Many-Body Dispersion Interactions in Molecular Crystal Polymorphism

Noa Marom[1], Robert A. DiStasio Jr.[2], Viktor Atalla[3], Sergey Levchenko[3], Anthony M. Reilly[3], James R. Chelikowsky[1], Leslie Leiserowitz[4], and Alexandre Tkatchenko[3]
[1]*Center for Computational Materials, Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX 78712, USA*
[2]*Department of Chemistry, Princeton University, Princeton, NJ 08544, USA*
[3]*Fritz–Haber–Institut der Max–Planck–Gesellschaft, Faradayweg 4-6, 14195, Berlin, Germany*
[4]*Department of Materials and Interfaces, Weizmann Institute of Science, Rehovoth 76100, Israel*
(Dated: April 12, 2013)

Polymorphs of molecular crystals are often very close in energy, yet they may possess markedly different physical and chemical properties. The understanding of polymorphism is therefore of paramount importance for a variety of applications, ranging from pharmaceutical design to non-linear optics and even hydrogen storage [1–3]. While the crystal structure blind tests conducted by the Cambridge Crystallographic Data Centre have shown steady progress toward reliable structure prediction for molecular crystals [4], several challenges still remain, including molecular salts, hydrates, and flexible molecules with several stable conformers. The ability to identify and rank all of the relevant polymorphs of a given molecular crystal hinges on an accurate description of their relative energetic stability. Hence a first-principles quantum mechanical method that can attain the required accuracy of approximately 0.1–0.2 kcal/mol per molecule would clearly be an indispensable tool for polymorph prediction. In this work, we show that accounting for the non-additive many-body dispersion (MBD) energy beyond the standard pairwise approximation is crucial for the correct qualitative and quantitative description of polymorphism in molecular crystals. We demonstrate this via three fundamental and stringent benchmark examples: glycine, oxalic acid, and tetrolic acid. These systems represent a broad class of molecular crystals, comprising hydrogen bonded networks of amino acids and carboxylic acids, and were also chosen because the standard pairwise methods for dispersion interactions are unable to correctly account for the observed stability of the different polymorphs in these molecular crystals [5, 6].

Among the first-principles methods available to date, density-functional theory (DFT) is the most frequently used approach in the study of polymorphism in molecular crystals. However, the most commonly utilized exchange-correlation functionals (including hybrid functionals) are based on semi-local electron correlation, and thereby fail to capture the contribution of dispersion interactions to the stabilization of molecular crystals. These ubiquitous non-covalent interactions are quantum mechanical in nature and physically correspond to the multipole moments that are induced in response to instantaneous fluctuations in the electron density. To incorporate these long-range electron correlation effects within the framework of DFT, significant progress has been made by utilizing the standard C_6/R^6 pairwise additive expression for the dispersion energy as derived from second-order perturbation theory [7–9]. Indeed, DFT with pairwise dispersion corrections often yields accurate results when the energy differences between molecular crystal polymorphs are sufficiently large [10–12]. Most notably, Neumann *et al.* have achieved the highest success rate in the last two blind tests using such methods [4, 13]. However, these pairwise dispersion energy approaches, even when used in conjunction with state-of-the-art functionals, are still unable to furnish the level of accuracy required to describe polymorphism in many relevant molecular crystals [14–17].

Recently, a novel and efficient method for describing the many-body dispersion (MBD) energy has been developed [18], building upon the Tkatchenko-Scheffler (TS) pairwise method [19]. Within the TS approach, the effective dispersion coefficients (C_6), which are proportional to the atomic polarizabilities, are calculated from the DFT electron density, hence the effect of the local environment of an atom in a molecule is accurately accounted for by construction. The MBD method presents a two-fold improvement over the TS approach by including: (*i*) the long-range electrodynamic screening via the self-consistent solution of the dipole–dipole electric-field coupling equations for the effective polarizability, and (*ii*) the non-pairwise-additive many-body dispersion energy to infinite order via diagonalization of the Hamiltonian corresponding to a system of coupled fluctuating dipoles. The inclusion of the MBD energy in DFT leads to a significant improvement in the binding energies between organic molecules, and for the cohesion of the benzene molecular crystal [18]. The MBD energy, like the TS energy, can be added to any DFT functional, requiring only a once-per-functional adjustment of a single range-separation parameter [18, 19].

We begin with a detailed analysis of the glycine (Gly) molecular crystal, which has three experimentally observed polymorphs: α-Gly, β-Gly, and γ-Gly, as illustrated in Figure 1. Figure 2 shows the performance of different DFT methods for the calculation of the unit cell volumes of these glycine polymorphs with respect to low temperature experiments. A complete account

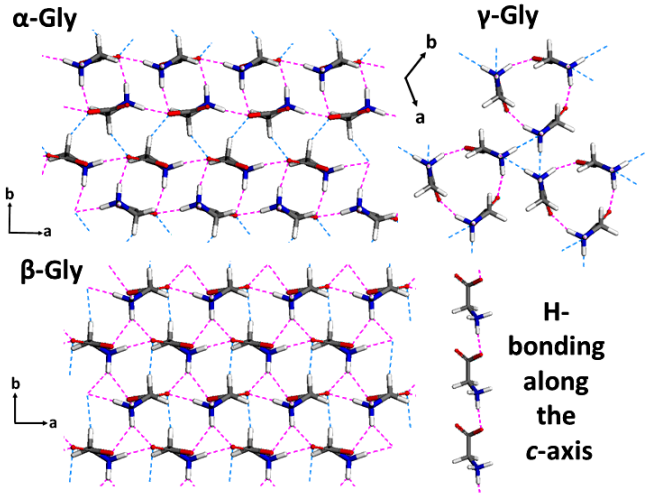


FIG. 1: Structures of the three polymorphs of glycine. H-bonds are indicated by dashed lines. The translation-related H-bonded chain along the c-axis, common to all three polymorphs, is also shown.

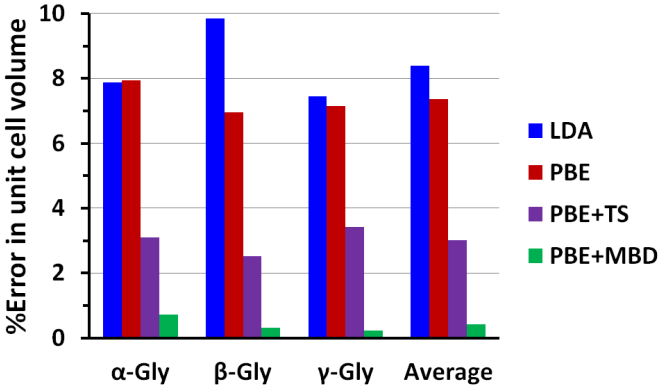


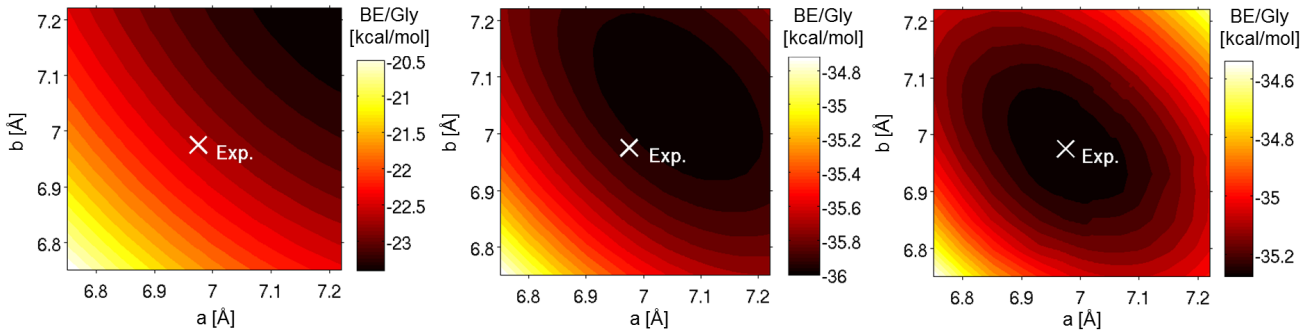
FIG. 2: Absolute percent error in the calculated unit cell volumes of the glycine polymorphs with respect to low-temperature experiments (α-Gly: Refs. [24, 25], β-Gly: Ref. [26], γ-Gly: Ref. [27]). The LDA and PBE data were taken from Ref. [20]. As described in the text, note that LDA underestimates the unit cell volumes, while PBE overestimates them.

of the computational details and of the cell parameters obtained with these different methods is provided in the supplementary material. As shown in Ref. [20], the local-density approximation (LDA) [21] underestimates the unit cell volumes by 7-10%, while the generalized gradient approximation of Perdew, Burke, and Ernzerhof (PBE) [22, 23] overestimates the unit cell volumes by 7-8% [20]. Adding the pairwise TS energy to the PBE functional reduces the error in the unit cell volumes to about 3%, which is already a significant improvement. PBE+MBD yields a further noticeable improvement with an accuracy of 0.3% for the unit cell volumes of β-Gly and γ-Gly and 0.8% for α-Gly compared to experimental values.

Both α-Gly and β-Gly consist of H-bonded sheets of molecular glycine in the a–c plane. The strong H-bonds within the glycine sheets (colored in magenta in Fig. 1), are described reasonably well by PBE even without accounting for dispersion interactions. This is not the case for the weaker interactions between the glycine sheets, along the b direction (colored in light blue in Fig. 1). For β-Gly, in which the glycine sheets are bound by bifurcated NH···O bonds, PBE overestimates b by 5%. PBE+TS reduces this overestimation to 1% and PBE+MBD yields excellent agreement with experiment. In α-Gly, the glycine sheets form a H-bonded (NH···O) bilayer, via the centers of inversion. The three-dimensional (3D) network is then completed by weaker CH···O interactions between the bilayers. These interactions determine the direction of the glide as well as the inter-bilayer distance along the b-axis. The weak interactions along the b-direction are reflected by a significant temperature dependence of the b parameter of α-Gly [25, 28]. PBE grossly overestimates b by 0.65 Å, which is significantly reduced to 0.16 Å at the PBE+TS level of theory. In the case of the b parameter, PBE+MBD does not yield further improvement over PBE+TS because the potential energy surface is very flat—the binding energy changing by only 0.01 eV per unit cell for 11.75 Å $< b <$ 12.15 Å.

The most stable γ-Gly polymorph has the same translation-related H-bonded chain motif as α-Gly and β-Gly along the c-axis. However, this is unique in the sense that the H-bonded chains form helices, related by three-fold screw symmetry, rather than sheets. The helices are held together by lateral NH···O H-bonds, forming a 3D network. The inter-helix H-bonds (colored in light blue in Fig. 1) are longer than the intra-helix H-bonds (colored in magenta in Fig. 1). The c parameter is reproduced correctly even by PBE, which clearly is able to successfully capture the strong intra-helix bonds. The a and b parameters are significantly improved by accounting for dispersion interactions. Figure 3 shows the change of the potential-energy landscape in the a–b plane of γ-Gly (with c fixed at 5.48 Å), resulting from including the dispersion contributions at different levels of approximation. The TS pairwise dispersion method significantly increases the binding energy and improves the position of the minimum, as compared to standard PBE. However, it is still insufficient for obtaining the correct experimental geometry. Accounting for the MBD interactions correctly captures the weak and complex inter-helix interactions, leading to a slight decrease in the crystal binding energy and yielding a minimum in excellent agreement with experiment.

We now proceed to discuss the relative stability of the glycine polymorphs. Experimentally, it has been deter-



(a) PBE (b) PBE+TS (c) PBE+MBD

FIG. 3: Potential energy surfaces for the a–b plane of γ-Gly [29]. Experimental lattice parameters are marked by a cross [20]. The experimental error bars are not visible on this scale.

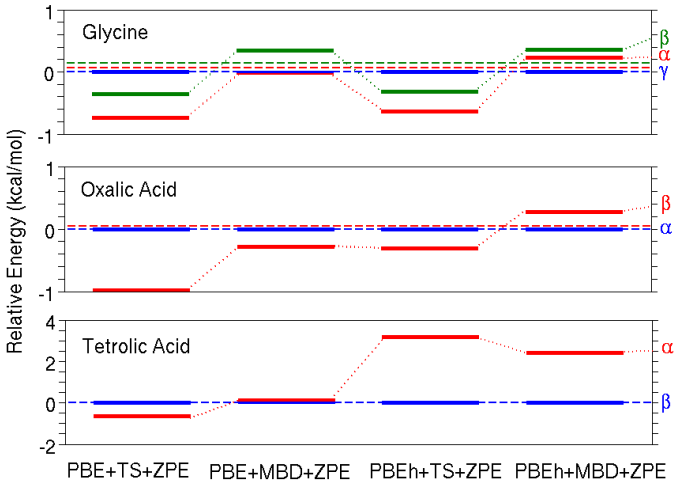


FIG. 4: Computed relative stabilities of the polymorphs of glycine (*top panel*), oxalic acid (*middle panel*), and tetrolic acid (*bottom panel*). Experimental data are shown as dashed lines.

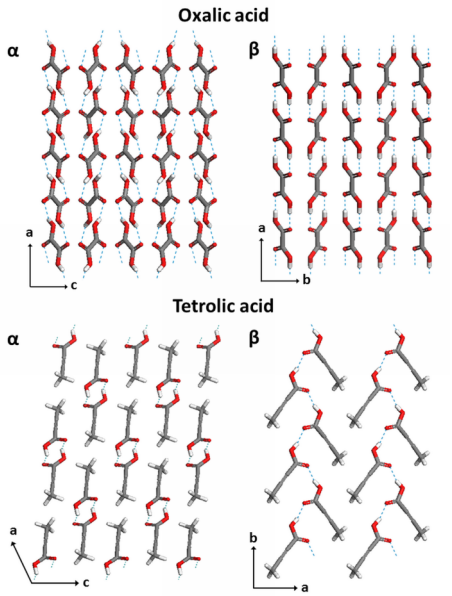


FIG. 5: Structures of the polymorphs of oxalic acid and tetrolic acid. α oxalic acid and β tetrolic acid comprise catemers. β oxalic acid and α tetrolic acid comprise cyclic dimers.

mined that γ-Gly is the most stable polymorph [28], although the energy difference between γ-Gly and α-Gly is very small. It is also well established that β-Gly is less stable than both γ-Gly and α-Gly [30]. The calculated relative energies, including zero-point energy (ZPE), are shown in Figure 4 and compared to the experimentally determined relative enthalpies from Ref. [28, 31]. Tabulated relative energies are given in the supplementary material. The PBE+TS method leads to the wrong order of stability: $\alpha > \beta > \gamma$ and the energy differences between the polymorphs are overestimated. Similar overstabilization of the α form has been reported for a different pairwise dispersion method [5]. Clearly, pairwise dispersion corrections fall short when the energy differences between polymorphs are small. Including the many-body dispersion effects via the PBE+MBD method significantly improves the agreement with experiment. In this case, the α and γ forms are nearly degenerate with the β form somewhat less stable, and the energy differences are also much closer to experiment. The PBE-based hybrid functional (PBEh), which includes 25% of exact exchange, has been shown to provide a more realible description of hydrogen bonds [32] and short-range vdW interactions [33] than PBE. Indeed, PBEh+MBD further improves upon the relative stability of the three polymorphs of glycine as shown in Figure 4, yielding the correct order of stability: $\gamma > \alpha > \beta$.

Our conclusions regarding the importance of MBD interactions also hold firmly for the polymorphs of oxalic

acid and tetrolic acid, as illustrated in Figure 5. Carboxylic acids have two modes of interlinking via OH···O hydrogen bonds: a cyclic dimer, as in β oxalic acid and α tetrolic acid, and a catemer, as in α oxalic acid and β tetrolic acid [34]. In both cases, the catemer structure is known to be more stable [35–37]. For oxalic acid, the enthalpies of sublimation of both polymorphs have been measured at room temperature [35, 36]. These measured enthalpies may be converted to a lattice energy difference of 0.05 kcal/mol (shown as dashed lines in Fig. 4) by using the PBE+TS vibrational quantities in the harmonic approximation, as described in Ref. [38]. The computed energy differences obtained using the PBE and PBEh functionals combined with the TS and MBD dispersion methods are shown in Fig. 4. Tabulated relative energies are given in the supplementary material. For both oxalic acid and tetrolic acid, PBE+TS overstabilizes the cyclic dimer with respect to the catemer and yields the wrong order of stability of the polymorphs. Similar overstabilization of the beta polymorph of oxalic acid has been reported for other pairwise dispersion methods [6]. The inclusion of exact exchange in the PBEh functional and the inclusion of MBD interactions both contribute to the stabilization of the catemer with respect to the cyclic dimer. For oxalic acid, similarly to glycine, only PBEh+MBD produces the correct energetic ordering of the polymorphs with α being more stable than β. For tetrolic acid, PBE+MBD already makes the catemer-based β form slightly more stable than the cyclic dimer-based α form and PBEh+MBD increases the energy difference between the two forms.

To summarize, we have demonstrated that an accurate description of the non-additive many-body dispersion energy with the DFT+MBD method reproduces the energetic ordering of polymorphs for three different molecular crystals, glycine, oxalic acid, and tetrolic acid, when compared to reliable experimental results. The improvement obtained with the MBD method as compared to the simple pairwise dipersion model is attributed to the sensitive dependence of the dispersion energy on the polymorph geometry and the dynamic internal electric fields produced within molecular crystals. The DFT+MBD method yields an unprecedented accuracy of 1% in the description of the structures of molecular crystal polymorphs and of 0.2 kcal/mol in their relative energies. Such accuracy is an essential ingredient for the reliable modeling of polymorphism in molecular crystals.

A.T. acknowledges support from ERC Starting Grant `VDW-CMAT`. Work at UT-Austin was supported under NSF grants DMR-0941645 and OCI-1047997, and DOE grant DE-SC0001878. Computational resources were provided by the Argonne Leadership Computing Facility (ALCF). We thank J. F. Hammond and O. A. von Lilienfeld from ALCF for their interest and help.

[1] J. Bernstein, in *Polymorphism in Molecular Crystals* (Oxford University Press, USA (New York), 2002).
[2] S. L. Price, Phys. Chem. Chem. Phys. **10**, 1996 (2008).
[3] S. L. Price, Acc. Chem. Res. **42**, 117 (2009).
[4] D. A. Bardwell *et al.*, Acta. Cryst. B **67**, 535 (2011).
[5] Q. Zhu, A. R. Oganov, C. W. Glass, and H. T. Stokes, Acta Crystallogr. Sect. B-Struct. Commun. **68**, 215 (2012).
[6] A. O. de-la Roza and E. R. Johnson, J. Chem. Phys. **137**, 054103 (2012).
[7] K. E. Riley, M. Pitonak, P. Jurecka, and P. Hobza, Chem. Rev. **110**, 5023 (2010).
[8] S. Grimme, Comput. Mol. Sci. **1**, 211 (2011).
[9] N. Marom, A. Tkatchenko, M. Rossi, V. V. Gobre, O. Hod, M. Scheffler, and L. Kronik, J. Chem. Theory Comput. **7**, 3944 (2011).
[10] H. C. S. Chan, J. Kendrick, and F. J. J. Leusen, Phys. Chem. Chem. Phys. **13**, 20361 (2011).
[11] N. Marom, A. Tkatchenko, S. Kapishnikov, L. Kronik, and L. Leiserowitz, Cryst. Growth Des. **11**, 3332 (2011).
[12] B. Schatschneider, J.-J. Liang, S. Jezowski, and A. Tkatchenko, Cryst. Eng. Comm. **14**, 4656 (2012).
[13] M. A. Neumann, F. J. J. Leusen, and J. Kendrick, Angew. Chem., Int. Ed. **47**, 2427 (2008).
[14] G. J. O. Beran and K. Nanda, J. Phys. Chem. Lett. **1**, 3480 (2010).
[15] K. Hongo, M. A. Watson, R. S. Sanchez-Carrera, T. Iitaka, and A. Aspuru-Guzik, J. Phys. Chem. Lett. **1**, 1789 (2010).
[16] O. A. von Lilienfeld and A. Tkatchenko, J. Chem. Phys. **132**, 234109 (2010).
[17] S. Wen, K. Nanda, Y. Huang, and G. Beran, Phys. Chem. Chem. Phys. **14**, 7578 (2012).
[18] A. Tkatchenko, R. A. DiStasio, Jr., R. Car, and M. Scheffler, Phys. Rev. Lett. **108**, 236402 (2012).
[19] A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102**, 073005 (2009).
[20] J. A. Chisholm, S. Motherwell, P. R. Tulip, S. Parsons, and S. J. Clark, Cryst. Growth Des. **5**, 1437 (2005).
[21] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).
[22] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[23] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **78**, 1396 (1997).
[24] J. Netzel, A. Hofmann, and S. van Smaalen, Cryst. Eng. Comm. **44**, 3226 (2008).
[25] R. Destro, P. Roversi, M. Barzaghi, and R. E. Marsh, J. Phys. Chem. A **104**, 1047 (2000).
[26] E. V. Boldyreva, T. N. Drebushchak, and E. S. Shutova, Z. Kristall. **218**, 366 (2003).
[27] A. Kvick, W. M. Canning, T. F. Koetzle, and G. J. B. Williams, Acta Crystallogr. Sect. B-Struct. Commun. **36**, 115 (1980).
[28] G. L. Perlovich, L. K. Hansen, and A. Bauer-Brandl, J. Therm. Anal **66**, 699 (2001).
[29] The binding energy per unit cell is referenced to a glycine zwitterion.
[30] I. Weissbuch, V. Y. Torbeev, L. Leiserowitz, and M. Lahav, Angew. Chem. Int. Ed. **44**, 3226 (2005).
[31] We assume that the solvation enthalpy of the three poly-

morphs in water is approximately equal, in which case the measured differences between the solution enthalpies are directly comparable to total energy differences.
[32] B. Santra, J. Klimes, D. Alfe, A. Tkatchenko, B. Slater, A. Michaelides, R. Car, and M. Scheffler, Phys. Rev. Lett. **107**, 185701 (2011).
[33] A. Tkatchenko and O. A. von Lilienfeld, Phys. Rev. B **78**, 045116 (2008).
[34] L. Leiserowitz, Acta Crystallogr. Sect. B-Struct. Commun. **32**, 775 (1976).
[35] H. G. M. de Wit, J. A. Bouwstra, J. G. Blok, and C. G. de Kruif, J. Chem. Phys. **78**, 1470 (1983).
[36] R. S. Bradley and S. Cotson, J. Chem. Soc., 1684(1953).
[37] V. Benghiat and L. Leiserowitz, J. Chem. Soc. Perkin Transactions II **12**, 1763 (1972).
[38] A. M. Reilly and A. Tkatchenko, J. Phys. Chem. Lett., in press, doi: 10.1021/jz400226x.