CrystEngComm

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/crystengcomm

Quantifying the symmetry preferences of intermolecular interactions in organic crystal structures[‡]

Robin Taylor,[*] Frank H. Allen[†] and Jason C. Cole

Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK. Email: robin@justmagnolia.co.uk

Abstract: The symmetry of a crystal structure and the types of intermolecular interactions found therein are related. We have quantified the extent to which this is the case for a wide variety of interactions using two different statistical approaches. The results show that symmetry effects can be huge; for example, over 67% of S(thiocarbonyl)…H(O,N) interactions involve molecules related by inversion but less than 10% of C(methyl)…O(ether) interactions. Unsurprisingly, the strongest symmetry preferences are often in favour of inversion, but there are many smaller but significant preferences for other symmetries. For example, carbonyl carbon atoms show their highest propensity to interact with carbon-bound halogen atoms when the interacting molecules are related by a screw axis. While many of the symmetry preferences can be understood, others are more difficult to rationalise.

Introduction

The critical role of close packing in determining the arrangements of molecules in crystal structures is widely recognised. Although other scientists contributed,[1-3] it is Kitaigorodsky who popularised the notion of molecules packing tightly by fitting the bumps of one into the hollows of another, thereby maximising attractive dispersion interactions.[4,5] Further, he demonstrated that certain symmetry

---

[‡] Electronic supplementary information (ESI) available: Table S1, containing symmetry-operator percentages, P(X|T_B…T_P), in comma-separated value format.

[†] Frank Allen died on 10th November 2014. He will be sadly missed. This paper is dedicated to his memory.

operators are particularly good at promoting close packing while others are lamentably ineffective. By their nature, centres of inversion often place bumps next to hollows and thus are favourable. At the other extreme, mirror planes result in bumps near to bumps and hollows facing hollows, a configuration inimical to the efficient use of space.

It is not just shape that matters. In their exposition on the “grammar of crystal packing”, Brock and Dunitz noted that the local orientation of bond dipoles is important.[6] The anti-parallel arrangement of bond dipoles that occurs across an inversion centre is electrostatically favourable, while the parallel dipole interactions across mirror planes are destabilising. In fact, so unfavourable are mirror planes that they only occur if occupied. Of the other common symmetry elements, Brock and Dunitz rated 2-fold rotation axes as neither particularly favourable nor particularly unfavourable for crystal packing; 2-fold screw axes as preferable to glide planes; and glide planes comparable to translations.

Ultimately, crystal packing must be rationalised in terms of interactions between complete molecules.[7] At a qualitative level, however, useful insights can be gained by focussing on the interactions between atoms or groups. One of the earliest examples was Leiserowitz’s study of monocarboxylic acids.[8] He observed that chiral acids have some tendency to form chains along 2_1 axes, but achiral and racemic acids almost invariably form cyclic dimers. Many years later, the excellent study of Eppel and Bernstein established that the acid dimer and many other cyclic hydrogen-bonded motifs have a pronounced tendency to be located on inversion centres in crystal structures of achiral molecules.[9] Of the 44 ring motifs they studied, almost all had >50% probability of forming on an inversion centre (with many >>50%). The probability of finding these ring motifs was much lower in structures of chiral molecules. That said, noncentrosymmetric structures with $Z'>1$ often show pseudosymmetry, allowing the possibility of cyclic hydrogen-bonded motifs at local centres of inversion.[10-12]

While inversion takes pride of place, other symmetry operators have been implicated in the formation of hydrogen-bond motifs. Achiral amide chains are most commonly propagated by glide planes or translation.[13] Steric factors make it difficult for monoalcohols to form closed dimers across inversion centres, but if the molecules are relatively “thin” they can form ...OH...OH... chains along 2_1 axes. Alcohols suffering from more severe steric constraints favour screw or rotation-inversion axes of order 3, 4 or 6, or packing arrangements with Z′>1.[14] Whatever the symmetry, strongly favoured hydrogen-bond motifs usually owe their stability to cooperative effects, almost invariably caused by polarisation. For example, donation of a hydrogen bond by a hydroxyl group increases the OH bond dipole, making the oxygen atom a better acceptor. This favours formation of ...OH...OH chains.[15,16] The anti-parallel arrangement of bond dipoles across inversion centres gives rise to permanent dipole...induced dipole polarisation, resulting in non-additivity of hydrogen bonds in ring motifs.[17,18]

Much less attention has been paid to the symmetry dependence of intermolecular interactions other than hydrogen bonds, but two studies are noteworthy. Lee *et al*. observed that the most common geometry for short nitrile...nitrile, carbonyl...carbonyl and C(sp^2)-F...C(sp^2)-F interactions is anti-parallel across an inversion centre.[19] Since this inevitably places other pairs of dipoles anti-parallel, the result cannot be attributed solely to the nitrile ...nitrile, carbonyl...carbonyl or C-F...C-F interaction. When the two interacting groups are chemically inequivalent, so cannot be related by symmetry (*e.g.* an aldehyde and ketone carbonyl), the tendency for anti-parallel orientation is very much reduced. This implies that the synergistic effect of having all dipoles anti-parallel across an inversion centre is important. Saha *et al*. found that halogen...halogen contact geometries depend on the symmetry element between the atoms: 2_1 axes or glide planes favour an L geometry and 2-fold rotations a V geometry.[20] They suggested that some control over crystal symmetry might therefore be exerted by choosing appropriate halo substitution, since I...I or Br...Br interactions are more likely to have L geometry while Cl...Cl favours V geometry.

In summary, published work indicates that some types of interactions are more likely to be found when the molecules containing the interacting groups are related by a particular symmetry operator; that the effect can be strong; and that it may have practical uses in crystal engineering. However, little has been done to quantify and establish the statistical significance of the symmetry dependence of intermolecular interactions, particularly those that are not hydrogen bonds. We address this deficiency here for a large number of interatomic interactions using statistically rigorous methodology.

Experimental

Data set

The study was based on 86,907 crystal structures taken from the Cambridge Structural Database (CSD), version 5.35.[21] All satisfied the following criteria: no elements other than H, C, N, O, F, P, S, Cl, Br or I; R-factor ≤ 7.5% (structures with no quoted R-factor were omitted); no missing atomic coordinates; no disorder (no disorder comment, no suppressed atoms); only one representative from each "refcode family". (A refcode family comprises independent determinations of the same chemical compound; the structure with the lowest R-factor was selected. *Inter alia*, this constraint will exclude all but one of any set of polymorphic structures.) In order to minimise the number of complicating factors, the data set was confined to structures with $Z'=1$, having no molecules on special positions, and with only one type of chemical residue present (so-called $Z^r=1$ structures[22]). Hence, solvates and ion pairs were excluded. Hydrogen atom positions were normalised by moving hydrogen atoms along observed X-H bond vectors so that XH distances were set to average neutron-diffraction values (CH = 1.089Å, NH = 1.015Å, OH = 0.993Å).[23,24] Molecules were classified as chiral or achiral using an in-house program.[25] Chirality due to restricted rotation was not taken into account, a molecule being assigned as chiral only if it had one or more carbon, phosphorus or sulfur stereocentres and was not a meso isomer.

Atom types and surface areas

Atoms were assigned types depending on their substructural environments. An ordered list of substructures was defined using SMARTS strings[26] and each atom assigned the type corresponding to the first substructure it matched (Table 1). Geometric tests were used to distinguish between planar, nonplanar, linear and nonlinear nitrogen atom types. Further atom types, described later, were used to aid analysis of some specific interactions.

The exposed surface area of each crystallographically independent atom in the data set was computed by placing points randomly on its van der Waals (vdw) surface and counting how many did not fall within the vdw envelope of any other atom in the same molecule.[27] 30,000 points per atom were used, sufficient to produce results to an uncertainty of <1%.

Interactions and their classification

The study was based on interatomic interactions obeying the following conditions: (a) the interaction is intermolecular; (b) it is line-of-sight (LoS), *i.e.* the interacting atoms "see" each other because no third atom intrudes between them; (c) it has $\Delta < 1$Å, where Δ is the interatomic distance minus the sum of the atoms' vdw radii. Except where otherwise stated below, only one interaction per crystallographically independent atom was allowed. For any such atom, B, this was chosen as the intermolecular, LoS interaction with the smallest value of Δ, *i.e.* the shortest interaction after correction for vdw radii. This is termed the *primary interaction* of B. The atom to which B (the *base atom*) forms its primary interaction is its *partner atom*, P. In some parts of the study, the *secondary* and *tertiary* interactions of each base atom were also included, these being the interactions with the second and third smallest values of Δ.[28]

Interactions were classified by the atom types of the base and partner atoms (denoted T_B and T_P, respectively) and by the symmetry operator relating the molecules containing these atoms. The symmetry operators considered were translation (T), inversion (I), glide (G), mirror (M), 2-fold screw (S) and 2-fold rotation (R).[29] Interactions between atoms in molecules related by operators of higher-symmetry were excluded. Interactions between molecules related by 2-fold rotation and a unit cell translation parallel to the axis were treated as 2-fold screws.

Analysis of interactions

The study was performed in two parts. The first part focussed on symmetry-operator percentages. For each combination of base and partner atom types, $T_B \ldots T_P$, we determined the numbers of primary interactions that involved atoms in molecules related by each of the operators T, I, G, M, S and R. If the total number of $T_B \ldots T_P$ primary interactions was N, of which M_X involved molecules related by operator X, the quantity $P_X = 100M_X/N$ is an estimate of the percentage probability $P(X|T_B \ldots T_P)$, *i.e.*

the probability that an interaction will involve atoms in molecules related by operator X, given that the base and partner atoms are of type T_B and T_P. The estimated standard error of P_X follows from the binomial distribution as:

$$\sigma(P_X) = \sqrt{[P_X(100-P_X)/N]} \qquad (1)$$

Using the normal approximation to the binomial distribution, the 95% confidence interval is $P_X-1.96\sigma(P_X)$ to $P_X+1.96\sigma(P_X)$. We also looked at symmetry-operator percentages for the entire sample of primary interactions, irrespective of atom types, and for primary interactions broken down by space group and the presence or absence of chirality.

The second part of the study was based on the R_F statistic.[28] For interactions of the type $T_B...T_P$, R_F is defined as:

$$R_F(T_B\ldots T_P) = \Sigma O(T_B\ldots T_P)_i \,/\, \Sigma E(T_B\ldots T_P)_i \qquad (2)$$

The summations are over all crystal structures in the data set that contain both types of atoms, $O(T_B\ldots T_P)_i$ is the observed number of $T_B\ldots T_P$ primary interactions in the ith structure, and $E(T_B\ldots T_P)_i$ is the expected number. The latter quantity is calculated assuming that the probability of a given base atom forming its primary interaction to an atom of type T_P is determined solely by surface area considerations, as follows. If there are $N(T_B)_i$ base atoms of type T_B in the ith structure (excluding any that do not form a LoS interaction shorter than our limit, *viz.* sum of vdw radii plus 1Å), the expected number of primary interactions to atoms of type T_P is binomially distributed. Each of the base atoms constitutes a “trial”. The probability of success in a trial (primary interaction is to an atom of type T_P) is $S(T_P)_i/S(\text{total})_i$, where $S(\text{total})_i$ is the total exposed surface area of all atoms in the asymmetric unit (*i.e.*, the molecular surface area) and $S(T_P)_i$ is the exposed surface area of atoms of type T_P. The mean of a binomial distribution is (number of trials) x (probability of success), so:

$$E(T_B\ldots T_P)_i = N(T_B)_i S(T_P)_i/S(\text{total})_i \qquad (3)$$

An interaction with R_F=x occurs x times more often than expected by chance.

In this work, we categorise interactions not only by atom types but also by symmetry. We therefore define the quantity:

$$R_F(T_B\ldots T_P|X) = \Sigma O(T_B\ldots T_P|X)_i \,/\, \Sigma E(T_B\ldots T_P|X)_i \qquad (4)$$

$R_F(T_B…T_P|X)$ is the R_F value of the interaction type $T_B…T_P$ *when the molecules containing the interacting atoms are related by symmetry operator X.* Only interactions obeying this symmetry constraint are included when evaluating $\Sigma O(T_B…T_P|X)_i$. The value of $E(T_B…T_P|X)_i$ is calculated from an expression analogous to (3) except that $N(T_B)_i$ is now the number of base atoms of type T_B in the ith structure that form their primary interaction to an atom in a molecule related by operator X. Given two operators X_1 and X_2, we can establish (see below) whether $R_F(T_B…T_P|X_1)$ and $R_F(T_B…T_P|X_2)$ are significantly different. If so, the propensity for T_B base atoms to form primary interactions to T_P is symmetry dependent. Uncertainties in R_F due to the choice of vdw radii were dealt with by performing calculations in triplicate using the radii values of Bondi,[30] Rowland and Taylor,[31] and Alvarez.[32] Unless otherwise stated, R_F values quoted below are averages from the three analyses.

Interactions were visualised and further analysed with the programs ConQuest[33] and Mercury.[34,35]

Confidence intervals and significance of R_F differences

Confidence intervals for R_F were determined by bootstrapping.[28] Let the crystal structures used to calculate $R_F(T_B…T_P|X)$ be C_1, C_2, …C_N. They will be the structures that contain atoms of types T_B and T_P and in which at least one of the T_B atoms forms it primary interaction to an atom in a molecule related by symmetry operator X. 100,000 estimates of R_F were made, each based on N structures chosen randomly from the set (C_1 … C_N) by sampling with replacement. The (x/2)th and (100-x/2)th percentiles of the resulting distribution could then be used as an estimate of the (100-x)% confidence interval of $R_F(T_B…T_P|X)$. In practice, three different estimates of the confidence interval were obtained, since calculations were done in triplicate using different sets of vdw radii. A combined confidence interval could be constructed by taking the smallest of the three lower limits and the largest of the upper limits.

We sought to use the confidence intervals to establish whether the probability of a base atom of type T_B forming its primary interaction to an atom of type T_P is symmetry dependent. For this to be so, there must be at least one pair of symmetry operators (X_1,X_2) for which $R_F(T_B…T_P|X_1)$ is significantly greater than $R_F(T_B…T_P|X_2)$. We were unwilling to accept a difference as significant unless the results from all three sets of vdw radii agreed on it. We therefore required that, with each set of radii, the lower limit of the (100-x)% confidence interval of $R_F(T_B\ldots T_P|X_1)$ exceeded the upper limit of the (100-x)% confidence interval of $R_F(T_B\ldots T_P|X_2)$, where x is the value that gives a test corresponding to the 99% confidence level.

The required value x is not 1, *i.e.* it is incorrect to use the (100-1)% = 99% confidence interval. That would be too conservative, both because non-overlapping 99% confidence intervals corresponds to a probability < 0.01[36] and because we performed our test in triplicate and required all three tests to agree. We estimated an appropriate value for x by trial and error, using R_F values and confidence intervals determined for a range of $T_B\ldots T_P$ types and symmetry operators. For a trial value of x, we found pairs of values, $R_F(T_B\ldots T_P|X_1)$ and $R_F(T_B\ldots T_P|X_2)$, whose (100-x)% confidence intervals did not overlap with any of the choices of vdw radii, and with $R_F(T_B\ldots T_P|X_1) > R_F(T_B\ldots T_P|X_2)$. Let the crystal structures used to calculate these values be $C_{1,1} \ldots C_{1,N1}$ and $C_{2,1} \ldots C_{2,N2}$, respectively. 100,000 estimates of $\Delta R_F = R_F(T_B\ldots T_P|X_1) - R_F(T_B\ldots T_P|X_2)$ were made from each set of vdw radii in turn, each estimate being based on random samples of N1 structures from the set ($C_{1,1} \ldots C_{1,N1}$) and N2 from the set ($C_{2,1} \ldots C_{2,N2}$), chosen by sampling with replacement. The proportion, p, of ΔR_F values that were positive was used as an estimate of the probability that $R_F(T_B\ldots T_P|X_1)$ is truly the higher R_F value. Trials with x = 5, 10, 15 and 20 resulted in, respectively, 94 out of 95, 118 out of 121, 132 out of 138 and 134 out of 156 pairs with p>0.99. On this basis, we selected x=15 (*i.e.* a requirement for non-overlapping 85% confidence intervals) as being appropriate. The data suggest that we can be (100x132)/138 ≈ 95% confident that two R_F values whose confidence intervals obey this condition are significantly different at the 99% confidence level.

Results and Discussion

Overall distribution of symmetry-operator percentages

The first column of Table 2 shows the percentage frequencies with which primary interactions, irrespective of atom types, involve the various symmetry operators. Inclusion of secondary and tertiary interactions makes little difference to the percentages (second column). The interactions were identified using Bondi vdw radii. Calculations using Rowland-Taylor and Alvarez radii gave very similar results, so all other data herein relating to symmetry-operator percentages pertain to primary interactions based on Bondi radii only.

A third of primary interactions are between atoms in molecules related by 2_1 screw axes. Inversion is the next most popular (25.7%), followed by translation (24.4%), glide (14.4%) and 2-fold rotation (1.0%). There are no primary interactions between atoms in molecules related by mirror planes. In fact, no structure in the data set has a space group that contains this symmetry element, which overwhelmingly endorses the dictum of Brock and Dunitz that mirror planes only occur if occupied.[6] The predominance of S over I goes against the widespread belief that the latter is the more favourable for close packing. However, if results are calculated only from primary interactions in structures of achiral molecules (third column of table), I does becomes the dominant operator (34.5%), with T, G and S all close to 20%. The last three columns of the table give results for primary interactions in all structures containing chiral molecules, and then subdivided according to whether the structures are enantiopure (*i.e.* crystallise in Sohnke groups) or racemic. Racemic structures have a higher I percentage and lower S percentage than structures of achiral molecules. Therefore, the high S percentage when all primary interactions are pooled together is due to the enantiopure structures. It is unknown how many of these were determined using crystals grown from enantiopure solutions, where crystallisation in a non-Sohnke group is impossible.

The first row of Table 3 compares the symmetry-operator percentages calculated from (a) all primary interactions and (b) very short primary interactions (*viz.* shorter than sum of vdw radii minus 0.6Å). Subsequent rows break the results down by space group. The space groups included are all those (except P1) for which the percentages are based on ≥100 structures. Almost all of the very short interactions are hydrogen bonds involving OH or NH donors, though there are also some halogen bonds. Restriction to these very short interactions consistently reduces the T percentage, both in the overall figure and in those for individual space groups. If more than one of I, G and S are available in a space group, I seems most likely to increase when switching from all to very short primary interactions, with S least likely to increase. Strong hydrogen bonds therefore show a preference (compared to other primary interactions) for I; for G in noncentrosymmetric, non-Sohnke space groups; and (to a small extent) for S in Sohnke groups. While the conclusion is based on a restricted set of space groups, they are the most common ones for small-molecule structures.

Symmetry-operator percentages naturally vary from one space group to another, depending on the types and numbers of symmetry elements in the unit cell. It is unsurprising, for example, that the S percentages tend to be higher in $P2_12_12_1$ than in $P2_1$; a molecule in the former group is related to three other molecules in its unit cell by screw axes, compared to only one in $P2_1$. Nevertheless, the percentages can show substantial variation from one structure to another in a given space group. Figs. 1 and 2 show two structures crystallising in $P2_12_12_1$. In the first, all of the primary interactions are to atoms in molecules related by S; in the other, 17 of 21 are to atoms in molecules related by T. Fig. 3 shows histograms of the S percentage in $P2_1$ and $P2_12_12_1$ structures. The distributions overlap substantially.

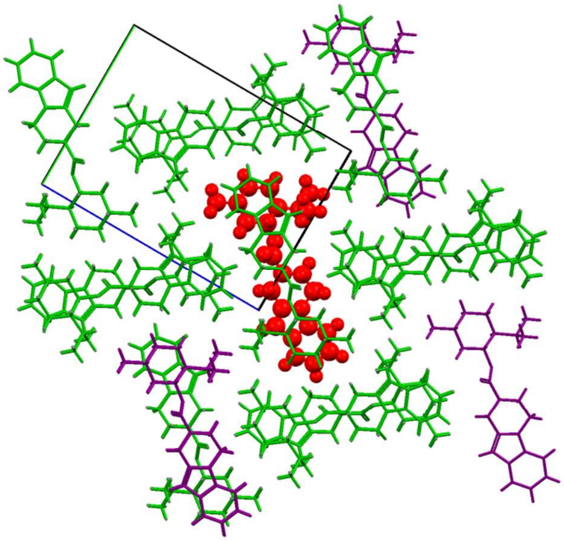


Fig. 1 Packing of CSD entry JUTXEB. Molecules in green are related to the red molecule by 2_1 axes, those in purple by translation.

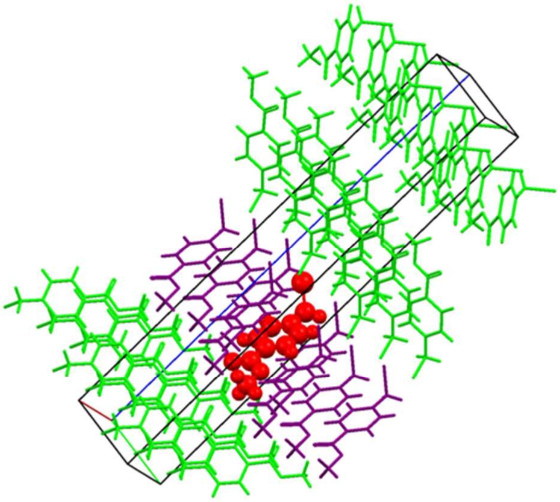


Fig. 2 Packing of CSD entry XEZBOU. Molecules in green are related to the red molecule by 2_1 axes, those in purple by translation.

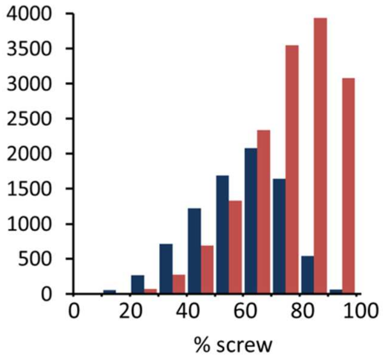


Fig. 3 Distribution of the percentage of primary interactions that occur between atoms in molecules related by screw axes for structures in $P2_1$ (blue bars) and $P2_12_12_1$ (red).

These results tell us something about the geometry of crystal packing. In $P\bar{1}$, for example, most primary interactions tend to be found - and hence the tightest packing tends to occur - between molecules related by I rather than T. But does this mean that the most of the lattice energy in $P\bar{1}$ structures tends to be contributed by interactions between inversion-related molecules? Our freedom to draw this inference depends to some extent on the assumption that shorter interatomic interactions tend to be more stabilising, the so-called length-strength relationship. Novoa, Braga and others have identified several exceptions to this correlation.[37-42] However, the exceptions are often found in interactions that involve ionic species, which are absent from our data set except for a small proportion of zwitterions. Moreover, the length-strength relationship, even though approximate, has been used successfully in many cases. For example, Desiraju correlated the acidity of haloalkane CH groups with the mean lengths of the CH…O interactions that they form.[43] At the very least, it seems reasonable to assume that the overwhelming proportion of very short primary interactions (those referred to in Table 3) are significantly stabilising. Further, the percentage of primary interactions in a structure that occur between molecules related by a given operator is an approximate measure of the surface area of contact between these molecules, which in turn should show some correlation with their interaction energy.

We compared our geometry-based results with those derived from energy calculations by Filippini and Gavezzotti.[44] They studied a set of about a thousand hydrocarbon, oxahydrocarbon and azahydrocarbon crystal structures. In each structure, the molecular coordination sphere around a central molecule was generated and the matrix of intermolecular interaction energies calculated. The importance of a given symmetry operator could be assessed by calculating the proportion of the total interaction energy contributed by molecule pairs related by that operator. In structures belonging to $P\bar{1}$, molecules related by I contributed, on average, more stabilization energy than those related by T. In $P2_1$ and $P2_12_12_1$, S dominated on average over T. In $P2_1/c$, the order was I>S>G>T, and in Pbca it was G>I>S>T. Their results are summarised by the red bars in Fig. 4, which show the percentage frequency with which each operator is the top-ranking contributor to interaction energy in each of five common space groups. The blue bars indicate the percentage frequency with which each operator contributes the largest number of primary interactions in each of those space groups. A qualitative concordance between the two sets of results is evident. Our percentages remain almost unchanged if they are calculated only from the types of structures used by Filippini and Gavezzotti, *viz.* hydrocarbons, oxahydrocarbons and azahydrocarbons.

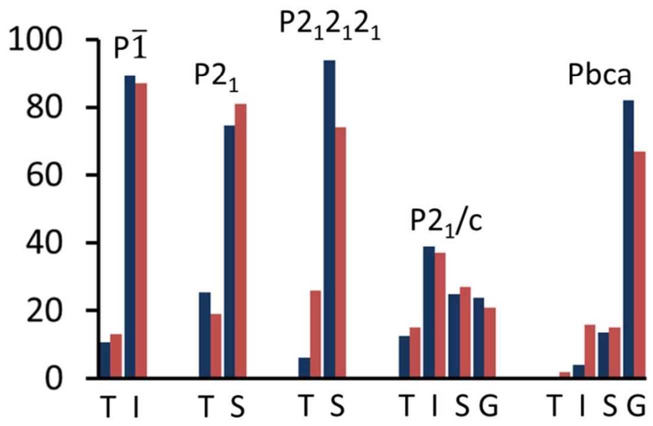


Fig. 4 Red bars: percentage frequency with which given operator (indicated on x axis, T=translation, I=inversion, S=2-fold screw, G=glide) makes the top-ranking contribution to lattice energy.[44] Blue bars: percentage frequency with which given operator contributes the largest number of primary interactions.

Symmetry-operator percentages broken down by atom types

We computed symmetry-operator percentages for the different types of primary interactions, as defined by base and partner atom types. Atom-type combinations with <500 primary interactions in the data set were discarded, together with the H[polar]…H[polar] interaction because of the high experimental uncertainties in polar hydrogen atom positions. Results for the remainder are deposited as supplementary information as Table S1. A χ^2 test established that the symmetry-operator percentages vary more with primary interaction type than would be expected by chance ($p<0.001$). This is not an important result in itself because, with such a large data set, even small effects can be statistically significant. But the effect is far from small. For example, the interaction with the smallest I percentage is C[methyl]…O[ether] at 9.8%; that with the largest is S[=C]…H[polar] at 67.4%, higher by a factor of over six (here and throughout, interactions are written with base atom first, partner atom second[45]). The other operators show similar variability. The interactions in Table 4 (sorted by I percentage) have been selected to illustrate this.

Although our analysis is couched in terms of atom…atom interactions, these are often surrogates for more complex motifs involving several atoms. For example, the S[=C]…H[polar] interactions almost invariably correspond to thiourea or thioamide cyclic dimer motifs (Fig. 5). Visual inspection of example atom…atom interactions was undertaken throughout the study to ensure that we had an adequate understanding of the motifs that they represented.

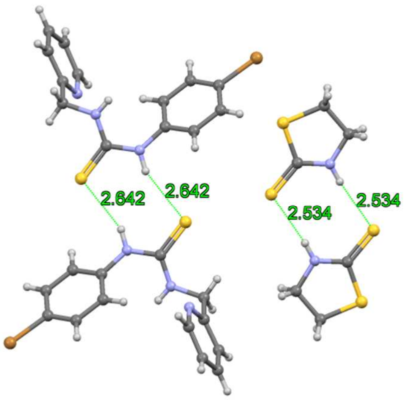


Fig. 5 Examples of the thiourea and thioamide dimer motifs that dominate the S[=C]…H[polar] interaction type (CSD entries ACAFUG, CASHOT03).

Several of the interactions with high I percentages are "like-atom interactions", in which the base and partner atoms have the same type, *e.g.* C[carbonyl]…C[carbonyl], N[planar]…N[planar]. The implications of this are discussed later. In view of the analysis of alcohol packing preferences by Brock and Duncan,[14] it is interesting that the interaction with the highest S percentage is O[hydroxyl]…H[polar] at 52.0%. The SMARTS definition of H[polar] includes both NH and OH groups. If they are separated into different atom types, the resulting S percentages are 53.8% and 44.1% for O[hydroxyl]…H[O] and O[hydroxyl]…H[N], respectively, confirming that it is the former interaction that is more likely to form along screw axes. The same tendency is seen in other hydrogen-bonding interactions involving H[O] and H[N], the former consistently yielding the higher S percentage (Table 5). This may partly be ascribed to the fact that molecules containing OH in our data set are more likely to be chiral than molecules containing NH. However, the same trend is found, though to an appreciably smaller extent, if the analysis is confined to achiral molecules.

There is an interesting feature of achiral molecules that crystallise in Sohnke groups. For a given pair of base and partner atom types, we define $N_{ACHIRAL}$ as the total number of primary interactions of that type in structures of achiral molecules, and $N_{ACHIRAL,SOHNKE}$ as the number in the subset that crystallise in Sohnke groups. When interactions for which $N_{ACHIRAL}<500$ are discarded and the remainder sorted on the ratio $N_{ACHIRAL,SOHNKE}/N_{ACHIRAL}$, the second highest ranking (*i.e.* second largest ratio) is O[hydroxyl]…H[polar]. This is consistent with Dey and Pidcock's conclusion that structures able to form OH...OH interactions are more likely to crystallise in Sohnke groups than those without hydroxyl groups, even when molecules are achiral.[13] (The largest ratio is for O[=X]…O[=X]; the reason for its association with crystallisation in Sohnke groups is unclear.)

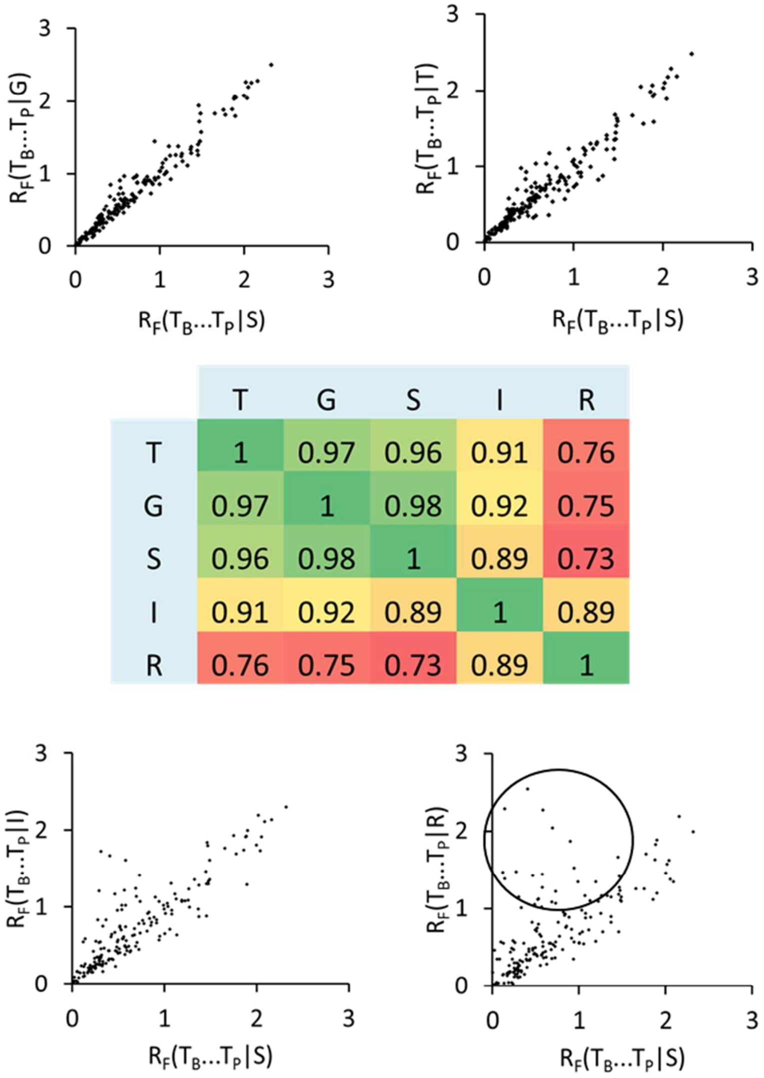
Comparison of R_F values with symmetry-operator percentages

We now turn our attention to the R_F statistic, considering first the example interaction O[ether]…H[polar]. The value of R_F(O[ether]…H[polar]|I) is 8.2; *i.e.*, when the interacting molecules are related by inversion, O[ether] forms about eight times more primary interactions to H[polar] than would be expected from surface area considerations. Other R_F(O[ether]…H[polar]|X) values are lower, *viz*. 4.3, 6.6, 5.3 and 4.8 for X = T, G, S and R, respectively.[46] The differences between these values are statistically significant, so the probability of the interaction forming is related to the symmetry element between the interacting molecules, peaking when X=I.

While R_F decreases in the order I>G>S>R>T, the corresponding symmetry-operator percentages (Table 4, third row from bottom) behave differently, *viz*. S>>T>I>G>R. We also find that, for a fixed operator X and variable interaction types $T_B…T_P$, $R_F(T_B…T_P|X)$ correlates poorly with $P(X|T_B…T_P)$. Both parameters measure the relationship between interaction type and symmetry, but in different ways. $R_F(T_B…T_P|X)$ measures the propensity of base atoms of type T_B to form primary interaction to atoms of type T_P when the interacting molecules are related by operator X. It takes no account of the intrinsic likelihood of X occurring, so an interaction might have a high R_F value for an unusual symmetry operator, such as rotation. $P(X|T_B…T_P)$ is the percentage of $T_B…T_P$ interactions occurring between molecules related by operator X and is therefore sensitive to the absolute frequencies with which the operators occur. The percentages are also sensitive to chemical bias in a way that R_F values are not. This is illustrated by our example interaction. Ether oxygen atoms occur in about 16% of structures in our data set but in over 30% of the subset of those structures that crystallise in Sohnke groups (many molecules containing ether groups are chiral). This at least partly accounts for the high S and T percentages observed for this interaction.

Overview of R_F trends

We calculated the Spearman correlation coefficient of $R_F(T_B…T_P|X_1)$ and $R_F(T_B…T_P|X_2)$ for all pairs of operators X_1,X_2. The calculations were based on the subset of primary interaction types (*i.e.* $T_B…T_P$ combinations) that had the potential to occur (*i.e.* both atom types and the symmetry operator were present) in ≥100 structures for each of the operators. Results are summarised in Fig. 6. The variation of R_F with interaction type is broadly similar for the three operators involving translation, but distinctly different for the non-translational operators I and, especially, R. Most notably, some interactions have a high propensity to occur when molecules are related by R but not when they are related by T, G or S (circled points in the bottom-right plot of Fig. 6).



	T	G	S	I	R
T	1	0.97	0.96	0.91	0.76
G	0.97	1	0.98	0.92	0.75
S	0.96	0.98	1	0.89	0.73
I	0.91	0.92	0.89	1	0.89
R	0.76	0.75	0.73	0.89	1

Fig. 6 Centre: Spearman correlation coefficients of $R_F(T_B…T_P|X_1)$ and $R_F(T_B…T_P|X_2)$ for all pairs of symmetry operators X_1,X_2 (cells colour-coded on correlation-coefficient magnitude). Outside: scatterplots of $R_F(T_B…T_P|S)$ against $R_F(T_B…T_P|X)$ where X is (clockwise from top left) G, T, R and I. In each scatterplot, interactions for which either R_F value exceeds 3 were omitted so that the region containing most observations can be seen more clearly. Circle shows area containing interactions with an appreciably higher propensity to occur with molecules related by R rather than S.

Interactions whose R_F values are significantly dependent on symmetry

We calculated $R_F(T_B…T_P|X)$, where X = T, I, G, S and R, for all combinations of base and partner atom types. Results based on <100 structures were discarded. $T_B…T_P$ combinations satisfying the following criteria were identified: (a) there are at least two symmetry operators (X_1,X_2) for which $R_F(T_B\ldots T_P|X_1)$ and $R_F(T_B\ldots T_P|X_2)$ are significantly different at the 99% confidence level; (b) the ratio $R_F(T_B\ldots T_P|X_1)/R_F(T_B\ldots T_P|X_2)$ exceeds 1.5 for at least one pair of operators X_1,X_2 (which need not be the same operators that fulfil criterion a); (c) at least one of the $R_F(T_B\ldots T_P|X)$ values exceeds 1.3, with the lower limit of its 85% confidence interval exceeding 1. The first criterion selects interactions whose R_F values are symmetry dependent at a high level of statistical significance. The second ensures that the symmetry dependence is not just statistically significant but also reasonably large. The third screens out interactions that are not common in crystal structures, even though they may show symmetry dependence.

Some 47 interactions satisfied these strict criteria, of which one, H[polar]…H[polar], was rejected because of the large experimental uncertainties in polar hydrogen-atom positions. The remainder are listed in Table 6. Many other interactions would be deemed symmetry dependent if less exacting standards were used; a few of these are included in the table as italicised entries. Interactions in the table are grouped by their chemical nature and, within each group, arranged in descending order of the ratio of the highest to lowest R_F value, *i.e.* roughly on the extent to which they are symmetry dependent. The last column lists all statistically significant differences. For example, the first row of the table indicates that, with 99% confidence, R_F(N[acc]…H[polar]|R) > R_F(N[acc]…H[polar]|X) for X = G, I, S and T, and R_F(N[acc]…H[polar]|X) > R_F(N[acc]…H[polar]|T) for X = G, I and S. N[acc] forms its primary interaction to H[polar] almost thirteen times more often than would be expected by chance when the interacting molecules are related by R, but only about five times more often when they are related by T.

A detailed analysis of the interactions in Table 6 is far beyond the scope of a single paper, but we discuss below some points of interest identified by visual inspection of example interactions and some limited geometrical analysis. Each subsection heading indicates the group of interactions in Table 6 under discussion.

Cyclic hydrogen bond motifs. Several hydrogen-bonding interactions have an elevated propensity to occur when they involve molecules related by I. They are invariably components of cyclic hydrogen-bond motifs, many of which are well known (*e.g.* Fig. 5), and their predilection for inversion centres is undoubtedly due to strengthening of the hydrogen bonds by cooperativity.[9,17,18] N[acc]…H[polar] and H[polar]…N[acc] have high R_F values when the interacting molecules are related by R.[45] Examples of this interaction are shown in Fig. 7. Again, cooperative effects are probably important.

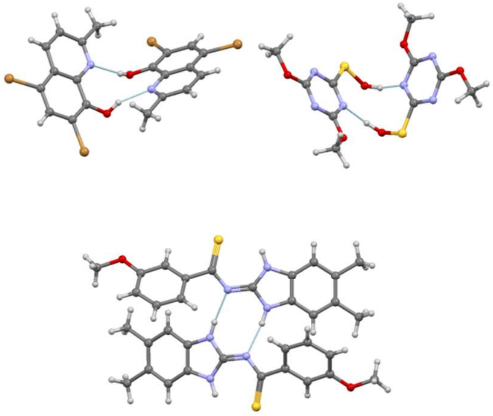


Fig. 7 Examples of cyclic motifs involving N[acc]…H[polar] hydrogen bonds with the interacting molecules related by 2-fold rotation (CSD entries UNUDOX, PEVSIS, MUZBIT).

Stacking interactions. The second group of interactions also show a particular propensity to occur between molecules related by I or R. Visual inspection of examples showed that interactions of the types N[planar]…C[carbonyl], C[unsat]…C[carbonyl] and N[planar]…C[unsat] usually occur as part of an extended stacked arrangement of planar or approximately planar systems (*e.g.* Fig. 8). Through-space electrostatic interactions between polarised atoms in such systems can be significantly stabilising,[47-51] and the stacking of planar systems is an efficient way of filling space.

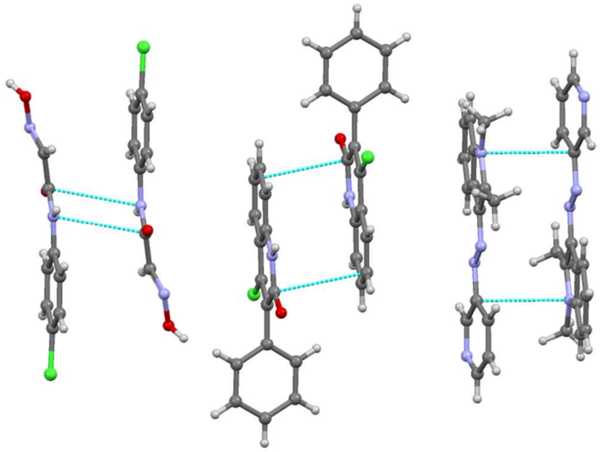


Fig. 8 Examples of centrosymmetric stacked arrangements that include N[planar]…C[carbonyl], C[unsat]…C[carbonyl] and N[planar]…C[unsat] primary interactions (CSD entries DUDJUI, ABABEL, IXUCIO).

Of the other interactions in this group, S[sulfoxide]…O[=X] almost invariably corresponds to a pair of stacked, antiparallel S=O bonds, often with some lateral displacement (Fig. 9, left). C[carbonyl]…O[carbonyl] corresponds to carbonyl…carbonyl dimers, which show more geometric variability but often adopt a similar stacked, antiparallel arrangement (Fig. 9, centre). The stabilisation energy in this geometry can be competitive with hydrogen bonding.[52] Finally, O[=X]…C[carbonyl] shows still more geometric variability and the interaction, in various geometries, sometimes appears as an apparent by-product of other proximal stacking interactions (Fig. 9, right).

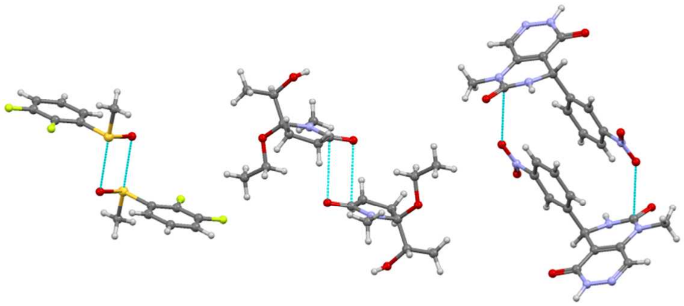


Fig. 9 Antiparallel stacking interactions between S=O bond dipoles (left, CSD entry KONJIG) and C=O bond dipoles (centre, ABEJIB). Nitro oxygen interaction with carbonyl carbon (right, EPATAR).

Like-atom interactions (stacking). These are similar to the interactions discussed directly above. They show a particular propensity for I and R, and are typically found either in extended stacked systems or as isolated stacked, antiparallel bond dipoles, usually with a substantial degree of lateral displacement.

Like-atom interactions (other). The interactions discussed so far in this section may well be sufficiently stabilising under I or R that they are significant in driving molecules to crystallise with one of these symmetries. That the former is much more common may then be explained because it is more effective at promoting close packing. In contrast, the second group of like-atom interactions in Table 6, which also have a propensity for I and R, are only weakly stabilising under these symmetries.

This statement perhaps requires some justification. It is clearly true for one of the interactions, C[methyl…C[methyl]: calculations on the methane dimer indicate an interaction energy of only about -0.4 kcal/mol at best.[53,54] The energetic preferences of the other interactions are more complicated. Like-atom interactions involving polarisable atoms such as the heavier halogens or divalent sulfur can be significantly stabilising due to the atoms' asymmetric electrostatic potentials. Given the correct interaction geometry, an electron-rich region of one atom can face an electron-deficient region ("sigma hole") of the other.[55] For halogens, this requires a "type II" geometry,[56] with one θ (*i.e.* C-Hal…Hal) angle about 90° and the other about 180°. However, most of the halogen…halogen interactions across inversion centres and 2-fold axes do not have this geometry. Indeed, they cannot if they involve inversion-related atoms, because the crystallographic symmetry requires both θ angles to be equal ("type I" geometry).[57] Capdevila-Cortada *et al.* have shown that the type I geometry with $\theta \approx 150°$ (commonly observed at inversion centres) lies at a saddle point on the potential energy surface and is only weakly stabilising.[58] The picture is similar for interactions between divalent sulfur atoms. Again, geometries commonly observed across inversion centres (Fig. 10) and 2-fold axes are not energetically optimum.[55,59] (Some S…S interactions occur as part of extended stacking arrangements and properly belong to the previous group.)

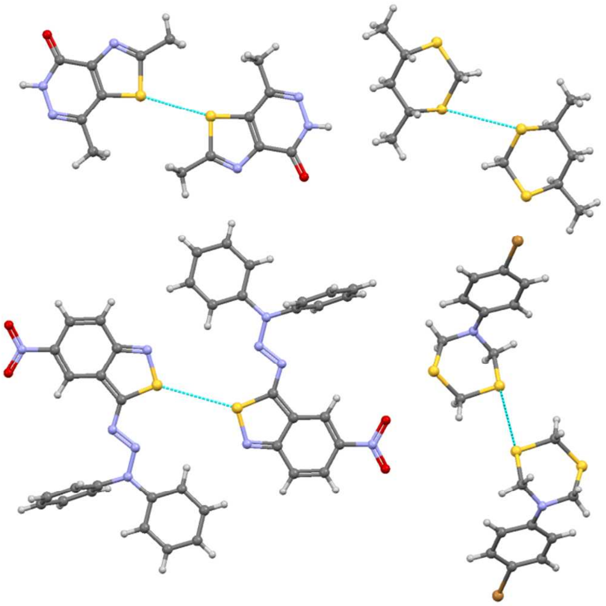


Fig. 10 Examples of S…S primary interactions across inversion centres (CSD entries AZAMIY, DMTHAN, BEHGAX, AYAMUJ).

It seems likely that the proclivity of this group of like-atom interactions for I or R is a geometrical by-product of dispersion-driven close packing. Specifically, when close packing is achieved with one of these non-translational operators, it inevitably results in at least one pair of symmetry equivalent atoms in reasonably close proximity to each other. Therefore, like-atom primary interactions are to be expected. Weakly stabilising like-atom interactions may often be the best available. It is striking that even when halogen…halogen interactions *can* adopt optimum or near-optimum geometries – *i.e.* with the other symmetry operators – their R_F values are low, implying that they do not have a dominant role, *per se*, in determining crystal packing arrangements.

Other hydrogen bonds. Some hydrogen bonds prefer S, G or T symmetry. Unsurprisingly, they include those accepted by hydroxyl oxygen, because steric factors impede formation of centrosymmetric cyclic hydrogen-bond motifs.[14] Hydrogen bonds involving anisole oxygen show similar behaviour. Oxygen atoms in conjugated environments (*e.g.* in esters) are known to be poor acceptors,[60,61] but nevertheless form more primary interactions to H[polar] than would be expected by chance when the interacting molecules are related by translation. Visual inspection suggested that some of these are genuine 2-centre hydrogen bonds, but most are the minor components of 3-centre (bifurcated) interactions (Fig. 11).

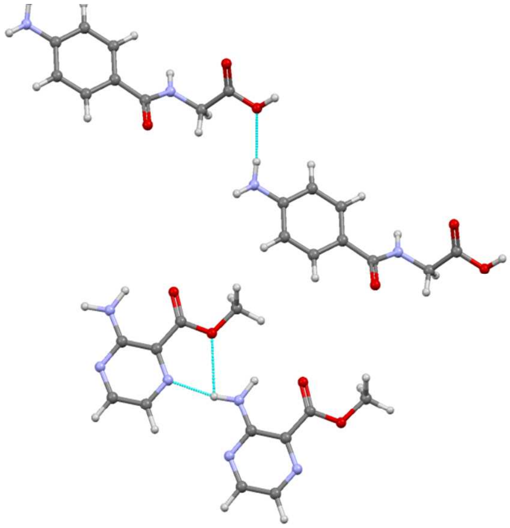


Fig. 11 Examples of 2-centre and 3-centre hydrogen bonds to conjugated oxygen acceptors involving molecules related by cell translation (CSD entries LAYCET01, ADUWUS01).

Halogen bonds and chalcogen bonds. All four types of Br…O interactions in Table 6 show a higher propensity to occur with G, S or (particularly) T than with I. The interactions with the former three symmetries tend to have geometries suggestive of halogen bonding. We performed an additional investigation into halogen bonds by determining the values R_F(Br/I…Acc|X), for X = T, I, G and S, where Br/I is any carbon-bound bromine or iodine, and Acc is any oxygen or acceptor nitrogen. The results were R_F = 1.8, 1.2, 1.8 and 1.5, respectively, with significant differences as follows: T>S,I; G>S,I; S>I. The chalcogen bond S[δ+]…O[carbonyl] also has a greater propensity to form under symmetries other than I, the R_F value being highest for G.

Orthogonal polar interactions of carbonyl carbon. The R_F values for these interactions tend to be highest for S or T and lowest for I. The interaction direction is usually within about 25° of the normal to the plane of the carbonyl carbon atom. Examples are shown in Fig. 12. Paulini *et al*. postulated that orthogonal polar interactions such as C-F…C=O and C-Cl…C=O are important in stabilizing crystal lattices and biological complexes.[62] Previous work by one of us cast some doubt on this, since the R_F values of these interactions could not be shown to be significantly greater than 1 when all symmetry operators were pooled together.[28] However, we now find that R_F(C[carbonyl]…F[C≠CF_3]|X) and R_F(C[carbonyl]…Cl[C]|X) significantly exceed 1 for X = T, G and S. If all carbon-bound halogen atoms are assigned the same atom type, Hal, the resulting values for R_F(C[carbonyl]…Hal|X) peak at 1.4 for X = S.

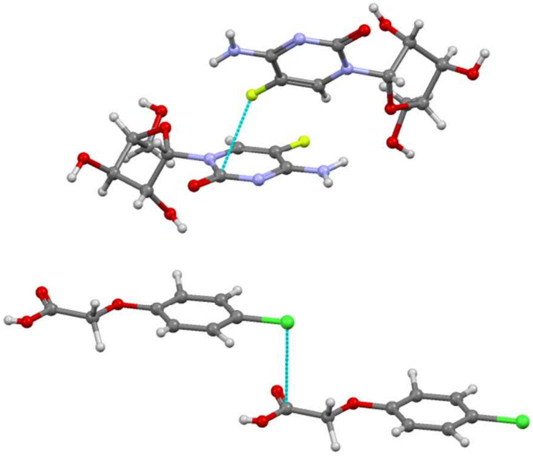


Fig. 12 Approximately orthogonal halogen contacts to carbonyl carbon atoms (CSD entries DOLSEC and CLPOAC).

Hydrophobic interactions associated with screw axes. The R_F values of N[planar]…H[C], C[carbonyl]…H[C], C[unsat]…H[C] and H[C]…C[C_6H_5] are significantly higher for S than for the other operators. The interactions are hydrophobic, formed by planar systems (often aromatic) to aliphatic or aromatic CH groups.

Interatomic distances

We wished to establish whether an interaction that shows an enhanced propensity to occur when the interacting molecules are related by a given symmetry operator also tends to be shorter under that symmetry. This might be expected if the interaction is stronger under the relevant symmetry (*e.g.* through a cooperative effect), leading both to shorter interatomic distance and increased probability of occurrence. Using primary interactions determined with the Bondi vdw radii, we calculated the mean interatomic distance, d(T_B…T_P|X), and its standard error for all possible T_B…T_P combinations when

X = T, I, G, S and R. We used the results to perform unpaired t-tests to find all pairs of operators (X_1,X_2) for which d(T_B…T_P|X_1) is significantly shorter than d(T_B…T_P|X_2) at the 99% confidence level. The results were compared with significant differences found in the R_F values (see above).

For a given T_B…T_P combination and pair of operators X_1,X_2, there are five possible outcomes. The first is that neither the d nor the R_F values are significantly different. The remaining outcomes are: (a) R_F values significantly different, d values not; (b) d values significantly different, R_F values not; (c) d(T_B…T_P|X_1) significantly shorter than d(T_B…T_P|X_2) and R_F(T_B…T_P|X_1) significantly greater than R_F(T_B…T_P|X_2); (d) d(T_B…T_P|X_1) significantly shorter than d(T_B…T_P|X_2) and R_F(T_B…T_P|X_1) significantly smaller than R_F(T_B…T_P|X_2). The numbers of outcomes falling in categories (a)-(d) were 734, 139, 230 and 52, respectively. Two main conclusions can be drawn. First, most significant R_F differences are not associated with a significant difference in mean distances. Second, when there is an association, the operator giving the higher R_F value usually gives the shorter distance: specifically, in (100x230)/(230+52) = 81.6% of cases. This is consistent with our initial conjecture, but the evidence is not conclusive.

Robustness of results

We investigated whether the symmetry dependence of R_F is similar in different subsets of the data, first comparing structures of chiral molecules (including racemates) with those of achiral molecules. Taking each subset in turn, we found all pairs of values R_F(T_B…T_P|X_1), R_F(T_B…T_P|X_2) – *i.e.* same atom types, different symmetry operators – which were significantly different at the 99% confidence level. We discarded all interactions except those for which at least one significant difference in both subsets were found and determined whether the differences were consistent. For example, the significant differences found for N[planar]…H[polar] with the achiral subset were R>S,G,I,T; S>I,T; G>I,T. With the chiral subset, they were R>G; I>G; S>G. Symmetry operator pairs which gave a significant difference for one subset but not the other (*e.g.* R>S) were ignored, since this might be due to insufficient data in the subset not giving a significant result. For the remainder, the number of significant differences that were consistent (*i.e.* in the same direction) in both subsets was counted, and the number that were inconsistent. In the example, there is one consistent difference, *viz.* R>G, and one inconsistent (achiral: G>I; chiral: I>G). In total, there were 400 consistent and only 14 inconsistent differences, showing that the symmetry trends are similar in the structures of chiral and achiral molecules.

We next compared structures in Sohnke groups and non-Sohnke groups, using the same methodology. Again, there was strong agreement, with 104 consistent differences and no inconsistent differences. Two thirds of the consistent differences involve the R operator. We have already seen that $R_F(T_B…T_P|R)$ values are often very different from $R_F(T_B…T_P|S)$ and $R_F(T_B…T_P|T)$, so it is perhaps unsurprising that the influence of this operator is similar in both structure subsets. However, the remaining consistent differences are of the types S>T or T>S. While statistically significant, many of these differences are small, as judged by the ratio of the larger to the smaller R_F value. It is therefore surprising they are found so consistently in both structure subsets. We had imagined that the absence of inversion centres and glide planes in Sohnke groups might affect the dependence of R_F values on the remaining symmetry operators, but have found no evidence to support this.

Whenever multiple hypothesis tests are performed, some ostensibly significant results are likely to arise by chance.[63] This is an issue here as we make many pairwise R_F comparisons. We could reduce the risk of type I errors (incorrectly concluding that two R_F values are different when in fact they are not) by applying a Bonferroni correction or something similar; this effectively demands that a very high level of significance must be reached before two values can be deemed different.[64] However, this would increase the risk of incorrectly concluding that two R_F values are not different when in fact they are (type II errors), which is undesirable in an exploratory project like this. Therefore, we chose not to apply the correction. However, given that our method for assessing statistical significance is approximate, we sought reassurance by performing a randomised control experiment. All primary interactions in the data set were found, as before, but instead of labelling each interaction by its true symmetry operator, we labelled it with a false one. The false operator was chosen at random, the probability of selecting operator X being determined by the percentage frequency with which X occurs in the data set (first row of Table 3). Only 13 "significant differences" were found amongst the 4057 R_F pairs tested, somewhat less than the 1% we might have expected given that the nominal confidence level of our hypothesis tests was 99%. We conclude that the vast majority of the (~1000) significant differences found in the true analysis are genuine.

Conclusions

The primary interaction of an atom (the "base atom") is defined as its shortest line-of-sight intermolecular interaction after correction for vdw radii. The atom to which the primary interaction is formed is the "partner atom". Each primary interaction is classified by: (a) its type, $T_B \ldots T_P$, where T_B and T_P are the atom types of the base and partner atoms; (b) the symmetry operator, X, relating the molecules containing the atoms. We have investigated the relationship between $T_B \ldots T_P$ and X using two measures, $P(X|T_B \ldots T_P)$ and $R_F(T_B \ldots T_P|X)$. The first is the percentage probability that the molecules containing the atoms of a primary interaction will be related by operator X, given that the interaction is of type $T_B \ldots T_P$. The second is a measure of the propensity of a base atom of type T_B to form its primary interaction to a partner of type T_P, given that the interacting atoms are in molecules related by operator X. $P(X|T_B \ldots T_P)$ is sensitive to the absolute likelihoods with which symmetry operators occur in crystal structures but $R_F(T_B \ldots T_P|X)$ is not.

The study shows that many different types of interactions are symmetry dependent, *i.e.* more likely to occur with some types of symmetries than others.[65] The effect is very often statistically significant and sometimes large. For example, approximately 25% of C=S…H-C interactions are between molecules related by cell translation but only about 6% of C=S…H(N,O) interactions. There are over twice as many S(sulfoxide)…S(sulfoxide) interactions as would be expected by chance when the sulfur atoms are in inversion-related molecules, but only a tenth as many when the molecules are related by 2_1 axes or translation. In common space groups, very short interactions (mainly hydrogen bonds) show enhanced propensity, compared to interactions in general, for (a) inversion in centrosymmetric space groups, (b) glide planes in noncentrosymmetric, non-Sohnke groups and (c) 2_1 axes in Sohnke groups. While the largest symmetry dependencies involve inversion or 2-fold rotation, there are many smaller but significant dependencies involving the other operators. The symmetry dependencies are consistent in different subsets of crystal structures, *e.g.* when structures of chiral and achiral molecules are compared, or Sohnke-group structures with non-Sohnke.

Some of the symmetry dependencies are well known, such as the tendency for cyclic hydrogen-bond motifs to form on inversion centres. Others are self-evident, given a little thought: for example, interactions in which the base and partner atoms have the same type are a natural geometrical consequence of symmetry operators that have no translational component. But many of the dependencies have not been previously reported. For example, interactions in which the base atom is bromine or iodine and the partner atom is oxygen or acceptor nitrogen have a higher propensity to occur with molecules related by translation or glide planes than with other symmetries. In many cases, the reasons for symmetry dependence are as yet unclear and will be an interesting subject for future study. Many symmetry effects are clearly interrelated. For example, the high propensity of thiocarbonyl sulfur to interact with polar hydrogen in an inversion-related molecule must lower its propensity to form contacts to other types of atoms under inversion symmetry.

Intermolecular interactions dictate crystal symmetry. Most importantly, dispersion interactions *en masse* favour symmetries that lead to close packing. However, a significant role may also be played by the symmetry preferences of individual interactions. The methods we have described quantify these preferences. We have used a rather general set of atom types but, subject to the availability of sufficient data, the methods can be used to probe the symmetry preferences of interactions involving much more precisely defined types of atoms. The results are objective and accompanied by realistic estimates of uncertainties, and could be a useful aid for the design of crystal structures with specific symmetries (*e.g.* noncentrosymmetric structures for second harmonic generation). However, a word of caution is appropriate. Establishing that there is a statistically significant relationship between a particular type of interaction and a particular symmetry does not prove that the relationship is causative. Of the many significant symmetry dependencies we have found, some will be capable of simple interpretation ("*this* interaction causes *that* symmetry") but others may be incidental consequences of factors that are difficult to discern.

Symmetry effects have possible implications for the use in drug design of interaction data from small-molecule crystal-structures.[66-68] For example, carbonyl groups have a greater than random propensity to form anti-parallel interactions with each other across inversion centres, whereas O=C…F has higher propensities with 2_1 axes and translation. The popularity of the centrosymmetric carbonyl…carbonyl interaction is not entirely due its strength *per se*, but partly to the synergistic effect of having many other bond dipoles anti-parallel across the inversion centre.[19] Therefore, O=C…F interactions are likely to be more important in situations where inversion symmetry is impossible, *e.g.* when substances are enantiopure. The best practice when drawing inferences for drug design may be to use only data from noncentrosymmetric groups, or data from all space groups, but omitting interactions between inversion-related molecules.

Acknowledgements

Colin Groom of the Cambridge Crystallographic Data Centre is thanked for critically reading the manuscript and Peter Wood of the same institution for helpful comments.

Notes and references

1. L. Pauling and M. Delbrück, *Science*, 1940, **92**, 77-79.
2. W. Nowacki, *Helv. Chim. Acta*, 1942, **25**, 863-878.
3. W. Nowacki, *Helv. Chim. Acta*, 1943, **26**, 459-462.
4. A. I. Kitaigorodsky, *Organic Chemical Crystallography*, Consultants Bureau, New York, 1961.
5. A. I. Kitaigorodsky, *Chem. Soc. Rev.*, 1978, **7**, 133-163.
6. C. P. Brock and J. D. Dunitz, *Chem. Mater.*, 1994, **6**, 1118-1127.
7. J. D. Dunitz and A. Gavezzotti, *Cryst. Growth Des.*, 2005, **5**, 2180-2189.
8. L. Leiserowitz, *Acta Crystallogr., Sect. B: Struct. Sci.,* 1976, **32**, 775-802.
9. S. Eppel and J. Bernstein, *Acta Crystallogr., Sect. B: Struct. Sci.,* 2008, **64**, 50-56.
10. R. E. Marsh, V. Schomaker and F. H. Herbstein, *Acta Crystallogr., Sect. B: Struct. Sci.,* 1998, **54**, 921-924.

11. K. M. Anderson, K. Afarinkia, H. Yu, A. E. Goeta and J. W. Steed, *Cryst. Growth Des.*, 2006, **6**, 2109-2113.
12. K. M. Anderson and J. W. Steed, *CrystEngComm*, 2007, **9**, 328-330.
13. A. Dey and E. Pidcock, *CrystEngComm*, 2008, **10**, 1258-1264.
14. C. P. Brock and L. L. Duncan, *Chem. Mater.,* 1994, **6**, 1307-1312.
15. G. A. Jeffrey and L. Lewis, *Carbohydr. Res*., 1978, **60**, 179-182.
16. G. A. Jeffrey and J. Mitra, *Acta Crystallogr., Sect. B: Struct. Sci.,* 1983, **39**, 469-480.
17. A. Karpfen, *Adv. Chem. Phys.,* 2002, **123**, 469-510.
18. L. E. Seijas, R. Almeida, A. J. Mora and G. E. Delgado, *J. Comp. Methods Sci. Eng.*, 2014, 1**4**, 5-16.
19. S. Lee, A. B. Mallik and D. C. Fredrickson, *Cryst. Growth Des*., 2004, **4**, 279-290.
20. B. K. Saha, A. Nangia and J.-F. Nicoud, *Cryst. Growth Des*., 2006, **6**, 1278-1281.
21. F. H. Allen, *Acta Crystallogr., Sect. B: Struct. Sci,* 2002, **58**, 380-388.
22. K. M. Anderson, M. R. Probert, C. N. Whiteley, A. M. Rowland, A. E. Goeta and J. W. Steed, *Cryst. Growth Des*., 2009, **9**, 1082-1087.
23. R. Taylor and O. Kennard, *Acta Crystallogr., Sect. B: Struct. Sci.*, 1983, **39**, 133-138.
24. F. H. Allen and I.J. Bruno, *Acta Crystallogr., Sect. B: Struct. Sci*, 2010, **66**, 380-386.
25. J. van de Streek and S. Motherwell, *CrystEngComm*, 2007, **9**, 55-64.
26. *SMARTS – A Language for Describing Molecular Patterns*, Daylight Chemical Information Systems, Inc., Laguna Niguel, CA, USA, http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html.
27. L. Infantes and S. Motherwell, *Struct. Chem.,* 2004, **15**, 173-184.
28. R. Taylor, *CrystEngComm*, 2014, **16**, 6852-6865.
29. We eschew the use of formal symbols such as C_2 and σ for the benefit of the general reader.
30. A. Bondi, *J. Phys. Chem.,* 1964, **68**, 441-451.
31. R. S. Rowland and R. Taylor, *J. Phys. Chem*., 1996, **100**, 7384-7391.
32. S. Alvarez, *Dalton Trans*., 2013, **42**, 8617-8636.

33. I. J. Bruno, J. C. Cole, P. R. Edgington, M. Kessler, C. F. Macrae, P. McCabe, J. Pearson and R. Taylor, *Acta Crystallogr., Sect. B: Struct. Sci.*, 2002, **58**, 389-397.
34. C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek and P. A. Wood, *J. Appl. Cryst*., 2008, **41**, 466-470.
35. R. A. Sykes, P. McCabe, F. H. Allen, G. M. Battle, I. J. Bruno and P. A. Wood, *J. Appl. Cryst*., 2011, **44**, 882-886.
36. M. J. Knol, W. R. Pestman and D. E. Grobbee, *Eur. J. Epidemiol*., 2011, **26**, 253-254.
37. T. Steiner, *Chem. Commun*., 1998, 313-314.
38. D. Braga, F. Grepioni and J. J. Novoa, *Chem. Commun*., 1998, 1959-1960.
39. D. Braga, F. Grepioni, E. Tagliavini, J. J. Novoa and F. Mota, *New J. Chem*., 1998, **22**, 755-757.
40. J. J. Novoa, I. Nobeli, F. Grepioni and D. Braga, *New J. Chem*., 2000, **24**, 5-8.
41. E. D'Oria and J. J. Novoa, *CrystEngComm*, 2004, **6**, 368-376.
42. J. J. Novoa, E. D'Oria and M. A. Carvajal, in *Making Crystals by Design: Methods, Techniques and Application*, ed. D. Braga and F. Grepioni, Wiley-VCH, Weinheim, 2007, pp. 25-57.
43. G. R. Desiraju, *Acc. Chem. Res*., 1991, **24**, 290-296.
44. G. Filippini and A. Gavezzotti, *Acta Crystallogr., Sect. B: Struct. Sci.,* 1992, **48**, 230-234.
45. It is not expected that R_F(X…Y) should equal R_F(Y…X).[28]
46. Our R_F values for this and other hydrogen-bonding interactions - *e.g.* involving acceptors such as O[carbonyl] and S[=C] - tend to be higher than those quoted in earlier work.[28] We ascribe this to the absence in our data set of anionic acceptor such as halide and nitrate. In their absence, uncharged acceptors show a greater propensity to form hydrogen bonds.
47. C. A. Hunter, K. R. Lawson, J. Perkins and C. J. Urch, J. *Chem. Soc., Perkin Trans. 2*, 2001, 651-669.

48. E. A. Meyer, R. K. Castellano and F. Diederich, Angew. Chem. Int. Ed., 2003, **42**, 1210-1250.

49. M. P. Waller, A. Robertazzi, J. A. Platts, D. E. Hibbs and P. A. Williams, *J. Comput. Chem*., 2006, **27**, 491-504.

50. S. E. Wheeler, *J. Am. Chem. Soc*., 2011, **133**, 10262-10274.

51. C. R. Martinez and B. L. Iverson, *Chem. Sci*., 2012, **3**, 2191-2201.

52. F. H. Allen, C. A. Baalham, J. P. M. Lommerse and P. R. Raithby, *Acta Crystallogr., Sect. B: Struct. Sci.*, 1998, **54**, 320-329.

53. R. L. Rowley and T. Pakkanen, *J. Chem. Phys*., 1999, **110**, 3368-3377.

54. A. H.-T. Li and S. D. Chao, *J. Chem. Phys*., 2006, **125**, 094312.

55. P. Politzer, J. S. Murray and M. C. Concha, *J. Mol. Model*., 2008, **14**, 659-665.

56. G. R. Desiraju and R. Parthasarathy, *J. Am. Chem. Soc*., 1989, **111**, 8725-8726.

57. The symmetry constraint is for the molecules containing the interacting atoms to be related by a specific symmetry operator, not necessarily the atoms themselves. But in the majority of Cl…Cl, Br…Br, I…I and S…S interactions between inversion- or rotation-related molecules, the atoms involved are indeed symmetry equivalent.

58. M. Capdevila-Cortada, J. Castelló and J. J. Novoa, *CrystEngComm*, 2014, **16**, 8232-8242.

59. T. N. Guru Row and R. Parthasarathy, *J. Am. Chem. Soc.,* 1981, **103**, 477-479.

60. H. –J. Böhm, S. Brode, U. Hesse and G. Klebe, *Chem. Eur. J.,* 1996, **2**, 1509-1513.

61. I. Nobeli, S. L. Price, J. P. M. Lommerse and R. Taylor, *J. Comp. Chem*, 1997, **18**, 2060-2074.

62. R. Paulini, K. Müller and F. Diederich, *Angew. Chem. Int. Ed*., 2005, **44**, 1788-1805.

63. J. W. Tukey, *Biometrics*, 1949, **5**, 99-114.

64. R. G. Miller Jr., *Simultaneous Statistical Inference*, Springer-Verlag, New York, 1981.

65. Our results cast no light on the relationship between interaction type and *pseudosymmetry*.

66. I. J. Bruno, J. C. Cole, J. P. M. Lommerse, R. S. Rowland, R. Taylor and M. L. Verdonk, *J. Comput.-Aided Molec. Design*, 1997, **11**, 525-537.

67. M. L. Verdonk, J. C. Cole and R. Taylor, *J. Mol. Biol.,* 1999, **289,** 1093-1108.

68. C. Bissantz, B. Kuhn and M. Stahl, *J. Med. Chem*., 2010, **53**, 5061-5084.

Table 1 Atom-typing scheme

SMARTS$^{a, b}$	Symbol	Description
[C,c]H	H[C]	Carbon-bound hydrogen
[N,n,O,o,F,P,p,S,s,Cl,Br,I]H	H[polar]	Polar hydrogen
[CX4H3]	C[methyl]	Methyl carbon
[CX4]	C[sat]	Saturated (sp^3) carbon, not methyl
[CX3]=[OX1], [CX3][OX1]	C[carbonyl]	Carbonyl, carboxylate, enolate carbon
[cH1]1[cH1][cH1][cH1][cH1]c1-!@[C,c], [cH1]1[cH1][cH1][cH1][cH1]c1-!@[C,c], [cH1]1[cH1][cH1][cH1][cH1]c1-!@[C,c], [cH1]1[cH1][cH1][cH1][cH1]c1-!@[C,c]	C[C_6H_5]	Phenyl carbon
[C!X4], c	C[unsat]	Unsaturated (sp, sp^2) carbon, not carbonyl, phenyl
[NX3],^{c} [nX3]^{c}	N[planar]	Planar nitrogen
[NX1], [NX2],^{d} [nX2], [NX3]^{e}	N[acc]	Nitrogen able to accept hydrogen bond (has localised lone pair)
O=C[OX2], C=COC=C, ccOC=C, ccOcc	O[conj]	2-Coordinate oxygen in conjugated environment, *e.g.* in furan, esters
[C,c][OX2H1]	O[hydroxyl]	Hydroxyl oxygen
c[OX2][CX4]	O[anisole]	Oxygen in anisole and other aryl-alkyl-ethers
[CX4][OX2][CX4]	O[ether]	Aliphatic ether oxygen
C=[OX1], C[OX1]	O[carbonyl]	Carbonyl, carboxylate, enolate

		oxygen
[N,n,P,p,S,s]=[OX1], [N,n,P,p,S,s][OX1]	O[=X]	Terminal oxygen bonded to nitrogen, phosphorus or sulfur
[CX4](F)(F)F	F[CF$_3$]	Trifluoromethyl fluorine
[C,c][FX1]	F[C≠CF$_3$]	Carbon-bound fluorine, not CF$_3$
[PX3]	P[3-coord]	3-Coordinate phosphorus
N=[C,N][SX2][C,N], [C,N]=N[SX2][C,N], [SX2]=N[SX2][C,N], [SX2][SX2]N=[C,N], [SX2][SX2]N=[SX2], [c,n][sX2]n, N[SX2][SX2], c[sX2][sX2]	S[δ+]	Sulfur in electron-withdrawing environment, *e.g.* in thiazole, isothiazole
C=[SX1], C[SX1]	S[=C]	Thiocarbonyl sulfur (also includes thiolate, thiocyanate, *etc*.)
[C,c][SX2][SX2][C,c]	S[disulfide]	Disulfide sulfur
[CX4][SX2][CX4]	S[thioether]	Aliphatic thioether sulfur
[C,c][SX3]([C,c])=[OX1], [C,c][SX3]([C,c])[OX1]	S[sulfoxide]	Sulfoxide sulfur
[c,C]-[ClX1]	Cl[C]	Carbon-bound chlorine
[c,C]-[BrX1]	Br[C]	Carbon-bound bromine
[c,C]-[IX1]	I[C]	Carbon-bound iodine

[a] Type assigned to underlined atom. If type has >1 SMARTS, a match against any is sufficient. If atom matches SMARTS in >1 row, type assigned is that corresponding to highest row. [b] An atom not matching any SMARTS string is assigned a miscellaneous type; interactions involving this type excluded from Tables 4, 6, 7, S1. [c] Must have sum of bond angles at N >355°. [d] Must have bond angle at N <140°. [e] Must have sum of bond angles at N <345°.

Table 2 Symmetry-operator percentages[a]

	primary	primary, secondary, tertiary	primary (achiral)	primary (chiral)	primary (chiral, Sohnke)	primary (chiral, non-Sohnke)
translation	24.4	24.8	21.7	26.9	31.8	21.3
inversion	25.7	26.5	34.5	17.7	-	38.0
glide	14.4	14.3	19.5	9.8	-	21.1
screw	33.3	32.3	22.6	43.1	64.8	18.2
rotation	1.0	1.0	1.0	1.0	1.0	1.0
other[b]	1.1	1.1	0.7	1.5	2.4	0.5

[a] All values in table have standard errors of <0.1. [b] Trigonal, tetragonal or hexagonal symmetry.

Table 3 Symmetry-operator percentages for all primary interactions and primary interactions shorter than sum of vdw radii minus 0.6Å[a,b]

	translation	inversion	glide	screw	rotation
all space groups[c]	24.4 (<0.1)	25.7 (<0.1)	14.4 (<0.1)	33.3 (<0.1)	1.0 (<0.1)
	19.1 (0.1)	*26.2 (0.2)*	*14.4 (0.1)*	*36.6 (0.2)*	*1.5 (<0.1)*
P$\bar{1}$	33.0 (0.1)	67.0 (0.1)	-	-	-
	22.0 (0.4)	*78.0 (0.4)*	-	-	-
P2_1	40.6 (0.1)	-	-	59.4 (0.1)	-
	39.2 (0.5)	-	-	*60.8 (0.5)*	-
C2	44.7 (0.3)	-	-	36.6 (0.3)	18.7 (0.2)
	38.3 (1.6)	-	-	*37.9 (1.5)*	*23.8 (1.4)*
Pc	44.5 (0.4)	-	55.5 (0.4)	-	-
	28.5 (2.7)	-	*71.5 (2.7)*	-	-
Cc	35.7 (0.2)	-	64.3 (0.2)	-	-
	24.1 (1.5)	-	*75.9 (1.5)*	-	-
P2_1/c	17.4 (<0.1)	32.6 (<0.1)	24.5 (<0.1)	25.4 (<0.1)	-
	10.3 (0.2)	*41.3 (0.3)*	*25.1 (0.3)*	*23.3 (0.3)*	-
C2/c	18.2 (0.1)	33.7 (0.1)	17.5 (0.1)	17.9 (0.1)	12.7 (0.1)
	11.4 (0.5)	*41.5 (0.8)*	*13.7 (0.6)*	*17.5 (0.7)*	*16.0 (0.6)*
P2_12_12	28.9 (0.4)	-	-	50.3 (0.4)	20.8 (0.3)
	26.5 (2.1)	-	-	*53.1 (2.3)*	*20.4 (1.9)*
P$2_12_12_1$	23.6 (0.1)	-	-	76.4 (0.1)	-
	21.1 (0.3)	-	-	*78.9 (0.3)*	-
Pca2_1	19.2 (0.2)	-	54.4 (0.3)	26.5 (0.3)	-
	12.9 (1.3)		*63.0 (1.9)*	*24.1 (1.7)*	
Pna2_1	21.2 (0.2)	-	49.5 (0.2)	29.3 (0.2)	-
	14.6 (1.0)	-	*58.1 (1.4)*	*27.3 (1.3)*	-

Fdd2	22.2 (0.5)	-	47.1 (0.5)	18.8 (0.4)	11.8 (0.4)
	11.9 (2.0)	-	*52.6 (3.1)*	*24.1 (2.7)*	*11.5 (2.0)*
Pbcn	7.0 (0.2)	16.1 (0.3)	44.7 (0.4)	17.2 (0.3)	15.0 (0.3)
	3.6 (1.0)	*19.7 (2.0)*	*36.6 (2.5)*	*17.9 (2.0)*	*22.1 (2.1)*
Pbca	5.3 (0.1)	16.4 (0.1)	50.0 (0.1)	28.3 (0.1)	-
	3.1 (0.3)	*20.8 (0.7)*	*49.5 (0.8)*	*26.6 (0.7)*	-

[a] Figures in normal type based on all primary interactions, those in italics on primary interactions shorter than sum of vdw radii minus 0.6Å. [b] Standard errors in parentheses. [c] Trigonal, tetragonal or hexagonal symmetry operators account for 1.1% of all primary interactions and 2.2% of the short interactions.

Table 4 Symmetry-operator percentages for selected primary interactions[a]

base atom	partner atom	translation	inversion	glide	screw	rotation
S[=C]	H[polar]	6.4 (0.7)	67.4 (1.3)	10.6 (0.8)	14.0 (1.0)	1.7 (0.4)
N[planar]	N[planar]	15.6 (1.3)	61.4 (1.7)	5.3 (0.8)	10.8 (1.1)	6.9 (0.9)
C[carbonyl]	C[carbonyl]	11.8 (0.7)	61.2 (1.0)	6.1 (0.5)	13.1 (0.7)	7.8 (0.5)
N[planar]	C[carbonyl]	21.9 (1.1)	54.1 (1.3)	9.7 (0.8)	10.8 (0.8)	3.4 (0.5)
C[C_6H_5]	C[C_6H_5]	23.4 (0.3)	44.3 (0.4)	9.9 (0.2)	20.2 (0.3)	2.2 (0.1)
N[acc]	H[polar]	12.2 (0.4)	31.3 (0.5)	21.7 (0.5)	32.1 (0.5)	2.7 (0.2)
S[=C]	H[C]	25.5 (1.1)	25.5 (1.1)	19.3 (1.0)	29.4 (1.2)	0.3 (0.1)
C[methyl]	F[C≠CF_3]	32.3 (1.9)	23.0 (1.7)	15.8 (1.5)	28.2 (1.9)	0.7 (0.3)
C[C_6H_5]	H[C]	24.1 (0.1)	22.0 (0.1)	14.4 (0.1)	38.8 (0.1)	0.8 (<0.1)
O[ether]	H[polar]	26.4 (1.0)	13.4 (0.8)	7.4 (0.6)	51.5 (1.1)	1.3 (0.3)
O[hydroxyl]	H[polar]	26.0 (0.5)	11.7 (0.4)	9.3 (0.3)	52.0 (0.6)	1.0 (0.1)
C[methyl]	O[ether]	31.3 (1.5)	9.8 (1.0)	7.1 (0.9)	50.7 (1.7)	1.1 (0.3)

[a] Standard errors in parentheses.

Table 5 Symmetry-operator percentages for primary interactions involving OH and NH groups[a,b]

base atom	partner atom	translation	inversion	glide	screw	rotation
N[acc]	H[O]	15.4 (0.7)	23.6 (0.8)	17.8 (0.7)	41.2 (0.9)	2.0 (0.3)
N[acc]	H[N]	10.5 (0.4)	35.4 (0.7)	23.8 (0.6)	27.3 (0.6)	3.1 (0.2)
O[=X]	H[O]	22.8 (1.1)	25.8 (1.2)	13.1 (0.9)	37.7 (1.3)	0.6 (0.2)
O[=X]	H[N]	17.5 (0.7)	32.6 (0.9)	20.2 (0.7)	28.5 (0.8)	1.2 (0.2)
O[anisole]	H[O]	22.7 (2.1)	21.6 (2.1)	16.0 (1.9)	37.6 (2.5)	2.1 (0.7)
O[anisole]	H[N]	24.2 (2.1)	26.5 (2.1)	20.1 (2.0)	28.7 (2.2)	0.5 (0.3)
O[carbonyl]	H[O]	18.4 (0.4)	28.9 (0.5)	10.7 (0.3)	40.7 (0.5)	1.3 (0.1)
O[carbonyl]	H[N]	21.6 (0.4)	30.6 (0.4)	17.6 (0.4)	28.8 (0.4)	1.4 (0.1)
O[ether]	H[O]	27.6 (1.1)	11.6 (0.8)	6.7 (0.6)	52.8 (1.3)	1.3 (0.3)
O[ether]	H[N]	22.0 (2.0)	20.1 (2.0)	9.9 (1.5)	46.5 (2.5)	1.5 (0.6)
O[hydroxyl]	H[O]	26.0 (0.6)	11.0 (0.4)	8.1 (0.4)	53.8 (0.6)	1.0 (0.1)
O[hydroxyl]	H[N]	26.0 (1.2)	14.6 (0.9)	14.2 (0.9)	44.1 (1.3)	1.1 (0.3)

[a] Standard errors in parentheses. [b] All rows based on >500 interactions except for O[anisole]…H[O] (based on 388), O[anisole]…H[N] (422), O[ether]…H[N] (413)

Table 6 R_F values of most significant symmetry-dependent interactions[a,b]

base atom	partner atom	R_F(T)	R_F(I)	R_F(G)	R_F(S)	R_F(R)	ratio[c]	significant differences
Cyclic hydrogen-bond motifs:								
N[acc]	H[polar]	5.1	8.3	8.3	8.3	12.8	2.5	R>G,I,S,T; G>T; I>T; S>T
S[=C]	H[polar]	7.2	16.6	9.7	10.2	[d]	2.3	I>S,G,T; S>T; G>T
H[polar]	N[acc]	3.5	4.9	4.7	4.9	6.9	2.0	R>S,I,G,T; S>T; I>T; G>T
O[ether]	H[polar]	4.3	8.2	6.6	5.3	4.8	1.9	I>G,S,R,T; G>S,T; S>T
H[polar]	S[=C]	2.7	4.9	3.0	3.3	[d]	1.8	I>S,G,T; S>T
O[=X]	H[polar]	6.9	11.1	8.1	8.4	[d]	1.6	I>S,G,T; S>T; G>T
O[carbonyl]	*H[polar]*	*11.6*	*16.6*	*14.4*	*12.5*	*15.6*	*1.4*	*I>G,S,T; R>S,T; G>S,T; S>T*
Stacking interactions:								
N[planar]	C[carbonyl]	1.0	1.7	0.7	0.4	2.5	6.2	R>I,T,G,S; I>T,G,S; T>G,S; G>S
C[unsat]	C[carbonyl]	1.0	1.6	1.0	0.6	1.4	2.8	I>T,G,S; R>T,G,S; T>S; G>S
N[planar]	C[unsat]	1.2	1.4	1.0	0.7	1.2	1.9	I>R,T,G,S; R>G,S; T>G,S; G>S
S[sulfoxide]	O[=X]	1.0	1.8	1.4	1.1	[d]	1.9	I>T
C[carbonyl]	O[carbonyl]	0.9	1.4	1.1	1.3	1.2	1.6	I>R,G,T; S>G,T; R>T; G>T
O[=X]	*C[carbonyl]*	*1.1*	*1.4*	*1.1*	*1.1*	[d]	*1.4*	*I>S,T*
Like-atom interactions (stacking):								
S[sulfoxide]	S[sulfoxide]	0.1	2.4	0.2	0.1	[d]	30.2	I>G,S,T
N[planar]	N[planar]	0.4	1.2	0.2	0.3	3.7	18.2	R>I,T,S,G; I>T,S,G;

								T>G
C[carbonyl]	C[carbonyl]	0.4	1.7	0.4	0.3	4.7	15.0	R>I,T,G,S; I>T,G,S; T>S
C[unsat]	C[unsat]	0.8	1.2	0.6	0.5	1.4	3.0	R>I,T,G,S; I>T,G,S; T>G,S; G>S
C[C$_{6}$H$_{5}$]	*C[C$_{6}$H$_{5}$]*	*0.5*	*0.7*	*0.4*	*0.3*	*1.0*	*3.0*	*R>I,T,G,S; I>T,G,S; T>G,S*
Like-atom interactions (other):								
P[3-coord]	*P[3-coord]*	*0.2*	*1.1*	*0.1*	*0.3*	^{d}	*14.5*	*I>S,T,G*
S[thioether]	S[thioether]	0.5	1.3	0.2	0.3	^{d}	6.4	I>T,S,G; T>G
S[δ+]	S[δ+]	0.4	1.8	0.3	0.6	^{d}	6.0	I>S,T,G; S>G
C[methyl]	C[methyl]	0.6	1.0	0.6	0.6	2.3	4.0	R>I,T,S,G; I>T,S,G
I[C]	I[C]	0.5	1.5	0.5	1.0	^{d}	3.1	I>S,G,T; S>G,T
Cl[C]	Cl[C]	0.7	1.1	0.7	0.7	2.0	3.0	R>I,T,S,G; I>T,S,G
Br[C]	*Br[C]*	*0.6*	*1.2*	*0.6*	*0.7*	^{d}	*1.9*	*I>S,G,T*
F[CF$_{3}$]	*F[CF$_{3}$]*	*0.8*	*1.2*	*0.9*	*0.9*	^{d}	*1.4*	*I>G,S,T*
Other hydrogen bonds:								
H[polar]	O[ether]	3.5	5.1	5.7	3.4	3.7	1.7	G>R,T,S; I>T,S
H[polar]	O[anisole]	2.6	1.6	2.5	2.4	^{d}	1.6	T>I; G>I; S>I
O[conj]	*H[polar]*	*1.2*	*0.6*	*0.9*	*1.1*	*0.9*	*1.9*	*T>I; S>I*
O[anisole]	*H[polar]*	*3.8*	*3.1*	*4.1*	*4.0*	^{d}	*1.3*	*G>I; S>I*
O[hydroxyl]	*H[polar]*	*5.1*	*5.8*	*5.3*	*6.3*	*5.3*	*1.2*	*S>G,T; I>T*
Halogen bonds:								
Br[C]	O[ether]	0.9	0.4	^{d}	1.7	^{d}	4.6	S>I
Br[C]	O[anisole]	2.3	0.8	1.7	1.7	^{d}	2.9	T>I
Br[C]	O[carbonyl]	1.7	1.1	1.7	1.6	^{d}	1.6	T>I; G>I; S>I
Br[C]	O[=X]	2.3	1.5	1.8	1.8	^{d}	1.5	T>I

Chalcogen bonds:								
S[δ+]	O[carbonyl]	1.1	1.0	2.1	1.5	^{d}	2.1	G>I
Orthogonal polar interactions of carbonyl carbon:								
C[carbonyl]	O[anisole]	1.1	0.9	1.3	1.5	0.7	2.1	S>T,I; G>I
C[carbonyl]	I[C]	1.5	0.8	^{d}	1.2	^{d}	2.0	T>I
C[carbonyl]	*F[CF$_3$]*	*1.1*	*0.6*	*1.1*	*1.1*	^{d}	*1.8*	*T>I; G>I; S>I*
C[carbonyl]	F[C≠CF$_3$]	1.4	0.9	1.4	1.6	^{d}	1.8	S>I; G>I; T>I
C[carbonyl]	Br[C]	1.2	1.0	1.1	1.6	^{d}	1.5	S>T,G,I
C[carbonyl]	*O[=X]*	*2.0*	*1.5*	*1.9*	*1.8*	^{d}	*1.3*	*T>I; G>I; S>I*
C[carbonyl]	*Cl[C]*	*1.4*	*1.1*	*1.1*	*1.3*	^{d}	*1.2*	*T>I*
Hydrophobic interactions associated with screw axes:								
N[planar]	H[C]	1.2	0.9	1.3	1.4	0.8	1.7	S>G,T,I,R; G>T,I,R; T>I,R
C[carbonyl]	H[C]	1.3	1.0	1.3	1.4	0.9	1.6	S>G,I,R; T>I,R; G>I,R
C[unsat]	H[C]	1.6	1.3	1.8	1.9	1.2	1.6	S>G,T,I,R; G>T,I,R; T>I,R; I>R
H[C]	*C[C$_6$H$_5$]*	*0.9*	*0.9*	*1.0*	*1.1*	*0.9*	*1.2*	*S>G,I,T,R; G>I,T,R*
Other:								
O[hydroxyl]	O[hydroxyl]	0.1	0.6	0.2	0.1	2.3	16.0	R>I,G,T,S; I>G,T,S
S[=C]	H[C]	2.2	0.9	2.0	1.9	^{d}	2.4	T>S,I; G>I; S>I
N[planar]	O[hydroxyl]	1.1	0.7	1.5	1.2	^{d}	2.1	G>I; S>I; T>I
N[planar]	F[CF$_3$]	1.4	0.8	1.0	1.1	^{d}	1.8	T>I; S>I
H[C]	F[C≠CF$_3$]	2.0	1.7	1.9	1.9	1.1	1.8	T>I,R; G>I,R; S>I,R; I>R
H[C]	O[hydroxyl]	1.6	1.6	1.8	1.5	1.1	1.7	G>T,I,S,R; T>S,R; I>S,R; S>R
H[C]	O[=X]	2.3	2.1	2.2	2.1	1.3	1.7	T>I,S,R; G>I,S,R; I>R;

								S>R
H[C]	S[=C]	2.0	1.9	1.9	1.8	1.2	1.6	T>G,S,R; I>S,R; G>R; S>R
N[acc]	H[C]	2.2	1.9	2.2	2.1	1.4	1.6	G>S,I,R; T>S,I,R; S>I,R; I>R
N[planar]	Br[C]	1.0	0.9	1.0	1.4	[d]	1.6	S>T,G,I
S[sulfoxide]	H[C]	1.6	1.1	1.8	1.6	[d]	1.6	G>I; T>I; S>I
O[hydroxyl]	H[C]	1.7	1.8	1.9	1.5	1.3	1.5	G>I,T,S,R; I>T,S,R; T>S,R; S>R

[a] Only R_F values based on at least 100 structures are given. [b] Entries in italics do not meet all of the criteria outlined in text. [c] Ratio of largest R_F to smallest R_F. [d] Insufficient interactions for R_F value to be calculated.

Many intermolecular interactions show preferences for particular crystallographic symmetries and the extent to which this is the case is quantifiable.