Monatshefte für Chemie 131, 47–54 (2000)

Monatshefte für Chemie
Chemical Monthly

Topological Studies of Dioxo Derivatives of Benzenoid Hydrocarbons: Number of *Kekulé* Structures of Dioxo Derivatives of Prolate Rectangles and a Connection to Hypericin

Ivan Gutman

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

Summary. Among the sixteen possible tautomers of hypericin, the 7,14-dioxo tautomer is thermodynamically the far most stable one. Earlier studies revealed that this is a consequence of the high degree of (stabilizing) cyclic conjugation of the π-electrons in the 7,14-dioxo tautomer, paralleled with a much greater number of *Kekulé* structures than in other tautomers or in the parent benzenoid hydrocarbon. In order to better understand this phenomenon, a generalization is examined: the *Kekulé* structure count in prolate rectangles $R(m,n)$ and their dioxo derivatives $R_{ij}(m,n)$. It is shown that $R(2,3) \equiv$ phenanthro[1,10,9,8-*opqra*]perylene, the parent hydrocarbon of hypericin, and its dioxo derivative $R_{2,2}(2,3) \equiv$ phenanthro[1,10,9,8-*opqra*]perylene-7,14-dione, are exceptional in some of their π-electron properties.

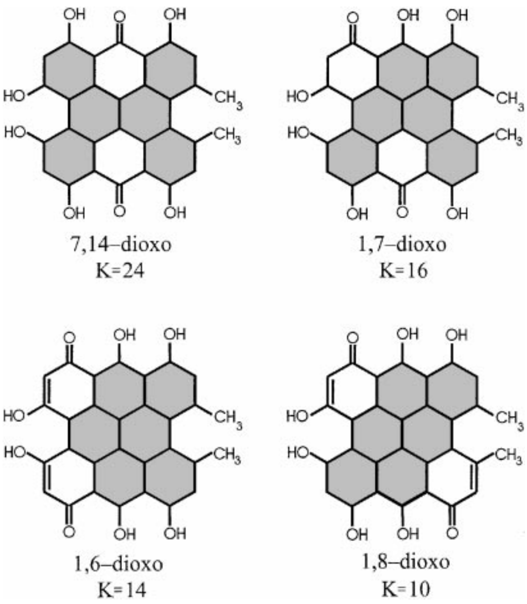
Keywords. Hypericin; *Kekulé* structures; Dioxo derivatives; Benzenoid hydrocarbons; Prolate rectangles.

Topologische Studien an Dioxoderivaten benzenoider Kohlenwasserstoffe: Anzahl der *Kekulé*-Strukturen von Dioxoderivaten gestreckter Rechtecke und eine Verbindung zu Hypericin

Zusammenfassung. Unter den sechzehn möglichen Tautomeren des Hypericins ist das 7,14-Dioxo-Tautomere das thermodynamisch bei weitem stabilste. Frühere Studien haben gezeigt, daß dies eine Folge des hohen (stabilisierenden) Anteils an zyklischer Konjugation der π-Elektronen im 7,14-Dioxo-Tautomer ist, welche von der höchsten Anzahl von *Kekulé*-Strukturen unter allen Tautomeren des benzolischen Grundkörpers begleitet wird. Um dieses Phänomen besser zu verstehen, wird eine Generalisierung untersucht: die Anzahl der *Kekulé*-Strukturen in gestreckten Rechtecken $R(m,n)$ und ihren Dioxoderivaten $R_{ij}(m,n)$. Es wird dabei gezeigt, daß $R(2,3) =$ Phenanthro[1,10,9,8-*opqra*]perylen, der Grundkörper des Hypericins, und sein Dioxoderivat $R_{2,2}(2,3) =$ Phenanthro[1,10,9,8-*opqra*]perylen-7,14-dion einzigartig hinsichtlich einiger ihrer π-elektronischen Eigenschaften sind.

Introduction

Tautomerism in hypericin (Fig. 1; for a review, see Ref. [1]) has recently attracted much interest, both experimentally [2–9] and theoretically [3, 10–13]; tautomerism



7,14–dioxo
K=24

1,7–dioxo
K=16

1,6–dioxo
K=14

1,8–dioxo
K=10

Fig. 1. Four of the 16 theoretically possible tautomers of hypericin [11]; shading indicates the conjugated domains; the most stable tautomer, with the oxo groups in positions 7 and 14, has 24 *Kekulé* structures; the significantly less stable 1,7-dioxo, 1,6-dioxo, and 1,8-dioxo tautomers [11, 12] have 6, 14, and 10 *Kekulé* structures, respectively; these *Kekulé* structure counts should be compared with $K = 16$ of the parent benzenoid hydrocarbon, phenanthro[1,10,9,8-*opqra*]perylene

and isomerism in some structurally closely related compounds have also been studied [14–16]. All these researchers lead to the conclusion that – if intermolecular interactions (including solvent effects) are neglected – the 7,14-dioxo species is the thermodynamically most stable tautomer of hypericin. For instance, the heat of formation of the 7,14-dioxo tautomer was found to be 45 and 83 kJ/mol below the heats of formation of the next-stable 1,7- and 1,6-dioxo tautomers, respectively [12].

It has been demonstrated that the stability order of the tautomers of hypericin is essentially the same as the stability order of the dioxo derivatives of the parent benzenoid hydrocarbon phenanthro[1,10,9,8-*opqra*]perylene [16]. In other words, the dominant factors determining the (relative) stability of hypericin tautomers are the same as the factors determining the (relative) stability of the dioxo derivatives of the respective benzenoid hydrocarbon. These latter factors are reasonably well understood [17] and can be traced back to the effects of cyclic conjugation of the π-electrons. A simple, yet not quite unreliable measure of the extent of cyclic conjugation in a benzenoid hydrocarbon is the number of its *Kekulé* structures [17, 18].

In fact, phenanthro[1,10,9,8-*opqra*]perylene-7,14-dione (cf. Ref. [19]) is the only dioxo derivative of phenanthro[1,10,9,8-*opqra*]perylene having more *Kekulé* structures ($K = 24$) than the parent hydrocarbon itself ($K = 16$). This observation led to the introduction of the concept of 'topological activation' [20, 21].

Let B be a benzenoid hydrocarbon and B_{ij} its dioxo derivative with the oxo groups in positions i and j. Let $K(B)$ and $K(B_{ij})$ be the respective *Kekulé* structure

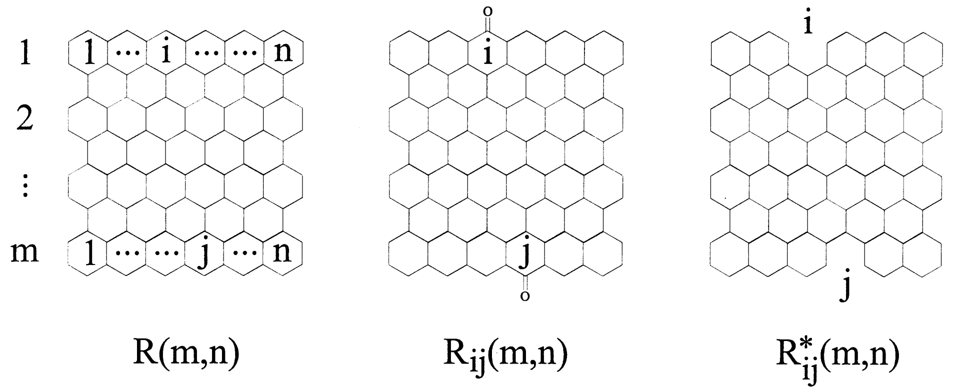


Fig. 2. The prolate rectangle $R(m,n)$, its dioxo derivative $R_{ij}(m,n)$, and the auxiliary benzenoid system $R_{ij}^{*}(m,n)$; note that $K(R_{ij}(m,n)) = (R_{ij}^{*}(m,n))$

counts. Then the sites i and j of B are said to be 'topologically activated, topologically indifferent', and 'topologically deactivated' if $K(B_{ij})/K(B) > 1$, $= 1$, and < 1, respectively [20, 21]. Oxidation, *Diels-Alder* addition, and similar reactions which interrupt the conjugated π-electron system of the molecule B usually occur at topologically activated positions; the most stable tautomer of hypericin is that with oxo groups in the (unique!) topologically activated positions. If a benzenoid hydrocarbon has no topologically activated sites, it is chemically very stable. If, in turn, it has several such sites, it is highly reactive, sensitive to air, *etc.* [20].

Phenanthro[1,10,9,8-*opqra*]perylene, the parent hydrocarbon of hypericin, is a member of the class of benzenoid hydrocarbons classified as 'prolate rectangles' [18], denoted here by $R(m,n)$, the structure of which is depicted in Fig. 2. In particular, phenanthro[1,10,9,8-*opqra*]perylene $\equiv R(2,3)$.

In the following the systems $R(m,n)$ and the topological activation/deactivation of their sites are examined. For this purpose it is necessary to know the number of *Kekulé* structures of $R(m,n)$ and the respective dioxo derivatives $R_{ij}(m,n)$ (see Fig. 2). The analysis will show that $R(2,3)$ is quite exceptional with regard to the topological activation/deactivation of its sites.

Methods

Kekulé structures have been enumerated for quite a few classes of benzenoid hydrocarbons, especially by the outstanding work of *Sven Cyvin* (see for instance Refs. [22–29], [18], and the references cited therein). The number of *Kekulé* structures of prolate rectangles has been known for some time [30]:

$$K(R(m,n)) = (n+1)^m \tag{1}$$

It, thus, remains to find an expression for $K(R_{ij}(m,n))$. In order to simplify the notation, $K_{ij}(m,n)$ shall be written instead of $K(R_{ij}(m,n))$ in the following. For reasons that will become evident below, $K_{ij}(m,n)$ will be considered as the (i,j)-entry of a matrix $\mathbf{K}(m,n)$. Clearly, $\mathbf{K}(m,n)$ is a symmetric square matrix of order n.

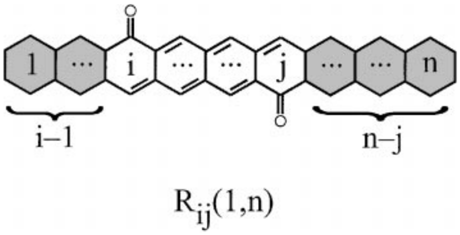


Fig. 3. The dioxo derivative $R_{ij}(1,n)$ and its conjugated domains

The first case considered will be that with $m = 1$. From Fig. 3 can be seen that $R_{ij}(1,n)$ has two disjoint conjugated domains, both being linear polyacenes. Because a linear polyacene with h hexagons has $h+1$ *Kekulé* structures, it follows that

$$K_{ij}(1,n) = \begin{cases} i(n-j+1) & \text{if} \quad i \le j \\ j(n-i+1) & \text{if} \quad i \ge j \end{cases} \tag{2}$$

It is easy to see that the dioxo derivative $R_{ij}(m,n)$ has the same number of *Kekulé* structures as the benzenoid system $R_{ij}^*(m,n)$ (Fig. 2). In order to find an expression for $K_{ij}(m,n)$, color, as usual [17, 18], the vertices of $R_{ij}^*(m,n)$ by two colors, say black and white, so that adjacent vertices have different colors (Fig. 4). Then any double bond in any *Kekulé* structure of $R_{ij}^*(m,n)$ lies between a black and a white vertex. Therefore, $R_{ij}^*(m,n)$ has equally many $(= 2mn + m-1)$ black and white vertices.

Consider the horizontal cut C depicted in Fig. 4 and observe that above C there is one black vertex in excess, whereas below C there is one white vertex in excess., Consequently, in every *Kekulé* structure of $B_{ij}^*(m,n)$ exactly one of the edges intersected by C must be a double bond.

Now those *Kekulé* structures of $R_{ij}^*(m,n)$ in which the double bond intersected by C is located between the l^{th} hexagon of the $(m-1)^{\text{th}}$ row and the l^{th} hexagon of the m^{th} row, with l being a fixed integer between 1 and n, are counted. Then the fragment of $R_{ij}^*(m,n)$ above C has $K_{il}(m-1,n)$ *Kekulé* structures, whereas the fragment below C has $K_{lj}(1,n)$ *Kekulé* structures. Therefore, the number of *Kekulé* structures of $B_{ij}(m,n)$ with the intersected double bond in position l is $K_{il}(m-1,n) \cdot K_{lj}(1,n)$, and the total number of *Kekulé* structures is

$$K_{ij}(m,n) = \sum_{l=1}^{n} K_{il}(m-1,n) \cdot K_{lj}(1,n).$$

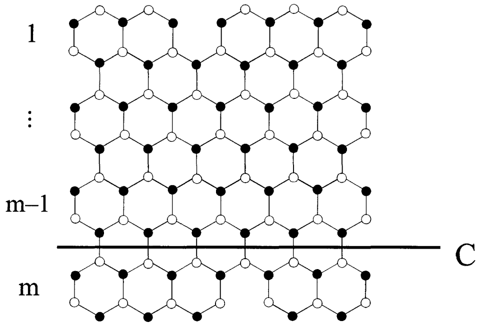


Fig. 4. The system $R_{ij}^*(m,n)$ with colored vertices; note that above the horizontal cut C there are $(2n+1)(m-1)$ black and $(2n+1)(m-1) - 1$ white vertices (hence one black vertex in excess), whereas below the cut C there are $2n$ black and $2n+1$ white vertices (hence one white vertex in excess)

Bearing in mind the rules for matrix multiplication we see that

$$\mathbf{K}(m,n) = \mathbf{K}(m-1,n) \cdot \mathbf{K}(1,n). \tag{3}$$

Identity (3) holds for all $n \geq 1$ and $m \geq 2$. By successively substituting $m = 2,3,\ldots$ into Eq. (3) we get

$$\begin{aligned}
\mathbf{K}(2,n) &= \mathbf{K}(1,n) \cdot \mathbf{K}(1,n) = \mathbf{K}(1,n)^2 \\
\mathbf{K}(3,n) &= \mathbf{K}(2,n) \cdot \mathbf{K}(1,n) = \mathbf{K}(1,n)^2 \cdot \mathbf{K}(1,n) = \mathbf{K}(1,n)^3 \\
\mathbf{K}(4,n) &= \mathbf{K}(3,n) \cdot \mathbf{K}(1,n) = \mathbf{K}(1,n)^3 \cdot \mathbf{K}(1,n) = \mathbf{K}(1,n)^4
\end{aligned}$$

etc. In the general case,

$$\mathbf{K}(m,n) = \mathbf{K}(1,n)^m. \tag{4}$$

By combining Eqs. (2) and (4) it is possible to easily compute the values of $K_{ij}(m,n)$ for any i, j, m, n. In the special cases of $m = 2$ and $m = 3$, the following explicit expressions for the number of *Kekulé* structures are obtained:

Table 1. Number of *Kekulé* structures of the dioxo derivatives $R_{ij}(m,n)$ of prolate rectangles (cf. Fig. 2) for $n = 2,3,4$ and $m = 2,3,4,5$; below each *Kekulé* structure count the respective quotient $K(R_{ij}(m,n))/K(R(m,n)) = K(R_{ij}(m,n))/(n+1)^m$ is given; note that because of the symmetry of the system $R(m,n), K_{ij}(m,n) = K_{ji}(m,n) = K_{n+1-i,n+1-j}(m,n)$ for all values of m and n; therefore, the data given in Table 1 cover all possible cases

	$m = 2$	$m = 3$	$m = 4$	$m = 5$
$n = 2$				
$i = 1, j = 1$	5	14	41	122
	0.56	0.52	0.51	0.50
$i = 1, j = 2$	4	13	40	121
	0.44	0.48	0.49	0.50
$n = 3$				
$i = 1, j = 1$	14	84	552	3728
	0.88	1.31	2.16	3.64
$i = 1, j = 2$	16	112	768	5248
	1.00	1.75	3.00	5.13
$i = 1, j = 3$	10	76	536	3696
	0.63	1.19	2.09	3.61
$i = 2, j = 2$	24	160	1088	7424
	1.50	2.50	4.25	7.25
$n = 4$				
$i = 1, j = 1$	30	330	4125	53350
	1.20	2.64	6.60	17.07
$i = 1, j = 2$	40	510	6600	86075
	1.60	4.08	10.56	27.54
$i = 1, j = 3$	35	490	6525	85800
	1.40	3.92	10.44	27.46
$i = 1, j = 4$	20	295	4000	52900
	0.80	2.36	6.40	16.93
$i = 2, j = 2$	65	820	10650	139150
	2.60	6.56	17.04	44.53
$i = 2, j = 3$	60	805	10600	138975
	2.40	6.44	16.96	44.47

If $i \leq j$, then

$$K_{ij}(2,n) = \frac{1}{6} i(n+1)(n+1-j)(1 - i^2 - j^2 + 2j(n+1))$$

$$\begin{aligned} K_{ij}(3,n) = \frac{1}{360} i(n+1)^2(n+1-j)(3i^4 - 15i^2 + 10i^2j^2 - 15j^2 + 3j^4 \\ + 12 + 2j(n+1)(15 - 10i^2 - 6j^2 + 4j(n+1) + 4(n+1)^2), \end{aligned} \qquad (5)$$

whereas for $i > j, K_{ij}(2,n) = K_{ji}(2,n)$ and $K_{ij}(3,n) = K_{ji}(3,n)$.

For $m \geq 4$, the analogous expressions are still more complicated.

Results and Discussion

Equation (4) together with Eq. (1) enables one to determine which pairs of sites i,j of prolate rectangles are topologically activated. In Tables 1 and 2, the K-values of the systems $R_{ij}(m,n)$ for $2 \leq n \leq 5$ and $2 \leq m \leq 5$ as well as the respective quotients $K(B_{ij})/K(B)$ are given. Recall that the sites i,j are topologically activated if this quotient is greater than unity.

An inspection of the data given in Tables 1 and 2 reveals the following:

1. If $n = 2$ and $m \geq 2$, then no two sites of $R(m,n)$ are topologically activated.
2. If $n = 3$ and $m \geq 3$ and also if $n \geq 4$ and $m \geq 2$, the system $R(m,n)$ has many topologically activated sites; more precisely: with only two exceptions, *all* pairs of sites i, j are topologically activated. Furthermore, the value of the quotient

Table 2. Number of *Kekulé* structures of the dioxo derivatives $R_{ij}(m,n)$ of prolate rectangles (cf. Fig. 2) for $n = 5$ and $m = 2,3,4,5$; for details, see Table 1

	$m = 2$	$m = 3$	$m = 4$	$m = 5$
$n = 5$				
$i = 1, j = 1$	55	1001	21307	471185
	1.53	4.63	16.44	60.59
$i = 1, j = 2$	80	1672	36608	814528
	2.22	7.74	28.25	104.75
$i = 1, j = 3$	81	1863	41877	938223
	2.25	8.63	32.31	120.66
$i = 1, j = 4$	64	1568	35968	810656
	1.78	7.26	27.75	104.25
$i = 1, j = 5$	35	889	20651	467281
	0.97	4.12	15.93	60.09
$i = 2, j = 2$	136	2864	63184	1409408
	3.78	13.26	48.75	181.25
$i = 2, j = 3$	144	3240	72576	1625184
	4.00	15.00	56.00	209.00
$i = 2, j = 4$	116	2752	62528	1405504
	3.22	12.74	48.25	180.75
$i = 2, j = 5$	64	1568	35968	810656
	1.78	7.26	27.75	104.25
$i = 3, j = 3$	171	3753	83835	1876689
	4.75	17.38	64.69	241.34

$K(R_{ij}(m,n))/K(R(m,n))$ is a rapidly increasing function of n and m and reaches its maximum when the oxo groups in $R_{ij}(m,n)$ adopt central positions. Exceptional cases are the sites $i=1, j=4$ in R(2,4) (which are topologically deactivated) and $i=1, j=5$ in R(2,5) (which are topologically indifferent).

3. Only in the case $n=3$, $m=2$ the system $R(m,n)$ has just one pair of topologically activated sites: $i=2$, $j=2$. In other words, among all prolate rectangles, only R(2,3), phenanthro[1,10,9,8-*opqra*]perylene, has a unique pair of topologically activated sites. This makes R(2,3) exceptional within the class of prolate rectangles, and $R_{2,2}$(2,3), phenanthro[1,10,9,8-*opqra*]perylene-7,14-dione, exceptional within the class of dioxo derivatives of R(2,3).

Whether the above observation is of any relevance for the chemistry of hypericin and/or for the understanding of its biological role and phylogenic origin is not clear. It, nevertheless, shows that hypericin possesses a π-electron system with distinguished and atypical topological properties.

Acknowledgements

The author thanks Prof. *Heinz Falk* (Linz, Austria) for valuable discussions on the physico-chemical properties of hypericin, and, in particular its tautomerism. Eq. (5) was suggested by the anonymous referee.

References

[1] Falk H (1999) Angew Chem Int Ed **38**: 3116

[2] Falk H, Schmitzberger W (1992) Monatsh Chem **123**: 731

[3] Etzlstorfer C, Falk H, Müller N, Schmitzberger W, Wagner UG (1993) Monatsh Chem **124**: 751

[4] Etzlstorfer C, Falk H, Oberreiter M (1993) Monatsh Chem **124**: 923

[5] English DS, Zhang W, Kraus GA, Petrich JW (1997) J Am Chem Soc **119**: 2980

[6] English DS, Das K, Ashby KD, Park J, Petrich JW, Castner EW (1997) J Am Chem Soc **119**: 11585

[7] Amer AM, Falk H, Tran HTN (1998) Monatsh Chem **129**: 1237

[8] Kapinus EI, Falk H, Tran HTN (1999) Monatsh Chem **130**: 623

[9] Dax T, Falk H, Kapinus EI (1999) Monatsh Chem **130**: 827

[10] Etzlstorfer C, Falk H (1994) Monatsh Chem **125**: 955

[11] Gutman I, Marković Z, Solujić S, Sukdolak S (1998) Monatsh Chem **129**: 481

[12] Etzlstorfer C, Falk H (1998) Monatsh Chem **129**: 855

[13] Petrich JW, Gordon MS, Cagle M (1998) J Phys Chem **A 102**: 1647

[14] Etzlstorfer C, Falk H, Müller N (1993) Monatsh Chem **124**: 431

[15] Etzlstorfer C, Falk H (1993) Monatsh Chem **124**: 1031

[16] Gutman I, Marković Z (1998) Monatsh Chem **129**: 1019

[17] Gutman I, Cyvin SJ (1989) Introduction to the Theory of Benzenoid Hydrocarbons. Springer, Berlin Heidelberg New York Tokyo

[18] Cyvin SJ, Gutman I (1988) Kekulé Structures in Benzenoid Hydrocarbons. Springer, Berlin Heidelberg New York Tokyo

[19] Falk H, Vaisburg AF (1995) Monatsh Chem **126**: 361

[20] Gutman I (1998) J Serb Chem Soc **63**: 987

[21] Gutman I (1999) J Mol Struct (Theochem) **460**: 47

[22] Cyvin SJ (1982) Monatsh Chem **113**: 1127

[23] Cyvin SJ (1983) Monatsh Chem **114**: 13
[24] Cyvin SJ (1986) Monatsh Chem **117**: 33
[25] Cyvin SJ (1987) Monatsh Chem **118**: 337
[26] Gutman I, Cyvin SJ (1987) Monatsh Chem **118**: 541
[27] Cyvin SJ (1989) Monatsh Chem **120**: 243
[28] Cyvin SJ, Cyvin BN, Brunvoll J, Gutman I (1991) Monatsh Chem **122**: 771
[29] Gutman I, Cyvin SJ, Chen RS, Lin KR (1993) Monatsh Chem **124**: 117
[30] Yeh TF (1971) Theor Chim Acta **20**: 399

Received June 4, 1999. Accepted June 28, 1999