Prediction of Molecular Geometries of Aromatic Six-membered Rings

R. NORRESTAM and L. SCHEPPER*

Structural Chemistry Group, Chemistry Department B, The Technical University of Denmark, DTH 301, DK-2800 Lyngby, Denmark

The structural effects of multiple substituents on bond angles and bond lengths in homo- and heterocyclic aromatic six-membered rings, have been analyzed for 14 different types of substituents. The data utilized, consists of the structural results from 48 published high quality crystal structure determinations. The parametrization chosen, three angular and two bond length deformation parameters per substituent type, and the obtained parameters are discussed in some detail. A possible method of estimating the angular parameters from calculated values of the σ and π electron donating properties of the substituents is described.

The accuracy obtained when using the present set of deformation parameters to predict molecular geometry, is comparable to that obtained by modern crystal structure analysis and applies to a large class of frequently used compounds as *e.g.* pollutants like halogenated phenols and biphenyls.

The amount of high quality information of molecular geometry of simple organic molecules has increased rather rapidly during the last few years, mainly due to the development of methods and techniques of crystal structure determination by diffraction methods. However, systematic interpretations of the irregularities observed within classes of related molecules have so far mostly been qualitative. Often, the observed bond angles and bond distances have only been compared to expected standard values e.g. bond angles of 120° and C-C distances of 1.39 Å in aromatic six-membered rings. The disagreement with standard values has sometimes been loosely discussed in terms such as

substitution effects, packing effects, overcrowding etc. without any deeper quantitative analysis. The present study has been undertaken in order to investigate the possibilities of a more rigorous and quantitative interpretation of the effects on aromatic six-membered rings.

Substitution of one or more of the hydrogen atoms attached to a benzene ring is known to give distortions from the 6/mmm symmetry of the ring.^{1,2} The substitution causes a redistribution of the electrons of the ring system and thus it affects both the intramolecular bond angles and bond distances, but in most cases the benzene rings remain planar.

In order to investigate such effects on the endocyclic bond angles Domenicano, Vaciago and Coulson 3,4 studied the angular deformations observed in mono- and para-disubstituted benzenes. If we let Δ_1 , Δ_2 , Δ_3 and Δ_4 denote the deviations from 120° of the endocyclic bond angles at the ipso, orto, meta and para positions from the substituent, (Fig. 1) the results obtained by Domenicano et al. indicated that Δ_1 was linearly dependent on the electronegativity. They also suggested that the value of Δ_2 was close to $-\Delta_1/2$ and that Δ_3 and Δ_4 were approximately 0°. Thus, the angular deformations induced by one substituent were expected to be unaffected by a second substituent in the para or meta position. Later Domenicano and Vaciago 5 noted that a more complex pattern of distortions might be obtained with substituents, such as NH₂, likely to interact strongly with the π -electron system of the benzene ring. These earlier assumptions (viz. $\Delta_2 = -\Delta_1/2$ and $\Delta_3 = \Delta_4 = 0$) were later found⁶ to be generally inadequate even for qualitative interpretations of the angular deformations ob-

^{*} Present address: Department of Reactor Technology, Risø National Laboratory, Risø, DK-4000 Roskilde, Denmark.

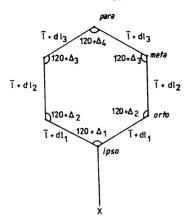


Fig. 1. Designations of the deformation parameters used in the present study.

served in accurate X-ray structural studies on substituted benzene rings, regardless whether the substituents were expected to interact strongly with the π -system of the rings or not.

In order to be able to analyze and predict endocyclic bond angle distortions both qualitatively and quantitatively Norrestam and Schepper developed a method,6 assuming unrestricted values of the angle distortions and additivity of the effects from several substituents on a ring. The preliminary studies 6 of the effects for 4 different substituent types were also 8,9 extended to 11 types and showed that reliable parameters could be derived, that were usable for predicting angular distortions also for multisubstituted benzene rings. In general, values significantly different from zero were obtained not only for the deviations Δ_1 and Δ_2 but also for Δ_3 and Δ_4 . Thus, the substituent effects were significant over the whole ring and not confined to the nearest neighbourhood of the substituted carbon atom. Recently, a closely related study of angular deformations in para disubstituted benzene derivatives has been performed by Domenicano and Murray-Rust. 10 Their results fully support the general pattern of the substituent effects described earlier.⁶ The differences encountered are probably due to their restriction to para disubstituted derivatives, which implies that a linear relationship between two of the deformation parameters (e.g. $\Delta_1 = a\Delta_2 + b$) has to be assumed.

The present paper describes extensions of the original method⁶ to include also heterocyclic sixmembered aromatics and a wider range of multi-

substituted derivatives, a comparison with results from molecular orbital calculations, and the formulation and preliminary applications of a related method for analyzing bond length deformation patterns. From the obtained parameters expectation values for bond angles and bond distances can easily be derived.

DATA COLLECTION

Substitution of the hydrogen atoms of an aromatic ring affects, as discussed above, both the bond angles and the bond distances. To be able to analyze the endocyclic geometrical effects of substitution in a completely general way, significant observations on both bond angles and bond length effects are needed. The observed distortions of the ring structure are, however, more significant in the endocyclic bond angles than in the bond distances. Typical values of the root mean square deviations (rms), as calculated from the data on neutral substituents selected for the present study, are 1.7° and 0.005 Å respectively. These data are selected as being the most accurate single crystal structural results available today, having estimated standard deviations (esd) of about 0.2° in the bond angles and of 0.003 Å in the bond distances. The distortions observed for the bond angles are thus statistically very significant, while the distortions of the bond distances (0.005 Å) are of about twice their esd's (0.003 Å). Thus, the quality of the single crystal structural data available today could not be expected to give very exact parameter values of bond distance effects for all types of substituents. More reliable values might, however, be derived for substituents like O⁻, which impose comparatively large and thus statistically significant bond length effects (rms of about 0.03 Å):

The data used are exclusively collected from high quality single crystal structure determinations by diffraction methods using X-ray or neutron radiation for which any possible systematic errors are comparatively small and most probably negligible. When available, structural results obtained at low temperature (usually about 110 K) have been preferred to room temperature investigations. The literature has been searched for structural data on small substituted benzene derivatives published until 1979, by utilizing e.g. the information in the Database of the Cambridge Crystallographic Data Centre. Among the structural results collected,

Table 1. The 48 structural results utilized in the present study. For COOH and NO_2 substituents, the cases where the planes of the substituents are twisted out of the ring plane by more than 30° are denoted by an asterisk.

	Substituent type at carbon atom number								
1	2	3	4	5	6	Ref.			
NH ₂					NH_2	14			
-	NH_2			NH_2	-	15			
	OH	ОН		•		16			
	OH			ОН		17			
	OH			0	ОН	18			
CH ₃	OH	CH ₃	CH ₃		CH ₃	19			
CH ₃		OH	C113	ОН	C113	12			
	ОН	OII		OH		13			
CH ₃	On			On		20			
C ₆ H ₅						20			
NH ‡			3.TTT+						
NH [‡]		G 11	NH ₃	G 17		22			
C_6H_5		C_6H_5		C_6H_5		23			
NH ⁺					NH_2	24			
SO ₃		_	NH_2			25			
COOH		NH_3^+				26			
OH		NO_2				27			
COOH		_	NO_2			28			
COOH*	SO_3^-		-			29			
COOH	COOH*	COOH				30			
C_6H_5	Cl	00011		Cl		31, 32			
C_6H_5	Ci			O1	Cl	31, 32			
C_6H_5	Cl	C1	Cl		Ci	31, 32			
C H	Cl	Ci	Cl	C1		31, 32			
C ₆ H ₅				Ci	CH				
COOH*	CH₃	г.	CH ₃	-	CH₃	33			
C ₆ H ₅	F	F	***	F	F	34			
C ₆ H ₅	F	F	F	F	F	35			
COOH*	F	F	F	F	${f F}$	33			
COOH	ОН			SO_3^-		36			
NO ₂		NO_2		NO [*]	OH	37			
$SO_{\overline{3}}$	NO [*]		NO_2		NO [*]	38			
NO [*] ₂	Cl -	NO_2	_	NO ₂	-	39			
OH	C_6H_5	-		-	C_6H_5	40			
COOH	-03		F		-03	41			
0-	NO ₂ *		NO ₂		NO_2	42			
ŏ-	NO ₂		NO ₂		NO_2	43			
ŏ-	NO ₂		NO ₂		NO_2	44			
N N	NU 2		CH ₃		1402	45			
	NH ₂		CH_3	CII		43			
N	NH_2	NIT T		CH ₃		46			
N		NH_2				47			
N	NH_2					48			
N			NH_2			49			
N	NH_2	N				50			
N	NH_2		N			51			
N	NH_2			Cl		52			
N	-	N				53			
N		N		CH ₃		53			
N	Cl	N		3		53			
N	NH ₂	N				53			
14	14115	1.4				33			

those representing complete structure determinations including hydrogen atom positioning, which yielded esd's in the endocyclic bond angles of 0.3° or less and crystallographic R values below 0.05 were selected to ensure that the substituent effects observed were significant. The esd's of the endocyclic bond distances are 0.004 Å or less. By including a few structure determinations on hydroxy substituted toluenes (Refs. 12-13) performed at our laboratory within the present project, structural data for altogether 48 different rings (Refs. 12-53) were collected. These results comprise e.g. 286 observations of endocyclic bond angles, for 4 types of common substituents (cf. Table 1).

To avoid erroneous data as far as possible, all bond distances and angles with esd's were recalculated and compared with the published results. In a few cases where discrepancies occur which could not be interpreted as misprints, the structural results were omitted.

PARAMETRIZATION OF BOND ANGLE EFFECTS

For a monosubstituted benzene derivative the endocyclic angular distortions caused by the substituent can be described by five parameters $\Delta_1, \ldots, \Delta_5$, representing the deviations by five of the bond angles from 120°. The sixth deviation, Δ_6 , is given by the condition $\Delta_1 + \ldots + \Delta_6 = 0$ provided that the ring is planar. This general case would be of relevance only for unsymmetrical substituents yielding a complex unsymmetrical pattern of angular distortions.

In most cases the main distortions can be properly described by using only three parameters Δ_1 , Δ_2 and Δ_3 (the deviations at the ipso, orto and meta carbon atoms), since $\Delta_1 + 2\Delta_2 + 2\Delta_3 + \Delta_4 = 0$ for a planar ring if $\Delta_6 = \Delta_2$ and $\Delta_5 = \Delta_3$ (cf. Fig. 1). Thus, it is assumed that the D_{6h} symmetry of the benzene molecule is lowered to $C_{2\nu}$ upon mono-substitution, with the 2-fold axis through the ipso and para carbon atoms. A similar symmetry lowering is assumed when one of the ring carbon atoms is exchanged with another atom giving a heterocyclic compound. The $C_{2\nu}$ symmetry of the ring could be expected only when the substituent also obeys the symmetry restrictions. However, experimental data and the outcome of the present investigation show that the C_{2v} symmetry of the deformation effects is valid to a good approximation also for "unsymmetrical" substituents such as coplanar COOH and OH groups.

As a second assumption the deformations caused by several substituents are considered to be additive. In this way linear relations, involving three parameters per substituent type, can be formulated for each molecule. The equations obtained for a set of molecules can then be solved by conventional least squares techniques (cf. Ref. 6). The relations can be properly weighted by the inverse of the estimated variance of each bond angle observation.

The validity of the assumption of additivity is supported by a simple analysis of the results from quantum chemical calculations. Thus, the electron redistribution effects upon substitution are approximately additive as indicated in Table 2, where the π -orbital charges calculated by *ab initio* techniques for various hydroxy toluenes ⁵⁴ are compared

Table 2. π orbital charges calculated ⁵⁴ by ab initio methods for monohydroxytoluenes, compared with those estimated by properly adding the corresponding calculated charges for toluene and phenol.

π orbital charge at at	om number					
	1	2	3	4	5	6
2-Hydroxytoluene						
Calc.	-0.025	0.007	-0.067	0.010	-0.029	0.005
Est.	-0.040	0.007	0.062	0.012	-0.033	0.006
3-Hydroxytoluene						
Calc.	0.049	-0.072	0.029	-0.085	0.028	-0.053
Est.	0.052	-0.086	0.031	-0.080	0.030	-0.057
4-Hydroxytoluene						
Calc.	-0.009	0.005	0.048	0.013	-0.067	0.005
Est.	-0.011	0.006	-0.062	0.013	-0.062	0.006

with those estimated by properly adding the calculated charges for toluene and phenol.⁵⁵

PARAMETRIZATION OF BOND LENGTH EFFECTS

For a monosubstituted derivative, we assume that the average, \bar{l} , of the endocyclic bond length can be written as $l=l_0+dl_0$, where dl_0 is the deviation from the ideal bond length, l_0 , in benzene. Thus, it is assumed that substitution gives rise to an *overall* change (dl_0) added to all bond lengths in the ring.

Apart from this overall effect there will be individual changes of the bond lengths depending on the position of the substituent. By again assuming C_{2v} symmetry for the effects of monosubstituted derivatives, there will be three independent bond length changes viz. at the ipso-orto, orto-meta and meta-para bonds. Let us denote these bond lengths by l_1 , l_2 and l_3 , respectively, and their derivations from the average endocyclic bond length, \overline{l} , by dl_1 , dl_2 and dl_3 . Thus, l_1 , l_2 and l_3 can be expressed by $l_1 = l_0 + dl_0 + dl_1$, $l_2 = l_0 + dl_0 + dl_2$ and $l_3 = l_0 + dl_0 + dl_3$ respectively. (cf. Fig. 1).

The value of dl_0 can be obtained from the ring average, \overline{l} , provided that a reliable value of the standard bond length l_0 is known. Independently of the values of l_0 and dl_0 , the individual changes dl_1 , dl_2 and dl_3 are obtained from the ring average as $dl_1 = l_1 - \overline{l}$, $dl_2 = l_2 - \overline{l}$ and $dl_3 = l_3 - \overline{l}$. To specify the deformations in a monosubstituted derivative completely, three different parameter values, e.g. dl_0 , dl_1 and dl_2 , are needed per substituent type, since the fourth, dl_3 , is given by the symmetry condition $dl_3 = -dl_1 - dl_2$.

For a multisubstituted derivative we assume, as before, that the bond length effects from several substituents are additive. Thus, simple linear relations between the deformation parameters can again be formulated and solved by conventional variance weighted least squares techniques. One set of equations are formed for the *individual* bond length variations to determine the dl_1 and dl_2 values, and another set for the *ring average* bond lengths to determine the dl_0 values. As an estimate of the standard bond length value, l_0 , we can possibly use either the total average bond length or a reasonable standard value of say 1.39 Å.

Obviously, endocyclic bond lengths and bond angles cannot be completely independent within a six-membered ring. As shown in the Appendix the

following relation is valid to a good approximation in a ring with $C_{2\nu}$ symmetry:

$$2\Delta_1 + 3\Delta_2 + \Delta_3 + \sqrt{3} (dl_1 - dl_3)/\overline{l} = 0$$

where the angular deformation parameters Δ_1 , Δ_2 and Δ_3 are given in radians. If we first determine the angular parameters, we can utilize this information in such a way that only two of the bond length parameters dl_0 and e.g. dl_1 has to be determined from the bond lengths. This case implies that only one independent parameter per substituent type has to be determined to describe the individual bond length variations.

ON THE EFFECT OF THERMAL MOTION ON BOND DISTANCES AND ANGLES

A possible major source of systematic errors in the structural data used, is due to the effects of thermal motion of the atoms.⁵⁶ The atomic positions are usually derived as the centroid of the shape of the atom averaged overall motions. The thermal effects can than lead to systematic errors in the derived bond lengths of up to a few hundreds of an Ångström, which is of the same magnitude as the effects of substitution which we intend to analyze. However, one might expect that the correlated motions between bonded atoms in a rigid aromatic ring are of approximately similar type. If this is true, the thermal motions will mainly affect the average bond length (l) of the ring, while the individual length deviations $(dl_1, dl_2 \text{ and } dl_3)$ from \bar{l} would remain largely unaffected. The thermal

insignificant. In few structural studies have the geometries of the molecules been corrected for thermal motion effects by utilizing the TLS-model ⁵⁶ for analyzing thermal motions. The bond length data obtained from such calculations ^{49,53} on eight molecules are given in Table 3. The results obtained support the assumption discussed above. Thus, the average bond lengths, \bar{l} , are indicated to be shortened by up to 0.013 Å due to thermal motion, while the individual bond length deviations, dl, are changes by less than 0.002 Å. The corrections of the bond angles are less than 0.1°.⁵³

effects on the bond angles would also be rather

Due to the parametrization of the substitution effects on the bond lengths, the use of structural data largely uncorrected for thermal motion effects will imply that rather unreliable values of the

Table 3. Differences between endocyclic bond length data with and without corrections for thermal effects. $\Delta \overline{l}$ denotes the differences between the average bond length (\overline{l}) in each ring, and Δdl the differences between the individual bond length deviations from \overline{l} . $\langle |\Delta dl| \rangle$ is the average of $|\Delta dl|$.

Compound	Temp. °C	$oldsymbol{\Delta}ar{ar{l}}$ Å	$\langle \Delta dl \rangle$	$\mathbf{Max}(\Delta \mathbf{d}l)$ Å
2-Aminopyridine	20	0.013	0.0012	0.002
3-Aminopyridine	20	0.011	0.0008	0.002
4-Aminopyridine	20	0.010	0.0008	0.002
2-Aminopyrazine	20	0.010	0.0008	0.002
Pyrimidine	107	0.003	0.0012	0.002
5-Methylpyrimidine	107	0.007	0.0005	0.001
2-Chloropyrimidine	107	0.003	0.0004	0.001
2-Aminopyrimidine	107	0.002	0.0000	0.000

overall changes, dl_0 , will be obtained. On the other hand the systematic errors in the observations of the individual bond length deviations, dl will probably be less than 0.002 Å. Thus, an analysis of substitution effects can be expected to give realistic results at least for those substituents that affect the dl-values by 0.01 Å or more. As an account for a possible systematic error in the observed dl-values, their e.s.d's have been increased by 0.002 Å throughout the present study.

In general, when comparing different bond length data for related rigid molecules or rigid molecular fragments, it is apparently safer to compare the relative bond length deviations from the proper averages of each fragment, than to compare the absolute values of the bond lengths.

BOND ANGLE PARAMETERS

The angular deformation parameters (Δ_1 , Δ_2 , Δ_3 and Δ_4) with e.s.d.'s, as obtained from the collected bond angle data (cf. Table 1), are listed in Table 4. During the initial calculations it became obvious that the deformations observed for molecules with COOH or NO₂ substituents could be classified into two different groups depending on the torsion angles between the plane through the substituent and the ring plane. One group with torsion angles below 30° (coplanar substituents) and the other group (denoted by an asterisk) consisting of the remaining NO₂ and COOH substituents which all had torsion angles considerably larger than 30° (roughly orthogonal positioned substituents). These two groups were treated as formally

Table 4. Angular deformation parameters in degrees. The value $d = dl_3 - dl_1$, in Ångström, is calculated from the expression given in the Appendix. Altogether 286 independent bond angle observations from 48 different molecules were used to determine the 42 independent angular parameters.

Substituent	Δ_1	Δ_2	Δ_3	Δ_4	d
NO ₂ *	4.59(28)	-3.07(15)	-0.02(15)	1.60(28)	-0.001(8)
NO ₂	2.92(18)	-1.85(11)	0.00(11)	0.79(19)	0.004(5)
F	2.49(24)	-1.80(17)	0.70(17)	-0.29(25)	0.004(7)
NH ⁺ ₃	1.79(22)	-1.16(16)	0.28(16)	-0.02(24)	0.005(7)
Cl	1.18(12)	-0.99(8)	0.60(8)	-0.40(12)	0.000(3)
SO ₃	0.70(14)	- 0.96(9)	0.48(9)	0.25(14)	-0.014(4)
OH	0.32(10)	-0.53(7)	0.60(7)	-0.46(11)	-0.005(3)
COOH*	-0.16(17)	-0.19(11)	0.12(11)	0.30(18)	-0.011(5)
COOH	-1.10(16)	-0.08(12)	0.94(13)	0.63(18)	-0.021(4)
NH ₂	-1.58(9)	0.38(6)	0.87(6)	-0.92(9)	-0.016(2)
CH ₃	-1.78(11)	0.95(6)	0.20(6)	-0.51(11)	-0.007(3)
Phenyl	-1.96(14)	0.87(8)	0.40(8)	-0.58(15)	-0.013(4)
N	-2.79(8)	3.49(5)	-1.47(5)	$-1.24(8)^{'}$	0.047(3)
O-	-5.27(28)	1.36(23)	2.02(24)	-1.47(30)	-0.062(6)

different substituents in the determination of the deformation parameters. No similar treatment appeared to be of relevance for any of the other types of substituents considered. In all, 42 independent angular deformation parameters (Δ_1 , Δ_2 and Δ_3) were determined from the 286 observations. As a check of the relevance of the results, a normal order statistic analysis was performed on the quantities $q_o = (\Delta_{obs} - \Delta_c)/(\sigma_{obs}^2 + \sigma_c^2)^{\frac{1}{2}}$. Here, Δ_{obs} are the observed angular deviations from 120° and $\sigma_{\rm obs}$ their observed e.s.d.'s, while $\Delta_{\rm c}$ are the deviations and σ_c their e.s.d.'s calculated from the obtained deformation parameters. The quantities $(\sigma_{obs}^2 +$ $\sigma_{\rm s}^2$ can be considered as the pooled e.s.d.'s. The full moment matrix was used in the calculations of $\sigma_{\rm c}$. A normal probability (n.p.) plot, ⁵⁷ e.g. a plot of q_o versus the expected q for a normal distribution with zero mean and unit variance, gave a correlation coefficient (r) of 0.997, a slope, s of 1.41, and an intercept t of -0.02. The correlation indicates ⁵⁸ that the deviations, q_0 , between observed and calculated bond angles are normally distributed. Due to the value, 1.41, of the slope the e.s.d.'s σ_{obs} and/or σ_c are slightly underestimated. The root mean square (rms) deviation $[1/N\Sigma(\Delta_{obs}-\Delta_c)^2]^{\frac{1}{2}}$ between observed and calculated angles was 0.38°.

As seen from Table 4, the two groups of the substituents NO2 and COOH respectively yield significantly different results e.g. the angular deformations caused by a coplanar COOH group (denoted COOH) are rather insignificant while those of a twisted COOH group (denoted COOH*) are highly significant. With the exceptions of the few cases where the deformation parameters are occasionally very small (below 0.30°), the analysis yields parameter values significantly different from 0°. Thus, the Δ_1 , Δ_2 , Δ_3 and Δ_4 values range from 0.53-4.6, -3.1 to 3.5, -1.5 to 2.0 and -1.6 to 0.8° , respectively. The $dl_3 - dl_1$ values, as derived by the expression given in the Appendix, yield as expected highly significant values only for a few of the substituents, viz. SO₃, COOH*, NH₂, hetero N and O⁻ (-0.014, -0.021, -0.016, 0.047) and -0.062 Å, respectively).

As indicated in the plot of the angular parameter Δ_2 against Δ_1 , shown in Fig. 2, there is an approximate linear correlation between these parameters for several of the substituents. However, the correlation is poor especially for O^- and hetero N, which both yield large deformation effects. By omitting these two substituents a correlation coefficient, r, of -0.988 for the least-squares line

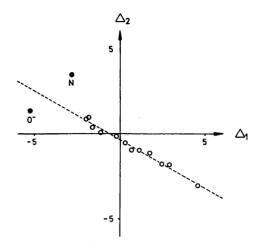


Fig. 2. Plot of the Δ_2 parameters versus the Δ_2 parameters. The dashed line indicates the least-squares line obtained when omitting the values (filled circles) for the O⁻ and htero N substituents.

 $(\Delta_2=0.58~\Delta_1-0.34)$ through the remaining twelve points is obtained (if all substituents are retained, r becomes -0.892 for $\Delta_2=-0.70~\Delta_1-0.29$). Thus, by omitting the O⁻ and N substituents, we can possibly assume a linear relation of the form $\Delta_2=a\Delta_1+b$ as a further approximation. If we utilize this approximation in the initial determination of the Δ 's, the number of independent variables will be reduced from 36 to 24 for the 192 observations of the 34 structures that are usable. In Table 5, the angular parameters are given that were obtained with such a linear relation used as a constraint in a least-squares fit of the parameters.

As discussed above Domenicano and Murray-Rust 10 have treated the angular deformations in para-substituted benzene derivatives, largely along the same lines as in our approach. However, their restriction to para-disubstituted derivatives implies that a linear relation must be assumed in order to to resolve the individual Δ 's. Thus, it should be expected that the parameters given in Table 5 agree with those obtained by Domenicano and Murray-Rust. Although the linear relations obtained, Δ_2 = $-0.60(2)\Delta_1 - 0.41(4)$ and $\Delta_2 = -0.54\Delta_1 - 0.30$ respectively, differ significantly, the correlation between e.g. the two sets of Δ_1 values are 0.983. However, the values obtained in the present study are generally somewhat smaller than that those found by Domenicano and Murray-Rust. Thus, the Δ_1 values are shifted by an average of 0.3°. An n.p. plot

Table 5. Angular deformation parameters in degrees. A linear dependence of Δ_2 on Δ_1 , viz. $\Delta_2 = a\Delta_1 + b$, was assumed. Altogether 192 observations from 34 different molecules were used to determine the 24 independent angular parameters and the 2 parameters a and b of the linear relation (a = -0.60(2), b = -0.41(4)).

Substituent	Δ_1	Δ_2	Δ_3	Δ_4
NO ₂ *	4.31(24)	-2.98(17)	0.06(13)	1.53(33)
NO ₂	2.66(16)	-2.00(12)	0.16(8)	1.01(22)
F	2.30(21)	-1.78(14)	0.70(13)	-0.14(28)
NH ⁺ ₃	1.59(19)	-1.36(12)	0.46(12)	0.21(27)
Cl	0.90(13)	-0.95(9) [°]	0.65(8)	-0.31(18)
SO ₃	0.58(16)	-0.76(10)	0.44(10)	0.05(23)
OH	0.21(9)	0.5 4 (7)	0.65(5)	-0.44(13)
COOH*	-0.34(14)	-0.21(9)	0.14(8)	0.47(18)
COOH	– 1.01(16)	0.19(11)	0.67(11)	-0.72(24)
NH ₂	-1.31(16)	0.37(11)	0.67(10)	-0.77(24)
CH ₃	- 2.29(11)	0.96(9)	0.39(6)	-0.39(18)
Phenyl	-2.12(12)	0.86(9)	0.47(7)	-0.52(18)

between the two sets of values of Δ_1 and Δ_3 yielded r=0.986 and s=1.68. The obtained intercept t=0.76 supports the conclusion that the two sets of values are systematically shifted with respect to each other. Possible sources for these systematic differences could be e.g. that in the present study data for any type of multi substituted derivatives were considered. Furthermore, the structural data were selected with somewhat more restricted criteria in the present study and the NO₂ and COOH substituents were also classified according to their twist angles to the rings.

An n.p. plot for the 192 bond angle deformations used in the present study gave r=0.988, s=1.149 and t=0.02. The r-value obtained in the non-restricted general case, including the large deformations occurring for O^- and heterocyclic N, was 0.997. Thus, we conclude that the general case (Table 4) without any linear constraints is a better model for describing angular deformations, especially since it is applicable also to heterocyclic nitrogen compounds (and to O^- substituted derivatives).

BOND DISTANCE PARAMETERS

The bond deformation parameters $(dl_0, dl_1, dl_2 \text{ and } dl_3)$ with e.s.d.'s, as obtained from the collected bond distance data (cf. Table 1), are listed in Table 6. For consistency with the analysis of the bond angles, the COOH and NO₂ substituents were divided into the same two groups as described

above. As discussed earlier, the value of the overall changes, dl_0 , can be determined independently of the individual bond length changes dl_1 , dl_2 and dl_3 . Only one independent individual bond parameter (e.g. dl_1) has to be determined for each substituent, as the $dl_3 - dl_1$ values can be derived from angular

Table 6. Individual $(dl_1, dl_2 \text{ and } dl_3)$ and overall (dl_0) bond deformation parameters in Ångström, multiplied by 10^3 . Altogether 264 observations were used to determine the 14 independent individual parameters. The dl_3-dl_1 values derived from angle deformations (cf. Table 4) were utilized to reduce the number of independent parameters from 28 to 14 (cf. also Table 7). The average bond length, 1.380 Å, was used as the standard bond length l_0 .

	dl_1	dl_2	dl_3	dl_0
NO [*]	-4(2)	10(2)	-5(2)	-4(2)
NO ₂	-4(1)	4(3)	0(1)	2(2)
F	-2(1)	0(2)	2(1)	1(1)
NH ⁺ ₃	-4(1)	2(2)	1(1)	3(2)
Cl	1(1)	-1(2)	1(1)	3(1)
SO ₃	6(1)	1(2)	-8(1)	10(2)
OH	3(1)	-1(1)	-2(1)	3(1)
COOH*	6(1)	-1(2)	-5(1)	3(2)
COOH	9(1)	3(2)	-12(1)	3(3)
NH ₂	12(1)	-7(1)	-4(1)	5 (1)
CH ₃	4 (1)	-2(1)	-3(1)	4(1)
Phenyl	8 (1)	-3(2)	-5(1)	1(2)
N	-31(1)	15(1)	16(1)	-15(1)
O-	47(1)	-32(3)	-15(1)	18(6)

deformation parameters and $dl_1 + dl_2 + dl_3 = 0$. The total average bond length calculated to 1.380 Å was used as the value for the standard bond length l_0 . The 14 overall parameters, dl_0 , were estimated from the 47 different ring average bond lengths, while the 14 independent dl_1 values were determined from altogether 270 different bond length observations.

An n.p. plot of the 47 overall bond length deviations gave r = 0.994, s = 0.79 and t = 0.03. The slope s indicates that the pooled e.s.d.'s are slightly underestimated, while the correlation r indicates normally distributed deviations. As a comparison an n.p. plot where no overall effects are assumed (i.e. all $dl_0 = 0$) gives r = 0.958, s = 3.82 and t = 0.04. In this latter case the correlation suggests that the deviations between the observed and calculated (all $dl_0 = 0$) values are probably significant ⁵⁸ and accordingly the hypothesis of normality of the deviations is probably contradicted. A more serious indication of an incorrect hypothesis is the value of the slope, which suggests that the pooled e.s.d.'s are underestimated by a factor as high as 3.84. Despite the fact that the introduction of the dl_0 parameters appears to be significantly justified, it should again be stressed that the use of data not properly corrected for e.g. thermal motion effects will give rather unreliable values of the overall parameters dlo compared to those obtainable for dl_1 , dl_2 and dl_3 .

An n.p. plot of the 270 individual bond length deviations gave r=0.981, s=1.39 and t=0.00. Thus, the e.s.d.'s are also in this case possibly somewhat underestimated. As a comparison, the n.p. plot, where no individual bond length deviations are assumed (all $dl_1 = dl_2 = dl_3 = 0$), gave r=0.963, s=9.01 and t=0.05. Again the slope is a very serious implication that this hypothesis is wrong, since it indicates that the pooled e.s.d.'s should be multiplied by a factor of 9.01, which would give unrealisticly high e.s.d.'s of 0.02 Å or more.

As a check of the consistency of the bond distance parameters, another calculation was performed in which no use was made of the dl_3-dl_1 values derived from the angular deformation parameters, *i.e.* the bond angle and bond distance effects were treated independently. Thus, the two independent individual parameters (*e.g.* dl_1 and dl_2) in this case were determined exclusively from the bond distance data. If parameters significantly different from those given in Table 6 were obtained, it would indicate that the derived bond distance parameters are

Table 7. Individual (dl_1 , dl_2 and dl_3) bond deformation parameters in Å, multiplied by 10^3 . Altogether 264 observations were used to determine the 28 independent parameters. No use was made of $dl_3 - dl_1$ values derived from bond angles.

	$\mathrm{d}l_1$	dl_2	dl_3	$dl_3 - dl_1$
NO [*]	-2(4)	10(4)	-7(4)	- 5(6)
NO ₂	-3(3)	4(3)	-1(3)	2(5)
F	-5(3)	0(2)	5(3)	10(4)
NH;	-4(2)	2(2)	2(3)	6(5)
Cl	-2(2)	-1(2)	3(2)	5(3)
SO ₃	5(2)	1(2)	-6(2)	-11(3)
OH	3(1)	-1(1)	-2(1)	-5(3)
COOH*	6(2)	-1(2)	-6(2)	-12(4)
COOH	7(2)	3(2)	-10(2)	-17(4)
NH ₂	11(1)	-7(1)	-4(1)	-15(2)
CH ₃	4(2)	-2(1)	-3(2)	-7(3)
Phenyl	7(2)	-3(2)	-4(2)	-11(4)
N	-31(1)	15(1)	16(1)	47(2)
O-	50(2)	-32(3)	-18(2)	- 68(4)

irrelevant and that the model for describing the bond distance effects is inadequate. The parameters obtained in this case are given in Table 7. An n.p. plot for a comparison of e.g. the dl_1 values from Tables 6 and 7 gives r=0.947, s=0.7 and t=0.2. With a sample size of 14, $r \ge 0.937$ indicates that the differences are not significant. 58 Thus, the consistency is good and the pooled e.s.d.'s are possibly somewhat overestimated. Perhaps a more obvious indication of the consistency would be a comparison (cf. Table 3) of the $dl_3 - dl_1$ values derived from bond angle data (Table 4) and from bond distance data (Table 7). As is already evident from Fig. 3, the agreement is quite good. An n.p. p,ot gives r=0.953, s=0.6 and t=0.2, and again indicates that the differences are not significant and that the pooled e.s.d's are somewhat overestimated.

As a result of those last analyses, we find that there are no serious contradictions to the model used for describing the bond length effects. Thus, the angular parameters derived from bond angle data are fully consistent with the bond length parameters derived from bond distance data. We might conclude that, by using the suggested model, it appears quite possible to quantitatively analyze and predict the effect of aromatic substitution on bond distances. Due to the limited quality of the structural data available today such an analysis does not give reliable estimates of the overall effects (dl_0) . However, significant parameters for the

individual effects (dl_1 , dl_2 and dl_3) have been obtained for most of the substituents listed in Table 6 e.g. NO₂*, SO₃*, COOH, COOH*, NH₂, phenyl, O⁻ and heterocyclic N. For the remaining substituents, the effects on the individual bond lengths seem to be very small (less than 0.005 Å).

ANALYSIS AND ESTIMATION OF PARAMETER VALUES

The study of angular deformations in mono- and para-disubstituted benzenes by Domenicano et al.^{3,4} indicated that the deformation, Δ_1 , that they observed at the ipso position was linearly dependent on the estimated group electronegativity of the substituent. As discussed earlier, they suggested that Δ_2 was $-\Delta_1/2$ and that $\Delta_3 = \Delta_4 = 0$. Thus, they could correlate all the bond angle effects to the electronegativity. In the present case, where no special conditions are applied to the different Δ -values, the Δ_1 and Δ_2 values are still correlated (r=0.88) to estimated 59 group electronegativities, but the correlations for the Δ_3 and Δ_4 parameters are not at all significant (r<0.65).

In the studies of substituent effects on properties as reactivities, IR-spectra, NMR-spectra, the correlations to semiempirical parameters within relations of the Hammet type have been used extensively. These parameters are assumed to reflect various effects on the electrons in the ring system, as inductive, exchange and charge-transfer effects. Some of these semiempirical parameters have been shown to be highly correlated to the calculated values of the σ and π electron charges in the ring. Accordingly, a very simple model which could possibly describe the observed Δ -values would be

$$\Delta_i = a_i q_\sigma + b_i q_\pi + c_i$$
 (i=1,2,3,4)

where q_{σ} and q_{π} are calculated values of the total number of σ and π electrons donated to the ring by

the substituent and where a_i , b_i and c_i are empirical constants to be determined. It should be noted that the sum $q_{\sigma} + q_{\pi}$ is closely related to the electronegativity of a substituent, and thus the possibilities for Δ_1 and Δ_2 to correlate with such a quantity is retained in the chosen expression. Values of q_{σ} and q_{π} from ab initio calculations have been published by Hehre et al.⁵⁵ for some substituted benzene derivatives. In all, q-values are accessible for 7 out of the 14 substituent types treated in the present study viz. NO₂, F, Cl, OH, COOH, NH₂ and CH₃. A least squares determination of the coefficients a, b and c yields the following expressions:

$$\Delta_1 = 21.6q_{\sigma} + 22.9q_{\pi} - 1.65$$

$$\Delta_2 = -12.7q_{\sigma} - 11.4q_{\pi} + 0.77$$

$$\Delta_3 = 0.0q_{\sigma} - 5.8q_{\pi} + 0.21$$

$$\Delta_4 = 3.5q_{\sigma} + 10.7q_{\pi} - 0.28$$

The correlations between observed and estimated Δ -values become 0.92, 0.93, 0.93 and 0.93, respectively, which is well above the critical value ⁵⁸ 0.899 for a sample size of 7, indicating that the differences are not significant.

According to the obtained expressions, both Δ_1 and Δ_2 appear to be approximately linearly dependent on the sum $q_{\sigma}+q_{\pi}$, as expected. The deformation, Δ_3 , occuring at the *meta* position is indicated to be dependent on changes of the number of π -electrons in the ring. The expression for Δ_4 is more complex, the effect at the *para* position appears to depend to 2/3 on q_{π} and 1/3 on q_{σ} . A possible explanation would be that the deformation occuring at the *para* position depends not only on induced changes to the electrons in the ring, but also on the planarity condition $\Delta_1 + 2\Delta_2 + 2\Delta_3 + \Delta_4 = 0$.

A check on the validity of the derived expressions and on their possible role as predictors of unknown Δ -values, can be obtained by comparing (cf. Table 8) estimated values for the substituents with the

Table 8. Comparison of some observed 10 Δ -values with those estimated according to the present study. The correlation r between the two sets is 0.966 and the rms deviation 0.22°.

Substituent	q_{σ}	q_{π}	Δ ₁ Obs.	Est.	Δ_2 Obs.	Est.	Δ ₃ Obs.	Est.	Δ ₄ Obs.	Est.
$CH_2 = CH -$	0.003	0.006	-1.8(2)	-1.5	0.8(1)	0.7	0.3(1)	0.2	-0.4(2)	-0.2
CH ₃ CO-	-0.007	0.028	-1.0(2)	- 1.2	0.4(1)	0.5	0.2(1)	0.1	-0.3(2)	0.0
CH ₃ O-	0.192	-0.105	0.2(2)	0.1	-0.6(1)	-0.5	1.1(1)	0.8	-1.1(2)	-0.7
NC-	0.104	0.022	1.1(2)	1.1	-0.8(1)	-0.8	0.3(1)	0.1	-0.1(2)	0.3

observations on these substituents made recently by Dominicano and Murray-Rust. ¹⁰ As seen from Table 8, the agreement is surprisingly good, keeping in mind the model used for estimating the Δ 's is most probably an oversimplification.

Due to the limited significance of the bond length effects for most of the seven substituents for which ab initio calculations where available, no similar analysis have been performed for the bond length effects.

As mentioned above, several other substituent effects, as different kinds of spectra, are most probably also due to changes in q_{σ} and q_{π} . Accordingly, it could be expected that such substituent effects can be correlated to the molecular geometries. An analysis of these types of relationship is beyond the scope of the present paper.

CONCLUDING REMARKS

The present study together with a previous one 6 have shown that substitution effects on endocyclic bond angles and bond lengths in multisubstituted homo- and heterocyclic aromatic six-membered rings, can be quantitatively predicted and analyzed by simple superposition models. As a further consequence of the type of analysis performed, the parameters given in Tables 4 and 6, define fairly accurately the geometries expected for monosubstituted derivatives with $C_{2\nu}$ symmetry.

The superposition model utilized, simply implies that the substitution effects in a multisubstituted derivative can be described by adding together the effects in the corresponding monosubstituted derivatives having the proper relative orientations. The accuracies obtained in the predictions of deformations of molecular geometries are expected to be

about 0.2° in bond angles and about 0.003 Å in bond lengths. This accuracy is comparable in magnitude to that obtained by single crystal X-ray diffraction techniques, as verified by recent applications ⁶¹ utilizing one of our preliminary sets ⁹ of parameters.

APPENDIX

With the designation shown in Fig. 3 the cosine of the angle $\alpha/2$ can be written $\cos(\alpha/2)=(a^2-ab\cos\beta+ac\cos(\beta+\gamma))/ad$ where the 2-fold axis of the C_{2v} symmetry coincides with the diagonal d. By using the approximation $\cos(2\pi/3+\Delta)\approx-1/2-\Delta\sqrt{3}/2$, valid for small Δ , and with the designations $\alpha=\Delta_1+2\pi/3$ ($\beta=\Delta_2+2\pi/3$ and $\gamma=\Delta_3+2\pi/3$) the expression can be written

$$\frac{d\sqrt{3}/4 \cdot \Delta_1 + (b+c)\sqrt{3}/2 \cdot \Delta_2 + c\sqrt{3}/2 \cdot \Delta_3 + (2a+b-c-d)/2 \approx 0}{(1)}$$

The diagonal d can be written as

$$d = ((b-c \cos \gamma - a \cos \beta)^2 + (c \cos \gamma - a \sin \beta)^2)^{\frac{1}{2}}$$

Using the same approximation as above we obtain

$$d\approx (a^2+b^2+c^2+ab+bc-ac+(ac+ab)\sqrt{3}\cdot\Delta_2+(ac+bc)\sqrt{3}\cdot\Delta_3)^{\frac{1}{2}}$$

By using the approximation $(1+\varepsilon)^{\frac{1}{2}} \approx 1+\varepsilon/2$, valid for small ε , and with the designations $a=\overline{l}+\mathrm{d}l_1$, $b=\overline{l}+\mathrm{d}l_2$ and $c=\overline{l}+\mathrm{d}l_3$, the expression for d can be written as

$$d \approx 2\bar{l} + dl_1/2 + dl_2 + dl_3/2 + \bar{l} \sqrt{3}/2 \cdot (dl_2 + dl_3)$$
 (2)

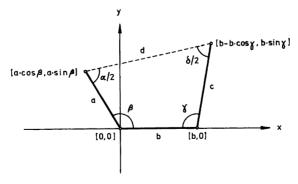


Fig. 3. Designations and the corresponding coordinate system used to derive a relation between Δ 's and d's.

Acta Chem. Scand. A 35 (1981) No. 2

where all higher order terms in $d\Gamma$ s and Δ 's are ignored. By combining (1) and (2), and again ignoring higher order terms, we obtain

$$2\Delta_1 + 3\Delta_2 + \Delta_3 + \sqrt{3}(dl_1 - dl_3)/\overline{l} \approx 0$$
 (3)

which expresses the dependence of the angular deformations (Δ) on the distance deformations (dl) and *vice versa*. An alternative derivation of the above type of relation has recently been published by Britton.⁶²

REFERENCES

- 1. Bent, H. A. Chem. Rev. 61 (1961) 275.
- Carter, O. L., McPhail, A. T. and Sim, G. A. J. Chem. Soc. A (1966) 822.
- 3. Domenicano, A., Vaciago, A. and Coulson, C. A. Acta Crystallogr. B 31 (1975) 221.
- Domenicano, A., Vaciago, A. and Coulson, C. A. Acta Crystallogr. B 31 (1975) 1630.
- Domenicano, A. and Vaciago, A. Third European Cryst. Meeting, Zürich (1976) 274.
- Norrestam, R. and Schepper, L. Acta Chem. Scand. A 32 (1978) 889.
- Norrestam, R. and Schepper, L. Acta Crystallogr. A 34 (1978) S111.
- Schepper, L. Geometriske Substituenteffekter i Benzenringe, Thesis, The Techn. Univ. of Denmark, Lyngby 1979.
- 9. Norrestam, R. and Schepper, L. Fifth European Cryst. Meeting, Copenhagen (1979) 122.
- Domenicano, A. and Murray-Rust, P. Tetrahedron Lett. 24 (1979) 2283.
- Kennard, O., Watson, D. G., Allen, F. H., Motherwell, W. D. S., Town, W. G. and Rodgers, J. Chem. Br. 11 (1975) 213.
- 12. Schepper, L. Acta Crystallogr. B 36 (1980). In press.
- Schepper, L. Acta Crystallogr. B 36 (1980). In press.
- 14. Stålhandske, C. Personal communication (1979).
- 15. Tsuchiga, H., Marumo, F. and Saito, Y. Acta Crystallogr. B 29 (1973) 659.
- Wunderlich, H. and Mootz, D. Acta Crystallogr. B 27 (1971) 1684.
- 17. Mahmoud, M. M. and Wallwork, S. C. Acta Crystallogr. B 31 (1975) 338.
- Bacon, G. E. and Jude, R. J. Z. Kristallogr. 138 (1973) 19.
- Bandour, J.-L. and Sanquer, M. Acta Crystallogr. B 30 (1974) 2371.
- Charbonneau, G. P. and Delugeard, Y. Acta Crystallogr. B 32 (1976) 1420.
- 21. Larsen, K. P. Acta Chem. Scand. A 28 (1974) 194.

- Domenicano, A., Seratoni, E. F. and Riva di Sanseverino, L. Acta Crystallogr. B 33 (1977) 1664.
- 23. Lin, Y. C. and Williams, D. E. Acta Crystallogr. B 31 (1975) 318.
- 24. Stålhandske, C. Acta Crystallogr. B 32 (1976) 2806.
- Bats, J. W. and Coppens, P. Acta Crystallogr. B 31 (1975) 1467.
- Arora, S. U., Sundaralingam, M., Dancz, J. S., Stanford, R. H. and Marsh, R. E. Acta Crystallogr. B 29 (1973) 1849.
- 27. Pandarese, F., Ungaretti, L. and Coda, A. Acta Crystallogr. B 31 (1975) 2671.
- Colapietro, M. and Domenicano, A. Acta Crystallogr. B 33 (1977) 2240.
- Attig, R. and Mootz, O. Acta Crystallogr. B 32 (1976) 435.
- 30. Mo, F. and Adman, E. Acta Crystallogr. B 31 (1975) 192.
- 31. Hovmöller, S., Norrestam, R. and Palm, T.-B. Acta Crystallogr. A 31 (1975) S116.
- 32. Palm, T.-B. To be published.
- 33. Benghiat, V. and Leiserowitz, L. J. Chem. Soc. Perkin Trans. 2 (1972) 1778.
- 34. Goodhand, N. and Hamor, T. A. Acta Crystallogr. B 34 (1978) 1644.
- Brock, C. P., Naae, D. G., Goodhand, N. and Hamor, T. A. Acta Crystallogr. B 34 (1978) 3691.
- Mootz, D. and Fayos, J. Acta Crystallogr. B 26 (1970) 2046.
- Herbstein, F. H. and Kaftory, M. Acta Crystallogr. B 32 (1976) 387.
- Lundgren, J.-O. Acta Crystallogr. B 28 (1972) 1684.
- Willis, J. S., Stewart, J. M., Ammon, H. L. and Preston, H. S. Acta Crystallogr. B 27 (1971) 786.
- 40. Nakatsu, K., Yoshioka, H. and Kunimoto, K. Acta Crystallogr. B 34 (1978) 2357.
- Colapietro, M., Domenicano, A. and Ceccarini,
 G. P. Acta Crystallogr. B 35 (1979) 890.
- Freeman, G. R. and Bugg, C. E. Acta Crystallogr. B 30 (1970) 431.
- 43. Bugg, C. E. and Thewalt, U. Acta Crystallogr. B 31 (1975) 121.
- 44. Palenik, G. J. Acta Crystallogr. B 28 (1972) 1633.
- 45. Kvick, Å. and Noordik, J. Acta Crystallogr. B 33 (1977) 2869.
- Nahringbauer, I. and Kvick, A. Acta Crystallogr. B 33 (1977) 2902.
- 47. Chao, M., Schempp, E. and Rosenstein, R. D. Acta Crystallogr. B 31 (1975) 2924.
- 48. Chao, M., Schempp, E. and Rosenstein, R. D. Acta Crystallogr. B 31 (1975) 2922.
- Chao, M. and Schempp, E. Acta Crystallogr. B 33 (1977) 1557.
- Scheinbeim, J. and Schempp, E. Acta Crystallogr. B 32 (1976) 607.

- 51. Chao, M., Schempp, E. and Rosenstein, R. D. Acta Crystallogr. B 32 (1976) 288.
- Kvick, A. and Backéus, M. Acta Crystallogr. B 30 (1974) 474.
- 53. Furberg, S., Grøgaard, J. and Smedsrud, B. Acta Chem. Scand. B 33 (1979) 715.
- Bursey, M. M., Greenberg, R. S. and Pedersen,
 L. G. Chem. Phys. Lett. 36 (1975) 470.
- Hehre, W. J., Radom, L. and Pople, J. A. J. Am. Chem. Soc. 94 (1972) 1496.
- Shoemaker, V. and Trueblood, K. N. Acta Crystallogr. B 24 (1968) 63.
- Abrahams, S. C. and Keve, E. T. Acta Crystallogr. A 27 (1971) 157.
- 58. Filliben, J. J. Technometrics 17 (1975) 111.
- Godfrey, M. J. Chem. Soc. Perkin Trans. 2 (1975) 1016.
- 60. Contributions in *Prog. Org. Chem.* Vols. 10-12. (1976-9).
- Fortier, S., Birnbaum, G. I., Buchanan, G. W. and Dawson, B. A. Can. J. Chem. 58 (1980) 191.
- 62. Britton, D. Acta Crystallogr. B 33 (1977) 3727.

Received August 19, 1980.