From dimers to the solid-state: Distributed intermolecular force-fields for pyridine

Alexander A. Aina, Alston J. Misquitta, and Sarah L. Price

Articles you may be interested in

A general intermolecular force field based on tight-binding quantum chemical calculations
The Journal of Chemical Physics **147**, 161708 (2017); 10.1063/1.4991798

Extension of the D3 dispersion coefficient model
The Journal of Chemical Physics **147**, 034112 (2017); 10.1063/1.4993215

Mapping the Drude polarizable force field onto a multipole and induced dipole model
The Journal of Chemical Physics **147**, 161702 (2017); 10.1063/1.4984113

Intermolecular interactions in the condensed phase: Evaluation of semi-empirical quantum mechanical methods
The Journal of Chemical Physics **147**, 161704 (2017); 10.1063/1.4985605

Toward chemical accuracy in the description of ion–water interactions through many-body representations. Alkali-water dimer potential energy surfaces
The Journal of Chemical Physics **147**, 161715 (2017); 10.1063/1.4993213

Interpolation of intermolecular potentials using Gaussian processes
The Journal of Chemical Physics **147**, 161706 (2017); 10.1063/1.4986489

From dimers to the solid-state: Distributed intermolecular force-fields for pyridine

Alexander A. Aina,[1] Alston J. Misquitta,[2] and Sarah L. Price[1]
[1]*Department of Chemistry, University College London, London WC1H 0AJ, United Kingdom*
[2]*School of Physics and Astronomy, Queen Mary, University of London, London E1 4NS, United Kingdom*

An anisotropic atom-atom force-field for pyridine, using distributed atomic multipoles, polarizabilities, and dispersion coefficients and an anisotropic atom-atom repulsion model derived from symmetry-adapted perturbation theory (density functional theory) dimer calculations, is used to model pyridine crystal structures. Here we show that this distributed intermolecular force-field (DIFF) models the experimental crystal structures as accurately as modelling all but the electrostatic term with an isotropic repulsion-dispersion potential that has been fitted to experimental crystal structures. In both cases, the differences are comparable to the changes in the crystal structure with temperature, pressure, or neglect of zero-point vibrational effects. A crystal structure prediction study has been carried out, and the observed polymorphs contrasted with hypothetical thermodynamically competitive crystal structures. The DIFF model was able to identify the structure of an unreported high pressure phase of pyridine, unlike the empirically fitted potential. The DIFF model approach therefore provides a model of the underlying pair potential energy surface that we have transferred to the crystalline phase with a considerable degree of success, though the treatment of the many-body terms needs improvement and the pair potential is slightly over-binding. Furthermore, this study of a system that exhibits isotopic polymorphism highlights that the use of an empirical potential has partially absorbed temperature and zero-point motion effects as well as the intermolecular forces not explicitly represented in the functional form. This study therefore highlights the complexity in modelling crystallization phenomena from a realistic pair potential energy surface. © *2017 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).* [http://dx.doi.org/10.1063/1.4999789]

I. INTRODUCTION

The true intermolecular potential energy surface of a molecule should be transferable between all phases, allowing simulation of all the physical properties that depend on the forces between the molecules. A pairwise approximation is suitable enough to describe the gas-phase properties but the many-body (non-additive) terms may affect the condensed phases.[1] This paper explores transferring a theoretically based many-body pair potential for pyridine[2] to the crystalline state.

The holy grail of defining a sufficiently accurate analytical pair potential to account for all the physical properties of the inert gases was achieved in the 1980's.[3] These potentials only included the three-body dispersion term for the condensed phases. Water and other small polyatomic molecules now have very accurate potentials available, though these are often not used in Molecular Dynamics (MD) simulations because of the need to compromise between accuracy, speed of evaluation, and functional forms assumed in the suitable codes. The condensed phases of organic molecules are an area of increasing interest, with the ability to simulate the differences in material properties between polymorphs being highly desirable for the specialty chemical industries.[4–6] Polymorphs, which are different crystal structures with identical liquid and vapor phases, usually have a number of differing physical properties, such

as solubility, melting point, morphology, shock and mechanical stress responses, surface chemistry, reactivity, and phase transformations. Over 50% of organic molecules are found to be polymorphic in the industrial searches for polymorphs which are routine in developing a pharmaceutical product.[7] Hence all organic specialty industries, from opto-electronic, energetic pigments, agrochemicals as well as pharmaceuticals, wish to establish the behavior of the organic molecular crystals under all functioning conditions. However, no current force-fields are accurate enough for use in predicting the crystal structures of weakly bound organic molecules: all successful methods in the recent blind test of crystal structure prediction[8] (CSP) relied on large numbers of electronic structure calculations for attempting to rank the relative energies of possible crystal structures. Thus, there is a need for force-fields that correctly model the thermodynamics and properties of the observed polymorphs and other hypothetical crystal structures, the liquid, and the gaseous phase if we are to be able to study the phase change behavior of organic molecules.

The intermolecular force-fields that have been most widely used in simulating organic condensed phases have been derived by empirical fitting to crystal structures and properties[9,10] with increasing the sophistication being used for the challenge of modelling crystal structures[6,11,12] to the extent that they rival popular periodic DFT+D calculations

in accuracy.[6] The parameterization of such empirically fitted force-fields inevitably absorbs, to some extent, errors, approximations, and assumptions in the experiments and computational methods, such as neglect of thermal effects, or assumed functional form. The problems with selecting the experimental data and the assumptions used in fitting a potential to the data can be avoided by fitting an analytical force-field to an *ab initio* potential energy surface.[13] This approach of using *ab initio* derived analytical potential energy surfaces (PESs) for spectra of gas-phase clusters has been extensively applied to water[14–18] and other small molecule complexes,[19–21] and various *ab initio* astrophysical applications.[22–25]

Pioneering studies of the organic solid-state using analytical fits to *ab initio* interaction energy calculations have mainly concentrated on energetic molecules,[26,27] where predictive modelling requires reliable extrapolation into the repulsive region that is not sampled in fitting empirical potentials. A non-empirical potential, based on an early variant of the Distributed Intermolecular Force-Field (DIFF) approach for $C_6Br_2ClFH_2$,[28] was successful in predicting its crystal structure in the fourth blind test. It is therefore very timely to assess the barriers to producing a sufficiently realistic model intermolecular potential from dimer calculations that can be used to predict the properties of a molecule in all phases and assess the extent to which the many-body terms are a limitation.

Considerable advances have been made in the ability to calculate intermolecular pair potential energy surfaces accurately.[2,13,29–33] Computing sufficient number of points to define the PES leads to two challenges:

(1) the computational cost incurred, particularly for large molecules, and
(2) the need for sufficient accuracy.

Both needs are satisfied by the symmetry-adapted perturbation theory based on density functional theory[11,34–37] SAPT(DFT), which has now become the method of choice for many applications involving weakly bound molecular interactions.[2,28,38–41]

However, errors are introduced by the choice of an analytical functional form and in the fitting process. Recently a method of automatically generating analytical intermolecular potential energy surfaces in an isotropic atom-atom functional form to SAPT(DFT) numerical points reported a typical fit error of about 0.8 kJ mol^{-1} in the negative energy region.[13] Also, recently, one of us has been involved in the development of the Slater-type Force-Field (Slater-FF) models, which exhibits average errors of 0.2 kJ mol^{-1} in the attractive region for a wide range of interacting dimers.[31] However, these model potentials have so far been based on an isotropic atom-atom form, primarily for the ease of use in simulation programs. Here we explore the alternative approach of using anisotropic atom-atom model potentials that are generated from distributed monomer properties and advanced fitting to SAPT(DFT) component energies. This approach, using the program CᴀᴍCASP, has been described in detail in an application to the pyridine dimer.[2] This type of non-empirical distributed intermolecular force-field (DIFF)

uses distributed multipolar models for the electrostatic, polarization, and dispersion interactions and anisotropic atom-atom exponential short range terms derived by fitting to SAPT(DFT) interaction energies via a distributed density-overlap model.

While all three of these models represent considerable progress in accurate force-fields for dimers, it is important to determine how well they perform for the crystalline state. There are two issues that arise here. The first is that the lattice energy differences between 80% of polymorph pairs[7] is less than 4 kJ mol^{-1}, with many observed polymorph energy differences being close to the reported errors in the dimer potentials. The second issue is that so far these models have only been tested on gas-phase and calculated dimer properties while there are many-body effects in condensed phases.

In this paper, we pioneer transferring this DIFF models to the solid-state phases of pyridine, seeking to explore the issues involved in defining and using analytical force-fields that can transfer between phases.

A. Pyridine

Pyridine has a melting point of 231.6 K,[42,43] and as a liquid at ambient conditions has relatively weak intermolecular interactions. The molecular structure varies so little in the condensed phases that pyridine is assumed to be rigid and to adopt the isolated static molecule structure. However, pyridine has a dipole moment of ~2.2 D[44] and can be considered capable of forming weak C–H$\cdots$N hydrogen bonds.[45] Thus, dispersion, electrostatic, and polarization forces are competitive with the repulsive forces in determining the structure of the different phases, as has already been shown in the design of an intermolecular potential for simulating liquid pyridine.[46]

The first crystal structure of pyridine to be determined is unusually complex,[42,43] with 4 independent molecules in the asymmetric unit (Z′ = 4) and 16 molecules in the unit cell (form I, P$na2_1$). After an early crystal structure prediction (CSP) study showed that simpler structures were thermodynamically competitive, a low-temperature form II (P$2_12_12_1$ Z′ = 1) was crystallized from pentane for deuterated-pyridine (*d*5) though not for protonated-pyridine (*h*5).[47] The two polymorphs are rather similar, with a large common coordination cluster (Fig. 1) but differ in the stacking so that a CH$\cdots\pi$ interaction in form I becomes a CH$\cdots$N interaction in the low-temperature, high-pressure form II.[48] A later CSP study, using DFT-D periodic electronic structure calculations, showed that the lattice energy of form II is slightly more stable than form I by less than 0.1 kJ mol^{-1}, though the difference increases with pressure, and alternative structures are within 0.1 kJ mol^{-1} of the observed structures.[49] The crystallization behavior of pyridine is isotope-dependent, as well as depending on pressure and temperature.[45,48] Additionally, towards the conclusion of this work, we became aware that a third form (*d*5-III) had been crystallized at around 2 GPa pressure[50] and were challenged to identify its structure, providing a test of how the intermolecular potential models the repulsive wall.
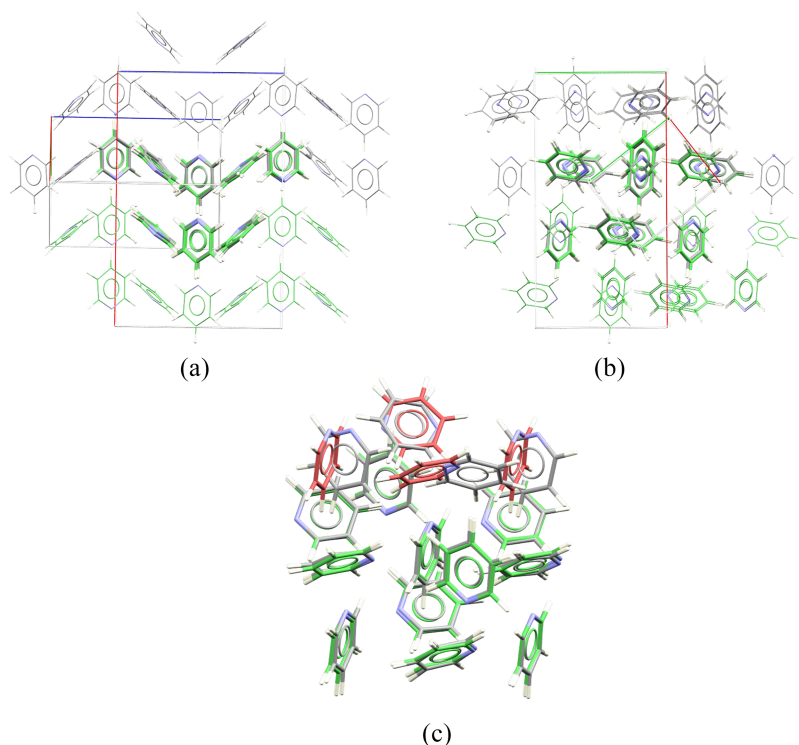
FIG. 1. Comparison of the crystal structures of form I (grey d5-I PYRDNA04[48]) and form II (green d5-II PYRDN05[48]) of pyridine, showing the optimum overlay of 11/15 molecules, where the root mean square deviation of the 11 molecules ($RMSD_{11}$) = 0.149 Å and additional molecules in wireframe (a) viewed showing the a (red) and c (blue) cell axes, (b) showing the form I and II overlay along the b (green) cell axes, and (c) the 15 molecule cluster including further unmatched molecules of form II in red.

II. METHOD

A. Analytical intermolecular potential models

This study used the most accurate version of the DIFF approach, referred to as model 3 in the paper which described its development using CᴀᴍCASP.[2] Here we focus on the aspects of the model that required modification and refitting for use in organic crystal structure modelling code DMACRYS,[51] and the checks used to ensure that the potentials were equivalent. The DIFF model has the following functional form, omitting the standard multipolar expansion of the electrostatic and polarization terms:[1]

$$U(R_{MN}, \Omega_{MN}) = \sum_{i \in M, k \in N} A_{ik} \exp[-B_{ik}(R_{ik} - \rho^{ik}(\Omega_{ik}))]$$
$$- f_6(\beta_{disp} R_{ik}) \frac{C_6^{ik}}{R_{ik}^6} - f_8(\beta_{disp} R_{ik}) \frac{C_8^{ik}}{R_{ik}^8}$$
$$- f_{10}(\beta_{disp} R_{ik}) \frac{C_{10}^{ik}}{R_{ik}^{10}}$$
$$+ \text{electrostatic}(Q_{lm}^i \, l \leq 4, \text{ISA}, \Psi)$$
$$+ \text{polarization}(Q_{lm}^i, \alpha_{lml'm'}^i, \beta_{ind}), \quad (1)$$

where the relative position (R_{ik}) and orientations (Ω_{ik}) of atom i in molecules M and k in molecule N are defined by the relative position and orientation (R_{MN}, Ω_{MN}) of the molecules M and N and the rigid molecular structure. The static isolated molecule structure was obtained by optimization using the Perdue-Burke-Emzerhof (PBE) general gradient approximation combined with a portion of exact exchange to form the PBE0 functional[52–54] and the cc-pVTZ basis set, and kept rigid in all simulations. The long range contributions are all modelled using distributed molecular properties[1] of the

asymptotically corrected[55] (PBE0/AC) charge density calculated with the d-aug-cc-pVTZ Dunning basis[56] and referred to a molecule-fixed axis system, with x along the C3–N symmetry axis and the molecule in the xy plane. The distributed multipoles (Q_{lm}^i) were derived using the basis-space iterated stockholder atoms (ISA)[57] analysis, which is a more effective partitioning method than GDMA[58] as it exhibits more natural charges, nearly spherical shapes, and is, overall, a better model of the atomic anisotropic electron densities. The distributed multipole expansion is evaluated up to rank 4 (hexadecapole) to include the contributions of the atomic anisotropy, such as the lone-pair (dipole) and π-orbital (quadrupole) features, which are important in organic crystals.[59] The atomic polarizabilities ($\alpha_{lm,l'm'}^i$) were limited to just the dipolar terms ($l = l' = 1$) and the atomic dispersion coefficients (C_n^{ik}) were isotropic, with both being derived using the Williams-Stone-Misquitta (WSM) polarizability model.[60–62] The C_8 and C_{10} dispersion terms are included for heavy atoms, to ensure the correct distance dependence for the dispersion, as the relative orientations of molecules in crystals sample the dispersion out to infinity. Both the polarization and dispersion terms are damped with a single atom-atom parameter ($\beta_{ind} = 1.25$ a.u., $\beta_{disp} = 1.67$ a.u.) Tang-Toennies damping function[63] to prevent unphysical behavior at close contacts.

The remaining short range contributions to the intermolecular potential were obtained using the CamCASP and ORIENT programs by fitting to a large number of SAPT(DFT) dimer energy calculations via a distributed density-overlap model. The fitting was performed in a hierarchical manner to obtain the atom-atom exponential repulsive model with isotropic pre-exponential (A_{ik}) and hardness (B_{ik}) coefficients.[2] The short range term is made anisotropic by the inclusion of the shape function ($\rho^{ik}(\Omega_{ik})$) in the exponential. The shape function

is a polynomial containing anisotropic coefficients that were defined relative to an atomic local axis system with the y axis normal to the plane of the molecule and the z axis along the C–H bond pointing out from the molecule. The changes to these specific molecule-fixed and local axes are required for most molecules (unlike for pyridine) by the conventions within the crystal structure modelling program DMACRYS. These conventions enable the multipole moments to be transformed to maintain a right-handed axis system when molecules are generated by inversion or other symmetry operators that would otherwise invert the axis definition. The change in axis systems meant that the repulsion part of the potential had to be refitted, but this DIFF model is essentially the same as model 3 from Ref. 2 except for the definitions of axis systems. The parameters of the DIFF model are given in full in the supplementary material in the ORIENT format so that the symmetry relationships are clear. The accuracy of the translation between the gas/dimer and solid state modeling was tested by the calculation of the second virial coefficients and by reproducing the previous ORIENT gas phase results[2] in DMACRYS calculations in which identical dimers were set up within cubic unit cells of side 100 Å. Two other non-empirical variants on this model were also tested (see the supplementary material). DIFF(no-pol) simply omits the polarization term, which is the only many-body contribution included in DIFF. The effect of the anisotropic repulsion was tested by refitting the short-range terms to an isotropic model ($\rho^{\iota\kappa}(\Omega_{ik}) = 0$) to give DIFF(iso-rep).

The empirically fitted potentials being used for comparison are representing all but the electrostatic contribution by an empirical isotropic atom-atom exp-6 potential,

$$U(R_{MN}, \Omega_{MN}) = \sum_{i \in M, k \in N} A_{\iota\kappa} \exp(-B_{\iota\kappa} R_{ik}) - \frac{C_6^{\iota\kappa}}{R_{ik}^6}$$
$$+ \text{electrostatic } (Q_{lm}^i \, l \le 4, \text{GDMA}, \Psi), \quad (2)$$

where atom i is of type ι (C/H/N) and atom k of type κ. The exp-6 FIT parameters[51] relevant to pyridine were originally fitted to aza-hydrocarbon crystal structures and a limited set of crystal energies[64] whereas the WILL01 set discriminates between different types of N hybridization.[12] The FIT model is combined with the following electrostatic models: the same ISA distributed multipoles as used in DIFF, the point charge only component, and the electrostatic model that has often been combined with FIT in crystal structure prediction studies[65] [i.e., distributed multipoles from the GDMA2[58] analysis of the PBE0 6-31G(d, p) molecular charge density, thus involving a change in the wavefunction ψ used to produce the electrostatic model]. Since the WILL01 potential is defined with the H interaction sites moved slightly into the C–H bond, to reflect the non-spherical nature of the charge distribution around hydrogen atoms, distributed multipoles for these non-atomic sites could only be obtained using GDMA2.[58]

B. Simulation methods

The gas phase dimer properties were simulated using ORIENT,[66] including the calculation of the classical second virial coefficients by numerical integration. The second virial coefficients are an appropriate test of the intermolecular pair potential in the gas phase and could be compared with experimental values[67,68] derived by vapor compressibility measurements in the 1950s, which showed some deviation from prior measurements, and gave a heat of vaporization of 35.2 kJ mol^{-1} at the boiling temperature of 115.26 °C and 40.4 kJ mol^{-1} at 25 °C.

The CSP study is based on lattice energies E_{latt}, the internal energy of the system relative to infinitely separated molecules in their lowest energy configuration: $E_{latt} = E_{inter}^{atom-atom} + \Delta E_{intra}^{molecule}$. The total lattice energy can be separated into atom-atom intermolecular interactions and intramolecular (geometry distortion) energies of the molecules, however, in this study, molecules are held rigid, hence $\Delta E_{intra}^{molecule} = 0$ and the lattice energy depends only on the atom-atom intermolecular interaction energy. The crystal structures were modelled using DMACRYS2.2.1.1,[51] with lattice summations being carried out to 15 Å followed by a 2 Å splined correction, with the charge-charge, charge-dipole, and dipole-dipole electrostatic contributions being evaluated by Ewald summation. Optimization used the analytical first derivatives apart from the forces due to the polarization term.

The polarization term is the only non-additive many-body term in the potential, and its use in the solid state produces new challenges. The induced moments in a crystal structure need to be solved iteratively to consistency[51] because the polarizing field depends on the polarized moments of all the molecules in the crystal. Hence calculating the polarization forces requires taking numerical derivatives of the iterated distributed induced moments. As smooth, numerical forces are needed, optimization with the polarization energy is very computationally demanding. This has been done for the experimental structures (see Table I of the supplementary material), where it was confirmed that the inclusion of polarization forces during optimization changed the lattice parameters from those optimized without the forces due to polarization by less than ±0.04 Å. Thus the geometric changes resulting from including polarization forces resulted in a negligible further change in the lattice energy. A significant effort was required to do the optimization including the polarization forces and since the calculations in Table I of the supplementary material show that the cell geometry with polarization forces included is so close to that obtained by just adding the polarization energy after optimization, all other optimizations and second derivative property calculations did not include the derivatives of the polarization contribution. A single-point polarization energy was added for the DIFF and DIFF(iso-rep) model potentials. This approximation may only be acceptable because the induced moments are small, as shown by a small induced moment electrostatic potential on the van der Waals surface (Fig. 2). However the induced moments do vary with the changes in crystalline environment, with all the four independent molecules in form I having a less symmetric environment, and hence greater polarization, than the single independent molecule in form II (Fig. 2).

The energy difference between the hypothetical static crystal at 0 K and a real crystal can be estimated by lattice dynamics within the harmonic approximation and
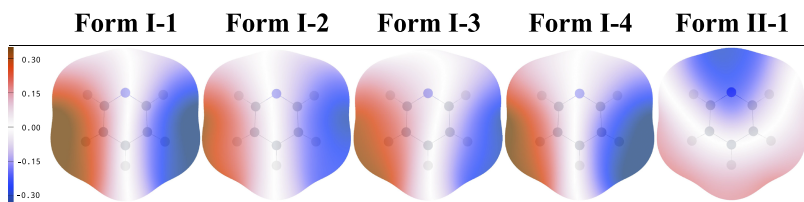
FIG. 2. The additional electrostatic potential around each pyridine molecule in the two crystals arising from the induced atomic dipole moments, after optimization with the DIFF model. The potential in eV is displayed on the van der Waals surface using ORIENT 4.9.08.[66] The lower symmetry environment in $Z' = 4$ form I produces a marked difference in the induced moments relative to $Z' = 1$ form II.

using the DMACRYS methodology and scripts developed by Nyman.[6,69] We sample a number of k-points in reciprocal space by computing the phonons frequencies[70] and elastic stiffness tensors[71] for a number of linearly elongated supercells of pyridine to sample the first Brillouin zone. The supercells are generated by selecting a default 0.12 Å^{-1} k-point spacing[6,69] which samples about 26 k-points. As pyridine is a small molecule, this does not require very large supercells. The Debye frequency contribution to the acoustic phonons and a Gaussian Kernel Density Estimate (KDE) of the optical density of states (using a bandwidth of 3 cm^{-1}) are calculated from the elastic tensors and phonons, respectively, and then used to calculate the free energy thermal correction F_{vib}, as the sum of the vibrational zero-point energy and the thermal energy at 190 K. The Helmholtz free energy $A_{free}(T) = E_{latt} + F_{vib}(T)$ was evaluated at 190 K as this is the temperature at which there is an experimental crystal structure of both forms I and II but this is only 40 K below the melting point of pyridine. Hence, the use of the harmonic approximation is unlikely to be realistic, as discussed later. The effects of pressure were calculated with DMACRYS by optimizing the cell geometry, including the PV contribution to the energy, to minimise the lattice enthalpy $H_{latt} = E_{latt} + PV$.

The CSP study used CrystalPredictor 1.8[72,73] to generate a million putative $Z' = 1$ crystal structures of pyridine within the 59 most probable space groups. Optimization using the FIT potential and ISA point charges generated just over five thousand unique structures. This set was re-optimized with the various anisotropic model potentials, any optimized structures whose second derivative properties showed that they were not true minima were discarded, and the remaining structures were re-clustered. The most stable structures were analyzed in detail for their similarities to each other and the isolated dimer structures. This was done using the similarity tool in Mercury[74] which determines how many molecules (n) of a maximum coordination cluster (15 for crystals, 2 for crystal/dimer comparisons) can be matched within a 20% distance in intermolecular atom-atom distances and 20° in interatomic intermolecular angles and reports the optimum RMSD$_n$ (root mean square deviation of n molecules) of the overlay, ignoring hydrogen or deuterium atoms.

III. RESULTS

A. Gas phase properties

The quality of the DIFF model is mainly supported by its theoretical origin and the quality of fit. The intermolecular potential energy surface has eight distinct minima[2] (Fig. 3), the most stable having two C–H···N weak hydrogen bonds (referred to as Hb1), others only one, some distorted T shaped geometries stabilized by C–H···π interactions, as well as a

trio of displaced stacked geometries (Fig. 3). It is notable that the dimers differ markedly in the dominant contribution to the binding energy; the electrostatic plus polarization contribution is similar to the total binding for the hydrogen bonded Hb1 dimer, but this contribution is less 1 kJ mol^{-1} for the almost iso-energetic stack S1, which has double the dispersion stabilization. There are three distorted T dimers with CH···π interactions and another C–H close to nitrogen, which are quite similar in energy and also in structure apart from the relative position of the nitrogen. These T-shaped dimers have all components intermediate between the hydrogen-bonded and stacked dimers.

The classical second virial coefficients (Fig. 4) differ markedly between the empirical and DIFF models, with the latter being closer to the experimental values. The empirical FIT potential underestimates the effect of the intermolecular forces on the gas-phase property by approximately 25%, whereas the DIFF model appears to slightly overestimate the effect by around 10%. The reasons why the DIFF model is slightly over-binding are discussed later.

Dimer Label	Structure	Dimer Energy Terms (kJ mol^{-1})				Total Dimer E
		Elst	Disp	St-Rg	Pol	
S1		-0.87	-28.88	14.85	-0.72	-15.63
S2		-0.29	-29.36	14.98	-0.72	-15.39
T1		-6.45	-19.99	12.26	-0.98	-15.17
T2		-6.47	-19.59	11.79	-1.02	-15.30
bT		-5.84	-19.19	11.38	-1.03	-14.67
Hb1		-12.94	-15.76	14.68	-2.53	-16.56
Hb2		-8.51	-15.72	10.68	-1.68	-15.22
Hb3		-8.23	-15.14	10.61	-1.57	-14.34

FIG. 3. The most stable minima in the pyridine intermolecular potential energy surface (gas phase dimers), with their DIFF energies broken down into components; multipolar electrostatic (Elst), damped dispersion (Disp), short-range (St–Rg) and damped multipolar polarisation (Pol). The short range terms include the exchange-repulsion, electrostatic penetration, and charge transfer energies.
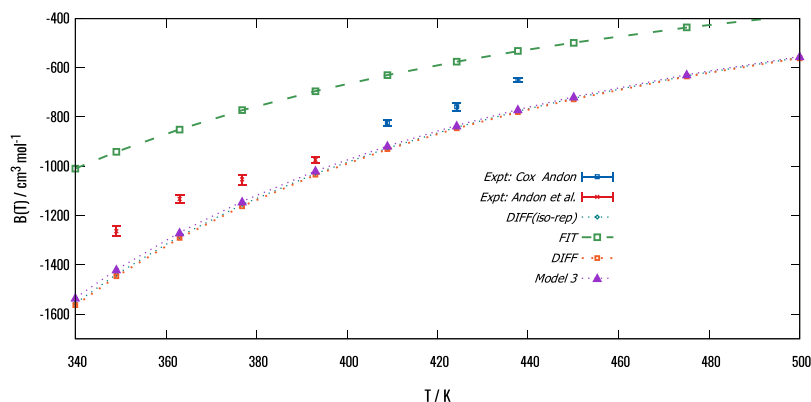
FIG. 4. The second virial coefficients of pyridine calculated from the DIFF and FIT+GDMA potentials, contrasted with the experimental values.[67,68] The values for the DIFF model differ from those previously reported[2] for model 3 which were affected by a coding error. The agreement between DIFF and model 3 tests the equivalence of the potentials.

B. Reproduction of known crystal structures of pyridine

A comparison of the cell parameters and densities of the lattice energy minima with the various analytical potentials with the three experimental determinations of form I at three temperatures between 190 and 5 K and the two determinations of form II at two pressures (but different temperatures) is given in Fig. 5, with a full set of comparisons in the supplementary material. The majority of the potentials give a satisfactory reproduction of the two crystal structures by lattice energy minimization, which corresponds to a static crystal, in comparison with the variation of structure with temperature. The

cell parameters of the non-empirical DIFF models are found to better match the 5 K form $h5$-I structure (PYRDNA04) than the empirical models, while the empirical models that include a distributed multipole electrostatic model (GDMA/ISA) better match the densities at higher temperatures Fig. 5(a). Accessible variations in pressure generally have a larger effect on the crystal structures than temperature changes, but since pressure experiments are usually done at ambient temperature, it is not possible to de-convolute temperature and pressure effects on pyridine which is only solid at ambient when under pressure. Allowing for this difference in temperatures, optimizing the lattice enthalpy seems to provide a sensible estimate of the structures under modest pressures [1.1 GPa Fig. 5(b),
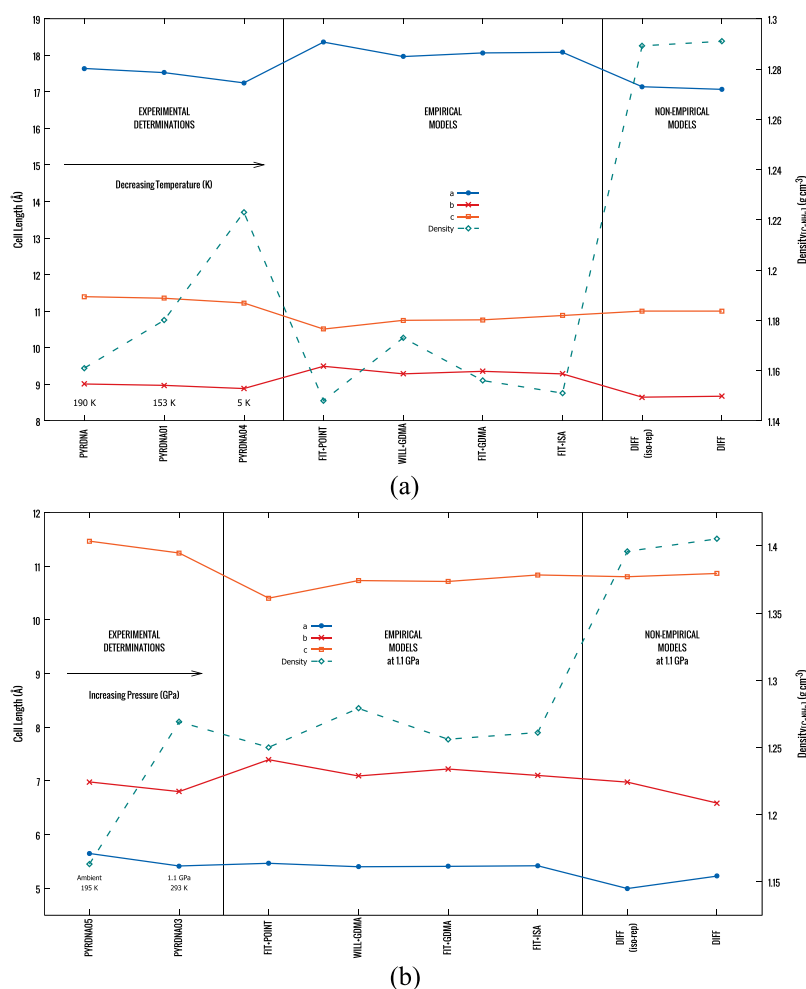

(a)

(b)

FIG. 5. Comparing the experimental determinations of the crystal structures of pyridine with the optimized structure computed with different empirical and non-empirical model potentials. (a) Form I lattice energy minima against structures determined at different temperatures. (b) Form II determinations as a function of pressure, against lattice enthalpies calculated including the PV energy at 1.1 GPa, the observed transition pressure.[48] Further comparisons can be found in the supplementary material. (All $d5$ structure densities have been converted to $h5$ equivalents.)
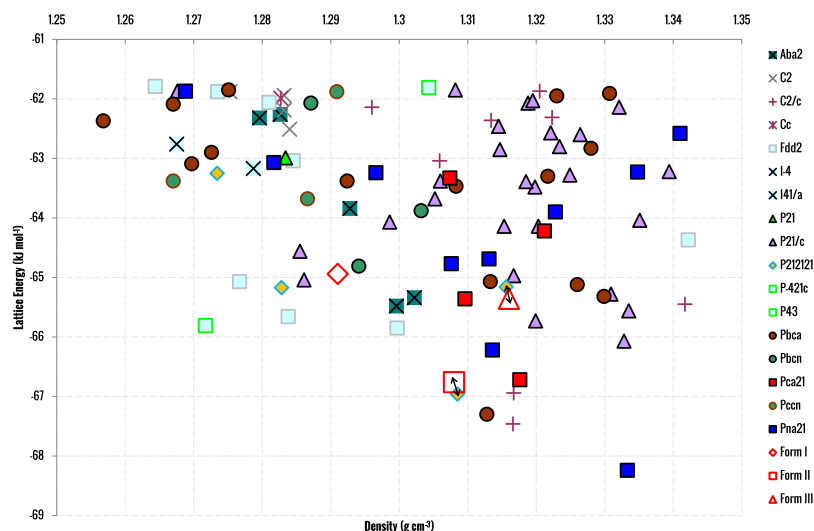
FIG. 6. CSP generated crystal structures of pyridine ranked by lattice energy calculated with the DIFF model. Each point represents a mechanically stable crystal structure, classified by its space group. The lattice energy minima obtained by minimizing the experimental structures with the DIFF model and fixed rigid molecular structure are shown by open red symbols and are joined by tie lines to the corresponding structure found in this $Z' = 1$ search (which is not capable of generating form I). The above figure shows the 100 most stable structures. The space group, lattice parameters, energy, and the corresponding dimer motif for the 30 most stable structures can be found in the supplementary material, linked to the structure identifier $_{pyr}$#.

1.23 GPa Fig. 1(b) of the supplementary material] for all the intermolecular potentials.

C. Stability of observed versus hypothetical structures

The CSP lattice energy landscape with DIFF generated many structures competitive in energy with forms I and II (Fig. 6), with there being 52 unique structures within 5 kJ mol^{-1} of the most stable. The lowest energy structure ($_{pyr}$2) is denser and 1.3 kJ mol^{-1} more stable than form II. DIFF therefore passes the test of predicting the known structures as sufficiently close to the global minimum to be thermodynamically plausible as polymorphs. Given the experimental uncertainty that the experimental structures are the most thermodynamically stable, and the neglect of temperature effects, as discussed further below, the DIFF model has transferred successfully to the crystalline state. It is therefore worth determining the importance of the different contributions to the relative stability ranking, in light of what parts of the potential surface are sampled by the CSP test.

The ranking of the low energy structures is very sensitive to the model potential (Fig. 7). Removing the polarization energy DIFF(no-pol) (Fig. 3 of the supplementary material) leads to considerable re-ranking, with form II becoming the

most stable, with the energy difference between the two forms increasing to nearly 5 kJ mol^{-1}. The model fitted with an isotropic repulsion model [DIFF(iso-rep)] reverses the stability order of forms I and II and has a different selection of structures (Fig. 3 of the supplementary material) that are slightly more stable than the known forms. The empirical FIT+GDMA potential gives the observed form II as very close in energy (0.14 kJ mol^{-1}) to form I, but there is just one lower energy structure, by 0.8 kJ mol^{-1}, $_{pyr}$83, which is significantly different (Fig. 3 of the supplementary material). The empirical potential produces structures that are less dense than with the DIFF model, as found for the observed structures (Fig. 5). This sensitivity to the energy model in ranking the structures requires understanding in terms of the variety of pairwise intermolecular interactions in the crystal structures which are thermodynamically plausible as polymorphs. The gas-phase dimer motifs[2] in the 30 lowest energy structures found in the DIFF search are given in Table IV of the supplementary material along with their lattice energies.

The analysis of the low energy crystal structures (Table IV of the supplementary material) showed that the coordination clusters are not dominated by the gas-phase dimer structures shown in Fig. 3. The most stable crystal structure ($_{pyr}$2) contains the T-shaped dimer T1 and a much distorted version
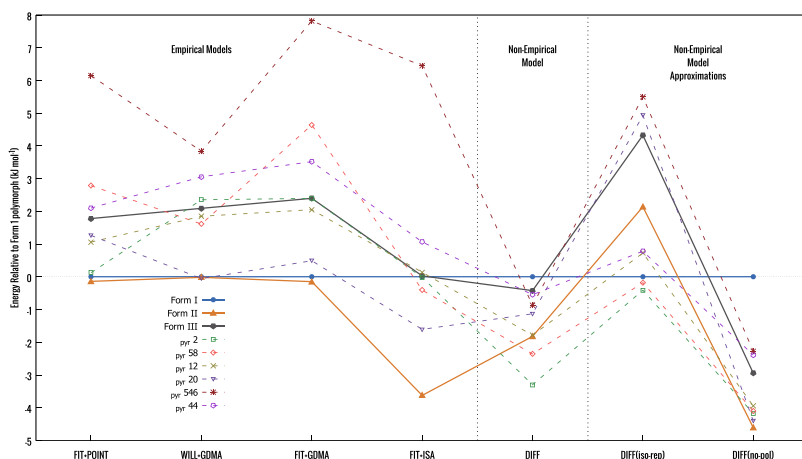

FIG. 7. Relative lattice energies of the observed and selected computer generated crystal structures of pyridine, relative to form I as a function of the model intermolecular potential. The labels for the CSP generated structures $_{pyr}$# refer to their ranking after the CrystalPredictor[72,73] stage, i.e., with the FIT+POINT model and are given in Table IV of the supplementary material.

TABLE I. Solid-state energy contributions (in kJ mol^{-1}) for DIFF in the observed and the most diverse hypothetical structures.

$_{pyr}$#	Form I	Form II	Form III	2 (GM)	58	12	20	546	44
Dimer motif	T2, bT, Hb3	T2, Hb3	None	T1	Hb1	Hb2	S1, T1	T2, Hb3	None
Electrostatic	−12.19	−16.83	−13.21	−11.74	−14.45	−13.58	−13.47	−15.86	−13.22
Dispersion	−94.37	−97.97	−100.53	−104.13	−99.57	−100.89	−102.50	−95.71	−99.51
Short-range	49.23	52.87	53.48	54.37	52.59	53.20	54.24	51.97	53.02
Polarization	−7.62	−4.83	−5.09	−6.74	−5.87	−5.46	−4.34	−6.21	−5.77
Lattice	−64.94	−66.76	−65.36	−68.24	−67.30	−66.72	−66.07	−65.81	−65.48

of the singly hydrogen bonded dimer Hb2, whereas the most stable dimer Hb1 occurs in a crystal structure ($_{pyr}$83) that is 1 kJ mol^{-1} less stable. Thus the low energy crystal structures differ in the relative contribution from each of the van der Waals contact dimers defining the structure, and each of these differs in the partitioning between the various contributions[75] in a similar manner to the dimers (Fig. 3). In Fig. 7, we display the relative energies of a selection of structures within 4 kJ mol^{-1} of the most stable, together with the three known polymorphs. The hypothetical structures were selected based on the diversity of the gas-phase-like dimers they contained, so as to allow us to more clearly explore the effects of the various energy models on the relative lattice energy and its components. The relative thermodynamic stability order varies significantly as a function of potential (Fig. 7). There is only a slight re-ranking of the structures in changing between the two empirical repulsion-dispersion models (FIT/WILL+GDMA), but there is a dramatic re-ranking caused by omitting the polarization energy or the repulsion anisotropy from the DIFF model. In addition, the global minimum from the DIFF ($_{pyr}$2) is less stable than forms I and II with the empirical models, and being denser, its stability may be an artifact of the DIFF model, as discussed later. There is also some variation with the electrostatic model used, with the re-ranking caused by using a point charge or distributed multipole representation of the same charge distribution (FIT+POINT vs FIT+GDMA) being less severe than that caused by changing the quality of charge distribution and its representation (FIT+GDMA vs FIT+ISA). This is unusual, as the use of distributed multipoles rather than the corresponding potential-derived charges usually makes a considerable improvement in CSP studies, particularly for hydrogen-bonded systems.[59]

The lattice energies of the selected crystal structures vary by less than 4 kJ mol^{-1} using the DIFF model (Table I); however, the various contributions to this energy vary considerably more. The dominant dispersion contribution varies by over 10 kJ mol^{-1} favoring denser structures that contain stacked or T-shaped dimers whose dimer energies are heavily dominated by the dispersion contribution, but this is often partially balanced out by an increase in the short range repulsion term. The electrostatic contribution varies by 5 kJ mol^{-1}, favoring the experimental forms and other structures containing the CH$\cdots$N "hydrogen bonds." The polarization energy is a small contribution but stabilizes form I (Z' = 4) the most, and the difference in the polarization energy between the two known polymorphs, at 2.8 kJ mol^{-1}, is larger than their total lattice energy difference. Although this difference is accounted for

by the lower symmetry within form I allowing larger induced moments (Fig. 2), the polarization term also stabilizes $_{pyr}$2 and all other structures that are lower in lattice energy than form II. Hence the polarization contribution is important, even for pyridine and the low energy structures re-rank substantially in the CSP which neglects this structure-dependent non-additive term by using DIFF(no-pol) (Fig. 3 of the supplementary material).

Thus the CSP study samples a wide range of relative orientations of the pyridine molecules and finds that the various contributions can balance to within the energy range of plausible polymorphism (<5 kJ mol^{-1}) in a variety of different structures.

D. Form III

On hearing of the existence of a new polymorph d5-III at around 2 GPa, the crystal energy landscape was recalculated at 2 GPa. The effect of the pressure-volume (PV) energy term significantly re-ordered the relative stability of the structures (Table V of the supplementary material) as well as increasing the density by 15%. One structure became significantly more favorable with pressure and was only 1.2 kJ mol^{-1} less stable than form II. Furthermore, this structure ($_{pyr}$35) had 7 coordinating molecules in in common with form II, and so it is very plausible as the result of a low-barrier transformation (Fig. 8), unlike the other low energy hypothetical structures at 2 GPa. This seemed likely to be form III on grounds of relative energy
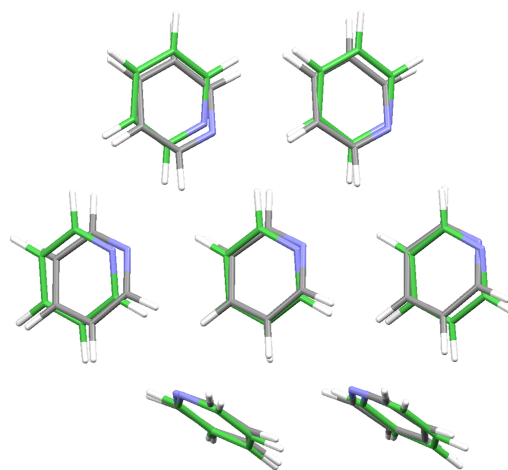

FIG. 8. Comparison of the experimental crystal structures of form II (grey h5-II PYRDNA05[48]) at 2 GPa and the form III of pyridine, refined from experimental powder data, at around 2 GPa, showing the optimum overlay of 7/15 molecules, with RMSD$_7$ = 0.364 Å.
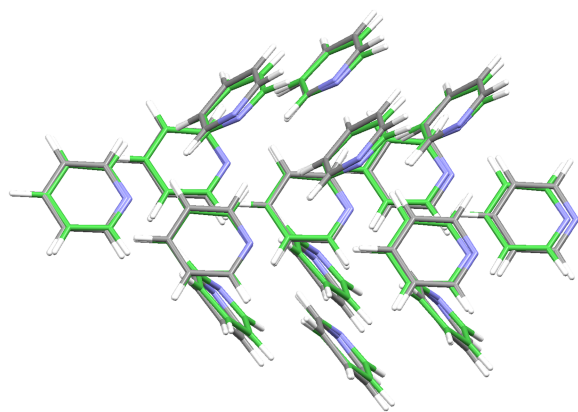
FIG. 9. Comparison of the crystal structures of the proposed form III (green $_{pyr}$35) calculated at 2 GPa and the neutron structure determined at approximately this pressure (grey)[50] showing the optimum overlay of 15/15 molecules, with RMSD$_{15}$ = 0.212 Å.

at 2 GPa and being the structure in the search with the most plausible degree of rearrangement from the known forms. The overlay of this structure ($_{pyr}$35) with the experimental structure of form III at approximately 2 GPa is shown in Fig. 9. In contrast, the structure of form III could not be identified with the empirical FIT potential: the corresponding structure $_{pyr}$35 is not amongst the most stable, independent of pressure (Fig. 3 + Table VI of the supplementary material), and is 3.3 kJ mol^{-1} less stable and less dense than form II at 2 GPa. Thus, it would not have been possible to propose a candidate structure for form II using the FIT potential solely from the knowledge of its existence at ca. 2 GPa.

IV. DISCUSSION

The non-empirical DIFF model for pyridine, which was derived from theoretical calculations on the monomer and dimer, gives a reasonable account of the structures of the polymorphs and their stability relative to hypothetical structures. The accuracy in the solid-state rivals, and arguably exceeds, that of empirical potentials that had been parametrized to experimental crystal structures and fails badly in reproducing the second virial coefficients or the behavior of the solid under pressure. This is a major advance, but this study does bring into focus both the advantages and disadvantages of having a genuine pair potential energy surface which it was hoped would be transferable between phases with the approximation that the polarization was the only significant many-body term for relative energies.

All the static lattice or enthalpy (P = 1.1 GPa and 2 GPa) structures are in good agreement with the experimental structures, given the variation between the structures that have been determined at different temperatures (5-190 K) and pressures ($\leq$1.1 GPa). The DIFF model reproduces the 5 K h5-I structure of pyridine well but overestimates the density by 5.6%. Nonetheless, perhaps this is not as serious an issue as it may seem as it has recently been noted that a quasi-harmonic modelling of crystalline imidazole using *ab initio* based potentials gave a 4% increase in molar volume on including just the zero-point energy.[76] Hence, a part of the error in the density at 5 K is probably due to neglecting the effect of the zero-point

TABLE II. The Helmholtz free energy at T = 190 K, p = 0 for both protonated and deuterated isotopes ($h5\,A_{free}$, $d5\,A_{free}$) calculated using the lattice dynamics and the DIFF model for the observed polymorphs and the global minimum (GM) in the CSP study (Fig. 6). Energies in kJ mol^{-1}.

$_{pyr}$#	FORM I	FORM II	FORM III	2 (GM)
Lattice E	**−64.939**	**−66.758**	**−65.359**	**−68.237**
h5 Zero-point E	2.027	2.163	2.146	2.114
F_{vib} (190 K)	−6.274	−5.561	−5.716	−6.282
h5 A_{free} (190 K)	**−71.213**	**−72.319**	**−71.075**	**−74.520**
d5 ZPE	1.912	2.039	2.020	1.991
F_{vib} (190 K)	−6.773	−6.050	−6.215	−6.783
d5 A_{free} (190 K)	**−71.712**	**−72.808**	**−71.574**	**−75.020**

expansion on the structure. Hence, to the extent that the static lattice energy (or enthalpy with a PV term) can model the crystal structures, DIFF models the experimentally unobservable static crystal structure, and the empirically fitted potentials implicitly include some average over zero-point energy effects and thermal expansion.

The lattice energies of the crystals are only an approximation to the thermal stability neglecting the isotope-dependent zero-point energy and thermal contributions. The free energy estimates in Table II show that including the zero-point energy and thermal corrections estimated by lattice dynamics reverses the stability of h5-forms I and III and significantly reduces the energy difference between forms I and II by over 0.7 kJ mol^{-1}. While, these harmonic, rigid-molecule lattice-dynamic estimates are approximate, they highlight that both the zero-point energy and thermal corrections are significant, and affected by deuteration, as seen in Table II. The effect of deuteration is virtually identical for every polymorph within this fixed rigid-molecule approximation. However, using experimental molecular structures can change the lattice energies by a few kJ mol^{-1}, showing that modelling small changes in molecular structure is important, even for pyridine. Most organic crystals have a significant thermal expansion, which can be very anisotropic and dependent on the specific crystal structure. Recent quasi-harmonic periodic electronic structure methods[77] and empirical potential estimates for a large dataset of organic crystals[78] show that the effects of thermal expansion are thermodynamically significant, producing an underestimation of heat capacities at high temperatures.[79] Moreover a divergence between harmonic-approximation phonon-modes and those in the crystal as modelled by Molecular Dynamics (MD) has been noted at quite low temperatures for imidazole and 5-azauracil.[80] For pyridine, which is a liquid at room temperature, the molecular motions are likely to be large in amplitude at the temperatures of the majority of experimental measurements and therefore should be more realistically modelled by a finite-temperature MD simulation. What Table II demonstrates is that both zero-point and thermal contributions are significant for the relative stability of the crystal structures, with the common observation that their inclusion reduces the energy differences between structures.[81]

The success of the DIFF model relative to potentials specifically developed for modelling the organic state derives from having a functional form based on the theory of intermolecular forces and not parameterization from the

experiment. The diversity of gas phase dimer motifs in the low energy hypothetical and experimental crystal structures emphasizes that a solid-state CSP samples the intermolecular potential energy surface much more extensively. A dimer structure only samples the potential around one configuration, whereas a crystal structure samples a wider range of relative orientations and close contacts from all the molecules in the nearest neighbor coordination sphere of typically fourteen molecules. The lattice summation samples the potential over a larger distance than a gas-phase dimer and the closest contacts in the crystals sample higher up the repulsive wall because of the attractive force from the second, and higher-order coordination spheres even at ambient pressure. The dispersion contribution is about 150% of the total lattice energy in the low energy crystals (Table I). Hence, the theoretically based distance-dependence of the long range forces is particularly important in the solid-state. Using the DIFF model with a spilt into C_6, C_8, C_{10} is more theoretically justified than using an empirical C_6 only potential where the higher dispersion coefficients have been absorbed in the parametrization. A small error in C_6 could result in a large overall error due to the lattice summation.

A similar advantage in the distance dependence is seen when the long range electrostatic terms are given by distributed multipoles over atomic point charge molecules, though the main advantage of the higher atomic multipole moments is to correctly model the orientation dependence of the electrostatic forces due to lone pair and π electron density. These non-spherical features in the charge distribution also determine the anisotropy in the repulsion. The slightly closer contacts in solid-state and diversity of contact configurations means that the relative energies of the crystal structures are very sensitive to the anisotropy in the repulsion, as shown by a considerable re-ranking of the crystal structures when this anisotropy is not explicitly fitted in the potential (Fig. 3 of the supplementary material).

The successful identification of form III as a structure which becomes relatively more stable under pressure is clear evidence of the importance of using a non-empirical anisotropic repulsion potential. It is not surprising that parameterizing an oversimplified functional form of the repulsive wall by fitting to ambient pressure crystal structures fails to extrapolate to the closer contacts sampled at moderate pressure. Applying pressure changes the intermolecular contact distances more than changing temperature and so probes interactions higher up the repulsive wall.[82,83] High-pressure recrystallization is a versatile route to generating new polymorphs, with structural properties modifying significantly with pressures around and above 1 GPa.[84–86] Hence more realistic, theory-based, non-empirical intermolecular potentials will be very important for more reliable exploration of structure-property relationships in organic crystals under pressure.

The second virial coefficients show that the DIFF model systematically over-estimates the intermolecular interactions (Fig. 4). We believe that these errors stem not from problems with the fitting procedure described by Misquitta and Stone[2] but from errors in the reference SAPT(DFT) interaction energies to which the DIFF model was fit. While a complete analysis of these problems would take us beyond the scientific scope of this paper, we can state that there are two main causes for the over-binding of the SAPT(DFT) interaction energies:

(1) Interaction energy contributions from third to infinite order in the intermolecular interaction operator are approximated by the δ_{int}^{HF} energy.[87,88] This energy includes mainly higher order induction and exchange-induction contributions, which are known to be important for capturing the effects of hydrogen bonding in water. However, this may not be suitable for configurations of pyridine in which the binding is primarily or dominantly from the dispersion interaction,[89–91] where this correction may lead to an overestimation of the binding.

(2) The choice of asymptotic correction to correct the charge density used for calculating the molecular properties, made by Misquitta and Stone,[2] may not be suitable for a strongly anisotropic (in shape) molecule such as pyridine. This is a subtle issue that is being researched in our group.

The first problem is a potentially serious one as the δ_{int}^{HF} term is needed for systems with hydrogen bonds[92] but is known to be inappropriate for dispersion-bound dimers for which it leads to an over-binding.[93,94] This term is sometimes excluded entirely[41,95] if the primary binding is from the dispersion energy as is the case for the benzene crystal. But it is as yet unclear what is to be done for a system such as pyridine which exhibits both hydrogen-bonding as well as dispersion-bonding. Work is currently underway in our groups, using CCSD(T) dimer calculations to better understand both problems, as it is essential for modelling pharmaceuticals where most crystal structures are a balance between hydrogen bonding, π-π stacking, and other dispersion interactions and conformational changes.

While a pairwise additive intermolecular potential is sufficient for a complete potential energy surface for the isolated dimer, it only accounts for the dominant contributions in condensed phases. The only explicitly non-pairwise terms included in the DIFF model are in the polarization term which leads to a net attractive many-body contribution to the lattice energy, However, in a weakly bound molecular crystal like pyridine, where most of the binding arises from the two-body dispersion energy, the many-body dispersion energy may be expected to be large and repulsive. For example, it has been shown in Refs. 38 and 96 that the three-body non-additive dispersion contributes repulsively to the benzene crystal by as much as 7%-14% of the total lattice energy. We expect a similar contribution to arise in the pyridine crystal. In particular, the dense, strongly dispersion bound global minimum structure ($_{pyr}2$) will be relatively destabilized by the three-body dispersion.

This study has shown the advantage of using non-empirical potentials for crystal structure prediction, particularly for form III which was obtained under pressure, where the repulsive wall is sampled in regions not used in fitting empirical potentials. However, it also shows how sensitive the relative thermodynamics are to the modelling assumptions. For example, $_{pyr}2$ is predicted to be more stable than form II, both at the level of the lattice energy Fig. 6 and taking into

account thermal effects Table II, though the difference is considerably reduced on application of pressure (Table V of the supplementary material) and its stability may be because of the neglect of the many-body dispersion term. However, what would be the consequence of a confident calculation that $_{pyr}2$ or any other structure was more thermodynamically stable than the observed polymorphs at accessible temperatures and pressures? The prediction of undiscovered more thermodynamically stable phases is a major justification[97] for the development and testing[8] of CSP methods, as this determines the risk of disappearing polymorphs,[98] helps justify devising experiments for finding new polymorphs[99] and can generally assist in the development of specialty organic materials. When polymorphs switch relative stability with temperature or pressure, it may often not be as obvious as a solid-state phase transformation but only apparent by recrystallization experiments in the presence of seeds of both polymorphs. It is difficult to obtain, let alone structurally characterize, the most stable crystalline forms of molecules that are liquid or gases under ambient conditions,[100] let alone change the experimental conditions sufficiently to vary the kinetics (already implicated in the isotopic polymorphism of pyridine), at low temperatures or at pressure[101] to compete with the nucleation of the known forms. The similarities between the known polymorphs of pyridine (Figs. 1 and 8) and the contrast to the different dimers in the coordination sphere in $_{pyr}2$ (Table I) emphasize that the observed structures may be kinetically favored. Hence, pyridine could well have alternative thermodynamically competitive polymorphs that have not yet been found. The DIFF model potential could be used for simulating liquid-phase pyridine, adding confidence to interpreting experimental data,[102,103] as recent simulations show that this is sensitive to the anisotropy of the electrostatic model.[46] Such simulations could suggest why the observed forms crystallize, by revealing a link between the liquid and solid state structures via the most readily formed nucleus.

Non-empirical potentials have considerable advantages over periodic electronic structure methods, or other advanced methods of calculating lattice energies of molecular crystals[104] in that they can be evaluated sufficiently readily to be used in CSP and for estimating the effect of temperature. Empirical potentials have the advantages of simpler functional forms, and the errors in the functional form, transferability assumptions, and method of simulation, such as neglect of zero-point and thermal effects, are partially absorbed into the potential. However, these advantages imply uncertainties in extrapolating to other conditions as required for solid state phase diagrams,[105] and empirical potentials cannot be expected to transfer to the gas phase, as demonstrated by the second virial coefficients (Fig. 4). In contrast, the DIFF and other non-empirical models[13,31] have the advantage that the approximations used are known and can be built upon. Although, the polarization energy is challenging to include in any force-field, this is being tackled in the development of the next generation force-fields and simulation codes.[106–108] The improved realism of intermolecular force-fields has to go alongside the further development of simulation methods and codes to include both zero-point effects and realistic temperature dependent dynamics.

In summary, there is clearly a conceptual advantage in using model intermolecular potentials where the approximations are known and controlled. The current potentials are not definitive but can act as a strong framework for the development of force-fields for organic molecules.

V. CONCLUSIONS

This paper has shown that a non-empirical distributed intermolecular pair potential (DIFF), derived from the theory of intermolecular forces and SAPT(DFT) calculations using the CAMCASP and ORIENT programs, can be used for modelling the solid-state of pyridine with a realism that exceeds that of transferable empirical potentials that have been derived for modelling the crystalline state. Furthermore, this study stresses the importance of crystal structure prediction as a comprehensive way of testing the robustness of intermolecular model potentials. The DIFF model was particularly effective for studying the effects of pressure on the relative stability of structures, as shown by the identification of form III. Nevertheless, empirical potentials do have the advantage of absorbing many errors, including approximations in the simulation methods, and being more readily implemented in existing codes. Hence, contrasting CAMCASP-derived potentials, and other models driven by the theory of intermolecular forces, with more approximate models, should produce a hierarchy of model force-fields that can be used with molecule-specific knowledge of the effects of various approximations.

SUPPLEMENTARY MATERIAL

See supplementary material for tables of CSP generated crystal structures, further calculations, and the potential models including parameter input files.

ACKNOWLEDGMENTS

[1] A. J. Stone, *The Theory of Intermolecular Forces* (Oxford University Press, Oxford, 2013), Vol. 2.
[2] A. J. Misquitta and A. J. Stone, J. Chem. Theory Comput. **12**, 4184 (2016).
[3] M. Rigby *et al.*, *The Forces Between Molecules* (Oxford Science Publication, 1986).
[4] C. C. Sun, J. Pharm. Sci. **98**, 1671 (2009).
[5] A. Y. Lee, D. Erdemir, and A. S. Myerson, Annu. Rev. Chem. Biomol. Eng. **2**(2), 259 (2011).
[6] J. Nyman, O. S. Pundyke, and G. M. Day, Phys. Chem. Chem. Phys. **18**, 15828 (2016).
[7] A. J. Cruz-Cabeza, S. M. Reutzel-Edens, and J. Bernstein, Chem. Soc. Rev. **44**, 8619 (2015).

[8]A. M. Reilly *et al.*, Acta Crystallogr., Sect. B **72**, 439 (2016).

[9]A. J. Pertsin and A. I. Kitaigorodsky, *The Atom-Atom Potential Method. Applications to Organic Molecular Solids*, Springer Series in Chemical Physics (Springer-Verlag, Berlin, 1987).

[10]S. Lifson, A. T. Hagler, and S. Dauber, J. Am. Chem. Soc. **101**, 5111 (1979).

[11]A. Hesselmann and G. Jansen, Chem. Phys. Lett. **367**, 778 (2003).

[12]D. E. Williams, J. Comput. Chem. **22**, 1154 (2001).

[13]M. Metz, K. Piszczatowski, and K. Szalewicz, J. Chem. Theory Comput. **12**, 5895 (2016).

[14]A. S. Tulegenov *et al.*, J. Chem. Phys. **126**, 094305 (2007).

[15]R. Bukowski *et al.*, J. Chem. Phys. **125**, 044301 (2006).

[16]M. P. Hodges *et al.*, J. Chem. Phys. **120**, 710 (2004).

[17]M. P. Hodges, R. J. Wheatley, and A. H. Harvey, J. Chem. Phys. **116**, 1397 (2002).

[18]P. Jankowski *et al.*, J. Phys. Chem. A **119**, 2940 (2015).

[19]G. Garberoglio *et al.*, J. Chem. Phys. **146**(5), 054304 (2017).

[20]G. Donoghue *et al.*, J. Mol. Spectrosc. **330**, 170 (2016).

[21]X. G. Wang, T. Carrington, and R. Dawes, J. Mol. Spectrosc. **330**, 179 (2016).

[22]J. Klos *et al.*, J. Chem. Phys. **146**(11), 114301 (2017).

[23]H. C. Schewe *et al.*, J. Chem. Phys. **142**(20), 204310 (2015).

[24]M. Kirste *et al.*, Science **338**, 1060 (2012).

[25]K. M. Walker *et al.*, Astrophys. J. **811**, 27 (2015).

[26]K. Szalewicz, Acc. Chem. Res. **47**, 3266 (2014).

[27]R. Podeszwa, B. M. Rice, and K. Szalewicz, Phys. Chem. Chem. Phys. **11**, 5512 (2009).

[28]A. J. Misquitta *et al.*, Chem. Phys. Lett. **456**, 105 (2008).

[29]J. Klimes and A. Michaelides, J. Chem. Phys. **137**, 120901 (2012).

[30]Y. Huang, Y. Shao, and G. J. Beran, J. Chem. Phys. **138**, 224112 (2013).

[31]M. Van Vleet *et al.*, J. Chem. Theory Comput. **12**, 3851 (2016).

[32]M. Tafipolsky and K. Ansorg, J. Chem. Theory Comput. **12**, 1267 (2016).

[33]R. Podeszwa and K. Szalewicz, J. Chem. Phys. **136**, 161102 (2012).

[34]A. J. Misquitta and K. Szalewicz, J. Chem. Phys. **122**, 214109 (2005).

[35]A. J. Misquitta *et al.*, J. Chem. Phys. **123**, 214103 (2005).

[36]K. Szalewicz, K. Patkowski, and B. Jeziorski, in *Intermolecular Forces and Clusters II*, edited by D. J. Wales (Springer-Verlag Berlin, Berlin, 2005), p. 43.

[37]A. Hesselmann, G. Jansen, and M. Schutz, J. Chem. Phys. **122**(1), 014103 (2005).

[38]R. Podeszwa, B. M. Rice, and K. Szalewicz, Phys. Rev. Lett. **101**, 115503 (2008).

[39]R. Podeszwa *et al.*, Phys. Chem. Chem. Phys. **9**, 5561 (2007).

[40]T. S. Totton, A. J. Misquitta, and M. Kraft, J. Chem. Theory Comput. **6**, 683 (2010).

[41]R. Podeszwa, R. Bukowski, and K. Szalewicz, J. Phys. Chem. A **110**, 10345 (2006).

[42]D. Mootz and H. G. Wussow, J. Chem. Phys. **75**, 1517 (1981).

[43]P. J. Wheatley, Acta Crystallogr. **13**, 80 (1960).

[44]B. A. Middleton and J. R. Partington, Nature **141**, 516 (1938).

[45]M. Podsiadlo, K. Jakobek, and A. Katrusiak, CrystEngComm **12**, 2561 (2010).

[46]C. Baker and G. Grant, J. Chem. Theory Comput. **3**, 530 (2007).

[47]A. T. Anghel, G. M. Day, and S. L. Price, CrystEngComm **4**, 348 (2002).

[48]S. Crawford *et al.*, Angew. Chem., Int. Ed. **48**, 755 (2009).

[49]J. van de Streek and M. A. Neumann, CrystEngComm **13**, 7135 (2011).

[50]N. P. Funnell, W. G. Marshall, and S. Parsons, "Experimental determination of pyridine form III" (unpublished).

[51]S. L. Price *et al.*, Phys. Chem. Chem. Phys. **12**, 8478 (2010).

[52]C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).

[53]J. P. Perdew, M. Emzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).

[54]J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[55]O. Gritsenko and E. J. Baerends, J. Chem. Phys. **121**, 655 (2004).

[56]T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

[57]A. J. Misquitta, A. J. Stone, and F. Fazeli, J. Chem. Theory Comput. **10**, 5405 (2014).

[58]A. J. Stone, GDMA code (University of Cambridge, Cambridge, United Kingdom, 2010), http://wwwstone.ch.cam.ac.uk/programs.html.

[59]G. M. Day, W. D. S. Motherwell, and W. Jones, Cryst. Growth Des. **5**, 1023 (2005).

[60]A. J. Misquitta and A. J. Stone, Mol. Phys. **106**, 1631 (2008).

[61]A. J. Misquitta and A. J. Stone, J. Chem. Phys. **124**, 024111 (2006).

[62]G. J. Williams and A. J. Stone, J. Chem. Phys. **119**, 4620 (2003).

[63]K. T. Tang and J. P. Toennies, J. Chem. Phys. **80**, 3726 (1984).

[64]D. E. Williams and S. R. Cox, Acta Crystallogr., Sect. B: Struct. Sci. **40**, 404 (1984).

[65]S. L. Price, Chem. Soc. Rev. **43**, 2098 (2014).

[66]A. J. Stone *et al.*, ORIENT and CamCASP, University of Cambridge, 2015, http://wwwstone.ch.cam.ac.uk/programs.html.

[67]J. D. Cox and R. J. L. Andon, Trans. Faraday Soc. **54**, 1622 (1958).

[68]R. J. L. Andon *et al.*, Trans. Faraday Soc. **53**, 1074 (1957).

[69]J. Nyman and G. M. Day, CrystEngComm **17**, 5154 (2015).

[70]G. M. Day, S. L. Price, and M. Leslie, J. Phys. Chem. B **107**, 10919 (2003).

[71]G. M. Day, S. L. Price, and M. Leslie, Cryst. Growth Des. **1**, 13 (2001).

[72]P. G. Karamertzanis and C. C. Pantelides, J. Comput. Chem. **26**, 304 (2005).

[73]P. G. Karamertzanis and C. C. Pantelides, Mol. Phys. **105**, 273 (2007).

[74]C. F. Macrae *et al.*, J. Appl. Crystallogr. **39**, 453 (2006).

[75]A. Gavezzotti, CrystEngComm **5**, 429 (2003).

[76]Y. Heit and G. Beran, Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater. **72**, 514 (2016).

[77]A. M. Reilly and A. Tkatchenko, J. Chem. Phys. **139**, 024705 (2013).

[78]J. Nyman and G. Day, Phys. Chem. Chem. Phys. **18**, 31132 (2016).

[79]H. Buchholz *et al.*, "Thermochemistry of Racemic and Enantiopure Organic Crystals for Predicting Enantiomer Separation," Cryst. Growth Des. (published online).

[80]A. E. Gray *et al.*, Mol. Phys. **102**, 1067 (2004).

[81]J. D. Dunitz, G. Filippini, and A. Gavezzotti, Helv. Chim. Acta **83**, 2317 (2000).

[82]N. P. Funnell *et al.*, CrystEngComm **15**, 1047 (2013).

[83]L. Budd *et al.*, Chem. Cent. J. **9**, 18 (2015).

[84]F. P. A. Fabbiani *et al.*, Cryst. Growth Des. **7**, 1115 (2007).

[85]F. P. A. Fabbiani *et al.*, J. Cryst. Growth **275**, 185 (2005).

[86]F. P. A. Fabbiani *et al.*, CrystEngComm **6**, 504 (2004).

[87]M. Jeziorska, B. Jeziorski, and J. Cizek, Int. J. Quantum Chem. **32**, 149 (1987).

[88]R. Moszynski, T. G. A. Heijmen, and B. Jeziorski, Mol. Phys. **88**, 741 (1996).

[89]A. J. Misquitta and A. J. Stone, J. Chem. Theory Comput. **4**, 7 (2008).

[90]R. Podeszwa and K. Szalewicz, Chem. Phys. Lett. **412**, 488 (2005).

[91]K. Patkowski, K. Szalewicz, and B. Jeziorski, J. Chem. Phys. **125**, 154107 (2006).

[92]U. Góra *et al.*, J. Chem. Phys. **135**, 224102 (2011).

[93]K. Pernal *et al.*, Phys. Rev. Lett. **103**, 263201 (2009).

[94]K. Patkowski, K. Szalewicz, and B. Jeziorski, Theor. Chem. Acc. **127**, 211 (2010).

[95]R. Podeszwa and K. Szalewicz, Phys. Chem. Chem. Phys. **10**, 2735 (2008).

[96]M. Kennedy *et al.*, J. Chem. Phys. **140**, 121104 (2014).

[97]S. L. Price, D. E. Braun, and S. M. Reutzel-Edens, Chem. Commun. **52**, 7065 (2016).

[98]D. K. Bucar, R. W. Lancaster, and J. Bernstein, Angew. Chem., Int. Ed. **54**, 6972 (2015).

[99]V. K. Srirambhatla *et al.*, Chem. Commun. **52**, 7384 (2016).

[100]M. Bujak *et al.*, Chem. Commun. **47**, 8769 (2011).

[101]M. A. Neumann *et al.*, Nat. Commun. **6**, 7793 (2015).

[102]E. Megiel *et al.*, J. Mol. Struct. **569**, 111 (2001).

[103]I. Bako, T. Radnai, and G. Palinkas, Z. Naturforsch. **51**, 859 (1996).

[104]G. J. O. Beran, Chem. Rev. **116**, 5567 (2016).

[105]D. Hofmann and L. Kuleshova, Cryst. Growth Des. **14**, 3929 (2014).

[106]A. Albaugh *et al.*, J. Phys. Chem. B **120**, 9811 (2016).

[107]T. Bereau, C. Kramer, and M. Meuwly, J. Chem. Theory Comput. **9**, 5450 (2013).

[108]I. Nessler, J. Litman, and M. Schnieders, Phys. Chem. Chem. Phys. **18**, 30313 (2016).