Zitierweise: *Angew. Chem. Int. Ed.* **2022,** *61,* e202205735
Internationale Ausgabe: doi.org/10.1002/anie.202205735
Deutsche Ausgabe: doi.org/10.1002/ange.202205735

Density Functional Theory

Best-Practice DFT Protocols for Basic Molecular Computational Chemistry**

Markus Bursch, Jan-Michael Mewes,* Andreas Hansen,* and Stefan Grimme**

Abstract: Nowadays, many chemical investigations are supported by routine calculations of molecular structures, reaction energies, barrier heights, and spectroscopic properties. The lion's share of these quantum-chemical calculations applies density functional theory (DFT) evaluated in atomic-orbital basis sets. This work provides best-practice guidance on the numerous methodological and technical aspects of DFT calculations in three parts: Firstly, we set the stage and introduce a step-by-step decision tree to choose a computational protocol that models the experiment as closely as possible. Secondly, we present a recommendation matrix to guide the choice of functional and basis set depending on the task at hand. A particular focus is on achieving an optimal balance between accuracy, robustness, and efficiency through multi-level approaches. Finally, we discuss selected representative examples to illustrate the recommended protocols and the effect of methodological choices.

1. Introduction

Chemistry and chemical synthesis are indispensable tools for humankind in addressing the most urgent current and future challenges, such as efficient energy storage and conversion, sustainable food supply, and affordable medication and health care. In all of these examples, a rational design of molecules and materials, e.g., new catalysts, electrolytes for batteries, hosts and emitters for organic electronics, and new drugs, takes a central role. Here, it is crucial to understand matter at the atomic and electronic-structure level, which is possible only through chemical synthesis, spectroscopy, and quantum chemical calculation. The latter of these three, computational chemistry, and specifically Kohn–Sham density functional theory (DFT), has firmly consolidated its position as a third workhorse besides synthesis and spectroscopy in recent decades. We argue that the general importance of computational chemistry and DFT, in particular, stems from its outstanding effort-to-insight and cost-to-accuracy ratios compared to related approaches, in other words, their efficiency (vide infra).

From a more fundamental perspective, DFT is a formally exact but practically empirical "first-principles" electronic-structure approach to solve the fermionic many-electron problem that underlies most of chemistry and large parts of biology and physics. When applied together with a mixed quantum-classical treatment for the nuclei using molecular dynamics (MD) or harmonic approximations for the potential energy surface (PES), DFT can address many problems in (bio)chemistry and physics with sufficient accuracy to derive meaningful insight. The subtle theoretical and technical details of DFT are well understood and have recently been discussed in an extensive open discussion type of review.[1] This open review addresses all more detailed and theoretical questions about DFT which are raised here but whose answer is beyond the scope of this practically oriented work. For a recent discussion of the grand challenges in theoretical and computational chemistry, see Refs. [2,3].

DFT offers an excellent compromise between required computation time and the quality of the results in comparison to the alternatives, which are less accurate and robust but much faster semi-empirical quantum mechanics[4–6] (often termed SQM) on the one hand, and on the other more accurate and robust but slower wavefunction-theory-based approaches such as coupled-cluster, (see Figure 1). Efficient approximations in the coupled-cluster framework such as DLPNO-CCSD(T)[7,8] can also be a noteworthy alternative to high-level DFT but are beyond the scope of this work. Further, human and power resources are spared. Moreover, and just as important, DFT can be considered a robust theory in that a breakdown in the form of entirely wrong results is scarce, even when applied to challenging molecules or exotic chemistry. This contrasts semi-empirical quantum mechanics and other, even empirical approaches,

[*] Dr. M. Bursch
Max-Planck-Institut für Kohlenforschung
Kaiser-Wilhelm-Platz 1
45470 Mülheim an der Ruhr (Germany)
E-mail: bursch@kofo.mpg.de

Dr. J.-M. Mewes, Dr. A. Hansen, Prof. Dr. S. Grimme
Mulliken Center for Theoretical Chemistry
Institut für Physikalische und Theoretische Chemie
Universität Bonn, Beringstraße 4, 53115 Bonn (Germany)
E-mail: janmewes@janmewes.de
hansen@thch.uni-bonn.de
grimme@thch.uni-bonn.de

[**] A previous version of this manuscript has been deposited on a preprint server (https://doi.org/10.26434/chemrxiv-2022-n304h).

which require much more careful sanity-checking by comparison to DFT. These properties grant DFT the role of a black-box method that non-experts can apply to many chemical problems. Such applications typically involve testing the plausibility of structures and reaction mechanisms suggested by synthetic chemists or visualizing the frontier orbitals and HOMO and LUMO energies to interpret electrochemical experiments in materials science.

Nevertheless, the choice of a reasonable, efficient yet accurate quantum chemistry treatment for a wide range of chemical problems is still a challenging task, even for experienced computational chemists. This is also due to a vast number of available method combinations that have been developed and presented in recent years. Already for the fundamental choice of the density functional/atomic orbital basis set combination, hundreds or even thousands of combinations are possible in typical programs, many of which are in common use. While this is uncritical for a small molecule where one can use a "sledgehammer to crack a nut" (that is, use expensive double-hybrid density functionals and large atomic orbital basis sets), the treatment of systems with 50–100 atoms or many relevant low-energy conformers demands critical compromises in methodological choices in order to keep the computational cost manageable.

Unfortunately, in some QM programs, the default methods are outdated, which may tempt inexperienced users to apply no longer recommended methods to circumvent these complications. A prominent example is the popular B3LYP[9,10]/6-31G* functional/atomic orbital basis set combination that is still frequently used even though it is known to perform poorly even for simple cases.[11–13] The knowledge that B3LYP/6-31G* suffers from severe inherent errors, namely missing London dispersion effects ("over-repulsiveness") and strong basis set superposition error (BSSE), seems to "diffuse" slowly from the theoretical to the computational chemist community.

Accordingly, educative tutorials that address these weaknesses and show alternatives such as the work by Peverati and Morgante[14] or Mobley[15] are appreciated, providing valuable general guidance in the field of computational chemistry. In the last 10–20 years, the availability of better functionals,[16,17] standardized dispersion corrections,[18] and empirical corrections for BSSE[19,20] made B3LYP/6-31G* computations obsolete. Today, much more accurate, robust, and sometimes even computationally cheaper alternatives exist, e.g., in the form of composite methods like B3LYP-3c,[21] r^{2}SCAN-3c,[22] B3LYP-D3-DCP,[23] or B97M-V/def2-SVPD/DFT-C[20] to name just a few. Such methods use new developments to eliminate the systematic errors of B3LYP/6-31G* without increasing the computational cost. Figure 1 illustrates the relative computational demands of the discussed approaches, while the examples in Sections 4.2 and 4.3 provide a clear demonstration of the shortcomings of B3LYP/6-31G*, and how they can be mitigated. The main goal of this work is to introduce these new methods and developments to interested non-experts and chemists with some background in theory.

Markus Bursch received his PhD in 2021 under the supervision of Stefan Grimme on the evaluation and application of efficient quantum chemical methods for inorganic molecular chemistry. He is currently a postdoc in Frank Neese's group at the Max-Planck Institut für Kohlenforschung, Mülheim an der Ruhr (Germany) and is working on quantum chemical challenges in inorganic main-group and organometallic chemistry. He further works on efficient methods and work-flows to enhance the description of structural, energetic, and spectroscopic properties of large inorganic molecules.

Jan-Michael Mewes studied chemistry in Frankfurt (Germany). In 2011, he followed his supervisor, Andreas Dreuw, to Heidelberg for a PhD in condensed-phase photochemistry. After defending his thesis in 2015 and founding a family in 2016, Jan joined Peter Schwerdtfeger's group in Auckland (NZ) as Feodor-Lynen Postdoc to explore periodic trends. In 2019, he obtained a Returning Scholarship (AvH) to join the group of Stefan Grimme (Bonn). Jan has since worked as a senior researcher to improve DFT for ground and excited states and develops lanthanide-based OLEDs.

Andreas Hansen received his PhD in 2012 on the development of local open-shell correlation methods under the supervision of Frank Neese. Since then, he has managed, taught, and conducted research as an academic chief counselor at the Mulliken Center of Theoretical Chemistry of the University of Bonn in collaboration with Stefan Grimme. His research focuses on developing quantum chemical methods and workflows, with special focus on compiling comprehensive benchmark sets and protocols for the reliable calculation of reference values for diverse chemical properties.

Stefan Grimme studied Chemistry and finished his PhD in 1991 in Physical Chemistry. He habilitated in Theoretical Chemistry in the group of Sigrid Peyerimhoff. In 2000, he got the C4 chair for Theoretical Organic Chemistry at the University of Munster. In 2011, he accepted an offer as the head of the Mulliken Center for Theoretical Chemistry at the University of Bonn. He is the recipient of the Schrodinger medal and the "Gottfried Wilhelm Leibniz-Preis". His main research interests are the development and application of quantum chemical methods for large molecules, density functional theory, and noncovalent interactions.

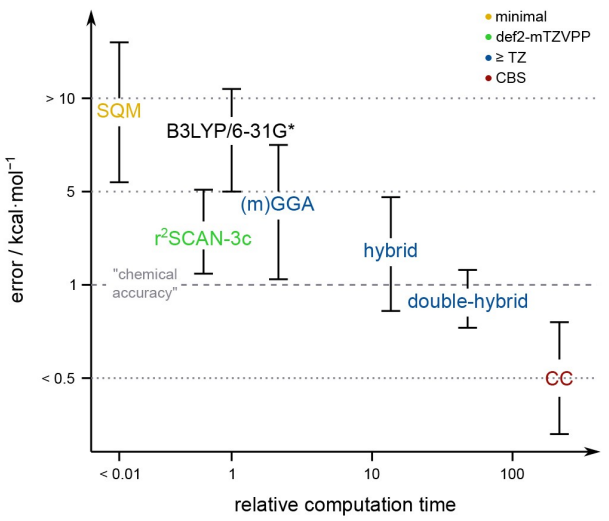

Figure 1. Accuracy in typical thermochemical applications vs. computation time (logarithmic scale) for common quantum chemistry methods. Wavefunction-based coupled-cluster (CC) methods like CCSD(T) (possibly in combination with local approximations) provide benchmark-quality results of better than 1 kcal mol^{-1} for common chemical energy changes of almost any well-behaved single-reference system (for details see text). Lower rungs of the DFT hierarchy of methods from small basis set composite approaches[31–33] (e.g., r^{2}SCAN-3c[22]), to (m)GGA or hybrid functionals provide systematically improved results when coupled with large atomic orbital basis sets. The most sophisticated double-hybrid functionals often yield results close to a coupled-cluster reference level. See Section 2.5 for a more detailed discussion of the density functional classes. SQM = semi-empirical quantum mechanical method. CBS = complete basis set limit.

To this end, best practice recommendations are provided for the most common workflows encountered in typical applications, i.e., structure and ensemble determination, computation of reaction energies, barriers, free energies, and solvation effects are illustrated with a few typical examples or case studies. We will also briefly discuss the "embedding" of properly conducted DFT calculations into multi-level workflows to determine solvated and thermally averaged conformer, protomer, or tautomer ensembles. However, due to quantum chemistry's broad range and complexity, some interesting but less common topics like theoretical spectroscopy, excited states, periodic systems, or heavy element chemistry are not considered here or touched only very briefly. Another important area not covered here is the analysis and understanding of atomic and fragment interactions by, e.g., energy decomposition analysis,[24,25] or wavefunction composition, e.g., by orbital interactions in the natural bonding orbital (NBO)[26] framework.

The conclusions herein are based on more than 25 years of experience in the field of DFT and functional development, thermochemical benchmarking, as well as hundreds of collaborative chemical applications in mechanistic or spectroscopic studies. Our recommendations are mostly based on hard evidence: They rely on large-scale comparisons of approximate DFT results for a wide range of chemical properties with those from experiments or highly accurate and robust coupled-cluster theory, the so-called gold standard in computational chemistry (benchmarking). However, some conclusions concerning the performance of a model, basis set, or particular density functional are numerically and statistically challenging to quantify. Hence, our recommendations have a personal-experience-based flavor. In this context, we put more emphasis on the robustness of a method than its "peak performance" reflected, for example, in the ranking in standard thermochemical benchmark sets such as the GMTKN55.[27] This is because in our experience in predictive applications, robustness and reliability, that is, avoiding large and unexpected errors, is more important than getting the numbers right to the last kcal mol^{-1}.

It should be clear, however, that density functional theory methods in their typically applied forms are not free of approximations and therefore, even choosing a robust functional does not guarantee perfect results for any arbitrary system.[28–30]

Finally, we want to point out that in computational chemistry, the wording "thermochemistry" also includes the calculation of reaction barrier heights, although they might be considered separately under the common term chemical kinetics.

2. General Considerations

2.1. Decision-Making in Computational Chemistry

Defining a suitable set of theoretical methods is the key to accurately describing a chemical system. This includes not only the selection of a quantum chemistry model for the basic electronic structure but also the choice of an appropriate model system to describe the physico-chemical totality of the problem. Generally, these choices comprise a complex set of fundamental decisions. An exemplary flowchart that illustrates this decision-making and is applicable to a large part of typical quantum chemical applications is shown in Figure 2.

2.2. Electronic Structure

The possibly most fundamental aspect that decides if common DFT is applicable is whether the system under consideration is well-represented by a single-determinant wavefunction and thus has single-reference character, or whether multiple determinants are required and the system has multi-reference character. Luckily, the most common examples fall into the first category. These systems possess a single-reference electronic structure and are thus readily describable by the common DFT methods discussed in this work. This holds true in particular for diamagnetic closed-shell (organic) molecules, the vast majority of which possess single-reference character. The few exceptions, such as biradicals, often have low-lying triplet states, which can be checked with an unrestricted broken-symmetry[34,35] DFT calculation.

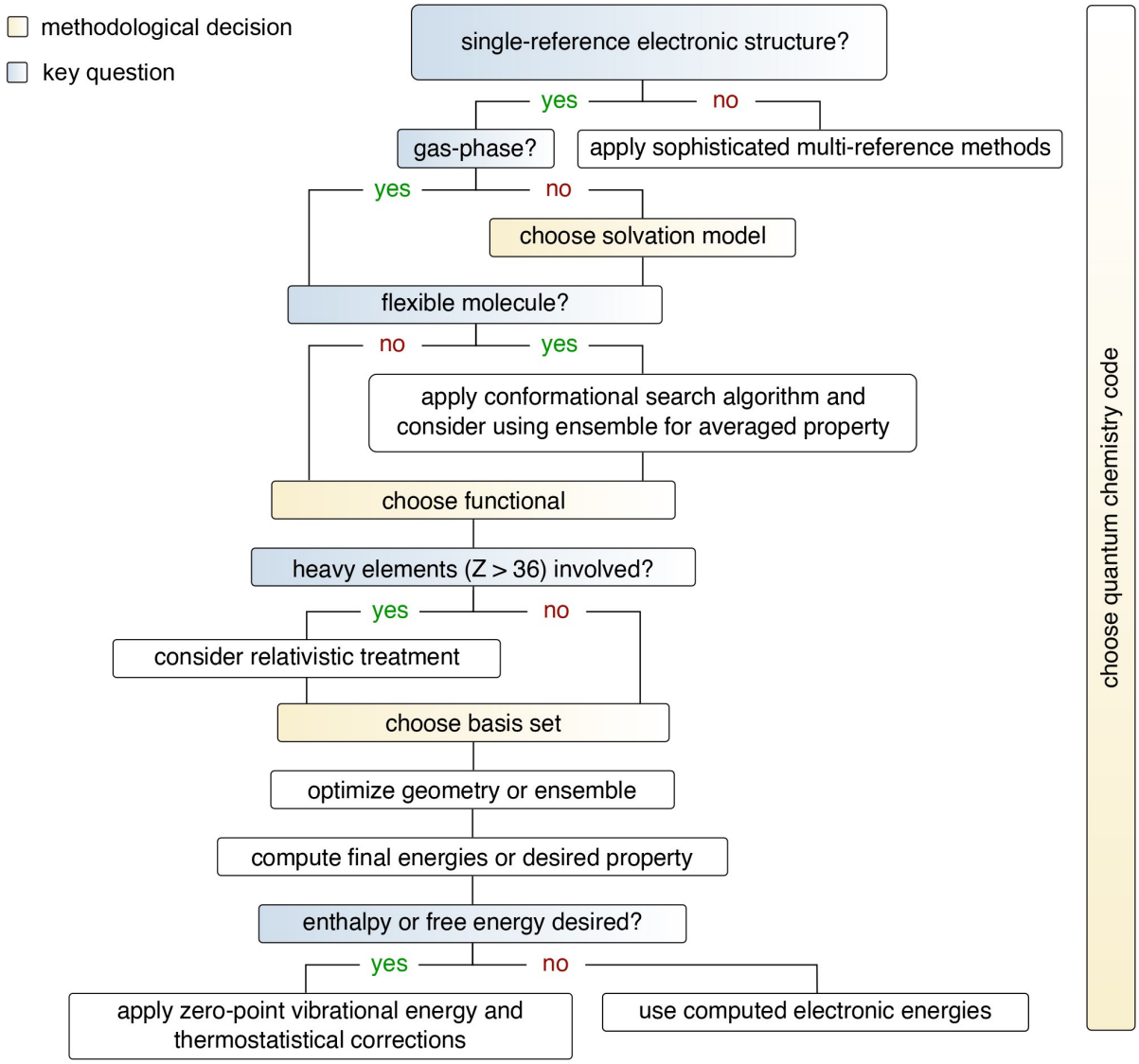

Figure 2. Conceptual flowchart of decision-making in elementary steps in typical molecular computational chemistry calculations.

Systems where multi-reference character should be expected are radicals, low band-gap systems (see below), transition states of open-shell dissociation processes, and transition metal complexes. Specifically, 3d metals with partially filled d-shells are prone to multi-reference situations since their ligand field is stronger than that of 4d and 5d metals. In all of these, multiple near-degenerate electronic configurations (with different orbital occupations) can be present, leading to complicated electronic structure. Accordingly, for these cases, it should be checked in advance whether a so-called multi-reference case is present. While this test is generally rather complicated, there are simple hints that allow an initial yet often sufficient assessment of possible multi-reference character. For example, a very small gap between the highest occupied and lowest unoccupied molecular orbitals, the so-called HOMO–LUMO gap, of <0.5–1 eV in a test GGA calculation (see below) and exceptionally slow self-consistent field convergence are first indications of unusual electronic complexity. If this first crude hint emerges, more sophisticated measures should be considered. Most available approaches are wave-function-theory-based[36–40] and thus limited in their applicability to small molecules. The alternative fractional-occupation-number-weighted density (FOD)[41–43] represents an easily applicable DFT-based estimate. Here, a moderate artificial increase in the electronic temperature can be used to populate and visualize low-lying, possibly problematic electronic states. With any indication of a significant multi-reference character, application of standard DFT methods is not recommended and experts for sophisticated multi-reference theory should be consulted. For a discussion of multi-reference vs. related multi-determinantal cases appearing in low-spin open-shell systems see Ref. [44].

Recommendations:

- Check for multi-reference character through simple indicators (HOMO–LUMO gap, fractional-occupation-number-weighted density).
- Be more careful with open-shell systems (low-spin in particular).
- Do not apply single-reference methods like DFT to multi-reference systems.

2.3. Solvation

The next fundamental question is that of the present state of aggregation or in which form of a substance mixture the molecule to be examined is present. The neighboring molecules in a solid or for a solute in solution can have drastic effects on the structure and properties of the entire system. Accordingly, for condensed-phase chemistry, a suitable solvent model should be applied in any case. The most common approach in a DFT context is to use continuum solvation models that include interaction of the molecule with the solvent implicitly via an effective potential in the Hamiltonian. This means that no actual solvent molecules are present in the calculation. Prominent representatives of this class include the conductor-like polarizable continuum model (CPCM),[45] the solvation model based on the molecular electron density (SMD),[46] the conductor-like screening model (COSMO),[47] the conductor-like screening model for real solvents (COSMO-RS),[48] and the direct conductor-like screening model for real solvents (DCOSMO-RS).[49] The mentioned methods differ in various aspects. CPCM and COSMO are purely electrostatic models, lacking contributions from cavity creation that cost energy in the solvent, and attractive van der Waals interactions with the solvent, which lead to substantial errors if the solvent-accessible surface area is changing significantly. SMD, COSMO-RS, and DCOSMO-RS include such contributions and are thus recommended (cf. Section 4.2). Nevertheless, it should be noted that COSMO-RS cannot be used in geometry optimizations or frequency calculations and has to be replaced by DCOSMO-RS for this purpose. Further noteworthy implicit solvation methods that can also be used with semi-empirical quantum mechanical and force-field methods are the generalized-Born model with solvent-accessible surface area (GBSA)[50,51] and the analytical linearized Poisson–Boltzmann model (ALPB).[52,53] Nevertheless, for specific cases, the inclusion of explicit solvent molecules may be necessary and implicit solvation models become insufficient.[54–62] In micro-solvation approaches, actual solvent molecules are placed at important, most strongly bound positions of a system.[63] Explicit solvent molecules should be included when they are strongly bound and/or are strongly involved in the chemical/physical process under consideration. This may be tested at a lower level of theory by exemplary geometry optimization and energy evaluation for a representative model system. However, explicit solvation also has its caveats as it can be very difficult to converge properties with the number of explicit solvent molecules. Moreover, the potential-energy surface of explicitly solvated systems is often flat and peppered with local minima of different solvent structures, making optimizations lengthy and tedious. Therefore, explicit solvation should be used with care. Neglecting solvation effects, specifically for polar or charged molecules, can result in large deviations in thermochemical calculations, and even fundamentally wrong electronic structures, e.g., for zwitterions.

Recommendations:

- Choose a model/state of aggregation close to the experiment.
- Apply implicit solvation models for a molecule in solution; best use physically complete models, such as COSMO-RS or SMD.
- Be careful with charged systems for which continuum models may be inaccurate (the higher the charge density, the more inaccurate).
- Consider explicit solvation if necessary.

2.4. Molecular Flexibility

Another important aspect is the structural flexibility of the system. For highly flexible structures, a molecular property, such as energy, nuclear magnetic resonance spectra, or optical rotation values may not be sufficiently described by a single structure. At finite temperatures, various conformers are populated and the overall property must be described as thermal average over the unique property values of each conformer. Accordingly, it is recommended to evaluate the flexibility and the accessibility of relevant conformers for any given system through an initial conformer search. This step should only be skipped in the most obvious cases (e.g., chlorobenzene, anthracene etc.) because in our experience, even rigid-looking molecules can have surprisingly many low-lying conformers.

The flexibility of a molecule may be roughly categorized by the number of conformers in an energy window of 3 kcal mol^{-1} with respect to the lowest energy conformer, which corresponds to five times the thermal energy at room temperature ($5RT$). Systems with only a few conformers ($\approx$1–3) in this window may be considered relatively rigid, those with dozens as intermediate cases, and with hundreds as very flexible. In any case, finding the overall lowest conformer (global minimum) in the given chemical environment is important,[64] and by no means a trivial task. For example, a conformer found in the solid, e.g., determined by X-ray crystallography, may not be the most favorable one in solution or in the gas phase.[65] This is another reason why an initial conformational search (with a solvent model) is strongly recommended even if X-ray structures are available. However, even for only medium-sized molecules (30–50 atoms), a sophisticated conformational search is not trivial: due to the vast size of the configuration space, the computational cost of the search is often prohibitive also at a pure DFT/(m)GGA level. Hence, multi-level approaches (see Section 3) involving efficient semi-empirical quantum mechanical or force-field (FF) methods are necessary. The

CREST[66]/*CENSO*[67] approach represents a valuable, easily applicable tool for semi-automated conformation sampling and subsequent energy ranking of conformer-rotamer ensembles (CREs). Further, the flexibility index given by the *CREST* program can be used as additional indicator for the molecular flexibility. Besides *CREST*, alternative, less general conformer generation procedures such as *Molassembler*,[68] *ConfGen*,[69] *MS-DOCK*,[70] *Frog2*,[71] *OMEGA*,[72] *RDkit*,[73] and others[74,75] are described in the literature.

Recommendations:

- Check for the role of structural flexibility/conformations by searching the conformational space.
- Apply automated conformation search algorithms, e.g., *CREST*.
- Try to find and verify the lowest energy conformer.
- Consider Boltzmann-averaged property calculations.

2.5. Choice of Functional

A critical issue in DFT is the choice of the exchange-correlation energy functional, often simply called functional. It aims to absorb the extremely complicated many-particle correlation and fermionic (exchange) effects into a seemingly simple but formally exact and theoretically existing mean-field electronic energy and potential. During the last decades, hundreds of different functionals have been constructed which vary in their conception, target application, and overall quality.[16,76] Accordingly, there are general-purpose as well as task-specific functionals, highly parameterized and fundamental first-principle-based ones. A detailed discussion of those aspects is beyond the scope of this work as we focus on the suitability for applications in thermochemical calculations. Perdew's "Jacob's ladder" represents the most prominent attempt to categorize density functionals based on their physical ingredients (Figure 3).[77] Here, functionals are ranked according to their degree of approximation as measured by the included electron density descriptors for the exchange-correlation term (occupied orbitals via the density, first derivative of the density, second derivative, occupied orbitals via Fock exchange, virtual orbitals via MP2, …) and thus the expected accuracy. The most relevant categories ordered by increasing accuracy include the generalized-gradient-approximations- (GGA), meta-GGA- (mGGA), hybrid-, and double-hybrid functionals.

Even though these categories are based on fundamental theoretical aspects, such that there is a general trend to improved accuracy with increasing rung, the performance variation between functionals of the same rung can be large. Accordingly, comprehensive benchmark studies that assess the performance of any functional with respect to the desired target property are indispensable. Nevertheless, the Jacob's ladder categorization allows a crude estimation of some systematic functional errors. Moreover, it can be used to categorize functionals by their numerical efficiency. In this regard, the perhaps most important distinction is made

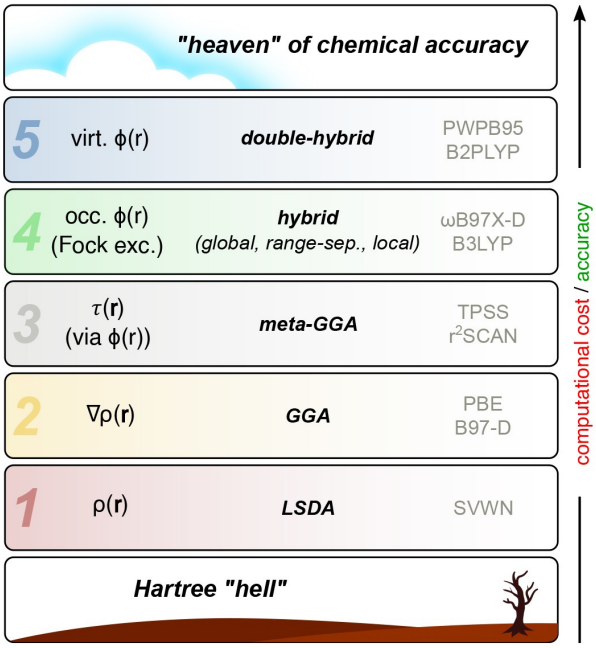

Figure 3. Functional categorization according to Perdew's "Jacob's ladder". ρ = electron density, τ = kinetic energy density, ϕ = molecular orbital, Fock exc. = Fock exchange.

based on the inclusion of Fock exchange, also termed non-local or exact exchange. The functionals of up to rung 3 do not include Fock exchange and are called local or semi-local functionals, whereas hybrid and double-hybrid functionals on rungs 4 and 5 include Fock exchange. Since the calculation of Fock exchange is a computational bottleneck, semi-local functionals are generally more efficient than (double-)hybrid functionals (see below).

Two of the most critical and thus prominent errors in actual DFT approximations are the so-called self-interaction error (SIE) and missing long-range correlation effects that give rise to London dispersion. Although the lacking description of long-range correlation is a fundamental shortcoming of DFT, it can nowadays easily be fixed by including one of several available proven dispersion corrections (we recommend D4, D3, or VV10).[18] We argue that nowadays, this is indispensable in any DFT treatment, and do not see any chemical context in which the dispersion correction should be left out (except when studying the influence of dispersion). See Section 3.7 for a discussion and example 4.2 and some illustrative numbers. In this context, it should be pointed out that while dispersion effects exert a crucial influence on calculated energies, the effect on the electron density appears to be negligible.[3,78,79] For example, it has been shown that excitation energies calculated with time-dependent DFT (TD-DFT) are hardly influenced by the self-consistent VV10 dispersion functional.[80]

The SIE[81–83] results from an artificial residual interaction of an electron with itself by the imperfect cancellation of the approximate exchange-correlation functional and the self-Coulomb interaction, and is more difficult to repair. In

contrast, Hartree–Fock theory is SIE-free, because of a mathematical relation between the integrals describing the Coulomb and (exact) exchange interaction: The two exactly cancel each other when an electron formally interacts with itself. In practice, SIE in DFT typically results in an over-delocalization of the electron density and artificial energy stabilization for delocalized electronic situations (with bond-stretched H_2^+ as a prime example). The SIE is present in all semi-local (m)GGA (meaning meta-GGA and GGA) functionals (rungs 1–3, e.g., PBE[84] or r^{2}SCAN). Hybrid functionals attempt to reduce the SIE by replacing a fraction of approximate DFT exchange with Fock exchange (e.g., B3LYP[9,10] with 20 % and PBE0[85] with 25 %). However, although this reduces SIE, it does not eliminate it. In turn, even functionals employing fractions of Fock exchange are still prone to SIE. Nevertheless, simply employing 100 % of Fock exchange also leads to a very poor performance. In general, Fock exchange admixtures of 5–25 % are typically considered small, 25–40 % moderate, and >40 % high. One way around this dilemma is provided by range-separated hybrid functionals (RSHs), which make the admixture dependent on the inter-electronic distance.[86,87] From a theoretical perspective, SIE can be viewed as a special form of the more general delocalization error of semi-local functionals resulting from the incorrect description of regions in a molecule with (effectively) fractional charges.[83,88,89] Regarding the understanding of these problems, some progress has been achieved recently by distinguishing density-driven errors caused by the erroneous (m)GGA exchange-correlation potential and inherent, usually smaller errors of the energy functional.[90]

Double-hybrid functionals[91–93] represent the highest rung and additionally introduce a wavefunction-theory-based correction to the correlation energy. This is most commonly achieved by perturbation theory methods,[94] such as second-order Møller–Plesset theory (MP2)[95] or its DFT variant.[94] Their typically very large amounts of Fock exchange (>50 %) make them particularly resilient towards SIE, and negative influences of large Fock exchange admixtures are balanced by the explicit, virtual-orbital-dependent correlation treatment. Nevertheless, these methods have an increased computational cost and further introduce some restrictions in their general applicability. For example, the vulnerability of the MP2 part for small gap systems may also be problematic in MP2-based double-hybrid functionals and hence, treating such systems in this way requires some caution.

Opposing the higher accuracy of fourth-rung hybrid and fifth-rung double-hybrid functionals is their increased computational cost and less favorable scaling, which become important specifically for large systems. (m)GGA treatments formally scale with the system size (N) as N^3 if the resolution of the identity (RI) approximation, also known as density-fitting,[96] is applied. For local (m)GGA functionals, RI can improve the computational speed by a factor of 5–30, depending on system and basis set. Hybrids already scale with N^4 and MP2-based double-hybrids with N^5. Accordingly, higher-rung functionals can become unfeasible for large systems (>200–300 atoms) on common hardware. This is mostly because hybrid calculations do not profit from RI as much since it cannot be applied efficiently to Fock exchange, which thus dominates the calculation time. However, there are semi-numerical integration techniques such as chain-of-spheres-exchange (COSX)[97] that also speed up hybrid calculations for basis sets of TZ quality or larger (factor of 3–10) when combined with RI. Thus, we strongly recommend using both RI and COSX if available (e.g., in *ORCA* as RIJCOSX and *TURBOMOLE* as SENEX[98]). The RI technique also speeds up the MP2 calculation part in double-hybrid calculations, which then become only slightly more costly than hybrids and hence this safe approximation is strongly recommended.

In addition to the consultation of comprehensive benchmark studies, consistency checks of the chosen functional may be recommended by comparing the results of a few functionals for a representative model system. These functionals should be chosen from a functional class that is conceptually suitable for the investigated problem, e.g., range-separated hybrid and double-hybrid functionals for barrier heights. This should be considered especially if the system under study differs greatly from those in corresponding benchmark studies.

A note on Minnesota functionals—Beginning in 2005 a series of functionals was developed by the Truhlar group in Minnesota, which are typically termed M*XX*-Y where *XX* stands for the year and Y informs about the purpose (e.g., HF for 100 % Fock exchange, "2X" for twice the Fock exchange, "L" for local).[16,99–104] These functionals are widely used and perform very well on large main-group benchmark sets like the GMTKN55, where M06-2X and M05-2X are actually some of the best-performing global hybrids, and M06-L is one of the best mGGAs. These methods represent a significant step forward in functional development compared to B3LYP, and solve some complicated electron correlation problems in DFT, such as alkane branching.[105] However, for several reasons, they should be used with care: First of all, they are often very sensitive to the size of the integration grid and the basis set,[14,106] which may lead to discontinuities in potential-energy surfaces and, in turn, problems with geometry optimizations.[14,107] Secondly, their performance strongly depends on the chemical system. Their very good performance for typical main-group chemistry relies on an extensive parameterization, which may lead to problems for less common systems reflected in the "mindless" benchmark[27] and for transition-metal chemistry.[43,108] Note that the performance for different classes of problems also strongly depends on the amount of Fock exchange, which varies between the different Minnesota functionals. This aspect is discussed in Ref. [16]. Thirdly, these functionals can be problematic for noncovalent interactions. Although they are designed to include dispersion effects at an electronic level and work quite well for weakly bound systems at their equilibrium distances (see example 4.2), they cannot recover the correct asymptotic behavior of London dispersion in the long intermolecular distance regime, e.g., for large molecules, solids, and liquids. To mitigate this, a dispersion correction needs to be added in

certain situations to Minnesota functionals, which, however, can also result in over-binding in others.[109]

Recommendations:

- Choose a functional with caution based on the chemical system under investigation and the task at hand and not based on popularity.
- Always include a dispersion correction.
- Check for reliable cost–benefit combinations and consider (m)GGAs. Hybrid functionals are more accurate but also much more expensive than ((m)GGAs).
- Check the consistency between different functional classes (e.g., compare a hybrid and a (m)GGA).
- In critical cases, test hybrids including different amounts of Fock exchange.
- Consider proven multi-level approaches and composite methods for larger systems (see Section 3).

2.6. Choice of Basis Set

Another important aspect regarding the computational speed/accuracy compromise is the applied atomic-orbital basis set. From a fundamental point of view, this is merely a technical aspect because DFT calculations can at least in principle be numerically converged to the complete basis set (CBS) limit where this influence is eliminated. In practice, however, this is rarely done and finite basis sets are applied, introducing some errors. However, basis set related errors in DFT are typically much smaller than for correlated wavefunction-theory-based methods. This weaker basis-set dependence compared to wavefunction theory is an important strong point of DFT. Nevertheless, even the best functional will yield bad results if evaluated in a small and insufficient basis set.

The most important characteristic of basis sets is their completeness, often referred to as basis set size. It reflects the number of functions to represent a given electron. An important error for too small basis sets is the so-called basis set incompleteness error (BSIE),[110] which results from an insufficient function space in the linear-combination-of-atomic-orbitals expansion. In short, BSIE arises when the employed basis set is not flexible enough to describe the fine details of the electron density. In most cases, the description of the valence electrons is most crucial and thus basis sets are usually categorized according to the so-called cardinal number that indicates the number of independent basis functions per occupied valence orbital. The corresponding size is usually referred to as double- (DZ), triple- (TZ), quadruple- (QZ),..., zeta, where the term zeta refers to the number of independent atomic functions per occupied (in the atomic ground state) valence orbital.

Another basis-set-related error arises if the basis set is too small: Spatially close atoms and fragments start to "borrow" basis functions from each other, resulting in an artificial energy lowering for more compact structures, which is known as the basis set superposition error (BSSE).[111] BSSE is the practically most relevant error and commonly only associated with weak interactions and non-covalently-bound intermolecular complexes, for which it can become relatively large (same magnitude as the interaction energy). However, it is important to recognize that BSSE is always present and also affects, e.g., conformational energies and even molecular structures if too small basis sets are used. Accordingly, knowing which computations are specifically sensitive to basis set size and where it is worth investing in the increased computational demand resulting from a larger basis set to significantly improve the result is fundamental. This aspect is the topic of Section 3 and Figure 5. For BSSE-prone systems with clearly separated fragments with no interfragment covalent bonds, the so-called counterpoise correction can be applied to correct for BSSE.[112] An efficient alternative to this computationally demanding correction is provided by approximate, empirical correction schemes that are based on the molecular structure, such as the geometric counterpoise correction (gCP), or employ specially adapted effective core potentials.[113] In contrast to the full counterpoise corrections, these are always applicable and computationally cheap, and thus can also be employed to correct for the intramolecular BSSE. Such approximate counterpoise corrections can repair the most drastic effects of BSSE, e.g., in geometry optimizations with small basis sets. Their application is similarly straightforward as that of dispersion corrections. Thus, we recommend the gCP approach of Kruse and Grimme[19] for geometry optimizations with small basis sets that supports HF and DFT as well as many basis sets. The related DFT-C approach of Witte and Head-Gordon is further recommended for accurate noncovalent interaction energy calculations (adapted specifically for DFT/def2-SVPD calculations, see example 4.2).[20] Other notable concepts are the proximity function of Faver and Merz for large biomolecules,[114] as well as the ACP-*n* approach of Jensen.[115] Even though such corrections can partially account for errors introduced by small basis sets, minimal basis sets should be avoided in general as the errors become uncontrollable.

The most commonly used Gaussian-type contracted basis sets belong to the Pople (e.g., 6-31G),[116] Dunning (cc-pVXZ),[117] Jensen (pc(seg)-X),[118,119] and Ahlrichs (def2-XVP)[120,121] families. As a technical side note, we want to mention that Pople-type basis sets such as 6-31G* or 6-311G** as well as the Dunning-type sets cc-pVXZ (X = D, T, ...) are not recommended here for standard DFT treatments, mainly because the basis sets by Ahlrichs and co-workers are more efficient and consistently available for a larger part of the periodic table.[122]

All basis sets may be extended by additional polarization functions that have a higher angular momentum or diffuse functions with small exponents to introduce more flexibility if necessary, e.g., for anions, dipole moments, or electric polarizabilities. Examples of the recommended Ahlrichs TZ basis set def2-TZVP are the def2-TZVPP basis set with added polarization functions and def2-TZVPD[123] or ma-TZVP[124] (in some contexts also denoted as ma-def2-TZVP) with added diffuse basis functions. In the following, basis sets with added (diffuse or polarization) functions will be referred to as large basis sets (e.g., def2-TZVPPD is a large

TZ basis), while cut-down variants of the def2 basis sets will be denoted small (e.g., mTZVPP derived from def-TZVP for the r^{2}SCAN-3c composite method is a small TZ basis). To test if the results of a calculation are converged with respect to the basis set size, it should be investigated how they change when the cardinal number is increased (def2-TZVP to def2-QZVP), and when polarization functions are added or removed (def2-TZVP to def2-TZVPP).

If heavy elements are involved, def2-XVP basis sets use the matching Stuttgart–Cologne effective core potentials (def2-ECPs for $Z>36$)[125] to replace the inner core electrons. This not only reduces the computation time but also increases the accuracy due to the implicit inclusion of (scalar) relativistic effects, which mainly affect the core electrons. The application of robust small-core effective-core potentials is typically sufficient for most thermochemical property calculations. Nevertheless, for certain properties of very heavy nuclei and properties involving core electrons, such as NMR shieldings, explicitly relativistic all-electron calculations with special Hamiltonians, such as X2C,[126,127] ZORA,[128] or DKH[129] may be necessary. For some approaches such as X2C, specially adapted versions of the Ahlrichs basis sets are available.[130] Further details on explicitly relativistic methods are beyond the scope of this work.

Recommendations:

- Be aware of BSIE and BSSE.
- Try to approach a reasonable basis set size ($\geq$TZ) for energy-related properties.
- Consider adding polarization functions for flexibility (def2-XVPP).
- Consider adding diffuse functions for anions, dipole moments, and polarizabilities (ma-XVP, def2-XVPD).
- Check for basis set convergence (increase/decrease cardinal number by one).
- If heavy atoms are present ($Z>36$), apply effective core potentials.

2.7. Comparing Apples with Apples

Finally, a remaining critical point is the consideration of finite temperature effects. In the Born–Oppenheimer approximation, most computational models per construction yield results valid for the absolute zero temperature ($T=$ 0 K, −273.15 °C), nuclear equilibrium scenario. Therefore, a systematic deviation compared to experimental data at finite temperatures is expected as standard equilibrium structure treatments, such as geometry optimizations, do not include nuclear zero-point vibrations (ZPV), vibrational (thermal) bond elongation, or entropy effects. In other words, a perfect agreement between experimental structures obtained at finite temperature and those calculated at $T=0$ K is not necessarily desirable since—sometimes substantial[131–133]—finite-temperature effects can cause a significant bias. However, for common covalent bonds between typical atoms, these effects are too small to have a significant influence.

Concerning calculated DFT energies, it is essential to recognize that the bare electronic energies (E) and differences thereof cannot be directly compared to measured (reaction) enthalpies (H) or free energies (G). This is because even measurements at very low temperature (assume 0 K) in the gas phase are affected by zero-point vibrations (ZPV), which are thus an inseparable part of the internal energy U^0 of any molecule. If bonds are formed or broken during a reaction, this will significantly change the ZPV energy by several kcal mol^{-1}, which is thus very important. Moreover, since most measurements are conducted at finite temperature (assume ambient conditions = 298 K), also thermostatistical corrections have to be included in the internal energy U^{298}, which again have a vibrational, translational, and rotational origin ($3kT$ for the translational and rotational part in an asymmetric molecule, typically smaller than the ZPV energy). Moving from U^{298} to the enthalpy H^{298}, the pressure-dependent volume-work pV has to be added under the usual assumption of ideal-gas-like behavior (done by most programs, but not entirely correct in the more common situation of a reaction in solution, see the discussion below).[134] Finally, the most common and relevant quantities are (Gibbs) free energies G^{298}, which include temperature-dependent entropic corrections (TS) for vibrational, rotational, and translational contributions ($G=H-TS$). For reactions in which the number of free particles changes, these entropic contributions can amount to tens of kcal mol^{-1} at ambient conditions (see examples in Section 4) and are thus also very important.[134]

For a reaction in solution as opposed to (an idealized) gas phase, these contributions have to be evaluated in the presence of a solvent model, and the free energy of solvation has to be included as detailed in Section 2.3. Moreover, the solvation process refers to a change of standard state (from the usually assumed gas phase at 1 atm to 1 mol L^{-1} in solution). Hence, a concentration-induced free-energy shift of $RT\ln V_M = 1.89$ kcal mol^{-1} (V_M: molar volume of an ideal gas, $T=298$ K) has to be added (in addition to the pV term in H^{298} discussed above) for each species when common solvent models (COSMO-RS, SMD, ALPB/GBSA) are used with default settings. This generally affects all non-unimolecular reactions modelled in solution. Other reference state definitions for gas or condensed phases are possible, and in these cases a detailed consideration of the applied solvation model is required.[134]

Recommendations:

- Try to model the experiment as closely as possible.
- Apply zero-point vibrational energy and thermostatistical corrections to enthalpy or free energy if necessary.
- For large systems with many small vibrational frequencies $<$50–100 cm^{-1}, consider the robust mRRHO[135–137] model for the entropy part of the free energy.

As many quantum chemical applications typically involve a manifold of methods based on various physical concepts, further reading on general computational chemistry as well as specific aspects is recommended.[5,33,93,110,138–141]

2.8. Quantum Chemistry Program Packages

In the last decades, various more or less specialized program packages have been developed. All of them provide the same basic DFT functionality and relevant differences concern their ease of use, their efficiency, and their availability (free or commercial). Further, many sophisticated program packages can make use of the *LibXC*[142] or *XCfun*[143] functional databases, drastically increasing the number of available density functionals. Nevertheless, in some cases, native functional implementations provide a larger technical functionality in some codes. In the following, a selection of common program packages with versatile functionality in molecular applications is given, and individual strengths and specialties are highlighted briefly. These programs proved to be reliable for a wide range of quantum chemistry applications in our group. Other noteworthy quantum chemistry program packages that are not explicitly discussed here as we lack explicit experience with them include *NWChem*, *CP2K*,[144], *DALTON*,[145] *MOLCAS*,[146] *PySCF*,[147] *DIRAC*,[148] *Gaussian*,[149] and many more.[150]

TURBOMOLE[151]
- Very fast and robust
- Technically advanced and state-of-the-art algorithms
- Advanced molecular symmetry handling
- Fast hybrid functional implementation by semi-numerical Fock-exchange approximation (SENEX)

ORCA[152–154]
- Efficient implementation of DFT and wavefunction theory
- Fast hybrid functional implementation by semi-numerical Fock-exchange approximation (RIJCOSX)
- Large toolkit for molecular spectroscopy
- Easy and intuitive input structure
- Free of charge for academic use

Q-Chem[155]
- Large number of natively implemented density functionals
- Large variety of specialized DFT treatments (time-dependent DFT, constrained DFT) and analysis tools (energy decomposition analysis, EDA)
- Special DFT methods for NCI (SAPT, DFT-C)

Psi4[156]
- Modular code with great interfacing/scripting capabilities
- Special DFT methods for NCI (SAPT)
- Free of charge for academic use
- Open source code

Molpro[157]
- Advanced DFT (RPA, ACFDT) and embedding (WFT-in-DFT) techniques
- Special DFT methods for noncovalent interactions (SAPT)

AMS[158]
- Use of Slater-type orbital (STO) basis sets up to QZ quality
- Good relativistic treatments
- Huge quantum chemical toolbox
- Great graphical user interface (GUI)

Specifically for large systems or very demanding computational tasks, DFT methods can be complemented by semi-empirical quantum mechanical methods. Even though some of these are implemented in the abovementioned quantum chemistry packages, several features are only available in the corresponding original codes. The most frequently used semi-empirical quantum mechanics programs are *xtb*,[159] *DFTB+*,[160] and *MOPAC*.[161] For example, the recent single-point hessian (SPH)[162] approach can be utilized with semi-empirical methods by using the native implementation in the program. Further, the *xtb* program can be used as a driver for other codes (using the *xtb* functionality with parts of other codes to employ, e.g., DFT in the SPH approach) such as *ORCA*. A detailed documentation can be found in Ref. [163].

3. The Right Tool for the Task

A central task in computational chemistry is to balance computational demands against methodological accuracy and robustness. To this end, it is important to be aware of the strengths and weaknesses of specific methods, to consider the system size, and to be aware of the targeted properties, such as structures, reaction energies, or conformational energies. Practically relevant systematic errors in this respect are the aforementioned self-interaction error (SIE), basis-set superposition (BSSE) and incompleteness errors (BSIE), as well as the lacking description of London dispersion by most functionals. The relevance of these aspects depends not only on the system under investigation but even more so on the task at hand. Thus, it is instructive to discuss methodological choices in the framework of the most typical steps of a computational investigation, which determines the targeted property. The aim of this section is thus to guide the choice of the methodological tool set, and to provide the means to adapt it to the task at hand. Simply put: We want to explain how one can cut corners where it does not hurt through clever methodological choices and the use of multi-level approaches.

An illustrative example for multi-level approaches is the use of efficient semi-local (GGA or mGGA, i.e., no Fock exchange) functionals or composite methods for structure optimizations and vibrational frequency calculations, which are then combined with single-point energy calculations at a higher (hybrid or double-hybrid) level in a larger basis set (e.g., QZ). The underlying idea is that although energies calculated with (m)GGA functionals are susceptible to SIE (amongst other shortcomings), they still provide reasonable structures at a small fraction of the computational cost of the more advanced non-local (hybrid and double-hybrid) functionals. This mostly holds true even for SIE-prone

systems, such that single-point calculations with hybrid functionals on (m)GGA structures often provide energies and properties as accurate as a fully optimized hybrid approach. One example for this is the anthracene-cyclopropenyl cation potential-energy surface from our recent article introducing r^{2}SCAN-3c[22] shown in Figure 4.

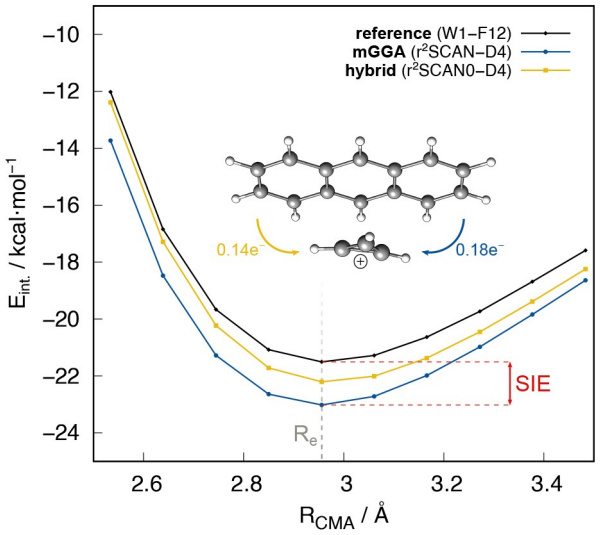

Figure 4. Potential energy surface along the cyclopropenyl-anthracene center-of-mass distance R_{CMA} for the mGGA r^{2}SCAN-D4, its hybrid variant with 25% Fock exchange r^{2}SCAN0-D4,[164] and the W1-F12 reference. W1-F12 denotes a highly accurate wavefunction theory-based reference level. All DFT data calculated with the def2-QZVPP basis set. R_e = equilibrium distance. Colored arrows indicate the charge transfer from anthracene to the cyclopropenyl cation for the respective theoretical level.

Although the mGGA-based composite method overestimates the interaction energy due to SIE, the equilibrium intermolecular distance is in very good agreement with the high-level coupled-cluster reference.

In more general terms, multi-level approaches exploit a very stable form of error cancellation: While low-level (e.g., PBE-D4/DZ) structures are certainly not as accurate as high-level (e.g., B2PLYP-D4/TZ) ones when comparing to reference data (e.g., bond length and angles), the deviations between high and low levels are typically highly systematic (for a detailed analysis see Ref. [165]). As a result, also the electronic energy differences between structures obtained at different levels are very systematic and therefore reliably cancel out between educts and products. Since such error cancellations can involve large energy differences of tens of kcal mol^{-1} (comparing the high-level energy of the low-level-optimized structure, here B2PLYP-D4/TZ//PBE-D4/DZ, to the fully optimized high-level structure, here B2PLYP-D4/TZ), it is of utmost importance to keep all details of the optimization (like functional, basis set, convergence criteria, used programs, and associated technical settings) consistent over all involved molecules. Moreover, it is important that the low-level method still provides a somewhat reasonable description of the system under investigation. If the structures are fundamentally wrong, e.g., because of a complete lack of dispersion effects (e.g., with plain B3LYP), the usage of a much too small (minimal) basis set, or a functional that yields the wrong spin/charge state of the fragments, even the most stable error cancellation will break eventually. Finally, we want to point out that the described (passive) error cancellation is not to be confused with (active) error compensation, the latter of which refers to the application of correction schemes for known sources of systematic errors, like using D4 to account for the lacking description of dispersion or the gCP correction to mitigate BSSE.

Recognizing that structure optimizations are generally much less sensitive to the level of theory than energy and many other property calculations, multi-level approaches enable large computational savings without any significant loss of accuracy. Generalizing from this example, we have summarized the most suitable functional/basis-set combinations for typical steps of a computational study in Figure 5. To guide the choices visually, we have marked the recommended level of theory with the best balance between computational effort and robustness in blue, accurate and robust but not necessarily efficient choices in green, and methods that should be avoided in red. Less clear-cut cases (in between red and green) are marked in yellow. These methods can be a good choice but are not as robust as green ones. Hence, results obtained at this level should be checked for systematic errors as indicated in the field, and often better alternatives (blue fields) are recommended instead. In all cases, the most severe systematic errors and drawbacks we expect based on our experience are shown as text. In the following, we will discuss this table column (task)-wise to motivate and explain our choices, and provide additional details. To streamline this discussion, we fully separate the theoretical tasks in the computational context. This means that, e.g., the discussion of methods for conformational energies only refers to calculation of electronic energies and excludes vibrational and entropic contributions, which we discuss separately.

3.1. Molecular Structure

Structure optimizations are the first step of most quantum-chemical investigations, either starting from an experimental reference or a low-level guess that may employ semi-empirical quantum mechanics or force-field methods. At this point, the most typical user errors concern the starting structure(s) and mistakes in the input resulting in wrong charge, spin, or protonation states, which drastically alter the results, whereas wrong choices of the theoretical model are less common and often less severe. Hence, we recommend carefully checking the input before setting the computational machinery in motion. Self-consistent field convergence issues are often a strong hint towards problems with the input.

As already discussed above, structures are much less sensitive to the functional and basis set than energies and properties.[168,169] Therefore, we argue that a TZ basis set

functional class / recommendation	basis size	structures	frequencies	conformers	reaction energies	barriers	NCI
(meta-)GGA B97M-V, r^{2}SCAN [a], TPSS	DZ		structure[b]	BSIE	SIE, BSIE	SIE, BSIE	SIE, BSIE
	TZ				SIE, BSSE	SIE, BSSE	SIE, BSSE
	QZ	excessive	excessive		SIE	SIE	SIE
hybrid PW6B95, PBE0	DZ		structure[b]	BSIE	BSIE	BSIE, SIE[c]	BSIE
	TZ				BSSE	BSSE, SIE[c]	BSSE
	QZ	excessive	excessive			SIE[c]	
range-separated hybrid ωB97X-V, ωB97X-D4, ωB97M-V, ωB97M-D4,	DZ		structure[b]	BSIE	BSIE	BSIE	BSIE
	TZ				BSSE	BSSE	BSSE
	QZ	excessive	excessive				
double-hybrid PWPB95, ωB97M(2), revDSD-PBEP86-D4	DZ		structure[b], excessive	BSIE	BSIE	BSIE	BSIE
	TZ	excessive	excessive	BSSE	BSSE	BSSE	BSSE
	QZ	excessive	excessive				
composite methods							
B97-3c	TZ				SIE	SIE	SIE
r^{2}SCAN-3c [a]	TZ				SIE	SIE	SIE
PBEh-3c [d]	DZ				BSIE	BSIE	BSIE
commonly used combinations							
B3LYP/6-31G*	DZ	BSSE, no-D	structure[b], no-D	BSSE, no-D	BSSE, no-D	BSSE, no-D	BSSE, no-D
BP86-D3	TZ				SIE, BSSE	SIE, BSSE	SIE, BSSE
M06-2X/6-311G** [a,d,e]	TZ						

[a] High dependency on the size of the numerical integration grid.

[b] If the structure is reasonable (DZ ≈ TZ), GGA/DZ or even semi-empirical frequencies are sufficient. If not, use SPH.

[c] Self-interaction error can limit the accuracy for hybrid functionals with small amounts of Fock exchange (< 25%).

[d] The large amount of Fock exchange can cause pronounced errors for transition metal complexes.

[e] 6-311G** behaves like a double-zeta basis set and should be used with caution.

not recommended | intermediate | accurate | recommended / good cost-benefit-ratio

Figure 5. Decision matrix to guide the choice for a method combination (functional class/basis set) for common computational tasks. All considerations imply the use of a dispersion correction. Red, yellow, and green indicate the accuracy and reliability of a method for the given task, while blue marks our recommendation based on a good cost/accuracy ratio. Text in the fields points out the most relevant, systematic errors we expect at this level (SIE, BSSE, BSIE) or if a theory level is unnecessarily demanding (excessive). Excessive method combinations probably do not yield any significant increase in accuracy justifying the much-increased computational cost in the respective application. Selected recommended methods are given with the respective method class. DZ = double-zeta; TZ = triple-zeta; QZ = quadruple-zeta basis set. BSIE = basis set incompleteness error; BSSE = basis set superposition error; no-D = missing London dispersion; SIE = self-interaction error; Fock exchange = non-local exchange (also termed exact exchange) from wavefunction theory. Regarding the limitations of the 6-311G** basis set see Refs. [166] and [167].

(def2-TZVP) is sufficient, and going higher to QZ level is generally a waste of computational resources. Even well-balanced DZ basis sets (def2-SVP) can provide useful results when combined with the empirical geometric counterpoise (gCP)[19] and dispersion corrections to mitigate the sometimes significant structural impact of BSSE. In fact,

combining small but well-balanced basis sets with the gCP and the D3 or D4 dispersion corrections is the basis of the PBEh-3c[32] (DZ, def2-mSVP) and r^{2}SCAN-3c (TZ, mTZVPP) composite methods, which are tailor-made for the task of structure optimization and thus strongly recommended. In our opinion, it is rarely required to go beyond this level for structure optimizations in standard thermochemical studies.

Concerning the functional, (m)GGAs are typically sufficient (e.g., r^{2}SCAN-D4,[170,171] TPSS[172]-D4, or even PBE-D4) as already discussed at the beginning of this section. They provide the additional advantage of fully exploiting the resolution-of-the-identity (RI) approximation,[96] also called density-fitting, which makes (m)GGA calculations much more affordable than using hybrids. As a result, (m)GGA functionals can be employed with larger basis sets, often at the same or even lower cost than hybrids with small basis sets like B3LYP/6-31G*. In turn, we argue that hybrid functionals should only be used for structural optimizations if there is a good reason, such as strong SIE, application to transition state searches, weakly bound electrons in anions, or the presence of heavy main-group elements. To investigate whether hybrid functionals have a significant influence, we suggest beginning by comparing results obtained with the r^{2}SCAN-3c and PBEh-3c composite methods. A very robust but also significantly more expensive choice is PBE0-D4/TZ, which can already be regarded as a benchmark level for structural optimizations. B3LYP-D4/TZ achieves a similar level of accuracy but is—in our experience—not quite as robust as PBE0-D4/TZ, in particular for transition-metal-containing systems. Also range-separated hybrids and double-hybrid functionals have been shown to provide very accurate structures.[173] However, the application of double-hybrids is typically excessive and not necessary for standard applications due to beneficial error-cancellation effects for structures. Even for very large systems, structural optimizations should not go below the polarized DZ basis set level. The functional should be at least of (m)GGA quality (PBE or TPSS) and always include a dispersion correction (D3/D4,[141,174,175] VV10[176]). Note that the prominent BLYP[177,178] functional is not recommended because the GGA-typical systematic overestimation of covalent bond lengths is more pronounced than with the recommended GGAs.

3.2. Vibrational Frequencies

The calculation of vibrational frequencies is indispensable to obtain zero-point vibration energies, thermostatistical corrections to enthalpy and free energy. It is also central for the prediction of IR/Raman spectra, which is, however, not in the focus here. Since the underlying calculation of second energy derivatives quickly becomes very demanding for larger systems, it is desirable to conduct such calculations at the lowest level that still provides reasonable results. Therefore, we recommend using efficient composite methods or, alternatively, (m)GGA functionals with a DZ basis set in combination with the gCP correction. Similar to structures, error cancellation is very stable and, since the thermostatistical correction to a reaction energy is typically much smaller than the electronic part from DFT, the impact of a lower level is naturally limited. In many cases, it is even sufficient to obtain thermostatistical corrections at a semi-empirical quantum mechanics level, e.g., with GFN2-xTB,[5,179] as shown in example 4.2. We note that the applicability of low-level semi-empirical methods should be carefully checked if exotic bonds, transition metals, or heavy main-group elements are present.[131]

The only complication that arises concerning the use of lower-level methods for frequencies is that the structure for which the vibrational frequencies are calculated has to be an energy minimum at this very level (fully optimized, vanishing atomic forces). Otherwise, the presence of many artificial imaginary frequencies severely limits the accuracy of the calculated thermostatistical corrections. For this reason, the same theory level has to be used for structure optimization and vibrational frequency calculations, or a second set of optimized structures obtained at this lower level has to be used. This limitation can be overcome with the recently proposed single-point Hessian (SPH) approach, which makes it possible to obtain frequencies and reasonable thermostatistical corrections for any non-equilibrium structure through the application of a specific biasing potential.[162] While originally developed for the semi-empirical GFNn-xTB method, the SPH approach can also be used for DFT if the *xtb* program is used as a driver for a quantum-chemistry program like *ORCA* (see *xtb* documentation).[163]

Note that it is quite typical for large systems to have a few low-energy imaginary frequencies < 50–$100\ cm^{-1}$, which are often technically related to a too small DFT integration grid. This is per se not a problem if properly dealt with, that is, the frequencies should be inverted (multiplied by $-i$) and treated as normal real frequencies, like in the *xtb* program using the mRRHO approach.[159,180] A prominent alternative consists of simply removing all low-lying imaginary modes (referred to as Truhlar's approach in the literature). However, since this can inconsistently change the number of degrees of freedom, it may lead to significant errors in the thermostatistical entropy corrections. In contrast, if a frequency calculation for fully optimized structures results in more than a few low-energy modes and/or higher imaginary modes ($> 100\ cm^{-1}$), this indicates that the given structure is not a minimum and requires further refinement, e.g., by manual distortion and re-optimization.

If individual vibrational frequencies and IR/Raman intensities are desired instead of just thermostatistical corrections, it may be useful to move to a hybrid level of theory and employ larger basis sets, as this improves the quality of the calculated spectral intensities. One functional with a proven track record for the computation of IR spectra is B3LYP, or its dispersion-corrected low-cost variant B3LYP-3c, see Refs. [9,21]. For a comprehensive analysis of the performance of various functionals, basis sets, and the influence of Fock exchange on vibrational frequencies, we refer to the work of Radom and co-workers.[181]

3.3. General Remarks on Energy Calculations

For the following four categories, conformers, reaction energies, barriers, and noncovalent interactions, the calculated electronic energies become the central quantity, which requires larger basis sets for converged results than structures or frequencies. Therefore, DZ basis sets (like 6-31G** or def2-SVP) are no longer sufficient, and we strongly advise against using them, except if they are part of purpose-made composite schemes. However, even in combination with full counterpoise corrections or gCP (as in composite schemes), the residual BSSE and BSIE of DZ basis sets can be substantial. Thus, we generally recommend at least TZ basis sets, which often yield results reasonably close to the basis set limit. Nevertheless, convergence should be checked for representative examples against QZ level (or larger TZ basis sets, e.g. def2-TZVP against def2-TZVPPD). In any case, before reducing the basis set size to the bare minimum, we suggest moving from hybrids to more efficient (m)GGA functionals or composite methods, which are purpose-made to perform well with smaller basis sets.

For double-hybrid functionals, the limitations of TZ basis are more severe since the MP2 component of the calculation typically has a stronger basis-set dependence than the (m)GGA or hybrid DFT part. Here, complete-basis-set-extrapolation (CBS) from TZ to QZ should be considered,[182,183] but the details are beyond the scope of this work. Lastly, we suggest deciding on one approach for all energy-related properties to retain a certain degree of consistency and comparability, meaning that reaction energies, barrier heights, and association energies of reactants in one reaction or reaction network should be obtained consistently with one method combination, e.g., ωB97M-V/QZ[184] single-point energies on r^{2}SCAN-3c structures (ωB97M-V/QZ//r^{2}SCAN-3c). Therefore, a compromise needs to be made considering all the requirements stated in the following section.

3.4. Conformers

Conformational energies refer to the electronic energy differences between various conformers of a given molecule, that is, differences between local minima for a fixed covalent bond topology. Due to the typically similar structures and fixed topology, conformational energies profit particularly from error cancellation effects. At the same time, however, conformational energies need to be accurate to within about 0.1–0.2 kcal mol^{-1} to predict Boltzmann populations at room temperature reasonably well. This is particularly important since properties can vary strongly and even qualitatively between populated conformers (see example 4.3.).

Thus, conformational energies should at least be obtained with high numerical precision settings (integration grids) and TZ basis sets, and for this reason, the PBEh-3c composite method (based on a DZ basis) is not recommended for the task. Conformational energies are particularly sensitive to mid- and long-range electron-correlation effects, such that the dispersion correction takes an important role. In particular, for metal–organic systems, density- or charge-dependent corrections like VV10 or D4 should be preferred over the charge-independent D3 scheme.

According to the conformational subsets of the GMTKN55 benchmark, r^{2}SCAN-D4 and the r^{2}SCAN-3c and B97-3c composite methods are particularly well-suited for predicting conformational energies with high accuracy. If the systems are small and higher-level calculations are affordable, hybrid functionals can be employed. Here, we recommend the ωB97X-V and ωB97M-V approaches of Mardirossian and Head-Gordon or the respective D4 or D3 analogues.[185,186] Also the common B3LYP functional with a dispersion correction and suitable basis set (e.g., B3LYP-D4/TZ or QZ) can provide accurate conformational energies, yet it should be noted that this method has been outperformed in a recent benchmark by the r^{2}SCAN-3c composite method at a small fraction of the computational cost (see also example 4.3.).

Double hybrid functionals can be used for small systems if maximum robustness and accuracy are desired. The most accurate functionals in the GMTKN55 benchmark are ωB97M(2)[187] of Mardirossian and Head-Gordon and revDSD-PBEP86-D4[188] of Martin and co-workers. A particularly robust and widely available double hybrid we also want to recommend is PWPB95-D4.[189] On a side note, we want to mention that semi-empirical quantum mechanics and force-field-based methods yield much less reliable conformational energies than DFT and hence can only be applied in the initial steps of multi-level workflows and by using a very conservative (large) energy selection window, and are best combined with DFT-based energy re-ranking (see workflow given in Figure 7).[67]

3.5. Reaction Energies

Reaction energies refer to the difference in the total electronic energies between reactants, products, and possible intermediates that constitute minima on the potential energy surface. Due to the larger differences in molecular geometries and electronic structures, reaction energies benefit less from error cancellation than conformational energy differences. However, the practically acceptable error for reaction energies is typically also larger than that for conformational energies. An accuracy of about 1–2 kcal mol^{-1}, which is difficult to obtain experimentally, is often considered sufficient. While all of this strongly depends on the investigated reaction, several general findings remain valid for the vast majority of examples: Regarding the basis set, reaction energies require large TZ (e.g. def2-TZVPPD) or QZ basis sets for converged results (see examples). The basis set dependence should be carefully investigated by comparing single-point energies obtained with the next-smaller basis set (e.g., def2-QZVP to def2-TZVP). Even if computational resources are at the limit, the basis set size should not be reduced below TZ quality.

Instead of reducing the basis set size, one should always consider switching to semi-local (m)GGA functionals if SIE is not a major concern. Surprisingly, one of the most efficient approaches for accurate reaction energies is the mGGA-based r^{2}SCAN-3c composite method, which outperforms the recommended B3LYP-D4/QZ and PW6B95/QZ hybrid methods on the reaction energy subset of the GMTKN55 benchmark (see Figure 4 in Ref. [22]).

In general, however, the accuracy and, in particular, the robustness of predicted reaction energies of nonmetallic systems profits from an admixture of Fock exchange. The optimum value for reaction energies with global hybrids is at 25 % Fock exchange. In contrast, barrier heights often profit from even higher amounts of 50 % depending on the reaction type (see details in Section 3.6.). Therefore, if both reaction energies and barrier heights are to be calculated, preferably with the same functional as discussed above, a compromise must be made with global hybrid functionals. In this regard, we recommend Truhlar's PW6B95 (28 % Fock exchange)[190] hybrid with D3 or D4 dispersion correction, which provides accurate and robust thermochemistry, and has been our default hybrid functional for mechanistic studies for many years.

An approach to solve the conflict between high/low Fock exchange for barriers/reaction energies is provided by range-separated functionals such as ωB97X-V and ωB97M-V. The variable admixture of Fock exchange in these functionals makes it possible to get the best of both worlds: good reaction energies and barrier heights. Accordingly, the best-performing hybrid functionals on the GMTKN55 benchmark set are the two above-mentioned range-separated functionals, which are the only hybrids with a WTMAD2 below 4 kcal mol^{-1}. Double hybrid functionals are even more accurate and robust choices but also the most computationally demanding because they require larger basis sets.

Lastly, to complete this discussion of the optimal amount of Fock exchange, we mention that this value also depends on the type of the studied reaction. While the arguments and values presented above apply to main-group chemistry, transition-metal compounds typically require lower amounts (roughly about half of the values given above). Accordingly, functionals with very large amounts of Fock exchange and range-separated hybrids should be applied with care. An example is the prominent M06-2X (54 %) functional and the PBEh-3c composite method, which showed larger errors in recent transition-metal reaction benchmarks.[43,108]

3.6. Barrier Heights

Barrier heights refer to the electronic energy difference between the transition state and the corresponding reactants and products. Many transition state structures typically involve at least one stretched bond and, in turn, near-degenerate orbitals with weakly bound electrons, which gives rise to particularly challenging electronic structures. As a result, transition states are typically prone to SIE, which often leads to a systematic underestimation of their electronic energy and, in turn, barrier heights by (m)GGA functionals. This error strongly depends on the one-electron character of the breaking bond(s) in the transition state. The largest errors are encountered for hydrogen-transfer or dissociation reactions. The errors are smaller e.g., for heavy-atom pericyclic reactions, and are almost absent for covalent-bond-conserving inversion or conformational processes. Hence, cheap (m)GGA functionals may be used in special cases such as conformational or inversion barriers after careful testing but, in general, barrier height calculations are the only category in which we advise against using semi-local (m)GGA functionals. Since the calculation of barrier heights is perhaps the most challenging task, we provide an extended discussion of the aspects mentioned above in the example shown in Section 4.4.

In general, to mitigate the errors related to SIE, range-separated hybrids are strongly recommended for barrier heights. Specifically, ωB97M-V (also ωB97X-V) performs very well on the barrier height subsets of the GMTKN55 and other benchmark sets. If global hybrids need to be employed for some reason, those with an increased Fock exchange (>30 %) should be preferred, and their results should be compared against those of range-separated hybrids. Global hybrid functionals that have been specifically designed for the prediction of barrier heights, like the BMK[191] or MPWB1K[192] functionals, typically use 40–50 % Fock exchange (BMK 42 %, MPWB1K 44 %). Note that such large amounts of Fock exchange typically deteriorate the performance of reaction energies. Also, double hybrid functionals are well suited for this task since they typically use a much larger fraction of Fock exchange (>50 %) than global hybrids and nevertheless provide very accurate reaction energies. Due to their challenging electronic structure, basis set convergence may also be slower for transition states/barrier heights than for other properties, and QZ basis sets should be considered. The London dispersion energy contribution to typical chemical reaction energies or barriers computed with standard functionals can be large (or even decisive, see Ref. [193]), especially for molecules with >20–30 atoms, and hence, its explicit consideration is very strongly recommended.

In addition to the final energy calculation for reaction barriers, already the initial search for transition states represents an important task. Even though technical and often program-specific aspects of transition state search are beyond the scope of this article, most of the given considerations also hold for, e.g., potential-energy surface scans and other transition state searches. Since (semi-)transition state searches are often computationally demanding due to the large number of energy and gradient evaluations, the use of lower-level methods cannot be avoided. For this purpose, hybrid-based composite methods with DZ basis sets (e.g., PBEh-3c) or even semi-empirical methods[194–199] may be used (the latter with care) to find a guess for transition states, which can then be refined at the recommended level of theory. For further detailed advice, we refer to the manuals of the respective quantum chemistry programs (e.g., the *ORCA* manual has a large section dealing with transition-state searches).

3.7. Noncovalent Interactions

Noncovalent interactions refer to the difference in electronic energies between a noncovalently interacting complex and its isolated molecular fragments. Since, by definition, no covalent bonds are changed upon association of the fragments, noncovalent interaction (NCI) energies strongly profit from error cancellation, even more so than conformational energies. However, at the same time, NCI energies are often relatively small on an atom pair-wise basis. Therefore the required accuracy and numerical precision are often higher than for reaction energies. Moreover, NCI energies may be of a similar magnitude to BSSE and SIE, whose influence should therefore be carefully investigated. Since London dispersion is usually a dominant contribution to the binding in NCI complexes, the dispersion correction is particularly important. While VV10 can have a slight edge over D4 in exotic and charged systems due to its dependence on the density, the inherently better C_6 coefficients[18,175,176] as well as the inclusion of three-body-terms in D4 (and D3-ATM in PBEh-3c),[200,201] make D3-ATM and D4 more accurate in highly polarizable, small-gap systems (e.g., buckyballs or graphene sheets, the L7 and S30L benchmarks[202,203]).

We recommend the "3c" composite methods for efficient calculations of noncovalent interaction energies since these methods were designed for this purpose. r^{2}SCAN-3c is the best choice for most systems. PBEh-3c can be superior for systems prone to SIE, which are typically highly polar. However, the small basis of PBEh-3c can be problematic in such cases due to BSSE and BSIE. Another composite approach has been developed by Head-Gordon and coworkers that combines the accurate B97M-V mGGA functional with the def2-SVPD basis set and a tailor-made gCP-derived correction termed DFT-C.[20] This approach provided NCI energies with an accuracy comparable to B97M-V/QZ results in their tests. We demonstrate this approach in the NCI example in Section 4.2.

Finally, we want to mention that due to the good error cancellation for noncovalent interaction energies, even semi-empirical quantum mechanical methods like GFN2-xTB or PM6-D3H4[204,205] can provide reasonable results for interactions including complex geometries at a fraction of the computation cost of a DFT calculation. Combining these very "low-cost" structures and frequencies with single-point calculations at a composite-DFT level is perhaps the best way to estimate noncovalent interaction energies for very large systems with hundreds or thousands of atoms.

On the high end of the methodological spectrum, ωB97M-V and ωB97X-V provide exceedingly accurate noncovalent interaction energies if combined with a TZ or QZ basis. Also, double hybrid functionals are very accurate and robust for noncovalent interaction energies, but also more basis set dependent and computationally demanding. Noncovalent interactions in or with very small gap systems (metals) are generally not well understood and require special treatment.[206]

4. Examples

The examples presented in this section are selected based on two considerations: Firstly, they should generally reflect our basic arguments from the previous discussions concerning the performance and suitability of functionals, basis sets, and multi-level workflows for certain tasks. Secondly, the examples furthermore show that a single example can never be as representative and instructive as a large and well-designed benchmark. Thus, it might be that, in some cases, a certain low-level method/functional performs surprisingly well, while a high-level approach shows a surprisingly large deviation of a few kcal mol^{-1}. Such outliers showing unsystematic behavior should be expected in any real-world application due to the approximate nature of DFT and efficient computational methods in general. We want to reiterate here our statement from the introduction: The aim of a best-practice protocol for DFT cannot be to get all the numbers right to the last kcal mol^{-1}. Instead, the aim is to avoid large and systematic errors by choosing robust and greatly reliable methods and protocols adapted to the physical reality of the studied systems.

4.1. Formation and Isomerization of [2.2]Paracyclophane

The dimerization of 3,6-dimethylidenecyclohexa-1,4-diene to [2.2]paracyclophane and its subsequent isomerization to [2.2]metacyclophane shown in Figure 6 represents a fundamental chemical transformation in organic chemistry. For both reactions, experimentally determined standard heats of formation can be used to derive the reaction enthalpies ($\Delta H_{\text{exptl.}}$).[207–209] Nevertheless, due to high uncertainties for the 3,6-dimethylidenecyclohexa-1,4-diene monomer, the reaction enthalpy of the dimerization amounts to -41.9 ± 10.6 kcal mol^{-1}. This error estimate is much smaller for the isomerization for which a reaction enthalpy of -17.9 ± 2.9 kcal mol^{-1} is derived (the numbers in Figure 6 have been rounded to significant digits). To obtain more precise reference values, electronic energies were calculated with the high-level W1-F12[210] coupled cluster method and combined with PBE0-D4/def2-TZVP zero-point vibrational energy and enthalpy corrections. The W1-F12 values have a much smaller estimated error window than the experiment and are thus used as the reference in the following. Geometries calculated on the same DFT level were used throughout (Figure 6). The calculated reference reaction enthalpies amount to -48.7 and -19.0 kcal mol^{-1}, and thus lie within the large error bars of the experimental values. Specifically, the dimerization reaction represents a challenging task for computational methods as two new strained single bonds are formed and the product includes pronounced intramolecular London dispersion as well as exchange-correlation interactions between the close aromatic rings.

Accordingly, a large deviation is obtained at the dispersion-uncorrected B3LYP/QZ level, which strongly underestimates the reaction enthalpy ($\Delta H_{\text{B3LYP/QZ}} = -25.2$ kcal mol^{-1}). Including the D4 dispersion correction

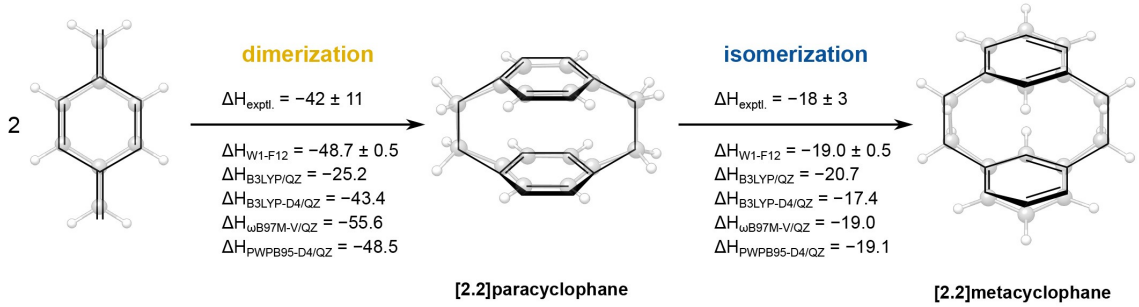

Figure 6. Formation of [2.2]paracyclophane[207] from 3,6-dimethylidenecyclohexa-1,4-diene[208] and subsequent isomerization to [2.2]metacyclophane.[209] All DFT reaction enthalpies employ PBE0-D4/def2-TZVP geometries, zero-point vibrational energy, and enthalpy corrections at $T = 298.15$ K. All values in kcal mol^{-1}. W1-F12 denotes a highly accurate wavefunction-theory-based reference level. QZ = def2-QZVP.

results in a much-improved value of $\Delta H_{\text{B3LYP-D4/QZ}} = -43.4$ kcal mol^{-1}. Nevertheless, the deviation from the W1-F12 reference value still amounts to 5.1 kcal mol^{-1}.

An even better agreement with the reference value is obtained by employing the PWPB95-D4/def2-QZVP double-hybrid. The respective reaction enthalpy is calculated to $\Delta H_{\text{PWPB95-D4/QZ}} = -48.5$ kcal mol^{-1} in near-perfect agreement with the high-level W1-F12 reference.

The strain-reducing isomerization is less prone to intrinsic functional errors, mainly due to beneficial error cancellation resulting from the chemical similarity of [2.2]paracyclophane and [2.2]metacyclophane. Accordingly, all tested DFT methods are in reasonable agreement and again, the PWPB95-D4/QZ result is in perfect agreement with the high-level W1-F12 reference value. In both cases, the energy to enthalpy corrections are small compared to the relative electronic energies, only contributing by 4.2 and 0.5 kcal mol^{-1}, respectively.

4.2. Noncovalent Interactions

Noncovalent interactions (NCIs) play an important role in chemistry, particularly in bio- and supramolecular systems.[211,212] The theoretical description of supramolecular complexes, and particularly the prediction of binding free energies (ΔG), is therefore of considerable importance.. Since chemically relevant systems are often quite large and also flexible, this task is challenging for computational chemistry. To compare calculated values and experimental data such as binding free energies measured in solution, the thermostatistical corrections must also include the entropy terms. This is specifically the case for bimolecular reactions. The common approach to calculating binding free energies for the formation of a NCI complex at a given temperature is shown in Equation (1),[180]

$$\Delta G = \Delta E + \Delta\delta G_{\text{solv}} + \Delta G_{\text{mRRHO}} \quad (1)$$

where ΔE refers to the difference of the total electronic gas-phase energies, $\Delta\delta G_{\text{solv}}$ to the difference in solvation free energies, and ΔG_{mRRHO} corresponds to the difference in the thermostatistical contributions. Depending on the details of the complex in question (e.g., charged vs. neutral, H-bonds vs. π–π stacking, solvent), these individual contributions may vary significantly in size and can be of different sign. Regardless, all three contributions to ΔG must be described fairly precisely, as these are typically large and only partially cancel each other to give the typically rather small experimental ΔG values (−1 to −15 kcal mol^{-1}).[203]

We have recently proposed a workflow[64] to compute the contributions to Equation (1), which is summarized in Figure 7 and showcased here. The example complex is a +4 charged macrocyclic host (termed CBPQT4+) with bromobenzene as guest, whose ΔG was experimentally determined to be −5 kcal mol^{-1} in aqueous solution.[213] The first challenge is to model the molecular structure in solution at finite temperature as realistically as possible. This is accomplished by a conformer search using the automatic

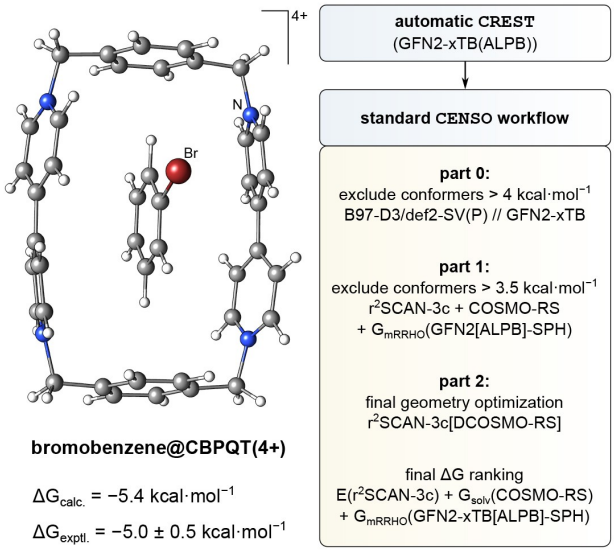

Figure 7. Suggested computational workflow to calculate the binding free energy (ΔG) of bromobenzene to CBPQT(4+) in water. For details on the *CREST/CENSO* workflow see Refs. [66, 67], and [64].

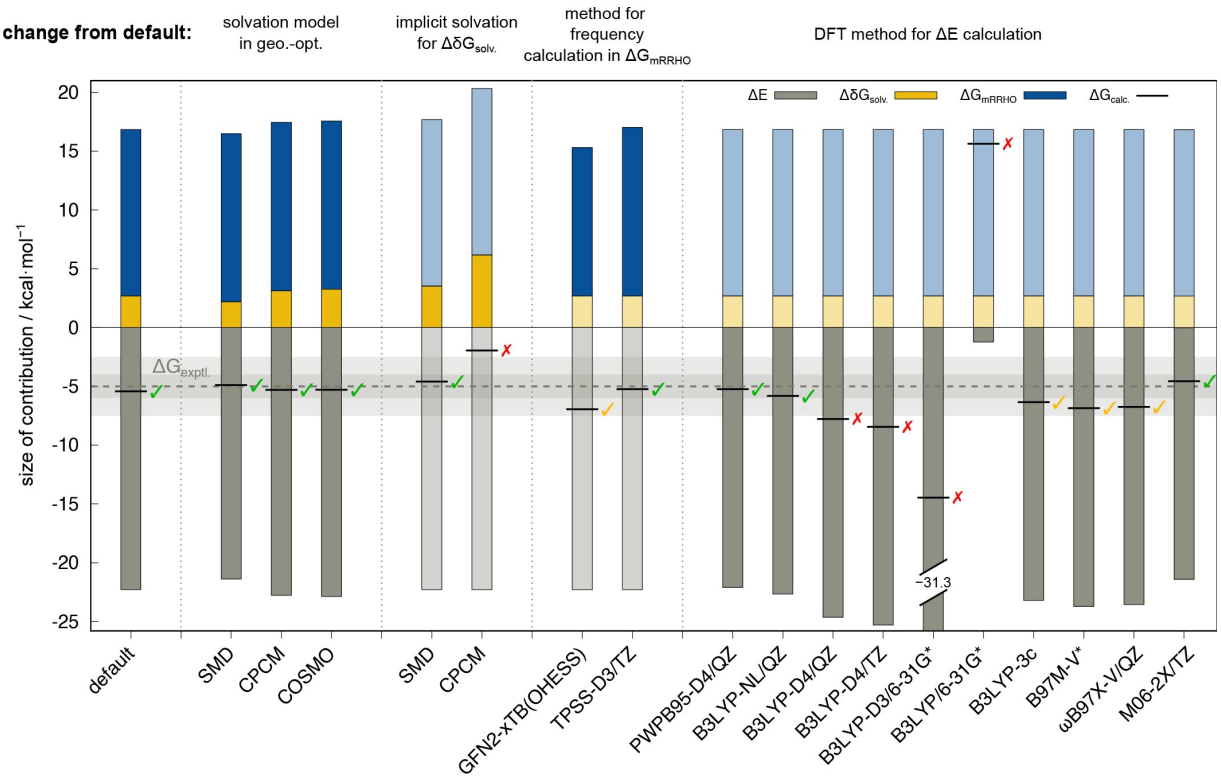

Figure 8. Breakdown of the contributions to the overall calculated $\Delta G_{calc.}$ with the final result represented by a black line, and experimental reference given as gray dashed line with the shaded areas marking the estimated error range of ± 1.0 kcal mol^{-1} and a more conservative error estimate of ± 2.5 kcal mol^{-1}. Methods that do not reach the ± 1.0 kcal mol^{-1} window but are within ± 2.5 kcal mol^{-1} of the experimental reference value are marked by yellow ticks. The leftmost bar represents the default approach (ΔE(r^{2}SCAN-3c) + $\Delta\delta G_{solv}$(COSMO-RS) + ΔG_{mRRHO}(GFN2-xTB[ALPB]-SPH) at r^{2}SCAN-3c[DCOSMO-RS] geometry), while the other bars illustrate the effect of selected method variations as indicated at the top (solvation, frequencies, electronic energy). Contributions that are not affected by these variations are depicted in brighter colors. B97M-V* = B97M-V/def2-SVPD/DFT-C. B3LYP-NL utilizes VV10 with refitted parameters.[78] OHESS = GFN2-xTB[ALPB] frequencies with a re-optimized structure instead of SPH with the DFT structure.

CREST and *CENSO* approaches. In Ref. [53], we demonstrated that the efficient semi-empirical extended tight-binding method GFN2-xTB with a specially adapted implicit solvation model (ALPB) is well-suited to conduct the initial search efficiently. The main goal of the conformer search is to find the energetically most favorable conformer, as this is typically sufficient in the case of rather rigid molecules like here. In contrast, if the host or guest structures are significantly more flexible, it is advisable to consider other low-lying conformers as well (note that already for this relatively rigid system, we find twelve conformers in an energy window of 2.5 kcal mol^{-1}). Contributions from other conformers can become important if there are large changes to the accessible configuration space upon complex formation, e.g., if a previously freely rotating alkyl chain is conformationally locked in the complex. Since a further discussion of this aspect is beyond the scope of this work, we refer interested readers to Refs. [67] and [214].

For the final geometry optimization (and conveniently also the final energy calculation, see below) of the most stable conformer, we recommend the composite DFT method r^{2}SCAN-3c with DCOSMO-RS as implicit solvation model. Alternatively, SMD, CPCM, and COSMO with descending preference from first to last can be used. Note that COSMO-RS cannot be used in geometry optimizations. For the example discussed here, all four mentioned implicit solvation models are suitable for the geometry optimization as evident from Figure 8. Moreover, we note that also our previous group-default for geometries, a TPSS-D3/def2-TZVP optimization in the gas phase, gives accurate results for this example and also in general. This is because results are often not very sensitive to the employed structures, as already discussed in Section 3.1.

The thermostatistical corrections ΔG_{mRRHO} typically provide a significant positive, repulsive contribution to the free binding energy (blue bars in Figure 8). We calculate them in the modified rigid-rotor harmonic oscillator (mRRHO) approximation,[180] which includes a special treatment for low-frequency modes, zero-point vibrational energy (ZPVE), and heat/volume work corrections, but neglects the conformational entropy. This can be a good approximation if the involved molecules are mostly rigid as discussed above. Our default is to compute vibrational frequencies with fast semi-empirical methods such as GFN2-xTB, which is typically in good agreement with DFT reference values (deviation $\lesssim$1–2 kcal mol^{-1}).[215] This requires a suitable implicit solvation model like ALPB as well as the single-point Hessian (SPH) ansatz,[162] which uses a

biasing potential approach to create an artificial minimum at the DFT structure. The comparison of the default GFN2-xTB[ALPB]-SPH $\Delta G_{\mathrm{mRRHO}}$ contributions with corresponding DFT values obtained in the gas phase at the TPSS-D3/TZ level shown in Figure 8 confirms that the semi-empirical approach yields practically identical results but in a few minutes vs. several hours of computation time for the DFT calculation. The alternative to SPH calculations, that is, a full re-optimization of the geometry with GFN2-xTB (xtb-keyword: *-OHESS*), yields slightly larger deviations from the gas-phase DFT reference (see Figure 8), but the SPH calculations benefit also from error cancellation due to the use of an implicit solvation model. Hence, we generally recommend calculating the frequencies with an implicit solvation model and the SPH algorithm.

To calculate the binding energy (ΔE) contribution to ΔG, our default protocol uses the r^{2}SCAN-3c single-point energy. In general, composite DFT methods such as r^{2}SCAN-3c or B97M-V/def2-SVPD/DFT-C are efficient alternatives to numerically converged QZ basis set DFT calculations and approximately 50 times faster for this system size. As evident from Figure 8, r^{2}SCAN-3c (default, leftmost bar), B3LYP-3c, and B97M-V/DFT-C are all in good agreement with the reference. However, it should be noted that when the systems are highly charged or feature exotic chemical interactions, it is generally advisable to compare ΔE obtained with composite methods to a more robust DFT/QZ calculation with an accurate hybrid (e.g., ωB97M-V) to be on the safe side. The remaining scatter of typically ± 2 kcal mol^{-1} for ΔE can usually be attributed in about half to the errors of the density functional and in the other half to the errors of the dispersion correction, which is particularly important in NCI examples (see the discussion in Section 3.7). For this example, VV10/NL is actually more accurate in combination with B3LYP than D4 (see Figure 8). Concerning the basis set for the DFT energy calculation, the residual basis set errors are smaller than the intrinsic functional error with QZ basis sets (often already with large TZ basis sets). However, this changes when the basis set size is further reduced to DZ, as evident from the hierarchy of B3LYP-based results summarized in Figure 8: With the TZ basis, the B3LYP-D4 results change only slightly compared to QZ (and also M06-2X provides very good agreement with the TZ basis), whereas the DZ basis 6-31G* with B3LYP-D3 exhibits strong BSSE-induced overbinding. In contrast, the still widely used B3LYP/6-31G* approach drastically underestimates the ΔE contribution due to the lack of a dispersion correction, and thus the resulting ΔG is off by more than 20 kcal mol^{-1}. This demonstrates that the often-assumed error cancellation between lacking London dispersion and BSSE in B3LYP/6-31G* can not be trusted. To show how the shortcomings of B3LYP/6-31G* can be fixed at essentially no extra cost, we included the value obtained with the B3LYP-3c approach, which combines B3LYP-D3 with a DZ basis (def2-SVP) and the default gCP correction for BSSE (see Figure 8). Evidently, this physically sound approach gives results very close to B3LYP-D4/QZ at the same cost as B3LYP/6-31G*, demonstrating once again that there is no reason for using this outdated but still popular method.[11]

The most challenging contribution in the entire workflow is the solvation free energy contribution ΔG_{solv}, especially for higher or negatively charged systems in polar, H-bonding solvents. Here, an error range of 2–3 kcal mol^{-1} is realistic due to a large $\Delta\delta G_{\mathrm{solv}}$ value, which has an estimated intrinsic error of 10–20 %, even with the best implicit solvation models available. Among them, according to our experience, COSMO-RS yields the most reliable results, while SMD can serve as a good alternative. In contrast, purely electrostatic models like COSMO and CPCM often perform worse because they neglect all non-electrostatic terms, which are particularly important if the solvent-accessible molecular surface changes during the reaction (as is the case here, see Figure 8).

Due to fortuitous error cancellation effects, the free binding energy produced by our standard workflow is typically accurate to within 1–3 kcal mol^{-1} from experimentally determined values,[203] which is much less than the sum of the maximum errors of the individual contributions would suggest. Also, for this example, the calculated ΔG value of -5.4 kcal mol^{-1} agrees very well with the experimentally determined one for a wide range of method combinations. Moreover, the presented workflow has the advantage of being computationally quite fast, requiring only about 30 hours in total on a common eight-core CPU for the shown example. This efficiency largely results from the use of fast semi-empirical methods for the frequencies and efficient mGGA-based composite DFT methods for geometries and single-point energies. With such an efficient multi-level scheme, reliable affinity predictions are possible for much larger complexes with up to 200–300 atoms in practical computation times. Finally, we want to mention that the protocol shown in Figure 4 is fully automated via the freely available *CREST* and *CENSO* programs, and can thus be invoked via two simple *UNIX* commands as described in the documentation.[216,217]

4.3. Optical Rotation of α-/β-D-Glucopyranose

The calculation of relative electronic energies is fundamental and can indirectly have a crucial impact on property calculations when ensemble-averaged treatments are used. This is the case for the example of computed specific optical rotation values of a mixture of α- and β-glucopyranose depicted in Figure 9.[218] The optical rotation strongly depends on details of the molecular structure, and for flexible systems, on a reliable calculation of the relevant populated conformers. Further, the values of α- and β-D-glucopyranose isomers differ strongly from each other, and thus the accurate description of the thermodynamic equilibrium between both forms is crucial.

As this example is typical for the complex decision-making in computational chemistry protocols, the application of the conceptual flowchart illustrated in Figure 2 will be demonstrated in the following.

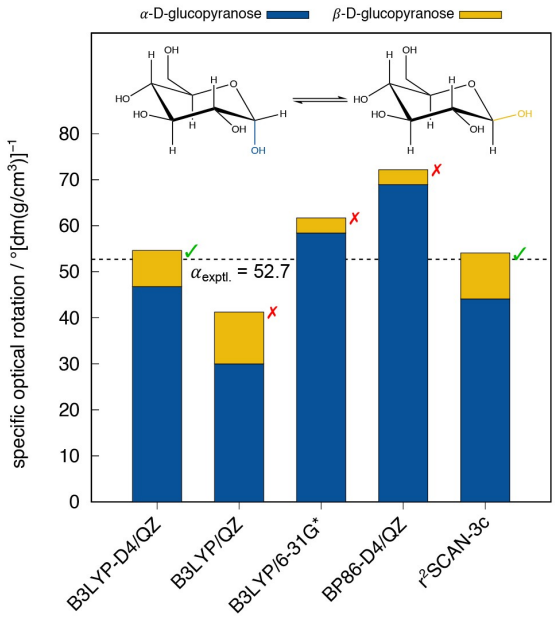

Figure 9. Calculated, Boltzmann-weighted optical rotation for the equilibrium of α- and β-D-glucopyranose in water at 20 °C. For each anomer, 15–18 conformers are considered in the example. QZ = def2-QZVP.

First, the investigated molecules are typical organic molecules that are known to bear no multi-reference character. Accordingly, we employed the single-reference electronic structure methods DFT in combination with semi-empirical methods (for the initial conformer search).

Secondly, since the optical rotation value was measured in solution, a solvation model was applied in all steps. In this case, the individual technical limitations (see below) lead to the choice of DCOSMO-RS for the geometry optimizations, COSMO for the optical rotation value calculation, and COSMO-RS for the free energy evaluation. COSMO-RS yields improved results for free energy calculations but is not available for geometry optimizations. Therefore, DCOSMO-RS, an improved variant of COSMO, is applied in the geometry optimization step. Since neither of the aforementioned is available for the calculation of optical rotation values, the most simple COSMO solvation model is the remaining choice.

Thirdly, we consider the role of conformers and molecular flexibility. Evidently, this is of particular importance here since α- and β-D-glucopyranose both bear a pronounced structural flexibility due to the large number of hydroxy substituents. Accordingly, a conformational search algorithm was used to identify the energetically lowest and thus relevant conformers. In this case, we employed the *CREST* in combination with *CENSO*. While *CREST* uses a fast semi-empirical level to find the low-lying conformers, *CENSO* automatically refines their energies and structures at a more accurate DFT level.

Next step is the choice of the functional. The r²SCAN-3c composite method has proven to yield both, good geometries and conformational free energies, as discussed in section 3.4. Therefore, it is chosen here as the default method. Further alternatives are evaluated and discussed in detail below. Since no heavy elements are present in the molecules, no relativistic treatment has to be considered. Also, the choice of the basis set is implicit since we are using a "3c" method, which comes with a built-in basis (def2-mTZVPP).

Finally, to optimize the ensembles, the *CREST*/*CENSO* approach with the r²SCAN-3c[COSMO-RS]//r²SCAN-3c[DCOSMO-RS] level was applied. Each ensemble includes 15–18 relevant (low-lying) conformers per anomer, which are considered in the Boltzmann-averaging of the target property. This target property is the specific optical rotation value, whose calculation requires specialized settings. It was found that for this task, the PBE GGA functional with the augmented double-zeta basis set def2-SVPD and the COSMO solvation model is sufficient. Note that usually hybrid functionals are recommended for this purpose, whereas PBE was chosen here for technical reasons (origin-independence of OR calculation) after careful testing against B3LYP results (see Ref. [218] for details). The double-zeta quality basis set is sufficient as the target property is generally not very basis set dependent. However, it was shown that the augmentation with diffuse functions is important to obtain reliable results.[218]

Even though this complex protocol involves many specific decisions, as illustrated above, the perhaps most obvious choice is that of the density functional for the calculation of the electronic energies. The impact of this choice on the final optical rotation values (via Boltzmann-weighting) is discussed in the following.

The conformer ensembles obtained from the aforementioned protocol are here re-ranked employing as an example B3LYP method combinations (B3LYP/QZ, B3LYP-D4/QZ, B3LYP/6-31G*, QZ = def2-QZVP), as well as BP86[177,220]-D4/QZ for the electronic energy contribution. All others (zero-point vibrational energy, thermostatistical, and solvation) were taken from the preceding r²SCAN-3c calculation. In agreement with results from benchmark studies, the r²SCAN-3c composite method yields a very good energetic ranking, such that the Boltzmann-weighted optical rotation value is in excellent agreement with the experimental value of $\alpha_{exptl.} = 52.7$ (all values in the following in the usual degree $dm(g/cm^3)^{-1}$ units). The same holds for the value that is based on a B3LYP-D4/QZ ranking amounting to $\alpha_{B3LYP\text{-}D4/QZ} = 54.6$. For the plain B3LYP/QZ level, a worse result of $\alpha_{B3LYP/QZ} = 41.2$ is obtained, underlining the indispensability of a London dispersion correction even for relatively small molecules. The frequently used B3LYP/6-31G* approach relies on a difficult-to-control error cancellation between neglected London dispersion and BSSE. Nevertheless, the contribution of α-D-glucopyranose is overestimated, resulting in a relatively bad value of $\alpha_{B3LYP/6\text{-}31G^*} = 61.7$. The pure GGA functional BP86-D4/QZ, which in general does not perform well for conformational energies, yields an even worse agreement with the experiment ($\alpha_{BP86\text{-}D4/QZ} = 72.2$). This example clearly demonstrates that the best results are obtained if all physically and technically

relevant effects, such as London dispersion and basis-set completeness, are properly taken into account. Furthermore, conformation-sensitive properties like optical rotation can indirectly be used to assess the quality of theoretical approximations.

4.4. Reaction Barriers

4.4.1. S_N2 and Diels–Alder Reactions

The reliable calculation of reaction barriers is crucial for the investigation of complex reaction mechanisms. They allow for a deeper understanding of key reactions and thus the computer-aided design of novel catalysts and targeted tuning of chemical reactions.[221] S_N2 and Diels–Alder reactions represent well-known reaction types in organic chemistry. Although these basic reaction types seem comparably simple, their theoretical description still requires a profound choice of the applied quantum chemical method. Accurate reference electronic activation energies (ΔE) are available for the [4+2] cycloaddition of ethylene to cyclopentadiene (Figure 10a) and the nucleophilic attack of OH^- to fluoromethane (Figure 10b).[27] For optimal comparability, the original geometries at the B3LYP/6-311G(2d,d,p) and QCISD/ MG3 were used. For both examples, the BP86-D3/QZ (QZ = def2-QZVP) GGA strongly underestimates the reaction barrier. Specifically, the reaction barrier of the S_N2 reaction is underestimated by 10.2 kcal mol^{-1}. Here, the pronounced charge-delocalization in the transition state causes SIE-related issues with the GGA method. The mGGA M06-L/QZ yields much improved results, which may be attributed to its highly empirical character, with the functional being also trained on reproducing barrier heights in similar systems. Nevertheless, hybrid functionals such as PBE0 are expected to yield improved results over the (m)GGA methods by an enhanced physical description and for both reactions. Accordingly, PBE0-D3/QZ yields a significant improvement over BP86-D3/QZ, while M06-2X is more accurate than M06-L. For M06-2X, the very large amount of Fock exchange is beneficial in this special case, reducing the SIE drastically. Nevertheless, functionals with very high amounts of Fock exchange should be used with caution (cf. Section 2.5). Finally, the range-separated hybrid functionals ωB97X-V/QZ and ωB97M-D4/QZ yield good results for both reactions, underlining their robustness. ωB97X-V/QZ underestimates the S_N2 reference reaction barrier only slightly by –1.8 kcal mol^{-1}. It also yields comparably good agreement for the Diels–Alder activation reaction, overestimating the reference value by only 1.8 kcal mol^{-1}.

4.4.2. Hydrogen-Atom-Transfer Reaction

Another challenging example reaction is depicted in Figure 11. Here, a hydrogen atom is transferred from a molybdenum hydride complex TpMo(CO)$_3$H (Tp = trispyrazolylborate) to a Gomberg-type (tBu-4-C$_6$H$_4$)$_3$C$^\bullet$ radical.[222] The experimental free energy activation barrier was determined as $\Delta G = 19.2$ kcal mol^{-1}. As discussed in the previous sections and the electronic activation energy examples, transition states with stretched covalent bonds are prone to SIE, and thus a significant underestimation of the reaction barrier by (m)GGA methods is expected. This especially applies to bonds involving hydrogen, where the one-electron character is usually large. Compared with the previous example, the desired property is an activation free energy in solution, involving additionally ZPVE, thermal, solvation, and entropy effects. Even though typical (m)GGA functionals should be used with care regarding the computation of electronic energies of transition states, they typically yield reasonable geometries and thermostatistical corrections even for such cases. Accordingly, the B97-3c composite method was used for geometry optimization and frequency calculation in this example. For this reaction, the BP86-D4/QZ

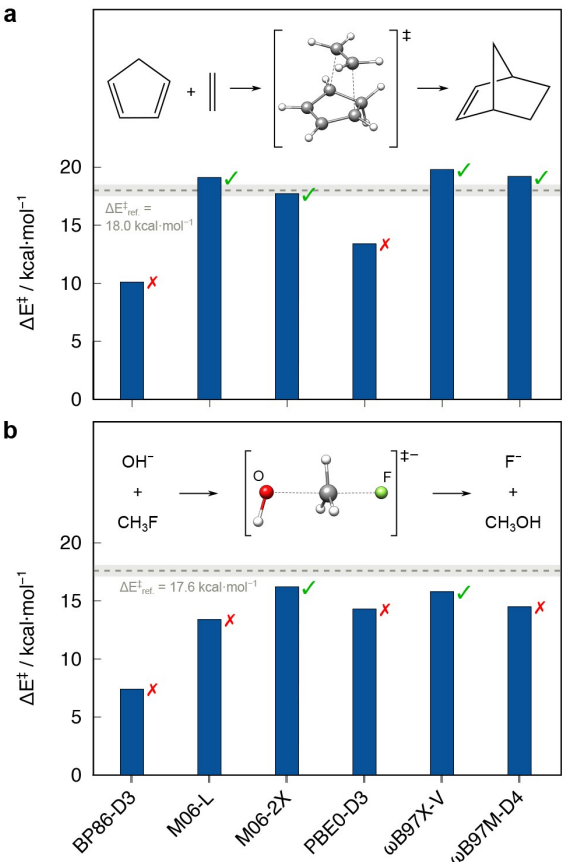

Figure 10. Calculated gas-phase electronic reaction barriers (ΔE) for selected functionals using the def2-QZVP basis set[27] for a) the Diels–Alder cycloaddition of cyclopentadiene and ethylene (reference level: accurate W1-F12 wavefunction theory at B3LYP/6-311G(2d,d,p) geometry), and b) the S_N2 reaction of fluoromethane with a hydroxide anion (reference level: very accurate W2-F12 wavefunction theory at QCISD/MG3 geometry. MG3 is a modified 6-311++G(3d2f,2df,2p) basis set). The light-gray area indicates the expected error range of the reference method.

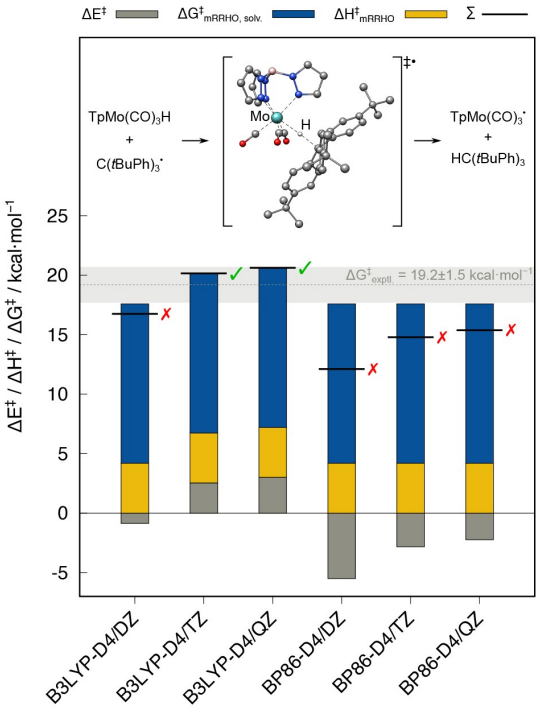

Figure 11. Calculated contributions to the activation free energy ΔG for a metal-centered hydrogen-transfer reaction based on hybrid (B3LYP-D4) and GGA (BP86-D4) electronic energies (gray bars), B97-3c thermostatistical contributions (yellow bars), and COSMO-RS solvation and entropic corrections (blue bars). Black lines represent the sum of all contributions yielding ΔG. DZ = def2-SVP; TZ = def2-TZVP; QZ = def2-QZVPP basis set. The light-gray area indicates the expected error-range of the reference method.

functional used for the electronic energy underestimates ΔG by almost 4 kcal mol^{-1}. This is even more pronounced upon decreasing the basis set size from QZ to TZ to DZ with a severe underestimation by over 7 kcal mol^{-1} for BP86-D4/DZ. With the B3LYP-D4 hybrid functional, ΔG is slightly overestimated by 1.4 kcal mol^{-1} yet close to the chemical accuracy window of 1 kcal mol^{-1}. Nevertheless, also for hybrid functionals, small basis sets result in an underestimation of the reaction barrier which seems to be rather sensitive here. The GGA functional even yields a negative electronic activation energy of $\Delta E = -2.2$ kcal mol^{-1} (BP86-D4/QZ) and the transition state does not represent a stationary point at the corresponding potential energy surface. B3LYP-D4/QZ yields a small, yet reasonable ΔE value of 3.0 kcal mol^{-1}. The small electronic activation energy further underlines the importance of solvation, enthalpy, and free energy corrections for a direct comparison to the experiment. The free activation barrier is only well-described upon inclusion of all relevant contributions (corrections computed at the geometry optimization level B97-3c, solvation corrections by COSMO-RS), leading to an increase of the barrier from 3.0 (ΔE) to 7.2 (ΔH) and finally 20.6 kcal mol^{-1} (ΔG).

5. Perspectives

The development of quantum chemistry over the last 20–30 years, and foremost DFT, is a great success story. The fact that nowadays, non-experts can do reasonable quantum chemistry calculations for large, chemically relevant systems on desktop computers is fantastic. The influence of having widely available computational tools on chemical research cannot be understated and will presumably grow even larger in the future. We hope that the guidelines and recommendations given in this work help to increase the reliability of DFT-based quantum chemistry predictions in the daily work of many chemists.

We strongly emphasize the aspect of finding the right methodological compromise between computational effort (speed) and desired accuracy, while still obtaining as close as possible "the right answer for the right reason". An often-overlooked aspect of this balance is that faster theoretical methods enable a more extensive—and hence more reliable—study of the system under investigation, for example regarding its conformational behavior, molecular dynamics, or explicit solvation issues. This is of particular importance since, in our experience, errors and deviations due to the neglect of important low-lying conformers (ensemble properties vs. individual molecule property) can be even larger than the errors in the electronic energy by the functional or basis set approximations. We thus want to motivate the reader to conduct systematic conformational searches, explore the dynamical behavior by means of MD simulations, consider explicit solvation treatments more routinely with currently developed methods, and employ efficient multi-level approaches for chemically realistic models.

However, also on the electronic structure side of the problem, there are challenges ahead. Although many very relevant chemical properties and problems nowadays can be solved by standard DFT treatments as described here, there are still problematic systems (e.g., open-shell transition metal complexes), open questions (e.g., how to treat strongly solvated, highly charged systems), and problems that are fundamentally difficult (entropy). For those aspects, non-standard treatments and expert knowledge are often required and should be involved.

In our opinion, the positive development of the basic density functionals over the last decade has slowed down and entered a kind of saturation regime in terms of the "peak" accuracy achieved. Nevertheless, we note some promising new developments that are already applicable such as local hybrid functionals[223–225] or non-self-consistent field treatments with cheap (m)GGAs to avoid exchange-correlation-potential-driven errors.[226–228] Currently, machine-learned functionals[229–231] are still in their infancy, but the potential of such extremely empirical (and in no way "cheap") methods could be high, at least for organic and main-group compounds. In turn, DFT is the method of choice for generating extremely large databases with millions of compound entries for machine-learning algorithms.[232–234] In this context, we still see a great potential to make the best-performing functionals

(which are sufficiently accurate for about >90% of all chemical applications) significantly faster but without losing robustness or numerical accuracy. Central here are accurate yet efficient approximations of the non-local Fock exchange of hybrid or double-hybrid density functionals and, for (m)GGA functionals, the numerical integration of the semi-local exchange-correlation energy. Because DFT is a rather general "first-principles" approach, we believe that most of our conclusions also hold true for computations of (molecular) solids and liquids under periodic boundary conditions, although we do not explicitly consider these here. For more special cases like low-band-gap systems, such as metallic solids, however, things may change more drastically, such that a consideration of special technical settings may be required to avoid a fundamental breakdown of approximations.

Acknowledgements

The German Science Foundation (DFG) is gratefully acknowledged for financial support through a Gottfried Wilhelm Leibniz prize to S.G. and the SPP 2363 "Utilization and Development of Machine Learning for Molecular Applications—Molecular Machine Learning". Further, S.G. and M.B. gratefully acknowledge financial support of the Max Planck Society through the Max Planck fellow program. The authors further thank the *ORCA*, *TURBOMOLE* (especially Uwe Huniar), and *AMS* developer teams for steady support, as well as Johannes Gorges for conducting some calculations and providing helpful corrections to the manuscript. We also thank Uwe Hohm, Tunga Salthammer, Wolf Palm, Sigrid Peyerimhoff, Marcel Bamberg, Christoph Plett, Matthias Wagner, Demyan Prokopchuk, and Julia Bursch-Pöppinghaus for comments and fruitful discussions. The thorough and valuable feedback from Susi Lehtola is especially acknowledged. Further, the authors thank the whole Grimme group for providing a cornucopia of diverse experiences and for proofreading. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

S.G. is associated with the *ORCA* and *TURBOMOLE* development teams and has a consulting contract with the FAccTs GmbH.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords: Computational Chemistry · Density Functional Calculations

[1] A. M. Teale, T. Helgaker, A. Savin, C. Adamo, B. Aradi, A. V. Arbuznikov, P. W. Ayers, E. J. Baerends, V. Barone, P. Calaminici, E. Cancès, E. A. Carter, P. K. Chattaraj, H. Chermette, I. Ciofini, T. D. Crawford, F. De Proft, J. F. Dobson, C. Draxl, T. Frauenheim, E. Fromager, P. Fuentealba, L. Gagliardi, G. Galli, J. Gao, P. Geerlings, N. Gidopoulos, P. M. W. Gill, P. Gori-Giorgi, A. Görling, T. Gould, S. Grimme, O. Gritsenko, H. J. A. Jensen, E. R. Johnson, R. O. Jones, M. Kaupp, A. M. Köster, L. Kronik, A. I. Krylov, S. Kvaal, A. Laestadius, M. Lewin, S. Liu, P.-F. Loos, N. T. Maitra, F. Neese, J. P. Perdew, K. Pernal, P. Pernot, P. Piecuch, E. Rebolini, L. Reining, P. Romaniello, A. Ruzsinszky, D. R. Salahub, M. Scheffler, P. Schwerdtfeger, V. N. Staroverov, J. Sun, E. Tellgren, D. J. Tozer, S. B. Trickey, C. A. Ullrich, A. Vela, G. Vignale, T. A. Wesolowski, X. Xu, W. Yang, *Phys. Chem. Chem. Phys.* **2022**, https://doi.org/10.1039/D2CP02827A.
[2] K. N. Houk, F. Liu, *Acc. Chem. Res.* **2017**, *50*, 539–543.
[3] S. Grimme, P. R. Schreiner, *Angew. Chem. Int. Ed.* **2018**, *57*, 4170–4176; *Angew. Chem.* **2018**, *130*, 4241–4248.
[4] T. Husch, A. C. Vaucher, M. Reiher, *Int. J. Quantum Chem.* **2018**, *118*, e25799.
[5] C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, S. Grimme, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2021**, *11*, e1493.
[6] M. Elstner, G. Seifert, *Philos. Trans. R. Soc. A* **2014**, *372*, 20120483.
[7] C. Riplinger, P. Pinski, U. Becker, E. F. Valeev, F. Neese, *J. Chem. Phys.* **2016**, *144*, 024109.
[8] C. Riplinger, B. Sandhoefer, A. Hansen, F. Neese, *J. Chem. Phys.* **2013**, *139*, 134101.
[9] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5656.
[10] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623–11627.
[11] A. Castro-Alvarez, H. Carneros, D. Sánchez, J. Vilarrasa, *J. Org. Chem.* **2015**, *80*, 11977–11985.
[12] P. R. Schreiner, A. A. Fokin, R. A. Pascal, A. De Meijere, *Org. Lett.* **2006**, *8*, 3635–3638.
[13] M. D. Wodrich, C. Corminboeuf, P. v. R. Schleyer, *Org. Lett.* **2006**, *8*, 3631–3634.
[14] P. Morgante, R. Peverati, *Int. J. Quantum Chem.* **2020**, *120*, e26332.
[15] E. Braun, J. Gilmer, H. B. Mayes, D. L. Mobley, J. I. Monroe, S. Prasad, D. M. Zuckerman, *Living J. Comp. Mol. Sci.* **2018**, *1*, 5957.
[16] N. Mardirossian, M. Head-Gordon, *Mol. Phys.* **2017**, *115*, 2315–2372.
[17] P. Verma, D. G. Truhlar, *Trends Chem.* **2020**, *2*, 302–318.
[18] S. Grimme, A. Hansen, J. G. Brandenburg, C. Bannwarth, *Chem. Rev.* **2016**, *116*, 5105–5154.
[19] H. Kruse, S. Grimme, *J. Chem. Phys.* **2012**, *136*, 154101.
[20] J. Witte, J. B. Neaton, M. Head-Gordon, *J. Chem. Phys.* **2017**, *146*, 234105.
[21] P. Pracht, D. F. Grant, S. Grimme, *J. Chem. Theory Comput.* **2020**, *16*, 7044–7060.
[22] S. Grimme, A. Hansen, S. Ehlert, J. M. Mewes, *J. Chem. Phys.* **2021**, *154*, 064103.
[23] J. A. Van Santen, G. A. Dilabio, *J. Phys. Chem. A* **2015**, *119*, 6703–6713.
[24] Y. Mao, M. Loipersberger, P. R. Horn, A. Das, O. Demerdash, D. S. Levine, S. Prasad Veccham, T. Head-Gordon, M. Head-Gordon, *Annu. Rev. Phys. Chem.* **2021**, *72*, 641–666 PMID: 33636998.
[25] M. von Hopffgarten, G. Frenking, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 43–62.

[26] E. D. Glendening, C. R. Landis, F. Weinhold, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 1–42.
[27] L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi, S. Grimme, *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184–32215.
[28] M. Korth, S. Grimme, *J. Chem. Theory Comput.* **2009**, *5*, 993–1003.
[29] J. Conradie, A. Ghosh, *J. Phys. Chem. B* **2007**, *111*, 12621–12624.
[30] N. Chéron, D. Jacquemin, P. Fleurat-Lessard, *Phys. Chem. Chem. Phys.* **2012**, *14*, 7170–7175.
[31] J. G. Brandenburg, C. Bannwarth, A. Hansen, S. Grimme, *J. Chem. Phys.* **2018**, *148*, 064104.
[32] S. Grimme, J. G. Brandenburg, C. Bannwarth, A. Hansen, *J. Chem. Phys.* **2015**, *143*, 054107.
[33] E. Caldeweyher, J. G. Brandenburg, *J. Phys. Condens. Matter* **2018**, *30*, 213001.
[34] A. Harju, E. Räsänen, H. Saarikoski, M. J. Puska, R. M. Nieminen, K. Niemelä, *Phys. Rev. B* **2004**, *69*, 153101.
[35] L. Noodleman, *J. Chem. Phys.* **1998**, *74*, 5737.
[36] C. L. Janssen, I. M. Nielsen, *Chem. Phys. Lett.* **1998**, *290*, 423–430.
[37] T. J. Lee, P. R. Taylor, *Int. J. Quantum Chem.* **1989**, *36*, 199–207.
[38] O. Tishchenko, J. Zheng, D. G. Truhlar, *J. Chem. Theory Comput.* **2008**, *4*, 1208–1219.
[39] A. Karton, S. Daon, J. M. Martin, *Chem. Phys. Lett.* **2011**, *510*, 165–178.
[40] S. R. Langhoff, E. R. Davidson, *Int. J. Quantum Chem.* **1974**, *8*, 61–72.
[41] C. A. Bauer, A. Hansen, S. Grimme, *Chem. Eur. J.* **2017**, *23*, 6150–6164.
[42] S. Grimme, A. Hansen, *Angew. Chem. Int. Ed.* **2015**, *54*, 12308–12313; *Angew. Chem.* **2015**, *127*, 12483–12488.
[43] L. R. Maurer, M. Bursch, S. Grimme, A. Hansen, *J. Chem. Theory Comput.* **2021**, *17*, 6134–6151.
[44] F. Neese, *J. Biol. Inorg. Chem.* **2006**, *11*, 702–711.
[45] V. Barone, M. Cossi, *J. Phys. Chem. A* **1998**, *102*, 1995–2001.
[46] A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B* **2009**, *113*, 6378–6396.
[47] A. Klamt, G. Schüürmann, *J. Chem. Soc. Perkin Trans. 2* **1993**, 799–805.
[48] A. Klamt, *J. Phys. Chem.* **1995**, *99*, 2224–2235.
[49] A. Klamt, M. Diedenhofen, *J. Phys. Chem. A* **2015**, *119*, 5439–5445.
[50] W. Clark Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.
[51] A. W. Lange, J. M. Herbert, *J. Chem. Theory Comput.* **2012**, *8*, 1999–2011.
[52] G. Sigalov, A. Fenley, A. Onufriev, *J. Chem. Phys.* **2006**, *124*, 124902.
[53] S. Ehlert, M. Stahn, S. Spicher, S. Grimme, *J. Chem. Theory Comput.* **2021**, *17*, 4250–4261.
[54] D. Sicinska, P. Paneth, D. G. Truhlar, *J. Phys. Chem. B* **2002**, *106*, 2708–2713.
[55] R. Glaves, G. Mathias, D. Marx, *J. Am. Chem. Soc.* **2012**, *134*, 6995–7000.
[56] C. Michel, F. Auneau, F. Delbecq, P. Sautet, *ACS Catal.* **2011**, *1*, 1430–1440.
[57] J. M. Park, A. Laio, M. Iannuzzi, M. Parrinello, *J. Am. Chem. Soc.* **2006**, *128*, 11318–11319.
[58] Y. Basdogan, J. A. Keith, *Chem. Sci.* **2018**, *9*, 5341–5346.
[59] G. N. Simm, P. L. Türtscher, M. Reiher, *J. Comput. Chem.* **2020**, *41*, 1144–1155.
[60] S. A. Katsyuba, S. Spicher, T. P. Gerasimova, S. Grimme, *J. Phys. Chem. B* **2020**, *124*, 6664–6670.
[61] M. Steiner, T. Holzknecht, M. Schauperl, M. Podewitz, *Molecules* **2021**, *26*, 1793.
[62] S. Spicher, C. Plett, P. Pracht, A. Hansen, S. Grimme, *J. Chem. Theory Comput.* **2022**, *18*, 3174.
[63] R. Sure, M. El Mahdali, A. Plajer, P. Deglmann, *J. Comput.-Aided Mol. Des.* **2021**, *35*, 473–492.
[64] S. Grimme, F. Bohle, A. Hansen, P. Pracht, S. Spicher, M. Stahn, *J. Phys. Chem. A* **2021**, *125*, 4039–4054.
[65] M. Bursch, A. Hansen, P. Pracht, J. T. Kohn, S. Grimme, *Phys. Chem. Chem. Phys.* **2021**, *23*, 287–299.
[66] S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 2847–2862.
[67] S. Grimme, F. Bohle, A. Hansen, P. Pracht, S. Spicher, M. Stahn, *J. Phys. Chem. A* **2021**, *125*, 4039–4054.
[68] J. G. Sobez, M. Reiher, *J. Chem. Inf. Model.* **2020**, *60*, 3884–3900.
[69] K. S. Watts, P. Dalal, R. B. Murphy, W. Sherman, R. A. Friesner, J. C. Shelley, *J. Chem. Inf. Model.* **2010**, *50*, 534–546.
[70] N. Sauton, D. Lagorce, B. O. Villoutreix, M. A. Miteva, *BMC Bioinf.* **2008**, *9*, 1–12.
[71] M. A. Miteva, F. Guyon, P. Tufféry, *Nucleic Acids Res.* **2010**, *38*, W622–W627.
[72] P. C. Hawkins, A. G. Skillman, G. L. Warren, B. A. Ellingson, M. T. Stahl, *J. Chem. Inf. Model.* **2010**, *50*, 572–584.
[73] RDKit: Open-source cheminformatics. https://www.rdkit.org.
[74] M. J. Vainio, M. S. Johnson, *J. Chem. Inf. Model.* **2007**, *47*, 2462–2474.
[75] S. Riniker, G. A. Landrum, *J. Chem. Inf. Model.* **2015**, *55*, 2562–2574.
[76] A. D. Becke, *J. Chem. Phys.* **2014**, *140*, 18A301.
[77] J. P. Perdew, K. Schmidt, *AIP Conf. Proc.* **2001**, *577*, 1–20.
[78] W. Hujo, S. Grimme, *J. Chem. Theory Comput.* **2011**, *7*, 3866–3871.
[79] S. Rösel, H. Quanz, C. Logemann, J. Becker, E. Mossou, L. Cañadillas-Delgado, E. Caldeweyher, S. Grimme, P. R. Schreiner, *J. Am. Chem. Soc.* **2017**, *139*, 7428–7431.
[80] J. Liang, X. Feng, D. Hait, M. Head-Gordon, *J. Chem. Theory Comput.* **2022**, *18*, 3460.
[81] J. P. Perdew, A. Zunger, *Phys. Rev. B* **1981**, *23*, 5048–5079.
[82] P. Mori-Sánchez, A. J. Cohen, W. Yang, *J. Chem. Phys.* **2006**, *125*, 201102.
[83] J. L. Bao, L. Gagliardi, D. G. Truhlar, *J. Phys. Chem. Lett.* **2018**, *9*, 2353–2358.
[84] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
[85] C. Adamo, V. Barone, *J. Chem. Phys.* **1999**, *110*, 6158–6170.
[86] T. M. Henderson, B. G. Janesko, G. E. Scuseria, *J. Phys. Chem. A* **2008**, *112*, 12530–12542.
[87] É. Brémond, Á. J. Pérez-Jiménez, J. C. Sancho-García, C. Adamo, *J. Chem. Phys.* **2019**, *150*, 201102.
[88] P. Mori-Sánchez, A. J. Cohen, W. Yang, *Phys. Rev. Lett.* **2008**, *100*, 146401.
[89] A. J. Cohen, P. Mori-Sánchez, W. Yang, *Science* **2008**, *321*, 792–794.
[90] S. Song, S. Vuckovic, E. Sim, K. Burke, *J. Chem. Theory Comput.* **2022**, *18*, 817–827.
[91] S. Grimme, *J. Chem. Phys.* **2006**, *124*, 034108.
[92] L. Goerigk, S. Grimme, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 576–600.
[93] J. C. Sancho-García, C. Adamo, *Phys. Chem. Chem. Phys.* **2013**, *15*, 14581–14594.
[94] A. Görling, M. Levy, *Phys. Rev. A* **1994**, *50*, 196–204.
[95] C. Møller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 618–622.
[96] O. Vahtras, J. Almlöf, M. W. Feyereisen, *Chem. Phys. Lett.* **1993**, *213*, 514–518.
[97] F. Neese, F. Wennmohs, A. Hansen, U. Becker, *Chem. Phys.* **2009**, *356*, 98–109.
[98] "Accelerating Hartree–Fock exchange calculation using the TURBOMOLE program system: different techniques for

different purposes": A. Hellweg, U. Huniar, **2016**, https://doi.org/10.48550/ARXIV.1610.07779.
[99] Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **2008**, *120*, 215–241.
[100] Y. Zhao, D. G. Truhlar, *J. Chem. Phys.* **2006**, *125*, 194101.
[101] R. Peverati, D. G. Truhlar, *J. Phys. Chem. Lett.* **2011**, *2*, 2810–2817.
[102] R. Peverati, D. G. Truhlar, *Phys. Chem. Chem. Phys.* **2012**, *14*, 13171–13174.
[103] H. S. Yu, X. He, S. L. Li, D. G. Truhlar, *Chem. Sci.* **2016**, *7*, 5032–5051.
[104] P. Verma, D. G. Truhlar, *Trends Chem.* **2020**, *2*, 302–318.
[105] D. H. Ess, S. Liu, F. De Proft, *J. Phys. Chem. A* **2010**, *114*, 12952–12957.
[106] N. Mardirossian, M. Head-Gordon, *J. Chem. Theory Comput.* **2013**, *9*, 4453–4461.
[107] L. Goerigk, *J. Phys. Chem. Lett.* **2015**, *6*, 3891–3896.
[108] S. Dohm, A. Hansen, M. Steinmetz, S. Grimme, M. P. Checinski, *J. Chem. Theory Comput.* **2018**, *14*, 2596–2608.
[109] L. Goerigk, H. Kruse, S. Grimme, *ChemPhysChem* **2011**, *12*, 3421–3433.
[110] F. Jensen, *Introduction to Computational Chemistry, Vol. 2*, Wiley, Hoboken, **2007**.
[111] R. M. Balabin, *J. Chem. Phys.* **2008**, *129*, 164101.
[112] F. B. van Duijneveldt, J. G. D. van de Rijdt, J. H. van Lenthe, *Chem. Rev.* **1994**, *94*, 1873–1885.
[113] V. K. Prasad, A. Otero-de-la Roza, G. A. DiLabio, *J. Chem. Theory Comput.* **2022**, *18*, 2208–2232.
[114] J. C. Faver, Z. Zheng, K. M. Merz, *Phys. Chem. Chem. Phys.* **2012**, *14*, 7795–7799.
[115] F. Jensen, *J. Chem. Theory Comput.* **2010**, *6*, 100–106.
[116] R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.* **1971**, *54*, 720–723.
[117] T. H. Dunning, *J. Chem. Phys.* **1989**, *90*, 1007–1023.
[118] F. Jensen, *J. Chem. Phys.* **2001**, *115*, 9113–9125.
[119] F. Jensen, *J. Chem. Phys.* **2002**, *116*, 7372–7379.
[120] R. Ahlrichs, F. Weigend, M. Häser, H. Patzelt, *Chem. Phys. Lett.* **1998**, *294*, 143–152.
[121] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
[122] F. Weigend, F. Furche, R. Ahlrichs, *J. Chem. Phys.* **2003**, *119*, 12753–12762.
[123] D. Rappoport, F. Furche, *J. Chem. Phys.* **2010**, *133*, 134105.
[124] J. Zheng, X. Xu, D. G. Truhlar, *Theor. Chem. Acc.* **2011**, *128*, 295–305.
[125] D. Andrae, U. Häußermann, M. Dolg, H. Stoll, H. Preuß, *Theor. Chim. Acta* **1990**, *77*, 123–141.
[126] W. Kutzelnigg, W. Liu, *J. Chem. Phys.* **2005**, *123*, 241102.
[127] T. Saue, *ChemPhysChem* **2011**, *12*, 3077–3094.
[128] E. van Lenthe, E. J. Baerends, J. G. Snijders, *J. Chem. Phys.* **1993**, *99*, 4597–4610.
[129] M. Reiher, *Theor. Chem. Acc.* **2006**, *116*, 241–252.
[130] P. Pollak, F. Weigend, *J. Chem. Theory Comput.* **2017**, *13*, 3696–3705.
[131] J.-M. Mewes, A. Hansen, S. Grimme, *Angew. Chem. Int. Ed.* **2021**, *60*, 13144–13149; *Angew. Chem.* **2021**, *133*, 13252–13257.
[132] U. F. Ndongmouo, M. S. Lee, R. Rousseau, F. Baletto, S. Scandolo, *J. Phys. Chem. A* **2007**, *111*, 12810–12815.
[133] G. Mercurio, R. J. Maurer, W. Liu, S. Hagen, F. Leyssner, P. Tegeder, J. Meyer, A. Tkatchenko, S. Soubatch, K. Reuter, F. S. Tautz, *Phys. Rev. B* **2013**, *88*, 035421.
[134] J. N. Harvey, F. Himo, F. Maseras, L. Perrin, *ACS Catal.* **2019**, *9*, 6803–6813.
[135] J. A. Pople, H. B. Schlegel, R. Krishnan, D. J. Defrees, J. S. Binkley, M. J. Frisch, R. A. Whiteside, R. F. Hout, W. J. Hehre, *Int. J. Quantum Chem.* **1981**, *20*, 269–278.
[136] R. E. Stratmann, J. C. Burant, G. E. Scuseria, M. J. Frisch, *J. Chem. Phys.* **1997**, *106*, 10175–10183.
[137] Y. P. Li, J. Gomes, S. M. Sharada, A. T. Bell, M. Head-Gordon, *J. Phys. Chem. C* **2015**, *119*, 1840–1850.
[138] W. Koch, M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, Wiley-VCH, Weinheim, **2001**.
[139] A. Szabo, N. S. Ostlund, *Modern Quantum Chemistry*, Dover Publications, Mineola, **1996**.
[140] T. Helgaker, P. Jørgensen, J. Olsen, *Molecular Electronic-Structure Theory*, Wiley, New York, **2000**.
[141] S. Grimme, A. Hansen, J. G. Brandenburg, C. Bannwarth, *Chem. Rev.* **2016**, *116*, 5105–5154.
[142] S. Lehtola, C. Steigemann, M. J. Oliveira, M. A. Marques, *SoftwareX* **2018**, *7*, 1–5.
[143] U. Ekström, L. Visscher, R. Bast, A. J. Thorvaldsen, K. Ruud, *J. Chem. Theory Comput.* **2010**, *6*, 1971–1980.
[144] J. Hutter, M. Iannuzzi, F. Schiffmann, J. VandeVondele, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 15–25.
[145] K. Aidas, et al., *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 269–284.
[146] F. Aquilante, et al., *J. Comput. Chem.* **2016**, *37*, 506–541.
[147] Q. Sun, et al., *J. Chem. Phys.* **2020**, *153*, 024109.
[148] DIRAC, a relativistic ab initio electronic structure program, Release DIRAC22 (2022), written by H. J. A. Jensen, R. Bast, A. S. P. Gomes, T. Saue and L. Visscher, with contributions from I. A. Aucar, V. Bakken, C. Chibueze, J. Creutzberg, K. G. Dyall, S. Dubillard, U. Ekström, E. Eliav, T. Enevoldsen, E. Faßhauer, T. Fleig, O. Fossgaard, L. Halbert, E. D. Hedegård, T. Helgaker, B. Helmich-Paris, J. Henriksson, M. van Horn, M. Iliaš, C. R. Jacob, S. Knecht, S. Komorovský, O. Kullie, J. K. Lærdahl, C. V. Larsen, Y. S. Lee, N. H. List, H. S. Nataraj, M. K. Nayak, P. Norman, G. Olejniczak, J. Olsen, J. M. H. Olsen, A. Papadopoulos, Y. C. Park, J. K. Pedersen, M. Pernpointner, J. V. Pototschnig, R. di Remigio, M. Repisky, K. Ruud, P. Sałek, B. Schimmelpfennig, B. Senjean, A. Shee, J. Sikkema, A. Sunaga, A. J. Thorvaldsen, J. Thyssen, J. van Stralen, M. L. Vidal, S. Villaume, O. Visser, T. Winther, S. Yamamoto and X. Yuan (available at https://doi.org/10.5281/zenodo.6010450, see also http://www.diracprogram.org).
[149] M. J. Frisch, et al. "Gaussian 16 Revision C.01", 2016 Gaussian Inc. Wallingford CT.
[150] S. Lehtola, A. J. Karttunen, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2022**, *12*, e1610.
[151] F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka, F. Weigend, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 91–100.
[152] F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 73–78.
[153] F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1327.
[154] F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2022**, *12*, e1606.
[155] E. Epifanovsky, et al., *J. Chem. Phys.* **2021**, *155*, 084801.
[156] D. G. Smith, et al., *J. Chem. Phys.* **2020**, *152*, 184108.
[157] H. J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 242–253.
[158] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. van Gisbergen, J. G. Snijders, T. Ziegler, *J. Comput. Chem.* **2001**, *22*, 931–967.
[159] *Semiempirical Extended Tight-Binding Program Package xtb*, Version 6.4.1., **2021**, https://github.com/grimme-lab/xtb.
[160] B. Hourahine, et al., *J. Chem. Phys.* **2020**, *152*, 124101.
[161] MOPAC2016, J. J. P. Stewart, Stewart Computational Chemistry, Colorado Springs, CO, USA https://OpenMOPAC.net (**2016**).

[162] S. Spicher, S. Grimme, *J. Chem. Theory Comput.* **2021**, *17*, 1701–1714.
[163] User Guide to Semiempirical Tight Binding, https://xtb-docs.readthedocs.io/en/latest/contents.html, access date: April 5, 2022.
[164] M. Bursch, H. Neugebauer, S. Ehlert, S. Grimme, *J. Chem. Phys.* **2022**, *156*, 134105.
[165] S. Vuckovic, K. Burke, *J. Phys. Chem. Lett.* **2020**, *11*, 9957–9964.
[166] R. S. Grev, H. F. Schaefer, *J. Chem. Phys.* **1989**, *91*, 7305–7306.
[167] D. Moran, A. C. Simmonett, F. E. Leach, W. D. Allen, P. V. Schleyer, H. F. Schaefer, *J. Am. Chem. Soc.* **2006**, *128*, 9342–9343.
[168] J. Witte, M. Goldey, J. B. Neaton, M. Head-Gordon, *J. Chem. Theory Comput.* **2015**, *11*, 1481–1492.
[169] M. Swart, J. G. Snijders, *Theor. Chem. Acc.* **2003**, *110*, 34–41.
[170] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, J. Sun, *J. Phys. Chem. Lett.* **2020**, *11*, 8208–8215.
[171] S. Ehlert, U. Huniar, J. Ning, J. W. Furness, J. Sun, A. D. Kaplan, J. P. Perdew, J. G. Brandenburg, *J. Chem. Phys.* **2021**, *154*, 061101.
[172] J. Tao, J. P. Perdew, V. N. Staroverov, G. E. Scuseria, *Phys. Rev. Lett.* **2003**, *91*, 146401.
[173] T. Risthaus, M. Steinmetz, S. Grimme, *J. Comput. Chem.* **2014**, *35*, 1509–1516.
[174] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
[175] E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth, S. Grimme, *J. Chem. Phys.* **2019**, *150*, 154122.
[176] O. A. Vydrov, T. Van Voorhis, *J. Chem. Phys.* **2010**, *133*, 244103.
[177] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.
[178] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.
[179] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.
[180] S. Grimme, *Chem. Eur. J.* **2012**, *18*, 9955–9964.
[181] J. P. Merrick, D. Moran, L. Radom, *J. Phys. Chem. A* **2007**, *111*, 11683–11700.
[182] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, J. Olsen, *Chem. Phys. Lett.* **1999**, *302*, 437–446.
[183] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, A. K. Wilson, *Chem. Phys. Lett.* **1998**, *286*, 243–252.
[184] N. Mardirossian, M. Head-Gordon, *J. Chem. Phys.* **2016**, *144*, 214110.
[185] Y. S. Lin, G. D. Li, S. P. Mao, J. D. Chai, *J. Chem. Theory Comput.* **2013**, *9*, 263–272.
[186] N. Mardirossian, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.
[187] J. D. Chai, M. Head-Gordon, *J. Chem. Phys.* **2009**, *131*, 174105.
[188] G. Santra, N. Sylvetsky, J. M. Martin, *J. Phys. Chem. A* **2019**, *123*, 5129.
[189] L. Goerigk, S. Grimme, *J. Chem. Theory Comput.* **2011**, *7*, 291–309.
[190] Y. Zhao, D. G. Truhlar, *J. Phys. Chem. A* **2005**, *109*, 5656–5667.
[191] A. D. Boese, J. M. Martin, *J. Chem. Phys.* **2004**, *121*, 3405–3416.
[192] Y. Zhao, D. G. Truhlar, *J. Phys. Chem. A* **2004**, *108*, 6908–6918.
[193] S. Grimme, P. R. Schreiner, *Angew. Chem. Int. Ed.* **2011**, *50*, 12639–12642; *Angew. Chem.* **2011**, *123*, 12849–12853.
[194] G. Henkelman, B. P. Uberuaga, H. Jónsson, *J. Chem. Phys.* **2000**, *113*, 9901–9904.
[195] L. D. Jacobson, A. D. Bochevarov, M. A. Watson, T. F. Hughes, D. Rinaldo, S. Ehrlich, T. B. Steinbrecher, S. Vaitheeswaran, D. M. Philipp, M. D. Halls, R. A. Friesner, *J. Chem. Theory Comput.* **2017**, *13*, 5780–5797.
[196] M. H. Rasmussen, J. H. Jensen, *PeerJ Phys. Chem.* **2020**, *2*, e15.
[197] P. M. Zimmerman, *J. Chem. Theory Comput.* **2013**, *9*, 3043–3050.
[198] P. M. Zimmerman, *J. Comput. Chem.* **2015**, *36*, 601–611.
[199] S. Dohm, M. Bursch, A. Hansen, S. Grimme, *J. Chem. Theory Comput.* **2020**, *16*, 2002–2012.
[200] B. M. Axilrod, E. Teller, *J. Chem. Phys.* **1943**, *11*, 299–300.
[201] Y. Muto, *Proc. Phys.-Math. Soc. Jpn.* **1943**, *17*, 629–631.
[202] R. Sedlak, T. Janowski, G. Pitopateanu, *Supramol. Chem.* **2008**, *20*, 473–477.
[203] R. Sure, S. Grimme, *J. Chem. Theory Comput.* **2015**, *11*, 3785–3801.
[204] J. J. P. Stewart, *J. Mol. Model.* **2007**, *13*, 1173.
[205] J. Řezáč, P. Hobza, *J. Chem. Theory Comput.* **2012**, *8*, 141–151.
[206] J. F. Dobson, *Int. J. Quantum Chem.* **2014**, *114*, 1157–1161.
[207] K. Nishiyama, N. Sakiyama, S. Seki, H. Horita, T. Otsubo, S. Misumi, *Bull. Chem. Soc. Jpn.* **1980**, *53*, 869–877.
[208] S. K. Pollack, B. C. Raine, W. J. Hehre, *J. Am. Chem. Soc.* **1981**, *103*, 6308–6313.
[209] C. F. Shieh, D. McNally, R. H. Boyd, *Tetrahedron* **1969**, *25*, 3653–3665.
[210] A. Karton, J. M. Martin, *J. Chem. Phys.* **2012**, *136*, 124114.
[211] I. V. Kolesnichenko, E. V. Anslyn, *Chem. Soc. Rev.* **2017**, *46*, 2385–2390.
[212] K. E. Riley, P. Hobza, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 3–17.
[213] P. I. Dron, S. Fourmentin, F. Cazier, D. Landy, G. Surpateanu, *Supramol. Chem.* **2008**, *20*, 473–477.
[214] P. Pracht, S. Grimme, *Chem. Sci.* **2021**, *12*, 6551–6568.
[215] S. Spicher, S. Grimme, *J. Phys. Chem. Lett.* **2020**, *11*, 6606–6611.
[216] Introduction to *CENSO—xtb* documentation, https://xtb-docs.readthedocs.io/en/latest/CENSO docs/censo usage.html#calcula free-energies-on-populated-dft-optimized-conformers, access date: April 5, 2022.
[217] Introduction to *CREST—xtb* documentation, https://xtb-docs.readthedocs.io/en/latest/crestxmpl.html#imtd-gc, access date: April 5, 2022.
[218] F. Bohle, J. Seibert, S. Grimme, *J. Org. Chem.* **2021**, *86*, 15522–15531.
[219] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.
[220] J. P. Perdew, *Phys. Rev. B* **1986**, *33*, 8822–8824.
[221] C. Poree, F. Schoenebeck, *Acc. Chem. Res.* **2017**, *50*, 605–608.
[222] H. B. Vibbert, H. Neugebauer, J. R. Norton, A. Hansen, M. Bursch, S. Grimme, *Can. J. Chem.* **2021**, *99*, 216–220.
[223] T. M. Maier, A. V. Arbuznikov, M. Kaupp, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2019**, *9*, e1378.
[224] H. Bahmann, M. Kaupp, *J. Chem. Theory Comput.* **2015**, *11*, 1540–1548.
[225] M. Haasler, T. M. Maier, R. Grotjahn, S. Gückel, A. V. Arbuznikov, M. Kaupp, *J. Chem. Theory Comput.* **2020**, *16*, 5645–5657.
[226] E. Sim, S. Song, K. Burke, *J. Phys. Chem. Lett.* **2018**, *9*, 6385–6392.
[227] M. C. Kim, E. Sim, K. Burke, *Phys. Rev. Lett.* **2013**, *111*, 073003.
[228] E. Sim, S. Song, S. Vuckovic, K. Burke, *J. Am. Chem. Soc.* **2022**, *144*, 6625–6639.
[229] J. Kirkpatrick, et al., *Science* **2021**, *374*, 1385–1389.
[230] B. Kalita, L. Li, R. J. McCarty, K. Burke, *Acc. Chem. Res.* **2021**, *54*, 818–826.

[231] R. Nagai, R. Akashi, O. Sugino, *npj Comput. Mater.* **2020**, *6*, 43.
[232] G. R. Schleder, A. C. Padilha, C. M. Acosta, M. Costa, A. Fazzio, *JPhys. Mater.* **2019**, *2*, 032001.
[233] C. E. Belle, V. Aksakalli, S. P. Russo, *J. Cheminf.* **2021**, *13*, 42.
[234] L. C. Blum, J. L. Reymond, *J. Am. Chem. Soc.* **2009**, *131*, 8732–8733.

Manuscript received: April 19, 2022
Version of record online: September 14, 2022