CRYSTAL GROWTH & DESIGN

pubs.acs.org/crystal

Article

Crystal Structure Predictions for 4-Amino-2,3,6-trinitrophenol Using a Tailor-Made First-Principles-Based Force Field

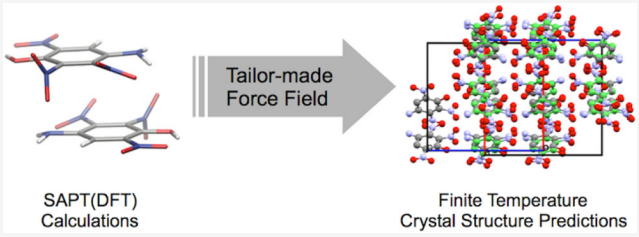
Michael P. Metz, Muhammad Shahbaz, Hongxing Song, Leslie Vogt-Maranto, Mark E. Tuckerman,* and Krzysztof Szalewicz*

Cite This: *Cryst. Growth Des.* 2022, 22, 1182−1195

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Predictions of crystal structures from first-principles electronic structure calculations and molecular simulations have been performed for an energetic molecule, 4-amino-2,3,6-trinitrophenol. This physics-based approach consists of a series of steps. First, a tailor-made two-body potential energy surface (PES) was constructed with recently developed software, autoPES, using symmetry-adapted perturbation theory based on a density-functional theory description of monomers [SAPT(DFT)]. The fitting procedure ensures asymptotic correctness of the PES by employing a rigorous asymptotic multipole expansion, which seamlessly integrates with SAPT(DFT) interaction energies. Next, crystal structure prediction (CSP) was performed by generating possible crystal structures with rigid molecules, minimizing these structures using the SAPT(DFT) force field, and running isothermal−isobaric molecular dynamics (MD) simulations with flexible molecules based on the tailor-made SAPT(DFT) intermolecular force field and a generic/SAPT(DFT) intramolecular one. This workflow led to the experimentally observed structure being identified as one of the forms with the lowest lattice energy, demonstrating the success of a first-principles, bottom-up approach to CSP. Importantly, we argue that the accuracy of the intermolecular potential, here the SAPT(DFT)-based potential, is determinative of the crystal structure, while generic/SAPT(DFT) force fields can be used to represent the intramolecular potential. This force field approach simplifies the CSP workflow, without significantly compromising the accuracy of the prediction.

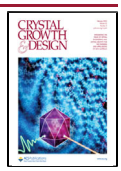


I. INTRODUCTION

Organic molecular crystals are materials of critical importance in numerous industries including pharmaceuticals, electronics, vector-borne disease management, and defense. In any of these areas, the performance of a given crystalline formulation depends on its crystal structure. If a given compound can crystallize into multiple forms—commonly referred to as *polymorphism*—then mapping out the landscape of possible polymorphs becomes the first step toward selecting a polymorph with desired properties. In particular, since transitions from metastable to stable polymorphs can cause drugs to fail,[1−4] affect the performance of organic semiconductors,[5] or alter the energetic properties of explosives,[6] predicting these structures and ranking them thermodynamically becomes all the more critical.

Polymorph landscapes can be generated using computational techniques tailored for crystal structure prediction (CSP). CSP is an essential step in the discovery of metastable polymorphs that may be sufficiently stable for a desired application. In CSP, the energy landscape of crystalline phases is explored to find a set of thermodynamically plausible polymorphs. Starting with information only about the molecule itself, CSP generates trial crystal structures and evaluates their lattice energies. While the idea is, in principle, straightforward, the practical implementation of CSP methods faces significant challenges in both the structure search and the calculation of accurate energies. Over the years, CSP has matured,[7−9] and its progress has been followed in a series of blind test competitions organized by the Cambridge Crystallographic Data Centre (CCDC).[10−15] Strategies for structure searches include random searches,[16,17] genetic algorithms,[18−20] simulated annealing,[21] or Monte Carlo methods.[22] For energy calculations, significant advances have been made in the development of reliable force fields,[23,24] and at the same time, the applications of electronic structure calculations with periodic boundary conditions to molecular crystals are also becoming increasingly important.[25−29] While the latter can provide very accurate energetics, they are also computationally much more expensive and have been mainly used for the final ranking of polymorphs after performing the structure search with a force field. However, a separate energy ranking step can

Received: September 27, 2021
Revised: January 5, 2022
Published: January 24, 2022



be avoided if the force field is of sufficient quality. Generic, empirical force fields are generally insufficiently accurate, and one has to resort to tailor-made force fields fitted to ab initio calculations. The first CSPs based on first-principles potential energy surfaces (PESs) have been performed by Podeszwa et al. in refs 30−32, using a workflow including crystal packing, lattice energy minimizations, and molecular dynamics (MD) simulations. A similar approach, except without the MD step, was used by Misquitta et al.[33] Both approaches fitted PESs to symmetry-adapted perturbation theory[34] based on density functional theory (SAPT(DFT)) calculations for dimers.[35−43] Some of us and coauthors have previously established a CSP workflow[15,44] combining a random search together with MD and enhanced sampling[45,46] to predict low free-energy crystal structures at experimentally relevant conditions.[15,44,47−49] This workflow is now being employed and adapted by other groups.[50] It has been shown, in particular, that accurate landscapes[15,24] can be predicted within this protocol using molecule-specific force fields generated using SAPT(DFT).

In this Article, we employ a version of our CSP workflow using a tailor-made SAPT(DFT) force field to determine the structure of an energetic material, 4-amino-2,3,6-trinitrophenol (**1**, see Figure 1). The key component of our workflow is an accurate intermolecular force field, with the total interaction energy determined based on the sum of pairwise interactions between molecules. Compound **1** poses a challenge for our dimer-based approach to generating the molecule-specific force field as it lacks internal symmetry, exhibits a large dipole moment, and is fairly polarizable. The outline of this paper follows our CSP workflow. First, an intermolecular PES for dimer configurations of **1** has been developed from first-principles quantum mechanical calculations, that is, by fitting the PES to ab initio interaction energies computed using SAPT(DFT). This step requires generation of suitable monomer geometries and subsequent calculation of ab initio interaction energies for sampled dimers (section II.A). The automated selection of dimers and fitting of the PES is controlled by the autoPES software package[24,51] (section II.B). Second, possible crystal structures are generated by randomly sampling monomer orientations and locations within random unit cells of a given space group. The crystals generated in this way are first optimized with rigid molecules using an empirical force field, followed by rigid molecule optimization using the SAPT(DFT) potential. Notice that this step is different from the workflow used by some of the present authors in the sixth CCDC blind test, where a simplified version of the SAPT(DFT) PES was used in the second stage lattice energy minimizations. The 0 K optimizations are followed by flexible-molecule MD simulations under ambient conditions to refine the structures using the SAPT(DFT) intermolecular force field plus generic/SAPT(DFT) intramolecular terms (section II.C). To evaluate the accuracy of our results, the generated potentials are analyzed with respect to SAPT(DFT) calculations (section III.A), and monomer and dimer geometries corresponding to the minima on the PES are compared to the experimental crystal structure (section III.B and III.C). Finally, we show that with a tailor-made intermolecular force field, the experimentally observed form is found to be one of the lowest energy crystal structures (section III.D).

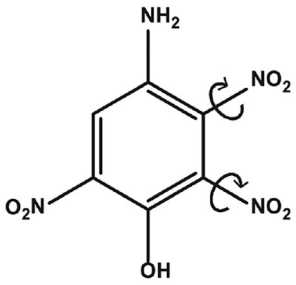


Figure 1. 4-Amino-2,3,6-trinitrophenol (**1**) molecule. Displacements of the two easily rotatable nitro groups (indicated by arrows) are included in the set of short-range dimer configurations used to fit the intermolecular potential.

II. METHODS

II.A. First-Principles Calculations. The first step in the CSP workflow is to generate an appropriate monomer geometry. The monomer geometry can be taken from an experimental structure if available or can be determined by optimizing the molecule with an electronic-structure method to obtain the so-called equilibrium (r_e) geometry. Although experimental structure is known for **1**,[52] we found the r_e geometry from scratch by performing DFT optimization using the ORCA electronic structure package,[53,54] the PBE0 functional,[55,56] and the aug-cc-pVTZ basis set.[57] This geometry was used as the reference for PES development and the subsequent structure generation steps.

The dimer PES has been fitted to two types of ab initio calculations: SAPT(DFT) at close range and to an expansion in inverse powers of intermonomer separation R, built from ab initio computed monomer multipole moments and polarizabilities, at long-range. SAPT(DFT) was used in the density-fitting version[39,58,59] coded in the SAPT2016 software package.[60] Monomer DFT calculations were performed using the ORCA electronic structure package.[53,54] The PBE functional[55] was used with the gradient-regulated asymptotic correction (GRAC).[61,62] This functional, as well as PBE0, were shown to be among the best in SAPT applications; see a recent discussion in ref 63. PBE was chosen since SAPT calculations with nonhybrid functionals are significantly faster than with hybrid ones. The ionization potential of the monomer, required for GRAC, was computed using separate PBE calculations. The aug-cc-pVDZ basis set[57] together with the corresponding correlation-energy fitted auxiliary bases from ref 64, were used in all cases. In the SAPT calculations, the basis was in the monomer-centered "plus" basis set (MC$^+$BS) form with midbond functions, see ref 65. Midbond function exponents and placement are as described in ref 66, with the midbond auxiliary basis set from ref 59. See Table S1 for values of some properties of **1**.

The component of the interaction energy arising from the Coulomb interaction of charge distributions of the monomers is called the electrostatic energy and is given by the SAPT correction $E^{(1)}_{\mathrm{elst}}$. The antisymmetrization of the total wave functions required by the fermionic nature of electrons leads to permutations (exchanges) of electrons between monomers and when applied to zeroth-order wave functions results in the first-order exchange energy, $E^{(1)}_{\mathrm{exch}}$. The electric field, due to the permanent charge distribution of monomer A (B), induces multipole moments on monomer B (A), and these multipole moments interact with the permanent multipoles of monomer A (B), resulting in the induction energy, $E^{(2)}_{\mathrm{ind}}$. The long-range correlation between electrons from different monomers results in the dispersion interaction, $E^{(2)}_{\mathrm{disp}}$. The antisymmetrization of the induction and dispersion wave functions leads to the exchange-induction and exchange-dispersion energies, denoted as $E^{(2)}_{\mathrm{exch\text{-}ind}}$ and $E^{(2)}_{\mathrm{exch\text{-}disp}}$, respectively. The components $E^{(2)}_{\mathrm{ind}}$ and $E^{(2)}_{\mathrm{disp}}$ were computed from coupled Kohn−Sham (CKS) frequency-dependent density susceptibilities (FDDS) (in the former case, only at zero frequency). The fast method of ref 67 was used to obtain FDDSs. When this approach is used, $E^{(2)}_{\mathrm{exch\text{-}disp}}$ can be computed only in the uncoupled form. It was then scaled to approximate the CKS value as described in ref 42. The term $E^{(2)}_{\mathrm{exch\text{-}ind}}$ was computed from coupled amplitudes.

https://doi.org/10.1021/acs.cgd.1c01117
Cryst. Growth Des. 2022, 22, 1182−1195

For significantly polar monomers, the $\delta E^{HF}_{int,resp}$ correction[68,69] should be included in addition to the standard SAPT(DFT) corrections. The term $\delta E^{HF}_{int,resp}$ accounts mostly for induction and exchange-induction effects beyond second order. The inclusion of $\delta E^{HF}_{int,resp}$ increases the accuracy of the SAPT(DFT) interaction energies for most systems but also increases the associated computational cost by about 60%. A criterion is automatically checked by autoPES to determine if $\delta E^{HF}_{int,resp}$ should be included for a given dimer based on the magnitude of the electric dipole moment of its monomers and the induction component of its interaction energy, see ref 24. This criterion was met, so the $\delta E^{HF}_{int,resp}$ correction was included in the total interaction energy for each of the close-range grid points.

The total second-order interaction energy, $E^{(2)}$, is the sum of the corrections discussed above

$$E^{(2)} = E^{(2)}_{ind} + E^{(2)}_{exch\text{-}ind} + E^{(2)}_{disp} + E^{(2)}_{exch\text{-}disp} \quad (1)$$

We will denote the sum of $E^{(2)}$ and $\delta E^{HF}_{int,resp}$ as $E^{(2)} + \delta$.

In the asymptotic region, the interaction energy is computed using the multipole expansion of the interaction potential operator rather than the much more computationally expensive SAPT(DFT) method. In the original autoPES approach, the origins of this expansion are placed at the centers of mass (COM) of monomers, that is, the standard asymptotic expansion described in refs 34 and 70 is used. Here, we have used instead an extension of the distributed asymptotic expansion approach of ref 71. The distributed expansion is a sum of expansions located on each of the atoms. This approach converges at shorter monomer separations than the COM-COM asymptotic approach, and so is appropriate for monomers of the size investigated here. The coefficients of this expansion are computed from ab initio distributed monomers' charge densities and FDDSs. These monomer properties are computed using the same basis set and level of theory as is used for the close-range calculations. Therefore, the resulting interaction energies connect seamlessly to those of the close-range calculations. While in ref 71 only the dispersion energies were computed, requiring distributed dynamic polarizabilities, we compute here also the electrostatic and induction energies, requiring also static distributed polarizabilities and multipole moments. Static distributed polarizabilities were computed by us using the same constrained density-fitting of FDDSs as is used for the dynamic polarizabilities, described in ref 71. The distributed multipole moments were computed from distributed monomer charge densities. Whereas the density-fitting constraint of ref 71 cannot be used for fitting charge densities, it is possible to devise constraints appropriate in this case. However, we have performed unconstrained fitting of charge densities. This approach is much simpler than other methods of distributing densities (see, for example, ref 72) but is completely adequate in our case for the following reasons. Our fitting represents the total monomer density $\rho(\mathbf{r})$ as the sum of atomic densities $\rho_a(\mathbf{r})$. As already stated, the latter densities are used to compute a set of multipole moments on each atom, which are subsequently used to compute electrostatic energies at asymptotic separations only; see section II.B. Such electrostatic energies agree very well in that region with those obtained from ab initio multipole moments located at the COM of each monomer. We then fit this set of electrostatic energies using only Coulomb interactions of optimized partial charges (this step is identical as in ref 24). Only the latter expansion is used at close-range separations.

II.B. PES Generation. The autoPES software[24,51] that we used fits dimer PESs assuming rigid-geometry monomers. Although a new version of this software, flex-autoPES, can include intramonomer degrees of freedom,[73] the rigid-geometry approximation was used here to reduce the complexity of the fit. If the two soft nitro group rotations were included explicitly as intramonomer degrees of freedom, the resulting PES would be 10-dimensional, with two additional degrees of freedom per monomer. Such a fit would include many more free parameters, and therefore would require a larger set of training data, leading to a much more expensive set of calculations than in the case of the present 6-dimensional PES. However, since the effects of varying selected internal degrees of freedom can be included using autoPES in an average way, we have done so in this case, as is discussed below.

The generation of potential energy surfaces is almost entirely automated in autoPES. The procedure can be roughly divided into five parts: asymptotic calculations, generation of a grid of close-range dimer configurations, calculation of interaction energies at each such grid point, fitting an analytic functional form to the data, and finally evaluation of the quality of the fit and iterative improvement. Each step is described in detail in ref 24. We give a brief outline below.

The purpose of the asymptotic calculations is to describe the interaction energy in as large a region of the dimer configuration space as possible using only monomer properties. This reduces the number of close-range dimer configurations required for the fit, and by extension, the overall computational cost of PES generation. This approach works when the monomers are sufficiently separated that charge overlap effects can be neglected. The region where this condition is met is determined by autoPES using geometrical criteria. A grid of about 10 000 long-range dimer configurations is selected and then the corresponding interaction energies are computed using the distributed multipole expansion.

The region where charge overlap effects cannot be neglected, which extends to about 1.5 times the radial van der Waals minimum separation and is referred to as the close-range grid region, must be adequately sampled by SAPT(DFT) calculations. The close-range grid is generated using a guided Monte Carlo procedure, which samples the entire relevant region of configuration space while placing a higher density of grid points near the most physically important regions. A guiding potential is used to determine the approximate interaction energy and locations of global and local minima. More grid points are placed in more energetically negative regions, and, in addition, a small percentage of points are placed specifically near each local minimum. In the initial iteration of grid generation, the OPLS-AA force field of Jorgensen and co-workers[74] is used as the guiding potential, while in later iterations the generated PES of the previous iteration is used. The interaction energy at each close-range grid point is computed as described in section II.A.

As discussed above, monomer **1** includes soft intramonomer degrees of freedom, that is, nitro-group rotation angles, which can change relatively significantly from their gas-phase values in the presence of other monomers in a crystal. Since in the MD stage of our CSPs the monomers are allowed to deform, we fitted rigid-monomer, six-dimensional form of PESs to grid points including some monomer deformations. For any interacting rigid monomers, each grid point is described by a set of six coordinates: the distance between the monomers' COMs R and five Euler angles describing the relative angular orientation of the monomers. However, the functional fit form used depends only on the atom−atom distances rab, where atom a (b) belongs to monomer A (B). The PESs therefore can account for internal deformations of the monomers to a limited degree, even if the training data contains only a single, fixed geometry for each monomer. In the present work, we went beyond this approximation and the usual rigid-monomer training set was extended by including monomers with explicitly varied coordinates that specify intramonomer deformations. In this way, dimer configurations including such deformations are better sampled. Four additional coordinates at each intermonomer grid point were used per dimer to represent rotation of the two nitro groups indicated in Figure 1. (The rotation of the third nitro group was not included because it is well constrained by a hydrogen bond to the neighboring hydroxyl group.) The ranges of the nitro group rotations were determined by the protocol of the flexible-monomer version of autoPES.[73] In this protocol, the range of sampling of an intramolecular degree of freedom is defined by an algorithm restricting the maximum extension or shortening of monomer's interatomic distances relative to the equilibrium ones, see eq 4 in ref 73. This resulted in about ±17° maximum rotations of the nitro groups. Within this range, the rotation angles were chosen at random, as described in ref 73. Thus, the equilibrium configuration was not in any way preferred. Note that this implicit inclusion of intramonomer degrees of freedom does not

https://doi.org/10.1021/acs.cgd.1c01117
Cryst. Growth Des. 2022, 22, 1182−1195

change the functional form of the fit: the same functional form as in the rigid-monomer case is simply optimized using a data set, which includes the intramonomer deformations, and the parameters of the fit assume values that are averages of those that would have been obtained for sets of training data, each with different rigid monomers. The resulting flexibilized rigid-monomer fit for the intermolecular PES was used with a generic/SAPT(DFT) intramonomer potential in MD simulations (described in more detail in section II.C).

After the asymptotic and close-range interaction energies are computed, the data is fit with an analytic function V of the form

$$V = \sum_{a\in A}\sum_{b\in B} u_{ab}(r_{ab}) \tag{2}$$

where a and b range over the sets of atoms in monomers A and B, respectively. The atom−atom function u_{ab} is of the form

$$u_{ab}(r_{ab}) = \left[1 + \sum_{i=1}^{2} a_i^{ab}(r_{ab})^i\right] e^{\alpha^{ab}-\beta^{ab}r_{ab}} + \frac{A_{12}^{ab}}{(r_{ab})^{12}} + \frac{q_a q_b}{r_{ab}} - \sum_{n=6,8} f_n(\delta_n^{ab}, r_{ab})\frac{C_n^{ab}}{(r_{ab})^n} \tag{3}$$

where f_n are the Tang−Toennies damping functions[75]

$$f_n(\delta_n^{ab}, r_{ab}) = 1 - e^{-\delta r}\sum_{m=0}^{n}\frac{(\delta r)^m}{m!} \tag{4}$$

We will refer to the sum of the first two terms in eq 3 as V_{exp}, the third term as V_{elst}, and the final term as $V_{\mathrm{asym}}^{(2)}$.

The parameters q_a and q_b of the V_{elst} term are fitted to the values of $E_{\mathrm{elst}}^{(1)}$ computed using the distributed asymptotic multipole expansion. In this fit, the charges were constrained based on the atomic charges computed using the CHELPG method.[76] Since no damping is used in this term, V_{elst} is expected to differ from $E_{\mathrm{elst}}^{(1)}$ at short separations. The C_n^{ab} coefficients in $V_{\mathrm{asymp}}^{(2)}$ (constrained to be positive and in a combination rule form, see below) are fit to the sum of the asymptotic multipole expansions of $E_{\mathrm{ind}}^{(2)}$ and $E_{\mathrm{disp}}^{(2)}$. The damping parameters δ_n^{ab} are fit to $E^{(2)} + \delta$ computed at close range. Note that this means that the damping is used to recover not only the charge-overlap effects, but also the second-order exchange effects. Fitting the sum of $E_{\mathrm{ind}}^{(2)}$ and $E_{\mathrm{exch\text{-}ind}}^{(2)}$ is necessary to trim the excessive values of $E_{\mathrm{ind}}^{(2)}$ at short separations resulting from unphysical tunneling of electrons between monomers due to the violation of Pauli's exclusion principle in Rayleigh−Schrödinger perturbation theory. This approach is not needed for the dispersion component, but we use it for consistency. The use of $\delta E_{\mathrm{int,resp}}^{\mathrm{HF}}$ in the fitting of the damping factors is a bit inconsistent since we neglect the asymptotic behavior of this term, but since it decays as $1/R^9$, this neglect is inconsequential. Finally, the free parameters α^{ab}, β^{ab}, a_i^{ab}, and A_{12}^{ab} of the V_{exp} term are fit to $E_{\mathrm{int}} - V_{\mathrm{elst}} - V_{\mathrm{asym}}^{(2)}$. This term mainly reproduces the $E_{\mathrm{exch}}^{(1)}$ component, but also takes care of a part of the short-range electrostatic interactions. For details of these fitting stages; see ref 24 section VI.

The coefficients A_{12}^{ab} are constrained to be positive and are included to ensure the correct repulsive behavior of the potential at very close range. We constrain the C_n^{ab} parameters by using the geometric mean combination rule $C_n^{ab} = \sqrt{C_n^a C_n^b}$, where the parameters C_n^a and C_n^b are defined per atom, rather than per atom-pair. Similarly, the parameters α^{ab} and β^{ab} are constrained using the arithmetic sum combination rule. We have found that such constraints have only a small effect on the accuracy of the resulting PES, while substantially reducing the number of free parameters.

For all cases in which some sites of a molecule are symmetrically equivalent, autoPES constrains all parameters associated with those sites to have the same value to preserve symmetry. To reduce fitting complexity, we also applied such constraints to some sites that are not exactly equivalent but are approximately so. In particular, all nitro groups are the same, even if they have stabilizing hydrogen bonds with a neighboring group. This additional constraint sacrifices some of the ability of the functional form to fit the ab initio data, but results in fewer free parameters and so fewer grid points are required. All atom equivalences are shown, together with the partial charges q_x, in Figure 2.

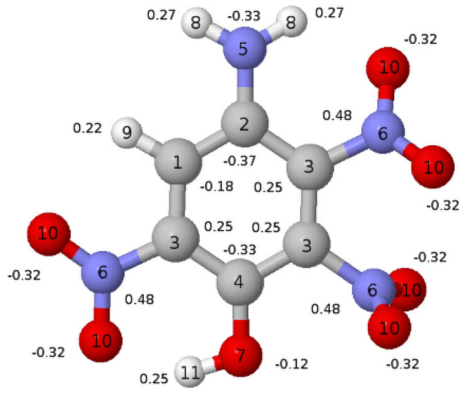


Figure 2. Equivalent atom types and partial charges of **1**. Atom equivalences are indicated by the atom numbering.

The PES was developed in an iterative way. Following each fitting stage, the quality of the PES is evaluated, and, if convergence criteria are not met, additional iterations are performed. There are two criteria used in this evaluation. The first criterion ensures that the repulsive wall does not contain "holes", that is, regions where the PES does not have the physically correct repulsive behavior at close range. Due to the cancellations between large positive and negative terms in the functional form used, such "holes" can occur if even a small region of the repulsive wall of the PES is insufficiently sampled. In such cases, additional grid points are placed at the appropriate configurations, and the potential is refit. The second criterion evaluates the general accuracy of the PES by computing its root-mean-square error (RMSE) on a set of grid points not used in the fit. Before each fitting iteration, the total set of grid points is split into a fitting set, consisting of 85% of the total number of grid points, and a test set consisting of the remaining 15%. If the RMSE of the PES evaluated on the test set is greater than 1.3 times the RMSE evaluated on the fitting set, then the PES is considered to be unconverged. In such cases, additional grid points are added using the Monte Carlo type grid generation described above. In the final fitting stage, the test set is not used and the PES is fit to all the available data. The subroutines computing the fit, as well as a text file with fit parameters and other information on the fit, are available in SI.

II.C. Crystal Structure Predictions. The pairwise PES fit described above is used as the intermolecular potential in CSP for the solid form of **1**. The details of CSP workflow can be found in ref 44 and are outlined below. First, the UPACK program suite[16] is employed to generate more than 1000 random $Z' = 1$ structures for each of the most common 13 space groups found in organic molecular crystals, including: $C2$, $C2/c$, Cc, $P2_12_12_1$, $Pbca$, $Pbcn$, $P2_1$, $P2_1/c$, $Pca2_1$, Pc, $Pna2_1$, $P1$, and $P\bar{1}$. During the generation step, rigid molecules may be randomly placed with interatomic distances smaller than the repulsive wall of the SAPT(DFT) potential, so the unit cells are first optimized using a combination of the standard OPLS-AA Lennard-Jones parametrization[74] and the SAPT(DFT) fitted charges to avoid ending up in holes in the potential (the $A_{12}^{ab}/(r_{ab})^{12}$ terms greatly reduce the severity of holes, but do not remove them completely). The final 0 K optimization uses the full SAPT(DFT) rigid-monomer intermolecular potential and hole-fixing procedure and a modified version of UPACK.[77] These energies are used to rank the possible rigid molecule structures.

Second, the top 30 structures are expanded into supercells that are at least 20 Å on each side and equilibrated at the target pressure and temperature using MD simulations implemented in PINY_MD package,[78] where all molecules in the simulation box are considered to be fully flexible. Note that the PINY_MD cannot be applied to many types of rigid monomers, and with flexible monomers, one needs an

https://doi.org/10.1021/acs.cgd.1c01117
Cryst. Growth Des. 2022, 22, 1182−1195

intramonomer force field. While the SAPT(DFT) PES described above is capable of describing the intermolecular interaction energy in the case of small internal deformations, it does not include a functional form to evaluate the intramolecular energy component. To allow for molecular flexibility, we used a mix of generic and SAPT(DFT) force fields to describe the intramolecular terms of the potential energy. While generic force fields are generally not accurate enough to produce reliable CSPs, we argue that the intramonomer component is adequate. For CSPs of the type performed here, the monomers do not deform very significantly from their gas-phase geometries and so the intramonomer force field needs only to describe a small region of configuration space near a single reference point. In the case of the intermolecular component of the force field, no such simplification is possible, so the CSP benefits from a tailor-made intermolecular PES.

The generic/SAPT(DFT) force field parameters used for intramolecular terms of the potential energy in the MD simulations are assigned using the ACPYPE[79] wrapper for AnteChamber.[80] The resulting force field uses bonds, bends, and dihedrals from the generalized AMBER force field (GAFF).[80] For the atom−atom interactions that are separated by three bonds, known as 1−4 interactions, OPLS-AA parameters are assigned,[74] with the standard approach of scaling both Coulomb and Lennard-Jones terms by 0.5. Atomic charges for all atoms, and intramolecular interactions for atoms separated by more than three bonds, use the SAPT(DFT) parametrization (see Table S2).

The molecules in each supercell are equilibrated at 100 K using a 10 ps constant volume and temperature (NVT) MD simulation, followed by a 10 ps flexible-cell constant pressure and temperature (NPT) equilibration at 100 K and 1 bar for the full crystalline system. A subsequent 20 ps flexible-cell NPT MD simulation is used to determine the average potential energy to rank the crystalline structures. The flexible-cell NPT averaged supercells are collapsed into unit cells, and the final symmetry and space group assignments are determined by the PLATON package.[81]

For all MD runs, the cutoff distance for the intermolecular interactions is taken to be 10 Å with an integration time step of 0.2 fs. In evaluating the long-range Coulomb potential, we use an Ewald screening factor of $\alpha = 0.35$ Å^{-1}, and the smooth particle mesh Ewald (SPME) summation[82] is used with the interpolation order of 10. The massive Nosé−Hoover chain (NHC) thermostat[83−85] is used for all system variables and each NHC has a length of 4 with a characteristic time scale of $\tau = 20$ fs. The NHC integrator uses the Suzuki−Yoshida[86,87] factorization scheme up to the 6-th order, or $n_{sy} = 7$, and a multiple time step factor $n_c = 4$.[88] In integrating the Martyna−Tobias−Klein (MTK) equations in the flexible-cell NPT ensemble,[88−90] a characteristic time scale $\tau = 0.5$ ps is used for the barostat and $\tau = 0.1$ ps for its NHC thermostat.

III. RESULTS

III.A. PES Analysis. The interaction energies from the PES are plotted versus ab initio interaction energies for all the fitting grid points in Figure 3. The large error in the positive energy region visible in these plots is expected due to the low fitting weight given to these grid points. Because these points are located on the steep repulsive wall of the PES, even a large error corresponds only to a very small shift in the position of the wall. The uncertainties of the fit are listed in Table 1. The RMSE in the most important region with $E_{int} < 0$, which is weighted more heavily in the PES fitting, is 0.3 kcal/mol. These are expected RMSEs for molecules of this size and fitting functions defined by eqs 2−4. The error relative to the depths of PESs is 2.4%. With fits of this type, one should be careful to avoid overfitting. Table 1 shows that the ratio of the number of grid points to the number of fit parameters, N_{grid}/N_{FP}, is 8.7. This ratio is fairly low since autoPES converged in only three iterations (the number of grid points in the first iteration is 6 times the number of free parameters and 20% of this number is added in each subsequent iteration), but it is definitely in the safe region.

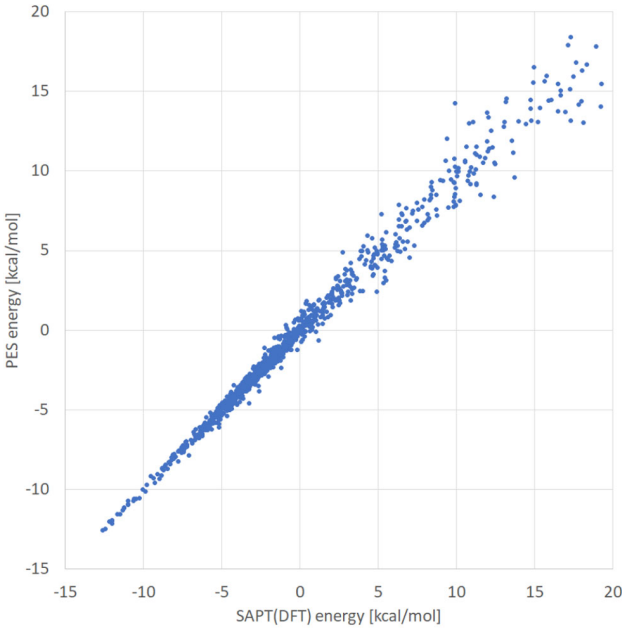


Figure 3. SAPT(DFT) interaction energy versus PES interaction energy for all fitting points of the **1**−**1** dimer.

Table 1. RMSEs of the Potential Evaluated on Subsets of the Close-Range Grid (All Energies in kcal/mol)[a]

N_{grid}	N_{FP}	N_{grid}/N_{FP}	N_{min}	RMSE $E_{int} < 0$	RMSE $E_{int} < 10$
1144	132	8.7	66	0.30 (733)	0.51 (1015)

[a]Numbers of grid points in the subsets are given in parentheses. The number of free parameters in the close-range fitting stage is denoted by N_{FP}, the number of detected local minima on the PES by N_{min}, and the total number of grid points by N_{grid}.

The number of minima found for **1**, 66, is larger than for other PESs that have been fitted so far with autoPES. The reason is the existence of approximate symmetries. For the benzene molecule, there exists a large number of symmetry-equivalent minima with exactly the same energy. When the benzene symmetry is broken by the addition of the nitro, hydroxyl, and amino groups, all these minima have different energies. The most energetically negative minima on the PES, plus the minima at dimer geometries that closely match those in the experimental crystal (discussed below), are listed in Table 2. All the minima, together with all parameters of the fit, are given in the SI. Table 2 shows that the low-lying minima are very closely spaced, reflecting that they all correspond to variations on the slipped-parallel dimer (shown in Figure S1). SAPT(DFT) energies computed at the geometries of the minima are also listed. All but two energy differences are within the 0.30 kcal/mol RMSE of the fit. The worst outlier is equal to 1 kcal/mol; the presence of such outliers are unavoidable with our global fitting strategy.

The uncertainties of the fit should be put in context of the uncertainties resulting from all other approximations of our approach. (1) The first such approximation is the use of SAPT(DFT). This method has been shown repeatedly to be similarly accurate to the coupled clusters method with single, double, and noniterated triple excitations, CCSD(T). As

https://doi.org/10.1021/acs.cgd.1c01117
Cryst. Growth Des. 2022, 22, 1182−1195

Table 2. Interaction Energies of the 10 Most Negative Minima on the PES and of the Minima Number 26, 32, 33, 54, and 58, Whose Geometries Most Closely Match the Dimers Found in the Experimental Crystal (Minimum 2 is Also in This Category)[a]

minimum	energy (kcal/mol)	minimum	energy (kcal/mol)
1	−12.57 (−12.66)	26	−8.36 (−8.18)
2	−12.05 (−12.13)	32	−7.74 (−7.53)
3	−11.68 (−11.50)	33	−7.68 (−7.26)
4	−11.66 (−10.66)	54	−6.24 (−6.08)
5	−11.60 (−11.68)	58	−5.38 (−5.16)
6	−11.17 (−11.18)		
7	−10.92 (−11.04)		
8	−10.79 (−10.72)		
9	−10.73 (−10.55)		
10	−10.67 (−10.43)		

[a]The SAPT(DFT) interaction energies are listed in parentheses.

shown in ref 91, the mean unsigned percentage error of SAPT(DFT) interaction energies computed using the aug-cc-pVTZ plus midbond basis set relative to the CCSD(T) interaction energies at the limit of a complete basis set (CBS) was 2.6% for a set of dimers with varying intermolecular separations, while the error of CCSD(T) in the aug-cc-pVTZ basis was 1.3%. Since CCSD(T)/CBS interaction energies are within about 1% of full configuration interaction energies at the CBS limit, SAPT(DFT) is certainly accurate enough method for crystal structure predictions. SAPT(DFT) is significantly more accurate than the so-called dispersion-corrected DFT approaches (DFT+D) (see ref 91).

(2) The next source of uncertainty results from the fact that we use the aug-cc-pVDZ plus midbond basis set in our calculations, rather than aug-cc-pVTZ plus midbond basis set used in ref 91. Thus, the uncertainties of our interaction energies are several times larger than the 2.6% found in ref 91, that is, they amount to a few percent. Note that with such uncertainties of ab initio interaction energies, the 2.4% error of the fit resulting from the assumed form of the fitting function should not affect our predictions in any significant way. If a higher accuracy of the fit is desirable, the simplest way would be to decrease the number of approximate symmetries (the monomer contains 21 atoms, but we have only 11 atoms with distinct parameters), although this would also require computing more grid points to avoid overfitting. One can further reduce the errors of autoPES fits to about 0.01 kcal/mol[92] by including additional off-atomic sites.

(3) Another uncertainty results from the limited number of grid points used to develop the PES. The PES of the **1** dimer is fairly complicated, as shown by the large number of minima, and may be undersampled with the 1144 grid points used. Undersampling can lead to regions in the fit, which may have uncertainties larger than the RMSEs listed in Table 1. For example, there is a fairly large discrepancy between the fit and the ab initio value of the interaction energy at minimum number 4 (see Table 2). A method for resolving this problem was proposed in ref 77. It consists of computing SAPT(DFT) interaction energies for additional grid points generated by near-neighbor dimers selected from a number of top-ranked polymorphs predicted with an initial version of the force field. If the present predictions were to be improved, this method offers the most reliable route.

(4) One more uncertainty of our force field results from neglect of pairwise-nonadditive many-body interactions. We approximate the total crystal lattice energy as a sum of two-body interactions, whereas the true energy includes contributions from pairwise-nonadditive effects, primarily in the form of polarization energy. Although the pairwise approximation is usually adequate for CSPs, the quality of predictions would certainly improve if many-body effects were included, particularly since **1** has a large dipole moment (5.9 D) and polarizability (152 au). In this case, the pairwise-nonadditive polarization effects could be large. AutoPES is capable of routinely developing polarizable PESs that recover many-body polarization effects, but the available CSP and MD programs cannot accommodate such PESs. In fact, there are only a couple of applications of polarizable force fields in CSP studies.[93−96]

(5) Additional uncertainties arise when the flexible-monomer force field is constructed. Unfortunately, these uncertainties are difficult to quantify because of the mixed nature of the force field used here. However, as will be discussed later, the effects of monomer flexibility on the predicted structures are modest and the impact of these uncertainties should be small.

Another way to improve our predictions would be to rerank the predicted structures using periodic DFT calculations with dispersion corrections (pDFT+D),[15,26−29] although the number of polymorphs that can be considered is limited due to the computational expense. It is unclear how the uncertainties of our fitted PES would compare to those from pDFT+D calculations. As we noted above, SAPT(DFT) calculations themselves are more accurate than available DFT+D methods, but due to the uncertainties of the PES fit, the same may not be true when comparing such fits with computed pDFT+D values. In addition to avoiding any fits, the pDFT+D approach includes pairwise-nonadditive polarization effects, and some dispersion functions include pairwise-nonadditive dispersion effects (although the latter are reproduced poorly[97,98]).

The total interaction energy and its components are plotted as functions of the COM-COM distance for the orientation corresponding to the global minimum in Figure 4. In addition to the fit and its components, we show the corresponding ab initio SAPT(DFT) quantities. The figure also shows the quality of our fits in the region of the van der Waals minimum. As one can see, the quality is excellent, as the ab initio total interaction energies lie almost ideally on the fit curve. For the components, the agreement for smaller R is much worse, but this is expected. The best agreement is for the dispersion plus induction components, again as expected since the damping factors in the PES are fitted to these ab initio values. The discrepancies are largest for the electrostatic energy which shows the expected failure of the asymptotic expansion at small R. Note that no short-range fitting, and therefore, no damping factor was used for electrostatics. The deviations of the exponential term are a consequence of the deviations for the electrostatics since the exponential term makes up for these discrepancies. In the asymptotic regime, where the exponential term is small, both the electrostatic and induction plus dispersion components are in good agreement with SAPT values. This can be seen in Figure 5, which displays interaction energies at the same dimer orientation as Figure 4 but at large separations.

https://doi.org/10.1021/acs.cgd.1c01117
Cryst. Growth Des. 2022, 22, 1182−1195

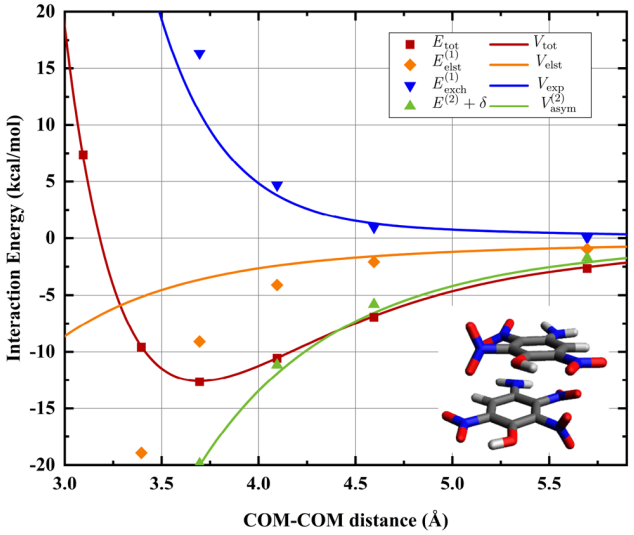


Figure 4. Radial dependence of the interaction energy and its components for the orientation corresponding to the global minimum of the **1**−**1** dimer. The geometry of the minimum is shown in the inset.

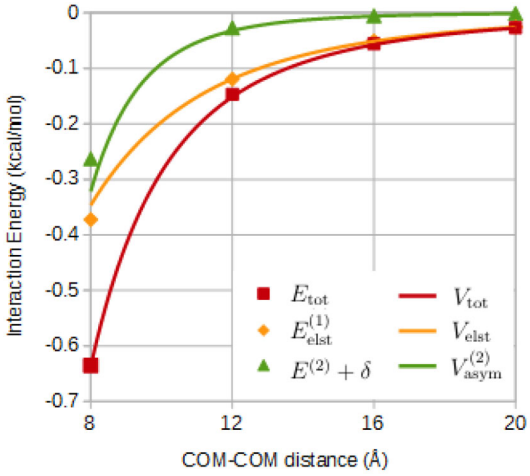


Figure 5. Radial dependence of the interaction energy and its components in the asymptotic regime for the orientation corresponding to the global minimum of the **1**−**1** dimer.

Since molecule **1** has a large polarizability because of π-electrons in the aromatic rings, one expects the dispersion interactions to be large. However, the dipole moment is also large and the quadrupole moments, shown in Table 3, are large as well. Thus, in general one expects substantial electrostatic and induction interactions as well. This results in a large interaction energy of −12.5 kcal/mol at the van der Waals minimum, compared to the corresponding benzene dimer interaction energy of only about −2.5 kcal/mol.[66] However, one cannot deduce the ratios of the attractive components from the knowledge of the monomer properties. Figure 4 shows that the dispersion plus induction energy (including their exchange counterparts) is about twice larger in magnitude than the electrostatic one. The quantity $E^{(2)} + \delta$ is dominated by the dispersion component (84%), hence the induction component, including $\delta E^{HF}_{int,resp}$, is only 16%. Thus, despite the molecule being strongly polar, the dispersion energy dominates. This is a general trend for larger molecules since the electrostatic interactions between different regions of monomers can be of either sign and therefore have a tendency to cancel.[99] For molecules of the size of **1**, the COM multipole expansion picture does not work well at distances near van der Waals minima and smaller, so only SAPT calculations can determine precise relations between interaction energy components. At large distances shown in Figure 5, the electrostatic interaction dominates because of its slower ($1/R^3$) decay than the dispersion and induction components ($1/R^6$).

Table 3. COM Quadrupole Moments of 1 Computed at the Same Level of Theory As Described in Section II.A, Given in Atomic Units[a]

xx	*yy*	*zz*	*xy*	*xz*	*yz*
−15.8	11.5	4.3	6.0	−4.9	−2.0

[a]The coordinate frame is given in SI.

The relative smallness of the electrostatic energy for the slipped parallel dimer configuration shown in Figure 4 may seem to clash with the well-known fact that for the benzene dimer, the slipped parallel configuration is stabilized relative to the sandwich configuration by the electrostatic interactions: negative for the former configuration and positive for the latter (note that these relations cannot be explained using only the leading quadrupole−quadrupole term[66]). However, the benzene-dimer electrostatic interactions are actually very small in magnitude relative to the dispersion interactions,[66] and the importance of the electrostatic energy is significantly larger for the dimer of **1**. This is because the quadrupole−quadrupole interactions, which decay as $1/R^5$, are the leading-order interactions for the benzene dimer, while for the **1** dimer the dipole−dipole and dipole−quadrupole interactions, which decay as $1/R^3$ and $1/R^4$, respectively, dominate. Since the dipole moments are in parallel planes and partly rotated with respect to each other at the minimum geometry, this dipole configuration corresponds to a smaller magnitude of the electrostatic energy than one could have expected based on its polarity.

At the distances of the van der Waals minimum and smaller, the charge-penetration effects strongly increase the magnitude of the electrostatic energy. This leads to an unexpected trend in the ratio of $E^{(1)}_{elst}/E^{(2)}$ for separations shown in the close-range figure. In the multipole approximation, this ratio should increase as $R^{6-3} = R^3$. Instead, this ratio slightly decreases with increasing R. The reason is, of course, charge-penetration effects. Here, these effects are unusually large, which has not been previously reported in the literature to our knowledge. This observation appears to be related to the slipped parallel configuration of the dimer. The reason for this phenomenon is as explained in the proceeding discussion. The electrostatic interactions consist of electron−electron, nuclear−nuclear, and electron−nuclear interactions. When the electron densities start to overlap significantly, the last interaction remains almost unaffected by this overlap since each nucleus of a monomer is still outside the region of substantial charge density of the interacting partner. Thus, only the positive electron−electron component is affected, and it becomes reduced by the overlap compared to what it would have been if the extent of the density were significantly reduced. This makes the electrostatic energy significantly more negative than the values predicted by the asymptotic expansion. Thus, this effect cannot be

https://doi.org/10.1021/acs.cgd.1c01117
Cryst. Growth Des. 2022, 22, 1182−1195

recovered using the usual multiplicative damping functions which range between 0 and 1 but requires the use of purely exponentially decaying terms resulting from the bipolar expansion of the interaction potential (see section V of ref 34). The effect is unusually large for dimers in a slipped parallel configuration, since charge penetration is obviously exceptionally strong.

III.B. Comparison of Monomer Minimum Energy Geometry with Those in the Experimental Crystal Structure. The experimental crystal structure has one molecule in the asymmetric unit cell,[52] but because of the the symmetry operations in the *Pbca* space group that invert the molecule, there are two monomer geometries present. These molecules differ primarily in the orientation of the nitro groups relative to the aromatic ring. A comparison of the r_e geometry with the two experimental geometries shows that the conformer used in fitting the PES is suitable to describe both forms (see Figure 6). To quantify the differences, the

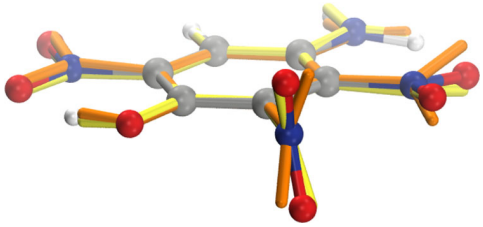


Figure 6. Geometries of **1** from DFT optimization (spheres), compared to the geometries of the two conformers from the experimental crystal structure[52] (orange and yellow). Two of the nitro groups in the experimental geometries deviate somewhat from the optimized gas-phase equilibrium positions.

experimental coordinates are aligned to the optimized geometry by minimizing the root-mean-square deviation (RMSD) of carbon atoms. After alignment, the $RMSD_C$ is 0.019 and 0.014 Å, while $RMSD_{O,N}$ is 0.151 and 0.427 Å for the two experimental geometries, respectively. The deviations are primarily due to slight rotations of two of the nitro groups indicated in Figure 1 (the third nitro group is in the plane of the ring due to the intramolecular hydrogen bond with the neighboring hydroxyl group). The discrepancies in the rotation angles of these two nitro groups between theory and experiment are to be expected, as the rotational potential is quite flat and therefore interactions in the crystal can easily lead to rotations relative to atomic positions in the isolated molecule.

III.C. Comparison of Dimer Minimum Geometries with Those of Dimers in the Experimental Crystal Structure. We compare the local minima of the PES to the dimer configurations found in the experimental crystal structure. An interesting question to answer is whether close-neighbor dimers in the crystal structures are in any way related to the minima on PESs. It is generally known that in many cases the crystal dimer configurations do not correspond to global minimum geometries of the isolated dimer, although such correspondence does exist for crystals with well-defined synthons such as hydrogen bonds. The question is, however, if a crystal's dimers are matched by isolated dimers corresponding to local minima. The other option is that dimer configurations in crystals are intermediate between the local minima of the isolated dimers.

We have identified the nearest-neighbor dimers from the *Pbca* experimental crystal[52] and overlapped them with all 66 dimers at the minimum geometries on our PES. The matches are determined by aligning the dimers based on minimizing the RMSD for the carbon and nitrogen atoms. One of the local minimum geometries, minimum 2, is the same as a closely spaced dimer from the crystal. The five other observed dimers from the crystal are related to local minima that are higher energy, with significantly less contact between the monomers in the molecular pair. We note that the first 22 minima are dimer configurations that belong to the group of slipped-parallel dimers similar to the global minimum. Therefore, it is reasonable that the next lowest energy match (minimum 26) to the experimental dimers would be so high in the energy ranking.

The first closely matching minimum is number 2 (interaction energy of −12.05 kcal/mol, 0.52 kcal/mol above the global minimum) with an $RMSD_{C,N}$ of only 0.17 Å. The other five minima, also reported in Table 2, are listed in energetic order (with $RMSD_{C,N}$ of their respective dimer in parentheses): number 26 (1.32 Å), number 32 (1.44 Å), number 33 (1.59 Å), number 54 (0.78 Å), and number 58 (1.18 Å). A section of the crystal cell highlighting the two most closely matched dimers is shown in Figure 7. The overlaps between these dimers and the corresponding minima on the PES (minimum 2 and 54) are shown in Figure 8 (the remaining matches are shown in SI).

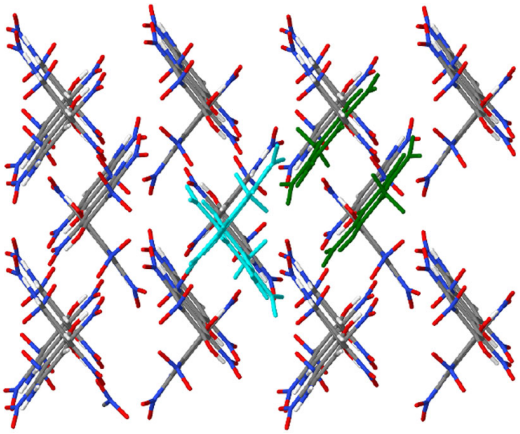


Figure 7. Experimental crystal structure of **1**. The dimer, which closely matches minimum 2, is shown in green, and the dimer, which closely matches minimum number 54, is shown in cyan.

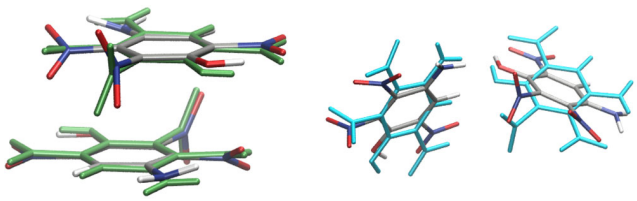


Figure 8. Overlap of the dimers at geometries corresponding to minima number 2 (left) and 54 (right) with dimers from the experimental crystal structure. The experimental dimers are shown in green and cyan (as in Figure 7), and the PES minima are shown colored by element. The dimers were aligned by minimizing the RMSD for all C and N atoms.

https://doi.org/10.1021/acs.cgd.1c01117
Cryst. Growth Des. 2022, 22, 1182−1195

Table 4. Cell Parameters, $RMSD_{20}$, and Ranks for Structures Matching the Observed Form

structure	inter potential	rigid?	rank	$RMSD_{20}$ (Å)	density (g/cm^3)	a (Å)	b (Å)	c (Å)
experimental[52]					1.839	12.5540	10.3306	13.5974
0 K	OPLS	rigid	53	0.301	1.796	12.9546	10.2355	13.6158
0 K	SAPT(DFT)	rigid	4	0.272	1.754	13.2466	10.2491	13.6153
100 K	SAPT(DFT)	flexible	5	0.350	1.684	13.4277	10.4279	13.7765

We can now answer the question asked in the beginning of this section: yes, dimers from crystal structures are similar to isolated dimers representing local minima. In particular, the experimental nearest-neighbor configuration is very close to the lowest-energy local minimum on the PES. Taking into account that 22 lowest-energy configurations are of slipped parallel type, the first nonparallel dimer matching experiment is also one of the lowest-energy dimers. Thus, the knowledge of the minima on the dimer PES of a notional compound can be helpful in crystal design. This may not be true for nonpolar monomers; for example, Aina et al.[96] did not find such correlations in their work on trinitrobenzene (TNB).

III.D. CSPs. Our CSP workflow ultimately generates structures that are equilibrated at experimentally relevant temperatures and pressures. However, we can learn more about predictions by examining the structures that match the reported form at each intermediate stage of the process. Such results are presented in Table 4. The rigid molecule packing step using the OPLS force field parameters yields a match to the experimental structure with an RMSD of 20 neighboring molecules ($RMSD_{20}$) of 0.301 Å, but the energy is of rank 53 and it is ∼12 kJ/mol above the lowest lattice energy. Optimizing the polymorphs with the SAPT(DFT) intermolecular potential, also with rigid molecules, significantly improves the energy ranking of the experimental form. The matching structure in this case is the fourth lowest lattice energy (see Figure 9) and has an $RMSD_{20}$ which is slightly improved over the OPLS version, despite having an even more elongated a-axis compared to experiment.

The density of the matching 0 K structure predicted by the SAPT(DFT) potential is lower than that of the reported finite temperature form. Since the rigid molecule geometry is quite close to the observed coordinates, this density error reflects the limits of the accuracy of our force field. As stated in section III.A, the main sources of uncertainties are (1) the use of SAPT(DFT), (2) the use of a fairly small aug-cc-pVDZ plus midbond basis set, (3) the limited number of grid points used in fitting, (4) neglect of pairwise-nonadditive many-body effects, and (5) use of an empirical force field for intramonomer interactions. Of these, reason (1) was argued to be accurate enough for CSPs, and reason (2) is not likely to affect density since the geometries of the dimer minima, the main factor dictating densities, converge in basis set much faster than the interaction energy. Reason (3) is unlikely to be problematic for densities since the RMSE of the fits, 0.3 kcal/mol, is fairly small compared to previously published fits of this type. Finally, reason (5) does not apply to the 0 K results. Thus, we are left with reason (4). Indeed, as already discussed, one can expect substantial pairwise-nonadditive polarization contributions to the lattice energy (two-body polarization effects are fully included in our force field). Very little is known about the importance of these effects in CSPs. Very recently, Aina et al.[96] included such effects in their CSPs for TNB, a system somewhat similar to **1**, but the polarization was only included in reranking the structures (i.e., not in the lattice minimization phase). Unfortunately, Aina et al. did not rigorously separate the two-body terms from pairwise-nonadditive ones. Since the reported overall polarization contribution for TNB is large, from 11% to 18% of the total lattice energy and several times larger than the induction energy of the equilibrium dimer, the implied relative magnitude of the nonadditive to additive polarization effects is enormous. In any case, even if the pairwise-nonadditive polarization contributions are much smaller for our system than for TNB (against expectations), they are expected to be attractive, so they should shift the minima to smaller separations and therefore increase densities after optimization. One should add, however, that the 0 K predicted unit cell differs from the reported form primarily by elongation along the a direction, making it difficult to discern why the pairwise-nonadditive effects should be directional.

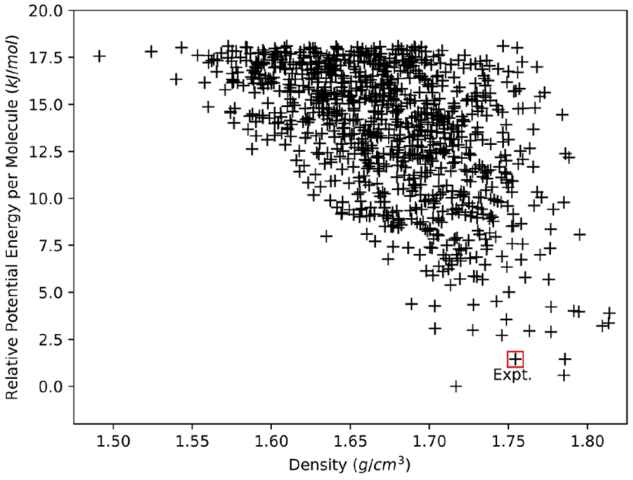


Figure 9. CSP results for **1** evaluated for predicted crystal structures optimized using rigid molecules and the tailor-made SAPT(DFT) intermolecular potential. The observed experimental structure is indicated with a red box.

In order to account for thermal expansion of the crystal unit cells, the predicted structures are equilibrated at 100 K and 1 bar through MD simulations with flexible molecules. Adding thermal fluctuations in the molecular structures increases the cell volume relative to the 0 K forms, so the predicted densities are slightly reduced (see Figure S4 for the change in density for each structure). The MD simulations also change the energy ranking (see Figure S5), resulting in the experimental structure as the fifth lowest energy form in the energy landscape, as shown in Figure 10.

After the MD step, the matching structure has an $RMSD_{20}$ of 0.350 Å when using an angle tolerance of 25° and the overlap of the predicted and experimental crystal structures is shown in Figure 11. The deviations from the reported structure arise from the slightly expanded a-axis for the unit cell and a slight mismatch in the orientation of the easily rotatable nitro groups.

https://doi.org/10.1021/acs.cgd.1c01117
Cryst. Growth Des. 2022, 22, 1182−1195

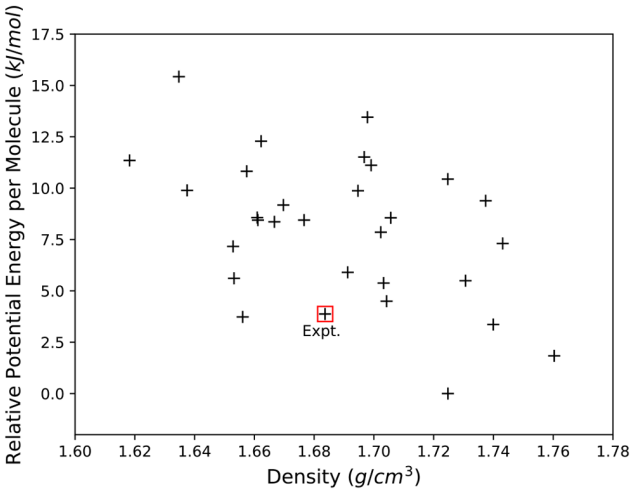


Figure 10. CSP results for **1** when equilibrated at 100 K and 1 bar. Average energy per molecule is shown relative to the energy of the most stable predicted polymorph. The experimentally observed *Pbca* structure has the fifth lowest average potential energy and is marked with a red box.

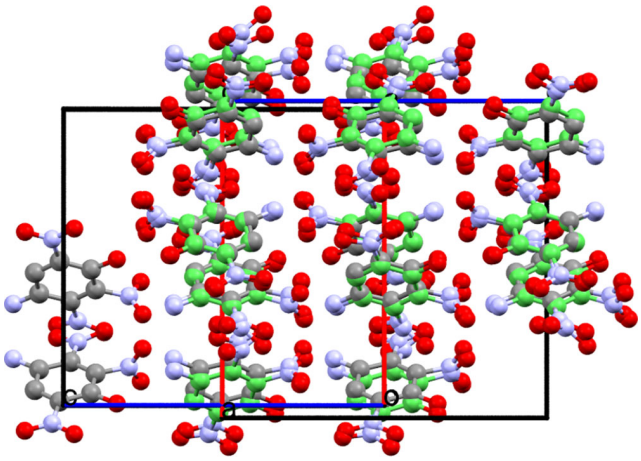


Figure 11. Overlap of the experimentally observed structure[52] with the predicted form equilibrated at 100 K and 1 bar using the tailor-made SAPT(DFT) force field for MD simulations. The experimental structure is shown colored by element, while the predicted structure is shown with green carbon atoms.

After alignment of the MD-averaged single molecules to the experimentally reported geometry, the average $RMSD_C$ is 0.037 Å, while the average $RMSD_{O,N}$ is 0.382 Å, compared to 0.289 Å for the rigid PES molecule. The resulting molecular structures, slightly different from the rigid-molecule CSP geometry, are clearly a reasonable match to the experimental structure. We find that the hybrid force field approach using the tailor-made SAPT(DFT) intermolecular terms with a generic/SAPT(DFT) intramolecular force field is adequate for **1**. Overall, we are encouraged that the computational time required to generate a molecule-specific intermolecular force field, as compared to the use of a generic intermolecular force field, is worth the effort to refine both the structural match and energy ranking of the observed form.

IV. CONCLUSIONS

In this Article, we demonstrated the application of a bottom-up protocol for crystal structure prediction that includes zero-temperature CSP and finite-temperature MD with a tailor-made molecule-specific force field, in the prediction of the landscape of an energetic compound, 4-amino-2,3,6-trinitrophenol. The custom force field includes an accurate intermolecular component built from dimers treated at the SAPT(DFT) level of electronic structure theory with a generic/SAPT(DFT) representation of the intramolecular potential. The landscape reveals that the experimentally observed form matches the predicted structure with an $RMSD_{20}$ of 0.35 Å, which lies well within the acceptable bounds prescribed by the CCDC blind structure prediction competition.[15]

A critical finding of our work is that the quality of the molecule-specific *intermolecular* potential appears to be determinitive of the accuracy of predicted structures. Using a generic/SAPT(DFT) representation of the *intramolecular* potential can be employed without significantly compromising that accuracy, even when some functional groups rotate easily within the crystalline environment. If greater accuracy of the final structure is required, several improvements of the present approach are possible depending on the computational effort permitted: increased size of the basis set for ab intio calculations, increased number of parameters in the SAPT fit, a more elaborate form of the fitting function, explicit inclusion of intramonomer degrees of freedom in the SAPT PESs, an iterative refinement of SAPT PESs by including specific dimers extracted from lowest-energy, highest-rank polymorphs within the fitting procedure, use of an intramonomer force field fitted to ab intio data, and by performing periodic boundary conditions dispersion-corrected DFT calculations for selected structures. However, reasonable CSP results have been obtained here for this challenging molecule without resorting to these more expensive approaches. The relative insensitivity of the landscape to the intramolecular potential, that is, small differences between the rigid- and flexible-monomer simulations, simplifies the CSP procedure, allowing for easy application of tools such as autoPES[51] for the creation of molecule-specific force fields. Whether the idea of using approximate intramolecular potentials in combination with tailor-made intermolecular potentials applies to crystals of larger, more flexible organic molecules merits further exploration and will be the subject of future work.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.cgd.1c01117.

Tables of molecular properties and force field components; fitted PES parameters and details on local dimer minima; figures of low energy dimers and those observed in the reported crystal structure; and figures comparing ranking of top 30 structures for 0 and 100 K conditions (PDF)

Supporting documents (ZIP)

■ AUTHOR INFORMATION

Corresponding Authors

Mark E. Tuckerman − *Department of Chemistry, New York University, New York, New York 10003, United States;*

https://doi.org/10.1021/acs.cgd.1c01117
Cryst. Growth Des. 2022, 22, 1182−1195

Courant Institute of Mathematical Sciences, New York University, New York, New York 10012, United States; NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai 200062, China; orcid.org/0000-0003-2194-9955; Email: marktuckerman@nyu.edu

Krzysztof Szalewicz − *Department of Physics and Astronomy, University of Delaware, Newark, Delaware 19716, United States;* orcid.org/0000-0002-2947-4694; Email: szalewic@udel.edu

Authors

Michael P. Metz − *Department of Physics and Astronomy, University of Delaware, Newark, Delaware 19716, United States; Applied Research Laboratories, University of Texas at Austin, Austin, Texas 78758, United States;* orcid.org/0000-0001-7083-7285

Muhammad Shahbaz − *Department of Physics and Astronomy, University of Delaware, Newark, Delaware 19716, United States; Department of Physics, University of Punjab, 05422 Lahore, Pakistan*

Hongxing Song − *Department of Chemistry, New York University, New York, New York 10003, United States*

Leslie Vogt-Maranto − *Department of Chemistry, New York University, New York, New York 10003, United States;* orcid.org/0000-0002-7006-4582

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.cgd.1c01117

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Rahul Nikhar for providing us with a modified UPACK program capable of using the SAPT(DFT) potential. This work was supported by the U.S. Army Research Laboratory and Army Research Office (Grant No. W911NF-19-1-0117) and by the NSF (Grant No. CHE-1900551).

■ REFERENCES

(1) Bauer, J.; Spanton, S.; Henry, R.; Quick, J.; Dziki, W.; Porter, W.; Morris, J. Ritonavir: An extraordinary example of conformational polymorphism. *Pharm. Res.* **2001**, *18*, 859.

(2) Chemburkar, S. R.; Bauer, J.; Deming, K.; Spiwek, H.; Patel, K.; Morris, J.; Henry, R.; Spanton, S.; Dziki, W.; Porter, W.; Quick, J.; Bauer, P.; Donaubauer, J.; Narayanan, B. A.; Soldani, M.; Riley, D.; McFarland, K. Dealing with the Impact of Ritonavir Polymorphs on the Late Stages of Bulk Drug Process Development. *Org. Proc. Res. Devel.* **2000**, *4*, 413−417.

(3) Rietveld, I. B.; Ceolin, R. Rotigotine: Unexpected Polymorphism with Predictable Overall Monotropic Behavior. *J. Pharm. Sci.* **2015**, *104*, 4117−4122.

(4) Mortazavi, M.; Hoja, J.; Aerts, L.; Quere, L.; van de Streek, J.; Neumann, M. A.; Tkatchenko, A. Computational polymorph screening reveals late-appearing and poorly-soluble form of rotigotine. *Comm. Chem.* **2019**, *2*, 70.

(5) Jurchescu, O. D.; Mourey, D. A.; Subramanian, S.; Parkin, S. R.; Vogel, B. M.; Anthony, J. E.; Jackson, T. N.; Gundlach, D. J. Effects of polymorphism on charge transport in organic semiconductors. *Phys. Rev. B* **2009**, *80*, 085201.

(6) Walley, S. M.; Field, J. E.; Greenaway, M. W. Crystal sensitivities of energetic materials. *Mater. Sci. Technol.* **2006**, *22*, 402−413.

(7) Price, S. L. Is zeroth order crystal structure prediction (CSP_0) coming to maturity? What should we aim for in an ideal crystal structure prediction code? *Faraday Discuss.* **2018**, *211*, 9−30.

(8) Price, S. L. Control and prediction of the organic solid state: A challenge to theory and experiment. *Proc. R. Soc. A* **2018**, *474*, 20180351.

(9) Nyman, J.; Reutzel-Edens, S. M. Crystal structure prediction is changing from basic science to applied technology. *Faraday Discuss.* **2018**, *211*, 459−476.

(10) Lommerse, J. P. M.; Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Mooij, W. T. M.; Price, S. L.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. A test of crystal structure prediction of small organic molecules. *Acta Cryst. B* **2000**, *56*, 697−714.

(11) Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Dzyabchenko, A.; Erk, P.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Lommerse, J. P. M.; Mooij, W. T. M.; Price, S. L.; Scheraga, H.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. Crystal structure prediction of small organic molecules: A second blind test. *Acta Cryst. B* **2002**, *58*, 647−661.

(12) Day, G. M.; Motherwell, W. D. S.; Ammon, H. L.; Boerrigter, S. X. M.; Della Valle, R. G.; Venuti, E.; Dzyabchenko, A.; Dunitz, J. D.; Schweizer, B.; van Eijck, B. P.; Erk, P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W. M.; Leusen, F. J. J.; Liang, C.; Pantelides, C. C.; Karamertzanis, P. G.; Price, S. L.; Lewis, T. C.; Nowell, H.; Torrisi, A.; Scheraga, H. A.; Arnautova, Y. A.; Schmidt, M. U.; Verwer, P. A third blind test of crystal structure prediction. *Acta Cryst. B* **2005**, *61*, 511−527.

(13) 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.

(14) 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 structure prediction of complex organic compounds − A report on the fifth blind test. *Acta Cryst. B* **2011**, *67*, 535−551.

(15) 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.; Cosburg, K.; Cuppen, H. M.; Curtis, F.; Day, G. M.; DiStasio, 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.; Hogmann, 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. Y.; Lohano, S.; Leusen, F. J. J.; Lund, A. M.; Lv, J.; Ma, Y. M.; Marom, N.; Masunov, A. E.; McCabe, P.; McMahon, D. P.; Meekes, H.; Metz, M. P.; Misquitta, A. J.; Mohamed, S.; Monserrat, B.; Nees, 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.; Vazques-Mayagoitia, A.; Vogt, L.; Want, Y. C.; Watson, R. E.; de Wijs, G. A.; Yang, J.; Ahu, Z.; Groom, C. R. Report on the sixth blind test of organic crystal structure prediction methods. *Acta Crystallogr.* **2016**, *B72*, 439−459.

(16) van Eijck, B. P.; Kroon, J. UPACK Program Package for Crystal Structure Prediction: Force Fields and Crystal Structure Generation for Small Carbohydrate Molecules. *J. Comput. Chem.* **1999**, *20*, 799−812.
(17) van Eijck, B. P.; Kroon, J. Structure predictions allowing more than one molecule in the asymmetric unit. *Acta Crystallogr. B Struct. Sci.* **2000**, *56*, 535−542.
(18) Kim, S.; Orendt, A. M.; Ferraro, M. B.; Facelli, J. C. Crystal structure prediction of flexible molecules using parallel genetic algorithms with a standard force field. *J. Comput. Chem.* **2009**, *30*, 1973−1985.
(19) Zhu, Q.; Oganov, A. R.; Glass, C. W.; Stokes, H. T. Constrained evolutionary algorithm for structure prediction of molecular crystals: methodology and applications. *Acta Crystallogr. B Struct. Sci.* **2012**, *68*, 215−226.
(20) Curtis, F.; Li, X.; Rose, T.; Vázquez-Mayagoitia, Á.; Bhattacharya, S.; Ghiringhelli, L. M.; Marom, N. GAtor: A First-Principles Genetic Algorithm for Molecular Crystal Structure Prediction. *J. Chem. Theory Comput.* **2018**, *14*, 2246−2264.
(21) Karfunkel, H. R.; Gdanitz, R. J. Ab Initio prediction of possible crystal structures for general organic molecules. *J. Comput. Chem.* **1992**, *13*, 1171−1183.
(22) Pillardy, J.; Arnautova, Y. A.; Czaplewski, C.; Gibson, K. D.; Scheraga, H. A. Conformation-family Monte Carlo: A new method for crystal structure prediction. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 12351−12356.
(23) Nyman, J.; Pundyke, O. S.; Day, G. M. Accurate force fields and methods for modelling organic molecular crystals at finite temperatures. *Phys. Chem. Chem. Phys.* **2016**, *18*, 15828−15837.
(24) Metz, M. P.; Piszczatowski, K.; Szalewicz, K. Automatic Generation of Intermolecular Potential Energy Surfaces. *J. Chem. Theory Comput.* **2016**, *12*, 5895−5919.
(25) Beran, G. J. O. Modeling Polymorphic Molecular Crystals with Electronic Structure Theory. *Chem. Rev.* **2016**, *116*, 5567−5613.
(26) Brandenburg, J. G.; Grimme, S. Organic crystal polymorphism: a benchmark for dispersion-corrected mean-field electronic structure methods. *Cryst. Acta B* **2016**, *52*, 502−513.
(27) 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.
(28) 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.
(29) Hoja, J.; Ko, H. Y.; Neumann, M. A.; Car, R.; DiStasio, R. A.; Tkatchenko, A. Reliable and practical computational description of molecular crystal polymorphs. *Sci. Adv.* **2019**, *5*, eaau3338.
(30) Podeszwa, R.; Bukowski, R.; Rice, B. M.; Szalewicz, K. Potential energy surface for cyclotrimethylene trinitramine dimer from symmetry-adapted perturbation theory. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5561−5569.
(31) Podeszwa, R.; Rice, B. M.; Szalewicz, K. On predicting structure of molecular crystals from first principles. *Phys. Rev. Lett.* **2008**, *101*, 115503.
(32) Podeszwa, R.; Rice, B. M.; Szalewicz, K. Crystal structure prediction for cyclotrimethylene trinitramine (RDX) from first principles. *Phys. Chem. Chem. Phys.* **2009**, *11*, 5512−5518.
(33) Misquitta, A. J.; Welch, G. W. A.; Stone, A. J.; Price, S. L. A First Principles Prediction of the Crystal Structure of $C_6Br_2ClFH_2$. *Chem. Phys. Lett.* **2008**, *456*, 105−109.
(34) Jeziorski, B.; Moszyński, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994**, *94*, 1887−1930.
(35) Williams, H. L.; Chabalowski, C. F.; Using, K. Using Kohn-Sham Orbitals in Symmetry-Adapted Perturbation Theory to Investigate Intermolecular Interactions. *J. Phys. Chem. A* **2001**, *105*, 646−659.
(36) Misquitta, A. J.; Szalewicz, K. Intermolecular forces from asymptotically corrected density functional description of monomers. *Chem. Phys. Lett.* **2002**, *357*, 301−306.
(37) Heßelmann, A.; Jansen, G. First-order intermolecular interaction energies from Kohn-Sham orbitals. *Chem. Phys. Lett.* **2002**, *357*, 464−470.
(38) Heßelmann, A.; Jansen, G. Intermolecular induction and exchange-induction energies from coupled-perturbed Kohn-Sham density functional theory. *Chem. Phys. Lett.* **2002**, *362*, 319−325.
(39) Misquitta, A. J.; Jeziorski, B.; Szalewicz, K. Dispersion energy from density-functional theory description of monomers. *Phys. Rev. Lett.* **2003**, *91*, 033201.
(40) Heßelmann, A.; Jansen, G. Intermolecular dispersion energies from time-dependent density functional theory. *Chem. Phys. Lett.* **2003**, *367*, 778−784.
(41) Misquitta, A. J.; Szalewicz, K. Symmetry-Adapted Perturbation Theory Calculations of Intermolecular Forces Employing Density Functional Description of Monomers. *J. Chem. Phys.* **2005**, *122*, 214109.
(42) Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. Intermolecular potentials based on symmetry-adapted perturbation theory with dispersion energies from time-dependent density-functional calculations. *J. Chem. Phys.* **2005**, *123*, 214103.
(43) Hesselmann, A.; Jansen, G.; Schütz, M. Density-functional theory-symmetry-adapted intermolecular perturbation theory with density fitting: A new efficient method to study intermolecular interaction energies. *J. Chem. Phys.* **2005**, *122*, 014103.
(44) Schneider, E.; Vogt, L.; Tuckerman, M. E. Exploring Polymorphism of Benzene and Naphthalene with Free Energy Based Enhanced Molecular Dynamics. *Acta Cryst. B* **2016**, *B72*, 542−550.
(45) Yu, T.-Q.; Tuckerman, M. E. Temperature-Accelerated Method for Exploring Polymorphism in Molecular Crystals Based on Free Energy. *Phys. Rev. Lett.* **2011**, *107*, 015701.
(46) Yu, T. Q.; Chen, P. Y.; Chen, M.; Samanta, A.; Vanden-Eijnden, E.; Tuckerman, M. E. Order-parameter-aided temperature-accelerated sampling for the exploration of crystal polymorphism and solid-liquid phase transitions. *J. Chem. Phys.* **2014**, *140*, 214109.
(47) Shtukenberg, A.; Zhu, Q.; Carter, D. J.; Vogt, L.; Hoja, J.; Schneider, E.; Song, H.; Pokroy, B.; Polishchuk, I.; Tkatchenko, A.; Oganov, A. R.; Rohl, A. L.; Tuckerman, M. E.; Kahr, B. Powder diffraction and crystal structure prediction identify four new coumarin polymorphs. *Chem. Sci.* **2017**, *8*, 4926−4940.
(48) Song, H.; Vogt-Maranto, L.; Wiscons, R.; Matzger, A. J.; Tuckerman, M. E. Generating Cocrystal Polymorphs with Information Entropy Driven by Molecular Dynamics-Based Enhanced Sampling. *J. Phys. Chem. Lett.* **2020**, *11*, 9751.
(49) Fellah, N.; Shtukenberg, A. G.; Chan, E. J.; Vogt-Maranto, L.; Xu, W.; Li, C.; Tuckerman, M. E.; Kahr, B.; Ward, M. D. Disorderly conduct of benzamide IV: Crystallogrphic and computational analysis of high entropy polymorphs of small molecules. *Cryst. Growth & Design* **2020**, *20*, 2670.
(50) Francia, N. F.; Price, L. S.; Nyman, J.; Price, S. L.; Salvalaglio, M. Systematic finite-temperature reduction of crystal energy landscapes. *Cryst. Growth & Design* **2020**, *20*, 6847.
(51) Metz, M. P.; Piszczatowski, K.; Szalewicz, K. *autoPES: Automatic Intermolecular Potential Energy Surface Generation Software*, 2016. https://www.physics.udel.edu/~szalewic/SAPT/index.html.
(52) Klapötke, T. M.; Preimesser, A.; Stierstorfer, J. Synthesis and Energetic Properties of 4-Diazo-2,6-dinitrophenol and 6-Diazo-3-hydroxy-2,4-dinitrophenol. *Eur. J. Org. Chem.* **2015**, *2015*, 4311−4315.
(53) Neese, F. The ORCA program system. *Wiley Interdis. Rev.: Comp. Mol. Sci.* **2012**, *2*, 73−78.
(54) Neese, F. An improvement of the resolution of the identity approximation for the formation of the Coulomb matrix. *J. Comput. Chem.* **2003**, *24*, 1740−1747.
(55) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.

https://doi.org/10.1021/acs.cgd.1c01117
Cryst. Growth Des. 2022, 22, 1182−1195

(56) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158−6170.
(57) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796−6806.
(58) Bukowski, R.; Podeszwa, R.; Szalewicz, K. Efficient calculation of coupled Kohn-Sham dynamic susceptibility functions and dispersion energies with density fitting. *Chem. Phys. Lett.* **2005**, *414*, 111−116.
(59) Podeszwa, R.; Bukowski, R.; Szalewicz, K. Density-fitting method in symmetry-adapted perturbation theory based on Kohn-Sham description of monomers. *J. Chem. Theory Comput.* **2006**, *2*, 400−412.
(60) Bukowski, R.; Cencek, W.; Jankowski, P.; Jeziorska, M.; Jeziorski, B.; Metz, M. P.; Moszynski, R.; Patkowski, K.; Podeszwa, R.; Rob, F.; Rybak, S.; Szalewicz, K.; Williams, H. L.; Wheatley, R. J.; Wormer, P. E. S.; Zuchowski, P. *SAPT2016: An Ab Initio Program for Symmetry-Adapted Perturbation Theory*, 2006. https://www.physics.udel.edu/~szalewic/SAPT/index.html.
(61) Grüning, M.; Gritsenko, O. V.; Van Gisbergen, S. J. A.; Baerends, E. J. Shape corrections to exchange-correlation potentials by gradient-regulated seamless connection of model potentials for inner and outer region. *J. Chem. Phys.* **2001**, *114*, 652−660.
(62) Cencek, W.; Szalewicz, K. On Asymptotic Behavior of Density Functional Theory. *J. Chem. Phys.* **2013**, *139*, No. 024104; Erratum: "On asymptotic behavior of density functional theory" [J. Chem. Phys. 139, 024104 (2013)]. *J. Chem. Phys.* **2014**, No. 149902.
(63) Garcia, J.; Podeszwa, R.; Szalewicz, K. SAPT codes for calculations of intermolecular interaction energies. *J. Chem. Phys.* **2020**, *152*, 184109.
(64) Weigend, F.; Köhn, A.; Hättig, C. Effcient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2002**, *116*, 3175−3183.
(65) Williams, H. L.; Mas, E. M.; Szalewicz, K.; Jeziorski, B. On the effectiveness of monomer-, dimer-, and bond-centered basis functions in calculations of intermolecular interaction energies. *J. Chem. Phys.* **1995**, *103*, 7374−7391.
(66) Podeszwa, R.; Bukowski, R.; Szalewicz, K. Potential energy surface for the benzene dimer and perturbational analysis of π-π interactions. *J. Phys. Chem. A* **2006**, *110*, 10345−10354.
(67) Podeszwa, R.; Cencek, W.; Szalewicz, K. Effcient calculations of dispersion energies for nanoscale systems from coupled density response functions. *J. Chem. Theor. Comp.* **2012**, *8*, 1963−1969.
(68) Jeziorska, M.; Jeziorski, B.; Cizek, J. Direct Calculation of the Hartree-Fock Interaction Energy via Exchange Perturbation Expansion - the He-He Interaction. *Int. J. Quantum Chem.* **1987**, *32*, 149−164.
(69) Patkowski, K.; Szalewicz, K.; Jeziorski, B. Third-order interactions in symmetry-adapted perturbation theory. *J. Chem. Phys.* **2006**, *125*, 154107.
(70) Stone, A. J. *The Theory of Intermolecular Forces*, 2nd ed.; Clarendon Press: Oxford, 2013.
(71) Rob, F.; Szalewicz, K. Distributed molecular polarisabilities and asymptotic intermolecular interaction energies. *Mol. Phys.* **2013**, *111*, 1430−1455.
(72) Misquitta, A. J.; Stone, A. J.; Fazeli, F. Distributed multipoles from a robust basis-space implementation of the iterated stockholder atoms procedure. *J. Chem. Theory Comput.* **2014**, *10*, 5405−5418.
(73) Metz, M. P.; Szalewicz, K. Automatic generation of flexible-monomer intermolecular potential energy surfaces. *J. Chem. Theory Comput.* **2020**, *16*, 2317−2339.
(74) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(75) Tang, K. T.; Toennies, J. P. An improved simple model for the van der Waals potential based on universal damping functions for the dispersion coefficients. *J. Chem. Phys.* **1984**, *80*, 3726−3741.
(76) Breneman, C. M.; Wiberg, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. *J. Comput. Chem.* **1990**, *11*, 361−373.
(77) Nikhar, R.; Szalewicz, K. Reliable crystal structure predictions from first principles. Submitted, 2021.
(78) Tuckerman, M. E.; Yarne, D. A.; Samuelson, S. O.; Hughes, A. L.; Martyna, G. J. Exploiting multiple levels of parallelism in Molecular Dynamics based calculations via modern techniques and software paradigms on distributed memory computers. *Comput. Phys. Commun.* **2000**, *128*, 333−376.
(79) Sousa da Silva, A. W.; Vranken, W. F. ACPYPE−AnteChamber PYthon Parser interfacE. *BMC Research Notes* **2012**, *5*, 367.
(80) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General AMBER Force Field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(81) Spek, A. L. Single-crystal structure validation with the program PLATON. *J. Appl. Crystallogr.* **2003**, *36*, 7−13.
(82) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(83) Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **1984**, *81*, 511−519.
(84) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695−1697.
(85) Martyna, G. J.; Klein, M. L.; Tuckerman, M. E. Nosé-Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **1992**, *97*, 2635−2643.
(86) Suzuki, M. General theory of fractal path integrals with applications to many-body theories and statistical physics. *J. Math. Phys.* **1991**, *32*, 400−407.
(87) Yoshida, H. Construction of higher order symplectic integrators. *Phys. Lett. A* **1990**, *150*, 262−268.
(88) Tuckerman, M.; Alejandre, J.; López-Rendón, R.; Jochim, A.; Martyna, G. J. A Liouville-operator derived measure-preserving integrator for molecular dynamics simulations in the isothermal-isobaric ensemble. *J. Phys. A: Math. Gen.* **2006**, *39*, 5629−5651.
(89) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **1994**, *101*, 4177−4189.
(90) Yu, T.-Q.; Alejandre, J.; López-Rendón, R.; Martyna, G. J.; Tuckerman, M. E. Measure-preserving integrators for molecular dynamics in the isothermal-isobaric ensemble derived from the Liouville operator. *Chem. Phys.* **2010**, *370*, 294−305.
(91) Taylor, D. C.; Angyan, J. G.; Galli, G.; Zhang, C.; Gygi, F.; Hirao, K.; Song, J. W.; Rahul, K.; von Lilienfeld, O. A.; Podeszwa, R.; Bulik, I. W.; Henderson, T. M.; Scuseria, G. E.; Toulouse, J.; Truhlar, D. G.; Peverati, R.; Szalewicz, K. Blind test of density-functional-based methods on intermolecular interaction energies. *J. Chem. Phys.* **2016**, *145*, 124105.
(92) Metz, M. P.; Szalewicz, K.; Sarka, J.; Tóbiás, R.; Császár, A. G.; Mátyus, E. Molecular dimers of methane clathrates: *ab initio* potential energy surfaces and variational (ro)vibrational states. *Phys. Chem. Chem. Phys.* **2019**, *21*, 13504−13525.
(93) Price, S. L.; Leslie, M.; Welch, G. W. A.; Habgood, M.; Price, L. S.; Karamertzanis, P. G.; Day, G. M. Modelling organic crystal structures using distributed multipole and polarizability-based model intermolecular potentials. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8478−8490.
(94) Aina, A. A.; Misquitta, A. J.; Price, S. L. From dimers to the solid-state: Distributed intermolecular force-fields for pyridine. *J. Chem. Phys.* **2017**, *147*, 161722.
(95) Jeong, K.-j.; McDaniel, J. G.; Yethiraj, A. A Transferable Polarizable Force Field for Urea Crystals and Aqueous Solutions. *J. Phys. Chem. B* **2020**, *124*, 7475−7483.
(96) Aina, A. A.; Misquitta, A. J.; Price, S. L. A non-empirical intermolecular force-field for trinitrobenzene and its application in crystal structure prediction. *J. Chem. Phys.* **2021**, *154*, 094123.
(97) Hapka, M.; Rajchel, L.; Modrzejewski, M.; Schaeffer, R.; Chalasinski, G.; Szczesniak, M. M. The nature of three-body

https://doi.org/10.1021/acs.cgd.1c01117
Cryst. Growth Des. 2022, 22, 1182−1195

interactions in DFT: Exchange and polarization effects. *J. Chem. Phys.* **2017**, *147*, 084106.
(98) Jankiewicz, W.; Podeszwa, R.; Witek, H. A. Dispersion-Corrected DFT Struggles with Predicting Three-Body Interaction Energie. *J. Chem. Theory Comput.* **2018**, *14*, 5079−5089.
(99) Bukowski, R.; Szalewicz, K.; Chabalowski, C. F. Ab initio interaction potentials for simulations of dimethylnitramine solutions in supercritical carbon dioxide with cosolvents. *J. Phys. Chem. A* **1999**, *103*, 7322−7340.

https://doi.org/10.1021/acs.cgd.1c01117
Cryst. Growth Des. 2022, 22, 1182−1195