

1. For total energies, FHI-AIMS with NAOs is able to reach more accurate results than NWChem with GTOs for basis sets of comparable size (e.g., “tier4” vs aug-cc-pV5Z).
2. For atomization energies, both NAOs and GTOs benefit from error cancellation to some extent. Such a cancellation is however much stronger for GTOs where the RMSD is lowered by a factor of 4–8 in most cases, whereas for NAOs, it is lowered only by a factor of 1.5–2. In both cases, the cancellation is more marked for the smallest bases. Despite the smaller cancellations, NAOs are still closer to the converged limit than GTOs, for comparable basis sets.
3. The two functionals considered (GGA-PBE and hybrid-PBE0) yield very similar results, and we therefore assume that our conclusions concerning the accuracy of the different approaches (NAOs and GTOs) will hold also for other functionals of the same type.
4. Dipole moments can be considered accurate if deviations are below 0.01 D.^{24,86} Only the largest basis sets in NAO and GTO used in our calculations achieve this target on average, but even such basis sets have outliers with errors close to 0.1 D.
5. Due to the cumbersome convergence of periodic DFT codes with respect to the box size, we did not include APW+lo results for the entire test set of molecules. Nevertheless, for a small subset of molecules for which the limit of the box size was reached, we found atomization energies with errors of about 1 kcal/mol (see the SI⁷⁹). Our experience suggests that it is technically challenging for APW-based codes to reach accuracies below 1 kcal/mol on atomization energies.

As a final remark, we stress that for a few atoms (Li, Be, Na, Mg), the aug-cc-pV5Z basis is not available, as previously mentioned in the atomic calculations part. Had it been available, GTOs might have yielded somewhat higher precision for the affected systems than in our benchmarks. However, considering the large size of our sample, the fact that only a few atoms in a molecule are affected, and the small improvement that can be inferred from the atomic calculations, our main conclusions still hold. On the other hand, such a de facto limitation of the availability of GTO basis sets illustrates how demanding it is to generate such basis sets. In contrast, MWs and NAOs are much less affected by such a limitation.

Considering that several families of GTO basis sets are available, we have also performed an additional set of calculations for the GGA-PBE functional in order to compare the performance across such sets. The results are displayed in Figure 3. In particular, we have considered the Pople basis sets 6-31+G** and 6-311++G(3df,3pd),^{87–90} the pc-2 and pc-3 basis sets^{91–94} (the augmented analogues were also considered, but here we encountered ill-conditioning problems), and finally def2-TZVPD and def2-QZVPD.⁹⁵ This is a selection of the basis sets considered by Mardirossian and Head-Gordon for the development of the B97M-V and ω B97M-V functionals.^{17,18} For comparison, one should keep in mind that pc-2 and def2-TZVPD, as well as the large Pople set, are comparable in size to aug-cc-pVTZ, while pc-3 and def2-QZVPD are comparable to aug-cc-pVQZ.

Although the detailed analysis of the performance of each basis set family is a relevant issue, it is outside of the scope of this Letter and will be considered in another study. Suffice it

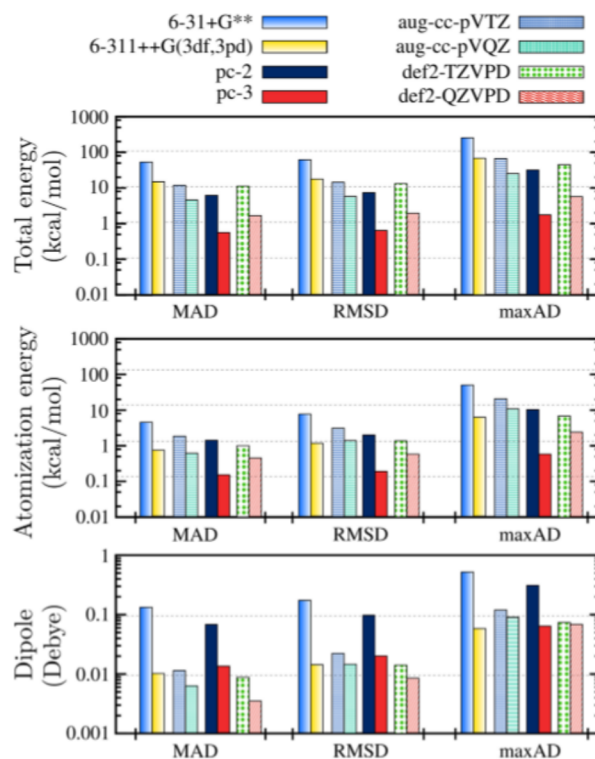


Figure 3. Statistics for the GGA-PBE total energy, atomization energy, and electrostatic dipole moment for the set of 211 molecules with respect to highly accurate values obtained using MRCHEM. Displayed values are the MAD, RMSD, and maxAD. The following basis sets have been considered: 6-31+G**, 6-311++G(3df,3pd), pc-2, pc-3, def2-TZVPD, def2-QZVPD, aug-cc-pVTZ, and aug-cc-pVQZ.

here to say that, among the considered basis sets, only pc-3 stands out when it comes to energy calculations, actually outperforming the much larger aug-cc-pV5Z basis and being competitive with NAOs. However, due to the lack of diffuse functions, the pc-*n* series suffers when dealing with dipole moments, where pc-3 performs only on par with the smaller aug-cc-pVTZ basis. The Pople sets are those that benefit the most from error cancellation when atomization energies are considered, while the def2-QZVPD basis yields better dipole moments, which is an indication of higher versatility in the basis. In practical calculations, the relatively weak performance of this broad range of production-quality GTO basis sets for DFT-based atomization energies can translate into real problems for subtle energy differences, for example, conformational energy differences of hydrogen-bonded systems.⁹⁶

To the best of our knowledge, this work presents the most accurate atomization energies calculated to date for a large benchmark set of molecules. We conclude that moderately sized GTO basis sets, frequently used in quantum chemistry applications, suffer from average total energy errors much larger than 10 kcal/mol, and while very large GTO basis sets yield the desired accuracy on average, there are still significant outliers. Moreover, it may not always be feasible to employ such basis sets for systems much larger or chemically more diverse than those included in this study. In addition to cost, ill-conditioning can be a practical limitation for large GTO basis sets. Applying large, high-accuracy GTO basis sets to all-electron calculations for elements beyond the lightest few (e.g., $Z = 1–18$ as covered here) is also not straightforward. In contrast, there is no