

Paper

Supramolecular synthon competition in organic sulfonates: A CSD survey

Delia A. Haynes,^a James A. Chisholm,^{a,b} William Jones^a and W. D. Samuel Motherwell^{*a,b}

^aPfizer Institute for Pharmaceutical Materials Science, Department of Chemistry, Lensfield Road, Cambridge CB2 1EW, UK. E-mail: motherwell@ccdc.cam.ac.uk

^bCambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK

Received 9th September 2004, Accepted 28th September 2004

First published as an Advance Article on the web 2nd November 2004

The hydrogen bonding motifs seen in sulfonate salts have been examined using the CSD. In a total sample of 1069 sulfonate salts, 594 structures contain NH-donors, and this subset of structures has been investigated in detail. Several robust hydrogen-bonding motifs have been identified. A particularly robust R(2,2)8 motif, here called the *bidentate* motif, has been investigated in more detail. This motif occurs with a probability of 75.4% in the CSD, and 78.6% in the sulfonate salts. A set of rules for the prediction of the occurrence of the *bidentate* motif in sulfonate salts has been developed.

1. Introduction

In the development of new pharmaceutical materials, it is often necessary to form the salt of a drug molecule in order to improve such physical properties as solubility and stability.^{1,2} A typical example of a group of counterions used for this purpose are the R-SO₃⁻ derivatives.³ In the hope of discovering preferred patterns of hydrogen bonding with specific functional groups on drug-like cationic molecules, the hydrogen-bonding motifs exhibited by the sulfonate anion have been investigated.

A previous study investigating hydrogen bonding in structures containing sulfate ions⁴ made several interesting observations regarding the characteristics of hydrogen bonding involving this anion. It was found that the sulfate anion exhibited no preference for either nitrogen or oxygen donors, but that hydrogen bonds from O-donors were shorter and stronger than those to N-donors. It was also found that N-donors often form two- or three-centred H-bonds, and also that in most cases a sulfate O accepted two or three (seldom one) H-bonds. The formation of hydrates, which occurred in around 60% of the structures studied, allowed the completion of the coordination sphere around the sulfate ion. It was concluded that the sulfate ion “will behave as a steric ‘tightener’, placing the attracted partners in close contact.”⁴ It was hoped that the present study would reveal whether the sulfonate anion exhibits similar characteristics.

Previous work on the geometrical features of hydrogen bonding of N-donors to the sulfonate group⁵ gave a mean value for the N...O distance of 2.946 Å. This value is slightly larger than those obtained for hydrogen bonds by N-donors to carboxylates and monohydrogen phosphates (2.836 and 2.810 Å, respectively), indicating that sulfonates form longer, and weaker, hydrogen bonds to N-donors than carboxylates and monohydrogen phosphates. This was attributed to the differences in electron densities on the O atoms of the various groups. This study⁵ also showed that the hydrogen bonds formed to sulfonates by N-donors were generally linear.

The Cambridge Structural Database (CSD) (version 5.24, November 2002),⁶ which contains structural data for 272 066 organic and organometallic compounds, was used in this study. As many in-depth studies have shown, the CSD can provide a useful tool for examining structural motifs and discovering new

supramolecular synthons.^{7,8} The identification of frequently occurring specific interactions in the CSD could indicate that these interactions are fairly robust, and may be useful in crystal design and structure prediction: if an interaction occurs with a high enough probability, it can be predicted to occur in unknown compounds containing the same groups. In this study, the CSD has been used to collect information on the hydrogen bond motifs formed in a sample of sulfonate salts.

2. Methodology

Searches to determine the number of CSD hits for each ion were performed using ConQuest,⁶ and hits were restricted to organic-only structures with 3D coordinates present. Variable bond types were allowed for in the sketches of SO₃⁻ groups. Unless otherwise stated, contact searches were performed using a van der Waals sum of 3.2 Å for N...O and O...O contacts (NH...O and OH...O hydrogen bonds were identified using these contact distances). This cutoff point of 3.2 Å is in agreement with values used previously in the literature.^{4,5} In all cases, water was defined as H₂O or H₃O⁺.

3. Results and discussion

3.1. Hydrogen bonding in sulfonates

There are 21146 S-containing compounds in the CSD. 7445 contain S bonded to an O atom, of which 1069 structures contain a sulfonate moiety. An investigation of the interactions of the SO₃⁻-group with several different hydrogen bond donors is summarised in Table 1. These results agree well with previous studies.⁴ This sample of 1069 structures contains a wide range of chemical environments for SO₃⁻. There are a small number of special cases such as S-SO₃(13) and N-SO₃(27), which perhaps should be treated separately. (Also, the SO₄²⁻ ion is a special case and has been excluded). Our main focus of interest is on C-SO₃, 1019 structures.

The 1069 sulfonate-containing structures can be sub-divided. We are interested in strong hydrogen bonding only, so we have eliminated the 268 structures that contain no N- or O-hydrogen bond donors, leaving 801 structures that can be

Table 1 Hydrogen bonds formed by SO_3^- to a selection of donors

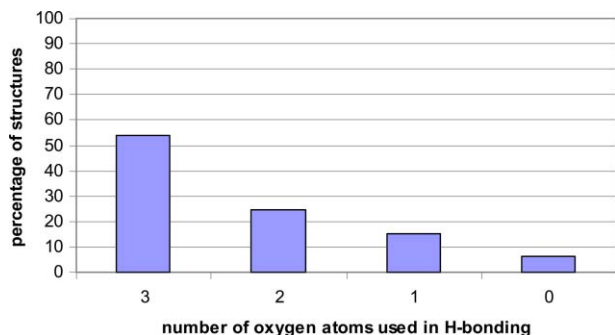
		RSO_3^-				
		$\text{R} \neq \text{O}$	SO_3OH	SO_3OX	SO_4	
Total hits		1069	63	58	162	
Hydrogen bond donors	NH	N_{pos}^a	594	39	25	146
		N_{occ}^b	522	35	23	140
		%	87.9	89.7	92.0	95.9
	OH	N_{pos}	488	29 ^c	35	125
		N_{occ}	433	26	34	115
		%	88.7	89.7	97.1	92.0
	HNXNH	N_{pos}	192	7	6	59
		N_{occ}	182	7	6	54
		%	94.8	100	100	91.5
	COOH	N_{pos}	78	6	0	43
		N_{occ}	39	4		36
		%	50.0	66.7		83.7
SH	N_{pos}	4	0	0	0	
	N_{occ}	0				
	%	0				

^a N_{pos} is the number of refcodes where an H-bond between two particular groups could possibly occur, *i.e.* the number of hits where both components (*e.g.*, SO_3^- and NH) occur in the same structure. ^b N_{occ} is the number of refcodes with an H-bond between the two components. ^c This number refers to OH groups other than SO_3OH .

subdivided according to donor types present in the structure as follows:

R- SO_3 , $\text{R} \neq \text{O}$	1069
R- SO_3 + NH or OH	801
R- SO_3 + NH only	313
R- SO_3 + NH + water only	100
R- SO_3 + OH only (no water)	47
R- SO_3 + OH + water only	64
R- SO_3 + NH + OH only (no water)	100
R- SO_3 + NH + OH + water	81
R- SO_3 + H_2O only	96

The overall picture of how the SO_3^- group interacts with other chemical groups in these 801 structures may be summarised by a simple count of the number of cases where we have at least one strong H-bond to each of the oxygens of the SO_3^- (Fig. 1). These data show that in more than half of the structures, all three sulfonate oxygen atoms are involved in hydrogen bonding to a strong donor. This indicates that the sulfonate moiety behaves as a 'steric tightener', *i.e.* brings several hydrogen bond donors into close contact, in the same way as the sulfate anion.⁴ There are 50 structures which show no sulfonate oxygen atoms are involved in H-bonding to a N- or O-donor. Of these structures, 18 are alkali metal salts, and often the cation forms short contacts to the anion. Also, there are a number of cases where there are no available N- or O-donors due to intramolecular H-bonding or complexation. There are also some cases where an existing H-bond is not detected in the search process, due to atoms not being present in the structure, or due to the H-bond being longer than 3.2 Å.

**Fig. 1** Percentage of structures using 0, 1, 2 or 3 oxygen atoms of SO_3^- in hydrogen bonding in a sample of 801 structures.

However, when a strong H-bond donor is available there is a 94% probability of a contact to the sulfonate moiety.

It must again be noted here that there is a large variety of chemical compounds in this sample, and this distribution is influenced by the 'random' acquisition of structures into the CSD. For example, there are cases of metal salts with no H-bonds but metal coordination, as in calcium methylsulfonate, CAMSUL; other cases have only one available H-donor (BAZRAV, DICKIJ). It is quite common to find when there are two donors available that they form H-bonds to two different oxygens of SO_3^- , *e.g.* CEBTIZIO, DETRID, FIDYIA. When we have 3 or more donors available we normally find H-bonds to all 3 oxygens of SO_3^- . Some structures have as many as 18 available donors, *e.g.* HEJRAP. This effect of increasing H-bond acceptance by a group with increasing donor/acceptor ratio for the compound has been studied in detail by Infantes and Motherwell.⁹ Further detailed analysis of the number of H-bonds accepted per SO_3^- group has not been reported in this paper, which is more concerned with identification of common H-bonding motifs.

As shown in Table 1, the sulfonate group forms H-bonds to N-donors in 87.9% of possible cases, and to O-donors in 88.7% of possible cases. Table 2 shows the number of structures containing $\text{SO}_3^- \cdots \text{HN}$ hydrogen bonds, $\text{SO}_3^- \cdots \text{HO}$ hydrogen bonds, or both, in the 181 structures containing both NH and OH donors. These results seem to imply that when both N- and O-donors are present in a structure which does not contain water, the SO_3^- group has a greater tendency to form an H-bond to the N-donor. However, when water is also present, there is a high tendency to form H-bonds to the water. In fact, 341 (31.8%) of the 1069 sulfonate structures contain water (which is considerably lower than the occurrence of hydrates of around 60% reported for sulfates⁴), and in 317 of these structures (93.0%), an H-bond forms between SO_3^- and water.

The 341 hydrates have been investigated further. 39 structures contain H_3O^+ , and in all cases this is the only cation. Removing these oxonium salts from the sample leaves 302 structures, of which 48% (145 structures) contain metal ions. In fact, of those organic structures in the CSD containing a metal and an SO_3^- group, 76% are hydrates (compared to 42% of structures containing any charge and any metal which are hydrates in the whole CSD). The high frequency of hydrates in metal salts has already been noted,¹⁰ but the 76% occurrence of hydrates in sulfonate metal salts is particularly high. Aside from this correlation however, no obvious reason for hydrate formation in the sulfonates could be identified.

3.2. Hydrogen bonding in NH-containing structures

There are a total of 594 sulfonate structures containing nitrogen donors, and these have been examined in detail. A variety of cyclic hydrogen bonding motifs have been identified in these structures. Some of these are represented diagrammatically in Fig. 2. We are conscious of the great variety of competing groups among these 594 structures, and the varying amounts of competing donors present, *e.g.* C-NH-C,

Table 2 Number of structures showing particular hydrogen bonds in sulfonates containing both NH and OH donors

H-bonds formed	Donors present in structure	
	NH and OH, no water 100 structures	NH, OH and water 81 structures
$\text{SO}_3^- \cdots \text{N}$	92 (92%)	67 (82.7%)
$\text{SO}_3^- \cdots \text{O}$	73 (73%)	79 (97.5%)
$\text{SO}_3^- \cdots \text{N}$ and $\text{SO}_3^- \cdots \text{O}$	66 (66%)	65 (80.2%)
$\text{SO}_3^- \cdots \text{N}$ not $\text{SO}_3^- \cdots \text{O}$	26 (26%)	2 (2.5%)
$\text{SO}_3^- \cdots \text{O}$ not $\text{SO}_3^- \cdots \text{N}$	7 (7%)	14 ^a (17.3%)
$\text{SO}_3^- \cdots \text{H}_2\text{O}$	—	76 (93.8%)

^a 13 of these are to water.

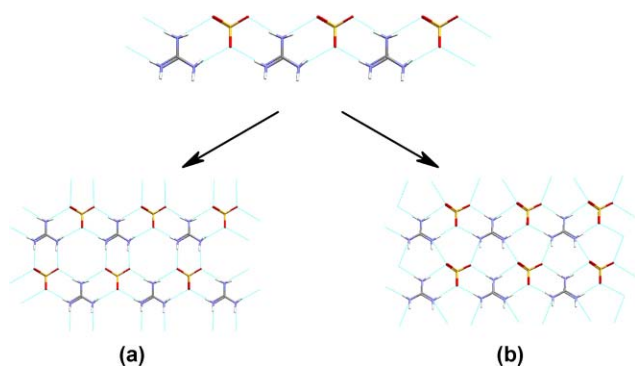


Fig. 3 Extended motifs formed in guanidinium sulfonate salts (a) From WETNEO and (b) From HEGQAL.

slipped manner *via* motif 1 (Fig. 3). We note that the structures that do not form the *bidentate* motif or ribbons of this motif all contain 2 or 3 NO₂ groups on the same aromatic ring as the sulfonate group. Only 1 of these 122 structures is a hydrate, and it is one of the structures with 2 NO₂ groups. The structures of these salts have been described in detail by Ward *et al.*,¹³ and exploited to predictively synthesize new crystals.

A preliminary analysis of the electrostatic potential of these compounds mapped onto their electron density reveals that it appears to be the electron-withdrawing effect of the nitro-groups on the rest of the molecule that influences the formation (or not) of the *bidentate* motif. A more positive potential on the sulfur atom combined with a less negative potential on the oxygen atoms could be the cause of the perturbation in the ribbon motif seen in all the other structures of this nature. Calculations indicate that the di- or tri-CN or -CF₃ derivatives would also exhibit perturbation of the *bidentate* motif.

3.2.2. Structures containing H_xNXNH_x. Aside from the guanidinium salts described above, 70 of the 594 NH-containing sulfonate salts contain an H_xNXNH_x moiety, of which 26 (*ca.* 37%) are hydrates. These 70 structures can be subdivided into those containing H₂NXNH₂ groups, those containing HNXNH₂ and those with HNXNH groups.

Fifteen structures contain H₂NXNH₂, and only 3 of these do not show the *bidentate* motif at all, but contain motif 3. (3 of those that do show the *bidentate* motif have N⋯O > 3.2 Å). The *bidentate* motifs are joined into more extended patterns in several ways: by a single O⋯H–N connection, or *via* motif 2 or 3 (Fig. 4), or the *bidentate* motifs are joined into ribbons as in the guanidinium salts. Two of these H₂NXNH₂-containing structures are uronium salts. One (UBESAV) forms flat sheets of connected *bidentate* motifs in a manner analogous to the sheets formed in the guanidinium salts. In the other uronium salt (BIKYUP), the H₂NCOH functionality successfully competes for the SO₃ functionality, and the *bidentate* motif is not seen.

28 structures contain an HNXNH₂ moiety, 7 of which contain water (*ca.* 25%). 24 of these structures have the *bidentate* motif (only 1 has N⋯O > 3.2 Å). In two cases, the *bidentate* motif is unable to form for steric reasons, and motif 2

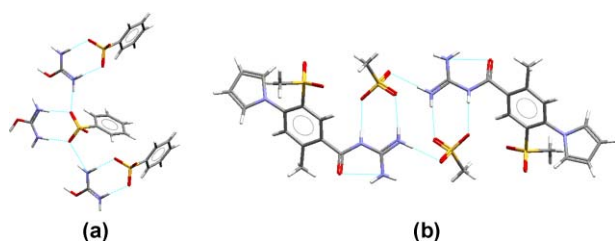


Fig. 4 Extended motifs seen in some sulfonate salts. (a) Two *bidentate* motifs joined by a single hydrogen bond (UBESAV) and (b) two *bidentate* motifs joined by motif 3, *i.e.* *bidentate-motif 3-bidentate* (NEG TIC).

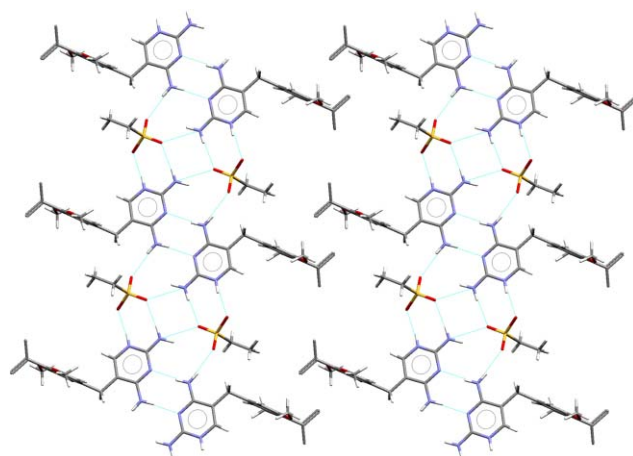


Fig. 5 'Hydrophobic' and 'hydrophilic' regions in CIDWOB, showing extended patterns of motifs, including the *bidentate-motif 3-bidentate* pattern.

or simple single NH⋯OS links form instead. Taking this into account, *ca.* 92% of these structures show the *bidentate* motif, which is almost as high an occurrence as in the guanidinium sulfonate salts (*ca.* 99%). In both structures where the *bidentate* motif could form but is not seen, there is a water molecule hydrogen-bonded to one of the N-donors, obstructing the formation of the expected motif.

A subgroup of these structures contain 2,4-diaminopyrimidine derivatives (or very similar functional groups such as triaminotriazine). This combination of functional groups always results in motifs 7, 8 or 9 (see Fig. 2) which, when combined with the *bidentate-motif 3-bidentate* pattern mentioned above, leads to layers of 'hydrophobic' and 'hydrophilic' regions in the crystal structure (Fig. 5). 9 structures contain this particular combination of functional groups, of which two are hydrates.

27 structures contain an HNXNH group, of which 14 (*ca.* 52%) are hydrates. In 17 of these structures the NH groups are cyclic, and several motifs are seen to occur. In three cases there are no NH⋯OS hydrogen bonds formed, and in all three of these structures X is C=O, and an alternative *bidentate*-type motif occurs (Fig. 6). In many of the remaining cases, two of the SO₃ oxygen atoms form hydrogen bonds to nitrogen donors, linking the molecules into chains (4), or other cyclic motifs such as motif 2 (2), motif 3 (1), motif 10 (1) and motif 14 (2).

Of the remaining 10 structures, 5 are sterically unable to form the *bidentate* motif. In these structures the molecules are linked by single NH⋯OS hydrogen bonds in four cases, and motif 2 in one case. Of the other 5 cases, 3 show the *bidentate* motif, and 2 structures contain a urea functionality, which forms the well-known urea hydrogen-bonding motif.

3.2.3. NH₄⁺- and NH₃⁺-containing structures. Once the structures containing H_xNXNH_x have been eliminated from the sample of 594 NH-containing sulfonates, 402 structures remain. We now discuss the 165 which contain only N-donors, and no O-donors.

8 structures contain NH₄⁺. 6 of these structures contain no other N-donors, and the structures form layers of NH₄⁺ and SO₃⁻ ions in an analogous manner to the guanidinium salts (see above). These layers are formed from ladders of various cyclic motifs (Fig. 7).

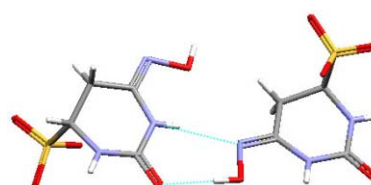


Fig. 6 Hydrogen bonding motif seen in BIFBOH.

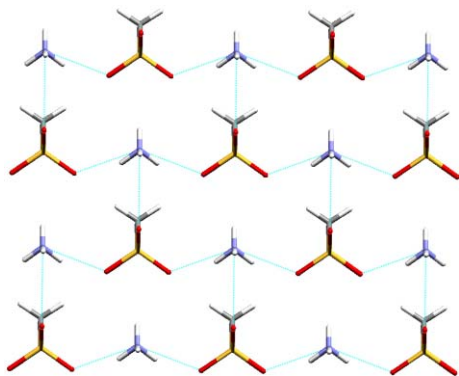


Fig. 7 Layers of ions in FASJUE.

30 structures contain NH_3^+ . Of these, 15 have no other N-donors, and once again layers of NH_3^+ and SO_3^- are formed. 10 structures contain other N-donors, and of these 9 structures show motif 2. In 6 of the 10 structures, there is some type of extended motif such as ladders or chains of motifs. However, as the complexity of the molecule increases, the degree of extended pattern formation decreases. The 5 remaining NH_3^+ -containing structures have O atoms (for example, in crown ethers) that accept hydrogen bonds from NH.

44 structures contain only NH_2 groups as hydrogen bond donors, and 83 have only NH groups. Very few of these structures show any motifs more complex than simple chains of hydrogen bonds, $\text{NH}_2 \cdots \text{SO}_3^- \cdots \text{NH}_2 \cdots \text{SO}_3^-$.

3.3. Structures containing only water as a donor

96 of the 1069 SO_3^- -containing structures have water as the only hydrogen-bond donor present in the structure. Approximately 47% of these structures are alkali or alkaline earth metal salts (see above), and in *ca.* 24% of the structures the cation is the oxonium ion (H_3O^+). Many of the hydrogen bonding motifs and patterns (ladders of motif 2, *etc.*) seen in the oxonium salts are similar to those seen in the ammonium salts. The oxonium salts are also similar to the ammonium salts in that often nets of ions form, separated by interstitial hydrophobic regions (Fig. 8). In fact, this separation of hydrophilic and hydrophobic regions into layers is characteristic of CSD structures where water is the only strong hydrogen-bond donor.

3.4. Prediction of the *bidentate* motif

Using the information described above, a set of rules have been developed for predicting the occurrence of the *bidentate* motif in sulfonate salts. These have then been tested against new structures added to the database since the collection of the 1069 sample. The rules, which apply in cases where the groups are sterically able to form the *bidentate* motif (*i.e.* not *trans* to one another) are as follows:

(i) All guanidinium sulfonate salts will show the *bidentate* motif, with the exception of those containing two or more strongly electron-withdrawing groups (NO_2 , CF_3 , CN) on the same molecule as the sulfonate group.

(ii) Sulfonate salts with no guanidinium cation, but containing an $\text{H}_x\text{N}_x\text{NH}_x$ group, will show the *bidentate* motif, with

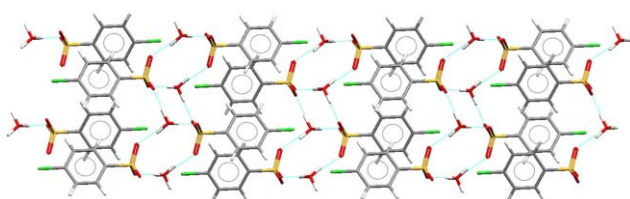


Fig. 8 'Hydrophilic' and 'hydrophobic' layers seen in JEHPAN.

the exception of structures where X is C=O or C=OH, in which case there will be strong competition for the formation of the *bidentate* motif. However, the O \cdots N distance in the motif may be greater than 3.2 Å.

Searching version 5.25 of the CSD, plus 1 update (January 2004) gave 20 new structures that contained R-SO_3^- and an $\text{H}_x\text{N}_x\text{NH}_x$ group. 13 of these are guanidinium salts, and it was predicted that all of these salts would exhibit the *bidentate* motif. In fact, 1 salt (OFOXAI) does not exhibit the *bidentate* motif. This could perhaps be rationalised by examining the electrostatic potential of this sulfonate ion. Of the remaining 7 structures, 2 were sterically unable to form the *bidentate* motif (MOTCON and WUYRUD), and so were removed from this analysis, leaving 5 structures. According to rule 2 above, all of these structures should contain the *bidentate* motif. In fact, only 1 structure (LUHWOA) does not show this motif.

4. Conclusions

This work describes the analysis of hydrogen bonding motifs formed in sulfonate salts, particularly with nitrogen donors. The robust nature of a particular R2,2(8) motif,¹¹ here called the *bidentate* motif, has been confirmed. It has been shown that certain functional groups will compete with HNXNH for the formation of this motif, particularly groups containing C=O or C=OH, and that certain motifs may form in favour of this motif, for example the urea hydrogen bonding motif. However, the *bidentate* motif can be predicted to occur with a fair degree of confidence.

Other cyclic motifs formed between the R-SO_3^- functionality and nitrogen donors have been described, and their occurrence noted. The nature of the functional groups involved obviously affects the motif that is seen in the structure. Extended networks of motifs have also been described. Guanidinium sulfonate salts are extremely likely to form extended motifs made up of ribbons of the *bidentate* motif, whereas NH_4^+ and NH_3^+ -containing sulfonate salts are more likely to form nets made up of ladders of motifs 1–5.

Acknowledgements

Discussions with Dr Neil Feeder and Dr Pete Marshall are appreciated. We also thank the Pfizer Institute for Pharmaceutical Materials Science for funding.

References

- 1 L. D. Bighley, S. M. Berge and D. C. Monkhouse, *Encyclopaedia of Pharmaceutical Technology Volume 13*, ed. J. Swarbrick and J. C. Boylan, Marcel Dekker Inc., New York, 1996, 453.
- 2 P. J. Gould, *Int. J. Pharm.*, 1986, **33**, 201.
- 3 See for example, *Handbook of Pharmaceutical Salts*, ed. P. H. Stahl and C. G. Wermuth, Wiley-VCH, Weinheim, 2002.
- 4 L. Chertanova and C. Pascard, *Acta Crystallogr., Sect. B*, 1996, **52**, 677.
- 5 B. Pirard, G. Baudoux and F. Durant, *Acta Crystallogr., Sect. B*, 1995, **51**, 103.
- 6 F. H. Allen, *Acta Crystallogr., Sect. B*, 2002, **58**, 380; I. J. Bruno, J. C. Cole, P. R. Edgington, M. Kessler, C. F. Macrae, P. McCabe, J. Pearson and R. Taylor, *Acta Crystallogr., Sect. B*, 2002, **58**, 389.
- 7 F. H. Allen and W. D. S. Motherwell, *Acta Crystallogr., Sect. B*, 2002, **58**, 407.
- 8 A. Nangia, *CrystEngComm*, 2002, **4**, 93.
- 9 L. Infantes and W. D. S. Motherwell, *Chem. Commun.*, 2004, 1166.
- 10 L. Infantes, J. Chisholm and S. Motherwell, *CrystEngComm*, 2003, **5**, 480.
- 11 F. H. Allen, P. R. Raithby, G. P. Shields and R. Taylor, *Chem. Commun.*, 1998, 1043.
- 12 J. A. Chisholm and S. Motherwell, *J. Appl. Crystallogr.*, 2004, **37**, 331.
- 13 See for example, V. A. Russell and M. D. Ward, *J. Mater. Chem.*, 1997, **7**, 1123 K; T. Holman, A. M. Pivovar and M. D. Ward, *Science (Washington, D. C.)*, 2001, **294**, 1907.