CROATICA CHEMICA ACTA CCACAA 56 (3) 357—363 (1983)

CCA-1382

YU ISSN 0011-1643
UDC 541;547.52
Original Scientific Paper

New Topological Indices for Alternant Polycyclic Aromatic Hydrocarbons*

Karl-Dietrich Gundermann, Claus Lohberger and Maximilian Zander

Department of Organic Chemistry, Technical University of Clausthal, 3392 Clausthal-Zellerfeld, and Laboratory of Rütgerswerke AG, 4620 Castrop-Rauxel, Federal Republic of Germany

Received August 30, 1982

Pairwise orthogonal topological indices Λ and W_{CG} derived from the characteristic graphs of alternant polycyclic aromatic hydrocarbons (kata-annellated and peri-condensed systems) correlate with resonance energy per electron, energy of the HOMO and total π electron energy respectively. For topological indices a high discrimination power is necessary but not sufficient to correlate with molecular properties. However, the indices and properties must be related similarly to topology and the size of the systems. Accordingly Type-A indices and properties which depend comparably strongly on both topology and size (e. g. Λ and resonance energy per electron) and Type — B indices and properties which depend predominantly on size (e. g. W_{CG} and total π electron energy) are distinguished.

Quantitative structure/property relationships in organic chemistry enable the numerical values of a given molecular property P_i to be predicted. Consequently, such relationships are of great importance for the organic chemist from the practical point of view, and in some cases they allow an understanding of experimental results in terms of molecular concepts.

In empirical correlations of the type described by eq. 1 (»quantitative analogy models«[1]) x_j are the predictor variables and $\varepsilon_{i,j}$ denotes the random variation not described by the mathematical model.

$$P_i = F(x_j) + \varepsilon_{i,j} \tag{1}$$

Predictor variables that are derived from invariants of some kind of molecular graphs are usually termed »topological indices« (TI's)[2]. They »translate« chemical constitution into numerical values.

In this paper we deal with TI's of alternant polycyclic aromatic hydrocarbons (a-PAH) and their correlations with the hydrocarbon π-electronic properties. If the small existing variances in geometry, i. e. variances in bond lengths and bond angles, are neglected, all representatives of the entire class of a-PAH differ only in two structural features viz. the number of carbon centres or »size« (S) and/or the connection of carbon centres or »topology« (T)

* Presented at the *IUPAC International Symposium on Theoretical Organic Chemistry,* held in Dubrovnik, Croatia, August 30 — September 3, 1982.

of the molecules. (Note: we do not use the term »topology« in this paper as usually defined, i. e. as the totality of information contained in the molecular graph). It then can be concluded directly that all numerical differences for any given property of a-PAH depend only on differences in S and T of the molecules. However, »simple« structure/property relationships can only be expected for a-PAH if many of the possible electronic interactions are apparently invariant to changes in S and T. In fact, this presupposition is well characterized by the postulates of Zeroth-order HMO theory, i. e. neglecting σ/π interaction, all Coulomb integrals $\alpha_r = \alpha$, all resonance integrals $\beta_{r,s} = \beta$, if r,s are connected carbon centres or otherwise $\beta_{r,s} = 0$, and all overlap integrals $S_{r,s} =$ const. From a pragmatic point of view, these postulates seem to be justified by the reasonable agreement between experimentally and theoretically obtained data of a-PAH.

To derive TI's for a-PAH the structural graph (SG) can be reduced to more simple graphs provided the new graphs still contain the relevant structural information. We found the »characteristic graphs« (CG) of a-PAH, introduced by Balaban and Harary[3], well-suited for modelling TI's. Although there is some loss of structural information when changing from SG's to CG's, useful TI's can be derived from the latter.

Inspection of the CG's of a-PAH, e. g. **I** (tetracene) and **II** (dibenzo[a,c]anthracene) readily reveals that the number n of vertices, the number s of edges and the number w of 120° angles are the graph invariants that can obviously be used for constructing TI's. In principle symmetry properties are also available but will not be considered in this context. The size S of the molecules, i. e. the number of carbon centres, is given by

$$S = 5n - s + 1 \tag{2}$$

This expression holds for both cata-annellated and peri-condensed a-PAH. As a measure of topology we define the »Angularity«

$$A = \frac{w + 1}{n} \tag{3}$$

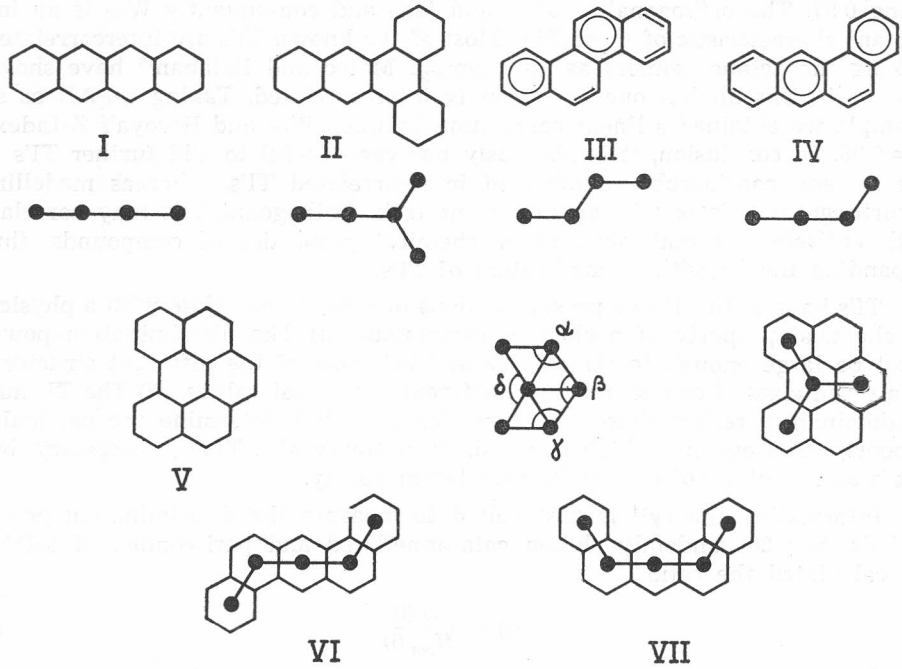
l/n has the function of a normalizing factor. Angularity is the specific type of »branching« for a-PAH. It does not result definitely from the vertex degrees, i. e. the number of vertices adjacent to each vertex.

Since angular annellation produces π-sextets, as defined by Clar[4], there is a correspondence between angularity and the number of π-sextets. However, angularity exhibits lower degeneracy than the number of π-sextets, as can be recognized easily by comparing the CG's and Clar's formula of e. g. chrysene (**III**) and benzo[a]anthracene (**IV**). While the number of π-sextets is identical in both compounds angularity is different. In fact these compounds differ markedly in such properties as the HOMO-LUMO transition energy or the resonance energy.

For peri-condensed a-PAH we define w as the sum of the »external« 120° angles w_e and the »internal« 120° angles w_i.

$$w = w_e + w_i \tag{4}$$

The definition and procedure of counting these angles can be demonstrated easily in the case of benzo[g,h,i]perylene (**V**)[5]. The number w_e of the external



120^0 angles can be obtained simply by counting the angles included in the perimeter of the graph, i. e. angle α, β, γ and δ, and so $w_e = 4$. The number w_i of the internal 120^0 angles is defined for any given peri-condensed a-PAH by

$$w_i = 3t - w_t \tag{5}$$

t is the number of triphenylene subgraphs and w_t the number of those 120^0 angles of the subgraphs that are already included in w_e. For benzo[g,h,i]perylene (**V**) $t = 1$, $w_t = 1$ is obtained and hence $w_i = 2$. (Note: triphenylene subgraphs with one edge in common will be counted only once.) While Balaban's code[6] of CG's (using digits 0, 1 and 2 for different annellation directions) has, as yet, proved not to be applicable to peri-condensed systems[7] and the application of Clar's π-sextets to derive quantitative correlations with molecular properties is also restricted to cata-annellated a-PAH[8], angularity, as defined above, successfully characterizes the topology of both cata-annellated and peri-condensed systems.[5]

The first topological index derived from the CG's of a-PAH is simply the ratio of angularity and size of the systems:

$$\Lambda = \frac{A}{S} = \frac{w+1}{(5n - s + 1)\,n} \tag{6}$$ [9]

As a further topological index W_{CG} the half-sum of the distance matrix elements is used, i. e. the »Wiener number«[10] of the CG's.[11]

Λ and W_{CG} proves to be orthogonal[12], while W_{CG} is linearly correlated with the Wiener number W_{SG} of the structural graphs, the correlation coefficient

being 0.97. The orthogonality of Λ and W_{CG} and consequently W_{SG} is an important characteristic of these TI's. Most of the known TI's are intercorrelated. Taking the octane isomers as an example Motoc and Balaban[13] have shown that of 17 TI's all but one are strongly intercorrelated. Taking a-PAH as an example we obtained a linear correlation between W_{CG} and Hosoya's Z-Index[2], $r = 0.98$. In conclusion, it is obviously not very useful to add further TI's to the already considerable number of intercorrelated TI's, whereas modelling of orthogonal TI's is still an important task. Orthogonal TI's may correlate with entirely different physical or chemical properties of compounds, thus expanding the breadth of application of TI's.

TI's have to fulfill two presuppositions in order to correlate with a physical or chemical property of a class of compounds: (*i*) The discrimination power must be large enough to »translate« at least most of the different structures of a given class of compounds into different numerical values. (*ii*) The TI must predominantly reflect those structural features that determine the particular property. Consequently, high discrimination power of a TI is a necessary, but not a sufficient condition for its correlation ability.

Information theory[14] is best suited to measure the discrimination power of TI's. For 26 randomly chosen cata-annellated and peri-condensed a-PAH we calculated the ratio

$$\eta\,(\mathrm{i}) = \frac{H\,(\mathrm{i})}{H_{\mathrm{pot}}\,(\mathrm{i})} \tag{7}$$

where H_{pot} (i) is the mean information content of the 26 systems (being $H_{\mathrm{pot}}\,(\mathrm{i}) = {}_2\log 26$) and H (i) is the mean information content of the respective TI. This ratio amounts to 1.0000 for W_{SG}, i. e. no information is lost during the »translation«, 0.9120 for Λ, i. e. there is an information loss of approx. 9%, and 0.8068 for W_{CG} corresponding to an information loss of approx. 19%[11]. The information loss observed in the TI's as derived from CG's, results mainly from the inability of these indices to discriminate between systems having identical numbers of vertices, edges and angles, but with different symmetry, e. g. systems **VI** (dibenzo[a,h]anthracene) and **VII** (dibenzo[a,j]anthracene). On the other hand, this failure is not serious, because systems interrelated in such a way also have similar physical and chemical properties.

As already stated above, all numerical differences in any given property of a-PAH depend only on differences in size (S) and topology (T) of the molecules. Using a regression analytical method that has been described elsewhere[11], the S,T dependence of the Zeroth-order HMO total π-electron energy E_π was calculated, i. e. the function E_π (S,T). In Figure 1 this function is represented as a surface in space using arbitrary units, where a and b differ in the visual angle. The wellknown fact[15], that the HMO total π-electron energy of a-PAH is almost independent of topology can be easily recognized in the Figure. Quite a different situation applies to Dewar resonance energies per electron (DREPE). Figure 1 c and d represent the function DREPE (S,T), viewed from different visual angles. DREPE depends comparably strongly on both topology and size of the systems. Thus, besides the possibility of calculating HMO data explicitely, one can also obtain these data by starting from a set of known data and using regression analysis, i. e. for a given size and topology one will find the desired information on the respective surface in space. While

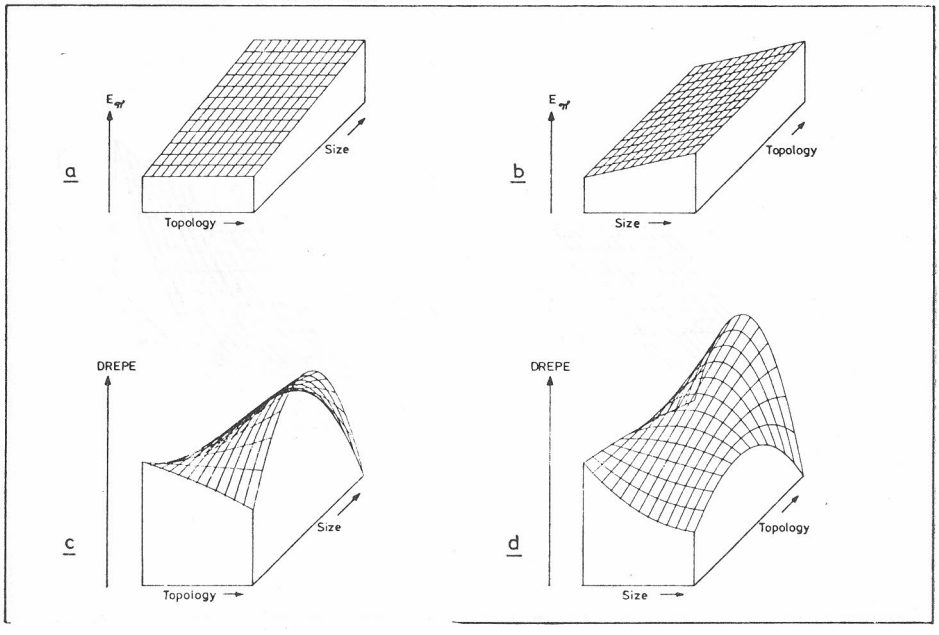


Figure 1. Size/topology dependence of HMO total π-electron energy E_π and Dewar resonance energy per electron DREPE of alternant polycyclic aromatic hydrocarbons (arbitrary units).

size of the systems is unambigously defined for a-PAH by the number of carbon centres, topology can be expressed in different ways. For the representations shown in Figure 1 a-d the sum of the 120° angles as defined above was used as a measure of topology.

Generalization of what has been said about the S,T dependence of total π-electron energy and resonance energy per electron respectively, leads to the conclusion that the properties of a-PAH can be divided into two groups viz. »Type-A-properties«, whose numerical values depend comparably strong on both S and T, and »Type-B-properties«, whose numerical values depend predominantly on S of the systems[11]. This reasoning straightforwardly yields an explanation of why properties of a-PAH can be linearely correlated, although they have obviously no common physical grounds. By way of example, the total π electron energy of a-PAH is linearely correlated with high pressure liquid chromatography retention data (k' values) ($r = 0.97$), because both properties are of the B-Type.[16]

For the empirical correlations found between properties (P_i) and topological indices (TI) of a-PAH the variation of P_i with TI, i. e. the first derivatives dP_i/dTI, are either independent of TI, proportional to TI, or inversely proportional to TI, i. e. the correlations are either linear, parabolic or logarithmic. In the first two cases the S,T surfaces of the property and the index are expected to be similar in shape (as follows from the respective polynomials in S and T). Parabolic correlations have been found between total π-electron energy E_π of a-PAH and the Wiener numbers W_{SG} and W_{CG} of the structural and characteristic graphs, respectively,[11] the correlation coefficients being 0.97 and 0.98. In Figure 2 the S,T surfaces corresponding to E_π (taken from

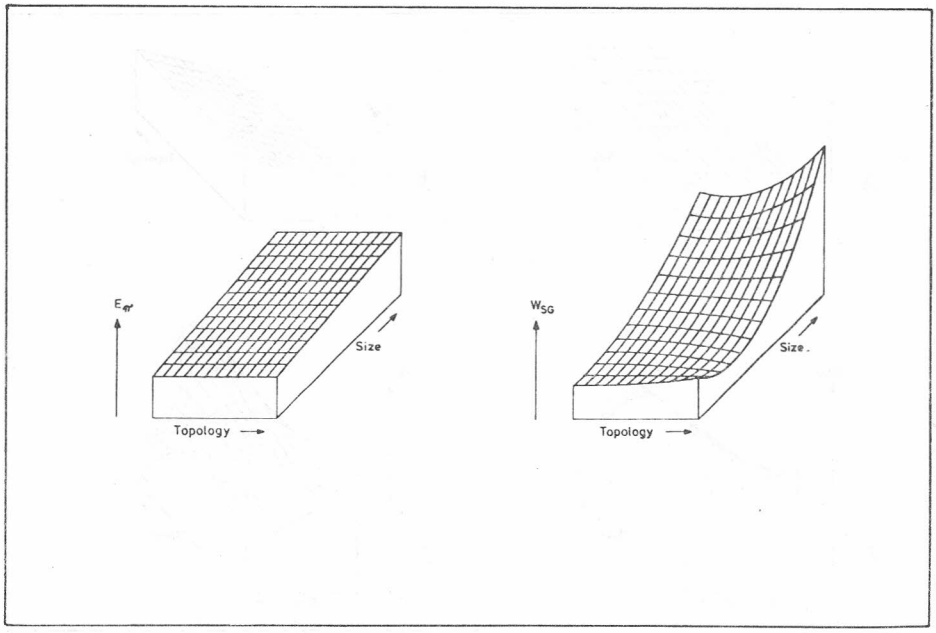


Figure 2. Size/topology dependence of E_π (see Figure 1) and the Wiener number W_{SG} of the structural graphs of alternant polycyclic aromatic hydrocarbons (arbitrary units).

Figure 1) and W_{SG} are represented (in arbitrary units). As expected these surfaces are very similar in shape, and both are characterized by being almost independent of the topology.

As with the properties of a-PAH two types of TI's can be distinguished viz. »Type-A-indices«, comparably strongly dependent on S and T, and »Type-B--indices«, predominantly dependent on S of the systems. The Wiener numbers are examples of Type-B-indices and correspondingly correlate with Type-B--properties, such as total π-electron energy or chromatographic retention data of a-PAH.

On the other hand, the numerical values of Λ are strongly dependent on both S and T of the systems, i. e. Λ in contrast to the Wiener numbers is an example of Type-A-indices. Accordingly correlations exist with Type-A-properties of a-PAH[5] and inversely no correlations have been found for Type-B-properties, e. g. total π-electron energy. In l. c.[5] correlations between Dewar resonance energy per electron (DREPE) and the energy ε_{HOMO} of the highest occupied orbital respectively with $\ln\Lambda$ have been reported. Eq. 8 holds for both cata-annellated and peri-condensed systems, the correlation coefficient being 0.96.

$$\text{DREPE}\,[\beta] = 0.00573\,\ln\Lambda + 0.0562 \tag{8}$$

Since DREPE and ε_{HOMO} are linearely correlated for cata-annellated and pericondensed a-PAH respectively, at which the slopes of the lines are identical but the intercepts are different, eq. 9 holds for both types of a-PAH ($r = 0.95$), the constant k being -0.6927 for peri-condensed systems and vanishing for cata-annellated systems.

$$\varepsilon_{\mathrm{HOMO}}\,[\beta] = 0.2651\ (\ln \Lambda + k) + 1.4143 \tag{9}$$

The application of relationships (8) and (9) enables the organic chemist to obtain information relevant for his practical work, i. e. information on stability, reactivity, para-band transition energy etc. of as yet unknown hydrocarbons, simply by using a pocket calculator. Generalizing this point of view, topological indices can be regarded as a useful aid for the practicing chemist. On the other hand it is still questionable whether TI's are suited for achieving a deeper understanding of the relationships between structure and the properties of organic molecules in terms of basic theory.

REFERENCES

1. S. Wold and M. Sjöström, in: *Correlation Analysis in Chemistry,* Edited by N. B. Chapman and J. Shorter, Plenum Press, New York-London 1978.
2. H. Hosoya, *Bull. Chem. Soc. Japan* **44** (1971) 2332.
3. A. T. Balaban and F. Harary, *Tetrahedron* **24** (1968) 2505.
4. E. Clar and M. Zander, *J. Chem. Soc. (London)* (1958) 1861; E. Clar, *The Aromatic Sextet,* Wiley, London 1972.
5. K. D. Gundermann, C. Lohberger, and M. Zander, *Z. Naturforsch.* **36a** (1981) 276.
6. A. T. Balaban, *Rev. Roum. Chim.* **15** (1970) 1243.
7. A. T. Balaban, *Pure Appl. Chem.* **54** (1982) 1075.
8. M. Zander, *Z. Naturforsch.* **33a** (1979) 1398.
9. In our previous papers (see ref. 5 and 11) we denoted the topological index defined by eq. 6 by »z«, but to avoid confusion with Hosoya's »Z-Index« (see ref. 2) we will denote it here by »Λ«.
10. H. Wiener, *J. Amer. Chem. Soc.* **69** (1947) 17; see also: D. Bonchev and O. Mekenyan, *Z. Naturforsch.* **35a** (1980) 739 and references given therein.
11. K. D. Gundermann, C. Lohberger, and M. Zander, *Z. Naturforsch.* **36a** (1981) 1217.
12. Note, that we do not use the term »orthogonal« in its strict, mathematical sense but rather colloquial to mean non-correlated.
13. I. Motoc and A. T. Balaban, *Rev. Roum. Chim.* **26** (1981) 593.
14. C. Shannon and W. Weaver, *The Mathematical Theory of Communication,* Illinois University Press, Urbana 1949.
15. See: A. Graovac, I. Gutman, and N. Trinajstić, *Topological Approach to the Chemistry of Conjugated Molecules,* Lecture Notes in Chemistry No. 4 Springer-Verlag, Berlin 1977.
16. For another interesting example (taken from NMR spectroscopy of PAH) for correlating properties that are obviously not causely connected in terms of physics see: R. B. Mallion *Org. Magnet. Reson.* **5** (1973) 91; K. D. Bartle, ibid. **5** (1973) 95.

SAŽETAK

Novi topološki indeksi za alternantne policikličke aromatske ugljikovodike

K.-D. Gundermann, C. Lohberger i M. Zander

Međusobno ortogonalni topološki indeksi Λ i W_{CG} izvedeni iz karakterističnog grafa za alternantne policikličke aromatske ugljikovodike koreliraju s rezonantnom energijom po elektronu, HOMO-energijom i ukupnom π-elektronskom energijom. Da bi topološki indeksi korelirali s molekularnim svojstvima, nužno je ali ne i dovoljno da pokazuju visoku moć diskriminiranja. Indeksi i svojstva moraju, naime, biti na sličan način povezani s topologijom i veličinom sistema. U skladu s tim razlikuju se indeksi i svojstva tipa A koja ovise razmjerno jako i o topologiji i o veličini (npr. Λ i rezonantna energija po elektronu), te indeksi i svojstva tipa B koja uglavnom ovise o veličini (npr. W_{CG} i ukupna π-elektronska energija).