Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

Published on 08 September 2026
Licensed under CC-BY-NC 4.0

Check for updates

Cite this: DOI: 10.1039/d6sc04941a

All publication charges for this article have been paid for by the Royal Society of Chemistry

Accessible hybrid DFT-quality NMR crystallography *via* gas-phase machine learning interatomic potentials

Shubha S. Gunaga,[a] Robert W. Schurko,[ab] Sean T. Holmes[ab] and Frederic Mentink-Vigier[*ab]

Received 12th June 2026
Accepted 6th September 2026

DOI: 10.1039/d6sc04941a

rsc.li/chemical-science

Nuclear magnetic resonance (NMR) crystallography is a robust method for structure determination, but its reliance on density functional theory (DFT) calculations for geometry refinement limits its speed and accessibility. Recent machine-learning predictors such as ShiftML3 can evaluate magnetic shieldings in seconds, but still need high-quality crystal geometries that are normally obtained from slow DFT optimisations. Here, we demonstrate that machine learning interatomic potentials (MLIPs) can eliminate this bottleneck for organic crystals. Interestingly, models trained on gas-phase molecules at the hybrid-DFT ωB97M-V level (OMol25 dataset)—such as UMA-omol and MACE-Polar-1—deliver structural quality rivaling hybrid periodic-DFT, without requiring large computational resources. The resulting MLIP/ShiftML3 workflow reduces computational costs by at least three orders of magnitude, making high-accuracy NMR crystallography accessible without HPC infrastructure, and opening the door to fast molecular dynamics simulations at the hybrid DFT level. Combined with sensitivity enhancement *via* dynamic nuclear polarization (DNP), we demonstrate the approach on L-histidine, extracting ^{13}C and ^{15}N chemical shift tensors and using ^{1}H chemical shifts at natural isotopic abundance to discriminate between its monoclinic and orthorhombic polymorphs.

Introduction

The determination of the structures of molecular crystals is both critical and challenging, notably for the pharmaceutical industry. Variations in crystal structure, including those associated with polymorphism, as well as the formation of multi-component solid forms such as salts, solvatomorphs, and cocrystals, directly impact vital physicochemical properties like stability, solubility, and bioavailability.[1–5] While X-ray diffraction is a powerful and frequently used characterization technique, its application is often limited by issues such as the inability to obtain high-quality crystals, intrinsic crystallographic disorder, and difficulty in precisely locating hydrogen atoms. Fully understanding structure–function relationships is one avenue for the rational design of new pharmaceutical solid forms with tailored physicochemical properties. Therefore, reliable identification of the correct structure of a pharmaceutical solid form is an essential step during both the drug screening and formulation processes.

In this context, nuclear magnetic resonance (NMR) spectroscopy is a very powerful method for characterizing organic solid forms, including pharmaceuticals and their respective polymorphs, salts, cocrystals, *etc.* A key feature of NMR is that it provides site-specific chemical data, providing insight into structure and bonding. Solid-state NMR experiments thus allow one to obtain high-resolution spectra of many of the NMR-active isotopes contained in a pharmaceutical. Most often, spin-$\frac{1}{2}$ nuclides such as ^{1}H, ^{13}C, and ^{15}N are used;[6–12] however, quadrupolar nuclides (with a spin greater than 1/2) such as ^{2}H, ^{14}N, ^{17}O, and ^{35}Cl provide additional complementary information.[13–17] This is all the more true with the sensitivity-enhanced method called magic angle spinning dynamic nuclear polarization (MAS-DNP). This technique not only enables quick one-dimensional experiments, but also full spectral assignment *via* multidimensional NMR at natural isotopic abundance,[18–24] and has been used to identify polymorphs in drug formulations.[23,25,26]

NMR crystallography (*i.e.*, determining crystal structures using NMR) is further hindered by the absence of straightforward relationships between NMR observables—such as chemical shifts, electric field gradients, and *J*-coupling constants—and crystal structure. Chemical shifts, and their anisotropic properties in particular, are very sensitive to molecular-level structure and bonding, including intermolecular noncovalent interactions such as hydrogen bonding, which produce readily detectable variations between materials.[27–29] Correlating the observed chemical shifts with the crystal structure often

[a]National High Magnetic Field Laboratory, Florida State University, 1800 E. Paul Dirac Dr, Tallahassee, FL, 32310, USA. E-mail: fmentink@fsu.edu

[b]Department of Chemistry and Biochemistry, Florida State University, Tallahassee, FL 32306, USA

© 2026 The Author(s). Published by the Royal Society of Chemistry

requires calculating the expected chemical shifts for a given structure using *ab initio* methods. Due to the periodic nature of crystals, this is often done with periodic density functional theory calculations (DFT). This process starts with the geometry optimization of each candidate crystal structure, followed by the calculation of the magnetic shieldings (also referred to as chemical shieldings), commonly *via* the gauge-including projector augmented-wave (GIPAW)[30,31] method. The magnetic shieldings are then converted to the chemical shift scale using linear regression and compared with experiment as a figure of merit. The accuracy of this approach is heavily dependent on the DFT functional used for both the geometry optimization and the prediction of magnetic shielding,[32,33] with both steps being time consuming due to the complexities of the tasks.

Recently, NMR crystallography has been accelerated by orders of magnitude through the use of machine learning to compute magnetic shieldings.[34–38] In this context, ShiftML[34] and its latest implementation, ShiftML3,[38] are promising for organic solids. Integration of ShiftML3 removes one major component of the NMR crystallography bottleneck, significantly reducing the computational cost; however, ShiftML3 still requires high quality structural models as inputs to predict magnetic shieldings accurately. Therefore, replacing DFT geometry optimization with an accurate low-cost refinement step is the remaining key step toward a fully accelerated NMR crystallography workflow.

The development of machine learning interatomic potentials (MLIPs) offers another promising avenue for NMR crystallography. MLIPs have recently shown potential to replace DFT in geometry optimizations,[39] as well as in more complex and larger-scale tasks,[40–42] for organic and inorganic systems. Yet, MLIPs can be trained on numerically noisy datasets,[43] which may impair their accuracy, and can sometimes fail at "simple" tasks.[44] Since magnetic shieldings are highly sensitive to subtle variations in molecular-level structure, it remains unclear how well MLIPs can refine crystal structures to the degree that subsequent calculations on them yield accurate chemical shifts. Recently, a number of MLIPs (*e.g.* Egret-1, Orb-v3, AIMNet2, MACE-OFF, PET-MOLS, MACE-Polar-1 (ref. 45–50)) have been introduced for organic molecular crystals. Two particular MLIPs trained on the OMol25 dataset, Universal Model for Atoms (UMA)[51] and MACE-Polar-1,[50] have attracted our attention as these aim at being versatile, *i.e.* applicable to molecules, crystals or surfaces, and rely on the FAIR Chemistry datasets, one of the largest and most accurate training datasets available.[43,52–56] In recent tests, UMA and MACE-polar-1 have performed admirably on molecular crystals, rivalling PBE-MBD,[57] r^{2}SCAN-3c,[58] and ωB97M-D3(BJ),[50] making them particularly interesting for crystal structure prediction (CSP) applications.[57,59,60] While these studies establish transferability at the level of energetics and bulk structural parameters, such benchmarks integrate over many atomic interactions and are relatively insensitive to local geometric errors. NMR chemical shifts—and chemical shift tensors in particular—provide a far more stringent test: sub-ångström variations in bond lengths or hydrogen bond geometries, which are invisible to most structural benchmarks, produce quantifiable shifts of several ppm, particularly for ^{15}N. Validating MLIP transferability through NMR observables therefore constitutes a very sensitive probe of whether gas-phase structural information genuinely captures the local atomic environments of the crystalline state. Beyond accuracy, accessibility matters: hybrid periodic-DFT geometry optimization depends on specialist codes and carries computational demands incompatible with routine or high-throughput work. Demonstrating that freely available gas-phase MLIPs can substitute for this step would therefore represent a practical advance for the broader NMR crystallography community, not only a computational one.

In this work, we systematically benchmark two UMA tasks—omc (trained at the PBE-D3 level on molecular crystals) and omol (trained at the hybrid ωB97M-V level on gas-phase molecules), alongside MACE-Polar-1 which is also trained at the ωB97M-V level and includes long-range electrostatic interactions—evaluating their accuracy for NMR crystallography and the degree to which high-level structural information transfers from the gas phase to the condensed phase. We first demonstrate that UMA/ShiftML3 achieves good overall agreement with the conventional RPBE-D2/RPBE[61,62] workflow across 65 organic structures, at about three orders of magnitude lower cost. We then use the benchmark set of Ramos *et al.*[63] to disentangle the contributions of geometry quality and shielding method, identify the optimal models and tasks for accurate chemical shift predictions, and show that OMol25-trained models yield structural refinements rivalling hybrid-DFT. Finally, the approach is extended to chemical shift tensors across all 65 structures, and combined with MAS-DNP to address a challenging application: distinguishing the orthorhombic and monoclinic polymorphs of L-histidine.[8]

Results and discussion

DFT optimization *vs.* UMA for magnetic shieldings

This section evaluates the use of UMA[51] for structural optimization and its performance for the prediction of isotropic chemical shifts when combined with ShiftML3.[38] Three sets of structures were generated from the 65 structures determined initially by X-ray diffraction: one with no optimization, one with plane-wave DFT optimization, and one with UMA optimization. Magnetic shieldings for each set were then calculated using either DFT GIPAW[30,31] (RPBE[62,64]) or ShiftML3, resulting in six datasets. These are labelled by the format "Optimization Method/Magnetic Shielding Method" (*e.g.*, UMA/ShiftML3, RPBE-D2 (ref. 65 and 66)/RPBE, *etc.*) as illustrated in Fig. 1.

The isotropic shieldings for each set are compared against the conventional RPBE-D2/RPBE reference with results presented in Fig. 2. Focusing first on the shieldings calculated *via* DFT GIPAW, Fig. 2(a) represents the correlation between values as derived from non-optimized X-ray structures *versus* those from DFT-optimized structures. A significant lack of agreement is observed mainly in the slope, which deviates from unity (*ca.* 0.9). This is a well-known consequence of experimental uncertainties in the atomic positions within X-ray structures, particularly the systematic underestimation of X–H bond lengths.[67] This systematic error/uncertainty significantly diminishes the

© 2026 The Author(s). Published by the Royal Society of Chemistry

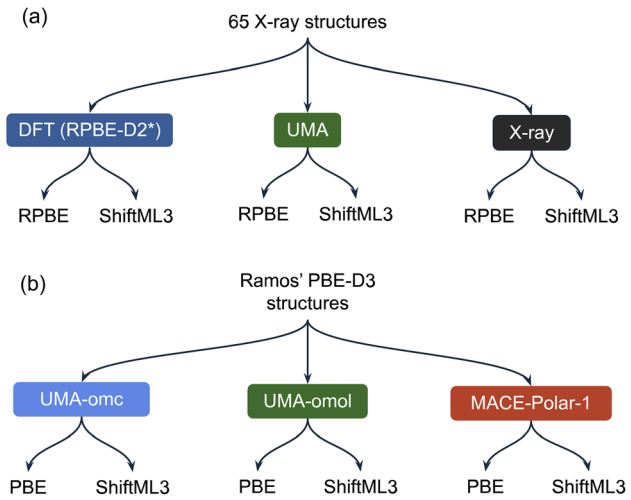


Fig. 1 Workflow used for the analysis of the ML tools discussed in this article. (a) Optimization of the 65 X-ray crystal structures with either DFT or UMA, or no optimization, followed by calculations of magnetic shielding tensors either *via* RPBE or ShiftML3. (b) For the set of Ramos *et al.*,[63] optimization of the reported PBE-D3 structures with either UMA (tasks omol or omc) or MACE-Polar-1, followed by calculations of magnetic shieldings with either PBE or ShiftML3.

accuracy of DFT predictions. In Fig. 2(b), the UMA/RPBE results are compared against the RPBE-D2/RPBE reference. Here, the agreement is substantially improved. The correlation slope deviates modestly from unity, and RPBE or UMA geometry optimizations appear to yield different results for some structures, notably uracil. These structural differences are sufficient to be reflected as differences in the GIPAW-calculated magnetic shieldings.

We next compare the results of the RPBE-D2/RPBE method with those of X-ray/ShiftML3 (Fig. 2(c)) and UMA/ShiftML3 (Fig. 2(d)). As expected, the agreement between RPBE-D2/RPBE and X-ray/ShiftML3 is poor. The comparison of RPBE-D2/RPBE with UMA/ShiftML3 is much more interesting because the agreement between the two protocols is again excellent, apart from the same few structures and the offset resulting from differences between the PBE (used to train ShiftML3) and RPBE functionals. Similar conclusions are drawn from ^{15}N, and to remain concise, these results are reported in the SI.

The excellent agreement is very promising and already demonstrates that UMA/ShiftML3 can effectively compete with the RPBE-D2/RPBE workflow, particularly given the dramatic difference in computational cost. In practice, optimizing all 65 structures and computing their magnetic shieldings with UMA/ShiftML3 required only *ca.* 10 minutes on a consumer-grade GPU, whereas the corresponding RPBE-D2/RPBE calculations required 2–3 weeks on a local HPC workstation. There are nonetheless outliers evident in Fig. 2(d), which could either

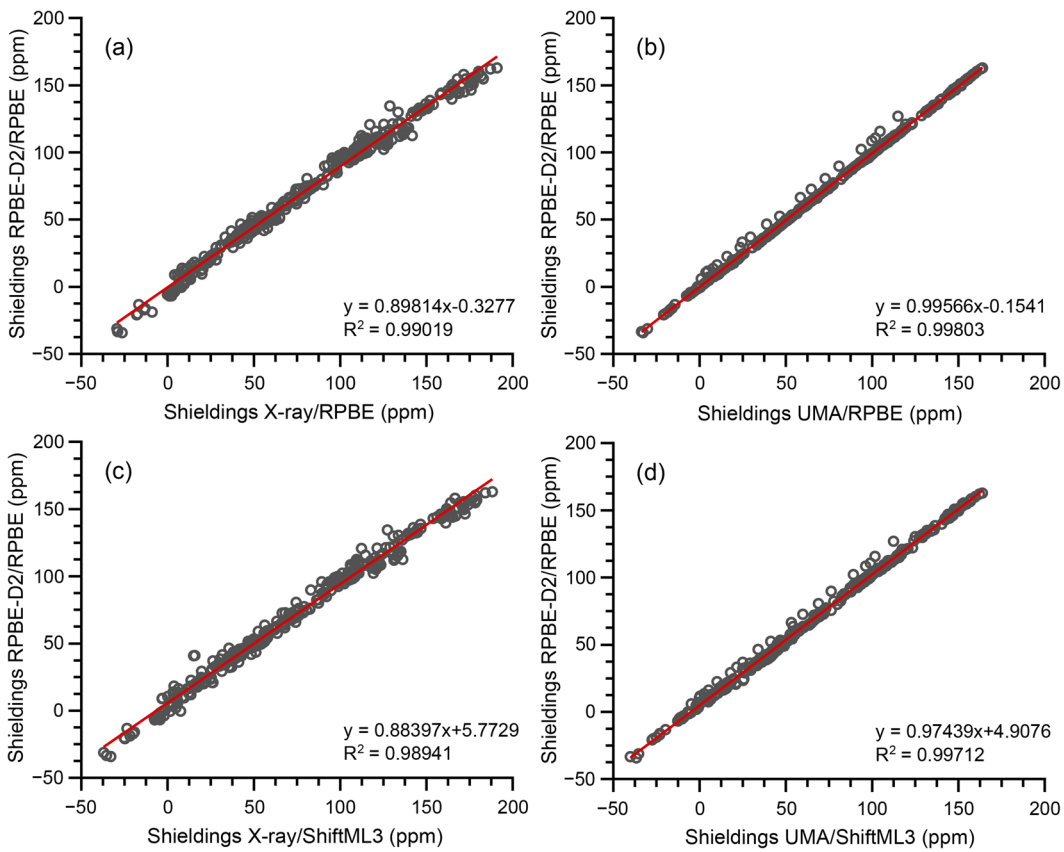


Fig. 2 Comparison between the ^{13}C isotropic magnetic shieldings obtained *via* (a) X-ray/RPBE *vs.* RPBE-D2/RPBE, (b) UMA/RPBE *vs.* RPBE-D2/RPBE, (c) X-ray/ShiftML3 *vs.* RPBE-D2/RPBE, (d) UMA/ShiftML3 *vs.* RPBE-D2/RPBE. UMA was used with the task omol and the uma-s-1p1 model.

© 2026 The Author(s). Published by the Royal Society of Chemistry

indicate a strength or a weakness of UMA/ShiftML3. This is further evaluated in the next section.

Influence of MLIP training set for magnetic shieldings with ShiftML3

UMA is available in three model sizes: two small variants (uma-s-1p1 and uma-s-1p2) and one medium variant (uma-m-1p1). These models have been trained on a wide range of datasets,[51] including OMC25,53 OMol25.[52] The uma-s-1p2 model includes OPoly26 as well.[56] UMA uses tasks to select the simulation type one wants to use. For molecular systems, two tasks are available: omc (Open Molecular Crystals) and omol (Open Molecules). The omc task emulates the PBE-D3 (ref. 61 and 64) level of theory—equivalent to the OMC25 benchmark set[53]—and is specifically designed for periodic crystal calculations, capturing the long-range many-body dispersion interactions that are essential for accurate crystal packing. The omol task targets gas-phase molecules, delivering energies and forces at the ωB97M-V/def2-TZVPD[68–70] level (a range-separated hybrid functional with dispersion correction and a triple-ζ basis set), which corresponds to the OMol25 benchmark set.[52] This combination of DFT functionnal and basis set is one of the most accurate DFT approaches currently available for both energies and geometries. Similarly, MACE-Polar-1 is trained on OMol25 but explicitly includes long range electrostatic interactions that appear to affect the quality of its predictions.[50] MACE-Polar-1 comes with three models: small, medium, and large, which offers tradeoff between accuracy and speed.

DFT methods for periodic solid and molecular systems have evolved into two distinct sub-fields, often utilizing different simulation packages and approximations. Intuitively, the omc task appears most appropriate for structural refinement in NMR crystallography, as its training set relies on periodic structures. However, a defining feature of modern MLIPs is their remarkable transferability between systems. Recent studies have demonstrated that potentials trained on molecular data are transferable to crystalline solids without loss of accuracy.[45,47,50,52,71,72] In the hierarchy of Jacob's Ladder, the OMol25 set—calculated at the ωB97M-V range-separated hybrid level—occupies a higher rung of chemical accuracy than the GGA-level PBE functional for OMC25 utilized in omc. This suggests that omol and MACE-Polar-1 may provide superior structural fidelity, a hypothesis supported by recent performance benchmarks.[73] Given that previous studies by Ramos *et al.*[63] and by Dracínský[29] have both shown that use of hybrid-level geometries can improve PBE GIPAW magnetic shielding predictions by approximately 20%, we evaluated the models trained on both OMC25 and OMol25 in combination with PBE GIPAW and ShiftML3, which is trained on PBE GIPAW magnetic shieldings. Using the benchmark set established by Ramos *et al.*,[63] we performed a comparative analysis of these two MLIP tasks and their respective agreement with experimental ^{13}C and ^{15}N chemical shifts (Fig. 3 and Table 1).

Fig. 3 reports the root mean square error (RMSE) and mean absolute error (MAE) for the Ramos *et al.* benchmark set, comparing experimental ^{13}C and ^{15}N chemical shifts with values calculated using ShiftML3 or PBE GIPAW on geometries refined by different UMA or MACE-Polar-1 models. Table 1 summarizes for the best models, their linear regression and error statistics. Common to all methods is that the linear regression slopes are consistently close to −1 (range: −0.997 to −1.023 for ^{13}C; Table S1). This near-unity scaling confirms that the calculated shielding scale is well-calibrated regardless of how the geometry was obtained, so the remaining discussion focuses on the residual errors.

As a reference point, we first applied ShiftML3 to structures optimized at the PBE-D3(BJ) and PBE0-D3(BJ) levels. On PBE-

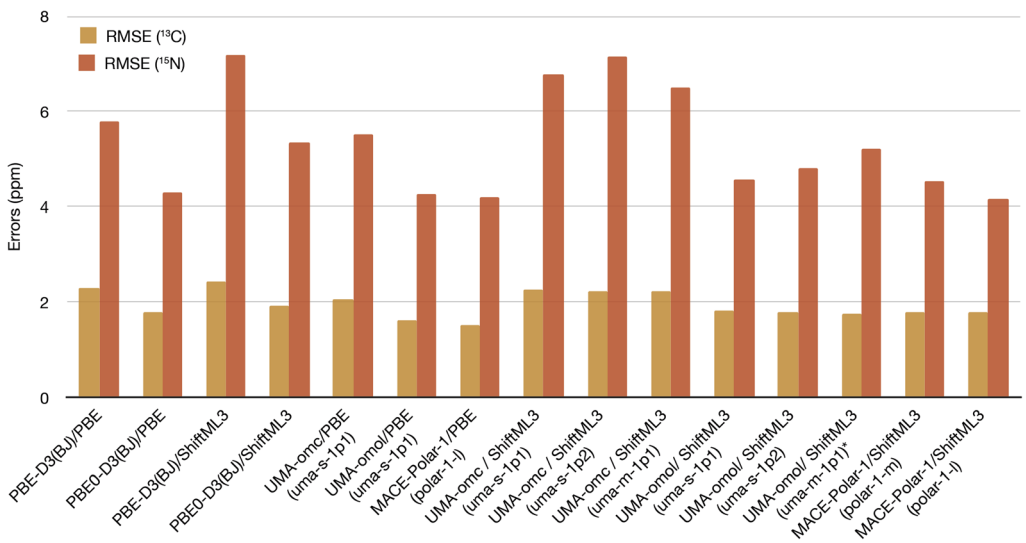


Fig. 3 RMSE values (ppm) on the ^{13}C and ^{15}N isotropic magnetic shielding plotted against the crystal-optimization workflow (GGA or hybrid DFT) and the MLIP model used (UMA-small, UMA-medium, MACE-Polar-1 medium/large). Shielding values were obtained either with PBE GIPAW or with ShiftML3. For PBE-D3(BJ)/PBE PBE0-D3(BJ)/PBE, RMSE values were reported in ref. 63. Note that UMA-omol with uma-m-1p1 did not fully converge all the structures and is thus unreliable.

© 2026 The Author(s). Published by the Royal Society of Chemistry

Table 1 Fitted linear regression parameters and comparison of the ShiftML3 and PBE magnetic shieldings relative to the experimental chemical shifts for the data set of 132 ^{13}C and 35 ^{15}N chemical shifts for crystals optimized by different methods. The root-mean-square error (RMSE), mean absolute error (MAE), and maximum absolute error (Max AE) are shown. Note, for ^{15}N we re-referenced the data so that NH_3(l) is used as the 0 ppm reference

	^{13}C					^{15}N				
	Regression		Error statistics (ppm)			Regression		Error statistics (ppm)		
Method	Slope	Intercept	RMSE	MAE	Max AE	Slope	Intercept	RMSE	MAE	Max AE
DFT benchmarks (GIPAW-PBE shieldings)										
PBE-D3(BJ)/ShiftML3	−1.024	170.47	2.44	1.88	12.49	−0.968	225.47	7.19	5.72	22.73
PBE0-D3(BJ)/ShiftML3	−1.019	172.85	1.92	1.41	10.65	−0.996	225.65	5.34	3.86	17.02
MLIP geometry—OMC25 training set (PBE-D3 level)										
UMA-omc/PBE (uma-s-1p1)	−1.011	170.54	2.08	1.67	6.75	−0.987	219.10	5.52	4.33	12.54
UMA-omc/ShiftML3 (uma-s-1p1)	−1.023	170.70	2.27	1.72	12.18	−0.980	219.33	6.79	5.37	18.57
MLIP geometry—OMol25 training set (ωB97M-V hybrid level)										
UMA-omol/PBE (uma-s-1p1)	−0.999	171.94	1.61	1.19	4.33	−1.012	187.65	4.28	2.73	17.17
UMA-omol/ShiftML3 (uma-s-1p1)	−1.012	171.89	1.84	1.33	10.30	−1.007	227.46	4.56	3.30	17.89
MACE-polar-1/PBE (polar-1-l)	−1.000	172.07	1.51	1.14	4.17	−1.012	226.57	4.20	2.94	14.93
MACE-polar-1/ShiftML3 (polar-1-l)	−1.013	172.03	1.78	1.28	9.97	−1.009	227.04	4.16	2.99	15.76

D3(BJ) geometries, ShiftML3 yields RMSEs of 2.4 ppm (^{13}C) and 7.2 ppm (^{15}N), in close agreement with the GIPAW-PBE values of 2.3 and 5.8 ppm reported by Ramos *et al.* Applying it to PBE0-D3(BJ) geometries reduces these errors by 20–30% to 1.9 ppm (^{13}C) and 5.3 ppm (^{15}N), again mirroring the improvement seen with GIPAW-PBE. The agreement between ShiftML3 and GIPAW-PBE across both geometry levels establishes that ShiftML3 faithfully captures the sensitivity of NMR shieldings to structural quality. The maximum absolute errors for both DFT levels are dominated by two structures—sucrose (SUCROS04) and the salt-cocrystal *N*-(pyridoxylidenium)tolylamine 2-nitrobenzoate, GEHHEH. It is worth noting that the GEHHEH structure is challenging for GIPAW-PBE as well.

Turning to MLIP-optimized geometries, structures refined with UMA-omc—trained on periodic crystals at the PBE-D3 level—yield GIPAW-PBE RMSEs of 2.1 ppm (^{13}C) and 5.5 ppm (^{15}N), essentially indistinguishable from the PBE-D3(BJ) DFT benchmark. This is an expected result: UMA-omc produces structures of equivalent quality to the DFT method it was trained on. Replacing GIPAW-PBE with ShiftML3 on the same geometries preserves this accuracy (RMSE 2.3 ppm, ^{13}C, 6.8 ppm, ^{15}N), establishing a fully ML-based UMA-omc/ShiftML3 workflow that matches the conventional PBE-D3/GIPAW pipeline at approximately three orders of magnitude lower computational cost.

The more striking result comes from models trained on the gas-phase OMol25 dataset. Structures optimized with UMA-omol or MACE-Polar-1 yield GIPAW-PBE RMSEs of *ca.* 1.5–1.6 ppm (^{13}C) and *ca.* 4.2–4.3 ppm (^{15}N), matching or slightly surpassing the PBE0-D3(BJ) benchmark, with ShiftML3 following closely at *ca.* 1.8 ppm (^{13}C) and *ca.* 4.2–4.6 ppm (^{15}N). Reaching PBE0-level accuracy matters because hybrid periodic-DFT calculations are substantially more expensive than their GGA counterparts and scale poorly to larger systems (typically $O(N^3)$ *vs.* $O(N^4)$)—in practice, they are incompatible with screening workflows. The OMol25-trained MLIPs deliver equivalent structural quality at orders-of-magnitude lower cost, effectively removing a barrier that has long restricted high-accuracy NMR crystallography. That this accuracy is achieved by models trained exclusively on isolated gas-phase molecules is a non-trivial finding. NMR is very sensitive to geometries, and such accuracy proves the reliability of the approach. It demonstrates that the hybrid ωB97M-V level structural information learned from molecular data transfers remarkably well to periodic crystalline environments, without any explicit treatment of crystal packing during training. The close agreement between UMA-omol and MACE-Polar-1—despite the latter incorporating explicit long-range electrostatics—suggests that for the organic systems studied here, the local atomic environment within UMA's embedding radius (6 Å for uma-s-1p1) captures the structurally relevant interactions. Within the UMA, model size has a marginal effect as shown in Fig. 3: uma-s-1p1 performs best overall, then uma-s-1p2 followed by uma-m-1p1, which failed to converge five structures when combined with the omol task and is therefore not recommended. The full error statistics, regression parameters, and model-size comparisons are provided in Table S5. For MACE-Polar-1, the polar-1-l (large model) performs better than the polar-1-m (medium) model.

All in all, despite a few specific cases, the performance of the UMA-omol/ShiftML3 or MACE-Polar-1/ShiftML3 workflows exceeded our initial expectations. The gain in speed unlocked by this workflow also makes molecular dynamics (MD)

© 2026 The Author(s). Published by the Royal Society of Chemistry

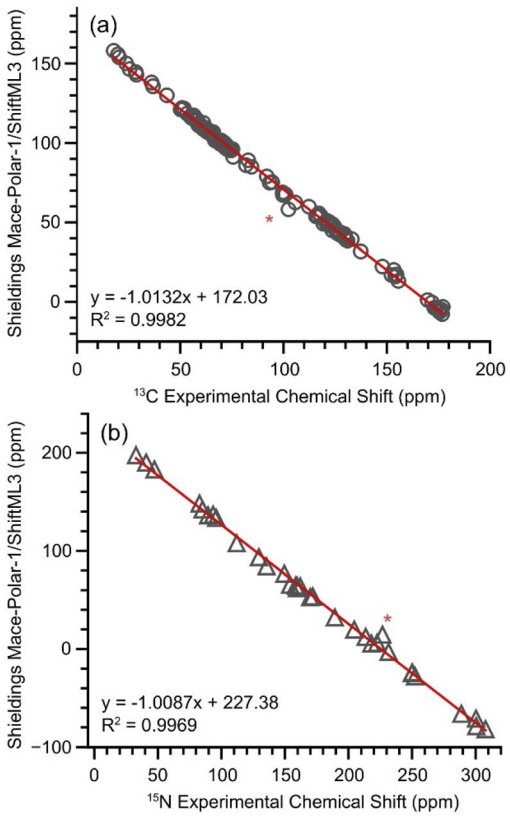


Fig. 4 Comparison between (a) the ^{13}C isotropic magnetic shieldings obtained *via* MACE-Polar-1/ShiftML3 and experimental ^{13}C chemical shifts (b) the ^{15}N isotropic magnetic shieldings obtained *via* MACE-Polar-1/ShiftML3 and experimental ^{15}N chemical shifts. *Denotes the C1 carbon in SUCROS04, and –NH in GEHHEH.

simulations at the ωB97M-V level computationally feasible on consumer grade hardware—a capability that is entirely out of reach with hybrid periodic DFT. While a systematic evaluation of MD averaging on the present benchmark set shows ambiguous improvements for the largely rigid structures considered here, this nonetheless reflects the nature of the test set rather than a limitation of the approach: dynamic averaging is expected to be most consequential for flexible or disordered systems, and this avenue is actively being explored.[49]

Fig. 4 reports the calculated magnetic shieldings obtained with MACE-Polar-1/ShiftML3 *vs.* the experimental chemical shifts for both the ^{13}C and ^{15}N set.[63]

Beyond isotropic chemical shifts: the accuracy of MLIPs/ShiftML3 for chemical shift tensors

One of the key features of ShiftML3 is its ability to predict the full magnetic shielding tensor rather than just the isotropic shielding. This facet is extremely valuable as chemical shift tensors are more sensitive to their environment and are inherently more useful for NMR crystallography applications. To be of full interest, the combination of MLIPs/ShiftML3 must be tested against DFT benchmarking methods that are known to give good agreement with experiment. For this purpose, we use the chemical shift anisotropies (CSAs) of 65 organic crystalline materials that have been reported in the literature. The chemical shift tensor, especially that of nitrogen, is very sensitive to molecular-level geometry, chemical bonding, and all types of noncovalent interactions (*e.g.*, hydrogen bonding).

The full magnetic shielding tensor can be obtained at the GIPAW (R)PBE level, using a cluster approach at the PBE0 level (see Methods),[74–76] or using ShiftML3. In general, the PBE0 cluster approach yields more accurate tensors than GGA GIPAW.[74–76] Here, we compare four workflows: (i) RPBE-D2 geometry optimization followed by RPBE GIPAW, (ii) RPBE-D2 geometry optimization followed by PBE0 calculations on molecular clusters to model short-range intermolecular interactions and COSMO to model long-range electrostatic effects, (iii) MACE-Polar-1/PBE GIPAW, and (iv) MACE-Polar-1/ShiftML3. The linear regression resulting from the principal values of the magnetic shielding tensors calculations are compared *versus* the experimental values in Table 2.

To assess how well the chemical shift tensors are reproduced, the distance, d, between the three principal components, δ_{11}, δ_{22}, and δ_{33}, of the experimental and the calculated shifts is given by the expression of Alderman *et al.*:[77]

$$d^2 = \frac{1}{15}\Big(3\Big(\big(\delta_{11}^{\mathrm{exp}} - \delta_{11}^{\mathrm{sim}}\big)^2 + \big(\delta_{22}^{\mathrm{exp}} - \delta_{11}^{\mathrm{sim}}\big)^2 + \big(\delta_{33}^{\mathrm{exp}} - \delta_{11}^{\mathrm{sim}}\big)^2\Big) + 2\big(\big(\delta_{11}^{\mathrm{exp}} - \delta_{11}^{\mathrm{sim}}\big)\big(\delta_{22}^{\mathrm{exp}} - \delta_{22}^{\mathrm{sim}}\big) + \big(\delta_{11}^{\mathrm{exp}} - \delta_{11}^{\mathrm{sim}}\big)\big(\delta_{33}^{\mathrm{exp}} - \delta_{33}^{\mathrm{sim}}\big) + \big(\delta_{22}^{\mathrm{exp}} - \delta_{22}^{\mathrm{sim}}\big)\big(\delta_{33}^{\mathrm{exp}} - \delta_{33}^{\mathrm{sim}}\big)\big)\Big)$$

where "exp" and "sim" relate to the experimental and simulated value. The average root mean square chemical shift distance (RMSD),

$$\mathrm{RMSD} = \frac{1}{n}\sqrt{\sum_{i=1}^{n} d_i^2}$$

is then calculated to assess the performance of the simulation methods.[77]

For the ^{13}C chemical shift tensors, relative to RPBE-D2/RPBE, the RMSD (3.5 ppm) decreases by ∼10% with MACE-Polar-1/PBE (RMSD = 3.26 ppm) and by *ca.* 30% with RPBE-D2/PBE0 cluster + COSMO (RMSD = 2.55 ppm). MACE-Polar-1/ShiftML3 performs on par with RPBE GIPAW (RMSD 3.3 ppm); specific outliers—SUCROS04, MALANAC02, TACETA01, and BUKVEI—contribute disproportionately to this error, likely reflecting limitations of ShiftML3's parent PBE functional rather than the geometry quality.

Considering ^{15}N, the trend differs slightly. When moving from RPBE GIPAW to the PBE0 cluster + COSMO, the RMSD is also improved from 7.81 to 6.67 ppm. Here, MACE-Polar-1/PBE GIPAW is significantly better with an RMSD of *ca.* 5.4 ppm, a trend still valid for MACE-Polar-1/ShiftML3, albeit with a slightly higher RMSD (5.9) ppm. For completeness, UMA-omol/ShiftML3 performs similarly to MACE-Polar-1, once again indicating that explicit inclusion of long-range electrostatic interaction has modest impact.

The results of the linear regression are reported in Fig. 5, and additional results are reported in the SI. The improvement of OMol25-trained MLIPs over RPBE-D2/RPBE for both ^{13}C and

© 2026 The Author(s). Published by the Royal Society of Chemistry

Table 2 Fitted linear regression parameters of the principal axis of ^{13}C anisotropic magnetic shieldings obtained *via* RPBE-D2/RPBE GIPAW, RPBE-D2/PBE0 cluster + COSMO, MACE-Polar-1/PBE GIPAW, MACE-Polar-1/ShiftML3 and UMA-omol/ShiftML3 *vs.* experimental ^{13}C principal axis component of the chemical shifts. The ^{15}N principal axis component of the anisotropic magnetic shieldings obtained RPBE-D2/RPBE GIPAW, RPBE-D2/PBE0 cluster + COSMO, MACE-Polar-1/PBE GIPAW, MACE-Polar-1/ShiftML3 and UMA-omol/ShiftML3 *vs.* the experimental ^{15}N principal axis component of the chemical shifts. For each, the linear regression parameters are given as well as the RMSD

	^{13}C			^{15}N		
	Regression			Regression		
Method	Slope	Intercept	RMSD	Slope	Intercept	RMSD
RPBE/RPBE (GIPAW)	−1.026	174.04	3.50	−1.056	231.62	7.81
RPBE/PBE0 + Clusters + COSMO	−1.061	187.43	2.55	−1.086	246.06	6.67
MACE-polar-1/PBE (polar-1-l)	−1.032	175.08	3.26	−1.056	233.43	5.42
MACE-polar-1/ShiftML3 (polar-1-l)	−1.035	174.28	3.31	−1.062	234.52	5.93
UMA-omol/ShiftML3 (uma-s-1p1)	−1.034	174.30	3.34	−1.092	234.57	5.89

^{15}N tensors reinforces the central conclusion: these models produce high-quality structures, and their combination with ShiftML3 yields shielding predictions that match or approach PBE GIPAW accuracy at a fraction of the cost.

Given the success of the MLIPs (OMol25)/ShiftML3 framework, the approach was subsequently applied to L-histidine.

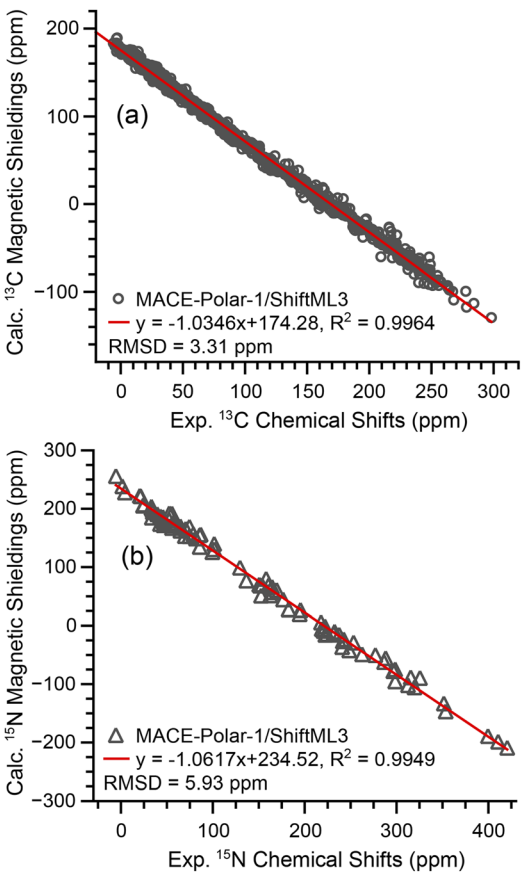


Fig. 5 Comparison between the principal axis of the anisotropic magnetic shieldings obtained *via* MACE-Polar-1/ShiftML3 *vs.* experimental principal axis component of the chemical shifts, for ^{13}C in (a) and ^{15}N in (b). For each, the linear regression parameters are given as well as the RMSD.

Application to MAS-DNP experiments on L-histidine

To illustrate the power of MLIPs/ShiftML3 in NMR crystallography applications, we applied the method to L-histidine. Here, we combine experimental ^{13}C and ^{15}N chemical shift tensors and ^{1}H isotropic chemical shifts with MACE-Polar-1/ShiftML3 predictions. L-Histidine has two known polymorphs, monoclinic and orthorhombic; these have similar conformations and crystal packing, which makes it particularly challenging to differentiate them.[8] To accelerate the characterization and assignment of L-histidine, we used MAS-DNP, which dramatically improves the signal-to-noise ratio of experiments carried out at natural isotopic abundance. This is particularly important for ^{15}N, which has a natural abundance of *ca.* 0.37%, severely impairing the sensitivity of the experiment. MAS-DNP is also carried out at low temperature, 100 K, which limits the impact of motion when extracting the anisotropic chemical shifts. Based on the benchmarks carried out so far, solid-state NMR is now the slowest step in the NMR crystallography protocol; however, use of MAS-DNP clearly accelerates this.

For L-histidine, the DNP enhancement was *ca.* 90 and the polarization time was a stretched exponential as commonly observed for particles with long intrinsic ^{1}H longitudinal relaxation times.[79] The optimal recycling delays (maximizing the signal to noise $S(t)/\sqrt{t}$) was *ca.* 20 s. The sensitivity allowed us to quickly measure the 1D ^{13}C and ^{15}N spectra in seconds to minutes but also run more complex correlations such as ^{1}H–^{13}C and ^{1}H–^{15}N, which are reported in the SI (Fig. S4 and S5), to assign the spectra.

Using the increased sensitivity afforded by MAS-DNP, we measured the isotropic chemical shift using MATPASS[80] and performed high signal-to-noise slow MAS experiments to extract the principal components of the chemical shift tensors. Due to the presence of multiple sites, the MATPASS experiment (Fig. 6) enabled accurate determination of the isotropic chemical shifts. This facilitates the fitting of the slow MAS experimental spectra and the accurate extraction of the CSAs. The values of the CSAs for ^{13}C and for ^{15}N are reported in Table 3. For ^{13}C, they are generally in agreement with previous work.[78]

For ^{13}C, the agreement between the experimental values and the MACE-Polar-1/ShiftML3 values is good when considering the isotropic shift with an MAE of 1.6 ppm. The CSA is also very

© 2026 The Author(s). Published by the Royal Society of Chemistry

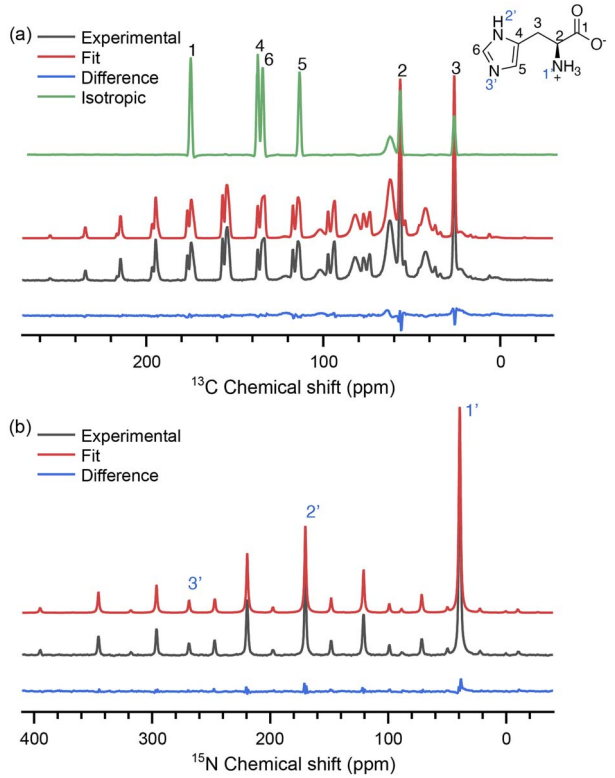


Fig. 6 Experimental CPMAS spectra of L-histidine collected at 3.0 kHz for (a) ^{13}C collected in 80 s and (b) ^{15}N collected in 28 minutes. For ^{13}C, the MATPASS spectrum is shown on top (green line), which corresponds to the isotropic shifts, while the fit of the slow MAS and the difference are shown in red and blue, respectively. Top right corner: molecular view of L-histidine with atomic numbering. The ethylene glycol signal appears at *ca.* 65 ppm in (a).

well reproduced with an RMSD of 3 ppm. For ^{15}N, the MAE for the isotropic shift is 1.4 ppm and an RMSD of 6.5 ppm. All in all, these results are excellent and in line with the statistics reported in the previous sections. This also confirms the ability to combine ML approaches with MAS-DNP for faster NMR crystallographic investigations. Including the heteronuclear correlation experiments, the complete set of experimental data was acquired in less than 5 hours. While the molecule studied here is simple, the same protocols can be applied to molecules of increasing size and complexity. It is likely that the ability to measure the CSA with high precision will be key to determining the molecular structures of active pharmaceutical ingredient solid forms.

However, the ^{13}C and ^{15}N CSAs alone are not sufficient to differentiate between the orthorhombic and monoclinic polymorphs of L-histidine. Indeed, these are nearly identical for the two structures (see Table S6). To have a clear differentiation, ^{1}H spectra[8] or PXRD are needed. The ^{1}H–^{13}C and ^{1}H–^{15}N HETCOR spectra can be used to extract the ^{1}H chemical shifts for each site. The ^{1}H NMR spectra have high resolution but present some anisotropic bulk magnetic susceptibility (ABMS) and are therefore slightly distorted. Nonetheless, we can resolve the seven chemical shifts (Table 4). The chemical shifts expected for both structures were calculated using MACE-Polar-1/ShiftML3 (shielding to shift calibration given in the SI). The RMSE calculated for all sites is not helpful because both structures produced very similar values for both polymorphs. Instead, only the RMSEs of the two differing sites, site #5 and site #6,8 are useful for differentiating between the polymorphs. In this case, the RMSE restricted to these two sites is 0.7 ppm for the orthorhombic and ∼1 ppm for the monoclinic one. This supports the identification of the orthorhombic crystal structure. Further confirmation of the crystal structure of L-histidine as the orthorhombic polymorph was provided by PXRD analysis (see Fig. S2).

The combination of MACE-Polar-1/ShiftML3 and MAS-DNP proved to be a valid approach to characterize L-histidine. It correctly predicted the CSAs for both ^{13}C and ^{15}N obtained by fitting the slow MAS data, while providing enough information on the ^{1}H spectra to differentiate between the orthorhombic and the monoclinic polymorphs (Table 4).

Computational

Structural models

The 65 crystal structures were obtained from the Cambridge Structural Database, or when unavailable, from previous literature reports. They aim at representing a diverse pool of

Table 3 Comparison between experimental CSA and simulated one *via* MACE-Polar-1/ShiftML3 ^{13}C and ^{15}N for L-histidine. Experimental isotropic shifts (δ_{iso}) were obtained from MAT-PASS (for ^{13}C) or 1D CP spectrum for ^{15}N collected at 10.5 kHz, and principal values (δ_{11}, δ_{22}, δ_{33}) were obtained from slow-MAS CSA lineshape fitting (see Methods). Predicted tensors were generated using MACE-Polar-1 optimized geometries with ShiftML3 for the orthorhombic phase. All values are chemical shifts in ppm, referenced as described in Methods. Principal values are reported in the Mehring convention and ordered such that $\delta_{11} \geq \delta_{22} \geq \delta_{33}$

Site ^{13}C	1	4	6	5	2	3	Site ^{15}N	3′	2′	1′
Experimental δ_{iso} (ppm)	174.9	137.1	134.2	113.3	56.6	26.1		247	170.2	39.3
δ_{11} (ppm)	244.5	210.5	205	185.9	67.8	33.5		397.1	271.3	57.6
δ_{22} (ppm)	176.1	143.5	138.1	116.4	64.8	31		325.1	187.3	43.7
δ_{33} (ppm)	104	57.1	59.7	37.8	37.2	13.7		18.3	51.9	16.5
Predicted δ_{iso} (ppm)	171.9	138.0	131.3	114	55.1	27.5		244.1	169.9	40.4
δ_{11} (ppm)	239.8	205.7	196.9	185.5	67.9	38.3		407.2	260.7	48.2
δ_{22} (ppm)	167.8	149.9	137.3	116.5	59.9	27.5		309.8	187.9	40.6
δ_{33} (ppm)	108.3	58.3	59.7	40.1	37.7	16.7		15.4	61.1	32.5
Error in δ_{iso} (ppm)	3.0	−0.9	2.9	−0.7	1.5	−1.4		2.9	0.3	−1.1
Distance (d, ppm)	4.4	3.0	3.8	1.0	2.1	2.7		7.1	5.1	6.9

© 2026 The Author(s). Published by the Royal Society of Chemistry

Table 4 Comparison between experimental ^{1}H shifts and simulated one *via* MACE-Polar-1/ShiftML3 for the L-histidine monoclinic and orthorhombic structures. Shielding calibration reported in the SI. Note that for the $-NH_3$ group (1′), there is not enough resolution to distinguish the three sites. As such, the values marked * are the average predicted values. Calibration of chemical shifts *vs.* magnetic shieldings is reported in Fig. S3, slope −1.0192, intercept, 30.95 ppm, RMSE = 0.54 ppm, MAE = 0.43 ppm, Max AE = 1.69 ppm

Observable sites	2′	1′	6	5	2	3	3	RMSE site 5 and 6 (ppm)
δ_{iso} ±0.2 (ppm)	14.0	8.9	5.5	5.0	3.80	2.9	2.7	
Predicted δ_{iso} (ppm) (orthorhombic)	14.4	8.2* (8.4, 8.2, 8.0)	**5.9**	**5.6**	3.5	2.8	2.7	0.72
Predicted δ_{iso} (ppm) (monoclinic)	14.3	8.3* (8.8, 8.3, 8.1)	**6.4**	**5.4**	3.4	2.7	2.6	0.99

chemical environments. The list of the crystal structures is reported in the SI. In addition, we used the set of calibrated chemical shifts recently reported in ref. 63.

Each of these structures was checked for missing hydrogen atoms. When hydrogen atoms positions were not reported, the atoms were instead placed in ideal positions using the Mercury software package. The structures were used either "as-is" or subject to different types of geometry optimization, either RPBE-D2, UMA[51] and MACE-Polar-1 (ref. 50) combined with atomic simulation environment (ASE).[81]

Structure optimization

When using UMA, either the task omol[52] or omc[53] was used in combination with the model uma-s-1p1 obtained from the FAIR Chemistry repository on Huggingface (version 2.16.0).[51] Note that in our tests, the uma-s-1p2 (ref. 56) and uma-m-1p1 models perform slightly worse. The geometry optimization with the MLIPs were carried out with limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS)[82] after testing FIRE (see SI).[83] For all optimizations, the cell parameters were kept constant.

Simulations with MACE-Polar-1 (ref. 50) were carried out similarly, and all model, polar-1-s, polar-1-m and polar-1-l have been tested. Note that we used a precision of Float32 after noting minimal loss in accuracy (*ca.* 10^{-2} on the RMSE). As reported in the SI, the large model outperforms the smaller ones in testing the Beran *et al.* set.[63]

Plane-wave DFT[84,85] geometry optimizations were performed using CASTEP.[86] These calculations used either the functionals PBE[64] or RPBE.[62] Ultrasoft pseudopotentials generated on-the-fly,[87] a plane-wave cutoff energy of 800 eV, a Monkhorst–Pack grid with a *k*-point spacing of 0.05 Å^{-1}. Dispersion corrections were introduced using the semi-empirical two-body (D2) or three-body (D3) dispersion force field corrections of Grimme.[61,65,66] Structural refinements employed the L-BFGS energy-minimizing scheme, in which the positions of all atoms were relaxed while holding unit cell parameters constant, with convergence thresholds including a maximum change in energy of 5×10^{-6} eV per atom, a maximum displacement of 5×10^{-4} Å per atom, and a maximum Cartesian force of 10^{-2} eV Å^{-1}.

The 65 structures were optimized using UMA on a NVIDIA RTX A4500 GPU and down to forces of 0.01 eV Å^{-1} in less than 10 minutes running on Ubuntu 24.04.1.

The cif files for orthorhombic and monoclinic L-histidine were obtained from the CSD (code CCDC 2097463 and 1206539). Both crystal structures have been optimized with UMA-omol, uma-s-1p1. The UMA-omol optimization took 8.5 s on a single GPU RTX A4500 while the CASTEP v25.1 CPU based approach (compiled with OpenMPI and executed on 20 cores of an Intel Xeon Gold 6258R) took 3360 s for optimizing the monoclinic structure at the PBE-D3 level.

The optimization with UMA-omol took 11.5 s while the same optimization took CASTEP 10550 s. This is 2–3 orders of magnitude and this difference scales further as the number of atoms increase (O(N) for UMA, O(N^4) for DFT).

For all the 15 structures containing ^{15}N in the Beran set, the optimization using UMA-omol, uma-s-1p1, took 265 s, while MACE-Polar-1-medium, took 307 s, and the large one 403 s.

Magnetic shielding tensors

^{13}C and ^{15}N magnetic shielding tensors were computed on all types of structures, X-ray, RPBE-D2, UMA and MACE-Polar-1 optimized, using either PBE, RPBE or r^{2}SCAN[88] GIPAW or ShiftML3.

GIPAW calculations were performed using CASTEP v25.1,86 using the same parameters and approximations as used in the geometry optimizations. Additional calculations of ^{13}C and ^{15}N magnetic shielding tensors were performed using the ADF suite within the Amsterdam Modelling Suite (ADF v.2021). These calculations used the more costly hybrid functional PBE0.[89] Large clusters of up to fifteen molecules were used to model the local lattice environment, as described in previous work.[74–76] For salts, these clusters had a non-zero net charge. A basis set partitioning scheme was applied in which the central molecule within the cluster was assigned the TZ2P basis set with Becke integration set to "excellent", whereas all other molecules in the cluster were assigned the DZ basis set with "normal" Becke integration.[90,91] Magnetic shielding tensors are calculated only for nuclei within the central molecule within the cluster. Electrostatic embedding to model longer range electrostatic interactions with the lattice was accomplished through use of the conductor-like screening model (COSMO),[92–95] as implemented in ADF.[96] Benzene or water were chosen as the solvent (for hydrocarbons and all other materials, respectively), using Delley-type cavity construction,[97] with atomic cavity radii based on the van der Waals radii of Allinger.[98]

© 2026 The Author(s). Published by the Royal Society of Chemistry

Magnetic shielding comparison with experimental chemical shifts

For the set of Beran *et al.*, the chemical shifts follow the assignment given in ref. 63. For the chemical shielding tensors, while the sites have been assigned, tracking these is not an easy task for such a large dataset. As such, we resorted to a different approach: for each crystal structure, and for each site, the experimental CSA was matched to the best anisotropic magnetic shielding calculated. This is done by reducing the overall RMSD of all sites of a given crystal structure at once using the Hungarian algorithm. This is done systematically for each crystal structure and across all three methods, ensuring a consistent comparison and representing the best-case scenario for each one.

Experimental

Powder X-ray diffraction

L-Histidine purchased from Sigma-Aldrich was ground in a mortar and pestle for 10 min with 4 μL of ethanol (EtOH) using liquid-assisted grinding (LAG). Powder X-ray diffraction (PXRD) data were collected on a Rigaku MiniFlex benchtop diffractometer using Cu Kα radiation ($\lambda = 1.5406$ Å) and a D/teX Ultra2 detector. Each sample was loaded into a zero-background silicon sample holder (well size 5 mm × 2 mm) and measured using the 8-position autosampler. Data were collected at 40 kV and 15 mA over 2θ spanning 2.5° to 60° with a 0.030° step size and a 5° min^{-1} scan rate (total scan time *ca.* 13 min). Experimental PXRD patterns were compared with simulated patterns to assess phase purity using HighScore and CrystalDiffract.

MAS-DNP

L-Histidine was uniformly impregnated with a 40 mM AsymPol-POK (https://pubs.acs.org/jacsat/article/140/35/11013/757696/Computationally-Assisted-Design-of-Polarizing) solution prepared in a 8/2 (% v/v) mixture of ethylene glycol-d_6 and ethylene glycol. The sample was mixed thoroughly, packed into a 3.2 mm sapphire rotor, and degassed by five flash-freeze/thaw cycles in liquid nitrogen outside the probe to remove dissolved oxygen prior to DNP measurements.

All spectra were collected at 100 K and a MAS frequency of 10.5 or 3.0 kHz using the NHMFL 600 MHz MAS-DNP instrument,[99] equipped with a NHMFL-built HCN probe with a 3.2 mm Bruker stator. Samples were packed into sapphire rotors for optimal efficiency.[100]

Except where otherwise specified, the nutation frequencies for 90 and 180° pulses used were 100 kHz for ^{1}H, and 50 kHz for ^{13}C and ^{15}N. SPINAL64 was used for ^{1}H decoupling[101] during acquisition. ^{15}N and ^{13}C spectra were obtained *via* cross-polarization (CP) using a 1 ms contact time and a ramp of 20% on the ^{1}H channel. For both, the ^{1}H CP nutation was set to 50 kHz at 100% of the ramp.

The L-histidine sample preparation yielded an enhancement of *ca.* 90. The buildup curve obtained by saturation recovery and ^{1}H → ^{13}C CP detection presents a non-exponential form that can be fitted with a biexponential function with time constants of 8.5 and 114 s, and respective weights of 0.2 and 0.8. The optimal recycling delay is thus *ca.* 27 s. A similar curve was obtained for all ^{13}C sites.

The ^{1}H–^{13}C HETCOR with FLGS homonuclear decoupling[102–104] spectra were collected with a cross-polarization contact time of 100 μs to minimize the effect of ^{1}H spin-diffusion. This generates lower intensities for sites without direct ^{1}H bonding. The indirect dimension was 100 points for ^{13}C and ^{15}N (*ca.* 1 ms evolution time). Both experiments took 67 minutes. The indirect dimension was referenced using ^{13}C-labelled α-glycine for ^{13}C, and NH_4Cl for ^{15}N.

Slow MAS spectra were collected at 3.0 kHz. For ^{13}C, the spectrum was acquired in 80 s (signal-to-noise ratio, SNR = 458) and 28 minutes for ^{15}N (SNR = 387). The NMR spectra were fitted in two steps to extract the CSA. First, a CP-MATPASS experiment was used to obtain the isotropic shifts,[80] and then these parameters were used to fit the 3.0 kHz data using SOLA module within TopSpin 4.1.14. For ^{15}N, the CP spectrum at 10.5 kHz provided the resolution needed to determine the isotropic chemical shifts.

Conclusions

This study demonstrates the capacity of machine learning to mitigate the heavy computational demands of NMR crystallography. Our analysis shows that the MLIP/ShiftML3 framework is not merely a high-speed substitute for the PBE/PBE GIPAW workflow; in many instances, it delivers superior structural and spectroscopic accuracy, while reducing computation time from days to minutes. UMA-omc, designed to reproduce PBE-D3 calculations, did precisely that—achieving parity with the conventional workflow at a fraction of the cost, and representing an excellent drop-in replacement for routine applications.

The performance of the two OMol25-trained models—UMA-omol and MACE-Polar-1—was especially notable: despite being trained exclusively on gas-phase molecular data, both produced structural refinements comparable to hybrid-DFT quality. When coupled with ShiftML3, the resulting magnetic shielding predictions rivaled PBE0/PBE GIPAW accuracy, representing a 20–30% improvement in experimental agreement over standard PBE methods. This demonstrates robust transfer of learning to the more complex environment of molecular crystals.

Beyond the practical implications for NMR crystallography, these results contribute to the emerging picture of gas-phase-to-crystal transferability of MLIPs. Whereas previous benchmarks have relied on energetic and structural metrics that average over many degrees of freedom, NMR chemical shifts—sensitive to sub-ångström variations in local geometry—provide a uniquely demanding validation metric. The fact that OMol25-trained models pass this test, matching hybrid-DFT structural quality as judged by one of the most sensitive local probes available, significantly strengthens the case for transferability and suggests that the relevant physics of intermolecular interactions in organic solids is well-represented within the local embedding of these models.

© 2026 The Author(s). Published by the Royal Society of Chemistry

The three-orders-of-magnitude gain in speed carries practical consequences worth emphasizing. Hybrid periodic-DFT optimization has historically required substantial HPC infrastructure, confining high-accuracy NMR crystallography to specialized groups. The present results show that OMol25-trained MLIPs, requiring only a consumer-grade GPU, provide a freely available alternative that democratizes hybrid-quality structural refinement. The same speed advantage also makes MD simulations at the ωB97M-V level tractable for the first time, enabling dynamic averaging of chemical shifts for flexible, disordered, or dynamically complex systems where static structures are insufficient.[27,28,49]

However, limitations still remain. ShiftML3 prediction errors appear to reflect the underlying PBE functional rather than the ML model itself; a delta-learning approach (using monomer or cluster correction) or r^{2}SCAN-trained iteration would likely improve agreement with experiment. Structural outliers in UMA were rare but warrant further investigation to separate force-field inaccuracies from reference functional limitations. Looking ahead, MLIPs trained on OMol25 are well-suited as geometry engines for CSP workflows such as FastCSP, Genarris 3.0, PackFlow, and MolCrystalFlow.[57,59,60,105,106] Combined with faster acquisition methods, such as fast MAS ^{1}H detection, or MAS-DNP as presented here, it proved to be efficient to not only characterize the chemical shift tensors, but also to distinguish between the two polymorphs of L-histidine.

In summary, the UMA/ShiftML3 or MACE-Polar-1/ShiftML3 synergy is a paradigm shift for NMR crystallography, offering a potent combination of PBE0-level accuracy and unprecedented computational speed.

Author contributions

The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript. F. M. V. performed conceptualization, formal analysis, investigation, methodology, software, validation, visualization, and supervision, and wrote, reviewed, and edited the original draft. S. T. H. performed conceptualization, formal analysis, investigation, methodology, and validation, and reviewed, and edited the original draft. S. S. G. performed formal analysis, investigation, and visualization, and reviewed, and edited the original draft. R. W. S. performed conceptualization, funding acquisition, methodology, visualization, and supervision.

Conflicts of interest

There are no conflicts to declare.

Data availability

The experimental data are available at 10.5281/zenodo.19058068 and the scripts are available at: https://github.com/fmentink/. UMA, MACE, ASE and ShiftML3 are freely available here: https://github.com/facebookresearch/fairchem, https://github.com/acesuit/mace, https://ase-lib.org/, https://github.com/lab-cosmo/shiftml.

Supplementary information (SI) is available. See DOI: https://doi.org/10.1039/d6sc04941a.

Acknowledgements

The National High Magnetic Field Laboratory is supported by the National Science Foundation under Cooperative Agreement No. DMR-2128556 and the State of Florida. This work also received partial support from the European Union's Horizon 2020 research and innovation program (Pan-European Solid-State NMR Infrastructure for Chemistry-Enabling Access, PANACEA, Grant Agreement 101008500), from the NIH National Resource for Advanced NMR Technology (NIH RM1-GM14867), and from the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Award Number DE-SC0022310, which covered expenses related to software for quantum-chemical computational methods used herein and in related projects.

References

1 S. Datta and D. J. W. Grant, Crystal structures of drugs: advances in determination, prediction and engineering, *Nat. Rev. Drug Discovery*, 2004, **3**, 42–57.
2 E. Pindelska, A. Sokal and W. Kolodziejski, Pharmaceutical cocrystals, salts and polymorphs: Advanced characterization techniques, *Adv. Drug Delivery Rev.*, 2017, **117**, 111–146.
3 S. R. Chemburkar, J. Bauer, K. Deming, H. Spiwek, K. Patel, J. Morris, R. Henry, S. Spanton, W. Dziki, W. Porter, J. Quick, P. Bauer, J. Donaubauer, B. A. Narayanan, M. Soldani, D. Riley and K. McFarland, Dealing with the Impact of Ritonavir Polymorphs on the Late Stages of Bulk Drug Process Development, *Org. Process Res. Dev.*, 2000, **4**, 413–417.
4 K. N. Houk and F. Liu, Holy Grails for Computational Organic Chemistry and Biochemistry, *Acc. Chem. Res.*, 2017, **50**, 539–543.
5 J. H. Van Drie, Computer-aided drug design: the next 20 years, *J. Comput.-Aided Mol. Des.*, 2007, **21**, 591–601.
6 E. Salager, G. M. Day, R. S. Stein, C. J. Pickard, B. Elena and L. Emsley, Powder Crystallography by Combined Crystal Structure Prediction and High-Resolution ^{1}H Solid-State NMR Spectroscopy, *J. Am. Chem. Soc.*, 2010, **132**, 2564–2566.
7 M. Baias, J.-N. Dumez, P. H. Svensson, S. Schantz, G. M. Day and L. Emsley, De Novo Determination of the Crystal Structure of a Large Drug Molecule by Crystal Structure Prediction-Based Powder NMR Crystallography, *J. Am. Chem. Soc.*, 2013, **135**, 17501–17507.
8 T. Oikawa, M. Okumura, T. Kimura and Y. Nishiyama, Solid-state NMR meets electron diffraction: determination of crystalline polymorphs of small organic microcrystalline samples, *Acta Crystallogr., Sect. C:Struct. Chem.*, 2017, **73**, 219–228.

© 2026 The Author(s). Published by the Royal Society of Chemistry

9 B. J. Wylie, C. D. Schwieters, E. Oldfield and C. M. Rienstra, Protein Structure Refinement Using $^{13}C\alpha$ Chemical Shift Tensors, *J. Am. Chem. Soc.*, 2009, **131**, 985–992.

10 S. Macholl, F. Börner and G. Buntkowsky, Revealing the Configuration and Crystal Packing of Organic Compounds by Solid-State NMR Spectroscopy: Methoxycarbonylurea, a Case Study, *Chem. - Eur. J.*, 2004, **10**, 4808–4816.

11 D. Sakellariou, S. P. Brown, A. Lesage, S. Hediger, M. Bardet, C. A. Meriles, A. Pines and L. Emsley, High-Resolution NMR Correlation Spectra of Disordered Solids, *J. Am. Chem. Soc.*, 2003, **125**, 4376–4380.

12 R. K. Harris, NMR studies of organic polymorphs & solvates, *Analyst*, 2006, **131**, 351.

13 H. Hamaed, J. M. Pawlowski, B. F. T. Cooper, R. Fu, S. H. Eichhorn and R. W. Schurko, Application of Solid-State ^{35}Cl NMR to the Structural Characterization of Hydrochloride Pharmaceuticals and their Polymorphs, *J. Am. Chem. Soc.*, 2008, **130**, 11056–11065.

14 S. T. Holmes, J. M. Hook and R. W. Schurko, Nutraceuticals in Bulk and Dosage Forms: Analysis by ^{35}Cl and ^{14}N Solid-State NMR and DFT Calculations, *Mol. Pharmaceutics*, 2022, **19**, 440–455.

15 S. T. Holmes, C. S. Vojvodin, N. Veinberg, E. M. Iacobelli, D. A. Hirsh and R. W. Schurko, Hydrates of active pharmaceutical ingredients: A ^{35}Cl and ^{2}H solid-state NMR and DFT study, *Solid State Nucl. Magn. Reson.*, 2022, **122**, 101837.

16 X. Kong, M. Shan, V. Terskikh, I. Hung, Z. Gan and G. Wu, Solid-State ^{17}O NMR of Pharmaceutical Compounds: Salicylic Acid and Aspirin, *J. Phys. Chem. B*, 2013, **117**, 9643–9654.

17 F. G. Vogt, H. Yin, R. G. Forcino and L. Wu, ^{17}O Solid-State NMR as a Sensitive Probe of Hydrogen Bonding in Crystalline and Amorphous Solid Forms of Diflunisal, *Mol. Pharmaceutics*, 2013, **10**, 3433–3446.

18 Q. Z. Ni, F. Yang, T. V. Can, I. V. Sergeyev, S. M. D'Addio, S. K. Jawla, Y. Li, M. P. Lipert, W. Xu, R. T. Williamson, A. Leone, R. G. Griffin and Y. Su, In Situ Characterization of Pharmaceutical Formulations by Dynamic Nuclear Polarization Enhanced MAS NMR, *J. Phys. Chem. B*, 2017, **121**, 8132–8141.

19 Q. Z. Ni, E. Daviso, T. V. Can, E. Markhasin, S. K. Jawla, T. M. Swager, R. J. Temkin, J. Herzfeld and R. G. Griffin, High Frequency Dynamic Nuclear Polarization, *Acc. Chem. Res.*, 2013, **46**, 1933–1941.

20 A. J. Rossini, A. Zagdoun, M. Lelli, A. Lesage, C. Copéret and L. Emsley, Dynamic Nuclear Polarization Surface Enhanced NMR Spectroscopy, *Acc. Chem. Res.*, 2013, **46**, 1942–1951.

21 D. Lee, S. Hediger and G. De Paëpe, Is solid-state NMR enhanced by dynamic nuclear polarization?, *Solid State Nucl. Magn. Reson.*, 2015, **66–67**, 6–20.

22 A. S. Lilly Thankamony, J. J. Wittmann, M. Kaushik and B. Corzilius, Dynamic nuclear polarization for sensitivity enhancement in modern solid-state NMR, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2017, **102–103**, 120–195.

23 A. J. Rossini, C. M. Widdifield, A. Zagdoun, M. Lelli, M. Schwarzwälder, C. Copéret, A. Lesage and L. Emsley, Dynamic Nuclear Polarization Enhanced NMR Spectroscopy for Pharmaceutical Formulations, *J. Am. Chem. Soc.*, 2014, **136**, 2324–2334.

24 W. Zhao, E. C. Thomas, D. Debnath, F. J. Scott, F. Mentink-Vigier, J. R. White, R. L. Cook and T. Wang, Enriched Molecular-Level View of Saline Wetland Soil Carbon by Sensitivity-Enhanced Solid-State NMR, *J. Am. Chem. Soc.*, 2025, **147**, 519–531.

25 A. C. Pinon, A. J. Rossini, C. M. Widdifield, D. Gajan and L. Emsley, Polymorphs of Theophylline Characterized by DNP Enhanced Solid-State NMR, *Mol. Pharmaceutics*, 2015, **12**, 4146–4153.

26 S. F. Cousin, C. E. Hughes, F. Ziarelli, S. Viel, G. Mollica, K. D. M. Harris, A. C. Pinon and P. Thureau, Exploiting solid-state dynamic nuclear polarization NMR spectroscopy to establish the spatial distribution of polymorphic phases in a solid material, *Chem. Sci.*, 2023, **14**, 10121–10128.

27 M. Dračínský, P. Bouř and P. Hodgkinson, Temperature Dependence of NMR Parameters Calculated from Path Integral Molecular Dynamics Simulations, *J. Chem. Theory Comput.*, 2016, **12**, 968–973.

28 M. Dračínský and P. Hodgkinson, Effects of Quantum Nuclear Delocalisation on NMR Parameters from Path Integral Molecular Dynamics, *Chem. - Eur. J.*, 2014, **20**, 2201–2207.

29 M. Dračínský, Analyzing Discrepancies in Chemical-Shift Predictions of Solid Pyridinium Fumarates, *Molecules*, 2026, **26**(13), 3857.

30 C. J. Pickard and F. Mauri, All-electron magnetic response with pseudopotentials: NMR chemical shifts, *Phys. Rev. B*, 2001, **63**, 245101.

31 C. Bonhomme, C. Gervais, F. Babonneau, C. Coelho, F. Pourpoint, T. Azaïs, S. E. Ashbrook, J. M. Griffin, J. R. Yates, F. Mauri and C. J. Pickard, First-Principles Calculation of NMR Parameters Using the Gauge Including Projector Augmented Wave Method: A Chemist's Point of View, *Chem. Rev.*, 2012, **112**, 5733–5779.

32 J. K. Harper, R. Iuliucci, M. Gruber and K. Kalakewich, Refining crystal structures with experimental 13C NMR shift tensors and lattice-including electronic structure methods, *CrystEngComm*, 2013, **15**, 8693.

33 E. Chaloupecká, O. Socha and M. Dračínský, Diverging errors: A comparison of DFT and machine-learning predictions of NMR shieldings, *Solid State Nucl. Magn. Reson.*, 2025, **138**, 102019.

34 F. M. Paruzzo, A. Hofstetter, F. Musil, S. De, M. Ceriotti and L. Emsley, Chemical shifts in molecular solids by machine learning, *Nat. Commun.*, 2018, **9**, 4501.

35 M. Balodis, M. Cordova, A. Hofstetter, G. M. Day and L. Emsley, De Novo Crystal Structure Determination from Machine Learned Chemical Shifts, *J. Am. Chem. Soc.*, 2022, **144**, 7215–7223.

36 M. Cordova, E. A. Engel, A. Stefaniuk, F. Paruzzo, A. Hofstetter, M. Ceriotti and L. Emsley, A Machine Learning Model of Chemical Shifts for Chemically and

© 2026 The Author(s). Published by the Royal Society of Chemistry

Structurally Diverse Molecular Solids, *J. Phys. Chem. C*, 2022, **126**, 16710–16720.

37 M. Bertani, A. Pedone, F. Faglioni and T. Charpentier, Accelerating NMR Shielding Calculations Through Machine Learning Methods: Application to Magnesium Sodium Silicate Glasses, *ChemPhysChem*, 2024, **25**, e202300782.

38 M. Kellner, J. B. Holmes, R. Rodriguez-Madrid, F. Viscosi, Y. Zhang, L. Emsley and M. Ceriotti, A Deep Learning Model for Chemical Shieldings in Molecular Organic Solids Including Anisotropy, *J. Phys. Chem. Lett.*, 2025, **16**, 8714–8722.

39 R. R. Brew, I. A. Nelson, M. Binayeva, A. S. Nayak, W. J. Simmons, J. J. Gair and C. C. Wagen, Wiggle150: Benchmarking Density Functionals and Neural Network Potentials on Highly Strained Conformers, *J. Chem. Theory Comput.*, 2025, **21**, 3922–3929.

40 S. Fayfar, R. Chahal, H. Williams, D. N. Gardner, G. Zheng, D. Sprouster, J. C. Neuefeind, D. Olds, A. Hwang, J. Mcfarlane, R. C. Gallagher, M. Asta, S. Lam, R. O. Scarlat and B. Khaykovich, Complex Structure of Molten (2LiF-BeF_2) Examined by Experimental Neutron Scattering, X-ray Scattering, and Deep-Neural-Network Based Molecular Dynamics, *PRX Energy*, 2024, **3**, 013001.

41 A. Mazitov, F. Bigi, M. Kellner, P. Pegolo, D. Tisi, G. Fraux, S. Pozdnyakov, P. Loche and M. Ceriotti, PET-MAD as a lightweight universal interatomic potential for advanced materials modeling, *Nat. Commun.*, 2025, **16**, 10653.

42 L. A. M. Rosset, D. A. Drabold and V. L. Deringer, Signatures of paracrystallinity in amorphous silicon from machine-learning-driven molecular dynamics, *Nat. Commun.*, 2025, **16**, 2360.

43 D. Kuryla, F. Berger, G. Csányi and A. Michaelides, How accurate are DFT forces? Unexpectedly large uncertainties in molecular datasets, *J. Chem. Phys.*, 2025, **163**, 224313.

44 S. Lee and A. B. Dieng, Are neural scaling laws leading quantum chemistry astray?, *arXiv*, 2025, preprint, arXiv:2509.26397, DOI: 10.48550/arXiv.2509.26397.

45 E. L. Mann, C. C. Wagen, J. E. Vandezande, A. M. Wagen and S. C. Schneider, Egret-1: Pretrained Neural Network Potentials for Efficient and Accurate Bioorganic Simulation, *arXiv*, 2025, preprint arXiv:2504.20955, DOI: 10.48550/ARXIV.2504.20955.

46 D. M. Anstine, R. Zubatyuk and O. Isayev, AIMNet2: a neural network potential to meet your neutral, charged, organic, and elemental-organic needs, *Chem. Sci.*, 2025, **16**, 10228–10244.

47 D. P. Kovács, J. H. Moore, N. J. Browning, I. Batatia, J. T. Horton, Y. Pu, V. Kapil, W. C. Witt, I.-B. Magdău, D. J. Cole and G. Csányi, MACE-OFF: Short-Range Transferable Machine Learning Force Fields for Organic Molecules, *J. Am. Chem. Soc.*, 2025, **147**, 17598–17611.

48 B. Rhodes, S. Vandenhaute, V. Šimkus, J. Gin, J. Godwin, T. Duignan and M. Neumann, Orb-v3: atomistic simulation at scale, *arXiv*, 2025, preprint arXiv:2504.06231, DOI: 10.48550/arXiv.2504.06231.

49 M. Kellner, R. Rodriguez-Madrid, J. B. Holmes, V. P. Principe, L. Emsley and M. Ceriotti, Quantum-corrected NMR crystallography at scale, *arXiv*, 2026, preprint arXiv:2603.06236, DOI: 10.48550/arXiv.2603.06236.

50 I. Batatia, W. J. Baldwin, D. Kuryla, J. Hart, E. Kasoar, A. M. Elena, H. Moore, M. J. Gawkowski, B. X. Shi, V. Kapil, P. Kourtis, I.-B. Magdău and G. Csányi, MACE-POLAR-1: A Polarisable Electrostatic Foundation Model for Molecular Chemistry, *arXiv*, 2026, preprint, arXiv:2602.19411, DOI: 10.48550/arXiv.2602.19411.

51 B. M. Wood, M. Dzamba, X. Fu, M. Gao, M. Shuaibi, L. Barroso-Luque, K. Abdelmaqsoud, V. Gharakhanyan, J. R. Kitchin, D. S. Levine, K. Michel, A. Sriram, T. Cohen, A. Das, A. Rizvi, S. J. Sahoo, Z. W. Ulissi and C. L. Zitnick, UMA: A Family of Universal Models for Atoms, *arXiv*, 2025, preprint arXiv:2506.23971, DOI: 10.48550/arXiv.2506.23971.

52 D. S. Levine, M. Shuaibi, E. W. C. Spotte-Smith, M. G. Taylor, M. R. Hasyim, K. Michel, I. Batatia, G. Csányi, M. Dzamba, P. Eastman, N. C. Frey, X. Fu, V. Gharakhanyan, A. S. Krishnapriyan, J. A. Rackers, S. Raja, A. Rizvi, A. S. Rosen, Z. Ulissi, S. Vargas, C. L. Zitnick, S. M. Blau and B. M. Wood, The Open Molecules 2025 (OMol25) Dataset, Evaluations, and Models, *arXiv*, 2025, preprint arXiv:2505.08762, DOI: 10.48550/arXiv.2505.08762.

53 V. Gharakhanyan, L. Barroso-Luque, Y. Yang, M. Shuaibi, K. Michel, D. S. Levine, M. Dzamba, X. Fu, M. Gao, X. Liu, H. Ni, K. Noori, B. M. Wood, M. Uyttendaele, A. Boromand, C. L. Zitnick, N. Marom, Z. W. Ulissi and A. Sriram, Open Molecular Crystals 2025 (OMC25) dataset and models, *Sci. Data*, 2026, **13**, 354.

54 K. S. Nayal, I. Cho and O. Isayev, Benchmarking Universal Machine-learned Interatomic Potentials for Intermolecular and Noncovalent Interactions, *ChemRxiv*, 2026, preprint, DOI: 10.26434/chemrxiv.15000203/v1.

55 L. Barroso-Luque, M. Shuaibi, X. Fu, B. M. Wood, M. Dzamba, M. Gao, A. Rizvi, C. L. Zitnick and Z. W. Ulissi, Open Materials 2024 (OMat24) Inorganic Materials Dataset and Models, *arXiv*, 2024, preprint arXiv:2410.12771, DOI: 10.48550/ARXIV.2410.12771.

56 D. S. Levine, N. Liesen, L. Chua, J. Diffenderfer, H. Ingolfsson, M. P. Kroonblawd, N. Kumar, A. Maiti, S. S. Mohottalalage, M. Shuaibi, B. V. Essen, B. M. Wood, C. L. Zitnick, S. M. Blau and E. R. Antoniuk, The Open Polymers 2026 (OPoly26) Dataset and Evaluations, *arXiv*, 2026, preprint arXiv:2512.23117, DOI: 10.48550/arXiv.2512.23117.

57 C. Zeng, H. W. Sullivan, T. Egg, M. M. Martirossyan, P. Höllmer, J. Jin, R. G. Hennig, A. Roitberg, S. Martiniani, E. B. Tadmor and M. Liu, MolCrystalFlow: Molecular Crystal Structure Prediction via Flow Matching, *arXiv*, 2026, preprint arXiv:2602.16020, DOI: 10.48550/arXiv.2602.16020.

© 2026 The Author(s). Published by the Royal Society of Chemistry

58 S. Grimme, A. Hansen, S. Ehlert and J.-M. Mewes, r^{2}SCAN-3c: A "Swiss army knife" composite electronic-structure method, *J. Chem. Phys.*, 2021, **154**, 064103.

59 A. Subramanian, E. Pan, J. Nam, M. Weiler, S. Qu, C. W. Park, T. S. Jaakkola, E. Olivetti and R. Gomez-Bombarelli, PackFlow: Generative Molecular Crystal Structure Prediction via Reinforcement Learning Alignment, *arXiv*, 2026, preprint arXiv:2602.20140, DOI: 10.48550/ARXIV.2602.20140.

60 V. Gharakhanyan, Y. Yang, L. Barroso-Luque, M. Shuaibi, D. S. Levine, K. Michel, V. Bernat, M. Dzamba, X. Fu, M. Gao, X. Liu, K. Noori, L. J. Purvis, T. Rao, B. M. Wood, A. Rizvi, M. Uyttendaele, A. J. Ouderkirk, C. Daraio, C. L. Zitnick, A. Boromand, N. Marom, Z. W. Ulissi and A. Sriram, FastCSP: Accelerated Molecular Crystal Structure Prediction with Universal Model for Atoms, *arXiv*, 2025, preprint arXiv:2508.02641, DOI: 10.48550/ARXIV.2508.02641.

61 S. Grimme, A. Hansen, J. G. Brandenburg and C. Bannwarth, Dispersion-Corrected Mean-Field Electronic Structure Methods, *Chem. Rev.*, 2016, **116**, 5105–5154.

62 B. Hammer, L. B. Hansen and J. K. Nørskov, Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals, *Phys. Rev. B*, 1999, **59**, 7413–7421.

63 S. A. Ramos, L. J. Mueller and G. J. O. Beran, The interplay of density functional selection and crystal structure for accurate NMR chemical shift predictions, *Faraday Discuss.*, 2025, **255**, 119–142.

64 J. P. Perdew, K. Burke and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.

65 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *J. Chem. Phys.*, 2010, **132**, 154104.

66 S. Grimme, Semiempirical GGA-type density functional constructed with a long-range dispersion correction, *J. Comput. Chem.*, 2006, **27**, 1787–1799.

67 D. M. Grant, F. Liu, R. J. Iuliucci, C. G. Phung, J. C. Facelli and D. W. Alderman, Relationship of ^{13}C NMR chemical shift tensors to diffraction structures, *Acta Crystallogr., Sect. B: Struct. Sci.*, 1995, **51**, 540–546.

68 F. Weigend, Accurate Coulomb-fitting basis sets for H to Rn, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1057.

69 F. Weigend and R. Ahlrichs, Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297.

70 N. Mardirossian and M. Head-Gordon, ωB97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation, *J. Chem. Phys.*, 2016, **144**, 214110.

71 V. Zaverkin, D. Holzmüller, R. Schuldt and J. Kästner, Predicting properties of periodic systems from cluster data: A case study of liquid water, *J. Chem. Phys.*, 2022, **156**, 114103.

72 J. Daru, H. Forbert, J. Behler and D. Marx, Coupled Cluster Molecular Dynamics of Condensed Phase Systems Enabled by Machine Learning Potentials: Liquid Water Benchmark, *Phys. Rev. Lett.*, 2022, **129**, 226001.

73 Rowan Benchmarks, (accessed March 12, 2026), https://benchmarks.rowansci.com/methods/uma.

74 S. T. Holmes, R. J. Iuliucci, K. T. Mueller and C. Dybowski, Density functional investigation of intermolecular effects on ^{13}C NMR chemical-shielding tensors modeled with molecular clusters, *J. Chem. Phys.*, 2014, **141**, 164121.

75 S. T. Holmes, R. J. Iuliucci, K. T. Mueller and C. Dybowski, Critical Analysis of Cluster Models and Exchange-Correlation Functionals for Calculating Magnetic Shielding in Molecular Solids, *J. Chem. Theory Comput.*, 2015, **11**, 5229–5241.

76 S. T. Holmes, O. G. Engl, M. N. Srnec, J. D. Madura, R. Quiñones, J. K. Harper, R. W. Schurko and R. J. Iuliucci, Chemical Shift Tensors of Cimetidine Form A Modeled with Density Functional Theory Calculations: Implications for NMR Crystallography, *J. Phys. Chem. A*, 2020, **124**, 3109–3119.

77 D. W. Alderman, M. H. Sherwood and D. M. Grant, Comparing, Modeling, and Assigning Chemical-Shift Tensors in the Cartesian, Irreducible Spherical, and Icosahedral Representations, *J. Magn. Reson., Ser. A*, 1993, **101**, 188–197.

78 C. Ye, R. Fu, J. Hu, L. Hou and A. Et, Carbon-13 chemical shift anisotropies of solid amino acids, *Magn. Reson. Chem.*, 1993, **31**(8), 699–704.

79 A. C. Pinon, J. Schlagnitweit, P. Berruyer, A. J. Rossini, M. Lelli, E. Socie, M. Tang, T. Pham, A. Lesage, S. Schantz and L. Emsley, Measuring Nano- to Microstructures from Relayed Dynamic Nuclear Polarization NMR, *J. Phys. Chem. C*, 2017, **121**, 15993–16005.

80 I. Hung and Z. Gan, On the magic-angle turning and phase-adjusted spinning sidebands experiments, *J. Magn. Reson.*, 2010, **204**, 150–154.

81 A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng and K. W. Jacobsen, The atomic simulation environment—a Python library for working with atoms, *J. Phys.: Condens. Matter*, 2017, **29**, 273002.

82 D. C. Liu and J. Nocedal, On the limited memory BFGS method for large scale optimization, *Math. Program.*, 1989, **45**, 503–528.

83 E. Bitzek, P. Koskinen, F. Gähler, M. Moseler and P. Gumbsch, Structural Relaxation Made Simple, *Phys. Rev. Lett.*, 2006, **97**, 170201.

© 2026 The Author(s). Published by the Royal Society of Chemistry

84 P. Hohenberg and W. Kohn, Inhomogeneous Electron Gas, *Phys. Rev.*, 1964, **136**, B864–B871.
85 W. Kohn and L. J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, *Phys. Rev.*, 1965, **140**, A1133–A1138.
86 S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. I. J. Probert, K. Refson and M. C. Payne, First principles methods using CASTEP, *Z. Kristallogr. Cryst. Mater.*, 2005, **220**, 567–570.
87 D. Vanderbilt, Soft self-consistent pseudopotentials in a generalized eigenvalue formalism, *Phys. Rev. B*, 1990, **41**, 7892–7895.
88 J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew and J. Sun, Accurate and Numerically Efficient r^2SCAN Meta-Generalized Gradient Approximation, *J. Phys. Chem. Lett.*, 2020, **11**, 8208–8215.
89 C. Adamo and V. Barone, Toward reliable density functional methods without adjustable parameters: The PBE0 model, *J. Chem. Phys.*, 1999, **110**, 6158–6170.
90 A. D. Becke, A multicenter numerical integration scheme for polyatomic molecules, *J. Chem. Phys.*, 1988, **88**, 2547–2553.
91 M. Franchini, P. H. T. Philipsen and L. Visscher, The Becke Fuzzy Cells Integration Scheme in the Amsterdam Density Functional Program Suite, *J. Comput. Chem.*, 2013, **34**, 1819–1827.
92 A. Klamt and G. Schüürmann, COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient, *J. Chem. Soc., Perkin Trans. 2*, 1993, 799–805.
93 A. Klamt, Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena, *J. Phys. Chem.*, 1995, **99**, 2224–2235.
94 A. Klamt and V. Jonas, Treatment of the outlying charge in continuum solvation models, *J. Chem. Phys.*, 1996, **105**, 9972–9981.
95 A. Klamt, C. Moya and J. Palomar, A Comprehensive Comparison of the IEFPCM and SS(V)PE Continuum Solvation Methods with the COSMO Approach, *J. Chem. Theory Comput.*, 2015, **11**, 4220–4225.
96 C. C. Pye and T. Ziegler, An implementation of the conductor-like screening model of solvation within the Amsterdam density functional package, *Theor. Chem. Acc.*, 1999, **101**, 396–408.
97 B. Delley, The conductor-like screening model for polymers and surfaces, *Mol. Simul.*, 2006, **32**, 117–123.
98 N. L. Allinger, X. Zhou and J. Bergsma, Molecular mechanics parameters, *J. Mol. Struct.:THEOCHEM*, 1994, **312**, 69–83.
99 T. Dubroca, A. N. Smith, K. J. Pike, S. Froud, R. Wylde, B. Trociewitz, J. McKay, F. Mentink-Vigier, J. van Tol, S. Wi, W. Brey, J. R. Long, L. Frydman and S. Hill, A quasi-optical and corrugated waveguide microwave transmission system for simultaneous dynamic nuclear polarization NMR on two separate 14.1 T spectrometers, *J. Magn. Reson.*, 2018, **289**, 35–44.
100 F. J. Scott, T. Dubroca, R. W. Schurko, S. Hill, J. R. Long and F. Mentink-Vigier, Characterization of dielectric properties and their impact on MAS-DNP NMR applications, *J. Magn. Reson.*, 2024, **365**, 107742.
101 B. M. Fung, A. K. Khitrin and K. Ermolaev, An Improved Broadband Decoupling Sequence for Liquid Crystals and Solids, *J. Magn. Reson.*, 2000, **142**, 97–101.
102 A. Lesage, D. Sakellariou, S. Hediger, B. Eléna, P. Charmont, S. Steuernagel and L. Emsley, Experimental aspects of proton NMR spectroscopy in solids using phase-modulated homonuclear dipolar decoupling, *J. Magn. Reson.*, 2003, **163**, 105–113.
103 B. Kumari, M. Brodrecht, T. Gutmann, H. Breitzke and G. Buntkowsky, Efficient Referencing of FSLG CPMAS HETCOR Spectra Using 2D ^{1}H–^{1}H MAS FSLG, *Appl. Magn. Reson.*, 2019, **50**, 1399–1407.
104 M. Leskes, P. K. Madhu and S. Vega, A broad-banded z-rotation windowed phase-modulated Lee–Goldburg pulse sequence for ^{1}H spectroscopy in solid-state NMR, *Chem. Phys. Lett.*, 2007, **447**, 370–374.
105 R. Hafizi, H. Gittins, C. R. Taylor and G. M. Day, Towards Foundation Models Trained from Crystal Structure Prediction of the Organic Molecular Solid State. *ChemRxiv*, 2026, preprint, DOI: 10.26434/chemrxiv.15000584/v1.
106 Y. Yang, R. Tom, J. A. G. L. Wui, J. E. Moussa and N. Marom, Genarris 3.0: Generating Close-Packed Molecular Crystal Structures with Rigid Press, *J. Chem. Theory Comput.*, 2025, **21**, 11318–11332.

© 2026 The Author(s). Published by the Royal Society of Chemistry