Angewandte
Highlights

Polymorphism

DOI: 10.1002/anie.201409823

A New Era for ab initio Molecular Crystal Lattice Energy Prediction**

Gregory J. O. Beran*

ab initio calculations · crystal growth ·
electronic structure · polymorphism ·
solid-state structures

Organic molecular crystal packing plays a critical role in determining the solubility and bioavailability of a pharmaceutical, the charge-carrier mobility in an organic semiconductor, and even the taste of chocolate. A given molecule can often adopt multiple crystal packing motifs, or polymorphs. Theoretical chemistry could play an important role in understanding crystal polymorphism or even engineering crystals for specific functions. To this end, Gavezzotti asked twenty years ago, "Are crystal structures predictable?", at which time the answer was "No".[1]

At the core, the challenge stems from the small, approximately 1 kJ mol^{-1}, energy differences which often separate different crystal polymorphs. Different crystal polymorphs of a given molecule frequently exhibit competition between intramolecular conformation and intermolecular interactions, and among noncovalent interactions, including multipolar interactions, polarization, hydrogen bonding, and van der Waals dispersion. Reliably predicting the most stable crystal packing motifs requires theoretical models which can describe all of these different types of interactions in condensed-phase systems with high and uniform accuracy, and manageable computational requirements.

Tremendous progress in crystal structure modeling has been made over the years using classical molecular mechanics force fields,[2] but the need for more reliable quantum mechanical descriptions of important interactions has become increasingly apparent. Indeed, quantum chemistry techniques have begun to play a major role in the field over the past five years or so. Periodic density functional theory (DFT) made its mark with the first successful prediction of all four molecules in the 2010 blind test of crystal structure prediction.[3] Subsequent improvements in the inclusion of van der Waals dispersion in DFT have led to many successful molecular crystal studies, with lattice energies often predicted with accuracies of 4–5 kJ mol^{-1}.[4]

Even higher accuracy can potentially be obtained using wave-function-based electronic structure methods like Møller–Plesset perturbation theory or coupled cluster methods. Such methods were traditionally computationally prohibitive for molecular crystals, but the past several years have witnessed substantial improvements both in periodic implementations of such techniques for infinite crystals[5] and in so-called fragment and/or embedding methods which decompose a crystal into a large number of more tractable fragments consisting of smaller numbers of molecules, sometimes embedded in a crystalline environment.[6]

By combining such a fragment approach with state-of-the-art electronic structure methods, Yang et al.[7] have now predicted the lattice energy for crystalline benzene to within sub-kilojoule per mole accuracy, thus obtaining a theoretical best estimate for the 0 K lattice energy of (55.9 ± 0.9) kJ mol^{-1}, compared to the corresponding experimental value of (55.3 ± 2.2) kJ mol^{-1} (Figure 1).

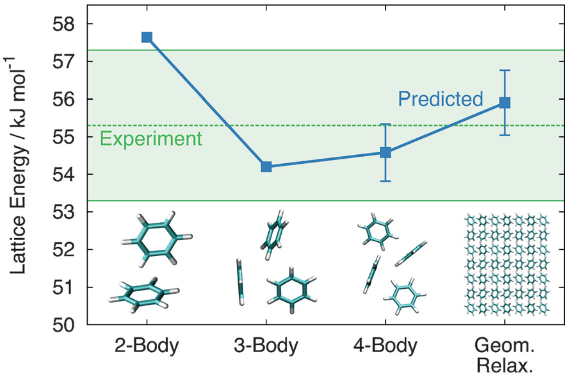

Figure 1. Convergence of the predicted benzene lattice energy as a function of the types of interactions included (dimers, trimers, etc.) and relaxation to the 0 K structure relative to the experimental value of (55.3 ± 2.2) kJ mol^{-1}.[7]

While Yang and et al. are not the first to strive for a high accuracy lattice energy prediction in benzene,[8] they made a particularly impressive effort to capture the large numbers of small energy contributions required with uniformly high levels of theory. They evaluated interaction energies in fragments containing up to four benzene molecules at a time with accurate coupled cluster models, they examined electron–electron correlations involving up to four electrons simultaneously, and they used a variety of strategies to

---

[*] Prof. G. J. O. Beran
Department of Chemistry, University of California at Riverside
Riverside, CA 92521 (USA)
E-mail: gregory.beran@ucr.edu

[**] Support from the National Science Foundation (CHE-1362465) is gratefully acknowledged.

estimate the errors arising from the remaining neglected contributions.

Performing high-level coupled cluster calculations on large numbers of benzene dimers, trimers, and tetramers in large basis sets is only feasible now thanks to years of research by many groups, thus resulting in 1) advances in local correlation approximations and arbitrary order coupled cluster techniques which allow one to compute the most important electron–electron correlations, 2) "explicitly correlated" techniques for achieving convergence with respect to the finite basis set used to represent the molecular orbitals, and 3) the ever-increasing capabilities and decreasing cost of computer hardware. Because they decompose a complex system into many simpler ones, fragment methods naturally scale well to thousands of computer processors, thus making them well-suited for modern multicore supercomputer architectures.

The new benzene lattice energy study also highlights another increasingly appreciated facet of predicting molecular crystal energetics: Given the miniscule energy differences between polymorphic structures, it is essential to estimate the inherent error bars on calculated polymorph energetics. Yang et al. expended considerable effort to estimate the contributions from small variations in the crystal geometry, longer-range interactions, higher-order electron–electron correlation terms, and the finite basis set. Such error analysis is the only way to assess whether a predicted $\approx 1\ \mathrm{kJ\,mol^{-1}}$ energy difference between polymorphs is meaningful within the computational approximations employed.

Given the achievement of Yang et al., are crystal structures now predictable? Momentum is clearly pointing toward "Yes", but we are not there yet. The fact that such coupled cluster calculations can now be performed on benzene trimers and tetramers is remarkable, but analogous calculations on a polymorphic drug like Lipitor ($C_{33}H_{35}FN_2O_5$) will remain cost-prohibitive for years to come. Promising electronic-structure methods with accuracy approaching that of coupled cluster theory at much lower cost are being developed,[9] though various limitations must still be overcome.

Equally problematic is the convergence of the many-body expansion (dimer, trimer, tetramer contributions, etc.), which often exhibits poor, oscillatory behavior in the higher-order terms.[10] The relatively rapid convergence (no more than tetramers) observed for benzene will not necessarily be observed for the cooperative hydrogen bonding found in ices, for instance. Electrostatic embedding can help decrease the importance of higher-order many-body terms, or the many-body contributions can be summed to infinite order using a simpler model.[6]

In addition, the crystal structure prediction problem encompasses much more than simply evaluating the lattice energy of putative structures. First, generating a short list of candidate structures is challenging itself, though substantial progress in hierarchical Monte Carlo searches[2,11] and global optimization algorithms[12] makes this tractable for molecules with up to a handful of degrees of freedom. Nevertheless, modern pharmaceuticals often contain ten or more flexible torsional degrees of freedom, thus pushing them well beyond current crystal structure prediction capabilities.

Second, traditional crystal structure prediction occurs at 0 K, while the real world operates at finite temperatures and pressures. Considerations of lattice expansion, thermal energy contributions, and entropy all play an important role in shaping the free-energy surface and phase diagram. At the same time, the kinetics of crystallization can be at least as important as the thermodynamics in determining which crystal structures are produced under given experimental conditions. Enhanced and rare-event molecular dynamics sampling techniques are making impressive progress in the study of crystal free-energy surfaces and crystallization processes,[13] but the need to sample so many configurations often means that the sampling is performed using comparatively crude force field models. Practical strategies for combining high quality energy models with these advanced statistical mechanical techniques are needed.

Finally, further interaction between theory and experiment should be encouraged. On the one hand, modelers are increasingly recognizing the value of predicting the mechanical, electronic, and spectroscopic properties of crystals to complement experimental vibrational/terahertz spectroscopy, nuclear magnetic resonance crystallography, organic materials design, etc. On the other hand, the effort required by Yang et al. to obtain a reliable experimental value for the benzene lattice energy is symptomatic of the relative dearth of modern experimental thermochemistry measurements for molecular crystals. Theorists developing models for molecular crystals desperately need more experimental thermochemical data on crystal polymorphs against which to test their models and to identify which approximations are most effective. The ability to predict the experimentally observed structure is a valuable but insufficient metric for the reliability of a model.

In summary, the impressive sub-kilojoule per mole accuracy obtained for the benzene lattice energy by Yang et al. reflects the culmination of many years of research in quantum chemistry. Indeed, we are reaching the stage where the thermodynamically most stable polymorphs for small, relatively rigid molecules can be predicted fairly reliably. The expansion of these successes to broader classes of molecular crystals over the next several years should be particularly exciting.

Received: October 7, 2014

---

[1] A. Gavezzotti, *Acc. Chem. Res.* **1994**, *27*, 309–314.

[2] a) S. L. Price, *Chem. Soc. Rev.* **2014**, *43*, 2098–2111; b) S. L. Price, M. Leslie, G. W. A. Welch, M. Habgood, L. S. Price, P. G. Karamertzanis, G. M. Day, *Phys. Chem. Chem. Phys.* **2010**, *12*, 8478–8490.

[3] M. A. Neumann, F. J. J. Leusen, J. Kendrik, *Angew. Chem.* **2008**, *120*, 2461–2464; *Angew. Chem. Int. Ed.* **2008**, *47*, 2427–2430.

[4] a) A. Otero-de-la-Roza, E. R. Johnson, *J. Chem. Phys.* **2012**, *137*, 054103; b) A. M. Reilly, A. Tkatchenko, *J. Chem. Phys. Lett.* **2013**, *4*, 1028–1033; c) J. G. Brandenburg, S. Grimme, *Top. Curr. Chem.* **2014**, *345*, 1–23.

[5] a) M. Marsman, A. Grueneis, J. Paier, G. Kresse, *J. Chem. Phys.* **2009**, *130*, 184103; b) L. Maschio, D. Usvyat, B. Civalleri, *CrystEngComm* **2010**, *12*, 2429–2435; c) M. Del Ben, J. Hutter, J. VandeVondele, *J. Chem. Theory Comput.* **2013**, *9*, 2654–2671.

[6] a) S. Wen, K. Nanda, Y. Huang, G. J. O. Beran, *Phys. Chem. Chem. Phys.* **2012**, *14*, 7578–7590; b) P. J. Bygrave, N. L. Allen, F. R. Manby, *J. Chem. Phys.* **2012**, *137*, 164102; c) S. Hirata, K. Gilliard, X. He, J. Li, O. Sode, *Acc. Chem. Res.* **2014**, *47*, 2721–2730.

[7] J. Yang, W. Hu, D. Usvyat, D. Matthews, M. Schutz, G. K.-L. Chan, *Science* **2014**, *345*, 640–643.

[8] a) W. B. Schweizer, J. D. Dunitz, *J. Chem. Theory Comput.* **2006**, *2*, 288–291; b) A. L. Ringer, C. D. Sherrill, *Chem. Eur. J.* **2008**, *14*, 2542–2547; c) M. R. Kennedy, A. R. McDonald, A. E. DePrince, M. S. Marshall, R. Podeszwa, C. D. Sherrill, *J. Chem. Phys.* **2014**, *140*, 121104; d) O. Bludsky, M. Rubes, P. Soldan, *Phys. Rev. B* **2008**, *77*, 092103; e) R. Podeszwa, B. M. Rice, K. Szalewicz, *Phys. Rev. Lett.* **2008**, *101*, 115503; f) S. Wen, G. J. O. Beran, *J. Chem. Theory Comput.* **2011**, *7*, 3733–3742.

[9] a) M. Pitoňák, A. Heßelmann, *J. Chem. Theory Comput.* **2010**, *6*, 168–178; b) T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno, C. D. Sherrill, *J. Chem. Phys.* **2014**, *140*, 094106.

[10] a) R. M. Richard, K. U. Lao, J. M. Herbert, *Acc. Chem. Res.* **2014**, *47*, 2828–2836; b) J. F. Ouyang, M. W. Cvitkovic, R. P. A. Bettens, *J. Chem. Theory Comput.* **2014**, *10*, 3699–3707.

[11] H. P. G. Thompson, G. M. Day, *Chem. Sci.* **2014**, *5*, 3173–3182.

[12] a) B. C. Revard, W. W. Tipton, R. G. Hennig, *Top. Curr. Chem.* **2014**, *345*, 181–222; b) Q. Zhu, A. R. Oganov, C. W. Glass, H. T. Stokes, *Acta Crystallogr. Sect. B* **2012**, *68*, 215–226.

[13] a) P. Raiteri, R. Martonak, M. Parrinello, *Angew. Chem. Int. Ed.* **2005**, *44*, 3769–3773; *Angew. Chem.* **2005**, *117*, 3835–3839; b) T.-Q. Yu, M. E. Tuckerman, *Phys. Rev. Lett.* **2011**, *107*, 3–6; c) J. Anwar, D. Zahn, *Angew. Chem. Int. Ed.* **2011**, *50*, 1996–2013; *Angew. Chem.* **2011**, *123*, 2042–2061; d) M. Salvalaglio, T. Vetter, F. Giberti, M. Mazzotti, M. Parrinello, *J. Am. Chem. Soc.* **2012**, *134*, 17221–17233.