computer programs

CrystalExplorer: a program for Hirshfeld surface analysis, visualization and quantitative analysis of molecular crystals

Peter R. Spackman,[a]* **Michael J. Turner,**[b] **Joshua J. McKinnon,**[b] **Stephen K. Wolff,**[b] **Daniel J. Grimwood,**[b] **Dylan Jayatilaka**[b] **and Mark A. Spackman**[b]

[a]School of Molecular and Life Sciences, Curtin University, PO Box U1987, Perth, WA 6845, Australia, and [b]School of Molecular Sciences, University of Western Australia, 35 Stirling Highway, Perth, WA 6009, Australia. *Correspondence e-mail: peter.spackman@curtin.edu.au

CrystalExplorer is a native cross-platform program supported on Windows, MacOS and Linux with the primary function of visualization and investigation of molecular crystal structures, especially through the decorated Hirshfeld surface and its corresponding two-dimensional fingerprint, and through the visualization of void spaces in the crystal via isosurfaces of the promolecule electron density. Over the past decade, significant changes and enhancements have been incorporated into the program, such as the capacity to accurately and quickly calculate and visualize quantitative intermolecular interactions and, perhaps most importantly, the ability to interface with the Gaussian and NWChem programs to calculate quantum-mechanical properties of molecules. The current version, CrystalExplorer21, incorporates these and other changes, and the software can be downloaded and used free of charge for academic research.

1. Introduction and background

Rational design of functional materials depends on our capacity to understand the driving forces of binding in the material, along with the mechanism which dictates the desired properties of the material. In the context of molecular crystals, this problem effectively reduces to understanding the strength and origin of intermolecular interactions and their relationship with crystal packing.

Hirshfeld surface (HS) analysis was initiated with the serendipitous discovery of the HS (Spackman & Byrom, 1997; McKinnon et al., 1998) as a novel means of partitioning space in molecular crystals into smooth non-overlapping molecular entities. Subsequent developments led to the invention of HS fingerprint plots (Spackman & McKinnon, 2002; McKinnon et al., 2004), or 'fingerprints' for short. These are a convenient means of summarizing the intermolecular contacts present in crystals via a two-dimensional plot, with the decomposition of this fingerprint plot into features in order to identify particular interactions (McKinnon, Jayatilaka & Spackman, 2007). The mapping of scalar properties via a colour scale onto the HS (McKinnon, Fabbiani & Spackman, 2007), known as decoration, has also proven to be a powerful approach to quickly and easily gain insight into molecular environments in the crystalline state.

1.1. Prior versions of CrystalExplorer

The CrystalExplorer software was initially developed to facilitate HS analysis (Spackman & Jayatilaka, 2009), and that
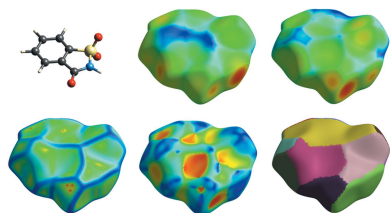
is still a primary purpose of the program which is currently widely used. However, the ability to perform quantum-chemical calculations, and subsequently generate isosurfaces and surface properties of quantum-mechanically derived quantities (Spackman *et al.*, 2008), as well as the capacity to identify and visualize void spaces (Turner *et al.*, 2011) in crystals, has contributed to the popularity of the tool over the past 15 years. These capabilities, made possible through an intuitive graphical user interface available on most desktop computers, have made *CrystalExplorer* into a valuable toolbox for advancing our comprehension and understanding of the structure and properties of molecular crystals.

1.2. Changes in *CrystalExplorer21*

Besides several important bug fixes, the latest version of *CrystalExplorer* now supports use of the open-source *NWChem* (Aprà *et al.*, 2020) package as a quantum-mechanical back end (through the use of a Molden input file), in addition to the existing support for the *Gaussian* (Frisch *et al.*, 2016) quantum-chemistry and *Tonto* (Jayatilaka & Grimwood, 2003) quantum-crystallography packages. Further improvements have been made in finer control over precise views down desired crystal directions, precision in energy tables and more.

2. Capabilities

The aim of this article is to provide a short overview of the capabilities of *CrystalExplorer*; the reader is referred to documentation available online at https://wiki.crystalexplorer.net for further details. *CrystalExplorer* reads a CIF (Hall *et al.*, 1991) and displays the molecules as ball-and-stick, van der Waals sphere or wireframe models that can be rotated or zoomed into using the mouse wheel.

2.1. Crystallographic tools and visualization

CrystalExplorer offers a number of options for building clusters of molecules, allowing the user to see as much or as little of the crystal structure as desired. These include (i) generating molecules close to those on the screen by clicking on nearest-neighbour 'ghost' atoms, (ii) generating the remaining atoms of incomplete molecules, (iii) generating all the atoms within a radius of a selection and (iv) generating 'slabs' of complete unit cells. The program further includes a selection of tools that are typically found in crystallographic packages, allowing, for example, the calculation of geometric parameters (distances, angles *etc.*), and identification and graphical depiction of short contacts and hydrogen bonds.

2.2. Hirshfeld and other surfaces

HSs, along with what has come to be called HS analysis, have been described in detail in

several publications (Spackman & Jayatilaka, 2009; McKinnon *et al.*, 1998, 2004; Spackman, 2013; Spackman & Byrom, 1997). They arose from an exploration of possible methods to define a region of the crystal as surrounding a specific molecule, for the purpose of integrating the electron density 'belonging' to that molecule.

In *CrystalExplorer*, the electron densities used to compute the HS surfaces come from tabulations of atomic wavefunctions expanded using exponential-type basis functions (Clementi & Roetti, 1974; Koga *et al.*, 1993, 2000). Inspired by Fred Hirshfeld's 'stockholder' partitioning of atoms in molecules (Hirshfeld, 1977), the resulting (spherical) atomic electron densities $\rho_A(\mathbf{r})$ are used to construct the promolecule electron density $\rho_{\mathrm{promol}}(\mathbf{r})$ and the procrystal electron density $\rho_{\mathrm{procryst}}(\mathbf{r})$, which are then used to define the weight function,

$$w_A(\mathbf{r}) = \frac{\rho_{\mathrm{promol}}(\mathbf{r})}{\rho_{\mathrm{procryst}}(\mathbf{r})} = \frac{\sum_{A \in \mathrm{molecule}} \rho_A(\mathbf{r})}{\sum_{C \in \mathrm{crystal}} \rho_C(\mathbf{r})}. \qquad (1)$$

This continuous scalar function has $0 < w_A(\mathbf{r}) < 1$ in all space, and the 0.5 isosurface of this function is the HS.

Surfaces of interest can be calculated in *CrystalExplorer* by selecting the molecule in question with the mouse and clicking on the HS icon. The user is subsequently prompted to choose the desired surface (and an optional property to decorate the surface) along with a desired surface resolution before the surface is calculated and presented in the three-dimensional graphical display. Fig. 1 shows the available molecular surfaces using as an example acetic acid [Cambridge Structural Database (CSD) refcode ACETAC01; Nahringbauer, 1970].

Molecular surfaces alone provide significant visual information when trying to understand crystal packing, but the addition of scalar 'surface properties' mapped to colours decorating the surface further aids in intuitive understanding.
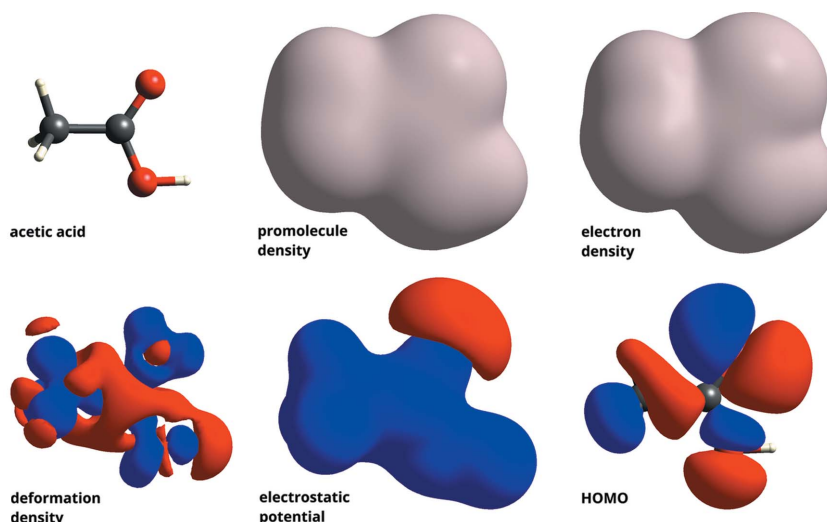

Figure 1
An ensemble of molecular surfaces available in *CrystalExplorer*, displayed using acetic acid (CSD refcode ACETAC01) as an example structure. HOMO refers to highest occupied molecular orbital. All density plots are calculated at the isovalue of 0.002 a.u., while the electrostatic potential surface is calculated at 0.05 a.u. A B3LYP/6-31G(*d,p*) wavefunction was used.

computer programs

Property	Description
d_i	Distance from HS to the nearest atom I internal to the surface
d_e	Distance from HS to the nearest atom E external to the surface
d_{norm}	Normalized sum of d_e and d_i, *i.e.* $(d_i - r_I^{vdw})/r_I^{vdw} + (d_e - r_E^{vdw})/r_E^{vdw}$
ρ_{promol}	Promolecule electron density
V_{elec}	Electrostatic potential
ρ	Electron density
$|\phi|^2$	Probability density of a given molecular orbital
ρ_{def}	Deformation density, *i.e.* $\rho - \rho_{promol}$
Fragment patch	Unique (coloured) region based on atoms external to the HS designed to indicate the nearest-neighbouring molecule
Curvedness	Function depending on the HS concavity or convexity
Shape index	Function depending on the HS flatness or curvature

Fig. 2 shows an example of an HS for saccharin, decorated with a variety of properties available in *CrystalExplorer*. The properties themselves are explained in Table 1. Only the properties on the last row of Fig. 2 require molecular quantum-mechanical wavefunctions for their evaluation; the other surfaces just make use of geometric distances to the HS, which itself is obtained using only the aforementioned tabulated atomic wavefunctions. Note how qualitatively similar the atom-generated promolecule-density isosurface is to the 'molecular' quantum-mechanical electron-density isosurface (Mitchell & Spackman, 2000).

2.3. Fingerprint plots

CrystalExplorer can interactively generate fingerprint plots for all HSs calculated at a sufficiently high resolution. *CrystalExplorer* 'links' the display of the fingerprint to its HS in that a mouse click on the fingerprint plot highlights the corresponding points on the HS (with red cones). Fingerprint plots can be exported in the PNG and PostScript formats, the latter being preferred for inclusion in some publications as it is an editable vector format unlike a rasterized PNG.

CrystalExplorer also provides options for decomposing a fingerprint plot into contributions from particular kinds of interactions. At present the user is limited to highlighting interactions on the basis of atom type (C, N, O *etc.*). Interactions are selected by choosing the atom type for the atoms inside the HS and for those outside it. The interactions with the inside and outside atoms switched can also be included by selecting the 'include reciprocal contacts' option.

An example of fingerprint decomposition for urea (CSD refcode UREAXX01; Sklar, Senko & Post, 1961) can be seen in Fig. 3. The contacts between oxygen atoms inside the HS and hydrogen atoms (*i.e.* O···H contacts) outside of it have been selected, including reciprocal contacts (*i.e.* H···O contacts). The parts of the fingerprint that correspond to these interactions remain coloured while the rest of the plot is greyed out, and simultaneously the displayed HS colouring is greyed out with the exception of the regions involving the interaction. The total area of these regions relative to the total surface area of the HS is used to report a percentage contribution of the interaction. We can see from the figure that O···H and H···O interactions make up 36.6% of the surface
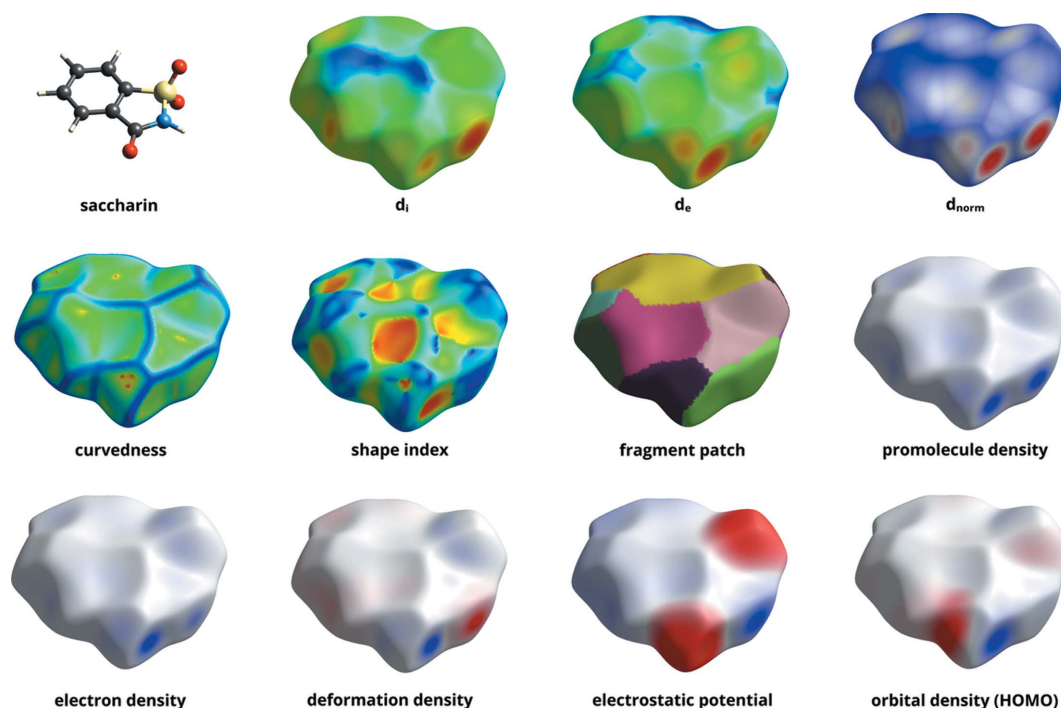

Figure 2
An ensemble of HSs for the saccharin molecule (CSD refcode SCCHRN02; Wardell, Low & Glidewell, 2005) decorated with different properties available in *CrystalExplorer*. Quantum-mechanical properties were calculated using B3LYP/6-31G(d,p) wavefunctions.

area of the HS. There are two fragment patches corresponding to these kinds of interactions and both of them are highlighted by the red-coloured d_{norm} property.

2.4. *CrystalExplorer* model energies

CrystalExplorer computes pairwise intermolecular interaction energies using an approach inspired by the PIXEL (Gavezzotti, 2005) method, but the approach used in *CrystalExplorer* differs in several important points. Full details of the *CrystalExplorer* models are available in the literature (Turner *et al.*, 2014; Mackenzie *et al.*, 2017) and only a brief outline is provided here. The basic expression follows that used by Gavezzotti and others, whereby the intermolecular interaction energy is decomposed into four physically motivated terms (explained below) but with the exception that each term is multiplied by its own scale factor:

$$E_{tot} = k_{ele}E_{ele} + k_{pol}E_{pol}$$
$$+ k_{dis}E_{dis} + k_{rep}E_{rep}. \quad (2)$$

Here E_{ele} is the classical electrostatic interaction energy between unperturbed quantum-mechanical charge distributions of the monomers, E_{pol} is the polarization energy calculated as a sum over isotropic atomic polarizabilities multiplied by the square of the local electric field caused by the other monomer, E_{dis} is Grimme's D2 dispersion correction (Grimme, 2006), and E_{rep} is the exchange-repulsion energy calculated between unperturbed quantum-mechanical charge distributions of the monomers according to the formula for the sum of the exchange and repulsion terms reported by Su & Li (2009) and earlier by Hayes & Stone (1984). The four scale factors in this expression have been optimized for two wavefunction sources: CE-HF (using HF/3-21G monomer wavefunctions) and CE-B3LYP [using B3LYP/6-31G(d,p) monomer wavefunctions]. These scale factors were optimized against reference intermolecular interaction energies: counterpoise-corrected B3LYP-D2/6-31G(d,p) interaction energies for 1794 molecule/ion pairs extracted from 171 experimental crystal structures. The resulting models have a mean absolute deviation of 4.7 kJ mol^{-1} for CE-HF and 2.5 kJ mol^{-1} for CE-B3LYP relative to the reference data across an energy range spanning 3.75 MJ mol^{-1} (Mackenzie *et al.*, 2017).

CrystalExplorer uses crystallographic symmetry to determine which interaction pairs are unique. Those that are identical by symmetry are coloured in the same way so that the user may easily see the symmetry-related pairs and their corresponding interactions, along with the crystallographic symmetry operation relating the two molecules. Fig. 4 shows an example interaction-
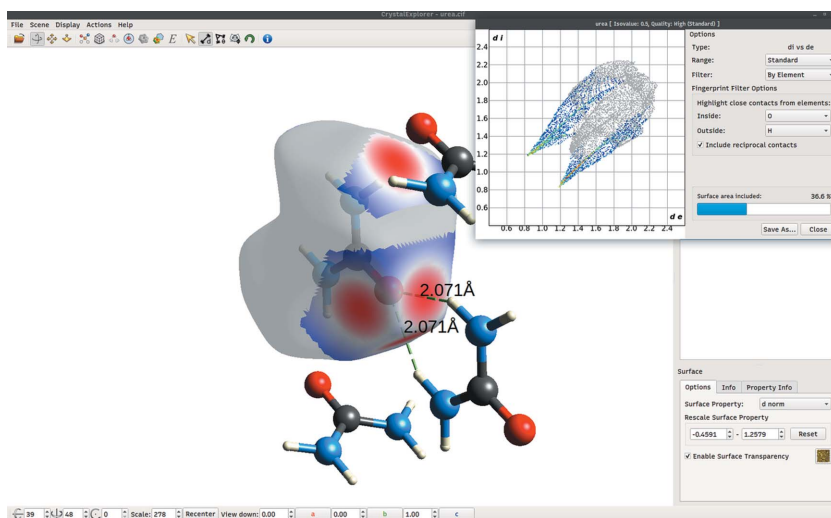

Figure 3
The *CrystalExplorer21* graphical window, showcasing an HS (transparent) surrounding a urea molecule in its crystal environment (CSD refcode UREAXX01), alongside its corresponding fingerprint plot. Neighbouring molecules, corresponding to the highlighted interactions, are also present in the view. O···H and H···O contacts have been highlighted, demonstrating the link between the regions of the HS and the corresponding values on the two-dimensional fingerprint plot. The coloured fragment patch is decorated with the d_{norm} property, and the vectors joining the oxygen atom inside the surface to the two hydrogen atoms outside of it are seen to go through the centre of the red-coloured d_{norm} spot.
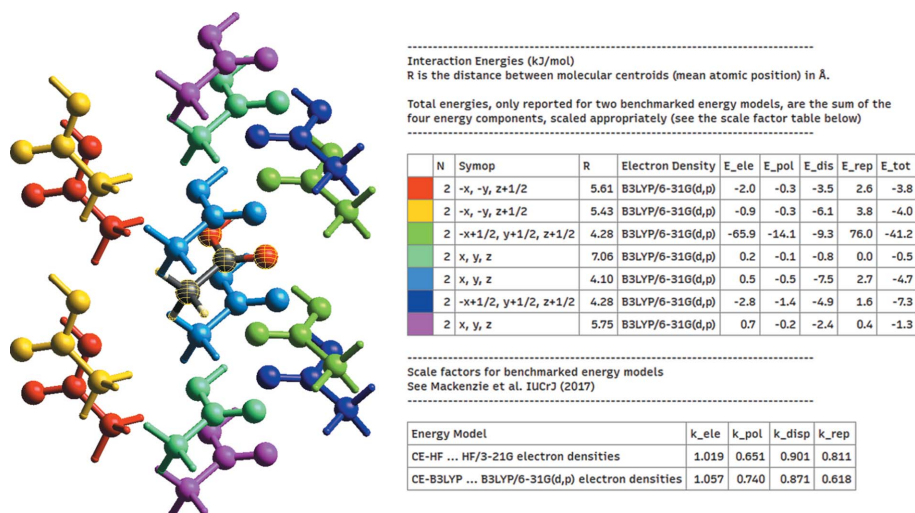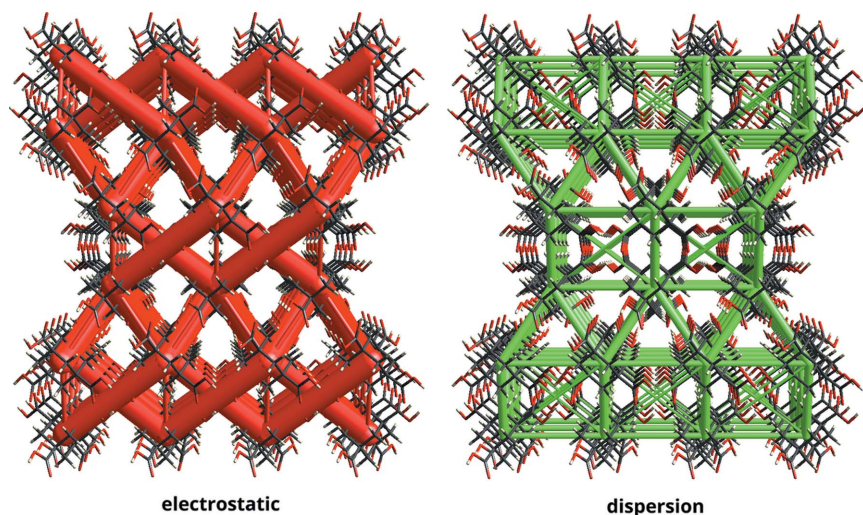

Figure 4
Interaction energies for acetic acid (CSD refcode ACETAC01) calculated with the CE-B3LYP model. It can be seen from the interaction-energies table that the catamer hydrogen-bonding motif between the central molecule (highlighted in yellow mesh) and the $-x + 1/2$, $y + 1/2$, $z + 1/2$ symmetry-related molecule (lime green) is by far the strongest interaction among near neighbours, with an interaction energy of -41.2 kJ mol^{-1}.

computer programs

Figure 5
Energy frameworks for a $3 \times 3 \times 1$ supercell, viewed down the crystallographic *b* axis for the fivefold interpenetrated adamantane tetracarboxylic acid (CSD refcode GEJVEW) crystal structure. Despite being a hydrogen-bonded organic framework structure, as demonstrated through the electrostatic framework, we can see a tight cohesive framework of dispersion interactions which also underpin the structure of the crystal.

energy output along with the corresponding graphical view for an acetic acid (CSD refcode ACETAC01) crystal.

The interaction energies described above provide quantitative insight into the binding (and non-binding) pairwise intermolecular interactions in the crystal but can be difficult to translate into the context of the actual crystal structure. *CrystalExplorer* alleviates this problem by providing a visualization technique to simplify the information, which we call 'energy frameworks' (Turner *et al.*, 2015). Energy frameworks permit us to examine 'cooperative effects' in these intermolecular interactions, across the crystal. Specifically, energy frameworks allowing separate views for the electrostatic (red), dispersion (green) and total energy (blue for binding, gold for non-binding) terms between pairs are available. Fig. 5 shows an energy-framework visualization for adamantane tetra-carboxylic acid (CSD refcode GEJVEW; Ermer, 1988), a molecular crystal comprising interpenetrated networks of hydrogen bonds.

Furthermore, these intermolecular interaction energies may be directly summed over a number of pairs to give a reasonable estimate of the lattice energy of a crystal (Thomas *et al.*, 2018; Spackman, 2018); the CE-B3LYP model energies in particular yield results comparable in accuracy to dispersion-corrected periodic density-functional-theory calculations.

3. Availability and installation

CrystalExplorer installers can be downloaded at https://crystalexplorer.net and used free of charge for academic purposes. Commercial licences are also available. *CrystalExplorer* is supported on Windows 10, MacOS and several GNU/Linux variants.

References

Aprà, E., Bylaska, E. J., de Jong, W. A., Govind, N., Kowalski, K., Straatsma, T. P., Valiev, M., van Dam, H. J. J., Alexeev, Y., Anchell, J., Anisimov, V., Aquino, F. W., Atta-Fynn, R., Autschbach, J., Bauman, N. P., Becca, J. C., Bernholdt, D. E., Bhaskaran-Nair, K., Bogatko, S., Borowski, P., Boschen, J., Brabec, J., Bruner, A., Cauët, E., Chen, Y., Chuev, G. N., Cramer, C. J., Daily, J., Deegan, M. J. O., Dunning, T. H. Jr, Dupuis, M., Dyall, K. G., Fann, G. I., Fischer, S. A., Fonari, A., Früchtl, H., Gagliardi, L., Garza, J., Gawande, N., Ghosh, S., Glaesemann, K., Götz, A. W., Hammond, J., Helms, V., Hermes, E. D., Hirao, K., Hirata, S., Jacquelin, M., Jensen, L., Johnson, B. G., Jónsson, H., Kendall, R. A., Klemm, M., Kobayashi, R., Konkov, V., Krishnamoorthy, S., Krishnan, M., Lin, Z., Lins, R. D., Littlefield, R. J., Logsdail, A. J., Lopata, K., Ma, W., Marenich, A. V., Martin del Campo, J., Mejia-Rodriguez, D., Moore, J. E., Mullin, J. M., Nakajima, T., Nascimento, D. R., Nichols, J. A., Nichols, P. J., Nieplocha, J., Otero-de-la-Roza, A., Palmer, B., Panyala, A., Pirojsirikul, T., Peng, B., Peverati, R., Pittner, J., Pollack, L., Richard, R. M., Sadayappan, P., Schatz, G. C., Shelton, W. A., Silverstein, D. W., Smith, D. M. A., Soares, T. A., Song, D., Swart, M., Taylor, H. L., Thomas, G. S., Tipparaju, V., Truhlar, D. G., Tsemekhman, K., Van Voorhis, T., Vázquez-Mayagoitia, Á, Verma, P., Villa, O., Vishnu, A., Vogiatzis, K. D., Wang, D., Weare, J. H., Williamson, M. J., Windus, T. L., Woliński, K., Wong, A. T., Wu, Q., Yang, C., Yu, Q., Zacharias, M., Zhang, Z., Zhao, Y. & Harrison, R. J. (2020). *J. Chem. Phys.* **152**, 184102.
Clementi, E. & Roetti, C. (1974). *At. Data Nucl. Data Tables*, **14**, 177–478.
Ermer, O. (1988). *J. Am. Chem. Soc.* **110**, 3747–3754.
Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A. V., Bloino, J., Janesko, B. G., Gomperts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., Izmaylov, A. F., Sonnenberg, J. L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V. G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery, J. A. Jr, Peralta, J. E., Ogliaro, F., Bearpark, M. J., Heyd, J. J., Brothers, E. N., Kudin, K. N., Staroverov, V. N., Keith, T. A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A. P., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Millam, J. M., Klene, M., Adamo, C., Cammi, R., Ochterski, J. W., Martin, R. L., Morokuma, K., Farkas, O., Foresman, J. B. & Fox, D. J. (2016). *Gaussian 16. Revision C.01.* Gaussian Inc., Wallingford, Connecticut, USA.
Gavezzotti, A. (2005). *Z. Kristallogr. Cryst. Mater.* **220**, 499–510.
Grimme, S. (2006). *J. Comput. Chem.* **27**, 1787–1799.
Hall, S. R., Allen, F. H. & Brown, I. D. (1991). *Acta Cryst.* A**47**, 655–685.
Hayes, I. & Stone, A. (1984). *Mol. Phys.* **53**, 83–105.
Hirshfeld, F. L. (1977). *Theor. Chim. Acta*, **44**, 129–138.
Jayatilaka, D. & Grimwood, D. J. (2003). *Computational Science – ICCS 2003*, Lecture Notes in Computer Science, Vol. 2660, edited

by P. M. A. Sloot, D. Abramson, A. V. Bogdanov, Y. E. Gorbachev, J. J. Dongarra & A. Y. Zomaya, pp. 142–151. Springer-Verlag Berlin Heidelberg.

Koenderink, J. J. (1990). *Solid Shape*. Cambridge: MIT Press

Koenderink, J. J. & van Doorn, A. J. (1992). *Image Vis. Comput.* **10**, 557–564.

Koga, T., Kanayama, K., Watanabe, T., Imai, T. & Thakkar, A. J. (2000). *Theor. Chem. Acc.* **104**, 411–413.

Koga, T., Tatewaki, H. & Thakkar, A. J. (1993). *Phys. Rev. A*, **47**, 4510–4512.

Mackenzie, C. F., Spackman, P. R., Jayatilaka, D. & Spackman, M. A. (2017). *IUCrJ*, **4**, 575–587.

McKinnon, J. J., Fabbiani, F. P. A. & Spackman, M. A. (2007). *Cryst. Growth Des.* **7**, 755–769.

McKinnon, J. J., Jayatilaka, D. & Spackman, M. A. (2007). *Chem. Commun.* pp. 3814.

McKinnon, J. J., Mitchell, A. S. & Spackman, M. A. (1998). *Chem. Eur. J.* **4**, 2136–2141.

McKinnon, J. J., Spackman, M. A. & Mitchell, A. S. (2004). *Acta Cryst.* B**60**, 627–668.

Mitchell, A. S. & Spackman, M. (2000). *J. Comput. Chem.* **21**, 933–942.

Nahringbauer, I. (1970). *Acta Chem. Scand.* **24**, 453–462.

Sklar, N., Senko, M. E. & Post, B. (1961). *Acta Cryst.* **14**, 716–720.

Spackman, M. A. (2013). *Phys. Scr.* **87**, 048103.

Spackman, M. A. (2018). *CrystEngComm*, **20**, 5340–5347.

Spackman, M. A. & Byrom, P. G. (1997). *Chem. Phys. Lett.* **267**, 215–220.

Spackman, M. A. & Jayatilaka, D. (2009). *CrystEngComm*, **11**, 19–32.

Spackman, M. A. & McKinnon, J. J. (2002). *CrystEngComm*, **4**, 378–392.

Spackman, M. A., McKinnon, J. J. & Jayatilaka, D. (2008). *CrystEngComm*, **10**, 377–388.

Su, P. & Li, H. (2009). *J. Chem. Phys.* **131**, 014102.

Thomas, S. P., Spackman, P. R., Jayatilaka, D. & Spackman, M. A. (2018). *J. Chem. Theory Comput.* **14**, 1614–1623.

Turner, M. J., Grabowsky, S., Jayatilaka, D. & Spackman, M. A. (2014). *J. Phys. Chem. Lett.* **5**, 4249–4255.

Turner, M. J., McKinnon, J. J., Jayatilaka, D. & Spackman, M. A. (2011). *CrystEngComm*, **13**, 1804–1813.

Turner, M. J., Thomas, S. P., Shi, M. W., Jayatilaka, D. & Spackman, M. A. (2015). *Chem. Commun.* **51**, 3735–3738.

Wardell, J. L., Low, J. N. & Glidewell, C. (2005). *Acta Cryst.* E**61**, o1944–o1946.