# QSPR correlations of half-wave reduction potentials of *cata*-condensed benzenoid hydrocarbons

**Jyoti Singh,[a] Shalini Singh,[b] Shahida Meer,[c] Vijay K. Agrawal,[a] Padmakar V. Khadikar,[d*] and Alexandru T. Balaban[e]**

*[a] QSAR and Computer Chemical Laboratories, A.P.S. University, Rewa 486 003, India*
*E-mail: jyoti_singh07@rediffmail.com; vijay-agrawal@lycos.com*
*[b] Department of Chemistry, Bareilly College, Bareilly (U.P.) 243 001 India*
*E-mail: shalini_singh15@yahoo.com*
*[c] Department of Chemistry, Holkar Model and Autonomus College, Indore 452 100, India*
*[d] Research Division, Laxmi Fumigation and Pest Control, Pvt. Ltd., 3, Khatipura, Indore 452 007, India*
*E-mail: pvkhadikar@rediffmail.com*
*[e] Texas A&M University at Galveston, 5007 Avenue U, Galveston, TX 77551, USA*
*E-mail: balabana@tamug.edu*

**Abstract**

Polarographic half-wave reduction potentials of 31 *cata*-condensed benzenoid hydrocarbons were estimated by using lowest unoccupied molecular orbital energies ($E_{LUMO}$), an indicator of the number of linearly-fused benzenoid rings ($n$), and distance-based topological indices. The first two parameters account for most of the variance. A parallelism between experimentally measured measured Diels-Alder reactivity and polarographic half-wave reduction potentials is illustrated by a correlation between these data.

**Keywords:** Benzenoid hydrocarbons, topological indices, regression analysis, molecular descriptor, LUMO, polarographic half-wave reduction potential, QSPR

## Introduction

The publication of the present article was prompted by the recent paper of Nikolić, Milicević, and Trinajstić about the QSPR study of polarographic half-wave reduction potentials ($E_{½}$ values) of benzenoid hydrocarbons.[1] These hydrocarbons included a selection of $E_{½}$ values for both *cata*-condensed and *peri*-condensed systems.

Benzenoid polycyclic aromatic hydrocarbons (benzenoids for brevity) are uniquely characterized by their dualist graphs consisting of vertices situated in the centers of benzenoid

rings, and of edges connecting vertices corresponding to condensed rings, i.e., rings sharing one C–C bond.[2,3] Benzenoids are classified as *cata*-condensed (catafusenes), *peri*-condensed (perifusenes) and *corona*-condensed (coronoids): catafusenes have no carbon atom common with three benzenoid rings, and their dualist graphs are acyclic; perifusenes have carbon atoms common with three benzenoid rings, and their dualist graphs have three-membered rings (and possibly also acyclic branches), whereas coronoids have dualist graphs with larger rings (and possibly also three-membered rings and/or acyclic branches). Long ago,[2,3] one of us argued that the notation introduced for catafusenes, indicating the geometry of their condensation by means of digits 0, 1, and 2 for their dualist graphs, contains essential information on many properties of these catafusenes. Digit 0 indicates linear annelation (angle of 180º in the dualist graph) giving rise to anthracene portions, whereas digits 1 and 2 denote kinked annelation (angles of 120° or 240° starting from one extremity of the catafusene, keeping the right → left direction of annelation corresponding to digits 1 and 2, respectively, and choosing among possible notations by the minimal number formed by digits 0, 1, and 2). In particular, the number of adjacent zeros (i.e., of linearly condensed benzenoid rings) is closely connected to the energy corresponding to the $^1L_b$ band (or *p*-band in Clar's nomenclature)[4,5] in the electronic absorption spectra.

The connection between absorption spectra and half-wave reduction potentials of benzenoids ($E_{½}$ ) has been repeatedly mentioned in the literature.[6–18] In both cases the lowest unoccupied molecular orbital (LUMO) is involved, and electron affinities or other quantum-chemical parameters are also derived from such experimental data.[19–28]

So far, practically all quantitative structure-property relationships (QSPR) involving $E_{½}$ (including ref.[1]) have resulted in correlations in terms of $E_{LUMO}$. The literature on $E_{½}$ data was reviewed in several papers.[29–33]

In the present study we find that a second strongly correlated parameter (not yet tested in QSPR studies of $E_{½}$) is the number of adjacent zeros in the code of catafusenes. We will denote this parameter by $n$ = 1 + the number of adjacent zeros in the code. More subtle topological differences between catafusenes (having the same numbers $h$ of benzenoid rings and numbers of anthracene portions) may be encoded in topological indices of benzenoids, possibly to be used as a third parameter in the correlations between structure and $E_{½}$ values. We restrict our study to all 31 catafusene values for $E_{½}$ reported in Bergman's paper,[8] whereas ref.[1] had included only 26 catafusene values. Because $E_{½}$ values (in Volts) are negative numbers, we use $-E_{½}$ values for convenience in the Tables and Figures. $E_{½}$ depends on the difference between $E_{HOMO}$ and $E_{LUMO}$; these potentials have been shown to be a linear function of both $E_{HOMO}$ and $E_{LUMO}$ and can be correlated with $E_{LUMO}$, because the "chemical non-crossing rule" is obeyed in this case.

## Results and Discussion

### Data on half-wave reduction potentials and molecular descriptors for catafusenes

Data of $E_{½}$ (in Volt) vs. the saturated calomel electrode using ethylene glycol monomethyl ether (cellosolve) with tetrabutylammonium iodide as supporting electrolyte as well as $E_{LUMO}$ values were taken from Bergman's paper.[8] An extended discussion of the reaction mechanism, the fate of the intermediate radical anion and its subsequent protonation, its ESR spectrum, and the solvents used in the polarography of benzenoids can be found in the literature..[29–33] Aprotic solvents such as acetonitrile, dioxane, tetrahydrofuran, ethylenediamine, dimethylformamide, or tetramethylurea present advantages over aqueous mixtures of ethanol, isopropyl alcohol, dioxane, tetrahydrofuran. Strongly acidic media such as sulfuric acid and the presence of oxygen gas considerably influence the mechanism of the electrolytic reduction, i.e., the number of electrons involved and the intermediate products. Many benzenoids have more than one (i.e., two or three) half-wave reduction steps. Cellosolve was preferred as solvent because it dissolves even pentacene.

The molecular descriptors used in the present study include the Wiener index (W),[34] the Szeged index (Sz),[35–38] the Ivan-Padmakar index (IP),[39–42] and the average distance-based connectivity index (J);[43] also several first and second order indices related to each other, namely the Randić-Kier-Hall indices ($^1\chi$ and $^2\chi$),[44] the valence Kier-Hall indices taking bond multiplicity into consideration ($^1\chi^v$ and $^2\chi^v$), and the differential connectivity indices ($\Delta^1$ and $\Delta^2$).[45,46] Zero-order indices were not included because of their high degeneracy for the present set of compounds. For catafusenes with $h$ benzenoid rings, their nominal molecular weight (NMW) is linearly correlated with $h$ by the relationship: NMW = 28 + 50$h$. All these indices were computed using DRAGON software.[47] Structural optimization was carried out using HyperChem[48] and ACD Labs software.[49] Calculations of topological indices were performed by stepwise regression using the method of maximum-$R^2$.[50,51]

### Results and discussion of QSPR for $E_{½}$ values

Figure 1 presents the structures, dualist graphs, and codes of the 31 catafusenes in the same order as in Bergman's paper[8] (which also contains 47 perifusenes that are not discussed in the present paper). Table 1 presents $-E_{½}$ values, codes, and the molecular descriptors. Table 2 contains the intercorrelations factors between the molecular descriptors and shows that all four χ-type connectivity indices have intercorrelation factors above 0.99.

Table 3 presents monoparametric and biparametric correlations for $-E_{½}$ values. By far the highest correlations are seen for $E_{LUMO}$ and $n$. Figure 2 indicates that the linear monoparametric correlations between $-E_{½}$ and $E_{LUMO}$ on one hand, and between $-E_{½}$ and $n$ (calculated as 1 + the number of adjacent zeros in the code) on the other hand, present the highest correlation coefficients ($R^2$ = 0.958 and 0.920, respectively). Although in this study $n$ takes only integer values between 1 and 4, there is practically no overlap between the ranges of $-E_{½}$ values for these four integers; thus, the average values for each range almost perfectly vary linearly with $n$.

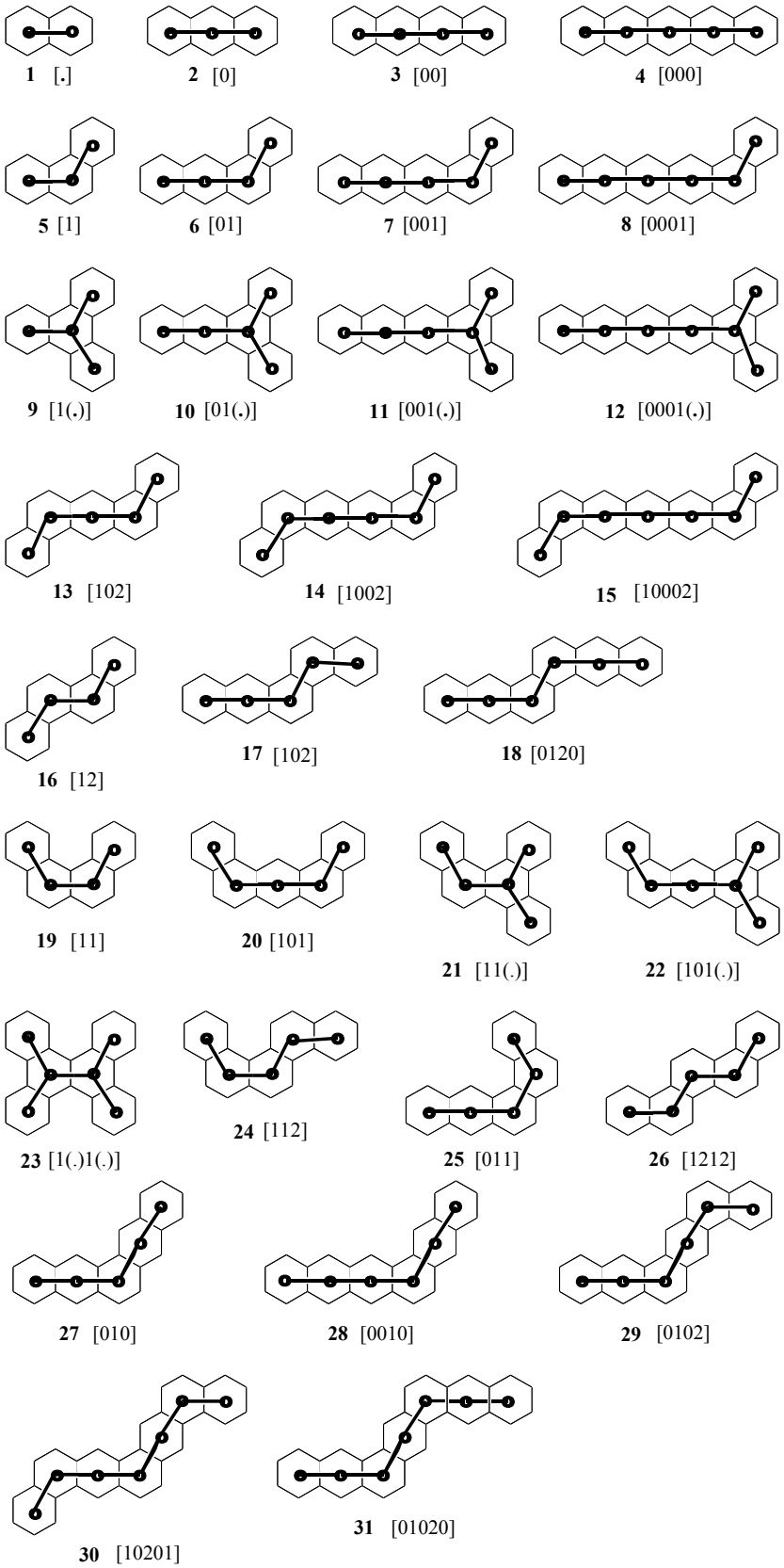


**Figure 1**. Structures of catafusenes **1**–**31**, their dualist graphs, and their codes in brackets.

**Table 1**. Values of $E_{1/2}$, $h$, $n$ , $E_{LUMO}$ and topological indices used

| Comp | $E_{1/2}$ | Code | $h$ | $n$ | $E_{LUMO}$ | W | Sz | PI | J | $^1\chi$ | $^2\chi$ | $^1\chi^v$ | $^2\chi^v$ | $^1\Delta$ | $^2\Delta$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1.** | 1.98 | [.] | 2 | 1 | 0.618 | 109 | 243 | 96 | 1.925 | 4.966 | 4.089 | 3.405 | 2.347 | 1.561 | 1.742 |
| **2.** | 1.46 | [0] | 3 | 2 | 0.414 | 279 | 656 | 216 | 1.682 | 6.933 | 6.081 | 4.809 | 3.547 | 2.124 | 2.534 |
| **3.** | 1.14 | [00] | 4 | 3 | 0.295 | 569 | 1381 | 404 | 1.465 | 8.899 | 8.072 | 6.214 | 4.746 | 2.685 | 3.326 |
| **4.** | 0.86 | [000] | 5 | 4 | 0.220 | 1011 | 2506 | 635 | 1.288 | 10.865 | 10.064 | 7.619 | 5.945 | 3.246 | 4.119 |
| **5.** | 1.94 | [1] | 3 | 1 | 0.605 | 271 | 677 | 218 | 1.740 | 6.949 | 5.994 | 5.095 | 3.804 | 1.854 | 2.190 |
| **6.** | 1.53 | [01] | 4 | 2 | 0.452 | 553 | 1325 | 388 | 1.512 | 8.916 | 7.986 | 6.500 | 5.003 | 2.416 | 2.983 |
| **7.** | 1.19 | [001] | 5 | 3 | 0.327 | 987 | 2410 | 606 | 1.321 | 10.882 | 9.977 | 7.904 | 6.203 | 2.978 | 3.774 |
| **8.** | 0.95 | [0001] | 6 | 4 | 0.244 | 1605 | 3975 | 587 | 1.169 | 12.848 | 11.969 | 9.309 | 7.402 | 3.539 | 4.567 |
| **9.** | 1.97 | [1(.)] | 4 | 1 | 0.684 | 513 | 1269 | 390 | 1.642 | 8.949 | 7.837 | 6.232 | 4.637 | 2.717 | 3.200 |
| **10.** | 1.54 | [01(.)] | 5 | 2 | 0.499 | 903 | 2290 | 610 | 1.451 | 10.916 | 9.828 | 8.196 | 6.418 | 2.720 | 3.410 |
| **11.** | 1.21 | [001(.)] | 6 | 3 | 0.356 | 1485 | 3783 | 593 | 1.271 | 12.882 | 11.82 | 9.601 | 7.618 | 3.281 | 4.202 |
| **12.** | 0.93 | [0001(.)] | 7 | 4 | 0.220 | 2279 | 5836 | 1194 | 1.240 | 14.848 | 13.811 | 11.005 | 8.817 | 3.843 | 4.994 |
| **13.** | 1.55 | [102] | 5 | 2 | 0.474 | 971 | 2354 | 418 | 1.346 | 10.899 | 9.891 | 7.910 | 6.164 | 2.989 | 3.727 |
| **14.** | 1.25 | [1002] | 6 | 3 | 0.348 | 1581 | 3879 | 878 | 1.188 | 12.865 | 11.882 | 9.315 | 7.363 | 3.550 | 4.519 |
| **15.** | 0.85 | [10002] | 7 | 4 | 0.186 | 2407 | 5964 | 1194 | 1.061 | 14.832 | 13.874 | 10.720 | 8.563 | 4.112 | 5.311 |
| **16.** | 1.81 | [12] | 4 | 1 | 0.520 | 545 | 1301 | 390 | 1.538 | 8.933 | 7.899 | 6.767 | 5.283 | 2.166 | 2.616 |
| **17.** | 1.44 | [012] | 5 | 2 | 0.405 | 971 | 2354 | 610 | 1.348 | 10.899 | 9.891 | 7.893 | 6.184 | 3.006 | 3.707 |
| **18.** | 1.33 | [0120] | 6 | 2 | 0.348 | 1573 | 3847 | 880 | 1.199 | 12.865 | 11.882 | 9.297 | 7.380 | 3.568 | 4.502 |
| **19.** | 1.75 | [11] | 4 | 1 | 0.568 | 529 | 1253 | 390 | 1.587 | 8.933 | 7.909 | 6.761 | 5.304 | 2.172 | 2.605 |
| **20.** | 1.57 | [101] | 5 | 2 | 0.492 | 955 | 2290 | 610 | 1.367 | 10.899 | 9.891 | 7.910 | 6.164 | 2.989 | 3.727 |
| **21.** | 1.65 | [11(.)] | 5 | 1 | 0.532 | 883 | 2202 | 420 | 1.492 | 10.933 | 9.751 | 8.202 | 6.383 | 2.731 | 3.368 |
| **22.** | 1.57 | [101(.)] | 6 | 2 | 0.522 | 1453 | 3647 | 597 | 1.300 | 12.899 | 11.733 | 9.607 | 7.579 | 3.292 | 4.154 |
| **23.** | 1.59 | [1(.)1(.)] | 6 | 1 | 0.512 | 1333 | 3447 | 599 | 1.426 | 12.933 | 11.603 | 9.619 | 7.515 | 3.314 | 4.088 |
| **24.** | 1.73 | [112] | 5 | 1 | 0.550 | 931 | 2234 | 420 | 1.408 | 10.916 | 9.814 | 7.899 | 6.148 | 3.017 | 3.666 |
| **25.** | 1.40 | [011] | 5 | 2 | 0.419 | 939 | 2258 | 610 | 1.395 | 10.899 | 9.900 | 7.893 | 6.189 | 3.006 | 3.711 |
| **26.** | 1.79 | [1212] | 5 | 1 | 0.502 | 963 | 2330 | 420 | 1.361 | 10.916 | 9.804 | 8.178 | 6.438 | 2.738 | 3.366 |
| **27.** | 1.53 | [010] | 5 | 2 | 0.437 | 979 | 2378 | 591 | 1.335 | 10.882 | 9.977 | 8.164 | 6.564 | 2.718 | 3.413 |
| **28.** | 1.22 | [0010] | 6 | 3 | 0.336 | 1589 | 3903 | 876 | 1.184 | 12.848 | 11.969 | 9.569 | 7.763 | 3.279 | 4.206 |
| **29.** | 1.50 | [0102] | 6 | 2 | 0.429 | 1573 | 3847 | 595 | 1.197 | 12.865 | 11.882 | 9.557 | 7.736 | 3.308 | 4.146 |
| **30.** | 1.52 | [10201] | 7 | 2 | 0.456 | 2375 | 5836 | 1206 | 1.080 | 14.848 | 13.787 | 10.967 | 8.897 | 3.881 | 4.890 |
| **31.** | 1.47 | [01020] | 7 | 2 | 0.435 | 2383 | 5860 | 1200 | 1.077 | 14.832 | 13.874 | 10.961 | 8.932 | 3.871 | 4.942 |

Catafusenes have been described in the literature with codes differing by interchanging digits 1 and 2; e.g., compounds **13** and **20**, **16** and **19**, **17** and **25**, **24** and **26** are isoarithmic, i.e., they have the same number of Kekulé structures.[52] Also their $-E_{½}$ values are very close.

The plot of $-E_{½}$ vs. $E_{LUMO}$ has a positive slope, the plot of $-E_{½}$ vs. $n$ has a negative slope. The former variation is due to an easier access of the available first LUMO. The latter variation is due to the fact that upon adding an electron to acene substructures the resulting radical anion has two Clar sextets (as in 9,10-dihydroanthracene), whereas the acene portion has only one Clar sextet. The longer the acene substructure the higher the delocalization of the π-electron system and the more stable the reaction product.

**Table 2**. Intercorrelation matrix for the parameters reported in Table 1

| | $E_{1/2}$ | $h$ | $E_{LUMO}$ | $n$ | W | Sz | PI | J | $^1\chi$ | $^2\chi$ | $^1\chi^v$ | $^2\chi^v$ | $^1\Delta$ | $^2\Delta$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $E_{1/2}$ | 1.000 | -0.577 | 0.979 | -0.960 | -0.580 | -0.583 | -0.591 | 0.671 | -0.573 | -0.599 | -0.531 | -0.530 | -0.670 | -0.701 |
| $h$ | -0.577 | 1.000 | -0.539 | 0.504 | 0.959 | 0.959 | 0.894 | -0.939 | 0.999 | 0.999 | 0.995 | 0.993 | 0.994 | 0.997 |
| $E_{LUMO}$ | 0.979 | -0.538 | 1.000 | -0.93 | -0.560 | -0.561 | -0.578 | 0.663 | -0.534 | -0.563 | -0.495 | -0.499 | -0.624 | -0.658 |
| $n$ | -0.960 | 0.504 | -0.934 | 1.000 | 0.540 | 0.545 | 0.553 | -0.611 | 0.499 | 0.529 | 0.458 | 0.458 | 0.605 | 0.642 |
| W | 0.580 | 0.959 | -0.560 | 0.540 | 1.000 | 0.999 | 0.943 | -0.901 | 0.957 | 0.963 | 0.947 | 0.949 | 0.941 | 0.947 |
| Sz | -0.583 | 0.959 | -0.561 | 0.545 | 0.999 | 1.000 | 0.940 | -0.893 | 0.959 | 0.963 | 0.9480 | 0.950 | 0.940 | 0.946 |
| PI | -0.591 | 0.894 | -0.578 | 0.553 | 0.943 | 0.940 | 1.000 | -0.855 | 0.893 | 0.901 | 0.877 | 0.881 | 0.894 | 0.899 |
| J | 0.671 | -0.939 | 0.663 | -0.611 | -0.901 | -0.893 | -0.855 | 1.000 | -0.938 | -0.949 | -0.923 | -0.929 | -0.936 | -0.946 |
| $^1\chi$ | -0.573 | 0.999 | -0.534 | 0.499 | 0.958 | 0.959 | 0.893 | -0.938 | 1.000 | 0.999 | 0.996 | 0.994 | 0.962 | 0.965 |
| $^2\chi$ | -0.599 | 0.995 | -0.563 | 0.529 | 0.963 | 0.963 | 0.901 | -0.949 | 0.999 | 1.000 | 0.993 | 0.991 | 0.968 | 0.971 |
| $^1\chi^v$ | -0.531 | 0.985 | -0.495 | 0.458 | 0.948 | 0.949 | 0.877 | -0.923 | 0.996 | 0.993 | 1.000 | 0.999 | 0.935 | 0.938 |
| $^2\chi^v$ | -0.530 | 0.993 | -0.499 | 0.458 | 0.950 | 0.950 | 0.881 | -0.929 | 0.994 | 0.991 | 0.999 | 1.000 | 0.929 | 0.934 |
| $^1\Delta$ | 0.670 | 0.994 | -0.624 | 0.605 | 0.941 | 0.940 | 0.894 | -0.936 | 0.963 | 0.967 | 0.935 | 0.929 | 1.000 | 0.998 |
| $^2\Delta$ | -0.701 | 0.997 | -0.658 | 0.642 | 0.947 | 0.947 | 0.899 | -0.946 | 0.965 | 0.971 | 0.938 | 0.934 | 0.998 | 1.000 |

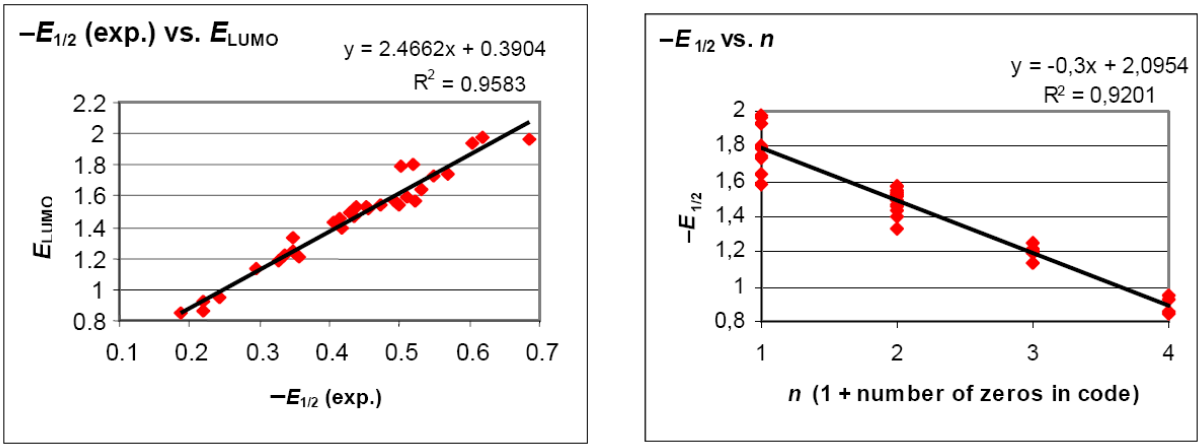


**Figure 2**. Monoparametric correlations of $-E_{½}$ vs. $E_{LUMO}$ (left plot) and $-E_{½}$ vs. $n$ (right plot).

Statistical parameters are the correlation coefficient ($R^2$), the standard error (*S.E.*), and the Fisher ratio (*F*). For validating the results we used the Pogliani quality parameter (*Q*).[53,54] $R^2_A$ denotes adjusted $R^2$. A higher correlation coefficient ($R^2$ = 0.974) is obtained in a biparametric correlation involving both $E_{LUMO}$ and *n* (model 15 according to Tables 3 and 4; cf. Figure 3).

**Table 3**. One- and two-variable modeling of half-wave potential ($E_{1/2}$)

| Model | Parameters used | *S.E.* | $R^2$ | $R^2_A$ | *F* | *Q* |
|---|---|---|---|---|---|---|
| 1 | $E_{LUMO}$ | 0.045 | 0.958 | 0.957 | 688.7 | 21.7 |
| 2 | *h* | 0.177 | 0.333 | 0.309 | 14.4 | 3.2 |
| 3 | *n* | 0.061 | 0.920 | 0.917 | 335.0 | 15.7 |
| 4 | W | 0.176 | 0.335 | 0.312 | 14.6 | 3.2 |
| 5 | Sz | 0.175 | 0.340 | 0.317 | 14.9 | 3.3 |
| 6 | PI | 0.174 | 0.349 | 0.327 | 15.5 | 3.3 |
| 7 | J | 0.160 | 0.450 | 0.431 | 23.7 | 4.1 |
| 8 | $^1\chi$ | 0.177 | 0.327 | 0.304 | 14.1 | 3.2 |
| 9 | $^2\chi$ | 0.173 | 0.358 | 0.336 | 16.2 | 3.4 |
| 10 | $^1\chi^v$ | 0.183 | 0.282 | 0.257 | 11.3 | 2.9 |
| 11 | $^2\chi^v$ | 0.183 | 0.280 | 0.258 | 11.3 | 2.8 |
| 12 | $^1\Delta$ | 0.160 | 0.447 | 0.428 | 23.5 | 4.1 |
| 13 | $^2\Delta$ | 0.154 | 0.492 | 0.474 | 28.1 | 4.5 |
| 14 | $E_{LUMO}$, *h* | 0.044 | 0.961 | 0.959 | 348.0 | 22.2 |
| 15 | $E_{LUMO}$, *n* | 0.036 | 0.974 | 0.971 | 518.3 | 27.4 |
| 16 | $E_{LUMO}$, *W* | 0.044 | 0.959 | 0.956 | 327.5 | 22.2 |
| 17 | $E_{LUMO}$, Sz | 0.044 | 0.959 | 0.956 | 329.9 | 22.2 |
| 18 | $E_{LUMO}$,PI | 0.044 | 0.958 | 0.955 | 324.5 | 22.2 |
| 19 | $E_{LUMO}$, *J* | 0.044 | 0.958 | 0.955 | 323.8 | 22.2 |
| 20 | $E_{LUMO}$, $^1\chi$ | 0.043 | 0.961 | 0.958 | 346.4 | 22.7 |
| 21 | $E_{LUMO}$, $^2\chi$ | 0.043 | 0.961 | 0.958 | 345.0 | 22.7 |
| 22 | $E_{LUMO}$, $^1\chi^v$ | 0.043 | 0.960 | 0.957 | 340.5 | 22.7 |
| 23 | $E_{LUMO}$, $^2\chi^v$ | 0.044 | 0.960 | 0.957 | 335.9 | 22.2 |
| 24 | $E_{LUMO}$, $^1\Delta$ | 0.042 | 0.963 | 0.960 | 366.5 | 23.3 |
| 25 | $E_{LUMO}$, $^2\Delta$ | 0.042 | 0.963 | 0.960 | 367.9 | 23.3 |
| 26 | *h, n* | 0.058 | 0.931 | 0.927 | 191.5 | 16.6 |
| 27 | *h,* W | 0.178 | 0.341 | 0.294 | 7.2 | 3.2 |
| 28 | *h*, Sz | 0.178 | 0.343 | 0.297 | 7.3 | 3.2 |
| 29 | *h*, PI | 0.176 | 0.361 | 0.315 | 7.9 | 3.4 |
| 30 | *h,* J | 0.159 | 0.475 | 0.438 | 12.6 | 4.3 |
| 31 | *h*, $^1\chi$ | 0.131 | 0.644 | 0.618 | 25.3 | 6.1 |
| 32 | *h*, $^2\chi$ | 0.123 | 0.685 | 0.663 | 30.5 | 6.7 |
| 33 | *h*, $^1\chi^v$ | 0.147 | 0.548 | 0.516 | 17.0 | 5.0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 34 | *h*, $^{2}\chi^{v}$ | 0.158 | 0.479 | 0.441 | 12.8 | 4.3 |
| 35 | *h*, $^{1}\Delta$ | 0.153 | 0.512 | 0.477 | 14.7 | 4.6 |
| 36 | *h*, $^{2}\Delta$ | 0.131 | 0.646 | 0.621 | 25.5 | 6.1 |
| 37 | *n, W* | 0.060 | 0.925 | 0.920 | 173.8 | 16.0 |
| 38 | *n*, Sz | 0.060 | 0.925 | 0.920 | 173.8 | 16.0 |
| 39 | *n*, PI | 0.060 | 0.925 | 0.920 | 174.3 | 16.0 |
| 40 | *n, J* | 0.057 | 0.931 | 0.927 | 191.4 | 16.9 |
| 41 | *n*, $^{1}\chi$ | 0.057 | 0.931 | 0.927 | 191.6 | 16.9 |
| 42 | *n*, $^{2}\chi$ | 0.057 | 0.931 | 0.926 | 191.2 | 16.9 |
| 43 | *n*, $^{1}\chi^{v}$ | 0.057 | 0.931 | 0.926 | 189.9 | 16.9 |
| 44 | *n*, $^{2}\chi^{v}$ | 0.057 | 0.931 | 0.926 | 188.8 | 16.9 |
| 45 | *n*,$^{1}\Delta$ | 0.057 | 0.932 | 0.927 | 193.8 | 16.9 |
| 46 | *2*, $^{2}\Delta$ | 0.057 | 0.932 | 0.927 | 193.8 | 16.9 |

Using the other molecular descriptors presented in Table 1, we attempted to explore triparametric correlations in order to see whether more subtle differences in the topology of catafusenes may be accounted for by such indices. The results of triparametric correlations in terms of $E_{LUMO}$, *n*, and one other molecular descriptor are displayed in Table 5. It can be seen that no distance-based topological index provides a better correlation than the number *h* of benzenoid rings. Since the best triparametric correlation in terms of $E_{LUMO}$, *n*, and *h* increases the value of $R^2$ only from 0.974 (model 15 in Table 3 and Figure 3) to 0.977, we consider that the biparametric model 15 is sufficiently satisfactory. The exceptionally high deviation for compound **23** is due to the exceptional branching (see next section).

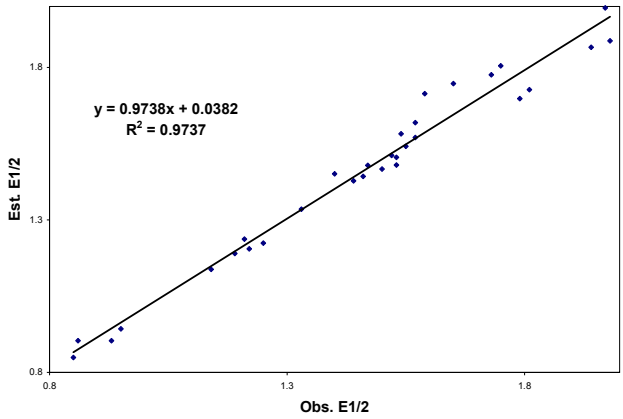


**Figure 3**. Correlation of observed and calculated $E_{1/2}$ using model 15 in terms of $E_{LUMO}$ and *n*.

**Table 4**. Comparison of observed and calculated $E_{1/2}$ using model 15, and the residuals

| Comp. | $E_{1/2}$ (exp.) | $E_{1/2}$ (calc.) | Residual |
|---|---|---|---|
| **1** | 1.980 | 1.887 | 0.093 |
| **2** | 1.460 | 1.443 | 0.017 |
| **3** | 1.140 | 1.138 | 0.002 |
| **4** | 0.860 | 0.904 | -0.044 |
| **5** | 1.940 | 1.866 | 0.074 |
| **6** | 1.530 | 1.505 | 0.025 |
| **7** | 1.190 | 1.190 | 0.000 |
| **8** | 0.950 | 0.943 | 0.007 |
| **9** | 1.970 | 1.995 | -0.025 |
| **10** | 1.540 | 1.582 | -0.042 |
| **11** | 1.210 | 1.237 | -0.027 |
| **12** | 0.930 | 0.904 | 0.026 |
| **13** | 1.550 | 1.541 | 0.009 |
| **14** | 1.250 | 1.224 | 0.026 |
| **15** | 0.850 | 0.849 | 0.001 |
| **16** | 1.810 | 1.727 | 0.083 |
| **17** | 1.440 | 1.428 | 0.012 |
| **18** | 1.330 | 1.335 | -0.005 |
| **19** | 1.750 | 1.805 | -0.055 |
| **20** | 1.570 | 1.570 | 0.000 |
| **21** | 1.650 | 1.747 | -0.097 |
| **22** | 1.570 | 1.619 | -0.049 |
| **23** | 1.590 | 1.714 | -0.124 |
| **24** | 1.730 | 1.776 | -0.046 |
| **25** | 1.400 | 1.451 | -0.051 |
| **26** | 1.790 | 1.698 | 0.092 |
| **27** | 1.530 | 1.480 | 0.050 |
| **28** | 1.220 | 1.205 | 0.015 |
| **29** | 1.500 | 1.467 | 0.033 |
| **30** | 1.520 | 1.511 | 0.009 |
| **31** | 1.470 | 1.479 | -0.009 |

**Table 5**. Three variable modeling of half-wave reduction potentials ($E_{1/2}$ )

| Model | Parameters used | *S. E.* | $R^2$ | $R^2_A$ | *F* | *Q* |
|---|---|---|---|---|---|---|
| 47 | $E_{LUMO}$, *h, n* | 0.033 | 0.977 | 0.975 | 384.1 | 29.9 |
| 48 | $E_{LUMO}$, *n, W* | 0.035 | 0.974 | 0.971 | 334.9 | 29.9 |
| 49 | $E_{LUMO}$, *n*, Sz | 0.035 | 0.974 | 0.971 | 346.9 | 28.1 |
| 50 | $E_{LUMO}$, *n*, PI | 0.035 | 0.974 | 0.971 | 342.1 | 28.1 |
| 51 | $E_{LUMO}$, *n, J* | 0.035 | 0.974 | 0.972 | 348.8 | 28.1 |
| 52 | *h, n*, W | 0.054 | 0.940 | 0.933 | 141.3 | 17.9 |
| 53 | *h, n,* Sz | 0.054 | 0.940 | 0.933 | 141.6 | 17.9 |
| 54 | *h, n,* PI | 0.058 | 0.933 | 0.925 | 125.7 | 16.6 |
| 55 | *h, n,* J | 0.058 | 0.932 | 0.924 | 123.9 | 16.6 |

**Parallelism between polarographic half-wave reduction potentials and Diels-Alder reactivity of catafusenes**

With samples of benzenoids provided by Clar in the 1980s, Biermann and Schmidt measured spectrophotometrically the rates of Diels-Alder reactions of polycyclic benzenoids with an excess of maleic anhydride under carefully controlled conditions in 1,2,4-trichlorobenzene solution at 91.5 ± 0.1 °C.[55–57] It was known that anthracene readily reacts with dienophiles or arynes, and that phenanthrene or other fibonacenes[58] do not. Fibonacenes have no zeros in their code. The reaction rate is presented in the following as *V*(diene) = 6 + log $k_2$ ($k_2$ is in [$LM^{-1}s^{-1}$]). If there are several sites for the cycloaddition, the measured rate may be divided by the number of sites (which can be 1, 2, or 3) to obtain *V*(diene)/site. In 1985, in a joint paper with one of us, the diene reactivity of benzenoids was rationalized using dualist graphs and equations derived from their topology.[59] The variables were the longest acene portion (parameter *n*) and a "form parameter" (*f*) depending on the annelation type at the ends of the acene moiety.[60] The values of this parameter *f* ranged from 1 (for no kinked annelation) to 9 (for four kinked ends). The straight lines calculated from equation (1) for the reaction rate (log $k_2$) meet at one point.[60]

$$(6 + \log k_2)_{calc} = (9.24 + 0.50f)\log n + 0.81 - 0.53f \qquad \text{Eq. (1)}$$

Values of the shape parameter *f* can be seen in Figure 4.

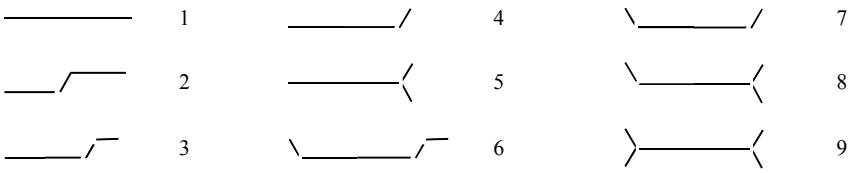


**Figure 4.** Shape parameter *f* for catafusenes. The long line symbolizes an acene part of the dualist graph, and the short lines a fibonacenic part.

Parameter *n* influences both the Diels-Alder reactivity and the polarographic reduction potentials of benzenoids; Table 6 and Figure 5 present data and correlations between these two sets of experimental data. With no anthracene moiety (no zeros in the code) the catafusene does not undergo a Diels-Alder addition with maleic anhydride. There is a significant correlation and not much of a difference between correlations involving the measured reaction rate or the measured reaction rate per site.

**Table 6**. Catafusenes with parameters, Diels-Alder reactivity and half-wave reduction potentials

| Comp. | Code | *h* | $E_{LUMO}$ | *f* | *n* | $V_{diene}$ | $V_{diene}$ /site | $E_{½}$ (V) |
|---|---|---|---|---|---|---|---|---|
| **1** | [.] | 2 | 0.6180 | 0 | 1 | - | - | 1.98 |
| **2** | [0] | 3 | 0.4142 | 1 | 2 | 3.36 | 3.36 | 1.46 |
| **3** | [00] | 4 | 0.2950 | 1 | 3 | 4.97 | 4.67 | 1.135 |
| **4** | [000] | 5 | 0.2197 | 1 | 4 | 6.22 | 6.22 | 0.86 |
| **5** | [1] | 3 | 0.6052 | 0 | 1 | - | - | 1.935 |
| **6** | [01] | 4 | 0.4523 | 4 | 2 | 2.13 | 2.13 | 1.53 |
| **7** | [001] | 5 | 0.3271 | 4 | 3 | 4.23 | 4.05 | 1.19 |
| **8** | [0001] | 6 | 0.2436 | 4 | 4 | 5.65 | 5.46 | 0.945 |
| **9** | [1(.)] | 4 | 0.6840 | 0 | 1 | - | - | 1.97 |
| **10** | [01(.)] | 5 | 0.4991 | 5 | 2 | 1.83 | 1.83 | 1.54 |
| **11** | [001(.)] | 6 | 0.3557 | 5 | 3 | 3.77 | 3.77 | 1.205 |
| **12** | [0001(.)] | 7 | 0.2200 | 5 | 4 | 5.45 | 5.45 | 0.93 |
| **13** | [102] | 5 | 0.4735 | 7 | 2 | 1.02 | 1.02 | 1.545 |
| **14** | [1002] | 6 | 0.3480 | 7 | 3 | 3.09 | 2.79 | 1.25 |
| **15** | [10002] | 7 | 0.1860 | 7 | 4 | 4.88 | 4.88 | 0.85 |
| **16** | [12] | 4 | 0.5201 | 0 | 1 | - | - | 1.805 |
| **17** | [012] | 5 | 0.4048 | 5 | 2 | 2.58 | 2.58 | 1.435 |
| **18** | [0120] | 6 | 0.3482 | 6 | 2 | 3.03 | 2.78 | 1.33 |
| **19** | [11] | 4 | 0.5676 | 0 | 1 | - | - | 1.745 |
| **20** | [101] | 5 | 0.4918 | 4 | 2 | 1.02 | 1.02 | 1.57 |
| **21** | [11(.)] | 5 | 0.5319 | 0 | 1 | - | - | 1.645 |
| **22** | [101(.)] | 6 | 0.5224 | 6 | 2 | 0.53 | 0.53 | 1.57 |
| **23** | [1(.)1(.)] | 6 | 0.5115 | 0 | 1 | - | - | 1.59 |
| **24** | [112] | 5 | 0.5498 | 0 | 1 | - | - | 1.73 |
| **25** | [011] | 5 | 0.4186 | 3 | 2 | - | - | 1.4 |
| **26** | [1212] | 5 | 0.5019 | 0 | 1 | - | - | 1.79 |
| **27** | [010] | 5 | 0.4372 | 5 | 2 | 1.82 | 1.52 | 1.525 |
| **28** | [0010] | 6 | 0.3357 | 6 | 3 | 3.82 | 3.82 | 1.22 |
| **29** | [0102] | 6 | 0.4287 | 6 | 2 | 1.74 | 1.74 | 1.495 |
| **30** | [10201] | 7 | 0.4560 | 7 | 2 | 0.91 | 0.91 | 1.515 |
| **31** | [01020] | 7 | 0.4350 | 7 | 2 | 2.00 | 1.69 | 1.465 |

Because all three experimentally determined properties of catafusenes (electronic absorption spectra, half-wave reduction potentials, and Diels-Alder reactivity) are determined by the same factors, namely the energy levels of HOMO-LUMO, it is logical that there should be correlations between these data. It is remarkable that parameters derived from a simple graph-theoretical model (the dualist graph of the catafusene) can provide satisfactory correlations with the data characterizing these phenomena.

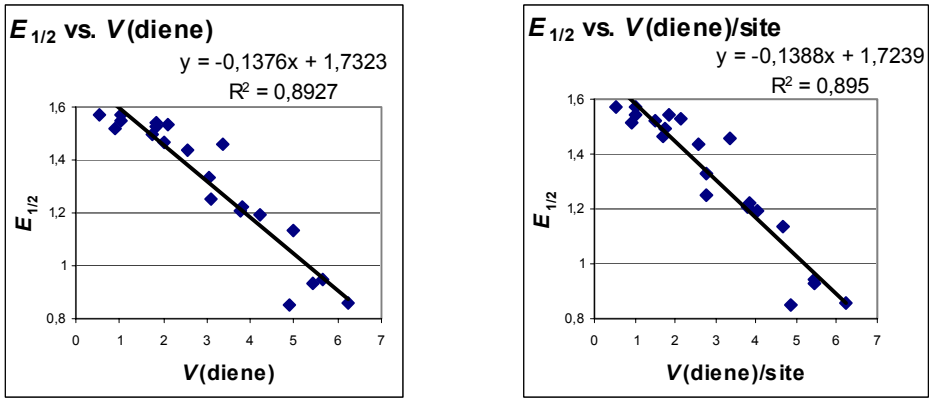


**Figure 5**. Correlation of Diels-Alder reaction rates *vs.* $E_{1/2}$ (left plot). Correlation of Diels-Alder reaction rates per site *vs.* $E_{1/2}$ (right plot).

**Attempted correlations between -$E_{1/2}$ and $E_{LUMO}$, $n$, and $f$**

The monoparametric dependence of $-E_{1/2}$ *versus* $E_{LUMO}$, illustrated by Fig. 2, becomes improved in the biparametric correlation in terms of $E_{LUMO}$ and $n$, displayed by Fig. 3. On adding the shape parameter $f$ for the dualist graph of the catafusene, the result is illustrated by Fig. 6.

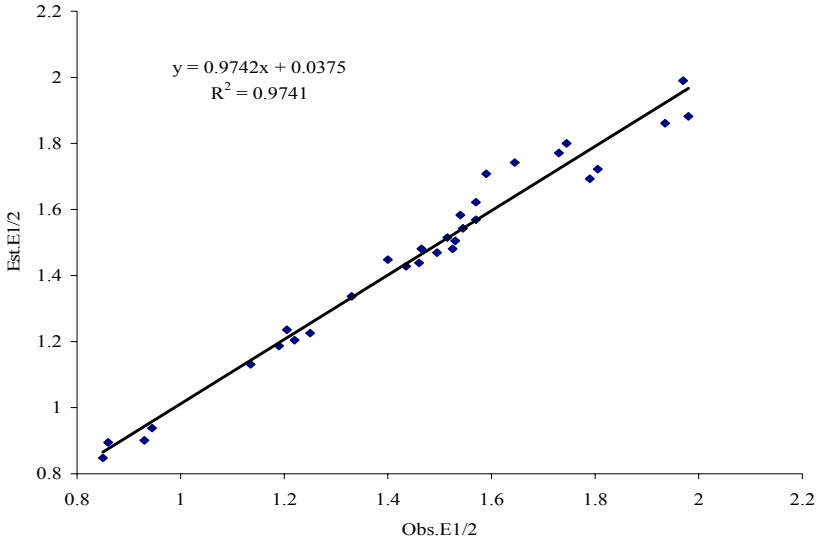


**Figure 6.** Correlation of observed and calculated $E_{½}$ using the model containing $E_{LUMO}$, $n$, and $f$ as the correlating parameters.

The regression parameters and quality of correlations for different models are presented below.

***Monoparametric regression***

$E_{1/2} = 0.390 + 2.466(\pm 0.098)E_{LUMO}$

$N = 31$, $S.E. = 0.044$, $R^2 = 0.958$, $R^2_A = 0.957$, $F = 666.1$

***Biparametric regression***

$E_{1/2} = 0.983 + 1.638(\pm 0.215)E_{LUMO} - 0.110(\pm 0.027)n$

$N = 31$, $S.E. = 0.035$, $R^2 = 0.974$, $R^2_A = 0.972$, $F = 525.9$

***Triparametric regression***

$E_{1/2} = 0.984 + 1.635(\pm 0.218)E_{LUMO} - 0.112(\pm 0.028)n + 0.001(\pm 0.004)f$

$N = 31$, $S.E. = 0.036$, $R^2 = 0.974$, $R^2_A = 0.971$, $f = 339.6$

It can be seen that upon adding $f$ as a third parameter (which for Diels-Alder reactivity had led to improved correlation), here there is no improvement of $R^2$, and the standard error even increases. Therefore triparametric correlations in terms of $E_{LUMO}$, $n$ and either $h$ or $f$ are practically equivalent and they do not add substantial improvements to the biparametric regression shown in Fig. 3 and in the Graphical Abstract in terms of $E_{LUMO}$ and $n$ ($R^2 = 0.974$, $S.E. = 0.036$).

## Conclusions

It is shown that dualist graphs of catafusenes encode sufficient information for explaining satisfactorily the variance of half-wave reduction potentials ($E_{½}$) of 31 *cata*-condensed benzenoids. For the same catafusenes, earlier work had outlined correlations between $E_{½}$ and the absorption maxima in the UV absorption spectra. Here we present a third type of phenomenon that is quantitatively related to $E_{½}$, namely reaction rates with dienophiles.

## References


1. Nikolić, S.; Milicević, A.; Trinajstić, N. *Croat. Chem. Acta* **2006**, *79*, 155.
2. Balaban, A. T.; Harary, F. *Tetrahedron* **1968**, *24*, 2505.
3. Balaban, A. T. *Tetrahedron* **1969**, *25*, 2949.
4. Clar, E. *Polycyclic Hydrocarbons*, Academic Press: New York, 1964.
5. Clar, E. *The Aromatic Sextet*, Wiley: London, 1972.
6. Pullman, A.; Pullman, B.; Berthier, G. *Bull. Soc. Chim. Fr*. **1950**, *17*, 591.
7. Watson, A. T.; Matsen, F. A. *J. Chem. Phys.* **1950**, *18*, 1305.
8. Bergman, I. *Trans. Faraday Soc*. **1954**, *50*, 829.
9. Hoijtink, G. J.; van Schooten, J. *Rec. Trav. Chim.* **1952**, *71*, 1089.
10. Hoijtink, G. J.; van Schooten, J. *Rec. Trav. Chim.* **1953**, *72*, 903.
11. Hoijtink, G. J.; van Schooten, J.; de Boer, E.; Aalbersberg, W. I. *Rec. Trav. Chim.* **1954**, *73*, 355.
12. Hoijtink, G. J. *Rec. Trav. Chim.* **1955**, *74*, 1525.
13. Bergman, I. *Trans. Faraday Soc*. **1956**, *52*, 690.
14. Basu, S.; Bhattacharya, R. *J. Chem. Phys.* **1956**, *25*, 596.
15. Fernandez-Alonso, J. I.; Domingo, R. *Nature* **1957**, *179*, 829.
16. Basu, S.; Bhattacharya, R. *Nature* **1957**, *180*, 143.
17. Bergman, I. *Experientia* **1962**, *18*, 46.
18. Loutfy, R. O.; Loutfy, R. O. *Can. J. Chem.* **1976**, *54*, 1454–1463.
19. Wawzonek, S.; Laitinen, H. A. *J. Am. Chem. Soc.* **1942**, *64*, 2365–2368.
20. Maccoll, A. Nature **1949**, *163*, 178.
21. Lyons, L. E. *Nature* **1950**, *166*, 193.
22. Given, P. H. *Nature* **1958**, *181*, 1001.
23. Given, P. H.; Peover, M. E. *J. Chem. Soc.* **1960**, 385.
24. Aalbersberg, W. I.; Mackor, E. L. *Trans. Faraday Soc.* **1960**, *56*, 1351.
25. Heilbronner, E.; Murrell, J. N. *Theor. Chim. Acta* **1963**, *1*, 235.
26. Peover, M. E. *Trans. Faraday Soc*. **1964**, *60*, 417.
27. Zahradník, R.; Párkányi, C. *Talanta* **1965**, *12*, 1289.

28. Shawali, A. S.; Elnadouli, B. E.; Párkányi, C.; Herndon, W. C. *Bull. Soc. Chim. Belg.* **1984**, *93*, 867–874.
29. Le Guillanton, G. *Bull. Soc. Chim. Fr*. **1963**, 2359.
30. Wawzonek, S. *Talanta*, **1965**, *12*, 1229–1235.
31. Peover, M. E. *Electroanalytical Chem.* **1967**, *2*, 1–51.
32. Zhdanov, S. I. *Russ. Chem. Rev*. **1969**, *38*, 608–625 (translated from *Usp. Khim*.**)**
33. Given, P. H.; Peover, M. E. *Advances in Polarography* **1960**, *3*, 948–964, Pergamon Press, Oxford.
34. Wiener, H. *J. Am. Chem. Soc.***1947,** *69***,** 17–20.
35. Gutman, I. *Graph Theory Notes* New York **1994,** *27,* 9–15.
36. Khadikar, P. V.; Kale, P. P.; Deshpande, N. V.; Karmarkar, S., Agrawal, V. K. *Commun. Math*. *Comput. Chem. (MATCH)***, 2001,** *43***,** 7–15.
37. Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Dobrynin, A.; Gutman, I.; Domotor, G. *J. Chem. Inf. Comput. Sci*. **1995,** *35,* 545–550.
38. Khadikar, P. V.; Karmarkar, S.; Agrawal, V. K.; Singh, J.; Shrivastava, A.; Lukovits, I.; Diudea, M. V. *Lett. Drug Design Discovery* **2005**, *2*, 606–624.
39. Khadikar, P. V. *Nat. Acad. Sci. Lett.* **2000**, *29,* 113–118.
40. Khadikar, P. V.; Karmarkar, S.; Agrawal, V. K. *J. Chem. Inf. Comput. Sci.* **2001**, *41,* 934–949.
41. Khadikar, P. V.; Kale, P. P.; Deshpande, N. V.; Karmarkar, S.; Agrawal, V. K. *J. Math. Chem.* **2001**, *29,* 143–150.
42. P. E. John, P. E.; Khadikar, P.V.; Singh, J. *J. Math. Chem.* **2006** (in press).
43. Balaban, A. T. *Chem. Phys. Lett.* **1982**, *89,* 399–404.
44. Randic, M. *J. Am. Chem. Soc.* **1975**, *97,* 6609–6615.
45. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure-Activity Analysis*, Research Studies Press-Wiley, Chichester, 1986.
46. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research,* Academic Press, New York. 1976.
47. DRAGON software for calculation of topological indices: http://www.disat.unimib.it
48. HyperChem-7 software for calculating molecular modeling parameters, http://www.hyper.com
49. ACD-Lab software for calculating physicochemical parameters; ChemSketch 3.0, http://www.acdlabs.com
50. Chatterjee, S.; Hadi, A. S.; Price, B. *Regression Analysis by Examples*, 3rd Ed., Wiley, New York, 2000.
51. NCSS, http://www.ncss.com.
52. Balaban, A. T; Tomescu, I. *Commun. Math. Comput. Chem. (MATCH)*, **1983**, *14*, 155–182.
53. Pogliani, L. *Amino Acids*, **1994,** *6,* 141–153; *J. Phys. Chem*. **1996**, *100,* 18065–18077.
54. Pogliani, L. *Chem. Rev.* **2000**, *100,* 3827–3858.
55. Biermann, D.; Schmidt, W. *J. Am. Chem. Soc.* **1980**, *102*, 3163–3173 ; 3173–3181.

56. Biermann, D.; Schmidt, W. *Isr. J. Chem.* **1980**, *20*, 312–318.
57. Hess, Jr, B.A.; Schaad, L. J.; Herndon, W. C.; Biermann, D.; Schmidt, W. *Tetrahedron* **1981**, *37*, 2983–2987.
58. Balaban, A. T. *Commun. Math. Comput. Chem. (MATCH)* **1989**, *24***,** 29–38
59. Balaban, A. T.; Biermann, D.; Schmidt, W. *Nouv. J. Chim.* **1985**, *9*, 443–449.
60. Balaban, A. T. *Pure Appl. Chem.* **1982**, *54*, 1075–1096.