match no. 3 pp. 87 - 96 1977

THE COMBINATORIAL ENUMERATION OF MESO-IONIC HETEROCYCLIC ANNULENES

C. W. HAIGH

Department of Chemistry
University College, Swansea,
Great Britain.

(received: May 1977)

A simple formula is derived for the number of meso-ionic heterocyclic [m] - annulenes obeying the Hückel rule. The treatment is extended in order to allow for restrictions on certain pairs of atoms being adjacent.

Balaban (with Tefeman)[1] enumerated the heterocyclic annulenes obeying the Hückel rule, and later Balaban and Harary[2] expressed the formulae more concisely by the use of Pólya's Theorem. Balaban[3] has recently pointed out that the enumeration of the subset of these compounds which are meso-ionic remains an unsolved problem: we now present a simple solution.

Let there be $\underline{m}$ atoms in the ring, which may consist of three types of atoms only, X, Y and Z. Then the number of distinct structural isomers with $\underline{m}_X$ of type X, $\underline{m}_Z$ of type Z (and $\underline{m}_Y$ of type Y) is given[2] by the coefficient of $\underline{x}^{\underline{m}_X} \underline{z}^{\underline{m}_Z}$ in:

$$F_m = Z (D_m ; 1 + x + z). \qquad (1)$$

Here $\underline{Z}\ (D_{\underline{m}})$ is Pólya's cycle index[4] for the dihedral group $D_{\underline{m}}$; and in the standard manner, the $\underline{p}$'th power of the indeterminate $\underline{s}_{\underline{k}}$ is replaced by $(1 + \underline{x}^{\underline{k}} + \underline{z}^{\underline{k}})^{\underline{p}}$. We shall represent this coefficient by the symbol $\underline{C}_{\underline{m}}\ (\underline{m}_X, \underline{m}_Z)$. It is obvious that

$$m_X + m_Y + m_Z = m \ . \qquad (2)$$

The formula (1) is quite general. But now [1-3] if we specify X to be an atom providing two π-electrons (such as pyrrolic nitrogen), Y one providing one π-electron (e.g. carbon) and Z no π-electrons (e.g. boron), and assume the Hückel rule to be fulfilled, we have[1] the further restriction

$$2 m_X + m_Y = 4 n + 2, \qquad (3)$$

with $\underline{n}$ an integer. In the Figure, we illustrate all the solutions of equations (1 - 3) for the simple but chemically interesting case $\underline{m} = 5$, classified in accordance with the concepts defined in the remainder of this paper.

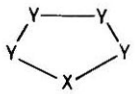
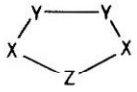
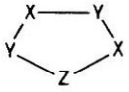
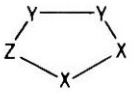
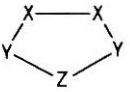
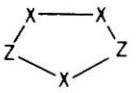
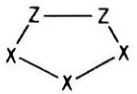
	Non-meso-ionic	Meso-ionic
I. Without X-X or Z-Z Adjacencies		
$m_X = 1,\ m_Z = 0$		—
$m_X = 2,\ m_Z = 1$		
II. With X-X but not Z-Z Adjacencies		
$m_X = 2,\ m_Z = 1$		
$m_X = 3,\ m_Z = 2$		—
III. With X-X and Z-Z Adjacencies		
$m_X = 3,\ m_Z = 2$		—

FIGURE — Classification of Heterocyclic [5]-Annulenes obeying the Hückel rule.

Balaban[1,3] defines a meso-ionic compound, in terms of the above symbolism, as one containing no odd-numbered chains of Y between X and/or Z. This is easily seen to be a useful generalisation of the more usual definition, which states that, if X, Y and Z and the compound are neutral, a meso-ionic compound has no canonical valence-bond structures which do not contain formal charges. Now it is easily seen[1,3] from (3) that $\underline{m}_Y$ must be even. It is accordingly convenient to turn the definition round, and state that a non-meso-ionic compound only contains even chains of Y. Therefore, in such compounds, Y atoms can only occur in adjacent pairs Y_2. In order to enumerate these non-meso-ionic compounds, our combinatorial question is now: 'In how many distinct ways can $\underline{m}_X$ atoms of type X, $\frac{1}{2}\underline{m}_Y$ groups Y_2 and $\underline{m}_Z$ atoms of type Z (total $\overline{\underline{m} - \frac{1}{2}\underline{m}_Y}$ atoms or groups) be arranged?' And the answer is immediately seen to be $C_{\underline{m} - \frac{1}{2}\underline{m}_Y}(\underline{m}_X, \underline{m}_Z)$. We have thus obtained the simple result that the number of meso-ionic compounds is given, with $\underline{m}$, $\underline{m}_X$, $\underline{m}_Z$ and $\underline{m}_Y$ as previously defined, by

$$C_m^{meso}(m_X, m_Z) = C_m(m_X, m_Z) - C_{m - \frac{1}{2}m_Y}(m_X, m_Z) . \quad (4)$$

This is illustrated in the Figure and the Table for some simple cases, which agree with earlier tabulations.[1,3]

Balaban[1,3] has also pointed out that, on grounds of likely chemical stability, it might be interesting to enumerate subsets of the heterocyclic annulenes obeying certain restrictions on which pairs of atoms could be adjacent (e.g. no Z - Z adjacency permitted). The general solution to this problem has recently been found by Lloyd.[5] Although he employs the picture-language

TABLE: Some Examples of the Enumeration of Heterocyclic Annulenes

				Unrestricted			Z-Z adjacency forbidden			X-X and Z-Z adjacency forbidden		
m	m_X	m_Y	m_Z	Total[×]	Non-meso-ionic[+]	Meso-ionic	Total[××]	Non-meso Ionic[++]	Meso-ionic	Total[×××]	Non-meso-ionic[+++]	Meso-ionic
4	2	2	0	2	1	1				1	0	1
5	1	4	0	1	1	0						
	2	2	1	4	2	2				2	1	1
6	1	4	1	3	2	1						
	2	2	2	11	4	7	7	2	5	5	1	4
7	0	6	1	1	1	0						
	1	4	2	9	4	5	6	2	4			
	2	2	3	18	6	12	4	1	3	3	1	2
8	0	6	2	4	2	2	3	1	2			
	1	4	3	19	6	13	6	1	5			
	2	2	4	33	9	24	2	0	2	2	0	2

[×] $C_m(m_X, m_Z)$ [+] $C_{m-\frac{1}{2}m_Y}(m_X, m_Z)$ [××] $D_m(m_X, m_Z)$ [++] $D_{m-\frac{1}{2}m_Y}(m_X, m_Z)$ [×××] $E_m(m_X, m_Z)$ [+++] $E_{m-\frac{1}{2}m_Y}(m_X, m_Z)$

of stringing coloured beads on a necklace, normal in the combinatorial literature in this context,[4] we shall here express his result in chemical terminology. Let m, m_X, m_Y and m_Z be as previously defined. Let there be certain restrictions on permissible adjacencies specified (vid. infra) by a 3 x 3 matrix $\underset{\sim}{T}$ and a vector $\underset{\sim}{b}$, of dimension 3. Then the number of distinct structural isomers is given[5] by the coefficient of $x_1^{m_X} x_2^{m_Y} x_3^{m_Z}$ in the expression

$$R_m = \tfrac{1}{2} Z(C_m ; s) + \tfrac{1}{2} \theta_m, \qquad (5)$$

where, for m odd, $\theta_m = \underset{\sim}{v}' (\underset{\sim}{T}_2)^{\frac{1}{2}(m-1)} \underset{\sim}{b}$, (5a)

and, for m even, $2\theta_m = \underset{\sim}{v}' (\underset{\sim}{T}_2)^{\frac{1}{2}(m-2)} \underset{\sim}{T}\underset{\sim}{e} + \underset{\sim}{d}' (\underset{\sim}{T}_2)^{\frac{1}{2}(m-2)} \underset{\sim}{b}$. (5b)

Here $Z(C_m)$ is the Pólya cycle-index[4] for the cyclic group C_m; and for the p'th power of the indeterminate s_k, the following non-standard substitution is made:

$$s_k^p \longrightarrow \text{tr}\,[(\underset{\sim}{T}_k)^p] \qquad (6)$$

where tr stands for the trace of a matrix. The unsymmetrical matrix $\underset{\sim}{T}$ has zero entries for $(\underset{\sim}{T})_{ij}$ if the i'th and j'th atom are not permitted to be adjacent, otherwise $(\underset{\sim}{T})_{ij} = x_j$. Similarly b_i is zero if atoms of type i may not be adjacent to each other, but otherwise $b_i = 1$. The matrices $\underset{\sim}{T}_k$ are identical in form to $\underset{\sim}{T}$, but all entries x_j are replaced by their k'th power, x_j^k. The row vector $\underset{\sim}{v}'$ has $v_i = x_i$. The column vector $\underset{\sim}{e}$ has all its elements equal to 1. The row vector $\underset{\sim}{d}'$ has $d_i = (\underset{\sim}{T}_2)_{ii}$.

If, for example, we apply the restriction that only Z - Z adjacency is forbidden, we have

$$\underset{\sim}{T} = \begin{bmatrix} x_1 & x_2 & x_3 \\ x_1 & x_2 & x_3 \\ x_1 & x_2 & 0 \end{bmatrix} \quad \text{and} \quad \underset{\sim}{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

We shall denote the coefficients in (5) obtained under these circumstances by $D_{\underline{m}}(\underline{m}_X, \underline{m}_Z)$. Lloyd[6] has calculated them for $\underline{m} = 7$; we have obtained them for the lower values of $\underline{m}$. If the restriction is that neither X - X nor Z - Z adjacencies are permitted, then

$$\underset{\sim}{T} = \begin{bmatrix} 0 & x_2 & x_3 \\ x_1 & x_2 & x_3 \\ x_1 & x_2 & 0 \end{bmatrix} \quad \text{and} \quad \underset{\sim}{b} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Such coefficients we denote by $\underline{E}_{\underline{m}}(\underline{m}_X, \underline{m}_Z)$. For example, our expansion for $\underline{m} = 6$ is

$$R_6 = x_1^3 x_3^3 + x_2(x_1^3 x_3^2 + x_1^2 x_3^3) + x_2^2(x_1^3 x_3 + 5 x_1^2 x_3^2 + x_1 x_3^3) +$$
$$+ x_2^3(x_1^3 + 4 x_1^2 x_3 + 4 x_1 x_3^2 + x_3^3) + x_2^4(2 x_1^2 + 3 x_1 x_3^2 + 2 x_3^2) + x_2^5(x_1 + x_3) + x_2^6.$$

Our numerical checks for both types of coefficient agree with earlier tabulations.[1,3]

If we now consider the enumeration of those heterocyclic annulenes which are subject to any adjacency restrictions and which are also meso-ionic, it is immediately clear that the previous logic is entirely applicable here also, so that, in an obvious symbolism, e.g.,

$$D_m^{meso}(m_X, m_Z) = D_m(m_X, m_Z) - D_{m-\frac{1}{2}m_Y}(m_X, m_Z), \tag{7}$$

and $$E_m^{meso}(m_X, m_Z) = E_m(m_X, m_Z) - E_{m-\frac{1}{2}m_Y}(m_X, m_Z). \tag{8}$$

Numerical checks agree with previous tabulations[1,3] and are illustrated in the Figure and Table.

The expressions (1) and (5), and our results (4), (7) and (8) can readily be generalised, and provide complete explicit solutions to these enumeration problems.†

† It is perhaps worth noting that (1) and (5) can also be used for saturated rings, where of course the limitation (3) is inapplicable, and X, Y, Z can stand for any desired type of atom. E.g. $E_6(2,2) = 5$ also means that there are five structurally isomeric diaza-disila-cyclohexanes which contain neither adjacent nitrogen atoms nor adjacent silicon atoms.

APPENDIX

We have found that, in practice, Lloyd's method[5] is laborious. We here present an alternative method of enumerating the cases where Z atoms are not allowed to be adjacent. Though neither so elegant or general, it is much quicker, especially for $\underline{m}_Z \leq 3$. It is based on the same idea used for meso-ionic compounds in the main part of this paper.

For the case $\underline{m}_Z = 2$, we obtain by exact analogy the result

$$D_m(m_X, 2) = C_m(m_X, 2) - C_{m-1}(m_X, 1).$$

When $\underline{m}_Z = 3$, we have two similar terms. But those isomers having triples (three adjacent Z atoms) will in general have been counted twice. We apply the Inclusion - Exclusion Principle, and add a term to enumerate them. However, a triple with a symmetry plane will not have been counted twice, and a final term corrects for this. Thus

$$D_m(m_X, 3) = C_m(m_X, 3) - G_{m-1}(m_X, 1, 1) + C_{m-2}(m_X, 1) - \left(\begin{matrix}\frac{1}{2}\sqrt{m-3} \\ \frac{1}{2}m_X\end{matrix}\right) - \left(\begin{matrix}\frac{1}{2}\sqrt{m-4} \\ \frac{1}{2}m_Y\end{matrix}\right) - \left(\begin{matrix}\frac{1}{2}\sqrt{m-4} \\ \frac{1}{2}m_X\end{matrix}\right).$$

Here $\underline{G}_{\underline{m}}(\underline{m}_X, \underline{m}_Z, \underline{m}_W)$ is the coefficient of $\underline{x}^{\underline{m}_X}\underline{z}^{\underline{m}_Z}\underline{w}^{\underline{m}_W}$ in

$$Z(D_m; 1 + x + z + w).$$

Here and henceforth, wherever a binomial coefficient or a factorial has a fractional argument, that term is to be set equal to zero.

When $\underline{m}_Z = 4$, pairs and triples are similarly enumerated. Quadruples (four adjacent Z atoms), whether general or symmetrical, are counted just once in this way, and no correction is required for them. But it is necessary to consider the occurrence also of two pairs of adjacent Z atoms. Straightforward application of the Inclusion-Exclusion Principle, avoiding double counting of quadruples as pairs of pairs, and allowing for the fact that isomers with pairs of adjacent pairs may have C_2, C_s or C_{2v} symmetry, leads finally to the following formula.

$$D_m(m_X, 4) = C_m(m_X, 4) - G_{m-1}(m_X, 2, 1) + G_{m-2}(m_X, 1, 1) +$$

$$+ C_{m-2}(m_X, 2) - C_{m-3}(m_X, 1) - \tfrac{1}{2}\frac{(\frac{1}{2}\overline{m-2})!}{(\frac{1}{2}m_X)!.(\frac{1}{2}m_Y)!} - \tfrac{1}{2}\frac{(\frac{1}{2}\overline{m-4})!}{(\frac{1}{2}m_X-2)!.(\frac{1}{2}m_Y)!} -$$

$$-\tfrac{1}{2}\frac{(\frac{1}{2}\overline{m-4})!}{(\frac{1}{2}m_X)!.(\frac{1}{2}\overline{m_Y-2})!} - \frac{(\frac{1}{2}\overline{m-4})!}{(\frac{1}{2}\overline{m_X-1})!.(\frac{1}{2}\overline{m_Y-1})!} - \frac{(\frac{1}{2}\overline{m-3})!}{(\frac{1}{2}\overline{m_X-1})!.(\frac{1}{2}m_Y)!} -$$

$$- \frac{(\frac{1}{2}\overline{m-3})!}{(\frac{1}{2}m_X)!.(\frac{1}{2}\overline{m_Y-1})!} + \binom{\frac{1}{2}\overline{m-5}}{\frac{1}{2}\overline{m_X-1}}\binom{\frac{1}{2}\overline{m-5}}{\frac{1}{2}\overline{m_Y-1}} - \tfrac{1}{2}\binom{\frac{1}{2}\overline{m-4}}{\frac{1}{2}m_X}.$$

It will be noticed that, for a given parity of m and of m_Z, at most four of the last nine terms are non-zero.

With these formulae, one may enumerate all isomers up to $m = 9$, and many cases for higher m. It is to be noted that in this method individual coefficients can readily be determined. All numerical checks[1,3,6] proved satisfactory. We also present the polynomial for $m = 9$. (We now prefer to use x, y, z rather than x_1, x_2, x_3. The expression is symmetrical in x and y, and we only explicitly give terms with $m_X \leq m_Y$.)

$$R_9 = (y^9 + xy^8 + 4x^2y^7 + 7x^3y^6 + 10x^4y^5) + z(y^8 + 4xy^7 + 16x^2y^6 + 28x^3y^5 + 38x^4y^4) +$$
$$+ z^2(3y^7 + 12xy^6 + 36x^2y^5 + 57x^3y^4) + z^3(3y^6 + 10xy^5 + 28x^2y^4 + 34x^3y^3) +$$
$$+ z^4(y^5 + 3xy^4 + 6x^2y^3) + \text{symmetrical terms.}$$

References

1. A. T. BALABAN, Studii şi Cercetări Chim. Acad. R.P. România, 7, 257 (1959); Appendix I in collaboration with S. Teleman.

2. A. T. BALABAN and F. HARARY, Rev. Roumaine Chim. 12, 1511 (1967).

3. A. T. BALABAN, in "Chemical Applications of Graph Theory", (A. T. Balaban, Ed.), Academic Press, London, 1976, p.70.

4. F. HARARY and E.M. PALMER, "Graphical Enumeration", Academic Press, New York, 1973, Chap. 2.

5. E. K. LLOYD, in "Combinatorics", Proceedings of the British Combinatorial Conference, Aberystwyth 1973 (London Mathematical Society Notes no.13), Cambridge University Press, 1974, p.97.

6. E. K. LLOYD, private communication.