PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2025, **27**, 15906

Combining double-hybrid functionals with rSCAN yields solid-state ^{13}C chemical shifts with sub-ppm accuracy†

Harry Brough, [a] Chris W. Cook, [ab] John M. Griffin [a] and Michael J. G. Peach *[a]

Ab initio calculations of NMR shieldings are often used to assign spectra and help refine crystal structures in the growing field of NMR crystallography. In periodic calculations, GGA exchange–correlation functionals such as PBE and BLYP are most often used, but a "monomer correction" has recently been proposed that incorporates a "higher quality" treatment of local electronic structure into calculated shieldings. The meta-GGA functional rSCAN reportedly generates improved geometries, particularly in systems with important dispersion interactions, but has scarcely been tested for its performance in periodic shielding calculations, with or without monomer corrections. Here, the performance of rSCAN is evaluated by comparing experimental chemical shifts from 75 diverse ^{13}C environments in 13 molecular solids, to chemical shifts calculated by rSCAN and PBE on geometries optimised by rSCAN, PBE and BLYP. We find rSCAN gives marginally improved geometries but produces less accurate chemical shifts than PBE. However, after a monomer correction is applied to the shieldings, corrected rSCAN consistently performs better than corrected PBE, indicating that rSCAN describes long-range contributions to shieldings more accurately than local effects. Monomer correction with a double-hybrid functional has previously been found to provide no additional benefit compared to correction with conventional hybrids. However, we show the double-hybrid mPW2PLYP predicts substantially improved chemical shifts when the monomer correction method is paired with an implicit solvation model, yielding better results than a correction with a cluster of molecules using a conventional hybrid functional. The method we find maximises agreement with experiment is a mPW2PLYP–CPCM correction to rSCAN periodic calculations on rSCAN-optimised geometries. When used on a larger set of organic crystals, with 132 ^{13}C environments, this method yields unprecedented accuracy, with root-mean-square error of 0.8 ppm and mean absolute error of 0.6 ppm.

Received 21st March 2025,
Accepted 7th July 2025

DOI: 10.1039/d5cp01111f

rsc.li/pccp

1. Introduction

The coupling of solid-state nuclear magnetic resonance (NMR) spectroscopy, X-ray diffraction (XRD) and quantum chemical calculations is known as NMR crystallography, and can provide additional insight into local structure, dynamics and interactions in solid materials.[1] In recent years, NMR crystallography has been applied diversely, from refining XRD-derived structures,[2–4] to characterising battery materials[5,6] and solving the structures of biological molecules.[7–9]

The success of NMR crystallography relies on a strong correlation between theoretical chemical shifts and experimental spectra, so increasing the accuracy of calculated nuclear shieldings is highly desirable. Solid-state shieldings are normally calculated by the gauge-including projector-augmented waves (GIPAW) approach[10] within Kohn–Sham density-functional theory (DFT),[11,12] which can provide an effective compromise between calculation accuracy and computational cost. A practical DFT calculation relies on a choice of density-functional approximation (DFA) to the unknown exchange–correlation functional, to incorporate quantum mechanical effects. For solid-state shielding calculations, the most widely used[1] is the generalised gradient approximation (GGA) by Perdew, Burke and Ernzerhof (PBE), which uses information from the local electron density and its gradient.[13]

1.1. Monomer corrections

Hybrid DFAs incorporate a proportion of the exact exchange expression from Hartree–Fock theory evaluated on the Kohn–Sham orbitals. These functionals have reduced self-interaction

[a] *Department of Chemistry, Lancaster University, Lancaster, LA1 4YB, UK. E-mail: m.peach@lancaster.ac.uk*
[b] *The Faraday Institution, Quad One, Didcot, OX11 0RA, UK*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d5cp01111f

error compared to GGA functionals and are often used for accurate molecular shielding calculations. However, so-called hybrids scale more steeply with the number of basis functions due to their explicit (occupied) orbital dependence, making them too computationally expensive to routinely use in a plane wave basis.[14,15]

To overcome this expense, a ''monomer correction'' approach to improve periodic shielding calculations has been proposed.[14] Here, GGA shielding calculations are performed on the whole crystal (using a plane-wave approach) and on a single molecule isolated from the periodic structure. A more reliable DFA—likely a hybrid functional—is also used to calculate shieldings on this isolated molecule. These molecular calculations are performed in an atom-centered basis, usually with gauge-including atomic orbitals (GIAO).[16]

The periodic shieldings are then ''corrected'' by

$$\sigma = \sigma_{\mathrm{GGA}}^{\mathrm{crystal}} - \sigma_{\mathrm{GGA}}^{\mathrm{molecule}} + \sigma_{\mathrm{high\text{-}level}}^{\mathrm{molecule}} \quad (1)$$

where σ is the isotropic shielding, the average of the three principal components of the shielding tensor. This correction combines the long-range influences of shielding from the GIPAW calculation with a higher-quality treatment of the local electronic structure.

This method, which assumes the difference in shieldings from two electronic structure methods is insensitive to the long-range geometry, has previously been shown to substantially increase agreement with experiment when correcting PBE shieldings with hybrid functionals.[14,17] Corrections to PBE have also been tested with ''double-hybrid'' DFAs, which incorporate a contribution from second-order Møller–Plesset perturbation theory within the resolution of the identity approximation (RI-MP2), but these were found to provide no additional benefit over the less costly conventional hybrids.[17]

1.2. The SCAN family of functionals

Exchange–correlation functionals that incorporate the non-interacting kinetic energy density, $\tau(\mathbf{r})$, in addition to the local density and its gradient, are known as meta-GGA (mGGA) functionals. The kinetic energy density provides implicit dependence on the curvature of the density *via* the gradients of occupied Kohn–Sham orbitals, $\varphi_i(\mathbf{r})$,[18]

$$\tau(\mathbf{r}) = \frac{1}{2}\sum_{i}^{\mathrm{occ}} |\nabla \phi_i(\mathbf{r})|^2. \quad (2)$$

Unlike computationally intensive (double-)hybrid functionals, mGGA functionals retain the $\mathcal{O}(N^3)$ scaling with basis functions of GGAs, making mGGAs potentially very useful for plane-wave calculations, providing they yield more accurate results.

However, an additional complication arises from incorporating $\tau(\mathbf{r})$. The magnetic field, $\mathbf{B}(\mathbf{r})$, is the curl of the vector potential, $\mathbf{A}(\mathbf{r})$, but since the curl of a gradient vanishes, there are many definitions of $\mathbf{A}(\mathbf{r})$ that produce the same $\mathbf{B}(\mathbf{r})$. The magnetic field and its related physical properties, are therefore said to be ''gauge invariant'', and physical equations should retain this invariance. This motivates the use of the GIPAW approach for periodic shielding calculations, and the use of GIAOs for calculations in an orbital basis.[19] However, the definition of $\tau(\mathbf{r})$ in eqn (2) is not invariant to the choice of gauge.[20]

Therefore, to calculate reliable shieldings using mGGAs, the form of $\tau(\mathbf{r})$ must be explicitly considered. The most widely used approach for dealing with this undesirable gauge-dependence was developed by Maximoff and Scuseria, $\tau_{\mathrm{MS}}(\mathbf{r})$.[21] However, this method has been shown to provide unphysical paramagnetic contributions to the shielding tensors of spherical atoms, and does not correctly generalise the constraint of exactness for one-electron systems to response properties.[20,22] A more costly, but physically meaningful approach $\tau_{\mathrm{D}}(\mathbf{r})$ developed by Dobson,[23] does not suffer from these problems, and has been shown to improve both nuclear shielding and time-dependent DFT calculations.[20,24]

The ''strongly constrained and appropriately normed'' (SCAN) functional is a mGGA designed to satisfy all of the known constraints of the (unknown) exact exchange–correlation functional.[25] The SCAN functional incorporates $\tau(\mathbf{r})$ *via* an ''iso-orbital indicator'', $\alpha(\mathbf{r})$, which identifies the bonding environment from the relationship between $\tau(\mathbf{r})$, the kinetic energy density of a one-orbital system, $\tau_{\mathrm{W}}(\mathbf{r})$,[26] and that of the uniform electron gas, $\tau_{\mathrm{U}}(\mathbf{r})$,[27]

$$\alpha(\mathbf{r}) = \frac{\tau(\mathbf{r}) - \tau_{\mathrm{W}}(\mathbf{r})}{\tau_{\mathrm{U}}(\mathbf{r})}. \quad (3)$$

The SCAN functional has been claimed to use $\alpha(\mathbf{r})$ to account for weak interactions,[25] demonstrating this ability by predicting the correct stability ordering for phases of ice—a task where PBE and the hybrid functional PBE0,[28] which incorporates 25% exact exchange, fail.[29]

However, SCAN was found to be numerically unstable, motivating the development of a regularised functional, rSCAN.[30] This new DFA was more efficient, but this came at the cost of breaking SCAN's exact constraints, leading to a large reduction in accuracy for atomisation energies.[27] The authors of SCAN have recently introduced a regularised and restored functional, r^{2}SCAN.[27] This was intended to bridge the gap between accuracy and stability, restoring all but one exact constraint of the parent functional. The DFA has been reported to match or exceed the accuracy of SCAN for several molecular properties.[31,32]

The rSCAN functional has recently been implemented into the widely used plane-wave DFT code, CASTEP,[33] but has seen few tests for its utility for NMR crystallography.[34–37] To our knowledge, there has never been a systematic review of the performance of rSCAN for GIPAW ^{13}C shielding calculations, or geometries in the context of solid-state NMR. Moreover, rSCAN has not before been tested with the monomer correction approach. Our overarching aim was therefore to assess whether the rSCAN functional could be a useful tool for NMR crystallography.

To do this, we calculated ^{13}C chemical shifts from a set of organic crystals to compare rSCAN's geometries to those from

PBE and the GGA by Becke, Lee, Yang and Parr (BLYP),[38,39] and to compare rSCAN's shieldings to those from PBE. Several monomer correction approaches were subsequently tested on the periodic shielding calculations, which were performed on each set of geometries. The performance of a method was quantified by agreement of theoretical chemical shifts with experimental chemical shifts determined from ^{13}C magic-angle spinning solid-state NMR spectra. We find that the most accurate method utilises monomer correction with the double-hybrid mPW2PLYP with a conductor-like polarisable continuum model (CPCM) solvent. Applying this correction to the rSCAN periodic shieldings, calculated from rSCAN-optimised geometries, yields highly accurate chemical shifts. On a larger test set of 132 ^{13}C environments,[17] this optimal method is found to yield an unprecedented accuracy with RMSE of 0.8 ppm.

2. Results and discussion

2.1. Periodic calculations

Structures of a diverse group of 13 molecular crystals, containing 75 carbon environments between them, were taken from the Cambridge Structural Database[40] (CSD). Fig. 1 outlines the molecules within these crystals, whose CSD reference codes can be found in Table S1 (ESI†). Structures were chosen to be easily obtainable, contain a range of different carbon environments and have minimal evidence of dynamic effects. For both the structural optimisations and subsequent periodic shielding calculations, a plane wave energy cut-off of 800 eV was used, and reciprocal space was sampled *via* a Monkhorst–Pack grid[41] with a *k*-point spacing of 0.03 Å^{-1}.

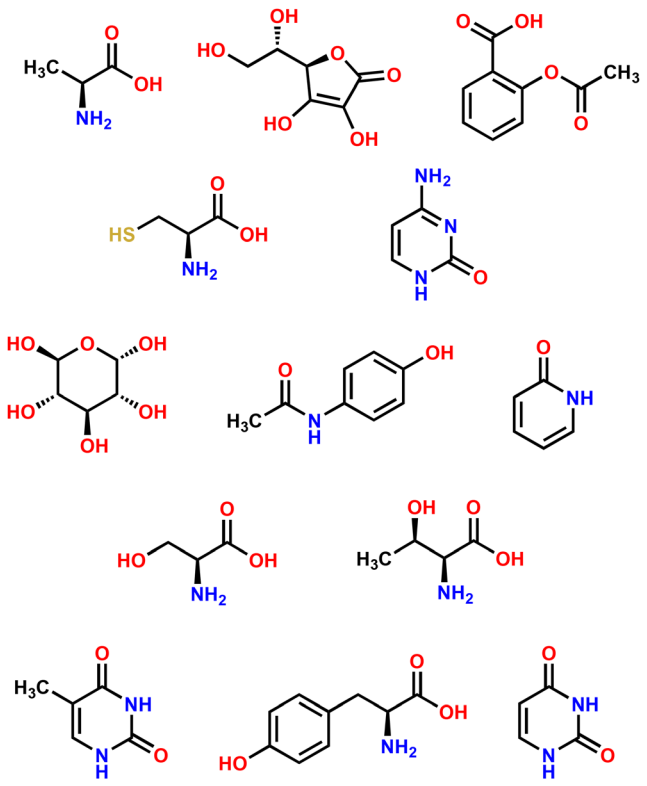

Fig. 1 Structures of molecules in our set of organic crystals.

The crystal structures were optimised with the PBE, BLYP and rSCAN functionals. Relaxation of unit cell parameters risks substantial increase of cell volumes if dispersion correction is not used[5] and does not incorporate the thermal expansion present at room temperature in the crystals on which solid-state NMR spectra were recorded. Therefore, unit cell parameters were fixed at their experimental values for optimisations with all three exchange–correlation functionals. This also facilitates more direct comparison to previously reported results where cell parameters were fixed.[17]

On the geometries computed with each of the three functionals, as well as on structures where only hydrogen atom postitions were relaxed, and on the diffraction-derived structures directly, nuclear shieldings were then calculated with PBE and rSCAN GIPAW. The BLYP functional has previously been shown to produce consistently inferior chemical shifts compared to PBE, including when monomer corrections are applied,[17] so shieldings with BLYP are not considered. Shieldings were converted to chemical shifts by the widely used linear regression procedure described in Section 3.2.1.

Fig. 2 and Table 1 show that a large reduction in root-mean-square-error (RMSE) and a narrowing of the error distribution results from allowing hydrogen atom positions to relax, with little sensitivity to the underlying functional, since their positions are poorly defined by XRD.[1] Fully relaxing the structures generally improves RMSE for our ambient temperature structures, but results in much higher sensitivity to the DFA used for geometry optimisation.

For both the PBE and rSCAN GIPAW calculations, the BLYP geometries produce the worst chemical shift agreement and the rSCAN structures produce the closest chemical shift agreement, although the difference between PBE and rSCAN geometries is small. This geometry sensitivity is most apparent in the rSCAN chemical shifts, which have a higher RMSE and interquartile range (IQR) than the PBE shifts. The best uncorrected results are produced by PBE GIPAW on the PBE or rSCAN geometries, with RMSEs of 2.1 ppm. The somewhat improved performance of rSCAN geometries for the rSCAN shielding calculations compared to PBE and BLYP should be taken in the context of the slightly higher computational cost for the mGGA calculations.

The apparent improvement of rSCAN for geometries—which becomes more convincing after monomer corrections are applied in Section 2.2—has been suggested to result from SCAN's incorporation of dispersion interactions using eqn (3).[25] We investigated this by calculating the energy profile over the interfacial bond of a dimer of planar uracil molecules with several electronic structure methods, described in Section S6 of the ESI.† Only the SCAN family of functionals and methods incorporating non-local MP2 correlation bound the dimer. The optimal binding distance was calculated to be 3.8 Å according to the consensus of DFAs corrected with the D4 semi-empirical dispersion correction scheme,[42] and the uncorrected

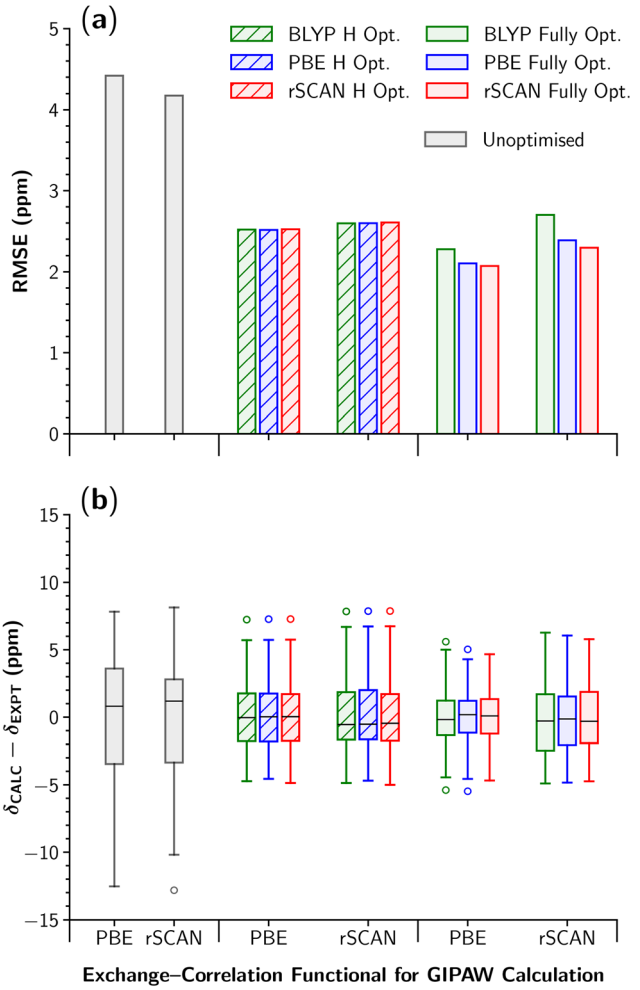

Fig. 2 Effect of underlying geometry (bar colours) and DFA for GIPAW calculation (*x*-axis) on (a) root-mean-square errors and (b) error distribution, comparing calculated to experimental chemical shifts. The central line on the box and whiskers plot indicates the median error, the top and bottom of the boxes are the first and third quartiles of the error distribution and the whiskers extend to 1.5× the interquartile range. Outliers are shown beyond this point as circles and are discussed in the Section S9 of the ESI.†

Table 1 Comparison of agreement with experiment of monomer corrections to ^{13}C shifts predicted by PBE and rSCAN GIPAW *versus* uncorrected GIPAW on rSCAN geometries. All values in ppm

Method	RMSE	MAE	IQR	MAX
PBE GIPAW	2.1	1.6	2.6	4.7
PBE0|PBE	1.5	1.2	2.0	4.3
mPW2PLYP|PBE	1.4	1.1	1.7	4.0
rSCAN GIPAW	2.3	1.9	3.9	5.8
PBE0|rSCAN	1.4	1.1	1.8	4.5
mPW2PLYP|rSCAN	1.2	0.9	1.4	4.2

SCAN family bound the dimer closer to this distance than non-local methods with steep $\mathcal{O}(N^5)$ scaling. This suggests that rSCAN's improvement over PBE for structures may result from the mGGA incorporating some treatment of dispersion forces.

We therefore calculated optimised geometries in CASTEP for our structures with PBE corrected with the D3 semi-empirical dispersion correction scheme[43] and rSCAN, with and without relaxing unit cell parameters. Interestingly, with unfixed cell parameters, rSCAN without dispersion correction performed as well as PBE–D3, with 0.2 ppm lower RMSE compared to conventional PBE. With fixed cell parameters, PBE GIPAW calculations on the PBE, PBE–D3 and rSCAN structures all gave RMSEs of 2.1 ppm. rSCAN has been shown to resemble the hybrid B3LYP[44] more closely than PBE for hydrogen bond lengths,[36] but in the optimisations where only hydrogen atom positions were relaxed, there was minimal sensitivity to the choice of functional used to optimise the structures. Therefore, rSCAN's performance when producing structures for ^{13}C NMR calculations results primarily from consideration of dispersion forces and other effects on heavy atoms.

2.2. Monomer corrections

Monomer corrections have previously demonstrated insensitivity to basis set.[14,17] To confirm this is true for our crystals, corrections with a number of basis sets, detailed in Tables S4–S6 (ESI†), to the PBE shieldings on the rSCAN geometries were calculated with the ORCA quantum chemistry package.[45] In these basis sets, corrections were calculated with the mGGA r^{2}SCAN, the hybrid functional PBE0, and MP2 within the resolution of the identity approximation.[46] The RI-MP2 corrections were the most sensitive to basis, performing best in aug-cc-pVTZ.[47] However, the cc-pVTZ basis set[48] was chosen due to its strong and consistent performance, lower computational cost, and to enable direct comparison with previously reported results.[14,17]

Multiple exchange–correlation functionals were chosen for monomer corrections to the PBE and rSCAN GIPAW calculations from each rung of Perdew's "Jacob's ladder" of density-functional approximations, representing increasing computational cost and reliability.[49] From the mGGA functionals, TPSS,[50] SCAN and r^{2}SCAN were tested. Functionals incorporating exact exchange included PBE0, B3LYP and the range-separated hybrid CAM-B3LYP.[51] Double-hybrid functionals included the physically-motivated PBE0-DH,[52] and the empirical B2PLYP and mPW2PLYP.[53,54] Corrections with RI-MP2 were also calculated. The Microsoft Excel document in the (ESI†) contains detailed statistics for all correction methods tested.

In total, we consider over fifteen monomer corrections each applied to two periodic calculations, themselves each performed on three sets of periodic geometries. For brevity, we therefore introduce notation to describe the corrected shieldings, of "correction| periodic calculation||geometry". For example, a PBE0 correction to shieldings calculated with PBE GIPAW on the rSCAN periodic geometries is denoted PBE0|PBE||rSCAN.

Fig. 3 and 4 show the effect of some of these monomer corrections to PBE and rSCAN GIPAW calculations, respectively. Corrections with mGGAs were highly sensitive to the geometry, particularly for the TPSS|rSCAN corrections, where using the rSCAN geometry lowers RMSE by 0.3 ppm compared to PBE and by 0.7 ppm compared to BLYP. Less geometry sensitivity is

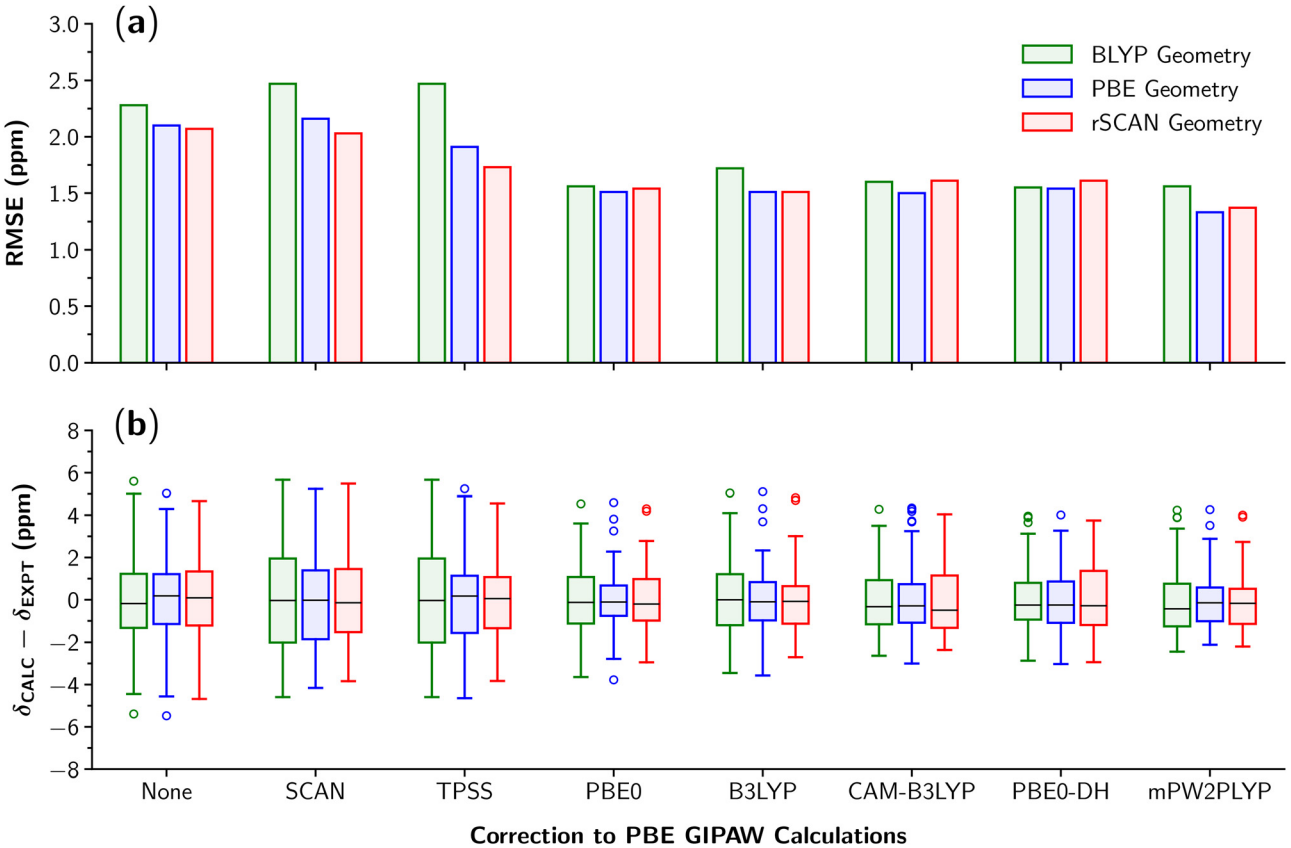

Fig. 3 (a) Root-mean-square errors and (b) error distribution of various monomer correction approaches to the PBE periodic shielding calculations, on geometries optimised with BLYP, PBE and rSCAN. Corrections with mGGA shieldings were calculated with $\tau_D(\mathbf{r})$.

observed for the (double-) hybrid corrections. Here, the rSCAN and PBE geometries perform similarly for corrections to PBE GIPAW, but the rSCAN geometries are consistently better than PBE for corrections to rSCAN GIPAW. Overall, the BLYP geometries perform the worst in thirteen cases, PBE performs the worst in two cases and rSCAN performs the worst only for the PBE0-DH|PBE|| rSCAN chemical shifts.

In contrast to previous results,[17] there is a small improvement in RMSE moving from corrections with PBE0 to those with mPW2PLYP for our crystals, from 1.4 to 1.2 ppm for PBE GIPAW. Moreover, the RMSE from PBE0|PBE||PBE is 1.5 ppm, 0.3 ppm higher than previously reported,[17] suggesting our set of molecules may be more difficult for the electronic structure methods to describe. These results, and the asymmetric distribution of outliers, are discussed in Section 2.4.

The rSCAN shieldings are slightly worsened by correction with SCAN, which may be due to SCAN's previously reported high dependence on the integration grid.[27] Corrections incorporating exact exchange substantially reduce RMSE, by at least 0.5 ppm for PBE and 0.8 ppm for rSCAN shieldings. Considering the low cost of these molecular calculations, the monomer correction method seems to be a simple and reliable strategy for NMR crystallography. Although the rSCAN geometry consistently performs well, and often slightly better than PBE, this reliability comes at a somewhat higher computational cost. There does not seem to be a benefit to using BLYP for geometries compared to PBE.

Table 1 shows the standard rSCAN GIPAW calculations benefit much more than PBE from correction, to the extent that using the (slightly) less reliable functional for GIPAW ends up a better choice if corrections are applied. This implies rSCAN's treatment of long-range effects is much better than its description of short-range effects.

Moreover, although corrected rSCAN outperforms corrected PBE, correcting PBE GIPAW with rSCAN does not improve results. This supports the hypothesis that rSCAN is picking up more long-range contributions to shieldings than PBE, but does not improve local effects. Therefore, rSCAN benefits greatly from monomer correction, where a good treatment of this local electronic structure is incorporated. rSCAN's improved long range performance may, like its geometries, result from incorporation of dispersion effects. The rSCAN functional not only incorporates dispersion into its energies, as do semi-empirical dispersion correction schemes, but also into its self-consistent electron densities, from which shieldings are calculated.

2.2.1 Gauge invariance corrections. Corrections with mGGAs were calculated using both the Maximoff–Scuseria and Dobson kinetic energy densities, $\tau_{MS}(\mathbf{r})$ and $\tau_D(\mathbf{r})$ respectively, and in most cases, subtraction of rSCAN shieldings using

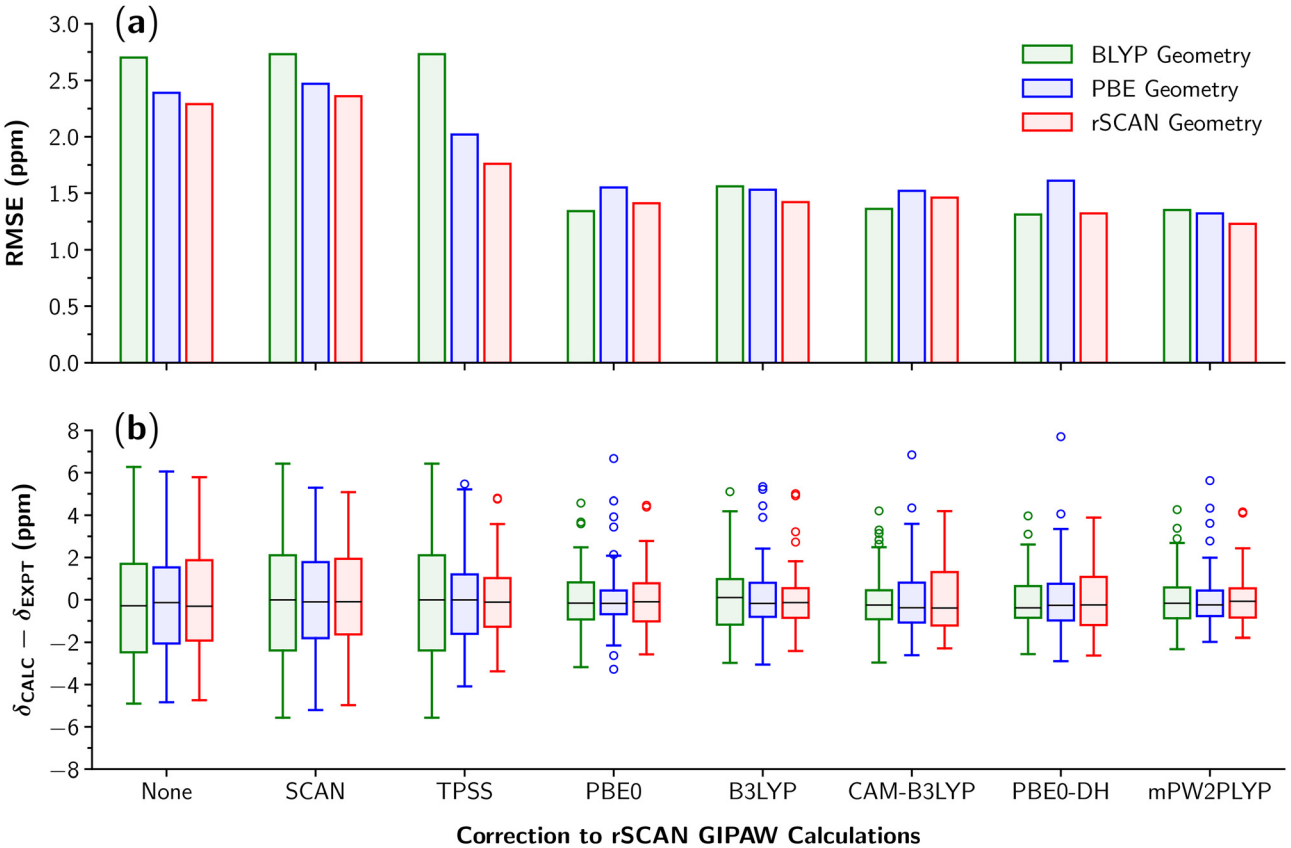

Fig. 4 (a) Root-mean-square errors and (b) error distribution of various monomer correction approaches to the rSCAN periodic shielding calculations, on geometries optimised with BLYP, PBE and rSCAN. The corrections to rSCAN shown were calculated by subtracting rSCAN $\tau_{MS}(\mathbf{r})$ shieldings and adding mGGA $\tau_{D}(\mathbf{r})$ shieldings.

$\tau_{MS}(\mathbf{r})$ led to better results than subtraction of $\tau_{D}(\mathbf{r})$ shieldings. We considered whether there may be a benefit to "correcting" the treatment of gauge invariance in rSCAN. To test this, a monomer correction was applied to the rSCAN GIPAW shieldings with $\tau_{D}(\mathbf{r})$, by

$$\sigma = \sigma^{\text{crystal}}_{\text{rSCAN}} - \sigma^{\text{molecule}}_{\text{rSCAN–MS}} + \sigma^{\text{molecule}}_{\text{rSCAN–D}}. \quad (4)$$

The result was a notable improvement, bringing RMSE down from 2.3 to 2.1 ppm, the same as PBE GIPAW, and reducing IQR by 0.9 ppm, as shown in Table 2.

When correcting rSCAN GIPAW, a version of $\tau(\mathbf{r})$ must be chosen with which to calculate the molecular rSCAN shieldings. Neglecting the gauge invariance of $\tau(\mathbf{r})$ completely in the molecular calculations leads to very disordered results, presumably due to a higher impact of basis set incompleteness in molecular calculations compared to in plane wave calculations. Therefore, we must either choose to subtract shieldings calculated with either $\tau_{MS}(\mathbf{r})$ or $\tau_{D}(\mathbf{r})$. For the vast majority of—but not all—correction methods, subtracting $\tau_{MS}(\mathbf{r})$ shieldings improves results compared to subtracting $\tau_{D}(\mathbf{r})$ shieldings. This evidence, along with the improvement from "correcting" the rSCAN shieldings with eqn (4), suggests that the Maximoff–Scuseria scheme somehow resembles the neglect of gauge invariance in the periodic calculations more closely than the Dobson scheme. Therefore, subtraction of $\tau_{MS}(\mathbf{r})$ shieldings generally improves results, removing some unphysical treatment of gauge invariance.

Table 2 Error statistics from gauge corrections with functionals from the SCAN family on rSCAN||rSCAN calculations. Here, −MS + D indicates subtraction of rSCAN−MS and addition of DFA−D, such as in eqn (4). Similarly, ±MS indicates subtraction of rSCAN−MS and addition of DFA−MS, and ±D indicates subtraction of rSCAN−D and addition of DFA−D. All values in ppm

DFA	Method	RMSE	MAE	IQR	MAX
rSCAN	GIPAW	2.3	1.9	3.9	5.8
	−MS + D	2.1	1.7	3.0	5.7
r^{2}SCAN	±MS	2.3	1.9	3.9	5.8
	±D	2.2	1.8	3.8	5.8
	−MS + D	2.0	1.6	3.0	5.7
SCAN	±MS	2.4	1.9	3.7	5.1
	±D	2.5	2.0	3.6	5.3
	−MS + D	2.1	1.8	3.5	5.1

We also considered whether its regularisation might be limiting rSCAN. Further corrections with SCAN and r^{2}SCAN were calculated, either subtracting and adding shieldings from $\tau_{MS}(\mathbf{r})$ and $\tau_{D}(\mathbf{r})$, or by a gauge correction as in eqn (4). While

correcting rSCAN GIPAW with r^{2}SCAN–MS does not improve results, applying the gauge correction does slightly, suggesting there may be more accuracy to be gained by a GIPAW calculation of r^{2}SCAN shieldings within the Dobson scheme. The improvement from rSCAN–D to r^{2}SCAN–D is less substantial than the improvement from rSCAN GIPAW to rSCAN–D, indicating the exact constraints broken by the neglect of the gauge invariance of $\tau(\mathbf{r})$ in CASTEP may be more important for shielding calculations than the constraints broken in the rSCAN functional itself.

Despite the fact that the gauge invariance of $\tau(\mathbf{r})$ is less crucial in periodic calculations compared to molecular calculations, the consistent improvement from "correcting" the kinetic energy density of rSCAN motivates the implementation of the Dobson kinetic energy density, and the more constrained r^{2}SCAN functional, into the CASTEP code.

2.3. Recreating the crystal environment

Placing the isolated molecules in a vacuum for the correction calculations may limit accuracy. If the correction procedure is valid, the corrected shieldings based on a subsystem should converge to those from calculating higher-level GIPAW directly on the full crystal, as the subsystem increases in size to more closely resemble the full system.

The simplest step towards mimicking the full crystal environment may be embedding the subsystem molecule in a dielectric medium. This confines the electron density to the vicinity of the molecule, potentially enhancing the separability of long- and short-range effects, and therefore aligning effectively with the assumptions underlying monomer correction.

2.3.1. Implicit solvation models. The conductor-like polarisable continuum model (CPCM) describes the bulk solvent environment primarily by its dielectric constant, ε.[56] Corrections within a CPCM have previously been shown to improve PBE0 corrections,[57] but have not before been reported with rSCAN or double-hybrids. We calculated corrections to PBE and rSCAN GIPAW with PBE0, B3LYP and the best performing DFA so far, mPW2PLYP, by

$$\sigma = \sigma^{\text{crystal}}_{\text{(m)GGA}} - \sigma^{\text{molecule–CPCM}}_{\text{(m)GGA}} + \sigma^{\text{molecule–CPCM}}_{\text{high-level}}$$

where $\sigma^{\text{crystal}}_{\text{(m)GGA}}$ are the periodic shieldings, $\sigma^{\text{molecule–CPCM}}_{\text{(m)GGA}}$ are the lower-level shieldings of the dielectric-embedded molecule, and $\sigma^{\text{molecule–CPCM}}_{\text{high-level}}$ are the higher-level shieldings of the dielectric-embedded molecule. A CPCM of ethanol was used which has a moderate dielectric constant of ε = 24.3. Fig. S4 and S5 (ESI†) show the effect of monomer corrections with a CPCM to PBE and rSCAN GIPAW calculations, respectively.

Conventional hybrids do not benefit particularly in terms of RMSE, but their error distributions. A more substantial improvement in IQR and RMSE was observed in the double-hybrid mPW2PLYP, suggesting that a closer resemblance to the crystal structure environment lets this high-level electronic structure method perform closer to its full potential. For rSCAN GIPAW, mPW2PLYP–CPCM correction decreases RMSE by over 0.3 ppm compared to PBE0 correction, down to 1.1 ppm. The mean absolute error (MAE) of mPW2PLYP–CPCM correction is 0.7 ppm, and the distribution of errors is very tight, with IQR of 1.1 ppm. Apart from a few pathological environments discussed in Section 2.4, mPW2PLYP–CPCM provides extremely good agreement with experimental chemical shifts.

The highly parameterised solvation model based on electron density[58] was also tested, and gave a marginally higher RMSE across GIPAW calculations and geometries. To determine the dielectric-dependence of the corrected shieldings, corrections were calculated with mPW2PLYP in an infinite dielectric constant, and CPCMs based on benzene (ε = 2.3), acetone (ε = 20.7), water (ε = 80.4), and formamide (ε = 111.4). No significant difference was observed where $\varepsilon > 20$, but the benzene CPCM gave a 0.1 ppm higher RMSE. The choice of dielectric therefore seems unimportant, in agreement with previously reported results which found that dielectric-dependence is largely eliminated in the process of converted shieldings to chemical shifts.[59]

Spectra are often assigned sequentially for NMR crystallography, so the correct ordering of calculated peaks is vital. However, conventional GIPAW calculations with PBE risk predicting the incorrect assignment for closely spaced peaks. On the rSCAN geometry of uracil, PBE GIPAW and a simple PBE0 correction predict the wrong ordering of peaks, but Fig. 5 shows that mPW2PLYP–CPCM correction gives an unambiguous and correctly ordered spectrum.

Fig. 6 compares methods used to correct rSCAN shieldings, showing how improvements in RMSE at each stage can sum to reduce the GIPAW error by over 50%. In fact, while a full optimisation is calculated routinely, the benefit of this is small compared to using a molecular correction with PBE0, and is competitive with the improvement moving from PBE0 to mPW2PLYP–CPCM corrections. The same pattern is seen for the PBE shieldings, although the best results are found in mPW2PLYP–CPCM|rSCAN||rSCAN. We also calculated PBE0|PBE monomer corrections on the unoptimised structures directly, but this did not significantly reduce RMSE from a conventional GIPAW calculation. Some kind of geometry optimisation therefore remains necessary for predicting accurate chemical shifts for ambient temperature structures.

Considering the low cost of monomer corrections compared to full plane wave optimisations, the technique has clearly demonstrated its utility for NMR crystallography. All of the calculations with mPW2PLYP–CPCM took less than 15 minutes of computational time running on eight parallel processes, with the exception of L-ascorbic acid where the asymmetric unit consists of two molecules. Because of the formal $\mathcal{O}(N^5)$ scaling of MP2 with number of basis functions, care should be taken applying this correction approach to large molecules if computational time is limited.

2.3.2. Cluster corrections. We also calculated corrections using a small cluster of molecules. This has been shown to improve agreement previously,[57] but has not been reported with rSCAN. Clusters were extracted from the rSCAN geometries such that a central molecule was completely surrounded by duplicates. Corrections to PBE and rSCAN GIPAW were then applied with PBE0, rSCAN–D and r^{2}SCAN–D, with and without an ethanol CPCM by

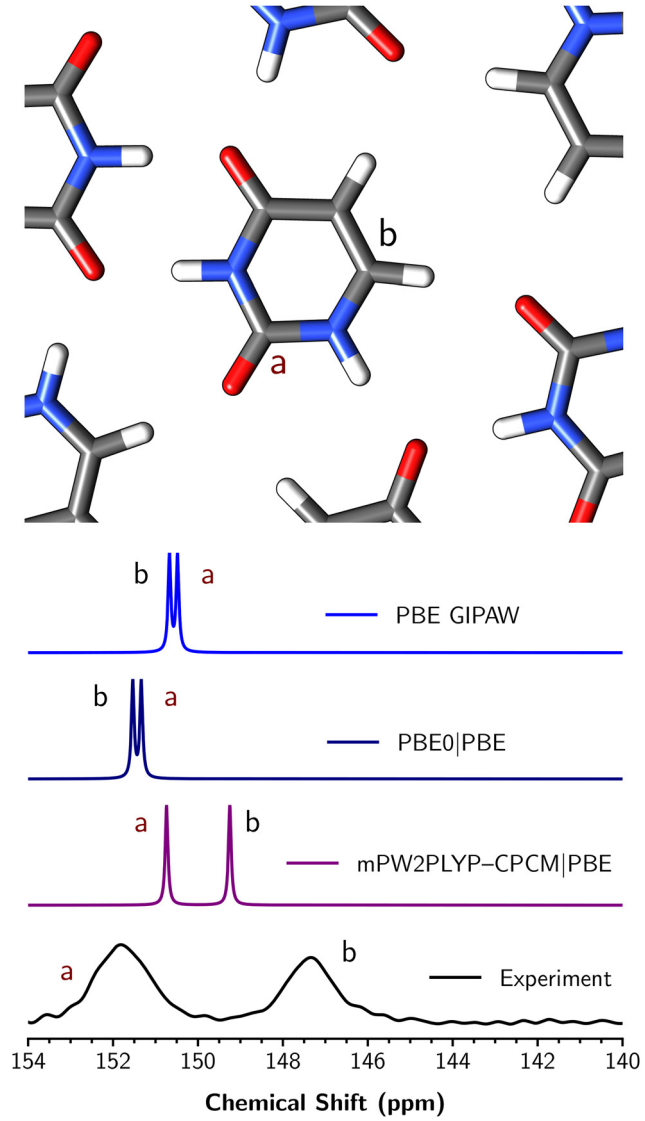

Fig. 5 Experimental ^{13}C spectrum of uracil, and theoretical spectra from PBE GIPAW on the rSCAN geometry and corrections with PBE0 and mPW2PLYP–CPCM. The experimental spectrum was recorded with 15 kHz magic-angle spinning and cross-polarisation on a spectrometer operating at 16.4 T. Simulated spectra were generated with Gaussian functions with full widths at half maximum of 0.1 ppm. The carbonyl peak, "a", should be observed at a higher chemical shift than the C–H peak, "b", around 152 ppm. A correction with mPW2PLYP–CPCM clearly predicts the correct ordering of these peaks, while PBE and PBE0 correction fail. The correct ordering of peaks was taken from previously reported solid-state NMR spectra of uracil.[55] and is confirmed by a dipolar dephasing experiment (spectrum shown in Fig. S19 of the ESI†), which shows that the lower chemical shift resonance corresponds to a C–H carbon.

$$\sigma = \sigma^{\text{crystal}}_{\text{(m)GGA}} - \sigma^{\text{cluster}}_{\text{(m)GGA}} + \sigma^{\text{cluster}}_{\text{high-level}}$$

In all cases, the CPCM marginally improved RMSE. Compared to the monomer corrections, these were very expensive, with the PBE0 cluster calculations taking several hours on eight parallel processes. The computational cost is highly dependent on the crystal geometry, as some clusters like α-D-glucose require many molecules to sufficiently "cover" the central

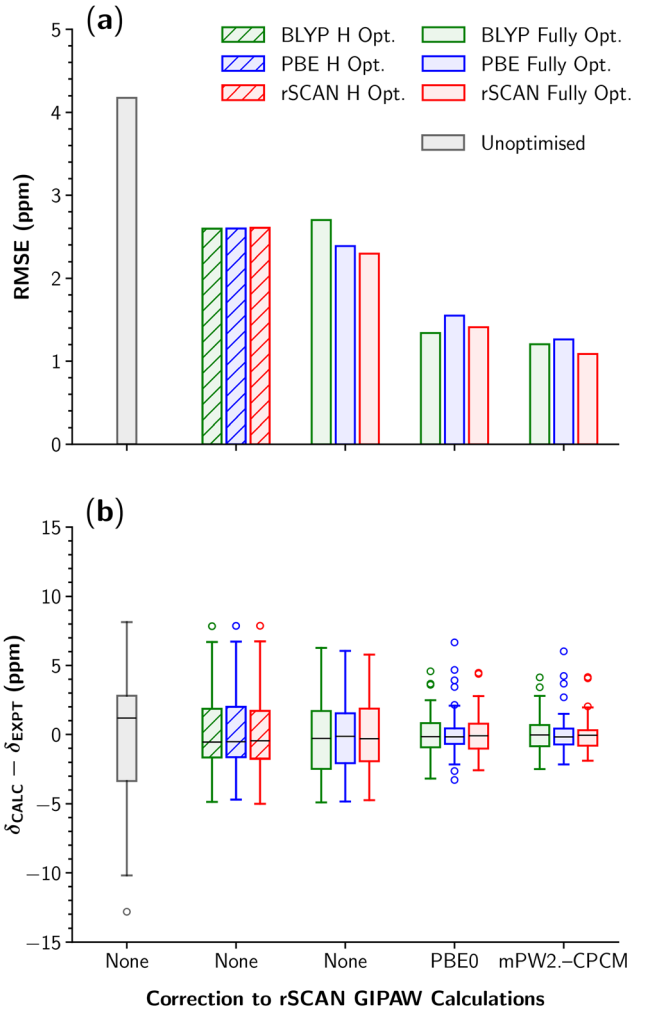

Fig. 6 (a) Root-mean-square errors and (b) box and whiskers plot of errors of performance of DFAs in optimisations and shielding calculations, from common approaches and mPW2PLYP–CPCM corrections, abbreviated as mPW2–CPCM. In the hashed bars, only hydrogen atoms were allowed to relax.

environments. The number of molecules in each cluster is shown in Table S8 (ESI†).

Table 3 shows that cluster corrections with PBE0 do not improve RMSE or MAE compared to monomer corrections, but increase maximum absolute error (MAX) and reduce IQR. This suggests that while most environments are predicted better, some outliers are predicted worse. There was no significant difference in the cluster gauge corrections compared to those on the isolated molecules. Chemical shifts from the PBE0 calculations directly on the cluster, without correcting the GIPAW calculations, were no better than PBE GIPAW, validating the molecular correction approach and highlighting how important long-range effects to shielding can be.

We considered that a remaining source of error may lie in the crystal geometries, which were only optimised with (m)GGA DFAs. Full optimisations on isolated molecules and clusters produced large deviations from the experimental

Table 3 Cluster and molecular corrections with PBE0, rSCAN–D and r^{2}SCAN–D to PBE and rSCAN GIPAW. Calculations were performed on the rSCAN geometries in an ethanol CPCM. In corrections to rSCAN, subtracted shieldings were calculated with $\tau_{MS}(\mathbf{r})$. All values in ppm

GIPAW	Correction	RMSE	MAE	IQR	MAX
PBE	None	2.1	1.6	2.6	4.7
	PBE0	1.5	1.1	2.0	4.3
	PBE0 cluster	1.5	1.1	1.4	4.9
rSCAN	None	2.3	1.9	3.9	5.7
	PBE0	1.4	1.1	1.5	4.5
	PBE0 cluster	1.5	1.0	1.3	5.1
	rSCAN–D	2.1	1.7	3.0	5.7
	rSCAN–D cluster	2.1	1.7	3.1	5.7
	r^{2}SCAN–D	2.0	1.6	3.5	5.7
	r^{2}SCAN–D cluster	2.1	1.7	3.1	5.7

structures, so constrained optimisations were performed on the rSCAN-optimised clusters. Here, the central molecule in each cluster was fully relaxed, but only hydrogen atoms in the other molecules were optimised. The optimisations were performed with PBE0–D4 with cc-pVTZ in an ethanol CPCM, before PBE0–CPCM shieldings were calculated.

When shieldings were corrected by

$$\sigma = \sigma_{\mathrm{PBE}}^{\mathrm{crystal}} - \sigma_{\mathrm{PBE-CPCM}}^{\mathrm{optimised\ cluster}} + \sigma_{\mathrm{PBE0-CPCM}}^{\mathrm{optimised\ cluster}}$$

an RMSE of 1.5 ppm was achieved, matching results from the cluster corrections without optimisation. Optimisation with PBE0–D4 therefore does not provide additional benefit, which may indicate the rSCAN geometry is already reliable. Residual sources of error which have been neglected in this investigation include nuclear quantum effects, the limited basis set, and influences from special relativity and thermal motion.[60] Another factor may be the lack of dependence on the current density in most exchange–correlation functionals.[61]

Overall, PBE0 cluster corrections do not come close to the 1.1 ppm RMSE from mPW2PLYP–CPCM monomer corrections, showing the importance of a high quality treatment of electronic structure in shielding calculations. These cluster corrections are also far more computationally expensive, so not recommended unless effects on chemical shifts from nearby molecules are thought to be particularly important.

2.4. Statistics and outliers

Previous studies on monomer corrections found larger reductions in RMSE, down to 1.1 ppm compared to 1.5 ppm here,[14,17] for PBE0|PBE||PBE corrections, suggesting the ^{13}C environments included in our crystal set are more challenging. Because our linear regressions were calculated by least-squares fitting, a pathological environment can skew the regression parameters, making the chemical shifts of all environments appear worse.

To identify the most challenging environments, the lowest absolute deviation from experiment of every environment was determined, across all correction methods tested. Of the 75 environments, only seven could never achieve an absolute deviation of less than half the average line width for our experimental spectra, 0.3 ppm. In four of these seven environments, the best method was a correction to the BLYP geometry, which generally performs poorly. Three of the environments correspond to amide C=O carbons, suggesting that these may be difficult for the electronic structure methods. When these seven outliers are excluded and chemical shifts are recalculated, the PBE0 correction matches the previously reported RMSE. The difference between PBE0 and mPW2PLYP correction decreases, suggesting the double-hybrid may be more useful in difficult cases, but the improvement in performance upon addition of a CPCM to mPW2PLYP remains clear, as MAE drops to an extremely low 0.6 ppm. Similar trends are observed for the other GIPAW calculations.

The charts in Section 2.2 show that chemical shifts are under-shielded for the vast majority of outliers, indicating the DFAs are overestimating the paramagnetic contribution in these environments. This contribution to shielding is reciprocal to the energy difference between the highest occupied and lowest unoccupied Kohn–Sham orbitals,[19] which is often underestimated by (m)GGAs.[62] In these cases, the poor performance of the GIPAW (m)GGA shieldings may not be recoverable with molecular corrections. An alternative explanation is that one of the three principal components of the shielding tensor, which are averaged to the isotropic shielding, is consistently poorly described.

To compare directly to previously reported results, we calculated chemical shifts with our best method, mPW2PLYP–CPCM|rSCAN|| rSCAN, on a previously used test set consisting of 132 ^{13}C environments, described in the Table S2 (ESI†).[17] On this test set, the best previously reported RMSE was 1.1 ppm, from PBE0|PBE||PBE. Our best correction method yields a substantially lower RMSE, of 0.8 ppm, and a MAE of 0.6 ppm. This is, to our knowledge, the best agreement between theoretical and experimental ^{13}C solid-state chemical shifts yet achieved on a large test set of structures. This confirms that our test set contains more challenging carbon environments. Since we found a small improvement by using double-hybrids for correction compared to conventional hybrids, the utility of the higher-level functionals is more pronounced for ''difficult'' cases.

2.4.1. Poor prediction in ascorbic acid. Despite the overall success of the monomer corrections, some molecules are still poorly described. The most concerning case by far is L-ascorbic acid, which achieves a minimum deviation of 2.3 ppm. The next worst ^{13}C environment has a 1.0 ppm minimum deviation, and is found in a pseudo-symmetry-related carbon in the L-ascorbic acid asymmetric unit, shown in Fig. 7. The best method for the 1.0 ppm deviation was a PBE0 calculation on the cluster directly, rather than a correction. After this, the best deviation is 2.7 ppm, substantially higher than the mean error of 2.1 ppm from a simple PBE GIPAW calculation.

Because these environments were so consistently badly described, we considered whether the geometry could be at fault. The structure used was from neutron diffraction at room temperature, so no poorly placed hydrogen atoms would be expected and thermal expansion is considered. The compound has also been reported to have no phase transition at ambient

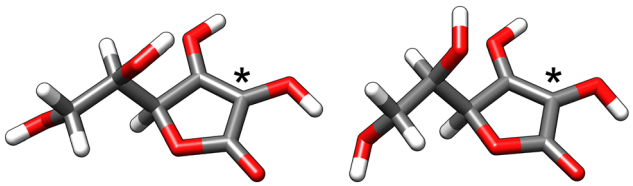

Fig. 7 Asymmetric unit of L-ascorbic acid from rSCAN-optimised geometry. The two molecules differ primarily by rotation of hydroxyl groups around the alkyl chain. Carbon atoms shown in grey, hydrogens in white and oxygens in red. The two challenging ^{13}C environments are highlighted with asterisks.

pressure under 440 K.[63] Periodic calculations with the very high plane wave energy cut-off of 1200 eV were performed, but gave no improvement. The compound is known to have a tautomer,[64] and we considered whether these could be affecting the chemical shift. A CASTEP optimisation was performed on the tautomer with rSCAN, but the final energy was around 100 kcal mol^{-1} higher than the original structure, giving it a negligible Boltzmann population at ambient temperature.

Despite how similar the molecules look in Fig. 7, the experimental difference in chemical shift between the ''difficult'' carbons was over 3 ppm. The main structural difference is in the overlying hydroxyl groups, but when these are removed from the structures, calculated chemical shifts still differ by around 3 ppm. We considered the environments could be very sensitive to the geometry, so a highly expensive B3LYP geometry optimisation in CASTEP was performed. This also produced no substantial benefit to the difficult environments.

Interestingly, L-ascorbic acid has the same conjugated, enone functionality as another organic crystal which has been reported as badly described even by monomer corrections, β-testosterone.[36] This might suggest that these shielding calculations could benefit from a method where multiple electronic configurations are explicitly considered, such as coupled-cluster theory. However, due to the asymmetric distribution of outliers, we suspect that the issues with these molecules result from an issue with the GIPAW calculations, rather than the corrections. The mPW2PLYP–CPCM corrections are generally so reliable that the breakdown in accuracy for these L-ascorbic acid environments may indict the underlying (m)GGAs, rather than the choice of corrections—the (m)GGA calculations seem to be describing these few environments so poorly that they are not able to be corrected.

3. Methodology

3.1. Experimental details

Molecular crystals for the test set were purchased from Fluorochem. Solid-state NMR experiments were performed on a Bruker Avance III HD Spectrometer operating at a magnetic field strength of 16.4 T. The ^{13}C chemical shifts were referenced to tetramethylsilane, using the methyl resonance of l-alanine at 20.5 ppm as a secondary reference.

Spectra were recorded at room temperature with a 2.5 mm probe at a magic-angle spinning rate of between 15 and 25 kHz. Cross-polarisation was used to transfer magnetisation from ^{1}H nuclei with contact time of between 1 and 5 ms. The cross-polarisation power was ramped linearly from 70% to 100%. ^{1}H heteronuclear decoupling using two-pulse phase modulation[65] with a pulse length of 4.8 μs and a radiofrequency field strength of 100 kHz was applied during acquisition. Spectra are the sum of between 112 and 1024 transients separated by a recycle interval of between 3 and 120 seconds. Fig. S6–S18 (ESI†) show the solid-state NMR spectra for our set of organic crystals. Structures were confirmed to be the polymorphs used for the theoretical calculations by comparing experimental powder XRD patterns to those calculated in Mercury.[66]

3.2. Computational details

3.2.1. Referencing chemical shifts. Density-functional theory calculations generate the absolute shielding tensor, $\boldsymbol{\sigma}$, in the crystal frame of reference. The isotropic shielding, σ, was calculated by $\frac{1}{3}\mathrm{Tr}\,\boldsymbol{\sigma}$, and is the average of the three principal components that arise from diagonalisation of the symmetric part of $\boldsymbol{\sigma}$.[19] The isotropic shieldings were then corrected as outlined in eqn (1).

To convert from these updated shieldings to isotropic chemical shifts, the widely used linear regression method was applied.[1] Here, the experimental spectra were sequentially assigned, matching the highest chemical shift with the lowest nuclear shielding calculated by PBE on the PBE geometries. A linear regression was then calculated by the least-squares algorithm in Microsoft Excel, fitting the calculated shieldings, σ_{calc}, and experimental shifts, δ_{exp}, to

$$\delta_{\mathrm{exp}} = \sigma_{\mathrm{ref}} - m\sigma_{\mathrm{calc}}. \quad (5)$$

This yielded the gradient, m, and reference shielding, σ_{ref}. The shieldings were then converted to calculated chemical shifts by

$$\delta_{\mathrm{calc}} = \sigma_{\mathrm{ref}} - m\sigma_{\mathrm{calc}}, \quad (6)$$

which incorporates a correction to the calculated shifts—regardless of the molecular correction approach—by allowing the value of m to deviate from one, limiting the effect of the deviation from a perfect correlation. As a result, the calculated shifts are more reliable than using a fixed reference shielding,[1] for instance from a DFT calculation on tetramethylsilane. Linear regressions were calculated separately for all methods tested, with regression parameters available in the Microsoft Excel document in the (ESI†).

3.2.2. Calculation of error statistics. Average agreement with experiment was evaluated by root-mean-square error and mean absolute error. The former is more dependent on outliers, so is more useful for assessing the reliability of a method to handle difficult situations, but can misrepresent the overall performance if there are a few very bad results.[1]

The RMSE is given by

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i}^{N}\left(\delta^{i}_{\mathrm{calc}} - \delta^{i}_{\mathrm{exp}}\right)^{2}}, \quad (7)$$

where N is the number of carbon environments. The MAE is another widely used error statistic, given by

$$\mathrm{MAE} = \frac{1}{N}\sum_{i}^{N}\left|\delta^{i}_{\mathrm{calc}} - \delta^{i}_{\mathrm{exp}}\right|. \quad (8)$$

This statistic is a simple mean, and does not account for the distribution of errors. In this study, the spread of the error distribution was quantified by the maximum absolute error and interquartile range. A small interquartile range indicates that shieldings are calculated consistently, but does not imply they are calculated consistently well, as the distribution may not be centered around an error of zero. Because of these limitations, a combination of error statistics was used to assess the utility of a method to calculate shieldings.

3.2.3. Periodic calculations. Plane wave DFT calculations were carried out with version 22.11 of the CASTEP code, incorporating the gauge-including projector-augmented waves algorithm. Core–valence interactions were described with ultrasoft pseudopotentials[67,68] for the pure exchange–correlation functionals PBE, BLYP and rSCAN, and with norm-conserving pseudopotentials for the hybrid functional B3LYP.[69] Calculations were performed on the Young UCL high-end computing cluster with 120 parallel processes *via* Intel Xeon Gold 6248 processors, with a clock speed of 2.50 GHz and 3–10 GB of memory per process. Crystallographic information files from the Cambridge Structural Database were converted to input geometry files for CASTEP with the ''cif2cell'' Python code.[70]

Several calculations on the l-alanine structure with plane wave kinetic energy cut-offs ranging from 200–1000 eV were performed with PBE and rSCAN to determine the convergence of the energy, forces and isotropic shieldings. For these calculations, a dense k-point spacing of 0.01 Å^{-1} was used *via* a Monkhorst–Pack grid. Calculations were then run at a 1000 eV plane wave energy cut-off with a k-point spacing ranging from 0.01–0.20 Å^{-1}. The parameters which converged shieldings to below 0.1 ppm and root-mean-square forces below 0.01 eV Å^{-1} were an energy cut-off of 800 eV and k-point spacing of 0.03 Å^{-1}. While only l-alanine was explicitly converged, these parameters are much stricter than those used in previous work on molecular corrections so were assumed to generalise to our other crystals.[14,17]

On diffraction-derived structures from the Cambridge Structural Database, geometry optimisations with PBE, BLYP and rSCAN with the converged energy cut-off and k-point spacing were performed. Optimisations with B3LYP were performed with an energy cut-off of 650 eV and a k-point spacing of 0.05 Å^{-1}. The unit cell parameters were fixed at their experimental values and the Broyden–Fletcher–Goldfarb–Shanno quasi-Newton algorithm[71] was used to relax the atoms in the unit cells. Optimisations were also performed where lattice parameters were allowed to relax with PBE, rSCAN and PBE–D3. Crystal structures were visualised using the VESTA program.[72]

On various geometries, nuclear shielding GIPAW calculations were performed with the PBE and rSCAN exchange–correlation functionals. All calculations in CASTEP were performed with an electronic energy convergence tolerance of 10^{-10} eV, a geometry convergence energy tolerance of 10^{-6} eV, and a fixed band occupancy. Increasing the geometry convergence tolerance has previously been shown to slightly increase agreement between experimental and calculated chemical shifts.[73] All other parameters used were CASTEP defaults, including the Koelling–Harmon relativistic treatment of pseudopotentials.[74]

3.2.4. Molecular calculations. Molecular calculations for corrections were performed with version 5.0.4 of ORCA. For basis set testing, cc-pVDZ, cc-pVTZ, cc-pVQZ, aug-cc-pVDZ, aug-cc-pVTZ, aug-cc-pwCVDZ, cc-pwCVDZ, cc-pwCVTZ, pcS-1, and pcS-2 were used. For all these basis sets, corrections were calculated to the PBE GIPAW shieldings with r^{2}SCAN, PBE0, and RI-MP2. The pcS-n and aug-cc-pwCVDZ basis sets were obtained from the basis set exchange.[75]

For the majority of the molecular corrections, the cc-pVTZ basis set was used. With this basis, molecular corrections to the PBE and rSCAN GIPAW calculations on all three sets of geometries were calculated with rSCAN, r^{2}SCAN, SCAN, TPSS, PBE0, B3LYP, CAM-B3LYP, PBE0-DH, B2PLYP, mPW2PLYP and RI-MP2. The PBE0, B3LYP and mPW2PLYP exchange–correlation functionals were also combined with the CPCM and SMD implicit solvation models. Corrections to the rSCAN GIPAW shieldings were calculated by subtracting the Maximoff–Scuseria and Dobson schemes for the gauge invariance of the kinetic energy density. Similarly, corrections with mGGA functionals to GIPAW shieldings were calculated by adding shieldings from the Maximoff–Scuseria and Dobson schemes.

For all calculations in ORCA, ''tight'' self-consistent field convergence criteria were used. For calculations involving RI-MP2 including all double-hybrid calculations, core electrons were correlated and the relaxed density used. The default approximations were used in the ORCA calculations, which have been shown to increase computational speed substantially while introducing minimal loss in accuracy.[46,76] The approximations include the RI-J approximation for the Coulomb interaction,[77] the chain of spheres approximation for Hartree–Fock exchange,[78] and the resolution of the identity approximation for MP2 and double-hybrid calculations.[46] For the RI-J approximation, the default auxiliary basis set by Weigend was used,[79] and for the RI-MP2 and double-hybrid calculations, the cc-pVTZ/C auxiliary basis set was used.

The coordinate scan calculation over the interfacial bond of the uracil dimer was performed with PBE, BLYP, rSCAN, r^{2}SCAN, SCAN, PBE0, B3LYP, CAM-B3LYP, B2PLYP and mPW2PLYP, in 0.1 Å increments from 3–5 Å. The calculations were performed in the aug-cc-pVTZ basis set in ORCA, with and without the D4 semi-empirical dispersion correction scheme. For the exact exchange corrections, exchange–correlation

functionals were defined based on PBE, SCAN and TPSS with the chain of spheres approximation, in line with the default implementation of hybrid functionals in ORCA.

Geometry optimisations of isolated molecules and molecular clusters were calculated with PBE0 and cc-pVTZ with a CPCM parameterised for ethanol and the D4 semi-empirical dispersion correction scheme. Corrections with geometry optimised structures were calculated both by subtracting the (m)GGA calculation on the unoptimised and optimised clusters. The corrections with implicit solvation models were determined with both the (m)GGA and high-level functional shieldings calculated in the same dielectric constant.

GaussView was used for visualisation of structures and to extract molecular geometries from the optimised crystal structures. For these extractions, the asymmetric unit was taken from the unit cell. This was always a single molecule, except in the case of the L-ascorbic acid dimer. For this structure, corrections were also calculated with the individual molecules rather than the asymmetric unit, but this did not improve RMSE.

Figures were rendered in version 1.14 of UCSF Chimera,[80] ChemDraw 21 and Matplotlib.[81] Simulated spectra were generated with a Python code, using Gaussian functions parameterised by full-width at half maximum.

4. Conclusions

At first glance of the initial data, neither periodic ^{13}C chemical shift calculations with rSCAN GIPAW nor monomer corrections with double-hybrids appear useful. However, more of rSCAN's potential is unlocked by monomer correction, and providing a more realistic structural model for a correction using mPW2PLYP with CPCM allows this (relatively) expensive electronic structure method to provide highly accurate results.

The very low RMSE of 0.8 ppm on a large test set of crystals, from monomer corrections with mPW2PLYP–CPCM to rSCAN GIPAW calculations on rSCAN-optimised geometries, shows how important a high quality treatment of electronic structure is for shielding calculations. For small to medium-sized molecules, these corrections take a few minutes of computational time, compared to several hours required for plane wave calculations. Moreover, the GIAO shieldings can be determined in parallel with the GIPAW calculation, and are much easier to perform than methods which do not provide as much benefit, such as sampling molecular dynamics simulations to incorporate finite temperature effects. The simplicity of these corrections and low cost for most molecules promises to increase confidence in NMR crystallography experiments and facilitate derivation of new structures.

The rSCAN functional consistently generates high quality geometries. Because of the higher geometry-dependence of mGGA functionals, rSCAN GIPAW should only be used on an rSCAN geometry. It is currently a better choice to use PBE than rSCAN for uncorrected GIPAW. However, when rSCAN's ability to account for long-range effects on shielding is combined with a high quality treatment of local electronic structure by monomer correction, the mGGA becomes the better starting choice. While the difference between corrected PBE and rSCAN is small, the surprising performance of corrected rSCAN strongly motivates the implementation of the more constrained r^{2}SCAN functional and the Dobson kinetic energy density into the CASTEP code.

Cluster corrections tighten the error distribution, and are worth considering when excellent predictions are required and computational time is not a constraint. However, the much higher cost means these cannot be recommended for routine use. The success of mPW2PLYP–CPCM indicates that more accuracy may be gained by a double-hybrid cluster correction. The poor scaling of RI-MP2 makes this intractable, but the asymptotically linear scaling "domain-based local pair natural orbital" (DLPNO) approximation[82,83] produces shieldings very similar to those with RI-MP2 for our isolated molecules. This method could be applied to the clusters. Another method to increase accuracy further could be within the geometries; DLPNO-double-hybrids could be used to optimise the central molecule in a cluster, before that structure is used to correct GIPAW shieldings.

Finally, our study should be expanded to the individual components of the shielding and electric field gradient tensors, comparing calculations to solid-state NMR powder patterns,[84–86] to test the success of mPW2PLYP–CPCM corrections more broadly. Likewise, this correction method should be applied to calculate chemical shifts for other relevant nuclei in NMR crystallography of molecular organics such as ^{15}N, ^{17}O and ^{1}H.

Author contributions

H. B.: conceptualization, investigation, formal analysis, writing – original draft, writing – review and editing. C. W. C.: investigation. J. M. G: conceptualization, formal analysis, writing – review and editing, funding acquisition, supervision. M. J. G. P.: conceptualization, formal analysis, writing – review and editing, funding acquisition, supervision.

Conflicts of interest

There are no conflicts of interest to declare.

Data availability

Research data supporting this publication can be accessed at **https://doi.org/10.17635/lancaster/researchdata/706**. The data supporting this article have been included as part of the ESI.†

Acknowledgements

We acknowledge support from the Collaborative Computational Project on NMR Crystallography (CCP-NC) funded by EPSRC (EP/T026642/1) and from the UK Materials and Molecular Modelling Hub (Young), which is partially funded by

EPSRC (EP/T022213/1, EP/W032260/1 and EP/P020194/1) for which access was obtained *via* the UKCP consortium and funded by EPSRC (EP/P022561/1). We also thank Dr Nathan R Halcovitch (Lancaster University) for helpful discussions. C. W. C. acknowledges the support of the Faraday Institution-funded FutureCat project (FIRG017).

References

1 P. Hodgkinson, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2020, **119**, 10–53.
2 A. Scarperi, G. Barcaro, A. Pajzderska, F. Martini, E. Carignani and M. Geppi, *Molecules*, 2021, **26**, 4577.
3 S. Neuberger, S. P. Culver, H. Eckert, W. G. Zeier and J. Schmedt auf der Günne, *Dalton Trans.*, 2018, **47**, 11691–11695.
4 A. L. Reviglio, F. A. Martínez, M. D. A. Montero, Y. Garro-Linck, G. A. Aucar, N. R. Sperandeo and G. A. Monti, *RSC Adv.*, 2021, **11**, 7644–7652.
5 T. Whewell, V. R. Seymour, K. Griffiths, N. R. Halcovitch, A. V. Desai, R. E. Morris, A. R. Armstrong and J. M. Griffin, *Magn. Reson. Chem.*, 2022, **60**, 489–503.
6 Y. Yasui, M. Tansho, K. Fujii, Y. Sakuda, A. Goto, S. Ohki, Y. Mogami, T. Iijima, S. Kobayashi, S. Kawaguchi, K. Osaka, K. Ikeda, T. Otomo and M. Yashima, *Nat. Commun.*, 2023, **14**, 2337.
7 D. A. Middleton, J. Griffin, M. Esmann and N. U. Fedosova, *RSC Adv.*, 2023, **13**, 34836–34846.
8 C. Shi, P. Fricke, L. Lin, V. Chevelkov, M. Wegstroth, K. Giller, S. Becker, M. Thanbichler and A. Lange, *Sci. Adv.*, 2015, **1**, e1501087.
9 P. Paluch, T. Pawlak, M. Oszajca, W. Lasocha and M. J. Potrzebowski, *Solid State Nucl. Magn. Reson.*, 2015, **65**, 2–11.
10 C. Pickard and F. Mauri, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2001, **63**, 245101.
11 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864–B871.
12 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133–A1138.
13 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
14 M. Dracínský, P. Unzueta and G. J. O. Beran, *Phys. Chem. Chem. Phys.*, 2019, **21**, 14992–15000.
15 M. Bursch, J.-M. Mewes, A. Hansen and S. Grimme, *Angew. Chem., Int. Ed.*, 2022, **61**, e202205735.
16 K. Wolinski, J. F. Hinton and P. Pulay, *J. Am. Chem. Soc.*, 1990, **112**, 8251–8260.
17 R. J. Iuliucci, J. D. Hartman and G. J. O. Beran, *J. Phys. Chem. A*, 2023, **127**, 2846–2858.
18 G. J. O. Beran, *Chem. Rev.*, 2016, **116**, 5567–5613.
19 J. C. Facelli, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **58**, 176–201.
20 J. E. Bates and F. Furche, *J. Chem. Phys.*, 2012, **137**, 164105.
21 S. N. Maximoff and G. E. Scuseria, *Chem. Phys. Lett.*, 2004, **390**, 408–412.
22 C. J. Schattenberg and M. Kaupp, *J. Chem. Theory Comput.*, 2021, **17**, 1469–1479.
23 J. F. Dobson, *J. Chem. Phys.*, 1993, **98**, 8870–8872.
24 R. Grotjahn, F. Furche and M. Kaupp, *J. Chem. Phys.*, 2022, **157**, 111102.
25 J. Sun, A. Ruzsinszky and J. P. Perdew, *Phys. Rev. Lett.*, 2015, **115**, 036402.
26 C. F. von Weizsäcker, *Z. Phys.*, 1935, **96**, 431–458.
27 J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew and J. Sun, *J. Chem. Phys. Lett.*, 2020, **11**, 8208–8215.
28 C. Adamo, M. Cossi and V. Barone, *J. Mol. Struct.*, 1999, **493**, 145–157.
29 P. M. Piaggi, A. Z. Panagiotopoulos, P. G. Debenedetti and R. Car, *J. Chem. Theory Comput.*, 2021, **17**, 3065–3077.
30 A. P. Bartók and J. R. Yates, *J. Chem. Phys.*, 2019, **150**, 161101.
31 M. Kothakonda, A. D. Kaplan, E. B. Isaacs, C. J. Bartel, J. W. Furness, J. Ning, C. Wolverton, J. P. Perdew and J. Sun, *ACS Mater. Au*, 2023, **3**, 102–111.
32 S. Ehlert, U. Huniar, J. Ning, J. W. Furness, J. Sun, A. D. Kaplan, J. P. Perdew and J. G. Brandenburg, *J. Chem. Phys.*, 2021, **154**, 061101.
33 S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson and M. C. Payne, *Z. Kristallogr.*, 2005, **220**, 567–570.
34 E. Chaloupecká, V. Tyrpekl, K. Bártová, Y. Nishiyama and M. Dracínský, *Solid State Nucl. Magn. Reson.*, 2024, **130**, 101921.
35 J. Blahut, J. R. Štocek, M. Šála and M. Dracínský, *J. Magn. Reson.*, 2022, **345**, 107334.
36 M. Dracínský, *Molecules*, 2021, **26**, 3857.
37 J. R. Yates and A. P. Bartók, *Faraday Discuss*, 2025, **255**, 192–202.
38 A. D. Becke, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1988, **38**, 3098–3100.
39 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785–789.
40 C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.*, 2016, **72**, 171–179.
41 H. J. Monkhorst and J. D. Pack, *Phys. Rev. B*, 1976, **13**, 5188–5192.
42 E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth and S. Grimme, *J. Chem. Phys.*, 2019, **150**, 154122.
43 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
44 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
45 F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2022, **12**, e1606.
46 F. Weigend, M. Häser, H. Patzelt and R. Ahlrichs, *Chem. Phys. Lett.*, 1998, **294**, 143–152.
47 R. A. Kendall, J. Dunning, H. Thom and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796–6806.
48 J. Dunning and H. Thom, *J. Chem. Phys.*, 1989, **90**, 1007–1023.
49 J. P. Perdew and K. Schmidt, *AIP Conf. Proc.*, 2001, **577**, 1–20.

50 J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.*, 2003, **91**, 146401.
51 T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51–57.
52 E. Brémond and C. Adamo, *J. Chem. Phys.*, 2011, **135**, 024106.
53 S. Grimme and F. Neese, *J. Chem. Phys.*, 2007, **127**, 154116.
54 T. Schwabe and S. Grimme, *Phys. Chem. Chem. Phys.*, 2006, **8**, 4398–4401.
55 S. K. Amini, H. Shaghaghi, A. D. Bain, A. Chabok and M. Tafazzoli, *Solid State Nucl. Magn. Reson.*, 2010, **37**, 13–20.
56 M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comput. Chem.*, 2003, **24**, 669–681.
57 J. D. Hartman and J. K. Harper, *Solid State Nucl. Magn. Reson.*, 2022, **122**, 101832.
58 A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B*, 2009, **113**, 6378–6396.
59 P. A. Unzueta and G. J. O. Beran, *J. Comput. Chem.*, 2020, **41**, 2251–2265.
60 E. A. Engel, V. Kapil and M. Ceriotti, *J. Phys. Chem. Lett.*, 2021, **12**, 7701–7707.
61 S. Reimann, U. Ekström, S. Stopkowicz, A. M. Teale, A. Borgoo and T. Helgaker, *Phys. Chem. Chem. Phys.*, 2015, **17**, 18834–18842.
62 Z. Wan, Q.-D. Wang, D. Liu and J. Liang, *Comput. Mater. Sci.*, 2021, **198**, 110699.
63 B. Nicolaï, M. Barrio, J. L. Tamarit, R. Céolin and I. B. Rietveld, *Eur. Phys. J.-Spec. Top.*, 2017, **226**, 905–912.
64 M. Eckert-Maksić, P. Bischof and Z. B. Maksić, *Croat. Chem. Acta*, 1985, **58**, 407–426.
65 A. E. Bennett, C. M. Rienstra, M. Auger, K. V. Lakshmi and R. G. Griffin, *J. Chem. Phys.*, 1995, **103**, 6951–6958.
66 C. F. Macrae, I. Sovago, S. J. Cottrell, P. T. A. Galek, P. McCabe, E. Pidcock, M. Platings, G. P. Shields, J. S. Stevens, M. Towler and P. A. Wood, *J. Appl. Crystallogr.*, 2020, **53**, 226–235.
67 D. Vanderbilt, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1990, **41**, 7892–7895.
68 J. R. Yates, C. J. Pickard and F. Mauri, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2007, **76**, 024401.
69 D. R. Hamann, M. Schlüter and C. Chiang, *Phys. Rev. Lett.*, 1979, **43**, 1494–1497.
70 T. Björkman, *Comput. Phys. Commun.*, 2011, **182**, 1183–1186.
71 B. G. Pfrommer, M. Cote, S. G. Louie and M. L. Cohen, *J. Comput. Phys.*, 1997, **131**, 233–240.
72 K. Momma and F. Izumi, *J. Appl. Crystallogr.*, 2008, **41**, 653–658.
73 H. E. Kerr, H. E. Mason, H. A. Sparkes and P. Hodgkinson, *CrystEngComm*, 2016, **18**, 6700–6707.
74 D. D. Koelling and B. N. Harmon, *J. Phys. C Solid State*, 1977, **10**, 3107.
75 B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibson and T. L. Windus, *J. Chem. Inf. Model.*, 2019, **59**, 4814–4820.
76 S. Kossmann and F. Neese, *Chem. Phys. Lett.*, 2009, **481**, 240–243.
77 F. Weigend, *Phys. Chem. Chem. Phys.*, 2002, **4**, 4285–4291.
78 F. Neese, F. Wennmohs, A. Hansen and U. Becker, *Chem. Phys.*, 2009, **356**, 98–109.
79 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
80 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.
81 J. D. Hunter, *Comput. Sci. Eng.*, 2007, **9**, 90–95.
82 G. L. Stoychev, A. A. Auer, J. Gauss and F. Neese, *J. Chem. Phys.*, 2021, **154**, 164110.
83 H. Neugebauer, P. Pinski, S. Grimme, F. Neese and M. Bursch, *J. Chem. Theory Comput.*, 2023, **19**, 7695–7703.
84 J. D. Hartman, L. E. Spock and J. K. Harper, *Magn. Reson. Chem.*, 2023, **61**, 253–267.
85 S. T. Holmes and C. Dybowski, *Solid State Nucl. Magn. Reson.*, 2015, **72**, 90–95.
86 J. D. Hartman and G. J. Beran, *Solid State Nucl. Magn. Reson.*, 2018, **96**, 10–18.