Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

Rules governing the crystal packing of mono- and dialcohols

Robin Taylor* and Clare F. Macrae

Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, England

Correspondence e-mail: taylor@ccdc.cam.ac.uk

Received 21 May 2001
Accepted 15 August 2001

In memory of Professor Jan Kroon, a pioneer in the exploitation of crystallographic databases.

A new program, *Mercury*, has been used to survey 144 monoalcohol (C_mH_nOH) and 101 dialcohol [$C_mH_n(OH)_2$] crystal structures. Results show that their hydrogen-bonding patterns are strongly correlated with steric effects. Primary monoalcohols have a strong preference to form infinite $\cdots OH\cdots OH\cdots$ chains. Secondary monoalcohols form chains and rings of hydrogen bonds with about equal facility. Tertiary monoalcohols very often form isolated $OH\cdots O$ hydrogen bonds or structures containing no $OH\cdots O$ hydrogen bonds at all. In the latter case, however, the structures almost invariably contain $CH\cdots O$ and/or $OH\cdots\pi$ interactions. Substitution on the β-carbon(s) of monoalcohols has a profound effect on packing patterns, with increased substitution disfavouring chains and rings. Dialcohols show a much stronger preference for chains of hydrogen bonds, compared with monoalcohols. This is particularly so when at least one of the hydroxyl groups is primary, in which case chains are overwhelmingly preferred. Once again, substitution on the β-C atoms is influential, heavy substitution tending to lead to packing arrangements that involve isolated or intramolecular $OH\cdots O$ hydrogen bonds. Dialcohols almost never crystallize without at least some $OH\cdots O$ hydrogen-bond formation. In both monoalcohols and dialcohols, chains show a stronger preference to be helical (usually threefold helices) as steric hindrance increases. Hydrogen-bonded rings usually contain four $OH\cdots O$ hydrogen bonds. It is possible that empirical observations such as these may aid crystal-structure prediction.

1. Introduction

Crystal structure prediction is a major scientific challenge and of immense potential interest to the materials, pharmaceutical and agrochemical industries. Some years ago, Gavezzotti (1994) wrote a well known paper entitled 'Are crystal structures predictable?' and gave the simple answer 'No'. Since then, great efforts have been made in the area by several groups, many of which participated in a recent competition to assess the state of the art (Lommerse *et al.*, 2000). The outcome was that all four of the test structures were predicted successfully by at least one of the groups. While no group was 100% successful and the test molecules were much smaller and less flexible than *e.g.* typical drug molecules, the study clearly showed that advances have been made. Nevertheless, reliable prediction of the types of structures in which practising chemists are interested is still far away.

A key complication is that crystallization may often be under kinetic control, which means that a successful prediction strategy needs to take account of dynamic processes in the liquid as well as the enthalpic stability of putative solid-state packing arrangements. In principle, molecular dynamics simulations offer a way forward (Gavezzotti & Filippini, 1998). The cost of such calculations, however, is prohibitive, certainly for routine prediction with current computers. We therefore need a cheap way of assessing the effects of the kinetics of crystal nucleation. One possibility is to look at the 230 000 crystal-packing arrangements that have been observed experimentally and are stored in the Cambridge Structural Database (CSD; Allen & Kennard, 1993). These do not tell us anything about the kinetics of crystal nucleation *per se*, but they do tell us something about their end results. If particular packing motifs are seen over and over again, they presumably reflect favourable kinetic processes as well as thermodynamically stable arrangements. Equally, there may be thermodynamically stable arrangements that are infrequently seen because they are kinetically disfavoured.

CSD information has already been used to develop pairwise potentials for use in crystal-structure prediction programs (*e.g. FlexCryst*; Hofmann & Lengauer, 1997). Our emphasis here, however, is on the use of the CSD to develop simple heuristics that can help discriminate between likely and unlikely packing arrangements, or can be used to reduce the search space that must be visited by a crystal-structure prediction program. Previous studies in the area, such as the pioneering work of Leiserowitz (1976), were reviewed recently by Brock (1999) and our present work follows on from one of the investigations mentioned in that review. Specifically, Brock & Duncan (1994) (hereinafter 'B&D') showed that monoalcohols (C_mH_nOH) are particularly likely to crystallize in high-symmetry space groups or with $Z' > 1$. Here we extend that work with the objective of elucidating as many useful rules as possible for aiding the prediction of the crystal packing of both monoalcohols and dialcohols [$C_mH_n(OH)_2$]. These systems were chosen because they are relatively simple and B&D have already had some success: if further rules cannot be discerned here, there can be little hope for more complex molecules.

An additional motive for this work was to test a new program for visualizing crystal structures and analysing packing arrangements. The program *Mercury* will soon be distributed as part of the standard CSD release.

2. Experimental

2.1. Database searches

All searches were performed on Version 5.20 of the CSD (October 2000) with the programs *ConQuest*1.1 and *QUEST*. Monoalcohols of the general formula C_mH_nOH were found subject to the following criteria: alkanols only (no phenols); no heavy-atom disorder; no structures flagged in the CSD as containing errors; no other molecules (*e.g.* solvates) in the structure; only one member of each CSD refcode family (chosen arbitrarily, but generally the first alphanumerically); all non-H atom coordinates available. The monoalcohols were divided into three subsets, depending on whether the hydroxyl group was primary, 1°, RCH$_2$OH, secondary, 2°, R_2CHOH, or tertiary, 3°, R_3COH.

Dialcohols of the general formula $C_mH_n(OH)_2$ were found subject to the same criteria as above plus the additional criteria: R factor < 5%; average e.s.d. for a C—C bond < 0.005 Å; hydroxyl H atom coordinates available. The additional criteria were possible because the CSD contains more dialcohols than monoalcohols, so we could afford to exert tighter control over experimental precision. (Having said that, it is possible that the use of more stringent criteria for the dialcohols has resulted in a bias, *e.g.* against structures with $Z' > 1$, with high molecular weight, or which are well determined except for the positions of disordered hydroxyl hydrogen atoms. We are grateful to a referee and the co-editor for pointing this out.) The dialcohols were divided into six subsets, depending on the nature (primary, secondary or tertiary) of each of the hydroxyl groups (hence primary–primary, 1°–1°; primary–secondary, 1°–2° *etc.*).

An additional search was performed, using the same criteria as in the dialcohol searches, to find primary monoalcohols containing at least one other O atom, *i.e.* molecules of the general formula $C_mH_nO_p$, where $p > 1$ and exactly one alkanol OH is present. A list of the CSD 'refcodes' of these and all other structures used in this study, together with the hydrogen-bond motifs they form, has been deposited as supplementary material.[1]

2.2. Mercury[2]

Analysis of the alcohol packing arrangements was performed with a pre-release version of a new program, *Mercury*, which owes much in its philosophy to an earlier program, *RPLUTO* (Motherwell & Shields, 2000). *Mercury* is an application for visualizing and analysing small-molecule crystal structures. It is intended for release with the CSD system in 2001, at which time it will be supported on Unix and Windows operating systems. The functionality it offers includes the location and display of intermolecular and/or intramolecular hydrogen bonds and other non-bonded contacts, and the ability to build a network of contacts using the methodology of Motherwell & Shields (2000).

2.3. Identification of hydrogen bonds

No fixed distance criteria were used to define hydrogen bonds. Rather, the environment of each hydroxyl group was studied in order to find the closest contacts it formed, relative to the sum of the van der Waals (vdw) radii of the atoms involved (Bondi, 1964). In the case of $OH\cdots O$ hydrogen bonding, the contact distances were, of course, almost invariably shorter than the sum of vdw radii. Our methodology, however, allowed us to find contacts longer than the sum of vdw radii if no shorter ones existed, in accordance with strictures against the use of fixed distance criteria for

---

[1] Supplementary data for this paper are available from the IUCr electronic archives (Reference: BK0097). Services for accessing these data are described at the back of the journal.

[2] See http://www.ccdc.cam.ac.uk/prods/mercury.

hydrogen bonds (Jeffrey & Saenger, 1991). In a very small minority of cases, the presence of a hydrogen bond was inferred from a short O···O distance even though the reported position of the hydroxyl hydrogen was inconsistent with this conclusion. For example, GUHBIU [2-(1-hydroxyprop-2-yl)-5-methylcyclohexanol; Korner *et al.*, 2000] contains an intermolecular O···O contact of 2.653 Å, which was held as indicative of a hydrogen bond even though the reported H-atom position implies a hydrogen-bond angle (O—H···O) of only 60°. To this extent, our approach for identifying hydrogen bonds might be criticized as judgmental, but the number of controversial cases is so small that the overall conclusions of our study cannot be significantly affected.

3. Results

3.1. Space group and *Z*′ statistics for the CSD

Space group and *Z*′ statistics for the entire CSD are required as a benchmark for interpreting corresponding statistics for the monoalcohol and dialcohol subsets. The 15 most common space groups in the CSD are 1 ($P1$), 2 ($P\bar{1}$), 4 ($P2_1$), 5 ($C2$), 9 (Cc), 11 ($P2_1/m$), 13 ($P2/c$), 14 ($P2_1/c$), 15 ($C2/c$), ($C2/c$), 19 ($P2_12_12_1$), 29 ($Pca2_1$), 33 ($Pna2_1$), 60 ($Pbcn$), 61 ($Pbca$), 62 ($Pnma$); see Cole (2000) and also Mighell *et al.* (1983). 91.0% (200 659) of entries in the CSD belong to one of these space groups. 9.0% (19735) crystallize in other space groups. These percentage figures neglect 4006 CSD entries for which space groups have not been unambiguously assigned.

91.5% (198 894) of CSD entries have Z' values of 1 or less. 8.5% (18 409) have $Z' > 1$. These percentage figures neglect 7097 entries whose Z' values are unknown.

3.2. Basic types of hydrogen-bond aggregates for monoalcohols

In agreement with the findings of B&D, the monoalcohol crystal structures in our study fall into four basic types: those with infinite ···OH···OH··· hydrogen-bonding chains; those containing rings of ···OH···OH··· hydrogen bonds; those containing isolated OH···O hydrogen bonds (called 'dimers' by B&D); and structures in which no OH···O hydrogen bonds are formed at all. Of 144 monoalcohols, all but two fall into one of these categories. The exceptions are WEYSUO [2-(9-fluorenyl)cyclohexanol; Rieger *et al.*, 1994] and NAMLAO (2-buta-1,3-diynyl-2-adamantanol; Keller *et al.*, 1995), both of which form finite hydrogen-bond chains. No monoalcohol structure contains more than one of the basic motifs (*i.e.* structures containing rings *and* chains or rings *and* isolated OH···O hydrogen bonds *etc.* are not found), except that WEYTAV [9-(2-hydroxy-1-phenylethyl)fluorene; Rieger *et al.*, 1994] forms a curious structure containing a ring of four OH···O hydrogen bonds with two additional OH···O hydrogen bonds above and below the plane of the ring at diagonally opposite corners.

In structures containing no OH···O hydrogen bonds, it is almost invariably the case that the hydroxyl hydrogen points towards a neighbouring π system and/or the hydroxyl oxygen accepts one or more CH···O interaction(s) (Desiraju & Steiner, 1999). There are only three reasonably convincing exceptions to this generalization, *viz.* CIPFEM, epi-β-amyrin (Reyes *et al.*, 1983); HOBPET, (1*Z*)-3-phenyl-1-(2′,6′,6′-trimethylcyclohex-1′-enyl)-1-buten-3-ol (Mori *et al.*, 1998); JAGJOQ, 14-taraxen-3α-ol (Chuah *et al.*, 1998). Few of the intermolecular CH···O and OH···π interactions are shorter than the sum of the van der Waals radii of carbon and oxygen, C···O distances typically being in the range 3.3–3.7 Å (intramolecular contacts are often shorter). The fact that these interactions are almost invariably present when OH···O hydrogen bonds are absent nevertheless suggests that, weak though they are, they must play a significant stabilizing role (Braga *et al.*, 1998). A counter-argument is that monoalcohols forming no OH···O hydrogen bonds are often highly sterically crowded, typically with aromatic rings, and the hydroxyl group often has little choice but to form CH···O and OH···π interactions.

Table 1
Frequency of occurrence of four basic packing motifs in 1°, 2° and 3° monoalcohols.

	Chains	Rings	Isolated OH···O	No OH···O	Other
1°	11	2	1	1	0
	73%	*13%*	*7%*	*7%*	*0%*
2°	19	19	5	6	1
	38%	*38%*	*10%*	*12%*	*2%*
3°	12	27	12	27	1
	15%	*34%*	*15%*	*34%*	*1%*
All	42	48	18	34	2
	29%	*33%*	*13%*	*24%*	*1%*

3.3. Steric effects on monoalcohol structures

Full utilization of hydrogen-bond donor and acceptor capacity in monoalcohols requires that the hydroxyl group both donates and accepts, which in turn should lead to a predisposition for chain and/or ring motifs. Furthermore, it has long been known that a hydroxyl oxygen that donates its hydrogen in a hydrogen bond becomes a better hydrogen-bond acceptor. Therefore, hydrogen bonds in a chain or ring are stronger than an isolated hydrogen bond of the same type (the so-called 'cooperative effect'; Del Bene & Pople, 1970; Ceccarelli *et al.*, 1981). This suggests that monoalcohols should preferentially form chains or rings. In fact, this is true only for 1° and 2° monoalcohols. Thus, Table 1 shows that the percentage of structures containing chains decreases dramatically along the series 1° (73%), 2° (38%), 3° (15%). Rings are most common for 2° monoalcohols (1° 13%, 2° 38%, 3° 34%). Structures containing isolated or no OH···O hydrogen bonds are most common for 3° monoalcohols (1° 13%, 2° 22%, 3° 49%). The obvious explanation is that chains form if steric factors allow (Braga *et al.*, 1997); rings are the next most preferred and isolated OH···O hydrogen bonds, or structures with no OH···O hydrogen bonds at all, which only form if steric hindrance prevents the formation of chains or rings.

The significance of steric effects is confirmed by the results in Table 2, which gives the frequency of occurrence of the four

Table 2
Occurrence of monoalcohol packing motifs as a function of the sum of substituents on β-carbon(s).

		Sum of substituents on β-carbon(s)									
		0	1	2	3	4	5	6	7	8	Av. (S.E.)
1°	Chains	1	3	7	0						1.6 (0.2)
	Rings	0	0	1	1						2.5 (0.5)
	Isolated OH···O	0	0	0	1						3.0 (–)
	No OH···O	0	0	1	0						2.0 (–)
2°	Chains	0	0	6	6	5	1	1			3.2 (0.3)
	Rings	0	0	2	4	9	2	2			3.9 (0.3)
	Isolated OH···O	0	0	0	1	3	1	0			4.0 (0.3)
	No OH···O	0	0	1	0	5	0	0			3.7 (0.3)
3°	Chains	1	1	1	3	3	2	1	0	0	3.3 (0.5)
	Rings	0	0	3	8	7	5	4	0	0	4.0 (0.2)
	Isolated OH···O	0	0	0	0	1	5	5	1	0	5.5 (0.2)
	No OH···O	0	0	0	1	1	13	8	1	3	5.6 (0.2)
All	Chains	2	4	14	9	8	3	2	0	0	2.8 (0.2)
	Rings	0	0	6	13	16	7	6	0	0	3.9 (0.2)
	Isolated OH···O	0	0	0	2	4	6	5	1	0	4.9 (0.3)
	No OH···O	0	0	2	1	6	13	8	1	3	5.1 (0.2)

Table 3
Space group and Z' statistics for 1°, 2° and 3° monoalcohols.

	Space groups		Molecules per asymmetric unit	
	Top 15	Other	$Z' = 1$	$Z' > 1$
1°	12	3	7	8
	80%	*20%*	*47%*	*53%*
2°	34	16	26	24
	68%	*32%*	*52%*	*48%*
3°	64	15	44	35
	81%	*19%*	*56%*	*44%*
3° alcohols forming chains or rings	27	12	13	26
	69%	*31%*	*33%*	*67%*
3° alcohols forming isolated OH···O	10	2	6	6
	83%	*17%*	*50%*	*50%*
3° alcohols forming no OH···O	26	1	25	2
	96%	*4%*	*93%*	*7%*

basic packing modes as a function of the quantity *SSBC*, the *sum of substituents on β-C atoms*. The *SSBC* value for a given alcohol is calculated by summing the number of non-H atoms (excluding the α-carbon) which are covalently bonded to the β-carbon atom(s) in the molecule. Thus, a β-carbon of the type $-CH_3$ will contribute zero to the *SSBC* value; $-CH_2R$ (R not hydrogen) will contribute one, and other types of β-C atoms will contribute as follows: $-CHR_2$ two, $-CR_3$ three, $-C(H){=}R$ one, $-C(R){=}R$ two, and $-C{\equiv}R$ one. For example, the molecule $(CH_3CH_2)_2CHOH$ has an *SSBC* count of two, being one for each of the two β-C atoms of the type $-CH_2R$. As another example, $PhCH_2OH$ has $SSBC = 2$, since the only β-carbon is in a phenyl ring and hence of the type $-C(R){=}R$.

Table 2 shows that the average *SSBC* value generally increases along the series: chains, rings, isolated OH···O, no OH···O. The only 1° monoalcohols not forming chains have $SSBC = 2$ or 3, *i.e.* the β-carbon is secondary or tertiary [and both of the molecules with secondary β-C atoms have very bulky substituents on this carbon, *viz.* TELZIT: 2-(7,9-diphenylcyclopenta(a)acenaphthadien-6*b*-yl)-2-phenylethanol; Repo *et al.*, 1996; and WEYTAV: 9-(2-hydroxy-1-phenylethyl)fluorene, Rieger *et al.*, 1994]. Amongst 2° monoalcohols, 18 out of 20 structures with $SSBC < 4$ form chains or rings, whereas the four types of packing arrangements are much more evenly split when *SSBC* is four or greater. The outlier 2° monoalcohol forming no OH···O hydrogen bonds but with *SSBC* equal to only two is the extremely hindered CIYSIM (1-hydroxyethylhexahelicene; van Meerssche *et al.*, 1984). For 3° monoalcohols, rings and (to a smaller extent) chains are predominant when $SSBC < 5$ (27 of 30 structures), whereas isolated or no OH···O hydrogen bonds are more common otherwise (36 out of 48 structures).

3.4. Space groups and Z' statistics for monoalcohols

This topic has already been well addressed by B&D and Gavezzotti & Filippini (1994); our work adds a few details. Results are summarized in Tables 3 and 4. As expected from the earlier work, unusual space groups (more specifically, trigonal and tetragonal groups) and Z' values exceeding one are far more common than in the CSD as a whole (see §3.1). This tendency is not shown, however, by 3° monoalcohols forming no OH···O hydrogen bonds, confirming the earlier workers' hypothesis that the unusual symmetries and Z' values are a direct result of the packing requirements for OH···O hydrogen-bond formation. Interestingly, even the formation of isolated OH···O hydrogen bonds in 3° monoalcohols often seems to require $Z' > 1$ (6 of 12 structures). Formation of chains or rings in 3° monoalcohol structures is almost precluded in low-symmetry space groups (orthorhombic or lower) with $Z' = 1$; there are only three such examples in our data set and two of these form ···OH···OH··· chains with very long O···O distances (*viz.* GUBJAO, $P2_12_12_1$, $Z' = 1$, O···O = 3.483 Å, 1-hydroxy-1,2,5,5,8a-pentamethylperhydronaphthalene; Drew *et al.*, 1999; MNPREY, $P2_12_12_1$, $Z' = 1$, O···O = 3.514 Å, 11β-methyl-19-nor-17α-pregn-4-en-20-yn-17β-ol; Rohrer *et al.*, 1976).

Table 4 shows that ring formation is not observed at all in our data set for structures crystallizing with $Z' = 1$ in one of the 15 most common space groups (all of which are triclinic, monoclinic or orthorhombic).

3.5. Geometries of hydrogen-bond aggregates in monoalcohols

The geometries of ···OH···OH··· chains in monoalcohols can be classified into four groups: planar zigzags; threefold helices; fourfold helices; and a broad category which we have

Table 4
Monoalcohol space group and Z' frequencies as a function of packing motif.

	Chains		Rings	
	$Z' = 1$	$Z' > 1$	$Z' = 1$	$Z' > 1$
Top 15 space groups	13	15	0	33
Other space groups	10	4	13	2

called ‘wave-like’ (not being aware of any more widely accepted term). Table 5 shows that helical chains are only common for 2° and 3° monoalcohols – implying that steric factors are relevant – and threefold helices are much more common than fourfold. Exact threefold helices are observed in trigonal space groups, but approximate helices are also seen frequently, formed by translation with $Z' = 3$. Some of the latter examples are sufficiently close to being exact that refinement in a higher symmetry space group might be successful. One of the four observed fourfold helices is exact, along a fourfold screw, while the other three are approximate, being formed by the arrangement of two crystallographically independent molecules along a 2_1 axis. Planar zigzag chains are usually formed by 2_1 axes or by translation with $Z' > 1$ (Fig. 1: VAFMUK, 9-anthrylmethanol; Sweeting & Rheingold, 1988). The geometry we have called ‘wave-like’, as it refers to monoalcohols, involves two molecules *A* and *B* arranged along a glide to give a chain of type $\cdots A\cdots B\cdots A'\cdots B'\cdots$, where *A* and A' (and *B* and B') are related to each other by the glide, and *A* and *B* are either crystallographically independent or related to each other by some other symmetry operation of the space group. The resulting ···O···O···O···O··· torsion angles have the sequence of signs + + − − and the motif can be regarded as a zigzag chain with the points of the zigzag pushed alternately in and out of the plane (Fig. 2: ETANOL, ethanol; Jonsson, 1976).

Virtually all ···OH···OH··· rings in our sample are 4-rings (*i.e.* involve four OH···O hydrogen bonds; 42 4-rings compared with three 3-rings and three 6-rings). The 4-rings are equally divided in geometry between planar and puckered, with the latter becoming more common along the sequence 1°, 2°, 3° (Table 5). All three 6-rings adopt a chair-like geometry. The 4-rings are formed in tetragonal space groups or in space groups of lower symmetry if there are two or four crystallographically independent molecules.

There has been some discussion in the literature regarding the nature of isolated OH···O hydrogen bonds in alcohols. B&D said that ‘closed dimers’ (non-disordered OH···HO contacts across a symmetry element) are highly improbable and that ‘open dimers’ (one dangling H-atom and one free acceptor) are favoured. Glidewell *et al.* (1996), however, concluded that the closed motif ‘is actually rather common in α-ferrocenyl alcohols’. A superficial examination of our data set suggests that both open and closed motifs occur (Table 5), but, given the possibilities of disorder and the uncertainty of H-atom positions, we draw no firm conclusion either way.

3.6. Basic types of hydrogen-bond aggregates for dialcohols

The vast majority of dialcohol crystal structures fall into one of three basic types: those with infinite ···OH···OH··· chains, those containing rings of ···OH···OH··· hydrogen bonds and those containing isolated OH···O hydrogen bonds (Table 6). Of 101 dialcohols, only five do not fall into one of these categories. One of the exceptions is the only dialcohol in our study that forms no OH···O hydrogen bonds (the highly hindered YIHJOO: α,α,α′,α′-tetraphenylcyclohexane-1,2-dimethanol; Ito *et al.*, 1994). The other four form structures containing either a mixture of the basic motifs (chains *and* isolated OH···O hydrogen bonds, one example), finite hydrogen-bond chains only (one example) or a mixture of finite and infinite hydrogen-bond chains (two examples). The virtual absence of dialcohols forming no OH···O hydrogen bonds, even when both hydroxyl groups are tertiary, is in marked contrast to the monoalcohols. It presumably indicates that

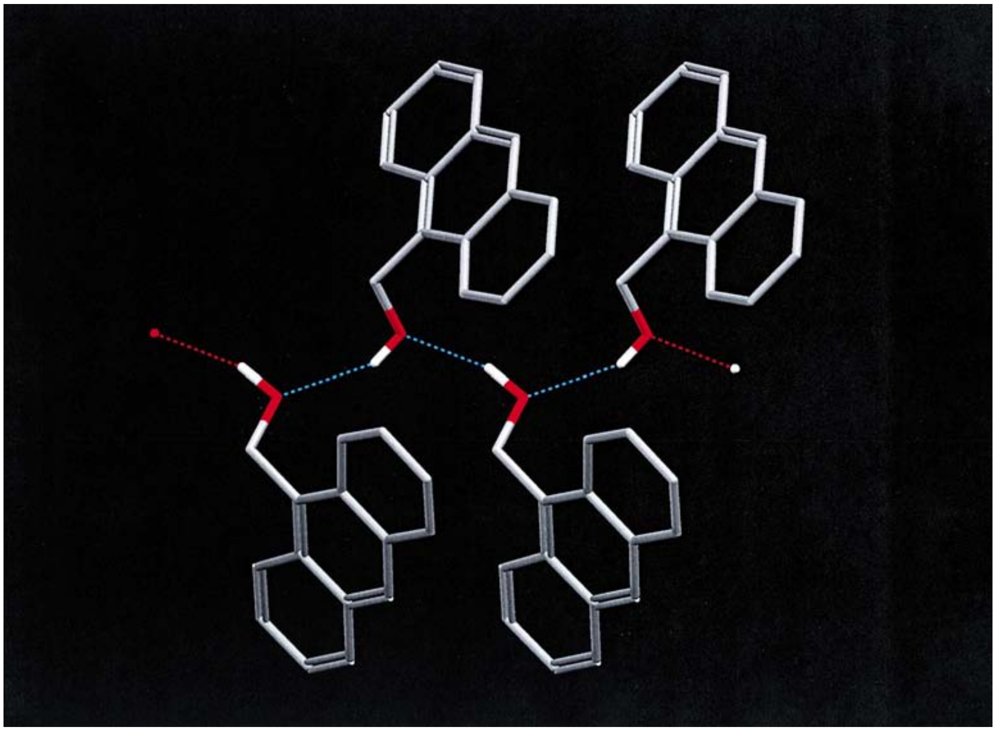


Figure 1
Planar zigzag ···OH···OH··· chain in VAFMUK (Sweeting & Rheingold, 1988).

Table 5
Geometries of chains, four-membered rings and dimers for 1°, 2° and 3° monoalcohols.

This table excludes five isolated OH···O hydrogen bonds for which no H-atom positions are available.

	Chains				4-Rings		Isolated OH···O	
	Zigzag	Wave	3-helix	4-helix	Planar	Puckered	Open	Closed
1°	4	4	2	1	2	0	0	0
2°	9	2	6	2	10	6	2	2
3°	3	2	6	1	9	15	5	4
All	16	8	14	4	21	21	7	6

dialcohols have more packing possibilities and hence more chance of finding an arrangement that satisfies hydrogen-bonding requirements.

Dialcohols have the possibility of forming intramolecular OH···O hydrogen bonds. This can occur in any of the three basic types of packing, *i.e.* some of the hydrogen bonds in an ···OH···OH··· chain or a hydrogen-bond ring may be intramolecular, or an isolated OH···O hydrogen bond might be intramolecular. All three of these situations occur, although comparatively rarely: 16 of the 60 dialcohol structures containing chains include intramolecular hydrogen bonds; corresponding figures for rings and isolated OH···O hydrogen bonds are two of 21 and three of 15, respectively.

If the two hydroxyl groups (A, B) of a dialcohol are not chemically equivalent (*i.e.* each is in a different substructural environment), then hydrogen bonds can, in principle, form between a hydroxyl group and its symmetry-generated equivalent (type $A\cdots A$ and $B\cdots B$) or between two non-equivalent hydroxyls (type $A\cdots B$). In fact, there is an extremely strong preference for the latter, irrespective of whether the hydrogen bond is isolated, in a ring or in a chain. Thus, in 38 dialcohol structures containing non-equivalent hydroxyl groups, 34 form hydrogen bonds of the type $A\cdots B$, and only four form hydrogen bonds of the type $A\cdots A$ and $B\cdots B$ (these figures exclude intramolecular hydrogen bonds, which must necessarily be of the type $A\cdots B$).

Infinite ···OH···OH··· chains in dialcohols may or may not contain dimers, by which is meant a pair of molecules linked by two hydrogen bonds, *e.g.* compare Fig. 3(*a*) (contains dimers; ZZZKPE01: *trans*-cyclohexane-1,2-diol; Sillanpaa *et al.*, 1984) and Fig. 3(*b*) (does not contain dimers; PUFJIJ: 2,5-hexanediol; Otten *et al.*, 1998). Overall, only a minority of chain structures contain dimers (10 of 60).

3.7. Steric effects on dialcohol structures

Table 6 shows that chains are overwhelmingly favoured for any dialcohol containing at least one primary hydroxyl group (25 chain arrangements in 27 structures). They are also predominant in 2°–2° and 2°–3° dialcohols, although to a smaller extent (27 chains, 10 rings, in a total of 46 structures). Only for 3°–3° dialcohols are structures with isolated OH···O hydrogen bonds as common as chain and ring structures. As with the monoalcohols, the trend is for chains to be formed if steric factors allow, with rings being the next most favoured and isolated OH···O hydrogen bonds only forming when steric hindrance is high. Compared with the monoalcohols, however, the proportion of chains is far higher and only a relatively small minority of structures (< 20%) crystallize without either chains or rings.

The importance of steric factors is confirmed by Table 7, which gives a more detailed breakdown of packing arrangements for the four most common types of dialcohols, *viz.* 1°–1°, 2°–2°, 2°–3°, 3°–3° (all five 1°–2° dialcohols form chains that involve neither dimers nor intramolecular hydrogen bonds; three 1°–3° dialcohols form chains of this type, one forms a chain containing intramolecular hydrogen bonds, the fifth forms

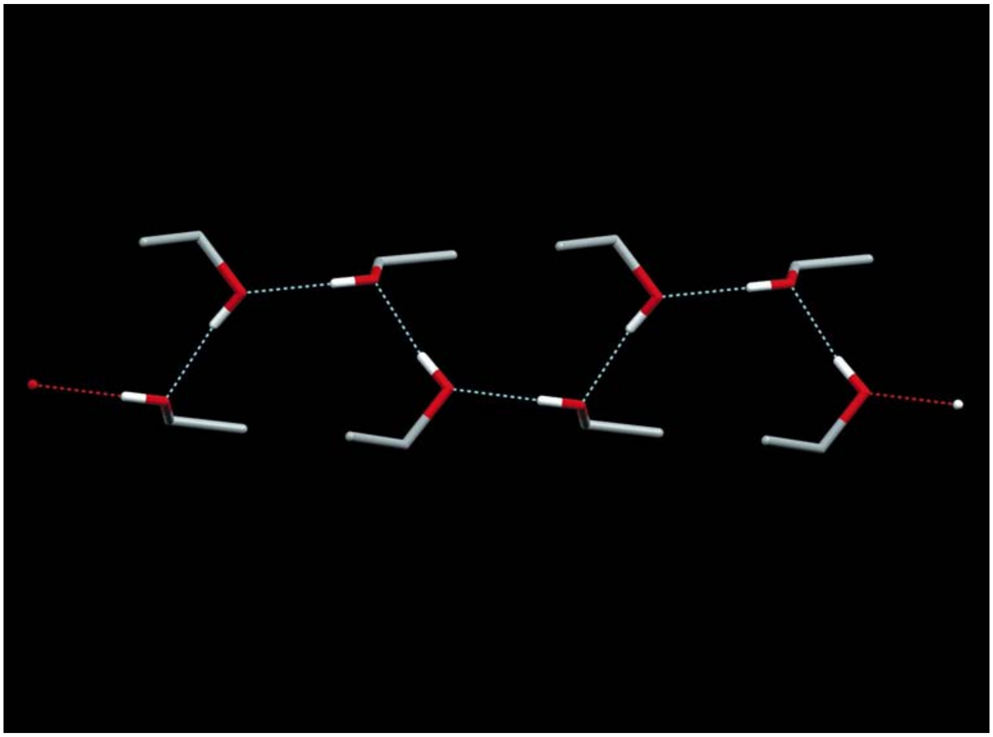


Figure 2
Wave-like ···OH···OH··· chain in ETANOL (Jonsson, 1976).

Table 6
Frequency of occurrence of four basic packing motifs in various types of dialcohols.

	Chains	Rings	Isolated OH···O	No OH···O	Other
1°–1°	16	1	0	0	0
	94%	*6%*	*0%*	*0%*	*0%*
1°–2°	5	0	0	0	0
	100%	*0%*	*0%*	*0%*	*0%*
1°–3°	4	1	0	0	0
	80%	*20%*	*0%*	*0%*	*0%*
2°–2°	17	7	1	0	3
	61%	*25%*	*4%*	*0%*	*11%*
2°–3°	10	3	4	0	1
	56%	*17%*	*22%*	*0%*	*6%*
3°–3°	8	9	10	1	0
	29%	*32%*	*36%*	*4%*	*0%*
All	60	21	15	1	4
	59%	*21%*	*15%*	*1%*	*4%*

a ring containing intermolecular hydrogen bonds only). The table also lists the *SSBC* values (*see above*) of all dialcohols in each category. *SSBC* is strongly correlated with packing arrangement. Thus, 40 of the 46 dialcohols in Table 7 with *SSBC* < 8 form intermolecular chains or rings, compared with only 2 of 24 with *SSBC* > 8. Clearly, steric hindrance leads not only to isolated OH···O hydrogen bonds, but also to packing arrangements that include intramolecular hydrogen bonds.

3.8. Space group and Z' statistics for dialcohols

Table 8 summarizes space group and Z' statistics for the dialcohols. The compounds show no abnormal preference for high-symmetry space groups or structures with $Z' > 1$, with the possible exception of 2°–2° and perhaps 3°–3° dialcohols. Of the dialcohols that have $Z' > 1$, three cannot be classified into any of the basic packing categories. Given that there are only four 'unclassified' dialcohols in total (Table 6), this figure is disproportionately high and presumably indicates that these three molecules have difficulty in packing (they are: CIZYOZ, perhydro-1,2-indenediol; Matoba *et al.*, 1984; HINZIN, 9,9-dimethyl-4-phenylbicyclo[4.4.0]decene-5,7-diol; Batey *et al.*, 1999; TETRDO01, *trans*-1,2,3,4-tetrahydro-2,3-naphthalenediol; Lloyd *et al.*, 1998).

3.9. Geometries of hydrogen-bond aggregates in dialcohols

Table 9 summarizes the geometries of intermolecular chains and rings in dialcohol crystal structures. As with monoalcohols, most of the chains can be assigned to one of four categories: planar zigzags, threefold helices, fourfold helices and wave-like chains. Helices are uncommon (six examples in all compared with 11 zigzags and 24 waves). Wave-like chains actually fall into two main groups: the type observed earlier for monoalcohols, generated by glide planes (§3.5); and a geometry generated by a screw axis which can best be described as a zigzag with flattened bottoms. Two slightly different examples are shown in Fig. 4 (FUNCIA, 6-hydroxypropyl-1,7,11,11-tetramethyltricyclo[8.4.0.0^{2,7}]tetradecan-5-ol; Manes *et al.*, 1988; HIYHAY, 1,11-undecanediol; Nakamura *et al.*, 1999).

Table 7
Occurrence of dialcohol packing motifs as a function of sum of substituents on β C atoms.

Table excludes three 2°–2° and one 2°–3° dialcohol structures that cannot be assigned to any of the motifs listed.

	Motif	Total No.	Sum of substituents on β C atoms	Av. (S.E.)
Motifs involving only intermolecular hydrogen bonds				
1°–1°	Chains (no dimers)	10	2, 2, 2, 2, 2, 2, 2, 2, 4, 4	2.4 (0.3)
	Chains (with dimers)	4	2, 4, 4, 6	4.0 (0.8)
	Rings	1	6	6.0 (–)
	Isolated OH···O	0	–	–
	No OH···O	0	–	–
2°–2°	Chains (no dimers)	8	2, 4, 4, 6, 6, 6, 7, 8	5.4 (0.7)
	Chains (with dimers)	5	6, 6, 7, 8, 8	7.0 (0.4)
	Rings	7	4, 6, 6, 8, 8, 8, 9	7.0 (0.7)
	Isolated OH···O	1	9	9.0 (–)
	No OH···O	0	–	–
2°–3°	Chains (no dimers)	5	5, 5, 5, 7, 8	6.0 (0.6)
	Chains (with dimers)	1	8	8.0 (–)
	Rings	3	6,7,8	7.0 (0.6)
	Isolated OH···O	4	5, 9, 11, 11	9.0 (1.4)
	No OH···O	0	–	–
3°–3°	Chains (no dimers)	3	6, 6, 8	6.7 (0.7)
	Chains (with dimers)	0	–	–
	Rings	7	2, 4, 6, 6, 8, 8, 12	6.6 (1.2)
	Isolated OH···O	7	4, 8, 10, 11, 14, 14, 16	11.0 (1.6)
	No OH···O	1	12	12.0 (–)
Motifs involving intramolecular hydrogen bonds				
1°–1°	Chains	2	5, 6	5.5 (0.5)
	Rings	0	–	–
	Isolated OH···O	0	–	–
2°–2°	Chains	4	6, 8, 8, 8	7.5 (0.5)
	Rings	0	–	–
	Isolated OH···O	0	–	–
2°–3°	Chains	4	9, 9, 9, 11	9.5 (0.5)
	Rings	0	–	–
	Isolated OH···O	0	–	–
3°–3°	Chains	5	6, 10, 12, 14, 14	11.2 (1.5)
	Rings	2	12, 12	12.0 (–)
	Isolated OH···O	3	8, 10, 16	11.3 (2.4)

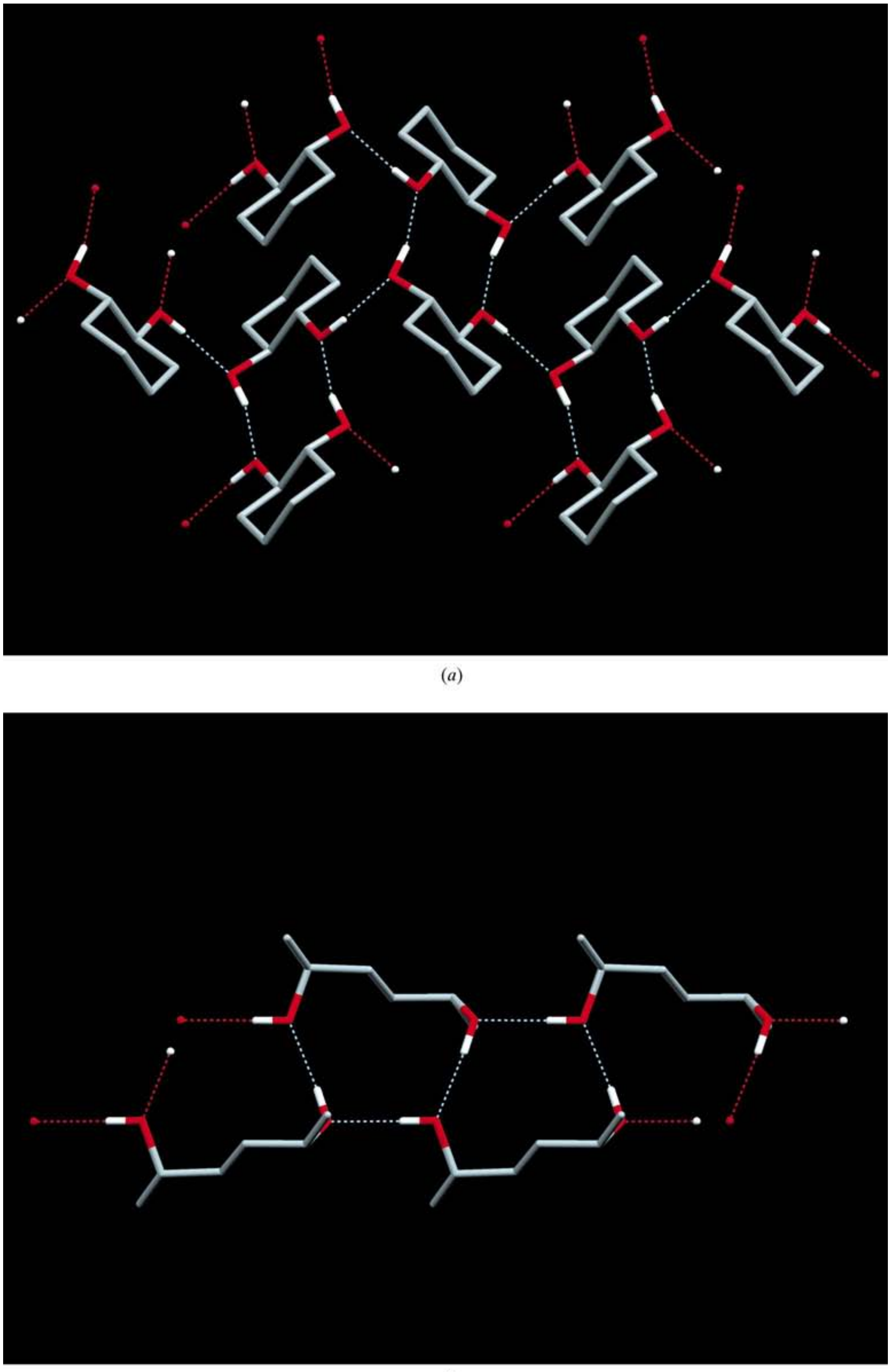
Virtually all the rings in dialcohol structures are 4-rings, *i.e.* contain four OH···O hydrogen bonds (20 of 21, the other is a 6-ring). The majority of these rings are planar, although puckering becomes more likely as steric effects increase.

3.10. Overall topology of hydrogen-bonding networks in dialcohol structures

An attempt was made to classify the overall topology of dialcohol networks by removing from the *Mercury* display all atoms other than those on the shortest path between the two hydroxyl groups and expanding the hydrogen-bonded network until an overall pattern emerged (see Batten & Robson, 1998, for a comprehensive review of network topologies). Fig. 5 shows example networks. In the first, the overall topology can be viewed as a network of 6-rings, four sides of each ring being hydrogen bonds, the other two being the covalently linked shortest paths between the hydroxyl groups

(CAVDEI, tetradecane-1,14-diol; Nakamura & Sato, 1999). The second is a network of 4- and 8-rings (BCHMOL, 5-*tert*-butylcyclohexane-1,3-dimethanol; van Koningsveld *et al.*, 1981).

One or two of the patterns thus found add insight to dialcohol packing (for example, the arrangement shown in Fig. 4*a* appears to be a common way in which dialcohols form infinite chains). Unfortunately, many overall topologies are impossible to represent in two dimensions and difficult to assimilate even with a three-dimensional graphics terminal. In general, we do not feel confident that they will aid the prediction of dialcohol packing, in contrast to the simpler rules identified in previous sections of this work. This is consistent with the view (Dunitz, 1999) that it is the short-range details that are important ('long-range periodicity is produced by directionally specific short-range interactions, nothing more').


(*a*)

(*b*)

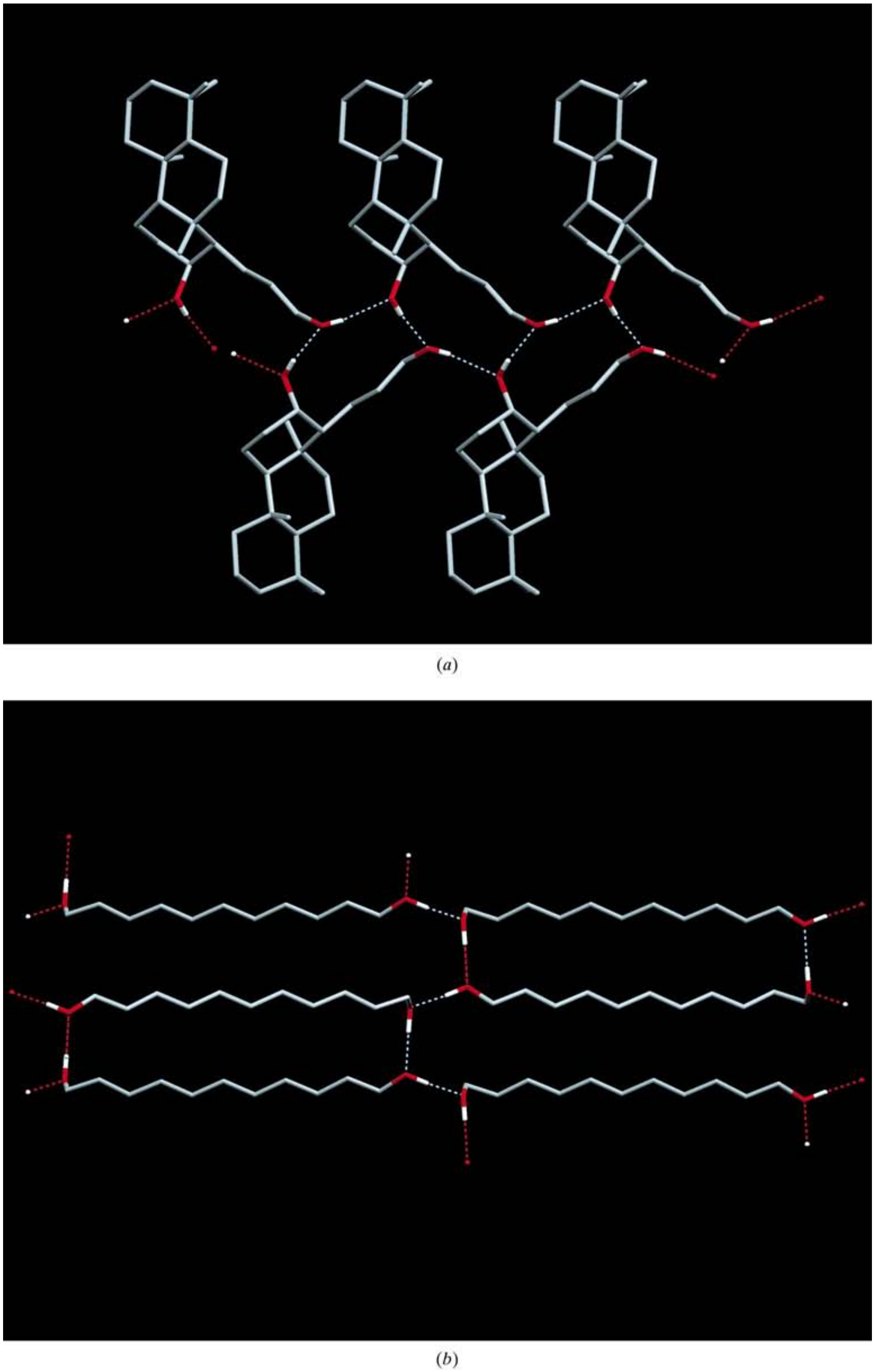
Figure 3
Infinite ···OH···OH··· chain in dialcohols, (*a*) containing dimers (ZZZKPE01; Sillanpaa *et al.*, 1984) and (*b*) not containing dimers (PUFJIJ; Otten *et al.*, 1998).

3.11. Packing of primary monoalcohols containing additional oxygen acceptors

This group of compounds was briefly examined in order to determine the extent to which the packing characteristics of simple monoalcohols, C_mH_nOH, are perturbed when one or more additional (non-hydroxyl) O atoms are present. A count was made of the numbers of different types of O atoms in each molecule, excluding the hydroxyl oxygen. Aromatic and ester (C—O—C) O atoms are also excluded because they are almost incapable of accepting hydrogen bonds (Boehm *et al.*, 1996; Nobeli *et al.*, 1997). The O atom in each structure which acts as the hydrogen-bond acceptor of the hydroxyl hydrogen was then identified. Details of these results have been deposited as supplementary data (see footnote **1**).

The analysis shows that hydrogen bonds between hydroxyl groups are rare in these structures, occurring in only four of the 20 examples. When they do occur, they form infinite ···OH···OH··· chains in the majority of cases (three out of four), as expected by analogy with the data in Table 1 on 1° monoalcohols. The presence of additional O atoms, however, represents a huge perturbation, since the hydroxyl hydrogen is normally donated to a non-hydroxyl oxygen (usually a carbonyl oxygen if one is present). Lessons learnt from our study of 'pure' monoalcohols, C_mH_nOH, are therefore largely irrelevant for these structures because different types of hydrogen bonds are present.

The space groups and Z' statistics of these compounds are not significantly different from those of the CSD as a whole.


(a)

(b)

Figure 4
Wave-like $\cdots$OH$\cdots$OH$\cdots$ chains in (*a*) FUNCIA (Manes *et al.*, 1988) and (*b*) HIYHAY (Nakamura *et al.*, 1999).

4. Conclusions

Rules governing the crystal packing of monoalcohols, C_mH_nOH, are:

(i) Virtually all (142 of 144) monoalcohols form one of four packing motifs, *viz.* infinite $\cdots$OH$\cdots$OH$\cdots$ chains, $\cdots$OH$\cdots$OH$\cdots$ rings, isolated OH$\cdots$O hydrogen bonds or no OH$\cdots$O hydrogen bonds.

(ii) If a monoalcohol structure contains no OH$\cdots$O hydrogen bonds, it will almost invariably be stabilized by CH$\cdots$O and/or OH$\cdots\pi$ interactions (only three structures found with no OH$\cdots$O, CH$\cdots$O or OH$\cdots\pi$ interactions).

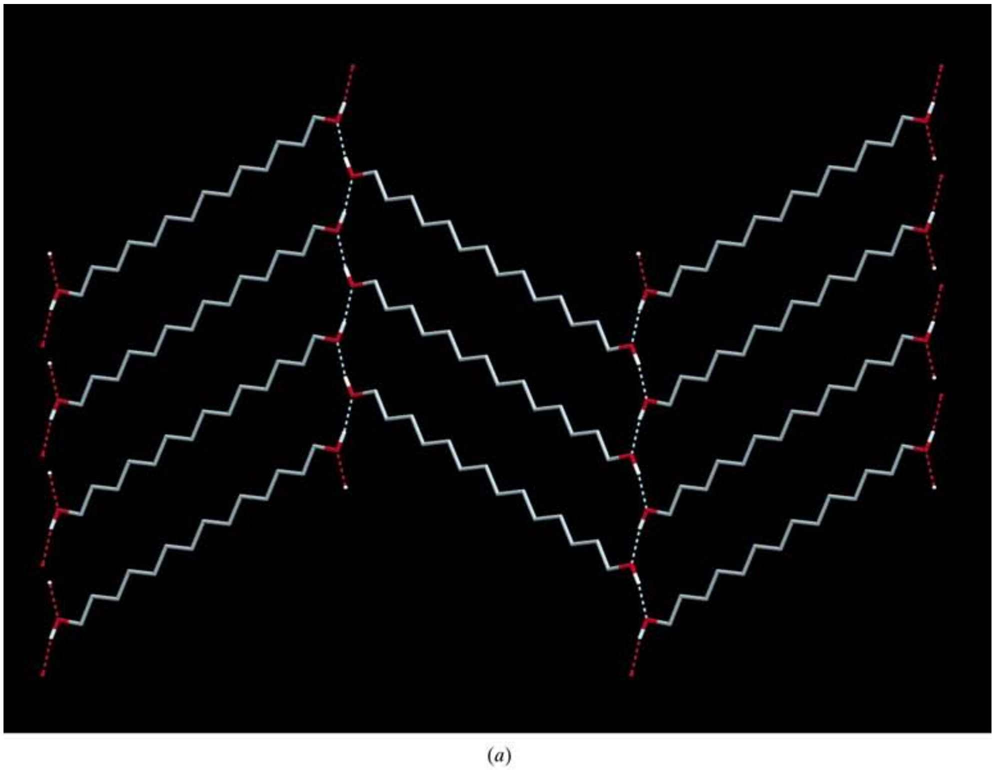
(iii) The occurrence of the basic motifs is strongly correlated with steric effects. Chains form if steric factors permit, rings are next most preferred and sterically hindered monoalcohols form isolated OH$\cdots$O hydrogen bonds, or no OH$\cdots$O hydrogen bonds at all. Thus:

(*a*) 1° monoalcohols always form chains (11 of 15) unless sterically hindered by tertiary substitution on the β carbon, or secondary substitution with very large substituents.

(*b*) 2° monoalcohols form all four of the basic motifs, but are mostly chains (19 of 50) or rings (19 of 50). Chains or rings are highly likely if *SSBC* < 4 (18 of 20).

(*c*) 3° monoalcohols usually form rings (27 of 79) or have no OH$\cdots$O hydrogen bonds at all

(27 of 79). Rings or chains are more likely if *SSBC* < 5 (27 of 30), whereas isolated/no OH···O hydrogen bonds are likely otherwise (36 of 48).



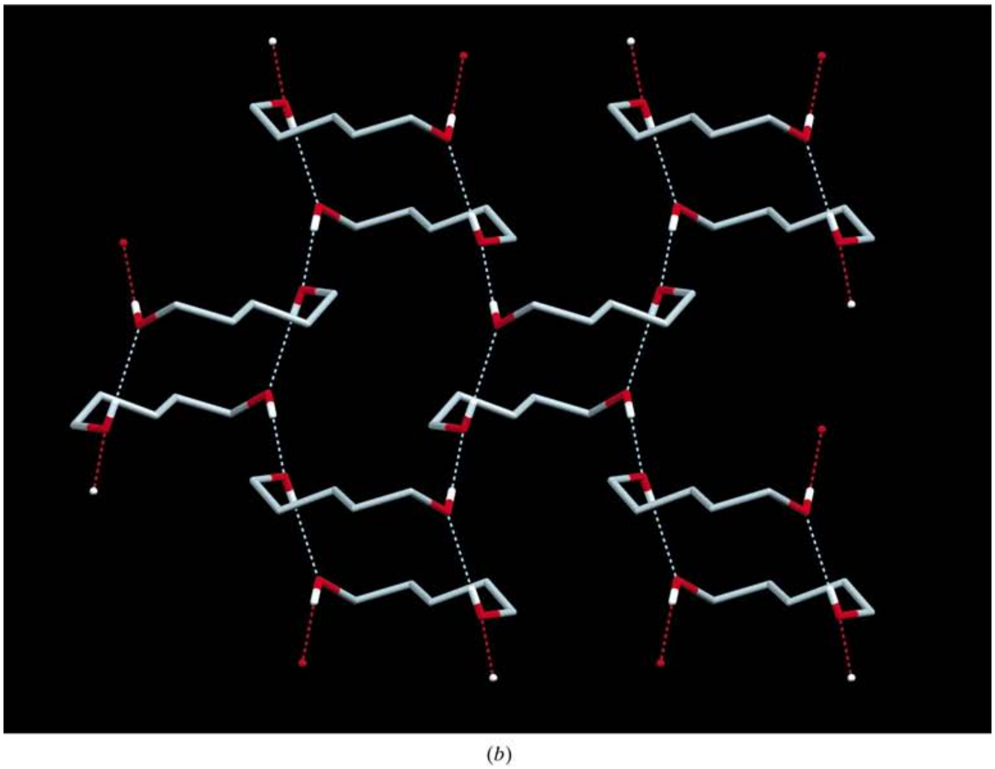


Figure 5
Extended hydrogen-bond networks in (*a*) CAVDEI (Nakamura & Sato, 1999) and (*b*) BCHMOL (van Koningsveld *et al.*, 1981).

(iv) Monoalcohols forming chains, rings or isolated OH···O hydrogen bonds have a strongly increased probability (compared with the rest of the CSD) of crystallizing in a high-symmetry (trigonal or tetragonal) space group or with $Z' > 1$. This tendency is not shown by monoalcohols that form no OH···O hydrogen bonds. In particular:

(*a*) It is very unlikely that a ring will occur if $Z' = 1$ and the space group is not high symmetry (no examples in 48 structures forming rings).

(*b*) It is very unlikely that a 3° monoalcohol structure will contain either a chain or a ring if the space group is not high symmetry and $Z' = 1$ (three examples only).

(v) Chains can be zigzag, wave-like or helical, the latter becoming somewhat more predominant as steric hindrance increases. Threefold helices are more common than fourfold (14 *versus* 4).

(vi) A large majority (42 of 48) of ···OH···OH··· rings are 4-rings (*i.e.* contain four OH···O hydrogen bonds). Considering all monoalcohols, such rings are equally likely to be planar or puckered. Puckering becomes more common as steric hindrance increases, *i.e.* for 3° monoalcohols.

Rules governing the crystal packing of dialcohols, $C_mH_n(OH)_2$, are:

(i) A large majority (96 of 101) of dialcohols form one of three packing motifs: infinite ···OH···OH··· chains; ···OH···OH··· rings; isolated OH···O hydrogen bonds. Each of these motifs can involve intramolecular OH···O hydrogen bonds, but in practice this occurs rather infrequently (in 21 of the 101 structures).

(ii) The occurrence of the basic motifs is strongly correlated with steric effects. Intermolecular chains form if steric factors permit, intermolecular rings are next most preferred

Table 8
Space group and Z' statistics for different types of dialcohols.

	Space groups		Molecules per au	
	Top 15	Other	$Z' = 1$	$Z' > 1$
1°–1°	16	1	17	0
	94%	*6%*	*100%*	*0%*
1°–2°	5	0	5	0
	100%	*0%*	*100%*	*0%*
1°–3°	4	1	4	1
	80%	*20%*	*80%*	*20%*
2°–2°	26	2	20	8
	93%	*7%*	*71%*	*29%*
2°–3°	15	3	17	1
	83%	*17%*	*94%*	*6%*
3°–3°	25	3	23	5
	89%	*11%*	*82%*	*18%*

Table 9
Geometries of chains and four-membered rings for different types of dialcohols.

	Chains (not involving intramolecular hydrogen bonds)					4-Rings	
	Zigzag	Wave	3-helix	4-helix	Other	Planar	Puckered
1°–1°	6	8	0	0	0	1	0
1°–2°	1	4	0	0	0	0	0
1°–3°	0	1	0	1	1	1	0
2°–2°	3	6	1	2	1	6	1
2°–3°	0	4	0	1	1	2	1
3°–3°	1	1	1	0	0	3	5
All	11	24	2	4	3	13	7

and only sterically hindered dialcohols form crystal structures involving intramolecular hydrogen bonds, isolated OH···O hydrogen bonds, or no OH···O hydrogen bonds at all. Thus:

(*a*) Dialcohols with at least one primary hydroxyl group overwhelmingly favour chains (25 of 27), of which the large majority (22) involve only intermolecular hydrogen bonds.

(*b*) Dialcohols containing at least one secondary hydroxyl group but no primary OH's usually form chains (27 examples out of 46) or rings (10 examples).

(*c*) Chains, rings and isolated OH···O hydrogn bonds are approximately equally likely in dialcohols in which both hydroxyls are tertiary.

(*d*) 2°–2°, 2°–3° and 3°–3° dialcohols with $SSBC < 8$ almost always form intermolecular chains or rings (25 of 29). The same types of dialcohols strongly favour isolated OH···O hydrogen bonds; no OH···O hydrogen bonds or motifs involving intramolecular hydrogen bonds if $SSBC > 8$ (22 of 24).

(iii) Intermolecular hydrogen bonds in the structures of dialcohol molecules containing topologically distinct hydroxyl groups (A, B) are almost invariably between the non-equivalent groups (*i.e.* $A\cdots B$ rather than $A\cdots A$ or $B\cdots B$; 34 of 38 structures). This is independent of whether the basic motif is a chain, ring or isolated OH···O hydrogen bond.

(iv) With the possible exception of 2°–2° and 3°–3° dialcohols, the compounds do not show an unusual predilection for crystallizing in high-symmetry space groups or with $Z' > 1$.

(v) Chains are usually zigzag or wave-like. Both threefold and fourfold helices are rare (two and four examples, respectively, in 44 structures forming intermolecular chains).

(vi) Virtually all (20 of 21) ···OH···OH··· rings are 4-rings (*i.e.* contain four OH···O hydrogen bonds). Planar rings are slightly favoured over puckered (13 *versus* 7), but puckering becomes more likely as steric hindrance increases.

Rules governing the crystal packing of primary monoalcohols containing at least one other O atom, $C_mH_n(O)_pCH_2OH$, are:

(i) Hydrogen bonds of the type OH···O=C will usually occur if the molecule contains at least one carbonyl group (seen in 10 of 14 examples).

(ii) Space group and Z' distributions are similar to those seen in the CSD as a whole.

5. Discussion

We have been successful in identifying rules governing the crystal packing of mono- and dialcohols. Moreover, the rules are based on simple quantities that can be rapidly determined from a trivial analysis of molecular connectivity. In this section we now consider whether the rules are of practical value.

In order to be useful, the rules must first add value to what is already achievable in crystal-structure prediction. Common experience suggests that using a state-of-the-art program to predict, say, the structure of a dialcohol will result in a rather large number of possible structures with similar calculated lattice energies. The experimental structure is likely to be amongst them, but cannot be distinguished (Mooij *et al.*, 1999; Lommerse *et al.*, 2000; van Eijck & Kroon, 2000). In this situation, if some of the hypothetical structures fail to match the empirical rules identified in this work, then they can be eliminated with confidence and the rules will have proved useful. Until the experiment is tried, however, we do not know whether this will be the case. If it is not (*i.e.* if all the putative structures obey the rules), then the only way our work can help is by restricting the search space that must be visited by the crystal-structure prediction program. This may allow more time to be spent on higher quality energy calculations which, in turn, may improve predictive ability. Then again, it may not, particularly if kinetic effects are important.

The next issue concerns transferability. Common empirical techniques such as molecular mechanics are successful because they are based on parameters that can be transferred from one situation to another. For example, the same stretching constant can be used for more or less any C—O single bond, irrespective of its molecular environment. In the case under discussion, it is regrettably clear that our rules are far from transferable. Even changing from monoalcohols to dialcohols alters the rules appreciably, and adding an extra type of oxygen acceptor renders them completely inapplicable. Our rules may be fine for the systems from which they were derived, but they are useless for anything else and therefore cannot assist in the solution of problems of real commercial interest.

This does not mean that the CSD itself is of no value in crystal-structure prediction; rather, it means that rules generated from the CSD must be tailored to the problem in hand. At first sight this seems impracticable and, indeed, this would be the case if rules had to be discerned by visual analysis of scores of crystal structures, as we have performed here. Fortunately, the architecture of *Mercury* should make it possible to automate a search for rules: the underlying C++ classes can be put together in a variety of ways to achieve this end. One can imagine the following methodology:

(i) Parse the structure of the molecule of interest to identify the chemical populations to which it belongs (there will normally be more than one – for example, a cyclopentenone with an oxime side chain belongs to the population *of all cyclopentenone derivatives* and the population of *all organic oximes*).

(ii) Find the relevant structures in the CSD and discern packing rules automatically.

(iii) Factor the rules into the crystal-structure prediction process.

The final issue concerns statistical significance and bias. Leusen *et al.* (1999) point out that scoring functions based on the CSD may be incapable of predicting anything novel. While true, this need not greatly concern us. We would be gratified if we could predict the majority of typical structures, never mind the few that are atypical. Bias in the CSD is, however, a problem. For example, the 15 primary monoalcohols used herein do not constitute a very diverse set of molecules. The answer lies in finding a good balance between the relevance of a sample and its size. As we increase the size of a sample by reducing the specificity of the substructure search from which it was generated, the diversity will tend to increase and biases will become less likely. On the other hand, the structures in the sample will, on average, become less similar to the molecule in which we are interested. Finding algorithms for optimizing this balance will be a major challenge in developing an effective methodology for using the CSD in crystal structure prediction.

Mercury was written by one of us (CFM). Significant contributions to the program were also made by Drs Jason Cole, Magnus Kessler and Jonathan Pearson, all present or former members of staff at the Cambridge Crystallographic Data Centre.

References

Allen, F. H. & Kennard, O. (1993). *Chem. Des. Autom. News*, **8**, 31–37.

Batey, R. A., Thadani, A. N. & Lough, A. J. (1999). *J. Am. Chem. Soc.* **121**, 450–451.

Batten, S. R. & Robson, R. (1998). *Angew. Chem. Int. Ed.* **37**, 1460–1494.

Boehm, H.-J., Brode, S., Hesse, U. & Klebe, G. (1996). *Chem. Eur. J.* **2**, 1509–1513.

Bondi, A. (1964). *J. Phys. Chem.* **68**, 441–451.

Braga, D., Grepioni, F. & Desiraju, G. R. (1998). *Chem. Rev.* **98**, 1375–1405.

Braga, D., Grepioni, F., Walther, D., Heubach, K., Schmidt, A., Imhof, W., Goerls, H. & Klettke, T. (1997). *Organometallics*, **16**, 4910–4919.

Brock, C. P. (1999). *Implications of Molecular and Materials Structure for New Technologies*, edited by J. A. K. Howard, F. H. Allen & G. P. Shields, pp. 251–262. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Brock, C. P. & Duncan, L. L. (1994). *Chem. Mater.* **6**, 1307–1312.

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

Chuah, C. H., Ng, S. W. & Yap, G. P. A. (1998). *Z. Kristallogr.* **213**, 493–494.

Cole, J. C. (2000). Private communication.

Del Bene, J. & Pople, J. A. (1970). *J. Chem. Phys.* **52**, 4858–4866.

Desiraju, G. R. & Steiner, T. (1999). *The Weak Hydrogen Bond.* Oxford University Press.

Drew, M. G. B., Harwood, L. M., Jahans, A., Robertson, J. & Swallow, S. (1999). *Synlett*, pp. 185–188.

Dunitz, J. D. (1999). *Implications of Molecular and Materials Structure for New Technologies*, edited by J. A. K. Howard, F. H. Allen & G. P. Shields, pp. 175–184. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Eijck, B. P. van & Kroon, J. (2000). *Acta Cryst.* B**56**, 535–542.

Gavezzotti, A. (1994). *Acc. Chem. Res.* **27**, 309–314.

Gavezzotti, A. & Filippini, G. (1994). *J. Phys. Chem.* **98**, 4831–4837.

Gavezzotti, A. & Filippini, G. (1998). *Chem. Commun.* pp. 287–294.

Glidewell, C., Klar, R. B., Lightfoot, P., Zakaria, C. M. & Ferguson, G. (1996). *Acta Cryst.* B**52**, 110–121.

Hofmann, D. W. M. & Lengauer, T. (1997). *Acta Cryst.* A**53**, 225–235.

Ito, Y. N., Ariza, X., Beck, A. K., Bohac, A., Ganter, C., Gawley, R. E., Kuhnle, F. N. M., Tuleja, J., Wang, Y. M. & Seebach, D. (1994). *Helv. Chim. Acta*, **77**, 2071–2110.

Jeffrey, G. A. & Saenger, W. (1991). *Hydrogen Bonding in Biological Structures*, pp. 29–33. Berlin: Springer-Verlag.

Jonsson, P.-G. (1976). *Acta Cryst.* B**32**, 232–235.

Keller, M. B., Rzepa, H. S., White, A. J. P. & Williams, D. J. (1995). *First Electronic Conf. Trends in Org. Chem.*, p. 50, http://www.ch.ic.ac.uk/ectoc.

Koningsveld, H. van, Baas, J. M. A., van de Graaf, B. & Steenks, A. (1981). *Cryst. Struct. Commun.* **10**, 691–697.

Korner, F., Schurmann, M., Preut, H. & Kreiser, W. (2000). *Acta Cryst.* C**56**, 74–75.

Leiserowitz, L. (1976). *Acta Cryst.* B**32**, 775–802.

Leusen, F. J. J., Wilke, S., Verwer, P. & Engel, G. E. (1999). *Implications of Molecular and Materials Structure for New Technologies*, edited by J. A. K. Howard, F. H. Allen & G. P. Shields, pp. 303–314. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Lloyd, M. A., Patterson, G. E. & Brock, C. P. (1998). *Acta Cryst.* C**54**, 1486–1488.

Lommerse, J. P. M., Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Mooij, W. T. M., Price, S. L., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P. & Williams, D. E. (2000). *Acta Cryst.* B**56**, 697–714.

Manes, L. V., Crews, P., Kernan, M. R., Faulkner, D. J., Fronczek, F. R. & Gandour, R. D. (1988). *J. Org. Chem.* **53**, 570–575.

Matoba, Y., Ishii, Y. & Ogawa, M. (1984). *Acta Cryst.* C**40**, 1735–1736.

Meerssche, M. van, Declercq, J. P., Germain, G., Soubrier-Payen, B., Lindner, H. J. & Kitschke, H. J. (1984). *Bull. Soc. Chim. Belg.* **93**, 445–448.

Mighell, A. D., Himes, V. L. & Rodgers, J. R. (1983). *Acta Cryst.* A**39**, 737–740.

Mooij, W. T. M., van Eijck, B. P. & Kroon, J. (1999). *J. Phys. Chem. A*, **103**, 9883–9890.

Mori, H., Ikoma, K., Isoe, S., Kitaura, K. & Katsumura, S. (1998). *J. Org. Chem.* **63**, 8704–8718.

Motherwell, W. D. S. & Shields, G. P. (2000). *RPLUTO*, http://www.ccdc.cam.ac.uk/prods/rpluto.

Nakamura, N. & Sato, T. (1999). *Acta Cryst.* C**55**, 1687–1689.

Nakamura, N., Setodoi, S. & Ikeya, T. (1999). *Acta Cryst.* C**55**, 789–791.

Nobeli, I., Price, S. L., Lommerse, J. P. M. & Taylor, R. (1997). *J. Comput. Chem.* **18**, 2060–2074.

Otten, S., Froehlich, R. & Haufe, G. (1998). *Tetrahedron Asym.* **9**, 189–191.

Repo, T., Klinga, M., Leskele, M., Polamo, M. & Rieger, B. (1996). *Acta Cryst.* C**52**, 2910–2912.

Reyes, J., Hidalgo, C. & Fronczek, F. R. (1983). *Rev. Latinoam. Quim.* **14**, 67–70.

Rieger, B., Jany, G., Steimann, M. & Fawzi, R. (1994). *Z. Naturforsch. Teil B*, **49**, 451–456.

Rohrer, D. C., Hazel, J. P., Duax, W. L. & Zeelen, F. J. (1976). *Cryst. Struct. Commun.* **5**, 543–546.

Sillanpaa, R., Leskele, M. & Hiltunen, L. (1984). *Acta Chem. Scand. Ser. B*, **38**, 249–254.

Sweeting, L. M. & Rheingold, A. L. (1988). *J. Phys. Chem.* **92**, 5648–5655.