Kamphuis, I. G., Kalk, K. H., Swarte, M. B. A. & Drenth, J. (1984). *J. Mol. Biol.* **179**, 233-256.

Kirkman, H. N. & Gaetani, G. F. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 4343-4348.

Konnert, J. H. & Hendrickson, W. A. (1980). *Acta Cryst.* A**36**, 344-350.

Lee, B. & Richards, F. M. (1971). *J. Mol. Biol.* **55**, 379-400.

Murthy, M. R. N., Reid, T. J. III, Sicignano, A., Tanaka, N. & Rossmann, M. G. (1981). *J. Mol. Biol.* **152**, 465-499.

Newcomer, M. E., Jones, T. A., Åqvist, J., Sundelin, J., Eriksson, U., Rask, L. & Peterson, P. A. (1984). *EMBO J.* **3**, 1451-1454.

Rao, S. T. & Rossmann, M. G. (1973). *J. Mol. Biol.* **76**, 241-256.

Reid, T. J. III, Murthy, M. R. N., Sicignano, A., Tanaka, N., Musick, W. D. L. & Rossmann, M. G. (1981). *Proc. Natl Acad. Sci. USA*, **78**, 4767-4771.

Richardson, J. S., Getzoff, E. D. & Richardson, D. C. (1978). *Proc. Natl Acad. Sci. USA*, **75**, 2574-2578.

Rossmann, M. G. (1976). *Acta Cryst.* A**32**, 774-777.

Rossmann, M. G., Abad-Zapatero, C., Hermodson, M. A. & Erickson, J. W. (1983). *J. Mol. Biol.* **166**, 37-83.

Rossmann, M. G. & Argos, P. (1975). *J. Biol. Chem.* **250**, 7525-7532.

Schonbaum, G. R. & Chance, B. (1976). *The Enzymes*, Vol. XIII, edited by P. D. Boyer, 3rd ed., pp. 363-408. New York: Academic Press.

Schroeder, W. A., Shelton, J. R., Shelton, J. B., Apell, G., Evans, L., Bonaventura, J. & Fang, R. S. (1982). *Arch. Biochem. Biophys.* **214**, 422-424.

Schroeder, W. A., Shelton, J. R., Shelton, J. B., Robberson, B., Apell, G., Fang, R. S. & Bonaventura, J. (1982). *Arch. Biochem. Biophys.* **214**, 397-421.

Sheriff, S., Hendrickson, W. A., Stenkamp, R. E., Sieker, L. C. & Jensen, L. H. (1985). *Proc. Natl Acad. Sci. USA*, **82**, 1104-1107.

Shrake, A. & Rupley, J. A. (1973). *J. Mol. Biol.* **79**, 351-371.

Ten Eyck, L. F. (1973). *Acta Cryst.* A**29**, 183-191.

Ten Eyck, L. F. (1977). *Acta Cryst.* A**33**, 486-492.

Vainshtein, B. K., Melik-Adamyan, W. R., Barynin, V. V., Vagin, A. A. & Grebenko, A. I. (1981). *Nature* (*London*), **293**, 411-412.

Vainshtein, B. K., Melik-Adamyan, W. R., Barynin, V. V., Vagin, A. A., Grebenko, A. I., Borisov, V. V., Bartels, K. S., Fita, I. & Rossmann, M. G. (1986). *J. Mol. Biol.* **188**, 49-61.

Acta Cryst. (1986). B**42**, 515-522

A Systematic Pairwise Comparison of Geometric Parameters Obtained by X-ray and Neutron Diffraction

By Frank H. Allen

*Crystallographic Data Centre, University Chemical Laboratory, Lensfield Road, Cambridge CB*2 1*EW, England*

(*Received* 11 *November* 1985; *accepted* 24 *March* 1986)

Abstract

The effects of X-ray atomic asphericity on derived molecular geometry have been examined for ten common substructures involving C, N, O. The Cambridge Structural Database has been used to locate pairs of X-ray (X) and neutron (N) structure determinations of the same compound. Rigid criteria based on experimental temperature, refinement procedures, crystalline form and precision have been applied in selecting suitable X, N structure pairs. Corresponding X and N values for derived parameters have been paired for 46 chemically unique subsets (39 bond-length sets and 7 valence-angle sets). The statistical significance of the distribution of $X-N$ differences (D_i) has been assessed in each case *via* a Wilcoxon matched-pairs signed-ranks test. The signed magnitude of the mean difference ($\bar{D}$) is large, as expected, for bonds to H, ranging from −0·096 (7) Å for C–H to −0·155 (10) Å for O–H. Some highly significant differences are also observed for bonds involving C, N, O alone. Representative $\bar{D}$ values (Å) are −0·0096 (9) for C=C (ethylenic), −0·0052 (9) for C=C (benzenoid), −0·0078 (16) for C≡N, +0·0035 (12) for C–OH (carboxyl), +0·0046 (10) for N=O (nitro), +0·0054 (8) for C–O (carbohydrate) and +0·0082 (21) for C–C (C–C≡N); the significance of the D_i distribution for C=O (keto) was marginal. No significant differences were observed for C=O (COOH and COO$^-$), nor for any of the valence angles tested. Some implications for the combination or comparison of structural results obtained from different techniques are noted.

Introduction

It is well known that the spherically symmetric atomic charge density distribution around H is subject to a relatively large distortion on formation of a bond A–H (Stewart, Davidson & Simpson, 1965). The H position obtained from X-ray data (X) with normal spherical scattering factors is displaced towards A along the bond vector when compared with H positions obtained from neutron data (N). Thus the mean $X-N$ shifts for C–H and O–H in sucrose are −0·13 (1) and −0·18 (1) Å, respectively (Hanson, Sieker & Jensen, 1973).

Measurable asphericity shifts are not, however, restricted to H (see for example Dawson, 1964). A significant $X-N$ positional shift of O(H) in oxalic

0108-7681/86/050515-08$01.50

acid dihydrate was noted by Coppens, Sabine, Delaplane & Ibers (1969); the C=C bond in tetracyanoethylene is foreshortened by some 0·01 Å in the *X* study (Little, Pautler & Coppens, 1971) by comparison with *N* (Becker, Coppens & Ross, 1973); *X*-foreshortening of phenyl-ring bond lengths are discussed by O'Connell, Rae & Maslen (1966). Coppens (1974) used *X–N* comparisons of structures available at that time to examine the magnitude and direction of *positional* shifts for first-row elements (C, N, O) in asymmetric bonding environments. Despite the fact that asphericity effects tend to be absorbed to some extent in (higher) anisotropic X-ray thermal parameters (Coppens, 1968), some highly significant *X–N* positional shifts were reported by Coppens (1974). Their absolute magnitude ranged from 0·005 (1) Å for carbonyl O in cyanuric acid (Coppens & Vos, 1971) to 0·018 (6) Å for pyramidal N in hexamethylenetetramine (Duckworth, Willis & Pawley, 1970). The atomic shift direction was towards conceptual charge-density features such as lone pairs or multiple bond density maxima.

The gradual increase in the use of combined X-ray and neutron studies of the same compound, to examine deformation density distributions or hydrogen-bonding patterns for example, has provided more quantitative evidence of asphericity shifts for non-H atoms. This evidence is frequently cited in terms of *X–N bond-length* differences, irrespective of whether the asphericity directions are coincident with a bond vector or not. Thus Poppleton, Jeffrey & Williams (1975) reported a mean *X–N* difference for 32 C–O bonds in four carbohydrates of +0·006 Å, equivalent to $\sim 2{\cdot}5\sigma$(C–O) in these structures. An identical result (+0·0058 Å) was observed for the C–O bonds in β-D,L-arabinopyranose (Takagi, Nordenson & Jeffrey, 1979) which represented 3σ(C–O) in this case.

These results indicate that asphericity effects must be taken into account in any comparison of X-ray bond lengths with internuclear separations obtained by other experimental or theoretical methods. Furthermore they suggest that, in certain cases, it may not be appropriate to average *X* and *N* results as part of the same population in the derivation of precise mean geometry for use in the study of molecular systematics. This paper presents a statistical pairwise comparison of geometric parameters for organic compounds which have been studied by both *X* and *N* methods. The primary objective is to establish the magnitudes of *X–N* differences in bond lengths involving C, N, O, but the possible effect of asphericity on valence angles is also examined.

Methodology

Selection of data

The January 1985 release of the Cambridge Structural Database (CSD) (Allen *et al.*, 1979) has been used in this study. Whilst CSD contains structural data for 21 583 distinct organic compounds (chemical classes 1-69), only 448 of these have been studied by neutron diffraction. This is not altogether surprising considering crystal-size requirements and the relative paucity of neutron sources. Furthermore the chemical spread of *N* studies is decidedly nonrandom due to its application in H or D atom location (H-bonding studies, problems concerned with protonation, and hydrido-H location in inorganics and complexes) and in combined *X–N* studies of charge-density distributions of selected small molecules. The organic neutron literature is dominated by simple and amino acids, carbohydrates, aromatics and charge-transfer complexes.

For this study strict criteria were applied in the selection of suitable *X*, *N* pairs of structure determinations. (i) Both studies must have been performed at the same temperature to minimize, as far as possible, the differential effects on bond lengths arising from large *X–N* differences in atomic vibrational parameters. (ii) X-ray coordinates must have resulted from refinements based on a full $\sin\theta/\lambda$ range of reflection intensities, since asphericity effects can be minimized by refinement based only on higher-order data. (iii) Both *X* and *N* determinations must refer to the same crystalline modification to avoid any possibility of geometric differences induced by variations in H-bonding patterns. (iv) For each compound *X*, *N* pairs must both satisfy precision criteria set at $R_{\max} \le 0{\cdot}09$ and $\mathrm{SIG}_{\max} \le 0{\cdot}015$ Å, where SIG is the mean e.s.d. reported for bond lengths involving C, N, O.

Final selection of suitable *X*, *N* structural pairs, according to the above criteria, was carried out manually with the aid of bibliographic and other listings generated from CSD. Exhaustive reference to the original literature was necessary to check experimental conditions, computational procedures, and SIG values. Searches for the substructures of Table 1 were carried out using program *CONNSER* whilst geometrical parameters were obtained *via* program *GEOM*78 (Allen *et al.*, 1979) or from original papers. Precise *X*, *N* pairing of individual parameters (see below) was not always possible from *GEOM*78 output for small symmetric substructures. For example, the two equivalent C–O bonds of $\mathrm{COO^-}$ have a 50% chance of mapping onto *positionally equivalent* O atoms in both *X* and *N* studies. Frequent reference to original coordinate lists was necessary to ensure exact pairing for the statistical tests.

Statistical assessment of X–N differences

For a given chemical population, *e.g.*, C–O bonds in carbohydrates, we may calculate the signed difference $D_i = d(X)_i - d(N)_i$ for each matched pair. The mean difference $\bar{D}$ may then be obtained for n observations $i = 1-n$, whence $(\bar{D}) =$

$[\sum_n (\bar{D}-D_i)^2/n(n-1)]^{1/2}$. In this work the statistical significance of the distribution of D_i values has been assessed by use of the Wilcoxon matched-pairs signed-ranks test (Wilcoxon, 1947; Siegel, 1956). Since we make no assumptions about the data, *i.e.* we do not predict in advance the direction or magnitude of $\bar{D}$, the test yields the two-tailed probability (P) of the null hypothesis H_o: there are no differences between the X and N parameters. The Wilcoxon test has been applied *via* subroutine *SHT*03 of the *CAMAL* subroutine library (Taylor, 1986).

It should be noted that the mean values cited in this paper are unweighted values, since individual e.s.d.'s for atomic coordinates have only been stored in CSD for the 1985 and subsequent literature. However, Taylor & Kennard (1983*a*, 1985) have shown by statistical simulations that unweighted means can be used in hypothesis testing with little difficulty. This is especially true if the range of SIG values (see above) is restricted to eliminate structures of lower precision. Interestingly Taylor & Kennard (1985) conclude that the use of the weighted mean in hypothesis testing cannot be recommended.

Presentation of results

All results are given in Table 1, collected under substructural headings. The number of contributing X, N structural pairs is given in parentheses following each heading. A complete list of paired references for each substructural group has been deposited.* For each individual parameter the following details are tabulated: n is the number of matched X, N pairs of results; (+) is the number of positive $X-N$ differences, $d(X) > d(N)$; ($-$) is the number of negative differences, $d(X) < d(N)$; (=) is the number of equalities $d(X) = d(N)$; $\bar{d}(X)$ and $\bar{d}(N)$ are mean values of the X and N results, respectively, with $\sigma(\bar{d})$ in parentheses, distances are in Å and angles in degrees; $\bar{D}$ is the signed mean difference for n matched pairs with $\sigma(\bar{D})$ in parentheses; P is the two-tailed probability of the null hypothesis (H_o) from the Wilcoxon test, values of $P < 0{\cdot}010$ are regarded as significant, *i.e.* H_o can be rejected at the 99% significance level. Special notes relating to individual substructures or parameters are given as footnotes to Table 1.

Results and discussion

Carbon-carbon bonds

The foreshortening of C=C double and benzenoid bonds in X studies is well known since the earliest X, N comparisons. Suitable data for ethylene derivatives are still limited [Table 1(*a*)], but all ten X, N pairs show consistent and highly significant X-shortening of the C=C bond. This is readily interpreted (Becker, Coppens & Ross, 1973; Druck & Guth, 1982) as movement of each C (by ~0·005 Å) towards the bonding density feature of the double bond. In C–C=C substructures such a movement of the central C along the C=C vector might be expected to lengthen C–C(X) by up to 0·002 Å. This small difference is not detected here even though single bonds from N≡C–C=C have been omitted for reasons noted below.

Early X-ray results for benzene (Cox, Cruickshank & Smith, 1958) gave a thermally corrected C⩦C distance of 1·392 (10) Å, whilst Bacon, Curry & Wilson (1964) obtained a value of 1·398 (7) Å from neutron data. A theoretical study of the electron density distribution in benzene was presented by O'Connell, Rae & Maslen (1966). They showed that asymmetry in the distribution about aromatic C with an H substituent led to a shift in the X-ray C position away from its true location. The direction of shift is towards the ring centre and the magnitude was estimated as 0·0075 Å. $X-N$ discrepancies of this order of magnitude have been noted by several authors over the years (see for example Rietveld, Maslen & Clews, 1970; James & Moore, 1975). The results of Table 1(*b*) for isolated (unfused) rings are in excellent agreement with the work of O'Connell, Rae & Maslen (1966). The X-shortening of C⩦C [$-0{\cdot}0052$ (9) Å] is taken over all C atoms, whether H-substituted or not. The corresponding shortening of the 1,4 C···C diagonals is almost exactly twice the bond shortening, as it should be for inwards radial C movement towards the ring centre. When the analysis is restricted to 64 HC···CH bonds the $X-N$ difference increases to $-0{\cdot}0064$ (17) Å, close to the $-0{\cdot}0075$ Å estimated from the theoretical work. The remaining 80 ring bonds have either one (69) or two (11) non-H substituents. In these cases radial C movement is reduced and $\bar{D}$ is $-0{\cdot}0042$ (10) Å. Further subdivision of this data set [Table 1(*b*)] is not convincing, and results are almost certainly affected by the wide variation in the electronic nature of the substituents. There is, however, a coherent set of 18 exocyclic C(ar.)–C bonds with lengths in the range 1·47–1·51 Å. Here the radial movement of C(ring) is reflected in C(ar.)–C(X) lengthening with $\bar{D} = +0{\cdot}0033$ (12) Å with >99% probability for the D_i distribution. Asphericity effects on C–C single bonds cannot be generalized since they are dependent on environment (as here). Further examples will therefore appear at various points in the ensuing discussion.

In any analysis which is concerned with small differences in bond lengths the possible effects of thermal motion must be considered. The selection of isothermal X, N structure pairs goes some way

---

* A full list of bibliographic references to X, N structure pairs used in this study has been deposited with the British Library Lending Division as Supplementary Publication No. SUP 42880 (22 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

Table 1. *Results of Wilcoxon test for $X-N$ difference distributions for a variety of geometric parameters (see text)*

	(n)	$(+)$	$(-)$	$(=)$	$\bar{d}(X)$	$\bar{d}(N)$	$\bar{D}$	$P(H_o)$
(a) C–C=C (ethylene) ($N=8$)								
C=C	10	0	10	0	1·338 (5)	1·347 (5)	−0·0096 (9)	<0·010
C–C	11	4	7	0	1·448 (2)	1·489 (2)	−0·007 (12)	>0·050
(b) Phenyl rings ($N=20$)								
C$\overset{\cdots}{-}$C	144	38	98	8	1·387 (1)	1·392 (1)	−0·0052 (9)	<0·001
C$\cdots$C(diag.)$^{(1)}$	71	12	56	3	2·771 (3)	2·782 (2)	−0·0108 (14)	<0·001
C$\overset{\cdots}{-}$C(corr.)$^{(2)}$	18	4	12	2	1·391 (2)	1·396 (1)	−0·0048 (15)	<0·010
HC$\overset{\cdots}{-}$CH$^{(3)}$	64	16	44	4	1·383 (2)	1·389 (1)	−0·0064 (17)	<0·001
XC$\overset{\cdots}{-}$C$^{(4)}$	80	22	55	3	1·390 (1)	1·394 (1)	−0·0042 (10)	<0·001
XC$\overset{\cdots}{-}$CH$^{(5)}$	69	15	53	1	1·388 (1)	1·393 (1)	−0·0055 (10)	<0·001
XC$\overset{\cdots}{-}$C$X^{(6)}$	11	6	3	2	1·401 (2)	1·400 (3)	0·0008 (11)	>0·050
C$_{ar.}$–C	18	14	3	1	1·495 (4)	1·492 (4)	0·0033 (12)	<0·010
C$_{ar.}$–H	76	3	73	0	0·974 (7)	1·076 (2)	−0·102 (6)	<0·001
(c) Cyano group C$_a$–C≡N ($N=5$)								
C≡N	10	1	9	0	1·135 (2)	1·143 (2)	−0·0078 (16)	<0·010
C$_a$–C	10	8	1	1	1·451 (5)	1·443 (5)	0·0082 (21)	<0·010
C$_a\cdots$N	10	7	3	0	2·585 (5)	2·585 (2)	0·0005 (15)	>0·050
C$_a$–C≡N	10	8	2	0	177·9 (4)°	177·8 (4)°	0·12 (11)°	>0·050
(d) C–NH$_3^+$ in amino acids ($N=15$)								
C–N	17	10	6	1	1·486 (2)	1·485 (2)	0·0012 (10)	>0·050
N–H	51	5	46	0	0·930 (11)	1·032 (3)	−0·103 (11)	<0·001
(e) Carbohydrates ($N=19$)								
C–O (all)	127	100	23	4	1·424 (1)	1·418 (1)	0·0054 (8)	<0·001
C–O (ether)	48	39	7	2	1·425 (2)	1·420 (2)	0·0049 (9)	<0·001
C–O (hydroxy)	79	61	16	2	1·423 (2)	1·417 (1)	0·0057 (12)	<0·001
C–C (all)	85	21	60	4	1·522 (1)	1·526 (1)	−0·0034 (8)	<0·001
C–C (cyclic)	63	16	45	2	1·523 (1)	1·526 (1)	−0·0035 (10)	<0·001
C–C (acyclic)	22	5	15	2	1·521 (2)	1·524 (2)	−0·0032 (14)	<0·010
C–H	68	4	64	0	0·989 (7)	1·085 (3)	−0·096 (7)	<0·001
O–H	34	0	34	0	0·808 (11)	0·964 (3)	−0·155 (10)	<0·001
Angles (all)	210	90	119	1	110·2 (1)°	110·3 (1)°	−0·08 (3)°	>0·050
C–O–C (ring)	15	5	9	1	113·5 (4)°	113·6 (3)°	−0·09 (11)°	>0·050
(C, O)–C–C (ring)	75	47	28	0	110·3 (2)°	110·2 (1)°	0·05 (5)°	>0·050
(f) C–C–O–(C, H) noncarbohydrate ($N=9$)								
C–O	13	12	1	0	1·425 (6)	1·418 (6)	0·0071 (12)	<0·010
C–C	11	3	6	2	1·515 (5)	1·516 (6)	−0·0014 (13)	>0·050
($e+f$) C–O (all C–OH and C–O–C) ($N=28$)								
C–O	140	112	24	4	1·424 (1)	1·418 (1)	0·0056 (7)	<0·001
(g) C=O (keto) in (i) (C, H)$_2$–C=O and (ii) (C),(N)–C=O ($N=21$)								
C=O$^{(i)}$	8	7	1	0	1·231 (10)	1·228 (11)	0·0033 (13)	0·050
C=O$^{(ii)}$	21	13	7	1	1·239 (5)	1·237 (5)	0·0022 (15)	>0·050
C=O$^{(i+ii)}$	29	20	8	1	1·237 (4)	1·234 (4)	0·0025 (11)	0·050
C–N$^{(ii)}$	23	8	12	3	1·353 (5)	1·356 (4)	−0·0029 (20)	>0·050
C–C$^{(i+ii)}$	29	12	15	2	1·475 (8)	1·476 (8)	−0·0011 (15)	>0·050
(h) C–COOH ($N=21$)								
C–OH	23	14	8	1	1·308 (2)	1·305 (2)	0·0035 (12)	0·010
C=O	23	11	11	1	1·209 (3)	1·210 (2)	−0·0016 (11)	>0·050
C–C	23	9	12	2	1·506 (5)	1·507 (5)	−0·0016 (16)	>0·050
O–H$^{(7)}$	21	5	16	0	0·909 (26)	1·025 (10)	−0·115 (24)	<0·010
O–C–OH$^{(8)}$	22	12	10	0	124·1 (4)°	124·0 (3)°	0·09 (9)°	>0·050
(i) C–COO$^-$ ($N=18$)								
C$\overset{\cdots}{-}$O	40	19	14	7	1·252 (2)	1·251 (2)	0·0006 (7)	>0·050
C–C	19	10	7	2	1·533 (3)	1·534 (2)	−0·0010 (16)	>0·050
O$\overset{\cdots}{-}$C$\overset{\cdots}{-}$O	20	13	7	0	125·6 (2)°	125·6 (2)°	0·03 (10)°	>0·050
(j) Nitro groups ($N=5$)								
N$\overset{\cdots}{-}$O	16	14	2	0	1·225 (4)	1·221 (4)	0·0046 (10)	<0·010
O$\overset{\cdots}{-}$N$\overset{\cdots}{-}$O	8	4	4	0	124·0 (5)°	124·0 (5)°	−0·04 (10)	>0·050

Notes: (1) 1,4 ring diagonals; (2) thermally corrected X, N pairs; (3) bonds with H substituents only; (4) bonds carrying at least one non-H substituent (X); (5) bonds carrying exactly one non-H substituent; (6) bonds carrying two non-H substituents; (7) long O–H bonds in symmetrical O–H–O systems were omitted; (8) restricted to systems with torsion angle O–C–O–H ~ 0°.

towards obviating this problem but, even then, differential effects may still occur. Coppens (1968) has shown that asphericity effects (owing to lone pairs, bonding density and atomic ionicity) tend to be absorbed to some extent into the anisotropic displacement parameters (U_{ij}) obtained in X-ray refinements. He reported averaged increases of $U_{ii}(X)$ over $U_{ii}(N)$ of 11·5% for *s*-triazine and 9·3% for α-deuterooxalic acid, whilst a value of 15% was recently reported for *trans*-2-cycloocten-1-yl 3,5-dinitrobenzoate by Ermer & Mason (1982). The CSD files do not contain U_{ij}, hence appropriate corrections were not applied in this work. However, a subset of aromatic C$\overset{\cdots}{-}$C bonds were located for which corrected bond lengths for both X and N studies were available in the literature. Although the subset is small the $\bar{D}$ value is $-0{\cdot}0048\,(15)$ for a D_i distribution which is significant at the 99% confidence level, and is virtually identical to the value obtained from the full uncorrected data set. This result implies that $U_{ij}(X, N)$ differences, minimized by selection of isothermal structure pairs, should not materially affect the results presented in this survey.

Carbon-nitrogen bonds

The cyano group [Table 1(*c*)] represents the only triply bonded functional group for which suitable X, N pairings were located, and it is the extended grouping $C_1–C_2\equiv N$ which illustrates most clearly the balance between lone-pair and bonding density effects. Table 1(*c*) shows that the nonbonded $C_1 \cdots N$ distance remains essentially equal in both X and N studies, but there is X-shortening of $C_2\equiv N$ and X-lengthening of equal magnitude for $C_1–C_2$. Within a linear system this represents a movement of the central C_2 towards N by 0·008 Å in the X-ray studies. This C asphericity is towards the triple-bond density and, as might be expected, it is larger in magnitude than the 0·005 Å noted above for C=C double bonds. The terminal N, however, remains static and this can be explained (Coppens, 1974) by an almost exact balance between the N lone-pair density on the one hand and the triple-bond density on the other. These results form the basis for an explanation of an apparent discrepancy (Fig. 1) between the microwave (MW) r_s-geometry recently obtained for cyanocyclopropane (Harmony, Nandi, Tietz, Choe, Getty & Staley, 1983) and the averaged X-ray geometry for the cyanocyclopropyl substructure (Allen, 1980, 1981). Here again the MW and X values for $C_1 \cdots N$ agree closely (within the error limits cited for each method), but C_2 shifts towards N by 0·02 Å in an $X–MW$ comparison. If we take into account the error limits associated with each determination and, to some extent, the problematical physical definition of r_s (Robiette, 1973), then it would be difficult to postulate a real statistical difference between the pair of results in Fig. 1.

The only other C–N substructure for which a coherent set of isothermal X, N pairs could be located were the amino acids. Here the zwitterionic $C–NH_3^+$ form is totally predominant and relevant results are in Table 1(*d*). The lack of lone-pair density precludes N asphericity in this case and $\bar{D}$ is close to zero as expected. Other occurrences of C–N(N) involve a variety of delocalized and conjugated systems from which it was impossible to obtain unique X, N chemical populations.

Carbon-oxygen bonds

Each of the 19 X, N carbohydrate pairs [Table 1(*e*)] yields at least six C–O single bonds for comparison. The overall $\bar{D}$ for 127 C–O bonds [Table 1(*e*)] is $+0{\cdot}0054\,(8)$ Å with a probability >99·9% that this is a real effect. This $\bar{D}$ value is virtually identical to the +0·006 Å reported by Poppleton, Jeffrey & Williams (1975) for a 32-bond subset. Lengthening of C–O(X) is readily explained in terms of O asphericity due to lone-pair density. The bond-length discrepancy corresponds to a positional movement of O(X) of 0·011 Å in a direction 120° to the C–O bond vector; an O positional shift of 0·008 Å in this (lone-pair) direction is cited by Coppens (1974) for C–OH in sucrose. The present bond-length data [Table 1(*e*)] yield $\bar{D} = +0{\cdot}0057\,(12)$ Å for C–OH, a value which is slightly, but not significantly, longer than the C–O–C ether value of $+0{\cdot}0049\,(9)$ Å. The $\bar{D}$ values published for sucrose alone (Hanson, Sieker & Jensen, 1973) at $+0{\cdot}0051\,(8)$ (C–OH) and $+0{\cdot}0037\,(8)$ Å (C–O–C) are within 3σ of the overall values of Table 1(*e*).

A number of additional C–OH and C–O–C bonds in noncarbohydrate X, N pairs are compared in Table 1(*f*). The sample size (13) is small but P is significant at >99·9% and $\bar{D}$ at $+0{\cdot}0071\,(12)$ Å is within 2σ of the overall value for 127 bonds in Table 1(*e*). Considerably more comparative results are required to establish whether there is a real difference between carbohydrate and noncarbohydrate results or, more particularly, between hydroxy and ether situations. The overall $\bar{D}$ of $+0{\cdot}0056\,(7)$ Å for 140 C–O bonds of Table 1(*e*+*f*) represents the best value that can currently be obtained.

A somewhat unexpected feature of the carbohydrate data is the consistent and statistically significant contraction of the X-ray C–C bond lengths [Table 1(*e*)]. The $\bar{D}$ value is $-0{\cdot}0034\,(8)$ Å for 85 bonds, and the effect is equally apparent for both cyclic and acyclic situations. A mean $X–N$ difference of $-0{\cdot}0024\,(7)$ Å was listed for C–C in sucrose (Hanson, Sieker & Jensen, 1973), but was regarded (quite justifiably) as being not significant owing to the small number of bond comparisons. The present significant results are not amenable to a simple rationalization, although it should be remembered that carbohydrate

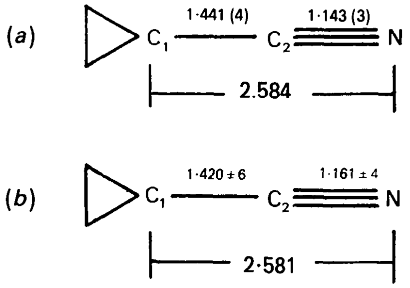

Fig. 1. Geometry of the C–C≡N group from (*a*) averaged X-ray values for cyanocyclopropyl fragment (Allen, 1980, 1981) located in CSD, and (*b*) microwave r_s values for parent cyanocyclopropane (Harmony, Nandi, Tietz, Choe, Getty & Staley, 1983).

C–C bonds are in a special environment, since a very high proportion carry at least two electronegative O substituents. This is borne out to some extent by Table 1(f), where C–C bonds associated with noncarbohydrate C–O bonds do not reveal significant X-shortening. Further careful study of experimental and theoretical deformation densities appears to be needed to establish the precise origins of this effect.

Results for other C–O bonds (of any type) are split between three substructural groupings (g, h, i) of Table 1. Data for noncarboxylic C=O bonds in Table 1(g) show X-lengthening of *ca* 0·003 Å, but the difference distribution is only significant at the 95% level. Early work on the noncarboxylic cyanuric acid (Coppens & Vos, 1971) showed $X-N$ positional shifts for O of 0·003 (1) and 0·006 (1) Å along the C=O vector for two N–C=O groupings. Even though O lone-pair hybridization is now sp^2 the mean atomic movement is very considerably less than the O positional shifts in C–O cited above. By analogy with C≡N it appears likely that asphericities of both C and O towards C=O bonding density cooperate to reduce $\bar{D}$, although movement of the central C in C–C=O and N–C=O has little effect on C–C or N–C bond lengths.

Within the carboxylate group [Table 1(h)] the X-lengthening of the formally single C–OH bond is reduced to +0·0035 (12) Å by comparison with the +0·0056 (7) Å obtained from the Table 1($e+f$) combination. Recent studies of low-temperature experimental and theoretical deformation densities for α-oxalic acid dihydrate (Stevens & Coppens, 1980; Stevens, 1980) show appreciable π density in the C–O bond, and also indicate the single sp^2-hybridized lone-pair on O. These factors suggest that the small reduction in $\bar{D}$ for C–OH (carboxyl) is in fact realistic, although it is of marginal statistical significance with current data sets. The carbonyl C=O bonds are shorter [1·210 (2) Å] in the carboxylates of Table 1(h) than for the noncarboxylates of Table 1(g) [1·235 (3) Å for 58 X, N bonds], a subset which is dominated by amido N–C=O systems. The near equality of X and N bond lengths for carboxylate C=O indicates negligible O asphericity. This may reflect increased (balancing) bond density and also slight differences in the disposition of O lone-pair density between carboxyl and amido C=O groups (Berkovitch-Yellin, Leiserowitz & Nader, 1977; Stevens, Rys & Coppens, 1978). This X, N equality is also shown by the delocalized C⩦O bonds of the carboxylate anion [Table 1(i)], when there are indications (Fuess, Bats, Dannöhl, Meyer & Schweig, 1982) that O lone-pair density is rather more diffuse than in neutral COOH groups.

Nitrogen–oxygen bonds

A qualitative picture of O asphericity in the nitro group was presented by Coppens & Coulson (1967). They predicted that X-displacement of O away from N owing to lone-pair density would considerably outweigh displacements in the opposite direction due to σ- and π-bonding density. Results for 16 N⩦O bonds [Table 1(j)] show a difference distribution significant at the 99% level and a mean $\bar{D}$ of +0·0046 (10) Å.

Valence angles

No systematic $X-N$ valence-angle discrepancies can be expected, of course, where X-displacement is along a bond vector, *e.g.* in C–C≡N; however, small effects are possible in situations where displacements are in lone-pair directions inclined to a bond vector. Thus an X-contraction of C–O–C valence angles in carbohydrates of up to 0·5°, and a slight X-expansion of O–C–OH (with torsion angle O–C–O–H ~ 0°), might be expected on this basis. Table 1 shows that no $X-N$ valence-angle difference distribution may be regarded as significant, and that difference magnitudes are all ~0·1°.

Bonds involving hydrogen

$X-N$ results for A–H bonds in relevant substructures are included in Table 1(b,d,e,h). Care was taken to ensure that X-ray H positions had been obtained by least-squares methods based on a full set of intensity data. Hydrogen bonding was not taken into account, except that long O–H distances in symmetrical O–H–O situations were excluded from the averaging. The individual $\bar{d}(X)$ and $\bar{d}(N)$ values should not be taken as definitive means since sample sizes are relatively small; however, trends in $\bar{d}(X, N)$ values are chemically sensible. The $X-N$ differences also vary with the electronegativity of A as expected, with $\bar{D}$(O–H) showing the largest magnitude of X-shortening [−0·155 (10) Å] and $\bar{D}$[C(sp^3)–H] the smallest [−0·096 (7) Å].

Summary and implications

Many of the results presented above have previously been noted in X, N comparison of individual structures, a context in which small differences in parameters have limited significance. This survey confirms that the effects of X-ray atomic asphericity on bond lengths are significant, not just for A–H bonds, but also for a variety of common bonding situations involving first-row elements. The statistical significance of $X-N$ bond-length difference distributions is established for $\bar{D}$ magnitudes in the range 0·003–0·010 Å, equivalent to several standard deviations in a modern crystal structure analysis. These differences result from $X-N$ positional shifts (Coppens, 1974) which may principally be ascribed to lone-pair or bonding density. In many cases the individual atomic asphericity represents a balance between these two factors (see *e.g.* Coppens &

Coulson, 1967), whilst the effect on bond lengths is often the resultant of two atomic asphericities; the results for the cyano group [Table 1(*c*)] are a clear example of these considerations. Furthermore it is not always possible either to predict or to rationalize these effects using simple concepts.

The purpose of this paper is to point out the dangers inherent in combining or comparing molecular dimensions obtained by different techniques. The apparent discrepancies depicted in Fig. 1, and discussed above, are a clear example of these problems. In recent years the systematic examination of structural results for series of related compounds has been used as a method for investigating chemical phenomena (see *e.g.* Allen, Kennard & Taylor, 1983; Bürgi & Dunitz, 1983). These studies have resulted from the ever-increasing volume of precise structural data in the literature (primarily X-ray results) and to improved availability *via* databases such as CSD. The averaging of molecular dimensions (Taylor & Kennard, 1983*a*, 1985), especially for common substructural fragments, is a common technique which is used to improve on the precision of results from individual structures. Often the object is to relate small differences in dimensions, either within or between substructures, to a variety of bonding effects (Allen, 1980, 1981; Allen, Kennard & Taylor, 1983). Since CSD is a primary source of information, the fact that it contains both X-ray *and* neutron results has some implications for very precise work. Firstly the uncritical use of CSD can, in special circumstances, introduce experimental bias into the resultant averages. Neutron studies constitute only 1·2% of CSD overall and $X-N$ discrepancies will typically be smoothed out during averaging. However, for reasons noted earlier, the chemical spread of neutron studies is far from random. For example, 15·2% of carboxylic acid salts with $R < 0{\cdot}07$ are neutron studies, and the corresponding figure for carbohydrates is 5·1%. Secondly, the results presented here show that asphericity effects must be taken into account when comparing X-ray bond lengths with results from other techniques. In their study of N—H···O=C hydrogen bonds, Taylor & Kennard (1983*b*) have discussed the importance of 'correcting' N–H(X) before X-ray and neutron results can be compared and combined. Table 1 shows that, even after due consideration of thermal motion, there are non-H asphericity corrections which should be applied before X-ray dimensions can be compared with spectroscopic or *ab initio* results. Finally, this work suggests that some of the geometrical parameters used in the construction of rigid groups for X-ray refinement should be chosen to reflect the most probable X-ray geometry. Thus the phenyl ring C=C distance commonly set at 1·390 or 1·395 Å should best be revised downwards to *ca* 1·387 Å as suggested by Table 1.

A referee has commented that there are three categories of results to be considered in a study of this type: (i) neutron results, (ii) results from refinement on complete X-ray data sets, and (iii) results based on high-angle X-ray data only. A comparison of (ii) and (iii) should mirror the results presented above, whilst a comparison of (i) and (iii) should satisfy the null hypothesis. At present, however, there is a lack of (i)-(iii) and (ii)-(iii) pairings which satisfy the selection criteria relating to temperature, crystalline modification, and precision (see *Methodology* above), and which generate suitably large geometric samples for application of the Wilcoxon test.

I thank Dr Robin Taylor for valuable discussion of the statistical procedures used in this paper.

References

Allen, F. H. (1980). *Acta Cryst.* **B36**, 81-96.

Allen, F. H. (1981). *Acta Cryst.* **B37**, 890-900.

Allen, F. H., Bellard, S., Brice, M. D., Cartwright, B. A., Doubleday, A., Higgs, H., Hummelink, T., Hummelink-Peters, B. G., Kennard, O., Motherwell, W. D. S., Rodgers, J. R. & Watson, D. G. (1979). *Acta Cryst.* **B35**, 2331-2339.

Allen, F. H., Kennard, O. & Taylor, R. (1983). *Acc. Chem. Res.* **16**, 146-155.

Bacon, G. E., Curry, N. A. & Wilson, S. A. (1964). *Proc. R. Soc. London Ser. A*, **279**, 98-110.

Becker, P., Coppens, P. & Ross, F. K. (1973). *J. Am. Chem. Soc.* **95**, 7604-7609.

Berkovitch-Yellin, Z., Leiserowitz, L. & Nader, F. (1977). *Acta Cryst.* **B33**, 3670-3677.

Bürgi, H.-B. & Dunitz, J. D. (1983). *Acc. Chem. Res.* **16**, 154-161.

Coppens, P. (1968). *Acta Cryst.* **B24**, 1272-1274.

Coppens, P. (1974). *Acta Cryst.* **B30**, 255-261.

Coppens, P. & Coulson, C. A. (1967). *Acta Cryst.* **23**, 718-720.

Coppens, P., Sabine, T. M., Delaplane, R. G. & Ibers, J. A. (1969). *Acta Cryst.* **B25**, 2451-2458.

Coppens, P. & Vos, A. (1971). *Acta Cryst.* **B27**, 146-158.

Cox, E. G., Cruickshank, D. W. J. & Smith, J. A. S. (1958). *Proc. R. Soc. London Ser. A*, **247**, 1-21.

Dawson, B. (1964). *Acta Cryst.* **17**, 990-996.

Druck, U. & Guth, H. (1982). *Z. Kristallogr.* **161**, 103-110.

Duckworth, J. A. K., Willis, B. T. M. & Pawley, G. S. (1970). *Acta Cryst.* **A26**, 263-271.

Ermer, O. & Mason, S. A. (1982). *Acta Cryst.* **B38**, 2200-2206.

Fuess, H., Bats, J. W., Dannöhl, H., Meyer, H. & Schweig, A. (1982). *Acta Cryst.* **B38**, 736-743.

Hanson, J. C., Sieker, L. C. & Jensen, L. H. (1973). *Acta Cryst.* **B29**, 797-808.

Harmony, M. D., Nandi, R. N., Tietz, J. V., Choe, J.-I., Getty, S. J. & Staley, S. W. (1983). *J. Am. Chem. Soc.* **105**, 3947-3951.

James, V. J. & Moore, F. H. (1975). *Acta Cryst.* **B31**, 1053-1058.

Little, R. G., Pautler, D. & Coppens, P. (1971). *Acta Cryst.* **B27**, 1493-1499.

O'Connell, A. M., Rae, A. I. M. & Maslen, E. N. (1966). *Acta Cryst.* **21**, 208-219.

Poppleton, B. J., Jeffrey, G. A. & Williams, G. J. B. (1975). *Acta Cryst.* **B31**, 2400-2404.

Rietveld, H. M., Maslen, E. N. & Clews, C. J. B. (1970). *Acta Cryst.* **B26**, 693-706.

ROBIETTE, A. G. (1973). *Molecular Structure by Diffraction Methods*, Vol. 1, edited by G. A. SIM & L. E. SUTTON, ch. 4. London: Chemical Society.
SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioural Sciences.* Tokyo: McGraw-Hill Kogakusha.
STEVENS, E. D. (1980). *Acta Cryst.* **B36**, 1876-1886.
STEVENS, E. D. & COPPENS, P. (1980). *Acta Cryst.* **B36**, 1864-1876.
STEVENS, E. D., RYS, J. & COPPENS, P. (1978). *J. Am. Chem. Soc.* **100**, 2324-2328.
STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175-3187.
TAKAGI, S., NORDENSON, S. & JEFFREY, G. A. (1979). *Acta Cryst.* **B35**, 991-995.
TAYLOR, R. (1986). *J. Appl. Cryst.* **19**, 90-91.
TAYLOR, R. & KENNARD, O. (1983*a*). *Acta Cryst.* **B39**, 517-525.
TAYLOR, R. & KENNARD, O. (1983*b*). *Acta Cryst.* **B39**, 133-138.
TAYLOR, R. & KENNARD, O. (1985). *Acta Cryst.* **A41**, 85-89.
WILCOXON, F. (1947). *Biometrics*, **3**, 119-122.

International Union of Crystallography

Acta Cryst. (1986). **B42**, 522

Commission on Journals
Author Grievance Procedure

The Commission on Journals has recently instituted a formal appeals procedure in which an author who believes his paper has been unjustifiably rejected by the Co-editor of an IUCr journal may appeal initially to the Editor of that journal for a new review and, finally, to the Editor of the other journal if the author is still aggrieved by the decision.

Acta Cryst. (1986). **B42**, 522

Commission on Journals
Equivalent Value of the Anisotropic Temperature Factor Coefficients

Anisotropic temperature factor coefficients have been published in *Acta Cryst.* since 1979 only if the table of values is very short, or they are necessary for understanding the paper, or they possess unusual features. In all other cases, the table of values has been deposited and a brief discussion of the deposited values given instead, including the maximum and minimum values found and the presence of any nonpositive-definite coefficients determined. In addition, the equivalent values of the anisotropic temperature factors have been published, together with the list of atomic coordinates and a definition of the equivalent values in terms of the individual coefficients with source reference, see *Notes for Authors* [*Acta Cryst.* (1983), **A39**, 174-186].

Authors have been encouraged to use their definition of choice. Among the acceptable definitions are those given by W. C. Hamilton [*Acta Cryst.* (1959), **12**, 609-610] and by B. T. M. Willis & A. W. Pryor [*Thermal Vibrations in Crystallography* (1975), pp. 101-102. Cambridge Univ. Press]. Arithmetic or geometric mean values for non-orthogonal crystal axes are correct only if derived from the principal axes of the thermal ellipsoid. Values of U_{eq} are to be preferred over B_{eq}.

Book Reviews

Works intended for notice in this column should be sent direct to the Book-Review Editor (J. H. Robertson, School of Chemistry, University of Leeds, Leeds LS2 9JT, England). As far as practicable books will be reviewed in a country different from that of publication.

Acta Cryst. (1986). **B42**, 522-524

Crystallographic computing 3: Data collection, structure determination, proteins, and databases. Edited by G. M. SHELDRICK, C. KRÜGER and R. GODDARD. Pp. ix+314. Clarendon Press, Oxford, 1985. Price £25.00.

Computing methods pervade all aspects of crystallographic research and it is essential that those working in this area be aware of all recent innovations in such methods. The need is tackled in this excellent volume that represents the proceedings of the Ninth International School on Crystallographic Computing under the auspices of the IUCr Computing Commission, that was held at the Max-Planck-Institut für Kohlenforschung, Mülheim an der Ruhr, Federal Republic of Germany, 30 July-8 August 1984, just before the International Union of Crystallography meeting in Hamburg, Germany. The international school was attended by 131 participants from 23 countries. Nine computers, ranging from a VAX 11/780 to a Rainbow personal computer, were available for use by participants at the practical sessions. There was also an Evans and Sutherland PS 300 color display unit on hand for use. Thus the stage was set for a state-of-the-art computing school, and the volume reviewed here reflects this great promise.

The range of subjects covered in this volume is divided into four sections. These consist of data collection and analysis of single crystal and powder samples, the use of data banks, program packages for maxi-, mini- and micro-computers and computing methods in protein crystallography.

The section on data collection and analysis starts with a well written article by Eric Gabe on random errors. When computing, it is essential, as Gabe points out, to avoid 'garbage in - garbage out'. It is excellent that, at the begin-