



## UvA-DARE (Digital Academic Repository)

### Hydrogen-bond landscapes, geometry and energetics of squaric acid and its mono- and dianions: a Cambridge Structural Database, IsoStar and computational study

Allen, F.H.; Cruz-Cabeza, A.J.; Wood, P.A.; Bardwell, D.A.

**DOI**

[10.1107/S2052519213020277](https://doi.org/10.1107/S2052519213020277)

**Publication date**

2013

**Document Version**

Final published version

**Published in**

Acta Crystallographica. Section B-Structural Science

[Link to publication](#)

**Citation for published version (APA):**

Allen, F. H., Cruz-Cabeza, A. J., Wood, P. A., & Bardwell, D. A. (2013). Hydrogen-bond landscapes, geometry and energetics of squaric acid and its mono- and dianions: a Cambridge Structural Database, IsoStar and computational study. *Acta Crystallographica. Section B-Structural Science*, 69(5), 514-523. <https://doi.org/10.1107/S2052519213020277>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the Library of the University of Amsterdam (<https://dare.uva.nl>)

Frank H. Allen,<sup>a\*</sup> Aurora J. Cruz-Cabeza,<sup>b</sup> Peter A. Wood<sup>a</sup> and David A. Bardwell<sup>a</sup><sup>a</sup>Cambridge Crystallographic Data Centre (CCDC), 12 Union Road, Cambridge CB2 1EZ, England, and <sup>b</sup>van't Hoff Institute for Molecular Sciences, Universiteit van Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

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

# Hydrogen-bond landscapes, geometry and energetics of squaric acid and its mono- and dianions: a Cambridge Structural Database, IsoStar and computational study

Received 3 June 2013  
Accepted 22 July 2013

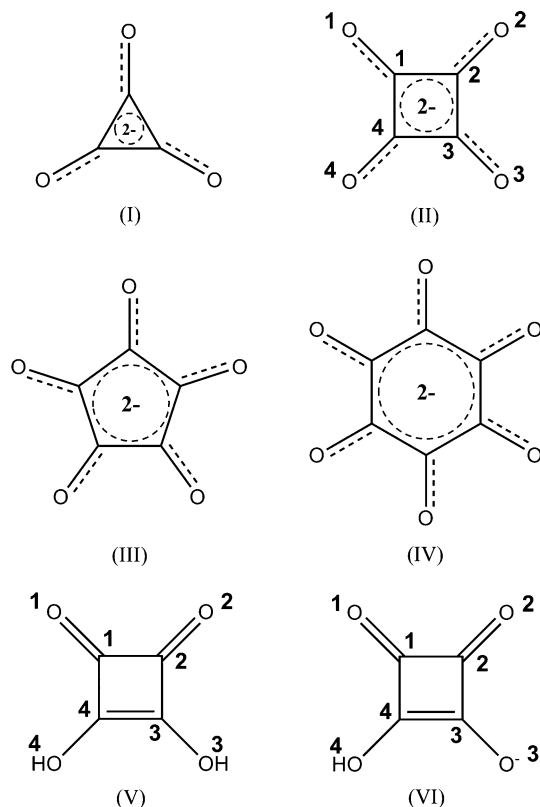
As part of a programme of work to extend central-group coverage in the Cambridge Crystallographic Data Centre's (CCDC) IsoStar knowledge base of intermolecular interactions, we have studied the hydrogen-bonding abilities of squaric acid ( $\text{H}_2\text{SQ}$ ) and its mono- and dianions ( $\text{HSQ}^-$  and  $\text{SQ}^{2-}$ ) using the Cambridge Structural Database (CSD) along with dispersion-corrected density functional theory (DFT-D) calculations for a range of hydrogen-bonded dimers. The  $-\text{OH}$  and  $-\text{C}=\text{O}$  groups of  $\text{H}_2\text{SQ}$ ,  $\text{HSQ}^-$  and  $\text{SQ}^{2-}$  are potent donors and acceptors, as indicated by their hydrogen-bond geometries in available crystal structures in the CSD, and by the attractive energies calculated for their dimers with acetone and methanol, which were used as model acceptors and donors. The two anions have sufficient examples in the CSD for their addition as new central groups in IsoStar. It is also shown that charge- and resonance-assisted hydrogen bonds involving  $\text{H}_2\text{SQ}$  and  $\text{HSQ}^-$  are similar in strength to those made by carboxylate  $\text{COO}^-$  acceptors, while hydrogen bonds made by the dianion  $\text{SQ}^{2-}$  are somewhat stronger. The study reinforces the value of squaric acid and its anions as cocrystal formers and their actual and potential importance as isosteric replacements for carboxylic acid and carboxylate functions.

## 1. Introduction

In the early 1960s West and co-workers (West *et al.*, 1960; West & Powell, 1963) identified and studied a general series of  $D_{nh}$  symmetric, electron-delocalized oxocarbon dianions of the general formula  $\text{C}_n\text{O}_n^{2-}$ , representing a new class of aromatic substances. Thus, deltate (I), squarate (II), croconate (III) and rhodizonate (IV) are the dianions for  $n = 3-6$ , respectively, of the parent fully protonated compounds deltic, squaric, croconic and rhodizonic acids. While all of these parent acids and their simple derivatives, as well as their mono- and dianions, have attracted crystallographic interest, this interest has been heavily concentrated on the  $n = 4$  homologue. Squaric acid [ $\text{H}_2\text{SQ}$ : 3,4-dihydroxycyclobut-3-ene-1,2-dione (V)] is a thermally stable, strained planar molecule and is highly acidic with  $\text{p}K_1 = 1.2-1.7$  and  $\text{p}K_2 = 3.2-3.5$  for the two O-H protons (Cohen *et al.*, 1959; MacDonald, 1968). In this paper we denote the monoanion, hydrogen squarate (VI), as  $\text{HSQ}^-$  and the squarate dianion (II) as  $\text{SQ}^{2-}$  and, where necessary, refer to the whole grouping as SQ.

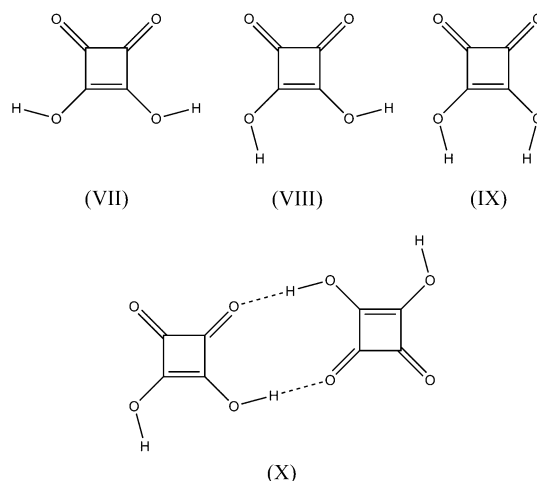
Our principal interest in squaric acid, its anions and derivatives arises from their applications, both actual and potential, as active pharmaceutical ingredients. Thus, the dibutylester of  $\text{H}_2\text{SQ}$  is used for the treatment of warts (see *e.g.* Silverberg *et al.*, 2000) and alopecia (Caserio, 1987), while Marson (2011) discusses the use of squarates as unusual

scaffolds in medicinal chemistry. He notes that derivatives of  $\text{H}_2\text{SQ}$  can act as isosteres of carboxylic acids, aminosquarate derivatives can mimic amino acids, while diamides of  $\text{H}_2\text{SQ}$  can act as phosphate isosteres and phosphate linkage mimics in DNA. Marson (2011) and Meanwell (2011) show several  $\text{H}_2\text{SQ}$  diamide derivatives that act as receptor antagonists, including the  $\text{H}_2\text{SQ}$  diamide derivative of a receptor antagonist (perzinfotel) that acts as a prodrug which increases the oral availability of perzinfotel by a factor of seven (Baudy *et al.*, 2009). Squaric acid and its anions are also the basis of squaraine and squarylium dyestuffs, due to the formation of extended delocalized systems with aniline or other electron-rich systems (see *e.g.* Sprenger & Ziegenbein, 1966), some of which are used as fluorescent biomarkers (Terpetschnig & Lakowicz, 1993). Some of these dyes present large second- and third-order nonlinear optical (NLO) responses and the NLO responses of these compounds and a number of other squarates, particularly pyridinium betaine derivatives, are an active research area (Spassova *et al.*, 2000; Kolev *et al.*, 2004). Squaric acid and its anions are also being investigated for their potential as organic ferroelectric materials (Horiuchi *et al.*, 2007).



Historically, the major features of crystallographic structural interest in  $\text{H}_2\text{SQ}$ ,  $\text{HSQ}^-$  and  $\text{SQ}^{2-}$  have been the effect of high strain and electron delocalization on the intramolecular geometry, and the patterns and geometry of the hydrogen bonds formed by the strong H-donors and O-acceptors (see *e.g.* Semmingsen *et al.*, 1977, 1995; Hollander *et al.*, 1977). It is the obvious ability of these molecules to form strong hydrogen bonds through charge- and resonance-assis-

tance (CAHB, RAHB; see Gilli & Gilli, 2009 for an overview) that has led to applications in crystal engineering and their occurrence as cocrystal formers (see, *e.g.*, Bertolasi *et al.*, 2001; Gilli *et al.*, 2001; Mathew *et al.*, 2002).



An extensive knowledge of the hydrogen-bond landscapes of chemical functional groups is vital in many areas of research, and especially in guiding studies of protein–ligand interactions and protein–ligand docking in medicinal chemistry applications. To assist in these studies, the CSD system contains the comprehensive IsoStar library of intermolecular interaction information (Bruno *et al.*, 1997). IsoStar provides three-dimensional scatterplots of the distribution of ‘contact groups’, particularly hydrogen-bond donors, around a set of selected central groups, *i.e.* chemical functional groups and heterocyclic rings that are common in organic systems. IsoStar contains extensive numerical information on these interactions together with computational energy minima for selected interactions calculated using intermolecular perturbation theory (IMPT: Hayes & Stone, 1984; Stone, 1993).

In this paper we use the Cambridge Structural Database (CSD: Allen, 2002), including the generation of IsoStar plots, to analyse and map the intramolecular and hydrogen-bond geometries observed in crystal structures that contain  $\text{H}_2\text{SQ}$ ,  $\text{HSQ}^-$  or  $\text{SQ}^{2-}$ . The CSD and IsoStar analyses are augmented by DFT-D calculations to study the attractive energies of  $\text{O} \cdots \text{H} \cdots \text{O}$  hydrogen bonds formed by all three SQ moieties and a comparison of these data with energy values obtained previously for other hydrogen-bonded systems [see, *e.g.*, the IsoStar library (Bruno *et al.*, 1997) and elsewhere (Allen *et al.*, 2010, 2012)]. As a result of this work, intermolecular interaction information for  $\text{HSQ}^-$  and  $\text{SQ}^{2-}$  has been included in the November 2012 release of the IsoStar library so as to improve its coverage of groups that are important in medicinal chemistry and materials design.

## 2. Methodology

### 2.1. Cambridge Structural Database (CSD) analyses

Database analyses were carried out using CSD Version 5.34 (November 2012) plus its first distributed update, a total of

**Table 1**

Mean intramolecular geometry for H<sub>2</sub>SQ, HSQ<sup>-</sup> and SQ<sup>2-</sup> determined from crystal structures retrieved from the CSD.

Atomic numbering is shown in the chemical formulae (II), (V) and (VI). Bond lengths (Å) and valence angles (°) are averaged over C<sub>s</sub> symmetry for H<sub>2</sub>SQ, C<sub>1</sub> symmetry for HSQ<sup>-</sup> and D<sub>4h</sub> symmetry for SQ<sup>2-</sup>. All numerical means ( $\bar{x}$ ) are given as  $\bar{x}(\sigma_s, n)$ , where  $\sigma_s$  is the sample e.s.d. and  $n$  is the number of observations contributing to the mean. Symmetry-related mean values are shown in brackets, e.g. as [ $\bar{x}$ ].

Parameter	H <sub>2</sub> SQ (C <sub>s</sub> )	HSQ <sup>-</sup> (C <sub>1</sub> )	SQ <sup>2-</sup> (D <sub>4h</sub> )
C1–C2	1.497(7,10)	1.489(13,78)	1.459(11,184)
C2–C3	1.459(12,20)	1.487(11,78)	[1.459]
C3–C4	1.403(12,10)	1.428(15,63)	[1.459]
C4–C1	[1.459]	1.433(13,78)	[1.459]
C1–O1	1.231(11,20)	1.247(16,78)	1.254(11,184)
C2–O2	[1.231]	1.226(8,78)	[1.259]
C3–O3	1.291(14,20)	1.252(13,78)	[1.259]
C4–O4	[1.291]	1.308(11,78)	[1.259]
C4–C1–C2	88.2(7,20)	89.0(9,78)	90.0(7,184)
C1–C2–C3	[88.2]	88.6(6,78)	[90.0]
C2–C3–C4	91.8(7,20)	89.2(8,78)	[90.0]
C3–C4–C1	[91.8]	93.1(6,78)	[90.0]
C2–C1–O1	135.8(6,20)	134.8(13,78)	135.0(11,368)
C4–C1–O1	136.4(7,20)	136.2(10,78)	[135.0]
C1–C2–O2	[136.4]	135.7(10,78)	[135.0]
C3–C2–O2	[135.8]	135.6(10,78)	[135.0]
C2–C3–O3	134.2(23,20)	134.5(11,78)	[135.0]
C4–C3–O3	134.1(23,20)	136.3(12,78)	[135.0]
C3–C4–O4	[134.1]	133.9(26,78)	[135.0]
C1–C4–O4	[134.2]	132.9(27,78)	[135.0]

586 977 structures. Substructure searches and the location of hydrogen bonds used the CSD system (CSDS) program *ConQuest* (Bruno *et al.*, 2002) with the following secondary search criteria: atomic coordinates error-free after CSD checks; no disorder in the crystal structure; no *catena* (polymeric) bonding; no powder studies; organic structures only according to CSD definitions; and a crystallographic *R* factor < 0.075 (unless stated otherwise).

The numbers of structures ( $N_s$ ) and independent chemical fragments ( $N_f$ ) available for detailed geometrical analysis were: H<sub>2</sub>SQ ( $N_s = 14$ ,  $N_f = 14$ ), HSQ<sup>-</sup> (69, 91) and SQ<sup>2-</sup> (47,47). Hydrogen bonds involving N–H and O–H donors were located and geometrically characterized using neutron-normalized H-atom positions (Allen, 1986) with N–H and O–H bond distances of 1.015 and 0.993 Å, respectively, as determined from accurate neutron structures in the CSD (Allen & Bruno, 2010). Geometrical parameters used to characterize hydrogen bonds of the general form  $D(\text{onor})\cdots H\cdots A(\text{ceptor})\cdots R$  were: the hydrogen-bond distance,  $d(\text{HA})$  (Å), the angle at H,  $\theta(\text{DHA})$  (°), and where relevant the angle of approach to the acceptor-O,  $\varphi(\text{HAR})$  (°). Hydrogen bonds were defined as having  $d(\text{HA})$  less than the sum of the van der Waals radii of H and A, and with  $\theta(\text{DHA}) > 120^\circ$  (Wood *et al.*, 2009). All data analyses were carried out using the CSDS program *Vista* and its successor code (Sykes *et al.*, 2011) incorporated within the CSDS *Mercury* program (Macrae *et al.*, 2008).

## 2.2. IsoStar maps of intermolecular interactions

IsoStar (Bruno *et al.*, 1997) contains information about non-bonded interactions derived from the CSD and from protein–ligand complexes in the Protein Data Bank (Berman *et al.*, 2000), together with IMPT energy minima for important interactions. The current distributed version of IsoStar (November 2012) covers more than 300 central groups and over 50 contact groups, which give rise to more than 20 000 scatterplots derived from the CSD and about 5500 scatterplots derived from PDB structures having a resolution better than 2 Å. *Ab initio* theoretical energy minima are included for more than 1500 key interactions. In this work we present IsoStar plots of the hydrogen-bond landscapes of HSQ<sup>-</sup> and SQ<sup>2-</sup> for which sufficient data exist in the CSD.

## 3. Mean intramolecular geometry

It is necessary to provide the computational procedures described later in this work with initial ground-state intramolecular geometry for model compounds, and this was determined for H<sub>2</sub>SQ, HSQ<sup>-</sup> and SQ<sup>2-</sup> from the crystal structures retrieved from the CSD. Mean geometrical data are presented in Table 1, with the bond lengths and angles averaged over C<sub>s</sub> symmetry for H<sub>2</sub>SQ, C<sub>1</sub> symmetry for HSQ<sup>-</sup> and D<sub>4h</sub> symmetry for SQ<sup>2-</sup>. The numbering scheme for the C and O atoms is that shown in chemical diagrams (II), (V) and (VI). The bond-length data clearly show the effects of electron delocalization throughout the three systems, with squaric acid itself being trapezium shaped and the dianion a perfect square. A variety of computational results are available (see e.g. Lunelli *et al.*, 1996; Zhou *et al.*, 2000; Spassova *et al.*, 2000), all of which show acceptable agreement with the experimental data in Table 1, but with variations due to the computational procedures or basis sets used. The hydroxy H atoms in parent H<sub>2</sub>SQ can adopt one of three isomeric relationships: *ZZ* (VII), *EZ* (VIII) or *EE* (IX). Published computational studies on isolated molecules all agree that the *ZZ* isomer is the most energetically favoured, but that the *EZ* isomer is also within about 6–10 kJ mol<sup>-1</sup> of the minimum. The *EE* isomer has a higher relative energy, around 28 kJ mol<sup>-1</sup> above the *ZZ* minimum (all energy data taken from Zhou *et al.*, 2000). In the CSD, where the positions of the H atoms will adjust to hydrogen-bonding requirements in the condensed phase, two of the ten examples are *ZZ*, seven are *EZ* (including four independent redeterminations of either polymorph of H<sub>2</sub>SQ itself), and only one example is *EE*.

## 4. Hydrogen bonding involving squaric acid and its anions

Parameter ranges and mean values derived from the CSD for  $d(\text{HA})$ ,  $\theta(\text{DHA})$  and  $\varphi(\text{HAR})$  for hydrogen bonds involving the O–H donors and the terminal-O acceptors of squaric acid and its mono- and dianions are collected in Table 2. While it is valid to derive mean values for  $d(\text{HA})$  and  $\varphi(\text{HAR})$  to assess the relative strengths of the various hydrogen bonds and their

**Table 2**

Mean geometry of hydrogen bonds involving O—H donors and terminal-O acceptors in squaric acid (H<sub>2</sub>SQ) and its anions (HSQ<sup>−</sup> and SQ<sup>2−</sup>).

Hydrogen bond lengths,  $d(\text{HA})$ , are in Å and relevant angles,  $\theta(\text{DHA})$  and  $\varphi(\text{HAR})$ , are in degrees. All numerical means ( $x$ ) are given as  $x(\sigma_x, n)$ , where  $\sigma_x$  is the sample e.s.d. and  $n$  is the number of observations contributing to the mean. Mean values for  $\theta(\text{DHA})$  are given as an indicator of the overall linearity of the hydrogen bonds included in the averaging process.

Hydrogen-bond type	$d(\text{HA})$ (Å)		$\theta(\text{DHA})$ (°)		$\varphi(\text{HAR})$ (°)	
	Range	Mean	Range	Mean	Range	Mean
Squaric acid (H <sub>2</sub> SQ)						
O—H(H <sub>2</sub> SQ)···O=C(sq)	1.42–1.62	1.54(5,18)	152–180	170(8,18)	112–136	121(7,18)
N <sup>+</sup> —H···O=C(H <sub>2</sub> SQ)	1.88–2.25	2.06(15,4)	121–171	149(20,4)	125–137	132(6,4)
Hydrogen squarate (HSQ <sup>−</sup> )						
O—H(HSQ <sup>−</sup> )···O=C(sq)	1.44–1.67	1.55(5,65)	143–180	166(9,65)	109–142	126(10,65)
O—H(HSQ <sup>−</sup> )···N(bipy)	1.59–1.65	1.63(2,6)	164–175	170(4,6)	—	—
O—H(water)···O=C(HSQ <sup>−</sup> )	1.65–2.30	1.89(17,25)	120–179	159(15,25)	107–165	123(13,25)
N <sup>+</sup> (ar)—H···O=C(HSQ <sup>−</sup> )	1.58–1.90	1.68(10,25)	144–177	165(9,25)	105–137	121(8,25)
N <sup>+</sup> —H···O=C(HSQ <sup>−</sup> )	1.71–2.41	1.98(18,79)	122–179	153(15,79)	95–147	123(11,79)
N <sup>0</sup> H···O=C(HSQ <sup>−</sup> )	1.70–2.56	2.04(20,45)	124–179	157(15,45)	92–162	126(16,45)
Squarate dianion (SQ <sup>2−</sup> )						
O—H(H <sub>2</sub> SQ)···O=C(SQ <sup>2−</sup> )	1.53–1.55	1.54(1,3)	157–162	160(2,3)	120–129	125(4,3)
C,N—O—H(not sq)···O=C(SQ <sup>2−</sup> )	1.60–2.01	1.70(2,7)	139–175	162(12,7)	108–135	121(10,7)
O—H(water)···O=C(SQ <sup>2−</sup> )	1.68–2.00	1.82(6,49)	147–179	166(8,49)	98–144	120(11,49)
N <sup>+</sup> (ar)—H···O=C(SQ <sup>2−</sup> )	1.52–1.73	1.62(6,11)	155–174	166(6,11)	110–132	119(7,11)
N <sup>+</sup> —H···O=C(SQ <sup>2−</sup> )	1.67–2.08	1.83(10,41)	134–175	161(11,41)	102–144	119(9,41)
N <sup>0</sup> —H···O=C(SQ <sup>2−</sup> )	1.65–2.28	1.87(14,39)	142–178	164(9,39)	88–175	124(15,39)
Comparison data						
O—H(COOH)···O=C(COO <sup>−</sup> )	1.44–1.91	1.57(7,1072)	119–180	169(8,1072)	95–177	121(9,1072)
O—H(COOH)···O=C(COOH)	1.48–2.19	1.67(5,2160)	122–180	171(7,2160)	105–176	124(8,2160)
O—H(water)···O=C(COO <sup>−</sup> )	1.59–2.30	1.84(9,1929)	120–180	164(9,1929)	80–170	124(8,1929)
N <sup>+</sup> (ar)—H···O=C(COO <sup>−</sup> )	1.50–1.93	1.67(9,410)	132–180	168(9,410)	83–167	117(9,410)
N <sup>+</sup> —H···O=C(COO <sup>−</sup> )	1.56–2.24	1.84(11,4090)	120–180	163(10,4090)	73–175	123(14,4090)
N <sup>0</sup> —H···O=C(COO <sup>−</sup> )	1.64–2.35	1.93(13,1144)	122–180	161(12,1144)	81–177	125(15,1144)

directionality of approach to O-acceptors, the mean values of  $\theta(\text{DHA})$  are included as an effective indicator of the overall hydrogen-bond linearity within a given distribution: the closer the mean  $\theta(\text{DHA})$  is to 180°, the greater the tendency for the distribution to approach linearity. A few of the O—H···O hydrogen bonds characterized in Table 2 involve the formal C=O acceptors of H<sub>2</sub>SQ, HSQ<sup>−</sup> and SQ<sup>2−</sup> taken together; these acceptors are denoted collectively as O=C(sq) in Table 2 and in the ensuing discussion.

#### 4.1. Squaric acid [H<sub>2</sub>SQ (V)]

The O—H donors of the nine structures containing the parent squaric that pass the acceptance criteria of §2 are always used in hydrogen-bond formation, and these hydrogen bonds universally form to the O=C acceptors of squaric acid itself or to those of co-crystallized HSQ<sup>−</sup> or SQ<sup>2−</sup> anions. Thus these O—H(H<sub>2</sub>SQ)···O=C(sq) hydrogen bonds are the only meaningful subset given for parent squaric acid in Table 2. As expected from the pK<sub>a</sub> values (see §1), these bonds are short [mean  $d(\text{HA}) = 1.54$  (5) Å], have a remarkably linear  $\theta(\text{DHA})$  distribution, and approach the O-acceptor along a lone-pair direction. These data should be compared with CSD averages for COOH···O=C(COO<sup>−</sup>) and COOH···O=C(COOH) hydrogen bonds presented in Table 2 for comparison

purposes. Clearly the O—H(H<sub>2</sub>SQ)···O=C(sq) hydrogen bonds are of comparable length (strength) to the COOH···O=C(COO<sup>−</sup>) subset and are shorter (stronger) than the COOH···O(OH)C bonds. While the O=C acceptors of squaric acid are largely occupied by O—H donors from H<sub>2</sub>SQ itself, they also form hydrogen bonds with four different N<sup>+</sup>—H donors and the mean geometry for these bonds is also included in Table 2. Additionally, the O(hydroxy) of H<sub>2</sub>SQ accepts hydrogen bonds from water (once, at a H···A distance of 2.18 Å) and from N<sup>+</sup>—H donors (twice, at 2.36 and 2.57 Å) and with rather non-linear  $\theta(\text{DHA})$  values. It would appear that these bonds to O(hydroxy) are secondary to the much stronger O—H(H<sub>2</sub>SQ)···O=C(sq) that define the structural frameworks in all squaric acid structures.

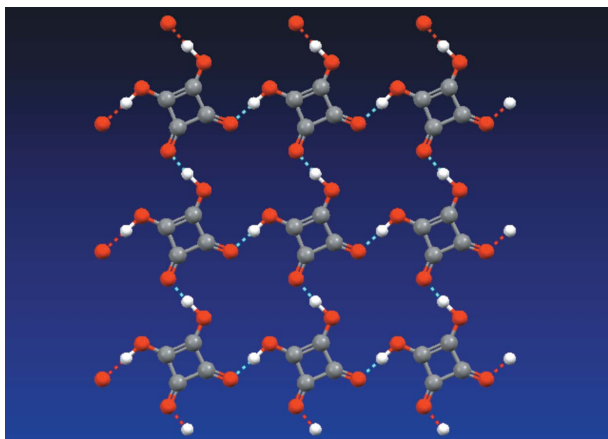
A 20-membered hydrogen-bonded ring motif, graph-set  $R_4^4(20)$ , is formed by four H<sub>2</sub>SQ molecules in both the monoclinic and tetragonal polymorphs of the parent acid, as shown in Fig. 1 [see *e.g.* CSD reference code KECYBU11 (Semmingsen, 1973) and others in the KECYBU family (Semmingsen *et al.*, 1977, 1995; Hollander *et al.*, 1977)]. This motif is preserved in SOCLEC (Kolev *et al.*, 2008) and TURPOL (Mathew *et al.*, 2002), which are both ammonium salts of a co-crystal of H<sub>2</sub>SQ with its monoanion (HSQ<sup>−</sup>); the ammonium ion, together with an additional water molecule in SOCLEC, serves to link the almost planar  $R_4^4(20)$  rings in these

structures. Other squaric acid structures also contain  $\text{HSQ}^-$  and form dimers that are then linked by hydrogen bonds from N-containing counterions.

#### 4.2. Hydrogen squarate monoanion [ $\text{HSQ}^-$ (VI)]

All of the O—H donors of the 78 independent  $\text{HSQ}^-$  fragments are involved in hydrogen bonding. The largest percentage (90%) form hydrogen bonds to O=C groups of  $\text{HSQ}^-$  ions or occasionally to O=C acceptors of  $\text{H}_2\text{SQ}$  or  $\text{SQ}^{2-}$  where they are also present in the structure. These 70 hydrogen bonds span a narrow distance range (1.42–1.62 Å) and are averaged together in Table 2 following removal of five obvious outliers. Two further O—H( $\text{HSQ}^-$ )···O bonds form to an amide O=C acceptor and to O-water. Only six hydrogen bonds are formed to N-acceptors, all to bipyridinium N atoms, for which the mean geometry is also shown in Table 2.

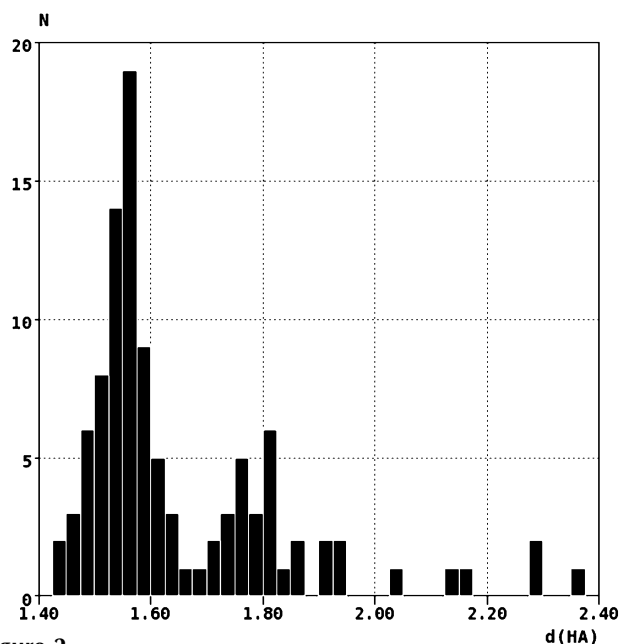
The C=O acceptors of  $\text{HSQ}^-$  are topologically and chemically different: O1 and O3 are vicinal to the OH substituent and are topologically equivalent while O2, diagonally opposite to the OH substituent, is unique. As noted by Bertolasi *et al.* (2001), O1 and O3 are both able to carry the negative charge in Kekulé resonance forms of  $\text{HSQ}^-$ , while this is not possible for O2. This factor is reflected in the CSD searches, where all 78 fragments form hydrogen bonds through one of the topologically equivalent O1 or O3 acceptors, while only 47 (60%) form hydrogen bonds using the unique O2 acceptor. However, these differences are not clearly reflected in the observed hydrogen-bond distances. The O1/O3 acceptors form 71 hydrogen bonds to squarate-OH ranging from 1.45–1.65 Å, with a mean  $d(\text{HA})$  of 1.54(5,71) Å (in the nomenclature of Table 1). There are only seven hydrogen bonds involving squarate-OH donors and O2 ( $\text{HSQ}^-$ ) acceptors. Their range (1.54–1.63 Å) lies within the O1/O3 acceptor range and average to 1.58(3,7) Å. Thus, the three C=O acceptors of  $\text{HSQ}^-$  are treated as equivalent in generating the statistical data in Table 2.



**Figure 1**  
The  $R_4^2(20)$  hydrogen-bonded ring motif observed in parent squaric acid (e.g. the monoclinic KECYBU11; Semmingsen, 1973) and in other polymorphs.

A histogram of general O—H···O=C ( $\text{HSQ}^-$ ) hydrogen bonds is shown in Fig. 2. The distribution is bimodal, with the larger peak at shorter  $d(\text{HA})$  (1.45–1.65 Å) arising from the O—H( $\text{HSQ}^-$ )···O=C( $\text{HSQ}^-$ ) hydrogen bonds, together with a few O—H( $\text{H}_2\text{SQ}$ )···O=C( $\text{HSQ}^-$ ) of very similar length, as discussed above. The mean H···O distance in these bonds is 1.55 (5) Å. This value is almost identical to the averages given in Table 2 for squaric acid O—H groups [O—H( $\text{H}_2\text{SQ}$ )···O=C(sq)] and for the unstrained comparator group O—H( $\text{COOH}$ )···O=C( $\text{COO}^-$ ). The second (and smaller) peak in Fig. 2 (bonds longer than 1.70 Å) arises from O—H(water)···O=C( $\text{HSQ}^-$ ) hydrogen bonds. This peak has a long tail as reflected in the H···O distance range shown in Table 2, but there is no good chemical or statistical reason to exclude these longer bonds from the overall average H···O distance of 1.89 (17) Å. This value is comparable to the mean value for O—H(water)···O=C( $\text{COO}^-$ ) hydrogen bonds of 1.84 (9) Å also given for comparison in Table 2.

In 95% of  $\text{HSQ}^-$  structures the counterion contains a protonated nitrogen ( $\text{N}^+\text{—H}$ ) donor; in addition 21 of these structures (30%) also contain amino-type  $\text{N}^0\text{—H}$  donors (three-coordinate N with zero charge). Hence there is significant N—H···O=C( $\text{HSQ}^-$ ) hydrogen bonding. These bonds form a broad overall distribution from 1.50 Å to around 2.40 Å, but inspection shows that those bonds involving  $\text{N}^+(\text{aromatic})\text{—H}$  donors occur at the shorter end of the distribution. The remainder of the distribution has been subdivided into bonds involving  $\text{N}^+(\text{non-aromatic})\text{—H}$  and amino-type  $\text{N}^0\text{—H}$  donors (as defined above). The geometrical data provided in Table 2 suggests that the  $\text{N}^+(\text{aromatic})\text{—H}\cdots\text{O}=\text{C}(\text{HSQ}^-)$  bonds are stronger than those involving the other N—H donor types, which have comparable mean



**Figure 2**  
Distances  $d(\text{HA})$  (Å) for O—H(any)···O=C ( $\text{HSQ}^-$ ) hydrogen bonds showing the bimodal distribution arising from the shorter O—H( $\text{HSQ}^-$ )···O=C( $\text{HSQ}^-$ ) bonds and the slightly longer O—H(water)···O=C( $\text{HSQ}^-$ ) bonds (see text and Table 1).

$d(\text{HA})$  values. This trend of hydrogen-bond distances agrees well with comparison data (Table 2) generated for hydrogen bonds formed to  $\text{COO}^-$  by these three  $\text{N}-\text{H}$  donor types.

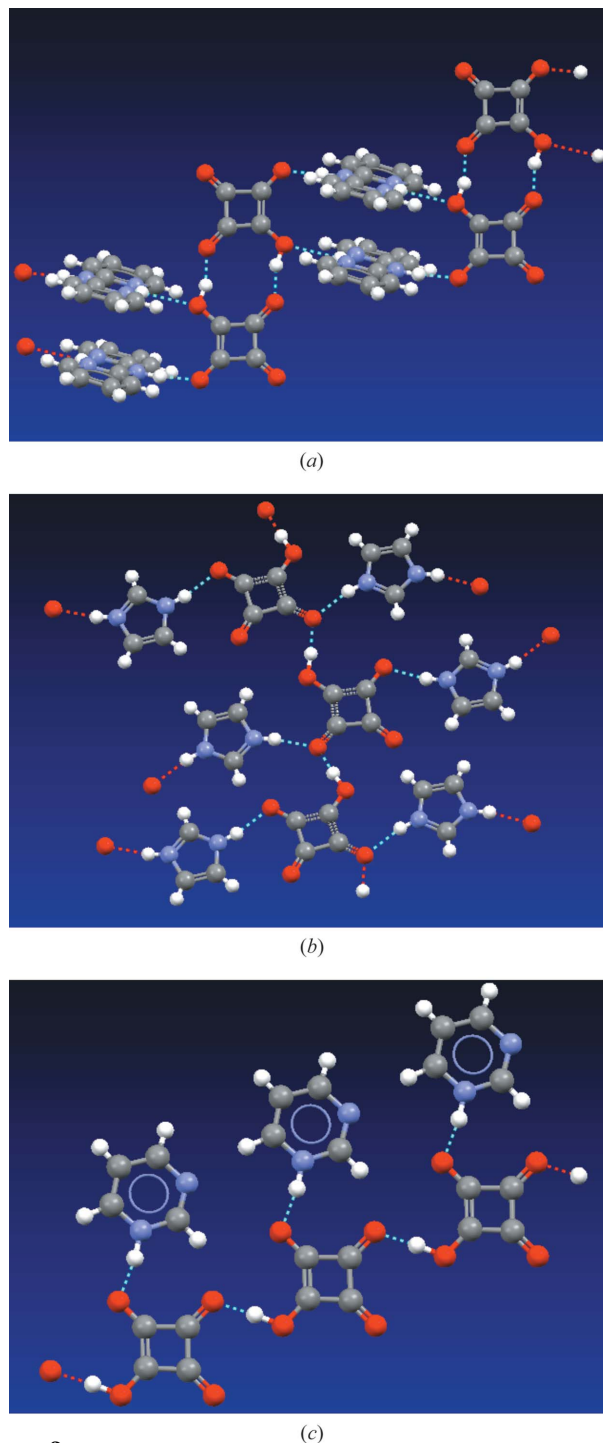
The hydroxy O atoms of  $\text{HSQ}^-$  also act as acceptors in a few structures, almost exclusively with  $\text{N}-\text{H}$  donors: only one  $\text{O}-\text{H}\cdots\text{O}(\text{hydroxy})$  hydrogen bond exists in structures with ordered  $\text{O}-\text{H}$  H atoms. This involves an  $\text{O}(\text{water})-\text{H}$  donor at  $d(\text{HA}) = 2.21 \text{ \AA}$ . The nine  $\text{N}-\text{H}\cdots\text{O}-\text{H}(\text{HSQ}^-)$  bonds involve both  $\text{N}^+-\text{H}$  and uncharged amino-type ( $\text{N}^0-\text{H}$ ) donors having a  $d(\text{HA})$  range from 2.03 to 2.70  $\text{ \AA}$ , and with  $\theta(\text{DHA})$  deviating quite significantly from linearity.

Almost half (47%) of the  $\text{HSQ}^-$  structures contain a dimer synthon (Desiraju, 1995). These dimers are almost always linked into chains or other extended structures *via*  $\text{N}-\text{H}\cdots\text{O}$  hydrogen bonds from counterions, as in BOLFUD (Bock *et al.*, 1998; Fig. 3*a*) where the counterion is pyrid-2-yl(2'-pyridinium)amine. As might be expected by analogy with the carboxylic acids, the other principal arrangement adopted by  $\text{HSQ}^-$  is the *catena* synthon (13/62 = 21%). Typical *catena* structures form in TIWWUR (Karle *et al.*, 1996) with imidazolium counterions (Fig. 3*b*) and in LUVTOL (Mattes *et al.*, 2003) with pyrimidinium counterions (Fig. 3*c*). All of these synthons use the  $\text{C}-\text{O}-\text{H}$  donor of  $\text{HSQ}^-$  and the dimers always use one of the topologically equivalent O1/O3 acceptor atoms (see above). The chains, though, can form using either an O1/O3 acceptor as in TIWWUR (Fig. 3*b*), or the topologically unique O2 acceptor (IV) as in LUVTOL (Fig. 3*c*). In fact, there are eight examples of chain formation *via* an O1/O3 acceptor and five examples of the use of O2. This factor contributes to the lower overall use of the O2 acceptor as discussed above. In a few structures that involve  $\text{HSQ}^-$  cocrystallized with either  $\text{H}_2\text{SQ}$  or  $\text{SQ}^{2-}$  the  $R_4^4(20)$  rings that are commonly found in  $\text{H}_2\text{SQ}$  structures (see above) are also observed.

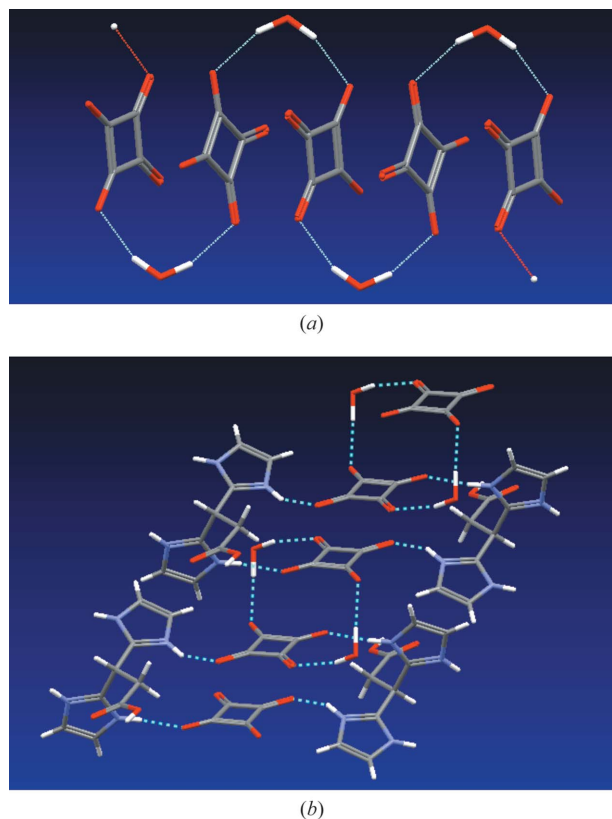
#### 4.3. Squarate dianion [ $\text{SQ}^{2-}$ , (II)]

The  $\text{SQ}^{2-}$  dianion with four equivalent  $\text{C}=\text{O}$  acceptors relies on other co-crystallized components to provide H-donors. Water molecules and molecules containing  $\text{N}^+-\text{H}$  and uncharged amino-type ( $\text{N}^0-\text{H}$ ) donors commonly occur so that all  $\text{SQ}^{2-}$  are able to form hydrogen bonds. Rather few of these hydrogen bonds involve  $\text{C}-\text{OH}$  donors, with just three examples arising from co-crystallized examples of the parent acid,  $\text{H}_2\text{SQ}$ . Their  $d(\text{HA})$  distances are closely comparable to other  $\text{C}-\text{O}-\text{H}(\text{sq})\cdots\text{O}=\text{C}(\text{sq})$  distances;  $d(\text{HA})$  values for  $\text{C}-\text{O}-\text{H}\cdots\text{O}=\text{C}(\text{SQ}^{2-})$  where  $\text{C}-\text{O}-\text{H}$  is 'non-squarate' are slightly but significantly longer. The largest distribution of  $\text{O}-\text{H}\cdots\text{O}=\text{C}(\text{SQ}^{2-})$  bonds arises from  $\text{O}-\text{H}$  (water) donors where the hydrogen-bond geometry (Table 2) is highly comparable to that occurring in  $\text{O}-\text{H}(\text{water})\cdots\text{O}=\text{C}(\text{HSQ}^-)$  and in  $\text{O}-\text{H}(\text{water})\cdots\text{O}=\text{C}(\text{COO}^-)$  systems. Hydrogen bonds from  $\text{N}^+-\text{H}$  and uncharged amino-type ( $\text{N}^0-\text{H}$ ; Table 2) have again been subdivided as for the monoanion ( $\text{HSQ}^-$ ) and for the acyclic comparator  $\text{COO}^-$ , and with similar overall results.

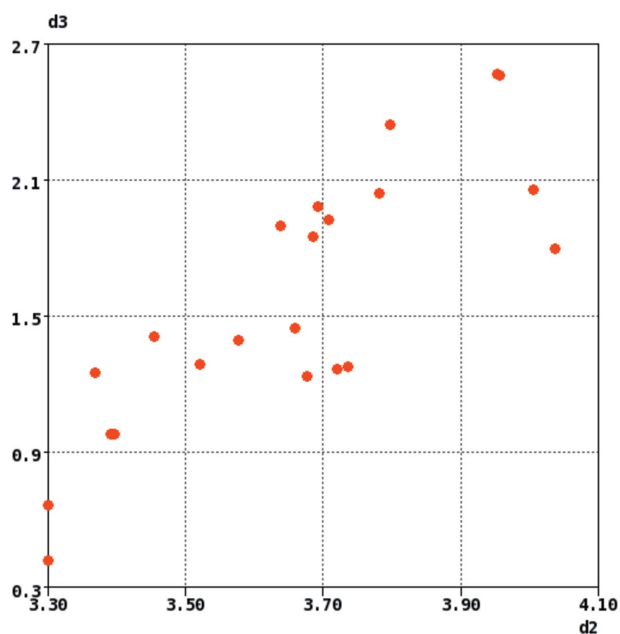
The most obvious feature of the packing of the aromatic  $\text{SQ}^{2-}$  dianions is the occurrence of stacked rings in 18 of the 46 structures. The stacking may be characterized by  $d_1$ , the perpendicular distance between the ring planes, and  $d_2$ , the distance between ring centroids, from which the offset from a



**Figure 3** Dimer and *catena* synthons in the crystal structures of  $\text{HSQ}^-$  compounds: (a) the  $\text{HSQ}^-$  dimer synthon in BOLFUD (Bock *et al.*, 1998) where the counterion is pyrid-2-yl(2'-pyridinium)amine; (b) the  $\text{HSQ}^-$  *catena* synthon in TIWWUR (Karle *et al.*, 1996) where the counterion is imidazolium; (c) the  $\text{HSQ}^-$  *catena* synthon in LUVTOL (Mattes *et al.*, 2003) where the counterion is pyrimidinium.



**Figure 4**  
Ring stacking in  $\text{SQ}^{2-}$  structures: (a) potassium diketocyclobutenediolate monohydrate (KKCUBO: Macintyre & Werkema, 1964) and (b) 3,3-bis(imidazolium-2-yl)propionic acid squarate monohydrate (NEHKUH: Akhriff *et al.*, 2006).



**Figure 5**  
Scatterplot of the inter-centroid distance ( $d_2$ ) and the offset from perfect centroid overlay ( $d_3$ ) in the 18  $\text{SQ}^{2-}$  structures which exhibit ring stacking.

**Table 3**

Propensities for single and multiple hydrogen-bond formation by the donor and acceptor O atoms of squaric acid ( $\text{H}_2\text{SQ}$ ) and its anions ( $\text{HSQ}^-$  and  $\text{SQ}^{2-}$ ).

These data have been compiled by analysis of the November 2012 release of the CSD but using the structural acceptance criteria in §2.  $N_s$  is the number of structures exhibiting the structural feature described and  $P$  is the corresponding propensity. All hydrogen bonds considered involve O or N as an acceptor or donor.

	$N_s$	$P$ (%)
<b><math>\text{H}_2\text{SQ}</math> (14 structures)</b>		
At least one O–H donor used	14	100
Both O–H donors used	14	100
At least one C=O acceptor used	11	79
Both C=O acceptors used	9	64
<b><math>\text{HSQ}^-</math> (70 structures)</b>		
O–H donor used	70	100
C=O1 or C=O3 acceptor used	69	98
C=O2 acceptor used	41	59
Both C=O1 and C=O3 acceptors used	57	81
Only C=O1 or C=O3 used	9	13
Only C=O2 used	0	0
C=O3 or C=O3 and C=O2 used	41	49
C=O3 and C=O3 both used but not C=O2	19	27
<b><math>\text{SQ}^{2-}</math> (44 structures)</b>		
C=O1 used ( <i>i.e.</i> any C=O)	44	100
C=O1 and C=O2 used	41	93
C=O1 and C=O3 used	43	98
C=O1, C=O2 and C=O3 used	40	91
All four C=O acceptors used	35	80
Exactly three C=O acceptors used	5	11
Exactly two C=O acceptors used	4	9
Exactly 1 C=O acceptor used	0	0

perfect overlay of ring centroids,  $d_3$ , can readily be calculated. The simplest example is KKCUBO (Macintyre & Werkema, 1964): potassium- $\text{SQ}^{2-} \cdot \text{H}_2\text{O}$  illustrated in Fig. 4(a). Here, the  $\text{SQ}^{2-}$  stack is stabilized by O(water)–H...O=C( $\text{SQ}^{2-}$ ) hydrogen bonds and there is almost perfect centroid overlay:  $d_1 = 3.27 \text{ \AA}$ ,  $d_2 = 3.30 \text{ \AA}$  and  $d_3 = 0.42 \text{ \AA}$ . However, in this system the stacked rings are rotated with respect to each other by exactly  $45^\circ$ . Apart from SADREW (Ivanova & Spittler, 2010), which has a ring rotation of  $19^\circ$ , all other systems have a stacked ring rotation of exactly  $0^\circ$ . A more complex structural example is provided by NEHKUH (Akhriff *et al.*, 2006): 3,3-bis(imidazolium-2-yl)propionic acid squarate monohydrate, where the  $\text{SQ}^{2-}$  stack is stabilized by hydrogen bonds from O(water)–H and N(imidazolium)–H, and is characterized by  $d_1 = 3.17 \text{ \AA}$ ,  $d_2 = 3.74 \text{ \AA}$  and  $d_3 = 1.98 \text{ \AA}$ . The scatterplot of Fig. 5 shows the extent of the inter-centroid distance ( $d_2$ ) and the offset from perfect centroid overlay ( $d_3$ ) in the 18  $\text{SQ}^{2-}$  structures which exhibit ring stacking.

#### 4.4. Hydrogen-bond propensities

The analysis is extended in Table 3, which records the propensities for hydrogen-bond formation for various individual donor and acceptor O atoms and their combinations. Table 3 confirms that the O–H donors in  $\text{H}_2\text{SQ}$  and  $\text{HSQ}^-$  are highly potent and donate hydrogen bonds in 100% of cases, but the C=O acceptors in  $\text{H}_2\text{SQ}$  are slightly less potent



than those in  $\text{HSQ}^-$  and  $\text{SQ}^{2-}$  where propensities are 100% or close to it. For  $\text{HSQ}^-$ , however, there is a clear decrease in the overall hydrogen-bond propensity of the unique  $\text{C}=\text{O}$  acceptor to 59%, and hydrogen bonds never form alone to this acceptor. These observations reflect the resonance effects referred to earlier. The  $\text{SQ}^{2-}$  dianion is notable for the multiple use of its  $\text{C}=\text{O}$  acceptors, with 80% of structures using all four acceptors in hydrogen-bond formation.

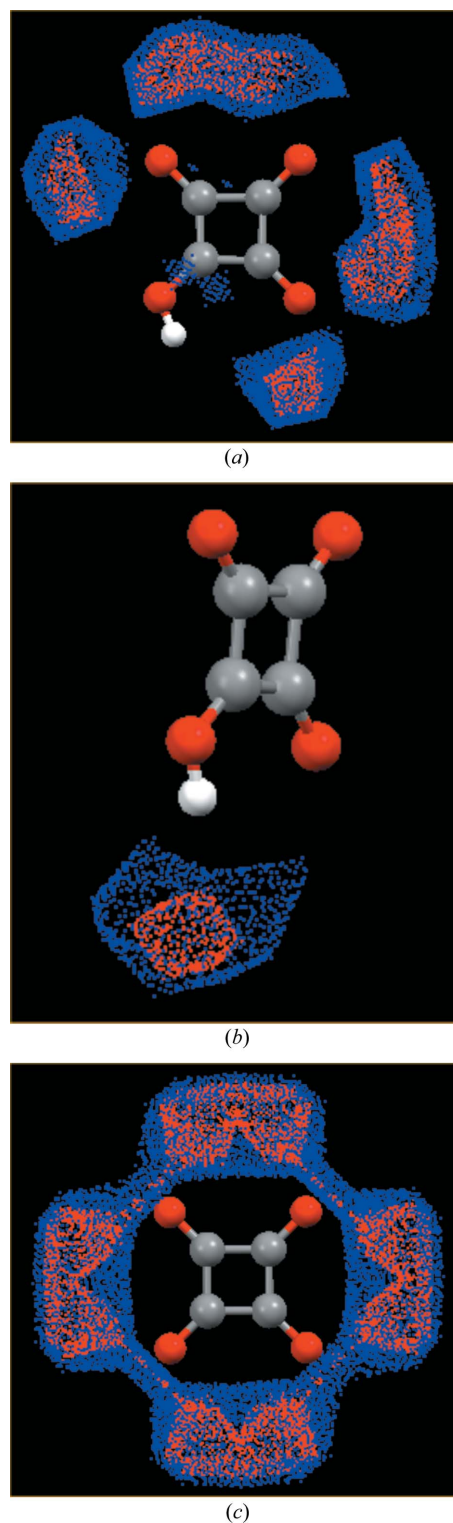
## 5. IsoStar representations of the hydrogen-bond landscapes

While insufficient data are available to derive useful plots using IsoGen for parent squaric acid (V), it has been possible to generate hydrogen-bond landscapes for both  $\text{HSQ}^-$  (VI) and  $\text{SQ}^{2-}$  (II). Fig. 6(a) shows the contoured distribution of H atoms from N–H and O–H donors around  $\text{HSQ}^-$ . The peaks show strong hydrogen-bond density approaching the terminal O-acceptors along the lone-pair directions. It is visually clear also that the density of hydrogen bonds around the O2 atom (diagonally opposite the –OH ring-substituent) is lower than that around O1 and O3. Fig. 6(b) shows a similar contoured distribution of acceptor O atoms (defined as terminal O atoms connected to C *via* a single, double or delocalized double bond) around the OH donor of  $\text{HSQ}^-$ . Finally, Fig. 6(c) shows the distribution of N–H and O–H donors around the four equivalent O acceptors of  $\text{SQ}^{2-}$ . The distribution has been symmetrized automatically within the IsoGen software and, again, the lone pair directionality of the N–H or O–H approach to acceptor-O is clearly displayed. We stress that the plots of Fig. 6 are selected two-dimensional static images: once generated by IsoGen these plots are fully interactive in terms of content, rotation and presentation. The plots in Fig. 6 do, however, show the importance of the IsoStar methodology in providing a rapid visual overview of information about non-bonded contacts that is vital in drug discovery and many other fields, and both  $\text{HSQ}^-$  and  $\text{SQ}^{2-}$  are now included in IsoStar as additional central groups from November 2012.

## 6. DFT-D and IMPT results

Initial models of the molecular structures of methanol, acetone,  $\text{HSQ}^-$ ,  $\text{SQ}^{2-}$  and the three isomers of  $\text{H}_2\text{SQ}$  [*ZZ* (VII), *EZ* (VIII), *EE* (IX)], were taken from X-ray crystal structures in the CSD. Isolated monomers were then fully optimized using the B97-D/cc-pVTZ model with *GAUSSIAN09* (Frisch *et al.*, 2009). The basis-set superposition error correction was not applied, as recommended for the B97-D functional (Grimme, 2006). The relative energies of the  $\text{H}_2\text{SQ}$  isomers were found to be (in  $\text{kJ mol}^{-1}$ ): *ZZ* (global minimum) = 0.0; *EZ* = 10.5, *EE* = 35.3, in good agreement with other published results (see *e.g.* Lunelli *et al.*, 1996; Zhou *et al.*, 2000; Spassova *et al.*, 2000). Dimer geometries were generated from the gas-phase optimized monomers. Dimers were constructed keeping the interacting molecules (the *EZ* isomer of  $\text{H}_2\text{SQ}$ ,  $\text{HSQ}^-$  or  $\text{SQ}^{2-}$  and methanol or acetone) coplanar and with  $\theta(\text{DHA})$  and  $\varphi(\text{HAR})$  at their optimal values of 180 and 120°,

respectively. These dimer geometries were then allowed to fully optimize in the gas phase at the same level of theory. Interaction energies ( $E_{\text{int}}$ ) for each *AB* dimer were calculated



**Figure 6** Contoured IsoStar plots of (a) the distribution of N–H and O–H donors around the terminal-O acceptors of  $\text{HSQ}^-$ , (b) the distribution of C=O acceptors around the O–H donor of  $\text{HSQ}^-$  and (c) the symmetrized distribution of N–H and O–H donors around the terminal-O acceptors of  $\text{SQ}^{2-}$ .

**Table 4**

DFT-D energies ( $E_{\text{int}}$ , kJ mol<sup>-1</sup>) of optimized hydrogen-bonded dimers formed between the *EZ* isomer of H<sub>2</sub>SQ, HSQ<sup>-</sup> or SQ<sup>2-</sup> (component *A*) and methanol or acetone (component *B*).

Data for the H<sub>2</sub>SQ:H<sub>2</sub>SQ cyclic dimer (*X*) are also included. Single-point IMPT energies (kJ mol<sup>-1</sup>) for hydrogen bonds involving CH<sub>3</sub>COO<sup>-</sup>, acetone and methanol (no values are available for CH<sub>3</sub>COOH) are taken from IsoStar (Bruno *et al.*, 1997) and values for urea are taken from Wood *et al.* (2008); these data are included for comparison purposes. Hydrogen-bond distances [*H* to acceptor, *d*(*HA*)] are in Å.

Component <i>A</i>	Component <i>B</i>	$E_{\text{int}}$	<i>d</i> ( <i>HA</i> )
H <sub>2</sub> SQ ( <i>E</i> hydroxy-H)	Acetone —C=O	-90	1.66
H <sub>2</sub> SQ ( <i>Z</i> hydroxy-H)	Acetone —C=O	-77	1.60
HSQ <sup>-</sup> (hydroxy-H)	Acetone —C=O	-71	1.75
H <sub>2</sub> SQ (carbonyl-O)	Methanol —OH	-59	2.07
HSQ <sup>-</sup> (carbonyl-O1,O3)	Methanol —OH	-66	1.77
HSQ <sup>-</sup> (carbonyl-O2)	Methanol —OH	-61	1.83
SQ <sup>2-</sup> (carbonyl-O)	Methanol —OH	-140	1.59
H <sub>2</sub> SQ	H <sub>2</sub> SQ	-124	1.57
CH <sub>3</sub> COO <sup>-</sup> (carbonyl-O)	Methanol —OH	-75	1.71†
CH <sub>3</sub> COO <sup>-</sup> (carbonyl-O)	Water —OH	-72	1.92†
Acetone (carbonyl-O)	Methanol —OH	-25	1.94†
Urea (carbonyl-O)	Methanol —OH	-33	1.90†

† IMPT values: dimer not optimized, *d*(*HA*) indicates the fixed point for the IMPT energy calculation.

as the difference between the optimized dimer energy ( $E_{AB}$ ) and the energies of the monomers ( $E_A$  and  $E_B$ ) using  $E_{\text{int}} = E_{AB} - E_A - E_B$ .

Interaction energies and optimized hydrogen-bond distances for a variety of dimers involving SQ species and acetone or methanol are collected in Table 4, which also includes similar data for the cyclic H<sub>2</sub>SQ:H<sub>2</sub>SQ dimer (*X*). Some comparison single-point IMPT hydrogen-bond energy data for CH<sub>3</sub>COO<sup>-</sup>, acetone and methanol taken from IsoStar (Bruno *et al.*, 1997), and for hydrogen bonds from methanol-OH to urea-C=O (Wood *et al.*, 2008), are also included in Table 4. Obviously one cannot view the optimized DFT-D dimer energies as arising solely from their respective C=O...H—O bonds, but this bond is likely to make a major contribution to the attractive energy. Further, the DFT-D dimer energies cannot be compared with the single-point IMPT energies in a fully quantitative manner.

Despite these constraints it is clear that the C=O...H—O bonds involving SQ species are very strong. Thus, the OH donors in H<sub>2</sub>SQ and HSQ<sup>-</sup> form bonds to acetone with energies that are, even on a qualitative basis, much stronger than those arising from methanol-OH to the —C=O groups of acetone or urea: the 100% propensities for hydrogen-bond formation by SQ-OH (Table 3) is clearly reflected in the dimer energies of Table 4. Similarly, the bonds formed by SQ-C=O groups are also strong. The DFT-D dimer energies for the SQ-C=O...H—O (methanol) increase just a little in moving from H<sub>2</sub>SQ to HSQ<sup>-</sup>, and the dimer energies here are at a level that is comparable with the IMPT energies obtained for COO<sup>-</sup>...H—O (methanol or water) hydrogen bonds. However, there is a much more dramatic increase in  $E_{\text{int}}$  in

moving to the SQ<sup>2-</sup> dianion, and the DFT-D calculations generate *d*(*HA*) values that decrease smoothly from 2.07 Å (H<sub>2</sub>SQ) to 1.77–1.83 Å (HSQ<sup>-</sup>) and then to the very short 1.59 Å for SQ<sup>2-</sup>. The difference between the  $E_{\text{int}}$  values for dimers formed through the ‘resonance-preferred’ O1,O3 of HSQ<sup>-</sup> and dimers formed through the unique carbonyl (C=O2) is quite small at 5 kJ mol<sup>-1</sup>, and there is only a very small concomitant elongation in the optimized hydrogen-bond distance, *d*(*HA*), from 1.77 to 1.83 Å. While resonance is clearly a factor here, it may be that steric effects also mitigate against the formation of hydrogen bonds to C=O2 and, hence, to its rather low hydrogen-bond propensity of 59% (Table 3). More structural data on HSQ<sup>-</sup> compounds will be needed for a proper analysis of this issue.

The data for the H<sub>2</sub>SQ:H<sub>2</sub>SQ cyclic dimer (*X*) are in good agreement with the other calculations, with  $E_{\text{int}}$  being about double the value for dimers formed through a single hydrogen-bonded interaction. The optimized *d*(*HA*) value of 1.57 Å for this dimer may reflect the increased potency of the H<sub>2</sub>SQ-OH donor by comparison with the methanol-OH used in the general calculations.

## 7. Conclusion

The primary objective of this work was to examine the available data for squaric acid and its anions, H<sub>2</sub>SQ, HSQ<sup>-</sup> and SQ<sup>2-</sup>, with a view to their inclusion as new central groups in IsoStar (Bruno *et al.*, 1997), the knowledge base of intermolecular interactions, and this objective was accomplished in the November 2012 release of IsoStar, for the more populous anion datasets. While the IsoStar maps, exemplified in Fig. 6, define the landscape of nonbonded interactions, particularly their hydrogen-bonded interactions, the extended study has confirmed that SQ compounds are potent donors and acceptors, as defined by their hydrogen-bond geometries in available crystal structures and by the attractive energies calculated for their dimers with acetone and methanol as model acceptors and donors. Hydrogen bonds involving H<sub>2</sub>SQ and HSQ<sup>-</sup> are of similar strengths to those made by carboxylate COO<sup>-</sup> acceptors, while hydrogen bonds made by the dianion SQ<sup>2-</sup> are somewhat stronger. The study reinforces the value of SQ compounds as cocrystal formers (see, *e.g.*, Bertolasi *et al.*, 2001; Gilli *et al.*, 2001; Mathew *et al.*, 2002) and, particularly, as isosteric replacements for carboxylic acid and carboxylate functions (Meanwell, 2011; Marson, 2011). In the latter case, it is likely that the anionic species with the strongest hydrogen-bonding abilities will predominate under physiological conditions.

AJCC thanks the Netherlands Organisation for Scientific Research (NWO) for the award of a VENI grant.

## References

- Akhriif, Y., Server-Carrió, J., García-Lozano, J., Folgado, J. V., Sancho, A., Escrivà, E., Vitoria, P. & Soto, L. (2006). *Cryst. Growth Des.* **6**, 1124–1133.

- Allen, F. H. (1986). *Acta Cryst.* **B42**, 515–522.
- Allen, F. H. (2002). *Acta Cryst.* **B58**, 380–388.
- Allen, F. H. & Bruno, I. J. (2010). *Acta Cryst.* **B66**, 380–386.
- Allen, F. H., Galek, P. T. & Wood, P. A. (2010). *Cryst. Rev.* **16**, 169–195.
- Allen, F. H., Groom, C. R., Liebeschuetz, J. W., Bardwell, D. A., Olsson, T. S. G. & Wood, P. A. (2012). *J. Chem. Inf. Model.* **52**, 857–866.
- Baudy, R. B. *et al.* (2009). *J. Med. Chem.* **52**, 771–778.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
- Bertolasi, V., Gilli, P., Ferretti, V. & Gilli, G. (2001). *Acta Cryst.* **B57**, 591–598.
- Bock, H., Schödel, H., Van, T. T. H., Dienelt, R. & Gluth, M. (1998). *J. Prakt. Chem.* **340**, 722–732.
- Bruno, I. J., Cole, J. C., Edgington, P. R., Kessler, M., Macrae, C. F., McCabe, P., Pearson, J. & Taylor, R. (2002). *Acta Cryst.* **B58**, 389–397.
- Bruno, I. J., Cole, J. C., Lommerse, J. P., Rowland, R. S., Taylor, R. & Verdonk, M. L. (1997). *J. Comput. Aided Mol. Des.* **11**, 525–537.
- Caserio, R. J. (1987). *Arch. Dermatol.* **123**, 1036–1041.
- Cohen, S., Lacher, J. R. & Park, J. D. (1959). *J. Am. Chem. Soc.* **81**, 3480.
- Desiraju, G. R. (1995). *Angew. Chem. Int. Ed.* **34**, 2311–2327.
- Frisch, M. J. *et al.* (2009). GAUSSIAN09, Revision A.1. Gaussian Inc., Wallingford, CT, USA.
- Gilli, G., Bertolasi, V., Gilli, P. & Ferretti, V. (2001). *Acta Cryst.* **B57**, 859–865.
- Gilli, G. & Gilli, P. (2009). *The Nature of the Hydrogen Bond*, pp. 251–252. Oxford University Press.
- Grimme, S. (2006). *J. Comput. Chem.* **27**, 1787–1799.
- Hayes, I. & Stone, A. (1984). *Mol. Phys.* **53**, 83–105.
- Hollander, F. J., Semmingsen, D. & Koetzle, T. F. (1977). *J. Chem. Phys.* **67**, 4825–4831.
- Horiuchi, S., Kumai, R. & Tokura, Y. (2007). *Chem. Commun.* pp. 2321–2329.
- Ivanova, B. B. & Spittelner, M. (2010). *J. Phys. Chem. A*, **114**, 5099–5103.
- Karle, I. L., Ranganathan, D. & Haridas, V. (1996). *J. Am. Chem. Soc.* **116**, 7128–7133.
- Kolev, T., Seidel, R. W., Koleva, B. B., Spittelner, M., Mayer-Figge, H. & Sheldrick, W. S. (2008). *Struct. Chem.* **19**, 101–107.
- Kolev, T., Wortmann, R., Spittelner, M., Sheldrick, W. S. & Mayer-Figge, H. (2004). *Acta Cryst.* **E60**, o1449–o1450.
- Lunelli, B., Roversi, P., Ortoleva, E. & Destro, R. (1996). *J. Chem. Soc. Faraday Trans.* **92**, 3611–3623.
- MacDonald, D. J. (1968). *J. Org. Chem.* **33**, 4559–4560.
- Macintyre, W. M. & Werkema, M. S. (1964). *J. Chem. Phys.* **40**, 3563–3568.
- Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., Rodriguez-Monge, L., Taylor, R., van de Streek, J. & Wood, P. A. (2008). *J. Appl. Cryst.* **41**, 466–470.
- Marson, C. M. (2011). *Chem. Soc. Rev.* **40**, 5514–5533.
- Mathew, S., Paul, G., Shivasankar, K., Choudhury, A. & Rao, C. (2002). *J. Mol. Struct.* **641**, 263–279.
- Mattes, R., Ebbing, J., Gruss, A., Koppe, J. & Majcher, K. (2003). *Z. Naturforsch. B*, **58**, 27–35.
- Meanwell, N. A. (2011). *J. Med. Chem.* **54**, 2529–2591.
- Semmingsen, D. (1973). *Acta Chem. Scand.* **27**, 3961–3969.
- Semmingsen, D., Hollander, F. J. & Koetzle, T. F. (1977). *J. Chem. Phys.* **66**, 4406–4412.
- Semmingsen, D., Tun, Z., Nelmes, R. J., McMullan, R. K. & Koetzle, T. F. (1995). *Z. Kristallogr.* **210**, 934–947.
- Silverberg, N. B., Lim, J. K., Paller, A. S. & Mancini, A. J. (2000). *J. Am. Acad. Dermatol.* **42**, 803–808.
- Spassova, M., Kolev, T., Kanev, I., Jacquemin, D. & Champagne, B. (2000). *J. Mol. Struct.* **528**, 151–159.
- Sprenger, H. E. & Ziegenbein, W. (1966). *Angew. Chem. Int. Ed.* **5**, 894.
- Stone, A. J. (1993). *Chem. Phys. Lett.* **211**, 101–109.
- Sykes, R. A., McCabe, P., Allen, F. H., Battle, G. M., Bruno, I. J. & Wood, P. A. (2011). *J. Appl. Cryst.* **44**, 882–886.
- Terpetschnig, E. & Lakowicz, J. R. (1993). *Dyes Pigments*, **21**, 227–234.
- West, R., Niu, H., Powell, D. L. & Evans, M. V. (1960). *J. Am. Chem. Soc.* **82**, 6204–6205.
- West, R. & Powell, D. L. (1963). *J. Am. Chem. Soc.* **85**, 2577–2579.
- Wood, P. A., Allen, F. H. & Pidcock, E. (2009). *CrystEngComm*, **11**, 1563–1571.
- Wood, P. A., Pidcock, E. & Allen, F. H. (2008). *Acta Cryst.* **B64**, 491–496.
- Zhou, L., Zhang, Y., Wu, L. & Li, J. (2000). *J. Mol. Struct.* **497**, 137–144.