

There are two cases where the normalized ion ASED-MO approach disagrees with experiment,  $\text{MgMn}_2\text{O}_4$  and  $\text{CoMn}_2\text{O}_4$ , both of which are mixed normal. However, the manganite family is extremely variable experimentally in its SPE. The crystal field approach also has trouble with these compounds.

Generally, the anion preference energy is important for compounds where the two cations have very different atomic numbers. It is least important for adjacent members of the periodic table. Cation preference energies are less simply characterized, being subject to both average d shift and splitting effects.

The d splitting site preference energies based on ASED-MO calculations are analogous to the empirically determined crystal field site preference energies of Dunitz and Orgel.<sup>6a</sup> A comparison of the two, given in Figure 7, shows the ASED-MO crystal field octahedral site preference energies overestimate the empirical values. The ASED-MO crystal field structure preference energies also overestimate the empirical ones, in a similar way, as shown in the case of the ferrites in the lower part of Figure 8. The other site preference energy components are required in order to make correct structure predictions in the cases of the Mg, Al, Mn, and Zn ferrites, as shown in the upper part of Figure 8. Comparison of the total structure preference energies with those of the Mott-Littleton potential modeling technique<sup>7a,b</sup> shows agreement in sign, but the ASED-MO values are larger by a factor of about 2. It is significant that the two approaches yield the same trend. The ASED-MO normalized ion approach details the physical reasons in terms of electronic structure. It is interesting that the two-body energy contributions to the structure preference energies,

which can be determined from the data in Table II by subtracting the anion and cation preference energies from the structure preference energies, are small, and in only one case,  $\text{TiMg}_2\text{O}_4$ , does this energy determine the structure preference.

### Conclusion

Predictions for the site preference energies of oxide spinels have been made for 50 compounds. The consideration of the ASED-MO cation and anion orbital energy levels reveals why a particular cation distribution is dominant over another. This is an improvement over the crystal field approach, which considers only the cation energies in an approximate way. The theoretical cation structure preference energy has been partitioned into two portions so that the crystal-field-like terms can be separately evaluated. The theoretical crystal field component is just as good (or bad) as crystal field theory for predicting whether an oxide spinel is normal or inverse. The two-body repulsion site preference energy is responsible for the structure in only one of the 50 compounds. The anion and two-body repulsion site preference energies provide the necessary corrections to the crystal field energies.

**Acknowledgment.** The initial stages of this work were funded by the National Science Foundation through Grant DRM81-19425 to the Materials Research Laboratory at Case Western Reserve University. R.W.G. is grateful to the Department of Chemistry and Computing Center at the University of Keele for the use of their facilities and to Professor C. R. A. Catlow for informative discussions.

## Aromaticity as a Quantitative Concept. 1. A Statistical Demonstration of the Orthogonality of "Classical" and "Magnetic" Aromaticity in Five- and Six-Membered Heterocycles

Alan R. Katritzky,<sup>\*,||</sup> Piotr Barczynski,<sup>||,‡</sup> Giuseppe Musumarra,<sup>§</sup> Danila Pisano,<sup>§</sup> and Mirosław Szafran<sup>||,‡</sup>

*Contribution from the Department of Chemistry, University of Florida, Gainesville, Florida 32611, Dipartimento di Scienze Chimiche, Università di Catania, Viale A. Doria, 8-95125 Catania, Italy, and the Department of Chemistry, A. Mickiewicz University, Poznań, Poland. Received January 19, 1988*

**Abstract:** Twelve characteristics comprising readily available geometrical, energetic, and magnetic data for the nine compounds benzene, pyridine, pyrimidine, pyrazine, thiophene, furan, pyrrole, pyrazole, and imidazole are assembled and treated by principal component analysis (PCA). Three principal components (PC) are found which account for 83% of the variance of the data. Values for the characteristics of the individual compounds recalculated from the scores and loadings are in good agreement with those used in the treatment. Scores are then estimated from the limited available data for an additional seven compounds (pyridazine, s-triazine, 1,2,4-triazine, thiazole, oxazole, isoxazole, and 1,2,4-triazole) by fitting them into the PC model; satisfying agreement is also found between the observed and recalculated values of the characteristics. This means that the scores and loadings can be used with some confidence to predict values of characteristics not available. The first and second PC scores for the whole group of 16 compounds divide them up into the four principal chemical groups of heterocyclic aromatics: (a) pyridine (and benzene), both positive; (b) other azines,  $t_1$  positive and  $t_2$  negative; (c) five-member heteroaromatics with one heteroatom,  $t_1$  negative and  $t_2$  positive; and (d) azoles,  $t_1$  and  $t_2$  both negative (except oxazole which lies in group (c)). The loadings for the characteristics divide them up into three groups: in group (a)  $I_{5,6}$ ,  $\Delta N$  and (to a somewhat lesser extent) DRE,  $^{15}\text{N}$ , and HSRE are dominated by  $t_1$ , whereas in group (b)  $\chi_m$  and  $\Lambda$  are almost independent of  $t_1$  but strongly dependent on  $t_2$  and  $t_3$ . This indicates that the "classical" and "magnetic" concepts of aromaticity are almost completely orthogonal. The other characteristics show hybrid dependence.

Aromaticity is arguably the most important general concept for the understanding of organic chemistry in general and of heterocyclic chemistry in particular.<sup>1-6</sup> Its influence is ubiquitous

in determining stability and reactivity, the nature of the reaction products to be expected, the symmetry and geometry of molecules,

<sup>||</sup> University of Florida.

<sup>‡</sup> Università di Catania.

<sup>§</sup> On leave from A. Mickiewicz University.

(1) Bergmann, E. D.; Pullman, B. Eds. *Aromaticity, Pseudo-Aromaticity, Anti-Aromaticity*; Israel Academy of Science and Humanities: Jerusalem, 1971; Jerusalem Symp. Quant. Chem. Biochem., Vol. III.

and numerous physicochemical properties. It is inconceivable to attempt to teach or practice organic chemistry without reference to the concept of aromaticity, and suggestions that it should be abandoned (because of difficulties to be discussed later) are quite unrealistic.

It is equally clear that aromaticity is not merely a yes/no concept: while most compounds are undisputably aromatic or nonaromatic, borderline cases do occur. Furthermore, few would dispute that, for example, furan is less aromatic than benzene. Hence, the concept of aromaticity is a quantitative, as well as a qualitative one, and many attempts have been made to define numerical scales or measures of aromaticity. It is here that the difficulties really start. Briefly, it has been found impossible to isolate any single well-defined, measurable or calculable, parameter that characterizes quantitatively the degree of aromaticity of a molecule or ion and which is even moderately acceptable.

The difficulties have arisen from a variety of grounds:

(i) Experimental measures of any one parameter are frequently not available for a full range of compounds.

(ii) Errors in experimental measurements arise from the need to use data obtained by different investigators at different locations, or from the fact that they often represent small differences between large measured values.

(iii) Parameters for heteroatoms for theoretical MO methods have been in the past, and in some cases still are, controversial.

(iv) A particular uncertainty, affecting theoretical and experimental methods alike, is the need for comparison of actual aromatic compounds with nonaromatic models, and the precise definition of these models is not easy.

It is because of such difficulties that the concept of aromaticity has remained in such an unsatisfactory state from a quantitative viewpoint. In the present paper we report the results of our examination of principal component analysis (PCA) as a possible method to clarify the situation.

#### Mathematical Foundation of Principal Component Analysis<sup>7</sup>

Principal component analysis (PCA) is a relatively straightforward method of transforming a given set of principal components (PC) that are orthogonal (uncorrelated) to each other. In contrast to multi-regression analysis (MRA), in PCA no particular assumption about the underlying structure of the variables is required.

The principal component model may be described by eq 1

$$x_{ik} = \bar{x}_{ik} + \sum_{a=1}^A t_{ia}p_{ak} + e_{ik} \quad (1)$$

where  $\bar{x}_{ik}$  is the mean scaled value of the experimental quantities (variables) (scaling weights,  $w_k$ , transfer  $\bar{x}_{ik}$  to unscaled data,  $\bar{x}'_{ik} = w_k^{-1} \bar{x}_{ik}$ );  $t_{ia}$  are scores;  $p_{ak}$  are loadings;  $e_{ik}$  are residuals;  $i$  is the chemical compound (object);  $k$  is the experimental measurement (variable); and  $a$  is the principal component.

The number of PC (scores) existing in characteristic vector space is equal to, or less than, the number of variables in the data set. The first principal component is defined as the best summary of linear relationship exhibited in the data. The second component may be viewed as the second best linear combination of variables that accounts for the most residual variance when the effect of

**Table I.** Characteristics 1–3: Aromaticity Structural Indexes for Compounds 1–16

no.	compd	1 $I_{S(6)}^a$	2 RC <sup>b</sup>	3 $\Delta N^c$
1	benzene	100	1.751	0
2	pyridine	85.7	1.731	0.09
3	pyrimidine	84.3	1.727	0.16
4	pyrazine	88.8	1.739	0.12
5	pyridazine	78.9	1.716	0.17
6	1,3,5-triazine	100	1.724	
7	1,2,4-triazine	86.1		0.14
8	thiophene	66		
9	furan	43	1.430	0.43
10	pyrrole	69	1.463	0.31
11	pyrazole	73	1.297	0.19
12	imidazole	64	1.423	0.28
13	thiazole	64		0.28
14	oxazole	38	1.392	
15	isoxazole	42	1.361	
16	1,2,4-triazole	81		0.14

<sup>a</sup> From ref 11a,b. <sup>b</sup> From ref 12. <sup>c</sup> From ref 10b.

**Table II.** Characteristics 4–8: Energetical Properties of Compounds 1–16

no.	compd	4 DRE <sup>a</sup>	5 HSRE <sup>b</sup>	6 $\Delta H_A^c$	7 $\Delta H_{F(\text{exp})}^d$	8 $\Delta H_{F(\text{MND})}^d$
1	benzene	22.6	0.390	57.16	19.8	21.2
2	pyridine	23.1	0.348	51.79	34.8	28.7
3	pyrimidine	20.2	0.297	46.99	47.0	35.5
4	pyrazine	17.1	0.293	46.44	46.9	38.3
5	pyridazine			45.59	66.4	44.2
6	1,3,5-triazine				40.8	40.8
7	1,2,4-triazine				52.4	52.4
8	thiophene	6.5	0.193	39.24	27.6	30.8
9	furan	4.3	0.044	41.69	-8.3	-8.7
10	pyrrole	5.3	0.233	44.77	25.9	32.4
11	pyrazole		0.330		43.3	45.4
12	imidazole	15.4	0.251	39.66	30.6	33.2
13	thiazole					
14	oxazole				-3.7	-8.3
15	isoxazole				18.8	19.2
16	1,2,4-triazole			34.29		44.6

<sup>a</sup> From ref 13a,b. <sup>b</sup> From ref 16a,b. <sup>c</sup> From ref 14a. <sup>d</sup> From ref 14b–e.

the first component is removed from the data. Subsequent components are defined similarly until all the variance in the data is exhausted.

Equation 1 was solved by using the SIMCA method.<sup>7b,e,k,l</sup> The

(2) (a) Breslow, R. *Pure Appl. Chem.* **1971**, *28*, 111. (b) Gerson, F. *Pure Appl. Chem.* **1971**, *28*, 131. (c) Hafner, K. *Pure Appl. Chem.* **1971**, *28*, 153. (d) Hoffmann, R. *Pure Appl. Chem.* **1971**, *28*, 181. (e) Johnson, A. W. *Pure Appl. Chem.* **1971**, *28*, 195. (f) Takeshi Nakajima *Pure Appl. Chem.* **1971**, *28*, 219. (g) Prinzbach, H. *Pure Appl. Chem.* **1971**, *28*, 281. (h) Sondheimer, F. *Pure Appl. Chem.* **1971**, *28*, 331. (i) Fogel, E. *Pure Appl. Chem.* **1971**, *28*, 355. (j) West, R. *Pure Appl. Chem.* **1971**, *28*, 379.

(3) (a) Badger, G. M. *Aromatic Character and Aromaticity*; University Press: Cambridge, 1969. (b) Garratt, P. J. *Aromaticity*; Wiley: New York, 1986.

(4) (a) Cook, M. J.; Katritzky, A. R.; Linda, P. *Adv. Heterocycl. Chem.* **1974**, *17*, 255. (b) Katritzky, A. R.; Lagowski, J. M. *The Principles of Heterocyclic Chemistry*; Chapman and Hall Ltd.: London, 1967. (c) Katritzky, A. R. *Handbook of Heterocyclic Chemistry*; Pergamon: Oxford 1985.

(5) Douben, H. J.; Wilson, J. D.; Laity, J. L. In *Non-Benzenoid Aromatics*; Snyder, J. P., Ed.; Academic: New York, 1971; Vol. II, p 167.

(6) Million, R. B. *Pure Appl. Chem.* **1980**, *52*, 1541.

(7) (a) Kowalski, B. R.; Bender, C. F. *J. Am. Chem. Soc.* **1972**, *94*, 5632. (b) Albano, C.; Dunn, W. J.; Edlund, U.; Johansson, E.; Norden, B.; Sjöström, M.; Wold, S. *Anal. Chim. Acta* **1978**, *103*, 429. (c) Strout, O. *Chemical Pattern Recognition*; Research Studies Press Ltd.: Letchworth, England and Wiley: New York, 1986. (d) Varmuza, K. *Pattern Recognition in Chemistry*; Springer-Verlag, Berlin 1980. (e) Albano, C.; Blomquist, G.; Coomans, D.; Dunn, W. J.; Edlund, U.; Eliasson, B.; Hellberg, S.; Johansson, E.; Norden, B.; Sjöström, M.; Söderström, B.; Wold, H.; Wold, S. *Pattern Recognition by Means of Disjunctive Principal Components Models: Philosophy and Methods*. In *Symposium i Anvendt Statistik*; Haskuldsen, A.; Esbensen, K., Eds.; Copenhagen, 1981; p 183. (f) Draper, N.; Smith, M. *Applied Regression Analysis*, 2nd ed.; Wiley-Interscience: New York, 1981. (g) Jurs, P. C.; Isenhour, T. L. *Chemical Applications of Pattern Recognition*; Wiley Interscience: New York, 1975. (h) Mather, P. M. *Computational Methods of Multivariate Analysis in Physical Geography*; Wiley: London, 1976. (i) Mardia, K. V.; Kent, J. T.; Biddy, J. M. *Multivariate Analysis*; Academic: London, 1979. (j) Malinowski, E. R.; Hovary, D. G. *Factor Analysis in Chemistry*; Wiley Interscience: New York, 1980. (k) Wold, S.; Sjöström, M. SIMCA: a Method for Analysing Chemical Data in Terms of Similarity and Analogy. In Kowalski, B. R. *Chemometrics: Theory and Application*; American Chemical Society: Washington, 1977; ASC Symp. Ser. No. 52, p 243. (l) Wold, S.; Albano, C.; Dunn, W. J.; Esbensen, K.; Hellberg, S.; Johansson, E.; Sjöström, M. *Pattern Recognition: Finding and Using Regularities in Analysis*; Martens, H.; Russwurm, H., Eds.; Applied Science: London, 1983; p 147. (m) Wold, S.; Albano, C.; Dunn, W. J.; Edlund, U.; Esbensen, K.; Hellberg, S.; Johansson, E.; Lindberg, W.; Sjöström, M. In *Chemometrics, Mathematics and Statistics in Chemistry*; Kowalski, B. R., Ed.; Reidel Publishing Company: Dordrecht, 1983; p 17–95. (n) Massart, D. L.; Vandeginste, B. G. M.; Deming, S. N.; Michotte, Y.; Kaufman, L. *Chemometrics: a textbook*; Elsevier: Amsterdam, 1988.

**Table III.** Characteristics 9–12: Magnetic Properties of Compounds 1–16

no.	compd	9 $\chi_M^a$	10 $\Delta^b$	11 $I_1^c$	12 $15N^d$
1	benzene	54.8	13.7	1.000	
2	pyridine	49.2	13.4	0.978	–63.5
3	pyrimidine	43.1			–84.5
4	pyrazine	37.6	7.1		–46.1
5	pyridazine	40.5			20.2
6	1,3,5-triazine	37.9			–98.5
7	1,2,4-triazine				40.0, 2.0, –62.0 (–6.66) <sup>e</sup>
8	thiophene	57.4	13.0	0.836	
9	furan	43.1	8.9	0.821	
10	pyrrole	47.6	10.2	0.866	–224.6
11	pyrazole	42.6	6.6	0.784	–173.1, 79.8 (–126.5) <sup>e</sup>
12	imidazole	44.3		0.751	–169.0
13	thiazole	50.6	12.3	0.749	–58.0
14	oxazole	39.2			–123.7
15	isoxazole	37.5			2.2
16	1,2,4-triazole				–127.4, –134.7 (–131.0) <sup>e</sup>

<sup>a</sup> From ref 4a, 5, and 20a–c. <sup>b</sup> From ref 4a. <sup>c</sup> From ref 28b. <sup>d</sup> From ref 29. <sup>e</sup> The average value reported in parentheses was used.

SIMCA method has already been applied in physical organic chemistry.<sup>8</sup>

**Objectives of the Present Study.** The objective behind the present study was to assemble for a representative number of heteroaromatic compounds as full as possible a set of characteristics associated with aromaticity (including both measured and calculated properties) and to determine the principal components. These components could then be tested for predictive power against the characteristics of other heterocycles not used in the original survey and in principle used for the estimation of values of the characteristics which were previously unavailable.

PCA allows the examination of a set of characteristics (variables) of a class of compounds (objects) to investigate the relations between them. It enables the identification of one, two, three, or more PC derived from the characteristics for the compounds examined; these components have defined values for each of the compounds ( $t_{1j}$ ,  $t_{2j}$ ,  $t_{3j}$ , the “scores”) and are taken in certain proportions ( $p_{1k}$ ,  $p_{2k}$ ,  $p_{3k}$ , etc., the “loadings”) for each type of characteristic.

Graphical representations of these values, the “scores” plot for the compounds and the “loadings” plot for the characteristics, provide pictures that allow the recognition of systematic patterns that are otherwise difficult to deduce from the original data matrix.

A multivariate statistical treatment is particularly suitable for heterocyclic compounds, where the presence of one or more heteroatoms in the ring adds new dimensions to the problems generally investigated in LFERs.

New examples of application of PC analysis in heterocyclic chemistry include the applicability of the Hammett equation to five-membered heterocycles<sup>81,m</sup> and also the simultaneous dependence of  $S_N2$  rates on alkyl group structure and leaving group nucleofugacity in nucleophilic displacements in which heterocycles act as leaving groups.<sup>8n</sup>

**Data Set Used for PCA.** The compounds (objects) were selected mainly on the criteria of availability of a rather full set of available characteristic data: benzene, pyridine, pyrimidine, pyrazine, thiophene, furan, pyrrole, pyrazole, and imidazole. The data used as characteristics in the PCA (variables) are collected in Table I (Geometrical Criteria), Table II (Energetical Criteria), and Table III (Magnetic Criteria). We now consider in turn available parameters of each type and justify the selection we made.

**Geometrical Criteria.** Cyclic conjugation leads to a leveling out of the differences in bond length between formal double and single bonds. A molecule is regarded as aromatic when the lengths of the carbon–carbon bonds are in the range 1.36–1.43 Å, i.e., are close to the value of 1.397 Å found in benzene. The deviation of the lengths of the ring bonds from such median values may in principle be used to evaluate the aromaticity of heterocycles.

Jug<sup>9</sup> proposed the aromaticity index  $A_1$ , which characterizes the extent of uniformity of the peripheral C–C bonds. Jug later introduced a new index  $A$ , which takes into consideration all the ring bonds:  $A = A_1 A_2$ . The correction term  $A_2$  accounts for the resistance to the cyclic conjugation due to  $\pi$ -electrons localized on charged atoms. Although the index  $A$  is rather frequently cited, it cannot be considered satisfactory. Different types of bonds such as C–C and C–N or C–O can be of the same length and yet have different bond orders, hence the use of simple bond lengths for evaluating heterocyclic aromaticity is flawed. Consideration of bond orders is more justified.

The sum of the differences in bond orders ( $\sum \Delta N$ ) was used as an aromaticity index in the case of five-membered heterocycles with one heteroatom.<sup>10a</sup> It was recently replaced by the  $\Delta N$  index which was proposed by Pozharskii.<sup>10b</sup> This index is the average of the fluctuations of all the ring bonds (eq 2) where  $n$  is the total number of differences between each possible pair of bond orders in the rings. We have selected  $\Delta N$  as a characteristic (no. 3).

$$\Delta \bar{N} = \sum \Delta N / n \quad (2)$$

A different approach was recently devised by Bird,<sup>11</sup> who proposed the calculation of aromaticity indexes  $I_5$  and  $I_6$ . These indexes also take into account the degree of statistical uniformity of the bond orders of the ring but use the differences between the actual bond orders and the arithmetic mean of these bond orders.

$$V = (100/N_0) \sqrt{\sum (N - N_0)^2 / n} \quad (3)$$

The coefficient of variation for the bond orders of a particular heterocycle are given by eq 3 where  $N_0$  is the arithmetic mean of the various bond orders  $N$  and  $n$  is the number of bonds. For a fully delocalized heterocycle,  $V$  will have the value 0. For a nondelocalized Kekulé form with alternating single and double bonds, the value of  $V$  depends upon the type of ring system. Thus for five-membered ring heterocycles  $V_K = 35$ , and for six-membered heterocycles  $V_K = 33.3$ . To place the values of  $V$  on a more convenient scale than, e.g., 0–35, the calculated  $V$  is substituted into eq 4. For ring systems of differing size, a subscript such as  $I_5$  or  $I_6$  was attached. We use this definition of  $I_{5(6)}$  as a characteristic (no. 1).

$$I = 100(1 - V/V_K) \quad (4)$$

An alternative approach to aromaticity, with a justification based on the ring current concept, was developed by Jug.<sup>12</sup> With use of the semiempirical SINDO-1 method, the bond orders of a large number of aromatic and heteroaromatic structures were calculated. Jug proposed that aromaticity should be evaluated relative not to the extent of leveling of all the bonds but to the bond order of the weakest ring bond. He offered this as a new

(8) (a) Sjöström, M.; Wold, W. *Acta Chem. Scand., Ser. B* **1981**, *35*, 537. (b) (a) Albano, C.; Wold, S. *J. Chem. Soc., Perkin Trans. 2* **1980**, 1447. (b) Alunni, S.; Clementi, S.; Edlund, U. *Acta Chem. Scand., Ser. B* **1983**, *37*, 47. (c) Johnels, D.; Clementi, S.; Dunn, W. J.; Edlund, U.; Grahn, H.; Hellberg, S.; Sjöström, M.; Wold, S. *J. Chem. Soc., Perkin Trans. 2* **1983**, 863. (d) Edlund, U.; Nostrom, A. *Org. Magn. Reson.* **1977**, *9*, 196. (e) Edlund, U.; Wold, S. *J. Magn. Reson.* **1980**, *37*, 183. (f) Musumarra, G.; Wold, S.; Gronowitz, S. *Org. Magn. Reson.* **1981**, *17*, 118. (g) Musumarra, G.; Scarlata, S.; Wold, S. *Gazz. Chim. Ital.* **1981**, *111*, 499. (h) Musumarra, G.; Carbone, D.; Johnels, D.; Edlund, U. *Magn. Reson. Chem.* **1986**, *24*, 209. (i) Sjöström, M.; Wold, S. *J. Chem. Soc., Perkin Trans. 2* **1981**, 104. (j) Musumarra, G.; Ballistreri, F. P.; Muratore, S.; Katritzky, A. R.; Wold, S. *J. Chem. Soc., Perkin Trans. 2* **1982**, 1049. (k) Wold, S. *Technometrics* **1978**, *20*, 397. (l) Alunni, S.; Clementi, S.; Ebert, C.; Linda, P.; Musumarra, G.; Sjöström, M.; Wold, S. *J. Chem. Soc., Perkin Trans. 2* **1985**, 485. (m) Musumarra, G. Side Chain Reactivity of Thiophenes. *Thienyl Derivatives. In Thiophene and its Derivatives*; Gronowitz, S., Ed.; J. Wiley: New York, 1986; Vol. 44, Part 3, Chapter V, pp 975–1153. (n) Musumarra, G.; Bruno, M.; Katritzky, A. R.; Sakizadeh, K.; Alunni, S.; Clementi, S. *J. Chem. Soc., Perkin Trans. 2* **1985**, 1887.

(9) Jug, A., ref 1, p 383.

(10) (a) Pozharskii, A. F. *Tieorieticheskiye osnovy khimii getierociclov*; Khimia: Moscow, 1985. (b) Pozharskii, A. F. *Chem. Heterocycl. Comp.* **1985**, *21*, 717.

(11) (a) Bird, C. W. *Tetrahedron* **1985**, *41*, 1409. (b) Bird, C. W. *Tetrahedron* **1986**, *42*, 89.

(12) Jug, K. E. *J. Org. Chem.* **1983**, *48*, 1344.

aromaticity index (RC) and justified this as the most significant breakdown in cyclic ring current occurs specifically in the vicinity of the weakest ring bond. Therefore, the aromaticity of a compound should be higher with increasing order of this bond. We use RC as characteristic no. 2.

The lack of alternation of bond lengths is not a single and sufficient condition for aromaticity of a cyclic conjugated system. Thus, the geometrical index, despite their considerable appeal, need to be considered along with the other aromaticity criteria.

**Energetic Criteria.** The different thermodynamic stability of aromatic, nonaromatic, and antiaromatic systems is the basis for the development of energy criteria. Quantitative estimates are based on the concept of resonance energy (RE). The RE is defined as the difference between the total  $\pi$ -electron energy  $E_\pi$  of a given conjugated molecule and that of the corresponding hypothetical reference structure with localized double bonds, respectively. Since this reference structure is hypothetical and does not exist, its choice is to some extent arbitrary.

The three most common variants of the RE are the empirical resonance energy,  $ERE^{4a}$  (the difference in energy between that calculated from the empirical bond energies and the observed heat of formation), the Dewar resonance energy,  $DRE^{13-15}$  (eq 5), and the Hess-Schaad resonance energy,  $HSRE^{16}$  (eq 6).

$$DRE = n(E_{CX}(s) - E_{CX}') + m(E_{CX}(d) - E_{CX}'') + E_{\pi b} \quad (5)$$

$$X = C, N, O, S$$

In eq 5,  $n$  is the total number of C-C and C-X bonds,  $m$  is the total number of C=C and C=X bonds;  $E_{CX}(s)$  and  $E_{CX}(d)$  are the  $\sigma$ -bond energy of "single" and "double" bonds in a polyene;  $E_{CX}'$  and  $E_{CX}''$  are the  $\sigma$ -bond energies of "single" and "double" bonds in cyclic polyene;  $E_{\pi b}$  is the  $\pi$ -binding energy.

$$HSRE = E_{HMO} - E_{loc} \quad (6)$$

In eq 6,  $E_{HMO}$  is the total HMO  $\pi$  energy and  $E_{loc}$  is the energy calculated from empirical  $\pi$ -bond energies.

The DRE can be approximated by the difference between the heats of formation ( $\Delta H_F$ ) of a compound and of its reference structure, or by the difference in the corresponding heats of atomization ( $\Delta H_A$ ). The HSRE is the difference between  $\Delta H_A$  and the calculated empirical bond energies.

We have chosen the DRE and HSRE as characteristics no. 4 and 5. With a few exceptions, the calculated resonance energies according to Dewar, and to Hess and Schaad, give the same order of heterocyclic aromaticity.<sup>17</sup>

Among the disadvantages of the DRE and HSRE definitions is that the necessary number of empirical parameters used in additive calculations for localized structures increases rapidly when heteroatoms are present. A new nonparametric variant of the DRE concept is directly related to the topology of the molecular  $\pi$  network. The definition of the topological resonance energy (TRE)<sup>18</sup> was based on the Hückel theory and on the graph theory<sup>19</sup> introduced in the middle of seventies. In the case of conjugated systems with Kekulé structures, TRE's and DRE energies lead to identical predictions, but for other cases the estimates of the TRE are unsatisfactory.<sup>18</sup>

Most of the experimental DRE values given in the literature for heterocycles were obtained from heats of combustion. The data of different authors for the same compounds (for a summary

of these values see ref 4a) often differ considerably, as a consequence of different assumptions for evaluating the heats of combustion of the localized reference structure. For this reason, we have chosen as characteristic no. 6-8 the experimental values of  $\Delta H_A$ ,<sup>14a</sup> the experimental values of  $\Delta H_F$ ,<sup>14b-e</sup> and the calculated values of  $\Delta H_{F(MNDO)}$ ,<sup>14b-e</sup> respectively. The heat of atomization differs from the heat of formation by an amount of bond dissociation enthalpies.

**Magnetic Criteria.** Compounds have been considered as aromatic when induced diamagnetic  $\pi$ -electron ring currents are maintained in the molecule. However, it must be emphasized that ring currents are themselves not physically observable and that other effects, besides the ring current, intrude upon the properties experimentally measurable (such as magnetic susceptibilities, anisotropies and exaltations, <sup>1</sup>H NMR chemical shifts, and Faraday effects). It is difficult to separate out these other effects, and consequently the use of magnetic parameters as criteria for aromaticity has been seriously criticized.<sup>6</sup>

$$\Lambda = \chi_M - \chi_M' \quad (7)$$

Diamagnetic susceptibility exaltation ( $\Lambda$ ) is defined as the difference between the measured molar susceptibility ( $\chi_M$ ) of the conjugated compound and that estimated ( $\chi_M'$ ) according to an additive scheme for a polyene of the same structure (eq 7).<sup>4a,20</sup> Especially large exaltations are observed for aromatic compounds and Palmer et al.<sup>20d,e</sup> proposed using the diamagnetic susceptibility exaltation as a criterion for aromaticity. We have taken  $\chi_M$  and  $\Lambda$  as characteristics no. 9 and 10.

$$\rho = Kn\Lambda/S^2[A] \quad (8)$$

Furthermore, it has been shown<sup>21</sup> that the magnetic exaltation  $\Lambda$  may be transformed into a quantitative aromatic index  $\rho$  defined by eq 8, in which  $n$  is the number of  $\pi$  electrons,  $S$  is the area of the ring, and  $K$  is a scaling constant that can be determined by reference to benzene.

The anisotropy of diamagnetic susceptibility is of great importance in aromatic and especially in heteroaromatic compounds. The heteroatoms within a  $\pi$ -system are magnetically anisotropic since their electron environment is not spherical. It is now accepted<sup>6</sup> that magnetic anisotropy does not itself provide a direct measure of aromaticity. Both the mean molecular magnetizability ( $\chi$ ) and the magnetic anisotropy depend on a fine balance between large, oppositely signed diamagnetic ( $\chi^d$ ,  $\Delta\chi^d$ ) and temperature-independent paramagnetic ( $\chi^p$ ,  $\Delta\chi^p$ ) contributions. It is only the latter that should correlate with electron delocalization.

Schmaltz, Morris, and Flygare<sup>22b</sup> demonstrated that a quantitative measure of electron delocalization in a planar, cyclic molecule can be obtained by comparing the measured out-of-plane magnetizability component, i.e., or the anisotropy component ( $\chi_{33}$ ,  $\Delta\chi$ ), with the values of these properties predicted for a hypothetical structure in which the electron distribution is completely localized ( $\chi_{33,loc}$ ,  $\Delta\chi_{loc}$ ). The differences between the observed and calculated values (i.e.,  $\chi_{33,loc}$ ) provide estimates of the extent of electron delocalization and, by inference, of relative aromaticity. This procedure was applied for determination of the relative aromaticity for several heteroaromatic compounds.<sup>22a</sup>

The Cotton-Mouton effect,<sup>22b,23,24</sup> the Kerr effect,<sup>23,24</sup> dielectric study,<sup>25</sup> the molecular Zeeman effect,<sup>26</sup> and high-field NMR

(13) (a) Dewar, M. J. S.; Harget, A. J.; Trinajstić, N.; Worley, S. D. *Tetrahedron* **1970**, *26*, 4505. (b) Dewar, M. J. S.; Trinajstić, N. *Theor. Chim. Acta* **1970**, *17*, 235.

(14) (a) Dewar, M. J. S.; Harget, A. I.; Trinajstić, N. *J. Am. Chem. Soc.* **1969**, *91*, 6321. (b) Fos, E.; Vilarrasa, J.; Fernandez, J. *J. Org. Chem.* **1985**, *50*, 4894. (c) Dewar, M. J. S.; McKee, M. L. *J. Comput. Chem.* **1983**, *4*, 84. (d) Olivella, S.; Vilarrasa, J. *J. Heterocycl. Chem.* **1981**, *18*, 1189. (e) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4907.

(15) Dewar, M. J. S. *Tetrahedron Suppl.* **1966**, *8*, 75.

(16) (a) Hess, B. S. A.; Schaad, L. S.; Holyoke, C. W. *Tetrahedron* **1975**, *31*, 295. (b) Hess, B. A.; Schaad, L. S. *J. Am. Chem. Soc.* **1973**, *95*, 3907.

(17) Reference 4, p 295.

(18) Rouvray, D. H. In *Chemical Application of Graph Theory*; Balaban, A. T., Ed.; Academic: London, 1976; Chapter 7.

(19) Biggs, N. L.; Lloyd, E. K.; Wilson, R. J. *Graph Theory 1736-1936*; Clarendon Press: Oxford, 1976.

(20) (a) Boudreaux, E. A.; Gupta, R. R. In *Physical Methods in Heterocyclic Chemistry*; Gupta, R. R., Ed.; Wiley: New York, 1984; p 281. (b) Calderbank, K. E.; Calvert, R. L.; Lukins, P. B.; Ritchie, G. L. D. *Aust. J. Chem.* **1981**, *34*, 1835. (c) Davidson, I. R.; Burnham, A. K.; Siegel, B.; Beak, P.; Flygare, W. H. *J. Am. Chem. Soc.* **1974**, *96*, 7394. (d) Palmer, M. H.; Findlay, R. H.; Moyes, W. J. *Chem. Soc., Perkin Trans. 2* **1975**, 841. (e) Palmer, H. M.; Findlay, L. H. *Tetrahedron Lett.* **1974**, 253.

(21) (a) Maouche, B.; Gayoso, J.; Ouamerli, O. *Rev. Roum. Chi.* **1984**, *29*, 613. (b) Gayoso, J.; Ouamerli, O. *Ibid.* **1981**, *26*, 1035.

(22) (a) Schmaltz, T. G.; Grieke, T. D.; Beak, P.; Flygare, W. M. *Tetrahedron Lett.* **1974**, 2885. (b) Schmaltz, T. G.; Morris, L.; Flygare, W. H. *J. Am. Chem. Soc.* **1973**, *95*, 7961.

(23) Battaglia, M. R.; Ritchie, G. L. D. *J. Chem. Soc., Perkin Trans. 2* **1977**, 897.

(24) Dalerband, K. E.; Calvert, R. L.; Lukins, P. B.; Ritchie, G. L. D. *Aust. J. Chem.* **1981**, *34*, 1834.

**Table IV.** Weights, Factors, Means, Principal Component Loadings, Residuals, and Modeling Powers for Aromatic Characteristics (Variables 1–12) in the PCA Model

variables ( <i>k</i> ) <sup>a</sup>	<i>w<sub>k</sub></i> <sup>b</sup>	$\bar{x}_{ik}$	first PC			second PC			third PC		
			<i>P<sub>1k</sub></i> <sup>d</sup>	<i>S<sub>k</sub></i> <sup>2(1)</sup> <sup>e</sup>	$\Psi_k(1)$ <sup>f</sup>	<i>p<sub>2k</sub></i> <sup>d</sup>	<i>S<sub>k</sub></i> <sup>2(2)</sup> <sup>g</sup>	$\Phi_k(2)$ <sup>f</sup>	<i>p<sub>3k</sub></i> <sup>d</sup>	<i>S<sub>k</sub></i> <sup>2(3)</sup> <sup>h</sup>	$\Phi_k(3)$ <sup>f</sup>
<i>I</i> <sub>5(6)</sub> (1)	0.0592	4.432	0.3574	0.0336	0.8168	−0.0088	0.0389	0.8028	−0.0133	0.0463	0.7849
RC (2)	5.411	8.496	0.2645	0.4891	0.3007	0.1917	0.4328	0.3421	0.5202	0.1885	0.5658
$\Delta\bar{N}$ (3)	7.296	1.441	−0.3431	0.0266	0.8369	−0.0175	0.0306	0.8249	0.0274	0.0373	0.8068
DRE (4)	0.1273	1.822	0.3066	0.2155	0.5358	−0.0318	0.2555	0.4945	0.1675	0.2459	0.5041
HSRE (5)	9.771	2.583	0.3362	0.1613	0.5984	−0.1203	0.1363	0.6307	−0.1501	0.1135	0.6631
6 $\Delta H_A$ (6)	0.1625	7.468	0.2785	0.3818	0.3821	0.2962	0.1868	0.5678	0.2252	0.1006	0.6828
$\Delta H_F$ (7)	0.0582	1.728	0.2219	0.7152	0.1543	−0.4383	0.1459	0.6180	−0.1506	0.1246	0.6470
$\Delta H_{F(MNDO)}$ (8)	0.0647	1.846	0.1799	0.8616	0.0718	−0.4465	0.2908	0.4608	−0.3599	0.0606	0.7538
$\chi_M$ (9)	0.1585	7.392	0.0508	1.120	0.0000	0.4072	0.7130	0.1556	−0.6116	0.0231	0.8480
$\Lambda$ (10)	0.3037	3.174	0.1075	1.098	0.0000	0.4106	0.4062	0.3626	−0.2841	0.1934	0.5602
<i>I</i> <sub>1</sub> (11)	10.61	9.151	0.2394	0.5647	0.2486	0.3613	0.1290	0.6409	0.0346	0.1684	0.5897
$\delta(^{15}\text{N})$ (12)	0.0147	−1.744	0.4969	0.1716	0.5857	−0.0567	0.2150	0.5363	0.1377	0.2923	0.4593

<sup>a</sup>See Tables I–III. <sup>b</sup>Factors required to autoscale; dividing each parameter by the corresponding weight makes the PC model comparable to unscaled data. <sup>c</sup>Mean values of the *k* autoscaled experimental quantities of *x<sub>i</sub>*. <sup>d</sup>First, second, and third principal-component loadings. <sup>e</sup>Variable residual variance, for *A* = 1. <sup>f</sup>Modeling powers. <sup>g</sup>Variable residual variance for *A* = 2. <sup>h</sup>Variable residual variance for *A* = 3.

techniques<sup>27</sup> have been employed to determine the anisotropy of electron-density distribution in aromatic molecules (the magnetic susceptibility tensor  $\chi_{ij}$ ). The determination of the value of the tensor  $\chi_{ij}$  from the Cotton–Mouton effect is based on another important characteristic of the electron distribution: the optical polarizability tensor  $\sigma_{ij}$  of a molecule, which also depends on the character of the delocalization of the electrons.<sup>3b</sup> For aromatic systems the highest polarizability is observed in the plane of the  $\pi$ -system and the lowest in the perpendicular direction.<sup>28a</sup>

Bulgarevich et al.<sup>28b</sup> proposed new aromaticity indexes based on the anisotropic polarizability of the molecules. If from the total polarizability of the molecule in the plane of the  $\pi$ -system, ( $b_{11} + b_{22}$ )<sub>mol</sub>, all the contributions in this plane associated with the  $\sigma$ -system, ( $b_{11} + b_{22}$ ) $\sigma$ , are subtracted then on the basis of the additivity approximation, the difference obtained, ( $b_{11} + b_{12}$ ) $\pi$ <sub>mol</sub>, should characterize the ease of displacement of electrons by an applied electric field. This effect must be greater as the delocalization of  $\pi$  electrons in the aromatic ring increases. However, the size of the ring and the type of atoms in it also influence the polarizability of the electron cloud. The quantity obtained is related to the number of bonds in the ring and, for convenience, normalized with respect to the similar quantity calculated for benzene. Thus, the first aromaticity index (our characteristic no. 11) was expressed by eq 9, where *n* is the number of endocyclic bonds.

$$I_1 = [(b_{11} + b_{22})\pi_{\text{mol}}]/n / [(b_{11} + b_{22})\pi_{\text{C}_6\text{H}_6}/6] \quad (9)$$

The second aromaticity index *I*<sub>2</sub> was based on bond polarizability anisotropy and is analogous to the structural index. It is defined as the ratio of the longitudinal polarizability of the formally single C—C bond of the heterocycle to the longitudinal polarizability of the formally double bond C=C. The second index is less universal and is already inapplicable for heterocycles with several heteroatoms, and we have not used it.

$$I_2 = b_L(\text{C—C})/b_L(\text{C=C}) \quad (10)$$

The generation of an induced ring current in aromatic systems leads to deshielding of the external ring protons and uniformity of the coupling constants of vicinal protons of adjacent bonds.

Proton chemical shifts and spin–spin coupling constants are now considered as a good test for aromaticity in general, but only in

**Table V.** Principal Component Scores (*t*<sub>11</sub>, *t*<sub>12</sub>, *t*<sub>13</sub>) for Compounds (Objects) Obtained from the PCA Model and Residual Standard Deviations (RSD)

no.	compd (i)	no. <sup>a</sup>	<i>t</i> <sub>11</sub>	<i>t</i> <sub>12</sub>	<i>t</i> <sub>13</sub>	RSD
1	benzene <sup>b</sup>	11	3.766	2.265	−0.168	0.412
2	pyridine <sup>b</sup>	12	2.490	1.044	−0.001	0.412
3	pyrimidine <sup>b</sup>	10	1.534	−0.8516	0.624	0.412
4	pyrazine <sup>b</sup>	11	1.739	−1.760	1.309	0.412
8	thiophene <sup>b</sup>	9	−1.643	0.7309	−2.433	0.412
9	furan <sup>b</sup>	11	−5.143	1.712	1.411	0.412
10	pyrrole <sup>b</sup>	11	−1.845	0.0235	−0.633	0.412
11	pyrazole <sup>b</sup>	10	−0.2844	−2.456	−0.805	0.412
12	imidazole <sup>b</sup>	11	−1.551	−1.398	−0.190	0.412
5	pyridazine <sup>c</sup>	8	2.419	−2.544	0.7146	0.500
6	1,3,5-triazine <sup>c</sup>	5	1.643	−1.792	1.070	0.600
7	1,2,4-triazine <sup>c</sup>	4	1.800	−5.369	3.268	0.013
13	thiazole <sup>c</sup>	6	0.0676	−1.773	−2.266	0.741
14	oxazole <sup>c</sup>	6	−3.755	1.042	1.945	1.123
15	isoxazole <sup>c</sup>	6	−1.260	−1.012	1.276	1.347
16	1,2,4-triazole <sup>c</sup>	5	−0.1683	−2.044	−1.823	0.571

<sup>a</sup>No of characteristics available. <sup>b</sup>PCA model. <sup>c</sup>Test objects.

individual special cases can they serve as a quantitative criterion for relative aromaticity. This is particularly true for heteroatomic systems, due to the non-uniform distribution of the electron density in these compounds, the change in the dihedral angle as a function of the ring geometry, and the effect of the anisotropy of the heteroatom, all of which give contributions that should not be neglected. We have included (as characteristic no. 12) nitrogen-15 chemical shifts,<sup>29</sup> as this is a simple and well-defined quantity. In compounds containing two or more N atoms, the average <sup>15</sup>N shift was taken.

**Results of the PCA.** Of the 16 compounds (objects) for which characteristics are listed in Tables I–III, we used only 9 in the PCA selecting those compounds for which the most characteristics were available. Results of the PC analyses on these 9 compounds are shown in Table IV. It was found that 83% of the variation of the characteristics was explained by three principal components of which the 1st, 2nd, and 3rd PC accounted for 47%, 22%, and 14%, respectively, of the variance.

The scores for each of the nine compounds used in the PCA are given in Table V (footnote b). We then calculated the scores for the first, second, and third PC for the other seven compounds (test set) by fitting them into the PCA model derived from the reference set. These scores are also reported in Table V (footnote c).

The RSD for compounds in the test set reported in Table V, apart from those of oxazole and isoxazole, are all comparable to

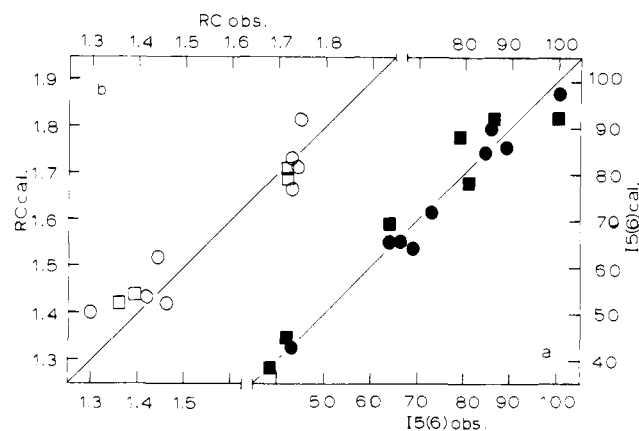
(25) (a) Le Fevre, R. J. W.; Williams, P. H.; Eckert, I. M. *Aust. J. Chem.* **1965**, *18*, 1133. (b) Le Fevre, C. G.; Le Fevre, R. J. W. The Kerr Effect. In *Techniques of Chemistry*; Weissberger, A., Ed.; Wiley Interscience: New York, 1972; Vol. 1, Part IIIC, Chapter 6, p 399.

(26) Davidson, J. R.; Burnham, A. K.; Siegel, B.; Beak, P.; Flyare, W. H. *J. Am. Chem. Soc.* **1974**, *96*, 7394.

(27) van Zijl, P. C. M.; MacLean, C.; Bothner-By, A. A. *J. Chem. Phys.* **1985**, *83*, 4978.

(28) (a) Bulgarevich, S. B.; Bolotnikov, V. S.; Scheinker, V. N.; Osipov, O. A.; Garnowski, A. D. *J. Org. Chem. USSR* **1985**, *12*, 191. (b) Bulgarevich, S. B.; Yusman, T. A.; Osipov, O. A. *J. Gen. Chem. USSR* **1984**, *54*, 1427.

(29) (a) Chinchin, B.; Von Philipsborn, W.; Nagarajan, K. *Helv. Chim. Acta* **1983**, *66*, 1537. (b) Stefaniak, L.; Roberts, I. D.; Witanowski, M.; Webb, G. A. *Org. Magn. Reson.* **1984**, *22*, 201. (c) Iovanovic, M. V. *Spectrochim. Acta* **1984**, *40A*, 637. (d) Fenn, M. D., in ref 20a, p 141.



**Figure 1.** Plots of (a)  $I_{5(6)}(\text{cal.})$  against  $I_{5(6)}(\text{obs.})$  and (b)  $RC(\text{cal.})$  against  $RC(\text{obs.})$ : (○, ●) compounds used in the PCA model (data from Table VI); (□, ■) compounds not used in the PCA model (data from Table VII).

that of the reference set (0.412). This indicates that predictions for these compounds should be almost as good as those for the reference ones, but rather less precise for oxazole and isoxazole.

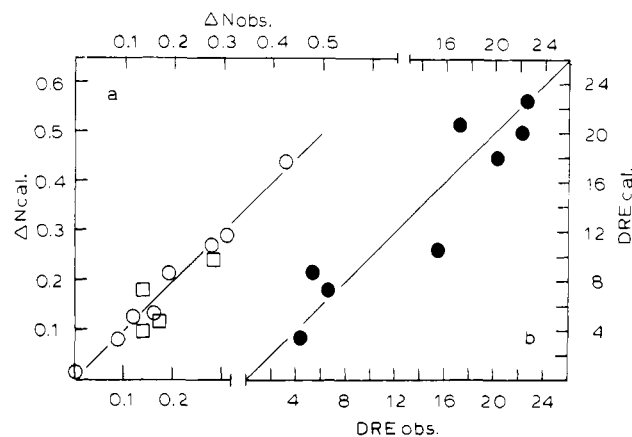
**Calculation of Characteristics (Tables VI and VII).** The next step was to recalculate the values of the characteristics from the appropriate PC parameters (loadings and scores listed in Tables IV and V) with eq 1. On dividing the  $x_{ik}$  values calculated from eq 1 by the corresponding scaling weight ( $w_k$ ), the unscaled values were obtained and are listed in Tables VI and VII.

The results given in Table VI are for the reference set of compounds used in the original PCA; comparison of the original and recalculated values provides a measure of the overall success of the PCA. For the test set of compounds not used in the PCA we can also recalculate the values for all the characteristics by inserting back in eq 1 the loadings and the scores listed in Tables IV and V. The results of such calculations are given in Table VII (footnote b) and show encouraging agreement with the experimental data.

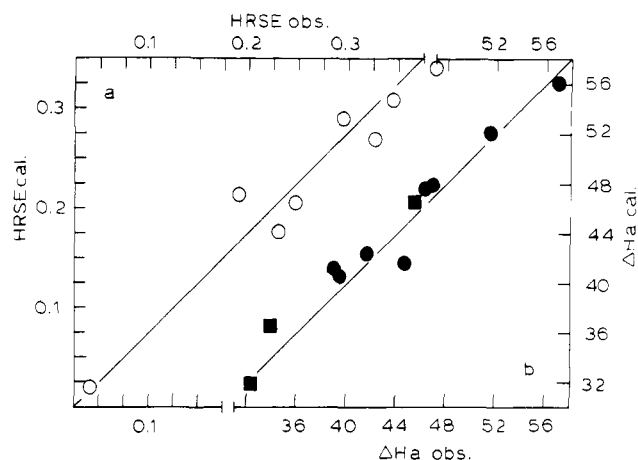
However, it can be argued that the above procedure to calculate characteristics by means of PC parameters for the test set, i.e., for compounds *not* included in the reference set of the model, is neither statistically sound nor does it utilize to the full the potential of the SIMCA approach, which provides empirical models of local validity for a given data set and consequently requires great caution in handling test objects not included in the reference set, especially when they exhibit RSD higher than that of the reference set, i.e., they are not within the domain of the model (cf. Table V; footnote c). We have therefore applied an alternative procedure for the calculation of the characteristics of the test set in Table VII.

In this alternative approach, we carried out PCA for 7 different matrices, each of which included all 12 variables and a total of 10 objects. The objects comprised the 9 compounds in the former reference set plus in each case the relevant one of the test set. This alternative procedure provided 7 different PCA models with different parameter values (for each of which the scores, loadings, and residuals are available as supplementary material) and resulted in even better predictions, as shown in the known characteristics in Table VII (footnote c).

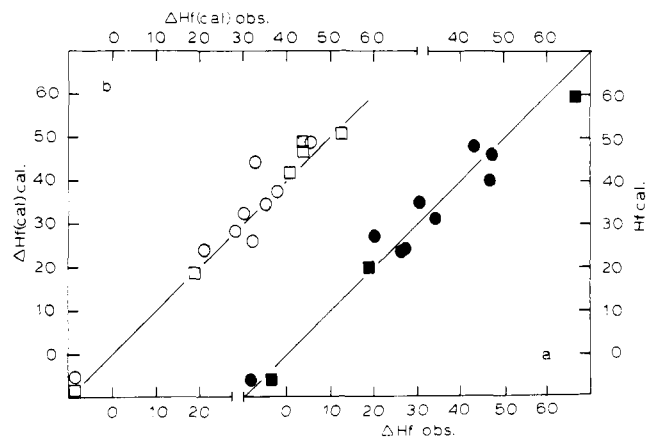
In this alternative procedure the RSD for each compound is always lower than that for the appropriate reference compound, in which the compound is now included. We realize that this alternative procedure is itself not statistically rigorous and is unconventional from the point of view of physical chemistry as it recalculates the characteristics by deriving a new model for each of the desired compounds rather than using a unique model of general validity. However it is, in our opinion, the best available approach for the estimation of unknown characteristics, when the lack of sufficient data prevents predictions by means of more rigorous statistical procedures, such as a PLS analysis including all 16 compounds in the same model, as occurs for the seven heteroaromatics reported in Table VII.



**Figure 2.** Plots of (a)  $\Delta N(\text{cal.})$  against  $\Delta N(\text{obs.})$  and (b)  $DRE(\text{cal.})$  against  $DRE(\text{obs.})$ : (○, ●) compounds used in the PCA model (data from Table VI); (□, ■) compounds not used in the PCA model (data from Table VII).



**Figure 3.** Plots of (a)  $HRSE(\text{cal.})$  against  $HRSE(\text{obs.})$  and (b)  $\Delta H_A(\text{cal.})$  against  $\Delta H_A(\text{obs.})$ : (○, ●) compounds used in the PCA model (data from Table VI); (□, ■) compounds not used in the PCA model (data from Table VII).



**Figure 4.** Plots of (a)  $\Delta H_F(\text{cal.})$  against  $\Delta H_F(\text{obs.})$  and (b)  $\Delta H_F(\text{cal.})$  against  $\Delta H_F(\text{cal.})$  (obs.): (○, ●) compounds used in the PCA model (data from Table VI); (□, ■) compounds not used in the PCA model (data from Table VII).

In Table VII, the results for the compounds *not* used in the PCA provide a stringent test of the whole procedure and a criterion to judge the likely success of the prediction of characteristics for which observed values are not available. For easier visual analysis, the values recalculated by the alternative procedure (c) are plotted against the original values in Figures 1–6. In these figures, the points for the nine compounds of the reference set used in the PC

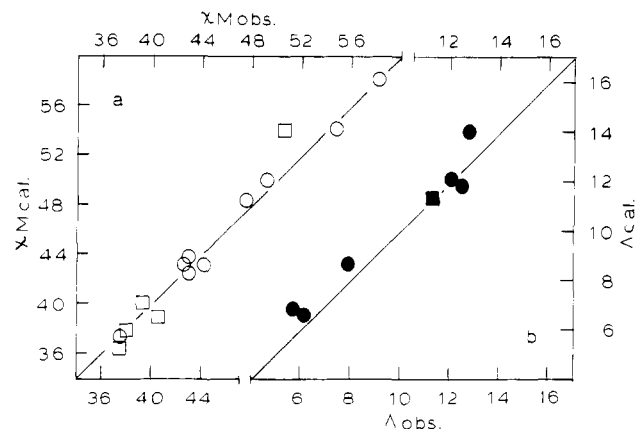
**Table VI.** Comparison of Recalculated with Observed Values of Characteristics for Compounds Used in the PCA Model.

no.	compd	1 $I_{S(6)}$	2 RC	3 $\Delta N$	4 DRE	5 HRSE	6 $\Delta H_A$	7 $\Delta H_F$	8 $\Delta H_{F(MNDO)}$	9 $\chi_M$	10 $\Delta$	11 $I_1$	12 $^{15}N$
1	benzene	100.0	1.75	0.000	22.6	0.390	57.2	19.8	21.2	54.8	13.7	1.00	
		97.3	1.82	0.014	22.6	0.368	56.3	27.4	24.3	54.3	15.0	1.02	
2	pyridine	85.7	1.73	0.090	23.1	0.348	51.8	34.6	28.7	49.2	13.4	0.978	-63.5
		89.7	1.73	0.078	20.0	0.337	52.1	31.3	28.3	50.1	12.7	0.954	-38.6
3	pyrimidine	84.3	1.73	0.160	20.2	0.297	47.0	47.0	35.5	43.1			-84.5
		84.1	1.67	0.130	19.0	0.318	47.9	40.4	35.2	42.5	9.26	0.870	-57.8
4	pyrazine	88.8	1.74	0.120	17.1	0.293	46.4	46.9	38.3	37.6	7.1		-46.1
		85.3	1.72	0.125	20.7	0.326	47.6	46.2	38.2	37.6	7.5	0.846	-40.9
8	thiophene	66.0			6.50	0.193	39.2	27.6	30.8	57.4	13.0	0.836	
		65.4	1.28	0.264	6.97	0.236	41.1	24.3	32.5	57.4	13.1	0.842	
9	furan	43.0	1.43	0.430	4.30	0.044	41.7	-8.30	-8.7	43.1	8.90	0.821	
		43.2	1.51	0.440	3.36	0.045	42.2	-6.48	-5.4	43.9	9.61	0.809	
10	pyrrole	69.0	1.46	0.310	5.3	0.233	44.8	25.9	32.4	47.6		0.866	-224.6
		63.9	1.42	0.282	9.0	0.210	42.0	24.1	26.8	48.5	10.4	0.819	-187.6
11	pyrazole	73.0	1.30	0.190		0.330		43.3	45.4	42.6	6.6	0.784	-127.0
		73.7	1.39	0.214	13.2	0.297	39.9	49.2	49.2	43.3	7.78	0.770	-126.7
12	imidazole	64.0	1.42	0.280	15.4	0.251	39.66	30.6	33.2	44.3		0.751	-169.0
		65.7	1.43	0.273	10.7	0.231	40.5	34.8	34.9	43.3	8.19	0.779	-168.0

**Table VII.** Comparison of (a) Observed Values of Characteristics for Compounds Not Used in the Statistical Study with (b) Those Calculated Using Equation 2 or (c) Calculated by the Alternative Procedure

no.	compds		1 $I_{S(6)}$	2 RC	3 $\Delta N$	4 DRE	5 HSRE	6 $\Delta H_A$	7 $\Delta H_F$	8 $\Delta H_{F(MNDO)}$	9 $\chi_M$	10 $\Delta$	11 $I_1$	12 $^{15}N$
5	pyridazine	a	78.9	1.72	0.170			45.6	66.4	44.2	40.5			+20.2
		b	89.7	1.67	0.092	21.1	0.368	46.4	56.2	48.8	38.1	7.20	0.833	-20.4
		c	86.2	1.69	0.118	20.6	0.355	46.4	59.5	47.0	39.0	8.16	0.845	3.78
6	s-triazine	a	100.0	1.72						40.8	37.9			-98.5
		b	84.8	1.69	0.128	20.1	0.326	47.0	46.7	39.5	38.4	7.61	0.842	-46.2
		c	91.9	1.71	0.109	20.3	0.339	47.4	47.8	42.1	37.9	8.22	0.854	-55.5
7	1,2,4-triazine	a	86.1		0.140					52.4				-6.66
		b	85.8	1.78	0.138	24.3	0.342	43.8	68.5	52.4	20.8	0.77	0.731	-7.96
		c	89.9	1.65	0.099	21.1	0.368	47.0	53.6	51.2	40.8	8.24	0.847	-12.32
13	thiazole	a	64.0		0.280						50.6	12.3	0.749	-58.0
		b	76.0	1.29	0.190	11.9	0.323	39.7	50.4	53.6	50.8	10.2	0.796	-130.7
		c	69.4	1.41	0.245	11.0	0.270	43.2	30.3	35.0	54.0	12.7	0.814	-83.1
14	oxazole	a	38.0	1.39					-3.7	-8.3	39.2			-123.7
		b	51.6	1.61	0.379	7.57	0.092	44.1	2.5	0.08	40.6	8.71	0.820	-231.4
		c	37.9	1.44	0.432	3.24	0.069	40.0	-5.9	-7.9	39.9	8.60	0.767	-126.8
15	isoxazole	a	42.0	1.36					18.8	19.2	37.5			+2.2
		b	67.1	1.59	0.264	13.2	0.214	43.7	29.2	24.9	38.7	7.44	0.804	-145.4
		c	44.5	1.39	0.384	6.3	0.126	38.5	19.1	18.4	36.3	6.64	0.735	-9.4
16	1,2,4-triazole	a	81.0		0.140			34.3		44.6				-131.0
		b	74.6	1.31	0.217	12.0	0.312	39.4	49.1	52.3	48.4	9.33	0.783	-133.5
		c	77.2	1.45	0.182	14.8	0.301	36.6	53.0	49.3	38.2	6.16	0.755	-116.6

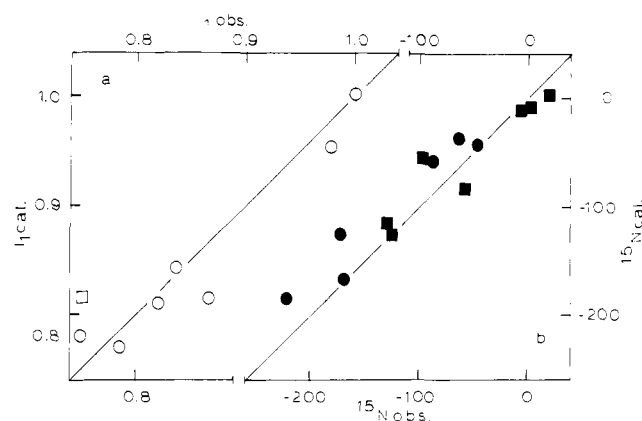
<sup>a</sup> Observed values. <sup>b</sup> Values calculated with eq 1 (see text). <sup>c</sup> Values calculated with the alternative procedure, i.e., by PCA of matrices including 10 objects, the 9 compounds reported in Table VII plus the appropriate compound in this table (see text).



**Figure 5.** Plots of (a)  $\chi_M(\text{cal.})$  against  $\chi_M(\text{obs.})$  and (b)  $\Delta(\text{cal.})$  against  $\Delta(\text{obs.})$ : (○, ●) compounds used in the PCA model (data from Table VI); (□, ■) compounds not used in the PCA model (data from Table VII).

analysis are distinguished from those for the other compounds of the test set.

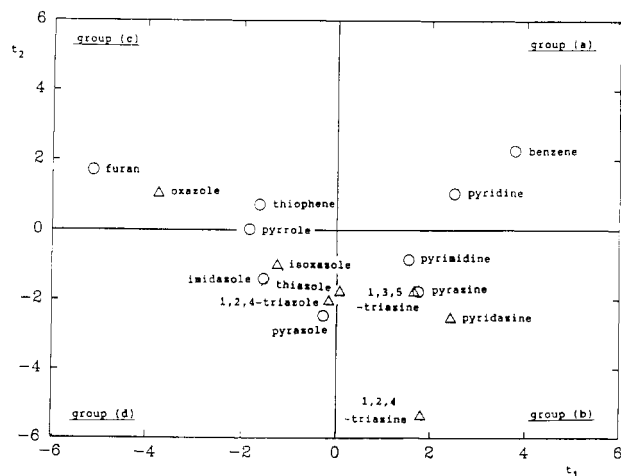
We will now proceed to discuss first the success or otherwise of these recalculations for the three classes of characteristic and



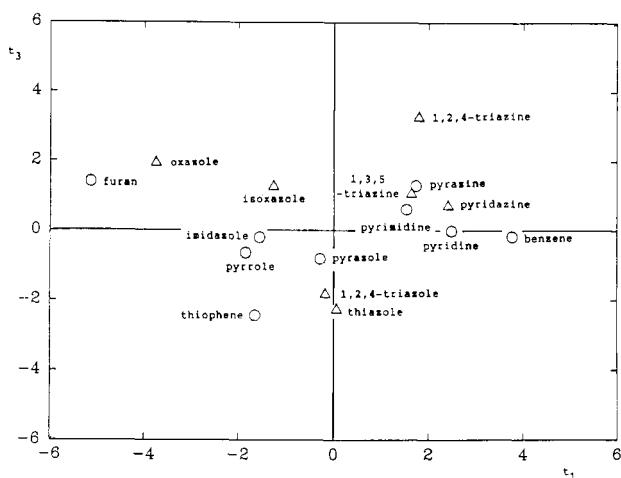
**Figure 6.** Plots of (a)  $I_1(\text{cal.})$  against  $I_1(\text{obs.})$  and (b)  $^{15}N \text{ NMR}(\text{cal.})$  against  $^{15}N \text{ NMR}(\text{obs.})$ : (○, ●) compounds used in the PCA model (data from Table VI); (□, ■) compounds not used in the PCA model (data from Table VII).

then the significance of the scores and loadings.

**Correlation and Prediction of Geometrical Properties.** For the compounds used in the statistical analysis, characteristics  $I_{S(6)}$  (Figure 1a) and  $\Delta N$  (Figure 2a) are recalculated with an excellent



**Figure 7.** Scores plot of  $t_2$  versus  $t_1$ . (○) compounds used in the PCA model; (Δ) compounds not used in the PCA model. Data are from Table V.



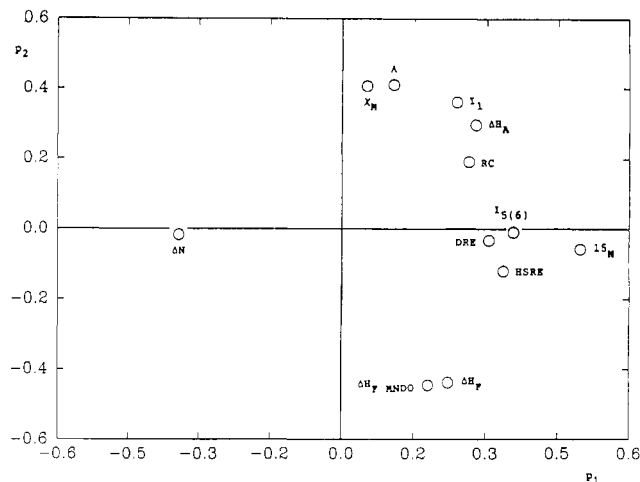
**Figure 8.** Scores plot of  $t_3$  versus  $t_1$ . (○) compounds used in the PCA model; (Δ) compounds not used in the PCA model. Data are from Table V.

fit to the experimental data from which they were derived. For RC (Figure 1b) the fit is only fair. This pattern repeats itself, with more scatter, for the second group of compounds, the fit is good for  $\Delta N$ , fair for  $I_{5,6}$ , and fair to poor for RC, the values of which are not equally distributed within the examined range (grouping of points) (cf. Figures 1a,b, and 2a).

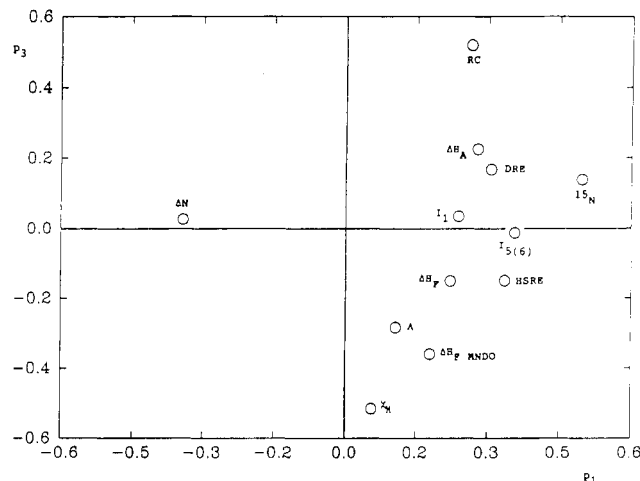
**The Correlation and Prediction of Energetical Properties.** The comparisons for the recalculated values of the first reference group are excellent for four (HSRE,  $\Delta H_A$ ,  $\Delta H_F$ ,  $\Delta H_{F(cal)}$ , cf. Figures 3a,b and 4a,b) of the five energetical characteristics and good for the remaining one, the DRE (Figure 2b). Relatively few values were available for these characteristics for the second test group of compounds: good fits were found for  $\Delta H_A$  and  $\Delta H_{F(cal)}$  (Figures 3b and 4b) and a fair fit for  $\Delta H_{F(exp)}$  (Figure 4a).

**The Correlation and Prediction of Magnetic Properties.** For three ( $\chi_M$ ,  $\Delta$ ,  $I_1$ ; Figures 5a,b and 6a) of the four magnetic parameters, again an excellent fit was found for the recalculated values, and for the remaining one ( $^{15}N$ ) a good fit was found (Figure 6b). For the predicted values for the second group of compounds, excellent agreement was found for  $\chi_M$  and  $I_1$  (Figures 5a and 6a), good for  $\Delta$  (Figure 5b), but somewhat poorer for  $^{15}N$  (Figure 6b).

**Significance of PC Scores.** The PC scores for all 16 compounds are plotted graphically in Figures 7 and 8. Examination of Figure 7 shows that principal components scores  $t_1$  and  $t_2$  divide the compounds into four groups: (a) for benzene and pyridine both values are positive; (b) for the other azines  $t_1$  is positive and  $t_2$  negative; (c) for the five rings with one hetero atom  $t_1$  is negative



**Figure 9.** Loadings plot of  $p_2$  versus  $p_1$  (data from Table IV).



**Figure 10.** Loadings plot of  $p_3$  versus  $p_1$  (data from Table IV).

and  $t_2$  positive; and (d) for the azoles both values are negative. The only exception to this generalization is for oxazole which lies in group (c) rather than group (d). (It should be remembered that, of all the compounds considered in the test set, it is oxazole and isoxazole which have high RSD: thus they lie on the border of the validity of the PCA model and their characterization by PC scores is less reliable.)

This is a most striking division into just those groups of compounds that have similar chemical properties (cf. discussion in ref 4b and 4c)!

Figure 8 shows clearly that the third PC scores differentiate the five-membered ring compounds into three groups: those containing oxygen, nitrogen, and sulfur heteroatoms, respectively. Within the six-membered rings, the  $t_3$  values increase (while  $t_1$  values decrease) on increasing the number of nitrogen atoms in the ring.

**Significance of the PC Loadings for the Various Characteristics.** The PC loadings for the 12 characteristics are given in Figures 9 and 10. For the interpretation of the loadings plots it is worth mentioning that variables which lie in the same direction with respect to the origin have similar information content. Moreover from the absolute numerical value of the loadings it is possible to estimate how much a single variable contributes to each component. In the three-dimensional space defined by their PC loadings, the characteristics can be divided into three main groups. Two of these groups (a and b) are orthogonal to each other and can be related to the concepts of "classical" and "magnetic" aromaticity. These groups are the following:

(a) The first group of five characteristics ( $I_{5,6}$ ,  $\Delta N$ , DRE, HSRE, and  $^{15}N$ ) is dominated by the  $p_1$  loadings; the  $p_2$  loading is small for all of these characteristics, and the  $p_3$  loadings are small to moderate. It is thus clear that  $p_1$  measures what may



be called "classical aromaticity". We can follow our situation best by reference to Figure 5a. Among this group of characteristics, it is  $I_{5,6}$  that provides the purest measure of classical aromaticity, closely followed by  $\Delta N$  (no. 3). The DRE (no. 4) and  $^{15}\text{N}$  (no. 12) have small  $p_2$  loadings but rather more significant  $p_3$  loadings whereas the HSRE (no. 5) is a somewhat more hybrid measure.

(b) The second group is comprised of the magnetic parameters  $\chi_m$  (no. 9) and  $\Lambda$  (no. 10). These are both marked by very small  $p_1$  components but large positive  $p_2$  and large negative  $p_3$  components. Clearly this second group of characteristics includes measurements of "magnetic aromaticity" which is almost completely orthogonal to the "classical aromaticity" measured by characteristics of group (a).

(c) The remaining group of five characteristics all show moderately positive  $p_1$  loading together with negative loadings for both  $p_2$  and  $p_3$  for  $\Delta H_F$  (calculated or experimental) and positive loadings for both  $p_2$  and  $p_3$  for RC,  $\Delta H_a$ , and  $I_1$ . The characteristics in this third group thus possess elements of both "classical" and "magnetic" aromaticity.

**Conclusions.** We believe that we have been able to a considerable extent to resolve the apparent impasse between classical and magnetic aromaticity and provide a firm basis for the consideration of aromaticity as a quantitative concept. In fact, there are at least two types of aromaticity: the best available measure of classical aromaticity is provided by the Bird  $I_{5,6}$  parameter and this parameter correlates well for  $\Delta N$  and DRE. The second type of

aromaticity is magnetic aromaticity, which is measured by  $\chi_m$  or  $\Lambda$ . These two types, classical and magnetic, of aromaticity are orthogonal. Other aromatic characteristics are influenced by both "classical" and "magnetic" aromaticity to varying extents. Hence, we believe that there are at least two quantitative aromaticity scales. Characteristics that depend on aromaticity can be determined by either or both of these scales. Indeed a case could be made for the existence of *three* types of aromaticity corresponding to the three PC found. At the present time we wish only to claim the existence of at least two types: further work is in hand which it is hoped will further illuminate the nature of the phenomenon of aromaticity. The methods used in this paper are capable of extension to a wide range of other aromatic, anti-aromatic, and non-aromatic compounds, ions, and radicals, and to numerous other chemical and physical properties. They offer the potential not only for the understanding and rationalization of known facts but also for the prediction and estimation of unknown properties. Work along these lines is ongoing in our laboratories.

**Acknowledgment.** We thank a Referee for helpful comments on this paper.

**Supplementary Material Available:** Predictions of aromatic characteristics for **5**, **6**, **7**, **13**, **14**, **15**, and **16** (65 pages). Ordering information is given on any current masthead page.

## Theoretical Studies on the Stability of Tl–C $\sigma$ -Bonds in Aliphatic Organothallium Compounds

P. Schwerdtfeger,<sup>\*,†,‡</sup> P. D. W. Boyd,<sup>‡</sup> G. A. Bowmaker,<sup>‡</sup> H. G. Mack,<sup>§</sup> and H. Oberhammer<sup>§</sup>

*Contribution from the Department of Chemistry, University of Auckland, Auckland, New Zealand, and the Institut für Physikalische und Theoretische Chemie, Universität Tübingen, D-7400 Tübingen, West Germany. Received March 7, 1988*

**Abstract:** Calculations of the electronic structure and bonding in organothallium compounds ( $\text{TlR}$ ,  $\text{TlR}^+$ ,  $\text{TlR}_2$ ,  $\text{TlR}_2^+$ ,  $\text{TlR}_3$ ,  $\text{TlR}_4^+$ ,  $\text{TlR}_2\text{Br}$ , and  $\text{TlR}'$  with  $\text{R} = \text{CH}_3$ ,  $\text{R}' = \text{C}_2\text{H}_3$ , or  $\text{C}_2\text{H}$ ), in the hydrides ( $\text{R} = \text{H}$ ), their dissociation products ( $\text{CH}_3$ ,  $\text{C}_2\text{H}_3$ ,  $\text{C}_2\text{H}$ ,  $\text{C}_2\text{H}_6$ , and  $\text{TlBr}$ ), and  $\text{Tl}_2$  have been carried out. A modified version of Pople's SCF program GAUSSIAN82 has been used, in which the core electrons on the Tl atom are described by a quasi-relativistic [Pt]-core pseudopotential. All bond distances and angles have been optimized with a Fletcher–Powell procedure. SCF–SW– $X\alpha$  calculations have been carried out to determine the extent of the Tl 5d core orbital participation in the Tl–C bond. Vibrational frequencies have been calculated from the harmonic diagonal SCF valence force field. As a result, the previously undetected molecule  $\text{TlCH}_3$  is found to be stable with respect to dissociation into Tl and  $\text{CH}_3$ . The isolated  $\text{Tl}(\text{CH}_3)_3$  molecule is found to possess the expected planar-trigonal ( $\text{TlC}_3$ ) structure. It is explained why organothallium chemistry is mainly the chemistry of Tl(III) in contrast to inorganic thallium chemistry and that Tl–Tl bonds are very weak or do not exist is a relativistic effect.

### I. Introduction

In the last 20 years organothallium chemistry has become more and more important in organic synthesis, and experimental studies on structure and bonding have been reported on many organothallium compounds.<sup>1–10</sup> Most of the stable organothallium compounds are of the type  $\text{R}_2\text{TlX}$  ( $\text{R} = \text{alkyl, aryl}$ ;  $\text{X} = \text{F, ClO}_4, \dots$ ). Trialkyl and triaryl compounds are unstable and very reactive; e.g.,  $\text{TlMe}_3$  ( $\text{Me} = \text{CH}_3$ )<sup>11</sup> is reported to be spontaneously inflammable in dry air.<sup>12</sup> In contrast,  $\text{Me}_3\text{M}$  compounds ( $\text{M} = \text{B, Al, Ga, In}$ ) are stable.<sup>10</sup> Monoalkyl compounds of the type  $\text{TlR}$  have never been isolated and the only known structures

containing this moiety are of the types  $\text{RTlX}_2$  and  $\text{RTlO}$  ( $\text{X} = \text{OAc, CN, Cl, Br, } \dots$ ).<sup>1,2,10,13</sup> Monoaryllthallium compounds are

(1) Lee, A. G. *The Chemistry of Thallium*; Elsevier: Amsterdam, The Netherlands, 1971.

(2) Abel, E. W.; Stone, F. G. A. *Organometallic Chemistry*; The Royal Society of Chemistry: London, 1971–1987; Vol. I–XV.

(3) (a) McKillop, G. A. *Acc. Chem. Rev.* **1970**, *3*, 138. (b) McKillop, G. A.; Taylor, E. C. *Adv. Organomet. Chem.* **1973**, *11*, 147. (c) McKillop, G. A.; Taylor, E. C. *Chem. Br.* **1973**, *9*, 4. (d) McKillop, G. A.; Taylor, E. C. *Comprehensive Organometallic Chemistry*; Pergamon: New York, 1982; Vol. 7, pp 765. (e) Davidson, G. *Coord. Chem. Rev.* **1984**, *56*, 113.

(4) (a) Kurosawa, H.; Kawasaki, Y. *Fundamental Research in Organometallic Chemistry*; Tsutsui, M., Ishii, Y., Yaozeng, H., Eds.; Van Nostrand Reinhold: London, 1982. (b) Kurosawa, H. *Comprehensive Organometallic Chemistry*; McKillop, A., Taylor, E. C., Eds.; Pergamon: New York, 1982; Vol. 3, pp 725.

<sup>†</sup> Alexander von Humboldt Feodor–Lynen fellow 1987/1988.

<sup>‡</sup> University of Auckland.

<sup>§</sup> Universität Tübingen.