Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

Bond angles and bond lengths in monosubstituted benzene and ethene derivatives: a comparison of computational and crystallographic results

Otto Exner[a] and Stanislav Böhm[b]*

[a]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, 16610 Praha 6, Czech Republic, and [b]Department of Organic Chemistry, Prague Institute of Chemical Technology, 16628 Praha 6, Czech Republic

Correspondence e-mail: bohm@vscht.cz

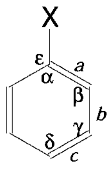
Received 28 March 2002
Accepted 12 June 2002

Bond angles and bond lengths in 29 monosubstituted benzene derivatives and in the same number of ethene derivatives were calculated at the B3LYP/6-311+G(*d*,*p*) level. Angle deformations in benzene derivatives agree reasonably with those derived statistically from the crystallographic data; in the case of small deformations, the calculated parameters are even more reliable. There is little correlation between geometry and reactivity parameters (σ-constants) in spite of some previous claims. Nevertheless, three components of the substitution effect can be distinguished: (*a*) strong deformation of the adjoining angles and bonds can be ascribed to changes of hybridization; (*b*) a weaker effect in the *meta* and *para* positions is only partially related to resonance; (*c*) in the case of unsymmetrical substituents, the symmetry of the benzene ring is also broken – the angular group-induced bond alternation (AGIBA) effect. The latter effect was also confirmed by searches in the Cambridge Structural Database for alkoxy, alkylthio, acyl and azo derivatives.

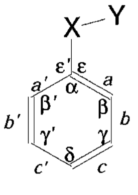
1. Introduction

Substituent effects in the ground state of neutral molecules are less regular and have been studied less (Topsom, 1987) than those in ions or in polar transition states. Particular attention has been focused on the geometric parameters: bond angles, bond lengths and dihedral angles. The restricted but not insignificant possibilities of this approach were appraised by Krygowski (1990). Domenicano and co-workers (Domenicano *et al.*, 1975*a*,*b*; Domenicano & Murray-Rust, 1979; Domenicano & Vaciago, 1979) and, independently, Norrestam & Schepper (1978, 1981) statistically evaluated the bond angles in benzene mono derivatives from selected X-ray structures. The result was a table of parameters $\Delta\alpha$, $\Delta\beta$, $\Delta\gamma$ and $\Delta\delta$ (Domenicano & Murray-Rust, 1979) denoting deviations of the angles α–δ (1) from 120° in the compounds C_6H_5X (see Scheme). Parameter $\Delta\gamma$ was related to the substituent resonance, and $\Delta\alpha$ depended mainly on its inductive effect or electronegativity (Domenicano *et al.*, 1975*b*). Within the framework of the STO-3G model, $\Delta\gamma$ was related to π-electron density and $\Delta\alpha$ to σ-electron density (Krygowski *et al.*, 1986). The main goal of these parameters should be to predict bond angles in benzene poly derivatives according to an additive scheme, but the estimated values are rather imprecise when *ortho* substituents are present [see for instance Tinant *et al.* (1993) or Císařová *et al.* (2000)]. The bond lengths *a*–*c* in (1) were also determined statistically, but the precision was low, mainly owing to thermal motion (Domenicano *et al.*, 1983).

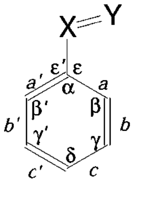
Within the accuracy achieved, it was assumed that the benzene ring in (1) has C_{2v} symmetry. More recently, Krygowski and co-workers (Krygowski, Anulewicz & Hiberty, 1996; Krygowski, Cyrański & Wisiorowski, 1996; Krygowski *et al.*, 1997; Howard *et al.*, 1998) discovered that unsymmetrical substituents coplanar with the ring plane also perturb the symmetry of the ring: internal bond angles α–δ, external angles ε and bond lengths *a*–*d*. This was called the AGIBA effect (angular group-induced bond alternation). This can be described as the prevalence of one Kekulé structure: substituents with a single bond favor structure (2); substituents with a double bond favor structure (3). The most convincing experimental proofs were obtained from certain polysubstituted benzenes and from the statistical treatment of the structures of azobenzenes (Krygowski, Anulewicz & Hiberty, 1996). The main theoretical support has been obtained from RHF calculations on 1,3,5-trisubstituted benzenes (Howard *et al.*, 1998). In all cases, attention was focused on the adjacent bonds: in (2), $a < a'$; in (3), $a > a'$. The geometry of monosubstituted benzenes was calculated by Bock *et al.* (1985) at the RHF/6-31G level.



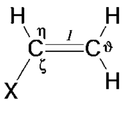
1



2



3



4

In our opinion, the problem is of sufficient importance to repeat the calculations on monosubstituted benzenes (1) within the framework of density functional theory (DFT) and for a larger set of 29 substituents. The same calculations were also carried out on monosubstituted ethenes (4). We believe that our calculations are on a sufficient level to fulfill two purposes: (*a*) provide a reliable and extended set of reference parameters related to isolated molecules, and (*b*) explore the relation between the geometry parameters and known scales of substituent effects. A possible minute systematic error of DFT in predicting bond lengths (Neugebauer & Häfelinger, 2002) is of no account when we are interested only in relative values.

2. Calculations

DFT calculations were performed at the B3LYP/6–311+G(*d*,*p*) level using the *GAUSSIAN*94 program of Frisch *et al.* (1995). Full geometry optimization and vibrational analysis were carried out in all cases. Minimum-energy conformations of some compounds have already been given in another connection (Exner & Böhm, 2001). In questionable cases, the optimization was started from two or more initial structures representing the anticipated less-populated conformations. The calculated geometrical parameters are listed in Table 1.

Searches of the Cambridge Structural Database (CSD) were carried out using version 5.22 (October 2001, 245 392 entries) and the program *ConQuest*, version 1.3. The partial structures searched were $C_6H_5CH{=}C$, C_6H_5CO-C, C_6H_5O-C, C_6H_5S-C and $(E)-C_6H_5N{=}N-C$. Two restrictions were introduced: *R* factor < 0.05, and all atoms in retrieved structures must have atomic weights less than 36. The molecules retrieved were subsequently restricted to approximately planar structures with the dihedral angle $C1-C2-X-Y$ between −20° and +20° and, in the case of the group C_6H_5O-C, with $C2-C1-O-C$ between −30° and +30°. The differences in bond lengths and bond angles were calculated; their median values are given in Table 2.

3. Results and discussion

3.1. Calculated and crystallographic bond angles

The substituent effects on geometrical parameters will be discussed in order of decreasing magnitude, beginning with the angular deformation in the benzene ring. Our parameters $\Delta\alpha$–$\Delta\delta$ (Table 1) calculated on isolated molecules of benzene mono derivatives were compared with the values determined from analysis of X-ray data of various derivatives. We expected differences due to packing forces and also due to considerable experimental errors. This has been confirmed. Of the angular parameters of Domenicano & Murray-Rust (1979), $\Delta\alpha$ and $\Delta\beta$ correlate with our values well (Table 3, lines 1 and 2), $\Delta\gamma$ reasonably (line 3) and $\Delta\delta$ poorly (line 4); this is understandable because the values of $\Delta\delta$ are small. The parameters of Norrestam & Schepper (1981) appear to be somewhat less reliable, but exact comparison is prevented by their small number. Moreover, there is an evident error in the values for the substituent COOH: probably the values for the coplanar and twisted substituents were exchanged. When this substituent is eliminated, the results are similar to those of Domenicano & Murray-Rust (1979), but there are only six degrees of freedom (see Table 3, line 5 as an example). Previous STO-3G calculations (Krygowski *et al.*, 1986) yielded $\Delta\alpha$–$\Delta\delta$ in much worse agreement with experiment (see Table 3, line 6 as an example). We can conclude that calculated and experimental values of angles $\Delta\alpha$–$\Delta\delta$ are in good agreement. Careful systematic selection of data can eliminate packing forces and yield reasonable values if they are not too small (*e.g.* not smaller than 0.2°).

3.2. Bond angles and reactivity parameters

In the next step we correlated the values $\Delta\alpha$–$\Delta\delta$ [and also the external angles at C(1), ε and ε', see (2) and (3)] with parameters expressing some defined or anticipated properties of substituents. The following parameters were tried: inductive (σ_I) and resonance (σ_R) constants (Charton, 1981), polarizability (σ_α) and electronegativity (σ_χ) constants (Taft & Topsom, 1987), steric constants υ (Charton, 1991), molar refraction (MR), and molecular weight; the last three were not significant in any case. We attempted many correlations by the methods of ascending regression and descending regression; in Table 3 we give only the best correlations for each response function and sometimes the correlations mentioned in the

Table 1
Calculated bond angles and bond lengths in (*a*) monosubstituted benzenes (1) and (*b*) monosubstituted ethenes (2)†.

(*a*) Substituent *X*	$\Delta\alpha$ C1	$\Delta\beta$ C2	$\Delta\beta'$ C6	$\Delta\gamma$ C3	$\Delta\gamma'$ C5	$\Delta\delta$ C4	$\Delta\varepsilon$	$\Delta\varepsilon'$	Δa C1—C2	$\Delta a'$ C1—C6	Δb C2—C3	$\Delta b'$ C5—C6	Δc C3—C4	$\Delta c'$ C4—C5
H	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CH_3	−1.8	1.0	1.0	0.2	0.2	−0.6	0.9	0.9	66	66	−20	−20	10	10
C_2H_5	−1.8	1.0	1.0	0.1	0.1	−0.5	0.9	0.9	53	53	−4	−4	−3	−3
$CH(CH_3)_2$	−2.0	1.1	1.2	0.3	0.1	−0.6	1.5	0.6	69	44	−14	6	5	−16
$C(CH_3)_3$	−2.7	1.6	1.6	0.3	0.3	−1.0	1.4	1.4	99	99	−28	−28	3	3
$CH{=}CH_2$	−2.1	0.9	1.3	0.4	0.0	−0.6	3.2	−1.1	108	90	−49	−14	25	−18
$C{\equiv}CH$	−1.0	0.3	0.3	0.3	0.3	−0.2	0.5	0.5	111	111	−36	−36	5	5
C_6H_5	−1.8	0.9	0.9	0.3	0.3	−0.6	0.9	0.9	86	86	−18	−18	−3	−3
CH_2Cl	−1.0	0.5	0.5	0.0	0.0	−0.2	0.5	0.5	46	46	−19	−19	0	0
CF_3	0.4	−0.4	−0.4	0.1	0.1	0.0	−0.2	−0.2	13	13	−18	−18	−4	−4
NH_2	−1.4	0.5	0.5	0.8	0.8	−1.1	0.7	0.7	87	87	−31	−31	3	3
$N(CH_3)_2$	−2.7	1.0	1.0	1.2	1.2	−1.6	1.4	1.4	181	181	−33	−33	−10	−10
$NHCOCH_3$	−0.8	0.4	0.1	0.3	0.6	−0.6	−1.2	2.0	63	47	−28	−12	0	−9
OH	0.1	−0.3	−0.4	0.5	0.8	−0.7	2.6	−2.7	16	18	−1	−34	−15	16
OCH_3	−0.2	−0.5	0.0	1.0	0.6	−0.8	4.5	−4.3	31	65	34	−64	−41	36
$OCOCH_3$	1.1	−0.8	−0.8	0.3	0.3	−0.2	−0.5	−0.5	−24	−24	−5	−5	2	2
SH	−0.5	0.0	0.1	0.6	0.6	−0.7	2.8	−2.3	41	50	−11	−23	−9	0
SCH_3	−0.9	0.1	0.4	0.8	0.5	−0.8	4.5	−3.6	41	90	15	−47	−30	19
F	2.6	−1.7	−1.7	0.5	0.5	−0.2	−1.3	−1.3	−80	−80	0	0	3	3
Cl	1.4	−1.0	−1.0	0.5	0.5	−0.3	−0.7	−0.7	−26	−26	−5	−5	−5	−5
Br	1.4	−1.1	−1.1	0.5	0.5	−0.2	−0.7	−0.7	−25	−25	2	2	−4	−4
CHO	−0.2	−0.1	0.2	0.0	−0.3	0.3	0.7	−0.6	74	48	−57	−18	40	−1
$COCH_3$	−1.0	0.5	0.4	0.0	0.0	0.0	−1.6	2.6	79	68	−51	−11	22	−8
COOH	−0.1	0.0	−0.2	0.0	0.1	0.1	−2.0	2.1	54	57	−40	−24	12	4
$COOCH_3$	−0.3	0.1	0.0	0.0	0.1	0.0	−2.1	2.3	53	57	−36	−22	10	4
CN	0.0	−0.3	−0.3	0.2	0.2	0.1	0.0	0.0	84	84	−38	−38	5	5
$N{=}NH$	0.4	−0.6	0.0	0.4	−0.2	0.2	4.3	−4.7	69	18	−60	−11	48	−12
NO_2	2.4	−1.5	−1.5	0.2	0.2	0.3	−1.2	−1.2	−31	−31	−27	−27	9	9
SO_2CH_3	1.7	−1.1	−1.1	0.1	0.1	0.3	−0.8	−0.8	−9	−9	−10	−10	6	6
Median absolute value	1.0	0.5	0.5	0.3	0.3	0.3	1.2	0.9	58	55	24	18	5	4
H (reference)	120	120	120	120	120	120	120	120	1.39425	1.39425	1.39425	1.39425	1.39425	1.39425

(*b*) Substituent *X*	$\Delta\zeta$	$\Delta\eta$	$\Delta\theta$	Δl C═C
H	0	0	0	0
CH_3	2.9	−2.6	0.1	18
C_2H_5	3.1	−2.4	0.1	23
$CH(CH_3)_2$	3.3	−2.2	0.1	24
$C(CH_3)_3$	4.2	−3.4	−0.3	38
$CH{=}CH_2$	2.2	−2.0	0.2	76
$C{\equiv}CH$	1.9	−1.8	0.9	80
C_6H_5	3.7	−2.8	0.1	66
CH_2Cl	1.4	−2.0	0.3	17
CF_3	1.0	1.1	1.0	−16
NH_2	4.5	−1.2	1.0	51
$N(CH_3)_2$	5.8	−1.9	0.6	99
$NHCOCH_3$	2.6	1.6	0.9	30
OH	5.3	1.2	1.1	12
OCH_3	0.4	1.0	1.3	−3
$OCOCH_3$	8.5	0.7	1.5	3
SH	1.1	−0.6	0.9	19
SCH_3	1.9	−0.8	0.9	34
F	0.5	4.2	2.6	−76
Cl	1.9	1.8	1.7	−32
Br	1.0	2.6	1.6	−32
CHO	−1.0	0.8	0.1	58
$COCH_3$	3.1	−0.6	−0.1	59
COOH	−1.5	1.0	1.5	24
$COOCH_3$	2.5	0.3	1.0	34
CN	0.6	−0.2	1.1	59
$N{=}NH$	1.9	1.4	1.8	30
NO_2	−0.8	5.5	2.6	−57
SO_2CH_3	−0.6	3.0	1.8	−10
Median absolute value	1.9	1.7	1.0	31
H (reference)	121.5	121.5	116.9	1.33544

† Relative values referenced to the unsubstituted hydrocarbon; bond lengths in 10^{-4} Å, bond angles in °. Reference values are given at the bottom.

Table 2
Krygowski AGIBA effect on bond lengths and bond angles in monosubstituted benzenes (1): comparison of calculated and crystallographic results†.

Substituent	$a - a'$	$b - b'$	$c - c'$	$\beta - \beta'$	$\gamma - \gamma'$	N‡
OH	−2	+33	−31	+0.1	−0.3	
OCH_3	−34	+98	−77	−0.55	+0.36	
O*R* CSD	**−25**	**+85**	**−70**	**−0.53**	**+0.30**	61
SH	−9	+12	−9	−0.1	0	
SCH_3	−49	+62	−49	−0.30	+0.24	
S*R* CSD	**−60**	**+50**	**0**	**−0.35**	**+0.27**	104
$CH{=}CH_2$	+18	−35	+43	−0.14	+0.38	
CH═C*XY* CSD	**+10**	**−10**	**+50**	**−0.30**	**+0.21**	435
CHO	+26	−39	+41	−0.3	+0.3	
$COCH_3$	+11	−40	+30	+0.06	+0.03	
CO*R* CSD	**+30**	**−50**	**0**	**+0.07**	**+0.05**	264
N═NH	+51	−49	+60	−0.61	+0.58	
N═N*R* CSD	**+10**§	**−10**	**+30**	**−0.48**	**+0.38**	84

† Bond lengths in 10^{-4} Å, bond angles in °. ‡ Number of hits selected from CSD. § Krygowski, Anulewicz & Hiberty (1996) obtained 66 × 10^{-4} Å from a set of 21 items.

literature. The correlations of $\Delta\alpha$ appear to be the most important since the values of $\Delta\alpha$ are greatest and have been most discussed. There is a relation to the inductive effect, but the correlation with σ_I is poor and is only slightly improved by including the polarizability constant σ_α (Table 3, line 7). We cannot confirm the claimed dependence on electronegativity σ_χ (Krygowski, 1990): the term with σ_χ is insignificant (Table 3, line 8). Most remarkably, correlation of $\Delta\beta$ is better than that of $\Delta\alpha$ (Table 3, line 9 as compared with line 7) while using the same explanatory variables; in particular, the term with σ_α is more significant.

Correlation of $\Delta\gamma$ with the resonance parameters was described as strongly non-linear, approaching a hyperbola (Domenicano & Murray-Rust, 1979). Fig. 1 reveals a very rough dependence with a negative slope but does not confirm a hyperbolic dependence. It also shows that the angles $\Delta\gamma$ and $\Delta\gamma'$ may sometimes be rather different, but their difference does not affect the overall correlation. A significant correlation of $\Delta\gamma$ requires three terms (Table 3, line 10), of which resonance is the most important. In this and in all subsequent cases when our results differ in a quantitative sense from the previous ones, we attribute the disagreements to our more extended data sets rather than to the difference between crystallographic and quantum chemical data.

The angles $\Delta\delta$ depend on the same reactivity parameters as $\Delta\gamma$, but the dependence on σ_R has a reversed slope (Table 3, line 11). This correlation is relatively good with respect to the generally small values of $\Delta\delta$. However, it cannot be simplified to σ_R only (Fig. 2); even a nonlinear regression would not help, as seen from an estimated curve. In conclusion, we can confirm the opinion (Domenicano *et al.*, 1983) that angles $\Delta\alpha$ and $\Delta\beta$ on the one hand and $\Delta\gamma$ and $\Delta\delta$ on the other are influenced differently by substitution; the former might be connected merely with the σ-electron density, the latter merely with the π-electron density (resonance). However, we cannot confirm that the former effect should have something in common with the substituent electronegativity since we have not found any relation to σ_χ; instead we would prefer the term inductive

Table 3
Correlations of geometrical parameters in monosubstituted benzenes (1) and monosubstituted ethenes (2) (angles in °, lengths in Å).

Number	Response function	Explanatory variables	b†	R†	s†	f†
1	$\Delta\alpha$ exp.‡	$\Delta\alpha$ calc.	1.08 (7)	0.969	0.43	16
2	$\Delta\beta$ exp.‡	$\Delta\beta$ calc.	1.08 (6)	0.981	0.21	16
3	$\Delta\gamma$ exp.‡	$\Delta\gamma$ calc.	1.02 (7)	0.944	0.12	16
4	$\Delta\delta$ exp.‡	$\Delta\delta$ calc.	0.93 (14)	0.848	0.32	16
5	$\Delta\alpha$ exp.§	$\Delta\alpha$ calc.¶	1.05 (5)	0.992	0.25	6
6	$\Delta\alpha$ exp.‡	$\Delta\alpha$ calc.	1.36 (23)	0.889	0.94	9
7	$\Delta\alpha$ calc.	σ_I, σ_α	5.8 (7), −1.8 (6)	0.886	0.68	26
8	$\Delta\alpha$ calc.	σ_I, σ_χ	6.0 (8), 0.7 (6)	0.843	0.79	26
9	$\Delta\beta$ calc.	σ_I, σ_α	−3.6 (4), −1.1 (3)	0.911	0.36	26
10	$\Delta\gamma$ calc.	σ_I, σ_R σ_α	0.38 (14), −1.09 (10), −0.34 (12)	0.915	0.14	25
11	$\Delta\delta$ calc.	σ_I, σ_R σ_α	0.51 (20), 1.51 (14), 0.69 (17)	0.930	0.19	25
12	a calc.	σ_I, σ_α	−0.014 (4), −0.009 (3)	0.669	0.0042	26
13	b calc.	σ_R	−0.0033 (14)	0.437	0.0020	27
14	c calc.	σ_R	0.0035 (10)	0.575	0.0014	27
15	$(a + a')/2$	σ_I, σ_R, σ_α	−0.016 (2), −0.027 (4), −0.013 (2)	0.964	0.0021	8
16	$(b + b')/2$	σ_I, σ_R, σ_α	0.0036 (6), 0.0033 (10), 0.0013 (6)	0.933	0.0005	8
17	$(b - b')$ calc.	σ_R	−0.0135 (29)	0.852	0.0027	7
18	$\Delta\beta$ calc.	$\Delta\alpha$ calc.	−0.578 (22)	0.982	0.16	27
19	$(b - a)$ calc.	$\Delta\alpha$ calc.	0.0037 (6)	0.764	0.0045	27
20	$\Delta\zeta$ calc.	σ_I, σ_χ	−6.25 (19), 3.9 (14)	0.591	1.85	26
21	$\Delta\eta$ calc.	σ_I, σ_α	8.78 (10), 3.2 (8)	0.896	1.00	26
22	$\Delta\theta$ calc.	σ_I, σ_χ	2.9 (4), 1.0 (3)	0.873	0.39	26
23	l calc.	σ_I, σ_α	−0.009 (3), −0.007 (3)	0.630	0.0032	26
24	$\Delta\eta$	$\Delta\alpha$ calc.	1.42 (11)	0.930	0.81	27
25	l calc.	Δa	0.70 (5)	0.942	0.0014	27

† Regression coefficients b with their standard deviations in parentheses, correlation coefficient R, standard deviation from the regression s and degrees of freedom f, respectively. ‡ Determined statistically by Domenicano & Murray-Rust (1979). § Determined statistically by Norrestam & Schepper (1981). ¶ Calculated at the STO-3G level by Krygowski *et al.* (1986); substituent COOH was discarded since the sum of all angles given is not 720°.

effect (σ_I). The inclusion of the polarizability term is new and rather surprising. It is never decisive and is sometimes of little significance, but it appears in all correlations (Table 3, lines 7, 9, 10 and 11). The dependence of bond angles on substitution is rather complex and cannot be described in a simplified way. This complex picture is not too much influenced by the AGIBA effect, which is of second-order importance for the angles (see for instance Fig. 1).

The external angles $\Delta\varepsilon$ and $\Delta\varepsilon'$ are of interest only with unsymmetrical substituents when they differ significantly from each other. Their difference should show some relation to the substituent steric effect, but only to the steric effect in one direction. Hence, it is not surprising that we have not found any correlation with steric parameters.

3.3. Bond lengths

Any significant correlation was not feasible with the crystallographic data (probably owing to their low precision) or with values calculated at the STO-3G level. Krygowski (1990) preferred interpreting the value $b - a$, in which the thermal-motion effects are partly compensated. We have not achieved any significant correlation with the bond lengths a, b or c (Table 3, lines 12–14) or with the difference $b - a$. Note that this failure is not connected with the AGIBA effect: practically the same results were obtained with the bonds a', b' and c' or with average values like $(a + a')/2$. We gave particular attention to a possible dependence on σ_R since simple representation of resonance by canonical formulae requires at least a significant shortening of the b bond. However, this shortening is expected for both acceptor and donor substituents, while the pertinent σ_R values are of opposite sign. We attempted either correlations with the absolute values $|\sigma_R|$ or separate correlations with acceptors and donors. Different behavior of donors and acceptors was observed many times [for instance, Exner & Böhm (2001)]. This can be explained by assuming that acceptors behave uniformly and possess negligible or zero resonance effects (Exner, 1966), while the effects of donors are graduated. In the case of bond lengths, we obtained significant correlations only with the actual donors, not including alkyl groups and using the average lengths $(a + a')/2$ and $(b + b')/2$ (Table 3, lines 15 and 16). The correlations seem good, but three terms are necessary (σ_I, σ_R and σ_α), and with the small number of data this result is not quite convincing.

3.4. The AGIBA effect

The effects of unsymmetrical substituents are presented in Table 2 as differences between the opposite bond lengths or opposite bond angles. Five of these eight substituents could also be defined more generally, with a partially variable structure. For instance, the methoxy substituent was redefined as alkoxyl OR with any substitution in R, and the pertinent geometrical parameters were obtained as median values from many structures retrieved from the CSD. This approach was already made in the case of azo compounds (Krygowski, Anulewicz & Hiberty, 1996). Agreement of the two approaches (one isolated molecule with the simplest substituent and many complex molecules in different crystal environments) is better than expected for such small quantities (Table 2); it may be classified as good for some 18 quantities out of 25 (Table 2). Nevertheless, the CSD values showed a wide variance, including many less reliable data. We restricted the search to the data with the lowest R (< 0.05), but it transpired that R is not always correlated with precision and many untrustworthy results appeared in the selection. In our opinion, it is not feasible to analyze all reports individually as to their precision: the only way out is using a robust statistic as used in Table 2.

The data of Table 2 confirm without any exception the statement by Krygowski *et al.* (1996) as expressed by (2) and (3). We can conclude that the AGIBA effect is firmly proven both from calculations and from crystal structures, irrespective of small differences between the two approaches. From

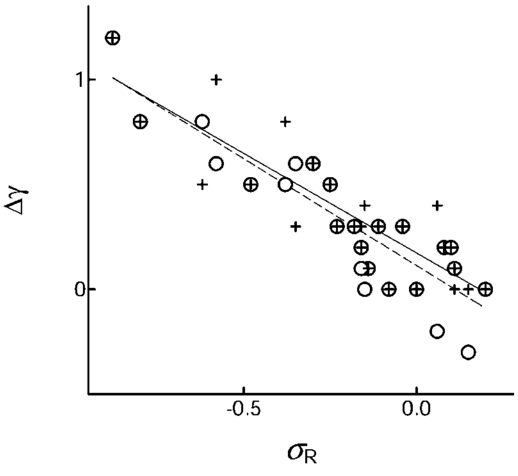


Figure 1
Calculated deformations of angles γ in benzene mono derivatives plotted *versus* the resonance constants σ_R: + position γ (full regression line), ○ position γ (broken regression line).

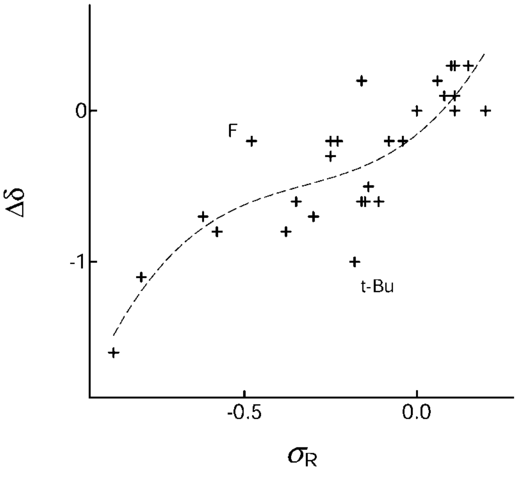


Figure 2
Calculated deformations of angles δ in benzene mono derivatives plotted *versus* the resonance constants σ_R; the curve is a computer LOWESS.

comparison of Table 2 with the mean values of Table 1 it is evident that AGIBA is more significant in the bond lengths, while in the bond angles it is only a second-order effect and can be neglected in a first approximation. Nevertheless, even the changes of bond angles are not negligible. Representation by the two Kekule structures in (2) and (3) must be regarded as a first approximation, since these structures do not require any angle alteration.

We have not found any significant correlation of the AGIBA parameters $(a - a')$, $(b - b')$ or $(c - c')$ with the reactivity parameters. In particular, possible relations with σ_R were tested in detail but the best one found was a barely significant correlation in Table 3, line 17.

3.5. Interrelations of geometrical parameters

The deformations of angles $\Delta\alpha$–$\Delta\delta$ and of bonds Δa–Δd are not independent when they relate to one hexagon. Any relation between them either can originate in a common physical principle or can have purely geometrical grounds. Britton (1980) derived three constraints valid for a common hexagon when the deformations are small; the most evident of them is the relation (valid for all hexagons):

$$\Delta\alpha + \Delta\beta + \Delta\beta' + \Delta\gamma + \Delta\gamma' + \Delta\delta = 0. \qquad (1)$$

When the hexagon has C_{2v} symmetry, the three constraints are simplified and their number is reduced to two. In these constraints, angle deformations of the order of 1° are related to bond deformations of approximately 0.02 Å. Changes of this order of magnitude are equivalent from the point of view of pure geometry but are not equivalent energetically: changes of bonds would require much greater energy (Krygowski, 1990). The bond deformations actually found are 0.002–0.005 Å (see Table 1, median values at the bottom). Hence, it is justified to describe the deformations in terms of two successive approximations: in the first, the bond lengths are constant and only the angles are deformed; in the second, even the bonds are stretched. In the first approximation and within the C_{2v} symmetry, the following constraints are valid:

$$\Delta\alpha - \Delta\delta = -\Delta\beta + \Delta\gamma, \qquad (2)$$

$$\Delta\alpha + \Delta\delta = -2\Delta\beta - 2\Delta\gamma. \qquad (3)$$

The total deformation can be described as a superposition of two limiting cases. When the hexagon is only stretched or compressed in the direction C(1)—C(4), both sides of equation (2) equal zero. When it is sheared, deformed at C(1) and C(4) in an opposite sense, both sides of equation (3) are zero. The actual relationship observed (Domenicano & Murray-Rust, 1979; Table 3, line 18) was only between $\Delta\alpha$ and $\Delta\beta$, approximately $\Delta\beta = -\Delta\alpha/2$. It can be described as a local deformation at C(1) that is only slightly transferred to adjoining atoms and would be in accordance with the interpretation by the Walsh rule (Krygowski, 1990) based on a change of hybridization on C(1). However, this interpretation has not been confirmed by a dependence on electronegativity constants σ_χ as described above.

The correlation between $\Delta\beta'$ and $\Delta\alpha$ (not given) is almost the same as between $\Delta\beta$ and $\Delta\alpha$. The deformations $\Delta\gamma$ and $\Delta\delta$ evidently have another cause: as mentioned, they can be connected with resonance (Domenicano & Murray-Rust, 1979), but only very roughly (Figs. 1 and 2). Further significant interrelations between bond angles or bond lengths were not revealed. Krygowski (1990) found a close relation of $(b - a)$ and $\Delta\alpha$ for ten substituents in *para*-homosubstituted benzenes. Our data gave a much worse correlation (Table 3, line 19), although we found a virtually equal regression coefficient.

Interrelations of geometrical parameters were also searched for by principal component analysis (PCA). The data matrix contained 12 columns ($\Delta\alpha$–$\Delta\delta$, Δa–Δc) and 29 rows (substituents). The correlation matrix revealed only the mentioned dependence of $\Delta\alpha$ and $\Delta\beta$ or $\Delta\beta'$, in addition to the expected relations of $\Delta\beta$ with $\Delta\beta'$, $\Delta\gamma$ with $\Delta\gamma'$ and Δa with $\Delta a'$. Four significant components were found explaining 43.8, 32.7, 11.0 and 8.3% of the variance, respectively (altogether 95.8%). The first component involves mainly $\Delta\alpha$, $\Delta\beta$, $\Delta\beta'$, Δa and $\Delta a'$ and, to a lesser extent, $\Delta\delta$; it can be connected with the hybridization on C(1). The second component involves $\Delta\gamma$, $\Delta\gamma'$, Δb and Δc. It represents the substituent effect transmitted through the ring; there might be some connection to resonance, but correlation with σ_R is poor. The third component effectively involves only $\Delta b'$ and $\Delta c'$ and can be understood as a correction for the AGIBA effect. The fourth component cannot be given any physical meaning. As expected, the components could not be correlated with any reactivity parameters. Our PCA thus differs from that of Domenicano *et al.* (1983) based on experimental data. These authors found only two components, rather similar to our first two.

The cluster analysis of substituents confirmed that the substituent effects are not of a common kind as known from the σ-constants. Any grouping of substituents into donors and acceptors was not observed. Substituents $N(CH_3)_2$ and F were found to have unique properties, and the pair of similar substituents OCH_3 and SCH_3 also form a separate cluster. Other clusters are difficult to understand, for instance the group CH_3, C_2H_5, $CH(CH_3)_2$, CH_2Cl, COOH, $COOCH_3$.

3.6. Ethene derivatives

The correlation of angles $\Delta\zeta$, $\Delta\eta$ and $\Delta\theta$ with reactivity parameters was as unsuccessful as that of $\Delta\alpha$–$\Delta\delta$. The best correlations from many attempts are given in Table 3 (lines 20–22). The dependence on σ_χ (lines 20 and 22) instead of on σ_α as in lines 7 and 9, is not relevant since the correlation is poor. Particular attention was given to the C═C bond length, but no relation to resonance was found, merely a certain dependence on the inductive effect (Table 3, line 23). The only regularity found was the essential similarity of the geometrical parameters in ethene and benzene derivatives (Table 3, lines 24 and 25). In our opinion, the geometry of ethene derivatives is still a worse model of substituent effects than in benzene derivatives since the substituent and the observed parameters are too close.

PCA with ethene and benzene derivatives together (data matrix 16 × 29) yielded results similar to those above with benzene derivatives only. Four components explained 91.9% of the variance; the angles $\Delta\eta$ and $\Delta\theta$ and the length Δl were grouped together with $\Delta\alpha$ and $\Delta\beta$ in the first component; $\Delta\zeta$ had no analogy.

4. Conclusions

In our opinion, the parameters of Table 1 can serve first as a reference to be compared with crystallographic data of derivatives of benzene and alkenes. The angles can be used even for estimating the geometry of benzene polyderivatives with a somewhat greater reliability than the original crystallographic parameters (Domenicano & Murray-Rust, 1979). Nevertheless, the essential agreement of both proves that statistical bond angles carefully selected from crystallographic data correspond well to the values from isolated molecules.

The geometrical parameters generally show only slight correlation to the known reactivity parameters; most of the contradictory claims in the literature are caused by imprecise values and small sets. Substituent effects on the geometry should, perhaps, be interpreted in terms other than the substituent-inductive effect, resonance or electronegativity.

References

Bock, C. W., Trachtman, M. & George, P. (1985). *J. Mol. Struct.* (*Theochem*), **122**, 155–172.

Britton, D. (1980). *Acta Cryst.* B**33**, 3727–3729.

Charton, M. (1981). *Prog. Phys. Org. Chem.* **13**, 119–251.

Charton, M. (1991). *Similarity Models in Organic Chemistry, Biochemistry and Related Fields*, edited by R. I. Zalewski, T. M. Krygowski & J. Shorter, pp. 629–687. Amsterdam: Elsevier.

Císařová, I., Podlaha, J., Böhm, S. & Exner, O. (2000). *Collect. Czech. Chem. Commun.* **65**, 216–226.

Domenicano, A. & Murray-Rust, P. (1979). *Tetrahedron Lett.* pp. 2283–2286.

Domenicano, A., Murray-Rust, P. & Vaciago, A. (1983). *Acta Cryst.* B**39**, 457–468.

Domenicano, A. & Vaciago, A. (1979). *Acta Cryst.* B**35**, 1382–1388.

Domenicano, A., Vaciago, A. & Coulson, C. A. (1975*a*). *Acta Cryst.* B**31**, 221–234.

Domenicano, A., Vaciago, A. & Coulson, C. A. (1975*b*). *Acta Cryst.* B**31**, 1630–1664.

Exner, O. (1966). *Collect. Czech. Chem. Commun.* **31**, 65–89.

Exner, O. & Böhm, S. (2001). *Collect. Czech. Chem. Commun.* **66**, 1623–1637.

Frisch, M. J. *et al.* (1995). *Gaussian*94. Revision C.3. Gaussian Inc., Pittsburgh, USA.

Howard, S. T., Krygowski, T. M., Ciesielski, A. & Wisiorowski, M. (1998). *Tetrahedron*, **54**, 3533–3548.

Krygowski, T. M. (1990). *Prog. Phys. Org. Chem.* **17**, 239–291.

Krygowski, T. M., Anulewicz, R. & Hiberty, P. C. (1996). *J. Org. Chem.* **61**, 8533–8535.

Krygowski, T. M., Cyrański, M. & Wisiorowski, M. (1996). *Pol. J. Chem.* **70**, 1351–1356.

Krygowski, T. M., Häfelinger, G. & Schüle, J. (1986). *Z. Naturforsch.* **41**b, 895–903.

Krygowski, T. M., Wisiorowski, M., Howard, S. T. & Stolarczyk, L. Z. (1997). *Tetrahedron*, **53**, 13027–13036.

Neugebauer, A. & Häfelinger, G. (2002). *J. Mol. Struct.* (*Theochem*), **578**, 229–247.

Norrestam, R. & Schepper, L. (1978). *Acta Chem. Scand.* A**33**, 889–890.

Norrestam, R. & Schepper, L. (1981). *Acta Chem. Scand.* A**35**, 91–103.

Taft, R. W. & Topsom, R. D. (1987). *Prog. Phys. Org. Chem.* **16**, 1–83.

Tinant, B., Declercq, J.-P. & Exner, O. (1993). *Collect. Czech. Chem. Commun.* **58**, 373–377.

Topsom, R. D. (1987). *Prog. Phys. Org. Chem.* **16**, 85–124.