Towards improved accuracy of Hirshfeld atom refinement with an alternative electron density partition

Michał Chodkiewicz* and Krzysztof Woźniak*

Biological and Chemical Research Centre, Department of Chemistry, University of Warsaw, Żwirki i Wigury 101, Warszawa 02-089, Poland. *Correspondence e-mail: ml.chodkiewicz@uw.edu.pl, kwozniak@chem.uw.edu.pl

Hirshfeld atom refinement (HAR) is generally the chosen method for obtaining accurate hydrogen atom parameters from X-ray diffraction data. Still, determination can prove challenging, especially in the case of atomic displacement parameters (ADPs). We demonstrate that such a situation can occur when the ADP values of the bonding partner of the hydrogen atom are not determined accurately. Atomic electron densities partially overlap and inaccuracies in the bonding neighbour ADPs can be partially compensated for with modifications to the hydrogen ADPs. We introduce a modified version of the original Hirshfeld partition: the exponential Hirshfeld partition, parameterized with an adjustable parameter (n) to allow control of the overlap level of the atomic electron densities which, for $n = 1$, is equivalent to the Hirshfeld partition. The accuracy of the HAR-like procedure using the new partition (expHAR) was tested on a set of organic structures using B3LYP and MP2 electron densities. Applying expHAR improved the hydrogen atom parameters in the majority of the structures (compared with HAR), especially in cases with the highest deviations from the reference neutron values. $X-H$ bond lengths and hydrogen ADPs improved for 9/10 of the structures for B3LYP-based refinement and 8/9 for MP2-based refinement when the ADPs were compared with a newly introduced scale-independent similarity measure.

1. Introduction

Hirshfeld atom refinement (HAR) is a method used for crystal structure determination from X-ray diffraction data (Jayatilaka & Dittrich, 2008; Capelli *et al.*, 2014). It uses atomic form factors based on atomic electron densities from quantum mechanical calculations for the system of interest. The Hirshfeld atom partitioning (Hirshfeld, 1977) of the electron density is applied to split the density into atomic contributions. HAR enables the determination of structural parameters much more accurately than the commonly used independent atom model (IAM), which is based on spherically symmetric atomic densities. The improvement is especially notable in the case of hydrogen atom parameters (Woińska *et al.*, 2016). Since X-rays are mostly scattered by electron density and hydrogen atoms have only a single electron, their structural parameters are relatively difficult to measure accurately as they are sensitive to small modifications of the model of electron density. IAM neglects the effects related to chemical bonding and interatomic interactions and provides quite poor bond lengths for hydrogen atoms and poor descriptions of hydrogen atomic displacement parameters (ADPs). The inclusion of information concerning the asphericity of atomic electron density usually greatly improves these descriptions. This was first demonstrated using a method based on multi-

pole expansion of the atomic densities in terms of the Hansen–Coppens model (Hansen & Coppens, 1978). It turned out that parameters of the atomic densities can be transferred between atoms of the same type in different structures (Brock *et al.*, 1991; Pichon-Pesme *et al.*, 1995; Koritsanszky *et al.*, 2002). This led to the development and application of databases of transferable multipole model parameters: ELMAM2 (Zarychta *et al.*, 2007; Domagała *et al.*, 2012; Nassour *et al.*, 2017), Invariom (Dittrich *et al.*, 2004, 2005, 2006, 2013), UBDB (Volkov *et al.*, 2007; Dominiak *et al.*, 2007; Jarzembska & Dominiak, 2012; Kumar *et al.*, 2019) and its successor MATTS (Jha *et al.*, 2022). The corresponding model is called the transferable aspherical atom model (TAAM). TAAM-based hydrogen atom parameters are much more accurate than those obtained using IAM (*e.g.* Bąk *et al.*, 2011; Jha *et al.*, 2020).
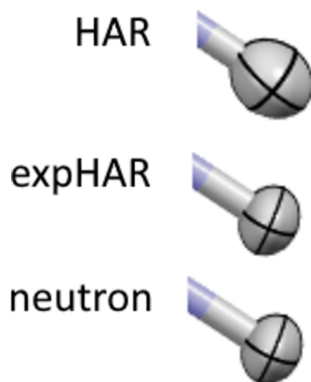
Further improvements were achieved with HAR, which relies on the calculation of the electron density for the system of interest. This is derived from the fact that HAR does not rely on transferability, can take into account the effects of intermolecular interactions and is usually based on a much more flexible model of the total electron density. Intermediate to HAR and TAAM are methods that use the Hirshfeld electron density partition and, like TAAM, utilize the idea of transferability: (1) HAR-ELMO (Malaspina *et al.*, 2019) couples HAR with libraries of extremely localized orbitals (Meyer & Genoni, 2018), (2) HAR employing fragmentation techniques in quantum mechanical calculations (Bergmann *et al.*, 2020; Chodkiewicz *et al.*, 2022a) and (3) a database of transferable atomic densities obtained with the Hirshfeld partition (Koritsanszky *et al.*, 2011; Chodkiewicz *et al.*, 2024). Though all of the above show accuracies similar to HAR, they are, in principle, less accurate than HAR due to applied approximations, but are also less time consuming.

Some systems are still challenging for HAR [*e.g.* those including hydrogen atoms bonded to heavy-element atoms (Woińska *et al.*, 2021, 2023)]. It was possible to obtain very good agreement between HAR- and neutron-measurement-derived hydrogen atom parameters for a number of organic molecules, for example, in urea (Chodkiewicz *et al.*, 2020; Ruth *et al.*, 2022) the N—H bond length could deviate from the neutron measurements by as little as 5 mÅ, although in certain cases the agreement was much worse, *e.g.* 40 mÅ for the O—H bonds in *N*-acetyl-L-4-hydroxyproline monohydrate (Chodkiewicz *et al.*, 2024), yet still much better than IAM. It seems that the parameters of carbon-bonded hydrogen atoms are more likely to be accurately derived with HAR than those of oxygen- or nitrogen-bonded hydrogen atoms [figure 2 in Woińska *et al.* (2016)]. Hydrogen atoms in COOH and OH groups appear to be prone to elongation of their thermal ellipsoids along bonding directions even if such a feature was not present in the neutron measurement structure.

Therefore, while HAR has significantly improved the structural description of hydrogen atoms compared with the spherical model, it is still not a reliable replacement for neutron measurements, even in the case of systems that contain only light elements.

The possibility of improving HAR results with the help of techniques that allow for the estimation of hydrogen atom ADPs, such as SHADE3 (Madsen & Hoser, 2014), TLS+INV (Lübben *et al.*, 2015) and Normal Mode Refinement (NoMoRe) (Hoser & Madsen, 2016, 2017) has been tested. Including estimated hydrogen ADPs has additionally been shown to be necessary (Dittrich *et al.*, 2017) in cases where the hydrogen atom ADPs become non-positive definite in free refinement. ADPs obtained with such methods appear to be more accurate than those obtained directly with HAR (Woińska *et al.*, 2024; Wanat *et al.*, 2021), yet SHADE3 estimates were not always accurate (Malaspina *et al.*, 2020), and NoMoRe was relatively expensive computationally and not applicable to disordered structures (in the current form). Fixing hydrogen ADPs to the value obtained with such methods did not improve X—H bond lengths (Woińska *et al.*, 2024; Malaspina *et al.*, 2020).

In this work, we focus on exploring the possibility of improving HAR. There are various ways refinement can be performed, including choice of quantum chemistry method, basis set and representation of intermolecular interactions [*e.g.* with multipoles mimicking the crystal field (Jayatilaka & Dittrich, 2008)] or using the periodic wavefunction (Wall, 2016; Ruth *et al.*, 2022). Finally, HAR can be generalized by the application of electron density partitions other than the Hirshfeld partition (Chodkiewicz *et al.*, 2020). All these aspects of HAR have been at least partially explored. It was observed that changes in various settings lead to systematic changes in the refined structure, yet no clear path to improve results has been found.

For example, application of the Hartree–Fock (HF) method usually leads to longer polar bonds to hydrogen than DFT. This phenomenon was studied by Landeros-Rivera *et al.* (2023) who linked the effect with the amount of Hartree–Fock exchange (E_x^{HF}) in the density functional (the larger the fraction of E_x^{HF}, the longer the polar X—H bonds). Though the HF method usually provides larger agreement factors (R factors) than DFT methods, it also seems to supply more accurate bond lengths. Among DFT functionals, those with about 25% admixture of E_x^{HF} gave the lowest R factors. Post-HF methods were also applied to HAR but no clear advantage was observed over less expensive methods (Wieduwilt *et al.*, 2020; Chodkiewicz *et al.*, 2020). In the case of a basis set choice, it is known that cc-pVDZ gives systematically different results than the larger basis set cc-pVTZ (cc-pVDZ gives smaller ADPs). It is however quite difficult to judge which ADPs are best, as the ADPs for heavier atoms derived from neutron and X-ray measurements are sometimes quite different (usually in terms of size rather than directionality) and therefore it is sometimes not clear how to choose reference values for assessment of hydrogen ADP accuracies. What appears to be clear in the case of quantum chemistry methods is that accounting for the effects of intermolecular interactions improves the accuracy of HAR, especially in the case of systems with strong hydrogen bonds. Methods that involve the calculation of periodic wavefunctions appear to be very promising (Ruth *et al.*, 2022) in this respect.

Unlike the choice of quantum chemistry method and basis set, the research does not suggest which partition should give better results in HAR, except for one study that explicitly tested some of the partitions in an HAR-like refinement (Chodkiewicz et al., 2020). However, no clear improvement over the Hirshfeld partition was observed. Results of similar accuracy were obtained with the Hirshfeld partition and iterative stockholder partition, even though the difference in atomic charge on the hydrogen atom reached values as large as 0.4 e (i.e. more than 40% of electrons in the atom).

Although obtaining accurate hydrogen atom parameters with HAR does not require high-resolution data, it is expected that higher resolutions should bring better results. Sometimes this leads to a reduced accuracy [figure 2 in Woińska et al. (2016)], especially for O—H bonds. C—H bond length accuracies appear to be much more stable to increasing data resolution. Furthermore, these parameters depend on the weighting scheme applied (Kleemiss et al., 2021), sometimes quite substantially in our experience.

Multiple factors influence the accuracy of HAR, which complicates the search for the optimal method to perform it. In this study, we try to shed light on some of the so-far unexplained trends and phenomena observed in HAR to make the search for optimal settings for HAR and HAR-derived methods easier. Such settings are also important for approaches relying on the transferability of atomic densities such as TAAM and THAM, since in this case the choice of method used for the generation of the corresponding data-banks of atomic electron densities also influences the accuracy of the methods (Chodkiewicz et al., 2024).

2. Test systems and methods

2.1. Test systems

Most of the datasets used in this work contain high-resolution data because there were matching neutron structures available for such types of data. High-resolution datasets are commonly used in HAR development, which has consequences for method assessment, discussed later. We have chosen the following datasets (see Fig. 1):

(1) Carbamazepine, form III – X-ray data (d_{min} = 0.42 Å) and neutron structure from Sovago et al. (2016), T = 100 K.

(2) Gly-L-Ala – X-ray data (d_{min} = 0.65 Å) and neutron structure from Capelli et al. (2014), T = 150 K.

(3) Ice VI, a disordered form of heavy ice (D_2O) – X-ray data (d_{min} = 0.62 Å) from Chodkiewicz et al. (2022b), p = 1.15 GPa, room temperature, in-house source data; neutron structure from Kuhs et al. (1989), p = 0.85 GPa, T = 296 K.

(4) L-Alanine (L-Ala) – X-ray data (d_{min} = 0.46 Å) from Destro et al. (1988), neutron structure from Malaspina et al. (2019), T = 23 K.
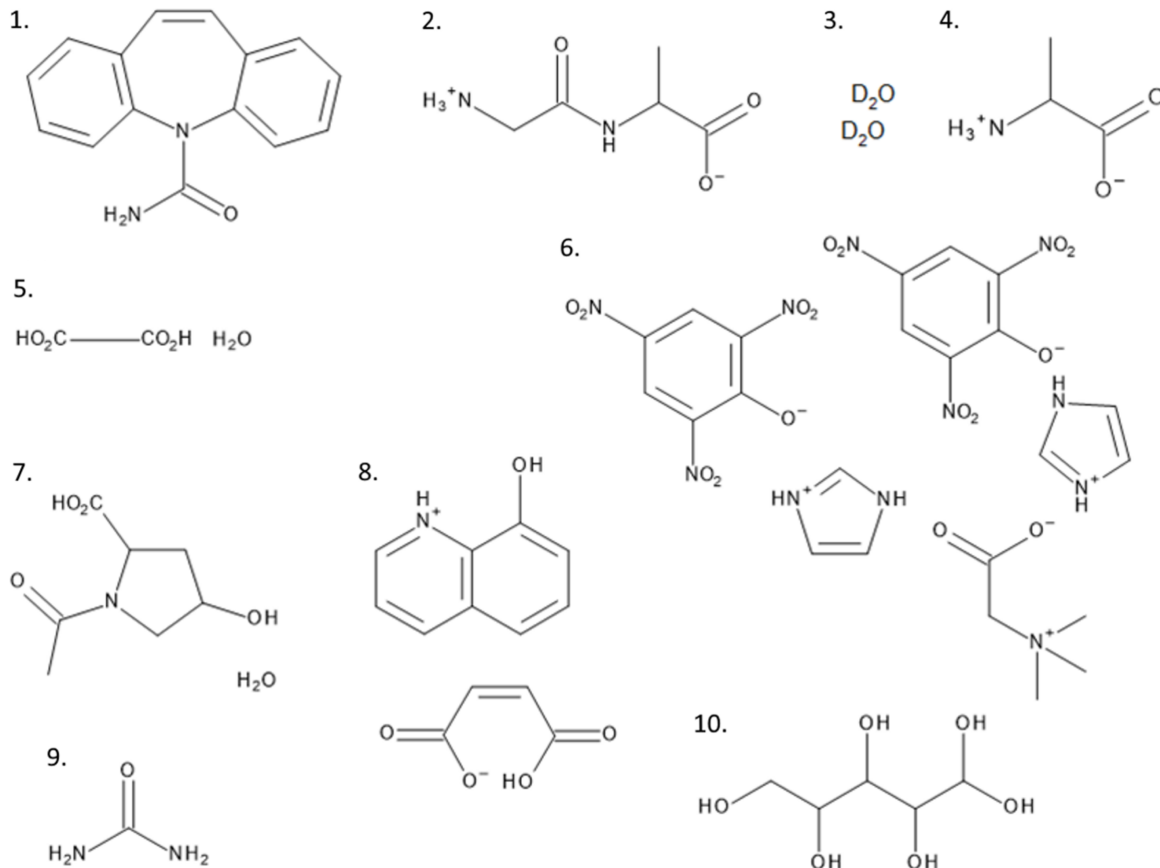

Figure 1
The test systems, symmetry independent molecules/ions shown: (1) carbamazepine, (2) Gly-L-Ala, (3) ice VI, (4) L-Ala, (5) Oxa·2H₂O, (6) BIPa, (7) NAC·H₂O, (8) 8HQ HM, (9) urea and (10) xylitol.

(5) Oxalic acid dihydrate (Oxa·2H$_2$O) – X-ray data (dataset oxa7, d_{min} = 0.43 Å) and neutron structure from Kamiński *et al.* (2014), T = 100 K.

(6) BIPa, a co-crystal of a betaine zwitterion, two imidazolium cations and two picrate anions – X-ray data (d_{min} = 0.42 Å; Overgaard *et al.*, 1999)] from Fugel *et al.* (2018) and neutron structure from Jørgensen *et al.* (2014), T = 100 K.

(7) *N*-acetyl-L-4-hydroxyproline monohydrate (NAC·H$_2$O) – X-ray data (d_{min} = 0.49 Å) and neutron structure from Lübben *et al.* (2014), T = 9 K.

(8) 8-hydroxyquinolinium hydrogen maleate (8HQ HM) – X-ray data (d = 0.43 Å) from Malaspina *et al.* (2020) T = 15 K; neutron data from Malaspina *et al.* (2017), T = 12 K.

(9) Urea – X-ray data (d_{min} = 0.36 Å) from Birkedal *et al.* (2004) and neutron structure from Swaminathan *et al.* (1984), T = 123 K.

(10) Xylitol – X-ray data (d_{min} = 0.42 Å) from Madsen *et al.* (2004) and neutron structure from Madsen *et al.* (2003), T = 122 K.

2.2. Refinements

HAR and a modified version of HAR with alternative partitions of the electron density (described later) were performed using a locally modified version of *Olex2* (Dolomanov *et al.*, 2009; Bourhis *et al.*, 2015). A program based on a development version of the DiSCaMB library (Chodkiewicz *et al.*, 2018) was used to generate files with atomic form factors in .tsc format (Kleemiss *et al.*, 2021; Midgley *et al.*, 2019). Such files were then imported into *Olex2* and used in the refinement conducted with *olex2.refine*. The HAR procedure was repeated until the maximum shift of a parameter divided by its standard deviation was lower than 0.1. Though we think the threshold is already sufficient, a tighter threshold (0.01) is commonly used in HAR (Capelli *et al.*, 2014).

HAR involves the calculation of a wavefunction. The most popular approach, also applied here, is based on the calculation of a wavefunction for a molecular system (in contrast to a periodic one). The wavefunction is calculated for molecular species extracted from the crystal structure. For the structures with Z' = 1 (where Z' denotes the number of symmetry-independent molecules in the asymmetric unit), the molecule from the asymmetric unit is used. For the other systems, the following subsystems have been used: Oxa·2H$_2$O – oxalic acid with the two closest water molecules (one subsystem); BIPa – five separate subsystems (*i.e.* wavefunctions calculated separately) corresponding to the five chemical units in the asymmetric unit; NAC·H$_2$O – the two chemical units treated as one subsystem; and 8HQ HM – wavefunctions for the two ions were calculated independently. The ice VI case is more complicated due to disorder; here, calculations of atomic form factors involved averaging atomic electron densities from multiple possible local conformations of a water molecule. Wavefunctions were calculated for clusters of five water molecules (the 'central' one and its closest neighbours). Six such clusters were used. The configurations of the neighbouring molecules were chosen randomly from a larger

number of possible configurations that fulfil the so-called ice rules (Bernal & Fowler, 1933). Other possible representations would differ only in the position of the hydrogen atom that is not directly involved in the hydrogen bond to the central molecule. The electron densities of the atoms in the central molecule were used for form-factor calculations.

Distributed multipoles up to dipoles were used to represent interactions in a crystal, for all structures except ice VI. The multipoles were located on chemical units up to 8 Å apart from the one for which the wavefunction is calculated, and they were calculated using the same electron density partition as the one used in the refinement.

The majority of quantum mechanical calculations were performed using DFT with the B3LYP functional and MP2 using the cc-pVTZ basis set. A few additional refinements used the HF method with the cc-pVTZ and cc-pVDZ basis sets. *ORCA* (version 5.0; Neese, 2012, 2022) was used for DFT and HF calculations, and *Gaussian16* (revision C. 01; Frisch *et al.*, 2016) was used for MP2 calculations.

Form factors were calculated via numerical integration with a 590-point Lebedev–Laikov grid for angular integration (Lebedev & Laikov, 1999) and a 75-point Mura–Knowles (Mura & Knowles, 1996) radial integration grid (the same grid size was used for all atoms).

2.3. Statistics

The accuracy of the refined structural parameters was assessed by comparison with reference neutron structures.

2.3.1. Bond lengths

The accuracy of the bond lengths was described with an average absolute difference between the value from the X-ray refinement and the reference one:

$$\langle|\Delta d|\rangle = \frac{1}{n}\sum_{k=1}^{n}\left|d_k^{\text{Xray}} - d_k^{\text{neutron}}\right|.$$

Similarly defined was the (non-absolute) average difference, $\langle\Delta d\rangle$, which was useful for detecting trends in bond lengths.

2.3.2. ADPs

ADPs for non-hydrogen atoms from neutron and X-ray measurements sometimes differed systematically. Comparison of hydrogen ADPs may involve some corrections that account for differences, *e.g.* a scaling procedure introduced by Blessing (1995). However, it did not fully solve the problem since the parameters of the scaling procedure were derived from ADPs of non-hydrogen atoms and, in principle, they may differ from those that should be used for hydrogen atoms.

To mitigate the problem of different scales for X-ray and neutron ADPs, a similarity indicator, MSD$_{corr}$, was introduced that is independent of the scale of the ADPs and compares the directionality of atomic displacements. It is defined in an analogous way as a correlation coefficient, though it misses statistical interpretation. It measures the 'correlation' of atomic mean square displacements (MSDs). The MSD of an atom is defined as the mean value of the square of the

projection of its displacement onto the chosen direction given by a unit vector $\mathbf{n}$ (Nelmes, 1969), related to ADP tensor $\mathbf{U}$ by the simple relation: $\text{MSD}_{\mathbf{n}}(\mathbf{U}) = \mathbf{n}^{\text{T}}\mathbf{U}\mathbf{n}$. A root mean square displacement $\text{RMSD}_{\mathbf{n}} = (\text{MSD}_{\mathbf{n}})^{1/2}$ was used to construct peanut plots (Hummel *et al.*, 1990). MSD_{corr} is defined similarly to a Pearson correlation coefficient between

$$\text{MSD}_{\text{corr}}(\mathbf{U}_A, \mathbf{U}_B) = \frac{\text{cov}\left[\text{MSD}(\mathbf{U}_A), \text{MSD}(\mathbf{U}_B)\right]}{\sigma_{\text{MSD}}(\mathbf{U}_A)\sigma_{\text{MSD}}(\mathbf{U}_B)},$$

where cov is the 'covariance', defined by

$$\text{cov}\left[\text{MSD}(\mathbf{U}_A), \text{MSD}(\mathbf{U}_B)\right] = \left\langle\left[\text{MSD}(\mathbf{U}_A) - \mu_{\text{MSD}}(\mathbf{U}_A)\right]\right.$$
$$\left.\times \left[\text{MSD}(\mathbf{U}_B) - \mu_{\text{MSD}}(\mathbf{U}_B)\right]\right\rangle_{\mathbf{n}}.$$

The expression in angle brackets is averaged over all possible displacement directions $\mathbf{n}$, μ_{MSD} is the averaged MSD (also averaged over all possible $\mathbf{n}$) and is equal to the equivalent isotropic ADP (U_{eq}), and $\sigma_{\text{MSD}}(\mathbf{U}_A)$ is the standard deviation of $\text{MSD}(\mathbf{U}_A)$: $\sigma_{\text{MSD}}(\mathbf{U}_A) = \left\{\text{cov}\left[\text{MSD}(\mathbf{U}_A), \text{MSD}(\mathbf{U}_A)\right]\right\}^{1/2}$. Since MSD_{corr} is a correlation coefficient, it can take values between -1 and 1. It is undefined for isotropic ADPs. Adding an identity matrix multiplied by a number to the ADP tensor did not change MSD_{corr}. To illustrate the properties of the index, its values for a few combinations of the diagonal ADP tensors are given: $\text{MSD}_{\text{corr}}[\text{diag}(2,1,1), \text{diag}(1,2,2)] = -1$, $\text{MSD}_{\text{corr}}[\text{diag}(2,1,1), \text{diag}(1,2,1)] = -0.5$, $\text{MSD}_{\text{corr}}[\text{diag}(2,1,1), \text{diag}(2,2,1)] = 0.5$, $\text{MSD}_{\text{corr}}[\text{diag}(2,1,1), \text{diag}(3,1,1)] = 1$, $\text{MSD}_{\text{corr}}[\text{diag}(3,3,2), \text{diag}(3,1,2)] = 0$. The calculation of MSD_{corr} was performed by numerical integration over a unit sphere using a Lebedev–Laikov grid with 5810 points (however, an analytical expression was also derived, see the supporting information).

Other scale-independent ADP similarity measures exist, *e.g.* the modified correlation coefficient (Kondrashov *et al.*, 2007) and the normalized correlation coefficient (Merritt, 1999). They are closely related to the correlation coefficient (CC) introduced by Merritt (1999) which, in turn, is closely related to the S_{12} index $[S_{12} = 100(1 - \text{CC})]$ introduced by Whitten & Spackman (2006). Another 'shape' similarity measure is the angle between the longest principal axes of the ellipsoids (Yang *et al.*, 2009; Mroz *et al.*, 2021). MSD_{corr} was chosen in this work because it is relatively easy to interpret the numerical values of the index since it is an analog of the correlation coefficient.

In addition to MSD_{corr}, the rescaled overlapping coefficient (Chodkiewicz *et al.*, 2024) was used. It is based on the overlapping coefficient (Inman & Bradley, 1989) which, for one-dimensional probability distributions, is the overlapping area of the distribution functions. The pair distribution functions (PDFs) $[p_1(\mathbf{u}), p_2(\mathbf{u})]$ of atomic displacements ($\mathbf{u}$), corresponding to the compared ADPs, can be written as

$$\eta = \int \min\left[p_1(\mathbf{u}), p_2(\mathbf{u})\right]d^3\mathbf{u}$$

or alternatively with the help of ℓ_1, a norm of the difference of PDFs as $\eta = 1 - 0.5||p_1(\mathbf{u}) - p_2(\mathbf{u})||_1$. The rescaled version $\eta_r = 100(1 - \eta)$ can be seen as the percentage difference between PDFs. This similarity index is scale-dependent so the

potential difference in the scale of ADPs from neutron and X-ray measurements would affect its value.

A ratio of equivalent isotropic ADPs (U_{eq}) was used to compare the extent of atomic displacements.

The computer program *compare_adps* for computing the ADP similarity indices is freely available at https://www.discamb.org/download.html; it is based on the DiSCaMB library (Chodkiewicz *et al.*, 2018).

3. Insight from the analysis of ADPs

An assessment of HAR as a refinement method usually involves a comparison of HAR-derived hydrogen atom structural parameters, mainly bond lengths and ADPs.

A bond-length comparison is relatively straightforward but its interpretation may be hindered by a cancellation of errors which may occur when two sources of error lead to the opposite effects on the bond lengths. For example, HAR for xylitol performed with the B3LYP functional using the cc-pVTZ basis set and distributed multipoles (to mimic the effect of the surrounding molecules) leads to $O-H$ bond lengths which differ from neutron diffraction results on average by 17 mÅ. In comparison, HAR with HF and a smaller basis set (cc-pVDZ) reduces the error to 4.5 mÅ. It might be surprising that the method that neglects correlation (HF) combined with a smaller basis set gives better results. But when the R factors are compared, it turns out that B3LYP-based HAR performs better ($R_1 = 1.56$) than HF-based refinement ($R_1 = 1.69$). Higher R factors but more accurate bond lengths are rather typical for HF-based refinements (*e.g.* Chodkiewicz *et al.*, 2024; Landeros-Rivera *et al.*, 2023; Wieduwilt *et al.*, 2020).

Clearly, analysis of the bond-length statistics alone might lead to incorrect conclusions. Therefore, it is important to also include ADPs in the analysis. Comparison of hydrogen ADPs from neutron and X-ray measurements is hindered by the fact that already the ADPs for heavier atoms sometimes differ considerably. There is no commonly accepted approach to tackle this problem. The differences between ADPs of non-hydrogen atoms can be approximately expressed in terms of simple adjustment schemes (Blessing, 1995, and references therein). Parameters of such a transformation fitted for non-hydrogen atoms can be applied to hydrogen ADPs (Blessing, 1995). These transformations can be useful (*e.g.* for deriving fixed hydrogen atom parameters in the X-ray analysis of electron density distribution) but sometimes they are also used to rescale ADPs before comparing neutron- and X-ray-derived structures. Such analysis should be applied with caution, since hydrogen atoms' contribution to high-angle scattering is very small. Their form factors are even a few hundred times smaller than those of heavier elements (like C, N and O) at the maximum resolution of some of the measurements used for HAR-testing purposes, see *e.g.* the form-factor ratio for aspherical form factors of atoms in urea (Fig. 2, calculation details are provided in the supporting information). Therefore, the effect of systematic errors on the hydrogen and non-hydrogen ADPs can be quite different and the application of a transformation fitted for non-hydrogen to
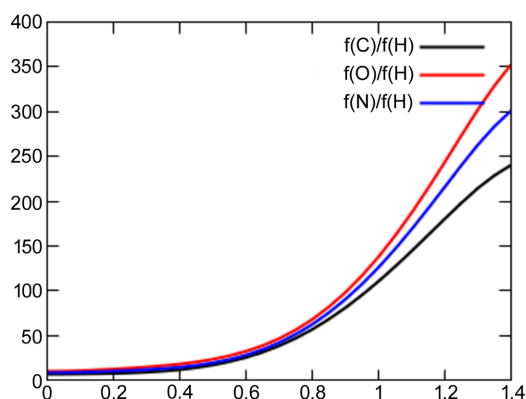
Figure 2
Ratio of non-hydrogen (C, N, O) to hydrogen atom form factors in urea as a function of $\sin(\theta)/\lambda$ (Å^{-1}).

hydrogen ADPs could be ineffective in correcting for systematic errors.

The simplest, isotropic correction, $U_{ij}^{X} = qU_{ij}^{N}$, uses only one parameter. The parameter q sometimes considerably changes when high-resolution X-ray data are omitted, suggesting that non-hydrogen ADPs determined with high-resolution data included ($U_{ij}^{X,high}$) can be systematically larger/smaller than those determined with limited resolution ($U_{ij}^{X,low}$). We can compare them using the same scheme: $U_{ij}^{X,low} = qU_{ij}^{X,high}$. For example, for the high-resolution carbamazepine structure [$\sin(\theta)/\lambda$ up to 1.19 Å^{-1} (Sovago *et al.*, 2016)], the q parameter is 0.865 (HAR with B3LYP, for low-resolution data, $d_{min} = 0.8$ Å), indicating that the use of high-resolution data leads to smaller ADPs in this case. Low-resolution refinement seems to be more similar to the neutron structure, which is especially visible in the case of the amide group (Fig. 3). MSD_{corr} for nitrogen-bonded
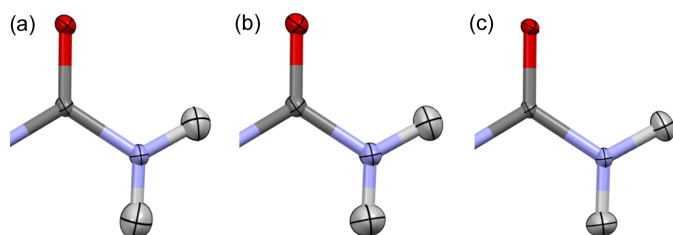

Figure 3
Amide group in carbamazepine, derived with (*a*) HAR with all data included ($d_{min} = 0.42$ Å), (*b*) HAR with limited data ($d_{min} = 0.8$ Å) and (*c*) neutron refinement.

hydrogen atoms increased from 0.87 to 0.98 and, in addition, $\langle|\Delta d|\rangle$ dropped from 20 to 9 mÅ. In the case of carbon-bonded hydrogen atoms, the MSD_{corr} increased from 0.89 to 0.94 and $\langle|\Delta d|\rangle$ stayed at the same level (5 mÅ).

Since hydrogen atoms' contribution to the scattering factor is tiny at high resolution, we would not expect that the contribution plays an important role in the change of hydrogen ADPs with resolution. One of the possible explanations for the observed inaccuracy of hydrogen ADPs in carbamazepine is that non-hydrogen ADPs refined for high-resolution data are not optimal for low-resolution data (they differ in size as quantified by the scaling factor $q = 0.865$), but the inaccuracies can be partially compensated with modifications to the hydrogen ADPs.

This hypothesis was tested with a dataset that provides a very good match between the neutron and X-ray measurement structures and also contains the urea amide group. In this case, the q parameter for the low-/high-resolution X-ray structure comparison is close to one (0.996). However, if nitrogen ADPs in urea are fixed to 90% of their original value (equivalent to $q = 0.9$ for N), the hydrogen ADPs become quite different from those from neutron measurement (Fig. 4). Similarly, as in the case of carbamazepine refinement for high-resolution data, one of the hydrogens in the amide group is elongated along the bond direction.

For the systems tested, use of high-resolution data increased the accuracy of polar hydrogen ADPs (as measured with MSD_{corr}) only when the scale factor q for non-hydrogen ADPs at low and high resolution is close to one (Fig. 5 and Table 1, low-resolution refinements were performed with $d_{min} = 0.8$ Å, results for B3LYP refinements are shown; MP2 refinements gave similar results, included in Section S10 of the supporting information). Otherwise, an increase in the resolution led to a decrease in the accuracy. This happened for four of the structures: carbamazepine, NAC·H$_2$O, L-Ala and BIPa.

4. Electron density partition with adjustable interatomic overlap

Inaccuracies in hydrogen ADPs from HAR appear to be a function of the ADP inaccuracies of their bonding partners. Assuming that a larger overlap of atomic electron densities makes the influence of the bonding partner larger, the inaccuracy can be at least partially eliminated by lowering the overlap. For that purpose, we have constructed an electron density partition based on the original Hirshfeld partition, defined as
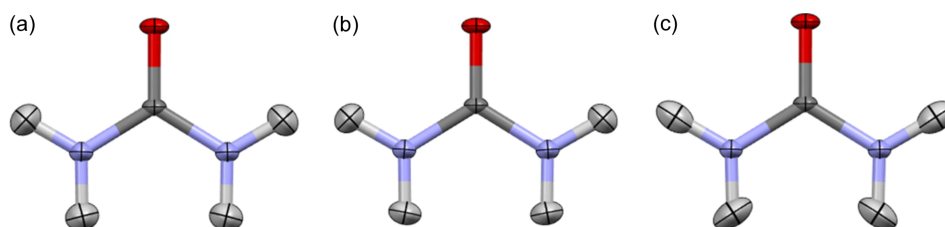

Figure 4
Urea structure derived from (*a*) HAR, (*b*) neutron experiment, (*c*) HAR with nitrogen ADPs fixed to 90% value of the values from HAR refinement.

Table 1
Scale parameter q for scaling low-resolution refinement ($d_{min} = 0.8$ Å) non-hydrogen ADPs to values from high-resolution refinement and the average atomic mean square displacement correlation for polar hydrogen atoms for the two refinements calculated with respect to neutron measurement data.

		$\langle MSD_{corr} \rangle$	
	q	Max $d = 0.8$	Max $d = d_{max}$
Carbamazepine	0.864	0.982	0.874
NAC·H_2O	0.884	0.731	0.080
L-Ala	0.924	0.631	0.271
BIPa	0.931	0.416	0.188
Oxa·$2H_2O$	0.968	0.685	0.744
8HQ HM	0.990	0.404	0.513
Ice VI	0.993	0.439	0.541
Urea	0.996	0.964	0.986
Xylitol	1.004	0.312	0.568
Gly-L-Ala	1.020	0.278	0.524

$$\rho_A(\mathbf{r}) = \frac{\left[\rho_A^0(\mathbf{r})\right]^n}{\sum_k \left[\rho_k^0(\mathbf{r})\right]^n} \rho(\mathbf{r}),$$

where $\rho_A(\mathbf{r})$ is the atomic electron density of atom A, $\rho_A^0(\mathbf{r})$ is the spherically averaged atomic density of the isolated atom A and $\rho(\mathbf{r})$ is total electron density. For $n = 1$, this is the original Hirshfeld partition. Hereafter, the partition will be referred to as the exponential Hirshfeld partition and the corresponding refinement as expHAR(n), where n is an adjustable parameter of the partition (exponent). Choosing parameter $n > 1$ lowers the overlap of atomic electron densities. The effect is illustrated using a simple model of the H_2^+ ion electron density, with the wavefunction $\psi(r) = N[\varphi A(r) + \varphi B(r)]$, where φA and φB are hydrogen atom $1s$ orbitals (Fig. 6).

An application of expHAR(2) to carbamazepine (with the B3LYP/cc-pVTZ level of theory) brings a clear improvement of the nitrogen-bonded hydrogen atom parameters (Fig. 7). Bond lengths and the shape of the thermal ellipsoids are now closer to those from the neutron experiment.
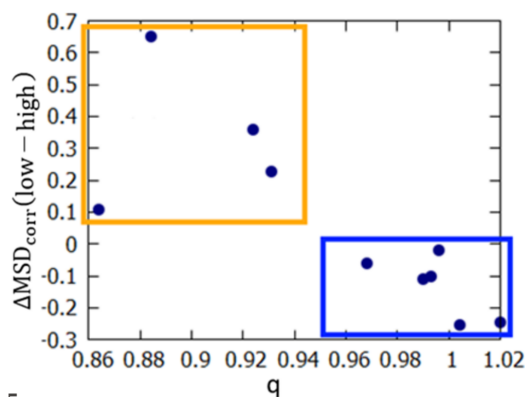
Figure 5
Change in polar hydrogen ADP accuracy (measured with average MSD_{corr}) when the data resolution is limited to $d = 0.8$ Å plotted against the scale parameter q for scaling low-resolution refinement ($d_{min} = 0.8$ Å) non-hydrogen ADPs to values from high-resolution refinement. Limiting the resolution deteriorated the accuracy of the ADPs when parameter q was close to 1 (points inside the blue box) and improved otherwise (points inside the orange box).
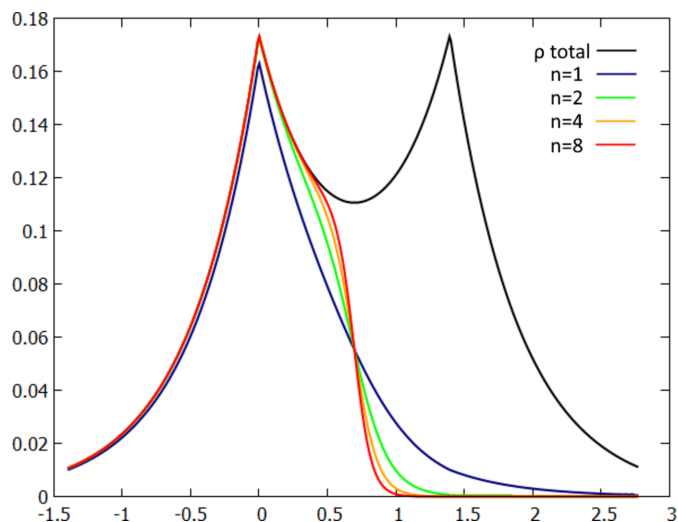
Figure 6
Dependence of atomic electron density on the parameter n in the exponential Hirshfeld partition for the model system: H_2^+ ion. Black line – total electron density, other lines – atomic electron densities for various values of the parameter n in the exponential Hirshfeld partition.
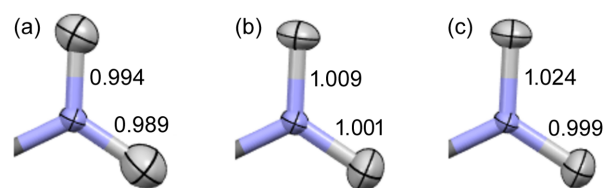
Figure 7
NH_2 group in carbamazepine, derived using (a) HAR, (b) expHAR(2) and (c) neutron diffraction.

Hydrogen atom ellipsoid elongation along the bond direction [as in Fig. 7(a)] seems much more common in the case of polar hydrogen atoms than in the case of carbon-bonded hydrogen atoms. This effect, as well as the smaller sensitivity of the C—H bond lengths and their ADPs to change of resolution, can be explained by a lower overlap of the atomic electron density. The overlap can be quantified with an overlap integral for the atomic densities:

$$o_{AB} = \int \rho_A(\mathbf{r})\rho_B(\mathbf{r})d^3\mathbf{r}.$$

The values of this parameter for C—H, N—H and O—H pairs in selected structures are given in Table 2. The overlap can be ordered in the following way: O—H > N—H > C—H – larger overlap appears to correspond to more frequent problems with the description of hydrogen ADPs. The overlap is much smaller for the exponential Hirshfeld partition with an exponent of 2, which is more than $2\times$ smaller than the overlap for the C—H pair for the Hirshfeld partition. The values of the overlap integral o_{AB} were calculated using the wavefunction from HAR, $i.e.$ the same geometry and wavefunction were used for both the Hirshfeld partition and the exponential Hirshfeld partition calculations.

Increasing the exponent in the exponential Hirshfeld partition usually leads to higher absolute values of atomic charges (see Section S11 of the supporting information).

Table 2
Average overlap coefficients for $X-H$ pairs in chosen systems for the Hirshfeld and exponential Hirshfeld partitions.

	Hirshfeld partition			Exponential Hirshfeld partition ($n = 2$)		
	O—H	N—H	C—H	O—H	N—H	C—H
Carbamazepine	–	0.0735	0.0510	–	0.0216	0.0186
Gly-L-Ala	–	0.0633	0.050	–	0.0184	0.0183
Xylitol	0.086	–	0.0498	0.0202	–	0.0185

Increasing the exponent to infinity limits the overlap of atomic electron densities to zero and introduces boundaries between atoms. There are few other non-overlapping partitions of electron density known. One of the most prominent ones was developed within the so-called quantum theory of atoms in molecules (QTAIM) by Bader (1990). There is also a group of partitions based on the concept of Voronoi polyhedra (Voronoi, 1908) proposed *e.g.* by Politzer & Harris (1970) and Rousseau *et al.* (2001). Becke (1988) introduced a fuzzy version of such partition as part of a scheme for numerical integration in DFT. Similar to the case of the exponential Hirshfeld partition, the level of overlap of atomic electron densities can be controlled with a parameter (which is an integer number in this case). Becke partition was already applied in HAR-like refinement (Chodkiewicz *et al.*, 2020) but it led to less accurate results than other partitions tested in that paper.

Since atoms in crystals constantly move, diffraction experiments 'see' dynamic electron density. It is common to approximate it as a sum of dynamic atomic electron densities represented as static atomic densities smeared out by atomic motion/displacement. Such dynamic atomic electron densities overlap even if the static atomic electron densities are described with a non-overlapping partition.

5. expHAR test calculations and discussion

Tests of expHAR were performed with ten test systems and using various exponents n (1.0, 1.25, 1.5, 2, 3 and 4). In addition to the B3LYP functional, second-order Møller–Plesset perturbation theory (MP2) is used for the electron density

calculation (except for BIPa, which is omitted from the MP2 calculations due to its size and the slow convergence of the HAR procedure). For the xylitol structure, HF-based HAR/expHAR refinements were also performed.

5.1. R factors

Exponential HAR with $n > 1$ usually leads to a slightly larger R factor than HAR. R_1 factors for B3LYP expHAR(2) were up to 0.02% larger than for HAR. A *SHELXL*-type (Sheldrick, 2008, 2015) weighting scheme was used during the refinements, *i.e.* the weights are defined as $w = 1/(\sigma^2(F_o^2) + (aP)^2 + bP$, where $P = [2|F_c|^2 + \max(|F_o|^2, 0)]$, and F_c and F_o are the calculated and observed structure factors. Therefore, R factors between the refinements are, in principle, not directly comparable since the a and b parameters may be different for different n in expHAR(n). For example, for L-alanine the a and b parameters change from 0.005 and 0 for expHAR(2) to 0.02 and 0.011 for expHAR(3) (MP2 refinements) and wR_2 changes significantly from 3.17 to 3.97%, yet if we used the same weighting scheme as in expHAR(2) in both refinements, then the difference in wR_2 was much smaller, only 0.02%, wR_2 = 3.19% for expHAR(3). For the B3LYP refinements, we did not observe such dramatic changes [Fig. 8(a)]. Changes in wR_2 might be caused by (1) different weighting schemes, (2) different capabilities of the partitions to describe dynamic electron density within convolution approximation and (3) different capabilities of the model to mimic the effects of experimental errors. It is unclear to what extent the factors listed contribute to the differences in wR_2. Though we can eliminate the influence of different weighting schemes, the effects of the two other factors cannot be easily distinguished. Therefore, we will not provide an explanation for the differences in wR^2.

5.2. X—H bond lengths

HAR-related methods provide much more accurate lengths of covalent bonds to hydrogen ($X-H$) than IAM in the case of organic structures. Still, there is a considerable variation in the accuracy of the HAR-derived $X-H$ bond lengths [see Fig. 9(a)]. Since C—H bond lengths appear to be reproduced with
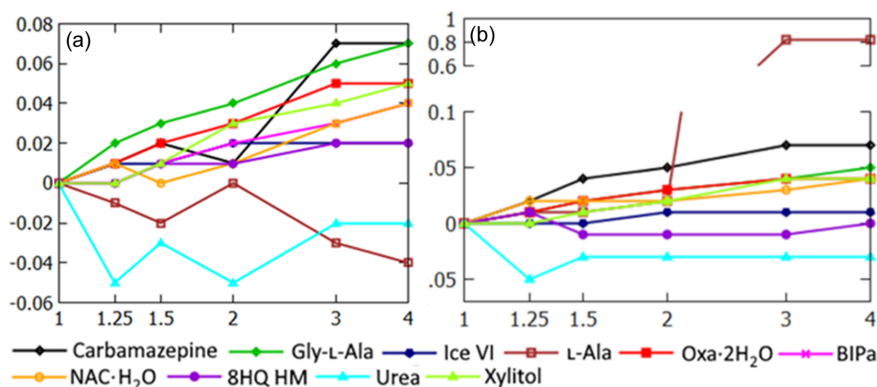
Figure 8
wR_2 (calculated for all data used in the refinement) for expHAR(n) as a function of n for refinements with (a) the B3LYP functional (b) using MP2. The large increase of wR_2 for L-Ala MP2 refinement is caused by the change of parameters in the weighting scheme (see the text for details). The lines connecting the datapoints serve as a guide to the eye only and cannot be used to predict wR_2 between the points.

HAR more accurately than O—H bonds (Woińska *et al.*, 2016), the bond-length analysis was performed separately for polar X—H bonds and for C—H bonds. Statistics calculated without dividing the hydrogen atoms into groups are included in the supporting information.

5.2.1. Polar X—H bond lengths

The average absolute difference ($\langle |\Delta d| \rangle$) between HAR and neutron-measurement-derived polar X—H bond lengths ranges from 3 mÅ for urea to 39 mÅ for NAC·H$_2$O for B3LYP refinements (up to 31 mÅ for MP2 refinements). This suggests that there might be room for considerable improvement in some cases. Indeed, for the structures with the largest $\langle |\Delta d| \rangle$ there is a very significant improvement when expHAR is applied [Figs. 9(a) and 9(b)] – *e.g.* for NAC·H$_2$O $\langle |\Delta d| \rangle$ decreases from 39 to 16 and for 8HQ HM from 27 to 14 mÅ when switching from HAR to expHAR(2) in the case of the B3LYP refinements. The X—H bond lengths increased with increasing n [Figs. 9(c) and 9(d)]. In the case of the B3LYP-based refinements, bond-length accuracy (measured through $\langle |\Delta d| \rangle$) showed an increased of n up to $n = 2$ for all structures except urea. The bond lengths for urea obtained with HAR were already very good and increasing n to 2 increased $\langle |\Delta d| \rangle$ from 3 to 8 mÅ in this case. In the case of MP2-based refinements [Fig. 9(b)], the situation is quite similar, however the improvement stops for lower n in some cases and, in addition to urea, expHAR does not improve polar X—H bond

lengths also in the case of Gly-L-Ala. It was observed that MP2 gives systematically longer polar X—H bonds than B3LYP (Wieduwilt *et al.*, 2020; Chodkiewicz *et al.*, 2020). This is notable in the average difference ($\langle \Delta d \rangle$) plots [Figs. 9(c) and 9(d)]. Therefore, a smaller elongation of the polar bonds was necessary to obtain optimal values in the case of MP2 refinements and as a consequence also smaller values of n led to optimized values.

Application of HF in HAR leads to relatively long polar X—H bond lengths (Chodkiewicz *et al.*, 2020; Capelli *et al.*, 2014; Wieduwilt *et al.*, 2020; Landeros-Rivera *et al.*, 2023). As a result, the application of HF in expHAR may also lead to long bonds. Such a possibility was tested with the xylitol structure. The O—H bonds in HF-based HAR structures were already too long – on average 9 mÅ longer than from the neutron measurement and for expHAR(2) this number increased to 21 mÅ. HF gave superior O—H bond lengths to B3LYP but larger R factors in the case of HAR, whereas in the case of expHAR(2) it was inferior in terms of both R factors and bond lengths.

5.2.2. C—H bond lengths

Non-polar bond lengths are also improved with exponential HAR in the majority of the tested cases [see Figs. 10(a) and 10(b)]. In this case, the discrepancies between HAR and the reference neutron values are on average smaller than in the case of the polar X—H bonds – the maximal $\langle |\Delta d| \rangle$ values
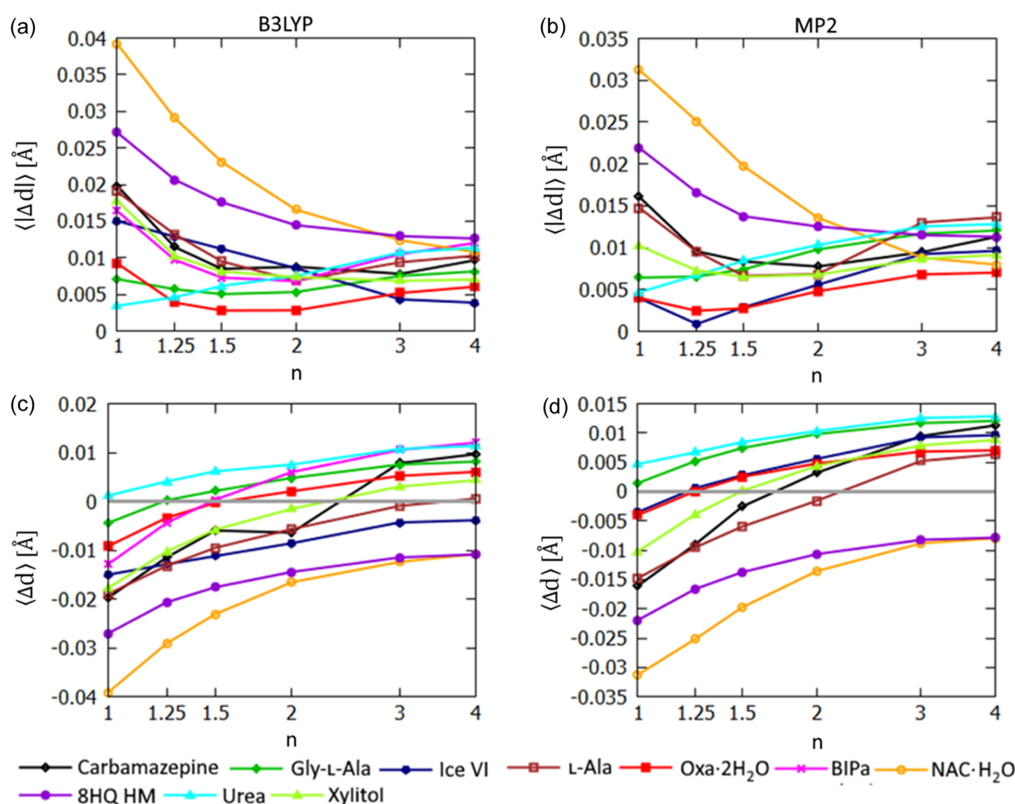

Figure 9
Polar X—H bond-length statistics for structures derived with expHAR(n) as a function of n. Average $\langle |\Delta d| \rangle$ of X—H bond lengths from neutron values for refinement with (*a*) B3LYP and (*b*) MP2. Average $\langle \Delta d \rangle$ for X—H bond lengths from neutron values for refinement with (*c*) B3LYP and (*d*) MP2.

Figure 10
C−H bond-length statistics for structures derived with expHAR(n) as a function of n. Average $\langle|\Delta d|\rangle$ of X−H bond lengths from neutron values for refinement with (a) B3LYP and (b) MP2. Average $\langle\Delta d\rangle$ of X−H bond lengths from neutron values for refinement with (c) B3LYP and (d) MP2.
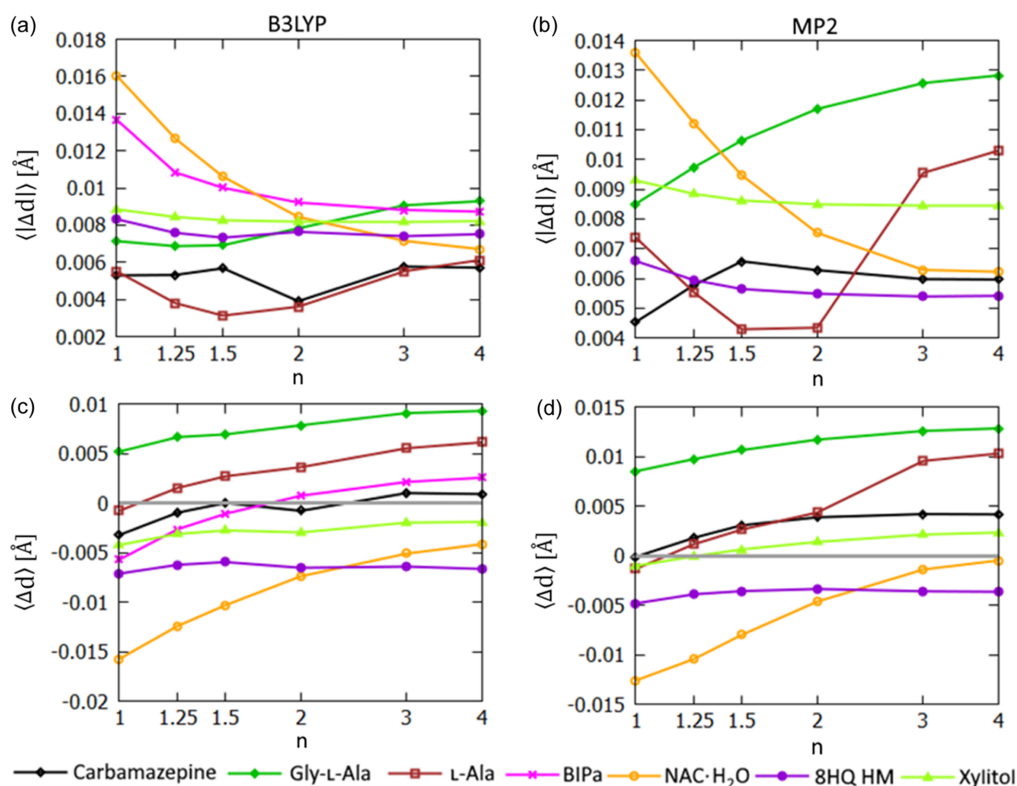
were 39 and 16 mÅ for polar and C−H bonds, respectively, for B3LYP refinements and 31 and 14 mÅ for MP2 refinements (NAC·H$_2$O structure in all cases). The C−H bond lengths were improved (in terms of $\langle|\Delta d|\rangle$) in 6/7 cases when switching from HAR to expHAR(1.5) for B3LYP-based refinements and in 4/6 cases for MP2-based refinements.

5.3. ADPs

An application of expHAR(n) with $n > 1$ improves the polar hydrogen ADPs in the majority of the test structures (Fig. 11). It increased the average mean square displacement correlation ($\langle\mathrm{MSD_{corr}}\rangle$) in 9/10 structures in the case of B3LYP-based refinements and in 8/9 structures in the case of MP2-based refinements. The average rescaled overlapping coefficient ($\langle\eta_r\rangle$) decreased in 8/10 structures (B3LYP) and 6/9 structures (MP2).

In HAR-derived structures, the non-polar hydrogen atoms had higher values of $\langle\mathrm{MSD_{corr}}\rangle$ [Figs. 12(a) and 12(b)] than the polar ones for all structures. Sometimes the difference was very large; for BIPa: 0.79 versus 0.19 (B3LYP refinements). The lowest $\langle\mathrm{MSD_{corr}}\rangle$ for polar hydrogen atoms was 0.08 for NAC·H$_2$O while for non-polar ones it was much closer to 1 (e.g. 0.71 for L-Ala). Clearly the polar atom ADPs were more challenging to reproduce with HAR. Also the non-polar hydrogen ADPs improved in a majority of the structures (as measured in $\langle\mathrm{MSD_{corr}}\rangle$ terms) when expHAR(n) with $n > 1$ was used in all but one structure (glycine) where it decreased slightly. The improvement was less obvious when the $\langle\eta_r\rangle$

indicator was used [Figs. 12(c) and 12(d)], but still $\langle\eta_r\rangle$ values were improved (decreased) in 5/7 cases (B3LYP) and 4/6 cases (MP2) when switching from HAR to expHAR(2)

A clear trend can be observed in the size of the hydrogen ADPs – an increase of n in expHAR(n) leads to smaller ADPs (smaller U_{eq}, see Fig. 13) for both polar and non-polar hydrogen atoms. Increasing the n parameter appears to make hydrogen atom thermal ellipsoids 'thinner' in the X−H bond direction (Fig. 14), especially when they were too elongated in that direction. Sometimes it led to a very significant improvement in the directionality of the corresponding atomic displacement, e.g. the MSD$_{\mathrm{corr}}$ for one of the hydrogens shown in Fig. 14 changed from -0.7 for HAR to 0.66 for expHAR(4). On the other hand, it may also potentially lead to overly small ADPs.

5.4. Optimal exponent parameter n

Replacing HAR with expHAR(n) with $n > 1$ improved the accuracy of hydrogen atom parameters in the majority of the structures tested. There was no improvement in the case of the urea structure – increasing n above 1 made the X−H bond lengths slightly too large in this case but practically did not change $\langle\mathrm{MSD_{corr}}\rangle$. This might suggest that if switching from HAR to expHAR($n > 1$) had only a minor impact on hydrogen ADPs (in terms of $\langle\mathrm{MSD_{corr}}\rangle$), then expHAR would not produce significant improvement in hydrogen atom parameters and hence n close to 1 was a good choice. For all structures tested, there was practically no deterioration of
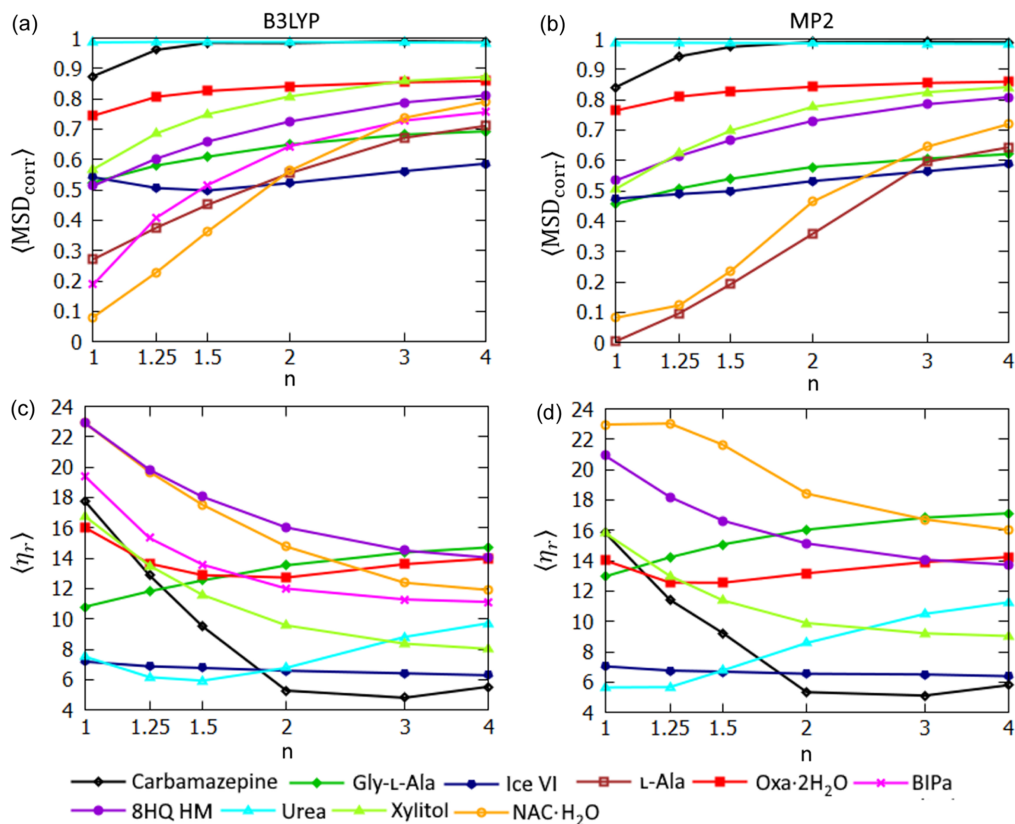
Figure 11
Polar hydrogen ADP statistics for structures derived with expHAR(n) as a function of n. Average $\langle MSD_{corr} \rangle$ calculated with the neutron measurement structure as a reference for refinement with (a) B3LYP and (b) MP2. Average $\langle \eta_r \rangle$ for atomic displacement probability distribution functions, calculated using the neutron measurement structure as a reference for refinement with (c) B3LYP and (d) MP2.
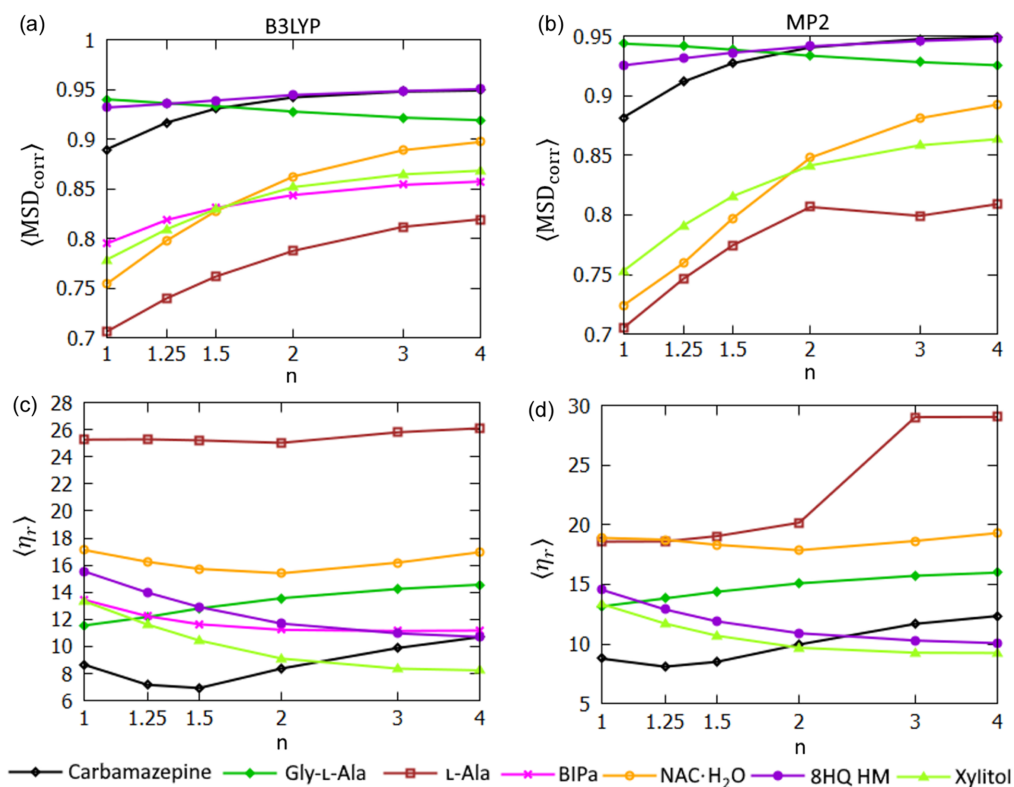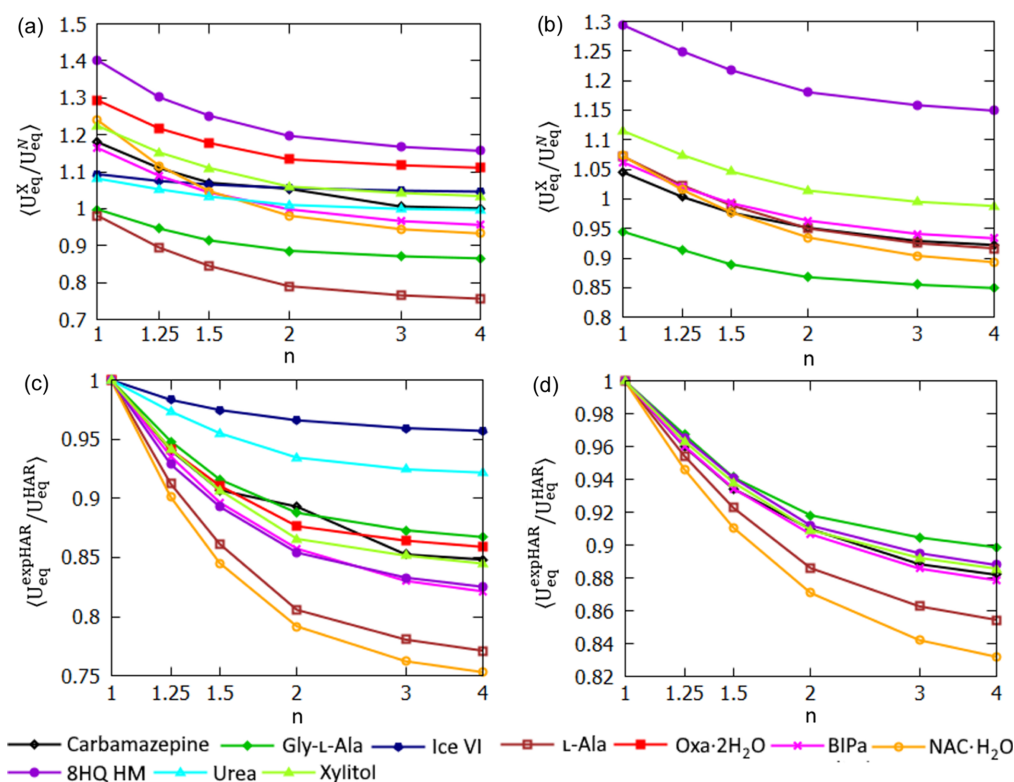

Figure 12
Non-polar hydrogen ADP statistics for structures derived with expHAR(n) as a function of n. Average $\langle MSD_{corr} \rangle$ calculated with neutron measurement structure as a reference for refinement with (a) B3LYP and (b) MP2. Average $\langle \eta_r \rangle$ for atomic displacement probability distribution functions, calculated using the neutron measurement structure as a reference for refinement with (c) B3LYP and (d) MP2.

Figure 13
Hydrogen atomic displacement extended for structures derived with expHAR(n) as a function of n. Average ratio of equivalent isotropic ADPs from expHAR(n) and neutron measurement (U_{eq}^X/U_{eq}^N) for (a) polar and (b) non-polar hydrogen atoms. Average ratio of equivalent isotropic ADPs from expHAR(n) and HAR ($U_{eq}^{expHAR}/U_{eq}^{HAR}$) for (c) polar and (d) non-polar hydrogen atoms.

$\langle MSD_{corr} \rangle$ observed with an increase of n. Therefore, if increasing n changes the ADPs, it was probably an improvement of the ADPs, at least in terms of $\langle MSD_{corr} \rangle$. The situation was different in the case of bond lengths – increasing n led to overly long $X-H$ bonds. In the case of B3LYP, $\langle |\Delta d| \rangle$ for polar $X-H$ bonds decreased with n up to $n = 1.5$, except in the urea structure for which it only increased. For $n > 1.5$, $\langle |\Delta d| \rangle$ started to rise in some cases, but for $n = 2$ the increase was below 0.0003 Å. It seems that, provisionally, $n = 2$ can be expected to be a safe choice that improves the accuracy of the refinements or is at least not detrimental in the case of B3LYP refinements. In the case of MP2, $n = 1.5$ is preferential. Note that there is no single choice of parameter n that is optimal for all the structures tested and the more general applicability of such rules remains to be tested.

6. Conclusions

In this work, we tried to identify factors limiting the accuracy of HAR. Many datasets used so far in testing HAR performance were collected to high resolution. It was observed that cutting the data resolution can sometimes improve the accuracy of HAR. This happened with structures in which non-hydrogen ADPs derived from high-resolution data and those derived from limited-resolution data differed most significantly. On the contrary, cutting the resolution led to lower accuracy when the difference was small. Different sizes of non-

hydrogen ADPs from high- and low-resolution refinements suggested that non-hydrogen ADPs refined for high-resolution data were not optimal for low-resolution data. This can be partially compensated for with modifications to the neighbouring hydrogen ADPs, leading to inaccurate hydrogen ADPs. A similar situation was observed when the non-hydrogen ADPs were artificially decreased.

It was assumed that the overlap of atomic electron densities facilitated the described mechanism of distortion of hydrogen ADPs. A new partition of the electron density that can lead to
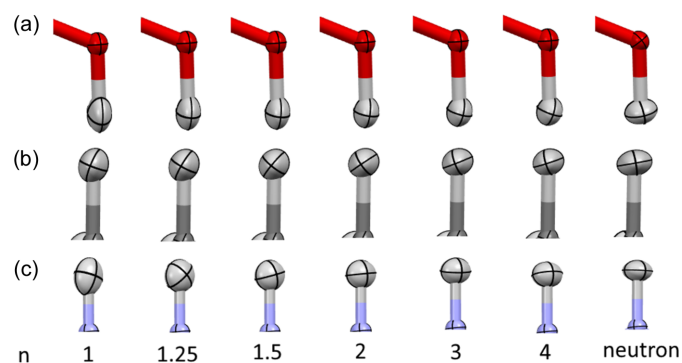
Figure 14
Dependence of thermal ellipsoids on the parameter n in expHAR(n) for selected hydrogen atoms: (a) oxygen-bonded hydrogen atom (H11) in xylitol, (b) carbon-bonded hydrogen atom (H5A) in xylitol, (c) nitrogen-bonded hydrogen atom (H1m) in BIPa. The last column shows the neutron measurement results.

a lower overlap of the atomic electron densities was proposed as a remedy. This 'exponential Hirshfeld' partitioning is based on the original Hirshfeld partition and it has an 'exponent' parameter n allowing for an adjustment of the overlap of atomic electron densities. Setting n to 1 makes both partitions identical and increasing n reduces the overlap of atomic electron densities.

The effect of applying the exponential Hirshfeld partition in HAR-like procedures (expHAR) was tested on a set of structures of polar organic molecules using electron density calculated with the B3LYP functional (ten structures), and with the MP2 method (nine structures) and a range of n parameters. The resulting structures were compared with reference neutron measurements. ExpHAR with $n > 1$ usually led to slightly larger wR_2 agreement factors than HAR (up to 0.02 per cent points for B3LYP refinements); however, comparison of R factors is hampered by the fact that the *SHELX*-type weighting scheme is used with adjustable parameters that differ from refinement to refinement, making the R factors in principle incomparable (which was especially clear for one of the refinements with the MP2 method). The differences between X-ray- and neutron-derived $X-H$ bond lengths were significantly reduced in the case of structures for which they were the largest. The average absolute differences were reduced for 9/10 structures in the case of B3LYP refinements (8/9 for MP2). Hydrogen ADPs also improved in 9/10 of the structures for B3LYP-based refinements and 8/9 for MP2-based refinement when the ADPs were compared with a newly introduced scale-independent similarity measure. The improvement was especially visible in the case of ADPs elongated along the $X-H$ bonds.

It was observed that oxygen- and nitrogen-bonded hydrogen atoms are usually less accurately described with HAR than carbon-bonded ones and that their atomic electron densities overlap to a greater extent with the electron density of their bonding partner. This observation is in line with the assumption that the lower overlap of atomic electron densities makes hydrogen atom parameters less sensitive to the potential inaccuracies of the parameters of their bonding partner. Under this assumption, the lower overlap is not an advantage when there are no such inaccuracies. In such a situation, we would expect that the most accurate structural parameters can be obtained with the electron density partition that most accurately reproduces the dynamic electron density when applied with the convolution approximation for vibrational smearing. In the case of expHAR, reducing the atomic overlap by increasing the parameter n leads at some point to too-long $X-H$ bonds. Setting n to infinity would make the atomic electron densities completely non-overlapping which does not seem to be an optimal choice. To use expHAR most effectively, a method for optimizing the exponent parameter n is needed. Yet such a method has not been introduced, only some provisional rules were proposed which appeared to be suitable for the tested systems.

An application of expHAR partition in HAR has led to significant improvements in the description of ADPs of hydrogen atoms in some test cases. We hope that further research will enable even greater accuracy improvements for the HAR-related methods.

References

Bader, R. F. W. (1990). *Atoms in Molecules*. Oxford University Press.

Bąk, J. M., Domagała, S., Hübschle, C., Jelsch, C., Dittrich, B. & Dominiak, P. M. (2011). *Acta Cryst.* A**67**, 141–153.

Becke, A. D. (1988). *J. Chem. Phys.* **88**, 2547–2553.

Bergmann, J., Davidson, M., Oksanen, E., Ryde, U. & Jayatilaka, D. (2020). *IUCrJ*, **7**, 158–165.

Bernal, J. D. & Fowler, R. H. (1933). *J. Chem. Phys.* **1**, 515–548.

Birkedal, H., Madsen, D., Mathiesen, R. H., Knudsen, K., Weber, H.-P., Pattison, P. & Schwarzenbach, D. (2004). *Acta Cryst.* A**60**, 371–381.

Blessing, R. H. (1995). *Acta Cryst.* B**51**, 816–823.

Bourhis, L. J., Dolomanov, O. V., Gildea, R. J., Howard, J. A. K. & Puschmann, H. (2015). *Acta Cryst.* A**71**, 59–75.

Brock, C. P., Dunitz, J. D. & Hirshfeld, F. L. (1991). *Acta Cryst.* B**47**, 789–797.

Capelli, S. C., Bürgi, H.-B., Dittrich, B., Grabowsky, S. & Jayatilaka, D. (2014). *IUCrJ*, **1**, 361–379.

Chodkiewicz, M., Patrikeev, L., Pawlędzio, S. & Woźniak, K. (2024). *IUCrJ*, **11**, 249–259.

Chodkiewicz, M., Pawlędzio, S., Woińska, M. & Woźniak, K. (2022*a*). *IUCrJ*, **9**, 298–315.

Chodkiewicz, M. L., Gajda, R., Lavina, B., Tkachev, S., Prakapenka, V. B., Dera, P. & Wozniak, K. (2022*b*). *IUCrJ*, **9**, 573–579.

Chodkiewicz, M. L., Migacz, S., Rudnicki, W., Makal, A., Kalinowski, J. A., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Adams, P. D. & Dominiak, P. M. (2018). *J. Appl. Cryst.* **51**, 193–199.

Chodkiewicz, M. L., Woińska, M. & Woźniak, K. (2020). *IUCrJ*, **7**, 1199–1215.

Destro, R., Marsh, R. E. & Bianchi, R. (1988). *J. Phys. Chem.* **92**, 966–973.

Dittrich, B., Hübschle, C. B., Luger, P. & Spackman, M. A. (2006). *Acta Cryst.* D**62**, 1325–1335.

Dittrich, B., Hübschle, C. B., Messerschmidt, M., Kalinowski, R., Girnt, D. & Luger, P. (2005). *Acta Cryst.* A**61**, 314–320.

Dittrich, B., Hübschle, C. B., Pröpper, K., Dietrich, F., Stolper, T. & Holstein, J. J. (2013). *Acta Cryst.* B**69**, 91–104.

Dittrich, B., Koritsánszky, T. & Luger, P. (2004). *Angew. Chem. Int. Ed.* **43**, 2718–2721.

Dittrich, B., Lübben, J., Mebs, S., Wagner, A., Luger, P. & Flaig, R. (2017). *Chem. A Eur. J.* **23**, 4605–4614.

Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K. & Puschmann, H. (2009). *J. Appl. Cryst.* **42**, 339–341.

Domagała, S., Fournier, B., Liebschner, D., Guillot, B. & Jelsch, C. (2012). *Acta Cryst.* A**68**, 337–351.

Dominiak, P. M., Volkov, A., Li, X., Messerschmidt, M. & Coppens, P. (2007). *J. Chem. Theory Comput.* **3**, 232–247.

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). *Gaussian16*, Revision C. 01. Gaussian, Inc., Wallingford, Connecticut, USA.

Fugel, M., Jayatilaka, D., Hupf, E., Overgaard, J., Hathwar, V. R., Macchi, P., Turner, M. J., Howard, J. A. K., Dolomanov, O. V., Puschmann, H., Iversen, B. B. & Grabowsky, S. (2018). *IUCrJ*, **5**, 32–44.

Hansen, N. K. & Coppens, P. (1978). *Acta Cryst.* A**34**, 909–921.

Hirshfeld, F. (1977). *Isr. J. Chem.* **16**, 198–201.

Hoser, A. A. & Madsen, A. Ø. (2016). *Acta Cryst.* A**72**, 206–214.

Hoser, A. A. & Madsen, A. Ø. (2017). *Acta Cryst.* A**73**, 102–114.

Hummel, W., Raselli, A. & Bürgi, H.-B. (1990). *Acta Cryst.* B**46**, 683–692.

Inman, H. F. & Bradley, E. L. Jr (1989). *Comm. Statist. Theory Methods*, **18**, 3851–3874.

Jarzembska, K. N. & Dominiak, P. M. (2012). *Acta Cryst.* A**68**, 139–147.

Jayatilaka, D. & Dittrich, B. (2008). *Acta Cryst.* A**64**, 383–393.

Jha, K. K., Gruza, B., Kumar, P., Chodkiewicz, M. L. & Dominiak, P. M. (2020). *Acta Cryst.* B**76**, 296–306.

Jha, K. K., Gruza, B., Sypko, A., Kumar, P., Chodkiewicz, M. L. & Dominiak, P. M. (2022). *J. Chem. Inf. Model.* **62**, 3752–3765.

Jørgensen, M. R. V., Hathwar, V. R., Sist, M., Wang, X., Hoffmann, C. M., Briseno, A. L., Overgaard, J. & Iversen, B. B. (2014). *Acta Cryst.* A**70**, 679–681.

Kamiński, R., Domagała, S., Jarzembska, K. N., Hoser, A. A., Sanjuan-Szklarz, W. F., Gutmann, M. J., Makal, A., Malińska, M., Bąk, J. M. & Woźniak, K. (2014). *Acta Cryst.* A**70**, 72–91.

Kleemiss, F., Dolomanov, O. V., Bodensteiner, M., Peyerimhoff, N., Midgley, L., Bourhis, L. J., Genoni, A., Malaspina, L. A., Jayatilaka, D., Spencer, J. L., White, F., Grundkötter-Stock, B., Steinhauer, S., Lentz, D., Puschmann, H. & Grabowsky, S. (2021). *Chem. Sci.* **12**, 1675–1692.

Kondrashov, D. A., Van Wynsberghe, A. W., Bannen, R. M., Cui, Q. & Phillips, G. N. (2007). *Structure*, **15**, 169–177.

Koritsanszky, T., Volkov, A. & Chodkiewicz, M. (2011). *New directions in pseudoatom-based X-ray charge density analysis*. In *Electron Density and Chemical Bonding II. Structure and Bonding*, Vol. 147. Berlin, Heidelberg: Springer.

Koritsanszky, T., Volkov, A. & Coppens, P. (2002). *Acta Cryst.* A**58**, 464–472.

Kuhs, W., Ahsbahs, H., Londono, D. & Finney, J. (1989). *Physica B*, **156–157**, 684–687.

Kumar, P., Gruza, B., Bojarowski, S. A. & Dominiak, P. M. (2019). *Acta Cryst.* A**75**, 398–408.

Landeros-Rivera, B., Ramírez-Palma, D., Cortés-Guzmán, F., Dominiak, P. M. & Contreras-García, J. (2023). *Phys. Chem. Chem. Phys.* **25**, 12702–12711.

Lebedev, V. I. & Laikov, D. N. (1999). *Dokl. Math.* **59**, 477–481.

Lübben, J., Bourhis, L. J. & Dittrich, B. (2015). *J. Appl. Cryst.* **48**, 1785–1793.

Lübben, J., Volkmann, C., Grabowsky, S., Edwards, A., Morgenroth, W., Fabbiani, F. P. A., Sheldrick, G. M. & Dittrich, B. (2014). *Acta Cryst.* A**70**, 309–316.

Madsen, A. Ø. & Hoser, A. A. (2014). *J. Appl. Cryst.* **47**, 2100–2104.

Madsen, A. Ø., Mason, S. & Larsen, S. (2003). *Acta Cryst.* B**59**, 653–663.

Madsen, A. Ø., Sørensen, H. O., Flensburg, C., Stewart, R. F. & Larsen, S. (2004). *Acta Cryst.* A**60**, 550–561.

Malaspina, L. A., Edwards, A. J., Woińska, M., Jayatilaka, D., Turner, M. J., Price, J. R., Herbst-Irmer, R., Sugimoto, K., Nishibori, E. & Grabowsky, S. (2017). *Cryst. Growth Des.* **17**, 3812–3825.

Malaspina, L. A., Hoser, A. A., Edwards, A. J., Woińska, M., Turner, M. J., Price, J. R., Sugimoto, K., Nishibori, E., Bürgi, H.-B., Jayatilaka, D. & Grabowsky, S. (2020). *CrystEngComm*, **22**, 4778–4789.

Malaspina, L. A., Wieduwilt, E. K., Bergmann, J., Kleemiss, F., Meyer, B., Ruiz-López, M. F., Pal, R., Hupf, E., Beckmann, J., Piltz, R. O., Edwards, A. J., Grabowsky, S. & Genoni, A. (2019). *J. Phys. Chem. Lett.* **10**, 6973–6982.

Merritt, E. A. (1999). *Acta Cryst.* D**55**, 1997–2004.

Meyer, B. & Genoni, A. (2018). *J. Phys. Chem. A*, **122**, 8965–8981.

Midgley, L., Bourhis, L. J., Dolomanov, O., Peyerimhoff, N. & Puschmann, H. (2019). *arXiv*:1911.08847.

Mroz, D., Wang, R., Englert, U. & Dronskowski, R. (2021). *CrystEngComm*, **23**, 1052–1058.

Mura, M. E. & Knowles, P. J. (1996). *J. Chem. Phys.* **104**, 9848–9858.

Nassour, A., Domagala, S., Guillot, B., Leduc, T., Lecomte, C. & Jelsch, C. (2017). *Acta Cryst.* B**73**, 610–625.

Neese, F. (2012). *WIREs Comput. Mol. Sci.* **2**, 73–78.

Neese, F. (2022). *WIREs Comput. Mol. Sci.* **12**, e1606.

Nelmes, R. J. (1969). *Acta Cryst.* A**25**, 523–526.

Overgaard, J., Schiøtt, B., Larsen, F. K., Schultz, A. J., MacDonald, J. C. & Iversen, B. B. (1999). *Angew. Chem. Int. Ed.* **38**, 1239–1242.

Pichon-Pesme, V., Lecomte, C. & Lachekar, H. (1995). *J. Phys. Chem.* **99**, 6242–6250.

Politzer, P. & Harris, R. R. (1970). *J. Am. Chem. Soc.* **92**, 6451–6454.

Rousseau, B., Peeters, A. & Van Alsenoy, C. (2001). *J. Mol. Struct. Theochem*, **538**, 235–238.

Ruth, P. N., Herbst-Irmer, R. & Stalke, D. (2022). *IUCrJ*, **9**, 286–297.

Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.

Sheldrick, G. M. (2015). *Acta Cryst.* C**71**, 3–8.

Sovago, I., Gutmann, M. J., Senn, H. M., Thomas, L. H., Wilson, C. C. & Farrugia, L. J. (2016). *Acta Cryst.* B**72**, 39–50.

Swaminathan, S., Craven, B. M. & McMullan, R. K. (1984). *Acta Cryst.* B**40**, 300–306.

Volkov, A., Messerschmidt, M. & Coppens, P. (2007). *Acta Cryst.* D**63**, 160–170.

Voronoi, G. F. (1908). *J. Die Reine Angew. Math. (Crelles J.)* **134**, 198–287.

Wall, M. E. (2016). *IUCrJ*, **3**, 237–246.

Wanat, M., Malinska, M., Hoser, A. A. & Woźniak, K. (2021). *Molecules*, **26**, 3730.

Whitten, A. E. & Spackman, M. A. (2006). *Acta Cryst.* B**62**, 875–888.

Wieduwilt, E. K., Macetti, G., Malaspina, L. A., Jayatilaka, D., Grabowsky, S. & Genoni, A. (2020). *J. Mol. Struct.* **1209**, 127934.

Woińska, M., Chodkiewicz, M. L. & Woźniak, K. (2021). *Chem. Commun.* **57**, 3652–3655.

Woińska, M., Grabowsky, S., Dominiak, P. M., Woźniak, K. & Jayatilaka, D. (2016). *Sci. Adv.* **2**, e1600192.

Woińska, M., Hoser, A. A., Chodkiewicz, M. L. & Woźniak, K. (2024). *IUCrJ*, **11**, 45–56.

Woińska, M., Pawlędzio, S., Chodkiewicz, M. L. & Woźniak, K. (2023). *J. Phys. Chem. A*, **127**, 3020–3035.

Yang, L., Song, G. & Jernigan, R. L. (2009). *Proteins*, **76**, 164–175.

Zarychta, B., Pichon-Pesme, V., Guillot, B., Lecomte, C. & Jelsch, C. (2007). *Acta Cryst.* A**63**, 108–125.