Some Relationships Between Bond Lengths and Angles in —COO⁻, —COOH and —COOCH₃ Groups

By Peter W. Borthwick

Department of Physics, The City University, St John Street, London EC1V 4PB, England

Abstract

Some systematic relationships between the bond lengths and angles of —COO⁻ and —COOH groups are proposed. From structural data for 70 such groups (occurring in a variety of crystallographic environments) the validity of the relations is demonstrated, statistical tests being applied where possible. The analysis is then extended (100 groups) to methoxy-carbonyl groups. It is concluded that although the groups considered undergo (often considerable) structural perturbation in their varied crystallographic environments these perturbations are not random but obey definite mathematical relationships.

1. Introduction

The isolated formate ion, $HCOO^-$, is expected to have C_{2v} group symmetry; *i.e.* $r_1 = r_2$ and $\varphi_1 = \varphi_2$.
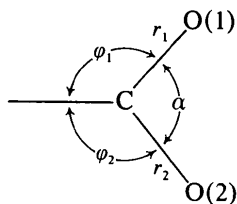
However, in a general situation a carboxyl group will be perturbed to only C_s symmetry by: (*a*) *intramolecular* perturbations, *e.g.* replacement of H by an asymmetric moiety or by protonation to —COOH; or (*b*) *intermolecular* perturbations, particularly in a crystalline environment, *e.g.* hydrogen bonding.

The possibility that such distortions of a carboxyl group may occur in a systematic fashion has been considered briefly by previous workers (Takusagawa, Hirotsu & Shimada, 1973; Dieterich, Paul & Curtin, 1974). In order to test this hypothesis in detail an initial study was undertaken with the results of 50 X-ray structure determinations of molecules associated with the γ-aminobutyric acid and L-glutamic acid neurotransmitter systems. Since such molecules contain —COO⁻ and —COOH groups in a wide variety of

crystalline environments they provide a suitable set of test data.

Analysis of the structural data (70 —COO⁻ or —COOH groups) showed the relationship

$$\varphi_1 - \varphi_2 = -100(r_1 - r_2) \quad (\varphi \text{ in degrees}, r \text{ in Å}) \quad (R1)$$

to be valid statistically with a correlation coefficient of 0·911 (Fig. 1 and Appendix). Random checks on carboxyl groups occurring in many other classes of molecules suggest that ($R1$) is a fundamental relationship which is true for —COO⁻ and —COOH groups in general.

In a study involving seven aromatic dicarboxylic acids, Takusagawa *et al.* (1973) have suggested the relation

$$\frac{1}{r_i} = k_1 \sin \varphi_j + k_2 \quad (i \neq j; i,j \in \{1,2\}). \quad (R2)$$

Our data set has been used to test the statistical significance of ($R2$); details of the analysis are shown in the Appendix. The fit of the data to ($R2$), although less good than ($R1$), is still significant with a correlation coefficient of 0·877.
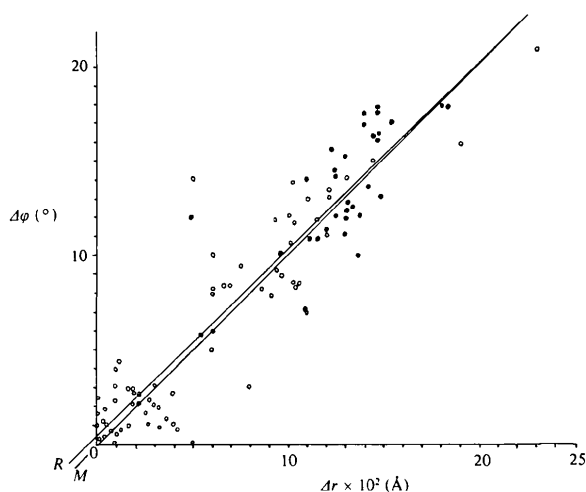

Fig. 1. A plot of difference in CCO angles ($\Delta\varphi$) against difference in CO bond lengths (Δr) for —COO⁻, —COOH (o) and —COOCH₃ (●) groups. The model (M) unit-slope line and regression (R) line are also shown.

Dieterich *et al.* (1974) briefly noted the relationship $|\Delta\varphi| = 100|\Delta r|$ in a discussion concerned with thermal disorder in 13 aromatic carboxylic acids: no sign dependence of the relation was noted.

2. Mathematical relationships

(R1) leads to an alternative to (R2) as a relation linking the individual r_i and φ_i. By writing the carboxyl group abstractly as
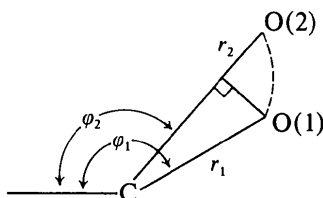
the relation to polar coordinates may be visualized:
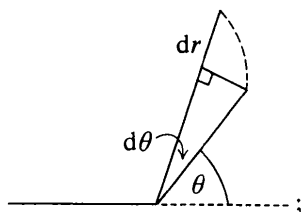
clearly $d\theta = -d\varphi$, thus from (R1), $d\theta = 100\, dr$ and $\theta = 100r + k_3$.

As a convenient method of estimating k_3 the mean values of r and φ (1·26 Å, 117·5°, respectively) from our data set [140 (r,φ) values] have been taken, giving

$$\varphi_i = 243 - 100r_i \quad (i \in \{1,2\}). \qquad (R3)$$

Since the data set contains ─∶O and an equal number of =O, ─O bonds the mean values quoted above should approximate to the value for a ─COO⁻ group. This is supported by *ab initio* LCAO SCF MO calculations (STO-3G basis) with *GAUSSIAN* 70 (Hehre, Latham, Ditchfield, Newton & Pople, 1971) in which a geometry optimization of the formate ion was carried out giving values of 1·25 Å, 117·2°, in good agreement with the experimental mean values. The form of (R3) (considered as a polar equation) is that of an Archimedean spiral and (R3) states that the O atoms of a ─COO⁻ or ─COOH group move on Archimedean spirals $r = \pm(0{\cdot}01\theta + 0{\cdot}63)$ referred to the C atom as the origin (r in Å, θ in degrees).

For groups which are well behaved with respect to (R1), (R2) and (R3) these relations are sufficient to generate the geometry of a carboxyl group from any one of its structural elements. Consider, for example, D,L-homocysteic acid (Clarke & Steward, 1977) ($r_1 = 1{\cdot}20$, $r_2 = 1{\cdot}31$ Å, $\varphi_1 = 124{\cdot}4$, $\varphi_2 = 112{\cdot}5°$); given r_1 it is possible to generate r_2, φ_1 and φ_2 as illustrated.
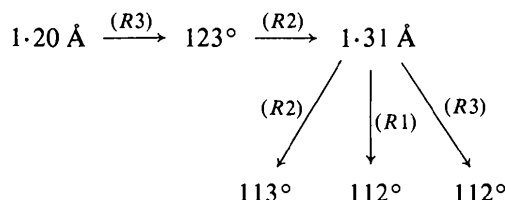
3. Error analysis

An inspection of points which lie off the unit-slope line indicates that they tend to be associated with structures having large R factors. This observation is linked to the fact that structure redeterminations (leading to improved accuracy) move points closer to the unit-slope line (Fig. 1). The validity of the model naturally requires that as measurements of the structural parameters of the ─COO⁻ and ─COOH groups improve in accuracy then the points so produced should approach closer to the unit-slope line. For example, the point (5,14) which lies well off the line corresponds to the structure of L-glutamic acid (β form) (Hirokawa, 1955); R for this work is 0·16. A subsequent redetermination (Lehmann, Koetzle & Hamilton, 1972) by neutron diffraction resulted in an R of 0·03 and a point (9,12) in much better agreement with (R1).

Another aspect of (R1) is illustrated by the point (8,3) which again lies some way from the unit-slope line and corresponds to the structure of DL-aspartic acid (Rao, Srinivasan & Valambal, 1968). One CO bond in this group ($r = 1{\cdot}22$ Å) has a corresponding φ of 111°, whereas (R3) predicts a value of 121°. We therefore recalculated the angle from the coordinates and found the value of 111° to be erroneous, the correct value being 119°. The new point for this group is now (8,5); R for this study is 0·18. A subsequent redetermination of DL-aspartic acid by Rao (1973) ($R = 0{\cdot}03$) further improved the position of the point to (9,8), in excellent agreement with (R1).

In overview it is suggested that (R1) [and (R3)] has practical value not only for assessing the relative inaccuracies in the structural parameters of ─COO⁻ and ─COOH groups [on the basis of $|(\Delta r - \Delta\varphi)|$],* but also in the identification of real errors occurring in published values.

4. A theoretical relation involving r_1, r_2 and α

Of theoretical interest is a relation which may be derived by combining (R1) and (R2). From (R2)

$$\frac{1}{r_1} = k_1 \sin \varphi_2 + k_2$$

* Henceforth, Δr, $\Delta\varphi$ are the positive differences of r and φ values respectively.

and

$$\frac{1}{r_2} = k_1 \sin \varphi_1 + k_2.$$

Therefore,

$$k_1(\sin \varphi_1 - \sin \varphi_2) = \frac{1}{r_2} - \frac{1}{r_1}$$

and

$$2k_1 \cos\left(\frac{\varphi_1 + \varphi_2}{2}\right) \sin\left(\frac{\varphi_1 - \varphi_2}{2}\right) = \frac{r_1 - r_2}{r_1 r_2}.$$

For the range of angles considered in this analysis it may be assumed that

$$\sin\left(\frac{\varphi_1 - \varphi_2}{2}\right) \sim \frac{\varphi_1 - \varphi_2}{2} \times \frac{180}{\pi}$$

which, from (R1),

$$\sim \frac{-100(r_1 - r_2)}{2} \times \frac{180}{\pi}.$$

Hence

$$r_1 r_2 \cos\left(\frac{\varphi_1 + \varphi_2}{2}\right) = \frac{\pi}{-100 \times k_1 \times 180}$$

if $\alpha = \widehat{OCO}$; then, since the carboxyl group is observed to be planar to a high degree,

$$\frac{\varphi_1 + \varphi_2}{2} = 180 - \frac{\alpha}{2};$$

hence

$$r_1 r_2 \cos\left(\frac{\alpha}{2}\right) = k_3.$$

Owing to the operations involved in establishing the above relation (e.g. division and differences of small quantities) the numerical aspects of the calculation are invalid. However, the analytic result that $r_1 r_2 \cos(\alpha/2)$ should equal a constant is valid and a separate *a posteriori* evaluation of that constant may be made. The mean values for r and φ (1·26 Å, 117·5°) lead to $k_3 = 0\cdot73$. Therefore, the relation

$$r_1 r_2 \cos\left(\frac{\alpha}{2}\right) \sim 0\cdot73 \qquad (R4)$$

is suggested, the constant k_3 (0·73) being a fundamental constant associated with –COO⁻ and –COOH groups.

5. Application to –COOCH₃ groups

It is evident that –COO⁻ groups are expected to lie in regions of the point population (Fig. 1) proximal to the origin, whereas –COOH groups will populate the more

distal regions. Another inhabitant of the more distal regions of the $(\Delta r, \Delta\varphi)$ plane might be the methoxycarbonyl group (–COOCH₃), the perturbing (steric) effect of –CH₃ being greater than that of a proton. From the Cambridge Crystallographic Data Centre file 20 structures with $R < 0\cdot10$ contributing 30 methoxycarbonyl groups have been added to the initial data set (total of 70 structures and 100 data points). Methoxycarbonyl groups lie on the unit-slope line* and it is noted that the new data set has (see Appendix for details) (a) a better correlation coefficient than the original, (b) a new regression line closer to the unit-slope line, and (c) a greater statistical significance.

It also seems that ethoxycarbonyl groups obey (R1) and conditions for the validity of (R1) to R–COOR′ in general are under study.

6. Typical geometries

Analysis of the data set enables the calculation of the mean structural parameters of typical –COO⁻, –COOH and –COOCH₃ groups.

(i) –COO⁻ group

(ii) –COOH group

(iii) –COOCH₃ group
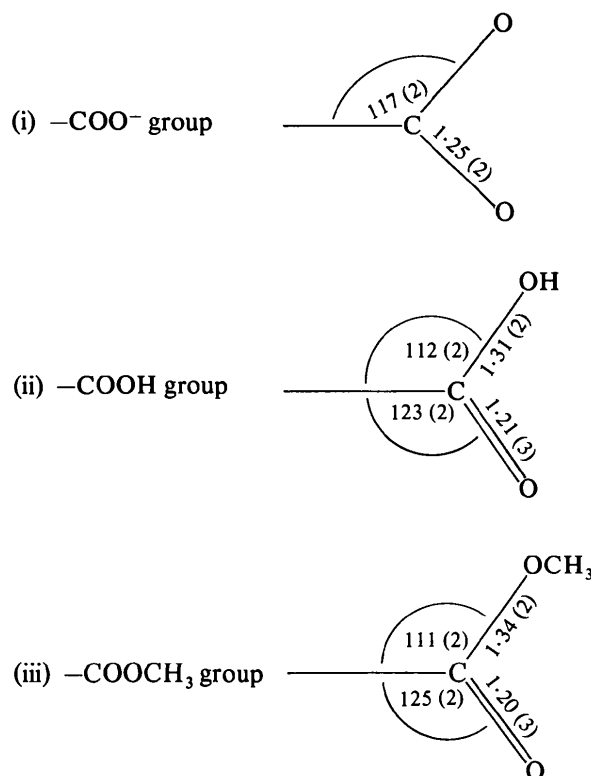
Note weakening of the CO single bond as H → CH₃.

---

* Attention is drawn to the methoxycarbonyl group in prostaglandin $F_{1\beta}$ ($R = 0\cdot11$) (Abrahamsson, 1963), the parameters for which are (1·10, 1·44 Å; 137, 104°) giving $(\Delta r, \Delta\varphi) = (0\cdot34$ Å. 33°)! This is the most extreme point known at present.

These typical groups can be used as models for (R1), (R3) and (R4). Plotted on a (Δr,Δφ) plane the groups appear as
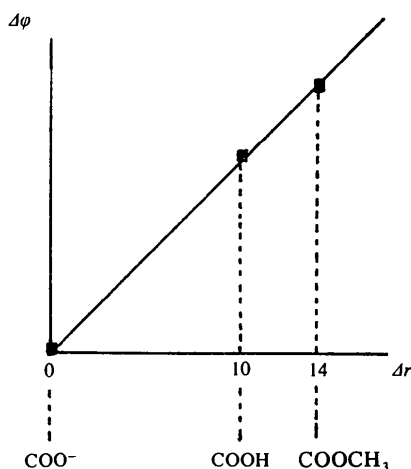
showing excellent agreement with (R1). The typical bond lengths from (R3) generate bond angles in good agreement with the typical values: 1·25 (118); 1·31 (112); 1·21 (122); 1·34 (109); 1·20 Å (123°). {The prediction is slightly less good for −COOCH₃; however, (R3) [and (R4)] were parameterized considering only averages for the data set of −COO⁻ and −COOH groups.} The values of k_3 in (R4) calculated for the typical −COO⁻, −COOH and −COOCH₃ groups are 0·73, 0·73 and 0·75, respectively. The latter values of k_3 confirm the figure adopted earlier.

7. Conclusion

It has long been realized that when one C and two O atoms combine in the following manner,

$$-C\begin{matrix} \nearrow O \\ \searrow O \end{matrix} \ominus ,$$
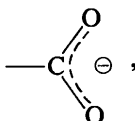
a particular set of chemical properties are imparted to the molecule containing this moiety. The invariant association of this structure and of its chemical effect allows us to think in terms of a carboxyl group. The term group is, however, more appropriate than may at first be realized since the −COO⁻ moiety is not a collection of atoms and bonds having relatively independent existences, but is in a very fundamental way structurally associated. Thus the structural variation of the group occurs in a well defined [by (R1)–(R4)] and correlated manner. The above comments have been shown to apply equally to −COOH and −COOCH₃; thus a set of closely related chemical groups exist, all obeying (R1)–(R4).

It is expected that correlated changes in bond lengths and angles will occur in other types of chemical groups and an analysis similar to that presented here may again be possible. The present study illustrates the possibilities (already demonstrated by Murray-Rust & Motherwell, 1978a,b) which now exist through the Cambridge Crystallographic Data Centre file to achieve an understanding of the conformations and configurations of molecules and molecular substructures in the crystalline phase.

APPENDIX

Relation (R1)

The regression equation is $\Delta\varphi = 0·94\Delta r + 0·45$, number of observations = 70, standard error = 0·05, correlation coefficient = 0·911, t value = 18·234.

The regression equation is close to the unit-slope line which is proposed in (R1). If the methoxycarbonyl groups are added, the new regression equation is $\Delta\varphi = 0·97\Delta r + 0·45$, number of observations = 100, standard error = 0·04, correlation coefficient = 0·923, and t value = 23·785.

This larger data set results in a regression line almost indistinguishable from the line $\Delta\varphi = \Delta r$ (Fig. 1), i.e. (R1).

Relation (R2)

The 70 groups comprising the initial data set give 140 (1/r, sin θ) values with which to test (R2). Strictly, the statistical tests and analyses applied here assume that the observations are independent. Hence, apart from use of the 140 values two sets of 70 have also been used, no one of which contains both (1/r, sin θ) pairs from the same −COO⁻ or −COOH group. The results of the three analyses are essentially the same, viz: the regression equation is $1/r = 0·65 \sin \varphi + 0·22$, number of observations = 140 (70, 70), standard error = 0·04, correlation coefficient = 0·877 (0·880, 0·864), and t value = 21·430 (15·231, 14·135).

It can be seen that correlation is good and (R2) is statistically significant, although (R1) remains the better relationship.

References

ABRAHAMSSON, S. (1963). *Acta Cryst.* **16**, 409–418.

CLARKE, G. R. & STEWARD, E. G. (1977). *J. Cryst. Mol. Struct.* **7**, 41–48.

DIETERICH, D. A., PAUL, I. C. & CURTIN, D. Y. (1974). *J. Am. Chem. Soc.* **96**, 72–80.

HEHRE, W. J., LATHAM, W. A., DITCHFIELD, R., NEWTON, M. D. & POPLE, J. A. (1971). *GAUSSIAN* 70 – Program 236, Quantum Chemistry Program Exchange, Indiana Univ., Indiana.

HIROKAWA, S. (1955). *Acta Cryst.* **8**, 637–641.

LEHMANN, M. S., KOETZLE, T. F. & HAMILTON, W. C. (1972). *J. Cryst. Mol. Struct.* **2**, 225–233.

MURRAY-RUST, P. & MOTHERWELL, S. (1978a). *Acta Cryst.* B**24**, 2518–2526.

MURRAY-RUST, P. & MOTHERWELL, S. (1978b). *Acta Cryst.* B**24**, 2534–2546.

RAO, S. T. (1973). *Acta Cryst.* B**29**, 1718–1720.

RAO, S. T., SRINIVASAN, R. & VALAMBAL, V. (1968). *Indian J. Pure Appl. Phys.* **6**, 523–526.

TAKUSAGAWA, F., HIROTSU, K. & SHIMADA, A. (1973). *Bull. Chem. Soc. Jpn*, **46**, 2020–2037.

Structural Aspects of Metacyclophanes.
Structures of 5,7,14,16-Tetramethoxy-1,2,3,10,11,12-hexathia[3.3]metacyclophane and 6,8,10,14,16,18-Hexamethyl-1,2,3,4,11,12-hexathia[4.2]metacyclophane

BY NEVENKA BRESCIANI-PAHOR, MARIO CALLIGARIS AND LUCIO RANDACCIO*

Istituto di Chimica, Università di Trieste, 34127 Trieste, Italy

Abstract

The crystal and molecular structures of the title compounds have been determined from three-dimensional X-ray data. The crystals of $C_{16}H_{16}O_4S_6$ (I) are triclinic, space group $P\bar{1}$, with $Z = 2$ and cell parameters $a = 13.86$ (1), $b = 10.86$ (1), $c = 8.155$ (7) Å, $\alpha = 93.83$ (8), $\beta = 96.12$ (8), $\gamma = 124.18$ (8)°, whereas those of $C_{18}H_{20}S_6$ (II) are orthorhombic, space group $Pca2_1$, with $Z = 8$ and cell parameters $a = 14.895$ (7), $b = 17.154$ (7), $c = 15.622$ (7) Å. The structures were refined to R values of 0.047 (I) and 0.033 (II), using 1940 and 3849 independent reflections respectively. Compound (I) is the [3.3] isomer in the *syn* form, whereas compound (II) is the [4.2] isomer in the *anti* form. No significant deviation from planarity of the bridged benzene rings was detected in either compound. The twelve-atom ring assumes a crown-like conformation, both in (I) and (II), suggesting a particular stability of this conformation which provokes the molecules to have either *syn* or *anti* forms depending upon the different lengths of the two polysulphur chains. Simple relationships of torsion angles around the S–S bonds with the chain length and S–S bond lengths are detected.

Introduction

Recently, the stereochemistries of metacyclophane systems have excited interest because of their peculiarities, such as deformations of the bridged rings, interconversion among different conformers, and the nature of the bridging chains. A great deal of work has been reported on systems having phenyl groups bridged by one or more aliphatic CH_2 chains (Vögtle & Newmann, 1972). Few structural data, however, have been reported for heterophanes (Bresciani-Pahor, Calligaris & Randaccio, 1978) or metacyclophanes having different types of bridges, such as polysulphide chains.

Thus, we have determined the structure of the title compounds to establish their geometries in the solid state. These compounds have been synthesized by Bottino, Foti & Pappalardo (1980), who suggested the hexathia[3.3] structure for (I). Preliminary results for compound (II) have already been reported (Bottino, Foti, Pappalardo & Bresciani-Pahor, 1979).

Experimental

Crystal data

(I) $C_{16}H_{16}O_4S_6$, $M_r = 464.7$, triclinic, $a = 13.86$ (1), $b = 10.86$ (1), $c = 8.155$ (7) Å, $\alpha = 93.83$ (8), $\beta = 96.12$ (8), $\gamma = 124.18$ (8)°, $U = 997.7$ Å³, $D_m = 1.56$