Assessing Density Functionals Using Many Body Theory for Hybrid Perovskites

Menno Bokdam,[*] Jonathan Lahnsteiner, Benjamin Ramberger, Tobias Schäfer, and Georg Kresse[†]
University of Vienna, Faculty of Physics and Center for Computational Materials Sciences, Sensengasse 8/12, A-1090 Vienna, Austria
(Received 29 June 2017; published 6 October 2017)

Which density functional is the "best" for structure simulations of a particular material? A concise, first principles, approach to answer this question is presented. The random phase approximation (RPA)—an accurate many body theory—is used to evaluate various density functionals. To demonstrate and verify the method, we apply it to the hybrid perovskite $MAPbI_3$, a promising new solar cell material. The evaluation is done by first creating finite temperature ensembles for small supercells using RPA molecular dynamics, and then evaluating the variance between the RPA and various approximate density functionals for these ensembles. We find that, contrary to recent suggestions, van der Waals functionals do not improve the description of the material, whereas hybrid functionals and the strongly constrained appropriately normed (SCAN) density functional yield very good agreement with the RPA. Finally, our study shows that in the room temperature tetragonal phase of $MAPbI_3$, the molecules are preferentially parallel to the shorter lattice vectors but reorientation on ps time scales is still possible.

There is no question that density functional theory (DFT) [1,2] has revolutionized our modeling of condensed matter and materials over the last five decades, and it is well understood why this is so: DFT offers a unique balance between computational efficiency and accuracy—a balance that is simply yet unreached by any other computational method. Equally important is that forces between the atoms are readily computable, so that relaxed ground state structures, vibrational properties, as well as finite temperature properties of any material are straightforwardly obtainable. But DFT is not without shortcomings. For instance, most semilocal functionals completely neglect van der Waals forces, and in many open structures such as perovskites the instabilities of the cage are not very well described. Quite generally, for materials modeling we are often left with the question of choosing the right density functional. This requires either some chemical intuition, or better, higher level reference calculations. In this Letter, we present a fully *ab initio* method based on the random phase approximation (RPA) to select the best functional for finite temperature structure predictions of a certain material (here hybrid perovskites).

The RPA is an approximate many-body technique for total energies summing the so-called bubble or ring diagrams to infinite order. Since it is closely related to the *GW* approximation of Hedin [3], it should describe band gap related properties very well. By comparison to experimental data and diffusion Monte Carlo calculations it has been shown that the RPA captures energy differences involving very diverse bonding types from covalent, over ionic, to van der Waals like very accurately [4–13]. The actual RPA calculations proceed in two steps. First, a DFT calculation is performed using an approximate density functional, in this case the PBE functional [14] and all occupied and unoccupied states are determined. Then the RPA correlation energy is evaluated as [5,15–19]

$$E_{RPA} = \frac{1}{2\pi} \int_0^\infty \mathrm{Tr}\{\ln[1 - \chi(i\omega)v] + \chi(i\omega)v\}d\omega,$$

where $\chi(i\omega)$ is the independent particle polarizability calculated using DFT orbitals, v is the Coulomb kernel, and $i\omega$ the imaginary frequency. The exact exchange energy (E_x) is calculated and added: $E_{XC} = E_x + E_{RPA}$.

If we trust RPA to be accurate for solids, the best strategy is obviously (i) to calculate finite temperature ensembles using the RPA, (ii) calculate total energies of the other functionals for these "RPA" distributed configurations, and (iii) rank the functionals based on the energy difference (ΔE) to the RPA. Calculating the ensembles using some approximate functional is also possible, but the explored structural phase space might not overlap with the correct higher accuracy RPA, biasing the final result.

This strategy is, in principle, straightforward, but cumbersome. With DFT, we can easily perform finite temperature simulations, yielding appropriately weighted microstates. In the absence of forces, this is rather difficult to do for high level methods, although energies alone suffice to perform Monte Carlo (MC) or hybrid Monte Carlo simulations [20]. To achieve high sampling efficiency, the MC moves need to be chosen judiciously, which is not a simple matter for systems where the dynamics is not well understood. Finite temperature molecular dynamics (MD) are conceptionally simpler, and this strategy has become possible with our recent implementation of forces within the RPA for solid state systems [21]. It is also the main strategy we adopt here to generate ensembles. How to obtain the RPA forces is elaborated in detail in Ref. [21]. In short, the key point is the insight that the first derivative of this functional with respect to the Green's function is just the self-energy (Σ) in the *GW* approximation [22]. The relation between the RPA

correlation energy and the *GW* self-energy is analogous to the relation between the Kohn-Sham energy and its potential [23].

As an example of the predictive power of this method, the experimentally difficult to determine atomic structure of hybrid perovskites is studied. This is a unique class of materials based on an organic donor, methyl-ammonium (MA) $CH_3NH_3^+$, trapped inside a $(PbI_3)^-$ perovskite cage [24]. These materials have reached a light to electricity conversion efficiency of 20% [25], approaching that of silicon based solar cells, but they can be produced cheaply using wet chemistry. Describing these systems with high precision using first principles methods is very challenging. Instabilities in the perovskite cage are often only well captured if band gap related properties and lattice constants are accurately described. In fact, the MA molecules rotate even at room temperature [26,27], coupled to concerted movements of PbI octahedra [28]. The situation is complicated by the molecules being allegedly bonded to the cage by van der Waals (vdW) interactions [29,30].

To study the influence of the specific functional on the ordering of the molecules and octahedra in MAPbI$_3$ we have performed long finite temperature parallel tempering *ab initio* MD calculations of $2 \times 2 \times 2$ pseudocubic supercells, containing 8 molecules in PbI$_3$ cages. The calculations were performed using the VASP code [31] with a plane-wave basis and the projector augmented wave (PAW) method [32] neglecting spin orbit coupling (details in Supplemental Material [33]). Similarly, as in Ref. [28], the orientations of all molecules inside the cages as well as the orientation of the PbI octahedra are extracted in polar coordinates $\{\phi, \theta\}$ as a function of time. These data are used to create the polar distributions of Fig. 1. The ordering of the molecules is highly dependent on the functional and

even amongst vdW functionals there are significant differences. The face-diagonal orientations are favored by PBEsol and TS, while optPBE and PBE-D3 favor the principle axes. The room-diagonal orientation is avoided by all considered functionals. The calculated relaxation times of the molecules are shown in Fig. 1(a) as white numbers. Most functionals give values in the order of picoseconds (ps) at room temperature in reasonable agreement with experiment [26]. The arrangement of the octahedra can be seen in Fig. 1(b). With PBEsol at 250 K the octahedra can tilt and rotate, but without a strong directional preference. For the other functionals, the octahedra tend to form zigzag patterns (or clockwise anticlockwise patterns), and avoid orientations exactly along the principle axis. At 400 K the orientations are more smeared out. Based on these results alone, it is very difficult to determine which density functional accurately describes the material.

To judge the quality of each functional, a well-distributed RPA canonical ensemble of structures is constructed. Because of the computational cost of the RPA, small $\sqrt{2} \times \sqrt{2}$ cubic cells are used. This cell allows for different orientations of neighboring MA molecules [see Fig. 2(a)]. As a starting point, we generate independent configurations using long DFT simulations at our desired temperature of 400 K. The internal geometry of the MA molecule is kept rigid allowing for a time step of up to 10 fs. For DFT simulations, results for rigid MA molecules are practically indistinguishable from unconstrained MD for the orientation of the molecules and dynamics of the cage. From this trajectory, 16 practically independent configurations are picked and finite temperature RPA-MD simulations at 400 K are performed. From the last 1.3 ps of each trajectory, 32 equally spaced configurations are picked yielding 512 partly correlated configurations. These form the RPA ensemble.
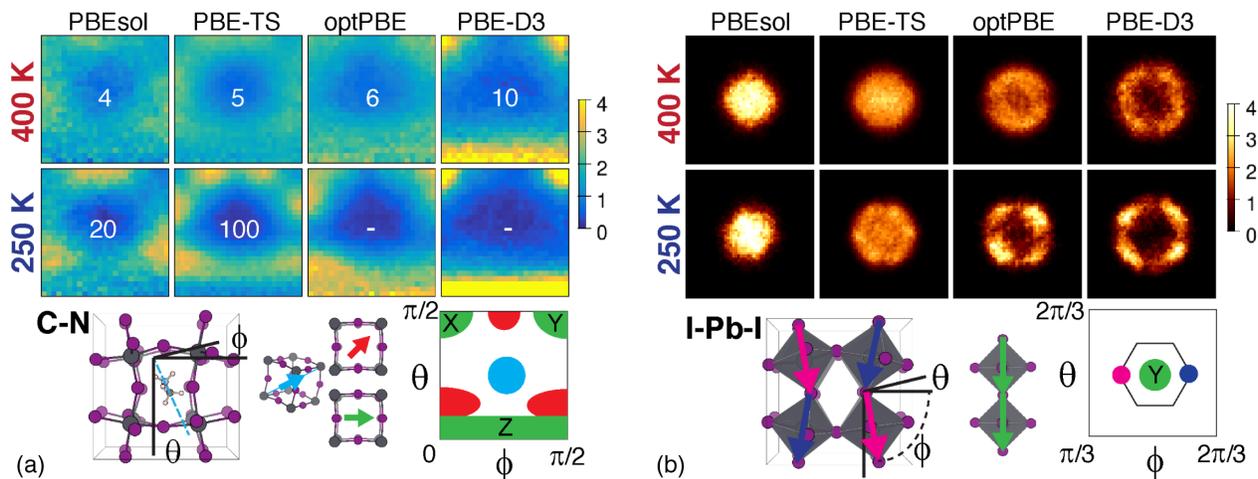

FIG. 1. Polar distribution of the orientation $\{\phi, \theta\}$ of the (a) molecules and (b) PbI octahedra (represented by C-N and I-I connecting vectors, respectively) in a $2 \times 2 \times 2$ supercell as function of temperature and the XC functional. For the molecules, cubic symmetry is applied to downfold the full polar distribution into a single octant. The molecular relaxation times are shown (in ps) by the white numbers. The legends below illustrate the orientations (a) of the MA molecules in the cage (principle axes, green; face diagonal, red; and room diagonal, blue) and (b) of the I-I connecting vectors (note the different values for the range of the angles).
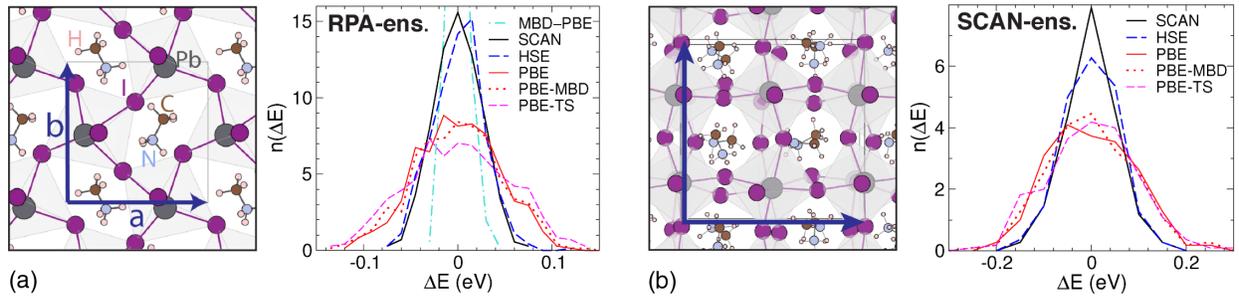
FIG. 2.   (a) Snapshot of the RPA-MD simulation of MAPbI$_3$, the $\sqrt{2} \times \sqrt{2}$ cell is indicated by the a and b lattice vectors. The histogram presents the energy difference between the RPA and selected density functionals for the RPA ensemble. In order to center the histograms the mean differences have been subtracted. The x axis shows the energy difference (in eV), whereas the y axis indicates the probability of finding a certain energy difference. (b) Same as in (a) but for the SCAN ensemble buildup with the larger $2 \times 2 \times 2$ cell. Additionally, the difference between the PBE-MBD and PBE energies [MBD − PBE] is shown in (a).

We evaluate the energy difference between each functional and the RPA for all configurations in the RPA ensemble and subtract the mean difference, the resulting number is hereafter referred to as ΔE. The histograms of ΔE are shown in Fig. 2(a). The curves follow roughly a Gaussian distribution. In order to rank the functionals, we have calculated the square root of the variance of the energy differences (σ) for many functionals and collect the values in Table I. Of all the considered functionals the HSE, HSE-D3, and SCAN functionals perform best, whereas the PBE functional is noticeably worse. To confirm that the present results carry over to larger supercells and that the short simulation times did not bias the results, we performed MD simulations for a larger $2 \times 2 \times 2$ supercell using the best and worst preforming functional (SCAN and PBEsol). For both functionals an ensemble of structures was made [33].

As before, the energy differences to the RPA are calculated, the σ values tabulated in Table I and the results for SCAN are plotted in Fig. 2(b). The SCAN and PBEsol ensemble yield a very similar ranking as the RPA ensemble, although for the PBE derived functionals the order varies somewhat between the ensembles, with the PBEsol ensemble clearly deviating more significantly from the RPA ensemble. For all ensembles, though, the SCAN functional is suggested to be the best choice. This indicates that using ensembles created with approximate density functionals is a cheap and viable alternative, if the

generation of ensembles using the high level functionals is not possible.

One intriguing result is that the approximate vdW corrected functionals do not improve systematically upon PBE. The many body dispersion (MBD) [39] correction of Tkatchenko *et al.* gives a hint why this is so. In MAPbI$_3$ the MBD corrections are small as indicated by the histogram for the energy differences [MBD − PBE, dashed dotted line in Fig. 2(a)]. The curve exhibits a very narrow peak with $\sigma = 10$ meV, whereas σ is approximately 30 meV between PBE and other dispersion corrected schemes for the RPA ensemble. This implies that the MBD corrections only add a constant shift to the energy but vary little between configurations. The MBD correction is based on the RPA total energy expression for the correlation energy. However, instead of calculating the polarizability from first principles, it assumes atomic dipole oscillators at each atomic site. Since the PbI$_3$ cage has a small optical gap and screens excitations fairly well [43], dispersion interactions are seemingly strongly washed out and show little structure dependence in MBD. The more conventional vdW corrections, such as TS and D3, which do not account for any screening of the cage, perform worse. For HSE-D3, the damping function in D3 is much stronger than for PBE, reducing the vdW interactions and the structure dependence. This explains why HSE-D3 is within the statistical error bars identical to the parent functional HSE. Since MBD is by construction

TABLE I.   Considered functionals, and the square root of the variance (σ, in meV) of the energy difference between the RPA and the functionals in the RPA|SCAN|PBEsol-ensemble, respectively.

SCAN	Strongly Constrained and Appropriately Normed [36]	25	53	67
HSE-D3	HSE [37] + Grimme D3 [38]	27	55	68
HSE	Heyd-Scuseria-Ernzerhof [37]	27	57	69
PBE	Perdew-Burke-Ernzerhof [14]	44	87	104
PBE-MBD	Many body dispersion [39]	47	89	109
PBE-D3	Grimme D3, vdW corrections [38]	51	88	111
PBE-TS	Tkatchenko-Scheffler, vdW corrections [40]	54	90	109
optPBE	vdW DFT functional based on PBE [41]	54	95	104
PBEsol	Revised PBE for solids [42]	55	96	106

more accurate than pairwise vdW interactions, the virtually perfect agreement between PBE and PBE-MBD shows that standard pairwise long range dispersion interactions should not be included for MAPbI$_3$ and, presumably, most other hybrid perovskites.

The best means to reduce the error of the semilocal functionals is to include exact exchange or use the SCAN functional. The SCAN functional fulfills all known constraints for DFT functionals and depends not only on the density but also on the kinetic energy density [36]. We believe that the main reason for the improved performance of these functionals is a more balanced description of the cage instabilities. The close agreement between HSE and SCAN is in line with a recent publication of Sun *et al.* indicating that instabilities of perovskites are described equally well by hybrid functionals and SCAN [44]. Remarkably, in the present case, the SCAN functional performs even slightly better than HSE. This is most likely related to the SCAN functional, including short range dispersion effects, which conventional hybrid functionals do not account for. We conclude that SCAN is the most suitable functional to study the atomic structure of hybrid perovskite materials.

Last, we use the SCAN functional to study the tetragonal distortion ($c/a > 1.0$) that is experimentally observed up to and about room temperature [45]. To show that this distortion is preferred at low temperatures, we performed parallel tempering MD simulations for different c/a ratios at a fixed volume ($V = 245$ Å^3 per formula unit) and calculate the average stress tensor σ. The results are shown in Fig. 3(a). For a stable cubic structure, one would expect that, as one elongates the cell in the c direction at fixed volume, a restoring stress opposing the distortion should develop [$\sigma_z - (\sigma_x + \sigma_y)/2 < 0$]. However, at the lowest temperature and $c/a = 1.01$, the MD suggests that even further expanding into the c direction is preferable, reaching a stable point at $c/a = 1.02$. Even at $T = 292$ K the system is soft, with a very small restoring stress that grows roughly quadratic with displacement. The conventional linear elastic behavior is only observed at higher temperatures. In agreement with experiment we therefore predict a tetragonal instability at low temperatures. Considering the small supercell, one certainly cannot expect perfect agreement with experiment for the transition temperature or c/a ratio. We note that the PBEsol functional, the only other functional we tested, does not yield a similar instability.

To study the origin of the distortion, Fig. 3(b) shows the polar distributions of the MA molecules and the PbI octahedra. In the tetragonal cell, the SCAN functional aligns the molecules along the principle axes. The fraction of the molecules aligned along the a, b, and c axes is indicated in white. At low temperatures, the molecules prefer to lie in the (a, b) plane, which is shorter than the c axis ($c/a = 1.01$). As the temperature rises this preference becomes less pronounced. This suggests that the molecules act like springs and contract the lattice along the direction
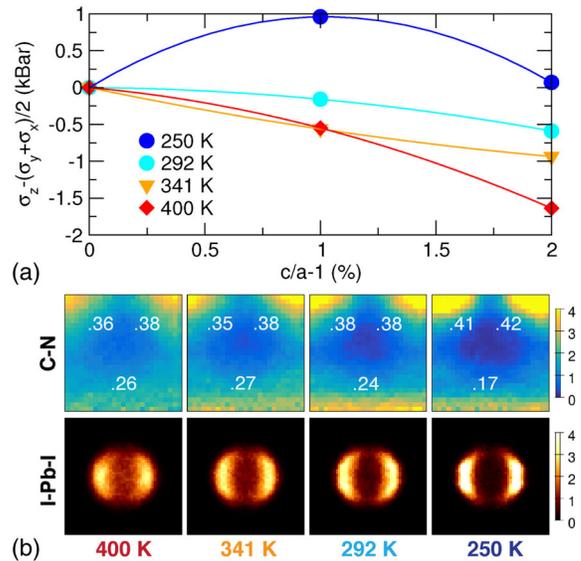

FIG. 3. (a) Stress difference in the tetragonal system as a function of c/a ratio and temperature. (b) Polar distributions of the MA molecules and PbI octahedra orientations obtained with $c/a = 1.01$. Only I-I connecting vectors in the ab-plane are included. Fraction of the molecules aligned along the a, b, and c axes are indicated by white numbers.

they are aligned to. Clearly this hints at a complex interplay between cage instabilities and molecule-cage interactions.

In summary, we have presented a general and predictive scheme to determine the optimal density functional for a specific materials problem. The strategy relies on the random phase approximation, which is very accurate for condensed matter systems. By creating RPA ensembles and evaluating the variance between the RPA and a selection of DFT functionals, one can choose the best functional for large scale structure predictions. This avoids the common trial and error approach where different functionals are explored until agreement with experiment is obtained. For the modeling of hybrid perovskites, the important finding of the present work is that the SCAN functional [36] works best compared to the RPA. Furthermore, many vdW functionals perform disappointingly, which is in contrast to previous suggestions that vdW interactions need to be included in these materials [29,30]. We have explained this by the strong screening of the cage. From a materials point of view, our present study sheds new light on the tetragonal instability of MAPbI$_3$ (Fig. 3). We confirm that the MA molecules are dynamically reorienting even in the tetragonal phase and show that they prefer to be aligned along the two shorter axes, a and b, thereby contracting the lattice along these directions.

[*] menno.bokdam@univie.ac.at
[†] georg.kresse@univie.ac.at

[1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[3] L. Hedin, Phys. Rev. **139**, A796 (1965).
[4] A. Marini, P. García-González, and A. Rubio, Phys. Rev. Lett. **96**, 136404 (2006).
[5] J. Harl and G. Kresse, Phys. Rev. Lett. **103**, 056401 (2009).
[6] S. Lebègue, J. Harl, T. Gould, J. G. Ángyán, G. Kresse, and J. F. Dobson, Phys. Rev. Lett. **105**, 196401 (2010).
[7] L. Schimka, J. Harl, A. Stroppa, A. Grüneis, M. Marsman, F. Mittendorfer, and G. Kresse, Nat. Mater. **9**, 741 (2010).
[8] T. Olsen, J. Yan, J. J. Mortensen, and K. S. Thygesen, Phys. Rev. Lett. **107**, 156401 (2011).
[9] T. Björkman, A. Gulans, A. V. Krasheninnikov, and R. M. Nieminen, Phys. Rev. Lett. **108**, 235502 (2012).
[10] X. Ren, P. Rinke, C. Joas, and M. Scheffler, J. Mater. Sci. **47**, 7447 (2012).
[11] L. Schimka, R. Gaudoin, J. Klimeš, M. Marsman, and G. Kresse, Phys. Rev. B **87**, 214102 (2013).
[12] M. Macher, J. Klimeš, C. Franchini, and G. Kresse, J. Chem. Phys. **140**, 084502 (2014).
[13] M. Kaltak, J. Klimeš, and G. Kresse, Phys. Rev. B **90**, 054115 (2014).
[14] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[15] P. Nozières and D. Pines, Phys. Rev. **109**, 741 (1958).
[16] D. C. Langreth and J. P. Perdew, Phys. Rev. B **15**, 2884 (1977).
[17] T. Miyake, F. Aryasetiawan, T. Kotani, M. van Schilfgaarde, M. Usuda, and K. Terakura, Phys. Rev. B **66**, 245103 (2002).
[18] M. Fuchs, Y.-M. Niquet, X. Gonze, and K. Burke, J. Chem. Phys. **122**, 094116 (2005).
[19] F. Furche, J. Chem. Phys. **129**, 114105 (2008).
[20] N. Akira, T. Tetsuya, and S. Motoyuki, Chem. Lett. **38**, 976 (2009).
[21] B. Ramberger, T. Schäfer, and G. Kresse, Phys. Rev. Lett. **118**, 106403 (2017).
[22] N. E. Dahlen, R. van Leeuwen, and U. von Barth, Phys. Rev. A **73**, 012511 (2006).
[23] We note that the computational complexity of RPA calculations including forces is linear in the number of k points and cubic in the number of plane waves. This is similar to DFT, however, the prefactors are typically a factor 10^2 (energies) to 10^3 (forces) larger than for standard DFT calculations. That calculating forces is about 5–10 times more expensive than calculating total energies is related to the computational costs for evaluating the self-energy $\Sigma(it)$. The steep cost means that one has to restrict simulations to fairly small supercells and few k points.
[24] M. A. Green, A. Ho-Baillie, and H. J. Snaith, Nat. Photonics **8**, 506 (2014).
[25] NREL, *Research Cell Efficiency Records* (National Renewable Energy Lab, Golden, CO, 2016).
[26] A. Poglitsch and D. Weber, J. Chem. Phys. **87**, 6373 (1987).
[27] F. Brivio, J. M. Frost, J. M. Skelton, A. J. Jackson, O. J. Weber, M. T. Weller, A. R. Goñi, A. M. A. Leguy, P. R. F. Barnes, and A. Walsh, Phys. Rev. B **92**, 144308 (2015); M. A. Carignano, A. Kachmar, and J. Hutter, J. Phys. Chem. C **119**, 8991 (2015); C. Quarti, E. Mosconi, and F. De Angelis, Phys. Chem. Chem. Phys. **17**, 9394 (2015); A. Mattoni, A. Filippetti, M. I. Saba, and P. Delugas, J. Phys. Chem. C **119**, 17421 (2015).
[28] J. Lahnsteiner, G. Kresse, A. Kumar, D. D. Sarma, C. Franchini, and M. Bokdam, Phys. Rev. B **94**, 214114 (2016).
[29] D. A. Egger and L. Kronik, J. Phys. Chem. Lett. **5**, 2728 (2014).
[30] C. Motta, F. El-Mellouhi, S. Kais, N. Tabet, F. Alharbi, and S. Sanvito, Nat. Commun. **6**, 7026 (2015).
[31] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993); G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996); G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
[32] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[33] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.119.145501 for a detailed description of the applied computational techniques and settings, which includes Refs. [34] and [35].
[34] Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **314**, 141 (1999).
[35] C. C. Stoumpos, C. D. Malliakas, and M. G. Kanatzidis, Inorg. Chem. **52**, 9019 (2013).
[36] J. Sun, A. Ruzsinszky, and J. P. Perdew, Phys. Rev. Lett. **115**, 036402 (2015).
[37] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, J. Chem. Phys. **125**, 224106 (2006).
[38] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).
[39] A. Tkatchenko, R. A. DiStasio, R. Car, and M. Scheffler, Phys. Rev. Lett. **108**, 236402 (2012).
[40] A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102**, 073005 (2009).
[41] J. Klimeš, D. R. Bowler, and A. Michaelides, Phys. Rev. B **83**, 195131 (2011).
[42] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, Phys. Rev. Lett. **100**, 136406 (2008).
[43] M. Bokdam, T. Sander, A. Stroppa, S. Picozzi, D. D. Sarma, C. Franchini, and G. Kresse, Sci. Rep. **6**, 28618 (2016).
[44] J. Sun, R. C. Remsing, Y. Zhang, Z. Sun, A. Ruzsinszky, H. Peng, Z. Yang, A. Paul, U. Waghmare, X. Wu, M. L. Klein, and J. P. Perdew, Nat. Chem. **8**, 831 (2016).
[45] P. Whitfield, N. Herron, W. E. Guise, K. Page, Y. Q. Cheng, I. Milas, and M. K. Crawford, Sci. Rep. **6**, 35685 (2016).