ChemComm

Pseudopolymorphism: occurrences of hydrogen bonding organic solvents in molecular crystals

Ashwini Nangia and Gautam R. Desiraju

School of Chemistry, University of Hyderabad, Hyderabad 500 046, India. E-mail: grdch@uohyd.ernet.in

Received (in Cambridge, UK) 15th December 1998, Accepted 24th February 1999

Multi-point recognition with strong and weak hydrogen bonds between solvent and solute molecules facilitates the retention of organic solvents in crystals

Pseudopolymorphism is the phenomenon wherein a compound is obtained in crystalline forms that differ in the nature or stoichiometry of included solvent molecules.[1] This subject has not been treated systematically though it is perceived to be of general importance, for example in the pharmaceutical industry.[2] In a supramolecular sense, pseudopolymorphs of a compound are different chemical systems and should be treated as such. Here, we examine the likelihood of several hydrogen bonding organic solvents in crystal structures and attempt to provide some rationalisation for these occurrences. We have not considered hydration. The presence of water in organic crystals is so widespread and the reasons for its inclusion so varied[1–3] that it would not be realistic or practicable to include it with the other solvents considered here.

The method of analysis uses the Cambridge Structural Database (CSD, version 5.15, April 1998, 181 309 entries).[4] Organic structures† containing solvent of crystallisation were retrieved and 20 common solvents found to occur at least 50 times. A few hits contain two solvents and these are included under both in Table 1. The raw frequencies of occurrence N are listed in Table 1, which also contains the numbers of fully ordered crystal structures in each case, N_{ord}.‡ Given that some solvents are used for recrystallisation more frequently than others, these N (or even N_{ord}) values cannot be taken to correspond directly to the abilities of the respective solvents to give solvated crystals. To perform an appropriate, though still approximate, correction for differences in solvent usage, all organic structures (solvated or otherwise) in the 1986 and 1996 issues of *Acta Crystallographica Section C* were examined. The number of times each solvent was used for recrystallisation alone or in combination with other solvents is given as N_{Acta} in Table 1. The quantity (N/N_{Acta}) for a particular solvent was then taken as a measure of its ability to be included in the crystal. This quantity was scaled appropriately to give the usage corrected occurrence (O_{corr}). A value of 1.00 for O_{corr} indicates a solvent that forms solvates to an extent commensurate with its usage. Values of O_{corr} much larger or much smaller than unity indicate solvents that have a greatly enhanced or a greatly reduced tendency to be included in crystals.

Table 1 Occurrences of 20 common solvents in organic crystal structures in decreasing order of O_{corr}

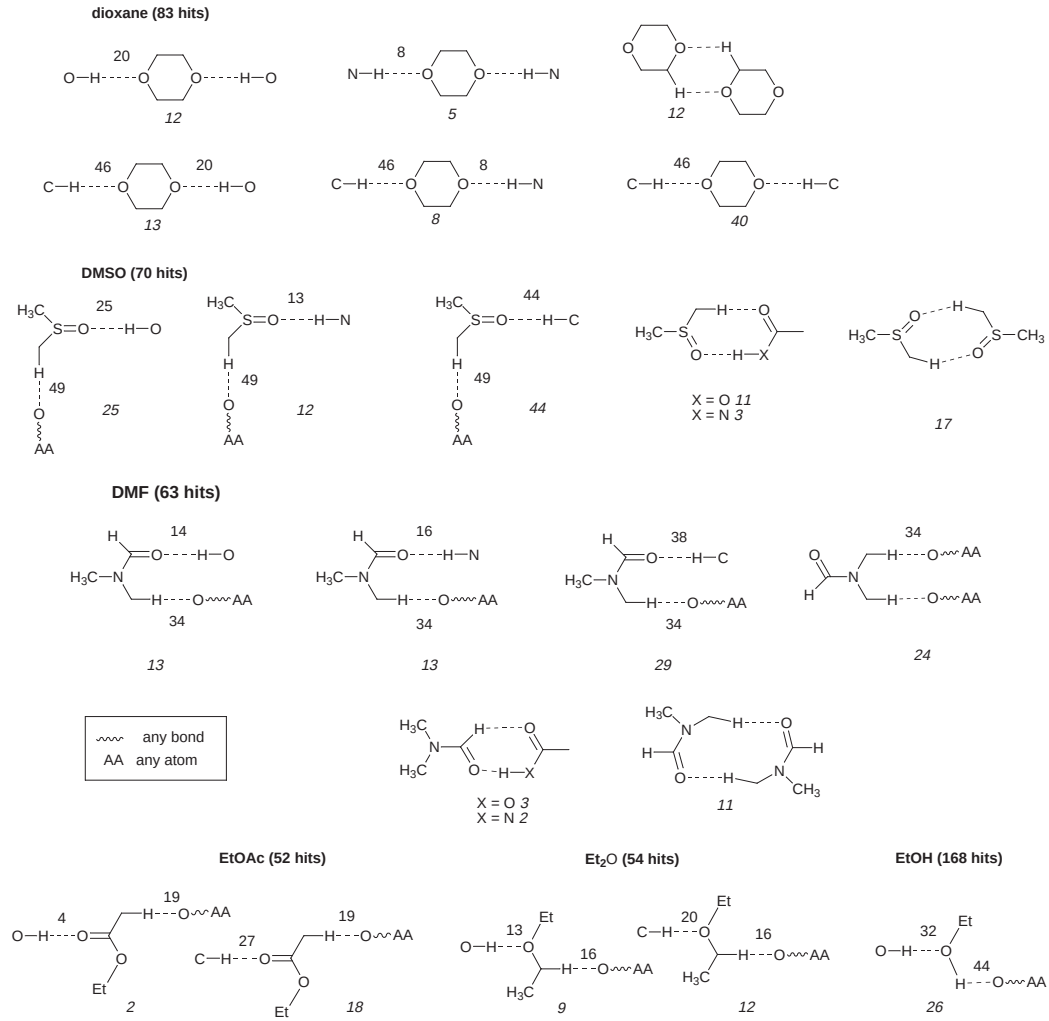
Solvent	N	N_{ord}	N_{Acta}	O_{corr}[a]
DMF	122	63	5	5.69
DMSO	142	70	7	4.73
Dioxane	161	83	8	4.70
p-Xylene	49	22	3	3.81
Benzene	471	271	32	3.43
THF	188	54	14	3.13
MeCN	396	181	31	2.98
AcOH	105	54	10	2.45
CCl_4	59	14	7	1.97
Toluene	266	65	36	1.72
CH_2Cl_2	455	183	77	1.38
$CHCl_3$	386	132	67	1.34
MeOH	795	433	158	1.17
Acetone	346	169	82	0.98
Pr^{i}OH	65	33	16	0.95
Cyclohexane	56	9	21	0.62
EtOAc	155	52	75	0.48
EtOH	406	168	206	0.46
Et_2O	144	54	96	0.35
Hexane	84	21	181	0.11
Total	4851	2131	1132	

[a] Calculated by the formula $(N/4851)/(N_{Acta}/1132) = 0.233\ (N/N_{Acta})$

The next step in the analysis focussed on the three solvents with the maximum values of O_{corr}. Let us consider the ordered dioxane solvates ($N_{ord} = 83$). Twenty hits contain an O–H···O hydrogen bond with specified metrics.§ In 12 of these cases, two O–H···O bonds approach a dioxane molecule as shown in the first synthon in Scheme 1. Similarly, 5 out of the 8 N–H···O hits correspond to two-point recognition with N–H groups.¶ This behaviour extends to hydrogen bonding by C–H groups. Forty six out of the 83 dioxane solvates contain a C–H···O bond to the solvent. When all the two-point cases are enumerated (13 with C–H, O–H; 8 with C–H, N–H; 40 with C–H, C–H), the total (61) is in excess of 46, indicating that in some of the cases, multiple donor groups approach the acceptor O-atom in the dioxane molecule. The overall situation is clear—when dioxane is included in an organic crystal, it is always hydrogen bonded and more importantly, to more than one donor. Generally, these donor groups (O–H, N–H, C–H) originate from different solute molecules though dioxane itself was noted to be the C–H donor in some cases (12 hits).

Multi-point recognition *via* strong and weak hydrogen bonding is also the favoured mode of association for DMSO (dimethyl sulfoxide) and DMF (dimethylformamide), both of which have a very high probability of being included in crystals. Whether the S=O group in DMSO is hydrogen bonded to an O–H, N–H or C–H group, the adjacent methyl group is invariably C–H···O bonded (25 out of 25 for O–H, 12 out of 13 for N–H, 44 out of 44 for C–H). It is interesting to note that all 44 hits that contain a C–H···O bond to DMSO as acceptor also contain a C–H···O bond from DMSO as donor. Like dioxane, DMSO shows a tendency to dimerise with itself (17 hits) or with acids (11 hits) and amides (3 hits). Many of these dimers are parts of extended hydrogen bonded networks and the various synthons shown in the scheme are not of an exclusive nature; any particular crystal structure usually contains more than one of the arrangements shown in Scheme 1. The situation with DMF is very similar. Multi-point recognition *via* C–H···O bonding from a methyl group of the solvent accounts for the majority of hits (13 out of 14 for O–H, 13 out of 16 for N–H and 29 out of 38 for C–H donors to DMF) while in 24 hits both methyl groups act as donors. As in DMSO, there is a tendency to form dimers and further these dimers extend into the general hydrogen bonding scheme. This tendency of dioxane, DMSO and DMF to self-associate as well as to form interactions with the solute improves the chances of solvent inclusion in crystals.

The final step of the analysis deals with solvates formed by EtOAc, EtOH and Et_2O. These three solvents are marked by an unusually reduced likelihood of being incorporated into solvate structures. The failure of these commonly-used solvents to be



Scheme 1 Multi-point supramolecular synthons formed by the frequently included solvents dioxane, DMSO and DMF. The numbers of strong and weak hydrogen bonds among the selected hits are given near the dotted lines. The italic numbers given below each synthon represent the incidence of two-point recognition. Notice the wide prevalence of C–H···O hydrogen bonding. The reduced incorporation of the solvents EtOH, Et_2O and EtOAc is reflected in the corresponding numbers for these solvents.

thus included is accounted for in terms of their inability to effectively participate in multi-point hydrogen bonded recognition schemes. Though EtOAc has a good acceptor group, the donors are largely unactivated. Et_2O has some acceptor ability but very poor donating capability. Unlike dioxane, neither of these solvents can act as multiple acceptors. To the extent, however, that the donors in EtOAc are more activated than in Et_2O, it is incorporated in crystals better. EtOH has a very good donor and a moderate acceptor group. However, it is generally included only in alcohol and acid structures and that too as part of a cooperative network, in which capacity it is also expendable. The numerical details in Scheme 1 bear out these qualitative generalisations. Of course, there must also be other reasons for the incorporation of organic solvents in crystals. Solvents like benzene, *p*-xylene and CCl_4 are included easily but generally in rigid framework structures where they are of the correct size to act as guests in a host lattice. Nonetheless, multipoint hydrogen bonding is clearly a dominant factor that governs the inclusion of the large number of solvents that are capable of exhibiting this phenomenon.

To conclude, why is multi-point recognition of the solvent so important in the formation of pseudopolymorphic structures? Crystallisation begins with solute–solvent aggregates that contain solute–solute, solute–solvent and solvent–solvent interactions. The entropic gain in eliminating solvent molecules from these aggregates into the bulk solution, and the simultaneous enthalpic gain in forming stable solute species that contain robust supramolecular synthons, provides an adequate driving force for nucleation and crystallisation with the result that most (85%) organic crystals are unsolvated. However, when solvent molecules are attached to solute molecules in a multi-point manner *via* either strong (O/N–H···O) or weak (C–H···O) hydrogen bonds, the extrusion of solvent from the aggregates into the bulk may become sufficiently disadvantageous from an enthalpic viewpoint with the result that the solvent remains an integral part of the nucleating crystal. The formation of a solvated crystal may therefore be likened to an interruption of the sequence of events that accompany the 'normal' crystallisation of an unsolvated form.

We thank the DST (SP/S1/G19/94) for financial assistance and Dr J. A. R. P. Sarma for discussions.

Notes and references

† By the term 'organic' we mean non-metal-atom-containing (screen 57 in the CSD). This group contains 84 562 entries.

‡ CSD screens 33 (error-free), 35 (no disorder), 85 (chemical/crystallographic connectivity match), 88 (*R*-factor ≤0.10) and 153 (atom coordinates present) were applied. The H-atoms are defined in these structures. This sub-group of organic structures contains 56 492 entries.

§ Hydrogen bond metrics: O–H···O, N–H···O, $1.5< d < 2.2$ Å, $140 < \theta < 180°$; C–H···O, $2.0 < d < 3.0$ Å, $110 < \theta < 180°$. All H-atom positions were neutron-normalised.

¶ No hits were found for two-point O–H, N–H recognition.

1 T. L. Threlfall, *Analyst*, 1995, **120**, 2435.
2 S. R. Byrn, *Solid-State Chemistry of Drugs*, Academic Press, New York, 1982.
3 G. R. Desiraju, *J. Chem. Soc., Chem. Commun.*, 1991, 426.
4 F. H. Allen, J. E. Davies, J. J. Galloy, O. Johnson, O. Kennard, C. F. Macrae and D. G. Watson, *J. Chem. Inf. Comput. Sci.*,1991, **31**, 204.

Communication 8/09755K