

CLAR FORMULAS AND KEKULÉ STRUCTURES

Ivan Gutman

Faculty of Science, University of Kragujevac, 34000 Kragujevac, Yugoslavia

Sigrid Obenland and Werner Schmidt

Institut für Organische Chemie der Universität, 8000 München 2, and Biochemisches Institut für Umweltcarcinogene, 2070 Ahrensburg, W. Germany

(received: October 1984)

Abstract

A method has been developed for the calculation of the number of those Kekulé structures of a benzoid hydrocarbon which are represented by its Clar formulas. Although the ratio between this number and the total number of Kekulé structures tends to be of the order of 0.5 for most existing benzoid hydrocarbons, special situations are indicated where it may assume any value between zero and one.

1. Introduction

Since its appearance in 1958 [1], Clar's theory of polycyclic aromatic hydrocarbons has attracted the attention of both experimentalists and theoreticians. A wealth of experimental data, supporting this theory, is collected in Clar's classical treatise [2]. For some more recent experimental work along these lines see Refs. [3 - 5].

Clar's theory stimulated a number of quantum chemical investigations whose aim was to examine whether this new concept is in agreement with (and can be deduced from) quantum theory. These investigations [6 - 9] revealed several, to some extent unexpected, connections between Clar's theory and quantum theory. Thus, in Ref. [7] it was demonstrated that those regions of a benzenoid hydrocarbon where the Clar formulas contain an aromatic sextet, have wave functions similar to benzene. The same regions also give the biggest contribution to the thermodynamic stability [8]. Finally, it has recently been established [9] that the Clar formulas are in good agreement with the electron densities arising from the occupied molecular orbitals of highest energy.

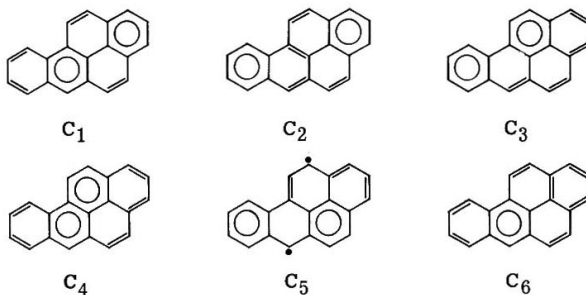
Another impetus to the study of Clar's theory came from Hosoya and Yamaguchi [10] who introduced the concept of generalized Clar formulas and discovered a far-reaching algebra behind them. The algebraic and combinatorial aspects of Clar's theory were thereafter the subject of numerous further studies [11, 12].

2. Clar structures

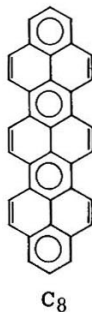
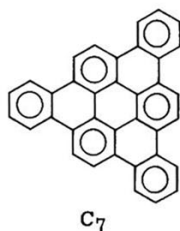
As is well known [2], Clar formulas of a benzenoid hydrocarbon are constructed by drawing circles, representing aromatic sextets, in some hexagons of the benzenoid system. The following three requirements must be obeyed.

- (a) Two circles must not be drawn in neighbouring hexagons.
- (b) The circles must be arranged in a way so that a Kekulé structure can be written for the rest of the molecule.
- (c) The formula must contain the maximum number of circles.

For example, the electronic structure of benzo[a]pyrene is represented by the three Clar formulas C_1 , C_2 and C_3 . The formulas C_4 , C_5 and C_6 are incorrect because they violate rules (a), (b) and (c), respectively.

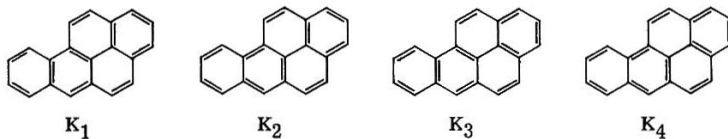


Many benzenoid hydrocarbons have a unique Clar formula. Two typical examples are provided by the recently synthesized tribenzo[a, g, m]coronene (C_7) [5] and teropyrene (C_8) [13]. In these cases the (unique) Clar formula is assumed to be a good representation of the real electronic structure and the chemical behaviour of the molecules.

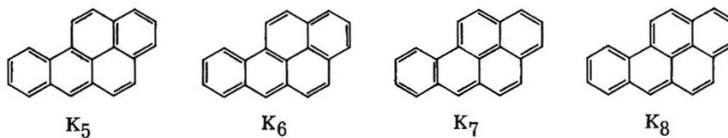


If, on the other hand, the Clar formula is not unique (as, for example, in the case of benzo[a]pyrene, C_1 , C_2 , C_3), then the real electronic configuration of the molecule is understood to be a resonance hybrid of the corresponding Clar-type structures. A quantitative resonance theoretical elaboration of this idea has recently been offered [14].

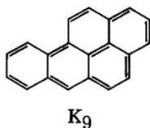
From the early days of Clar's theory [1, 15] it was clear that a Clar formula is a shorthand notation for a group of Kekulé structures. Thus, for example, formula C_1 is a simultaneous representation of the four Kekulé structures K_1 , K_2 , K_3 and K_4 :



In a similar manner, C_2 represents the Kekulé structures K_1 , K_3 , K_5 and K_6 , whereas C_3 represents K_1 , K_5 , K_7 and K_8 :



On the other hand, neither C_1 nor C_2 nor C_3 represent the Kekulé structure K_9 :



Hence, the Clar formulas of benzo[a]pyrene are equivalent to the representation of this molecule by means of eight Kekulé structures.

It is evident that each Clar formula with n sextets represents 2^n Kekulé structures. Therefore, if the Clar formula is unique (e.g. C_7 , C_8), then one immediately concludes that it corresponds to 2^n Kekulé structures.

From the above example it is seen that if the Clar formula is not unique, then different Clar formulas may represent the same Kekulé structures. Consequently, it is in general not easy to find the number of Kekulé structures which are contained in the Clar representation of a given polycyclic aromatic hydrocarbon.

Let us denote a polycyclic aromatic hydrocarbon by H , its number of Kekulé structures by $K(H)$, and the number of Kekulé structures contained in the Clar formulas by $C(H)$.

Whereas the enumeration of the Kekulé structures is a classical problem of the topological theory of benzenoid hydrocarbons [16, 17], methods for the calculation of $C(H)$ seem to be lacking in the literature.

The first attempt towards the determination of $C(H)$ was made by one of the present authors [5]. In this work, we shall communicate a general method for the calculation of this quantity.

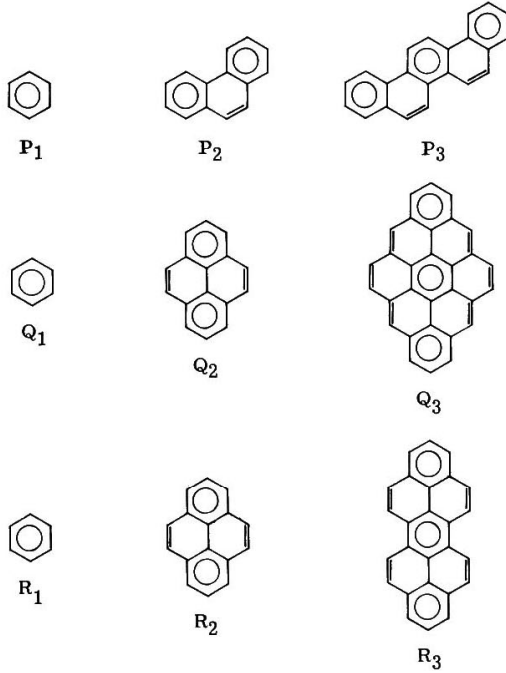
3. On the 50 % rule

It is evident from the preceding considerations that the Clar formulas may, but need not correspond to all Kekulé structures of a polycyclic aromatic hydrocarbon. Consequently, the Clar and the Kekulé structure representation are not fully equivalent. Moreover, Clar's theory can be interpreted as a resonance theory based on a selected number of (significant) Kekulé structures in which the less significant ones are fully neglected (This, however, is not the only possible interpretation of Clar's theory [10, 14]; the real relation of Clar's formalism to resonance and molecular orbital theory is not yet fully understood). Anyway, from the earliest days of Clar's theory [1, 15, 18], the question has been posed how big is the ratio between the Kekulé structures contained in the Clar formulas and the total structure count.

It is an empirical fact that in all existing benzenoid hydrocarbons having a unique Clar formula, this formula represents some 50 % or more of the total number of Kekulé structures [1, 5, 15]. This observation was recently named the "50 % rule".

We wish to point out here that in special situations, when the number of aromatic sextets increases, the 50 % rule may sometimes fail.

As an example consider the following three series of benzenoid molecules: P_n , Q_n , R_n , $n = 1, 2, 3, \dots$, each having a unique Clar formula with n aromatic sextets.



It is known that [16],

$$K(P_n) = \frac{\sqrt{5} + 1}{2\sqrt{5}} \left(\frac{3 + \sqrt{5}}{2}\right)^n + \frac{\sqrt{5} - 1}{2\sqrt{5}} \left(\frac{3 - \sqrt{5}}{2}\right)^n$$

$$K(Q_n) = 2 \cdot 3^{n-1}$$

$$K(R_n) = \binom{2n}{n}$$

Consequently, for sufficiently large values of n ,

$$C(P_n)/K(P_n) \approx \frac{5 - \sqrt{5}}{2} \left(\frac{3 + \sqrt{5}}{4}\right)^{-n}$$

$$C(Q_n)/K(Q_n) \approx (3/2)^{-n+1}$$

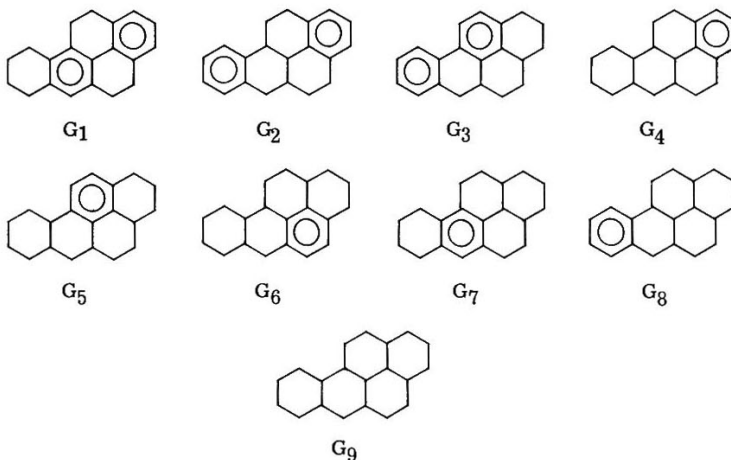
$$C(R_n)/K(R_n) = \sqrt{\pi n} \cdot 2^{-n}$$

In all three cases the ratio $C(H)/K(H)$ decreases exponentially with increasing n . Hence, if n is chosen sufficiently large, the (unique) Clar structure of P_n , Q_n or R_n represents only a negligibly small fraction of the total number of Kekulé structures.

4. Generalized Clar formulas

Hosoya and Yamaguchi [10] considered generalized Clar structures which are constructed by drawing circles in the benzenoid system so that the conditions (a) and (b) (given in section 2) are fulfilled. Condition (c) needs, however, not be fulfilled.

The generalized Clar formulas of benzo[a]pyrene are $G_1 - G_9$. Note that the "empty" formula G_9 is also to be counted among the generalized Clar formulas.



It has been demonstrated [10, 12] that the number of generalized Clar for-

mulas coincides with the Kekulé structure count.* We shall refer to this as to the Hosoya-Yamaguchi rule.

The generalized Clar formulas G_1 , G_2 and G_3 coincide in an obvious manner with the "normal" Clar formulas C_1 , C_2 and C_3 . Some of the other generalized Clar formulas are obtained from G_1 , G_2 and/or G_3 by deleting some of their circles.

We shall say that G_j is a subformula of G_i if G_j can be obtained from G_i by deleting some of its circles. By definition G_i is a subformula of G_i .

Thus, G_4 , G_7 and G_9 are subformulas of G_1 , whereas G_4 , G_8 and G_9 are subformulas of G_2 , etc. The empty formula is a subformula of each of G_1, \dots, G_9 . Note, however, that G_6 is not a subformula of either G_1 or G_2 or G_3 .

An important special case is the following. Suppose that for every hexagon of a benzenoid system one can find a Clar formula having an aromatic sextet in that hexagon. Then all generalized Clar formulas of that benzenoid system are subformulas of the generalized Clar formulas with maximum number of circles. The following result is then obvious from the Hosoya-Yamaguchi rule.

Proposition 1. If all generalized Clar formulas are subformulas of the "normal" Clar formulas, then the Kekulé structure count is equal to the number of subformulas of the "normal" Clar formulas.

* Coronene and related systems are exceptions to this rule. They will be discussed later in this paper.

We can reformulate Proposition 1 in a more formal manner. Label the hexagons of a benzenoid system by A, B, C, \dots . Then there is a one-to-one correspondence between a generalized Clar formula and a set of symbols A, B, C, \dots .

For example, if we label the hexagons of benzo[a]pyrene by A, B, C, D, E as indicated below,



then the set $\{B, D\}$ corresponds to G_1 , the set $\{A, C\}$ to G_3 , the set $\{E\}$ to G_6 etc. The empty set \emptyset corresponds to G_9 .

Let the hexagons of a benzenoid system be labeled by A, B, C, \dots . Let this system have m Clar formulas. Let S_1, S_2, \dots, S_m be the sets of symbols A, B, C, \dots , corresponding to the Clar formulas.

Proposition 2. If all generalized Clar formulas are subformulas of the "normal" Clar formulas, then the Kekulé structure count is equal to the number of subsets of S_1, S_2, \dots, S_m .

Note that neither Proposition 1 nor 2 apply to benzo[a]pyrene. This is because G_6 is not a subformula of either G_1 or G_2 or G_3 . In other words, the set $\{E\}$ is not a subset of either $\{B, D\}$ or $\{B, C\}$ or $\{A, C\}$.

Let us illustrate Proposition 2 (and thus also Proposition 1) using benzo[5]helicene (H_1) as an example.



H_1

The five Clar structures correspond to the sets $\{A, C, E\}$, $\{A, C, F\}$, $\{A, D, E\}$, $\{A, D, F\}$ and $\{B, D, F\}$. It is immediately seen that there are ten two-element subsets: $\{A, C\}$, $\{A, D\}$, $\{A, E\}$, $\{A, F\}$, $\{B, D\}$, $\{B, F\}$, $\{C, E\}$, $\{C, F\}$, $\{D, E\}$ and $\{D, F\}$ and, of course, six one-element subsets: $\{A\}$, $\{B\}$, $\{C\}$, $\{D\}$, $\{E\}$ and $\{F\}$. Together with the empty set \emptyset this gives a total of $5 + 10 + 6 + 1 = 22$.

According to Proposition 2 we claim that H_1 has 22 Kekulé structures.

5. A method for calculating $C(H)$

Consider a benzenoid system H . Let its Clar formulas be C_1, C_2, \dots, C_m . Observe that some of the bonds of H are single in all Clar formulas C_1, C_2, \dots, C_m . Similarly, some of the bonds of H may be double in all Clar formulas. For obvious reasons these single and double bonds are fixed in all Kekulé structures which are contained in C_1, C_2, \dots, C_m and thus play no role in the determination of $C(H)$. Therefore these bonds can be disregarded as long as we are concerned with the calculation of $C(H)$.

Delete from H all these single and double bonds and obtain the (not necessarily connected) benzenoid system H_{red} . We will call it the reduced form of the benzenoid system H .

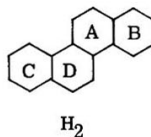
H_{red} has obviously the same Clar formulas as H . Moreover, by dele-

ting all the fixed double and single bonds we have eliminated also all hexagons which are empty in the sense of Clar's theory.* Therefore, for every hexagon of H_{red} one can find a Clar formula having an aromatic sextet in that hexagon. This means that Proposition 2 is applicable to H_{red} .

By deleting all fixed double and single bonds, we have disregarded all Kekulé structures which are not contained in the Clar formulas C_1, C_2, \dots, C_m . Therefrom we arrive at the following important conclusion.

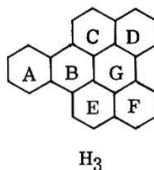
Proposition 3.* The number of Kekulé structures of H_{red} is equal to the number of those Kekulé structures of H which are represented by the Clar formulas. In other words, $C(H) = K(H_{\text{red}})$.

For example, the reduced form of benzo[a]pyrene is chrysene (H_2):



According to Proposition 3, $C(\text{benzo[a]pyrene}) = K(\text{chrysene}) = 8$.

As another illustration consider naphtho[1, 2, 3, 4-ghi]perylene (H_3),



whose reduced form is just the benzoid hydrocarbon H_1 , examined after

* Coronene and related compounds are exceptions, see the subsequent section.

Proposition 2. It has been deduced that H_1 has 22 Kekulé structures.

Therefore $C(H_3) = 22$.

As a matter of fact there are numerous general and efficient methods for the enumeration of Kekulé structures of polycyclic aromatic hydrocarbons [16, 17]. Furthermore, for a large number of benzenoid systems, there exist explicit combinatorial expressions for $K(H)$. Hence, by means of Proposition 3 the finding of $C(H)$ is reduced to a previously solved problem of the finding of $K(H_{red})$.

Combining Propositions 2 and 3 we arrive at the main result of the present paper.

Let H be a benzenoid system and C_1, C_2, \dots, C_m be its Clar formulas. Let each Clar formula contain n aromatic sextets. Label the hexagons of H by A, B, C, \dots . Each Clar formula C_i of H can then be represented by an n -element set S_i of symbols A, B, C, \dots ($i = 1, 2, \dots, m$).

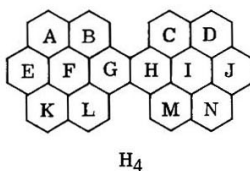
Proposition 4. The number of $C(H)$ of Kekulé structures which are contained in the Clar formulas of H is equal to the number of subsets S_1, S_2, \dots, S_m .

The above proposition provides an efficient and fully general procedure for the calculation of $C(H)$. In order to illustrate it, we shall determine $C(\text{benzo}[a]\text{pyrene})$.

The sets S_1, S_2, S_3 , corresponding to the Clar formulas C_1, C_2, C_3 are $\{B, D\}$, $\{B, C\}$ and $\{A, C\}$. The respective one-element subsets are $\{A\}$, $\{B\}$, $\{C\}$ and $\{D\}$, and the zero-element subset is \emptyset . This gives a total of $3 + 4 + 1 = 8$ subsets, which, of course, coincides with the value

determined in the example after Proposition 3.

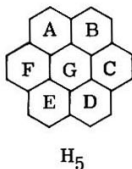
As another example we determine the number $C(H)$ for coroneno-coronene (H_4).



The two Clar formulas of H_4 are described by the 6-element sets $\{A, C, G, J, K, M\}$ and $\{B, D, E, H, L, N\}$. Each set has 2^6 subsets (including the empty set). Since the two sets have no common elements, their only common subset is the empty set. Therefore there are $2^6 + 2^6 - 1$ distinct subsets of $\{A, C, G, J, K, M\}$ and $\{B, D, E, H, L, N\}$ which means that $C(H_4) = 2^6 + 2^6 - 1 = 127$. It is worth noting that $K(H_4) = 364$.

6. Coronene - an exception

As already mentioned, coronene (H_5) and molecules related to it (e.g. H_4) are exceptions to Proposition 3.



This occurs because the two Clar formulas for coronene are $\{A, C, E\}$ and $\{B, D, F\}$, and thus the hexagon G is empty. Nevertheless, H_5 has neither fixed single nor fixed double bonds. Therefore the reduced form of H_5

should coincide with H_5 , what contradicts the fact that hexagon G is empty.

Thus, in the case of coronene-like systems, $C(H)$ cannot be determined by means of Proposition 3. On the other hand, Proposition 4 applies to all benzoid systems without any restriction.

An illustration of the calculation of $C(H)$ of a coronene-like polycyclic aromatic hydrocarbon has already been given in the previous section.

The case of coronene is fully analogous. Since the sets $\{A, C, E\}$ and $\{B, D, F\}$ have no common elements, $C(H_5) = 2^3 + 2^3 - 1 = 15$.

References

- [1] E. Clar and M. Zander, *J. Chem. Soc.*, 1861 (1958).
- [2] E. Clar, *Polycyclic Hydrocarbons*, Academic Press, New York, 1964;
E. Clar, *The Aromatic Sextet*, Wiley, London, 1972.
- [3] E. Clar and W. Schmidt, *Tetrahedron* **35**, 1027 (1979), and earlier papers quoted therein.
- [4] D. Biermann and W. Schmidt, *J. Am. Chem. Soc.* **102**, 3163, 3173 (1980);
D. Biermann and W. Schmidt, *Isr. J. Chem.* **20**, 312 (1980).
- [5] S. Obenland, *Ph.D. Thesis*, University of Munich, 1984.
- [6] R. Pauncz, *J. Chem. Soc.*, 3288 (1960).
- [7] O. E. Polansky and G. Derflinger, *Int. J. Quantum Chem.* **1**, 379 (1967).
- [8] I. Gutman and S. Bosanac, *Tetrahedron* **33**, 1809 (1977).
- [9] M. Aida and H. Hosoya, *Tetrahedron* **36**, 1317 (1980); H. Hosoya, M. Shobu, K. Takano and Y. Fujii, *Pure Appl. Chem.* **55**, 269 (1983).
- [10] H. Hosoya and T. Yamaguchi, *Tetrahedron Letters*, 4659 (1975).

- [11] J. Aihara, *Bull. Chem. Soc. Japan* 49, 1429 (1976); I. Gutman, *Theor. Chim. Acta* 45, 309 (1977); *Z. Naturforsch.* 37a, 69 (1982); S. El-Basil, *Chem. Phys. Letters* 89, 145 (1982); S. El-Basil and I. Gutman, *Chem. Phys. Letters* 94, 188 (1983).
- [12] N. Ohkami, A. Motoyama, T. Yamaguchi, H. Hosoya and I. Gutman, *Tetrahedron* 37, 1113 (1981); I. Gutman, *Match* 11, 127 (1981); N. Okhami and H. Hosoya, *Theor. Chim. Acta* 64, 153 (1983).
- [13] T. Umemoto, T. Kawashima, Y. Sakata and S. Misumi, *Tetrahedron Letters*, 1005 (1975).
- [14] W. C. Herndon and H. Hosoya, to be published.
- [15] E. Clar, C. T. Ironside and M. Zander, *J. Chem. Soc.*, 142 (1959).
- [16] M. Gordon and W. H. T. Davison, *J. Chem. Phys.* 20, 428 (1952); I. Gutman, *Match*, in press.
- [17] T. F. Yen, *Theor. Chim. Acta* 20, 399 (1971); W. C. Herndon, *J. Chem. Educ.* 51, 10 (1974); M. Randić, *J. Chem. Soc., Faraday Trans. 2*, 72, 232 (1976); O. E. Polansky and I. Gutman, *Match* 8, 269 (1980).
- [18] E. Clar, W. Kemp and D. G. Stewart, *Tetrahedron* 3, 325 (1958).