CrystEngComm

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *CrystEngComm*, 2025, **27**, 6314

Sources of symmetry in 'Blind Tests' crystal structures

Simon N. Black

Molecules create crystals *via* intermolecular interactions that embody long-range order. Previously, close approaches between the aromatic rings of halobenzene molecules, designated 'Symthons', were found to be highly effective in creating long-range order in their crystal structures. Here, analysis of 42 'blind test' crystal structures extends the application of 'Symthons' to fused rings, heterocycles and five-membered rings. This concept is further adapted to identify hydrogen-bonded approaches which create symmetry. Most of the 42 crystal structures contain at least one Symthon, some of which create long-range order. Other aromatic approaches and hydrogen bonds help to create symmetrical interactions, providing an alternative route to long-range order. In this dataset, hydrogen bonds were both less frequent and less effective at creating long-range order. However, hydrogen bonds were more effective at linking different molecules in salts, hydrates and cocrystals. Implications for nucleation, supramolecular synthons and graph sets are discussed briefly.

Received 17th June 2025,
Accepted 1st September 2025

DOI: 10.1039/d5ce00618j

rsc.li/crystengcomm

Introduction

Synthons are molecular fragments used to design and construct larger molecules in chemical synthesis. Supramolecular synthons[1,2] are energetically favourable close approaches between identical or different functional groups that help to create intermolecular interactions. Supramolecular synthons may exist in liquids, solutions, amorphous solids or crystals. Some supramolecular synthons link different molecules, some link identical molecules, and some may incorporate local symmetry. More recently, analysis of halobenzene crystal structures revealed 'Symthons' which are close approaches between aromatic rings that embody symmetry and are building blocks for crystal structures.[3–5]

Close aromatic approaches are characterized by heavy (non-hydrogen) atoms occupying hollows above and below aromatic rings.[6] In crystal structures of halobenzenes containing chlorine, bromine and/or iodine, the dominant approaches have specific geometries and embody translational symmetry in a variation of π–π stacking called 'Symthon I'.[3,4] Two other Symthons and their derivatives were more frequent in crystal structures of halobenzenes containing fluorine.[5] These three Symthons are displayed in Fig. 1. In Symthon I the molecules are parallel, embodying translation (grey) symmetry directly. Symthon II embodies either screw (green) or glide (magenta) symmetries, which repeat in 'zig-zag' motifs to give chains. In Symthon III the molecules are anti-parallel, embodying inversion (yellow) symmetry, without creating chains.

Symthons accounted for long-range order in at least one direction in 100 out of 111 halobenzene crystal structures. The associated intermolecular interaction energies increased with the number of halogen substituents.[5] Fig. 2 compares these interaction energies, calculated using Crystal Explorer,[7,8] with interaction energies for supramolecular synthons calculated previously using PIXEL.[2] This suggests a simple energy hierarchy: H-bond pairs > single H-bonds ~ Symthons I and III > C–H···O/Cl pairs ~ Symthon II.

The success of Symthons in explaining how molecules create long-range order in halobenzene crystal structures was qualified because there are few alternative ways in which halobenzene molecules can interact. A fairer test of the utility of Symthons was sought, requiring a small dataset of crystal structures, ideally selected independently, that better reflected the diversity of crystal structures within the Cambridge Structural Database

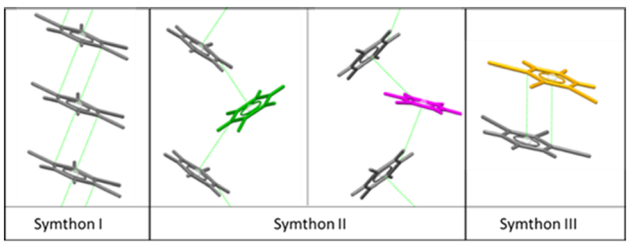

Fig. 1 Symthons, with their translation, screw, glide and inversion symmetries.

Department of Chemistry, Durham University, Durham, DH1 3LE, UK.
E-mail: simon.n.black@durham.ac.uk

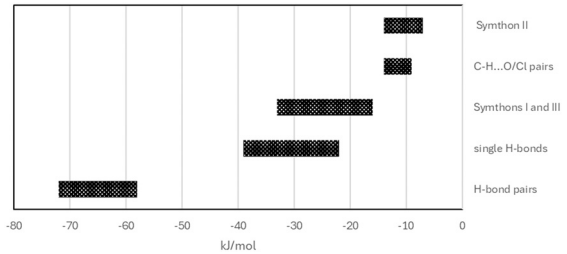

Fig. 2 Typical interaction energies for supramolecular synthons[2] and aromatic Symthons.[5]

(CSD). It was noted previously[5] that one of the mixed fluorobenzenes in the halobenzene dataset was molecule XIII in the fourth 'blind test' for crystal structure prediction.[9] This prompted exploration of all fifty experimental crystal structures that were determined as targets for the seven 'Blind Tests' and subsequently included in the CSD.[10–15]

The 'Blind Test' dataset

Table 1 gives all the crystal structures that constitute the blind test dataset - 'the dataset' used in this study. Eight crystal structures had $Z' > 1$ and were set aside for future study, with the consequence that compounds XI and XXIX are not represented. The remaining 31 compounds include molecules with a variety of H-bond donors and acceptors, as well as a variety of aromatic rings, sometimes within the same molecule. There are also several multi-component crystals, including salts, hydrates and co-crystals. Several compounds were also polymorphic.

The 42 crystal structures in the dataset were examined systematically for Symthons. The method used to search the crystal structures was based on that used previously for halobenzenes.[4] Hollows above and below each aromatic ring were considered 'occupied' if a heavy (non-carbon) atom was within 4 Å of the ring centroid and directly above the ring when viewed exactly perpendicular to the ring plane. In this dataset, some of these heavy atoms belonged to different aromatic rings or other functional groups, as discussed in more detail below. The contribution of these aromatic approaches to the creation of long-range order was also assessed – see SI for further details.

Aromatic approaches in the dataset

36 out of the 42 crystal structures contain a total of 82 aromatic rings, which is broadly consistent with the distribution of aromatic rings in the entire CSD.[16] Hence there are 82 × 2 = 164 hollows, each being a potential site for accepting an aromatic approach such as a Symthon. In halobenzenes studied previously, each molecule contains only one aromatic ring, so any approach between aromatic rings in crystal structures with $Z' = 1$ necessarily involved symmetry. In the dataset studied here there were many more options for occupying each hollow. Several molecules contain multiple aromatic rings, including five- and six-membered rings, heterocycles and fused rings - so close approaches could be between different aromatic rings, without embodying symmetry. Asymmetric approaches may also involve other functional groups – for example methyl···π approaches. In the four co-crystals in this dataset, both components contain aromatic rings: close approaches between them are necessarily asymmetric.

Symthons are defined by geometry and symmetry.[5] These definitions were also suitable without alteration for heterocycles, five-membered rings and individual rings that were fused together. This broader applicability greatly enhances the potential utility of Symthons. 26 Symthons were found, distributed across 21 crystal structures in the dataset. A summary is presented in Table 2. Selected Symthons are illustrated in Fig. 3–6 together with the corresponding interaction energies, $E_{\_tot}$ calculated using Crystal explorer.[7,8,17] Further details of these calculations are given in the SI.

Fig. 3 shows how the fused rings adopt Symthon I in XATMIP. The six fused aromatic rings are in a '3 × 2' configuration and this approach occupies all 12 hollows above and below the plane of the fused rings. Five other crystal structures also contain Symthon I, including the halobenzene SOXLEX01 as noted previously.[5]

Table 1 The blind test dataset, with compound numbers and REFCODES

Blind test (year published)	Compound number	Crystal structures	REFCODES/polymorphs
First (2000)	I–III, VII	5	XULDUD/XULDUD01, GUFJOG, QAMTAZ, JAYDUI
Second (2002)	IV–VI	4	BOQQUT, BOQWIN, UJIRIO/UJIRIO05
Third (2005)	VIII–X	3	PAHYON01, XATMIP, HAMTIZ01
Fourth (2009)	XII–XV	4	AXOSOW01, SOXLEX01, WIDBAO, WICZUF (co-crystal)
Fifth (2011)	XVI–XXI	9	OBEQUJ (zwitterion), OBEQOD, OBEQET, XATJOT (salt), OBEQIX KONTIQ/KONTIQ01/KONTIQ05/KONTIQ 06 (monohydrates)
Sixth (2016)	XXII–XXVI	7	NACJAF, XAFPAY/XAFPAY01/XAFPAY03, XAFQON (monohydrated salt), XAFQAZ (co-crystal), XAFQIH
Seventh (2024)	XXVII–XXVII, XXX–XXXIII	10	XIFZOF, OJIGOG01, MIVZIE and MIVZEA (1:1 and 2:1 co-crystals), ZEHFUR/ZEHFUR01/ZEHZUR02, JEKVII, ZEGWAN/ZEGWAN01 (salts)
Totals	31	42	

Table 2 Symthons in the dataset

Symthon	Frequency	Examples	Repeat (Å)
I	6	XATMIP (Fig. 3)	3.8–4.3
I (offset)	2	XIFZOF (Fig. 4)	9.3
		HAMTIZ01 (Fig. 5)	4.9
II	4	XULDUD (Fig. 6)	5.3–7.7
III	2	GUFJOG (×2, Fig. 6)	7.5
	8	XIFZOF (Fig. 4)	—
III (offset)	4	UJIRIO05	—

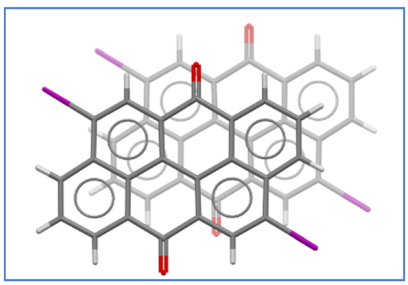

Fig. 3 Symthon I in XATMIP (−77.3 kJ mol^{-1}).

Fig. 4 Symthon III (left, −113.8 kJ mol^{-1}) and offset translation (right, −66.3 kJ mol^{-1}) in XIFZOF.

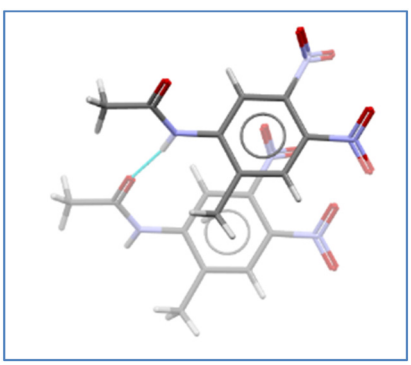

Fig. 5 Offset translation and H-bond in HAMTIZ01 (−46.0 kJ mol^{-1}).

XIFZOF (Fig. 4) contains a molecule with a row of five fused aromatic rings which adopts two different aromatic approaches. The approach on the right occupies two hollows below the central molecule and embodies an offset translation, giving a = 9.296. The next molecule in this chain (omitted for clarity) occupies the three hollows on the left above the central

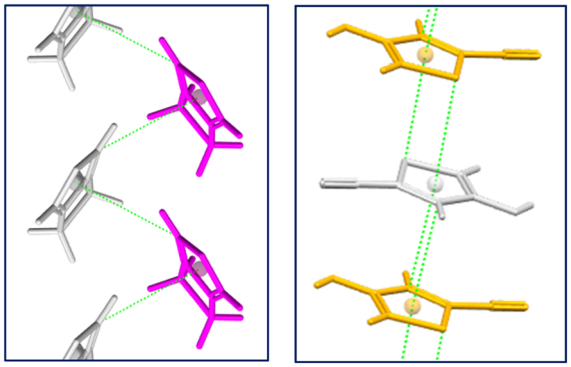

Fig. 6 Symthon II in XULDUD (left, −12.5 kJ mol^{-1}), 2 × Symthon III in GUFJOG (right, −29.0 kJ mol^{-1}, −21.7 kJ mol^{-1}).

molecule. The approach shown on the left below the central molecule occupies two hollows and embodies inversion; the terminal rings overlap each other and their neighbouring rings. Attached hydrogen atoms (centre, far left) occupy the next aromatic rings in the fused rows.

In HAMTIZ01 (Fig. 5), the two hollows above and below the aromatic ring are occupied by a methyl group and an oxygen atom from a nitro group respectively. The approach embodies offset translation, giving b = 4.853 Å. This resembles the offset translation of 1-fluoro-4-iodobenzene in FACQEF01,[5] in which fluorine and iodine atoms occupy hollows, giving b = 4.893 Å. In HAMTIZ01 this approach is supported by an H-bond between the *trans*-amide groups.

All four examples of Symthon II embody short repeats of in the range 5.3–7.7 Å which are generated by repeated screw or glide approaches. Fig. 6(left) shows this motif in XULDUD. Carbon atoms from one edge of the five-membered heterocycle occupy hollows in corresponding rings of neighbouring molecules. The centroid···C distance is 3.575 Å; repetition of this motif gives a translation repeat of 5.31 Å in the vertical direction of Fig. 6.

GUFJOG contains two subtly different face···face approaches embodying Symthon III, as shown in Fig. 6 (right). In the upper approach, sulphur atoms occupy the hollows of neighbouring rings, with centroid···S distances of 3.554 Å. In the lower approach, carbon atoms occupy the hollows of neighbouring rings, with centroid···C distances of 3.358 Å. These two approaches combine to give a translation of 7.52 Å in the vertical direction of Fig. 6. Similar repeats were noted previously for this variant of π–π stacking.[5]

Symthons create long-range order in thirteen crystal structures in this dataset. They form centrosymmetric motifs in a further eight crystal structures, without creating long-range order. Overall, 54 of the 164 hollows accept Symthons. 56 hollows are either unoccupied or only occupied by hydrogen atoms. The other 54 hollows are occupied by heavy (non-hydrogen) atoms but do not possess the symmetry of Symthons. This is often because the heavy atom (usually carbon) belongs to a different aromatic ring. Some examples are presented in Fig. 7–9, together with the corresponding interaction energies calculated using Crytal Explorer.[7,8,17]

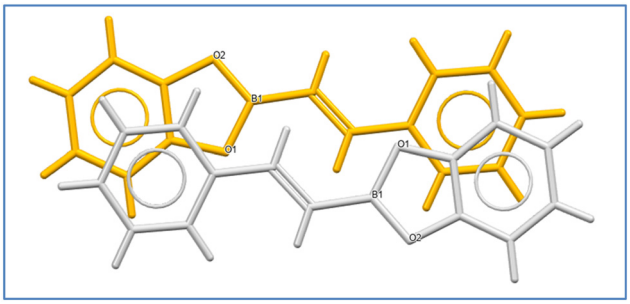

Fig. 7 Aromatic approaches in QAMTAZ (−37.8 kJ mol^{-1}).

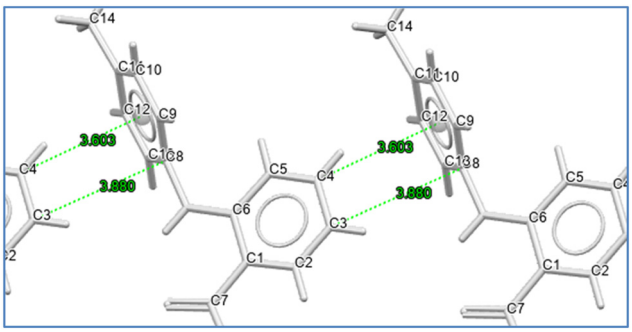

Fig. 8 Aromatic approaches in XAFPAY01 (−16.8 kJ mol^{-1}).

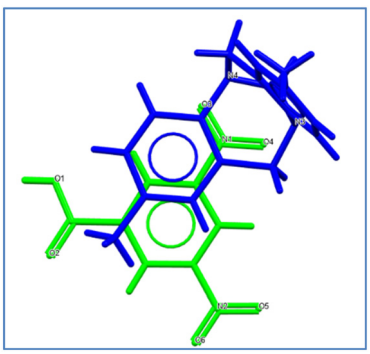

Fig. 9 Aromatic approach in XAFQAZ (−32.5 kJ mol^{-1}).

In QAMTAZ (Fig. 7), phenyl rings occupying both hollows on one face of the fused heterocycle. These are asymmetric face···face *approaches*; the aromatic planes are not quite parallel, with interplanar angles of 7.1°. These asymmetric approaches assist the formation of an intermolecular *interaction* with inversion symmetry, occupying three of the six hollows of each molecule. Overall, fifteen hollows in the dataset are occupied by asymmetric approaches that assist in centrosymmetric interactions.

Fig. 8 shows a close edge···face approach in XAFPAY01, occupying one hollow. This approach has Symthon II geometry (as indicated by the distances shown) but is between different aromatic rings (as indicated by the atom numbering). This asymmetric approach assists in forming an interaction with translational symmetry, giving a repeat of 7.805 Å = *b*, in the horizontal direction in Fig. 8. Overall, 29 hollows in the dataset are occupied by asymmetric approaches that assist in translation, screw or glide interactions leading to long-range order.

Fig. 9 illustrates another option, in which an asymmetric *approach* between two different molecules forms part of an asymmetric *interaction.* The two components of the cocrystal in XAFQAZ are depicted in blue and green; both contain C6 aromatic rings. There is an asymmetric face···face approach between 'green' and 'blue' aromatic rings with an interplanar angle of 1.1°. Two hollows are occupied. Similar asymmetric face···face approaches between co-crystal components are seen in two of the three other cocrystal structures (WICZUF and MIVZEA). Overall, 14 hollows in the dataset are occupied by asymmetric approaches that assist in necessarily asymmetric interactions between different molecules.

Further details of all these aromatic approaches are given in the SI. This analysis identifies three different types of aromatic approaches: Symthons, symmetry assistants, and approaches between different molecules. The final step of this analysis was to ascertain how these aromatic approaches, individually or in combination, contribute to long range order. Long-range order may be achieved by individual approaches such as Symthon I or Symthon II, or by assisted interactions embodying symmetry, or by combinations of other aromatic approaches – further details are in the SI. Long-range order may be created in one direction (chains), two directions (sheets) or three directions (networks).

The results were obtained by careful inspection of the aromatic approaches and their combinations in each crystal structure. This was compared with the number of aromatic rings in each compound, to test the expectation that more aromatic rings in a compound would lead to higher dimensionality of resulting long-range order in its crystal structure(s). The results are presented in Table 3.

For example, HAMTIZ01 (one aromatic ring) and XATMIP (six fused aromatic rings) both adopt Symthon I which occupies all the hollows. This precludes further aromatic interactions, giving one-dimensional chains. Three crystal structures that contain Symthon II (XULDUD as shown in Fig. 6(left), OBEQUJ, and OBEQIX) form two-dimensional sheets of aromatic interactions in a similar way to benzene and monohalobenzenes.[3] Two of the three aromatic rings in XAFPAY01 are almost mutually perpendicular, as shown in Fig. 8. A combination of edge···face and face···face aromatic approaches involving all six aromatic hollows facilitates a 3D

Table 3 Crystal structures grouped by number of aromatic rings and dimensionality of aromatic approaches

Dimensionality	Number of aromatic rings					Total frequency
	0	**1**	**2**	**3**	**4–6**	
Zero	6	7	0	1	1	15
1 – Chains	—	7	4	1	4	16
2 – Sheets	—	3	0	1	1	5
3 – Network	—	—	2	3	1	6

Table 4 Symmetry-forming H-bonds in the dataset

Symmetry	REFCODES	Functional group	Repeat (Å)	Fig.
Translation	HAMTIZ01	*trans*-Amide chain	4.9	10
Screw	KONTIQ06	Water	3.6	10
Glide	BOQQUT	*cis*-Amide chain	7.7	10
Inversion	KONTIQ/KONTIQ06, XAFPAY/XAFPAY01/XAFPAY03	Carboxylic acid rings	—	10
	WICZUF	Pyrimidine ring	—	13
Inversion × 2	PAHYON01	2 *cis*-amide rings	12.2	11
Inversion × 2	OJIGOG01	···Cl-Cu-N-H···	5.3	11

network or aromatic approaches in XAFPAY01 – further details are in the SI.

Overall, aromatic approaches were found to create long-range order in 27 of the 42 crystal structures in the dataset. As expected, they are less dominant than in halobenzenes (100 out 111 crystal structures[5]) but still significant. What approaches are responsible for the creation of long-range order in other directions and in other crystal structures in the dataset? Given the pre-eminence of H-bonds in supramolecular synthons,[1,2] it was expected that H-bonds would provide the answer.

Strong H-bonds in the blind test dataset

The search for H-bonds was greatly simplified by using the default definition of H-bonds in the Mercury software provided by CCDC.[18] It was noted that version 2024.3.0 of this software includes revised van der Waals radii,[19] causing some inconsistencies with earlier analyses. Further details are given in the SI. 25 of the 42 crystal structures contained between one and seven H-bond donors each, with an excess of H-bond acceptors - broadly consistent with the entire CSD.[16] There are 71 H-bond donors in this dataset, all of which are satisfied. Each crystallographically independent H-bond was identified by the atom label of the hydrogen atom.

The next step was to identify which H-bonds created symmetry, in the same way as aromatic Symthons. A starting point was provided by the graph set classification system[20] in which H-bonds are either intramolecular (S), link different molecules (D) or create rings (R) or chains (C). From the perspective of symmetry creation, only types R and C can be symmetry-forming, and only type C gives long-range order. This classification system was modified here for comparison with aromatic Symthons, in which one hollow of an aromatic ring accepts an aromatic approach donated from another, equivalent aromatic ring. These donor/acceptor links for Symthons I, II and III are shown in Fig. 1. Hence an inherent property of aromatic Symthons is that the donor and acceptor belong to the same functional group - an aromatic ring.

Extending this concept to H-bonding requires functional groups that accept and donate H-bonds. This includes amides, carboxylic acids and pyrimidines, in which donors and acceptors are chemically bonded to the same atom. It also includes hydroxyl groups, primary and secondary amines and water molecules. The dataset was searched for examples of these functional groups that formed H-bonds with themselves. These were classified by symmetry, as shown in Table 4.

The ten crystal structures in Table 4 contain a total of 13 H-bond approaches that embody symmetry. In five of these crystal structures, H-bonds embody long-range order, each in a different way, as illustrated in Fig. 10 and 11.

HAMTIZ01 is the only example in which H-bonds give rise directly to translation; accommodated with an offset translation of aromatic rings as shown in grey in Fig. 10(left). The H-bonds are parallel and coplanar with the *trans*-amide group, fixing the translational repeat at ~4.9 Å (see Fig. 5). In KONTIQ06 the H-bonds from water molecules form a 'zig-zag' chain of water molecules which embodies 2-fold screw symmetry, as shown in

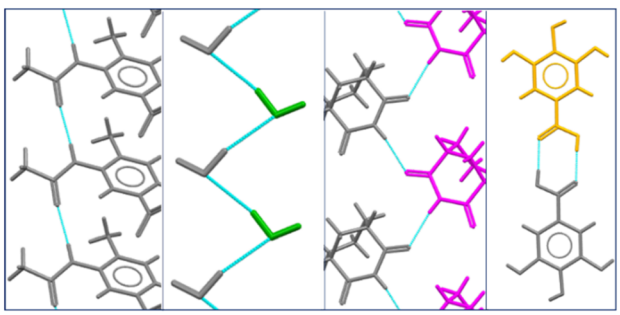

Fig. 10 H-Bonds embodying symmetry in (from left to right) HAMTIZ01 (–46.0 kJ mol^{-1}), water in KONTIQ06 (–10.8 kJ mol^{-1}), BOQQUT (–31.7 kJ mol^{-1}) and KONTIQ (–73.3 kJ mol^{-1}).

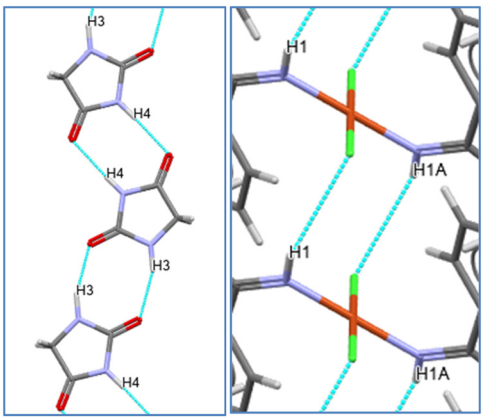

Fig. 11 Inversion pairs creating H-bonded chains in PAHYON01 (left, –61.2 kJ mol^{-1}, –50.6 kJ mol^{-1}) and OJIGOG01 (right, –153.3 kJ mol^{-1}).

green in Fig. 10(centre left). Here the hydrogen atoms lie outside the plane of the oxygen atoms, accommodating a translational repeat of 3.6 Å. This matches the π–π stacking of the gallic acid molecules in this crystal structure, which create Symthon I in the same direction. In BOQQUT the H-bonds form a 'zig-zag' chain and are coplanar with the cis-amide group, embodying glide symmetry with a fixed translational repeat of ~7.7 Å, as shown in magenta in Fig. 10(centre right). Six other H-bonded approaches embody only inversion symmetry – five of which are carboxylic acids featuring anti-parallel H-bonds as exemplified by gallic acid in KONTIQ in yellow in Fig. 10(right). These six approaches do not generate long-range order.

However, combinations of inversion interactions can create long-range order, as shown in Fig. 11. In PAHYON01 (Fig. 11 left) two different antiparallel H-bonded pairs (H3, H4) each embody inversion symmetry, combining to give infinite chains. In OJIGOG01 (Fig. 11 right), the centrosymmetric molecule itself occupies a crystallographic inversion centre. Antiparallel H-bonds (H1, H1A) create an H-bond ring embodying inversion, which combines with the molecular inversion centre to give an infinite chain.

The remaining 58 H-bonds in the dataset do not create symmetry directly, showing three different behaviours. There were five intramolecular H-bonds – making no contribution to crystal symmetry. Fig. 12(top) shows one of many examples of an extended ring; here amines H-bond to carboxylic acid groups.

This is the H-bond equivalent of the ring built from aromatic approaches in Fig. 7 - asymmetric *approaches* between different functional groups assist the creation of a centrosymmetric *interaction*. This motif occurs seven times in this dataset. Fig. 12(bottom) shows an alternative arrangement in which H-bonds between different functional groups (–OH and –CN) combine in GUFJOG to give chains. This is similar to the chain built from asymmetric aromatic approaches in Fig. 8. There are

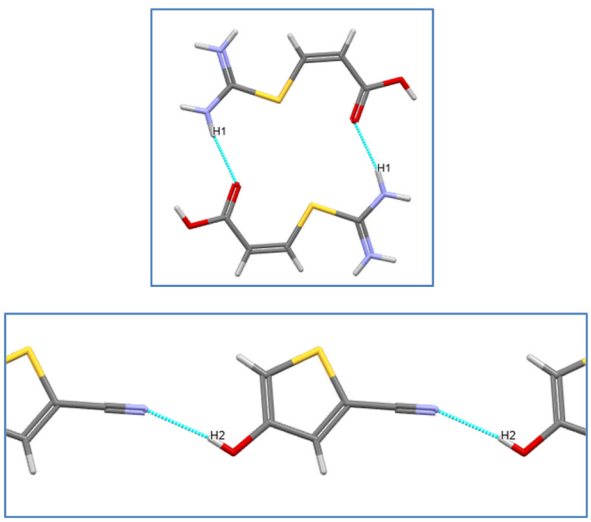

Fig. 12 H-Bonds assisting in symmetry creation in XAFQON (top, –138.9 kJ mol^{-1}) and GUFJOG (bottom, –34.3 kJ mol^{-1}).

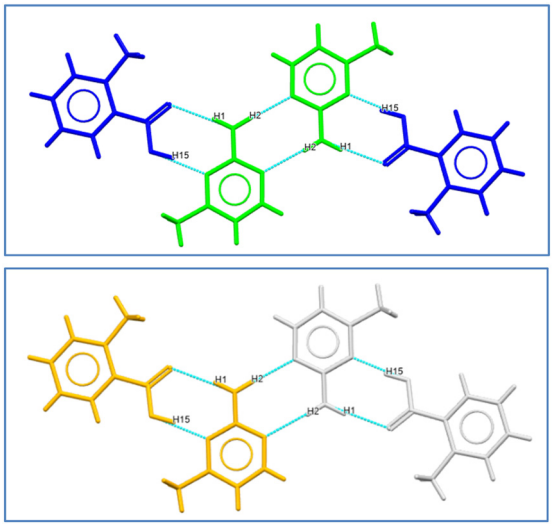

Fig. 13 H-Bonding in WICZUF, coloured by symmetry equivalence (top) and symmetry operation (bottom), showing asymmetric (–71.4 kJ mol^{-1}) and inversion (–52.6 kJ mol^{-1}) H-bonded pairs.

fifteen examples in this dataset of H-bonded chains with translation, screw or glide symmetry assisted by asymmetric H-bonds between different functional groups.

Fig. 13 shows three H-bonds in the co-crystal WICZUF. In the middle of the molecular complex, the centrosymmetric pyrimidine···pyrimidine pair of anti-parallel H-bonds is created by H2. This is similar to the centrosymmetric approach of carboxylic acid groups in KONTIQ (Fig. 10 right) and is the H-bond analogue of Symthon III. There are also two pyrimidine···carboxylic acid H-bonded pairs created by H1 and H15. These are necessarily asymmetric as they link different (green/blue) molecules. The net result of these three H-bonds (H1, H2 and H15) is a centrosymmetric complex of four molecules, but still no long-range order. The twelve multi-component crystal structures in the dataset contain 31 asymmetric H-bonds between different components, 16 of which are within asymmetric units. Further details of all 71 H-bonds are given in the SI.

The contribution of H-bonds to long-range order in this dataset was established rapidly by examining each structure and using the 'expand all' option once all H-bonds had been selected. The results are presented in Table 5, in analogous format to Table 3 for ease of comparison. 17 crystal structures contained no H-bond donors and a further ten contained one, two or three H-bonds that were all either

Table 5 Crystal structures grouped by number of H-bonds and dimensionality of H-bonding

	Number of H-bonds					Total
Dimensionality	**0**	**1**	**2**	**3**	**4–7**	frequency
Zero	17	4	5	1	0	27
1 – Chains	—	3	3	2	0	8
2 – Sheets	—	0	0	0	0	0
3 – Network	—	—	0	0	7	7

intramolecular, asymmetric or embodied or assisted in inversion symmetry. Eight crystal structures, including five crystal structures listed in Table 4, contained only chains of H-bonds. There were no crystal structures in this dataset that included only H-bonded sheets. The seven crystal structures with three-dimensional H-bonded networks included five hydrates.

This analysis revealed four different roles for H-bonds in crystal structures - intramolecular, embodying symmetry, assisting symmetry, or linking different crystal components. The *trans*-amide group is the *only* H-bonding functionality in this dataset that directly embodies a specific crystallographic repeat (4.9 Å). During preparatory work for the current study, two analogous *trans*-sulphonamide chains, embodying slightly longer translations of 5.5 Å, were found in crystal structures of the free acid of the salt in ZEGWAN.

The contrasting roles played by H-bonds are further illustrated here by XATJOT which is a *hydrogen* fumarate salt of a dicarboxylic acid. One proton is transferred to the cation and takes part in an asymmetric charge-assisted H-bond with the anion to create this salt. The other H-bond connects hydrogen fumarate anions in chains to create long-range order. This exemplifies how this dicarboxylic acid creates *crystalline* salts.

Eight crystal structures were studied in more detail to rank strong interactions and compare them with lattice energies. The eight crystal structures (XULDUD, XATMIP, SOXLEX01, QAMTAZ, GUFJOG, HAMTIZ01, UJIRIO and UJIRIO05) from the first four blind tests[9–12] all contained aromatic Symthons. The energies of intermolecular interactions in each crystal structure were ranked systematically and expressed as a % of the total lattice energy. Hitherto, interaction energies have been expressed per mole of *interactions*. For ranking and lattice energy purposes, the energies must be expressed per mole of *molecules*.[17] Here this meant halving the energies of interactions embodying inversion.

Symthons were either part of the strongest interactions, or of the next strongest after H-bonds, in seven of these eight crystal structures. The exception was QAMTAZ, where the strongest interaction was the assisted inversion shown in Fig. 7. The next four interactions in QAMTAZ, in order of decreasing strength, were a side···side interaction, Symthon III, a 'step' interaction and Symthon II. Similar 'step' interactions were noted previously in halobenzenes[3] and are being investigated separately. In the absence of H-bonds, interactions containing Symthons contributed 24–34% of the lattice energy; in the presence of H-bonds the contributions were in the range 13–28%. Further details are given in the SI.

Discussion

Comparison of Fig. 1 and 10 reveals underlying similarities between aromatic Symthons and symmetry - forming H-bond approaches. The chains formed by parallel H-bonds embodying translation (Fig. 10 left) could be designated 'H-bond Symthon I'. The zig-zag chains of H-bonds chains embodying screw and glide symmetry (Fig. 10 centre) could be designated 'H-bond Symthon II'. The antiparallel H-bond pairs embodying inversion (Fig. 10 right) could be designated 'H-bond Symthon III'. Further research is underway to test the broader applicability of this proposal.

Table 6 Sources of long-range order in the dataset

Source of long-range order	Number of crystal structures
Aromatic approaches & H-bonds	8
Aromatic approaches only	19
H-bonds only	7
Other	8
Total	42

Table 6 summarises the sources of long-range order in the 42 crystal structures in the dataset, based on the data presented previously in Tables 3 and 5. Aromatic approaches create long-range order in 27 (=19 + 8) of the 42 crystal structures in the dataset. H-Bonds create long-range order in 15 (=7 + 8) of the 42 crystal structures in this dataset. This finding, that aromatic approaches were more effective than H-bonds at creating long-range order in this dataset, was contrary to expectations.

There are three reasons for this surprising finding. Firstly, although the total numbers of H-bonds (71) and aromatic rings (82) are similar, each aromatic ring can accept two approaches, one on each face – giving over twice as many options for aromatic approaches. Secondly, seventeen crystal structures contain no H-bonds at all, whereas only six contain no aromatic rings. Thirdly, as noted previously, there are ten crystal structures in which H-bonds are present but do not create long-range order.

31 out of 71 H-bonds in the dataset were asymmetric, linking different molecules in all twelve multi-component (salts, cocrystals and hydrates) crystal structures. This compares with only 14/164 asymmetric aromatic approaches linking different molecules, mostly (12/14) in cocrystals. This difference arises because all the co-crystal formers in this dataset contain aromatic rings, whereas the counter-ions (and water!) do not. This is indicative of the more general situation that most solvate-forming solvents and common counterions can take part in H-bonds but do not contain aromatic rings.

H-bonds and aromatic approaches were treated differently in previous discussions of these 42 crystal structures,[9–15] which identified all of the H-bonded chains and networks shown in Table 5. Descriptions of aromatic approaches included 'layers' of molecules in QAMTAZ,[10] PAHYON01[12] and XATMIP,[12] 'offset face···face interactions/stacking' in HAMTIZ01[12] and SOXLEX01[9] and 'herring-bone motifs with f···f arene···arene interactions' in WICZUF.[9] This illustrates how the general agreement on the precise geometric definitions of H-bonds, as enshrined systematically in the Mercury software,[18] makes H-bonds easier to recognise. The systematic use of geometric definitions, symmetry relationships and consistent terminology

in this study gave a very different perspective on aromatic approaches in this dataset. A recent systematic study of aromatic approaches in co-crystals[21] found that they were at least as important as H-bonds in co-crystal formation. Similar systematic studies of aromatic approaches in other structures are necessary, to test the possibility that these conclusions about their importance may be more widely valid.

H-Bonds show a slight advantage over aromatic approaches (seven v. six crystal structures) in creating networks with three-dimensional order. As noted above, five of these seven H-bonded networks are hydrates. The five water molecules all donate two H-bonds and accept two H-bonds – corresponding to 'type 6' hydrates.[22] The tetrahedral disposition of these bonds helps to create the three-dimensional H-bonded networks in these five crystal structures. In contrast, compounds containing several *fused* (Fig. 3 and 4) or coplanar aromatic rings favour the creation of chains, as also seen in phthalocyanines.[23]

Comparison of the H-bonding in Fig. 12 illustrates one further difference between complexation and crystallization. Both approaches yield one H-bond per molecule in the crystal structure. Fig. 12(top) shows a complex of two molecules featuring *two* H-bonds. In Fig. 12(bottom), a complex between two molecules contains *one* H-bond. As the chain in Fig. 12(bottom) grows, the number of H-bonds per molecule increases towards unity. Complexes of the type shown in Fig. 12(top) with inversion symmetry may be more prevalent in solutions. This consideration also applies to centrosymmetric H-bonded pairs formed by carboxylic acids and pyrimidines. The presence of these centrosymmetric complexes in solutions/melts/amorphous phases may aid nucleation of crystal structures that incorporate these centrosymmetric complexes on inversion centres and may hinder nucleation of alternative crystal structures that do not. For example, H-bonded pairs of tetrolic acid molecules were found by IR spectroscopy in chloroform solution, leading to the initial crystallisation of metastable α-tetrolic acid ($P\bar{1}$), which contains H-bonded centrosymmetric pairs like those shown in Fig. 10(right). This polymorph later transforms to the more stable β-tetrolic acid, ($P2_1$), in which the carboxylic acid groups are linked by chains of H-bonds.[24] Similarly, H-bonded pairs of cis-amide groups in carbamazepine were found by NMR spectroscopy in chloroform solution.[25,26] Centrosymmetric cis-amide pairs, like those in PAHYON01 (Fig. 11 left), are found in four polymorphs of carbamazepine. *cis*-Amide chains, like those in BOQQUT (Fig. 10), were present in a predicted polymorph of carbamazepine. This form V could not be prepared from solution but was made by templating from the vapour phase.[27,28]

Graph set analysis[20] is also available within the Mercury software[18] and was mentioned for three of the 42 crystal structures in the 'blind tests'.[11,14] The H-bonded rings in WICZUF (Fig. 13) are both classified as "R2,2(8)". Full graph set analysis identifies one H-bond ring as "R2,2(8) > b > b 1st order", and the other as "D1,1(2)a 1st order + D1,1(2)c 1st order = R2,2(8) > a > c 2nd order". The ready availability of atom numbering and colouring by symmetry equivalence/operation within Mercury, as shown in Fig. 13, offers an alternative way of explaining the differences between the symmetric (H2) and the asymmetric (H1, H15) rings.

The interaction energies presented here (Fig. 3–13) are broadly consistent with the simple energy hierarchy displayed in Fig. 2: H-bond pairs > single H-bonds ~ Symthons I and III > Symthon II. Most molecules in this dataset contain more than one functional group, so interaction energies between whole molecules will sometimes be larger than the energies of close approaches of functional groups that form part of these interactions. For example, the proximity of copper atoms in OJIGOG01 (Fig. 11) contributes to an interaction energy of −153.3 kJ mol^{-1}, twice the energy typically associated with an H-bond pair. Other larger interaction energies are associated with salt formation (XAFQON, Fig. 12, −138.9 kJ mol^{-1}) and extended face···face approaches of fused aromatic rings (XATMIP, Fig. 3, −77.3 kJ mol^{-1}; XIFZOF, Fig. 4, −66.3 kJ mol^{-1}, −113.8 kJ mol^{-1}). The energy of the interaction shown in Fig. 5, (HAMTIZ01, −46.0 kJ mol^{-1}) is consistent with a combination of a single H-bond and an offset Symthon I. The H-bond chain of water molecules in KONTIQ06 (Fig. 10, −10.6 kJ mol^{-1}) is weaker than typical H-bonds.[2] This H-bond geometry may be sub-optimal, as a compromise with Symthon I between gallic acid molecules (−18.6 kJ mol^{-1}) in the same direction, combining to embody b = 3.641 Å.

For a further perspective on the results obtained here, compatible symmetries for the 18 common H-bonded supramolecular synthons identified previously[2] were determined by inspection. *trans*-Amides are the only supramolecular synthon in that set that embody translation directly. Two other supramolecular synthons, both alcohols, only specify long-range order. Pyrazoles, carboxylic acids and cis-amides (Fig. 10 and 11) can form either centrosymmetric H-bond pairs or chains of single H-bonds. Nine supramolecular symthons contain H-bond pairs, five of which are centrosymmetric and four are between different molecules (Fig. 13). This is consistent with the finding here that H-bonds frequently create asymmetric and centrosymmetric complexes, without necessarily creating long-range order.

This study shows that Symthons are not confined to aromatic rings containing six carbon atoms in halobenzenes. They are also formed by five-membered rings, heterocycles and fused aromatic rings. They occur for aromatic rings with a wide variety of substituents. This concept of 'symmetry forming close approaches' can also be extended to certain H-bonds which embody translation, screw/glide and inversion symmetries in H-bond versions of Symthons I, II and III respectively, as shown in Fig. 10. Only one out of the 42 crystal structures in this dataset contains a halobenzene; this investigation revealed aromatic or H-bond Symthons in 28 of these 42 crystal structures.

Conclusions

From the perspective of symmetry, there are three distinct types of close approaches; those that embody symmetry directly,

those lacking local symmetry that assist interactions embodying symmetry, and those that are part of asymmetric interactions. This classification applies to both aromatic approaches and H-bonds. In this dataset, aromatic approaches and H-bonds that embody symmetry are in a significant minority. Many embody inversion symmetry and could occur in amorphous phases, solution or in melts as well as in crystals. In this dataset, aromatic Symthons were more prevalent than their H-bonded analogues. Aromatic approaches were better than H-bonds at forming long-range order, either directly through Symthons or indirectly by assisting in the creation of chains, sheets or networks. H-bonds were more prominent in multi-component crystals, particularly salts and hydrates. This is consistent with the dominant role of H-bonds in creating supramolecular synthons, leaving fewer H-bonds available to create crystals.

Conflicts of interest

There are no conflicts to declare.

Data availability

Supplementary information (SI) available: Molecular structures for 31 compounds in the dataset, finding and classifying aromatic approaches and H-bonds, complete lists of all 164 aromatic hollows and all 71 H-bonds, data sources for Fig. 2, intermolecular interaction energies, lattice energies and ranking. See DOI: **https://doi.org/10.1039/D5CE00618J**.

The data supporting this article have been included as part of the SI.

Acknowledgements

The reviewers are thanked for their insights. The author is grateful for the wisdom and encouragement from Roger Davey.

References

1 G. R. Desiraju, *Angew. Chem., Int. Ed.*, 1995, **34**, 2311–2327.
2 J. D. Dunitz and A. Gavezotti, *Cryst. Growth Des.*, 2012, **12**, 5873–5877.
3 S. N. Black, *Cryst. Growth Des.*, 2021, **21**, 6981–6991.
4 S. N. Black, *CrystEngComm*, 2023, **25**, 3079–3087.
5 S. N. Black and R. J. Davey, *CrystEngComm*, 2024, **26**, 4498–4508.
6 A. I. Kitaigorodskii, *Molecular crystals and molecules*, Academic Press, 1973.
7 A. J. Edwards, C. F. Mackenzie, P. R. Spackman, D. Jayatilaka and M. A. Spackman, *Faraday Discuss.*, 2017, **203**, 93–112.
8 C. F. Mackenzie, P. R. Spackman, P. Jayatilaka and M. A. Spackman, *IUCrJ*, 2017, **4**, 575–587.
9 G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. S. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmaenn, M. A. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. van de Streek, A. K. Wolf and B. Schweizer, *Acta Crystallogr., Sect. B*, 2009, **65**, 107–125.
10 J. P. M. Lommerse, W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, W. T. M. Mooij, S. L. Price, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer and D. E. Williams, *Acta Crystallogr., Sect. B*, 2000, **56**, 697–714.
11 W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Dzyabchenko, P. Erk, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, W. T. M. Mooij, S. L. Price, B. Schweizer, H. Scheraga, M. U. Schmidt, B. P. van Eijck, P. Verwerj and D. E. Williams, *Acta Crystallogr., Sect. B*, 2002, **58**, 647–661.
12 G. M. Day, W. D. S. Motherwell, H. L. Ammon, S. X. M. Boerrigter, R. G. Della Valle, E. Venuti, A. Dzyabchenko, J. D. Dunitz, B. Schweizer, B. P. van Eijck, P. Erk, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, F. J. J. Leusen, C. Liang, C. C. Pantelides, P. G. Karamertzanis, S. L. Price, T. C. Lewis, H. Nowell, A. Torrisi, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt and P. Verwer, *Acta Crystallogr., Sect. B*, 2005, **61**, 511–527.
13 D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova, E. Bartashevich, S. X. M. Boerrigter, D. E. Braun, A. J. Cruz-Cabeza, G. M. Day, R. G. Della Valle, G. R. Desiraju, B. P. van Eijck, J. C. Facelli, M. B. Ferraro, D. Grillo, M. Habgood, D. W. M. Hofmann, F. Hofmann, K. V. Jovan Jose, P. G. Karamertzanis, A. V. Kazantsev, J. Kendrick, L. N. Kuleshova, F. J. J. Leusen, A. V. Maleev, A. J. Misquitta, S. Mohamed, R. J. Needs, M. A. Neumann, D. Nikylov, A. M. Orendt, R. Pal, C. C. Pantelides, C. J. Pickard, L. S. Price, S. L. Price, H. A. Scheraga, J. van de Streek, T. S. Thakur, S. Tiwari, E. Venutij and I. K. Zhitkovu, *Acta Crystallogr., Sect. B*, 2011, **67**, 535–551.
14 A. M. Reilly, R. I. Cooper, C. S. Adjiman, S. Bhattacharya, A. D. Boese, J. G. Brandenburg, P. J. Bygrave, R. Bylsma, J. E. Campbell, R. Car, D. H. Case, R. Chadha, J. C. Cole, K. Cosburn, H. M. Cuppen, F. Curtis, G. M. Day, R. A. DiStasio Jr, A. Dzyabchenko, B. P. van Eijck, D. M. Elking, J. A. van den Ende, J. C. Facelli, M. B. Ferraro, L. Fusti-Molnar, C.-A. Gatsiou, T. S. Gee, R. de Gelder, L. M. Ghiringhelli, H. Goto, S. Grimme, R. Guo, D. W. M. Hofmann, J. Hoja, R. K. Hylton, L. Iuzzolino, W. Jankiewicz, D. T. de Jong, J. Kendrick, N. J. J. de Klerk, H. Sin-Yu, L. N. Kuleshova, X. Li, S. Lohani, F. J. J. Leusen, A. M. Lund, J. Lv, Y. Ma, N. Marom, A. E. Masunov, P. McCabe, D. P. McMahon, H. Meekes, M. P. Metz, A. J. Misquitta, S. Mohamed, B. Monserrat, R. J. Needs, M. A. Neumann, J. Nyman, S. Obata, H. Oberhofer, A. R. Oganov, A. M. Orendt, G. I. Pagola, C. C. Pantelides, C. J. Pickard, R. Podeszwa, L. S. Price, S. L. Price, A. Pulido, M. G. Read, K. Reuter, E. Schneider, C. Schober, G. P. Shields, P. Singh, I. J. Sugden, K. Szalewicz, C. R. Taylor, A. Tkatchenko, M. E. Tuckerman, F. Vacarro, M. Vasileiadis, A. Vazquez-Mayagoitia, L. Vogt, Y. Wang, R. E. Watson, G. A. deWijs, J. Yang, Q. Zhuqq and C. R. Groom, *Acta Crystallogr., Sect. B*, 2016, **72**, 439–459.

15 L. M. Hunnisett, J. Nyman, N. Francia, N. S. Abraham, C. S. Adjiman, S. Aitipamula, T. Alkhidir, M. Almehairbi, A. Anelli, D. M. Anstine, J. E. Anthony, J. E. Arnold, F. Bahrami, M. A. Bellucci, R. M. Bhardwaj, I. Bier, J. A. Bis, A. D. Boese, D. H. Bowskill, J. Bramley, J. G. Brandenburg, D. E. Braun, P. W. V. Butler, J. Cadden, S. Carino, E. J. Chan, C. Chang, B. Cheng, S. M. Clarke, S. J. Coles, R. I. Cooper, R. Couch, R. Cuadrado, T. Darden, G. M. Day, H. Dietrich, Y. Ding, A. DiPasquale, B. Dhokale, B. P. van Eijck, M. R. J. Elsegood, D. Firaha, W. Fu, K. Fukuzawa, J. Glover, H. Goto, C. Greenwell, R. Guo, J. Harter, J. Helfferich, D. W. M. Hofmann, J. Hoja, J. Hone, R. Hong, G. Hutchison, Y. Ikabata, O. Isayev, O. Ishaque, V. Jain, Y. Jin, A. Jing, E. R. Johnson, I. Jones, K. V. J. Jose, E. A. Kabova, A. Keates, P. F. Kelly, D. Khakimov, S. Konstantinopoulos, L. N. Kuleshova, H. Li, X. Lin, A. List, C. Liu, Y. M. Liu, Z. Liu, Z.-P. Liu, J. W. Lubach, N. Marom, A. A. Maryewski, H. Matsui, A. Mattei, R. A. Mayo, J. W. Melkumov, S. Mohamed, Z. Momenzadeh Abardeh, H. S. Muddana, N. Nakayama, K. S. Nayal, M. A. Neumann, R. Nikhar, S. Obata, D. O'Connor, A. R. Oganov, K. Okuwaki, A. Otero-de-la-Roza, C. C. Pantelides, S. Parkin, C. J. Pickard, L. Pilia, T. Pivina, R. Podeszwa, A. J. A. Price, L. S. Price, S. L. Price, M. R. Probert, A. Pulido, G. R. Ramteke, A. U. Rehman, S. M. Reutzel-Edens, J. Rogal, M. J. Ross, A. F. Rumson, G. Sadiq, Z. M. Saeed, A. Salimi, M. Salvalaglio, L. S. de Almada, K. Sasikumar, S. Sekharan, C. Shang, K. Shankland, K. Shinohara, B. Shi, X. Shi, A. G. Skillman, H. Song, N. Strasser, J. van de Streek, I. J. Sugden, G. Sun, K. Szalewicz, B. I. Tan, L. Tan, F. Tarczynski, C. R. Taylor, A. Tkatchenko, R. Tom, M. E. Tuckerman, Y. Utsumi, L. Vogt-Maranto, J. Weatherston, L. J. Wilkinson, R. D. Willacy, L. Wojtas, G. R. Woollam, Z. Yang, E. Yonemochi, X. Yue, Q. Zeng, Y. Zhang, T. Zhou, Y. Zhou, R. Zubatyuk and J. C. Cole, *Acta Crystallogr., Sect. B*, 2024, **80**, 517–547.
16 M. J. Bryant, S. N. Black, R. Docherty, A. G. P. Maloney and S. C. Taylor, *J. Pharm. Sci.*, 2019, **108**, 1655–1662.
17 **https://crystalexplorer.net**, downloaded 23/07/2025.
18 C. F. Macrae, I. Sovago, S. J. Cottrell, P. T. A. Galek, P. McCabe, E. Pidcock, M. Platings, G. P. Shields, J. S. Stevens, M. Towler and P. A. Wood, *J. Appl. Crystallogr.*, 2020, **53**(1), 226–235.
19 S. Alvarez, *Dalton Trans.*, 2013, **42**, 8617–8638.
20 M. C. Etter, *J. Phys. Chem.*, 1991, **95**, 4601–4610.
21 A. J. Cruz-Cabeza, P. R. Spackman and A. V. Hall, *Commun. Chem.*, 2024, **7**, 284.
22 A. L. Gillon, N. Feeder, R. J. Davey and R. Storey, *Cryst. Growth Des.*, 2003, **3**, 663–673.
23 J. M. Robertson, *J. Chem. Soc.*, 1935, 615–621.
24 R. J. Davey, S. L. M. Schroeder and J. H. ter Horst, *Angew. Chem., Int. Ed.*, 2013, **52**, 2166–2179.
25 C. A. Hunter, J. F. McCabe and A. Spitaleri, *CrystEngComm*, 2012, **14**, 7115–7117.
26 R. J. Davey, G. Dent, R. K. Mughal and S. Parveen, *Cryst. Growth Des.*, 2006, **6**, 1788–1796.
27 J.-B. Arlin, L. S. Price, S. L. Price and A. J. Florence, *Chem. Commun.*, 2011, **47**, 7074–7076.
28 A. V. Hall, A. J. Cruz-Cabeza and J. W. Steed, *Cryst. Growth Des.*, 2024, **24**, 7342–7360.