

The enrichment ratio of atomic contacts in crystals, an indicator derived from the Hirshfeld surface analysis

Christian Jelsch,^{a*} Krzysztof Ejsmont^b and Loïc Huder^a

Received 30 December 2013

Accepted 13 February 2014

^aCRM², CNRS, Institut Jean Barriol, Université de Lorraine, BP 70239, 54506 Vandoeuvre les Nancy, France, and

^bFaculty of Chemistry, Opole University, ul. Oleska 48, 45-052 Opole, Poland. *Correspondence e-mail:

christian.jelsch@univ-lorraine.fr

Edited by M. Eddaoudi, King Abdullah University, Saudi Arabia

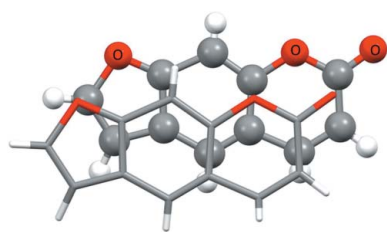
Keywords: enrichment ratio; Hirshfeld surface analysis; crystal packing; fingerprint plots

Supporting information: this article has supporting information at www.iucrj.org

The partitioning of space with Hirshfeld surfaces enables the analysis of fingerprint molecular interactions in crystalline environments. This study uses the decomposition of the crystal contact surface between pairs of interacting chemical species to derive an enrichment ratio. This quantity enables the analysis of the propensity of chemical species to form intermolecular interactions with themselves and other species. The enrichment ratio is obtained by comparing the actual contacts in the crystal with those computed as if all types of contacts had the same probability to form. The enrichments and contact tendencies were analyzed in several families of compounds, based on chemical composition and aromatic character. As expected, the polar contacts of the type H···N, H···O and H···S, which are generally hydrogen bonds, show enrichment values larger than unity. O···O and N···N contacts are impoverished while H···H interactions display enrichment ratios which are generally close to unity or slightly lower. In aromatic compounds, C···C contacts can display large enrichment ratios due to extensive π ··· π stacking in the crystal packings of heterocyclic compounds. C···C contacts are, however, less enriched in pure (C,H) hydrocarbons as π ··· π stacking is not so favourable from the electrostatic point of view compared with heterocycles. C···H contacts are favoured in (C,H) aromatics, but these interactions occur less in compounds containing O, N or S as some H atoms are then involved in hydrogen bonds. The study also highlights the fact that hydrogen is a preferred interaction partner for fluorine.

1. Introduction

The involvement of functional groups in crystal packing contacts is widely investigated in the literature in order to better understand the principles of crystal formation. With the crystallographic analysis tools available using Hirshfeld surfaces, the crystal packing and contacts of chemical compounds can be studied extensively (Spackman *et al.*, 2002). Fingerprint plots can be made by evaluating the pairs of d_i and d_e distances from the surface to the nearest atom interior/exterior to the surface, respectively. *CrystalExplorer* software (McKinnon *et al.*, 2004; Wolff *et al.*, 2012) is now a tool widely used for visualizing and exploring intermolecular interactions in crystals. The program notably enables fingerprinting of intermolecular interactions by pairs of chemical species. In this manner different interaction types such as hydrogen bonding, van der Waals contacts, C—H··· π and π ··· π stacking can be identified in the fingerprint plots. The d_{norm} quantity, which was introduced by McKinnon *et al.* (2007), is negative where contacts are shorter than the van der Waals separation, and



positive for longer contacts. This quantity plotted on a Hirshfeld surface provides a very informative tool to analyze crystal packings.

The *CrystalExplorer* software is also a powerful tool for the visualization and characterization of voids in crystalline materials (Turner *et al.*, 2011). When electrostatic potentials of molecules are mapped onto their Hirshfeld surfaces and displayed within a crystal packing diagram, the molecular contacts can be analyzed in terms of the electrostatic complementarity of touching surface patches (Spackman *et al.*, 2008).

In the current paper a new descriptor is proposed to analyze contacts in molecular crystals. The surface contact data derived from the Hirshfeld surface analysis are used to derive enrichment ratios which relate to the propensity of pairs of chemical species to form crystal-packing interactions.

Several series of compounds composed of the same chemical species were analyzed to identify trends in crystal-packing contacts. A selection of CIF files containing the structural coordinates of the molecules were retrieved from the Cambridge Structural Database (CSD; Allen, 2002). For the crystal structures collected, the contacts are analyzed in terms of chemical types involved at the Hirshfeld surface. The propensity of two chemical species to be in contact are analyzed and discussed for several sets of crystal structures containing a limited number of elements.

2. Materials and methods

2.1. Hirshfeld surface and fingerprint plots

Hirshfeld surfaces were introduced by Spackman & Byrom (1997) to partition the space in molecular crystals for electron-density integration purposes. The Hirshfeld surface is a region of crystal space around the molecule that is defined by $W(\mathbf{r}) = 0.5$, where W is a weight function derived from the atomic electron densities $\rho_i(\mathbf{r})$

$$W(\mathbf{r}) = \sum_{i \in \text{molecule}} \rho_i(\mathbf{r}) / \sum_{i \in \text{crystal}} \rho_i(\mathbf{r}). \quad (1)$$

Thus, the Hirshfeld surface can be considered as the frontier between regions where the electron distribution is dominated by the contribution of the reference molecule (interior) and of the neighbouring molecules in the crystal (exterior).

The program *CrystalExplorer* (Wolff *et al.*, 2012) enables the analysis and visualization of crystal contacts through the Hirshfeld surface. Using the CIF files *CrystalExplorer* can compute the Hirshfeld surface and then the fingerprint plots. The contact surface decomposition in these plots by specific pairs of chemical elements is the interesting resulting information. This allows the calculation of derived properties such as the percentage of surface that involves a particular element.

2.2. Definition of the enrichment ratio

The percentage of contacts between one ($X \cdots X$) or two ($X \cdots Y$) chemical elements in a crystal packing is information given by *CrystalExplorer* which can be used to indirectly calculate the enrichment ratios.

The proportion of Hirshfeld surface contacts involving the (X, Y) pair of elements is referred to as C_{XY} . The proportion S_X of chemical type X on the molecular surface is obtained by the summation

$$S_X = C_{XX} + \frac{1}{2} \sum_{Y \neq X} C_{XY}. \quad (2)$$

The value C_{XY} includes both $X \cdots Y$ and $Y \cdots X$ contacts in the S_X sum, where the first and second atoms are interior and exterior to the Hirshfeld surface. The factor $\frac{1}{2}$ relates to the fact that C_{XY} contributes to both S_X and S_Y summations. The summation of all the surface proportions is equal to unity

$$\sum_X S_X = 1. \quad (3)$$

The ratio of random contacts R_{XY} between the two chemical elements X and Y is then introduced. The R_{XY} values are defined as if all contact types $X \cdots Y$ in the crystal packing were equi-distributed between all chemical types and are obtained by probability products

$$R_{XX} = S_X S_X \quad \text{and} \quad R_{XY} = 2 S_X S_Y. \quad (4)$$

The R_{XY} proportion again includes both $X \cdots Y$ and $Y \cdots X$ contacts, which explains the presence of factor 2 in this definition.

The sum of all random contact proportions also yields unity, by definition

$$\sum_X R_{XX} + \sum_{Y \neq X} R_{XY} = 1. \quad (5)$$

The enrichment ratio E_{XY} of a pair of elements (X, Y) is then defined as the ratio between the proportion of actual contacts in the crystal and the theoretical proportion of random contacts

$$E_{XY} = C_{XY} / R_{XY}. \quad (6)$$

The enrichment ratio is expected to be generally larger than unity for pairs of elements which have a high propensity to form contacts in crystals, while pairs which tend to avoid contacts with each other should yield an E value lower than unity.

2.3. Selection of molecules

The enrichment ratios E_{XY} were computed for sets of molecules belonging to several classes of molecules and were analyzed as a function of S_X values. In order to limit the number of E_{XY} values and S_X variables present in the contacts data, compounds containing only two or three different chemical elements were selected and grouped in their respective set: CH, CHN, CHS, CHO, CHF.

Aliphatic molecules are defined here as being without aromatic rings and are generally devoid of double or triple bonds, except carbonyl and carboxylate groups which were accepted. These compounds generally have a large number of H atoms on their surface (S_H large), while C atoms are rarely present at the molecular surface ($S_C \simeq 0$), as they generally form four bonds with other atoms.

On the other hand, 'aromatic sets' of molecules were selected for having mainly planar rings and double bonds. In these molecules, carbon is significantly present on the mole-

cular surface ($S_C > 0$). Due to this major difference in the molecular surface composition, aliphatic and aromatic molecules were separated into different groups. The molecules were selected in the CSD in order to avoid ambiguities and were either purely (or mostly) aromatic or largely aliphatic.

3. Results and discussion

3.1. Example of enrichment ratio calculation

Privileged and disfavoured contacts from a chemical elements point of view can be highlighted in a crystal structure *via* the enrichment ratios. An example of a calculation for the coumarin 102 molecule ($C_{16}H_{17}NO_2$) is given in Table 1. This molecule is composed of two aromatic cycles (one of which is heterocyclic), two aliphatic cycles and a methyl group. More than three quarters of the molecular surface is generated by H atoms, while the two other significantly contributing elements are carbon and oxygen, slightly above 10%. C atoms are exposed to the molecular surface only in molecules where they form fewer than four covalent bonds, which is the case in aromatic rings and more generally for Csp^2 atoms. Among the actual contacts, $H\cdots H$ accounts for more than half of the interaction surface, while $C\cdots H$ and $O\cdots H$ contacts are the only other significant contributions ($> 15\%$). $C\cdots H$ contacts represent $C-H\cdots\pi$ interactions on the two sides of the aromatic cycles and $O\cdots H-C$ contacts are weak hydrogen bonds, in the case of coumarin 102.

The list of enrichment ratios highlights the $O\cdots H$ contacts ($E_{OH} = 1.24$, Table 1) which turn out to be favoured in the crystal packing, notably the $O=C$ oxygen atoms form three weak hydrogen bonds of the type $O\cdots H-C$. The $C\cdots H$ and $H\cdots H$ contacts are, respectively, slightly enriched and disfavoured ($E_{CH} = 1.04$ and $E_{HH} = 0.96$). The enrichment ratios of other types of contacts shall be analyzed cautiously as resulting from quotients of two small numbers. Their contributions are very small, below 2.6% if equiprobable contacts R_{XY} are considered. It can still be observed that $C\cdots C$ contacts are favoured ($E_{CC} = 1.5$) and that $O\cdots O$ contacts are completely avoided in coumarin 102 with $E_{OO} = 0.0$.

3.2. CH aromatic compounds

The crystal contacts of a series of aromatic compounds containing only H and C atoms are analyzed in Fig. 1. The crystal packings of polynuclear aromatic hydrocarbons have been classified as stack, layer, glide or herringbone and their occurrences have been rationalized on the relative importance of $C\cdots C$ and $C\cdots H$ interactions, which depend on the H/C composition and positioning in the molecule (Desiraju & Gavezzotti, 1989).

The $H\cdots H$ contacts appear with enrichment ratios close to unity and necessarily constitute most of the interaction surface as S_H is larger than 65% in this set of molecules. Matta *et al.* (2003) analyzed ubiquitous hydrogen \cdots hydrogen bonding by the quantum theory of atoms in molecules and found that they are stabilizing interactions in crystal packings. Interestingly, $H\cdots H$ contacts were found to become extremely prominent at

Table 1

Example of Hirshfeld contact surfaces and derived 'random contact' and 'enrichment ratios' for the coumarin-102 crystal (Bibila Mayaya Bisseyou *et al.*, 2012).

Values in italics at the top of the table are the data obtained from *CrystalExplorer*. The enrichment ratios were not computed when the 'random contacts' were lower than 0.9%, as they are not meaningful.

Atoms	H	C	N	O
H	<i>56.4</i>	–	Contacts	(%)
C	<i>19.9</i>	2.3	–	–
N	<i>0.9</i>	<i>0.0</i>	<i>0.0</i>	–
O	<i>19.5</i>	<i>0.5</i>	<i>0.5</i>	<i>0.0</i>
Surface %	76.6	12.5	0.7	10.3
H	58.6	Random	Contacts	(%)
C	19.1	1.6	–	–
N	1.1	0.2	0.0	–
O	15.7	2.6	0.1	1.1
H	0.96	–	–	Enrichment
C	1.04	1.5	–	–
N	0.8	/	/	–
O	1.24	0.20	/	0.0

high pressures in surface fingerprint plots (of compounds containing C, H, N and O atoms) and did not appear to compress below a $H\cdots H$ distance of 1.7 Å in a recent study (Wood *et al.*, 2008).

On the other hand, $H\cdots C$ contacts are generally slightly favoured in a sample of CH aromatic molecules with an average E_{HC} value of 1.1. The $C\cdots C$ contacts are in most cases very disfavoured with $E_{CC} < 0.5$. The case of the benzene crystal with extensive $C-H\cdots\pi$ stacking is characteristic of the CH aromatic class of compounds, as it displays a E_{CC} value of zero (Fig. 1), $E_{CH} = 1.21$ and $E_{HH} = 0.96$.

However, three compounds relatively rich in carbon on the Hirshfeld surface ($S_c > 0.24$) show slight E_{CC} enrichments around 1.2. These molecules have large aromatics parts

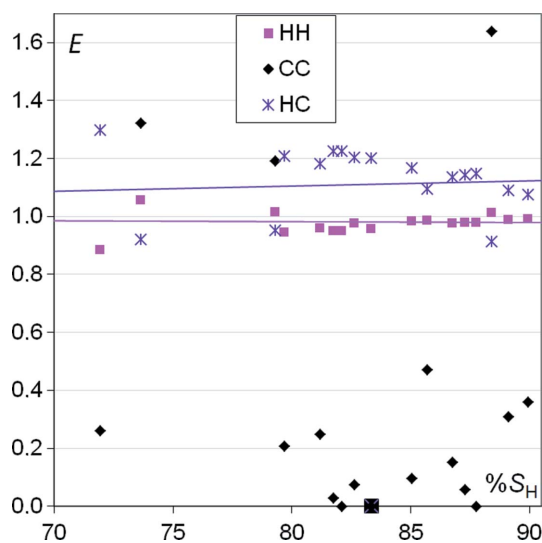


Figure 1

Contact enrichment ratios in crystals of aromatic CH compounds as a function of hydrogen proportion on the Hirshfeld surface. The case of benzene is highlighted.

composed of three or four adjacent rings and are involved in $\pi \cdots \pi$ stacking.

The geometry of aromatic interactions has been investigated notably by Hunter (1994; Hunter *et al.*, 2001), who explained that $\pi \cdots \pi$ stacking can be associated with repulsive electrostatic interactions and is therefore less favourable than could be expected on the basis of the van der Waals interactions. Due to π -bonding in aromatic rings, the deformation electron-density peaks on the C—C bonds are elongated in the direction perpendicular to the aromatic ring and a global quadrupole is generated in that direction. On the other hand, π -facial hydrogen bonds (C \cdots H interactions) are favoured by the charge distribution.

3.3. CHN aromatic compounds

Enrichment ratios of the different types of contacts are shown for a sample of aromatic CHN molecules in Fig. 2. The set of CHN aromatic molecules is composed of molecules with (C,N) heterocycles, but also contains nitrile and planar N atoms (—CN, —NN, —NH₂). Fig. 2 and several subsequent figures display the different enrichment ratios as a function of the molecular surface percentage due to H atoms. This element was chosen instead of others because it shows a large range of surface proportion values, making the graphs more straightforward and easier to read. Moreover, hydrogen is the most frequent chemical type occurring at the surface of organic compounds, therefore the proportion of H surface is an important indicator. Other representations of the data, with a different quantity chosen for the *X* axis (for example *S_N*), are

shown in the supporting information and yield another perspective and additional information on the crystal contacts.

In Fig. 2 the enrichments of the H \cdots H contacts are stable around unity, generally slightly below. The H \cdots N contacts, which can be N—H \cdots N or C—H \cdots N hydrogen bonds or H \cdots π contacts, are favoured as they consistently show enrichments between 1.0 and 1.6, but they are slightly more enriched for molecules with the lowest content of hydrogen. The N \cdots N contacts generally show the lowest enrichment with all values below 0.9 and half of the compounds actually displaying no N \cdots N contact.

On the other hand, C \cdots C contacts show a wide range of enrichments from 0 to 3. The general tendency for *E_{CC}* is to be high for molecules with low hydrogen content, but small, even reaching zero values, for compounds rich in hydrogen. On the contrary, H \cdots C interactions are more likely to occur when *S_H* is high. As a consequence, C—H \cdots π and N—H \cdots π interactions are found to be favoured when H atoms are highly available, but replaced by some C \cdots C contacts, which are π stackings, when H atoms are less abundant.

The graph of *E* ratios as a function of *S_N* values (Fig. S1) shows that, in CHN aromatic molecules, the propensity for H \cdots H and H \cdots N interactions remains stable. On the other hand, H \cdots C contacts are more enriched for compounds with poor nitrogen content and the *E_{HC}* coefficient is then similar to that found for CH aromatic compounds (Fig. 1). Compounds rich in nitrogen display generally lower *E_{HC}* values as more H atoms are involved in H \cdots N contacts which can be hydrogen bonds. N \cdots N contacts are disfavoured compared with C \cdots N interactions which have themselves *E_{CN}*

values generally lower than unity. Both contact types show increasing enrichments as nitrogen becomes more abundant on the molecular surface. This is to be related to the higher propensity of heterocycles to form π stacking as δ^+ and δ^- polarized atoms can be interacting partners (Salonen *et al.*, 2011).

3.4. CHN aliphatic nitrile compounds

A series of CH aliphatic molecules containing additional nitrile groups was also analyzed. The proportion of carbon on the Hirshfeld surface is generally small in these compounds and originates from the nitrile groups. The trends are similar to CHN aromatic molecules for H \cdots C, C \cdots N and N \cdots N contacts (Fig. 2*b*). *E_{HN}* enrichments are found to vary more with the *S_H* proportion.

The behaviour of H \cdots H contacts is different as *E_{HH}* is found to

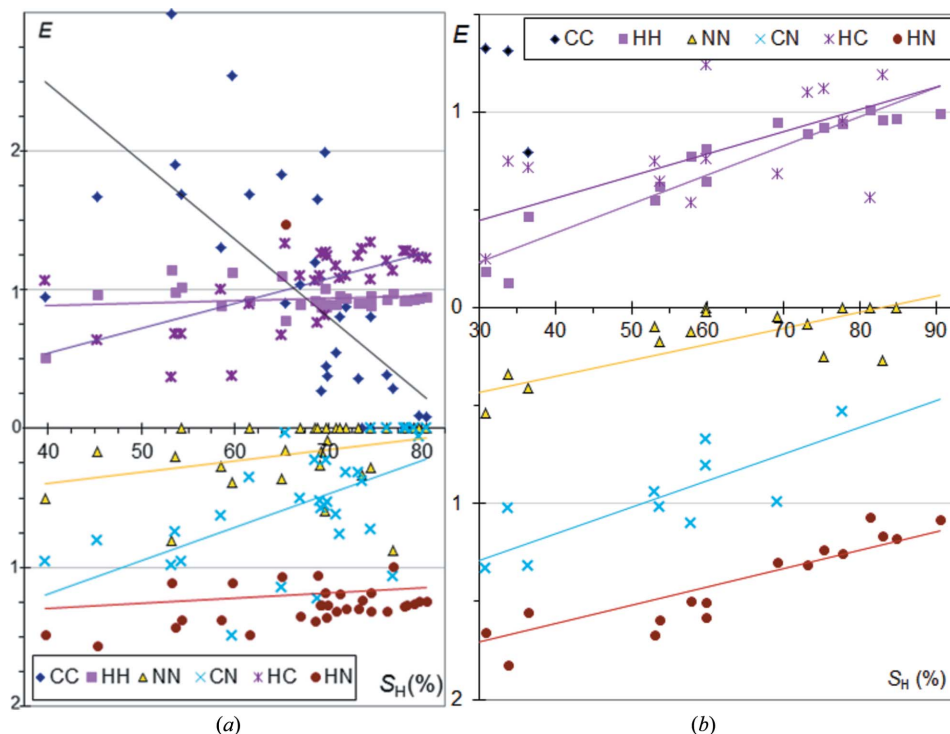


Figure 2

Contact enrichment ratios in crystals of CHN compounds as a function of hydrogen proportion on the Hirshfeld surface. (a) Aromatics; (b) CH aliphatics containing nitrile groups.

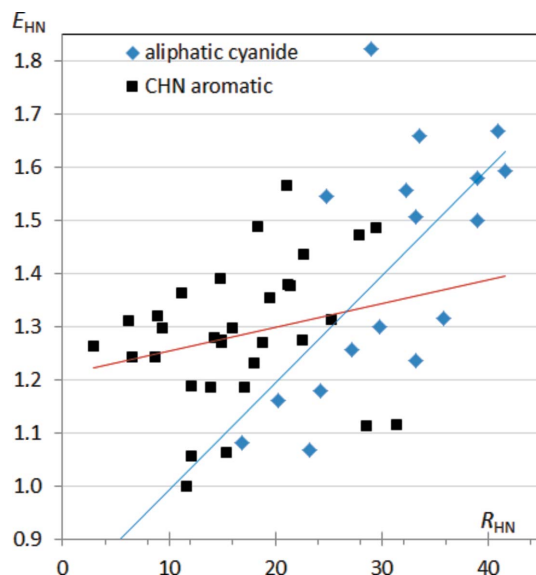


Figure 3
Enrichment of H \cdots N contacts in CHN compounds as a function of the proportion R_{HN} (%) in CHN-type molecules.

decrease significantly below unity for molecules with poor hydrogen content. There is a very clear depletion of H \cdots H contacts presumably due to the involvement of H atoms in C–H \cdots N hydrogen bonds. For compounds with a very high hydrogen content on the surface, the enrichment E_{HH} tends, by definition, towards a value close to 1.

The C \cdots N contacts tend to be enriched for molecules with a low content in hydrogen but are very impoverished when S_{H} is large, presumably as the N atoms are then involved in C–H \cdots N attractive hydrogen bonds.

The same data viewed as a function of S_{N} show systematically significant slopes for the fitted curves of the enrich-

ment ratios (Fig. S1). A global increase of H \cdots N and C \cdots N contact enrichments is observed as nitrogen becomes more abundant. The N \cdots N contacts also show an increase of enrichments with S_{N} , but the values always stay well below unity. A major difference with CHN aromatic compounds is that E_{HN} and E_{HH} values are not stable but increase/decrease significantly, respectively, with S_{N} .

Enrichment ratios can also be analyzed as a function of other variables, such as the proportion of random (or effective) contact surface in the crystal packing. Fig. 3 shows the enrichment ratio of H \cdots N contacts as a function of R_{HN} random contact proportion, which depends on the content of C and N atoms on the Hirshfeld surface. The E_{HN} enrichments generally increase with R_{HN} for both types of CHN molecule. However, aliphatic nitrile compounds show a greater slope of E_{CN} versus R_{CN} compared with aromatic molecules.

3.5. CHO compounds

In CHO aromatic compounds (Fig. 4a) the O \cdots O contacts are quite impoverished, generally with E_{OO} below 0.5. On the other hand, H \cdots O contacts with E_{HO} in the [1.2, 1.6] range are, as expected, quite enriched due to the formation of energetically favourable hydrogen bonds. The enrichment of H \cdots H contacts shown as a function of S_{H} remains stable around unity, but decreases slightly when S_{O} augments (Fig. S2), due to competition with H \cdots O hydrogen bonds.

The proportion of carbon on the Hirshfeld surface for this set of molecules is between 16 and 24%. C \cdots O and H \cdots C contacts are generally impoverished with average values of $\langle E_{\text{CO}} \rangle = 0.8$ and $\langle E_{\text{HC}} \rangle = 0.7$, which are below unity. Despite a few exceptions, the C \cdots C contacts are the most enriched ones, with many molecules overcoming the 2.5 mark, especially the compounds low in hydrogen at the surface. The E_{CC} enrichment is generally higher in CHO (and CHN) compounds compared with pure aromatic CH hydrocarbons. This is in accordance with reports that π -stacking between aromatic heterocycles is favourable in face-to-face geometry where there is a good superimposition of the positively and negatively charged atoms (Salonen *et al.*, 2011; Martinez & Iverson, 2012). An example of CHO aromatic compounds displaying extensive $\pi\cdots\pi$ stacking is shown in Fig. 5 for furo[3,2-*g*]coumarin (Bideau *et al.*, 1979), a molecule which is also present in Fig. 4(a) with $S_{\text{H}} = 53\%$ and $E_{\text{CC}} = 2.91$. In $\pi\cdots\pi$ stacking the O atoms interact with positively charged C atoms which are themselves covalently bonded to O atoms. In the other directions, the interactions are of the type H \cdots H and H \cdots O,

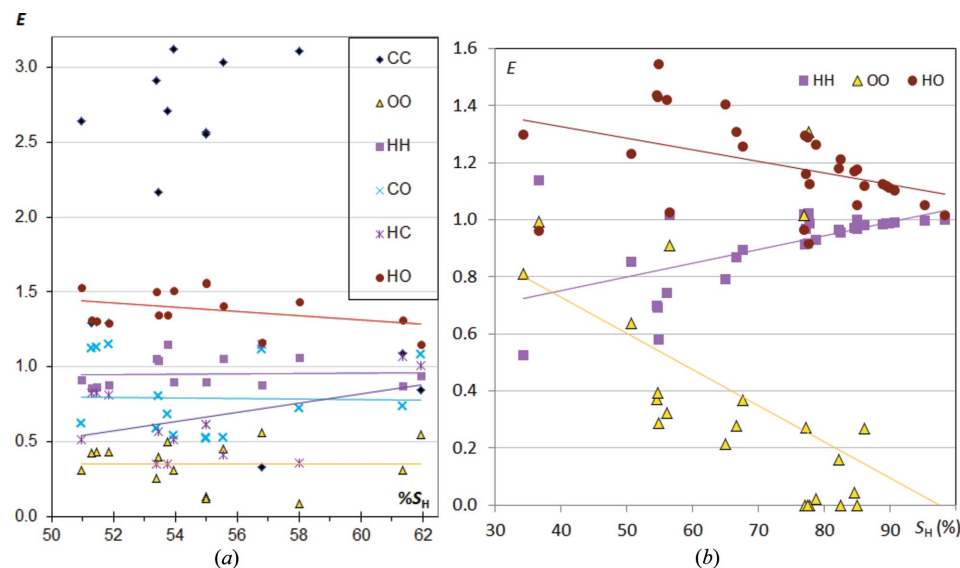


Figure 4
Contact enrichment ratios in crystals of CHO compounds as a function of hydrogen on the Hirshfeld surface: (a) aromatics; (b) aliphatics.

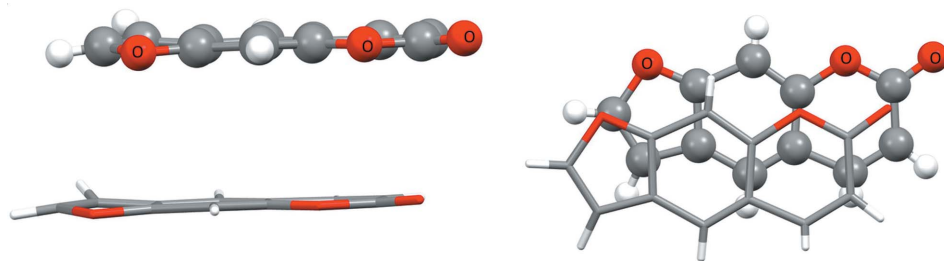


Figure 5
Example of furo[3,2-g]coumarin, a CHO aromatic molecule with heterocycles displaying $\pi \cdots \pi$ stacking. The two figures showing a dimer are rotated by 90° around the horizontal direction. The two interacting molecules are related by translation along the c axis in the triclinic $P1$ space group.

with one strong $C-H \cdots O=C$ hydrogen bond.

Globally the scatterplots for CHO and CHN molecules (Figs. 2 and 4) show many similarities if nitrogen and oxygen are considered as equivalent, both chemical types being electronegative and prone to make polar contacts such as hydrogen bonds. When the contacts in CHO aromatics are analyzed as a function of S_O (Fig. 2), the proportion of oxygen on the molecular surface varies over a narrow range (21–32%). The different enrichments show generally stable values, with no clear increasing or declining trends, except for $H \cdots H$ interactions which decrease with S_O . The equivalent graph (S_O , E) for CHO aliphatics is redundant: as C atoms are scarcely present at the molecular surface, the proportions S_H and S_O are nearly complementary ($S_H + S_O \simeq 1$).

In CHO aliphatics (Fig. 4b), some molecules have a high proportion of H atoms on their surface. The most striking trend appearing with the S_H increase is the convergence of both $H \cdots O$ and $H \cdots H$ contacts to enrichments close to 100%, the first one by decreasing and the second one by increasing. The main difference with CHO aromatics is that E_{OO} strongly decreases when the proportion of hydrogen on the molecular

surface decreases. When H atoms are abundant, $O \cdots O$ contacts are not favoured as O atoms tend to be involved in $O \cdots H$ hydrogen bonding.

The enrichment ratios of $H \cdots O$ contacts are displayed as a function of R_{HO} in Fig. 6. These contacts are generally enriched, but there are a few exceptions with E_{HO} slightly smaller than 1. The graph highlights the fact that for CHO aliphatics, the $H \cdots O$ contacts have an increased likelihood to form when the molecules are rich in both oxygen and hydrogen (large R_{HO} values).

The R_{HO} values are restricted to a small range of values in the sample of studied CHO aromatic molecules, but the E_{HO} enrichments are comparatively higher than for aliphatic molecules.

In both subsets of CHO compounds, the $O \cdots O$ contacts are globally impoverished, but their likelihood to occur increases strongly as oxygen becomes more abundant at the molecular surface (Fig. S3 shows E_{OO} as a function of S_O). The E_{OO} ratios are very small, close or equal to zero, in CHO aliphatic molecules when the surface contact is poor in oxygen ($S_O < 20\%$). On the other hand, $O \cdots O$ contacts can have E_{OO} ratios close or even larger than one for some molecules richer in oxygen. Aromatic molecules show a narrow range of S_O values, but the corresponding E_{OO} ratios are comparable to those of aliphatic molecules with a similar oxygen content on the surface. Oxygen–oxygen contacts in crystals have been discussed from an energetic point of view (Gavezzotti, 2010); they generally do not occur in a standalone way but are secondary to stronger interactions.

3.6. CHS compounds

The set of CHS aromatic compounds selected contain 10–32% carbon and 3–66% sulfur on the molecular surface. The $H \cdots S$ contacts have generally stable E_S values in the [1.0, 1.5] interval, generally above unity as sulfur is a weak hydrogen-bond acceptor (Fig. 7a). The $S \cdots S$ contacts are actually totally absent for several compounds rich in hydrogen or low in sulfur (Fig. S4). For CHS compounds with a lower content of hydrogen on the surface, the $S \cdots S$ contacts can be moderately enriched, as E_S values are between 0.6 and 1.6. On the other hand, $C \cdots S$ contacts are generally under-represented with E_{CS} ratios taking, in general, values between 0 and 1.

On the contrary, both $H \cdots C$ and $H \cdots H$ contacts tend to increase with S_H ; these contacts have enrichment ratios which are generally between 0.5 and 1.4. $C \cdots C$ contacts often show high enrichments between 1.5 and 3.5 for molecules with low hydrogen content, while molecules with a high S_H proportion have a much lower propensity to form $C \cdots C$ interactions, with E_{CC} close to zero for several crystal packings. This is to be related to the higher propensity of C atoms in contact with hydrogen when S_H is high. E_{CC} ratios larger than 2.5 are

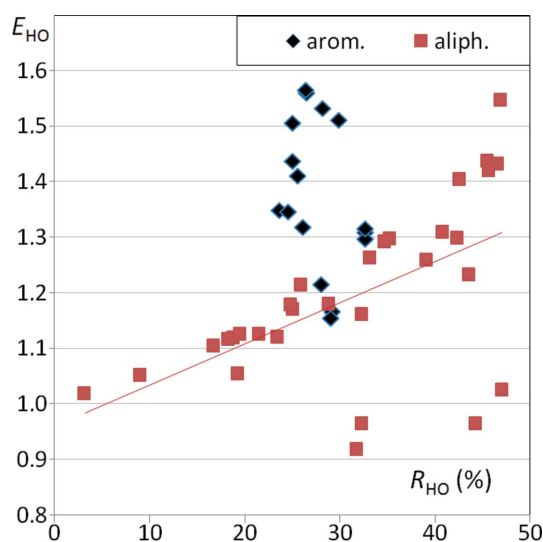


Figure 6
Contact enrichment ratios E_{HO} in crystals of CHO compounds as a function of R_{HO} random contact surface.

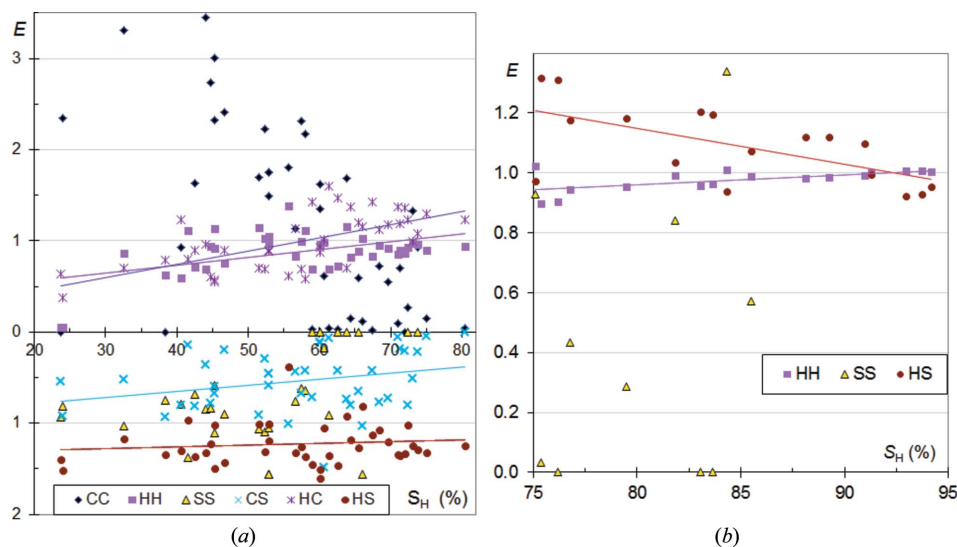


Figure 7
Contact enrichment ratios in crystals of CHS compounds as a function of % of hydrogen on the Hirshfeld surface: (a) aromatics; (b) aliphatics.

related to molecules with carbon content $S_H < 17\%$ (or $R_{HH} < 2.8\%$) on their surface and their significance is limited as issued from ratios with small divisors.

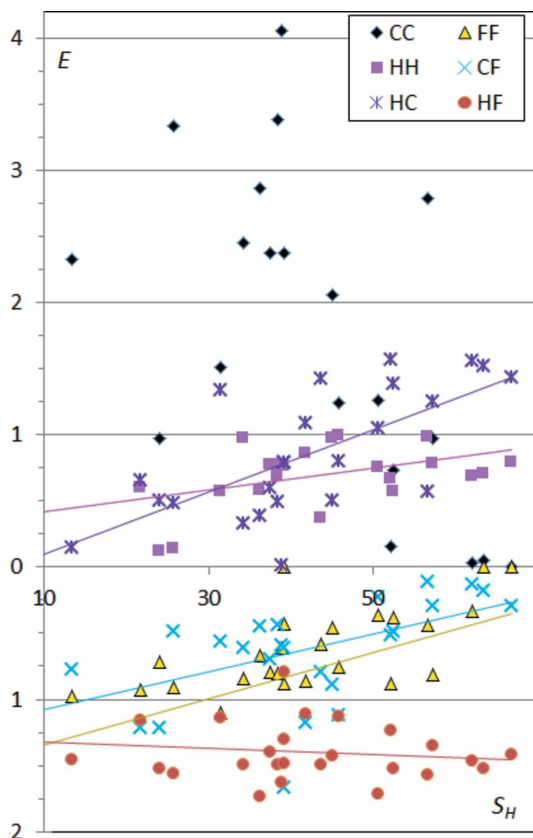


Figure 8
Contact enrichment ratios in crystals of CHF compounds as a function of % of hydrogen on the Hirshfeld surface.

Similarly, $H \cdots H$ interactions tend to be impoverished for compounds poor in hydrogen, or rich in sulfur (Fig. S4). $H \cdots S$ interactions show E_{HS} values around 1.2 for compounds with low sulfur content, *i.e.* which contain mostly H and C; this was also observed for CHN and CH compounds. The graph of contact enrichment as a function of S_S (Fig. S5) shows a strong increase of E_{CC} and E_S with the proportion of sulfur at the molecular surface, while E_{HH} and E_{HC} tend to decrease.

In CHS aliphatic molecules, the proportion of Hirshfeld surface related to C atoms is generally zero or very small, therefore contacts involving C atoms are not shown in Fig. 7(b). For molecules rich in hydrogen on the surface ($S_H > 88\%$

implying $S_S < 12\%$), R_S values are consequently very small and the E_S data are discarded in the graph.

$H \cdots H$ and $H \cdots S$ contacts show the same behaviour as for aromatic CHS compounds and their enrichment ratios tend to unity when the proportion of H atoms on the Hirshfeld surface becomes large. The $S \cdots S$ contacts show a large range of E_S enrichments from 0.0 to 1.3.

3.7. CHF compounds

Very few purely aliphatic CHF-type molecules were found, therefore, the analysis was made on a sample containing molecules with aromatic rings which were more abundant in the CSD. In addition, a few relatively small molecules (up to four C atoms) containing $C=C$ double bonds were also considered, which similarly display carbon on their surface.

Interactions involving organic (C–F) fluorine are generally considered as weak, although fluorine is electronegative and displays an electron accumulation of torus shape around the C–F axis (Bach *et al.*, 2001). The enrichment of the $F \cdots F$ contacts is generally lower than unity (Fig. 8) and is globally decreasing with S_H from unity to values as low as zero. The $C \cdots F$ contacts show a similar trend; the E_{CF} ratios are generally lower than one and are steadily decreasing as the content of hydrogen becomes high. This is to be related to the fact that $C-F \cdots \pi$ interactions occur mostly in perfluorinated aromatic rings which are electron depleted (Reichenbacher *et al.*, 2005; Berger *et al.*, 2011). This last statement is clearly confirmed as E_{FC} ratios increase as compounds become rich in fluorine (Fig. S5).

The $H \cdots F$ contacts appear with high enrichment values stable around 1.5 ± 0.4 , indicating that hydrogen is a preferred partner in contacts made by F atoms, as the two atom types are attractive from an electrostatic point of view.

The role of organic fluorine in crystal engineering and notably the occurrence of $H \cdots F$ and $F \cdots F$ interactions in crystals was recently reviewed. Choprah & Row (2011) and Thalladi *et al.* (1998) studied $C-H \cdots F$ interactions in the crystal structures of several fluorobenzenes. Comparing the crystal packings of similar compounds they found that $C-H \cdots F$ interactions resemble $C-H \cdots N$ and $C-H \cdots O$ interactions and can be considered as weak hydrogen bonds. In molecules with a high fluorine content, $C-H$ groups have increased acidity and the $C-H \cdots F$ interactions become stronger. They also showed that the $C-F$ group prefers to form $C-H \cdots F$ interactions rather than $F \cdots F$ contacts, which is in accordance with the related enrichment ratios found in this study.

The $H \cdots H$ contact enrichment ratios globally increase in CHF compounds with the percentage of H at the molecular surface, but remain generally lower than unity (Fig. 8). E_{HH} can be quite low for CHF compounds low in hydrogen due to competition with $H \cdots F$ contacts.

The $C \cdots C$ contacts are, in many cases, very enriched. A few compounds with high S_H display, on the contrary, an absence or strong under-representation of these contacts and correlatively $H \cdots F$ are over-represented.

The same data displayed as a function of S_F show that the $F \cdots F$ contacts are generally impoverished, but increase and tend to a unitary enrichment for molecules with a large fluorine content. The $C \cdots F$ contacts show a similar increasing trend with S_F . $H \cdots F$ contacts are again stable around $E_{HF} = 1.5$, while $H \cdots H$ contacts have a decreasing likelihood of occurring for fluorine-rich molecules as H atoms are attracted by F atoms.

3.8. $C \cdots C$ contacts

The $C \cdots C$ interactions were further investigated as they show the highest enrichment values. The E_{CC} ratios are

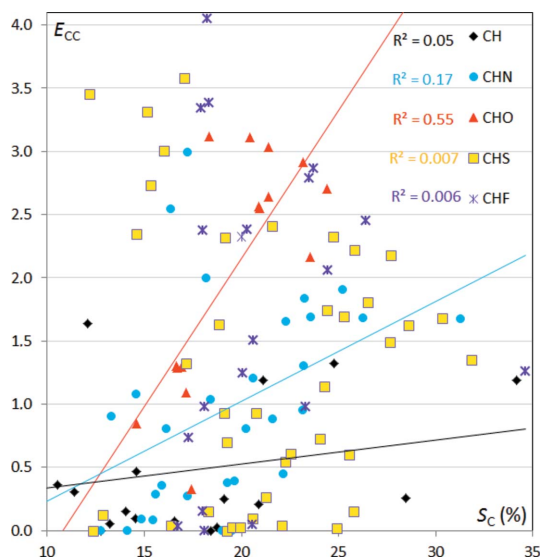


Figure 9
Contact enrichment ratios E_{CC} versus proportion of carbon S_C in the different sets of aromatic compounds.

displayed as a function of S_C in the different sets of aromatic compounds in Fig. 9; they show a wide range of values from 0 to 4. The E_{CC} values increase globally with S_C for CHO compounds and the two variables are 55% correlated. The same trends hold for CHN compounds, but the slope of the linear fit and the correlation coefficient are reduced. For CH, CHS and CHF compounds, the E_{CC} and S_C variables show no correlation. The E_{CC} ratios larger than 3 originate from molecules with less than 21% carbon on the surface, *i.e.* the divisor $R_{CC} < 4.4\%$ is small in the enrichment quotient.

Enrichment values related to different interactions can also be analyzed together. The graph of the enrichment ratios E_{CC} versus E_{HC} (Fig. 10) shows that the two variables are in all cases highly anti-correlated ($|\rho| > 71\%$). This suggests that $C \cdots C$ and $H \cdots C$ interactions are strongly in competition with each other in the crystal contacts. The highest anti-correlations are found for the CHO and CHN compounds, *i.e.* which contain the polar atoms O or N.

4. Conclusions

The contacts between chemical species were analyzed in several sets of crystal structures containing at most three different chemical types. The percentage of the surface of different elements which compose each molecule was derived from the analysis with *CrystalExplorer*. Subsequently, contact enrichment ratios were computed for all pairs of chemical elements. The fraction of enrichment is obtained by dividing the effective surface contacts in the crystal by the respective ‘random contact surfaces’ computed as if all the contact types occurred in an equiprobable way. The enrichment ratio is a novel and powerful tool derived from the Hirshfeld contact

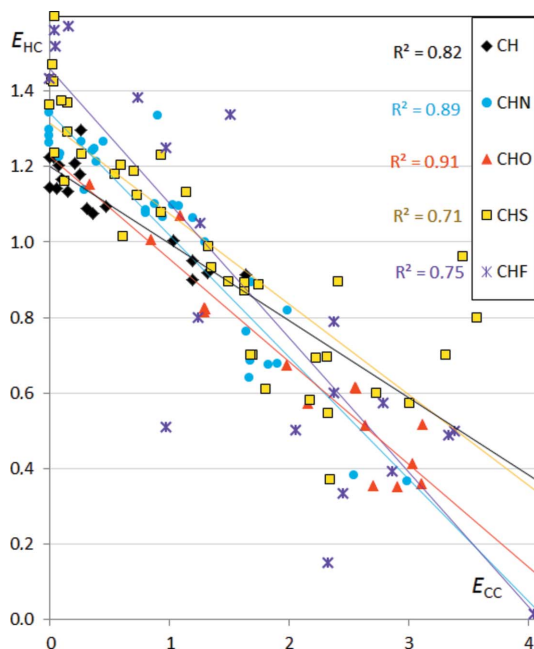


Figure 10
Contact enrichment ratios E_{HC} versus E_{CC} in the different sets of aromatic compounds.

Table 2

 Recapitulation of the clear general tendencies of the enrichment ratios E for the different sets of molecules.

 The variation (increase, decrease, stability) of E as a function of increasing percentage of hydrogen at the molecular surface, S_H , is indicated by arrows.

Set	Aromatics		Aliphatics	
CH	$E_{HH} \sim 1$	→	–	
	$E_{CH} \sim 1.1$	→	–	
CHN	$E_{HH} \sim 0.9$	→	$E_{HH} < 1$	↘
	$E_{NN} < 1$	↘	$E_{NN} < 0.6$	↘
	$E_{HN} > 1$	→	$E_{HN} > 1$	↘
	E_{CN}	↘	E_{CN}	↘
	E_{HC}	↗	E_{HC}	↗
CHO	E_{CC}	↘	–	
	$E_{HH} \sim 0.9$	→	E_{HH}	↗
	$E_{OO} < 0.6$	→	$E_{OO} < 1$	↘
	$E_{HO} > 1$	↘	$E_{HO} > 1$	↘
	E_{CO}	→	–	
CHS	$E_{HC} < 1$	↗	–	
	$E_{HH} \sim 1$	↗	$E_{HH} \sim 1$	→
	$E_{HS} > 1$	→	E_{HS}	↘
	$E_{CS} < 1$	–	–	
	E_{HC}	↗	–	
CHF	E_{CC}	↘	–	
	$E_{HH} < 1$	↗	–	
	$E_{FF} < 1$	↘	–	
	$E_{HF} > 1$	→	–	
	E_{HC}	↗	–	
	E_{CF}	↘	–	

surface analysis to determine which type of contacts are over- or under-represented in crystal packings. A synthesis of the general tendencies found for the different molecular sets is shown in Table 2.

For instance, analysis of CHO- and CHN-type compounds has confirmed that H···O and H···N contacts are generally favoured as they form hydrogen bonds. H···H contacts are often found to have enrichment ratios close to or slightly lower than unity for the different kinds of molecules. However, in CHN and CHO aliphatic compounds H···H contacts are more under-represented in molecules with low hydrogen content, due to competition with hydrogen acceptors. The H···N and H···O contact enrichments tend to increase in aliphatic compounds when S_H becomes low; this is less the case in aromatic compounds where H atoms are also attracted by C atoms.

Aromatic compounds, which have significant amounts of carbon on their surface, can display C···C interactions that can be very enriched when they have heterocycles containing N, O, S or F atoms. Stacking between aromatic units is favoured when electron-rich aromatics interact with electron-deficient aromatics. The C···C contacts are particularly favoured, with E_{CC} ratios reaching values in the [2, 4] interval, in aromatic molecules which have a low or moderate proportion of hydrogen on their surface. However, it has to be remembered that in compounds with low S_C , the occurrence of very large E_{CC} ratios is facilitated by numerical effects. When hydrogen is very abundant at the molecular surface, the higher propensity of C···H interactions to occur lowers as a consequence of C···C contacts.

In CH compounds C···C contacts are also more favoured when S_H is small, but the E_{CC} values are much smaller as π ··· π stacking is less favourable when there is no heterocycle.

The enrichment ratios are a useful tool which can be applied to specific classes of molecules to identify new relationships between their structure, crystal contacts and properties or point out some outliers. The most relevant non-covalent interactions (or synthons) in crystals are π ··· π stacking interactions and hydrogen bonds (Choprah & Row, 2011), which have been extensively studied in the literature and can display high enrichment ratios (H···O, C···C). For instance, correlations between different E_{XY} ratios could be discovered within families of molecules (Fig. 10). Crystal engineering consists of using the understanding of the formation of intermolecular interactions in crystal packings to design new solids with desirable chemical or physical properties (Desiraju, 1989). The formation of the crystal packing is a compromise between all the strong and weak interactions occurring between the neighbouring molecules. Analysis of crystal contact-forming propensities depending on the composition of the molecular surface by chemical elements gives knowledge which can contribute significantly to crystal engineering purposes. Enrichment ratios could also be computed for protein/ligand contact surfaces and yield interesting new information for the understanding of molecular recognition and drug design.

Acknowledgements

KE thanks the University of Lorraine for an invited assistant professor position for one month.

References

- Allen, F. H. (2002). *Acta Cryst.* **B58**, 380–388.
- Bach, A., Lentz, D. & Luger, P. (2001). *J. Phys. Chem. A*, **105**, 7405–7412.
- Berger, R., Resnati, G., Metrangolo, P., Weber, E. & Hulliger, J. (2011). *Chem. Soc. Rev.* **40**, 496–3508.
- Bideau, J.-P., Bravic, G. & Desvergne, J.-P. (1979). *Cryst. Struct. Commun.* pp. 695–698.
- Bibila Mayaya Bisseyou, Y., Bouhmaid, N., Guillot, B., Lecomte, C., Lugan, N., Ghermani, N. & Jelsch, C. (2012). *Acta Cryst.* **B68**, 646–660.
- Choprah, D. & Row, T. N. G. (2011). *CrystEngComm*, **13**, 2175–2186.
- Desiraju, G. R. (1989). *Crystal Engineering: The Design of Organic Solids*. Amsterdam: Elsevier.
- Desiraju, G. R. & Gavezzotti, A. (1989). *Acta Cryst.* **B45**, 473–482.
- Gavezzotti, A. (2010). *Acta Cryst.* **B66**, 396–406.
- Hunter, C. A. (1994). *Chem. Soc. Rev.* **23**, 101–109.
- Hunter, C. A., Lawson, K. R., Perkins, J. & Urch, C. J. (2001). *J. Chem. Soc. Perkin Trans. 2*, pp. 651–669.
- Martinez, C. R. & Iverson, B. L. (2012). *Chem. Sci.* **3**, 2191–2201.
- Matta, C. F., Hernández-Trujillo, J., Tang, T.-H. & Bader, R. F. W. (2003). *Chem. Eur. J.* **9**, 1940–1951.
- McKinnon, J. J., Jayatilaka, D. & Spackman, M. A. (2007). *Chem. Commun.* pp. 3814–3816.
- McKinnon, J. J., Spackman, M. A. & Mitchell, A. S. (2004). *Acta Cryst.* **B60**, 627–668.

- Reichenbacher, K., Süss, H. I. & Hulliger, J. (2005). *Chem. Soc. Rev.* **34**, 22–30.
- Salonen, L. M., Ellermann, M. & Diederich, F. (2011). *Angew. Chem. Int. Ed.* **50**, 4808–4842.
- Spackman, M. A. & Byrom, P. G. (1997). *Chem. Phys. Lett.* **267**, 215–220.
- Spackman, M. A., Joshua, J. & McKinnon, J. J. (2002). *CrystEngComm*, **4**, 378–392.
- Spackman, M. A., McKinnon, J. J. & Jayatilaka, D. (2008). *CrystEngComm*, **10**, 377–388.
- Thalladi, V. R., Weiss, H. C., Bläser, D., Boese, R., Nangia, A. & Desiraju, G. R. (1998). *J. Am. Chem. Soc.* **120**, 8702–8710.
- Turner, M. J., McKinnon, J. J., Jayatilaka, D. & Spackman, M. A. (2011). *CrystEngComm*, **13**, 1804–1813.
- Wolff, S. K., Grimwood, D. J., McKinnon, J. J., Turner, M. J., Jayatilaka, D. & Spackman, M. A. (2012). *CrystalExplorer*, Version 3.0. University of Western Australia.
- Wood, P. A., McKinnon, J. J., Parsons, S., Pidcock, E. & Spackman, M. A. (2008). *CrystEngComm*, **10**, 368–376.