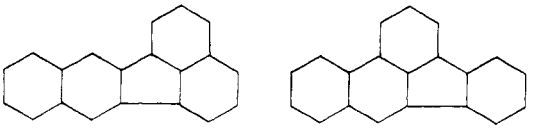
Table II. Equations Applicable to Totally Fused PAHs

no.	equation
1	$N_c = N_{Pc} + N_{Ic} + N_H$
2	$N_{Pc} = N_H - 6$
3	$q = q_I + q_p$
4	$q_P = q - q_I = 2N_H - 6$
5	$d_s = N_{Pc} - r$
6	$2q = \Sigma d_i = \Sigma d_2 + \Sigma d_3 = 2N_H + 3(N_c - N_H) = 3N_c - N_H$
7	$p_3 = q_I + 1 + d_s = N_c - N_H = N_{Pc} + N_{Ic}$
8	faces − edges + vertices = 2 $(r + 1) - q + p = 2$ or $r = q + 1 - N_c = (N_c + 2 - N_H)/2$
9	$r_{5max} \leq N_{Ic} = N_c - 2N_H + 6$
10	$r_{4max} \leq N_{ic}/2 = [N_c - 2N_H + 6]/2$

in the same row has the same N_{Ic} value. Member compounds in the same column with the same d_s belong to the same column series ($N_c = 3N_H - 14 - 2d_s$, where d_s is specified), and member compounds in the same row with the same N_{Ic} value belong to the same row series ($N_c = 2N_H - 6 + N_{Ic}$, where N_{Ic} is specified). For example, all PAH6 isomers with the formulas $C_{16}H_{10}$ and $C_{20}H_{12}$ have the same value of $N_{Ic} = 2$, and, similarly, all PAH6 isomers with the formulas $C_{20}H_{12}$ and $C_{26}H_{14}$ have the same value of $d_s = 1$. Table III extends infinitely in three directions, horizontally to the right, vertically to the bottom, and in a slanting direction to the left. Any formula found on Table III is defined as a PAH6 formula, and any structure composed of totally fused hexagonal rings is defined as a PAH6 structure and has a formula belonging to Table III.

In Table III, the $N_c = 2N_H - 6$ ($N_{Ic} = 0$) row series are the formulas of the cata-condensed PAH6. The formulas of the $N_c = 2N_H - 4$ and $N_c = 2N_H - 2$ row series correspond to peri-condensed PAH6. The corona-condensed PAH6s have formulas which range to the left of the $N_c = 3N_H - 14$ column series (i.e., belong to the column series $N_c = 3N_H - 12$, $N_c = 3N_H - 10,...$). Starting at the $N_c = 3N_H - 14$ column series but below the $N_c = 2N_H - 2$ row series, the Table III formulas on the right-hand side can be represented either by peri- or corona-condensed PAH6 structures. In a strict sense, the PAH6 structures corresponding to formulas located to the extreme right (and below the $N_c = 2N_H - 2$ row series) of Table III can be regarded as corona- or peri-condensed PAH6 compounds with cata-condensed branching. Thus Table III provides further definition of the cata-, peri-, and corona-condensed classification.[4]

Fused PAHs Containing Pentagonal Rings. Benzo[*b*]-fluoranthene and benzo[*k*]fluoranthene are isomers of ben-



zo[*a*]pyrene ($C_{20}H_{12}$) that contain a single pentagonal ring. Since the number of third degree vertices ($p_3 = N_{Ic} + N_{Pc}$ = constant) and the total number of edges (σ C–C bonds) must be constant, going from benzo[*a*]pyrene to either benzo[*b*]-fluoranthene or benzo[*k*]fluoranthene converts one six-membered ring to a five-membered ring while one internal carbon atom becomes a perpheral carbon atom with a corresponding increase in disconnections per $d_s = N_{Pc} - r$ and a decrease in the number of internal edges per $q_I + 1 + d_s = N_{Pc} + N_{Ic}$ = constant. Thus eq 2 must be modified to $N_{Pc} = N_H - 6 + r_5$ and eq 4 to $q_P = 2N_H - 6 + r_5$, where r_5 equals the number of five-membered rings possessed by the PAH. Since the cata-condensed PAH6s ($N_c = 2N_H - 6$ row series) have no internal third degree carbon atoms, no PAH having five-membered rings can exist for the formulas having a C/H

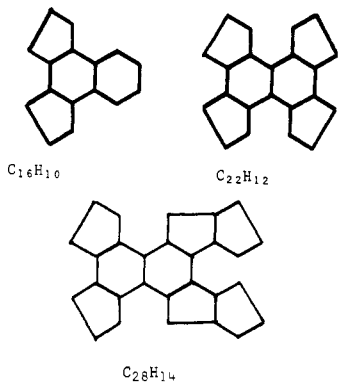


Figure 2. Some structures corresponding to PAH6 formulas in the $N_c = 3N_H - 14$ column series believed to possess the maximum number of possible five-membered (pentagonal) rings.

composition such that $N_c = 2N_H - 6$. In general, if all the internal third degree carbon atoms can be converted to peripheral third degree carbon atoms (i.e., $r_{5max} \leq N_{Ic}$ for a PAH6 formula) with corresponding simultaneous conversion of six-membered rings to five-membered rings, then the maximum number of five-membered rings (r_{5max}) that a corresponding PAH6 formula can possess is given by eq 9. There may be other structural and graphical constraints which limit the maximum number of five-membered rings that a particular PAH6 formula can possess to values below this upper limit. Figure 2 presents structures that contain the maximum number of five-membered rings for the PAH6 formulas from the $N_c = 3N_H - 14$ column series. Thus, the PAH6 formula for pyrene ($C_{16}H_{10}$) can be alternatively represented by a structure containing at most two five-membered rings. Similarly, the PAH6 formula for benzo[*g*,*h*,*i*]perylene ($C_{22}H_{12}$) can be alternatively represented by a structure having at most four five-membered rings. For the PAH6 formulas of $C_{16}H_{10}$ ($N_{Ic} = 2$ for PAH6 structures) and $C_{22}H_{12}$ ($N_{Ic} = 4$ for PAH6 structures) all the internal third degree carbon atoms possible have been transformed to peripheral third degree carbon atoms with the associated conversion of six-membered rings to five-membered rings in the structures of Figure 2 as predicted by the above equation for r_{5max}. The fused PAH6 isomers having the formula $C_{36}H_{18}$ contain six internal third degree carbon atoms ($N_{Ic} = 6$) that can be converted to peripheral third degree carbon atoms with the corresponding conversion of six six-membered rings to five-membered rings to give the following structure.

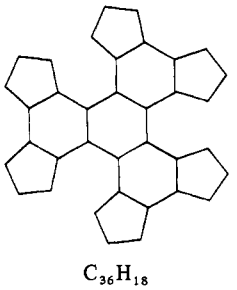


Table IV expands the formula PAH periodical table (Table III) into the diagonal region of the upper left-hand side where the formulas above the zigzag dashed line can be represented by structures that must contain one or more pentagonal rings. Many structures of these formulas can be derived by successive attachment of C=C units to the nonconcave regions of the diagonally located PAH6 structures having the same number of hydrogens. This is illustrated by derivatives of pyrene ($C_{16}H_{10}$) and 7-coronene ($C_{24}H_{12}$) in Figure 3. Pentalene (C_8H_6) has a formula placing it in the forbidden PAH6 formula region of the $N_c = 2N_H - 4$ row series, and eq 9 still

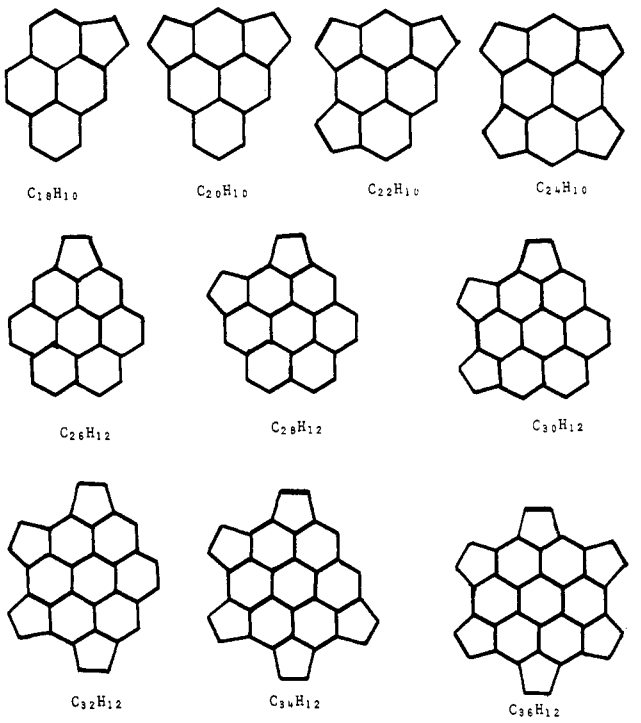


Figure 3. Structures containing pentagonal rings that belong in the formula region forbidden for PAH6.

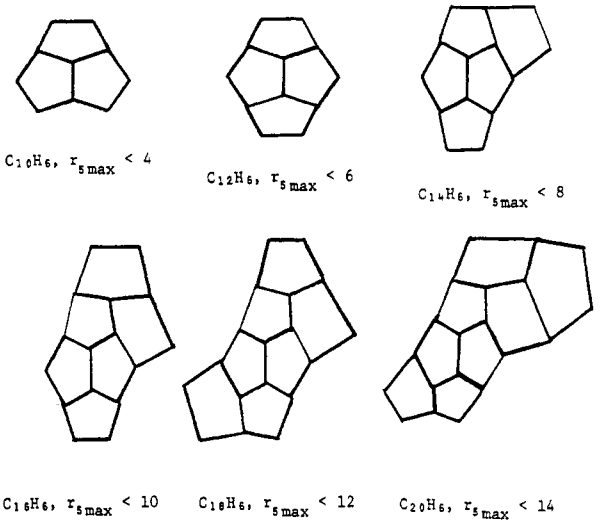
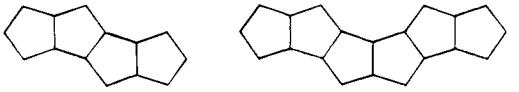


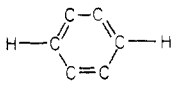
Figure 4. Structures of fused five-membered rings whose formulas define the upper diagonal boundary of Table IV and where the number of five-membered rings plus the number of internal carbons (N_{Ic}) is given by $r_{5max} \leq N_c - 2N_H + 6 = r_5 + N_{Ic}$.

correctly predicts that it can have at most only two pentagonal rings; similarly, the formulas $C_{14}H_8$ or $C_{20}H_{10}$ are predicted



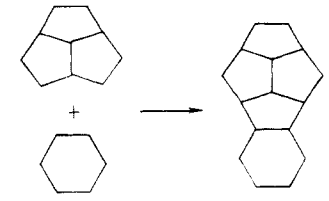
by eq 9 to have at most four or six pentagonal rings, respectively, and are consistent with the structures shown. Figure 4 presents structures composed of exclusively fused pentagonal rings whose formulas define the upper limit of the diagonal in Table IV. These structures in Figure 4 have internal carbons (N_{Ic}) that cannot be converted to peripheral carbons (N_{Pc}) simultaneous to the conversion of six-membered rings to five-membered rings because they contain no six-membered rings; eq 9 is still valid for the formulas of these structures and

is equal to the number of five-membered rings plus the number of internal carbons (N_{Ic}). Other formulas in Table IV can be recursively generated by attachment of I units to the structures

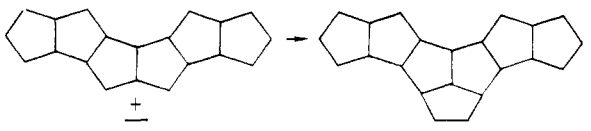


I

of Figure 4 or C=C units to the linear pentalene-related structures of the $N_c = 3N_H - 10$ column series ($C_{8+6n}H_{6+2n}$; $n = 0, 1, 2, \ldots$) as illustrated below.

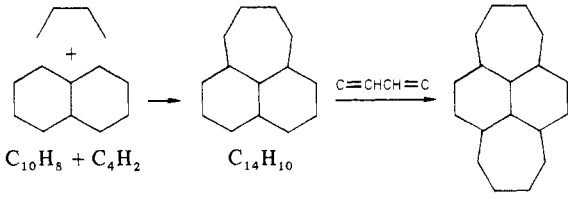


$C_{10}H_6 + C_6H_2$ $C_{16}H_8$ ($r_{5max} \leq 6$)



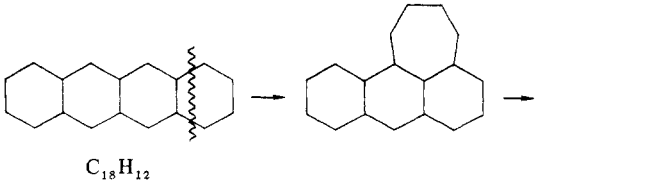
$C_{20}H_{10} + C_2$ $C_{22}H_{10}$ ($r_{5max} < 8$)

Fused PAHs Containing Heptagonal Rings. Seven-membered ring containing polycylic aromatic hydrocarbons having a formula corresponding to the PAH6 compounds of Table III can be generated by attaching C=CHCH=C units to nonconcave regions (i.e., only to three carbon atom segments as shown below) of PAH6 structures. This is illustrated with the attachment of C=CHCH=C units to naphthalene per the following to give PAH6 seven-membered-ring analogues

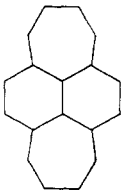


$C_{10}H_8 + C_4H_2$ $C_{14}H_{10}$ $C_{18}H_{12}$

of the $N_c = 2N_H - 6$ row series. Figure 5 presents similar analogues for the $N_c = 2N_H - 6$, $N_c = 2N_H - 4$ and $N_c = 2N_H - 2$ row series. An alternative construction of these seven-membered-ring analogues from PAH6 compounds would be simply transferring the C=CHCH=C units from a cata-condensed region to a nonconcave region as illustrated with naphthacene ($C_{18}H_{12}$) in the following.



$C_{18}H_{12}$



$C_{18}H_{12}$

In these examples for the first three-row series of Table III, the number of disconnections goes to zero and the number of

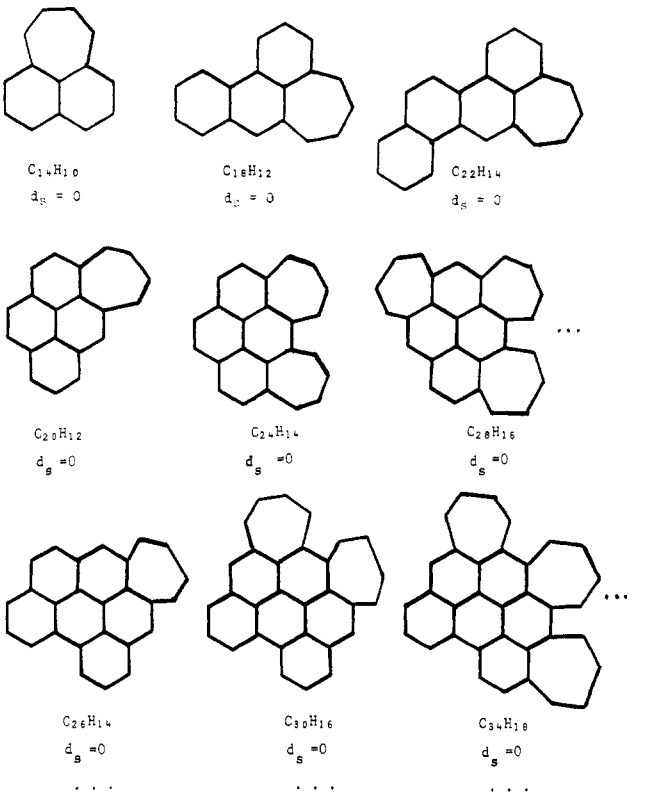


Figure 5. Structures containing heptagonal rings that belong to the $N_c = 2N_H - 6$, $N_c = 2N_H - 4$ and $N_c = 2N_H - 2$ row series.

internal carbons (N_{Ic}) increases in the conversion of the PAH6 structure to the analogue structure containing the maximum number of possible seven-membered rings (Figure 5); thus for these three-row series the maximum number of seven-membered rings that a structure corresponding to a PAH6 formula can possess is given by $r'_{7max} = d_s(\text{PAH6}) = [-N_c + 3N_H - 14]/2$.

For the row series $N_c = 2N_H$ and $N_c = 2N_H + 2$ in Table III, attachment of C=CHCH=C units leads to structures based on coronene having seven-membered rings and is illustrated in Figure 6. Again, as in the first three series discussed above, the number of disconnections decreases, but except for these lower row series it decreases to $d_s = -1$; thus, for the row series $N_c = 2N_H$ and $N_c = 2N_H + 2$, the maximum number of seven-membered rings that a structure corresponding to a PAH6 formula can possess is given by $r_{7max} = d_s(\text{PAH6}) + 1 = [-N_c + 3N_H - 12]/2$. This process can be continued so that, in general, $r_{7max} = d_s(\text{PAH6}) + M$, where $M = 0, 1, 2, \ldots$ corresponds to $|d_s|$ of the first member of the associated row series in Table III.

Pairwise attachment of CH=C—CH—C=CH units to the cata-condensed PAH6 structures of the $N_c = 2N_H - 6$ row series leads to the representative structures shown in Figure 7 which the author believes have the maximum number of heptagonal rings possible for these formulas where the only other ring size permitted is hexagonal. These structures in Figure 7 have formulas not contained in Table III but are present in the upper right-hand region of Table V, above the $N_c = 2N_H - 6$ series; the formulas of the cata-condensed even-carbon polycyclic heptagonal hydrocarbons ($C_{12+10n}H_{10+6n}$) in the top rows of Figure 7 define the upper diagonal edge of Table V. Note that the structures in Figure 7 have internal bonds that are totally disconnected ($d_s \geq 0$) and that as one goes to structures corresponding to formulas belonging to lower row series, the d_s becomes more negative. Figure 8 presents other structures formed by the pairwise attachment of CH=C—CH—C=CH units; note that this

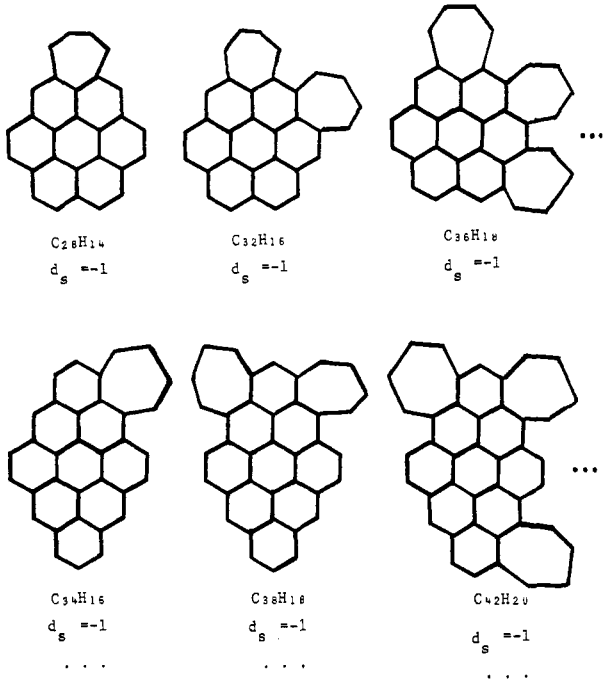


Figure 6. Structures containing heptagonal rings that belong to the $N_c = 2N_H$ and $N_c = 2N_H + 2$ row series.

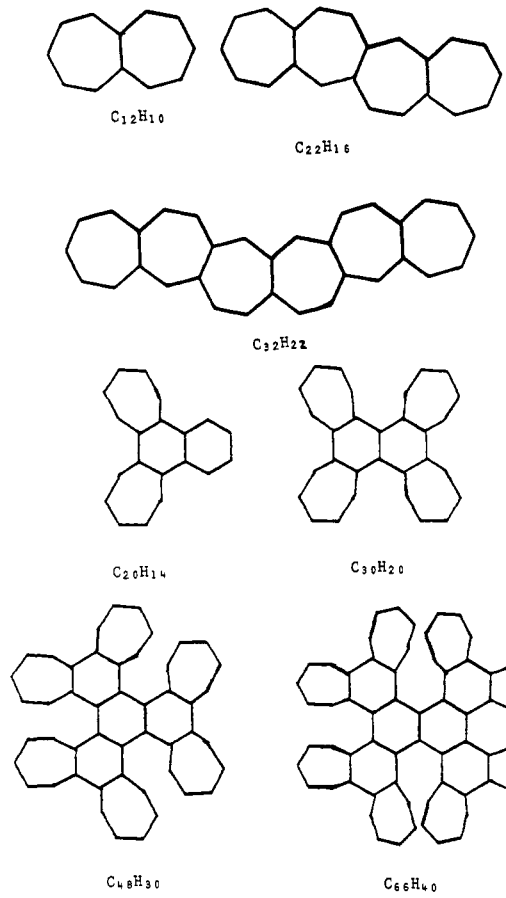


Figure 7. Structures with formulas above the $N_c = 2N_H - 6$ row series that are believed to possess the maximum number of heptagonal rings.

structures that have formulas one row series higher than the precursor formula; e.g., $C_{26}H_{16}$ belongs in the $N_c = 2N_H - 6$ row series and derives from the attachment of two C_5H_3 units to pyrene ($C_{16}H_{10}$), belonging to the $N_c = 2N_H - 4$ row series. Alternatively, pairwise attachment of C=CH—C· units to structures containing two or more concave regions (defined

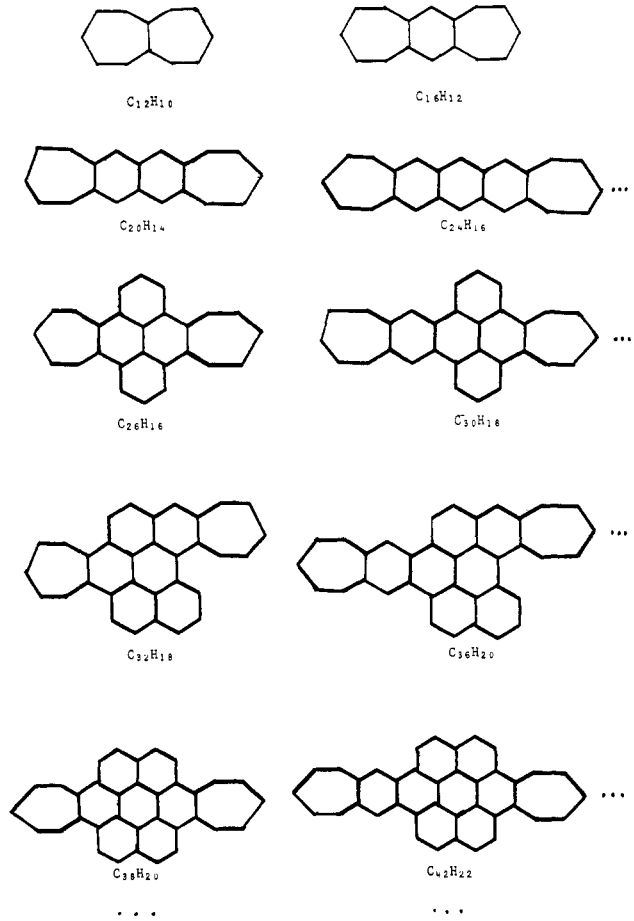
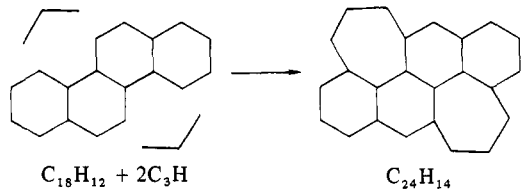


Figure 8. Structures that are derived from PAH6 structures by fusing on two heptagonal rings.

in Figure 1) lead to structures that have formulas one row series lower than the precursor formula as shown by the following.



Equations 1, 3, and 5–8 are still applicable without modification to PAHs having both hexagonal and heptagonal rings. Equations 2 and 4 must be modified by substracting r_7 (the number of seven-membered rings) from the right-hand side of these equations to give $N_{Pc} = N_H - 6 - r_7$ and $q_p = 2N_H - 6 - r_7$, respectively.

Azulenoid Fused PAH. Azulenoid PAHs can be generated from PAH6s by either of two methods. A simple 1,2 hydrogen migration can move the fusion juncture with the simultaneous formation of an adjacent five-membered ring to a seven-membered ring as illustrated below. Alternatively, a C=C

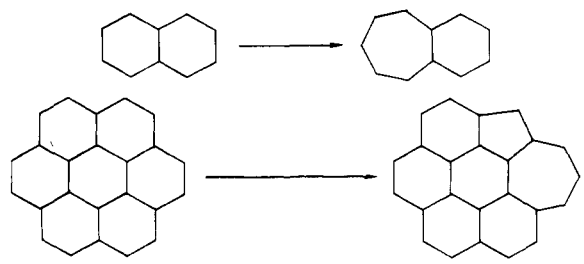


Table III. All the Formulas Possible for Polycyclic Aromatic Hydrocarbons Having Nonradical Hexagonal Rings PAH Periodic Table[a]

$\cdots N_c =$ $3N_H + 60$	$\cdots$	$N_c =$ $3N_H + 26$	$N_c =$ $3N_H + 24$	$N_c =$ $3N_H + 22$	$N_c =$ $3N_H + 20$	$N_c =$ $3N_H + 18$	$N_c =$ $3N_H + 16$	$N_c =$ $3N_H + 14$	$N_c =$ $3N_H + 12$	$N_c =$ $3N_H + 10$	$N_c =$ $3N_H + 8$	$N_c =$ $3N_H + 6$
												$C_{66}H_{20}$ $C_{72}H_{22}$ $C_{78}H_{24}$ ...
										$C_{74}H_{22}$ $C_{80}H_{24}$ $C_{86}H_{26}$ ...		
									$C_{76}H_{22}$ $C_{82}H_{24}$ $C_{88}H_{26}$ ...			
								$C_{78}H_{22}$ $C_{84}H_{24}$ $C_{90}H_{26}$ ...				
							$C_{80}H_{22}$ $C_{86}H_{24}$ $C_{92}H_{26}$ ...					
						$C_{88}H_{24}$ $C_{94}H_{26}$ $C_{100}H_{28}$ ...						
					$C_{90}H_{24}$ $C_{96}H_{26}$ $C_{102}H_{28}$ ...							
				$C_{92}H_{24}$ $C_{98}H_{26}$ $C_{108}H_{28}$ ...								
			$C_{94}H_{24}$ $C_{100}H_{26}$ $C_{106}H_{28}$ ...									
		$C_{96}H_{24}$ $C_{102}H_{26}$ $C_{108}H_{28}$ $C_{114}H_{30}$ $C_{120}H_{32}$ $C_{126}H_{34}$ $C_{132}H_{36}$ ...										
	$C_{104}H_{26}$ $C_{110}H_{28}$ $C_{116}H_{30}$ ...											
...												
$\cdots C_{150}H_{30}$ ...												

[a] $N_c = 3N_H - 14 - 2d_s$ and $N_c = 2N_H - 6 + N_{Ic}$ are the general formulas for the column series and the row series, respectively.

Table IV. Formulas of PAHs Containing Pentagonal Rings in Addition to Hexagonal Rings

$\cdots N_c = 3N_H$	$N_c =$ $3N_H - 2$	$N_c =$ $3N_H - 4$	$N_c =$ $3N_H - 6$	$N_c =$ $3N_H - 8$	$N_c =$ $3N_H - 10$	$N_c =$ $3N_H - 12$	$N_c = 3N_H - 14$	series
					C_8H_6	$C_{12}H_8$	$C_{16}H_{10}\cdots$	$N_c = 2N_H - 4$
				$C_{10}H_6$	$C_{14}H_8$	$C_{18}H_{10}$	$C_{22}H_{12}\cdots$	$N_c = 2N_H - 2$
			$C_{12}H_6$	$C_{16}H_8$	$C_{20}H_{10}$	$C_{24}H_{12}$	$C_{28}H_{14}\cdots$	$N_c = 2N_H$
		$C_{14}H_6$	$C_{18}H_8$	$C_{22}H_{10}$	$C_{26}H_{12}$	$C_{30}H_{14}$	$C_{34}H_{16}\cdots$	$N_c = 2N_H + 2$
	$C_{16}H_6$	$C_{20}H_8$	$C_{24}H_{10}$	$C_{28}H_{12}$	$C_{32}H_{14}$	$C_{36}H_{16}$	$C_{40}H_{18}\cdots$	$N_c = 2N_H + 4$
$\cdots C_{18}H_6$	$C_{22}H_8$	$C_{26}H_{10}$	$C_{30}H_{12}$	$C_{34}H_{14}$	$C_{38}H_{16}$	$C_{42}H_{18}$	$C_{46}H_{20}\cdots$	$N_c = 2N_H + 6$
$C_{24}H_8$	$C_{28}H_{10}$	$C_{32}H_{12}$	$C_{36}H_{14}$	$C_{40}H_{16}$	$C_{44}H_{18}$	...	$C_{52}H_{22}\cdots$	$N_c = 2N_H + 8$
$C_{30}H_{10}$	$C_{34}H_{12}$	$C_{38}H_{14}$	$C_{42}H_{16}$	$C_{46}H_{18}$	$C_{50}H_{20}$		$C_{58}H_{24}\cdots$	$N_c = 2N_H + 10$
$C_{36}H_{12}$	$C_{40}H_{14}$	$C_{44}H_{16}$	$C_{48}H_{18}$	$C_{52}H_{20}$	...		$C_{64}H_{26}\cdots$	$N_c = 2N_H + 12$
$C_{42}H_{14}$	$C_{46}H_{16}$	$C_{50}H_{18}$	$C_{54}H_{20}$	...			$C_{70}H_{28}\cdots$	$N_c = 2N_H + 14$
$C_{48}H_{16}$	$C_{52}H_{18}$	$C_{56}H_{20}$	...				$C_{76}H_{30}\cdots$	$N_c = 2N_H + 16$
$C_{54}H_{18}$	$C_{58}H_{20}$	$C_{62}H_{22}$					$C_{82}H_{32}\cdots$	$N_c = 2N_H + 18$
$C_{60}H_{20}$	$C_{64}H_{22}$	...					$C_{88}H_{34}\cdots$	$N_c = 2N_H + 20$
...	...						...	...

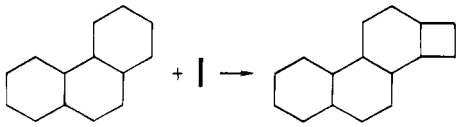
unit can be detached from a PAH6 structure to form a concave region and reattached elsewhere to form a five-membered ring followed by detaching a C=CH—C unit and reattaching it to the concave region formed in the first step or detaching and attaching a C=CHCH=C unit per the Scheme I.

From the previous discussion on formation pentagonal and heptagonal containing PAH compounds, one can generalize eq 2 and 4 for totally fused PAHs to be $N_{Pc} = N_H - 6 + r_5 - r_7$ and $q_P = 2N_H - 6 + r_5 - r_7$, respectively, where the plus sign before r_5 corresponds to an increase in the number of peripheral third degree carbon sites (i.e., a decrease in $N_{Ic} = p_3 - N_{Pc}$) and the minus sign before r_7 corresponds to a decrease in the number of peripheral third degree carbon sites (i.e., an increase in $N_{Ic} = p_3 - N_{Pc}$). Thus, polycylic aromatic

$N_c = 3N_H + 4$	$N_c = 3N_H + 2$	$N_c = 3N_H$	$N_c = 3N_H - 2$	$N_c = 3N_H - 4$	$N_c = 3N_H - 6$	$N_c = 3N_H - 8$	$N_c = 3N_H - 10$	$N_c = 3N_H - 12$	$N_c = 3N_H - 14$	$N_c = 3N_H - 16$	series
									$C_{10}H_8$	$C_{14}H_{10}\cdots$	$N_c = 2N_H - 6$
									$C_{16}H_{10}$	$C_{20}H_{12}\cdots$	$N_c = 2N_H - 4$
									$C_{22}H_{12}$	$C_{26}H_{14}\cdots$	$N_c = 2N_H - 2$
								$C_{24}H_{12}$	$C_{28}H_{14}$	$C_{32}H_{16}\cdots$	$N_c = 2N_H$
								$C_{30}H_{14}$	$C_{34}H_{16}$	$C_{38}H_{18}\cdots$	$N_c = 2N_H + 2$
							$C_{32}H_{14}$	$C_{36}H_{16}$	$C_{40}H_{18}$	$C_{44}H_{20}\cdots$	$N_c = 2N_H + 4$
							$C_{38}H_{16}$	$C_{42}H_{18}$	$C_{46}H_{20}$	$C_{50}H_{22}\cdots$	$N_c = 2N_H + 6$
							$C_{44}H_{18}$	...	$C_{52}H_{22}$	$C_{56}H_{24}\cdots$	$N_c = 2N_H + 8$
							$C_{50}H_{20}$		$C_{58}H_{24}$	$C_{62}H_{26}\cdots$	$N_c = 2N_H + 10$
							...		$C_{64}H_{26}$	$C_{68}H_{28}\cdots$	$N_c = 2N_H + 12$
									$C_{70}H_{28}$	$C_{74}H_{30}\cdots$	$N_c = 2N_H + 14$
									$C_{76}H_{30}$	$C_{80}H_{32}\cdots$	$N_c = 2N_H + 16$
									$C_{82}H_{32}$	$C_{86}H_{34}\cdots$	$N_c = 2N_H + 18$
									$C_{88}H_{34}$	$C_{92}H_{36}\cdots$	$N_c = 2N_H + 20$
									$C_{94}H_{36}$	$C_{98}H_{38}\cdots$	$N_c = 2N_H + 22$
									$C_{100}H_{38}$	$C_{104}H_{40}\cdots$	$N_c = 2N_H + 24$
									$C_{106}H_{40}$	$C_{110}H_{42}\cdots$	$N_c = 2N_H + 26$
									$C_{112}H_{42}$	$C_{116}H_{44}\cdots$	$N_c = 2N_H + 28$
									...	...	$N_c = 2N_H + 30$
											$N_c = 2N_H + 32$
											$N_c = 2N_H + 34$
											$N_c = 2N_H + 36$
											$N_c = 2N_H + 38$
											$N_c = 2N_H + 40$
											$N_c = 2N_H + 42$
											$N_c = 2N_H + 44$
											$N_c = 2N_H + 46$
											$N_c = 2N_H + 48$
											$N_c = 2N_H + 50$
											$N_c = 2N_H + 52$
											$N_c = 2N_H + 54$
											$N_c = 2N_H + 56$
											$N_c = 2N_H + 58$
											$N_c = 2N_H + 60$
											$N_c = 2N_H + 62$
											$N_c = 2N_H + 64$
											$N_c = 2N_H + 66$
											$N_c = 2N_H + 68$
											$N_c = 2N_H + 70$
											$N_c = 2N_H + 72$
											$N_c = 2N_H + 74$
											$N_c = 2N_H + 76$
											...
											$N_c = 2N_H + 90$
											...
$C_{64}H_{20}$	$C_{62}H_{20}$	$C_{54}H_{18}$	$C_{52}H_{18}$	$C_{50}H_{18}$	$C_{42}H_{16}$	$C_{40}H_{16}$					
$C_{70}H_{22}$	$C_{68}H_{22}$	$C_{60}H_{20}$	$C_{58}H_{20}$	$C_{56}H_{20}$	$C_{48}H_{18}$	$C_{46}H_{18}$					
$C_{76}H_{24}$	$C_{74}H_{24}$	$C_{66}H_{22}$	$C_{64}H_{22}$	$C_{62}H_{22}$	$C_{54}H_{20}$	$C_{52}H_{20}$					
...	...	$C_{72}H_{24}$	$C_{70}H_{24}$	...	...	...					
		$C_{78}H_{26}$	$C_{76}H_{26}$								
		$C_{84}H_{28}$	...								
		$C_{90}H_{30}$									
		...									

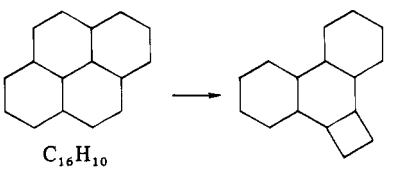
hydrocarbons composed of fused hexagonal rings and/or fused azulene units[5] (i.e., PAHs that contain an equal number of five- and seven-membered rings) will obey the same equations in Table II (except the last equation) and will have the same formula PAH periodic table (Table III) as do the PAH6 structures since r_5 and r_7 cancel each other out.

Fused PAHs Containing Tetragonal Rings. Tetragonal ring containing PAH structures can be recursively generated from PAH6 structures by two basic methods. A C=C unit can be attached to any ethylene edge (-CH=CH-) of a PAH6 structure as shown by the following example with phenanthrene. This process produces a formula one row series lower

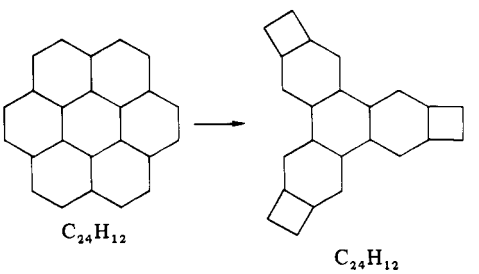


$C_{14}H_{10} + C_2$ $C_{16}H_{10}$

and one column series to the left of the precursor PAH6 formula in Table III. Other examples are presented in Figures 9 and 10. Alternatively, a C=C unit can be detached and reattached to a ethylene edge as illustrated by the following C=C unit transfer in pyrene and 7-coronene. Similarly the

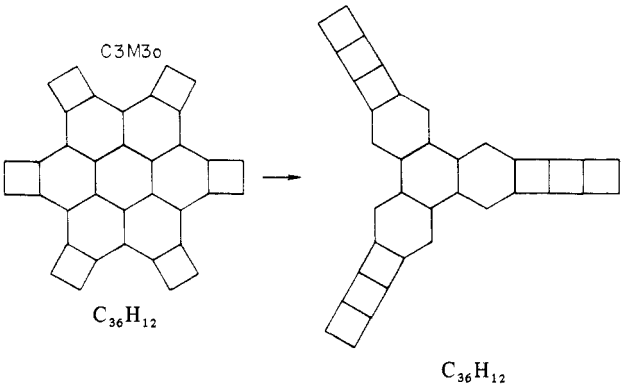


$C_{16}H_{10}$ $C_{16}H_{10}$



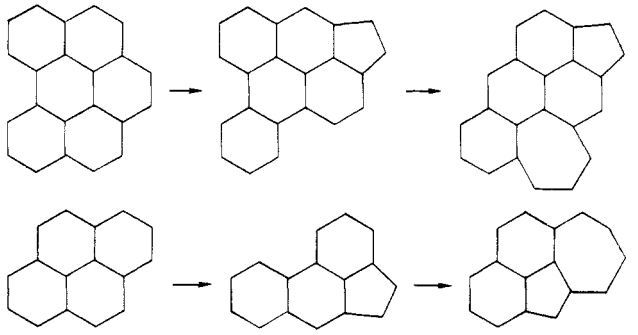
$C_{24}H_{12}$ $C_{24}H_{12}$

$C_{36}H_{12}$ structure in Figure 10 can be converted by a similar process to give a structure having nine tetragonal rings ($C_{36}H_{12}$). In these latter transformations, with each hexagonal ring interconverted to a tetragonal ring two of the internal third degree carbons (N_{Ic}) were simultaneously converted to two peripheral third degree carbons (N_{Pc}). Since the number of

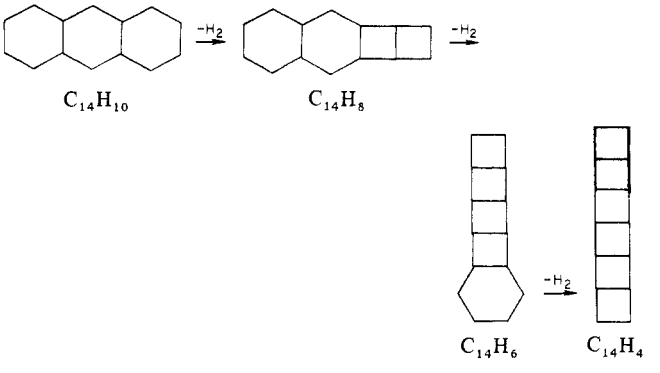


third degree vertices ($p_3 = N_{Ic} + N_{Pc}$ = constant) and the total number of edges (σ C–C bonds) must be constant in any conversion of a PAH6 structure to a PAH structure containing tetragonal rings, the number of disconnections ($d_s = N_{Pc} - r$) increases while the number of internal edges ($q_I + 1 + d_s = N_{Pc} + N_{Ic}$ = constant) decreases. Furthermore, these examples demonstrate two generalizations. First, since the cata-condensed PAH6 have no internal carbons ($N_{Ic} = 0$) for simultaneous conversion to peripheral carbons of third degree with the interconversion of hexagonal rings to tetragonal rings and attachment of C=C units to them give a formula belonging one row series lower, there can be no PAH structures containing tetragonal rings that have a formula ($C_{10+4n}H_{8+2n}$; n = 0, 1, 2,...) corresponding to the cata-condensed PAH6s. Second, if all the internal carbons (internal vertices of degree 3) can be converted to peripheral carbons with simultaneous interconversion of hexagonal rings to tetragonal rings, then the maximum number of four-membered rings (r_{4max}) in addition to six-membered rings that a structure corresponding to a PAH formula can possess is given by $r_{4max} = N_{Ic}/2 \leq [N_c - 2N_H + 6]/2$. This conjecture is analogous to the one made previously for the maximum number of five-membered rings ($r_{5max} = N_{Ic} \leq [N_c - 2N_H + 6]$) that a PAH structure can contain.

Scheme I



Scheme II



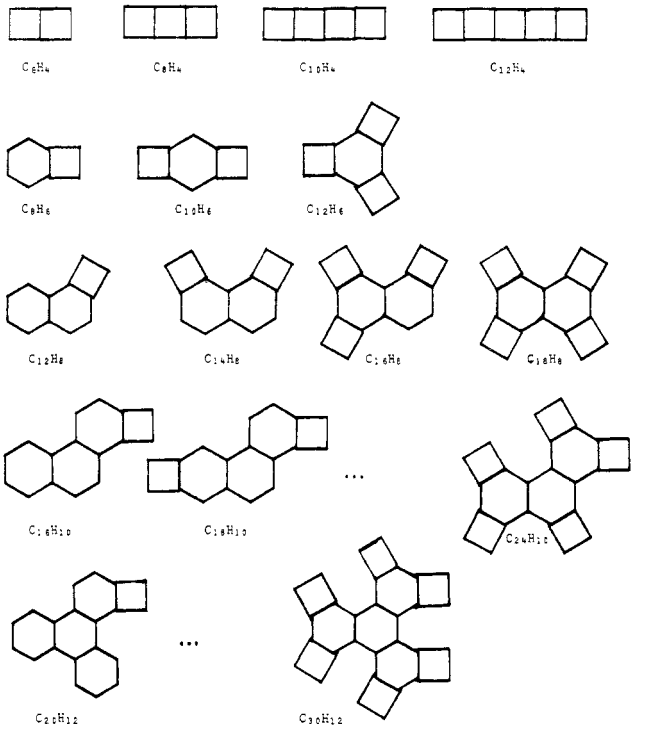


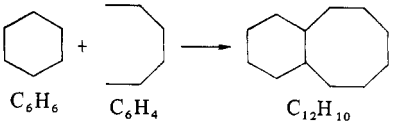
Figure 9. Structures possessing tetragonal rings.

The possible formulas of PAH containing tetragonal rings in addition to hexagonal rings is presented in Table VI. Note that the upper left-hand diagonal of formulas in Table VI is defined by $C_{6+n}H_4$ (where n = 0, 1, 2,...), and the formulas are the fused cata-condensed polycyclic antiaromatic hydrocarbons shown in the first row of Figure 9. One way that these cata-condensed polycyclic antiaromatic hydrocarbons can be formally generated is by successive dehydrogenation of a linear cata-condensed PAH6 as in Scheme II.

Equations 1, 3, and 5–8 in Table II are still valid for tetragonal ring containing PAHs. However, eq 2 and 4 must be modified by adding $2r_4$ to the right-hand side so that they become $N_{Pc} = N_H - 6 + 2r_4$ and $q_P = 2N_H - 6 + 2r_4$, respectively.

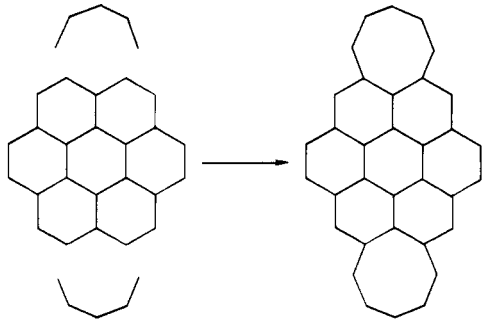
PAHs containing polyalkenyl side chains can be generated by appropriately attaching C_2 units to PAH6 structures. If a C_2 unit is attached to a PAH6 structure, then the formula produced is shifted one row series lower and one column series to the left of the precursor formula. Successive attachment of more than one C_2 unit results in a formula table nearly identical to Table VI but with the diagonal formulas having four and six hydrogens (i.e., C_nH_4 and C_nH_6) omitted.

Octagonal Ring Containing PAHs. Attachment of C=C-HCH=CHCH=C units to an ethylene edge of a PAH6 structure leads to an octagonal ring containing PAH with a formula one row series higher and three column series to the right of the precursor PAH6 formula. The attachment of a C=CHCH=CHCH=C unit to benzene and naphthalene is illustrated by the following compounds belonging to the N_c

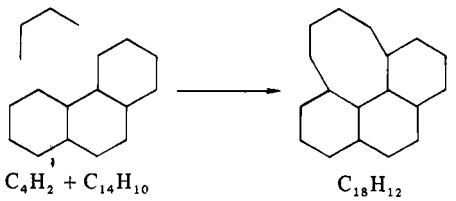


$= 2N_H - 8$ row series in Table VII. The formulas above the dashed line in Table VII cannot be represented by a PAH6 structure whereas those below can. The cata-condensed polycyclic antiaromatic hydrocarbons shown in the first row of Figure 11 defines the upper diagonal formula boundary above the dashed line in Table VII. Other examples of octagonal ring containing PAHs generated by this process is shown in Figures 11 and 12.

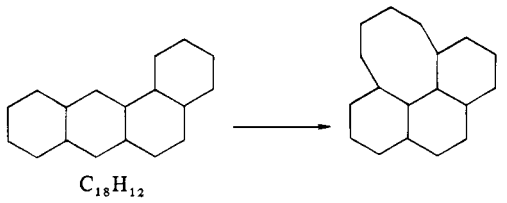
Pairwise attachment of two C=CH—ĊH—CH=C radical units to PAH6 structures leads to PAHs having two octagonal rings as illustrated with 7-coronene.



Transformation of PAH6 structures into octagonal ring containing PAHs can be achieved by two limited processes. Structures possessing a four-carbon concave region (defined in Figure 1) can accept the attachment of a C=CHCH=C unit as follows.



The other process involves the interconversion of a PAH6 structure via detachment and reattachment of a C=CHCH=C unit as shown for $C_{18}H_{12}$. This latter example suggests



that PAHs containing octagonal rings with a corresponding PAH6 formula will give a structure with more internal third degree carbons and less disconnections than the analogous PAH6 structure as observed with heptagonal ring containing PAHs.

The formulas in Table II are valid for PAHs containing octagonal and hexagonal rings except that eq 2 and 4 must be modified by subtracting $2r_8$ from the right-hand side of each to give $N_{Pc} = N_H - 6 - 2r_8$ and $q_P = 2N_H - 6 - 2r_8$, respectively.

PAHs Containing an Equal Number of Tetragonal and Octagonal Rings. Figure 13 presents a number of PAHs containing an equal number of tetragonal and octagonal rings. It becomes evident from these structures and others that the absence of hexagonal rings can occur only in those formulas

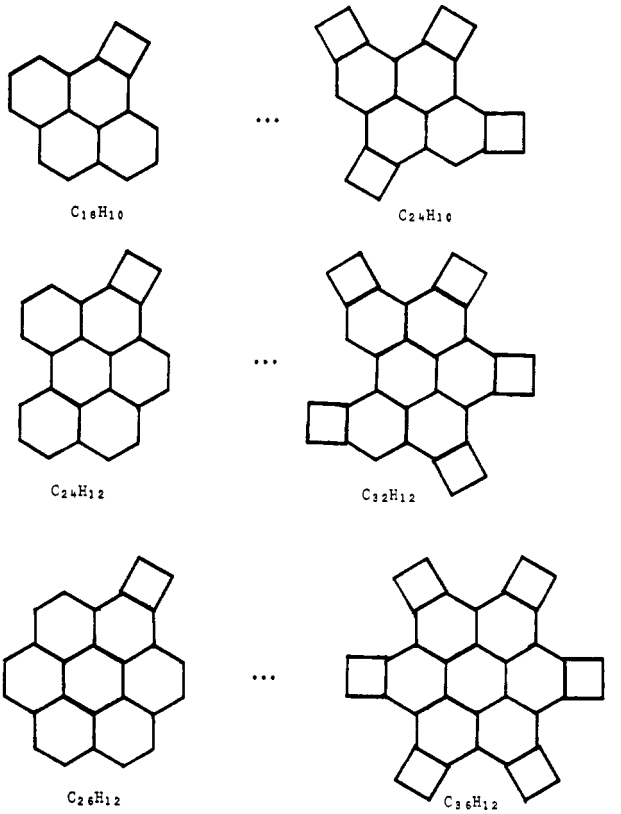


Figure 10. PAH structures possessing tetragonal rings generated by attachment of C=C units to PAH6 compounds.

belonging to column series having an even number of disconnections (d_8). From the previous sections on the tetragonal and octagonal ring containing PAHs, one can generalize eq 2 and 4 for fused PAHs possessing tetragonal, hexagonal, and octagonal rings to obtain $N_{Pc} = N_H - 6 + 2r_4 - 2r_8$ and $q_P = 2N_H - 6 + 2r_4 + 2r_8$, respectively. Thus, if PAH structures have an equal number of tetragonal and octagonal rings, r_4 and r_7 cancel each other out, and these structures will obey the same equations in Table II and will have the same formula PAH periodic table (Table III) as do the PAH6 structures. These results are analogous to those reached for azulenoids.

PAHs Containing a Mix of Ring Sizes. From the previous discussions on PAHs containing an equal number of tetragonal and octagonal rings or an equal number of pentagonal and heptagonal rings in addition to hexagonal rings where it was shown that the PAH6 table (Table III) and the PAH6 equations were equally applicable, one can now conjecture the following. Table VI may also be applicable for PAHs having one more tetragonal ring than the number of octagonal rings or for PAHs having an equal number of pentagonal and heptagonal rings and one more tetragonal ring than the number of octagonal rings, etc. Similarly, other analogous combinations of ring sizes may result in the applicability of Table VII or the corresponding tables for pentagonal and heptagonal ring containing PAHs, depending upon which ring size is in excess over its negating counterpart. For PAHs having two pentagonal rings and one octagonal ring, the r_5 and r_8 terms in eq 2 and 4 cancel each other out, and Table III becomes applicable; Figure 14 gives sample PAH structures consistent with this. Similarly, for PAHs having two tetragonal rings and one tetragonal ring, the r_7 and r_4 terms in eq 2 and 4 cancel, and Table III becomes applicable; consistent PAH structures are presented in Figure 14. Full confirmation of this conjecture awaits computer enumeration.

Phenyl-Substituted Fused PAHs. If phenyl-substituted PAH6s are considered, then the formula PAH periodic table (Table III) becomes expanded at the upper right-hand region above the $N_c = 2N_H - 6$ row series. This expansion is given

Table V. Formulas of Heptagonal Ring Containing PAHs

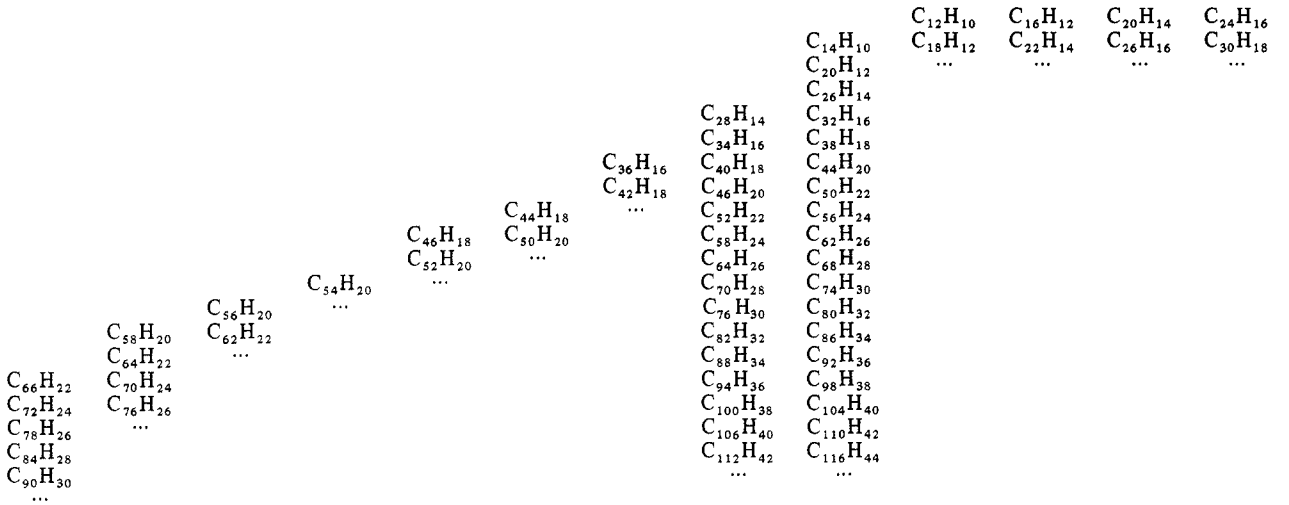
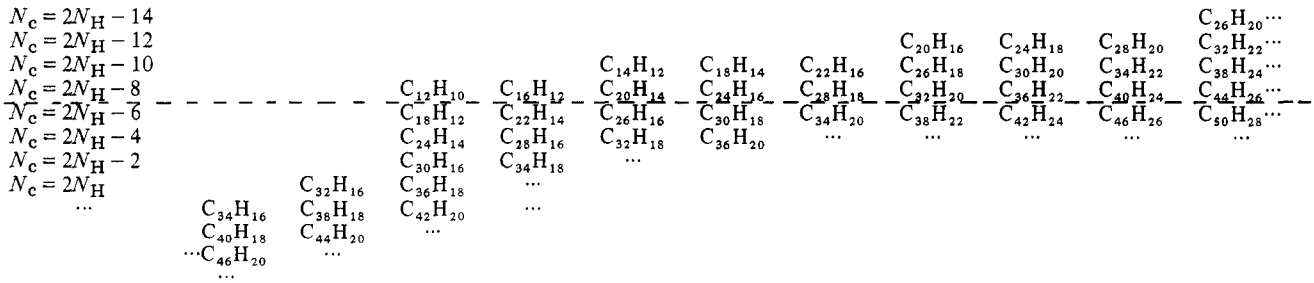


Table VI. Formulas of PAHs Containing Tetragonal Rings in Addition to Hexagonal Rings

$\cdots N_c = 3N_H$	$N_c = 3N_H - 2$	$N_c = 3N_H - 4$	$N_c = 3N_H - 6$	$N_c = 3N_H - 8$	$N_c = 3N_H - 10$	$N_c = 3N_H - 12$	$N_c = 3N_H - 14$	series
$\cdots C_{12}H_4$	$C_{10}H_4$	C_8H_4	C_6H_4	$C_{10}H_6$	C_8H_6	$C_{12}H_8$	$C_{16}H_{10}\cdots$	$N_c = 2N_H - 4$
$C_{18}H_6$	$C_{16}H_6$	$C_{14}H_6$	$C_{12}H_6$	$C_{16}H_8$	$C_{14}H_8$	$C_{18}H_{10}$	$C_{22}H_{12}\cdots$	$N_c = 2N_H - 2$
$C_{24}H_8$	$C_{22}H_8$	$C_{20}H_8$	$C_{18}H_8$	$C_{22}H_{10}$	$C_{20}H_{10}$	$C_{24}H_{12}$	$C_{28}H_{14}\cdots$	$N_c = 2N_H$
$C_{30}H_{10}$	$C_{28}H_{10}$	$C_{26}H_{10}$	$C_{24}H_{10}$	$C_{28}H_{12}$	$C_{26}H_{12}$	$C_{30}H_{14}$	$C_{34}H_{16}\cdots$	$N_c = 2N_H + 2$
$C_{36}H_{12}$	$C_{34}H_{12}$	$C_{32}H_{12}$	$C_{30}H_{12}$	$C_{34}H_{14}$	$C_{32}H_{14}$	$C_{36}H_{16}$	$C_{40}H_{18}\cdots$	$N_c = 2N_H + 4$
$C_{42}H_{14}$	$C_{40}H_{14}$	$C_{38}H_{14}$	$C_{36}H_{14}$	$C_{40}H_{16}$	$C_{38}H_{16}$	$C_{42}H_{18}$	$C_{46}H_{20}\cdots$	$N_c = 2N_H + 6$
$C_{48}H_{16}$	$C_{46}H_{16}$	$C_{44}H_{16}$	$C_{42}H_{16}$	$C_{46}H_{18}$	$C_{44}H_{18}$	...	$C_{52}H_{22}\cdots$	$N_c = 2N_H + 8$
$C_{54}H_{18}$	$C_{52}H_{18}$	$C_{50}H_{18}$	$C_{48}H_{18}$	$C_{52}H_{20}$	$C_{50}H_{20}$		$C_{58}H_{24}\cdots$	$N_c = 2N_H + 10$
$C_{60}H_{20}$	$C_{58}H_{20}$	$C_{56}H_{20}$	$C_{54}H_{20}$	...	...		$C_{64}H_{26}\cdots$	$N_c = 2N_H + 12$
...	$C_{64}H_{22}$	$C_{62}H_{22}$	...				$C_{70}H_{28}\cdots$	$N_c = 2N_H + 14$
	...	...					$C_{76}H_{30}\cdots$	$N_c = 2N_H + 16$
							$C_{82}H_{32}\cdots$	$N_c = 2N_H + 18$
							$C_{88}H_{34}\cdots$	$N_c = 2N_H + 20$
							...	...

Table VII. Formulas of Octagonal Rings Containing PAHs



in Table VIII. The formulas underlined in Table VI are polyphenyl chains like biphenyl ($C_{12}H_{10}$), terphenyl ($C_{18}H_{14}$), etc.; these polyphenyl formulas define the upper left boundary for this table. The first row ($N_c = 2N_H - 8$ row series) above the $N_c = 2N_H - 6$ row series (above the dashed line) can be generated by adding a phenylene group (a C_6H_4 unit) to each member of the $N_c = 2N_H - 6$ row series; e.g., adding a phenylene unit to any isomer of $C_{18}H_{12}$ gives a $C_{24}H_{16}$ compound, like 1-phenyltetracene. Similarly, adding two phenylene groups to each member of the $N_c = 2N_H - 6$ row series gives the $N_c = 2N_H - 10$ row series, adding three phenylene groups gives the $N_c = 2N_H - 12$ row series, and so forth. Formally, biphenylene is formed by adding a phenylene unit to benzene. Tetraphenylene has a formula of $C_{24}H_{16}$. Adding one phenylene unit to benzo[*g*,*h*,*i*]perylene or its isomer ($C_{22}H_{12}$) gives $C_{28}H_{16}$, adding another C_6H_4 unit gives $C_{34}H_{20}$ which isomeric to octacene, adding still another phenylene unit gives $C_{40}H_{24}$, and finally adding enough phenylene units to totally substitute all the hydrogens (12 phenyl groups) of benzo[*g*,*h*,*i*]perylene gives $C_{100}H_{64}$. The formula of $C_{28}H_{18}$ can be generated by

								$C_{42}H_{28}$	$C_{46}H_{30}$	...
				$C_{32}H_{22}$	$C_{36}H_{24}$	$C_{40}H_{26}$	$C_{44}H_{28}$	$C_{48}H_{30}$	$C_{52}H_{32}$	$C_{56}H_{34}$
$C_{22}H_{16}$	$C_{26}H_{18}$	$C_{30}H_{20}$	$C_{34}H_{22}$	$C_{38}H_{24}$	$C_{42}H_{26}$	$C_{46}H_{28}$	$C_{50}H_{30}$	$C_{54}H_{32}$	...	...
$C_{28}H_{18}$	$C_{32}H_{20}$	$C_{36}H_{22}$	$C_{40}H_{24}$	$C_{44}H_{26}$	$C_{48}H_{28}$	$C_{52}H_{30}$	$C_{56}H_{32}$	$C_{60}H_{34}$	...	
$C_{34}H_{20}$	$C_{38}H_{22}$	$C_{42}H_{24}$	$C_{46}H_{26}$	$C_{50}H_{28}$	$C_{54}H_{30}$	$C_{58}H_{32}$	$C_{62}H_{34}$	$C_{66}H_{36}$	...	
...	...	...	...	...	...	...	...	...		

$N_c = 2N_H - 14$
$N_c = 2N_H - 12$
$N_c = 2N_H - 10$
$N_c = 2N_H - 8$
$N_c = 2N_H - 6$
$N_c = 2N_H - 4$
$N_c = 2N_H - 2$
$N_c = 2N_H$
$N_c = 2N_H + 2$
$N_c = 2N_H + 4$
$N_c = 2N_H + 6$
$N_c = 2N_H + 8$
$N_c = 2N_H + 10$
$N_c = 2N_H + 12$
$N_c = 2N_H + 14$
$N_c = 2N_H + 16$
$N_c = 2N_H + 18$
$N_c = 2N_H + 20$
$N_c = 2N_H + 22$
$N_c = 2N_H + 24$
$N_c = 2N_H + 26$
$N_c = 2N_H + 28$
$N_c = 2N_H + 30$
...

Table VIII. Formulas of Phenyl-Substituted PAH6s

...													
$N_c = 2N_H - 14$												$C_{30}H_{22}$[a]	$C_{34}H_{24}$...
$N_c = 2N_H - 12$											$C_{24}H_{18}$[b]	$C_{28}H_{20}$	$C_{32}H_{22}$	$C_{36}H_{24}$	$C_{40}H_{26}$...
$N_c = 2N_H - 10$							$C_{18}H_{14}$[c]	$C_{22}H_{16}$	$C_{26}H_{18}$	$C_{30}H_{20}$	$C_{34}H_{22}$	$C_{38}H_{24}$	$C_{42}H_{26}$	$C_{46}H_{28}$...
$N_c = 2N_H - 8$				$C_{12}H_{10}$[d]	$C_{16}H_{12}$	$C_{20}H_{14}$	$C_{24}H_{16}$	$C_{28}H_{18}$	$C_{32}H_{20}$	$C_{36}H_{22}$	$C_{40}H_{24}$	$C_{44}H_{26}$	$C_{48}H_{28}$	$C_{52}H_{30}$...
$N_c = 2N_H - 6$					$C_{22}H_{14}$	$C_{26}H_{16}$	$C_{30}H_{18}$	$C_{34}H_{20}$	$C_{38}H_{22}$	$C_{42}H_{24}$	$C_{46}H_{26}$	$C_{50}H_{28}$	$C_{54}H_{30}$	$C_{58}H_{32}$...
$N_c = 2N_H - 4$					$C_{28}H_{16}$	...	...	...	...	...	...	...	...	...
$N_c = 2N_H - 2$				$C_{30}H_{16}$	$C_{34}H_{18}$	...								
$N_c = 2N_H$				$C_{36}H_{18}$	$C_{40}H_{20}$	...								
$N_c = 2N_H + 2$			$C_{38}H_{18}$	$C_{42}H_{20}$	$C_{46}H_{22}$	...								
...			$C_{44}H_{20}$	$C_{48}H_{22}$	$C_{52}H_{24}$	...								
		...$C_{46}H_{20}$	$C_{50}H_{22}$	$C_{54}H_{24}$	...								
		...	...	...									

[a] Tetraphenyl. [b] Triphenyl. [c] Diphenyl. [d] Monophenyl.

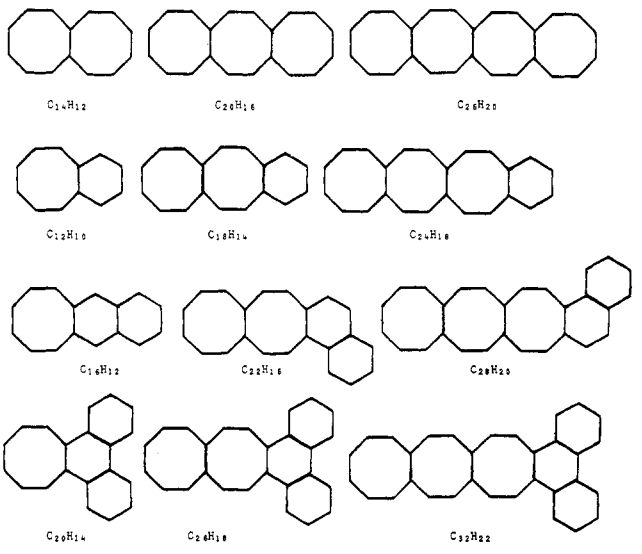
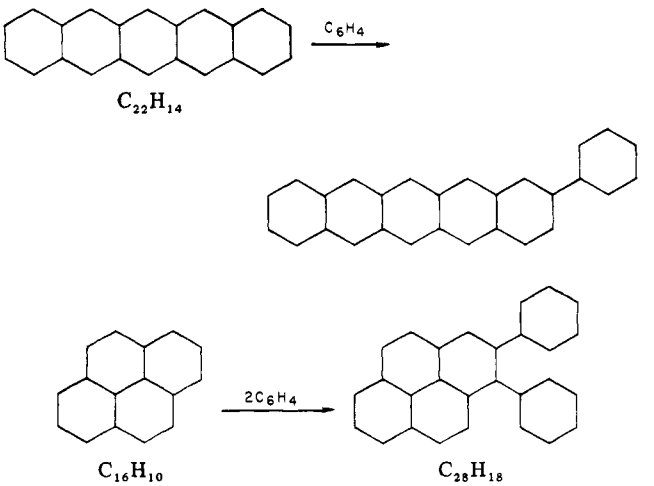


Figure 11. Structures containing fused octagonal rings.

adding a phenylene unit to pentacene ($C_{22}H_{14}$) or two phenylene units to pyrene ($C_{16}H_{10}$) and is illustrated by the following.



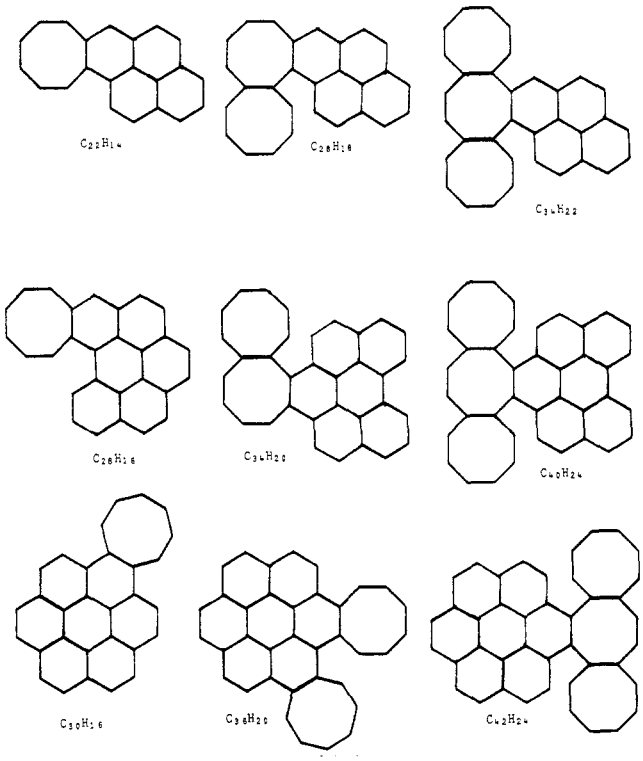


Figure 12. PAH structures containing fused octagonal rings.

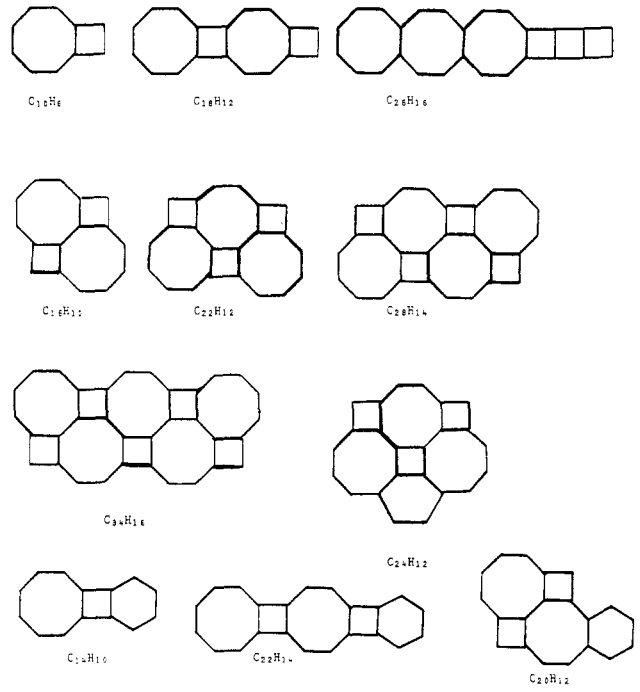


Figure 13. Structures containing both fused octagonal and tetragonal rings.

The lower left-hand boundary of Table VIII is determined by substracting a single C_6H_4 unit from the test formula and confirming that the resulting formula exists in Table III; if it does not, then it cannot be a phenyl-substituted PAH6 structure. For example, substracting C_6H_4 from $C_{22}H_{14}$ gives $C_{16}H_{10}$ which is a formula existing on the left-hand diagonal boundary of Table III; therefore, the $C_{22}H_{14}$ formula can be represented by a phenyl-substituted pyrene. Thus the left-hand boundary of Table VIII is generated by adding C_6H_4 to each formula on the left-hand diagonal boundary of Table III.

An algorithm for determining the maximum number of phenyl substituents that any structure corresponding to a

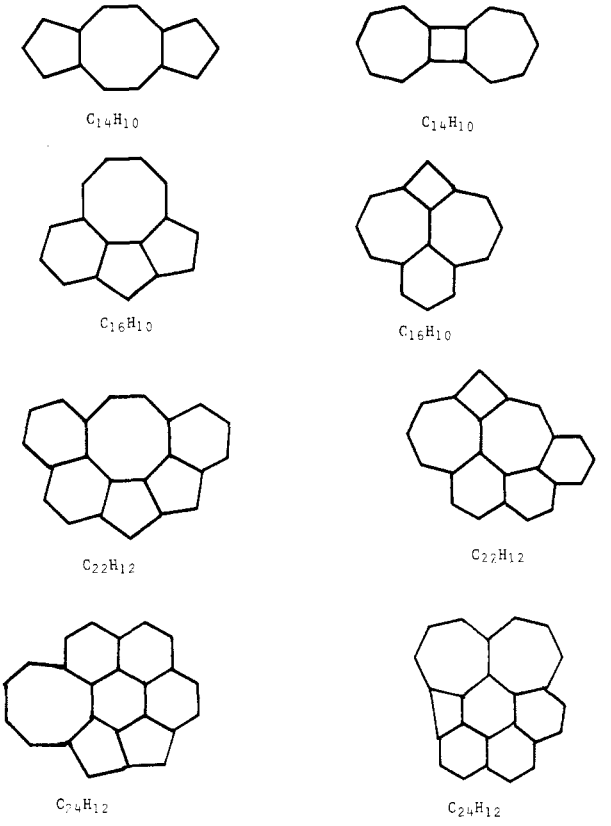


Figure 14. Sample PAH structures having one octagonal and two pentagonal rings ($r_5 - 2r_8 = 0$) or one tetragonal and two heptagonal rings ($2r_4 - r_7 = 0$).

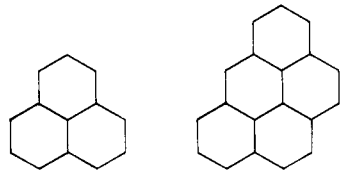
particular PAH6 formula can possibly contain is given by the following. Successively substract C_6H_4 from a PAH formula verified to be in Table III and/or Table VIII until the resulting formula is not longer observed in Table III (i.e., goes off into the upper left-hand domain of Table III where PAH6 structures do not exist); the maximum number of phenyl substituents is one less than the number of C_6H_4 units substracted. For example, substracting three C_6H_4 units from $C_{58}H_{28}$ gives $C_{40}H_{16}$ which would give a nonexistent PAH6 formula of $C_{34}H_{12}$ upon substracting another C_6H_4; thus, a PAH structure corresponding to $C_{58}H_{28}$ cannot possess more than three phenyl substituents.

Equations 1, 6, 8, and 9 in Table II are still applicable to PAHs containing phenyl substituents. However, in going from strictly PAHs composed of exclusively fused hexagonal rings to PAHs having only fused hexagonal rings and phenyl substituents, one must modify eq 2, 4, and 7 by adding a $2q_b$ term to the left-hand side of each of these equations to give $N_{Pc} + 2q_b = N_H - 6$, $q_P + 2q_b = 2N_H - 6$, and $q_I + 1 + d_s + 2q_b = N_c - N_H = N_{Ic} + N_{Pc}$, respectively; similarly, eq 3 and 5 become $q = q_I + q_P + q_b$ and $d_s = N_{Pc} - r - q_b$, respectively.

Although, the focus of the above discussion has been on phenyl substituted PAH6 structures, the term q_b really represents a different kind of bond whether it is between a phenyl group and a PAH6 compound or between two PAH6 groups. Thus the formula $C_{20}H_{14}$ has one such bond whether it is between a phenyl and phenanthrenyl groups, two naphthalenyl groups, etc.

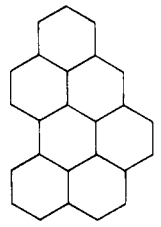
Radical-Containing PAHs. The PAH6 formula periodic table (Table III) excluded odd carbon formulas since these would invariably lead to radical PAHs of lesser likelihood. However, from a theoretical point of view they are now considered in this section since they also could exist as ions. Consider pernaphthalene ($C_{13}H_9$). It has one internal third degree carbon ($N_{Ic} = 1$) with a formula intermediate between

those of naphthalene ($C_{10}H_8$, $N_{Ic} = 0$) and pyrene ($C_{16}H_{10}$, $N_{Ic} = 2$). Successive fusion of C=CHCH=C units onto pernaphthalene would produce the $N_c = 2N_H - 5$ row series. Similarly, other odd internal third degree carbon containing row series can be generated at $N_c = 2N_H - 3$, $N_c = 2N_H - 1, \ldots$. The first members of these rows series are presented below. Extension of Table III to include the row series having



perinaphthalene $C_{13}H_9, N_{Ic} = 1$

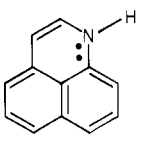
peripyrene $C_{19}H_{11}, N_{Ic} = 3$



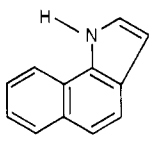
peri[*c,d*]benzo[*g,h,i*]perylene $C_{25}H_{13}, N_{Ic} = 5$

an odd number of internal third degree carbon atoms should be obvious from these examples. Note that all the equations in Table II are still valid for these radical PAH6s.

PAHs Containing Heteroatoms. Substitution of nitrogen (or phosphorus, arsenic, or antimony) for C–H units in a PAH found in Tables III–VII can be easily accommodated in this treatment. However, if nitrogen replaces a third degree internal or peripheral carbon atom, then the PAH must be a positive ion. Thus, the formulas $C_{15}H_9N$ and $C_{15}H_{10}N^+$ would correspond to these two different replacements of CH or third degree carbon in pyrene ($C_{16}H_{10}$) by nitrogen. Oxygen or a chalcogen can only replace a CH unit, and this leads to a positive ion PAH analogue. Another type of nitrogen replacement is demonstrated by indole (C_8H_7N) which results from the substitution of a CH=CH unit in naphthalene ($C_{10}H_8$) by an NH unit; benzofurans (C_8H_6O) result from the substitution of CH=CH in naphthalene by O. In odd carbon-radical PAH6s, replacement of a CH unit, at which the odd electron can be located, by a NH unit or oxygen gives a nonradical PAH analogue. This latter replacement generates a structure which can be isomeric with a structure produced by replacement of a CH=CH unit in an even-carbon PAH by a NH unit or oxygen. This is illustrated by the following structures formed by substituting an N into perinaphthalene and by substitution of a CH=CH unit in phenanthrene by a NH unit, respectively.



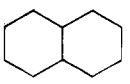
$C_{12}H_9N$



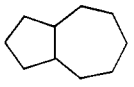
$C_{12}H_9N$

Obviously multiple heteroatom substitution will be more complex, but from the above discussion, the following algorithm for classifying structure–formula relationships of heteratom-containing formulas that can be represented by PAH analogue structures may be summarized as follows. (1) In the given molecular formula substract out NH and O for each NH_2 and OH substituent known to be present. (2) After consideration of the possibility of quinones, replace all Ns and Os by CHs. (3) For each N replaced in step 2 successively add a C, and for each O successively add a CH_2; for each even carbon/even hydrogen or odd carbon/odd hydrogen formula obtained at each step in this process search the PAH formula tables for the existence of the formula thus far obtained. As an illustration of this algorithm, consider 9-aminoacridine ($C_{13}H_{10}N_2$) and quinalizarin ($C_{14}H_8O_6$). For 9-aminoacridine the three steps give (1) $C_{13}H_9N$, (2) $C_{14}H_{10}$, and (3) $C_{15}H_{10}$, of which step 2 results in an acceptable solution corresponding to anthracene or phenanthrene. Application of the three steps to quinalizarin led to (1) $C_{14}H_8O_2$, (2) $C_{14}H_{10}$ and $C_{16}H_{10}$, and (3) $C_{17}H_{11}$ and $C_{18}H_{12}$, of which steps 2 and 3 give acceptable solutions. Needless to say, a unique determination of a solution in the latter will require other chemical information beyond the molecular formula.

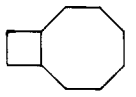
Effect of Ring Sizes on Resonance Energies of PAHs. The comparison of pentagonal and heptagonal ring containing PAHs with tetragonal and octagonal ring containing PAHs led to consistent graph-theoretical results. However, one must consider the probable formation of the polyantiaromatic hydrocarbons containing tetragonal and octagonal rings. This is made dramatically evident by comparison of the relative stability of the isomers of naphthalene. Use of either a corrected structure count (CSC)[6] or a conjugated circuit resonance energy (RE)[7] determination gives the data shown below.



CSC = 3
RE = 1.323 eV



CSC = 2
RE = 0.246 eV



CSC = 1
RE = −1.203 eV

Although azulene is a well-known compound, its RE is significantly below that of naphthalene. The azulenoid isomers of pyrene have been shown to contain various degrees of reduced RE.[8] However, pyrene (CSC = 6, RE = 2.133 eV) and fluoranthene (CSC = 6, RE = 2.192 eV) have very nearly identical resonance energies. Thus it appears that versions of the formula periodic table for tetragonal and octagonal rings containing PAHs will have mainly graph-theoretical importance and limited application to formula–structure relationships of PAHs, i.e., will only find application to PAHs containing a preponderance of hexagonal rings and only a few tetragonal or octagonal rings. If some of the compounds in Figure 13 could be synthesized, they no doubt would possess some interesting properties such as exceptional crystal conductivity.

CONCLUSION

Formula–structure relationships for PAHs have been adduced by the unique application of graph theoretical concepts. This intuitive approach presumes that the PAHs can be recursively generated. The maximum number of pentagonal rings that a PAH6 formula can contain along with hexagonal rings is given by $r_{5max} \leq N_c - 2N_H + 6$. Similarly, the maximum number of heptagonal rings that a PAH6 formula can contain along with hexagonal rings is given by $r_{7max} \leq d_s$-(PAH6) + M (where $M = 0, 1, 2, \ldots$) and corresponds to $|d_s|$ of the first member of the associated row series in Table III. Analogously, the maximum number of tetragonal rings that a structure having a PAH6 formula can contain along with hexagonal rings is given by $r_{4max} \leq [N_c - 2N_H + 6]/2$. It has been semiempirically derived (cf. Appendix) that for PAHs the generalized version of eq 2 becomes

$$N_{Pc} = N_H - 6 + 6c + 2r_4 + r_5 - r_7 - 2r_8 - 2q_b$$

An algorithm for determining the maximum number of phenyllike substituents that a structure having a PAH6 formula can possess has been presented.

The original formula PAH periodic table (Table III) becomes expanded in the left-hand region when smaller ring sizes are permitted and in the upper right-hand region when larger ring sizes are allowed; there is a tendency for fewer internal third degree carbons (N_{Ic}) in the former case and for more internal third degree carbon atoms in the latter case. When $c = 0$ and $q_b = 0$, the following have been postulated: (a) Table III is applicable when $2r_4 + r_5 - r_7 - 2r_8 = 0$; (b) Table IV is applicable when $2r_4 + r_5 - r_7 - 2r_8 = 1$; (c) Table V is applicable when $2r_4 + r_5 - r_7 - 2r_8 = -1$; (d) Table VI is applicable when $2r_4 + r_5 - r_7 - 2r_8 = 2$; (e) Table VII is applicable when $2r_4 + r_5 - r_7 - 2r_8 = -2$.

The formula subscript numbers in Tables III–VII are simply subsets of the universal even pair set ($N \times N$). The universal even pair set $N \times N$ is an *equivalence relation in N*, where N is the set of all even, positive, nonzero integers (i.e., natural numbers). Thus $N \times N = \{\langle a,b\rangle: a, b \epsilon N, a > 0, b > 0\}$ and

$$\{\text{Table VI}\} = R_4^* = \{\langle a,b\rangle: aR_4b\}$$

$$\{\text{Table IV}\} = R_5^* = \{\langle a,b\rangle: aR_5b\}$$

$$\{\text{Table III}\} = R_6^* = \{\langle a,b\rangle: aR_6b\}$$

$$\{\text{Table V}\} = R_7^* = \{\langle a,b\rangle: aR_7b\}$$

$$\{\text{Table VII}\} = R_8^* = \{\langle a,b\rangle: aR_8b\}$$

$$R_4^* \subset N \times N,\ R_5^* \subset N \times N,\ R_6^* \subset N \times N,\ R_7^* \subset N \times N,\ R_8^* \subset N \times N$$

$$R_5^* \subset R_4^*,\ R_6^* \subset \{R_5^*,\ N_c = 2N_H + 6 \text{ row series}\} \subset \{R_4^*,\ N_c = 2N_H + 6 \text{ row series}\}$$

$$\{\text{Table VIII}\} = R_p^* \subset R_8^*$$

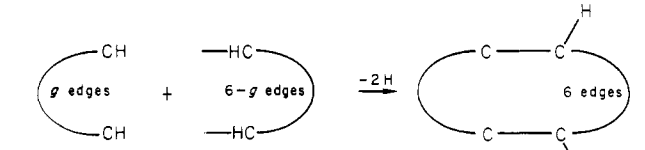
APPENDIX. INDUCTIVE PROOF OF EQUATION 2

Theorem: For exclusively fused polycyclic aromatic hydrocarbons composed of only hexagonal rings; $N_{Pc} = N_H - 6$.

Proof: The proof is by induction. This proposition is true for $r = 1$ and 2; i.e., for benzene ($0 = 6 - 6$) and naphthalene ($2 = 8 - 6$) let the graphs be recursively built up from the $r = 1$ case by successive fusion of additional hexagonal rings.

It will be shown that at each step N_{Pc} and N_H are incremented by the same amount. In any single step, a new hexagon will be formed by taking g existing adjacent peripheral edges ($g = 1$–5) of which the two extreme carbon vertices contain hydrogen atoms (the remaining $g - 1$ carbon vertices are peripheral third degree carbon atoms without adjacent hydrogen atoms) and adding $6 - g$ edges, i.e., $5 - g$ additional carbon vertices each possessing a hydrogen.

Since two hydrogens are lost, N_H increases by $3 - g$. Also, $g - 1$ peripheral carbon vertices become internal carbon vertices, and two former hydrogen-bearing carbons become peripheral ones, and thus N_{Pc} also increases by $2 - (g - 1) = 3 - g$.



$(g + 1)$ C atoms + $(6 - g - 1)$ C atoms − 2 H atoms + $(6 - g - 1)$ H atoms

6 C atoms
$(3 - g)$ H atoms

This completes the inductive proof of eq 2.

REFERENCES AND NOTES

(1) Part 1: Dias, J. R. *J. Chem. Inf. Comput. Sci.* **1982**, *22*, 15.
(2) Balaban, A. T.; Harary, F. *Tetrahedron* **1966**, *24*, 1097.
(3) Harary, F.; Palmer, E. "Graphical Enumeration"; Academic Press: New York, 1973. Balaban, A. T. "Chemical Applications of Graph Theory"; Academic Press: New York, 1976.
(4) Balaban, A. T. *Rev. Roum. Chim.* **1981**, *26*, 407.
(5) Randić, M. *J. Am. Chem. Soc.* **1977**, *99*, 444.
(6) Herndon, W. *Tetrahedron* **1973**, *29*, 3.
(7) Randić, *Chem. Phys. Lett.* **1976**, *38*, 68.
(8) Randić, M. *Tetrahedron* **1977**, *33*, 1905.

WISENOM. A Formal Organic Chemical Nomenclature System

E. V. KRISHNAMURTHY

Department of Applied Mathematics, Indian Institute of Science, Bangalore 560012, India

Received April 28, 1981

A formal chemical nomenclature system WISENOM based on a context-free grammar and graph coding is described. The system is unique, unambiguous, easily pronounceable, encodable, and decodable for organic compounds. Being a formal system, every name is provable as a theorem or derivable as a terminal sentence by using the basic axioms and rewrite rules. The syntax in Backus–Naur form, examples of name derivations, and the corresponding derivation trees are provided. Encoding procedures to convert connectivity tables to WISENOM, parsing, and decoding are described.

I. INTRODUCTION

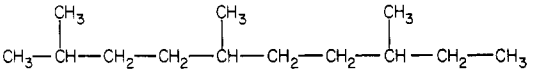
There are many systems for naming organic compounds.[1,2] Although there are several advantages in these systems, e.g., human comprehension, mechanical encoding and decoding, these names are difficult, and in some cases impossible, due to the lack of a formal grammar with a well-defined syntax.

Many names cannot be decoded into a structural formula, especially trivial names, and others that are highly context dependent require intuition and human skill. For example, consider

(i) 2,5,8-trioxadecane

$$CH_3\text{–}O\text{–}CH_2\text{–}CH_2\text{–}O\text{–}CH_2\text{–}CH_2\text{–}O\text{–}CH_2\text{–}CH_3$$

and (ii) 2,5,8-trimethyldecane



In (i), the oxygen atoms are along the main chain of 10 atoms, and in (ii), methyl groups are not in the main chain but are substituents.

0095-2338/82/1622-0152$01.25/0 © 1982 American Chemical Society