Acta Cryst. (1995). **D51**, 93–97

Water Molecules Which Apparently Accept No Hydrogen Bonds are Systematically Involved in C—H···O Interactions

By Thomas Steiner

Institut für Kristallographie, Freie Universität Berlin, Takustrasse 6, *D*-14159 *Berlin, Germany*

(*Received* 17 *March* 1994; *accepted* 29 *June* 1994)

Abstract

The hydrogen-bond acceptor functions are analyzed for 672 water molecules in crystal structures of hydrated small biological molecules. Of the 672 water molecules, 26 neither accept hydrogen bonds from O—H or N—H donors, nor coordinate to metal ions. As compensation, they satisfy their acceptor potential with up to four C—H···O hydrogen bonds. They are found in hydrated amino acids, purines and pyrimidines, and alkaloids, but not in carbohydrates, nucleosides, nucleotides and steroids. As for conventional hydrogen bonds, the preferred coordination geometry is tetrahedral. No example of a true non-accepting water molecule is found. A particularly strong C—H hydrogen-bond donor occurring in many biological molecules is C—H neighbouring protonated N.

Introduction

Water molecules which cannot arrange in a more or less ideal O—H···O or N—H···O hydrogen-bond configuration or coordinate to metal cations in crystal structures often satisfy apparently 'free' hydrogen-bond acceptor potential by C—H···O$_W$ interactions [Steiner & Saenger (1993*a*). For background on C—H···O hydrogen bonds, see Green (1974); Berkovitch-Yellin & Leiserowitz (1984); Jeffrey & Saenger (1991); Desiraju (1991)]. Using C—H donors to fill gaps in incomplete coordination shells, water molecules can find stable positions also in relatively apolar surroundings. This is evidently of importance for hydrated biological molecules, which often have surface regions with poor hydrogen-bond donor capacities facing the surrounding solvent. For systems of biological relevance, stabilization of extended water arrangements by C—H···O$_W$ interactions was discussed on the basis of neutron diffraction data for cyclodextrin inclusion complexes (Steiner & Saenger, 1992) and vitamin B$_{12}$ hydrate (Steiner & Saenger, 1993*b*; structure data from Bouquiere, Finney, Lehmann, Lindley & Savage, 1993).

Occasionally, crystal structures are reported, in which water O atoms neither accept any hydrogen bonds, nor coordinate to metal cations. According to Jeffrey & Maluszynska (1990), such cases amount to a significant percentage of the water molecules in the hydrates of amino acids and peptides (see also Jeffrey & Saenger, 1991), and therefore are of potential interest for biological structures. As in these studies, possible C—H···O$_W$ interactions were generally not taken into account, it should be checked whether these peculiar non-acceptors are in fact accepting hydrogen bonds from C—H donors. On a phenomenological basis of contact and coordination geometries, the present study presents such an analysis for the hydrates of small biological molecules.

Table 1. *Number of analyzed water molecules*

Compounds	Total No.	Water accepting only C—H···O$_W$
Amino acids and peptides	222	12 (= 5%)
Nucleosides and nucleotides	133	0
Carbohydrates*	124	0
Pyrimidines and purines	114	6 (= 5%)
Alkaloids	51	8 (= 16%)
Steroids	28	0

* Including cyclodextrin inclusion complexes.

Data analysis

X-ray and neutron crystal structures of six classes of hydrated biologically relevant small molecules were retrieved from the Cambridge Structural Database (Allen *et al.*, 1979. Quality criteria: structures must not be disordered, $R < 0.07$, H-atom coordinates given, relevant C—H and O—H bond lengths between 0.6 and 1.15 Å, water H—O—H angles between 90 and 120°). Since in X-ray crystal structures, *X*—H bond lengths are systematically too short by ~0.1 Å or more (Allen, 1986), they were 'normalized' by shifting the H atom along the covalent bond to bond lengths of 0.97 Å for O—H, 1.04 Å for N—H (all types) and 1.09 Å for C—H (all types) (see also Jeffrey & Maluszynska, 1990). Water molecules forming unrealistic H···H contacts < 1.8 Å (after normalization) were excluded. The data set obtained contains 672 water molecules, Table 1, for which the acceptor functions were then determined as in the previous study (Steiner & Saenger, 1993*a*).

Acta Crystallographica Section D
ISSN 0907-4449 ©1995

Table 2. *The water molecules that accept no other than* C—H···O *hydrogen bonds* (*and no metal ion contacts*)

Charge of the main molecule: + cation; ± zwitterion; 0 uncharged.

Compound	Charge	Water	No. of C—H···O_W*	Shortest H···O_W†	Shortest C···O_W†	Reference
Amino acids and peptides						
PAIPPR‡ (monohydrate)	0	*W*2	3	2.45	3.34	Aubry *et al.* (1980)
PAIPPR‡ (monohydrate)	0	*W*1	1	2.67	3.60	Aubry *et al.* (1980)
BGPLGQ§ (dihydrate)	0	*W*9	4	2.43	3.29	Bando *et al.* (1978)
cyclo-(Gly-L-Pro-L-Pro-Gly-L-Pro-L-Pro) trihydrate	0	*W*2	4	2.43	3.36	Czugler *et al.* (1982)
cyclo-(Gly-Sar-Sar-Sar) monohydrate	0	*W*	3	2.26	3.21	Declercq *et al.* (1975)
Gly-L-Ala-L-Ile hemihydrate (form 2)	±	*W*	2	2.40	3.26	Go *et al.* (1992)
cyclo-(Sar)$_7$ monohydrate	0	*W*	2	2.38	3.12	Groth (1975)
cyclo-[L-Pro-(Bzl)Gly-D-Pro] monohydrate	0	*W*	2	2.45	3.27	Kessler *et al.* (1983)
Betaine monohydrate	±	*W*	3	2.39	3.48	Mak (1990)
cyclo-(L-His-L-Asp) trihydrate	±	*W*3	1	2.31	3.13	Ramani *et al.* (1978)
L-Ornithine L-aspartate hemihydrate	±	*W*	2	2.44	3.46	Salunke & Vijayan (1983)
cyclo-(-L-Pro-Gly-]$_2$ trihydrate	0	*W*2	3	2.19	3.19	Shoham *et al.* (1989)
Purines and pyrimidines						
Pipemidic acid trihydrate	±	*W*3	3	2.39	3.44	Fonesca *et al.* (1986)
FUCWOP¶ (chloride dihydrate)	+	*W*1	2	2.17	3.25	Jaskólski *et al.* (1987)
GEXTUY** (dichloride monohydrate)	+	*W*	1	2.23	3.18	Jaskólski *et al.* (1988)
Caffeine hydrochloride diydrate	+	*W*2	2	2.29	3.15	Mercer & Trotter (1978)
9-Methyl-8-phenyl-thiopurine hemihydrate	0	*W*	3	2.45	3.51	Nygjerd & Sletten (1973)
Hypoxanthinium nitrate monohydrate	+	*W*	1	2.05	3.13	Rosenstein *et al.* (1982)††
Alkaloids						
Isomalindine chloride dihydrate	+	*W*1	2	2.56	3.37	Arbain *et al.* (1993)
Clivorine monohydrate	0	*W*	3	2.31	3.36	Birnbaum (1972)
(−)-Mesembrane hydrochloride monohydrate	+	*W*	2	2.31	3.39	Capps *et al.* (1977)
ent-Norsecurinine hydrochloride monohydrate	+	*W*	4	2.41	3.48	Joshi *et al.* (1986)
Naloxon hydrochloride dihydrate	+	*W*5	2	2.28	3.34	Karle (1974)
Naloxon hydrochloride dihydrate at 90 K	+	*W*5	2	2.39	3.31	Klein *et al.* (1987)
Crispatine monohydrate	0	*W*	2	2.55	3.58	Mackay *et al.* (1984)
VUXFOJ‡‡ (hydrobromide monohydrate)	+	*W*	3	2.43	3.45	Verlinde *et al.* (1988)

* With H···$O_W < 2.8$ Å.

† For normalized H-atom positions (Å). The shortest H···O_W contact is not necessarily associated with the shortest C···O_W separation.

‡ Pivaloyl-D-alanyl-*N*-isopropyl-D-prolinamide monohydrate.

§ *N*-Benzyloxycarbonyl-Gly-L-Pro-L-Leu-Gly-L-Pro dihydrate.

¶ *N*-(2-Amino-6-purinyl)pyridinium chloride dihydrate.

** *N*,*N*′-(2,4-Pyrimidinediyl)dipyridinum dichloride monohydrate.

†† Repeated and verified by Schmalle, Hänggi & Dubler (1990).

‡‡ (−)-(1*R*,5*R*,9*R*)-2′-Hydroxy-2-methoxyethyl-5,9-dimethyl-6,7-benzomorphan hydrobromide monohydrate.

Results

General

Of the 672 water molecules, 646 accept hydrogen bonds from O—H or N—H donors (and often additional ones from C—H donors), or are in contact with metal cations, and can be disregarded in the present context. The remaining 26 water molecules are found in hydrates of amino acids and peptides (12 of 222 water molecules = 5%), of pyrimidines and purines (6 of 114 = 5%) and of alkaloids (8 of 51, = 16%), but no examples were found in carbohydrate, nucleoside and nucleotide, or in steroid crystal structures, Table 1. These water molecules all accept C—H···O_W hydrogen bonds, Table 2. The typical H···O_W distances are around 2.4 Å (down to 2.1 Å), and the typical C···O_W separations are around 3.3 Å (down to 3.1 Å), Table 2. The water molecule for which the closest C—H donor is most far away, is *W*1 in the peptide PAIPPR (Table 2); since H···O_W = 2.67 Å, this can still be regarded as a long C—H···O hydrogen bond. No example of a true non-accepting water molecule is found.

Water coordination

The number of accepted C—H···O hydrogen bonds is given in Table 2 for the 26 individual water molecules. The most frequent acceptor function is, as expected, twofold (11 molecules of 26). Tetrahedral water coordination is hence possible even in the absence of conventional hydrogen-bond donors, two typical examples with short H···O_W distances around 2.4 Å are shown in Fig. 1. A threefold acceptor function is found for eight of the 26 water molecules; a typical example with resulting fivefold coordination (associated with irregular geometry) is shown in Fig. 2(*a*). Fourfold and onefold acceptor functions are rarer (three and four molecules of the 26, respectively); a fourfold acceptor is shown in Fig. 2(*b*), and a onefold acceptor (*W*3) in Fig. 2(*c*). The water-coordination behaviour thus essentially parallels the one with stronger O—H and N—H donors (except for the longer donor–acceptor separations).

The C—H *donor molecules*

Water molecules which accept only C—H···O hydrogen bonds are found in hydrates of only certain substance classes, Table 1. The common feature of these classes is that a lack of strong O—H and N—H donors is often combined with the occurrence of acidic C—H groups. This is most pronounced for the alkaloids, which are usually crystallized as cations: they often carry only the single strong donor N^+—H, and at the same time, have several acidic C—H donors close to the positive charge [the C—H donor strength is increased in the vicinity of electron-withdrawing groups and correlates with the carbon acidity. (Green, 1974; Pedireddi & Desiraju, 1992)]. The consequences for some hydration water molecules are illustrated in an example in Fig. 2(*b*): in norsecurinine hydrochloride monohydrate (Joshi, Gawad, Pelletier, Kartha & Bhandary, 1986), the only strong hydrogen-bond donors are N^+—H and the water molecule, all of which donate hydrogen bonds to (symmetry-related) chloride anions. As no strong donor is left for the water aceptor, it can only resort to C—H···O_W interactions. In ionic purine and pyrimidine hydrates, water molecules sometimes are similarly unable to compete with the anion for the few strong donors; a typical example is caffeine hydrochloride dihydrate (Mercer & Trotter, 1978) shown in Fig. 2(*d*); note that the caffeine molecule is protonated at N9.

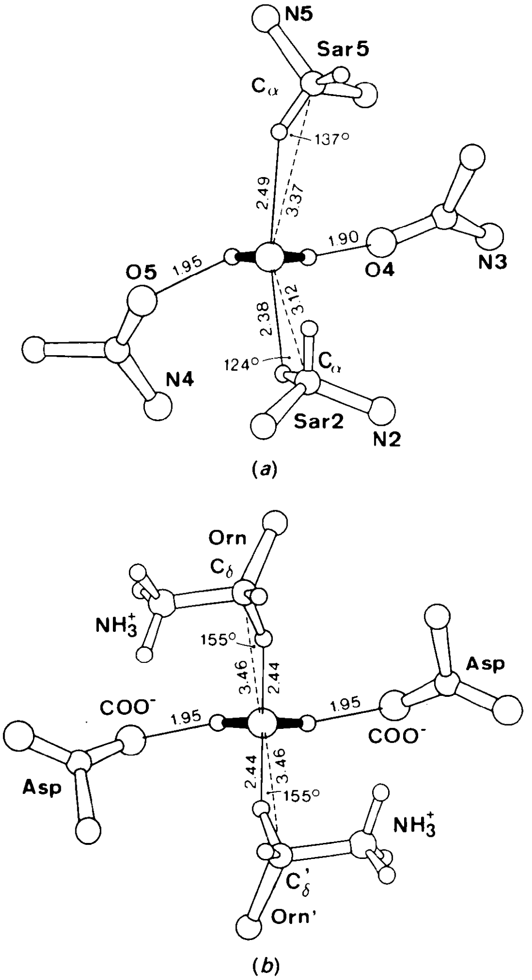

Fig. 1. Examples for water molecules accepting two C—H···O_W hydrogen bonds with almost ideal tetrahedral coordination geometries, projections are along the water bisectors. Geometries for normalized H-atom positions (Å). (*a*) In *cyclo*-$(Sar)_7$ (cycloheptasarcosyl) monohydrate (Groth, 1975). (*b*) In L-ornithine L-aspartate hemihydrate (Salunke & Vijayan, 1983) [ornithine is related to lysine, with $R = (CH_2)_3$—NH_3^+].

Of particular importance are the hydrates of amino acids and peptides. These molecules frequently form zwitterions, and then have increased C—H donor strengths near the positive charge(s). Also uncharged peptides carry strong C—H donors, such as C_α—H (compare Jeffrey & Maluszynska, 1982). Of the 12 water molecules from amino acid and peptide structures in Table 2, six accept their shortest C—H···O_W interaction from a C_α—H donor (four from uncharged molecules, and two with adjacent N^+), two typical examples are shown in Figs. 1(*a*) and 2(*a*). The other six accept their shortest contact from side-chain C—H donors (two Pro C_δ, two other uncharged residues, and two charged residues), two instructive examples with protonated side-chains are shown in Figs. 1(*b*) and 2(*c*). Note that in Fig. 2(*c*) all H atoms of the protonated histidine ring are involved in short hydrogen bonds; the water arrangement around this residue definitely cannot be understood without taking C—H···O interactions into account.

Discussion

In systems in which a relative lack of O—H and N—H donors is combined with the occurrence of acidic C—H groups water molecules which accept no other near-neighbor interactions than C—H···O_W hydrogen bonds are found occasionally. In this way, water molecules can find stable positions even in the complete absence of conventional hydrogen-bond donors (still, the rareness of this configuration suggests that it is not a very favourable one). In biologically relevant systems strong C—H donors are abundant, for example C_α—H in peptides and C—H neighboring to protonated N. As there is no reason why the observations above should be restricted to small biological molecules and to the solid

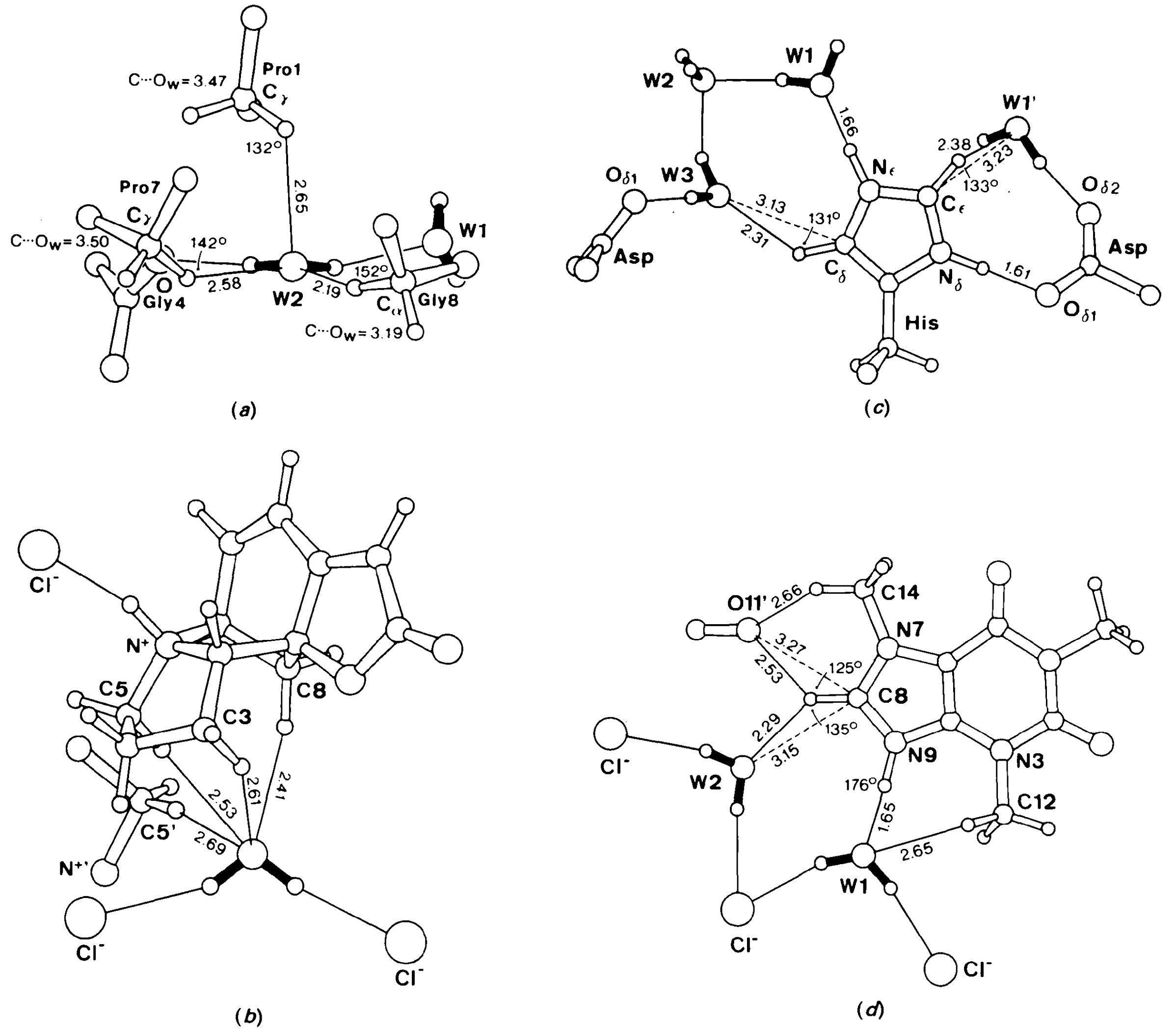

Fig. 2. Examples for water molecules accepting only C—H···O_w hydrogen bonds. Geometries for normalized H-atom positions (Å). (*a*) In the uncharged peptide *cyclo*-(L-Pro-Gly)$_2$ trihydrate (Shoham, Burley & Lipscomb, 1989). The coordination is irregular fivefold. (*b*) In norsecurinine hydrochloride monohydrate (Joshi *et al.*, 1986). Note the chelation of the water molecule in a concave part of the alkaloid. C3···O_w = 3.55 Å (angle at H = 144°), C5···O_w = 3.48 Å (angle at H = 144°), C8···O_w = 3.48 Å (angle at H = 169°), C5′···O_w = 3.77 Å (angle at H = 172°). (*c*) In *cyclo*-(L-His-L-Asp) trihydrate (Ramani, Venkatesan & Marsh, 1978), the histidine ring is protonated at N_ε. For *W*3, the interaction with C_δ—H is the only accepted hydrogen bond. (*d*) In caffeine hydrochloride dihydrate (Mercer & Trotter, 1978) the caffeine is protonated at N9. A long C—H···O contact of *W*2 with a symmetry-related caffeine cation is omitted for clarity (H···O = 2.75 Å).

state, it is assumed that they are also valid for structurally related macromolecules and for molecules in solution. Due to the lower hydrogen-bond energy (normally 1 to 2 kcal mol^{-1}), however, the lifetime of C—H···O hydrogen-bonded arrangements should then be shorter than for other hydrogen bonds.

In the present data set, no example of a real non-accepting water molecule is observed. This does not necessarily mean that such water molecules do not exist in crystal structures: if they do occur, they would certainly be prone to vibrate extensively or to disorder, and would then be excluded by the rigid structural quality criteria of this study.

I thank Wolfram Saenger for giving me the opportunity to carry out this study in his laboratory.

References

ALLEN, F. H. (1986). *Acta Cryst.* **B42**, 515–522.

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.

Arbain, D., Lajis, N. H., Putra, D. P., Sargent, M. V., Skelton, B. W. & White, A. H. (1993). *Aust. J. Chem.* **46**, 969–976.

Aubry, A., Protas, J., Boussard, G. & Marraud, M. (1980). *Acta Cryst.* B**36**, 321–326.

Bando, S., Tanaka, N., Ashida, T. & Kakudo, M. (1978). *Acta Cryst.* B**34**, 3447–3449.

Berkovitch-Yellin, Z. & Leiserowitz, L. (1984). *Acta Cryst.* B**40**, 150–165.

Birnbaum, K. B. (1972). *Acta Cryst.* B**28**, 2825–2833.

Bouquiere, J. P., Finney, J. L., Lehmann, M. S., Lindley, P. F. & Savage, H. F. J. (1993). *Acta Cryst.* B**49**, 79–89.

Capps, T. M., Hargrave, K. D., Jeffs, P. W. & McPhail, A. T. (1977). *J. Chem. Soc. Perkin Trans.* II, pp. 1098–1104.

Czugler, M., Sasvárin, K. & Hollósi, M. (1982). *J. Am. Chem. Soc.* **104**, 4465–4469.

Declercq, J.-P., Germain, G., van Meerssche, M., Debraerdemaeker, T., Dale, K. & Titlestad, K. (1975). *Bull. Soc. Chim. Belg.* **84**, 275–287.

Desiraju, G. R. (1991). *Acc. Chem. Res.* **24**, 290–296.

Fonesca, I., Martinez-Carrera, S. & Garcia-Blanco, S. (1986). *Acta Cryst.* C**42**, 1618–1621.

Go, K., Chaturvedi, S. & Parthasarathy, R. (1992). *Biopolymers*, **32**, 107–117.

Green, R. D. (1974). *Hydrogen Bonding by* C—H *Groups*. New York: Wiley Interscience.

Groth, P. (1975). *Acta Chem. Scand.* A**29**, 38–44.

Gylbert, L. (1973). *Acta Cryst.* B**29**, 1630–1635.

Jaskólski, M., Skalski, B. & Adamiak, D. A. (1988). *Acta Cryst.* C**44**, 1409–1412.

Jaskólski, M., Skalski, B., Adamiak, D. A. & Adamiak, R. W. (1987). *Acta Cryst.* C**43**, 2110–2113.

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

Jeffrey, G. A. & Maluszynska, H. (1990). *Acta Cryst.* B**46**, 546–549.

Jeffrey, G. A. & Saenger, W. (1991). *Hydrogen Bonding in Biological Structures*. Berlin: Springer-Verlag.

Joshi, B. S., Gawad, D. P., Pelletier, S. W., Kartha, G. & Bhandary, K. (1986). *J. Nat. Prod.* **49**, 614–620.

Karle, I. (1974). *Acta Cryst.* B**30**, 1682–1686.

Kessler, H., Bermel, W., Krack, G., Bats, J. W., Fuess, H. & Hull, W. E. (1983). *Chem. Ber.* **116**, 3164–3181.

Klein, C. L., Majeste, R. J. & Stevens, E. D. (1987). *J. Am. Chem. Soc.* **109**, 6675–6681.

Mackay, M. F., Sadek, M. & Culvenor, C. C. J. (1984). *Acta Cryst.* C**40**, 470–472.

Mak, T. C. W. (1990). *J. Mol. Struct.* **220**, 13–18.

Mercer, A. & Trotter, J. (1978). *Acta Cryst.* B**34**, 450–453.

Nygjerd, G. & Sletten, E. (1973). *Acta Chem. Scand.* **27**, 2902–2908.

Pedireddi, V. R. & Desiraju, G. R. (1992). *J. Chem. Soc. Chem. Commun.* pp. 988–990.

Ramani, R., Venkatesan, K. & Marsh, R. E. (1978). *J. Am. Chem. Soc.* **100**, 949–953.

Rosenstein, R. D., Oberding, M., Hyde, J. R., Zubieta, J., Karlin, K. D. & Seeman, N. C. (1982). *Cryst. Struct. Commun.* **11**, 1507–1513.

Salunke, D. M. & Vijayan, M. (1983). *Int. J. Pept. Protein Res.* **22**, 154–160.

Schmalle, H., Hänggi, G. & Dubler, E. (1990). *Acta Cryst.* C**46**, 340–342.

Shoham, G., Burley, S. K. & Lipscomb, W. N. (1989). *Acta Cryst.* C**45**, 1944–1948.

Steiner, T. & Saenger, W. (1992). *J. Am. Chem. Soc.* **114**, 10146–10154.

Steiner, T. & Saenger, W. (1993*a*). *J. Am. Chem. Soc.* **115**, 4540–4547.

Steiner, T. & Saenger, W. (1993*b*). *Acta Cryst.* D**49**, 592–593.

Verlinde, C. L., Blaton, N. M., Peeters, O. M. & de Ranter, C. J. (1988). *Acta Cryst.* C**44**, 1792–1794.