RIAND, J., CHENON, M.-T. & LUMBROSO-BADER, N. (1980). *Can. J. Chem.* **58**, 466–471.

ROSENFIELD, R. E., TRUEBLOOD, K. N. & DUNITZ, J. D. (1978). *Acta Cryst.* A**34**, 828–829.

RUSH, J. J. (1967). *J. Chem. Phys.* **47**, 3936–3943.

SCHLEMPER, E. O., HAMILTON, W. C. & LAPLACA, S. J. (1971). *J. Chem. Phys.* **54**, 3990–4000.

SCHOMAKER, V. & TRUEBLOOD, K. N. (1968). *Acta Cryst.* B**24**, 63–76.

SCHWEIZER, W. B., PROCTER, G., KAFTORY, M. & DUNITZ, J. D. (1978). *Helv. Chim. Acta*, **61**, 2783–2808.

SEILER, P. & DUNITZ, J. D. (1979). *Acta Cryst.* B**35**, 2020–2032.

SEILER, P. & DUNITZ, J. D. (1980). *Acta Cryst.* B**36**, 2946–2950.

SPEAKMAN, J. C. (1975). *Mol. Struct. Diffr. Methods*, **3**, 93.

SUNNERS, B., PIETTE, L. H. & SCHNEIDER, W. G. (1960). *Can. J. Chem.* **38**, 681–688.

THOMAS, L. P., TRUE, N. S. & BOHN, R. K. (1980). *J. Phys. Chem.* **84**, 1785–1789.

TRUEBLOOD, K. N. (1978). *Acta Cryst.* A**34**, 950–954.

TRUEBLOOD, K. N., GOLDISH, E. & DONOHUE, J. (1961). *Acta Cryst.* **14**, 1009–1017.

TRUEBLOOD, K. N., KNOBLER, C. B., LAWRENCE, D. S. & STEVENS, R. V. (1982). *J. Am. Chem. Soc.* **104**, 1355–1362.

WILLIS, B. T. M. & PRYOR, A. W. (1975). *Thermal Vibrations in Crystallography*. Cambridge Univ. Press.

YODER, C. H. & GARDNER, R. D. (1981). *J. Org. Chem.* **46**, 64–66.

Acta Cryst. (1983). B**39**, 133–138

Comparison of X-ray and Neutron Diffraction Results for the N—H···O=C Hydrogen Bond

BY ROBIN TAYLOR AND OLGA KENNARD

Crystallographic Data Centre, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, England

(*Received* 7 *September* 1982; *accepted* 28 *September* 1982)

Abstract

The errors associated with X-ray determinations of N—H···O=C hydrogen-bond geometries have been analysed. The analysis was based on a comparison of X-ray and neutron diffraction results for 57 hydrogen bonds which have been determined by both methods. The random errors in X-ray N—H and H···O distances were found to vary between 0·02–0·17 Å, the mean value being 0·065 Å. These errors exceed the e.s.d.'s quoted in the literature by a factor of approximately 1·6. The average systematic error in the X-ray N—H and H···O distances is about 0·1 Å. However, systematic errors can be corrected by 'normalizing' the N—H bond lengths to a standard value (1·030 Å).

Introduction

The N—H···O=C hydrogen bond plays a major role in determining the conformations of proteins and nucleic acids (Pimentel & McClellan, 1960). Consequently, its geometry has been investigated extensively by crystallographic techniques. Early X-ray diffraction studies of amino acids and nucleosides provided detailed information about the distribution of N···O distances (Ramakrishnan & Prasad, 1971). However, the values of the true hydrogen-bond lengths (*i.e.* H···O) and N—H···O angles were obscured by the large experimental uncertainties in the hydrogen-atom positions. Unfortunately, the N···O distance is an unreliable guide to the strength of a hydrogen bond; *e.g.* the six N—H···O bonds in Table 1 (determined by neutron diffraction) have almost identical N···O distances but H···O distances ranging from 1·844 (2) to 2·083 (6) Å.

Recent reviews of N—H···O hydrogen-bond geometries have therefore relied extensively on the results of neutron diffraction studies, which provide precise information about the H···O distances and N—H···O angles (*e.g.* Jeffrey & Maluszynska, 1982; Koetzle & Lehmann, 1976; Olovsson & Jönsson, 1976; see also Ceccarelli, Jeffrey & Taylor, 1981; Taylor & Kennard, 1982, for studies of other types of hydrogen bonds using neutron data). However, neutron studies con-

Table 1. *Selected hydrogen-bond geometries determined by neutron diffraction*

Structure	Hydrogen bond[a]	N···O (Å)	H···O (Å)	Reference number in Table 2
L-Lysine monohydrochloride dihydrate	N(2)–H(4)···O(2)	2·887 (3)	2·083 (6)	22
Hydrazinium hydrogen oxalate	N(2)–H(2)···O(1)	2·884 (1)	1·935 (2)	12
9-Methyladenine-1-methylthymine	N(10′)–H(4)···O(8)	2·872 (3)	1·932 (5)	23
L-Glutamic acid	N(1)–H(2)···O(11)	2·895 (1)	1·926 (2)	18
L-Tyrosine	N(1)–H(3)···O(2)	2·884 (2)	1·853 (4)	21
DL-Serine	N(1)–H(2)···O(2)	2·876 (1)	1·844 (2)	6

(*a*) Atom labels are those used in the Cambridge Structural Database.

0567-7408/83/010133-06$01.50

stitute only about 1% of the structure determinations in the Cambridge Structural Database (Allen *et al.*, 1979). Thus, although it is important to study the true hydrogen-bond length, H···O, rather than the N···O distance, it is also desirable to use the available X-ray data if possible. It is therefore of interest to evaluate the errors associated with X-ray determinations of hydrogen-bond geometries. We present here an analysis of these errors, based on a comparison of X-ray and neutron diffraction results for 57 N—H···O=C hydrogen bonds which have been determined by both methods.

Crystallographic data

All crystallographic data were retrieved from the Cambridge Structural Database. The analysis was based on 26 organic crystal structures (Table 2), each of which has been determined by X-ray and neutron diffraction at approximately the same temperature. Structure determinations with $R > 0{\cdot}10$, or with serious disorder problems, were excluded from the study. The X-ray determinations vary in quality from early, relatively imprecise studies to very precise determinations performed in connection with charge-density analyses. The structures contain a total of 57 N—H···O=C hydrogen bonds for which both neutron and X-ray atomic coordinates are available. The X-ray hydrogen-atom positions were all determined experimentally rather than calculated geometrically.

Error analysis

Terminology

Since there is some confusion over the use of the terms *random* and *systematic error* (Hamilton, 1969),

Table 2. *Bibliography*

(1) *N*-Acetylglycine. *X*: J. Donohue, R. E. Marsh, *Acta Cryst.*, 15, 941, 1962; *N*: M. F. Mackay, *Cryst. Struct. Commun.*, 4, 225, 1975
(2) Ammonium oxalate monohydrate. *X*: J. C. Taylor, T. M. Sabine, *Acta Cryst.*, B28, 3340, 1972; *N*: J. C. Taylor, T. M. Sabine, *Acta Cryst.*, B28, 3340, 1972
(3) L-Cysteic acid monohydrate. *X*: W. A. Hendrickson, J. Karle, *Acta Cryst.*, B27, 427, 1971; *N*: M. Ramanadham, S. K. Sikka, R. Chidambaram, *Acta Cryst.*, B29, 1167, 1973
(4) Cyanuric acid. *X*: G. C. Verschoor, E. Keulen, *Acta Cryst.*, B27, 134, 1971; *N*: P. Coppens, A. Vos, *Acta Cryst.*, B27, 146, 1971
(5) Dialuric acid monohydrate. *X*: W. Bolton, *Acta Cryst.*, 19, 1051, 1965; *N*: B. M. Craven, T. M. Sabine, *Acta Cryst.*, B25, 1970, 1969
(6) DL-Serine. *X*: T. J. Kistenmacher, G. A. Rand, R. E. Marsh, *Acta Cryst.*, B30, 2573, 1974; *N*: M. N. Frey, M. S. Lehmann, T. F. Koetzle, W. C. Hamilton, *Acta Cryst.*, B29, 876, 1973
(7) Dimethylammonium hydrogen oxalate. *X*: J. O. Thomas, S. Pramatus, *Acta Cryst.*, B31, 2159, 1975; *N*: J. O. Thomas, *Acta Cryst.*, B33, 2867, 1977
(8) Ethylenediammonium D-tartrate. *X*: S. Perez, *Acta Cryst.*, B32, 2064, 1976; *N*: C. K. Fair, E. O. Schlemper, *Acta Cryst.*, B33, 1337, 1977
(9) Glycine. *X*: J. Almlöf, Å. Kvick, J. O. Thomas, *J. Chem. Phys.*, 59, 3901, 1973; *N*: P.-G. Jonsson, Å. Kvick, *Acta Cryst.*, B28, 1827, 1972
(10) Glycylglycine. *X*: Å. Kvick, T. F. Koetzle, E. D. Stevens, *J. Chem. Phys.*, 71, 173, 1979; *N*: Å. Kvick, A. R. Al-Karaghouli, T. F. Koetzle, *Acta Cryst.*, B33, 3796, 1977
(11) Glycine hydrochloride. *X*: B. di Blasio, V. Pavone, C. Pedone, *Cryst. Struct. Commun.*, 6, 745, 1977; *N*: A. R. Al-Karaghouli, F. E. Cole, M. S. Lehmann, C. F. Miskell, J. J. Verbist, T. F. Koetzle, *J. Chem. Phys.*, 63, 1360, 1975
(12) Hydrazinium hydrogen oxalate. *X*: J. O. Thomas, *Acta Cryst.*, B29, 1767, 1973; *N*: J. O. Thomas, R. Liminga, *Acta Cryst.*, B34, 3686, 1978
(13) Hippuric acid. *X*: W. Harrison, S. Rettig, J. Trotter, *J. Chem. Soc. Perkin Trans.* 2, 1972, 1036; *N*: M. Currie, A. L. Macdonald, *J. Chem. Soc. Perkin Trans.* 2, 1974, 784
(14) L-Histidine hydrochloride monohydrate. *X*: K. Oda, H. Koyama, *Acta Cryst.*, B28, 639, 1972; *N*: H. Fuess, D. Hohlwein, S. A. Mason, *Acta Cryst.*, B33, 654, 1977
(15) Imidazolium maleate. *X*: M. N. G. James, M. Matsushima, *Acta Cryst.*, B32, 1708, 1976; *N*: B. Hsu, E. O. Schlemper, *Acta Cryst.*, B36, 3017, 1980
(16) L-Alanine. *X*: H. J. Simpson, R. E. Marsh, *Acta Cryst.*, 20, 550, 1966; *N*: M. S. Lehmann, T. F. Koetzle, W. C. Hamilton, *J. Am. Chem. Soc.*, 94, 2657, 1972
(17) L-Cysteine. *X*: K. A. Kerr, J. P. Ashmore, *Acta Cryst.*, B29, 2124, 1973; *N*: K. A. Kerr, J. P. Ashmore, T. F. Koetzle, *Acta Cryst.*, B31, 2022, 1975
(18) L-Glutamic acid. *X*: N. Hirayama, K. Shirahata, Y. Ohashi, Y. Sasada, *Bull. Chem. Soc. Jpn.*, 53, 30, 1980; *N*: M. S. Lehmann, A. C. Nunes, *Acta Cryst.*, B36, 1621, 1980
(19) L-Histidine. *X*: J. J. Madden, E. L. McGandy, N. C. Seeman, *Acta Cryst.*, B28, 2377, 1972; *N*: M. S. Lehmann, T. F. Koetzle, W. C. Hamilton, *Int. J. Pept. Protein Res.*, 4, 229, 1972
(20) L-Threonine. *X*: M. Ramanadham, S. K. Sikka, R. Chidambaram, *Pramana*, 1, 247, 1973; *N*: M. Ramanadham, S. K. Sikka, R. Chidambaram, *Pramana*, 1, 247, 1973
(21) L-Tyrosine. *X*: A. Mostad, H. M. Nissen, C. Rømming, *Acta Chem. Scand.*, 26, 3819, 1972; *N*: M. N. Frey, T. F. Koetzle, M. S. Lehmann, W. C. Hamilton, *J. Chem. Phys.*, 58, 2547, 1973
(22) L-Lysine monohydrochloride dihydrate. *X*: D. A. Wright, R. E. Marsh, *Acta Cryst.*, 15, 54, 1962; *N*: T. F. Koetzle, M. S. Lehmann, J. J. Verbist, W. C. Hamilton, *Acta Cryst.*, B28, 3207, 1972
(23) 9-Methyladenine-1-methylthymine. *X*: K. Hoogsteen, *Acta Cryst.*, 16, 907, 1963; *N*: M. N. Frey, T. F. Koetzle, M. S. Lehmann, W. C. Hamilton, *J. Chem. Phys.*, 59, 915, 1973
(24) Quinolinic acid. *X*: F. Takusagawa, K. Hirotsu, A. Shimada, *Bull. Chem. Soc. Jpn.*, 46, 2372, 1973; *N*: Å. Kvick, T. F. Koetzle, R. Thomas, F. Takusagawa, *J. Chem. Phys.*, 60, 3866, 1974
(25) *N*-Hydroxyurea. *X*: N. Armagan, J. P. G. Richards, A. A. Uraz, *Acta Cryst.*, B32, 1042, 1976; *N*: W. E. Thiessen, H. A. Levy, B. D. Flaig, *Acta Cryst.*, B34, 2495, 1978
(26) Urea. *X*: D. Mullen, E. Hellner, *Acta Cryst.*, B34, 1624, 1978; *N*: H. Guth, G. Heger, S. Klein, W. Treutmann, C. Scheringer, *Z. Kristallogr.*, 153, 237, 1980

The symbols *X* and *N* denote the X-ray and neutron diffraction determination, respectively.

their meanings in the present context will be defined here. Let X be the true value of a given parameter, of which n observations are made. Let the ith observation, x_i, be expressed as:

$$x_i = \bar{x} + \Delta x_i \tag{1}$$

where $\bar{x}$ is the mean of the population to which the n observations belong. The *systematic error* is defined as the quantity $|X - \bar{x}|$, *i.e.* the bias in the observations. In the case of X-ray determinations of hydrogen-bond geometries, the principal cause of systematic error will be defects in the theoretical model used in least-squares refinement. In particular, the observed N—H bond lengths will tend to be shorter than the true mean internuclear separations because of total or partial neglect of the N—H bonding electron density. The *random error* is defined as the standard deviation of the Δx_i, *i.e.* the unbiased component of the error. We consider below the magnitudes of the random and systematic errors in X-ray hydrogen-bond geometries, and the possibility of correcting for the latter.

Magnitude of random errors

The published e.s.d.'s of the X-ray N—H bond lengths range from 0·014 to 0·106 Å, with a mean of 0·041 Å. However, previous work (*e.g.* Abrahams, 1974; Verbist, Lehmann, Koetzle & Hamilton, 1972) has shown that standard deviations derived from least-squares covariance matrices are usually underestimated. We wished to examine this possibility by the normal-probability-plot method (Abrahams & Keve, 1971; Hamilton & Abrahams, 1972), *i.e.* by plotting the ranked weighted differences between the neutron and X-ray N—H bond lengths against the quantiles expected for a normal distribution. The ith weighted difference, Δ_i, is defined as:

$$\Delta_i = [l(N)_i - l(X)_i]/[\sigma^2(N)_i + \sigma^2(X)_i]^{1/2} \tag{2}$$

where $l(N)_i$ and $l(X)_i$ are, respectively, the observed neutron and X-ray values of the ith N—H bond length, and $\sigma(N)_i$ and $\sigma(X)_i$ are the corresponding e.s.d.'s. Unfortunately, the application of this method is complicated by the systematic differences between the neutron and X-ray distances. If the true X-ray value of the ith bond length [*i.e.* the value observed in the absence of random error; equivalent to $\bar{x}$ in (1)] is $\bar{l}(X)_i$ and the corresponding neutron value is $\bar{l}(N)_i$, (2) may be written as:

$$\Delta_i = [\bar{l}(N)_i + z(N)_i - \bar{l}(X)_i - z(X)_i]/[\sigma^2(N)_i + \sigma^2(X)_i]^{1/2} \tag{3}$$

i.e.

$$\Delta_i = [\bar{l}(N)_i - \bar{l}(X)_i]/[\sigma^2(N)_i + \sigma^2(X)_i]^{1/2} + [z(N)_i - z(X)_i]/[\sigma^2(N)_i + \sigma^2(X)_i]^{1/2}. \tag{4}$$

Assuming that the random errors are normally distributed, $z(N)_i$ and $z(X)_i$ are normal random variables with mean values of zero. If there were no systematic differences between the X-ray and neutron distances, the first term in (4) would vanish. The extent to which the e.s.d.'s are over- or underestimated could then be inferred from the slope of the normal probability plot (Abrahams & Keve, 1971). However, since there *are* systematic differences, *i.e.*

$$\bar{l}(N)_i \neq \bar{l}(X)_i \tag{5}$$

the first term in (4) has a non-zero value which varies from one hydrogen bond to another. Consequently, the probability plot for the complete sample of 57 N—H···O bonds (shown in Fig. 1) cannot be simply interpreted.

However, since:

$$\bar{l}(N)_i - \bar{l}(X)_i \simeq \bar{l}(N)_j - \bar{l}(X)_j, \quad i \neq j \tag{6}$$

and:

$$\sigma^2(X)_i \gg \sigma^2(N)_i \tag{7}$$

variations in the first term of (4) are mainly due to differences between the $\sigma(X)_i$. Thus, useful results can be obtained by restricting the probability plot to a subset of hydrogen bonds with similar $\sigma(X)_i$. The first term in (4) is then approximately constant, and the probability plot should have a non-zero intercept and a gradient that can be meaningfully interpreted in terms of the accuracy of the e.s.d.'s. Fig. 2 shows the

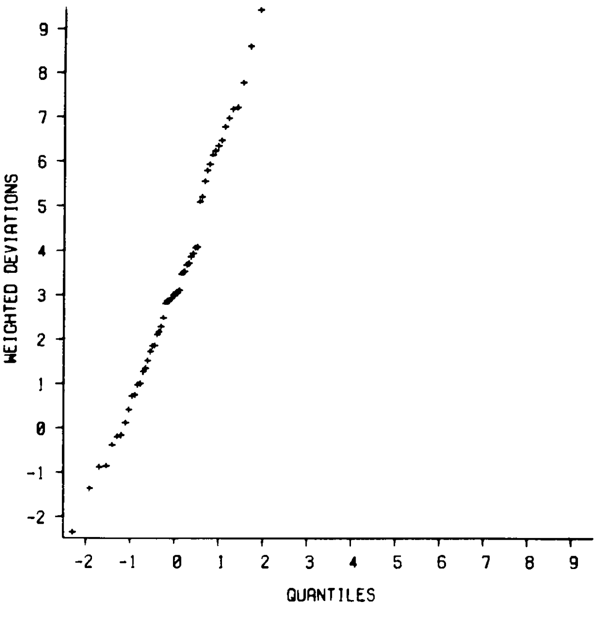


Fig. 1. Normal probability plot for complete sample. For reasons of clarity, one extreme point with coordinates (2·30,14·53) has been omitted.

probability plot for 24 N—H···O bonds with $0{\cdot}04 \le \sigma(X)_i \le 0{\cdot}06$ Å. This range, chosen empirically, is small enough to keep the first term of (4) approximately constant, but large enough to include a reasonable number of hydrogen bonds. The plot is approximately linear, indicating that the random errors are normally distributed. As expected, the intercept is non-zero. The gradient is about 1·6; this is much smaller than the slope of the plot for the complete sample ($\simeq 2{\cdot}9$). We conclude that the pooled e.s.d.'s – and hence the e.s.d.'s of the X-ray N—H distances [since $\sigma^2(X)_i \gg \sigma^2(N)_i$] – are underestimated by a factor of about 1·6.

This result was substantiated by a slightly different method (Kroon, Kanters, Van Duijneveldt-Van de Rijdt, Van Duijneveldt & Vliegenthart, 1975). The sample variance of the 57 X-ray N—H distances ($\sigma^2_{\rm sample}$) is 0·005679 Å². The variance is composed of a 'physical' contribution ($\sigma^2_{\rm phys}$), due to genuine differences between the various N—H bond lengths, and an 'experimental' contribution ($\sigma^2_{\rm expt}$), due to the random experimental errors. Since the two terms are statistically independent:

$$\sigma^2_{\rm sample} = \sigma^2_{\rm phys} + \sigma^2_{\rm expt}. \tag{8}$$

The sample variance of the neutron N—H distances (= 0·000216 Å²) can be taken as an upper limit to $\sigma^2_{\rm phys}$. Thus:

$$0{\cdot}005679 \ge \sigma^2_{\rm expt} \ge (0{\cdot}005679 - 0{\cdot}000216)\ \text{Å}^2 \tag{9}$$

i.e.

$$0{\cdot}0754 \ge \sigma_{\rm expt} \ge 0{\cdot}0739\ \text{Å}. \tag{10}$$

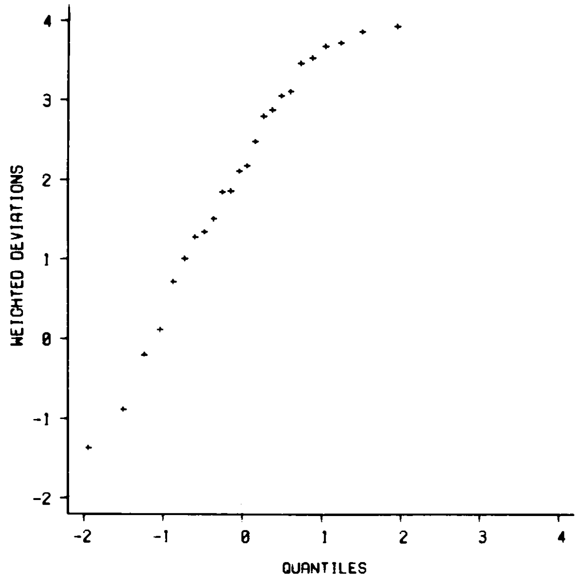


Fig. 2. Normal probability plot for hydrogen bonds with $0{\cdot}04 \le \sigma(X)_i \le 0{\cdot}06$ Å (see text).

Table 3. *Means and standard deviations of distributions*

Parameter	Neutron Mean[b]	Neutron S.d.[c]	Uncorrected X-ray Mean	Uncorrected X-ray S.d.	Corrected X-ray[a] Mean	Corrected X-ray[a] S.d.
N—H (Å)	1·030 (2)	0·015	0·938 (10)	0·075	1·030	0·000
H···O (Å)	1·865 (16)	0·122	1·956 (16)	0·119	1·869 (14)	0·108
N···O (Å)	2·849 (12)	0·090	2·850 (12)	0·089	2·850 (12)	0·089
N—H···O (°)	161·5 (15)	11·2	162·2 (15)	11·4	161·3 (15)	11·6
O···N—H (°)	12·0 (10)	7·6	12·2 (10)	7·8	12·2 (10)	7·8

(*a*) See text for details of correction procedure.

(*b*) Number in brackets is standard error of mean $\{=[\sum(x_i-\bar{x})^2/N(N-1)]^{1/2}\}$ in the units of the least significant digit of the preceding number.

(*c*) Sample standard deviation $\{=[\sum(x_i-\bar{x})^2/(N-1)]^{1/2}\}$.

The root-mean-square value of the e.s.d.'s of the X-ray N—H bond lengths is 0·0458 Å. Since 0·0754/0·0458 = 1·65, and 0·0739/0·0458 = 1·61, this again suggests that the X-ray e.s.d.'s are underestimated by a factor of about 1·6. Multiplication of the $\sigma(X)_i$ by this factor shows that the random errors in the X-ray N—H distances vary between 0·02–0·17 Å, the mean value being 0·065 Å. Since oxygen and nitrogen have similar atomic weights, the random errors in the X-ray H···O distances will be about the same.

Magnitude of systematic errors

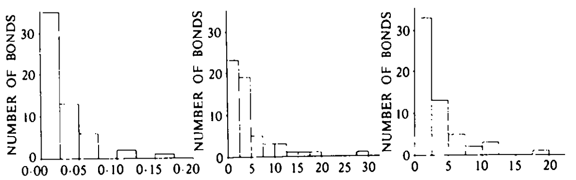
The mean values of the X-ray and neutron N—H bond-length distributions are 0·938 (10) and 1·030 (2) Å, respectively (Table 3), which suggests that the average systematic error in the X-ray distances is approximately 0·1 Å. Underestimation of the N—H bond length will produce systematic errors in two related parameters, the H···O distance and the N—H···O angle. Both are likely to be overestimated, because the N—H···O angle is invariably greater than 90°. Application of a one-tailed Wilcoxon test (Siegel, 1956) showed that the X-ray H···O distances are stochastically larger than the corresponding neutron distances at the 99·9% confidence level. The mean values [X-ray: 1·956 (16) Å; neutron: 1·865 (16) Å] indicate that the average systematic error in the X-ray H···O distances is approximately 0·1 Å. As expected, the mean value of the X-ray N—H···O angles [162·2 (1·5)°] is larger than the mean neutron value [161·5 (1·5)°]. However, the difference between them is not statistically significant.

Correction of systematic errors

The simplest method of correcting for the systematic errors is to 'normalize' the X-ray hydrogen-atom positions (Jeffrey & Lewis, 1978). This involves moving the hydrogen-atom position along the observed N—H bond direction until the N—H distance is equal to the mean neutron value. Although this procedure ignores the effect of crystal packing forces on the N—H distance, an analogous technique has been shown to give good results when applied to the O—H···O hydrogen bond (Jeffrey & Lewis, 1978; Jeffrey & Takagi, 1978).

The mean neutron N—H distance was found to be 1·030 (2) Å. This is an unweighted average of all 57 bond lengths in our sample. Omission of the eight bond lengths determined at low temperature, or weighting of the various bond lengths according to their e.s.d.'s (Hamilton, 1964), did not produce significantly different results. The X-ray N—H distances were therefore normalized to 1·030 Å and the $H\cdots O$ and $N—H\cdots O$ distributions redetermined. The results (Table 3) indicate that the correction procedure is satisfactory. In particular, the mean $H\cdots O$ distance of the 'corrected' X-ray distribution, 1·869 (14) Å, is not significantly different from the mean neutron value, 1·865 (16) Å. The average $N—H\cdots O$ angles are also in good agreement [corrected X-ray: 161·3 (1·5)°; neutron: 161·5 (1·5)°]. The histograms in Fig. 3 show the distributions of the following parameters: $|(H\cdots O)_C - (H\cdots O)_N|$; $|(N—H\cdots O)_C - (N—H\cdots O)_N|$; $|(O\cdots N—H)_C - (O\cdots N—H)_N|$. Here, the subscript '$N$' denotes a value determined by neutron diffraction and the subscript 'C' denotes a corrected X-ray value (owing to the nature of the normalization procedure, the corrected and uncorrected values of the $O\cdots N—H$ angle are identical). The histograms confirm that the fit between the corrected X-ray and neutron hydrogen-bond geometries is good. *E.g.* the difference between the corrected X-ray and neutron $H\cdots O$ distances exceeds 0·05 Å in only nine out of 57 cases; the parameter $|(N—H\cdots O)_C - (N—H\cdots O)_N|$ is less than 10° for all but seven of the hydrogen bonds.*

The success of the normalization procedure is based on the fact that the mean neutron N—H distance, 1·030 Å, is usually a more accurate estimate of the true value of a given N—H bond length than the observed X-ray distance. We considered the possibility that even better estimates might be obtained by the use of statistical techniques such as regression analysis. However, the X-ray and neutron N—H bond lengths in our sample have a correlation coefficient of only 0·048 (the scatterplot is shown in Fig. 4); this can be ascribed to the relatively large random errors in the X-ray values. Consequently, the use of regression analysis to estimate 'true' (*i.e.* neutron) N—H distances from observed X-ray values is unlikely to be worthwhile.



(*a*) (*b*) (*c*)

Fig. 3. Distributions of: (*a*) $|(H\cdots O)_C - (H\cdots O)_N|$; (*b*) $|(N—H\cdots O)_C - (N—H\cdots O)_N|$; and (*c*) $|(O\cdots N—H)_C - (O\cdots N—H)_N|$. The subscript '$N$' denotes a value determined by neutron diffraction, 'C' denotes a corrected X-ray value.

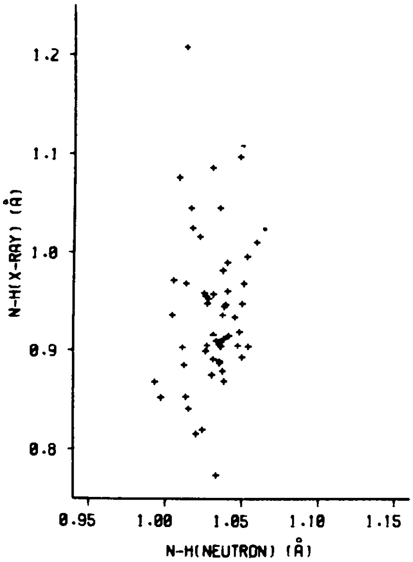


Fig. 4. Scatterplot of X-ray and neutron N—H distances.

Conclusions

Precision of X-ray hydrogen-bond geometries

The results of this study show that estimates of mean hydrogen-bond lengths and angles derived from corrected X-ray data are likely to be more precise than estimates based on neutron diffraction results. This is because the large number of X-ray determinations in the literature more than compensates for the relatively low precision of the X-ray results. The following example illustrates this point.

In a recent survey of $O—H\cdots O$ hydrogen-bond geometries determined by neutron diffraction (Ceccarelli *et al.*, 1981) it was found that the sample standard deviation of the $H\cdots O$ hydrogen-bond and O—H valence-bond distributions were 0·0774 and 0·0112 Å, respectively. The difference between these figures is due to the greater sensitivity of the $H\cdots O$ distance to its environment. From (8), the physical contribution to the standard deviation of the $H\cdots O$ distribution (*i.e.* σ_{phys}) can be estimated to be at least $(0·0774^2 - 0·0112^2)^{1/2}$, *i.e.* 0·0766 Å. If the $H\cdots O$ distances were to be redetermined by X-ray diffraction, the results of our error analysis suggest that the average random experimental error would be approximately 0·065 Å. The expected sample standard deviation of the X-ray $H\cdots O$ distances would therefore be about 0·1005 Å $[= (0·065^2 + 0·0766^2)^{1/2}]$.

---

* *Note added in proof:* A referee suggests that the directions of N—H bonds are accurately determined by X-ray diffraction because the hydrogen electrons have predominantly *s* character and are therefore cylindrically symmetrized around the N—H vector.

The mean value of a sample can be estimated with a standard error (σ_{mean}) of:

$$\sigma_{\text{mean}} = \sigma_{\text{sample}}/\sqrt{N} \qquad (11)$$

where σ_{sample} is the sample standard deviation and N is the number of observations in the sample (Hamilton, 1964). If the figures calculated above are typical, they suggest that the mean proton···acceptor distance of a given class of hydrogen bond (*e.g.* H_3N^+—H···O=C or Me_3N^+—H···O=C) can be estimated more precisely from X-ray data than from neutron data, provided that:

$$N_x/N_n > (0{\cdot}1005/0{\cdot}0774)^2 \qquad (12)$$

i.e.

$$N_x/N_n > \sim 1{\cdot}7 \qquad (13)$$

where N_x = number of available X-ray observations, and N_n = number of available neutron observations. Since the number of X-ray determinations in the literature exceeds the number of neutron determinations by two orders of magnitude, inequality (13) is always likely to be satisfied.

The details of this argument are altered if the X-ray N—H bond lengths are normalized, but the overall conclusion remains the same. The effect of normalization is to make the uncertainties in the X-ray hydrogen-atom positions anisotropic. However, the general trend will be to *reduce* the errors rather than increase them, because the standard deviation of the neutron N—H bond-length distribution (0·015 Å) is smaller than the average random error in the X-ray hydrogen-atom positions.*

Accuracy of X-ray hydrogen-bond geometries

The results of the error analysis show that X-ray data can provide accurate information about N—H···O hydrogen-bond geometries, provided that systematic errors are corrected by normalization of the N—H bond lengths. In the present study, the mean H···O distance and N—H···O angle calculated from the corrected X-ray data are almost identical to the corresponding neutron diffraction estimates. In many cases, the discrepancies between the neutron and corrected X-ray hydrogen-bond geometries are no larger than errors introduced by neglecting, or inadequately correcting for, thermal-motion effects. Librational corrections to neutron N—H distances can easily be as large as 0·02 Å and are very sensitive to the model used for the thermal motion (Koetzle, Lehmann, Verbist & Hamilton, 1972). The removal of systematic errors by normalization means that X-ray and neutron data can be used together in statistical surveys of hydrogen-bond geometries.

Olga Kennard is a member of the external staff of the Medical Research Council.

References

Abrahams, S. C. (1974). *Acta Cryst.* **B30**, 261–268.

Abrahams, S. C. & Keve, E. T. (1971). *Acta Cryst.* **A27**, 157–165.

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.

Ceccarelli, C., Jeffrey, G. A. & Taylor, R. (1981). *J. Mol. Struct.* **70**, 255–271.

Hamilton, W. C. (1964). In *Statistics in Physical Science*. New York: Ronald Press.

Hamilton, W. C. (1969). *Acta Cryst.* **A25**, 194–206.

Hamilton, W. C. & Abrahams, S. C. (1972). *Acta Cryst.* **A28**, 215–218.

Jeffrey, G. A. & Lewis, L. (1978). *Carbohydr. Res.* **60**, 179–182.

Jeffrey, G. A. & Maluszynska, H. (1982). *Int. J. Biol. Macromol.* **4**, 173–185.

Jeffrey, G. A. & Takagi, S. (1978). *Acc. Chem. Res.* **11**, 264–270.

Koetzle, T. F. & Lehmann, M. S. (1976). In *The Hydrogen Bond*, Vol. II, edited by P. Schuster, G. Zundel & C. Sandorfy. Amsterdam: North-Holland.

Koetzle, T. F., Lehmann, M. S., Verbist, J. J. & Hamilton, W. C. (1972). *Acta Cryst.* **B28**, 3207–3214.

Kroon, J., Kanters, J. A., Van Duijneveldt-Van de Rijdt, J. G. C. M., Van Duijneveldt, F. B. & Vliegenthart, J. A. (1975). *J. Mol. Struct.* **24**, 109–129.

Olovsson, I. & Jönsson, P.-G. (1976). In *The Hydrogen Bond*, Vol. II, edited by P. Schuster, G. Zundel & C. Sandorfy. Amsterdam: North-Holland.

Pimentel, G. C. & McClellan, A. L. (1960). In *The Hydrogen Bond*. San Francisco: W. H. Freeman.

Ramakrishnan, C. & Prasad, N. (1971). *Int. J. Protein Res.* **3**, 209–231.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences*, pp. 75–83. Tokyo: McGraw-Hill Kogakusha.

Taylor, R. & Kennard, O. (1982). *J. Am. Chem. Soc.* **104**, 5063–5070.

Verbist, J. J., Lehmann, M. S., Koetzle, T. F. & Hamilton, W. C. (1972). *Acta Cryst.* **B28**, 3006–3013.

---

* Surprisingly, the sample standard deviation of the corrected X-ray H···O distances in the present study is less than that of the neutron distances (Table 3). We regard this as a freak result, because the random errors in the X-ray distances are certainly larger than those in the neutron distances.