ACCOUNTS of chemical research

Article

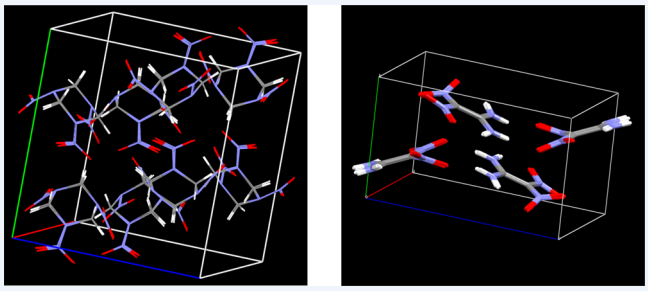
pubs.acs.org/accounts

Determination of Structure and Properties of Molecular Crystals from First Principles

Krzysztof Szalewicz*

Department of Physics and Astronomy, University of Delaware, Newark, Delaware 19716, United States

CONSPECTUS: Until recently, it had been impossible to predict structures of molecular crystals just from the knowledge of the chemical formula for the constituent molecule(s). A solution of this problem has been achieved using intermolecular force fields computed from first principles. These fields were developed by calculating interaction energies of molecular dimers and trimers using an *ab initio* method called symmetry-adapted perturbation theory (SAPT) based on density-functional theory (DFT) description of monomers [SAPT(DFT)]. For clusters containing up to a dozen or so atoms, interaction energies computed using SAPT(DFT) are comparable in accuracy to the results of the best wave function-based methods, whereas the former approach can be applied to systems an order of magnitude larger than the latter. In fact, for monomers with a couple dozen atoms, SAPT(DFT) is about equally time-consuming as the supermolecular DFT approach. To develop a force field, SAPT(DFT) calculations are performed for a large number of dimer and possibly also trimer configurations (grid points in intermolecular coordinates), and the interaction energies are then fitted by analytic functions. The resulting force fields can be used to determine crystal structures and properties by applying them in molecular packing, lattice energy minimization, and molecular dynamics calculations. In this way, some of the first successful determinations of crystal structures were achieved from first principles, with crystal densities and lattice parameters agreeing with experimental values to within about 1%. Crystal properties obtained using similar procedures but empirical force fields fitted to crystal data have typical errors of several percent due to low sensitivity of empirical fits to interactions beyond those of the nearest neighbors. The first-principles approach has additional advantages over the empirical approach for notional crystals and cocrystals since empirical force fields can only be extrapolated to such cases.



As an alternative to applying SAPT(DFT) in crystal structure calculations, one can use supermolecular DFT interaction energies combined with scaled dispersion energies computed from simple atom−atom functions, that is, use the so-called DFT+D approach. Whereas the standard DFT methods fail for intermolecular interactions, DFT+D performs reasonably well since the dispersion correction is used not only to provide the missing dispersion contribution but also to fix other deficiencies of DFT. The latter cancellation of errors is unphysical and can be avoided by applying the so-called dispersionless density functional, dlDF. In this case, the dispersion energies are added without any scaling. The dlDF+D method is also one of the best performing DFT+D methods. The SAPT(DFT)-based approach has been applied so far only to crystals with rigid monomers. It can be extended to partly flexible monomers, that is, to monomers with only a few internal coordinates allowed to vary. However, the costs will increase relative to rigid monomer cases since the number of grid points increases exponentially with the number of dimensions. One way around this problem is to construct force fields with approximate couplings between inter- and intramonomer degrees of freedom. Another way is to calculate interaction energies (and possibly forces) "on the fly", i.e., in each step of lattice energy minimization procedure. Such an approach would be prohibitively expensive if it replaced analytic force fields at all stages of the crystal predictions procedure, but it can be used to optimize a few dozen candidate structures determined by other methods.

1. INTRODUCTION

Predictions and microscopic-level understanding of properties of molecular crystals is critical for many branches of science and technology. The key element for achievement of these goals is the knowledge of the forces between molecules (monomers), which are the building blocks of such materials. These forces can be either obtained by fitting measured properties of crystals, resulting in the so-called *empirical* force fields or obtained by solving equations of quantum mechanics, resulting in *ab initio* force fields. With the latter fields, one potentially may predict crystal structures and properties fully from first principles, that is, without utilizing any experimental data. Thus, this method is of particular interest in crystal engineering, which aims at creating new types of crystals.

For a long time, theory was unable to predict crystal structures from structural chemical formulas of monomers. This was considered to be a serious failure, and in 1988 Maddox[1] wrote "One of the continuing scandals in the physical sciences is that it remains in general impossible to predict the structure of even the

Special Issue: DFT Elucidation of Materials Properties

Received: July 31, 2014
Published: October 29, 2014

simplest crystalline solids from a knowledge of their chemical composition". Fourteen years later, Desiraju[2] concluded that "A method for predicting crystal structures from just molecular formulae has eluded scientists for more than 50 years". As will be discussed later on, these authors did not appreciate how sensitive crystal structures are to details of the intermolecular force fields, that is, to the intermolecular interactions between monomers building a crystal. In the 1980s, reliable *ab initio* calculations of intermolecular interaction energies near van der Waals minima were virtually impossible: in a 1988 review, Buckingham, Fowler, and Hutson[3] referred to this part of an intermolecular potential using the phrase "theoretically intractable 'intermediate region'". Thus, the only option then was to use empirical potentials, which, as we know now, are insufficiently accurate for predicting crystal structures. The problems of such methods were exposed by a series of "blind tests" organized by the Cambridge Crystallographic Data Centre (CCDC): for the first time in 1999,[4] followed by four more tests, the last one published in 2011.[5]

Reliable *ab initio* calculations of intermolecular potentials for few-atom dimers became possible in the 1990s, for example, see refs 6−8, and currently such potentials can be computed for dimers containing close to 100 atoms. This progress has been made possible partly by increases of computer power, by about 8 orders of magnitude between the 1980s and the present time, but to a larger extent by the development of new computational methods. The latter progress has been more critical for the field of intermolecular interactions since interactions energies include large contributions from electron correlation effects, that is, from the solutions of Schrödingers's equation that go beyond the independent-particle Hartree−Fock (HF) approximation. Older methods for calculating correlation effects, such as the configuration interaction (CI) method, scale as $N!$ with system size N (expressed as the number of electrons). Such methods were applicable in the 1980s only to diatomic few-electron dimers.[9] The 8 orders of magnitude increase in computer power would have extended applicability of such methods to systems with only a dozen or so electrons. More significant progress was achieved by the development of the many-body perturbation methods based on the Møller−Plesset partition of the Hamiltonian (called MP methods), of coupled-cluster (CC) approaches,[10] and of symmetry-adapted perturbation theory (SAPT).[11−13] These methods scale as powers of N, and methods with N^7 scaling have to be used to obtain reliable interaction energies. This growth is much slower than $N!$, and therefore calculations of interaction energies for systems with hundreds of electrons are now possible. An even better scaling, N^4, is achieved by density-functional theory (DFT) in the Kohn−Sham (KS) implementation.[14] Although standard versions of DFT fail badly for intermolecular interactions,[15,16] DFT can be used within SAPT formalism[17,18] to obtain accurate interaction energies at low costs. This method, dubbed SAPT(DFT), can be used to compute interaction energies for dimers with hundreds of atoms,[19] that is, for systems for which the costs of N^7 calculations would be prohibitive.

Ab initio-derived potentials have two main advantages over empirical ones: (a) uniform and potentially arbitrarily high accuracy in all regions, whereas empirical force fields fitted to crystal data are not sensitive to configurations beyond nearest-neighbor arrangements; (b) applicability to notional (i.e., hypothetical) crystals for which no experimental data exist, whereas empirical potentials have to be extrapolated to such cases.

Results computed by *ab initio* electronic structure methods are independent of whether experimental values of these properties are known. Empirical approaches fit parameters to reproduce some experimental results, so these methods model rather than predict these properties. Furthermore, whereas the literal meaning of "*ab initio*" is just "from first principles", we will use the former term for electronic structure calculations and the latter one for work involving also crystal packing, minimization, and molecular dynamics (MD) simulations. The DFT approaches will be included in the *ab initio* category, although some variants of DFT are fitted to experimental data and are therefore semiempirical.

This Account will discuss first-principle methods of crystal structure determination. It will concentrate on construction of force fields, while other steps of the process, i.e., methods of crystal packing, lattice optimizations, and MD simulations will be described only briefly. Discussions of the latter steps can be found in the CCDC blind test papers and, for example, in refs 20−23. In section 2, the basic concepts of the SAPT approach to investigations of intermolecular forces will be described. Section 3 will present various DFT-based methods adapted to calculations of interaction energies, including SAPT(DFT). In section 4, the methodology of crystal-structure determination will be outlined and results of some investigations presented. Section 5 will discuss perspectives of extending first-principle methods to crystals with flexible monomers. Section 6 contains a summary.

2. SYMMETRY-ADAPTED PERTURBATION THEORY

The simplest way to obtain interaction energies is by subtracting from the total energies of the whole cluster the total energies of the constituent monomers. Any electronic structure method can be used, but reliable calculations of interaction energies require a high-level description of electron correlation effects. This scheme is called the *supermolecular* approach to intermolecular interactions. In contrast, SAPT is a perturbation theory starting from isolated monomers and computing interaction energies directly, without computations of the total cluster energies.[13,24,25] The total Hamiltonian, H, for a dimer consisting of monomers A and B is decomposed into the unperturbed Hamiltonian, H_0, which is the sum of the monomer Hamiltonians, $H_0 = H_A + H_B$, and the perturbation operator V containing all Coulomb interactions of particles of monomer A with those of B. One first solves Schrödingers's equations for monomers and then uses perturbation theory to calculate contributions to interaction energies as an expansion in powers of V

$$E_{\mathrm{int}} = E^{(1)} + E^{(2)} + \ldots$$

The simplest, Rayleigh−Schrödinger (RS) perturbation expansion can be applied, but it gives unphysical interaction energies at small intermonomer separations. This is because the RS wave functions are not fully antisymmetric with respect to electron exchanges (in other words, the Pauli exclusion principle is violated). The symmetry adaptation utilized in SAPT fixes this problem by properly projecting the wave functions to make them fully antisymmetric.

SAPT provides the standard way of understanding intermolecular interactions in terms of the fundamental physical components: the electrostatic, induction (polarization plus charge transfer), dispersion, and exchange (valence repulsion) contributions. These contributions are not only computed by SAPT in a quantitative way but seamlessly connected to the same

dx.doi.org/10.1021/ar500275m | Acc. Chem. Res. 2014, 47, 3266−3274

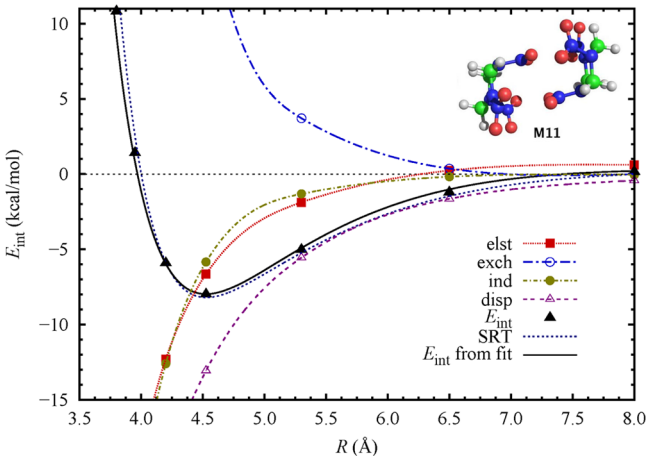


Figure 1. SAPT components for the RDX dimer.[26] The curve denoted as SRT is the empirical potential of Sorescu et al.[27] Adapted with permission from ref 26. Copyright 2007 by Royal Society of Chemistry.

components in the multipole expansion of interaction energies at large intermonomer separations R. Figure 1 shows the decomposition of the interaction energy for the cyclotrimethylene trinitramine (RDX) dimer.[26] The configuration chosen is that of the nearest neighbors in the RDX crystal.

Exact solutions of Schrödinger's equations for the monomers, the starting points for SAPT expansions, are possible only for small atoms such as helium[28] or lithium.[29] For larger systems, one has to further decompose the monomers' Hamiltonians into Fock operators, F_X, and MP potentials, W_X. Thus, SAPT becomes a triple perturbation theory with respect to the operators V, W_A, and W_B.[11−13] If all the corrections available in the SAPT codes are included, accuracy of SAPT interaction energies is comparable to that of the CC method with single, double, and noniterative triple excitations, CCSD(T). A comparison of these two methods is shown in Figure 2 on the example of the helium dimer, the only dimer for which virtually exact interaction energies are known. As one can see, SAPT performs on this system overall slightly better than CCSD(T), but differences are very small. Several computer packages can be used to perform SAPT calculations.[31−34]

3. DENSITY FUNCTIONAL CALCULATIONS OF INTERACTION ENERGIES

Since DFT scales much better than wave function-based methods for computing electron correlation effects, it should be in principle the method of choice for calculations of interaction energies for large systems. Unfortunately, as already mentioned, the standard DFT methods are unsuitable for this task. In particular, these methods fail badly for systems with large dispersion interactions, such as the argon dimer. The performance of various DFT methods for this dimer is shown in Figure 3.

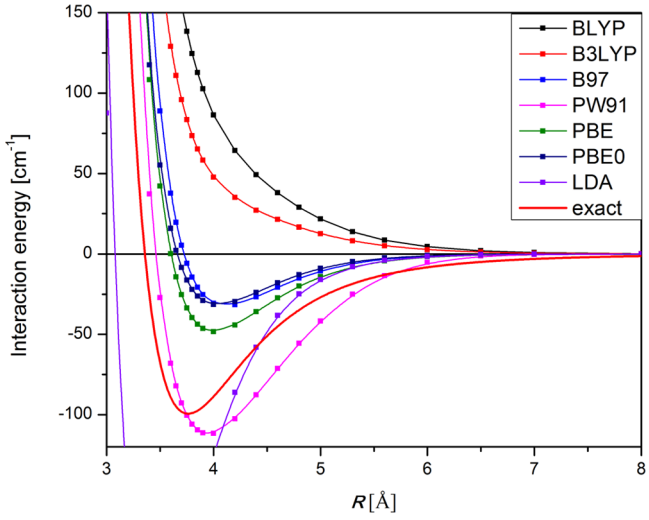


Figure 3. Comparison of performance of various DFT methods on the argon dimer. The benchmark curve[35] is accurate to about its line width. The acronyms define various DFT methods, for example, see, ref 36.

Clearly, none of the methods comes even close to the exact results. This failure of DFT is generally attributed to its inability to recover dispersion energies, which result from interactions between electrons separated by distances of several angstroms, whereas standard DFT models such interactions only for separations of the order of 1 Å. However, this is not the only problem of DFT methods, as clearly seen in Figure 3. If it were,

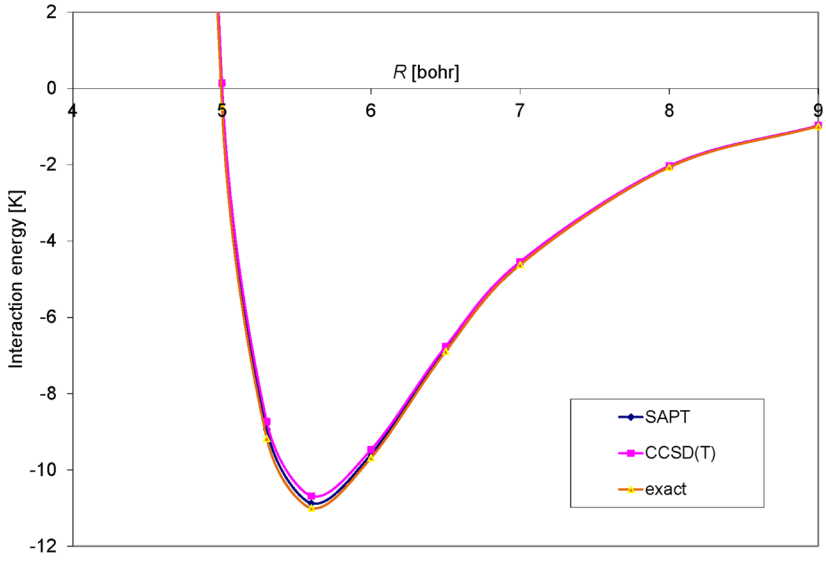


Figure 2. Comparison of performance of SAPT and CCSD(T) methods for the helium dimer.[30] All the results were extrapolated to the complete basis set limit. The "exact" potential energy curve is accurate to better than its width.

dx.doi.org/10.1021/ar500275m | Acc. Chem. Res. 2014, 47, 3266−3274

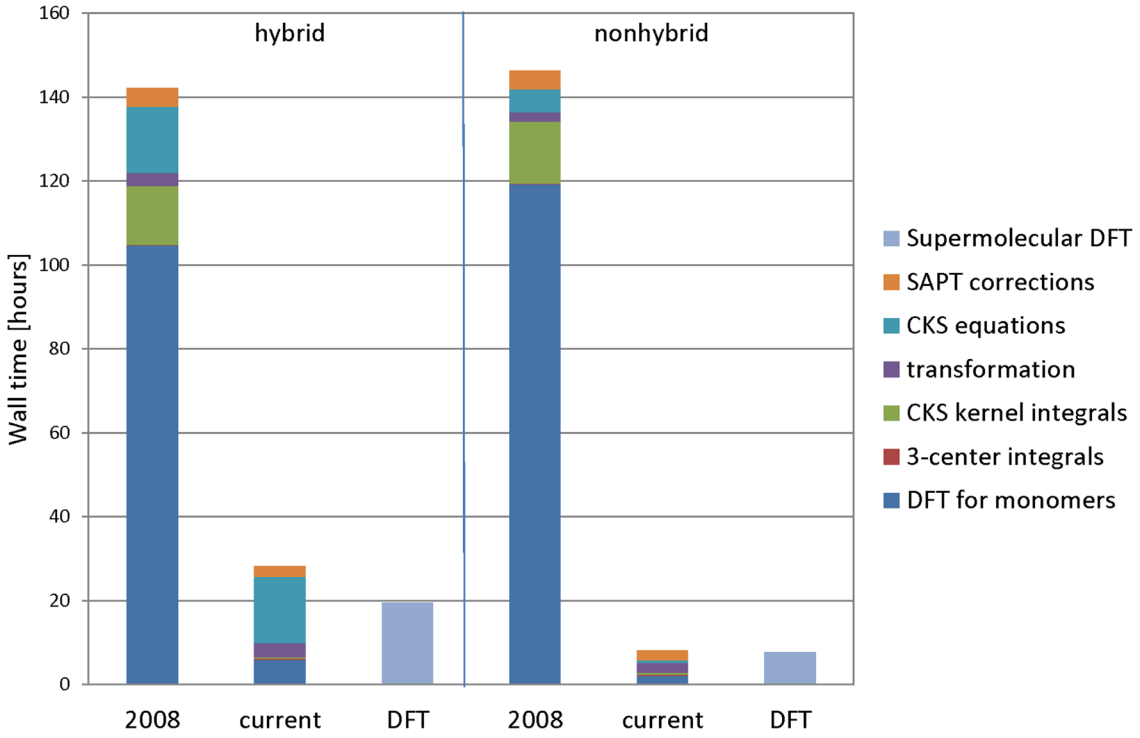


Figure 4. Density-fitted SAPT(DFT) wall times (on a single core of a 2.2 GHz Opteron processor) for calculations on the 42-atom RDX dimer. The aug-cc-pVDZ plus 3s3p2d2f midbond basis set in the monomer-centered "plus" (MC^+) form[47] and the pure and hybrid versions of the PBE functional[48] were used. The set of SAPT(DFT) corrections was identical as in ref 26. Bars marked as 2008 were computed with the SAPT2008 codes and Dalton 2.0 front-end, whereas the current calculations used Orca 3.0.1 front-end. The unexpectedly shorter PBE0 time compared with PBE is due to smaller number of iterations in the former case. The "DFT" bars are supermolecular counterpoise-corrected DFT calculations.

all the methods would had given wrong but similar interaction energies, which is not the case. The other problem is that the electron densities given by DFT methods decay too slowly for large distances from nuclei.[37,38]

3.1. SAPT(DFT)

There are three main methods of making DFT usable for intermolecular interactions. The earliest solution was SAPT(DFT).[17,18,39−44] This method overcomes both the dispersion and asymptotic problems of DFT by utilizing it only for calculations on monomers while computing all components of interaction energies from SAPT expressions. The simplest version of such approach[39] is to utilize the isomorphism of HF and KS one-electron solutions and replace HF orbitals and orbital energies by KS ones in the SAPT formulas that neglect operators W_A and W_B. Such a simplistic approach works poorly, but if the KS densities are asymptotically corrected[40,41] and the dispersion energies are computed at the coupled KS (CKS) instead of uncoupled KS level,[42,43] the accuracy of SAPT(DFT) interaction energies is comparable to that of the regular SAPT. The advantage of using SAPT(DFT) rather than the regular SAPT is that the former method scales as N^5 (N^4 if some small terms are neglected), whereas the regular SAPT at the level of intramonomer electron correlation high enough to be competitive to the CCSD(T) method scales as N^7 [same scaling as that of CCSD(T)]. The reason for the speedup is that the most time-consuming corrections in regular SAPT, accounting for the intramonomer electron correlation effects, are not needed in SAPT(DFT).

The scalings of costs mentioned above assume the use of density-fitting techniques.[18,45,46] The efficiency of density-fitted SAPT(DFT) significantly improved in recent years. Figure 4 shows the timings of the original codes[45,46] compared with the current version in the SAPT2012 package.[31] The main speedup was actually achieved by interfacing SAPT(DFT) with a more efficient front-end package, but other speedups are also significant. With current timings and with availability of a few hundred computer cores, one can perform all runs needed to develop an RDX dimer potential similar to that of ref 26 in about a day. This means in practice that the range of applicability of SAPT(DFT) to molecular crystals has been extended from monomers containing about 20 atoms to monomers with about 40 atoms. At the present, the majority of effort needed to develop the potential is actually the human effort at the stage of fitting the computed interaction energies. However, an automatic program performing such fitting is now under development in our group. Figure 4 shows that if a nonhybrid functional is used, SAPT(DFT) calculations are equally costly as supermolecular DFT calculations. At the same time, the latter calculations give nonsensical interaction energies (−1.25 and −1.54 kcal/mol using PBE and PBE0, respectively), whereas the former ones give −8.00 kcal/mol using PBE0, the minimum value in Figure 1, and −7.94 kcal/mol using PBE.

3.2. Functionals Optimized on Intermolecular Interactions

One possible approach to improve the performance of DFT methods on interaction energies is to optimize parameters in the functionals to better reproduce these energies.[49,50] However, one should note that the form of the functionals remains the same so that these functionals still cannot recover dispersion energies for the reasons mentioned earlier. This fact is visible at intermolecular separations larger than about two times the minimum distance where DFT methods of this type perform poorly.[51] In particular, for systems dominated by dispersion

dx.doi.org/10.1021/ar500275m | Acc. Chem. Res. 2014, 47, 3266−3274

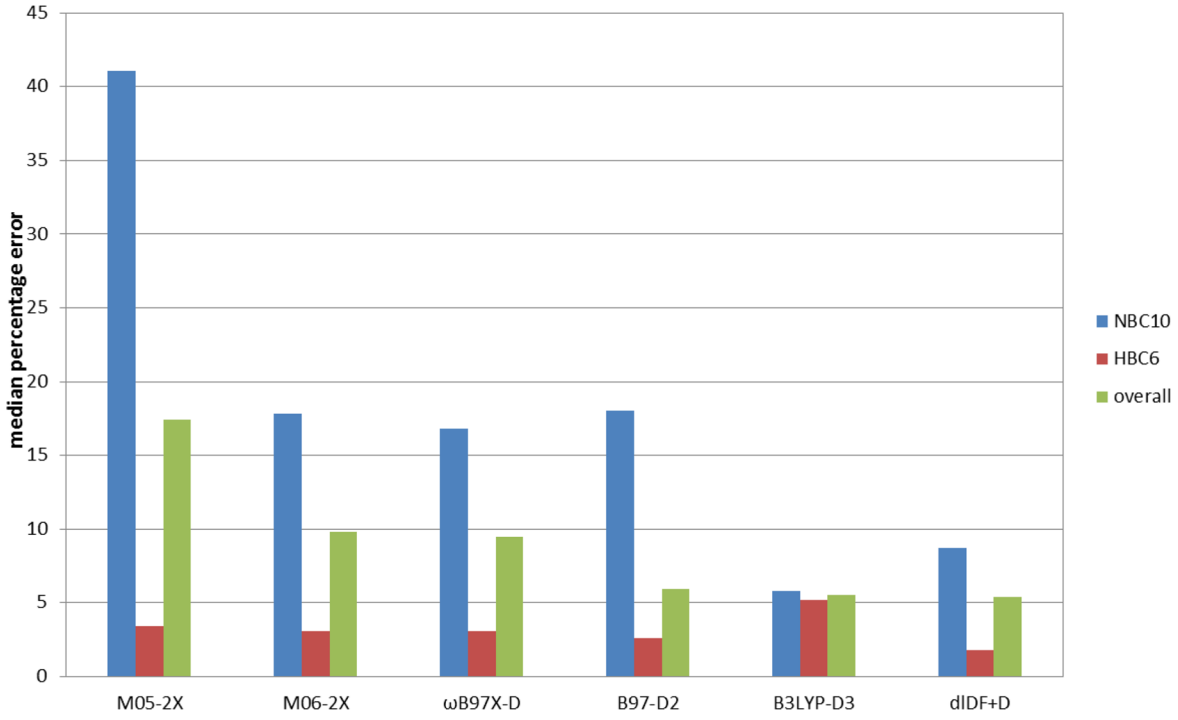


Figure 5. Comparison of performance of various DFT methods on *R*-dependent benchmarks from refs 58 and 59. The dlDF+D results are from ref 57. See ref 57 for acronyms defining various DFT methods.

forces, the interaction energies do not decay as the sixth inverse power of *R*.

3.3. DFT Supplemented by Dispersion Energies

Another solution to the problems of DFT in calculations of intermolecular forces is to supplement DFT interaction energies with *scaled* dispersion energies, most often computed from simple atom−atom functions analogous to those used in empirical biomolecular force fields, leading to the so-called DFT+D approaches.[52−56] Separation-dependent scaling is needed since, as seen in Figure 3, the DFT interaction energy for each density functional needs to be corrected by a different amount. Thus, the dispersion energy term is used not only to add this missing component, which is physically reasonable, but also to remove the asymptotic density errors, which cannot be justified on physical grounds.

To eliminate the necessity of using scaled dispersion energies to correct errors of DFT methods unrelated to dispersion interactions, Pernal et al.[51] proposed a version of DFT+D including a novel "dispersionless" density functional, denoted as the dlDF+D method. This functional was optimized to reproduce dispersionless interaction energies, that is, interaction energies from which the dispersion and exchange-dispersion energies (computed using SAPT) were subtracted. Since standard density functionals are capable of describing all interactions contained in dispersionless interaction energies, this approach has a solid physical foundation. The dispersion energies can then be added to the dlDF interaction energies without any scaling. The dlDF+D method is not only better justified than other DFT+D variants, it also is one of the most accurate ones.[57] This performance is shown in Figure 5. The bars show errors on dimers dominated by dispersion forces (NBC10), on hydrogen-bonded dimers (HBC6), and the overall error. As one can see, the dlDF+D approach gives the smallest overall error, although this error is only slightly smaller than in the case of B3LYP+D3 method. dlDF+D performs better on hydrogen-bonded systems, but B3LYP+D3 performs better on the NBC10 set. The reason for the latter is that 9 out of the 10 NBC10 dimers contain monomers with aromatic rings and no such systems were used in training of the dlDF method.

One challenge in the DFT+D approach is to replace the simple atom−atom asymptotic dispersion functions by *ab initio* dispersion energies, which are sufficiently inexpensive for applications to very large systems. For monomers with more than a few atoms, one has to use distributed asymptotic expansions, that is, expansions containing inverse powers of interatomic distances rather than inverse powers of the distance between centers of mass of monomers. A new distribution method was recently developed in refs 60 and 61. The charge-overlap effects can then be included using the localized-overlap dispersion energy algorithm,[62] which speeds up calculations of dispersion energies for large molecules by 2 orders of magnitude compared with the best standard SAPT(DFT) algorithm.

4. DETERMINATION OF CRYSTAL STRUCTURES

The process of crystal structure determination consists of several steps, listed in Figure 6. Assuming that the monomers are given only by their structural chemical formula, one has to start from an optimization of monomer (gas-phase) geometry. Of course, in a crystalline environment, the monomers will always deform, but for a large class of crystals, these deformations are so small that their effect on the interaction potential can be neglected.

The next step is to choose a set of grid points and perform SAPT(DFT) calculations for all these points. The grid points have usually been selected based on physical intuition, but a new method under development in our group uses a guided Monte Carlo algorithm. The computed interaction energies are then fitted by an analytic function. For rigid monomers, such functions usually are sums of isotropic pair potentials between the atoms of monomer A and those of monomer B and include Coulomb interactions of partial charges, exponential terms modeling

dx.doi.org/10.1021/ar500275m | *Acc. Chem. Res.* 2014, 47, 3266−3274

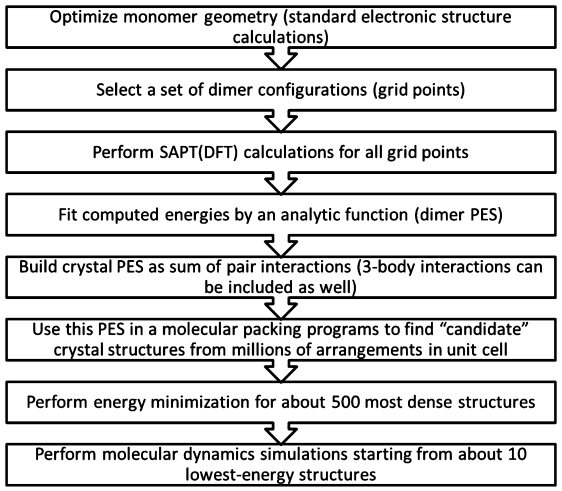


Figure 6. Steps in first-principles determination of crystal structures.

exchange repulsion, and terms proportional to the sixth and higher inverse powers of interatomic distances. Often an explicit polarization term containing induced dipoles is also included, which does not improve appreciably two-body fits but provides an approximate account of pairwise nonadditive many-body effects. The fit can be made more accurate by using off-atomic sites in addition to the atomic ones. The quality of the potential fit is validated by checking its accuracy relative to *ab initio* values on a number of points not used in fitting. If the accuracy of reproducing these points is too low, additional *ab initio* points are computed and the fit is repeated. Typically, for systems of the size of the RDX dimer, one can obtain fits that are accurate to a fraction of a kilocalorie per mole for negative interaction energies. This accuracy is commensurate with the accuracy of SAPT(DFT) calculations and significantly higher than can be achieved with any other method for systems of this size. The total force field of the crystal is constructed as a sum of pair interactions. One can also compute *ab initio* three-body effects beyond the polarization model.[63−65]

The force field can now be applied in a search of the minimum-energy crystal structure. The first step of such search, called molecular packing, can even be performed without using any potentials but just by representing monomers as a set of hard spheres with radii equal to the van der Waals radii of atoms. All the commonly observed space symmetry groups and a large number of monomer arrangements in the crystal cell of each group are considered. The crystal is first built with large intermonomer separations and then contracted until some spheres touch. The density of the crystal is then computed. This procedure is performed for several hundred thousand initial configurations and a few hundred configurations with highest densities are selected for further processing. The next step is a minimization of the lattice energy (usually within the space symmetry group constrains) using the SAPT(DFT) force field starting from each of the structures selected in the previous step. The final step of the procedure is MD calculations for a dozen or so lowest-energy crystals. Almost all published determinations of crystal structures were based on internal energies, whereas ordering of crystal structures is determined by more difficult to compute free energies. Several methods are available for calculations of the latter energies, for example, the metadynamics approach[66] or temperature-accelerated sampling,[67] the latter one recently applied to crystal polymorphism problems.[68]

The first crystal-structure determinations using the method described above were published for the RDX crystal in 2008.[65] While these were not blind predictions, first-principle approaches give the same result irrespective whether they are blind or not, as mentioned in the Introduction. A similar SAPT(DFT)-based method was developed by Misquitta et al.[69] and successfully used in blind predictions of the crystal structure of $C_6Br_2ClFH_2$. Around the same time, Neumann et al.[70] successfully applied a method based on empirical force fields and DFT+D. Later, SAPT(DFT)-based calculations were performed by Taylor et al.[71] for the 1,1-diamino-2,2-dinitroethylene (FOX-7) crystal. This work was oriented at thermal properties of the crystal, so no search of the minimum structure was performed, and MD simulations were started from the experimental one. Figure 7 shows the computed and experimental crystal structures superimposed on each other. Clearly, the two types of structures are very close to each other. The corresponding numerical results are presented in Table 1. For comparison, Sorescu et al.[27] computed the RDX crystal density with an error of 3.8% despite applying an empirical potential fitted to RDX crystal data (and using a very similar crystal structure determination procedure to that used in ref 65). As is shown in Figure 1, the empirical potential agrees very well with the SAPT(DFT) potential for the nearest neighbor configuration shown in this figure. Clearly, this is the configuration most sensitive to experimental data. Other regions of the potential

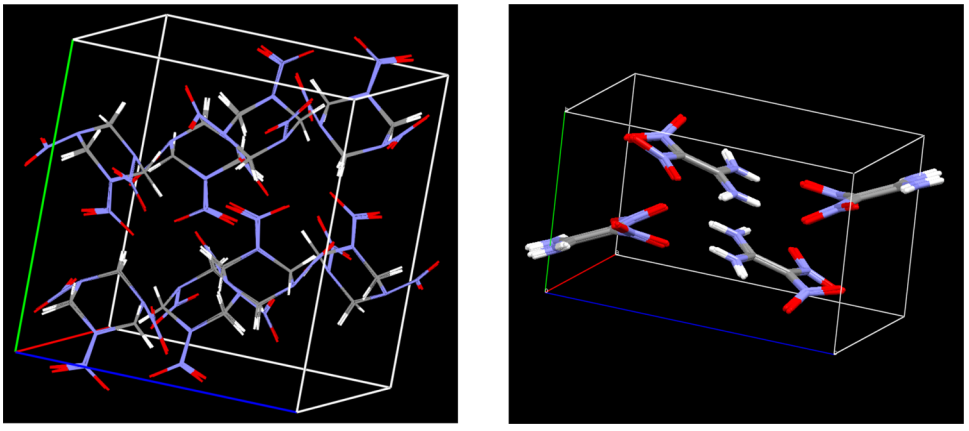


Figure 7. Comparisons of the computed and experimental structures for the RDX (left panel) and FOX-7 crystals. Reproduced with permission from ref 65, copyright 2008 by American Physical Society, and from ref 71, copyright 2011 by Royal Society of Chemistry.

dx.doi.org/10.1021/ar500275m | *Acc. Chem. Res.* 2014, 47, 3266−3274

Table 1. Errors of Computed Properties of the RDX and FOX-7 Crystals Relative to Experimental Values

	RDX	FOX
density	0.8%	0.4%
cell dimensions	<0.4%	<0.8%
monomer positions	0.2%	2%
monomer angles	2.5°	4°

surface are less well represented by the empirical potential,[26] which leads to a larger density error. Taylor et al.[71] calculated several properties of the FOX-7 crystal, in particular the dependence of volume on pressure, agreeing much better with experiment[72] than with the calculations[73] based on empirical potentials fitted to measurements on FOX-7 and on similar crystals.

5. CRYSTALS WITH FLEXIBLE MONOMERS

First-principle methods have been applied so far only to crystals with approximately rigid monomers, where rigid-monomer force fields can be expected to be adequate. The SAPT-based approach can be extended in its present form to crystals with partly flexible monomers, that is, to monomers with a few of the internal coordinates expected to vary significantly without producing large changes in internal energies compared to lattice energies. However, the calculations will be more time-consuming than for rigid monomers since the potential energy surface will depend on additional coordinates and the number of grid points scales as k^D, with D dimensions and k points per dimension. For any rigid monomer, $D = 6$ taking, and since taking k between 3 and 4 seems to be sufficient, this generates on the order of 10^3 grid points. With three soft internal degrees of freedom per monomer, D becomes equal to 12 and one has to use on the order of 10^5 grid points, which is still doable but costly.

The difficulties of constructing force fields with flexible monomers from first principles are often underappreciated since most empirical force fields, in particular biomolecular ones, include all degrees of freedom. However, the simplicity of the latter fields is achieved by assuming intermonomer parameters to be independent of intramonomer degrees of freedom. Flexible-monomer potentials of this type can be constructed by adding to a first-principles rigid-monomer intermolecular potential an intramonomer potential, both potentials constructed independently. If the former potential is represented by a sum of atom−atom interactions, it will change with monomers' deformations. Such atom-following representation was found to reproduce only 30% of infrared shifts of intramonomer vibrations in a small dimer (unpublished result from ref 74). However, such shifts are purely due to the intermolecular−intramolecular couplings, so that the performance of this method on other properties may be more satisfactory.

The atom-following approximation becomes progressively worse as monomers deform more and more from the reference configuration. Therefore, one strategy that can be tried is to develop several rigid-monomer potentials, with monomers' geometries corresponding to a number of local low-lying minima, and then make each such potential fully flexible by using the atom-following approach. Of course, for monomers with soft degrees of freedom, the monomer geometry in a crystal usually (but not always) will be appreciably distorted compared with any gas-phase local minima geometries, but hopefully the former geometry will be close enough to one of the latter to make the atom-following approach sufficiently reliable.

A radically different approach is to abandon analytic force fields and calculate potential energies (and possibly forces) "on the fly", that is, in each step of a lattice energy minimization procedure. Such minimizations can usually be performed with just thousands of steps almost independently of including the internal degrees of freedom. Thus, finding the minimum in this way requires a much smaller number of energy calculations than in the development of a flexible-monomer potential. Still, such approach would be prohibitively expensive with any *ab initio* method including DFT, if used at all stages of the crystal structure determination procedure. However, it can be used to optimize a few dozen candidate structures obtained by other methods, as implemented by Neumann and coauthors.[70,75]

6. SUMMARY

The results presented herein show that the problems of crystal structure predictions discussed in refs 1 and 2 have been resolved and that the main reason for these problems was an insufficient accuracy of empirical force fields. Currently, several *ab initio* wave functions-based methods can be used to compute sufficiently accurate interaction energies but these methods are too time-consuming for applications to crystals with monomers containing more than a dozen or so atoms. In contrast, the SAPT(DFT) approach, which combines wave function and DFT methodologies, determines interaction energies equally accurately and can be applied to crystals with monomers containing about 40 atoms. The first-principles approach has additional advantages over the empirical approach for notional crystals and cocrystals since empirical force fields can only be extrapolated to such cases. Reliable interaction energies can also be computed using DFT+D methods, in particular dlDF+D.

AUTHOR INFORMATION

Corresponding Author

*E-mail: szalewic@udel.edu.

Notes

The authors declare no competing financial interest.

Biography

Krzysztof Szalewicz obtained his M.S. (1973), Ph.D. (1977), and D.Sc. (1984) degrees from the University of Warsaw. His initial faculty appointments were at the University of Warsaw and at the University of Florida. Since 1988, he has been a faculty member at the University Delaware. He is a member of the International Academy of Quantum Molecular Science and of the Franklin Award Committee, as well as a fellow of the American Physical Society. His research interests are in theory of atoms, molecules, and solids.

ACKNOWLEDGMENTS

This work was supported by the U.S. Army Research Laboratory and the Army Research Office under Grant W911NF-13-1-0387 and by the National Science Foundation Grant CHE-1152899.

REFERENCES

(1) Maddox, J. Crystals from First Principles. *Nature* **1988**, *335*, 201.
(2) Desiraju, G. R. Cryptic Crystallography. *Nat. Mater.* **2002**, *1*, 77−79.
(3) Buckingham, A. D.; Fowler, P. W.; Hutson, J. M. Theoretical Studies of van der Waals Molecules and Intermolecular Forces. *Chem. Rev.* **1988**, *88*, 963−988.
(4) 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 Crystallogr.* **2000**, *B56*, 697−714.
(5) 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 Crystallogr.* **2011**, *B67*, 535−551.
(6) Williams, H. L.; Szalewicz, K.; Jeziorski, B.; Moszynski, R.; Rybak, S. Symmetry-Adapted Perturbation Theory Calculation of the Ar-H_2 Intermolecular Potential Energy Surface. *J. Chem. Phys.* **1993**, *98*, 1279−1292.
(7) Lotrich, V. F.; Williams, H. L.; Szalewicz, K.; Jeziorski, B.; Moszynski, R.; Wormer, P. E. S.; van der Avoird, A. Intermolecular Potential and Rovibrational Levels of Ar-HF from Symmetry-Adapted Perturbation Theory. *J. Chem. Phys.* **1995**, *103*, 6076−6092.
(8) Mas, E. M.; Szalewicz, K.; Bukowski, R.; Jeziorski, B. Pair Potential for Water from Symmetry-Adapted Perturbation Theory. *J. Chem. Phys.* **1997**, *107*, 4207−4218.
(9) Liu, B.; McLean, A. D. The Interacting Correlated Fragments Model for Weak Interactions, Basis Set Superposition Error, and the Helium Dimer Potential. *J. Chem. Phys.* **1989**, *91*, 2348−2359.
(10) Bartlett, R. J. Coupled-Cluster Approach to Molecular Structure and Spectra - A Step Toward Predictive Quantum Chemistry. *J. Phys. Chem.* **1989**, *93*, 1697−1708.
(11) Szalewicz, K.; Jeziorski, B. Symmetry-Adapted Double-Perturbation Analysis of Intramolecular Correlation Effects in Weak Intermolecular Interactions. *Mol. Phys.* **1979**, *38*, 191−208.
(12) Rybak, S.; Jeziorski, B.; Szalewicz, K. Many-Body Symmetry-Adapted Perturbation Theory of Intermolecular Interactions - H_2O and HF Dimers. *J. Chem. Phys.* **1991**, *95*, 6576−6601.
(13) 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.
(14) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, 1133−1138.
(15) Kristyan, S.; Pulay, P. Can (Semi)Local Density Functional Theory Account for the London Dispersion Forces? *Chem. Phys. Lett.* **1994**, *229*, 175−180.
(16) Wu, X.; Vargas, M. C.; Nayak, S.; Lotrich, V. L.; Scoles, G. Towards Extending the Applicability of Density Functional Theory to Weakly Bound Systems. *J. Chem. Phys.* **2001**, *115*, 8748−8757.
(17) Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. Intermolecular Potentials Based on Symmetry-Adapted Perturbation Theory Including Dispersion Energies from Time-Dependent Density Functional Calculations. *J. Chem. Phys.* **2005**, *123*, No. 214103.
(18) 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*, No. 014103.
(19) Podeszwa, R.; Cencek, W.; Szalewicz, K. Efficient Calculations of Dispersion Energies for Nanoscale Systems from Coupled Density Response Functions. *J. Chem. Theory Comput.* **2012**, *8*, 1963−1969.
(20) Holden, J. R.; Du, Z.; Ammon, H. L. Prediction of Possible Crystal Structures for C-, H-, N-, O-, and F-containing Organic Compounds. *J. Comput. Chem.* **1993**, *14*, 422−437.
(21) Woodley, S. M.; Catlow, R. Crystal Structure Prediction from First Principles. *Nat. Mater.* **2008**, *7*, 937−946.
(22) Price, S. L. Computed Crystal Energy Landscapes for Understanding and Predicting Organic Crystal Structures and Polymorphism. *Acc. Chem. Res.* **2009**, *42*, 117−126.
(23) Price, S. L. Predicting Crystal Structures of Organic Compounds. *Chem. Soc. Rev.* **2014**, *43*, 2098−2111.
(24) Stone, A. J. *The Theory of Intermolecular Forces*, 2nd ed.; Clarendon Press: Oxford, U.K., 2013.
(25) Szalewicz, K. Symmetry-Adapted Perturbation Theory of Intermolecular Forces. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 254−272.
(26) 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.
(27) Sorescu, D. C.; Rice, B. M.; Thompson, D. L. Intermolecular Potential for the Hexahydro-1,3,5-trinitro-1,3,5-s-triazine Crystal (RDX): A Crystal Packing, Monte Carlo, and Molecular Dynamics Study. *J. Phys. Chem. B* **1997**, *101*, 798−808.
(28) Korona, T.; Williams, H. L.; Bukowski, R.; Jeziorski, B.; Szalewicz, K. Symmetry-Adapted Perturbation Theory Calculation of He−He Interaction Energy. *J. Chem. Phys.* **1997**, *106*, 5109−5122.
(29) Patkowski, K.; Jeziorski, B.; Szalewicz, K. Unified Treatment of Chemical and van der Waals Forces via Symmetry-Adapted Perturbation Expansion. *J. Chem. Phys.* **2004**, *120*, 6849−6862.
(30) Jeziorska, M.; Cencek, W.; Patkowski, K.; Jeziorski, B.; Szalewicz, K. Pair Potential for Helium from Symmetry-Adapted Perturbation Theory Calculations and from Supermolecular Data. *J. Chem. Phys.* **2007**, *127*, No. 124303.
(31) Bukowski, R.; Cencek, W.; Jankowski, P.; Jeziorska, M.; Jeziorski, B.; Kucharski, S. A.; Lotrich, V. F.; Misquitta, A. J.; Moszyn′ski, R.; Patkowski, K.; Podeszwa, R.; Rob, F.; Rybak, S.; Szalewicz, K.; Williams, H. L.; Wheatley, R. J.; Wormer, P. E. S.; Żuchowski, P. S. SAPT2012: An *Ab Initio* Program for Many-Body Symmetry-Adapted Perturbation Theory Calculations of Intermolecular Interaction Energies. University of Delaware and University of Warsaw, 2012; http://www.physics.udel.edu/~szalewic/SAPT/SAPT.html.
(32) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Schütz, M.; Celani, P.; Korona, T.; Manby, F. R.; Rauhut, G.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Hampel, C.; Hetzer, G.; Lloyd, A. W.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pitzer, R.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T. MOLPRO, version 2009.1, A Package of Ab Initio Programs. 2009; See http://www.molpro.net.
(33) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. PSI4: An Open-Source Ab Initio Electronic Structure Program. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 556−565.
(34) Misquitta, A. J.; Stone, A. J. CamCASP: A Program for Studying Intermolecular Interactions and for Calculations of Molecular Properies in Distributed Form. University of Cambridge, UK, 2010; http://www-stone.ch.cam.ac.uk/programs/camcasp.html.
(35) Patkowski, K.; Szalewicz, K. Argon Pair Potential at Basis and Excitation Limits. *J. Chem. Phys.* **2010**, *133*, No. 094304.
(36) Jensen, F. *Introduction to Computational Chemistry*, 2nd ed.; Wiley, Chichester, U.K., 2007.
(37) Cencek, W.; Szalewicz, K. On Asymptotic Behavior of Density Functional Theory. *J. Chem. Phys.* **2013**, *139*, No. 024104.
(38) Cencek, W.; Szalewicz, K. Erratum "On Asymptotic Behavior of Density Functional Theory" [J. Chem. Phys. 139, 024104 (2013)]. *J. Chem. Phys.* **2014**, *140*, No. 149902.
(39) Williams, H. L.; Chabalowski, C. F. Using Kohn-Sham Orbitals in Symmetry-Adapted Perturbation Theory to Investigate Intermolecular Interactions. *J. Phys. Chem. A* **2001**, *105*, 646−659.
(40) Misquitta, A. J.; Szalewicz, K. Intermolecular Forces From Asymptotically Corrected Density Functional Description of Monomers. *Chem. Phys. Lett.* **2002**, *357*, 301−306.
(41) Hesselmann, A.; Jansen, G. First-Order Intermolecular Interaction Energies from Kohn-Sham Orbitals. *Chem. Phys. Lett.* **2002**, *357*, 464−470.

(42) Misquitta, A. J.; Jeziorski, B.; Szalewicz, K. Dispersion Energy from Density-Functional Theory Description of Monomers. *Phys. Rev. Lett.* **2003**, *91*, No. 033201.
(43) Hesselmann, A.; Jansen, G. Intermolecular Dispersion Energies from Time-Dependent Density Functional Theory. *Chem. Phys. Lett.* **2003**, *367*, 778−784.
(44) Misquitta, A. J.; Szalewicz, K. Symmetry-Adapted Perturbation Theory Calculations of Intermolecular Forces Employing Density Functional Description of Monomers. *J. Chem. Phys.* **2005**, *122*, No. 214109.
(45) Bukowski, R.; Podeszwa, R.; Szalewicz, K. Efficient Calculations of Coupled Kohn-Sham Dynamic Susceptibility Functions and Dispersion Energies with Density Fitting. *Chem. Phys. Lett.* **2005**, *414*, 111−116.
(46) Podeszwa, R.; Bukowski, R.; Szalewicz, K. Density Fitting Methods in Symmetry-Adapted Perturbation Theory Based on Kohn-Sham Description of Monomers. *J. Chem. Theory Comput.* **2006**, *2*, 400−412.
(47) 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.
(48) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(49) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. Design of Density Functionals by Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2006**, *2*, 364−382.
(50) Peverati, R.; Truhlar, D. G. Improving the Accuracy of Hybrid Meta-GGA Density Functionals by Range Separation. *J. Phys. Chem. Lett.* **2011**, *2*, 2810−2817.
(51) Pernal, K.; Podeszwa, R.; Patkowski, K.; Szalewicz, K. Dispersionless Density Functional Theory. *Phys. Rev. Lett.* **2009**, *103*, No. 263201.
(52) Gianturco, F. A.; Paesani, F.; Laranjeira, M. F.; Vassilenko, V.; Cunha, M. A.; Shashkov, A. G.; Zolotoukhina, A. F. Computed and Measured Transport Coefficients for CO-He Mixtures: Testing a Density Functional Approach. *Mol. Phys.* **1998**, *94*, 605−622.
(53) Wu, Q.; Yang, W. Empirical Correction to Density Functional Theory for van der Waals Interactions. *J. Chem. Phys.* **2002**, *116*, 515−524.
(54) Grimme, S. Accurate Description of van der Waals Complexes by Density Functional Theory Including Empirical Corrections. *J. Comput. Chem.* **2004**, *25*, 1463−1473.
(55) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(56) Grimme, S.; Antony, J.; Elrich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, No. 154104.
(57) Podeszwa, R.; Szalewicz, K. Density Functional Theory Overcomes the Failure of Predicting Intermolecular Interaction Energies. *J. Chem. Phys.* **2012**, *136*, No. 161102.
(58) Thanthiriwatte, K. S.; Hohenstein, E. G.; Burns, L. A.; Sherrill, C. D. Assessment of the Performance of DFT and DFT-D Methods for Describing Distance Dependence of Hydrogen-Bonded Interactions. *J. Chem. Theory Comput.* **2011**, *7*, 88−96.
(59) Burns, L. A.; Vázquez-Mayagoitia, A.; Sumpter, B. G.; Sherrill, C. D. Density-Functional Approaches to Noncovalent Interactions: A Comparison of Dispersion Corrections (DFT-D), Exchange-Hole Dipole Moment (XDM) Theory, and Specialized Functionals. *J. Chem. Phys.* **2011**, *134*, No. 084107.
(60) Rob, F.; Szalewicz, K. Asymptotic Dispersion Energies from Distributed Polarizabilities. *Chem. Phys. Lett.* **2013**, *572*, 146−149.
(61) Rob, F.; Szalewicz, K. Distributed Molecular Polarizabilities and Asymptotic Intermolecular Interaction Energies. *Mol. Phys.* **2013**, *111*, 1430−1455.
(62) Rob, F.; Misquitta, A. J.; Podeszwa, R.; Szalewicz, K. Localized-Overlap Algorithm for Unexpanded Dispersion Energies. *J. Chem. Phys.* **2014**, *140*, No. 114304.
(63) Lotrich, V. F.; Szalewicz, K. Symmetry-Adapted Perturbation Theory of Three-Body Nonadditivity of Intermolecular Interaction Energy. *J. Chem. Phys.* **1997**, *106*, 9668−9687.
(64) Podeszwa, R.; Szalewicz, K. Three-Body Symmetry-Adapted Perturbation Theory Based on Kohn-Sham Description of the Monomers. *J. Chem. Phys.* **2007**, *126*, No. 194101.
(65) Podeszwa, R.; Rice, B. M.; Szalewicz, K. On Predicting Structure of Molecular Crystals from First Principles. *Phys. Rev. Lett.* **2008**, *101*, No. 115503.
(66) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.
(67) Maragliano, L.; Vanden-Eijnden, E. A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations. *Chem. Phys. Lett.* **2006**, *426*, 168−175.
(68) Yu, T.-Q.; Chen, P.-Y.; Chen, M.; Samanta, A.; Vanden-Eijnden, E.; Tuckerman, M. Order-Parameter-Aided Temperature-Accelerated Sampling for the Exploration of Crystal Polymorphism and Solid-Liquid Phase Transitions. *J. Chem. Phys.* **2014**, *140*, No. 214109.
(69) 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.
(70) Neumann, M. A.; Leusen, F. J. J.; Kendrick, J. A Major Advance in Crystal Structure Prediction. *Angew. Chem., Int. Ed.* **2008**, *47*, 1−5.
(71) Taylor, D.; Rice, B. M.; Podeszwa, R.; Rob, F.; Szalewicz, K. A Molecular Dynamics Study of 1,1-Diamino-2,2-Dinitroethylene (FOX-7) Crystal Using a Symmetry Adapted Perturbation Theory-Based Intermolecular Force Field. *Phys. Chem. Chem. Phys.* **2011**, *13*, 16629−16636.
(72) Peiris, S. M.; Wong, C. P.; Zerilli, F. J. Equation of State and Structural Changes in Diaminodinitroethylene Under Compression. *J. Chem. Phys.* **2004**, *120*, 8060−8066.
(73) Sorescu, D.; Boatz, J.; Thompson, D. Classical and Quantum-Mechanical Studies of Crystalline FOX-7 (1,1-diamino-2,2-dinitroethylene). *J. Phys. Chem. A* **2001**, *105*, 5010−5021.
(74) Murdachaew, G.; Bukowski, R.; Szalewicz, K. Efficient Generation of Flexible-Monomer Intermolecular Potential Energy Surfaces. *Phys. Rev. Lett.* **2002**, *88*, No. 123202.
(75) Neumann, M. A.; Perrin, M. Energy Ranking of Molecular Crystals Using Density Functional Theory Calculations and an Empirical van der Waals Correction. *J. Phys. Chem. B* **2005**, *109*, 15531−15541.