

Hydrogen-Bond Acceptor and Donor Properties of Divalent Sulfur ($Y-S-Z$ and $R-S-H$)

FRANK H. ALLEN,^{a*} CRAIG M. BIRD,^a R. SCOTT ROWLAND^a AND PAUL R. RAITHBY^b

^aCambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, England, and ^bDepartment of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, England. E-mail: allen@ccdc.cam.ac.uk

(Received 12 December 1995; accepted 10 February 1997)

Abstract

The hydrogen-bond acceptor ability of divalent sulfur in $Y-S-Z$ systems, $Y, Z = C, N, O$ or S , and the donor ability of thiol $S-H$ have been studied using crystallographic data retrieved from the Cambridge Structural Database. Of 1811 $Y-S-Z$ substructures that co-occur with $N-H$ or $O-H$ donors, only 86 (4.75%) form $S \cdots H-N, O$ bonds within $S \cdots H < 2.9 \text{ \AA}$. In dialkylthioethers, the frequency of $S \cdots H$ bond formation is 6.24%, but drops below 3% when the alkyl groups are successively replaced by Csp^2 centres. This parallels an increasing δ -positivity of S as calculated using *ab initio* methods. A similar frequency trend is observed for $O \cdots H-N, O$ bond formation by analogous oxyethers. Mean intermolecular $>S \cdots H$ distances for $O-H$ [2.67 (3) \AA] and $N-H$ [2.75 (2) \AA] donors (with H positions normalized to neutron values) are *ca* 0.25 \AA longer than in $C=S \cdots H-N, O$ systems, indicative of very weak hydrogen bonding to $>S$. Intramolecular $>S \cdots H$ are slightly more frequent (8.56%), with $S \cdots H$ slightly shorter than for the intermolecular case. In contrast, 26 (70.3%) out of 37 $S-H$ donors that co-occur with suitable acceptors form $X \cdots H-S$ bonds. The $C=O \cdots H-S$ system is predominant with a mean $O \cdots H$ distance of 2.34 (4) \AA , considerably longer (weaker) than in $C=O \cdots H-O$ systems.

1. Introduction

Crystallographic data on the molecular recognition properties of sulfur, in its terminal ($X=S$) or divalent ($Y-S-Z$) states, have rarely been the subject of detailed analyses. This situation pertains, despite the fact that sulfur is relatively common in biological systems (amino acids, disulfide bridges, thiamine, *etc.*), and is a well established constituent of therapeutic agents (penicillins, thionucleosides, *etc.*) and agrochemicals. In particular, little attention has been paid to the most fundamental recognition mechanisms, those involving $D-H \cdots S$ hydrogen bonds from $D = N$ or O donors. An earlier paper (Allen, Bird, Rowland & Raithby, 1997) addressed the hydrogen-bond acceptor characteristics of sulfur

in $>C=S$ systems, as observed in crystal structures retrieved from the Cambridge Structural Database (CSD; Allen *et al.*, 1991). It was shown that the small formal electronegativity difference between C and S can be enhanced by resonance effects in suitable chemical environments, thus enhancing the acceptor ability of $=S$ in, *e.g.* thioureido species ($N_2C=S$) compared with pure thiones ($C_2C=S$), which are extremely poor acceptors.

In this paper we address the hydrogen-bond acceptor capability of divalent sulfur in $Y-S-Z$ systems, particularly in thioethers $C-S-C$. Here, despite the minimal electronegativity difference between C and S [2.55 for C , 2.58 for S , Pauling (1960)], the chemical analogy to oxyethers $C-O-C$ is a powerful one and some degree of hydrogen-bond acceptor ability for $-S-$ might be inferred.

Most existing studies of non-covalent interactions at sulfur have concentrated on the divalent $Y-S-Z$ case ($Y, Z =$ any atom, except H). Rosenfield, Parthasarathy & Dunitz (1977) showed that short non-bonded $YZS \cdots X$ contacts ($X =$ any atom, H included) fell into two groups: (a) nucleophilic X approaching S in the $Y-S-Z$ plane along the posterior extensions of the $Y-S$ or $Z-S$ bonds, interpreted as an X interaction with the LUMO [$\sigma^*(S-Y)$ or $\sigma^*(S-Z)$], and (b) electrophilic X approaching S at angles $< 40^\circ$ from the perpendicular to the $Y-S-Z$ plane, interpreted as an $X \cdots HOMO$ (S lone pair) interaction. It was noted that $H^{\delta+}$ fell into this latter category. A more recent, but similar, analysis (Desiraju & Nalini, 1991) addressed the implications of $YZS \cdots C$ interactions for crystal engineering by subdividing the (now much larger) set of available data according to the length of the shortest unit-cell axis.

Other studies that followed the original paper of Rosenfield, Parthasarathy & Dunitz (1977) have concentrated on sulfur-nucleophile interactions (Kuczman & Kapovits, 1985; Iwasaki, 1986). Burling & Goldstein (1992, 1993) have carried out *ab initio* computations and a crystallographic database analysis of short intramolecular $S(\text{thiazole}) \cdots O(\text{furan})$ interactions in thiazofurin and analogous thiazole nucleosides. They show that $S \cdots O$ contacts in the range 2.77–3.16 \AA , much shorter than

the van der Waals radii sum of 3.3 Å (Bondi, 1964), result from an electrostatic interaction between $S^{\delta+}$ and $O^{\delta-}$, with the residual charge on S computed to be +0.273e using 6-31G* basis sets. The implications of this result for hydrogen-bond acceptance by Y—S—Z systems would appear to be significant.

2. Methodology

The April 1995 release of the Cambridge Structural Database (5.09, 140,236 entries) was used in this work. Searches for bonded substructures and short non-covalent inter- and intramolecular contacts were carried out using the program *QUEST3D* (Cambridge Structural Database, 1994). Data display and analyses were performed with *VISTA2.0* (Cambridge Structural Database, 1995). The chemical constitutions of various generic thioether and oxyether search fragments are defined in subsequent sections. Here, the CSD search constraint on atomic *total coordination number* (T) is used extensively. Thus, S^2, O^2 indicate divalent S, O respectively, while C^4, C^3 are used to indicate sp^3 - and sp^2 -hybridized carbons. Substructures involving covalent and/or non-covalent linkages were only accepted for data analysis if they occurred in CSD entries that satisfied the secondary search criteria: (a) organic compounds within CSD chemical class criteria, (b) error-free coordinate sets in CSD check procedures, (c) no disorder present in the structure, (d) not polymeric and (e) had $R \leq 0.10$ (thioethers) or $R \leq 0.05$ (oxyethers). All H atoms involved in non-bonded contact searches were placed in neutron-normalized positions, as described elsewhere (Allen, Bird, Rowland & Raithby, 1997).

Geometrical analyses for the generalized hydrogen-bonded fragment $YZ(S \text{ or } O) \cdots H-D$ (where D is N or O) were carried out in terms of: the hydrogen-bond length $d(SH)$ or $d(OH)$, the acceptor-donor distance $d(SD)$ or $d(OD)$ and the hydrogen-bond angle ($\rho_H = S \text{ or } O \cdots H-D$). The directionality of approach of H to the S or O centre was assessed using the angle of elevation (θ) and angle of rotation (φ) shown in Fig. 1. These angles are calculated with respect to the plane that is perpendicular to the YSZ or YOZ plane and contains the bisector of the YSZ or YOZ angle. Thus, this plane would contain the S or O lone pairs, assuming local tetrahedral sp^3 geometry at S or O and angles of $|\theta| \simeq 0$, $|\varphi| \simeq 120^\circ$ would indicate the colinearity of the S, O lone pair and S, O $\cdots H$ vectors. Searches for hydrogen-bonded units required only that the hydrogen-bond lengths $d(SH)$ or $d(OH)$ were less than the appropriate sums of van der Waals radii using $v(S) = 1.80$, $v(O) = 1.52$ (Bondi, 1964) and $v(H) = 1.10$ Å (Rowland & Taylor, 1996). *Ab initio* molecular orbital (MO) calculations for model compounds were carried out to obtain residual atomic charges using the *GAMESS-UK* package (Guest *et al.*, 1993). Closed-shell self-consistent field (restricted Hartree-Fock) calculations

were performed using the 6-31G* basis set with full geometry optimization. Point charges, q_j , were obtained by Mulliken population analysis.

2.1. Statistical descriptors of geometrical parameter distributions

Mean values (m) of geometrical parameters given in the text and tables throughout this paper are cited with the e.s.d. (σ) of that mean in parentheses, together with the number of observations (N_{obs}) that contribute to these statistics. The sample standard deviation (S) is simply derived as $\sigma(N_{\text{obs}})^{1/2}$ and limiting values of $ca m \pm 3.0S$ may then be used as a guide to the numerical spread in the underlying parameter distribution.

3. Results and discussion

3.1. Relative frequency of intermolecular hydrogen-bond formation at —S— acceptors

Preliminary searches for Y—S—Z fragments showed that restriction of Y, Z to the elements C or N, or O or S was appropriate: only a few (four) examples with Y, Z = Se or Te were omitted. To establish the relative frequency of intermolecular hydrogen-bond formation by divalent sulfur, two types of searches were carried out. Searches (a) established the number of CSD entries, $N_c(a)$, that contained both Y—S—Z and a suitable N—H or O—H donor. Here, the numbers of unique Y—S—Z substructures $N_s(a)$ and the numbers of donor H atoms $N_H(a)$ were also determined together with the ratio $h = N_H(a)/N_s(a)$. Searches (b) located hydrogen-bonded fragments $S \cdots H-N, O$ within $S \cdots H < 2.9$ Å. The numbers of entries $N_c(b)$ and numbers of substructures $N_s(b)$ were established. The relative frequencies of hydrogen-bond formation [$f_H = 100N_s(b)/N_s(a)$] are reported in Table 1 for the complete Y—S—Z dataset and for various Y, Z subdivisions. Analogous data for C—O—C subdivisions are also included in Table 1.

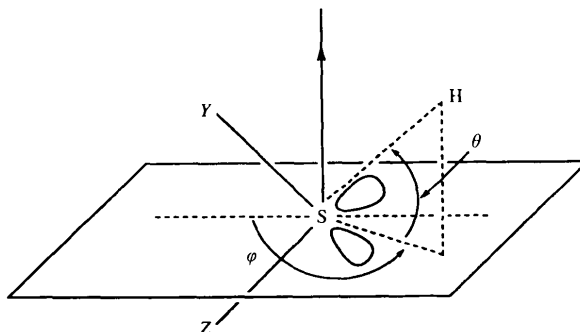


Fig. 1. Definition of the hydrogen-bond directionality parameters θ , the angle of elevation of the S, O $\cdots H$ vector from the S or O sp^3 lone-pair plane, and φ , the angle of rotation of that vector when projected onto that plane.

Table 1. Relative frequencies of intermolecular hydrogen-bond formation by divalent sulfur and divalent oxygen (the column headings are defined in the text)

Y	Z	$N_e(a)$	$N_s(a)$	$N_H(a)$	h	$N_e(b)$	$N_s(b)$	f_H (%)
(i) Divalent sulfur Y—S—Z								
C,N,O,S	C,N,O,S	1223	1811	3070	1.70	74	86	4.75
C,N,O,S	N,O,S	175	354	528	1.49	18	29	8.19
C	S*	94	214	335	1.57	7	11	5.14
C sp^3	C sp^3	332	449	862	1.92	27	28	6.24
C sp^3	C sp^2	372	464	836	1.80	13	13	2.80
C sp^2	C sp^2	392	544	972	1.79	15	16	2.94
(ii) Divalent oxygen Y—O—Z								
C sp^3	C sp^3	1901	3788	6477	1.71	710	1123	29.65
C sp^3	C sp^2	2400	4102	5173	1.26	209	267	6.51
C sp^2	C sp^2	334	399	678	1.70	18	19	4.76

* C—S—S—C systems.

Table 2. Mean geometries of YZS···H—N,O hydrogen bonds

Parameter designations are given in the *Methodology*. Distances are in Å, angles in °, with e.s.d.'s of mean values in parentheses. N_{obs} is the number of observations.

Y	Z	Donor	N_{obs}	$\langle d(SH) \rangle$	$\langle d(SD) \rangle$	$\langle \rho_H \rangle$	$\langle \theta \rangle$	$\langle \phi \rangle$
(a) Intermolecular								
C,N,O,S	C,N,O,S	O—H	39	2.67 (5)	3.39 (4)	134 (3)	23 (3)	118 (5)
		N—H	47	2.75 (2)	3.58 (2)	142 (2)	18 (2)	123 (4)
C,N,O,S	N,O,S	O—H	8	2.83 (3)	3.46 (4)	123 (4)	32 (11)	93 (8)
		N—H	21	2.76 (2)	3.57 (3)	139 (3)	19 (3)	116 (5)
C	C	O—H	31	2.63 (4)	3.37 (5)	137 (4)	20 (3)	125 (5)
		N—H	26	2.74 (2)	3.58 (3)	145 (3)	17 (3)	129 (5)
(b) Intramolecular								
C,N,O,S	C,N,O,S	O—H	40	2.55 (4)	3.10 (4)	118 (3)	21 (2)	85 (4)
		N—H	115	2.60 (2)	3.12 (2)	113 (1)	23 (2)	92 (3)

The results of Table 1 show unequivocally that divalent sulfur is a very poor acceptor of hydrogen bonds. The overall frequency of formation of 4.75% across a statistically large sample, and one in which the number of potential donor—H is in clear excess, represents a very low success rate. The vast majority of S atoms located in this study are from C—S—C environments and, hence, the result is perhaps unsurprising given the ready polarizability of S and the almost negligible electronegativity difference between C and S.

There is a noticeable decrease in the frequency of hydrogen-bond formation (f_H , Table 1) as the C sp^3 atoms of dialkyl thioethers are replaced by one or two C sp^2 centres. Here, simple Mulliken charge analyses (see the *Methodology*) show that the S atom has an increasingly positive residual charge (q_S) from *ab initio* calculations for a series of appropriate model molecules: dimethylthioether, $q_S = +0.112e$; vinylmethylthioether, $q_S = +0.169e$; divinylthioether, $q_S = +0.240e$. Despite the expected much higher f_H value for the analogous dialkyl oxyethers (29.65%), the decrease in f_H on C sp^2 substitution is even more dramatic for the C—O—C system. Obviously, a conjugative interaction involving the S or O lone pairs and delocalized systems associated with C sp^2 substituents is responsible for the low f_H values in the unsaturated thio- and oxyethers. This is

essentially the same mechanism as that established by Burling & Goldstein (1992, 1993) in their study of unsaturated ring (S)···O interactions. Recently, Klebe (1995) kindly informed us of an extensive *ab initio* study of the electrostatic potential around O in C—O—C systems that is in accordance with these crystallographic results. Hence, we have not further pursued the computational work suggested by the data of Table 1.

3.2. Intermolecular hydrogen-bond geometry at S acceptors

The obvious paucity (Table 1) of YZS···H—N,O bonds makes it difficult to present statistically meaningful average geometries for all except the most general YZS subsets, particularly if we wish to subdivide the bonds in terms of N—H or O—H donors. Table 2 (a) shows the mean geometry for all S···H—O and S···H—N systems, together with data for two YZS subsets that have acceptable numbers of observations. Fig. 2 shows histograms of $d(SH)$, ρ_H , $|\theta|$ and $|\phi|$ (see Fig. 1), together with the $d(SH)$ — ρ_H scattergram, for all 86 S···H—N,O bonds.

The overall mean geometry [Table 2 (a), lines 1 and 2] shows that $d(SH)$ for N—H acceptors is some 0.08 Å longer than that for O—H. This result reflects the relative

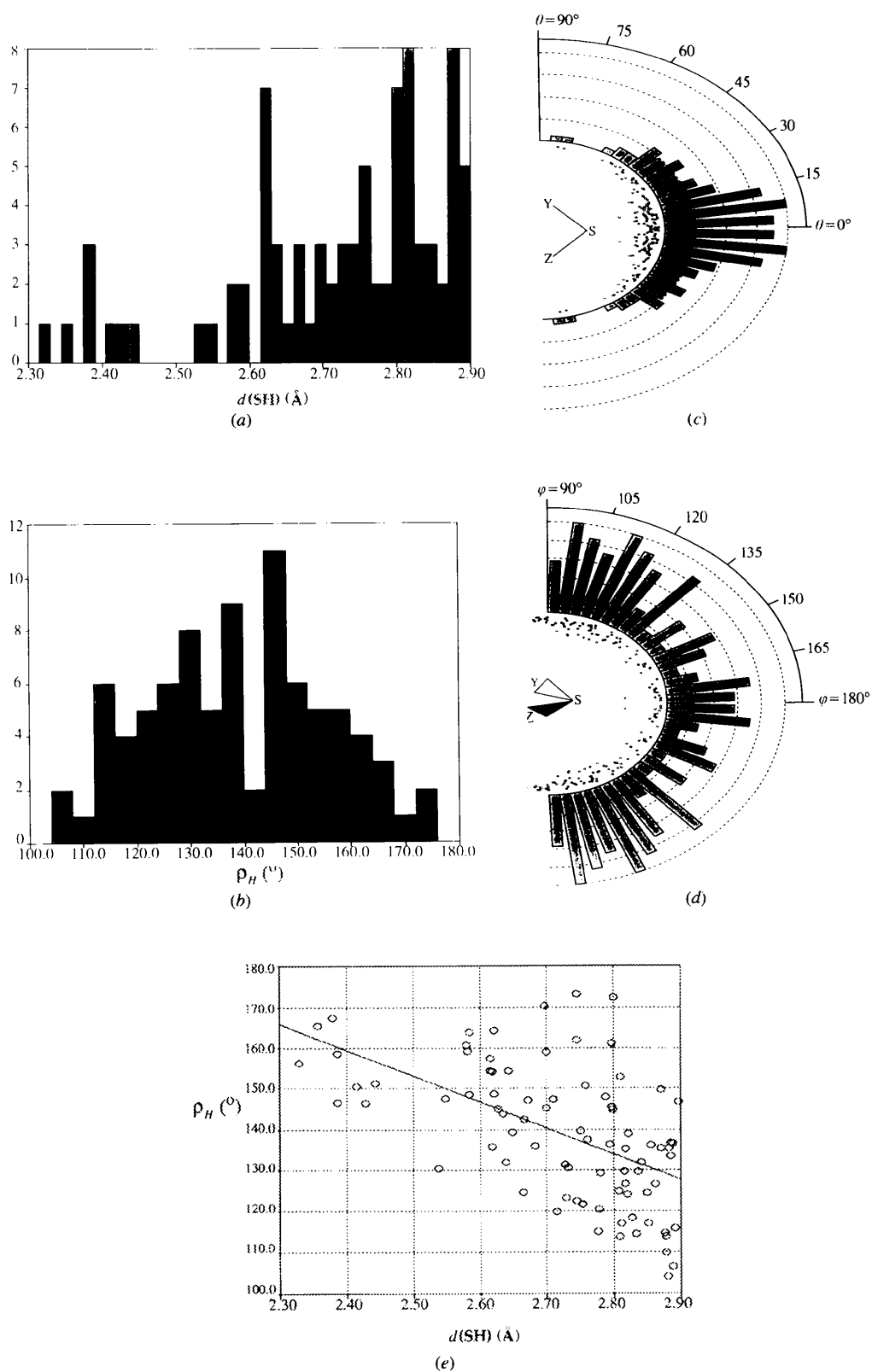


Fig. 2. YZS...H...N,O hydrogen bonds: histograms of (a) $d(\text{SH})$, (b) ρ_H , (c) $|\theta|$, (d) $|\phi|$, together with (e) the scatterplot of $d(\text{SH})$ versus ρ_H with the least-squares regression line indicated.

donor abilities of N—H and O—H (Jeffrey & Saenger, 1991) and the distance difference is comparable to that observed for =S acceptors (Allen, Bird, Rowland & Raithby, 1997). However, by comparison with the =S case, $d(\text{SH})$ values involving divalent S are elongated by *ca* 0.25 Å, indicative of the much weaker hydrogen bonding that is suggested by the frequency data of Table 1. The $d(\text{SH})$ histogram (Fig. 2*a*) does show a few bonds that are comparable in length to those formed by =S acceptors. These arise from dialkylthioethers in which the S atom is the only realistic acceptor. Examination of the chemical environments of other $\text{YZS} \cdots \text{H}$ bonds does not reveal any systematic patterns or local features that might be responsible for bond formation.

The overall mean hydrogen-bond angle (ρ_H) is 139(3)°, much lower than the 157(1)° obtained for =S acceptors (Allen, Bird, Rowland & Raithby, 1997). However, the $d(\text{SH})$ – ρ_H scatterplot (Fig. 2*e*) shows a relatively high linear correlation coefficient (–0.55), with the shortest (strongest) hydrogen bonds being associated with the largest (most linear) ρ_H values. Both these observations are, of course, typical of hydrogen-bond behaviour.

The overall mean $|\theta|$ and $|\varphi|$ values for 86 hydrogen bonds are 20(2) and 120(3)°, respectively. Thus, H approaches S close to the lone-pair plane and at a $|\varphi|$ angle that would be closely aligned with an $\text{S}(sp^3)$ lone-pair direction. These results are in agreement with the early study of Rosenfield, Parthasarathy & Dunitz (1977), although the geometrical constraints used differ slightly. The symmetrized φ distribution (Fig. 2*d*) shows a preference for values in the 80–130° range, with a smaller, but distinct, population reaching out to the limiting value of 180°. Thus, the approach of donor-H towards the conventional S lone-pair density is less marked for —S— than it is for =S (see the φ distributions in Allen, Bird, Rowland & Raithby, 1997), where a much clearer separation between ‘lone-pair’ φ peaks is observed. This comparative situation for —S— and =S closely parallels the case of —O— versus =O (Murray-Rust & Glusker, 1984). Here, however, the separation of ‘lone-pair’ φ peaks for —O— acceptors is much less clear than for —S—, with an almost continuous density of approaching donor-H to ether-O spanning the conventional lone-pair regions.

3.3. Intramolecular hydrogen-bond geometry at S acceptors

Whilst the primary focus of this study is the intermolecular hydrogen-bonding capability of divalent sulfur, we find that the frequency of intramolecular bond formation is nearly double that of intermolecular bonds: 155 (8.56%) of the 1811 —S— atoms (Table 1, line 1) forms intramolecular $>\text{S} \cdots \text{H}—\text{N}, \text{O}$ contacts < 2.9 Å. These CSD searches were carried out under the additional constraint that the acceptor S and donor H atoms must be separated by at least four contiguous

covalent bonds within the molecule, thus eliminating 1,3- and 1,4-interactions from consideration. The mean geometry for the 155 located bonds, subdivided by the O—H or N—H donor, is given in Table 2 (*b*). Here, the mean $d(\text{SH})$ for N—H donors is again longer than for O—H donors, as observed for intermolecular bonds, but this result is less statistically significant. However, the overall lengths of the intramolecular bonds are shorter (by 0.12–0.15 Å) than for the intermolecular case. While steric constraints are, typically, affecting the ρ_H and $|\varphi|$ values, we note that the mean $|\theta|$ value for intramolecular bonds again suggests lone-pair involvement.

3.4. The S—H donor

The CSD contains 43 thiol groups in 35 crystal structures located using the secondary search criteria listed in the *Methodology*; 37 of these —S—H groups occur in 29 structures that also contain suitable acceptor groups. A search for —S—H $\cdots X$ contacts (X = any element except C or H) was carried out using the criteria: (*a*) the H $\cdots X$ distance was less than the appropriate sum of van der Waals radii [$\text{H} = 1.1$ Å (Rowland & Taylor, 1996), X taken from Bondi (1964)]; (*b*) inter- and intramolecular contacts were located using the ‘four covalent bond’ constraint (see above) for the latter search; (*c*) the S—H distance was normalized to 1.34 Å, a value taken from the neutron diffraction study of *N*-acetyl-L-cysteine by Takusagawa, Koetzle, Kou & Parthasarathy (1981); the mean S—H distance from the X-ray structures was 1.23(2) Å.

A total of 29 S—H $\cdots X$ bonds were located, involving 26 (70.3%) of the available S—H donors, with 15, 10, 2, 1, 1 bonds involving O, S, N, Cl, F acceptors, respectively. Three bonds (one to O, two to S) were intramolecular. The O acceptors were principally from C=O in ketones, carboxylates and amides, the S acceptors were always from C=S groups. The mean intermolecular H $\cdots \text{O}$ distance was 2.34(4) Å, which was considerably shorter than the van der Waals radius sum (2.62 Å), but much longer (weaker) than typical H $\cdots \text{O}=\text{C}$ bonds [1.884(3) and 2.011(2) Å for O—H and N—H donors, respectively (Allen, Bird, Rowland & Raithby, 1997)]. The SH $\cdots \text{S}=\text{C}$ hydrogen bonds average 2.51(5) Å for the eight intermolecular examples, a value which is very similar to those involving N—H or O—H donors (Allen, Bird, Rowland & Raithby, 1997). The mean S—H $\cdots \text{O}, \text{S}$ angle (ρ_H) is 151(5)° over all 22 intermolecular examples.

4. Conclusions

The statistical data presented in this paper clearly indicate that divalent sulfur is a poor hydrogen-bond acceptor. Only in a few cases, involving dialkyl thioethers as the sole potential acceptor, do these bonds appear to contribute significantly to molecular aggregation in organic crystal structures. Evidently,

their role in biological systems is almost certainly minimal. The answer 'no' is often important in science and the negative result obtained here is not without significance. It indicates that other, and probably more significant, non-covalent interactions are fundamental to the molecular recognition properties of divalent sulfur. The directional in-plane approach of $X^{\delta-}$ to YSZ systems (Rosenfield, Parthasarathy & Dunitz, 1977), the short S...O contacts observed by Iwasaki (1986) and Burling & Goldstein (1993), and the increasing δ -positivity of —S— (particularly in delocalized systems) as calculated here and by Burling & Goldstein (1992) all indicate that non-hydrogen-mediated interactions are more significant than hydrogen-bond acceptance for divalent sulfur. Thus, we are now carrying out extensive database surveys and high-level *ab initio* calculations to ascertain the geometries, directional properties and energies of a variety of intermolecular $>S \cdots X^{\delta-}$ interactions.

We thank Dr R. Taylor and Dr J. P. M. Lommerse (CCDC) for valuable discussions.

References

- Allen, F. H., Bird, C. M., Rowland, R. S. & Raithby, P. R. (1977). *Acta Cryst.* **B53**, 680–695.
- Allen, F. H., Davies, J. E., Galloy, J. J., Johnson, O., Kennard, O., Macrae, C. F., Mitchell, E. M., Mitchell, G. F., Smith, J. M. & Watson, D. G. (1991). *J. Chem. Inf. Comput. Sci.* **31**, 187–204.
- Bondi, A. (1964). *J. Phys. Chem.* **68**, 441–451.
- Burling, F. T. & Goldstein, B. M. (1992). *J. Am. Chem. Soc.* **114**, 2313–2320.
- Burling, F. T. & Goldstein, B. M. (1993). *Acta Cryst.* **B49**, 738–744.
- Cambridge Structural Database (1994). *User's Manual. Getting Started with the CSD System*. Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, England.
- Cambridge Structural Database (1995). *User's Manual. VISTA2.0*. Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, England.
- Desiraju, G. R. & Nalini, V. (1991). *J. Mater. Chem.* **1**, 201–203.
- Guest, M. F., van Lenthe, J. H., Kendrick, J., Schoeffel, K., Sherwood, P. & Harrison, R. J. (1993). *GAMESS-UK User's Guide and Reference Manual*. Computing for Science Ltd, Daresbury Laboratory, Warrington, England.
- Iwasaki, F. (1986). *Acta Cryst.* **C42**, 121–124.
- Jeffrey, G. A. & Saenger, W. (1991). *Hydrogen Bonding in Biological Systems*. Berlin: Springer-Verlag.
- Klebe, G. (1995). Personal communication.
- Kuczman, A. & Kapovits, I. (1985). In *Organic Sulfur Chemistry*, edited by F. Bernardi, I. G. Csizmadia & A. Mangui, pp. 191–245. Amsterdam: Elsevier.
- Murray-Rust, P. & Glusker, J. P. (1984). *J. Am. Chem. Soc.* **106**, 1018–1025.
- Pauling, L. (1960). *The Nature of the Chemical Bond*, 3rd ed. Ithaca: Cornell University Press.
- Rosenfield, R. E., Parthasarathy, R. & Dunitz, J. D. (1977). *J. Am. Chem. Soc.* **99**, 4860–4862.
- Rowland, R. S. & Taylor, R. (1996). *J. Phys. Chem.* **100**, 7384–7391.
- Takusagawa, F., Koetzle, T. F., Kou, W. W. H. & Parthasarathy, R. (1981). *Acta Cryst.* **B37**, 1591–1597.