

Enumeration and Classification of Coronoid Hydrocarbons

J. BRUNVOLL,* B. N. CYVIN, and S. J. CYVIN

Division of Physical Chemistry, The University of Trondheim, N-7034 Trondheim-NTH, Norway

Received July 3, 1986

A coronoid is defined as a planar system of identical regular hexagons (no overlapping) with one or more holes of a size not less than two hexagons. The numbers of coronoids with h (number of hexagons) = 8, 9, 10, 11, and 12 were found to be 1, 5, 43, 283, and 1954, respectively. A detailed report is given on a classification of the coronoids up to $h = 11$. They are Kekuléan or non-Kekuléan. The former class is divided into catacondensed (normal), normal pericondensed, and essentially disconnected (pericondensed). Among the catacondensed systems the primitive coronoids are distinguished; they consist of one circular single chain. A further subdivision is represented by the hollow hexagons, which possess exactly six segments each. The Δ value (absolute magnitude of the difference between the numbers of peaks and valleys) plays an important role in the classification. The different classes are also treated in general. The non-Kekuléans (always pericondensed) are divided into obvious ($\Delta > 0$) and concealed ($\Delta = 0$) categories. It is pointed out that all integer values of Δ are possible. The smallest concealed non-Kekuléan coronoids that were detected have 15 hexagons each. The smallest double coronoid (with two holes) has $h = 13$. Two concealed non-Kekuléan double coronoids with $h = 18$ are depicted. Finally the principles of a computer program in several modifications are mentioned briefly.

Benzenoid (polycyclic aromatic) hydrocarbons have been enumerated and classified by Balaban et al.¹⁻⁴ and others. In more recent studies computerized enumerations and computer-aided classifications have been executed with success.⁵⁻¹¹ A pioneering work in this area by Balasubramanian et al.⁵ stimulated the Düsseldorf-Zagreb group⁶⁻⁹ to more extensive studies. A useful review is available.⁸

In the mentioned work of Balasubramanian et al.⁵ the catacondensed benzenoids (i.e., without internal vertices) were enumerated, and the numbers of carcinogenic bay regions were determined. Knop et al.⁶ extended this work to all benzenoids (also pericondensed with internal vertices). Altogether they established the numbers of all benzenoids with h (the number of hexagons) up to 10. For $h = 10$ their number is 30 086. Recently Doroslovački and Tošić¹² obtained the number 141 229 for all benzenoids with $h = 11$.

Much less work of this kind has been done for coronoids (corona-condensed benzenoids, circulenes, cycloarenes;^{1,13,14} for precise definitions see below). These hydrocarbons are of considerable interest in organic chemistry. It is true that only one chemically pure coronoid hydrocarbon has been synthesized so far, viz., kekulene.¹⁵⁻¹⁸ However, attempts to synthesize others have been announced and are referred to in a recent theoretical work by Vogler¹⁹ and references cited therein. Furthermore, the skeleton of another coronoid has been isolated in the form of a dihydroderivative.²⁰

With regard to the enumeration of coronoids we refer to Knop et al.⁸ The definition of "true" circulenes used by these researchers seems to coincide with our coronoid concept (see below). The authors⁸ have reported the existence of 1, 5, and 48 such systems for h (number of hexagons) = 8, 9, and 10, respectively. We have confirmed the two first numbers, but found 43 coronoids for $h = 10$.²¹ A complete account of the existing coronoids with $h = 11$ is also given here, in addition to a discussion of certain aspects of still larger systems.

RESULTS AND DISCUSSION

Definitions. The Düsseldorf-Zagreb school⁶⁻⁹ enumerated primarily the systems referred to as benzenoids²² in Gutman's review.²³ This definition is also used as a guideline for the present definition of a coronoid. We consider planar systems consisting of identical regular hexagons. Consequently, we exclude "heli-circulenes" in the terminology of Knop et al.,⁸ i.e., hexahelicene derivatives and similar systems with coalescing edges if drawn in a plane. The system should have a

"hole" of the size of more than one hexagon. Allowance is made for more than one hole by introducing the terms double-coronoid, triple-coronoid, etc.

There has been a slight controversy over whether coronoids (referred to as corona-condensed systems) should be regarded as pericondensed¹ or catacondensed,¹³ depending on whether the emphasis is on chemical or structural properties. The controversy is eliminated first by our proposal to abandon the designation "corona-condensed". Second, we wish to speak about catacondensed coronoids and pericondensed coronoids in the same spirit as these useful designations are defined for benzenoids: in a catacondensed coronoid all vertices are on either the outer or inner perimeter, while a pericondensed coronoid has at least one vertex belonging to three hexagons. These properties manifest themselves by the absence or presence of triangles in the dualist graph for catacondensed or pericondensed coronoids as well as benzenoids. In a dualist graph¹ each hexagon is represented by a point.²⁴

Also in complete analogy with benzenoids we define the following terms: A Kekuléan coronoid possesses Kekulé structures, a non-Kekuléan does not. An essentially disconnected coronoid has some bonds fixed; they are single or double in all Kekulé structures. The simplest essentially disconnected benzenoid is perylene, where the two naphthalene fragments are joined by fixed single bonds. In zethrene there are both single and double fixed bonds. A normal coronoid is defined as Kekuléan and not essentially disconnected.

Values of Δ . The number Δ is a positive integer or zero designating the absolute magnitude of the difference between the numbers of peaks and valleys.²⁵⁻²⁷ Suppose the coronoid is oriented so that some of its edges are vertical. Then a peak is a vertex on the perimeter characterized by having both its neighboring vertices below it. Similarly, a valley is a vertex on the perimeter below both its neighbors. The value of Δ is independent of the orientation of the system and may be proved to be equal to the absolute magnitude of the difference between the numbers of black and white vertices.^{23,28,29} The condition $\Delta > 0$ assures that the coronoid is non-Kekuléan. On the other hand, if $\Delta = 0$, the coronoid is not necessarily Kekuléan. Non-Kekuléan coronoids with $\Delta = 0$ are presently referred to as concealed, in contrast to obvious non-Kekuléans when $\Delta > 0$. All these properties are completely analogous to corresponding properties of benzenoids.

All Coronoids with $h = 8, 9$, and 10. Figure 1 shows the unique coronoid of $h = 8$ and the five systems with $h = 9$. They are ordered according to Δ values and the number of

Table I. Number of Coronoids with Different Δ Values and h up to 12

h	Kekuléan, $\Delta = 0$	non-Kekuléan			total
		$\Delta = 1$	$\Delta = 2$	$\Delta = 3$	
8	1	0	0	0	1
9	3	2	0	0	5
10	24	16	3	0	43
11	128	131	23	1	283
12	854	906	182	12	1954

Kekulé structures (K). In Figure 2 the 43 coronoids with $h = 10$ are displayed in the same way. Isoarithmic³⁰ systems (framed in Figure 2) differ only in the way the kinks go in an annelated single chain; this does not affect the number of Kekulé structures.

All Coronoids with $h = 11$. The studies were extended to include the coronoids with 11 hexagons. Two hundred eighty three systems were identified as constituting the complete set. In the following the systems are presented in terms of a practical classification.

Kekuléan Derivatives of the $h = 8$ Coronoid. The coronoid of eight hexagons (cf. Figure 1) may be supplied with catacondensed appendages to form catacondensed coronoids as shown in Figure 3. There are three possibilities, viz., (1) annelation of a chain of three hexagons, (2) separated annelation of two hexagons and one hexagon, and (3) three hexagons annelated in separated positions. The cases are referred to as 3-0, 2-1, and 1-1-1 types of annelation, respectively.

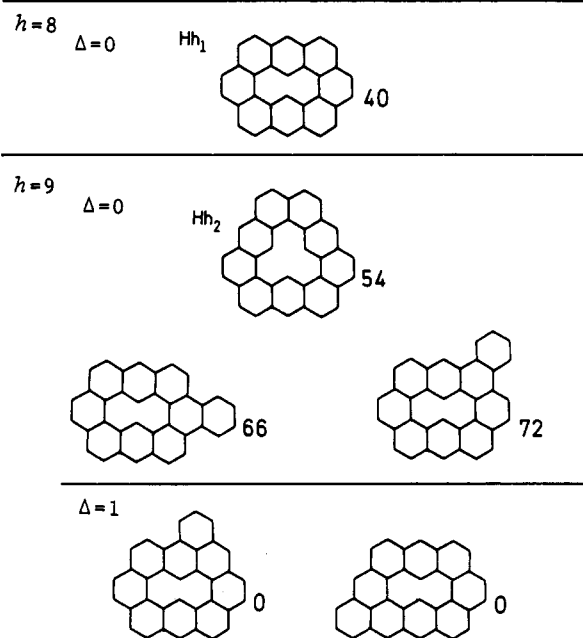
The next class (Figure 4, upper part) consists of pericondensed coronoids with one hexagon annelated to each. Before annelation the system is derived from the $h = 8$ coronoid by adding two hexagons, either in contact with each other or separated. These two cases are identified by (2) and (1,1), respectively.

Next we have the types without catacondensed appendages; three hexagons are condensed with the original coronoid, either connected as (3) or partly disconnected as (2,1) (cf. Figure 4, lower part).

Finally, in the category of coronoids considered (Kekuléans with $h = 11$ derived from the $h = 8$ coronoid) we have the two smallest essentially disconnected coronoids (see Figure 5).

Non-Kekuléan Derivatives of the $h = 8$ Coronoid. All non-Kekuléan coronoids (as well as benzenoids) are pericondensed. The derivatives of the $h = 8$ coronoid were classified in the same way as the Kekuléans. The results for those with $\Delta = 1$ are summarized in Figure 6, where dualist graphs^{1,24} are employed for the sake of brevity. In the bottom line of the figure the coronoids may be considered a fusion of the two indicated units. Otherwise, the symbolism follows the principles used for the Kekuléans above. Figure 7 shows in the same way the coronoids with $\Delta = 2$. For $\Delta = 3$ a unique coronoid with 11 hexagons exists. It is depicted in Figure 8.

Kekuléan Derivatives of the $h = 9$ Circular Single Chain. The next group of coronoids with $h = 11$ was derived from the addition of hexagons to the 9-hexagon circular single chain

**Figure 1.** All coronoids with h (number of hexagons) = 8 and $h = 9$. Numbers of Kekulé structures (K) are indicated. For the designations "Hh" (hollow hexagon), see below.

(Figure 1). Additions inside the hole are not considered because they already are covered in the above treatment. Figure 9 shows the possibilities, catacondensed in the upper part and pericondensed in the lower part of the figure.

Non-Kekuléan Derivatives of the $h = 9$ Circular Single Chain. In Figure 10 the non-Kekuléans in question are displayed. They have Δ values of 1 (11 systems) and 2 (4 systems).

Derivatives of the Three $h = 10$ Circular Single Chains. The eight catacondensed coronoids along with the unique Kekuléan pericondensed coronoid of Figure 11 are obtained on adding one hexagon every time to one of the three existing 10-hexagon circular single chains (Figure 2). Similarly, one obtains 10 non-Kekuléans (Figure 12) on adding one hexagon in a pericondensed mode. All of these non-Kekuléans have $\Delta = 1$.

Primitive Coronoids. A coronoid consisting of a circular single chain will presently be referred to as primitive. There exist two primitive coronoids with 11 hexagons (cf. Figure 13). This completes the listing of coronoids with $h = 11$.

Survey of the Coronoids with h up to 12. Table I summarizes the number of coronoids with h up to 12 according to the values of Δ . For detailed classifications within each group the preceding paragraphs should be consulted. Some of these details for the Kekuléan coronoids ($\Delta = 0$) are summarized in Table II.

Primitive Coronoids. Hollow Hexagons. Among the primitive coronoids we distinguish primarily those with exactly six linear segments, which therefore have no corners protruding

Table II. Number of Kekuléan Coronoids with h up to 11, Classified into Some Groups

h	primitive	catacondensed					pericondensed			
		derived from					normal derived from			
										essentially disconnected
8	1	0	0	0	0	0	0	0	0	0
9	1	2	0	0	0	0	0	0	0	0
10	3	11	1	0	0	0	9	0	0	0
11	2	45	7	2	2	4	58	5	1	2

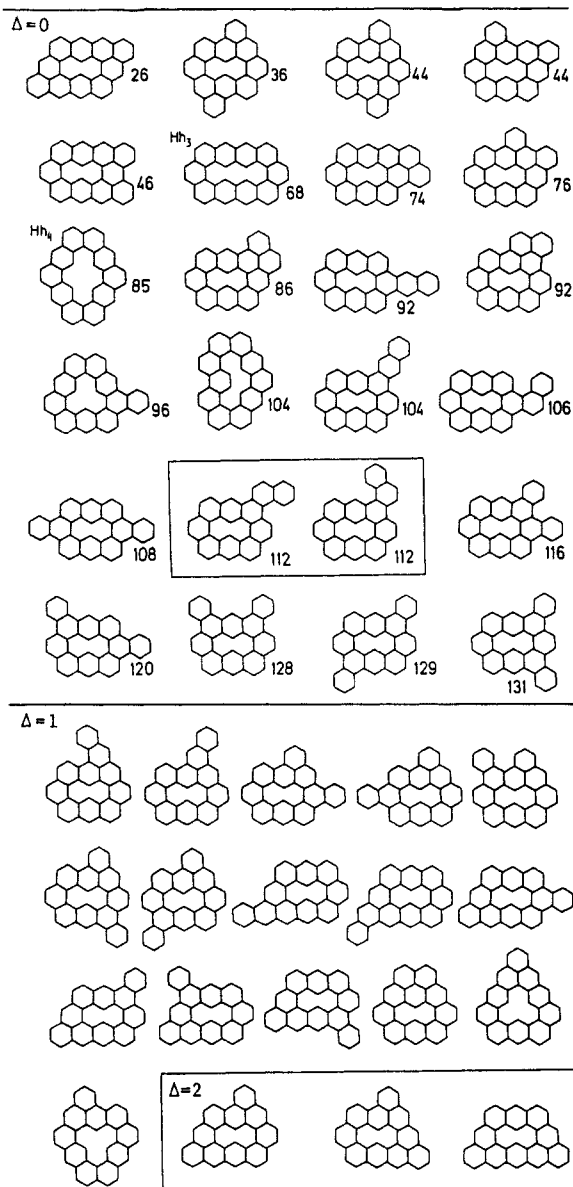


Figure 2. All coronoids with $h = 10$. Numbers of K for the Kekuléans are indicated. Framed structures are isoarithmic.

into the hole. They are presently designated hollow hexagons. A system was devised to determine all such coronoids with up to a given h value. Figure 14 shows the idea, which does not seem to be necessary to explain here in detail. The graphs are drawn in a standardized manner, where the bottom row is the longest linear segment or one of the longest if more than one have equal lengths. The integer t designates the number of horizontal rows. The numbers of hollow hexagons with $h = 8, 9, \dots, 20$ were found to be 1, 1, 2, 1, 4, 2, 5, 4, 7, 5, 11, 7, and 13, respectively.

All Primitive Coronoids. All the primitive coronoids may be derived from the hollow hexagons by making allowance for corners protruding into the hole. Such coronoids have more than six linear segments. This feature is not possible unless $h \geq 10$. For $h = 10$ and $h = 11$ there exists one system each of this kind (cf. Figures 12 and 13). Figure 15 shows the 11 existing primitive coronoids with $h = 12$. The four hollow hexagons are indicated as Hh_6 – Hh_9 . Two of them have been modified in different ways to produce intruding corners.

Figure 16 shows the 12 primitive coronoids with $h = 13$ and 40 with $h = 14$. Notice especially the last system derived from the hollow hexagon Hh_{15} . The system Hh_{15} itself has five rows, while the derived system has four.

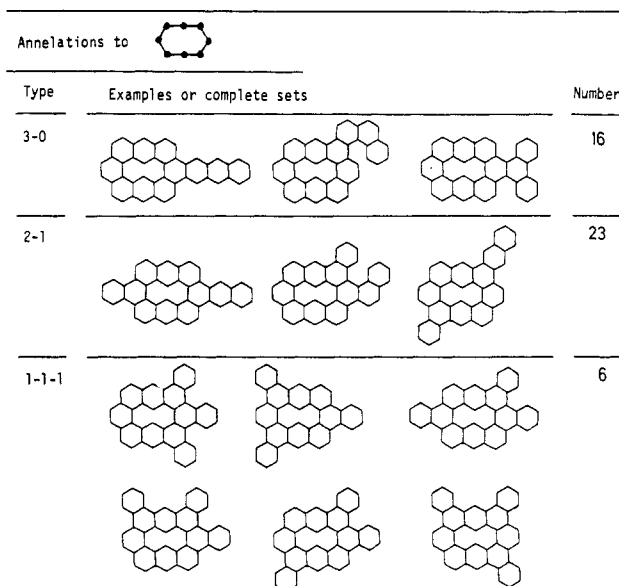


Figure 3. Catacondensed coronoids with $h = 11$ derived from the $h = 8$ coronoid (cf. the dualist graph).

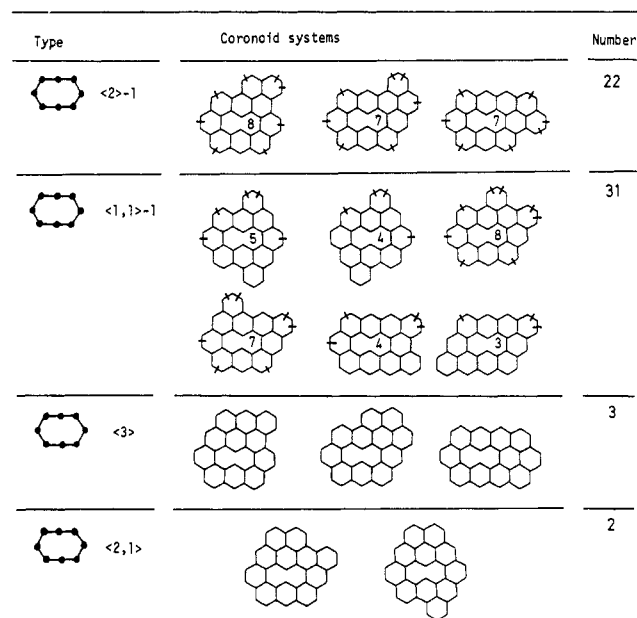


Figure 4. Normal pericondensed $h = 11$ coronoids derived from the $h = 8$ coronoid. Positions of annelation are indicated by strokes.

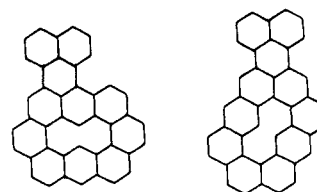


Figure 5. Two essentially disconnected coronoids with $h = 11$. In both cases $K = 3 \times 40 = 120$.

Coronoids with Features inside the Hole. Figure 17 shows the addition of a hexagon inside the hole of a primitive coronoid. The hexagon was added in a pericondensed mode (left) or annelated (right). For larger holes the inside features may of course be more elaborate: branched, pericondensed units, etc.

Essentially Disconnected Coronoids. The two smallest ($h = 11$) essentially disconnected coronoids are found in Figure 5. They are rather trivial inasmuch as two independent units


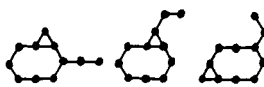

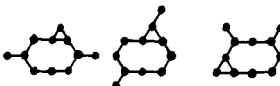



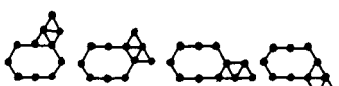

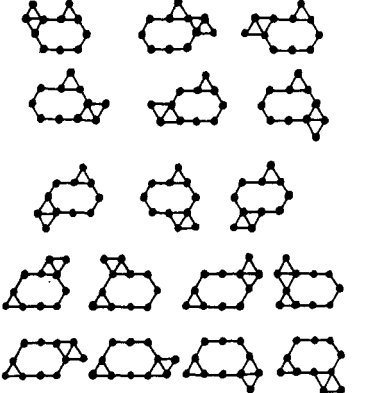

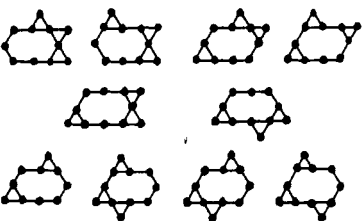

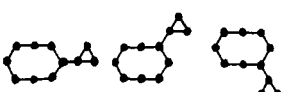
Type	Examples or complete set	Number
 $\langle 1 \rangle - 2$		39
 $\langle 1 \rangle - 1 - 1$		34
 $\langle 2 \rangle - 1$		3
 $\langle 3 \rangle$		4
 $\langle 2, 1 \rangle$		17
 $\langle 1, 1, 1 \rangle$		10
 Δ		3

Figure 6. Dualist graphs of the coronoids with $h = 11$ and $\Delta = 1$, derived from the $h = 8$ coronoid.

are joined in a way that is well-known from benzenoids.³¹ The actual cases (Figure 5) are substituted perylenes. Figure 18a shows also a kind of substituted perylene, but with features inside the hole. The other examples of Figure 18 demonstrate that the hole may be partly or completely surrounded by fixed bonds. In parts e and f it should be noticed in particular that no fixed double bonds are present.

Non-Kekuléan Coronoids. Obvious Non-Kekuléans. Figure 1 shows the two smallest ($h = 9$) non-Kekuléan coronoids, which have $\Delta = 1$ and therefore are obvious. Similarly, the three smallest ($h = 10$) coronoids with $\Delta = 2$ are found in Figure 2, while Figure 8 displays the unique $h = 11$ coronoid with $\Delta = 3$. It is not difficult to show that any integer value of Δ is possible for the coronoids as well as for benzenoids.³² It is sufficient to demonstrate the validity of the statement for triangle-shaped coronoids. Figure 19 shows some illustrative examples.

Concealed Non-Kekuléans. All the non-Kekuléan coronoids encountered in the preceding paragraphs are of the obvious type ($\Delta > 0$). Also, concealed non-Kekuléan ($\Delta = 0$) coronoids exist. Figure 20 shows some examples. Notice the principle


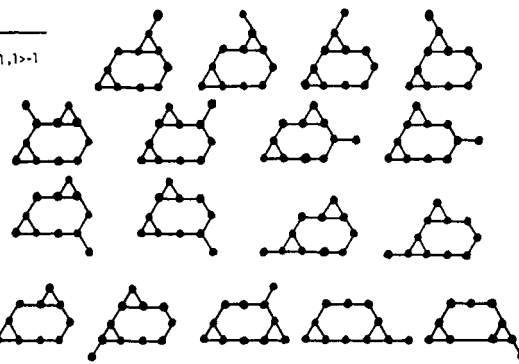




Type	Examples or complete set	Number
 $\langle 1, 1 \rangle - 1$		
 $\langle 2, 1 \rangle$		
 $\langle 3 \rangle$		

Figure 7. Dualist graphs of the $17 + 1 + 1 = 19$ coronoids with $h = 11$ and $\Delta = 2$, derived from the $h = 8$ coronoid.

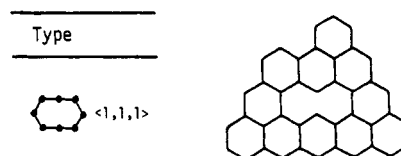


Figure 8. Unique coronoid with $h = 11$ and $\Delta = 3$.


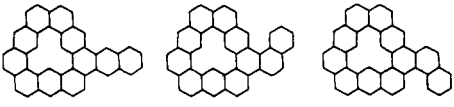
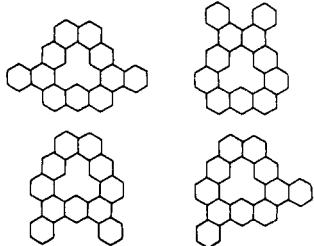

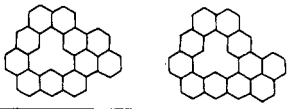
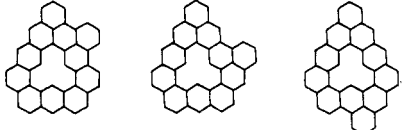
Catacondensed derivatives of 		
Type	Complete sets	Number
2-0		3
1-1		4
Pericondensed derivatives of 		
Type	Complete sets	Number
$\langle 2 \rangle$		2
$\langle 1, 1 \rangle$		3

Figure 9. Catacondensed and pericondensed Kekuléan $h = 11$ coronoids derived from the $h = 9$ circular single chain (cf. the dualist graph).

difference between Figure 18d and Figure 20a. The former system is essentially disconnected, the latter non-Kekuléan. The first three examples of Figure 20 have $h = 22$, but this is not the minimum value of h . In systems e and d advantage


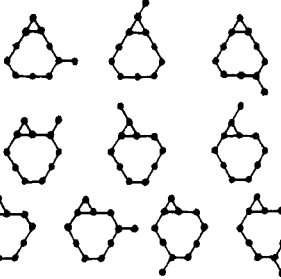




Δ	Type	Complete sets	Number
1	 $\langle 1, > 1 \rangle$		10
	 $\langle 2 \rangle$		1
2	 $\langle 1, 1 \rangle$		4

Figure 10. Dualist graphs of the $h = 11$ non-Kekuléan coronoids derived from the $h = 9$ circular single chain.


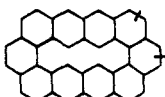

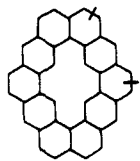

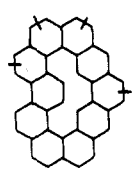

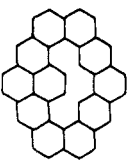
Type	Coronoid systems	Number
 1-0		2
 1-0		2
 1-0		4
 $\langle 1 \rangle$		1

Figure 11. $h = 11$ Kekuléan coronoids derived from the four systems given as dualist graphs. Positions of annelation are indicated by strokes.

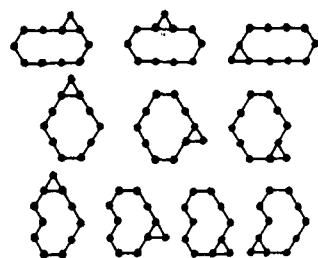


Figure 12. Dualist graphs of the $h = 11$ non-Kekuléan coronoids ($\Delta = 1$) derived from $h = 10$ circular single chains.

was taken of the well-known fragments of concealed non-Kekuléan benzenoids,³² and the number of hexagons was reduced to 15. These four examples are virtually substituted perylenes. There are several more variants of concealed

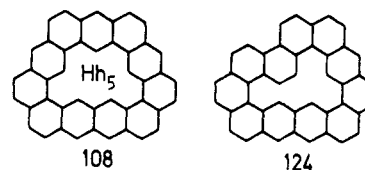


Figure 13. Two primitive (circular single chain) coronoids with $h = 11$. K numbers are given.







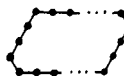





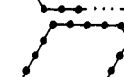



$t=3$		$h = 8, 10, 12, 14, 16, 18, \dots$
$t=4$		$h = 9, 11, 13, 15, 17, \dots$
		$h = 10, 12, 14, 16, 18, \dots$
$t=5$		$h = 12, 14, 16, 18, \dots$
		$h = 12, 14, 16, 18, \dots$
		$h = 13, 15, 17, \dots$
		$h = 14, 16, 18, \dots$
$t=6$		$h = 15, 17, \dots$
		$h = 15, 17, \dots$
		$h = 16, 18, \dots$
		$h = 16, 18, \dots$
		$h = 17$
		$h = 18$
$t=7$		$h = 18$
		$h = 18$
		$h = 18$

Figure 14. Dualist graphs of hollow hexagons (primitive coronoids); they cover all systems up to $h = 18$.

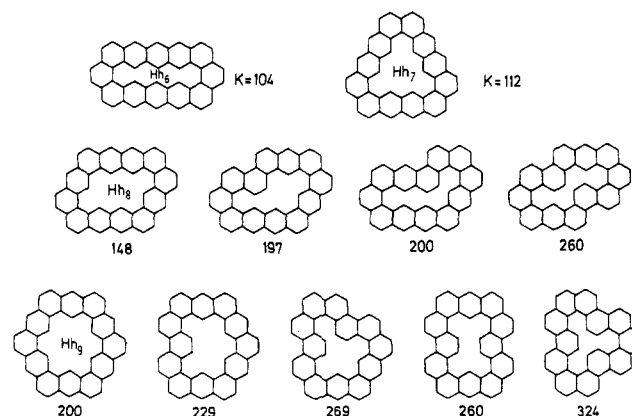


Figure 15. Eleven primitive coronoids with $h = 12$. K values are given. "Hh" designates a hollow hexagon. For Hh_1, \dots, Hh_5 , see Figures 1, 2, and 5.

non-Kekuléan coronoids with $h = 15$, but no smaller ($h < 15$) variants were detected by trial and error. This does not mean that they do not possibly exist.

Double Coronoids. The smallest double coronoid has been identified by Dias³³ and others. It is depicted in Figure 21 along with other examples of normal double coronoids.

Figure 22 shows examples of two essentially disconnected and four concealed non-Kekuléan coronoids. Examples a, b,

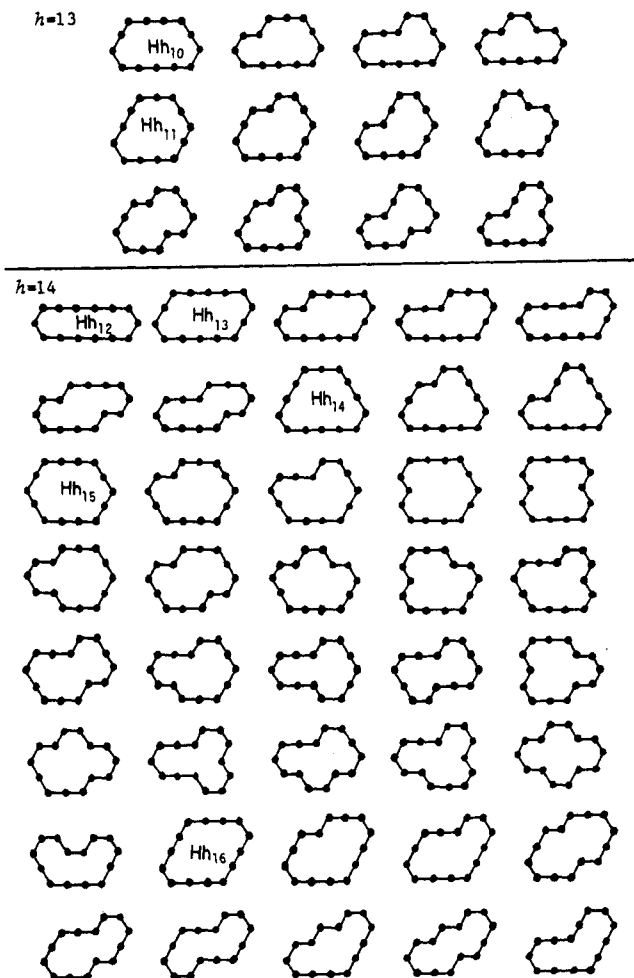


Figure 16. Dualist graphs of the primitive coronoids with $h = 13$ and $h = 14$. Hh_{10} , ..., Hh_{16} are the hollow hexagons.

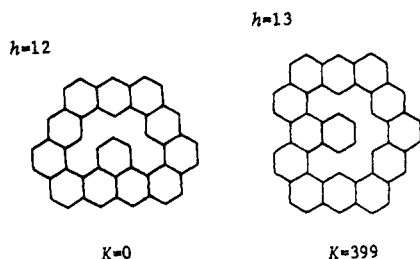


Figure 17. One non-Kekuléan (left) and one Kekuléan (right) coronoid with features inside the hole.

and d are substituted perylenes. Notice that the smallest ($h = 18$) concealed non-Kekuléan double coronoids, which were found by trial and error, viz., e and f, are not substituted perylenes. It has not been proved that still smaller ($h < 18$) systems of this type do not exist.

Computer Programming. The enumerations of the present work were automatized by computer programming. Any coronoid (or benzenoid) is identified by a pair of coordinates for every hexagon. A skew coordinate system is employed. This identification is substantially different from the system of the Düsseldorf-Zagreb group,⁶⁻⁹ who made use of a boundary code. It is also different from the He-matrix formalism of He and He.¹⁰ A description of the program and its application to benzenoids is published elsewhere.¹¹ The flexibility of this program, which allows for a number of options, made it relatively easy to adapt it to the analysis of coronoids. Only a short supplement to the previous description will be sufficient here.

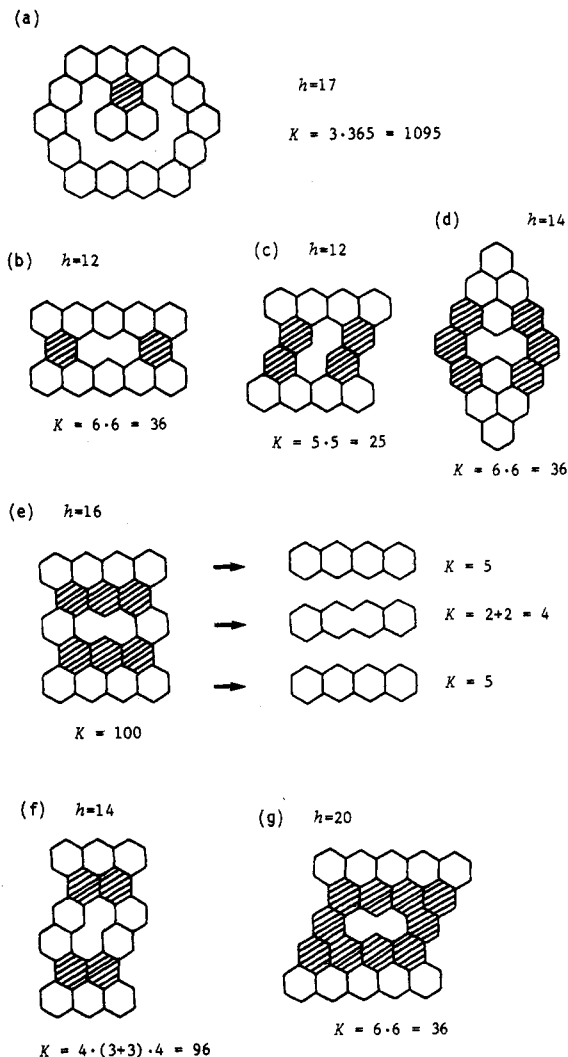


Figure 18. Examples of essentially disconnected coronoids. Hexagons that contain fixed bonds are shaded. K values are given, including hints on how to determine them.

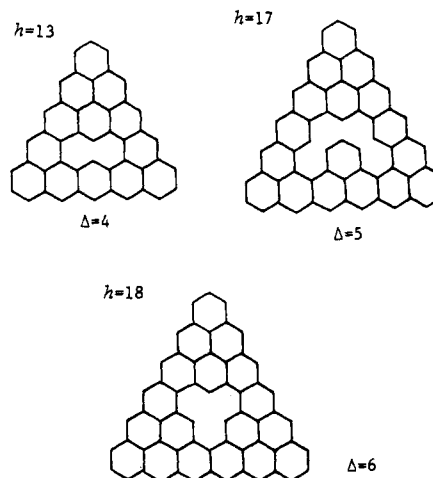


Figure 19. Examples of obvious non-Kekuléan coronoids.

A series of coronoids with $h + 1$ hexagons are generated by adding one hexagon every time to one or more coronoids with h hexagons. A procedure takes care of eliminating the newly generated coronoid if it is identical with one generated before. A special feature of our program is the counting of peaks and valleys to obtain Δ values. That gives us automatically a useful classification as shown in Table I.

The coordinates of the primitive coronoids had to be used as input for the generation of new coronoids. The primitive

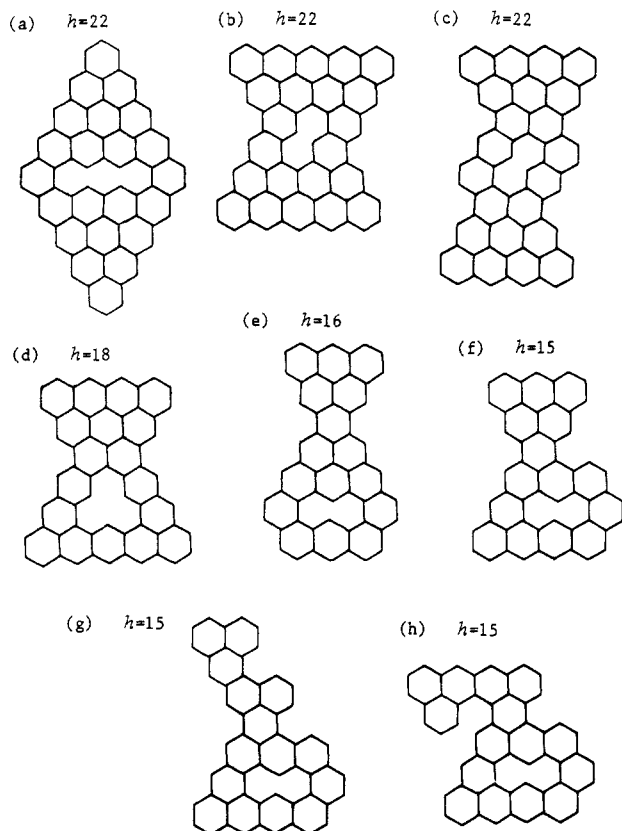


Figure 20. Examples of concealed non-Kekuléan coronoids.

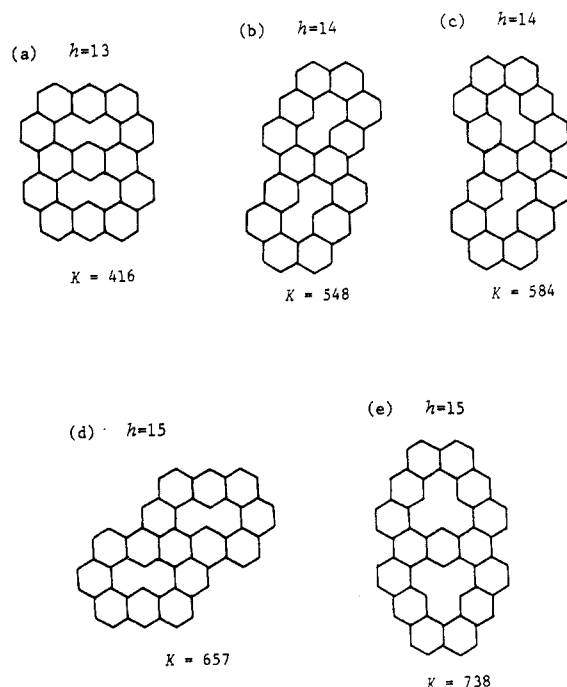


Figure 21. Examples of normal double coronoids. K numbers are given.

systems themselves were obtained by several methods: (a) benzenoids were circumvented by hexagons and then deleted to form the hole of the coronoids; (b) derivation on the basis of Figure 14; and (c) computerized procedure by systematic addition of hexagons to the two ends of unbranched cata-condensed benzenoids until the two ends meet. Primarily this program also produced successively all the unbranched cata-condensed benzenoids. A special version was designed for the detection of primitive coronoids for a given upper limit of h in such a way that many of the benzenoids were discarded

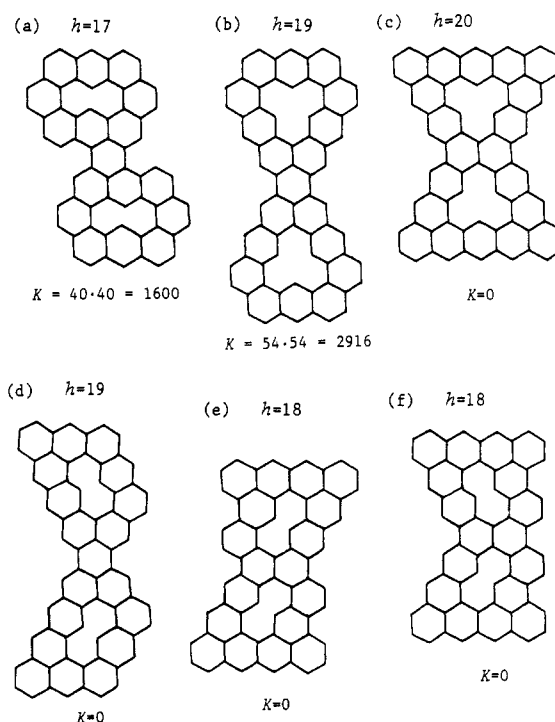


Figure 22. Examples of (a, b) essentially disconnected and (c-f) concealed non-Kekuléan double coronoids.

during the process if the two ends were too far apart to have a chance to meet eventually.

A picture of every generated coronoid may be displayed on the screen. This facilitates a convenient inspection of the systems.

CONCLUSION

The numbers of all (true) coronoids up to $h = 12$ are summarized in Table I. A coarse classification in terms of Δ values is also given therein. Up to $h = 11$ a more detailed classification is described in the text. The Kekuléan coronoids (cf. Table II) are divided into catacondensed (primitive or annelated), normal pericondensed, or essentially disconnected systems. We have all the material necessary for a corresponding classification of the 1954 coronoids with 12 hexagons. For the sake of brevity this classification is not treated here.

The primitive coronoids, and in particular the hollow hexagons among them (cf. Figure 14), are treated in some detail. Also, some general properties of essentially disconnected and non-Kekuléan (obvious and concealed) coronoids are pointed out. Finally some different kinds of double coronoids are exemplified.

The numbers of Kekulé structures (K) are reported for selected systems, including all systems with $h = 8, 9$, and 10 . Two approaches were used for determining the K numbers: (a) analytically by the Randić method of fragmentation³⁴ and by the fragmentation method in a new symmetry-adapted version³⁵ and (b) by a computerized method employing the determinant of the adjacency matrix, basically following the principles of Brown.³⁶ A detailed analysis of the K numbers of the 128 Kekuléan coronoids with 11 hexagons is in progress.

ACKNOWLEDGMENT

Financial support to B.N.C. from the Norwegian Research Council for Science and the Humanities is gratefully acknowledged.

REFERENCES AND NOTES

- (1) Balaban, A. T.; Harary, F. *Tetrahedron* **1968**, *24*, 2505.
- (2) Balaban, A. T. *Tetrahedron* **1969**, *25*, 2949.

- (3) Balaban, A. T. *MATCH* 1976, 2, 51.
- (4) Balaban, A. T. *Rev. Roum. Chim.* 1981, 26, 407.
- (5) Balasubramanian, K.; Kaufman, J. J.; Koski, W. S.; Balaban, A. T. *J. Comput. Chem.* 1980, 1, 149.
- (6) Knop, J. V.; Szymanski, K.; Jeričević, Ž.; Trinajstić, N. *J. Comput. Chem.* 1983, 4, 23.
- (7) Trinajstić, N.; Jeričević, Ž.; Knop, J. V.; Müller, W. R.; Szymanski, K. *Pure Appl. Chem.* 1983, 55, 379.
- (8) Knop, J. V.; Szymanski, K.; Jeričević, Ž.; Trinajstić, N. *MATCH* 1984, 16, 119.
- (9) Knop, J. V.; Müller, W. R.; Szymanski, K.; Trinajstić, N. *Computer Generation of Certain Classes of Molecules*; Association of Chemists and Technologists of Croatia: Zagreb, 1985.
- (10) He, W.; He, W. *Theor. Chim. Acta* 1985, 68, 301.
- (11) Brunvoll, J.; Cyvin, S. J.; Cyvin, B. N. *J. Comput. Chem.*, in press.
- (12) Doroslovački, R.; Tošić, R., private communication, Novi Sad, Yugoslavia.
- (13) Polansky, O. E.; Rouvray, D. H. *MATCH* 1976, 2, 63.
- (14) Polansky, O. E.; Rouvray, D. H. *MATCH* 1977, 3, 97.
- (15) Vögtle, F.; Staab, H. A. *Chem. Ber.* 1968, 101, 2709.
- (16) Jenny, W.; Baumgarten, P.; Paioni, R. *Proceedings of the Symposium on the Nonbenzenoid Aromatic Compounds*; Sendai, Japan, 1970; p 183.
- (17) Diederich, F.; Staab, H. A. *Angew. Chem.* 1978, 90, 383.
- (18) Staab, H. A.; Diederich, F. *Chem. Ber.* 1983, 116, 3487.
- (19) Vogler, H. *THEOCHEM* 1985, 122, 333.
- (20) Jenny, H.; Peter, R. *Angew. Chem.* 1965, 77, 1027.
- (21) A large selection of names has been suggested for this concept: hexagonal animal, hexanimal, hexagonal polyomino, polyhex, PAH-6 (PAH = polycyclic aromatic hydrocarbon), fusene (catafusene, perifusene). We are using the term benzenoid, but no standard terminology seems to prevail at present.
- (22) Professor N. Trinajstić has informed us privately that the wrong number (48) is due to a typing error in ref 8. It was unfortunately repeated in ref 9. The error has been noticed by several researchers.
- (23) Gutman, I. *Bull. Soc. Chim., Beograd* 1982, 47, 453.
- (24) Smith, F. T. *J. Chem. Phys.* 1961, 34, 793.
- (25) Gordon, M.; Davison, W. H. T. *J. Chem. Phys.* 1952, 20, 428.
- (26) Hall, G. G. *Proc. R. Soc. London, A* 1955, 229, 251.
- (27) Gutman, I.; Cyvin, S. J. *THEOCHEM* 1986, 138, 325.
- (28) Coulson, C. A.; Longuet-Higgins, H. C. *Proc. R. Soc. London, A* 1947, 129, 16.
- (29) Longuet-Higgins, H. C. *J. Chem. Phys.* 1950, 18, 265.
- (30) Balaban, A. T.; Tomescu, I. *MATCH* 1983, 14, 155.
- (31) Cyvin, S. J.; Gutman, I. *J. Serb. Chem. Soc.* 1985, 50, 443.
- (32) Gutman, I. *Croat. Chem. Acta* 1974, 46, 209.
- (33) Dias, J. R. *J. Chem. Inf. Comput. Sci.* 1984, 24, 124.
- (34) Randić, M. *J. Chem. Soc., Faraday Trans. 2* 1976, 72, 232.
- (35) Cyvin, S. J.; Bergan, J. L.; Cyvin, B. N. *Acta Chim. Hung.*, in press.
- (36) Brown, R. L. *J. Comput. Chem.* 1983, 4, 556.

Atomic Physicochemical Parameters for Three-Dimensional-Structure-Directed Quantitative Structure-Activity Relationships. 2. Modeling Dispersive and Hydrophobic Interactions

ARUP K. GHOSE* and GORDON M. CRIPPEN*

College of Pharmacy, University of Michigan, Ann Arbor, Michigan 48109

Received July 22, 1986

In an earlier paper (Ghose A. K.; Crippen, G. M. *J. Comput. Chem.* 1986, 7, 565) the need of atomic physicochemical properties for three-dimensional-structure-directed quantitative structure-activity relationships was demonstrated, and it was shown how atomic parameters can be developed to successfully evaluate the molecular water-1-octanol partition coefficient, which is a measure of hydrophobicity. In the present work the atomic values of molar refractivity are reported. Carbon, hydrogen, oxygen, nitrogen, sulfur, and halogens are divided into 110 atom types of which 93 atomic values are evaluated from 504 molecules by using a constrained least-squares technique. These values gave a standard deviation of 1.269 and a correlation coefficient of 0.994. The parameters were used to predict the molar refractivities of 78 compounds. The predicted values have a standard deviation of 1.614 and a correlation coefficient of 0.994. The degree of closeness of the linear relationship between the atomic water-1-octanol partition coefficients and molar refractivities has been checked by the correlation coefficient of 89 atom types used for both the properties. The correlation coefficient has been found to be 0.322. The low value suggests that both parameters can be used to model the intermolecular interaction. The origin of these physicochemical properties and the types of interaction that can be modeled by these properties have been critically analyzed.

INTRODUCTION

In the process of drug design, medicinal chemists evaluate the binding energy of some closely related ligands with a biological receptor. The explicit structure of the receptor in most cases is unknown. The ultimate objective of any quantitative structure-activity relationship (QSAR) is to portray the receptor by the structural, physicochemical, and biological properties of the ligand. Not only is the task difficult but the inherent weakness of the approach ought to make the portrait misty. Explanation of the simplest biological data, namely, the binding energy of the ligand on the purified receptor, involves (1) the three-dimensional structure of the biological receptor¹ and its conformational flexibility,² (2) knowledge of the active site,¹ (3) the conformational behavior of the ligand,^{3,4} (4) the interaction of the biophase⁵ with the ligand/receptor, and, most important, (5) the interaction of the ligand with the receptor. Each process has its energetic (enthalpic) and entropic contribution. The energetic contribution often is easier to model than the entropic part. Entropy is

related to the flexibility of the ligand and the receptor as well as the structural randomness of the biophase around the ligand and the receptor before and after binding. The complexity of these processes leads to very slow development along this line and urges some method that can allow us a rough estimate of the active site.

Most QSAR approaches therefore correlate the binding energy of the ligand with different physicochemical properties for different parts of the ligand. If these physicochemical properties represent the different types of molecular forces, one can guess the nature of interaction at different regions. The first problem is therefore to identify the possible types of forces in the biomolecular interaction and next to identify the physicochemical properties that can model these forces. Unlike the intermolecular interaction between simple molecules, the biochemical interaction of a drug involves a macromolecule on one side. The macromolecule is assumed to have low flexibility under physiological conditions, and hence the steric fit of the ligand structure at the active site often con-