Improving Predicted Nuclear Magnetic Resonance Chemical Shifts Using the Quasi-Harmonic Approximation

Jessica L. McKinley and Gregory J. O. Beran*

Department of Chemistry, University of California, Riverside, California 92521, United States

E-mail: gregory.beran@ucr.edu

Phone: +1 951 827-7869

Abstract

Ab initio nuclear magnetic resonance (NMR) chemical shift prediction plays an important role in the determination or validation of crystal structures. The ability to predict chemical shifts more accurately can translate to increased confidence in the resulting chemical shift or structural assignments. Standard electronic structure predictions for molecular crystal structures neglect thermal expansion, which can lead to an appreciable underestimation of the molar volumes. This study examines this volume error and its impact on 68 ^{13}C and 28 ^{15}N predicted chemical shifts taken from twenty molecular crystals. It assesses the ability to recover more realistic room-temperature crystal structures using the quasi-harmonic approximation and how refining the structures impacts the chemical shifts. Several pharmaceutical molecular crystals are also examined in more detail. On the whole, accounting for quasi-harmonic expansion changes the ^{13}C and ^{15}N chemical shifts by 0.5 and 1.0 ppm on average. This in turn reduces the root-mean-square errors relative to experiment by 0.3 ppm for ^{13}C and 0.7

ppm for ^{15}N. While the statistical impacts are modest, changes in individual chemical shifts can reach multiple ppm. Accounting for thermal expansion in molecular crystal chemical shift prediction may not be needed routinely, but the systematic trend toward improved accuracy with experiment could be useful in cases where discrimination between structural candidates is challenging, as in the pharmaceutical theophylline.

1 Introduction

Structural characterization of molecular crystals is fundamental in chemistry. Single-crystal diffraction techniques represent the gold standard for structure determination, but alternative strategies are needed for cases where suitable single crystals are not obtainable. Nuclear magnetic resonance (NMR) crystallography represents one such alternative approach. It combines solid-state NMR, powder x-ray diffraction, and *ab initio* chemical shielding predictions to solve crystal structures.[1–9] Diffraction methods characterize the long-range order effectively, but they can sometimes have difficulty resolving local features such as protonation states. Solid-state NMR complements powder x-ray diffraction by providing detailed information about the local chemical environments. However, mapping from the chemical shifts in an NMR spectrum to a three-dimensional crystal structure can be difficult.

Computational chemical shift predictions can facilitate NMR-driven structure determination. However, doing so requires one or more candidate crystal structures on which to perform the chemical shift predictions. One might obtain such structures from partial solutions from ambiguous powder x-ray diffraction data.[10] Alternatively, crystal structure prediction can also provide a powerful and increasingly reliable candidate structures, as evidenced by recent blind tests of crystal structure prediction.[11–13] Given a set of candidate structures, the central computational challenge in NMR crystallography becomes the discrimination between correct and incorrect structures. The more accurately the chemical shifts can be predicted, the greater the discrimination between candidate structures. For example, switching from a generalized gradient approximation (GGA) density functional like PBE to a hybrid func-

tional like B3LYP or PBE0 reduces the chemical shift errors by 20–30%[14,15] and increases the structural discrimination.[16]

The electronic structure model used to predict the chemical shifts is not the only source of error, however. The quality of the predicted geometry also matters. If the experimental structure is known, DFT refinement of the atomic positions (both hydrogen and heavy atoms) while constraining the experimental lattice parameters frequently leads to better agreement with neutron diffraction structures[17] and smaller differences between the measured and predicted chemical shifts.[14,18] The situation becomes more challenging when the experimental structure is unknown and has been predicted via crystal structure prediction or other similar lattice energy minimization modeling techniques. Whereas typical solid state NMR experiments are performed at room temperature, most crystal structure prediction studies are performed with classical force fields or electronic structure models that predict the electronic energy, rather than the free energy. By doing so, they neglect the zero-point vibrational and finite temperature effects that lead to thermal expansion. While a typical molecular crystal might expand its molar volume by ∼3–4% between the electronic minimum energy structure and the experimental room temperature structure, systems dominated by weaker intermolecular interactions such as van der Waals or π–π stacking can exhibit volume expansion of nearly ∼9% (Figure 1). Surveys by Nyman and Day found that accounting for these free energy effects can reorder the relative energetic stabilities of molecular crystal polymorphs in up to 21% of cases.[19,20]

Various strategies exist for incorporating free energy effects into the simulation. Molecular dynamics techniques naturally account for finite-temperature effects,[9] though standard classical dynamics omits zero-point vibrational contributions that impact crystal volumes[21] and chemical shifts[22] appreciably. Configurational/dynamical sampling is essential in biomolecules,[23–25] though it often has smaller impacts in densely packed molecular crystals.[5,26,27] Disordered crystals[5] and nuclear quantum effects[28] can represent important exceptions where configurational sampling in crystals is important, however. Unfortunately,

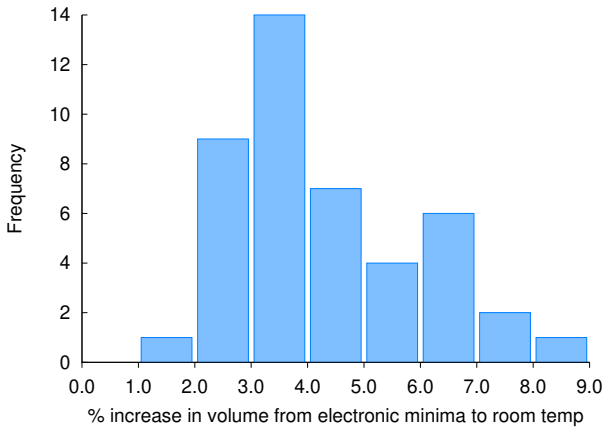

Figure 1: Total volume expansion of 44 molecular crystals. On average the unit cell will expand by 4.3% of its total volume. See SI Section 8 for the structures used to generate this graph.

empirical force fields are not always sufficiently accurate to generate appropriate ensembles,[29] and the challenges of developing effective force fields that can capture subtleties such as the balance of inter- and intramolecular forces in molecular crystals are well-known.[30–32] *Ab initio* molecular dynamics can overcome these challenges in principle, though the computational cost is often prohibitive in practice.

Alternatively, the quasi-harmonic approximation (QHA) provides a computationally practical strategy for estimating free energy contributions and capturing thermal expansion and its impact on crystal properties. The QHA models how the Gibbs free energy depends on temperature and pressure by incorporating a dependence of the harmonic phonons with volume. A number of studies have shown that use of the QHA with DFT and other correlated wavefunction approaches allows the prediction of small-molecule structures and their associated properties in excellent agreement with experiment.[21,33–42] Of course, the QHA model has its own limitations. Its static harmonic approach neglects vibrational averaging/dynamical motions,[22] is less well-suited to disordered crystals, and is expected to break down at higher temperatures, for example.

Here we investigate how accounting for thermal expansion in molecular crystals via the quasi-harmonic approximation refines isotropic ^{13}C and ^{15}N NMR chemical shift predictions.

First, using a set of twenty small-molecule molecular crystals, we demonstrate that accounting for thermal expansion effects increases the accuracy of the chemical shift predictions moderately, and that the chemical shifts computed using quasi-harmonic room-temperature structures rival those based on the experimental unit cell parameters. While the statistical reductions in error are modest on the whole, selected cases are found where the improvements to individual chemical shifts are appreciable. In other words, properly accounting for thermal expansion may be helpful in NMR crystallography applications where discriminating between different candidate structures proves challenging. Second, the capabilities of this QHA technique are demonstrated by applying the approach to several pharmaceutical crystals. Finally, we conclude by examining how QHA-based chemical shift refinement helps in the context of discriminating among different candidate structures of theophylline, which represents a challenging case for NMR crystallography.

2 Theory

2.1 Quasi-Harmonic Approximation

To model molecular crystals at finite temperatures and pressures the Gibbs free energy must be computed. The Gibbs free energy is composed of the electronic internal energy U_{el}, the Helmholtz vibrational free energy F_{vib}, and a pressure-volume (PV) contribution,

$$G(T, P) = U_{el} + F_{vib}(T) + PV \tag{1}$$

At ambient pressure, the PV term contributes negligibly to the total energy and hence will be neglected in this work. In that situation, the Gibbs free energy reduces to the Helmholtz free energy F.

The internal electronic energy U_{el} is computed here via dispersion-corrected planewave DFT. The Helmholtz vibrational free energy is computed from standard harmonic oscillator

vibrational partition functions as,

$$F_{vib}(T) = N_a \sum_i \left(\frac{\hbar\omega_i}{2} + k_b T \ln \left[1 - \exp\left(-\frac{\hbar\omega_i}{k_b T} \right) \right] \right) \tag{2}$$

where N_a is Avogadro's number, $\hbar$ is Plank's constant, k_b is the Boltzmann constant, and ω_i is the vibrational frequency of mode i. The first term corresponds to the zero-point vibrational contribution, while the second term describes the temperature dependence. For computational expedience, the phonons are computed only at the Γ point in the present work.

The crystal structures which minimize the electronic energy U_{el} (i.e. those obtained from relaxing both atomic positions and unit cell parameters) are referred to as the reference structure. We compute harmonic phonons for the reference crystal (ω_i^{ref}) to ensure the optimized structure lies at a minimum on the potential energy surface. Electronic energy versus volume curves $U_{el}(V)$ are then mapped out by minimizing the DFT energy of the reference cell under positive (cell compression) or negative pressure (cell expansion). This approach allows the cell to deform anisotropically along the lowest-energy path, which can be important for reliable modeling of the free energies.[43,44]

In principle, one ought to re-evaluate the phonons at every volume along this potential energy surface slice, but that is computationally impractical given the cost of typical electronic structure theory calculations. Instead we employ mode-specific Grüneisen parameters γ_i to estimate how individual phonon modes vary with unit cell volume,

$$\gamma_i = -\left(\frac{\partial \ln \omega_i}{\partial \ln V} \right) \tag{3}$$

Integrating Eq 3 yields,

$$\omega_i = \omega_i^{ref} \left(\frac{V}{V^{ref}} \right)^{-\gamma_i} \tag{4}$$

which allows the reference phonons ω_i^{ref} to be scaled to any given new volume V.

To compute these Grüneisen parameters, harmonic phonons are computed at one expanded and one compressed volume selected from those geometries obtained along the $E(V)$ curve. The selected structures were typically chosen to be about ± 10 Å^3 away from the reference volume here.[21,37] The Grüneisen parameter for each mode is then obtained via finite difference of these phonons computed on the expanded and compressed structures.

Using the calculated reference phonons and the mode-specific Grüneisen parameters, the Helmholtz vibrational free energy is evaluated at each volume in the electronic energy curve over a range of temperatures. Summing these two curves together gives the free energy. Since the particular volume which minimizes the free energy for a given temperature is unlikely to correspond to one of the sampled volumes, each free energy curve is fitted to a weighted double-Murnaghan equation of state. The Murnaghan equation of state is given by,

$$F(V) = F_0 + \frac{B_0 V}{B_0'} \left[\frac{(V_0/V)^{B_0'}}{B_0' - 1} + 1 \right] - \frac{B_0 V_0}{B_0' - 1} \tag{5}$$

where F_0, V_0, B_0, and B_0' are the fit parameters. F_0 gives the free energy at the minimum, V_0 is the molar volume at the minimum energy, B_0 is the bulk modulus, and B_0' is the first derivative of the bulk modulus with respect to pressure. This method identifies the free-energy minimum effectively while avoiding artifacts that can be caused by simpler equation of state fits or splines.[43]

Once the free energy equation of state $F(V)$ has been obtained at a chosen temperature, the free energy can be minimized to find the optimal molar volume. The lattice constants and atomic positions at this optimal volume are interpolated based on the explicitly optimized structures that were obtaining in generating $U_{el}(V)$. The atomic positions are then relaxed subject to those lattice constants being fixed to ensure the structure is at a minimum on the QHA free energy surface.

The computational cost of these QHA structure determinations is considerably higher than that of the conventional approach, which would typically involve only a single relax-

ation of the crystal structure followed by a single chemical shift prediction. In the QHA approach used here, the crystal was relaxed at ∼15-20 different pressures to map out $U_{el}(V)$. Furthermore, the harmonic phonon frequencies were computed three volumes. Each Hessian was evaluated via finite difference of the forces. Each force calculation that contributes to evaluation of the Hessian exhibits the same formal scaling with system size as conventional DFT optimization, but the number of displacements required for the finite difference depends on both the number of atoms and symmetry of the system. For systems of the sort studied here, evaluation of the Hessian has a computational cost of ∼2-5 times that of a geometry optimization in practice. Lastly, one final fixed-cell optimization is employed on the predicted structure at each given temperature. Overall, the computational cost of this QHA approach is ∼20–40 times higher than a conventional structure optimization. It might be possible to reduce the number of structure optimizations performed to construct U_{el}, but the QHA costs will remain at least an order of magnitude higher than that of a conventional calculation regardless.

Note that once the electronic energy curve and phonons have been computed, one can readily predict the structure at multiple different temperatures. Determining the QHA cell volume/lattice parameters at each temperature has trivial computational cost. However, at each different temperature, a single fixed-cell geometry optimization is performed to ensure the correct atomic positions are obtained. The computational cost of the subsequent chemical shielding calculation on the final structure is independent of how one obtains that structure (with or without QHA), and it typically comprises only a small fraction of the overall computational costs.

2.2 *Ab initio* Fragment-based Shielding Calculations

Periodic density functional theory (DFT) has long been used for modeling chemical shift calculations as they are well suited to describe extended crystal systems. In particular the planewave DFT-based gauge-inducing projector augmented wave (GIPAW)[45,46] is one

of the more commonly used models and has shown great success in NMR crystallography.[5,9] Modern fragment- and cluster-based methods[47–49] provide a competitive alternative to planewave methods for chemical shift prediction. Due primarily to their ability to use hybrid density functionals with much lower computational cost, fragment methods can frequently out-perform the more commonly used GIPAW method.[15,16,18] Switching from the GGA functional PBE to the hybrid PBE0 improves the accuracy of the chemical shifts by ∼20-30%.

Fragment-based methods decompose the total energy of the system into a many-body expansion,

$$E_{total} = \sum_i E_i + \sum_{ij} \Delta^2 E_{ij} + \sum_{ijk} \Delta^3 E_{ijk} + ... \tag{6}$$

where E_i represents the energy of the isolated molecule, $\Delta^2 E_{ij}$ is the pairwise interaction energy between molecules i and j, and $\Delta^3 E_{ijk}$ is the three-body interaction energy between molecules i, j, and k. Differentiating this expression with respect to the α-th component of the external magnetic field B and the nuclear magnetic moment μ produces a many-body expansion for the chemical shielding tensor σ on atom A,

$$\tilde{\sigma}^A = \sum_i \sigma_i^A + \sum_{ij} \Delta^2 \sigma_{ij}^A + \sum_{ijk} \Delta^3 \sigma_{ijk}^A + ... \tag{7}$$

This expression corresponds to expressing the chemical shielding of atom A in the crystal, $\tilde{\sigma}^A$, in terms of the chemical shielding σ^A computed for the isolated monomer plus a series of corrections to that shielding due that monomer's interactions with other molecules in the lattice. Given the high computational cost of computing three-body (trimer) interactions, Eq 7 is truncated after the two-body terms. To account for the polarization/electrostatic effects that are neglected by this truncation, the monomer and dimer calculations are electrostatically embedded in a set of self-consistent point charges designed to mimic the crystalline lattice,

$$\tilde{\sigma_i}^A = \sum_i \sigma_i^{A,emb} + \sum_{ij} \Delta^2 \sigma_{ij}^{A,emb} \tag{8}$$

See Ref 15 for details. The fragment approach generally captures the impact of the crystalline environment on chemical shieldings well.[14,15] Furthermore, it reduces the overall computational cost of the chemical shift calculation and can take advantage of highly parallel software implementations which enable the chemical shifts of even complicated crystals to be computed within a few hours.[14]

3 Computational details

3.1 Systems studied

To evaluate the effect thermal expansion has on predicted chemical shifts, 20 molecular crystals with 68 ^{13}C and 28 ^{15}N experimentally measured chemical shifts were chosen (Figure 2). These represent a subset of crystals from our earlier benchmark sets.[14,18] The CSD reference codes and experimental isotropic chemical shifts for these two test sets are:

- Carbon (12 structures, 68 shifts): HXACAN13,[50] GLYCIN03,[51] LALNIN12,[52] LSERIN01,[53] LTYROS11,[54] SUCROS04,[55] NAPHTA36[56] ADENOS12,[54] LTHREO01,[54] GLUTAM01,[54] ASPARM03,[57] and LCYSTN21.[54]

- Nitrogen (14 structures, 28 shifts): HXACAN13,[58] GLYCIN03,[59] ALUCAL04,[14] LSERIN01,[14] GLUTAM01,[14] ASPARM03,[57] LCYSTN21,[14] FUSVAQ01,[60] URACIL,[60] BAPLOT01,[61] LHISTD02,[62] BITZAF,[63] CYTSIN,[60] and THYMIN01.[60]

Several updates/modifications were made to these structures compared to the earlier works. Most importantly, key hydrogen positions were corrected for acetaminophen (HXACAN13) and theophylline (FUSVAQ01). Additionally, we now utilize experimental ^{13}C isotropic chemical shifts for ADENOS12, GLUTAM01, LCYSTN21, LTHREO01 and LTYROS11 which were recently re-measured by Dračínský et al,[54] correcting errors found in previously published literature. See SI Section 9 for the complete list of changes to the structures and reference data.

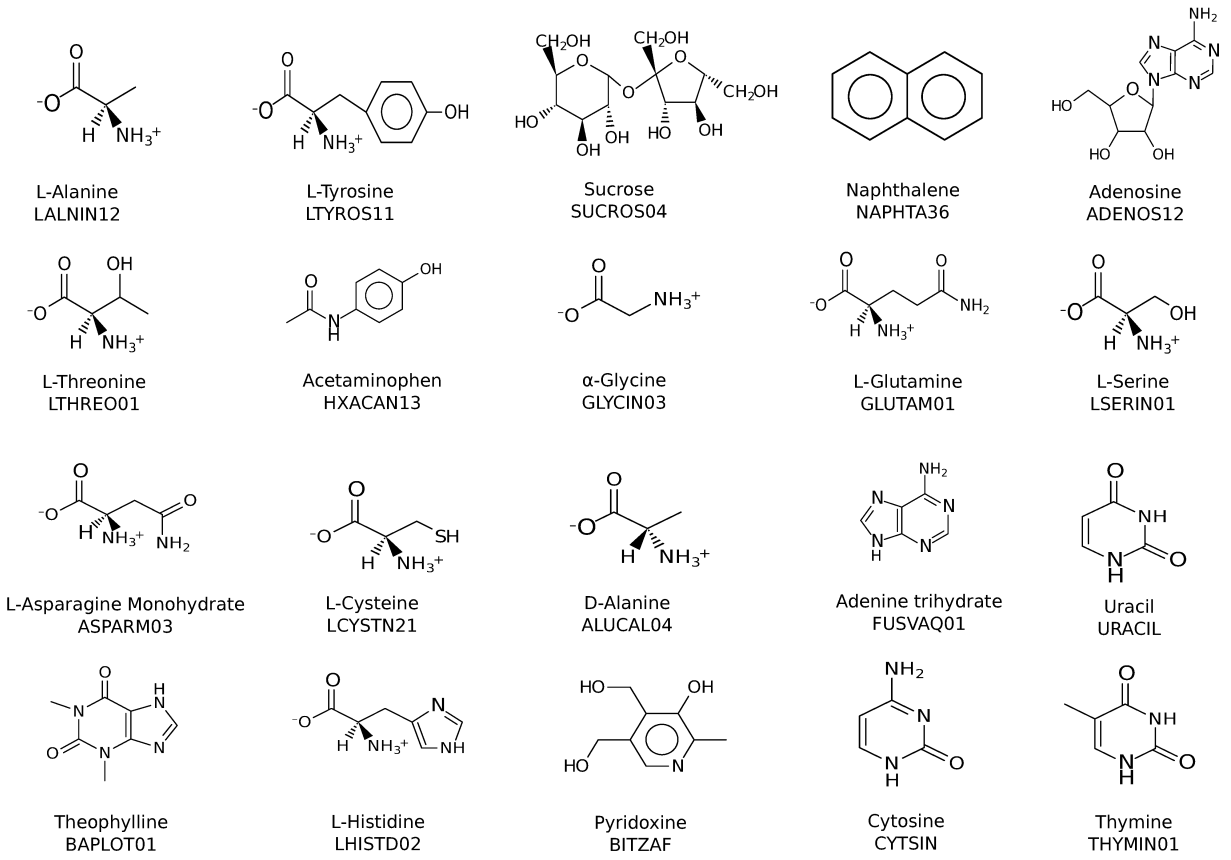

Figure 2: The twenty molecular crystals studied here, indicated by their species name and the CSD RefCode.

Finally, we investigate five pharmaceutical crystals more closely: acetaminophen form I (HXACAN13),[50] ibuprofen form I (IBPRAC16),[64] theophylline form II (BAPLOT01),[61] carbamazepine form III (CBMZPN23),[65] and enantiopure (*S*)-naproxen (COYRUD12).[66] Fixed-cell geometry optimizations and chemical shift calculations were performed for benchmark comparison on room temperature structures IBPRAC06, BAPLOT01, CBMZPN14, and COYRUD. Again, the acetaminophen room temperature structure was generated as described in SI section 9. Theophylline candidate structures generated via crystal structure prediction for Section 4.5 were obtained from the authors of Ref 61.

3.2 DFT structure optimizations

Three different optimizations were carried out for each of the structures mentioned here. The first is a full electronic energy optimization of both the atomic positions and the unit cell (labeled "No QHA"). The second uses the QHA to thermally expand the fully optimized structure to its free energy minimum at 0 K (labeled "0 K") and 300 K (labeled "300 K"). For the 0 K structure, the thermal expansion comes solely from zero-point energy (ZPE) contributions which is known to contribute up to 30% of the total expansion of the crystal at ambient temperature and pressure.[21] Finally, to benchmark against the known experimental structure, results were taken from a previous study[15] which performed fixed-cell optimizations on these structures (labeled "Expt. Cell"). In the pharmaceutical crystals for which earlier fixed-cell optimizations were not available, they were performed here.

For the calculations performed here, the crystals were first optimized with periodic boundary planewave DFT in Quantum Espresso v6.1,[67] using the B86bPBE density functional[68,69] and the exchange-hole dipole method (XDM) dispersion correction.[70] Core electrons were treated according to the projector augmented wave (PAW) approach, and PAW potentials for H, C, N, O, F, S, Cl, and Br were produced using A. Dal Corso's Atomic code v6.1.[71] Optimizations were carried out using a 50 Ry planewave energy cutoff. Reciprocal space k-points were placed with a Monkhorst-Pack grid[72] with a typical k-point grid spacing of 0.04 $Å^{-1}$ between nearest k-points and maximum spacings of no more than 0.09 $Å^{-1}$. See SI Section 3.3 for the specific k-point grid used for each structure. Γ-point harmonic vibrational frequencies were computed using finite difference with Phonopy v1.11.2.[73] Predicted room-temperature molar volumes are compared against the available experimental volumes from the CSD (SI Section 4).

The structures for the fixed-cell optimizations from Ref 15 were optimized using the PBE density functional,[69] the D2 dispersion correction,[74] and ultrasoft pseudopotentials. See Ref 15 for details. Although the specific modeling parameters used there differ from those here, test calculations suggest that re-optimizing the structures with the B86bPBE-XDM

protocol used here would impact the chemical shifts minimally. In other words, constraining the optimization to retain the experimental lattice parameters reduces the impact of the optimization protocol on the resulting structure. See SI Section 2.2 for details.

3.3 Chemical shielding calculations

Fragment-based NMR chemical shift calculations were carried out using the Hybrid Many-Body Interaction model (HMBI) v2.0 (available on github https://github.com/gberan/HMBI).[75–77] Chemical shielding calculations for individual fragments were carried out using Gaussian 09 revision E.01[78] with the PBE0 density functional.[79] This functional was chosen based on its excellent performance in earlier benchmarks.[14,15]

The chemical shielding calculations were carried out using the gauge-including atomic orbital (GIAO) approximation[80] and a locally-dense basis set.[81,82] All calculations employed a 6-311+G(2d,p) basis for atoms in the asymmetric unit, 6-311G(d,p) basis for neighboring atoms out to 4 Å, and a 6-31G basis for all atoms beyond 4 Å.[83–88] Previous work showed this mixed basis approach provides a noticeable speed-up on the computational cost with a minimal loss in chemical shift accuracy.[18] Prior testing indicates that basis sets of this size provide reasonably well-converged chemical shift predictions, especially when paired with the linear regression-based chemical shift referencing described in Section 3.4.[18,89,90] As described in previous work, a large DFT integration grid consisting of 150 radial and 974 Lebedev angular points was used to approach rotational invariance and mitigate numerical noise in the fragment calculations.[18]

The CHELPG atom-centered charges[91] used to construct the electrostatic embedding environment were computed using Gaussian 09 revision E.01.[78] Atom-centered charges were computed using the same PBE0 functional and 6-311+G(2dp,p) basis set. Point charges were placed on all molecules within 30 Å of any atom in the asymmetric unit, and these were surrounded by an additional set of point charges designed to mimic the Madelung potential of the full crystal at the nuclear centers in the asymmetric unit, as described in Ref 15.

3.4 Chemical shift referencing

The chemical shifts reported here are referenced relative to neat TMS under magic angle spinning (MAS) conditions for ^{13}C and solid NH_4Cl under MAS for ^{15}N. A linear regression scheme is used to map between the computed absolute chemical shieldings σ_i and the experimentally observed chemical shifts δ_i,

$$\delta_i = A\sigma_i + B \tag{9}$$

The regression parameters A and B were taken from an earlier study[15] which employed fixed, experimental lattice parameter optimizations on a set that includes all twenty crystals in the benchmark here plus many additional structures. For ^{13}C, the regression line is $\delta_i = -0.9658\sigma_i + 179.48$, while for ^{15}N, it is $\delta_i = -1.0106\sigma_i + 197.46$. Note that these same regression parameters determined using the experimental unit cells were employed for the No QHA, 0 K, and 300 K QHA structures.

3.5 Statistical testing

The paired Student t statistic is used to assess whether the changes in the error distributions for a given pair of models is statistically significant. The null hypothesis that states that the true mean difference between the paired samples is zero (i.e. the accuracy of the two models does not differ meaningfully at the selected significance level α). A rejection of the null hypothesis would indicate that the means of the two error distributions are statistically different from one another at the chosen significance level. The one-sided version of the test is used to assess whether the No QHA errors are significantly worse than the experimental cell ones and whether the 300 K QHA model provides significant reduction in error relative to the No QHA one. The two-sided version is employed to determine whether the 300 K QHA results differ meaningfully from the experimental cell ones.

Because a useful predictive model should be both accurate and precise, the analysis was

performed on the difference distributions obtained from the absolute (unsigned) values of the chemical shift errors e_i in each data set relative to experiment,

$$d_i = |e_i^{Model\ 2}| - |e_i^{Model\ 1}| \tag{10}$$

This ensures that the final t statistic effectively incorporates both shifting of the signed error distribution toward zero as well as any narrowing of the error distribution (reduction in the standard deviation). This choice primarily impacts the analysis of the carbon chemical shifts, for which all error distributions reported here are centered near zero, but the better-performing models exhibit smaller standard deviations.

If μ is the mean value of the d_i, $\mu = \sum_i d_i/N$, and the variance is given by $s^2 = \sum_i (d_i - \mu)^2/(N-1)$, the t statistic is computed as,

$$t = \sqrt{N}\frac{\mu}{s} \tag{11}$$

where $N = 68$ shifts for the ^{13}C test set and $N = 28$ shifts for the ^{15}N test set. The t statistic, the probability of obtaining such a large t statistic when the null hypothesis is true (p value), and critical value t_{crit} at the 0.05 significance level (i.e. 5% chance of rejecting the null hypothesis incorrectly) are reported. There are limitations to this analysis. For simplicity, the error in each chemical shift is assumed to be independent of the other errors, though one might argue that the errors for different atoms in a given crystal might be correlated due to systematic errors in the electronic structure modeling or the experimental measurement (incorrect shift referencing, for example). In addition, the crystal structures included in the benchmark calculations were selected based on the availability of experimental data, rather than representing a truly random sample of chemical space. The set includes a disproportionate number of amino acid crystals, for example.

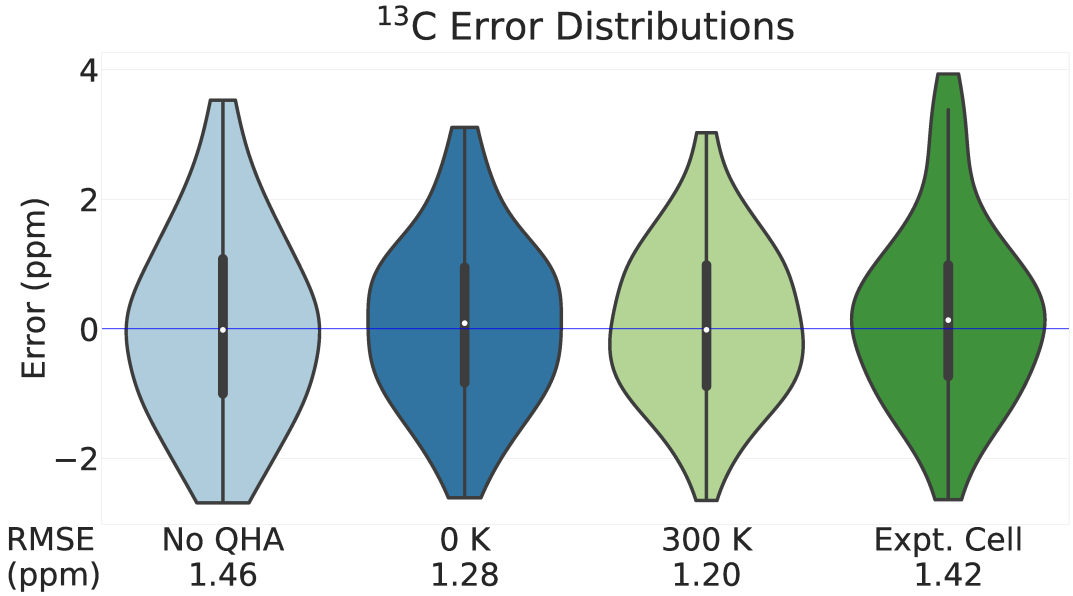

Figure 3: Distribution of errors relative to experiment over 68 ^{13}C chemical shifts for 12 molecular crystals. Overall RMS errors are displayed below each violin plot.

4 Results and Discussion

Section 4.1 examines how structure relaxation and subsequent thermal expansion impact the chemical shifts in the test set as a whole. Section 4.2 examines the performance of the QHA for reproducing experimental room temperature crystal volumes and how that relates to the accuracy of the chemical shifts. Section 4.3 considers the impact of thermal expansion on chemical shifts associated with different functional groups. Section 4.4 examines the impact of including thermal expansion in five pharmaceutical crystals: acetaminophen, ibuprofen, theophylline, carbamazepine, and naproxen, and Section 4.5 investigates how thermal expansion alters the ability to discriminate among several crystral structure prediction candidates for theophylline.

4.1 Accuracy of predicted chemical shifts

To begin, we first compare the distributions of errors in the predicted chemical shifts relative to experiment for the fixed-cell optimized (Expt Cell) and fully optimized (No QHA)

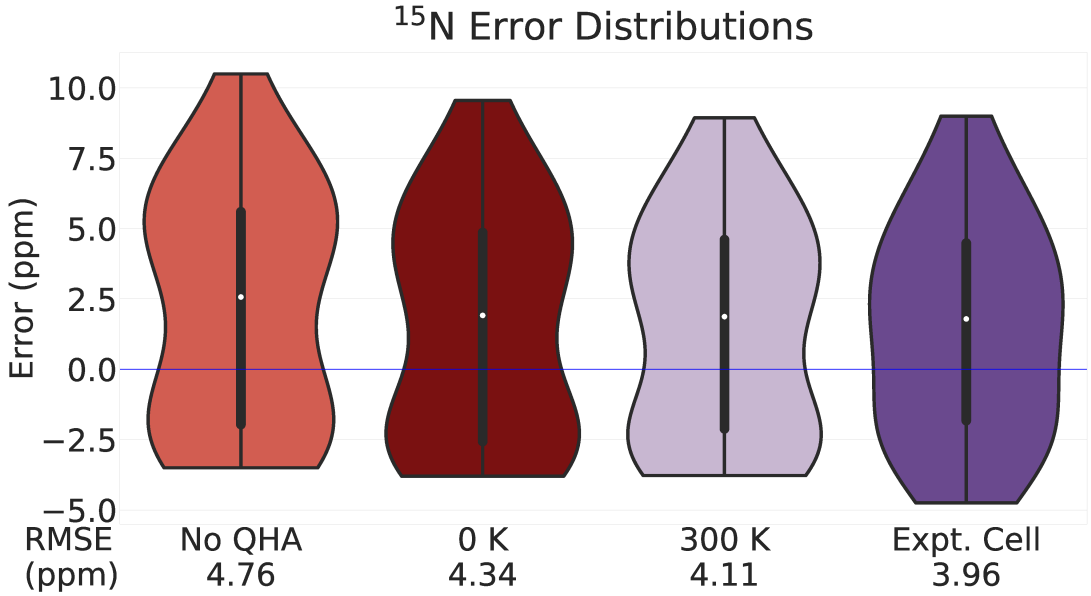

Figure 4: Distribution of errors relative to experiment over 28 ^{15}N chemical shifts for 14 molecular crystals. Overall RMS errors are displayed below each violin plot.

geometries shown in Figures 3 and 4. The data in these figures is represented using violin plots. Each violin curve represents a kernel density estimate for the chemical shift error distribution associated with a given model. Within each violin lies a box plot, in which the white dot denotes the median error, the surrounding black box indicates the middle 50th percentile, and the protruding whiskers highlight the extent of the largest errors (which also correspond to the ends of the violins). For the ^{13}C chemical shifts, using the experimental cell gives an overall root-mean-square (rms) error of 1.42 ppm. The error distribution exhibits a single maximum near zero error. For ^{15}N, the rms error is larger at 4.0 ppm, and the error distribution is slightly bimodal. This bimodal shape is retained across all four sets of structural models and probably reflects the smaller number of data points in the nitrogen set (28 for nitrogen versus 68 for carbon). The larger errors obtained for nitrogen shifts than carbon ones are consistent with the larger ^{15}N chemical shift range, its sensitivity to electrostatic environment, and prior chemical shift benchmarks.[14,92,93]

Fully relaxing the unit cells (i.e. without constraining the experimental lattice parame-

Table 1: Summary of computed paired t statistics, p values, and critical values t_{crit} at the $\alpha = 0.05$ significance level for the chemical shift errors. One-sided tests are used to assess whether the No QHA results are significantly worse than the Expt Cell ones and whether the 300 K QHA model provides significant improvement over the No QHA one. A two-sided test assesses whether the 300 K QHA errors differ meaningfully from the Expt Cell ones. Rejecting the null hypothesis and determining that the error distributions differ significantly requires $|t| > |t_{crit}|$.

	Expt Cell vs No QHA (one-sided)	No QHA vs 300 K (one-sided)	300 K vs Expt Cell (two-sided)
^{13}C t statistic	1.18 ($p = 0.12$)	-3.23 ($p < 0.001$)	1.48 ($p = 0.14$)
excluding LCYSTIN21	2.04 ($p = 0.023$)	-2.93 ($p = 0.002$)	0.72 ($p = 0.47$)
$t_{crit}(\alpha = 0.05)$, 67 d.o.f.	1.67	-1.67	2.00
^{15}N t statistic	2.34 ($p = 0.013$)	-1.86 ($p = 0.037$)	1.30 ($p = 0.20$)
$t_{crit}(\alpha = 0.05)$, 27 d.o.f.	1.70	-1.70	2.05

ters) leads to a 3.9% volume decrease on average, with several crystals decreasing by 6–8% (Tables 2 and 3). Surprisingly, this full relaxation only increases the ^{13}C isotropic shift error versus experiment trivially, from 1.42 to 1.46 ppm. However, this small change in the statistical error relative to experiment belies more substantial changes to the individual ^{13}C shifts. The statistical error relative to experiment conflates many factors, including structural errors, density functional approximations, basis set limitations, the fragmentation scheme, etc. In contrast, changes in the predicted chemical shifts highlight the impact of thermal expansion more directly. The average individual ^{13}C chemical shift changes by half a ppm upon cell relaxation, and the largest change is 2.4 ppm. In other words, the statistically similar ^{13}C rms chemical shift errors for the fixed cell and fully optimized structures is fortuitous (see statistical discussion below). For ^{15}N, full relaxation of the cell increases the errors more noticeably, from 4.0 to 4.8 ppm. The average individual ^{15}N shift change is 1.1 ppm, with a maximum change of 3.7 ppm.

Accounting for the expansion arising from zero-point vibrational energy (ZPVE) contributions (0 K) and thermal expansion (300 K) via the QHA expands the crystals appreciably. It also has noticeable impacts on the chemical shifts. By 300 K, the average ^{13}C chemical shift has changed by half a ppm, and chemical shifts in the upper quartile have changed

Table 2: For each crystal, the percent volume change and rmsd15 relative to the experimental structure are reported, along with the RMS error in predicted ^{13}C isotropic chemical shifts. These are reported for the fully optimized crystal (No QHA), the 0 K QHA structure, the 300 K QHA structure, and the structure optimized with the fixed experimental cell (Expt. Cell). For the Expt. Cell there is no change in volume, hence the percent volume change is not reported.

	No QHA			0 K QHA			300 K QHA			Expt. Cell	
Species	% Volume change	rmsd15 (Å)	RMSE (ppm)	% Volume change	rmsd15 (Å)	RMSE (ppm)	% Volume change	rmsd15 (Å)	RMSE (ppm)	rmsd15 (Å)	RMSE (ppm)
HXACAN13	-6.27	0.211	1.50	-2.82	0.143	1.31	0.34	0.123	1.21	0.055	1.08
GLYCIN03	-1.83	0.069	1.69	1.32	0.086	1.43	3.74	0.124	1.30	0.067	1.98
LALNIN12	-2.63	0.107	2.04	0.43	0.116	1.90	2.02	0.153	1.91	0.027	2.38
LSERIN01	-2.31	0.106	1.37	0.24	0.116	1.15	1.74	0.133	1.07	0.063	1.77
LTYROS11	-3.12	0.127	1.51	-0.18	0.058	1.25	1.43	0.061	1.15	0.024	1.38
SUCROS04	-3.24	0.154	1.36	-0.45	0.152	1.07	0.95	0.170	0.97	0.079	0.86
NAPHTA36	-8.13	0.171	1.16	-3.89	0.089	1.02	1.21	0.078	0.95	0.007	0.82
ADENOS12	-5.23	0.134	1.35	-1.65	0.079	1.28	0.10	0.078	1.19	0.049	1.29
LTHREO01	-2.94	0.090	1.25	-0.38	0.092	1.03	1.27	0.121	0.95	0.068	1.31
GLUTAM01	-2.52	0.103	1.63	0.06	0.084	1.56	1.44	0.090	1.54	0.068	1.83
ASPARM03	-2.82	0.120	1.66	-0.22	0.114	1.55	1.30	0.123	1.52	0.030	1.57
LCYSTN21	-2.50	0.151	1.27	-0.06	0.130	0.96	1.06	0.123	0.85	0.106	2.46

Table 3: For each crystal, the percent volume change and rmsd15 relative to the experimental structure are reported, along with the RMS error in predicted ^{15}N isotropic chemical shifts.

Species	No QHA			0 K QHA			300 K QHA			Expt. Cell	
	% Volume change	rmsd15 (Å)	RMSE (ppm)	% Volume change	rmsd15 (Å)	RMSE (ppm)	% Volume change	rmsd15 (Å)	RMSE (ppm)	rmsd15 (Å)	RMSE (ppm)
LSERIN01	-2.31	0.106	2.21	0.24	0.116	2.86	1.74	0.133	3.12	0.063	2.08
GLYCIN03	-1.83	0.069	2.21	1.32	0.086	3.07	3.74	0.124	3.56	0.067	1.49
FUSVAQ01	-6.74	0.148	3.55	-4.13	0.118	2.93	-2.32	0.092	2.52	0.048	2.08
URACIL	-3.74	0.080	7.53	-0.28	0.069	6.85	2.08	0.102	6.41	0.030	6.51
BAPLOT01	-4.66	0.193	5.83	-1.50	0.192	5.29	1.72	0.212	5.05	0.131	4.66
HXACAN13	-6.27	0.211	7.04	-2.82	0.143	7.52	0.34	0.123	7.31	0.055	7.07
LHISTD02	-3.32	0.102	5.03	-0.28	0.063	4.46	1.26	0.079	4.32	0.052	4.10
BITZAF	-4.04	0.191	4.48	-0.94	0.208	3.80	2.25	0.274	3.20	0.142	4.74
CYTSIN	-3.77	0.091	3.86	-0.41	0.094	3.83	1.45	0.093	3.83	0.061	3.45
LCYSTN21	-2.50	0.151	1.88	-0.06	0.130	2.75	1.06	0.123	3.13	0.106	4.56
GLUTAM01	-2.52	0.103	6.00	0.06	0.084	4.50	1.44	0.090	3.81	0.068	3.41
THYMIN01	-6.33	0.235	3.91	-2.29	0.166	3.55	4.13	0.153	3.05	0.056	3.43
ASPARM03	-2.82	0.120	3.86	-0.22	0.114	3.16	1.30	0.123	2.92	0.030	3.15
ALUCAL04	-3.08	0.110	3.47	0.08	0.113	3.73	1.95	0.157	3.77	0.028	3.11

by 0.7–1.2 ppm. The average ^{15}N chemical shift has changed by 1.0 ppm, and those in the upper quartile have changed by 1.5–3.3 ppm. Notably, accounting for thermal expansion eliminates some of the largest error shifts in Figures 3 and 4. For example, whereas 18% of the fully relaxed No QHA ^{13}C shifts exhibit errors greater than 2 ppm, only 10% are greater than 2 ppm in the 300 K QHA structures, and the maximum error drops from 3.5 ppm to 3.0 ppm. For ^{15}N, the number of shift errors larger than 5 ppm decreases from 32% in the fully relaxed structures to 18% in the 300 K QHA ones. These improvements manifest in the smaller rms errors of 1.2 ppm for ^{13}C and 4.1 ppm for ^{15}N at 300 K, which are competitive with or better than the errors obtained with the experimental unit cells. Even just accounting for the ZPVE expansion captures a sizable fraction of the impact of the overall chemical shift correction due to expansion, which is consistent with earlier QHA studies[21,39] that found that ZPVE contributes a sizable fraction of the overall expansion that separates the fully relaxed structures with no vibrational contributions and those modeled at room temperature via the QHA.

The fact that the 300 K QHA ^{13}C rms error of 1.2 ppm is smaller than the fixed experimental cell error of 1.4 ppm deserves further examination. The largest improvement in the predicted ^{13}C chemical shifts comes from L-cysteine. The experimental LCYSTN21 crystal structure used omits a hydrogen atom on the thiol group. In our previous work, this missing hydrogen atom was placed manually, and the structure was relaxed under fixed experimental cell conditions. However, upon full cell relaxation here, the hydrogen bonding network reconfigures to a more optimal arrangement. Apparently the combination of poor manual placement and constraining the unit cell parameters in the earlier relaxation prevented the necessary relaxation. This new arrangement is maintained in the 300 K QHA structures, and this reduces the error on the cysteine α and side-chain carbons by 2.5 and 1.5 ppm, respectively.

A series of paired t statistic tests were performed to assess the significance of these changes in the chemical shift error distributions between different models. Table 1 summarizes the

t statistics, p values, and critical t values at the 0.05 probability level. Focus first on the ^{15}N results, whose chemical shifts are more sensitive to crystal packing details. Paired comparison of the No QHA versus Expt Cell chemical shift errors results in a t statistic of 2.34, which exceeds the one-sided t_{crit} of 1.70 for 27 degrees of freedom (d.o.f.). This is consistent with the hypothesis that the ^{15}N No QHA shift errors are significantly worse than those obtained using experimental unit cells ($p = 0.013$). Comparing the No QHA and 300 K QHA results, the t statistic of -1.86 exceeds the one-sided critical value of -1.70 in magnitude ($|t| > |t_{crit}|$), indicating that accounting for thermal expansion results in a statistically significant reduction in the chemical shift errors ($p = 0.037$). Furthermore, the small t statistic ($t < t_{crit}$) obtained when comparing the 300 K shift errors to those found when using the experimental cell indicates that the differences between those two models are not statistically significant ($p = 0.20$).

The statistical analysis for the carbon chemical shift errors with 67 d.o.f produces a similar story: the shift errors for the QHA 300 K structures are meaningfully smaller than those from the No QHA ones ($p < 0.001$), and the 300 K QHA chemical shift error distribution is consistent with the one obtained using the experimental unit cell ($p = 0.14$). On the other hand, one finds that the No QHA and experimental cell shift errors do not differ significantly ($p = 0.12$), contrary to what was seen for the ^{15}N data. However, if one removes LCYSTN21 from the analysis based on the discussion regarding its thiol hydrogen above, the difference between the No QHA and experimental cell shift errors also become significant ($p = 0.023$). Overall, this statistical analysis is consistent with the argument that while fully relaxing the unit cell degrades the accuracy of the predicted chemical shifts, performing the QHA thermal expansion improves the quality of the chemical shift prediction and achieves an accuracy level comparable to what one obtains when using the experimental unit cell.

Finally, as described in Section 3.4, the regression parameters used here to map from chemical shielding to chemical shift are those determined from a larger set of fixed-cell ^{13}C and ^{15}N benchmarks used in an earlier study.[15] One might wonder how different the results

would be if the regressions were determined directly by fitting the data for these specific systems against experiment separately for each structure set (No QHA, 0 K, 300 K and experimental cell). Fitting the regressions to fixed-cell structures could conceivably lead to beneficial error cancellation in the fixed cell structures here that is lost in the fully relaxed and QHA structure sets. In practice, however, the changes in the ^{13}C rms errors upon re-fitting the regression parameters specifically for each set are a negligible 0.01 ppm. For nitrogen, whose chemical shifts tend to be more sensitive to intermolecular separation and other packing details, refitting the regression parameters would decrease the chemical shift errors by a few tenths of a ppm, but it would not substantially affect the general trends. In that scenario, the rms errors become 4.3, 3.9, and 3.7 ppm for the No QHA, 0 K, and 300 K ^{15}N sets, respectively, versus 3.6 ppm for the experimental cell data. The rather weak dependence of the rms errors on the specific origin of the structures used to determine the regression parameters supports the use the regression parameters derived from our earlier work as described in Section 3.4.

4.2 Assessment of QHA thermal expansion

Having seen that accounting for thermal expansion improves the quality of the predicted chemical shifts relative to those obtained from fully optimized crystal structures, we now investigate the impact of expansion in more detail. In particular, how well do the room-temperature QHA structures reproduce the experimental ones, and to what extent are the error reductions correlated with the magnitude of the volume correction between the No QHA and 300 K structures? Table 2 lists the percent volume changes relative to the experimental structure, the rmsd15 metric, and the rms error in the chemical shifts relative to experiment for that structure. The rmsd15 metric measures the root-mean square deviations in the non-hydrogen atom positions between the optimized and experimental structure using a 15-molecule cluster.[94] Plots visualizing the QHA volume changes relative to the experimental volume are provided in SI Section 5.

As noted earlier, the full No QHA relaxations reduce the molar volumes of the 20 molecular crystals tested here by an average of 4.0%. The unit cell volumes for five crystals, naphthalene, acetaminophen, thymine, adenine trihydrate, and adenosine, contract by 5% or more. The relatively weak, non-specific van der Waals interactions between naphthalene molecules lead to a large 8.1% volume change between the fully optimized and experimental naphthalene crystal structures. Acetaminophen (6.7% change) and thymine (6.8%) form layered structures with low-dimensional hydrogen bond networks that allow facile expansion between layers. Adenosine exhibits a three-dimensional hydrogen bonding network, but the highly flexible intramolecular degrees of freedom may allow it to change volume more readily (5.2%), while the aforementioned issues with the room-temperature adenine trihydrate crystal structure creates the appearance of a large contraction upon relaxation.

Employing the QHA approximation allows the fully optimized crystals to expand back toward the room temperature structures. At 0 K, the predicted crystal volumes are on average 0.9% smaller than the room-temperature experimental ones (excluding adenine trihydrate). By 300 K, the predicted volumes have expanded too much, exceeding the experimental volumes by 1.5% on average. Paired one-sided t statistic comparisons indicate that both QHA sets represent a significant reduction in the volume errors relative to the No QHA results at the 0.05 significance level: $p < 10^{-5}$ for 0 K vs No QHA and $p = 0.018$ for 300 K vs No QHA. On the other hand, the 300 K QHA volume errors are significantly worse ($p = 0.043$) than the 0 K ones, as expected from the mean errors. See SI Section 5 for details.

For adenine trihydrate, the QHA structure with the energetically preferred hydrogen bonding arrangement remains 2.3% smaller than the experimental FUSVAQ structure at room temperature. The outlier nature of FUSVAQ suggests that it might be worthwhile to redetermine its room-temperature crystal structure experimentally. The tendency for the B86bPBE-XDM quasi-harmonic protocol employed here to over-estimate the molar volume is consistent with our earlier work on several small-molecule crystals.[39] That study also found that further refining the crystal structures with higher-level wave function models frequently

decreased the molar volumes relative to the B86bPBE-XDM calculations, bringing them into closer agreement with experiment. However, such wave function-based calculations are somewhat more computationally expensive,[95] so they are not pursued here.

Looking at the rmsd15 values, one unsurprisingly finds excellent agreement between the fixed-cell relaxations and the experimental structures (mean rmsd15 value of 0.06 Å). The rmsd15 values are roughly twice as large for the fully optimized, 0 K QHA, and 300 K QHA structures, with mean values ranging from 0.11-0.13 Å. While these values represent generally good agreement between the predicted and experimental structures, the fact that the rmsd15 values do not improve monotonically along the progression from No QHA to 0 K and 300 K structures further highlights the systematic over-estimation of the room-temperature volumes by the QHA model. In other words, employing the QHA to refine the fully optimized structures clearly improves the packing densities of the crystal structures relative to experiment. However, the improvement one would expect from heating the QHA model from 0 K to 300 K is not obvious.

Despite the 300 K QHA crystal structures being less accurate than the 0 K ones relative to experiment, the 300 K structures are clearly superior to the 0 K ones in terms of the accuracy of the predicted chemical shifts, as discussed in Section 4.1. Furthermore, the statistical similarity in the accuracy of the chemical shifts predicted for the 300 K QHA and fixed experimental cell structures suggests that, despite differences in the crystal structures, the QHA structures are reproducing the local chemical environments well. Examining the rms errors on a crystal-by-crystal basis in Tables 2 and 3, one sees that in a number of cases (especially for the ^{13}C), the shift errors for the 300 K QHA structures are smaller than those for the fixed experimental cells. One might expect to find a direct relationship between good structural agreement and small errors in the predicted chemical shifts, but no such obvious correlation exists.

Several factors might contribute to absence of a clear correlation between the quality of the structure and the quality of the chemical shifts. First, chemical shifts are inherently

governed by local electronic structure which is most sensitive to the local atomic environment. Perhaps the gross structural metrics employed here fail to capture more subtle local geometric refinements. Second, the experimental crystal structures may be imperfect. DFT can help refine or correct X-ray diffraction structures,[17,96] for example, and predicted chemical shifts from even high-quality neutron diffraction structures are improved by relaxing the atomic positions beforehand.[97] Third, even the reported experimental crystal unit cell volumes can vary considerably between different diffraction studies. For example, the CSD contains eight different crystal structures for form I acetaminophen at ambient temperature and pressure. These molar volumes for these structures span a range of 3%, with a standard deviation of ±1.2%. Fourth, experimentally observed chemical shifts reflect an ensemble average over dynamically moving atoms, while the calculations here assume a single static structure for each crystal/temperature. Finally, there are of course the errors inherent in density functional theory and the basis sets used here, including perhaps the use of different functionals for the geometry optimization (the GGA B86bPBE-XDM) and chemical shift prediction (the hybrid functional PBE0) steps.

4.3 Functional group analysis

For further insight into how accounting for thermal expansion improves the predicted chemical shifts, Figures 5 and 6 plot how the error associated with each predicted chemical shift changes from the fully relaxed structure (arrow tail) to the 300 K QHA structure (arrow head) for ^{13}C and ^{15}N chemical shifts, respectively. The shift changes are color-coded by functional group. See SI Section 3.1 for assignment details. Ideally, accounting for the expansion would shift the errors toward zero.

Consider first the ^{13}C chemical shifts in Figure 5. The magnitude of the average ^{13}C chemical shift changes by 0.5 ppm between the No QHA and 300 K QHA structures. The smallest changes occur for alkyl and amide carbons (∼0.2-0.3 ppm), while the remaining functional groups exhibit average changes ranging ∼0.4–0.6 ppm, and the largest individual

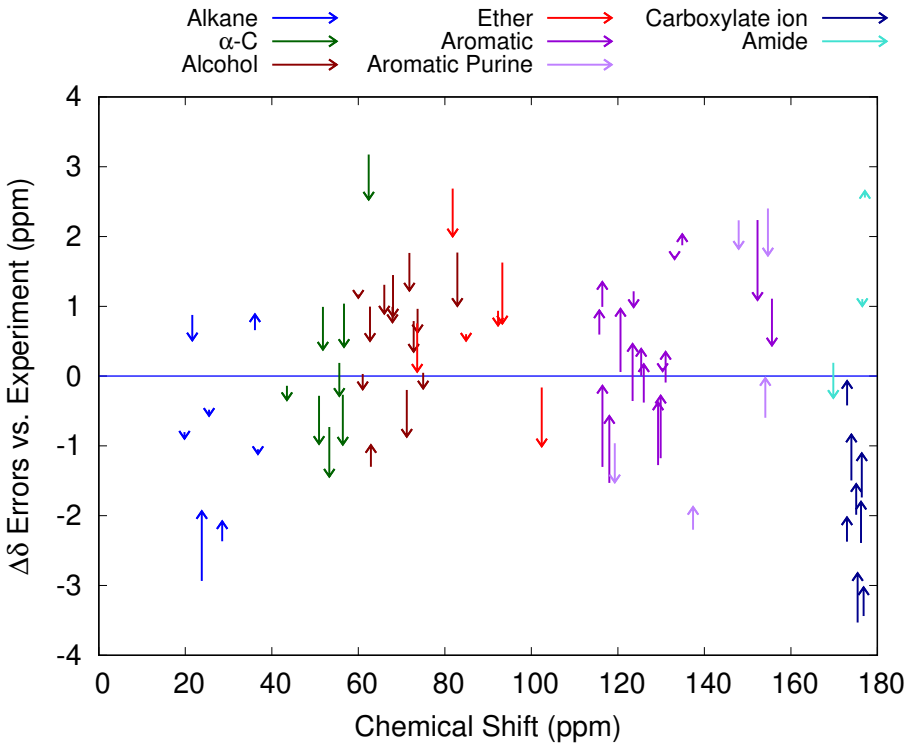

Figure 5: Impact of QHA expansion on the errors in individual predicted ^{13}C chemical shifts relative to experiment. Arrows indicate the change in the error from the fully optimized geometry to the 300 K QHA one. The chemical shifts are color coded by their functional group.

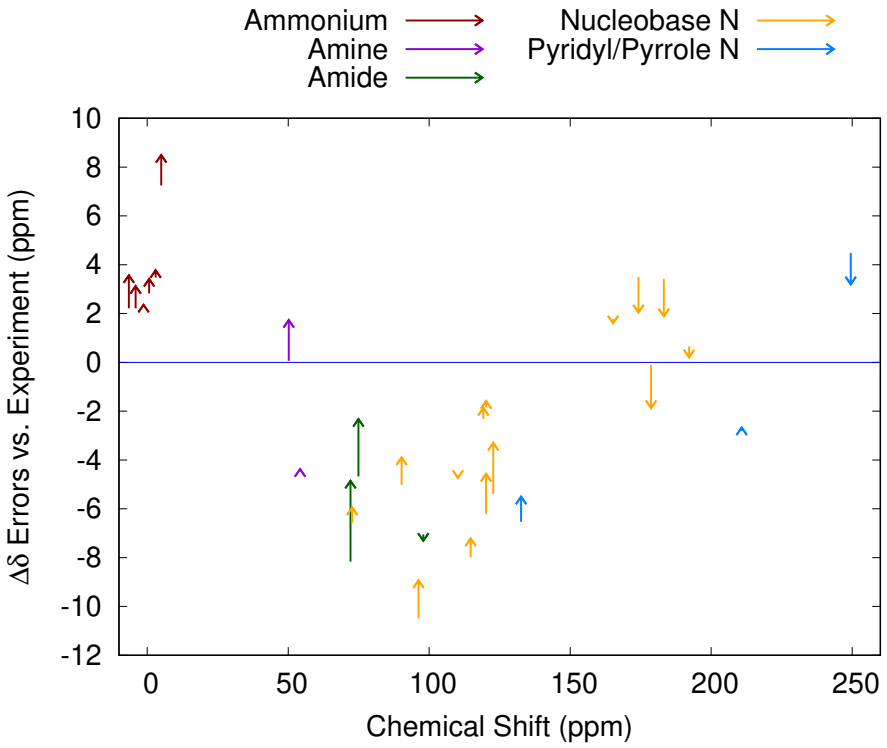

Figure 6: Impact of QHA expansion on the errors in individual predicted ^{15}N chemical shifts relative to experiment. Arrows indicate the change in the error from the fully optimized geometry to the 300 K QHA one. The chemical shifts are color coded by their functional group.

shift changes exceed 1 ppm. As can be seen from Figure 5, however, the shift changes do not always correspond to improved agreement with experiment. Defining the (arbitrary) threshold of 0.3 ppm for an appreciable change in the error relative to experiment, one finds that accounting or thermal expansion improves 47% of the ^{13}C shifts by 0.3 ppm or more relative to experiment, it has little impact on 37% of the shifts (i.e. changes the error by less than 0.3 ppm), and it increases the errors by more than 0.3 ppm for 16% of the shifts.

With the exception of the methyl carbon in acetaminophen, whose shift is improved by 1.0 ppm, the changes in the alkyl shifts upon thermal expansion are small and do not appreciably impact the agreement with experiment. These flexible groups are perhaps more impacted by factors such as dynamical motions that are not captured in the simple QHA model here. Thermal expansion corrections to the amide carbons have similarly small impacts on the agreement with experiment. The test set contains quite a few amino acids, and interestingly, thermal expansion consistently pushes all of the α-carbon shifts downfield by an average 0.6 ppm. However, this correction does not always prove beneficial, and the net agreement with experiment becomes negligibly worse (by 0.1 ppm). Mixed behavior is also seen for carbons adjacent to ether oxygens and in aromatic rings, but the net effects are toward improved agreement with experiment. The most interesting impact of thermal expansion occurs for the amino acid carboxylate carbons, which are uniformly improved by half a ppm on average. Experience suggests that the chemical shifts of carboxylate carbons are very sensitive to geometry, and the data here indicates that accounting for thermal expansion proves helpful.

Figure 6 plots the shift and error changes upon thermal expansion for the ^{15}N chemical shifts. As noted earlier, accounting for thermal expansion has an average 1.0 ppm impact on the nitrogen chemical shifts, double what was observed for ^{13}C. Using the same 0.3 ppm criterion as above, accounting for thermal expansion improves 50% of the 28 nitrogen shifts relative to experiment, has little impact on 29% of the shifts, and makes 21% of the shifts worse. Because the overall root-mean-square errors of the predicted chemical shifts are about three times larger for ^{15}N than for ^{13}C in these crystal test sets (e.g. 3.96 ppm vs 1.42 ppm for

the experimental cells), one might argue that a higher threshold should be used to indicate a notable change for ^{15}N chemical shifts. Tripling the selected threshold from 0.3 ppm to 0.9 ppm would classify 36% of the ^{15}N shifts as improving, 46% as not changing significantly, and 18% as becoming worse upon thermal expansion.

Interestingly, the shift changes for ammonium nitrogens are relatively modest and below average for the set (average 0.7 ppm), and they systematically shift toward worse agreement with experiment. These functional groups occur in the amino acid crystals and hydrogen bond with the carboxylate groups, whose ^{13}C shifts were systematically improved via the inclusion of thermal expansion (though they still remain some of the largest errors in the ^{13}C set). In reality, these functional groups are likely involved in a fast quantum mechanical proton exchange[28] between the carboxylate and ammonium groups that is not captured in the models here. It has also been demonstrated that DFT artificially stabilizes crystalline salts due to delocalization error,[98] and that might also lead to problematic treatment of these ammonium-carboxylate interactions.

Aside from the ammonium nitrogens, thermal expansion improves the remaining ^{15}N shifts by an average 0.8 ppm relative to experiment. Most of the shifts that change appreciably with thermal expansion shift toward better agreement with experiment. The largest gains occur for the terminal amides in L-glutamine and L-asparagine monohydrate near 72–75 ppm. The two most notable exceptions to the general improvement occur for the primary amine in cytosine at 50.2 ppm and the imine nitrogen in theophylline at 178.7 ppm.

4.4 Refining Chemical Shift Predictions in Pharmaceutical Crystals

Next, we examine the impact of accounting for thermal expansion in predicting the chemical shifts for five pharmaceutical crystals: acetaminophen form I, ibuprofen form I, theophylline form II, carbamazepine form III, and enantiopure (*S*)-naproxen (Figure 7). As in the previous sections, the crystal structures were fully optimized (No QHA) and then subsequentally

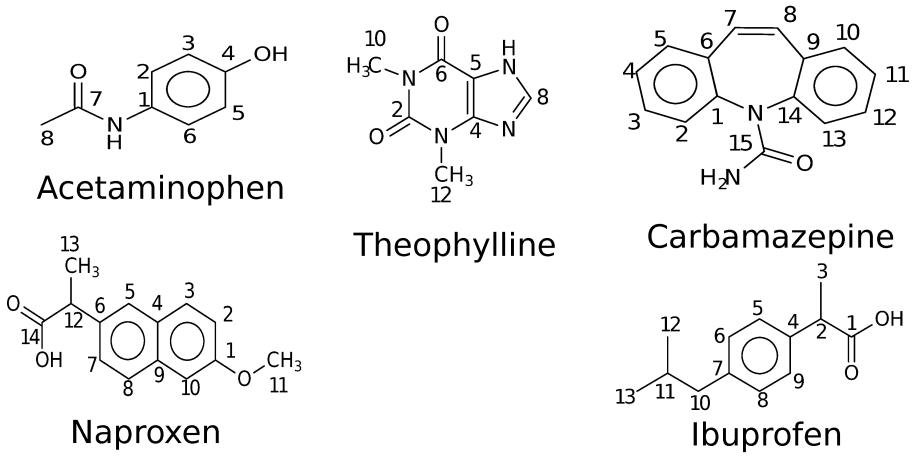

Figure 7: The five pharmaceutical species tested here and the carbon atom numbering for each.

expanded to 300 K using the QHA. Fixed-cell geometry optimizations and chemical shift calculations were performed for benchmark comparison. Note that while acetaminophen and theophylline were both included in the results presented above, ibuprofen, carbamazepine, and naproxen were not.

Figure 8 compares the predicted ^{13}C isotropic chemical shifts to the experimental solid-state NMR spectra. See Figure 7 for the atom numbering in each species. Volume expansion curves for each structure are available in SI Section 4 and the raw chemical shift data is available in SI Section 3.2. Root-mean-square errors (RMSE) between the predicted and experimental shifts are provided in each case. As was found for the crystals discussed in Section 4.2, the RMS error for each crystal generally improves from No QHA to 300 K. Once again, fully optimizing the crystal structure results in less accurate chemical shifts than using the fixed experimental cell structure. However more interesting details can be teased out by examining how each individual predicted chemical shift changes.

Acetaminophen, ibuprofen, and naproxen all show modest statistical improvement going from the fully optimized structure down to the fixed cell. For acetaminophen, all shifts except C3/C5 near 115 ppm improve progressively toward the experimental peaks upon application of the QHA. The fixed cell resolves the problem with the C3/C5, but it predicts a much worse

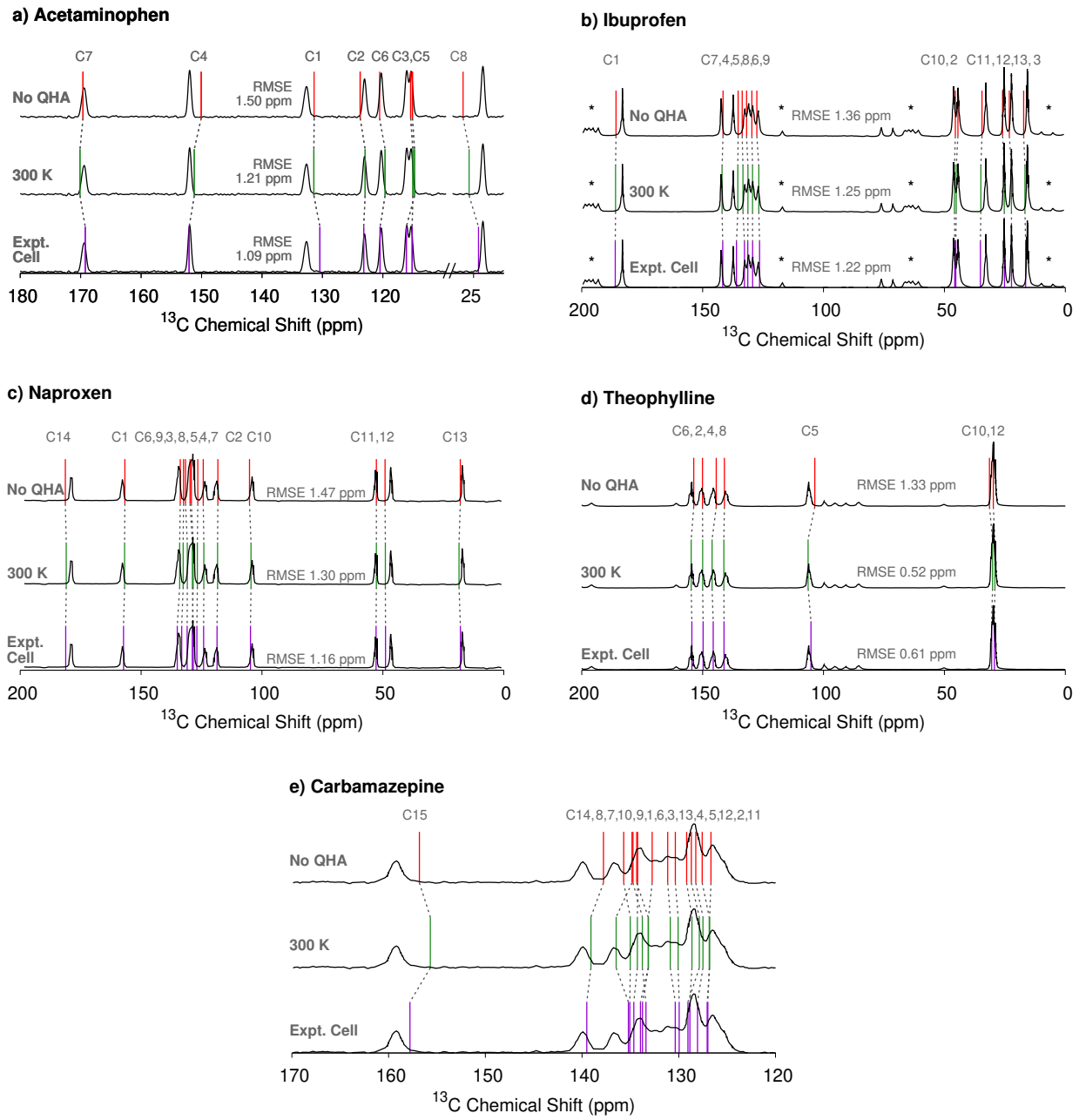

Figure 8: Impact of QHA structure refinement on the predicted ^{13} isotropic chemical shifts for a) acetaminophen, b) ibuprofen, c) naproxen, d) theophylline, and e) carbamazepine. Spinning side bands in the ibuprofen spectrum are marked with asterisks.

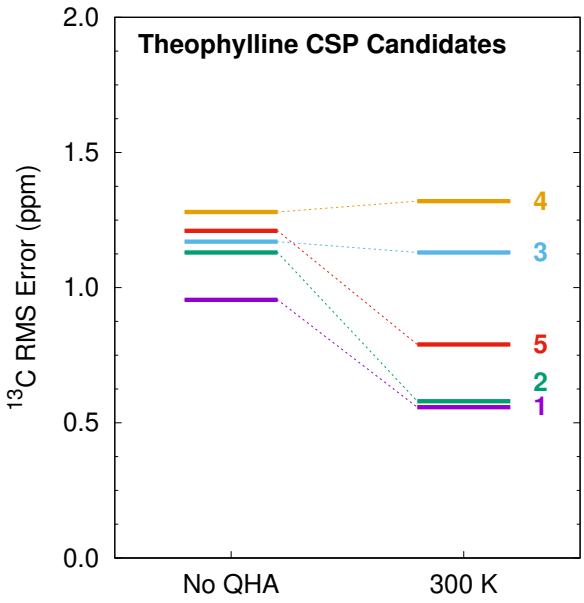

Figure 9: Impact of QHA structure refinement on the predicted ^{13}C isotropic chemical shifts for five candidate structures of theophylline.

shift for C1. Ibuprofen and naproxen exhibit relatively little temperature dependence in their spectra, though including thermal expansion does slightly improve the level of agreement between theory and experiment. For theophylline, the agreement is already quite good with the No QHA structure. However, performing QHA expansion to 300 K reduces the RMS error over half to only an exceptional 0.5 ppm, with every shift predicted in excellent agreement with experiment.

Finally, carbamazepine represents an interesting case. Most of the chemical shifts shown are concentrated in the aromatic region between ∼125–140 ppm. In that region, accounting for thermal expansion substantially refines and even re-orders many of the chemical shifts. In the absence of cross-correlation experiments or other clarifying information, re-ordering of the predicted chemical shifts might change how one assigns the experimental spectrum in this region. Unfortunately, the relatively low resolution of the experimental spectrum prevents clear assignment of specific peaks or calculation of an rms error. Nevertheless, it is clear visually that thermal expansion noticeably improves agreement with all peaks save the

amide group (C15). Indeed, the 300 K QHA ^{13}C spectrum is arguably in better agreement with experiment than the shifts predicted from the fixed cell structure.

Overall these five pharmaceutical species re-iterate how QHA thermal expansion generally improves the quality of the predicted chemical shifts and brings them to a level of agreement that is competitive with what is obtained from the experimental fixed-cell structures. In cases like theophylline or carbamazepine, the QHA expansion has a large impact on the agreement and even the potential chemical shift assignment. In others such as ibuprofen or naproxen, the impact is quite small. Since the QHA treatment is computationally expensive and it is not currently obvious *a priori* in which cases the effects of thermal expansion on the chemical shifts will be large, it appears that one should likely consider performing QHA refinement of the structures primarily in cases where the agreement with the experimental chemical shifts is unsatisfactory or where one is having difficulty discriminating between structural candidates.

4.5 Improved resolution of crystal candidates

Finally, an important potential use case for QHA refinement of structures would be in the context of NMR crystallography, where one might try to discriminate among a number of predicted crystal structures. In 2013, Baias et al examined this approach for four different species, including theophylline.[61] While the combination of structure and NMR chemical shift prediction worked well for most of the crystals, neither ^{13}C nor ^{1}H chemical shifts provided clear discrimination among the 44 candidate structures of theophylline, with many different crystal structures exhibiting similar errors in the predicted chemical shifts relative to experiment. Given the sizable improvement in the theophylline chemical shifts upon accounting for thermal expansion observed in Section 4.4, it is interesting to revisit this discrimination problem here. Specifically, we perform a proof-of-concept investigation of how accounting for QHA thermal expansion impacts the chemical shifts for the five lowest-energy structures from Ref 61.

Table 4: Relative electronic and quasi-harmonic Gibbs free energies for the five candidate theophylline crystal structures after B86bPBE-XDM refinement. Energies are relative to Structure 1.

Candidate Structure	Electronic Energy (kJ/mol)	Relative ΔG(300 K) (kJ/mol)
1	0.0	0.0
2	3.5	2.3
3	-0.6	2.1
4	6.7	5.6
5	0.8	1.1

The structures in Ref 61 were ranked using a hybrid B3LYP/6-31G(d,p) intramolecular energies and a classical intermolecular potential. Relaxing the five lowest-energy structures with periodic B86bPBE-XDM here leads to a considerable energetic re-ranking (Table 4). Structure 1 corresponds to the experimental structure, with an rmsd15 value of 0.212 Å relative to the experimental structure BAPLOT01. At the electronic energy level, structure 3 actually becomes slightly more stable than structure 1 after DFT refinement. However, the QHA Gibbs free energy restores structure 1 as the most stable. Structure 5 is also stabilized appreciably, becoming the second most stable structure, while structures 2 and 3 are nearly degenerate in free energy. Structure 4 lies nearly 6 kJ/mol above structure 1, suggesting it is a less likely candidate.

Next, consider the chemical shifts. Ideally, accounting for thermal expansion would improve the accuracy of the chemical shifts predicted for the correct structure, while increasing the errors for those from the incorrect structures. Figure 9 shows how the rms errors in the ^{13}C chemical shifts change with the inclusion of thermal expansion. Without any thermal expansion, the rms errors on all five structures span a range of only about 0.3 ppm, with structure 1 exhibiting marginally smaller errors than several others. Accounting for thermal expansion reduces the errors appreciably for structures 1, 2, and 5, while it has little impact on those for structures 3 and 4. This more than doubles the spread of rms chemical shift to 0.8 ppm. While the rms error for structure 5 is reduced, the difference between structures 1 and 5 remains about the same.

Interestingly, however, the 300 K QHA models predict nearly identical rms errors for structures 1 and 2. Examination of the crystal structures reveals that both crystal structures exhibit identical stacks of one-dimensional hydrogen bonded chains. The primary difference between the two structures lies in how these adjacent stacks of one-dimensional chains are oriented relative to one another in the third dimension. In structure 1, they are oriented at roughly 45° relative to one another, while in structure 2 they are oriented at 180°. In both cases, the inter-stack distances exceed 3 Å, suggesting a relatively weak interaction between the adjacent stacks. In other words, the difficulty in discriminating those two structures based on ^{13}C reflects the high similarity in crystal packing motifs and local chemical environments of these two structures. Overall, although identifying the correct theophylline crystal structure based on ^{13}C chemical shifts alone remains difficult, accounting for thermal expansion does generally increase the differentiation among these structures.

5 Conclusions

This work has demonstrated how the quasi-harmonic approximation can be used to refine isotropic NMR chemical shift predictions for molecular crystals. Employing the QHA recovers most or all of the chemical shift error accrued by neglecting the thermal expansion of the crystal. Indeed, the accuracy of chemical shifts computed using quasi-harmonic room-temperature structures rivals what is obtained for shifts computed with the experimental unit cell parameters. In select cases, using the QHA to thermally expand a fully relaxed structure led to slightly better fidelity with experimental chemical shifts than a fixed-cell simulation.

The impact of QHA thermal expansion on chemical shifts can be quite variable, however. No clear relationship between the quality of the structural agreement and the chemical shift errors was found, nor does the magnitude of the change of the chemical shifts upon thermal expansion seem to correlate strongly with the amount of thermal expansion that occurs. We

did find, however, that certain functional groups appeared to benefit more from the QHA treatment. In particular, functional groups which exhibit dynamic flexibility such as methyl or alkyl groups tended to benefit less from the QHA treatment than those associated with aromatic or other more rigid functionalities.

Although the QHA expansion calculations are considerably more expensive than traditional optimization and chemical shift prediction approaches, they are feasible for small-molecule pharmaceutical crystals and other chemically interesting species. The five pharmaceutical examples considered here provide a sampling of the diverse behaviors that can be observed with QHA refinement. In cases like ibuprofen or naproxen, thermal expansion has very little effect on the predicted chemical shifts, while in others like theophylline and carbamazepine, the QHA expansion noticeably improves the accuracy of the shifts and can even alter how one might assign them. Finally, it was also demonstrated how employing QHA thermal expansion could help increase the discrimination between candidate structures in a proof-of-concept examination of several theophylline structures that were generated via crystal structure prediction.

Given the modest statistical improvements to the chemical shifts resulting from the QHA thermal expansion models and the higher computational costs compared to conventional approaches, it may not worthwhile performing QHA expansion prior to predicting the chemical shifts on a routine basis. If the experimental lattice parameters are known, it is generally preferable to use that information in the chemical shift prediction. However, in situations where the experimental lattice parameters are unknown and better agreement with experiment is required, as in the context of crystal structure prediction, QHA refinement may prove helpful.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgement

Funding for this work from the National Science Foundation (CHE-1665212 to G.B. and Graduate Research Fellowship Grant 1326120 to J.M.), and supercomputer time from XSEDE (TG-CHE110064 to G.B. and TG-CHE170098 to J.M.) are gratefully acknowledged. Additional computational resources were provided by the UC Riverside high-performance computing center, which was funded by grants from the National Science Foundation (MRI-1429826) and National Institutes of Health (1S10OD016290-01A1). We thank Dr. Joshua Hartman for helpful discussions and Dr. Martin Dračínský for re-investigating the experimental spectra for many of these systems in Ref 54.

Supporting Information Available

Electronic Supplementary Information (ESI) available: [Full-opt vs. fixed-cell opt chemical shifts, 300 K fit, B86bPBE-XDM vs PBE-D2 fit, raw chemical shifts for all modeled species, volume expansion curves, plots correlating volumes to rmsd15 and rms error, and histograms detailing the effects of QHA on individual shifts, raw chemical shifts and volume expansion curves for theophylline CSP candidates, the species whose volumes were used to create Figure 1, and changes made to the crystal structures and reference data compared to some earlier studies.]

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Rajeswaran, M.; Blanton, T. N.; Zumbulyadis, N.; Giesen, D. J.; Conesa-Moratilla, C.; Misture, S. T.; Stephens, P. W.; Huq, A. Three-Dimensional Structure Determination of N -(p -Tolyl)-dodecylsulfonamide from Powder Diffraction Data and Validation of Structure Using Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* **2002**, *124*, 14450–14459.

(2) Harris, R. K. NMR crystallography: the use of chemical shifts. *Solid State Sci.* **2004**, *6*, 1025–1037.

(3) Harper, J. K.; Grant, D. M. Enhancing Crystal-Structure Prediction with NMR Tensor Data. *Cryst. Growth Des.* **2006**, *6*, 2315–2321.

(4) Salager, E.; Day, G. M.; Stein, R. S.; Pickard, C. J.; Elena, B.; Emsley, L. Powder Crystallography by Combined Crystal Structure Prediction and High-Resolution 1 H Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* **2010**, *132*, 2564–2566.

(5) Bonhomme, C.; Gervais, C.; Babonneau, F.; Coelho, C.; Pourpoint, F.; Azaïs, T.; Ashbrook, S. E.; Griffin, J. M.; Yates, J. R.; Mauri, F.; Pickard, C. J. 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.

(6) Baias, M.; Dumez, J.-N.; Svensson, P. H.; Schantz, S.; Day, G. M.; Emsley, L. 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.

(7) Kalakewich, K.; Iuliucci, R.; Harper, J. K. Establishing accurate high-resolution crystal structures in the absence of diffraction data and single crystals - An NMR approach. *Cryst. Growth Des.* **2013**, *13*, 5391–5396.

(8) Mueller, L. J.; Dunn, M. F. NMR Crystallography of Enzyme Active Sites: Probing Chemically Detailed, Three-Dimensional Structure in Tryptophan Synthase. *Acc. Chem. Res.* **2013**, 2008–2017.

(9) Ashbrook, S. E.; McKay, D. Combining Solid-State NMR Spectroscopy with First-Principles Calculations A Guide to NMR Crystallography. *Chem. Commun.* **2016**, *52*, 7186–7204.

(10) Yang, C.; Zhu, L.; Kudla, R. A.; Hartman, J. D.; Al-Kaysi, R. O.; Monaco, S.; Schatschneider, B.; Magalhaes, A.; Beran, G. J. O.; Bardeen, C. J.; Mueller, L. J. Crystal structure of the meta-stable intermediate in the photomechanical, crystal-to-crystal reaction of 9-tertbutyl anthracene ester. *CrystEngComm* **2016**, *18*, 7319–7329.

(11) Day, G. M.; Cooper, T. G.; Cruz-Cabeza, A. J.; Hejczyk, K. E.; Ammon, H. L.; Boerrigter, S. X. M.; Tan, J. S.; Della Valle, R. G.; Venuti, E.; Jose, J.; Gadre, S. R.; Desiraju, G. R.; Thakur, T. S.; van Eijck, B. P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W. M.; Neumann, M. A.; Leusen, F. J. J.; Kendrick, J.; Price, S. L.; Misquitta, A. J.; Karamertzanis, P. G.; Welch, G. W. A.; Scheraga, H. A.; Arnautova, Y. A.; Schmidt, M. U.; van de Streek, J.; Wolf, A. K.; Schweizer, B. Significant progress in predicting the crystal structures of small organic molecules—A report on the fourth blind test. *Acta Cryst. B* **2009**, *65*, 107–125.

(12) Bardwell, D. A.; Adjiman, C. S.; Arnautova, Y. A.; Bartashevich, E.; Boerrigter, S. X. M.; Braun, D. E.; Cruz-Cabeza, A. J.; Day, G. M.; Della Valle, R. G.; Desiraju, G. R.; van Eijck, B. P.; Facelli, J. C.; Ferraro, M. B.; Grillo, D.; Habgood, M.; Hofmann, D. W. M.; Hofmann, F.; Jose, K. V. J.; Karamertzanis, P. G.; Kazantsev, A. V.; Kendrick, J.; Kuleshova, L. N.; Leusen, F. J. J.; Maleev, A. V.; Misquitta, A. J.; Mohamed, S.; Needs, R. J.; Neumann, M. A.; Nikylov, D.; Orendt, A. M.; Pal, R.; Pantelides, C. C.; Pickard, C. J.; Price, L. S.; Price, S. L.; Scheraga, H. A.; van de Streek, J.; Thakur, T. S.; Tiwari, S.; Venuti, E.; Zhitkov, I. K. Towards crystal struc-

ture prediction of complex organic compounds—A report on the fifth blind test. *Acta Cryst. B* **2011**, *67*, 535–551.

(13) Reilly, A. M.; Cooper, R. I.; Adjiman, C. S.; Bhattacharya, S.; Boese, A. D.; Brandenburg, J. G.; Bygrave, P. J.; Bylsma, R.; Campbell, J. E.; Car, R.; Case, D. H.; Chadha, R.; Cole, J. C.; Cosburn, K.; Cuppen, H. M.; Curtis, F.; Day, G. M.; DiStasio Jr, R. A.; Dzyabchenko, A.; van Eijck, B. P.; Elking, D. M.; van den Ende, J. A.; Facelli, J. C.; Ferraro, M. B.; Fusti-Molnar, L.; Gatsiou, C.-A.; Gee, T. S.; de Gelder, R.; Ghiringhelli, L. M.; Goto, H.; Grimme, S.; Guo, R.; Hofmann, D. W. M.; Hoja, J.; Hylton, R. K.; Iuzzolino, L.; Jankiewicz, W.; de Jong, D. T.; Kendrick, J.; de Klerk, N. J. J.; Ko, H.-Y.; Kuleshova, L. N.; Li, X.; Lohani, S.; Leusen, F. J. J.; Lund, A. M.; Lv, J.; Ma, Y.; Marom, N.; Masunov, A. E.; McCabe, P.; McMahon, D. P.; Meekes, H.; Metz, M. P.; Misquitta, A. J.; Mohamed, S.; Monserrat, B.; Needs, R. J.; Neumann, M. A.; Nyman, J.; Obata, S.; Oberhofer, H.; Oganov, A. R.; Orendt, A. M.; Pagola, G. I.; Pantelides, C. C.; Pickard, C. J.; Podeszwa, R.; Price, L. S.; Price, S. L.; Pulido, A.; Read, M. G.; Reuter, K.; Schneider, E.; Schober, C.; Shields, G. P.; Singh, P.; Sugden, I. J.; Szalewicz, K.; Taylor, C. R.; Tkatchenko, A.; Tuckerman, M. E.; Vacarro, F.; Vasileiadis, M.; Vazquez-Mayagoitia, A.; Vogt, L.; Wang, Y.; Watson, R. E.; de Wijs, G. A.; Yang, J.; Zhu, Q.; Groom, C. R. Report on the sixth blind test of organic crystal structure prediction methods. *Acta Cryst. B* **2016**, *72*, 439–459.

(14) Hartman, J. D.; Kudla, R. A.; Day, G. M.; Mueller, L. J.; Beran, G. J. O. Benchmark fragment-based ^{1}H, ^{13}C, ^{15}N and ^{17}O chemical shift predictions in molecular crystals. *Phys. Chem. Chem. Phys.* **2016**, *18*, 21686–21709.

(15) Hartman, J. D.; Balaji, A.; Beran, G. J. Improved Electrostatic Embedding for Fragment-Based Chemical Shift Calculations in Molecular Crystals. *J. Chem. Theory Comput.* **2017**, *13*, 6043–6051.

(16) Hartman, J. D.; Day, G. M.; Beran, G. J. Enhanced NMR Discrimination of Pharmaceutically Relevant Molecular Crystal Forms through Fragment-Based Ab Initio Chemical Shift Predictions. *Cryst. Growth Des.* **2016**, *16*, 6479–6493.

(17) Deringer, V. L.; Hoepfner, V.; Dronskowski, R. Accurate Hydrogen Positions in Organic Crystals: Assessing a Quantum-Chemical Aide. *Cryst. Growth Des.* **2012**, *12*, 1014–1021.

(18) Hartman, J. D.; Monaco, S.; Schatschneider, B.; Beran, G. J. O. Fragment-based ^{13}C nuclear magnetic resonance chemical shift predictions in molecular crystals: An alternative to planewave methods . *J. Chem. Phys.* **2015**, *143*, 102809.

(19) Nyman, J.; Day, G. M. Static and lattice vibrational energy differences between polymorphs. *CrystEngComm* **2015**, *17*, 5154–5165.

(20) Nyman, J.; Day, G. M. Modelling temperature-dependent properties of polymorphic organic molecular crystals. *Phys. Chem. Chem. Phys.* **2016**, *18*, 31132–31143.

(21) Heit, Y. N.; Beran, G. J. O. How important is thermal expansion for predicting molecular crystal structures and thermochemistry at finite temperatures? *Acta Cryst. B* **2016**, *72*, 514–529.

(22) Dumez, J.-N.; Pickard, C. J. Calculation of NMR chemical shifts in organic solids: Accounting for motional effects. *J. Chem. Phys.* **2009**, *130*, 104701.

(23) Exner, T. E.; Frank, A.; Onila, I.; Möller, H. M. Toward the Quantum Chemical Calculation of NMR Chemical Shifts of Proteins. 3. Conformational Sampling and Explicit Solvents Model. *J. Chem. Theory Comput.* **2012**, *8*, 4818–4827.

(24) Lehtivarjo, J.; Tuppurainen, K.; Hassinen, T.; Laatikainen, R.; Peräkylä, M. Combining NMR ensembles and molecular dynamics simulations provides more realistic models of

protein structures in solution and leads to better chemical shift prediction. *J. Biomol. NMR* **2012**, *52*, 257–267.

(25) Zhang, H.; Hou, G.; Lu, M.; Ahn, J.; Byeon, I.-J. L.; Langmead, C. J.; Perilla, J. R.; Hung, I.; Gor'kov, P. L.; Gan, Z.; Brey, W. W.; Case, D. A.; Schulten, K.; Gronenborn, A. M.; Polenova, T. HIV-1 Capsid Function Is Regulated by Dynamics: Quantitative Atomic-Resolution Insights by Integrating Magic-Angle-Spinning NMR, QM/MM, and MD. *J. Am. Chem. Soc.* **2016**, *138*, 14066–14075.

(26) Beran, G. J. O. Modeling polymorphic molecular crystals with electronic structure theory. *Chem. Rev.* **2016**, *116*, 5567–5613.

(27) Li, X.; Neumann, M. A.; van de Streek, J. The application of tailor-made force fields and molecular dynamics for NMR crystallography: a case study of free base cocaine. *IUCrJ* **2017**, *4*, 175–184.

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

(29) Dračínský, M.; Möller, H. M.; Exner, T. E. Conformational Sampling by Ab Initio Molecular Dynamics Simulations Improves NMR Chemical Shift Predictions. *J. Chem. Theory Comput.* **2013**, *9*, 3806–3815.

(30) Price, S. L. Computational prediction of organic crystal structures and polymorphism. *Int. Rev. Phys. Chem.* **2008**, *27*, 541–568.

(31) Price, S. L.; Price, L. S. In *Solid State Characterization of Pharmaceuticals*, 1st ed.; Storey, R., Ymen, I., Eds.; Blackwell, 2011; pp 427–450.

(32) Price, S. L. Predicting crystal structures of organic compounds. *Chem. Soc. Rev.* **2014**, *43*, 2098–111.

(33) Stoffel, R. P.; Wessel, C.; Lumey, M.-W.; Dronskowski, R. Ab Initio Thermochemistry of Solid-State Materials. *Angew. Chem. Int. Ed.* **2010**, *49*, 5242–5266.

(34) Červinka, C.; Fulem, M.; Stoffel, R. P.; Dronskowski, R. Thermodynamic Properties of Molecular Crystals Calculated within the Quasi-Harmonic Approximation. *J. Phys. Chem. A* **2016**, *120*, 2022–2034.

(35) Červinka, C.; Fulem, M. State-of-The-Art Calculations of Sublimation Enthalpies for Selected Molecular Crystals and Their Computational Uncertainty. *J. Chem. Theory Comput.* **2017**, *13*, 2840–2850.

(36) Hoja, J.; Reilly, A. M.; Tkatchenko, A. First-principles modeling of molecular crystals: structures and stabilities, temperature and pressure. *WIREs Comput. Mol. Sci.* **2017**, *7*, e1294.

(37) Heit, Y. N.; Nanda, K. D.; Beran, G. J. Predicting finite-temperature properties of crystalline carbon dioxide from first principles with quantitative accuracy. *Chem. Sci.* **2016**, *7*, 246–255.

(38) Iuzzolino, L.; McCabe, P.; Price, S. L.; Brandenburg, J. G. Crystal structure prediction of flexible pharmaceutical-like molecules: density functional tight-binding as an intermediate optimisation method and for free energy estimation. *Faraday Disc.* **2018**, *211*, 275–296.

(39) McKinley, J. L.; Beran, G. J. O. Identifying pragmatic quasi-harmonic electronic structure approaches for modeling molecular crystal thermal expansion. *Faraday Disc.* **2018**, *211*, 181–207.

(40) Whittleton, S. R.; Otero-de-la Roza, A.; Johnson, E. R. Exchange-Hole Dipole Dispersion Model for Accurate Energy Ranking in Molecular Crystal Structure Prediction. *J. Chem. Theory Comput.* **2017**, *13*, 441–450.

(41) Erba, A. On combining temperature and pressure effects on structural properties of crystals with standard ab initio techniques. *J. Chem. Phys.* **2014**, *141*, 124115.

(42) Filippini, G.; Gramaccioli, C. M. Deriving the equilibrium conformation in molecular crystals by the quasi-harmonic procedure: some critical remarks. *Acta Crys. A* **1981**, *37*, 335–342.

(43) Červinka, C.; Beran, G. J. Ab initio thermodynamic properties and their uncertainties for crystalline α-methanol. *Phys. Chem. Chem. Phys.* **2017**, *19*, 29940–29953.

(44) Červinka, C.; Beran, G. J. Ab initio prediction of the polymorph phase diagram for crystalline methanol. *Chem. Sci.* **2018**, *9*, 4622–4629.

(45) Pickard, C. J.; Mauri, F. *Phys. Rev. B* **2001**, *63*, 245101.

(46) Yates, J. R.; Pickard, C. J.; Mauri, F. *Phys. Rev. B* **2007**, *76*, 024401.

(47) Hartman, J. D.; Beran, G. J. O. Fragment-based electronic structure approach for computing nuclear magnetic resonance chemical shifts in molecular crystals. *J. Chem. Theory Comput.* **2014**, *10*, 4862–4872.

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

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

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

(51) Malkin, V. G.; Malkina, O. L.; Salahub, D. R. Influence of Intermolecular Interactions on the ^{13}C NMR Shielding Tensor in Solid .alpha.-Glycine. *J. Am. Chem. Soc.* **1995**, *117*, 3294–3295.

(52) Naito, A.; Ganapathy, S.; Akasaka, K.; McDowell, C. A. Chemical shielding tensor and ^{13}C–^{14}N dipolar splitting in single crystals of L alanine. *J. Chem. Phys.* **1981**, *74*, 3190–3197.

(53) Chen, X.; Zhan, C.-G. First-principles studies of C-13 NMR chemical shift tensors of amino acids in crystal state. *J. Mol. Struct. (THEOCHEM)* **2004**, *682*, 73–82.

(54) Dračínský, M.; Unzueta, P.; Beran, G. J. O. Improving the accuracy of solid-state nuclear magnetic resonance chemical shift prediction with a simple molecular correction. *Phys. Chem. Chem. Phys.* **2019**, *21*, 14992–15000.

(55) Sherwood, M.; Alderman, D.; Grant, D. Assignment of Carbon-13 Chemical-Shift Tensors in Single-Crystal Sucrose. *J. Magn. Reson. A* **1993**, *104*, 132–145.

(56) Sherwood, M. H.; Facelli, J. C.; Alderman, D. W.; Grant, D. M. Carbon-13 chemical shift tensors in polycyclic aromatic compounds. 2. Single-crystal study of naphthalene. *J. Am. Chem. Soc.* **1991**, *113*, 750–753.

(57) Naito, A.; McDowell, C. A. Determination of the 1 4 N quadrupole coupling tensors and the 1 3 C chemical shielding tensors in a single crystal of L asparagine monohydrate. *J. Chem. Phys.* **1984**, *81*, 4795–4803.

(58) Burley, J. C.; Duer, M. J.; Stein, R. S.; Vrcelj, R. M. Enforcing Ostwald's rule of stages: Isolation of paracetamol forms III and II. *Eur. J. Pharm. Sci.* **2007**, *31*, 271–276.

(59) Gervais, C.; Dupree, R.; Pike, K. J.; Bonhomme, C.; Profeta, M.; Pickard, C. J.; Mauri, F. Combined First-Principles Computational and Experimental Multinuclear

Solid-State NMR Investigation of Amino Acids. *J. Phys. Chem. A* **2005**, *109*, 6960–6969.

(60) Hu, J. Z.; Facelli, J. C.; Alderman, D. W.; Pugmire, R. J.; Grant, D. M. 15 N Chemical Shift Tensors in Nucleic Acid Bases. *J. Am. Chem. Soc.* **1998**, *120*, 9863–9869.

(61) Baias, M.; Widdifield, C. M.; Dumez, J. N.; Thompson, H. P.; Cooper, T. G.; Salager, E.; Bassil, S.; Stein, R. S.; Lesage, A.; Day, G. M.; Emsley, L. Powder crystallography of pharmaceutical materials by combined crystal structure prediction and solid-state ^{1}H NMR spectroscopy. *Phys. Chem. Chem. Phys.* **2013**, 8069–8080.

(62) Wei, Y.; de Dios, A. C.; McDermott, A. E. Solid-State 15 N NMR Chemical Shift Anisotropy of Histidines: Experimental and Theoretical Studies of Hydrogen Bonding. *J. Am. Chem. Soc.* **1999**, *121*, 10389–10394.

(63) Sharif, S.; Schagen, D.; Toney, M. D.; Limbach, H.-H. Coupling of Functional Hydrogen Bonds in Pyridoxal-5-phosphateEnzyme Model Systems Observed by Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* **2007**, *129*, 4440–4455.

(64) Azaïs, T.; Hartmeyer, G.; Quignard, S.; Laurent, G.; Tourné-Péteilh, C.; Devoisselle, J.-M.; Babonneau, F. Solid-state NMR characterization of drug-model molecules encapsulated in MCM-41 silica. *Pure Appl. Chem.* **2009**, *81*, 1345–1355.

(65) Limwikrant, W.; Nagai, A.; Hagiwara, Y.; Higashi, K.; Yamamoto, K.; Moribe, K. Formation mechanism of a new carbamazepine/malonic acid cocrystal polymorph. *International Journal of Pharmaceutics* **2012**, *431*, 237–240.

(66) Czernek, J. On the solid-state NMR spectra of naproxen. *Chem. Phys. Lett.* **2015**, *619*, 230–235.

(67) Giannozzi, P.; Andreussi, O.; Brumme, T.; Bunau, O.; Buongiorno Nardelli, M.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Cococcioni, M.; Colonna, N.; Carn-

imeo, I.; Dal Corso, A.; de Gironcoli, S.; Delugas, P.; DiStasio, R. A.; Ferretti, A.; Floris, A.; Fratesi, G.; Fugallo, G.; Gebauer, R.; Gerstmann, U.; Giustino, F.; Gorni, T.; Jia, J.; Kawamura, M.; Ko, H.-Y.; Kokalj, A.; Küçükbenli, E.; Lazzeri, M.; Marsili, M.; Marzari, N.; Mauri, F.; Nguyen, N. L.; Nguyen, H.-V.; Otero-de-la Roza, A.; Paulatto, L.; Poncé, S.; Rocca, D.; Sabatini, R.; Santra, B.; Schlipf, M.; Seitsonen, A. P.; Smogunov, A.; Timrov, I.; Thonhauser, T.; Umari, P.; Vast, N.; Wu, X.; Baroni, S. Advanced capabilities for materials modelling with Quantum ESPRESSO. *J. Phys. Condens. Mat.* **2017**, *29*, 465901.

(68) Becke, A. D. Variable temperature study of the crystal structure of paracetamol (p-hydroxyacetanilide), by single crystal neutron diffraction. *J. Chem. Phys.* **1986**, *38*, 7184–7187.

(69) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(70) Otero-De-La-Roza, A.; Johnson, E. R. Van der Waals interactions in solids using the exchange-hole dipole moment model. *J. Chem. Phys.* **2012**, 174109.

(71) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Dal Corso, A.; De Gironcoli, S.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Mat.* **2009**, 395502.

(72) Monkhorst, H. J.; Pack, J. D. *Phys. Rev. B* **1976**, *13*, 5188–5192.

(73) Togo, A.; Tanaka, I. First principles phonon calculations in materials science. *Scr. Mater.* **2015**, *108*, 1–5.

(74) Antony, J.; Grimme, S. Density functional theory including dispersion corrections for intermolecular interactions in a large benchmark set of biologically relevant molecules. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5287.

(75) Beran, G. J. Approximating quantum many-body intermolecular interactions in molecular clusters using classical polarizable force fields. *J. Chem. Phys.* **2009**, *130*, 1–9.

(76) Beran, G. J. O.; Nanda, K. Predicting organic crystal lattice energies with chemical accuracy. *J. Phys. Chem. Lett* **2010**, *1*, 3480–3487.

(77) Wen, S.; Beran, G. J. O. Accurate molecular crystal lattice energies from a fragment QM/MM approach with on-the-fly Ab initio force field parametrization. *J. Chem. Theory Comput.* **2011**, *7*, 3733–3742.

(78) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Montgomery Jr., J. A.; Peralta, J. E.;

Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision D.01. *Gaussian Inc.* **2013**,

(79) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158.

(80) Ditchfield, R. Self-consistent perturbation theory of diamagnetism. *Mol. Phys.* **1974**, *27*, 789–807.

(81) Chesnut, D. B.; Moore, K. D. Locally dense basis sets for chemical shift calculations. *J. Comp. Chem.* **1989**, 648–659.

(82) Chesnut, D. B.; Rusiloski, B. E.; Moore, K. D.; Egolf, D. A. Use of locally dense basis sets for nuclear magnetic resonance shielding calculations. *J. Comp. Chem.* **1993**, 1364–1375.

(83) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Selfconsistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72*, 650–654.

(84) McLean, A. D.; Chandler, G. S. Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, Z =1118. *J. Chem. Phys.* **1980**, *72*, 5639–5648.

(85) Frisch, M. J.; Pople, J. A.; Binkley, J. S. Selfconsistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets. *J. Chem. Phys.* **1984**, *80*, 3265–3269.

(86) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements, Li-F. *J. Comp. Chem.* **1983**, *4*, 294–301.

(87) Hehre, W. J.; Ditchfield, R.; Pople, J. A. SelfConsistent Molecular Orbital Methods. XII. Further Extensions of GaussianType Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261.

(88) Hariharan, P. C.; Pople, J. A. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta* **1973**, *28*, 213–222.

(89) Hartman, J. D.; Neubauer, T. J.; Caulkins, B. G.; Mueller, L. J.; Beran, G. J. O. Converging nuclear magnetic shielding calculations with respect to basis and system size in protein systems. *J. Biomol. NMR* **2015**, *62*, 327–340.

(90) Beran, G. J. O. Calculating Nuclear Magnetic Resonance Chemical Shifts from Density Functional Theory: A Primer. *eMagRes* **2019**, *in press*.

(91) Breneman, C. M.; Wiberg, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comp. Chem.* **1990**, *11*, 361–373.

(92) Blanco, F.; Alkorta, I.; Elguero, J. Statistical analysis of ^{13}C and ^{15}N NMR chemical shifts from GIAO/B3LYP/6-311++G** calculated absolute shieldings. *Magn. Reson. Chem.* **2007**, 797–800.

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

(94) Chisholm, J. A.; Motherwell, W. D. S. COMPACK: A program for identifying crystal structure similarity using distances. *J. Appl. Crystall.* **2005**, *38*, 228–231.

(95) Addicoat, M.; Adjiman, C. S.; Arhangelskis, M.; Beran, G. J. O.; Bowskill, D.; Brandenburg, J. G.; Braun, D. E.; Burger, V.; Cole, J.; Cruz-Cabeza, A. J.; Day, G. M.; Deringer, V. L.; Guo, R.; Hare, A.; Helfferich, J.; Hoja, J.; Iuzzolino, L.; Jobbins, S.; Marom, N.; McKay, D.; Mitchell, J. B. O.; Mohamed, S.; Neumann, M.; Nilsson Lill, S.; Nyman, J.; Oganov, A. R.; Piaggi, P.; Price, S. L.; Reutzel-Edens, S.; Rietveld, I.; Ruggiero, M.; Ryder, M. R.; Sastre, G.; Schön, J. C.; Taylor, C.; Tkatchenko, A.; Tsuzuki, S.; van den Ende, J.; Woodley, S. M.; Woollam, G.; Zhu, Q. Crystal structure evaluation: calculating relative stabilities and other criteria: general discussion. *Faraday Disc.* **2018**, *211*, 325–381.

(96) van de Streek, J.; Neumann, M. A. Validation of molecular crystal structures from powder diffraction data with dispersion-corrected density functional theory (DFT-D). *Acta Cryst. B* **2014**, *70*, 1020–1032.

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

(98) LeBlanc, L. M.; Dale, S. G.; Taylor, C. R.; Becke, A. D.; Day, G. M.; Johnson, E. R. Pervasive Delocalisation Error Causes Spurious Proton Transfer in Organic Acid-Base Co-Crystals. *Angew. Chem. Int. Ed.* **2018**, *57*, 14906–14910.

Graphical TOC Entry