A second-generation range-separated hybrid van der Waals density functional

Vivekanand Shukla, Yang Jiao, Jung-Hoon Lee, and Per Hyldgaard

1Department of Microtechnology and Nanoscience – MC2, Chalmers University of Technology, SE-41296 Gothenburg, Sweden
2Computational Science Research Center, Korea Institute of Science and Technology, Seoul 02792, Republic of Korea
3Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720-7300, USA
4Department of Physics, University of California, Berkeley, California 94720-7300, USA

(Dated: March 15, 2022)

We define, test, and illustrate use of a range-separated (screened) hybrid, termed vdW-DF2-ahbr and abbreviated AHBR, that uses the second-generation nonlocal-correlation formulation [PRB 82, 081101(R) (2010)] within the van der Waals density functional (vdW-DF) method [ROPP 78, 066501 (2015)]. It supplements the vdW-DF-ahcx [JPCM 34, 025902 (2022)] (abbreviated AHCX) which relies on the first general-geometry formulation of the nonlocal correlation energy [PRL 92, 246401 (2004)]. We build the AHBR off a new analytical-hole analysis of the exchange in vdW-DF2-b86r [PRB 89, 121103 (2014)]. Like AHCX, the AHBR uses an exchange enhancement with a large-gradient behavior that prevents spurious exchange binding and with a small-gradient form that is set from many-body perturbation analysis. Unlike AHCX, however, the AHBR relies on a slightly different interpretation of this input, moving the resulting exchange description closer to PBEsol exchange [PRL 100, 136406 (2008)]. We validate that the range-separated hybrid (RSH) vdW-DF AHBR retains and exceeds the strong AHCX performance both on bulk and in a complete planewave benchmarking across the full GMTKN55 suite on broad molecular properties [PCCP 19, 32184 (2017)]. Interestingly, the AHBR accuracy gains are most pronounced for the class of barrier-heights benchmarks, suggesting that the AHBR is even more robust for transition-state and binding-energies than we show that AHBR correctly predicts DNA base-pair stacking energies, the binding-site preference for CO/Pt(111), and CO₂ adsorption in Mg-MOF74.

I. INTRODUCTION

Accurate density functional theory (DFT) descriptions of molecular properties are important in chemistry and biochemistry. We seek a good exchange-correlation (XC) energy functional that allows computationally efficient calculations and that brings us close to the (1 kcal/mol) chemical-accuracy goal on the energies that define key chemistry processes. This accuracy is needed for understanding and optimizing reactions, charge-transfer processes, and barriers. An even better accuracy is needed for understanding molecular-energy storage, CO₂ capture, batteries, and biochemistry in general, for example, Refs. 24–27. These are cases where we must understand the role of the weak inter- as well as intramolecular noncovalent (NOC) interactions. Furthermore, we want that XC functional to also have a general-purpose character, for example, that we can use this very same functional both for molecules and for understanding surface and interface properties. Having such genuine XC-functional transferability means that we can predict and model chemistry under both homogeneous and heterogeneous conditions with one and the same DFT approximation.

Nonlocal-correlation functionals of the van der Waals (vdW) density functional (vdW-DF) method fulfill a necessary condition for serving as a general-purpose DFT. They incorporate the ubiquitous van der Waals (vdW) interactions that dominate intermolecular interactions. The vdW-DF method is a systematic way to design a truly nonlocal-correlation functional from the insight that we have already gained in making semilocal functionals within the generalized gradient approximation (GGA). The two first method releases, termed vdW-DF1 (or vdW-DF) and vdW-DF2, have the same overall structure but differ in how a plasmon-response model sets the details of the nonlocal-correlation energy and in choice of the exchange component. The exchange component is in both releases set to ideally eliminate (without overcompensating) spurious weak-system binding by exchange contributions. In the former (latter) vdW-DF method release, the gradient-corrected exchange form was not picked as the PBE GGA-exchange but rather as the more-repulsive revPBE variant (as a refitted PW86 form), following analysis of binding of noble-gas and molecular systems, Refs. 30 and 55.

Broader method impact and benefits as general-purpose tools came upon further inspection of the role of gradient-corrected exchange in the overall design for a balanced exchange-correlation (XC) energy functional design. This approach produced both variants that reflect a practice-driven analysis as well as a focus on seeking nonempirical designs with an XC balance that respects constraints set by many-body perturbation theory (MBPT). The nonempirical design focus led to the definition of a class of consistent vdW-DFs where Lindhard screening is employed to balance the XC terms and ensure compliance with current-conservation criteria both on the
FIG. 1. Performance comparison of vdW-DFs as averaged over the 6 groups of benchmarks in the GMTKN55 suite [5] (left panel) and as tracked for transition-state problems (right panel): The barrier height benchmarks for small and cyclic molecules (BH76 & BHPERI), in diverse, inversion and rotation processes (BHDIV10, INV24, & BHROT27), as well as in proton exchange and transfer problems (PX13 & WCPT18). We report mean-absolute deviations (MADs,) in kcal/mol, compared with reference energies, at reference geometries.

Regular functionals, exemplified by the three overall best molecule performers, namely vdW-DF-cx (abbreviated CX), vdW-DF-obk8 (abbreviated OBK8) and vdW-DF2-b86r (or rev-vdW-DF2, abbreviated B86R) are challenged by this type of problems, where the interaction contributions are typically diffuse, i.e., distributed in space. Dispersion-corrected HSE [5,57–59] and the RSH vdW-DF-ahcx [42] (abbreviated AHCX) improve the accuracy, but the simple unscreened hybrid vdW-DF2-br0 (abbreviated DF2-BR0 and defined as a vdW-DF+0 form based on B86R) holds the performance edge in general as well as for transition-state problems. Accordingly, we here launch an associated RSH vdW-DF, termed vdW-DF2-ahbr (abbreviated AHBR), that leverages screening and that therefore (like AHCX) permits broad application to adsorption studies [42].

Table S I of the supplementary information (SI) materials provides an overview of the abbreviations that we presently use in discussing the many regular- and hybrid-vdW-DF releases and variants.

We have very recently suggested a focus on systematically exploring use of a tool chain of related vdW-DF versions [27,42,43]. This tool chain comprises the consistent-exchange regular vdW-DF-cx version [35] (abbreviated CX), the unscreened hybrid vdW-DF-cx0p [40] (here abbreviated CX0P), and the RSH vdW-DF-ahcx [42] (abbreviated AHCX). All members use the original general-geometry formulation of the nonlocal-correlation energy [31,33]. The CX0P belongs in the generic vdW-DF+0 class [39] defined by an analogy to PBE0 [66,67]; It is used at a 0.20 fraction of Fock exchange following a coupling-constant analysis of the CX description for a set of atomization energies [40]. The AHCX is a range-separated and thus screened CX-based hybrid defined in analogy with the HSE [57,58] but it relies on an analytical-hole (AH) analysis [69] of the nature of exchange in the CX design [31,42].

Together, CX/CX0P/AHCX form a tool chain that aims to maximize the leverage of a Lindhard screening logic [31,35,40] for securing robustness (that is, minimal presence of performance outliers). Having the small-gradient exchange form set from the early Langreth-Vosko (LV) MBPT analysis of gradient-corrections to exchange [33,48,49], this CX-based tool chain is set up to likely succeed at bulk and surfaces [43]. Good CX performance for molecular properties is not a priori given, although a welcome bonus [27,42,43]. The RSH AHCX is the tool-chain cornerstone for it allows a mapping across both molecules and generic substrates (including metals), i.e., for the important heterogeneous-materials problems [42,70–75]. The overall idea of using this CX-based tool chain is that we can map the interplay between truly nonlocal exchange and truly nonlocal correlation and, hopefully, thus understand how to balance them also when used in combination [27,42,43].

The present paper contains several mutually supporting analysis, design, and computational results that lead us to the formulation of a second-generation RSH vdW-DF (termed vdW-DF2-ahbr and abbreviated AHBR), as well as an associated XC tool chain. Specifically, we now start with the vdW-DF2 nonlocal-correlation energy formulation [9]. We note that, with the AHCX launch [42] we also provided an analytical-hole (AH) analysis [69]...
The exchange choice for vdW-DF2, implicitly permitting the formulation of a corresponding RSH vdW-DF2-ah, abbreviated DF2-AH. The DF2-AH hybrid is not, however, a robust design (neither for bulk nor molecules) – just like we confirm that the performance of vdW-DF2 itself varies dramatically among different types of material challenges. Worse, the design of rPW86 and thus vdW-DF2-ah exchange does not reflect the diagrammatic MBPT input, summarized in Refs. [44, 48, and 49] on gradient-corrected exchange (except for the smallest of density gradients [48]). That MBPT input is instead prioritized in the PBEsol [53] and LA [55] exchange, and hence in CX, where exchange is set from an LV-rPW86 interpolation [55]. The vdW-DF2-ah is not easy to use as a tool-chain complement in our XC development focus, for it is not sufficiently related to AHCX that we can draw trusted conclusions in a performance contrast.

Instead, the new vdW-DF2-based tool chain comprises vdW-DF2-b86r [55] (or rev-vdW-DF2, here abbreviated B86R), the associated simple vdW-DF-br0 hybrid (abbreviated DF2-BR0), being of the vdW-DF+0 class [53], as well as the AHBR. The AHBR works as the desired AHCX supplement because it relies on well-understood reinterpretations of diagrammatic-MBPT input [44, 48, and 49] on gradient-corrected exchange (except for the smallest of density gradients [48]). That MBPT input is instead prioritized in the PBEsol [53] and LA [55] exchange, and hence in CX, where exchange is set from an LV-rPW86 interpolation [55]. The new tool chain uses different input to set the assumed plasmon dispersion, Ref. 9 versus Refs. 31 and 33. Both RSH vdW-DFs, AHCX and AHBR, can be used when screening is essential, for example, for descriptions of adsorption at metal surfaces. The new B86R-based tool chain can assist us in systematic vdW-DF tool-chain explorations [27]. Contrasting AHCX and AHBR may, for example, help to illuminate the nature of binding in systems that vdW-DF usage identifies as outstanding material challenges. Such discussions, for example, in Refs. 19, 75, and 76 are essential to help drive the vdW-DF development program forward.

Our specific contributions in this paper can be summarized as follows. First, we observe that B86R (like CX) reflects the MBPT-diagram input on gradient-corrected exchange but (unlike CX) interprets it as it was used in PBEsol. This fact suggests a potential advantage that is different from but supplements the Lindhard advantage of the CX-based tool chain [27]. Second, we extend our AH analysis [49] to cover B86R and show that it is moved closer to PBEsol exchange, while still retaining an ability to counteract spurious exchange-only binding [53, 55] in molecular cases. Third, we complete a robust procedure for plane-wave assessments across the full-GMTKN55 benchmark suite [53], documenting that DF2-BR0 provides the best performance on molecular properties: This is true among all of the vdW-DFs, including the AHCX. Fourth, since the DF2-BR0 performance edge exists evenly across all of the GMTKN55 benchmark groups, we proceed to also define the AHBR, using our here-provided AH characterization for B86R. Fifth, we use our GMTKN55 characterization to suggest that the AHBR be used with a default 0.25 fraction of Fock exchange mixing (and the standard HSE/AHCX choice for the range-separation screening parameter [42]).

In turn, we validate that this generic AHBR retains a strong performance also for bulk structures. Finally, we illustrate the usefulness of the new B86R-based tool chains for DNA assembly and molecular adsorption problems, finding good agreement on quantum-chemistry reference calculations, the correct site preference for the CO/Pt(111) problem, and a strong performance for characterization of CO2 uptake in a MOF [18].

Concerning the molecular benchmarking, we note that our use of the plane-wave-code QuantumEspresso (QE) suite [77] gives us the prerequisites for delivering a high-quality (in principle, complete-basis-set) assessment. We can secure a robust characterization as long as we also include all relevant electrons in our pseudopotentials (PPs) and compensate or control spurious electrostatic and vdW-type intercell interactions in our periodic-cell calculations. There are challenges with plane-wave benchmarking related to self-interaction errors (SIEs) in negatively charged atoms and radicals [15, 17]. In appendix A, however, we document that use of a dielectric-environment extension [81] permits us to circumvent these challenges and reliably complete the functional assessment even for the G21EA and WATER27 benchmark sets [82] of the GMTKN55 suite.

Figure 1 summarizes the main results for the full GMTKN55 plane-wave benchmarking (left panel), while also illustrating that even the best vdW-DF performers are generally challenged by the class of transition-state

![Figure 2](image-url)
Quantitative performance comparisons between the AHBR and all of the vdW-DFs (that are implemented in the present version of QChem) as well as for dispersion-corrected, revPBE, and HSE are reported in the SI material. Here, Table S II (Tables S III and S IV) reports comparisons on the group-average WTMA9 values (other averaged-MAD measures for comparison) while Tables S V through S X report MAD characterizations for the individual GMTKN55 benchmarks, as sorted into the six GMTKN55 groups. A more traditional performance overview (reflected group performance) for the vdW-DF hybrids is presented in the abstract figure [See Fig. 13]. Here, we restrict the initial discussion to the three best overall-performing regular vdW-DFs – CX, vdW-DF-okb, and B86R – and compare their performance to those of AHCX, DF2-BR0, AHBR and dispersion-corrected HSE. Figure I highlights two important take-away messages of this paper, namely that: 1) use of range-separated and simple hybrid vdW-DFs provides substantial improvements on the accuracy for computing molecular properties, and 2) use of the B86R hybrids, i.e., the DF2-BR0 and the here-defined AHBR, provides an evenly robust performance. The AHBR avoids weakness for the important challenge defined by barrier, transition-state, and some isomerization problems. In this hybrid status assessment, Fig. 1 illustrates the impact of switching between the two (standard, 0.20 and 0.25) choices for the extent of Fock-exchange mixing, as marked by subscripts (for example, AHBR0.20).

The rest of the paper is organized as follows. The theory section II presents an overview of the vdW-DF method, analysis of the B86R exchange hole, and the formulation of the new AHBR. Section III, supported by an appendix, presents computational details, including the electrostatic-environment approach used to complete planewave benchmarking across the full GMTKN55 suite. Section IV contains results and discussion, including demonstrators for DNA, for CO/PT(111) adsorption, and for a CO2-MOF adsorption problem. Finally, Section V contains a summary and outlook.

II. THEORY

Central in MBPT and in the electron-gas foundation of DF is the screened density response \( \delta n(\omega) \) to some external-potential change \( \delta \hat{\Phi}^{\text{ext}} \), oscillating at frequency \( \omega \). In MBPT we can, at least in principle, compute the nonlocal response function \( \chi(\mathbf{r}, \mathbf{r}'; \omega) = \delta n(\mathbf{r})/\delta \hat{\Phi}^{\text{ext}}(\mathbf{r}') \), often as a function of a complex frequency \( \omega = i\epsilon \) and at a range of an assumed reduced strength \( 0 < \lambda < 1 \) of the electron-electron interaction \( V_{\lambda} \). Assuming such insight, the adiabatic connection formula (ACF) permits an exact determination

\[
E_{\text{xc}} = -\int_0^1 d\lambda \int_0^\infty \frac{du}{2\pi} \text{Tr}\{\chi(\lambda) V\} - E_{\text{self}},
\]

of the XC energy functional \( E_{\text{xc}} \). Here \( V = |\mathbf{r} - \mathbf{r}'|^{-1} \) denotes the matrix element of the electron-electron interaction \( V \). \( u \) denotes an imaginary frequency argument in the response description, while the last term is the electron self energy \( E_{\text{self}} = \text{Tr}(\hat{V} V) / 2 \). The expressions for \( E_{\text{xc}} \) and \( E_{\text{self}} \) involve Coulomb-weighted traces, that is, integrations in spatial coordinates of \( |\mathbf{r} - \mathbf{r}'|^{-1} \) times \( \chi(\mathbf{r}', \mathbf{r}; \omega) \) and times the electron density \( n(\mathbf{r}) \). Also we have (at every coupling-constant value \( \lambda \)) added an auxiliary potential that keeps the electron density \( n(\mathbf{r}) \) unchanged across the implied adiabatic turn on of the electron-electron interaction \( V_{\lambda} = \lambda V \). The actual XC potential used in the Kohn-Sham (KS) scheme for efficient DFT calculations is simply the \( \lambda \to 0 \) limit of this auxiliary potential. It is given by a functional derivative of the XC energy, as discussed many places elsewhere.

In MBPT, we compute the response functions \( \chi(\omega) \) as a ground-state expectation value of correlations between density fluctuations. As such, \( \chi(\omega) \) is directly reflecting the Lindhard-type screening that exists in the electron gas at assumed coupling constant \( \lambda \). The screening is given by the dielectric function \( \kappa(\omega) = (1 + \lambda V \chi)^{-1} \).

For practical DFT, we seek XC energy-functional approximations and these can be obtained by a formal analysis of the response function \( \chi(i\epsilon) \). In the electron-gas tradition, and underpins both the constraint-based GGA designs and the vdW-DF method, we define an XC hole \( n_{\text{xc}}(\mathbf{r}; \mathbf{r}') \) that contains the \( \chi(i\epsilon) \) information while permitting an ACF.
FIG. 3. Radial variation in the scaled exchange hole \( J(s, y = kF|\mathbf{r} - \mathbf{r}'|) = n_e(\mathbf{r}, \mathbf{r}')/n(r) \) for PBE, PBEsol, and B86R, all as described in an analytical-hole (AH) model parameterization. The shapes of these holes define the exchange components of PBE, PBEsol, and B86R, respectively. Using the latter exchange-hole model, we define the here-released RSH AHBR, following the design steps that we recently documented in crafting AHCX.

**A. The vdW-DF framework**

To discuss the XC energy in the vdW-DF method it is convenient to rely on a formal recast,

\[
E_{xc} = \frac{1}{2} \int_r \int_{r'} \frac{n(r)n_{xc}(r, r')}{|r - r'|} \ . \tag{2}
\]

It has the appearance of electrostatics, but the XC hole reflects the impact of zero-point vibrations, i.e., plasmon excitations in the electron distributions. The exchange-hole component \( n_{xc}(r, r') \) of this total XC hole describes the impact of Pauli exclusion.

All of these inputs are used in the vdW-DF method to provide an effective (MBPT-guided) approximation to the \( \lambda \)-averaged response description, Eq. (3), as discussed in Refs. 31, 33, 34, 37, 38, and 43.

In the vdW-DF method, we furthermore use a so-called internal functional \( E_{xc}^{\text{in}} \) to first set a lowest-order approximation for screening defined by a dielectric function \( \epsilon(\omega) \).

\[
E_{xc}^{\text{in}} = \int_0^\infty \frac{du}{2\pi} \text{Tr} \{ \ln(\kappa_{\text{ACF}}(iu)) \} - E_{\text{self}} , \tag{4}
\]

and we use \( S_{xc}(\omega) = \ln(\epsilon(\omega)) \) as a (starting) approximation for the plasmon propagator \( S_{xc}(\omega) \), as discussed elsewhere. By construction, the \( S_{xc}(\omega) \) is also an approximation for an effective electron-gas susceptibility \( \alpha(\omega) = (\epsilon(\omega) - 1)/4\pi \).

For \( S_{xc}(\omega) \), the vdW-DF method relies on an explicit two-pole approximation reflecting plausible assumptions for plasmon dispersion. The practical use of Eq. (4) is to set the details of this dispersion so that the \( S_{xc}(\omega) \) starting point reflects the density-density response (and hence screening) that is already captured in the internal functional \( E_{xc}^{\text{in}} \).

Finally, in the vdW-DF, we explicitly enforce a longitudinal projection of the response in the dielectrics approximation function

\[
\kappa_{\text{ACF}}(\omega) = -\nabla \cdot \epsilon(\omega) \nabla V/4\pi . \tag{5}
\]

We note that having this projection inside the ACF recast, Eq. (4), produces terms that capture the vdW...
attraction as described in the presence of electron-gas screening. Moreover, the use of \( \epsilon(\omega) = \exp(S_{xc}(\omega)) \) implies a cumulant-/cluster-expansion logic in the response description. This allows the vdw-DF method to also pick up higher-order susceptibility and screening effects, i.e., to balance the vdw attraction by other nonlocal-correlation effects.

**B. General-geometry vdw-DF releases**

For a computationally efficient evaluation, the standard general-geometry formulation expands the ACF recast, Eq. (3), to second order in the (nonlocal) plasmon propagator \( S_{xc}(\omega) = \ln(\epsilon(\omega)) \). Formally, the expansion is written

\[
E_{xc}^{DFs} = E_{xc}^{in} + E_{xc}^{nl}
\]

where again \( E_{xc}^{in} \) determines the \( S_{xc}(iu) \) details through Eq. (1).

In all present general-geometry vdw-DFs, the internal \( E_{xc}^{in} \) functional is chosen semi-local (GGA-like), comprising LDA and a simple choice of physics-motivated gradient-corrected exchange. This is done to avoid double-counting of nonlocal correlations. The gradient-exchange choices used in \( E_{xc}^{in} \) are defined by the MBPT input as interpreted by \( \delta E_{xc}^{nl} \) in the first general-geometry release, vdw-DF1, and by the Schwinger exact-exchange scaling analysis in the second, vdw-DF2.

**C. Choice of vdw-DF exchange**

The top panel of Fig. 2 compares the so-called exchange-enhancement factor \( F_x \) that defines the nature of the gradient-corrected exchange in all GGA descriptions and in present vdw-DFs. It is plotted as a function of the scaled density gradient \( s = \nabla n / (2nk_F) \), where \( k_F = (3\pi^2n)^{1/3} \). The exchange energy in any semilocal approximation (that the present vdw-DFs also use) must take the form

\[
E_{xc}^{GA} = \int n(r)c_{x}^{LDA}(n(r))F_x(s(r)) dr.
\]

Here \( c_{x}^{LDA} = -3k_F/4\pi \) denotes the local-density approximation (LDA) exchange result, i.e., the exchange energy-per-particle value that characterizes a homogeneous system (at density \( n = \langle n(r) \rangle \)). The variations in gradient corrected exchange is thus set alone by the enhancement form of factor \( F_x(s) \), which for the internal expansion \( E_{xc}^{in} \) is set simply as a quadratic expansion, \( F_x^{in}(s) = 1 + \mu s^2 \), Ref. [9] and [33].

The ACF foundation, Eqs. (3) through (5), does motivate the use of Eq. (6) to pick the exchange. Such a pick finalizes the design so that the ACF and Lindhard screening logic balance the exchange and the correlation terms and ensure current conservation. This is, we should ideally use the internal-exchange formulation, and hence the \( S_{xc}(iu) \) form, to also define the actual exchange through an exchange-enhancement factor \( F_x(s) \approx 1 + \mu s^2 \) (as in the LV exchange design). Having such a quadratic form would certainly eliminate the risk of any spurious exchange-only binding but the repulsion by this gradient-corrected exchange is excessive, impossible to use for molecular systems.

Instead, the actual exchange design, set by \( F_x(s) \) in \( E_{xc}^{nl} \), must at larger \( s \) values be guided by other considerations. Practical and popular exchange descriptions arise when we impose an exchange-hole-conservation criterion, for example, as implemented in the (revised) PW86, PBE, and PBEsol while also preventing the exchange-hole depth from dramatically exceeding the local-electron density. For the vdw-DFs we should also, ideally, craft an asymptotic \( F_x(s) \) behavior that produces an adequate but not excessive repulsion by gradient-corrected exchange for weakly interacting molecules.

Aiming to satisfy primarily the latter condition, we set the exchange in the first (untruncated but layered-geometry) vdw-DF-lg and in vdw-DF to revPBE (since it has the most rapidly rising \( F_x(s) \) form among the PBE variants). Meanwhile, the analysis in Ref. [55] shows that the asymptotic form a GGA-type exchange component for a vdw-DF should ideally have an \( F_x(s) \approx s^{2/7} \) asymptotic behavior (to mimic a trusted Hartree-Fock type repulsion in weak molecular-dimer binding). This behavior is found in PW86 and in Becke86. The related rPW86 was crafted (with this asymptote) for the vdw-DF2 release, and the asymptotic \( F_x(s) \sim s^{2/7} \) form was also kept in CX.

In summary and accordingly, it is not possible to fully align \( F_x \) and \( E_{xc}^{nl} \) and thus directly enforce compliance with the formal expansion Eq. (6). Instead the general-geometry vdw-DFs are defined with a looser connection between the XC terms,

\[
E_{xc}^{DFs} = E_{xc}^{0} + E_{xc}^{nl}.
\]

Here \( E_{xc}^{0} = E_{xc}^{in} + \delta E_{xc}^{0} \) is set as LDA plus some reasonable gradient-corrected exchange, in which we permit a cross-over term \( \delta E_{xc}^{0} \) (containing nothing but gradient-corrected exchange). Above we gave the logic for the exchange choice (and hence XC balance) in the first and second general-geometry versions. Variants, for example, the OBK82 the vdw-DF-C09 (abbreviated C09,) the vdw-DF-optB86 (abbreviated OB86,) and the B86R exploit flexibility of the \( \delta E_{xc}^{0} \) term.

On the one hand, the introduction of such variants has enhanced the range of applications that can easily be addressed with the vdw-DF method (compared with reach of vdw-DF1 and vdw-DF2). Some of these variants compensate for possible \( E_{xc}^{nl} \) limitations by fitting the choice of \( \delta E_{xc}^{0} \) to a target or expected representative application. The variant flexibility has helped further il-
illustrate the potential of the vdW-DF method, helping it succeed at numerous applications as summarized in review,[18,19,101] as well as perspectives.[12,75]

On the other hand, a practice-driven approach to picking the vdW-DF exchange is implicitly putting significant trust in the assumed robustness of the $E^{\text{nl}}$ description. Also, having too much flexibility can diffuse the underlying drive for seeking increasingly more versatile XC functionals by hiding an actual limitation in the vdW-DF design logic. As vdW-DF developers, we seek a systematic approach to leverage MBPT input in describing also truly nonlocal-correlation effects[93,113,114,43] and we must therefore restrict the perceived exchange freedom.[27,33,43]

We note that, in principle, the ACF leaves no arbitrariness on how to balance the XC components.[13]

D. MBPT input to set the vdW-DF exchange

We have in the past decade tried to keep to the core vdW-DF focus on improving the accuracy and transferability of the correlation term in balanced, truly nonlocal XC designs. To that end, we have released the consistent-robustness of the correlation term in balanced, truly nonlocal 43

We note that, in principle, the ACF leaves no arbitrariness by hiding an actual limitation in the vdW-DF design logic. As vdW-DF developers, we seek a systematic approach to leverage MBPT input in describing also truly nonlocal-correlation effects[93,113,114,43] and we must therefore restrict the perceived exchange freedom.[27,33,43]

We did this noting that use of the Lindhard-screening logic (and current conservation) sets the XC balance,[27], i.e., the success or failure is then a direct consequence of the quality of approximation for the plasmon dispersion that is used inside $S_{\text{xc}}(\omega)$. There are no other modeling choices that impact the class of consistent vdW-DFs,[27,33,43] The same focus of facilitating XC progress also underpins the present work to define the second-generation RSH AHBR and thus an additional nonlocal-correlation tool chain (comprising B86R, DF2-BR0, and AHBR).

Our suggestion to use the CX-based tool chain,[27] reflects our wish to check if there is already sufficient accuracy in the vdW-DF1/CX/AHCX plasmon modeling[26,69,100] to realize a general-purpose approach for materials and molecules; The CX-based tool chain works as a fair vdW-DF status indicator[27] because it eliminates arbitrariness in the XC balance, as far as it goes.[26,69,100] Of course, strictly enforcing $\delta E^{\text{nl}} = 0$ is impossible[69] but CX uses an exchange form (termed LV-rPW86[26]) that has full $F^{\text{nl}}_x - F^{\text{nl}}_s$ alignment up to $s \sim 2-3$ while it also retains the asymptotic $s$ behavior that is needed for a proper repulsion in that gradient-corrected exchange.[69,62] Importantly, we find the binding contributions from the $E^{\text{nl}}$ term essentially always saturate at small $s < 3$ values. The Lindhard screening logic is still effectively in place to balance the nonlocal-correlation attraction with gradient-corrected repulsion, in the CX-based design.[33,34,43]

For our overall XC development goals, however, we also need a supplement (AHBR) that can contrast the AHCX, and thus point to ideas for improvements. When using the AHCX (or AHBR) tool chain we often find a robust performance, but there are challenges, for example, molecular barrier heights, Fig. 1. As XC developers we desire to learn in which way we should then go forward. We pursue the AHBR design to supplement a broad materials characterization with an option for a controlled ‘functional-derivative’ or ‘functional-contrast’ analysis. Here, getting a ‘derivative’ means being able to interpret observations of performance variations in terms of differences in the nature of the XC-functional designs. A similar idea of making a functional-contrast analysis was also explored for adsorption studies in Refs. [95] and [94]; but the present suggestion to use the AHBR design for a performance contrast should work better. This is because (as we document below) the AHCX and the AHBR have well-defined differences in their design nature.

In practice, our AHBR development work starts from the observation that the B86R design makes it an important, well-motivated, alternative to the CX design logic. Like vdW-DF2 and CX, the B86R respects the observation[29] that the asymptotic form of the exchange-enhancement factor should rise as $s^{2/3}$ asymptotically to appropriately counteract errors in the exchange contributions to (weak) binding.[43,48] The B86R accomplishes that by relying on a modified-Becke86ol[26,100] exchange-enhancement factor,

$$F^{\text{B86R}}_x(s) = 1 + \frac{\mu_{\text{GEA}} s^2}{(1 + \mu_{\text{GEA}} s^2 / \kappa)^{4/5}},$$

where $\mu_{\text{GEA}} = 10/81$ is the small-$s$ expansion coefficient. This low-$s$ form is aligned with correct gradient-expansion result from a diagrammatic MBPT analysis.[43,49,51,101] The B86R does not have a strong argument to set the value of $\kappa$ in Eq. (10) whereas all parameters in the interpolation LV-rPW86 (that defines CX exchange) are set from MBPT input[43] and enforcement of exchange-hole conservation.[52] Also, being a vdW-DF2-type design, the exchange enhancement of the internal functional is set as $F_x(s) = 1 + 0.2097 s^2$, while the expansion of Eq. (10) is given by $\mu_{\text{GEA}} = 10/81$; The Lindhard consistency of CX is therefore lost. However, the use of $\mu_{\text{GEA}}$ in Eq. (10) and B86R does reflect the very same diagrammatic-MBPT input[43] that underpins LV and hence CX.[33,43] This input was reinterpreted[29] between LV analysis[49,48] and the design of exchange in PBEsol.[52] This implies that the net effect of using the $\mu_{\text{GEA}} = 10/81$ value brings the B86R closer to the PBEsol (exchange), without giving up the large-$s$ asymptotic behavior that is also necessary.[55]

In summary, switching between CX-AHCX and B86R-AHBR means using different assumptions when setting the exchange impact on both the plasmon modeling[26,113] and on the actual XC balance. Importantly, this switching for vdW-DFs can be done while still staying within the framework of MBPT analysis.[33,34,43,55,101,113,100] We also note that a switch between the PBE and PBEsol versions of constrained-based GGA can also be done without stepping outside the MBPT guidance.[33] That is, in both the vdW-DF and in the GGA design for XC en-
ergy approximations, one must prioritize among MBPT lessons and implications.

Beyond these general observations, we note that robustness and potential of associated hybrid designs are additional motivations for now crafting a second-generation tool chain off of the B86R. The logic of the vdW-DF+0 hybrid class also leads to the formulation of DF2-BR0, a simple unscreened hybrid that has a 0.25 Fock-exchange mixing factor; Some of us coded also this simple-hybrid formulation in QE, but we have not, before now, documented its properties. Doing a complete-GMTKN55 survey, summarized in Fig. 1 and in the SI, reveals that B86R supports meaningful hybrid formations as there are very few outliers in the DF2-BR0 benchmark results on the GMTKN55 suite. In fact, we find that the DF2-BR0 is the best performer on molecular properties among regular and hybrid vdW-DFs.

Importantly, the performance gain of DF2-BR0 is particularly good for the important transition-state problems, right panel of Fig. 1. Such systems are challenges where AHCX is helping but not bringing a clear resolution either. As XC developers, we can clearly learn by also considering the results of B86R and DF2-BR0 studies. For such learning, we propose to rely on the more general RSH framework where we can interpret AHBRAHCX performance contrasts and ideally channel this information into ideas for future XC development.

E. Analytical-hole design of AHBR

The bottom panel of Fig. 2 compares the key exponent form \( H(s) \) that defines the long-separation shape of the exchange hole. These forms can be asserted in an AH model: The \( H(s) \) form is given by a rational form with parameters fitted subject to constraints so as to accurately describe the B86R exchange behavior without introducing any spurious variation (that cannot be ascribed any physical meaning). In turn, this input is sufficient to recoup not only the exchange enhancement \( F_x(s) \) but also the RSH forms. The procedure, used previously to discuss and understand the CX and AHCX exchange in Ref. 42, is here simply repeated; Table S XI of the SI material reports the parameterizations for the \( H(s) \) forms.

Figure 3 shows spatial variations of the exchange holes at a set of increasing values for the scaled density gradient \( s \) values and as obtained in the AH modeling: The hole variations are completely set by the \( H(s) \) parameterizations. The panels contrast the dependence of the exchange hole on the local electron-density environment for PBE, PBEsol, B86R exchange. Initially (at small \( s \) values,) the B86R exchange hole form follows the PBEsol-exchange nature but does gradually roll over to a more PBE-exchange type behavior; It also eventually approaches a CX-like behavior at large \( s \) values where it respects the lessons of the analysis in Ref. 53. Interestingly, the B86R exchange does, for \( s \lesssim 3 \), perform better than (CX and) PBE exchange in terms of avoiding the formation of deep exchange holes: The local suppression in the B86R exchange hole is not modelled as being larger than the local value of electron-density variation (at \( s \lesssim 3 \)).

We compare this AH analysis for the B86R exchange hole variation also with that for CX and AHCX using the lower panel of Fig. 2 and Fig 1 of Ref. 42. First, it is clear that setting the exchange enhancement by \( \mu_{\text{REA}} \) brings the B86R closer to PBEsol exchange than the CX exchange design. Since the B86R exchange is still constrained by the input from Ref. 55, the large-s behavior rolls over towards that of CX where hole conservations is set by a current-conservation mechanism. Moreover, being an intermediate of the PBEsol and of the CX exchange-hole modeling (Fig. 1 of Ref. 42) the B86R has an intermediate-s behavior (around \( s \approx 3 \)) that is close to the PBE hole form. This is a more generally trusted behavior than that for rPW86, and hence for CX and AHCX.

Taken together, the B86R has exchange characteristics that makes it an exciting alternative to CX for seeking development, in our view: An AHBR-AHCX performance contrast can be interpreted as a consequence of prioritizing different types of well-understood physics input. Also, from the AH analysis of B86R, Figs 2 and 3, it is a small step to complete the RSH RSH functional AHBR. The procedure is given in Refs. 42 and 69 yielding the B86R parameters shown in Table S XI of the SI material. In practice, we simply add these parameters in our previously documented QE subroutine to determine the value (as well as derivatives) of RSH XC functionals (like the DF2-AH and AHCX) that are based on the AH exchange modeling.

III. COMPUTATIONAL DETAILS

All calculations are carried out using the QE code suite, using an in-house coding for the AHBR design. The AHBR code will be released to QE once the ideas of the paper have been accepted for publication.

For hybrids, we used the adaptive-compressed exchange (ACE) implementation to speed up the Fock-exchange evaluation (except in a few cases – ionized Li, Na, and K atoms – where it seemed to prevent an easy convergence). We use the spin vdW-DF formulation (for regular and hybrid vdW-DF, when relevant).

We systematically use the ONCV-SG15 set of PPs at a 160 Ry wavefunction-energy cutoff for all of the here-reported molecular and bulk benchmarking as well as for demonstrator work on base-pair stepping energies in DNA assembly and for a description of CO/Pt(111) adsorption. To document the PP sensitivity, we also provide a characterization of the base-pair stepping energies using the ultrasoft GBRV PP set at 50 Ry wavefunction energy cutoff and 400 Ry density cutoff. For our final RSH vdW-DF demonstration and assessment, we calcu-
TABLE I. Size-convergence of functional-performance assessment: Pilot-study comparison of characteristic WTMAD1 values (in kcal/mol) as obtained for GMTKN53, a 53-benchmark part of the GMTKN55 suite. We contrast the planewave performance characterizations that result when using a standard 10 Å and a larger 15 Å choice of vacuum padding (see text). The convergence tests are done for the AbInit PP set (109) at 80 Ry wavefunction cutoff. We focus on the ‘easily accessible GMTKN53’ subset (33) that excludes (as indicated by asterisks ‘*’) the G21EA benchmark set from the GMTKN55 group 1 and the WATER27 set from the GMTKN55 group 4. This is done because these sets require a separate dielectric handling (33) see Appendix A.

| Tests                      | CX(10 Å) | CX(15 Å) |
|----------------------------|----------|----------|
| Group 1*                   | 5.19     | 5.20     |
| Group 2                    | 4.93     | 4.92     |
| Group 3                    | 7.22     | 7.22     |
| Group 4*                   | 3.36     | 3.35     |
| Group 5                    | 3.78     | 3.79     |
| Group 6 (4* and 5)         | 3.55     | 3.55     |
| GMTKN53 (GMTKN55*)         | 4.79     | 4.79     |

We also provide illustrations of new-tool-chain usage in both a biomolecular and a carbon-capture setting. For the biomolecular-usage illustration, we compare with coupled-cluster reference calculations for a model that replaces the DNA backbone with hydrogen terminations on the individual bases (24,116,117) Here, we perform Γ-point
TABLE II. Comparison of tool-chain performance on Pt- and noble-metal structure and elastic response: Lattice constants $a$ (in Å) and bulk moduli $B_0$ (in GPa). The SI material gives full listings for our computed $a$ and $B_0$ results, as well as a listing of cohesive-energies $E_{\text{coh}}$, for these transition metals and 9 other materials. Experimental values, back-corrected for vibrational effects, are taken from Ref. [115].

|       | CX   | AHCX | AHCX$_{e25}$ | B86R  | AHBR$_{e20}$ | AHBR   | Exper.* |
|-------|------|------|--------------|-------|--------------|--------|---------|
| Cu    | 3.576| 3.587| 3.592        | 3.602 | 3.613        | 3.617  | 3.599   |
| $B_0$ | 163  | 148  | 146          | 151   | 141          | 136    | 144     |
| Ag    | 4.065| 4.078| 4.082        | 4.104 | 4.115        | 4.118  | 4.070   |
| $B_0$ | 115  | 105  | 104          | 102   | 95           | 95     | 106     |
| Au    | 4.101| 4.098| 4.097        | 4.134 | 4.127        | 4.126  | 4.067   |
| $B_0$ | 171  | 168  | 167          | 153   | 152          | 151    | 182     |
| Pt    | 3.929| 3.910| 3.906        | 3.952 | 3.929        | 3.925  | 3.917   |
| $B_0$ | 284  | 298  | 298          | 264   | 278          | 279    | 286     |

Calculations for the set of 10 different two-base-pair combinations (at representative geometries identified in Ref. [117]) using a 30 Å-cubed unit cell and electrostatic decoupling. In effect, we thus assert the AHBR on energies that characterize the DNA assembly.

Finally, for the carbon-capture-usage illustration, we characterize CO$_2$-in-Mg-MOF-74 adsorption enthalpies, using a Γ point $k$-point sampling for the Brillouin-zone integration. We first use CX to compute the unit-cell structure and identify the the adsorption geometry. Next, we proceed to determine relaxations for the CO$_2$ molecules, using the access to forces. We characterize the CO$_2$ vibrations, in a finite-distortion setup using PHONOPY [118] and thus complete predictions of the room-temperature heat of adsorption.

IV. RESULTS AND DISCUSSION

A. Bulk-structure performance

Figure 4 documents a robust bulk-system performance of the new RSH AHBR. More broadly, Fig. 4 contrasts the performance for bulk of the second- and first-generation vdW-DF tool chains (and of DF2-AH). We show only regular functional and screened-hybrid data because the use of the unscreened hybrids, DF2-Br0 and CX0/CX0P, is not motivated for metals [32].

The violin plots summarize deviations in percentage of our computed results for lattice constants $a$, cohesive energies $E_{\text{coh}}$, and bulk moduli $B_0$ from back-corrected experimental values for five transition metals (Cu, Ag, Au, Pt, Rh), one simple metal (Al), four semiconductors (Si, C, SiC, GaAs) and three ionic insulators (LiF, MgO, NaCl) as in Ref. [42] Tables S XIII, XIV, and XV of the SI present a full quantitative presentation, contrasting values computed in the two tool chains with back-corrected experimental values and with those we obtain for the RSH DF2-AH [32]. The subscript on one AHBR-data (and on one AHCX-data) label identifies the extent of Fock-exchange mixing; A corresponding specification is suppressed for ‘AHBR=AHBR$_{e20}$’ (‘AHCX=AHCX$_{e25}$’) since this mixing reflects a recommended default, as explained in the following subsection.

In making a contrast of performance, we compare the position of the mean (median) deviation, shown by a central bar (diamond) and the so-called interquartile range, shown as a bar. This bar reflects the difference of positions for the first and third quartile of the performance distribution (for each functional). We also consider the presence or absence of outliers (open circles), by which we mean a performance that lies beyond markers (wiskers) that identify 1.5 times the interquartile range. Lattice-constant outliers are Au (Au and Ag) for AHBR$_{e20}$ (AHBR) while cohesive-energy outliers are Au and Rh for B86R. There are more outliers for the bulk modulus: Rh for AHCX/AHCX$_{e25}$, GaAs for B86R, and Au for AHBR and AHBR$_{e20}$.

Figure 4 shows that there is a systematic AHBR improvement for lattice constants and bulk moduli compared to B86R, at both 0.20 and 0.25 Fock-exchange mixings. However, the B86R is more accurate than AHBR for predictions of the bulk cohesive energies.

More broadly, Fig. 4 shows that the RSH AHCX is overall a better performer for bulk than AHBR$_{e20}/e25$. This is consistent with our finding that CX has a small bulk-performance edge over B86R. The status is clear, for example, in terms of comparing accuracy for lattice constant. However, the B86R-AHBR tool chain remains accurate on structure and hence useful also for substrate descriptions in heterogeneous system. For example, the lattice-constant accuracy on noble-metal and Pt metals remains within 0.5% deviations relative to (back-corrected) experimental values, Table II (and Table S XIII of the SI).

As an interesting aside, we also note that RSH DF2-AH (implicitly defined in Ref. [42]) is not well suited for bulk-system use. Tables S XIII, XIV, and XV of the SI document that there are large deviations between DF2-AH (and vdW-DF2) predictions and back-corrected experimental values for all of the investigated bulk properties. In the violin-plot Fig. 4 we do not even depict
FIG. 4. Bulk-system performance as asserted in percentage deviations for the CX-AHCX and B86R-AHBR tool chains on descriptions of lattice constants $a$, bulk cohesive energies $E_{\text{coh}}$, and bulk moduli $B_0$. We also show the impact of setting the Fock-exchange mixing: The default AHCX and AHBR mixing value is set at 0.20 and 0.25, respectively, while a subscript identifies an adjustment. We compare our results (listed in the SI), computed in the Born-Oppenheimer approximation, with back-corrected experimental values. The violin plots summarize result-statistics data for 13 solids (5 transition metals, 1 simple metal, 4 semiconductors, and 3 ionic insulators), with the CX and AHCX results repeated from Ref. 42. The set of horizontal bars (of diamonds) reflects the mean (median) deviation in the distributions while the boxes identify the so-called interquartile range, see text. Outliers (identified in the text) are shown by open circles.

the full extent of the interquartile range which for DF2-AH is set by large first/third-quartile relative deviations (0.78% and 3.5% for $a$, 11% and 29% for $E_{\text{coh}}$ and -3% and -34% for $B_0$). Unlike AHBR, the DF2-AH is simply not a reliable option for bulk (and hence for substrates in adsorption studies).

Finally, Fig. 4 shows that the AHBR and AHCX bulk-structure characterizations are relatively insensitive to the choice of the Fock-exchange mixing. The AHCX design is set with a default 0.20 Fock exchange mixing but the overall bulk performance does improve slightly by going from AHCX=AHCX$_{0.20}$ to AHCX$_{0.25}$.

B. Fock-exchange mixing in AHBR: Lessons from large and small systems

We set a default recommendation for the AHBR Fock-exchange mixing, AHBR=AHBR$_{0.25}$, by directly relying on the broad GMTKN55 molecular benchmarking for this second generation RSH vdW-DF. We shall illustrate such usage at a standard HSE-type choice for the Coulomb range-separation or screening parameter, 0.106 inverse Bohr. For individual problems and benchmarks we should ideally proceed to make a system-specific analysis to establish a plausible choice of the Fock-exchange mixing and screening, for example; as pursued in Refs. 119–128. However, our present goal is to facilitate a practical new XC functional and associated second-generation vdW-DF tool chain. To that end, for AHBR, we let convenience set the default Fock-exchange mixing: The use of AHBR$_{0.25}$ is systematically more accurate than AHBR$_{0.20}$ on molecular properties.

We note that the suggested default AHCX 0.20 mixing came from a coupling-constant analysis of the contribution of CX correlation to the atomization energies. The logic of that specification needs not hold for CX0 and...
FIG. 6. Performance of original vdW-DF releases (left panel), of recent regular vdW-DFs (middle panel) and of hybrid vdW-DFs (right panel) on a set of C$_{60}$ isomerization problems. The left (right) panel shows the regular (highly distorted) configuration. The middle panel depicts one of the intermediate configurations in the transformation, as tracked in the C60ISO benchmarks set. The bottom left (middle) panel shows MAD values (in kcal/mol) characterizing the performance of the original vdW-DF releases (recent regular vdW-DFs) on describing energy differences between such configurations. The bottom right panel shows that the hybrid vdW-DFs are needed to substantially improve the description.

AHCX when it comes to large systems (or bulk), let alone for AHBR. Looking at the full survey, in the SI material, we find that moving the AHCX to 0.25 Fock-exchange mixing gives a small performance gain both overall and for all but the NOC-interaction benchmark groups 4 and 5. The impact is in any case limited.

Figure 5 summarizes the inter- and intra-molecular NOC-interaction parts of our functional-performance comparison. The figure concentrates on the benchmark groups that give different indications for a convenient AHCX and AHBR choice of default Fock-exchange mixing. For the vdW-DFs and for rVV10, as well as dispersion-corrected revPBE and HSE, the group-averaged performance is represented in a scatter plot that relies on taking a raw, so-called TMAD, performance indicator that is defined by taking a simple average over the MAD values that we obtain for the individual benchmark. We use the intermolecular (intra-molecular) NOC TMAD value for group 4 (group 5) to set the abscissa (ordinate). Table S IV of the SI materials contains a listing of the TMAD values that are reported in the panels of Fig. 5 as well as for those evaluated for the other GMTKN55 benchmark groups.

Figure 5 also highlights the key impact of excluding (top panel) or including (bottom panel) the impact of the WATER27 benchmark. This benchmark is often excluded in functional comparisons on vdW problems, because it contains the negatively charged small radical OH$^-$ and it is therefore not accessible in a simple benchmarking. Like for the systems in the G21EA benchmark set, this radical has pronounced self-interaction errors as well as convergence challenges that prompted us to pursue the more general planewave benchmarking procedure defined in the appendix. Both insight on water in general and accuracy in WATER27 benchmarking are essential on fundamental science grounds and for practical reasons when pursuing XC development.

We find that the correct inclusion of the WATER27 benchmark set, Appendix A and Fig. 5 has dramatic impact of what we consider XC functional promise for NOC interactions. The assessment map is clearly affected as we switch from the top to the bottom panel of Fig. 5. In fact, it is alone the consistent-exchange class (CX, CX0P/CX0 and AHCX), the unscreened DF2-BR0, and the new AHBR that remain good options for this challenge, at least as asserted by our planewave benchmarking of the GMTKN55 suite. The impact of including the WATER27 challenge on the first-generation tool chain (comprising CX-CX0/CX0P-AHCX) is, relatively speaking, smaller than that on the B86R design (which is moved out of the figure range for intermolecular-NOC assessment values). In contrast, the corresponding B86R-based hybrids, DF2-BR0 and the new AHBR, remain exceptionally well suited to meet the full set of NOC challenges, at least as presently asserted. Looking also at the quantitative measures, Table S IV of the SI material, we find that AHBR$_{0.25}$ appears to perform slightly better than AHBR$_{0.20}$ on intermolecular NOC interactions in the approximate assessment (top panel), but the actual status is different (bottom panel).

Overall, we see the robustness of CX/CX0P/AHCX and DF2-BR0/AHBR as an indication of value in keeping both tool chains as we seek to map for and understand outstanding DFT challenges.

C. Need for vdW-DF hybrids: Molecular examples

Figure 6 considers the problems that exist for the vdW-DFs (and XC functionals in general) in accurate descriptions of some isomerization problems. The top panels of Fig. 6 illustrate the nature of the C60ISO benchmark set in the large-system isomerization group 2 of the GMTKN55 suite. The benchmark set considers the energy differences among 10 meta-stable forms of C$_{60}$ ($E_n$) and the GMTKN55 suite provides reference data for 9 isomerization energy-difference problems, denoted ‘in’ (for $n = 1$ through $n = 9$). These C60ISO problems are specified by reference values for the set of differences $E_{n+1} - E_1$ ($n = 1$ through 9). The top middle panel shows the atomic geometry configuration for the frustrated structure ‘in’ $n = 8$, i.e., a form that determines isomerization problem ‘17’. This is one problem that presents a hard challenge for the vdW-DFs, as here asserted. Meanwhile, for this class of
problems, the HSE+D3 provides an accurate description. The set of bottom panels provide a radar-plot comparison that reveals both vdW-DF limitations and vdW-DF promise. Specifically, the bottom left (middle) panel shows that vdW-DF1 and vdW-DF2 both fail (that recent releases and variants including CX and B86R offer little improvements). However, the bottom-right panels shows that the move to simple, unscreened hybrid and RSH vdW-DFs improves the description to the HSE+D3 level.

We interpret the general nonhybrid-vdW-DF performance issues as arising because we must here describe stretched and frustrated binding. The set of meta-stable C_{60} configurations can also be seen as configurations that define some effective deformation paths taking us from configuration ‘1’ to ‘10’. As such, this type of DFT challenge can be viewed as another transition-state challenge. On the one hand it is therefore surprising that the vdW-DF1 and vdW-DF2 are not adequate, whereas it is expected that the DF2-BR0 and AHBR succeed to the extent that they actually do (lower-right panel). On the other hand, it is clear that there are more than one type of challenges in describing transition states. We can certainly still learn from vd-W-DF2, but it is also clear that we need to cast a wider net, for example, in the form of the present AHBR design work.

Transition-state problems stand out because they involve a comparison of energy terms that must simultaneously reflect several different types of binding, for example, relaxed and stretched or diffusive. Some of us have recently characterized functional performance in the description of CO_{2}-MOF adsorption that happens in concert with a site-specific reaction and resulting CO_{2}-insertion in a diamine-appended MOF. There, the vdW-DF1 and vdW-DF2 also had an accuracy edge but it did not apply for CO_{2} adsorption in the simpler Mg-MOF-74 system. The transition-state problems, in general, are key drivers for XC development.

Figure 1 provides a broad illustration that a general-purpose vdW-DF tool bag must include also the set of hybrid vdW-DFs (including DF2-BR0, AHCH and AHBR). The set of recent regular vdW-DFs (from C09 to B86R and CX) do not remain uniformly accurate when tracked across the 7 individual barrier-height benchmarks of GMTKN55, Fig. 1 (right panel) and SI material. Inspecting SI Tables S V through S X makes it clear that problems of a transition-state nature (such as we deem those of the C60ISO and MB16-43 benchmarks) are in general challenges for nonhybrid vdW-DFs. Nevertheless, we find that the DF2-BR0 performance stands out in terms of both accuracy and cross-benchmark resilience even at transition-state problems. As noted in the introduction, in turn, the DF2-BR0 resilience and transferability for broad types of molecular problems motivate the AHBR design.

Meanwhile, the underlying benchmarking data, Tables S II through S X of the SI material, shows that the vdW-DF2-ah hybrid is not a reliable option for the study of broad molecular properties (just like it failed for bulk). Like the vdW-DF2, it has an acceptable performance for NOC-interaction and barrier-height problems. However, vdW-DF2 and vdW-DF2-ah have real shortcomings when the focus is instead moved to general properties of small- and large-molecule systems.

More generally, our GMTKN55 survey suggests that we cannot hope for one gradient-corrected exchange form (as used in the present type of nonhybrid vdW-DF designs) to always succeed. We find that the assessments on GMTKN55 performance show a spread in performance – for all nonhybrid vdW-DFs – among different classes of problems, see SI material. The same is true, but to a significantly smaller extent, of the CX and B86R-based hybrids. There are many problems, and it is not trivial to secure a good XC balance in the present type of vdW-DF designs, across all of such problems. Specifically, in the vdW-DFs, we add an attractive nonlocal-correlation term to the overbinding that is already in the LDA XC description and we must then rely on nothing but gradient-corrected exchange for stabilization.

The fact that the strength of the vdW attraction may enhance in select situations suggests an additional role for vdW-DF hybrids. They can help us correlate the nature of specific challenges with design choice that we may implement in XC developments. The plan is simply to use and contrast both of the vdW-DF tool chains for explorations, for example, in the presence of difusive interactions when interactions compete and when phase-transformations compete. In the latter case, the presence of an actual or incipient ferroelectric transformation will itself affect the magnitude of the dielectric constant. For example, while the unscreened
hybrid CX0P is highly accurate for the BaZrO$_3$,\textsuperscript{131} we need more to address the SrTiO$_3$.\textsuperscript{22} The complexity of matter has plenty to teach us still but we are getting tools to reveal the details.

In Fig. \textsuperscript{7} we identify a key set molecular-type challenges that we think can be used to drive XC development, without going directly to more complex (but technologically relevant) cases. For example, an accurate description of some transition-state problems (benchmark sets ‘I’, ‘II’, ‘IV’, and ‘V’ in the left panel) requires a unique XC balance that vdW-DF2 occasionally provides, yet vdW-DF2 often fails spectacularly, as seen in the SI material. Ref. \textsuperscript{19} provides an adsorption case where vdW-DF2 also shines by delivering an exceptionally large repulsion by gradient-corrected exchange, but such repulsion is not always needed.\textsuperscript{19} It is therefore good to seek a simpler way to survey for XC issues, as in Fig. \textsuperscript{7}.

Our proposal is based on the experience that we have gained by planewave benchmarking across both the entire GMTKN55 and for CO$_2$ uptake in MOFs, for example, in Refs. \textsuperscript{17, 19, 27, 30} and \textsuperscript{42}. We mostly echo, but also simplify, the logic that led to the definition of the full GMTKN55 suite.\textsuperscript{9} In making this identification, we are assuming that we are working with a vdW-inclusive functional, like the vdW-DFs, so that the design is also capable of dealing with the group 4 and group 5 types of NOC-interaction problems.

Well known challenges for any DFT are, of course, still found in the SIE4x4 set (on self-interaction errors in of neutral and positively charged systems), the MB16-43 (mindless benchmarking) set and the DC13 (difficult-for-DFT) set. To these sets we add the G21EA and WATER27 sets because these two sets suffer from pronounced SIE effects, and in the latter case will also significantly impacts the performance of any given candidate on the group 4 problems. Additionally, we include the IDISP set as it is almost always the challenge that dominates in setting the performance of an XC functional on the group 5 of intramolecular NOC interactions.

From the barrier-height class, we find that the performance varies prominently when inspecting the BHPERI, PX13, and WCPT18 sets. At least it is clear that these benchmarks allow the hybrid benefits to directly manifest themselves. Finally, we point to the above-discussed C60ISO set as a supplement to the focus on barrier-height problems: It clearly reflects both a transition-state nature but it also points to an additional, yet to be identified aspect. This is because the vdW-DF1 and vdW-DF2 barrier-height successes, at the WCPT18 and PX13 benchmark sets, do not port well to this C60ISO set.

Finally, Fig. \textsuperscript{7} illustrates the usefulness of simplifying functional comparisons while focusing on where we can learn more. The figure compares the performance of the first- (CX-based) and second-generation (B86R-based) tool chains in two types of radar plot. The left panel makes the comparison based on G21EA as well as on the sets where there are often massive deviations between DFT results and reference data from quantum-chemistry calculation.\textsuperscript{4} The right panel makes the comparison on the selection of barrier problems and of NOC-interaction problems. An overall impression is that the robustness of the DF2-BR0 and AHBR is confirmed from the testing summary presented in Fig. \textsuperscript{1}. A detailed analysis (combining also with the input on bulk structure characterizations) shows that in a continued development focus we may learn most from: a) the B86R tool-chain logic when problems have a SIE4x4 or barrier character and b) the CX tool-chain logic for a G21EA-type and bulk/substrate problem. For the remainder, both tool-chain approaches seem to have potential.

### Table III. Comparison of the CX-based and B86R-based tool chain performance for DNA assembly

| XC study       | $\Delta E_{WC-step}$ | $\Delta E_{\text{pair}}$ |
|---------------|----------------------|-------------------------|
| B3LYP+D3\textsuperscript{a} | 0.89                 | -                       |
| CX\textsuperscript{b}         | 1.73                 | 0.44                    |
| CX\textsuperscript{GBRV}      | 1.17                 | 1.49                    |
| CX            | 1.48                 | 1.75                    |
| AHCX          | 3.06                 | 3.30                    |
| AHCX$_{0.25}$ | 3.47                 | 3.69                    |
| B86R          | 0.38                 | 0.82                    |
| AHBR$_{0.20}$ | 0.12                 | 0.60                    |
| AHBR          | 0.08                 | 0.52                    |

\textsuperscript{a}Ref. \textsuperscript{117}
\textsuperscript{b}Ref. \textsuperscript{124}

### D. Base-pair stacking in a DNA model

DNA can be seen as a stacking of Watson-Crick (WC) base pairs that are essentially flat and therefore have a significant (eV-scale) vdW attraction\textsuperscript{22,23,113,114} from one base pair to the next. The WC pairs are steps in the resulting double-helix DNA structure. There are in total 10 possible combinations for 2 steps, i.e., base-pair combinations that are here denoted ApA, ApT, ApC, ApG, CpC, CpG, GpC, TpT, TpG, and TpG: Ref. \textsuperscript{117} identifies and illustrates a set of possible atomic positions for these base-pair combinations. We seek to compute such base-pair stepping energies since the mutual vdW attraction might have driven the DNA self-assembly in the first place, as life emerged.\textsuperscript{131,132}

Some of the DNA cohesion comes, of course, from the

\textsuperscript{X}
presence of the sugar-phosphate backbones that incorporate and organize the WC bases into two strands, with the sequence of WC bases, adenine (A), thymine (T), cytosine (C), or guanine (G), setting the genetic code. The strands are complementary in the sense that the WC-base sequence must be exactly matched, with each of the individual bases having only one suitable counter part, i.e., forming steps that must have one of the A-T, T-A, C-G, or G-C forms. Our base-pair-combination nomenclature, ApA though TpG, reflects the observation that it is sufficient to track the code sequence on one of the DNA strands. The DNA strands are mutually bonded, by a combination of hydrogen and vdW binding, but the energy of the WC pairings (A-T, T-A, C-G, G-C, among one base and its counter part directly across) is not our present focus. Instead, we seek to understand the extent that the mutual step-binding energies contribute to the DNA cohesion, using a DNA model that ignores the backbone but instead relies on hydrogen terminations.

To set us up for future, more general DNA explorations, we contrast the performance of the first- and second-generation vdW-DF tool chains relative to the reference descriptions provided in Ref. [117]. That is, we compare with so-called domain-based pair natural orbital couple-cluster (‘DLPNO-CCSD(T)’) calculations at fixed reference base-pair combination structures. The DLPNO-CCST(T) method is also used for setting reference energies of the GMTKN55 suite. We compute the base-pair stepping energies as total-energy differences between the full system and the two WC base pairs, for example, in the case of the ApC combination:

$$\Delta E_{W_C-step}^\text{ApC} = E_{\text{ApC}} - E_{\text{A-T}} - E_{\text{C-G}}.$$  \hspace{1cm} (11)

The reference work [117] also provides data for the sum of pair interactions among the four bases, $\Delta E_{B-pair}$, excluding the two WC pairings (as indicated by the prime). This pair summation is illustrated in the abstract figures of Refs. [117] and [134].

Tables S XVI and S XVII of the SI material compares the performance of CX/AHCX/AHCX0.25 and of B86R/AHBR0.20/AHBR for each of the base-pair combinations, reporting (in kcal/mol) the $\Delta E_{W_C-step}$ and $\Delta E_{B-pair}$, respectively. We also report mean deviation (MD) and MAD values relative to the DLPNO-CCSD(T) calculations. Table III summarises the performance comparison (in terms of 10-base-pair-combination averages) and makes it clear that the AHBR is a strong performer.

We find that B86R performs better than CX and that the AHBR functional design is in fact very accurate also for descriptions of the DNA stepping energies. This is especially true when it is used in the suggested default mode with a 0.25 Fock-exchange mixing. The description is significantly more accurate than a standard choice of dispersion-corrected hybrid DFT provides, as also listed in Table III. This finding is in itself encouraging.

At the same time, we also note that the AHBR is somewhat less accurate when it is instead used to study energies from the sum of pair contributions, $\Delta E_{B-pair}$. As in Ref. [134] we find that the vdW-DF based descriptions of the stepping energies are benefitting from a cancellation of errors that affect the descriptions of pairing between individual bases.

E. CO adsorption on Pt

Figure 8 presents a schematics of Co/Pt(111), contrasting the atomic configurations in two competing adsorption sites, TOP and FCC. The atom positions are set from a CX characterization of the adsorption-induced relaxations. We note that the CX has a lattice constant that is in close agreement with experimental characterizations of Pt. Table IV summarises our comparison of tool chain performance for the classic CO/Pt(111) problems. All calculations are performed at the CX-lattice constant and adsorption induced geometries to simplify comparisons among the functionals and with experimental observations, as summarized in the caption. We note that the B86R lattice constant is slightly larger (and further from experiment). We also note that a AHBR adsorption study at the AHBR lattice constant should have a more narrow Pt $d$ band and therefore smaller adsorption energies.

Table IV summarizes our comparison of tool chain performance for the classic CO/Pt(111) problems. All calculations are performed at the CX-lattice constant and adsorption induced geometries to simplify comparisons among the functionals and with experimental observations, as summarized in the caption. We note that the B86R lattice constant is slightly larger (and further from experiment). We also note that a AHBR adsorption study at the AHBR lattice constant should have a more narrow Pt $d$ band and therefore smaller adsorption energies.

We find that neither CX nor the default AHCX description (using a 0.20 fraction of Fock-exchange mixing) offers an improvement in the description of the site-preference challenge. In contrast the new AHBR works (as do AHBR0.20 and AHCX0.25,) bringing the site-preference description in alignment with experimental observations. Of course, the AHBR use does not offer...
FIG. 8. Schematics of the CO-on-Pt(111) problem: Competition between TOP- and FCC-site adsorption. The image is generated using the VESTA program.  

TABLE IV. Comparison of TOP- and FCC-site CO/Pt(111) binding energies $E_{\text{bind}}$, site-preference energies $\Delta E_{\text{site}} = E_{\text{TOP}}^{\text{bind}} - E_{\text{FCC}}^{\text{bind}}$, as well as of molecular-gap results $E_{\text{CO}}^{\text{gap}}$ (all in eV). The first and second block are for the CX-AHCX and the here-defined B86R-AHBR tool chain, respectively. For both tool chains we also illustrate the impact of the choice of the Fock-mixing (as identified in the functional-label subscript). All results are provided for the CX provided substrate lattice constant $a_0 = 3.929 \, \text{Å}$, molecular structure, and adsorption-induced deformations. Experimental observations of CO adsorption find TOP site adsorption with binding energy $\Delta E_{\text{site}} = -1.32 \, \text{eV}$. 

|               | CX AHCX | AHCX0.25 | B86R AHBR0.20 | AHBR |
|---------------|---------|----------|---------------|------|
| $E_{\text{TOP}}^{\text{bind}}$ | -1.830  | -1.949   | -1.683        | -1.824 |
| $E_{\text{FCC}}^{\text{bind}}$  | -1.966  | -1.954   | -1.780        | -1.752 |
| $\Delta E_{\text{site}}$       | 0.133   | 0.005    | 0.097         | -0.034 |
| $E_{\text{CO}}^{\text{gap}}$    | 7.048   | 8.768    | 7.081         | 8.798 |

FIG. 9. Correlation between results for the CO molecular gap $E_{\text{gap}}^{\text{CO}}$ and for the CO/Pt(111) site-preference energy $\Delta E_{\text{site}}$ in both the CX-based tool chain and in the new second-generation B86R tool chain. We keep the adsorption geometry fixed in all studies.

FIG. 10. Optimized crystal structure of CO$_2$ adsorbed Mg-MOF-74. The image is generated using the VESTA program.

a complete resolution of the long-standing CO/Pt(111) puzzle because it is still over-estimating the actual adsorption energy $E_{\text{bind}}$.

Figure 9 presents a mapping of correlation between the CO gap and the site-preference energy, as obtained in the CX-based tool chain and in the new B86R tool chain. The adsorption is often discussed in terms of the Blyholder model and, as such, controlled by the substrate electronic structure (which we must accurately characterize to correctly describe the molecule-to-substrate charge transfer) and the molecular gap (that we must accurately characterize to correctly describe the back donation). We keep the adsorption geometry fixed in all calculations, and thus contrast the direct effects that the functionals have on both the molecule gap and substrate electronic structure.

We make several observations from Fig. 9. First, there is a strong impact on the molecular gap by moving from a regular function to the associated RSH. Second, this gap variation effectively controls the prediction of site preference within a given tool chain. Third, there is also a systematic effect of switching between the tool chains and thus changing the substrate electronic structure. In effect, the figure documents that the well-defined AHCX-AHBR differences (in terms of nonlocal-correlation-and-exchange-design details) have a direct impact on the Pt(111) electronic description. This impact is making the AHBR better at reflecting the true site variation in the CO/Pt(111) adsorption energy.

F. MOF adsorption in Mg-MOF-74

Figure 10 shows a schematic of CO$_2$ adsorption in Mg$_2$(dobdc), known as Mg-MOF-74. We further as-
TABLE V. Comparison of the CX-based and B86R-based tool chain performance for CO₂ adsorption energies Mg–O bond distance in Mg-MOF-74; For the former (latter) we compare computed results in kJ/mol (in Å) to facilitate an easy comparison with other MOF studies. All calculations are performed at the CX characterizations for the MOF geometry but permitting the CO₂ to relax. The subscript ‘BO’ identifies Born-Oppenheimer results, while a superscript ‘room’ identifies adsorption-free-energy values that include the effects of vibrations as described at room temperature. The experimental adsorption energies reflect room-temperature observations. Values in parentheses indicate deviations (in %) from reported experimental values.

|              | CX      | AHCX    | AHCX_{0.25} | AHBR    | Exper. |
|--------------|---------|---------|-------------|---------|--------|
| Mg–O         | 2.29 (1.1) | 2.27 (-0.1) | 2.26 (-0.4) | 2.25 (-1.0) | 2.27 |
| E_{ads,BO}   | -53.7 | -58.5 | 59.7 | -50.3 | - |
| ZPE          | 2.8 | 2.4 | 2.5 | 2.6 | |
| TE           | 1.1 | 1.6 | 1.3 | 1.1 | |
| H′_{ads}^{room} | -49.7 (14) | -54.5 (25) | -55.9 (29) | -46.6 (7) | -43.5 |

To compare with the experimental data for CO₂ heat of adsorption ΔH_{ads}, we also compute the zero-point energy (ZPE) and thermal-energy (TE) corrections. To that end, we calculate and contrast the vibrational frequencies of adsorbed CO₂ and free CO₂ using a finite difference approach with the Phonopy package. Specifically, we displace each atom of CO₂ in twelve random directions with a constant displacement distance (0.03 bohr) to extract corrections that hold at 298 K. Following Ref. we add these values to ΔE_{ads} to get the heat of adsorption:

\[ \Delta H_{ads} = \Delta E_{ads} + \Delta ZPE + \Delta TE. \] (15)

Table V compares our results, contrasting the vdW-DF tool chain descriptions of CO₂ uptake in Mg-MOF-74 with experiments. We find that the AHCX descriptions (for both a 0.20 and 0.25 choice of assumed Fock exchange mixing) actually worsens the good accuracy that CX itself has for characterizing the adsorption energies. Their deviations of the CO₂ adsorption enthalpies from the experiment are 25% and 29%, respectively. However, the AHBR (investigated here at the default 0.25 Fock-exchange mixing) provides a characterization that is more accurate than the CX results with the 7% deviation for H_{ads}.

For the Mg–O bond distance, all four nonlocal-correlation functionals generally show good performance, giving structure predictions within 1% deviation from the experiment. Among these, AHCX with 0.20 Fock exchange mixing most accurately predicts the experimental Mg–O bond distance.

V. SUMMARY AND OUTLOOK

We have launched a second-generation-vdW-DF tool chain comprising B86R, DF2-BR0, and a here-defined new RSH vdW-DF, termed AHBR. Being based on the vdW-DF2-formulation of the nonlocal-correlation term, it supplements the recently-defined CX-CX0P-AHCX tool chain. The B86R functional uses some of the MBPT input on gradient-corrected exchange that also defines CX (and the first-generation tool chain). However, the B86R is based on a slightly different interpretation of the MBPT input for exchange in the weakly perturbed electron gas. This adjustment towards an often-used exchange comes at the expense of breaking current conservation and compliance with the Lindhard screening logic. The AHBR/new tool chain supplements the first vdW-DF based tool chain by limited, well-controlled and documented adjustments. Accordingly, we expect that, using a contrast focus, their combined use 1) can help identify the nature of problems that presently challenge the vdW-DF method and 2) provide analysis that suggests ideas for a systematic vdW-DF development.

To detail and validate the AHBR design idea, and to illustrate potential usefulness for intensifying a tool-chain exploration, this paper also has several theory contributions that may have interest in their own right. First, we extend the AH analysis of exchange from our previous CX to the present B86R focus. Second, we illustrate how the use of an electrostatic-environment handling allows us to complete full planewave-based functional assessment on the GMTKN55 benchmark suite; Unlike in Refs. and we pursue benchmarking for negatively charged ions and radicals as a limiting process, adiabatically removing a fictitious dielectric environment. Third, looking at bulk properties (and NOC interaction
systems in GMTKN55) we illustrate and discuss the impact on adjusting the extent of Fock exchange mixing in both AHCX and AHBR, suggesting a default value \( \alpha = 0.25 \) for this new RSH vdW-DF. Fourth, we contrast and discuss a full GMTKN55-suite benchmarking for the vdW-DFs that QE supports. Fifth, we document that the six XC functional members of the two vdW-DF tool chains are robust and top performers both as asserted on average and on various types of material properties. The B86R-based vdW-DF tool chain has a performance edge for broad molecular properties while it also remains accurate for bulk. In contrast, the CX-based tool chain has a small edge for bulk properties while it remains in the performance lead for most of the types of molecular properties that the GMTKN55 benchmark suite tests.

As an exciting sixth point, we find that DF2-BR0 and AHBR have the strongest performance gains (over B86R itself and over AHCX) for the transition-state-barriers and some isomerization problems that often challenge DFT. As a result, the AHBR may be most useful when (as in catalysis) we must simultaneously account for both stable binding configurations and for the chemical reactions between such materials forms. The finding suggests that the B86R uses the MBPT input slightly better that what the CX-design does for molecules. These observations inspire us to continue vdW-DF method development.

Finally, we proceed to test the AHBR on demonstrators, i.e., applications concerning the elementary steps of DNA assembly as well as the CO/Pt(111) and CO2-Mg-MOF-74 adsorption problems. This is done to reflect the key importance of biochemistry, heterogeneous catalysis, carbon capture, and molecular-energy systems. We provide a double-tool-chain exploration to reveal relative strengths and weaknesses and we deliberately kept the structure fixed in the comparison. That is, for the DNA assembly problem we rely on reference geometries, in the CO/Pt adsorption problem we keep the atoms as given by a CX characterization of adsorption-induced relaxations, and in the CO2-capture study we set a fixed MOF structure by a CX characterization. We also consider AHCX/AHBR characterizations at 0.20 and 0.25 Fock-exchange mixing for completeness.

We show that AHBR provides a highly accurate description of the DNA stepping energies and a Pt adsorption characterization that is in good agreement with the observed site-energy preference and experimental observation (although overshooting the magnitude of the actual CO binding compared to experiment). The AHBR results for CO2 uptake is found to be in excellent agreement with measurements once vibrational corrections are accounted for.

Figure 7 supplements this above summary with an outlook, identifying a set of molecular benchmarks that can drive developments of the vdW-DF method. This is true even if the access to simple and range-separated hybrid vdW-DFs, including AHCX and AHBR, have helped. There are obviously also many extended-system challenges, like CO2 adsorption in the diamine-extended Zn2(dobpdc) MOF19 and a select range of small-molecule adsorption systems. However, for XC developments it is not always wise to go for complex systems directly. Instead, we have learned to keep attention on some indicators, and these computationally simpler molecular cases, Fig. 7, are judged to be fair indicators of overall robustness (or lack thereof). They are in any case problems that a truly general-purpose XC functional must also be able to navigate and master.

ACKNOWLEDGEMENT

We thank Carl M. Frostenson for useful discussions. Work supported by the Swedish Research Foundation (VR.) through grants 2018-03964 and 2020-04997, the Swedish Foundation of Strategic research, through grant IMF17-0324, and Sweden’s Innovation Agency (Vinnova,) through project No. 2020-05179, as well as the Chalmers Area-of-Advance-Production theory activities, and U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award DESC0019992, KIST Institutional Program (Project No. 2E31201), the program of Future Hydrogen Original Technology Development (2021M3IA1083946), through the National Research Foundation of Korea (NRF), funded by the Korean government (Ministry of Science and ICT (MSIT)). The authors also acknowledge computer allocations from the Swedish National Infrastructure for Computing (SNIC) under Contracts SNIC2019-2-19, SNIC 2020-3-13, SNIC2021-3-18, the Chalmers Centre for Computing, Science and Engineering (C3SE) as well as KISTI Supercomputing Centre (Project No. KSC-2020-CRE-0189).

Appendix A: Planewave molecular benchmarking

This appendix summarizes our strategy for completing a plane-wave molecular benchmarking across the full GMTKN55 suite. In particular, we explore the nature and impact of convergence-related factors that make benchmarking challenging, but which, as we show, can also systematically be handled and circumvented. As such, the appendix motivates and validates the molecular benchmarking that we also report in this paper.

1. Impact of spurious vdW attraction

Our focus on nonlocal-correlation functionals presents, in principle, a problem for plane-wave benchmarking: The vdW attraction is long-ranged and causes at least some spurious coupling between the periodically repeated images. However, we can document that the net impact by false, intercell vdW coupling on each of the GMTKN55 benchmarks is vanishing, well within our
overall benchmarking target of 0.01 kcal/mol\textsuperscript{114} To that end, we consider the asymptotic limit of the vdW-DF method\textsuperscript{31113114}

We proceed as follows. First, for the atomic configurations of molecules or cluster (in each of the roughly 2450 different unit cells that define the GMTKN55 suite\textsuperscript{23} we evaluate an effective molecular/cluster $C_{6}^{\text{mol}}$ interaction coefficient. These coefficients characterize the spurious vdW couplings in the asymptotic vdW-DF limit that is defined and used in Refs. \textsuperscript{112} and \textsuperscript{113} Here we use an in-house extension of the QE code suite, starting on a set of CX calculations for the self-consistent electron density variation. This initial step was done in a pilot study, using the more electron-sparse AbInit PP\textsuperscript{129} at 80 Ry cutoff, but typically in our standard benchmarking setup, i.e., with a minimum of 10 Å vacuum padding in the choice of unit cells. However, we used smaller unit-cell sizes in the G21EA set and for the negatively charged OH$^{-}$ radical in the WATER27 set\textsuperscript{27} for reasons that are further discussed below.

We also compute (for each and every unit-cell problem in GMTKN55) a corresponding interaction-energy contribution to a given unit cell $i$\textsuperscript{113}

$$E_{vdW,i} = - \sum_{j \neq i} \frac{C_{6}^{\text{mol}}}{|\mathbf{R}_{j} - \mathbf{R}_{i}|^{6}}. \quad (A1)$$

Here, $\mathbf{R}_{j}$ denotes a Bravais vector of the periodic computational setup in our QE calculations. We observe that this measure of asymptotic vdW attraction, Eq. (A1), is different from a full vdW-DF method study\textsuperscript{31113} on fundamental reasons that are discussed in Refs. \textsuperscript{34} and \textsuperscript{113} We would, for example, not get exactly the same values if we tracked the spurious coupling in a super-cell study. However, we are setting up of molecular benchmarking with a large vacuum padding, see Section III. The use of Eq. (A1) is therefore an acceptable approximation that can gauge the expected magnitude of spurious vdW coupling in our plane-wave calculations.

Table S XII illustrates how we use Eq. (A1) to validate that the spurious-vdW impact on our benchmarking is truly neglectable. From Eq. (A1) we determine the extent that the spurious intercell-vdW coupling is causing an offset on the periodic-cell vdW-DF results, for each of the roughly 2450 molecular-process energies in the GMTKN55\textsuperscript{23} Next, we define a set of 55 per-benchmark MAD-offset values by tracking the offset relative to the reference energies. Table S XII reports these benchmark-specific error estimates in descending order (truncated to five sets). The largest impact seems to appear for the G21EA set, but that impact-estimate is also an exaggeration. This is because we dumped the pilot study to a reduced unit-cell size (8 Å total) whenever we faced convergence issues for the negatively charged ions and small radicals in the G21EA set\textsuperscript{27,42,80} Table S XII makes it clear that the false-vdW impact on our plane-wave molecular benchmarking is significantly less that 0.01 kcal/mol.

2. Impact of spurious electrostatic coupling

Table I of the main text provides documentation that we use a sufficiently large plane-wave-benchmarking setup. The table summarizes the following testing that our use of a standard 10 Å vacuum padding (in our automatic scripting of QE input files) is sufficient to ensure an electrostatic decoupling among the images in the periodic-cell setup, bringing the assessment error down to a desired 0.01 kcal/mol limit\textsuperscript{114} for averaged measures. This validation of our benchmarking strategy was done brute-force, that is, by repeating a standard CX testing with one where instead we use a 15 Å vacuum padding. Again we use the AbInit-PP setup (at 80 Ry) but otherwise following the regular problem specifications for all of the individual problems. However, the focus is here limited to the easily-accessible ‘GMTKN53’ subset of the GMTKN55, excluding the G21EA and WATER27 sets, as was also done in Ref. \textsuperscript{122} There are fundamental reasons why it is sometimes impossible -- and generally not practical – to converge the electronic-structure of small negative ions and small negative radicals in QE using large unit-cell sizes\textsuperscript{27,42,80}

We find, Table I that the net impact of going to a truly large vacuum padding in the set up, is bounded by 0.01 kcal/mol. Here the assessment is done in terms of the weighted WTMAD1 measures as resolved on the GMTKN55 bench groups. As indicated by asterisks $\ast$, the reported WTMAD1 measures are slightly adjusted (as described in Ref. \textsuperscript{42}) due to the ‘GMTKN53’ focus. We deem use of our standard (10 Å) vacuum padding setup, Section III, validated also in terms of securing sufficient electrostatic decoupling\textsuperscript{114}

3. Instabilities driven by self-interaction errors

A planewave benchmarking setup like ours can technically only reach a true complete-basis limit\textsuperscript{114} when we check the size convergence, at least in principle, to the infinite-size limit. Effectively, the QE plane-wave code sets the average potential to zero\textsuperscript{117} This means that, for any finite unit-cell size, the potential value in regions far from the nuclei (in isolated molecule problems) lies slightly above the true vacuum floor. A SIE impact, that arises in small negative ions will, in part, be masked by this potential offset in the plane-wave code; For sufficiently small unit-cell sizes one can even craft a fictitious electron trapping in general XC functionals, even if non-hybrids may not actually be able to trap it at all in a fully converged description\textsuperscript{27,42,80} Unfortunately, a direct discussion of convergence with unit-cell size for all of the GMTKN55 is of limited meaning.

Figure I1 reports a successful (but highly tedious and difficult) electronic-structure convergence that we provide for CX (in a plain QE) for the OH$^{-}$ system, as described in our setup for characterizing it in the WATER27 set. Here the use of our automatic (benchmark-
specific) setting of the unit-cell size (Section III) leads to use of a 22.8 Å cell size. This is a highly challenging choice of the unit cell (in terms of convergence) for a negatively charged system, as represented in a planewave code. The figure shows the variation of the electrostatic potential along the O-H axis, together with the energy position of the highest occupied molecular (HOMO) level of the negatively charged radical. The insert focuses on the near-molecule regions and tracking differences between the neutral OH lowest-unoccupied molecular level (broken red bar) and the OH$^-$ HOMO level.

FIG. 11. Self-interaction error (SIE) driven instability for convergence in the OH$^-$ radical, here illustrated by noting a closeness of the HOMO energy position and the vacuum level. The figure reports a CX study, showing the self-consistently computed electrostatic potential along a ray containing the nuclei (potential dips) and OH$^-$ bonding region. The QUANTUM ESPRESSO approximation for the vacuum level is found at the plateau value, here located just above the OH$^-$ HOMO level (solid blue horizontal bar). The insert provides a closer look, showing the potential barriers surrounding the molecular regions and tracking differences between the neutral OH lowest-unoccupied molecular level and the OH$^-$ HOMO level.

We note that these SIE-impact problems are important for DFT usage, yet a general planewave-DFT characterization appears impossible. We need to track and understand the problems as they relate to charge-transfer processes and they point to potentially deeper SIE problems of XC functionals. However, to quantify SIE impact on the G21EA and WATER27 benchmark sets, we do need to actually trap the HOMO level — and that is not directly possible in a complete-basis approach for some negative ions and small radicals.

However, there is a good solution available from analysis. There is an electrostatic nature to the SIE-impact problems in G21EA and WATER27 benchmarking. Trapping the HOMO level in a negatively charged small system is difficult, see Fig. because the electrostatic potential, on the one hand, must overshoot to positive values and, on the other hand, will be set by an (unphysical) exponential decay in the cross-over to that asymptotic-repulsion region. As a consequence, the trapping region becomes excessively narrow, pushing the would-be HOMO level above the (true) vacuum level, in some cases. Meanwhile, planewave DFT sets the average potential to zero, see for example Ref. giving an asymptotic potential variation with a floor that will approach the true vacuum (and thus cause QE problems) as we push for size convergence. Frustratingly, our planewave benchmarking appears to be foiled by its very strength, namely that we can fairly easily approach the complete-basis-set limit.

Fortunately, the electrostatic nature of the problem also gives us a way to leverage the planewave advantages for size-converged G21EA, WATER27 benchmarking, and hence for complete GMTKN55 assessments. The idea is to 1) introduce a control knob that systematically affects the vacuum position in QE, 2) obtain well-defined G21EA/WATER27 assessments as a function of the control-knob value, and 3) extrapolate these estimates as we turn off the controlled stabilization. We simply do perturbation theory in reverse.

In practice, we rely on an electrostatic-environment extension of the QE code suite. Use of a fictitious dielectric-constant $\varepsilon_\infty > 1$ softens the repulsion in the electrostatic potential for the HOMO-level trapping. For sufficiently large $\varepsilon_\infty$ values we reduce the spurious quantum-confinements effect and can thus obtain a SIE-robust G21EA/WATER27 assessment that works at general unit-cell sizes.

Figure shows the procedure used to assess the OH electron affinity (as described for the WATER27 setup) for both CX (left panel) and AHBR (right panel). Not-
The choice of a fictitious vacuum-dielectric constant $\varepsilon_\infty$ counteracts the SIE-driven instability. The resulting electron-affinity energies $E_{\text{EA}}$ (shown as dots) vary with $\varepsilon_\infty$ but have a rapid convergence with the choice of the cell dimensions.

**FIG. 12.** Robust, environment-extended quantum Espresso determination of the OH$^-$ electron-affinity energies in CX (left panel) and hybrid AHBR (right panel). The choice of a fictitious vacuum-dielectric constant $\varepsilon_\infty > 1$ counteracts the SIE-driven instability. The resulting (environment-adjusted) OH$^-$ electron-affinity energies $E_{\text{EA}}$ (shown as dots) vary with $\varepsilon_\infty$ but have a rapid convergence with the choice of the cell dimensions.

*hyldgaar@chalmers.se*

1 Equal contribution

1. A. D. Becke, J. Chem. Phys. 140, 18A301 (2014).
2. K. Burke, J. Chem. Phys. 136, 150901 (2012).
3. J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, K. Burke, J. Chem. Phys. 2, 1 (1992).
4. A. D. Becke, J. Chem. Phys. 96, 246 (1992).
5. J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, K. Burke, J. Chem. Phys. 102, 11822 (1995).
6. N. Rosi, J. Eckert, M. Eddaoudi, D. Vodak, J. Kim, M. O’Keefe, and O. Yaghi, Science 300, 1127 (2003).
7. D. C. Langreth, B. I. Lundqvist, S. D. Chakarova-Käck, V. R. Cooper, M. Dian, P. Hyldgaard, A. Kelkkanen, J. Kleis, L. Kong, S. Li, P. G. Moses, E. Murray, A. Puzder, H. Rydberg, E. Schröder, and T. Thonhauser, J. Phys.: Condens. Matter 21, 084203 (2009).
8. L. Kong, V. R. Cooper, N. Nijem, K. Li, J. Li, Y. J. Chabal, and D. C. Langreth, Phys. Rev. B 79, 081407 (2009).
9. K. Lee, E. D. Murray, L. Kong, B. I. Lundqvist, and D. C. Langreth, Phys. Rev. B 82, 081101 (2010).
10. Q. Li and T. Thonhauser, J. Phys.: Condens. Matter 24, 422404 (2012).
11. V. R. Cooper, Y. Ihm, and J. R. Morris, Phys. Proc. 34, 32 (2012).
12. A. R. Millward and O. M. Yaghi, J. Am. Chem. Soc. 127, 17998 (2005).
13. T. M. McDonald, W. Lee, J. A. Mason, B. M. Wiers, C. S. Hong, and J. R. Long, J. Am. Chem. Soc. 134, 7056 (2012).
14. T.-H. Bae and J. R. Long, Energy Environ. Sci. 6, 3565 (2013).
15. T. M. McDonald, J. A. Mason, X. Kong, E. D. Bloch, D. Gygi, A. Dani, V. Crocella, F. Giordano, S. O. Odoh, W. S. Drisdell, B. Vlassisavljevich, A. L. Dzubak, R. Poloni, S. K. Schnell, N. Planas, K. Lee, T. Pascal, L. F. Wan, D. Prendergast, J. B. Neaton, B. Smit, J. B. Kortright, L. Gagliardi, S. Bordiga, J. A. Reimer, and J. R. Long, Nature 519, 303 (2015).
16. S. Zuliuaga, P. Canepa, K. Tan, Y. J. Chabal, and T. Thonhauser, J. Phys.: Condens. Matter 26, 133002 (2014).
17. R. Poloni, K. Lee, R. F. Berger, B. Smit, and J. B. Neaton, J. Phys. Chem. Lett. 5, 861 (2014).
18. W. L. Queen, M. R. Hudson, E. D. Bloch, J. A. Mason, M. I. Gonzalez, J. S. Lee, D. Gygi, J. D. Howe, K. Lee, T. A. Darwish, M. James, V. K. Peterson, S. J. Teat, B. Smit, J. B. Neaton, J. R. Long, and C. M. Brown, Chem. Sci. 5, 4560 (2014).
19. J.-H. Lee, P. Hyldgaard, and J. B. Neaton, Submitted to J. Chem. Phys. (2022).
20. E. Lee and K. A. Persson, Nano letters 12, 4624 (2012).
21. K. Persson, Y. Himuna, Y. S. Meng, A. Van der Ven, and G. Ceder, Physical Review B 82, 125416 (2010).
22. V. Shukla, N. K. Jena, S. R. Naqvi, W. Luo, and R. Ahuja, Nano Energy 58, 877 (2019).
23. V. Shukla, R. B. Araujo, N. K. Jena, and R. Ahuja, Physical Chemistry Chemical Physics 20, 22008 (2018).
24. V. R. Cooper, T. Thonhauser, A. Puzder, E. Schröder, B. I. Lundqvist, and D. C. Langreth, J. Am. Chem. Soc. 130, 1304 (2008).
25. D. Le, A. Kara, E. Schröder, P. Hyldgaard, and T. S. Rahman, J. Phys.: Condens. Matter 24, 424210 (2012).
26. S. Umrao, A. Maurya, V. Shukla, A. Grigoriev, R. Ahuja, M. Vinayak, R. Srivastava, P. Saxena, I.-K. Oh, and A. Srivastava, Materials Today Bio 1, 100001 (2019).
27. C. M. Frostenson, E. J. Granhed, V. Shukla, A. Van der Ven, and J. B. Neaton, Physical Review B 82, 125416 (2010).
28. V. Shukla, N. K. Jena, S. R. Naqvi, W. Luo, and R. Ahuja, Nano Energy 58, 877 (2019).
29. V. Shukla, R. B. Araujo, N. K. Jena, and R. Ahuja, Physical Chemistry Chemical Physics 20, 22008 (2018).
30. V. R. Cooper, T. Thonhauser, A. Puzder, E. Schröder, B. I. Lundqvist, and D. C. Langreth, J. Am. Chem. Soc. 130, 1304 (2008).
31. D. Le, A. Kara, E. Schröder, P. Hyldgaard, and T. S. Rahman, J. Phys.: Condens. Matter 24, 424210 (2012).
32. S. Umrao, A. Maurya, V. Shukla, A. Grigoriev, R. Ahuja, M. Vinayak, R. Srivastava, P. Saxena, I.-K. Oh, and A. Srivastava, Materials Today Bio 1, 100001 (2019).
33. C. M. Frostenson, E. J. Granhed, V. Shukla, A. Van der Ven, and J. B. Neaton, Physical Review B 82, 125416 (2010).
34. V. Shukla, N. K. Jena, S. R. Naqvi, W. Luo, and R. Ahuja, Nano Energy 58, 877 (2019).
35. V. Shukla, R. B. Araujo, N. K. Jena, and R. Ahuja, Physical Chemistry Chemical Physics 20, 22008 (2018).
36. V. R. Cooper, T. Thonhauser, A. Puzder, E. Schröder, B. I. Lundqvist, and D. C. Langreth, J. Am. Chem. Soc. 130, 1304 (2008).
37. D. Le, A. Kara, E. Schröder, P. Hyldgaard, and T. S. Rahman, J. Phys.: Condens. Matter 24, 424210 (2012).
38. S. Umrao, A. Maurya, V. Shukla, A. Grigoriev, R. Ahuja, M. Vinayak, R. Srivastava, P. Saxena, I.-K. Oh, and A. Srivastava, Materials Today Bio 1, 100001 (2019).
The AHB21 set also contains the OH radical, but sitting in a benchmark set that can use a small unit cell, the convergence problems seldom manifest themselves. We have still in present GMTKN55 survey used the electrostatic-environment approach to reliably assert also the performance for the AHB21 benchmark set; However, finding no relevant correction for AHB21 relative to the default brute-force (that is, native or environment-free) QE exploration (within the present selection for unit-cell sizes), we focus the discussion of self-interaction-error impact on G21EA and WATER27.

The vdW-DF-BEEF performance is not asserted, even if it is now also implemented in the QE code suite. The idea of that design is to seek system-optimized exchange partners for the nonlocal-correlation description. This design idea does not lend itself easily to broad testing and a fair comparison in the fixed-parameter benchmarking strategy that we use.

The AHB21 set (and, to a lesser extent, the DIPCS10 and MB16-43 sets) forms an exception to this general observation.

set (and, to a lesser extent, the DIPCS10 and MB16-43 sets) forms an exception to this general observation.
The functional choice also has an indirect effect on the adsorption description and the competition between sites. This is because the XC functional choice sets the lattice constant, which, in turn, adjusts the electronic structure.

The functional choice also has an indirect effect on the adsorption description and the competition between sites. This is because the XC functional choice sets the lattice constant, which, in turn, adjusts the electronic structure.

FIG. 13. [Abstract-figure] Full GMTKN55 benchmark-suite performance comparison among unscreened hybrid vdW-DFs (CX0P, DF2-BR0) and RSH vdW-DFs (AHCX and AHBR).
Supplementary Materials for:
A second-generation range-separated hybrid van der Waals density functional

TABLE S I. XC functionals: abbreviations, code-nature or QE-input specifications, and literature overview. Where relevant, we also list the Fock-exchange mixing that we used in the hybrid benchmarking. All benchmark results that (in the following tables and in the main text discussion) are marked ‘(OB)’ are taken from the orbital-based-DFT assessment summarized in Ref. 5. All other benchmarks are provided here, using the stated QE ‘input_dft’ specification (and the benchmarking strategy and setup that are defined and discussed in Section III and Appendix A).

| Abbreviation     | (DFT-type) XC name                  | Code/QE input       | Mixing | Literature                  |
|------------------|------------------------------------|---------------------|--------|-----------------------------|
| PBE              | PBE                                | pbe                 | -      | Refs. 51 and 52             |
| revPBE+D3        | revPBE+D3                          | revpbe/Grimme-D3    | -      | Refs. 51 and 59             |
| SCAN+D3(OB)      | SCAN+D3 (Orbital-based)            | SCAN+D3 (From GMTKN55 paper) | - | Refs. 5, 59, and 160        |
| HSE+D3           | HSE+D3                             | hse/Grimme-D3       | 0.25   | Refs. 57, 59                |
| HSE+D3(OB)       | HSE+D3 (Orbital-based)             | HSE+D3 (From GMTKN55 paper) | 0.25 | Refs. 5, 57, 59             |
| vdW-DF1          | vdW-DF                             | vdw-df              | -      | Refs. 31 and 33             |
| vdW-DF2          | vdW-DF2                            | vdw-df2             | -      | Ref. 9 and 51               |
| rVV10            | revised VV10                       | rVV10               | -      | Refs. 129 and 130           |
| C09              | vdW-DF-C09                         | vdw-df-c09          | -      | Refs. 31 and 33             |
| OB86             | vdW-optB86                        | vdw-df-ob86         | -      | Refs. 31 and 41             |
| DF3-opt1         | vdW-DF3-opt1                       | vdw-df3-opt1        | -      | Refs. 31 and 41             |
| DF3-opt2         | vdW-DF3-opt2                       | vdw-df3-opt2        | -      | Refs. 31 and 41             |
| OBK8             | vdW-optB88                        | vdw-df-obk8         | -      | Refs. 31, 33, and 62        |
| CX               | vdW-DF-cx                          | vdw-df-cx           | -      | Refs. 31 and 33             |
| B86R             | rev-vdW-DF2                        | revvdw-df2-b86r     | -      | Refs. 9, 31, and 60         |
| CX0              | vdW-DF-cx+0                        | vdw-df-cx0          | 0.25   | Refs. 31, 35, and 39        |
| CX0P             | Zero-param. vdW-DF-cx+0            | vdw-df-cx0p         | 0.20   | Ref. 35, 39, and 40         |
| DF2-BR0          | rev-vdW-DF2+0                      | revvdw-df2-br0      | 0.25   | Refs. 9, 39, and 60 and this work. |
| AHCX             | vdW-DF-ahcx                        | vdw-df-ahcx         | 0.20   | Refs. 9, 39, and 42         |
| AHCX25           | vdW-DF-ahcx                        | vdw-df-ahcx         | 0.25   | Refs. 9, 39, and 42         |
| DF2-AH           | vdW-DF2-ah                         | vdw-df2-ah          | 0.20   | Refs. 9 and 42              |
| AHBRR20          | vdW-DF2-ahbr                       | vdw-df2-ahbr        | 0.20   | Refs. 9, 42 and 60 and this work. |
| AHBRR            | vdW-DF2-ahbr                       | vdw-df2-ahbr        | 0.25   | Refs. 9, 42 and 60 and this work. |
1. Planewave benchmarking on GMTKN55

TABLE S II. Comparison of functional performance of regular (density explicit) vdW-DFs and associated unscreened hybrid and RSH vdW-DFs, as asserted by our planewave DFT benchmarking across the full GMTKN55 suite on broad molecular properties; The benchmark groups 1-6 are defined and introduced in Ref. 5. We report computed values (in kcal/mol) for the weighted-mean-absolute-deviation measure 'WTMAD1' that is introduced and discussed in Ref. 5. For comparison, we also report our performance assessment for PBE, for the best-performing dispersion-corrected GGA (revPBE+D3), of a meta-GGA (SCAN+D3), and of the PBE-based hybrid (HSE-D3). We have used the electrostatic-environment handling to accurately characterize the performance on the G21EA, WATER21, and AHB21 benchmark sets, in spite of the presence of negative ions and radicals. Entries marked ‘OB’ are literature orbital-based DFT results that we include to illustrate that there is some but also not an excessive dependence of the code nature (and basis-set choice) in this broad molecular benchmarking.

| XC functional | Group 1 | Group 2 | Group 3 | Group 4 | Group 5 | Group 6 (4&5) | GMTKN55 |
|---------------|---------|---------|---------|---------|---------|---------------|---------|
| PBE           | 5.27    | 7.37    | 7.09    | 10.29   | 11.90   | 10.98         | 8.03    |
| revPBE+D3     | 4.91    | 5.55    | 6.30    | 3.41    | 4.27    | 3.78          | 4.76    |
| SCAN+D3(OB)   | 4.81    | 4.55    | 6.29    | 4.40    | 3.61    | 4.06          | 4.67    |
| HSE+D3        | 3.94    | 4.42    | 4.47    | 4.08    | 4.26    | 4.15          | 4.17    |
| HSE+D3(OB)    | 3.56    | 4.21    | 4.14    | 4.26    | 3.73    | 4.03          | 3.92    |
| vdw-DF1       | 6.99    | 7.83    | 3.48    | 4.31    | 6.04    | 5.05          | 5.94    |
| vdw-DF2       | 8.59    | 9.52    | 3.53    | 3.17    | 4.27    | 3.64          | 6.21    |
| rVV10         | 5.15    | 6.17    | 6.20    | 4.95    | 5.07    | 5.00          | 5.39    |
| C09 (vdW-DF-c09) | 5.20 | 5.53    | 8.44    | 4.29    | 6.10    | 5.07          | 5.62    |
| OB86 (vdW-DF-optB86r) | 4.74 | 4.65    | 7.18    | 3.81    | 4.63    | 4.16          | 4.81    |
| DF3-opt1      | 5.31    | 5.22    | 8.90    | 5.40    | 6.13    | 5.72          | 5.91    |
| DF3-opt2      | 4.87    | 4.66    | 7.67    | 4.69    | 4.60    | 4.65          | 5.11    |
| OBK8 (vdW-DF-optB88) | 5.03 | 4.77    | 6.23    | 3.48    | 4.56    | 3.94          | 4.73    |
| CX (vdW-DF-cx) | 4.95    | 4.99    | 7.56    | 3.18    | 4.04    | 3.55          | 4.75    |
| B86R (rev-vdW-DF2) | 4.95 | 4.82    | 6.99    | 3.22    | 3.77    | 3.45          | 4.62    |
| CX0 (vdW-DF-cx0,0.25) | 3.55 | 3.40    | 3.88    | 3.38    | 3.04    | 3.24          | 3.45    |
| CX0P (vdW-DF-cxp0,0.20) | 3.68 | 3.61    | 4.52    | 3.09    | 3.06    | 3.07          | 3.54    |
| DF2-BR0 (B86R+0.25) | 3.51 | 2.61    | 3.11    | 1.99    | 1.87    | 1.94          | 2.71    |
| AHCX (vdW-DF-ahcx,0.20) | 3.68 | 3.64    | 4.54    | 2.97    | 2.93    | 2.96          | 3.51    |
| AHCX 0.25 (0.25) | 3.55 | 3.43    | 3.89    | 3.25    | 2.86    | 3.09          | 3.40    |
| DF2-AH (vdW-DF2-ah,0.20) | 6.90 | 7.27    | 3.12    | 2.41    | 3.92    | 3.05          | 5.01    |
| AHBR 0.20 (0.20) | 3.67 | 3.02    | 3.82    | 2.22    | 2.17    | 2.20          | 3.02    |
| AHBR (vdW-DF2-ahbr,0.25) | 3.52 | 2.72    | 3.15    | 2.18    | 1.91    | 2.06          | 2.79    |
TABLE S III. Comparison of functional performance of regular vdW-DFs and associated unscreened hybrid and RSH vdW-DFs across the full GMTKN55 suite. Same underlying benchmark data as in Table S II but values (in kcal/mol) are here reported for a different weighted-mean-absolute-deviation measure, WTMAD2, that is also introduced and discussed in Ref. 5.

| XC functional         | Group 1 | Group 2 | Group 3 | Group 4 | Group 5 | Group 6 (4&5) | GMTKN55 |
|-----------------------|---------|---------|---------|---------|---------|--------------|---------|
| PBE                   | 6.23    | 15.72   | 16.36   | 16.61   | 20.42   | 18.48        | 13.91   |
| revPBE+D3             | 5.82    | 10.30   | 14.88   | 6.36    | 8.08    | 7.20         | 8.26    |
| SCAN+D3(0B)           | 5.31    | 7.86    | 14.94   | 8.50    | 6.61    | 7.58         | 7.86    |
| HSE+D3                | 4.73    | 8.68    | 9.10    | 7.26    | 7.98    | 7.61         | 7.07    |
| HSE+D3(0B)            | 4.14    | 8.67    | 9.47    | 7.79    | 7.07    | 7.44         | 6.80    |
| vdW-DF1               | 8.11    | 14.25   | 10.54   | 10.38   | 11.17   | 10.77        | 10.47   |
| vdW-DF2               | 10.05   | 17.24   | 11.38   | 6.71    | 7.07    | 6.88         | 10.13   |
| rVV10                 | 6.04    | 11.89   | 15.52   | 10.24   | 9.99    | 10.12        | 9.82    |
| C09 (vdW-DF-c09)      | 6.61    | 10.82   | 20.10   | 8.05    | 11.39   | 9.69         | 10.25   |
| OB86 (vdW-DF-optB86r) | 5.93    | 8.91    | 17.80   | 7.38    | 8.47    | 7.92         | 8.73    |
| DF3-opt1              | 6.98    | 9.91    | 21.35   | 10.12   | 12.43   | 11.25        | 11.00   |
| DF3-opt2              | 6.23    | 9.16    | 19.12   | 9.34    | 9.01    | 9.18         | 9.53    |
| OBK8 (vdW-DF-optB88)  | 6.03    | 8.72    | 16.14   | 6.22    | 8.26    | 7.22         | 8.24    |
| CX (vdW-DF-cx)        | 6.19    | 9.60    | 18.23   | 7.83    | 7.37    | 7.60         | 8.85    |
| B86R (rev-vdW-DF2)    | 6.08    | 9.15    | 17.57   | 6.41    | 7.18    | 6.79         | 8.34    |
| CX0 (CX+0,0.25)       | 4.44    | 5.64    | 8.74    | 7.80    | 5.57    | 6.71         | 6.09    |
| CX0P (vdW-DF-cx0p,0.20)| 4.46   | 6.24    | 10.40   | 7.38    | 5.58    | 6.50         | 6.32    |
| DF2-BR0 (B86R+0,0.25) | 4.34    | 4.49    | 7.56    | 4.78    | 3.76    | 4.28         | 4.76    |
| AHCX (vdW-DF-ahcx,0.20)| 4.47   | 6.25    | 10.47   | 7.14    | 5.38    | 6.28         | 6.24    |
| AHCX25 (0.25)         | 4.45    | 5.64    | 8.81    | 7.68    | 5.27    | 6.50         | 6.01    |
| DF2-AH (vdW-DF2-ah,0.20)| 8.55  | 13.39   | 8.02    | 4.79    | 6.52    | 5.63         | 8.11    |
| AHB20 (0.20)          | 4.39    | 5.32    | 9.46    | 5.04    | 4.26    | 4.65         | 5.30    |
| AHB25 (vdW-DF2-ahbr,0.25)| 4.35 | 4.70    | 7.70    | 5.67    | 3.83    | 4.77         | 5.00    |
TABLE S IV. Comparison of functional performance of regular vdW-DFs and associated unscreened hybrid and RSH vdW-DFs across the full GMTKN55 suite. Same underlying benchmark data as in Table S II but values (in kcal/mol) are here reported for so-called total weighted-mean-absolute-deviation measure, TMAD, that simply averages MAD values over the number of benchmarks in each group. For Group 4 we also include the TMAD estimates (marked with an asterisk *) that results when ignoring the impact of WATER27.

| XC functional       | Group 1 | Group 2 | Group 3 | Group 4 | Group 4 | Group 5 | Group 6 (4&5) |
|---------------------|---------|---------|---------|---------|---------|---------|---------------|
| PBE                 | 5.44    | 6.92    | 6.41    | 1.37    | 1.55    | 2.28    | 1.86          |
| revPBE+D3           | 5.20    | 5.74    | 5.68    | 0.58    | 0.75    | 0.75    | 0.75          |
| HSE+D3              | 4.40    | 3.58    | 3.64    | 0.69    | 1.10    | 0.72    | 0.94          |
| vDW-DF1             | 8.02    | 11.69   | 2.91    | 0.93    | 1.50    | 1.36    | 1.44          |
| vDW-DF2             | 10.28   | 14.87   | 2.91    | 0.69    | 0.77    | 1.20    | 0.96          |
| rVV10               | 5.34    | 6.37    | 5.55    | 0.74    | 1.61    | 0.90    | 1.31          |
| C09 (vdW-DF-c09)    | 5.72    | 6.17    | 7.81    | 0.66    | 1.27    | 0.94    | 1.13          |
| OB86 (vdW-DF-optB86r)| 4.99   | 5.01    | 6.60    | 0.55    | 0.97    | 0.72    | 0.86          |
| DF3-opt1            | 5.91    | 5.53    | 8.22    | 0.88    | 1.95    | 0.89    | 1.49          |
| DF3-opt2            | 5.25    | 4.96    | 7.06    | 0.69    | 1.47    | 0.75    | 1.16          |
| OBK8 (vdW-DF-optB88)| 5.28    | 5.64    | 5.65    | 0.52    | 0.92    | 0.81    | 0.87          |
| CX (vdW-DF-cx)      | 5.36    | 5.42    | 6.98    | 0.47    | 0.67    | 0.63    | 0.65          |
| B86R (rev-vdW-DF2)  | 5.24    | 5.34    | 6.41    | 0.46    | 0.84    | 0.65    | 0.76          |
| CX0 (CX+0,0.25)     | 4.18    | 3.16    | 3.17    | 0.50    | 0.71    | 0.50    | 0.62          |
| CX0P (vdW-DF-cx0p,0.20) | 4.08 | 3.40    | 3.84    | 0.45    | 0.65    | 0.49    | 0.58          |
| DF2-BR0 (B86R+0,0.25) | 4.36  | 3.05    | 2.46    | 0.34    | 0.52    | 0.31    | 0.43          |
| AHCX (vdW-DF-ahcx,0.20) | 4.09 | 3.43    | 3.85    | 0.44    | 0.62    | 0.45    | 0.55          |
| AHCX25 (0.25)       | 4.17    | 3.18    | 3.16    | 0.49    | 0.68    | 0.46    | 0.58          |
| DF2-AH (vdW-DF2-ah,0.20) | 9.31 | 11.77   | 2.59    | 0.53    | 0.56    | 1.02    | 0.75          |
| AHBR20 (0.20)       | 4.19    | 3.38    | 3.18    | 0.35    | 0.56    | 0.39    | 0.49          |
| AHBR (vdW-DF2-ahbr,0.25) | 4.33 | 3.15    | 2.48    | 0.36    | 0.54    | 0.34    | 0.45          |
TABLE S V. Functional performance of vdW-DFs and associated unscreened hybrid and RSH vdW-DFs for individual benchmark sets of the GMTKN55 group 3: Barriers heights. The nature and importance of these benchmark sets are discussed in the main text. All entries are MAD values in kcal/mol.

| XC functional     | BH76 | BHPERI | BHDIV10 | INV24 | BIROT27 | PX13 | WCPT18 |
|-------------------|------|--------|---------|-------|---------|------|--------|
| PBE               | 8.46 | 3.85   | 8.10    | 3.07  | 0.54    | 12.12| 8.71   |
| PBE (OB)          | 9.15 | 3.95   | 8.23    | 2.67  | 0.47    | 11.54| 8.61   |
| revPBE+D3         | 7.38 | 5.74   | 7.47    | 2.20  | 0.48    | 9.29 | 7.22   |
| revPBE+D3(OB)     | 8.32 | 6.29   | 7.83    | 2.18  | 0.37    | 8.75 | 7.22   |
| SCAN+D3(OB)       | 7.77 | 5.50   | 6.62    | 1.16  | 0.84    | 8.34 | 6.22   |
| HSE+D3            | 4.21 | 2.83   | 4.75    | 1.13  | 0.64    | 7.38 | 4.57   |
| HSE+D3(OB)        | 4.58 | 2.52   | 4.56    | 1.16  | 0.61    | 6.02 | 4.01   |
| vDW-DF1           | 6.33 | 2.37   | 4.35    | 2.07  | 0.44    | 2.36 | 2.45   |
| vDW-DF2           | 6.90 | 3.08   | 4.67    | 2.32  | 0.49    | 1.14 | 1.75   |
| rVV10             | 8.26 | 4.57   | 6.82    | 2.48  | 0.51    | 9.23 | 6.96   |
| C09 (vdW-DF-c09)  | 9.86 | 8.68   | 9.06    | 1.95  | 0.49    | 14.36| 10.30  |
| OB86 (vdW-DF-optB86r) | 9.11 | 6.91   | 7.69    | 1.93  | 0.45    | 11.63| 8.48   |
| DF3-opt1          | 10.60| 8.83   | 9.46    | 2.10  | 0.53    | 15.18| 10.82  |
| DF3-opt2          | 9.83 | 7.39   | 8.11    | 2.01  | 0.47    | 12.43| 9.19   |
| OBK8 (vdW-DF-optB88) | 8.60 | 5.64   | 6.56    | 1.91  | 0.46    | 9.58 | 6.79   |
| CX (vdW-DF-cx)    | 9.15 | 7.20   | 8.19    | 1.96  | 0.45    | 12.80| 9.14   |
| B86R (rev-vdW-DF2) | 9.22 | 6.08   | 7.46    | 2.06  | 0.45    | 11.36| 8.27   |
| CX0 (CX+0.0.25)   | 4.19 | 3.38   | 3.57    | 0.97  | 0.55    | 6.04 | 3.46   |
| CX0P (vdW-DF-cx0p,0.20) | 5.09 | 4.07   | 4.46    | 0.83  | 0.53    | 7.35 | 4.57   |
| DF2-BR0 (B86R+0.0.25) | 4.05 | 1.81   | 2.77    | 0.90  | 0.51    | 4.58 | 2.56   |
| AHCX (vdW-DF-ahcx,0.20) | 5.15 | 4.05   | 4.51    | 0.84  | 0.54    | 7.30 | 4.54   |
| AHCX25 (0.25)     | 4.26 | 3.33   | 3.62    | 0.95  | 0.57    | 5.98 | 3.43   |
| DF2-AH (vdW-DF2-ah,0.20) | 4.14 | 3.48   | 4.00    | 1.08  | 0.41    | 2.53 | 2.50   |
| AHBR20 (0.20)     | 5.08 | 2.65   | 3.74    | 0.84  | 0.50    | 5.88 | 3.58   |
| AHBR (vdW-DF2-ahbr,0.25) | 4.14 | 1.84   | 2.84    | 0.89  | 0.53    | 4.56 | 2.54   |
TABLE S VI. Functional performance of regular vdW-DFs and associated unscreened hybrid and RSH vdW-DFs for individual benchmark sets of the GMTKN55 group 1: small-system properties. This table reports on benchmarks containing charged systems. The calculations of negatively charged ions in the G21EA set are dramatically affected by self-interaction errors and requires the here-discussed electrostatic environment assessment procedure for a meaningful planewave assessment, see appendix and Refs. 42 and 80. All entries are MAD values in kcal/mol.

| XC functional | G21EA | G21IP | DIPCS10 | PA26 | SIE4x4 | ALK8 | RC21 | BH76RC |
|---------------|-------|-------|---------|------|--------|------|------|--------|
| PBE           | 3.07  | 4.45  | 5.12    | 1.84 | 21.61  | 2.67 | 4.49 | 3.37   |
| PBE (OB)      | 3.43  | 3.85  | 4.51    | 1.97 | 23.44  | 2.78 | 5.48 | 4.09   |
| revPBE+D3     | 2.83  | 4.72  | 4.89    | 4.41 | 21.67  | 3.70 | 4.02 | 2.87   |
| revPBE+D3(OB) | 2.75  | 4.20  | 4.81    | 4.73 | 23.43  | 3.61 | 4.85 | 2.76   |
| SCAN+D3(OB)   | 3.64  | 4.69  | 4.92    | 3.18 | 17.99  | 3.45 | 6.69 | 3.38   |
| HSE+D3        | 3.40  | 4.31  | 3.19    | 2.66 | 13.58  | 4.67 | 4.77 | 1.79   |
| HSE+D3(OB)    | 2.68  | 3.70  | 3.05    | 2.65 | 14.49  | 4.67 | 4.75 | 2.35   |
| vDW-DF1       | 5.68  | 5.19  | 9.11    | 5.06 | 22.31  | 5.93 | 3.23 | 4.27   |
| vDW-DF2       | 9.66  | 8.11  | 16.38   | 4.46 | 21.73  | 7.60 | 5.55 | 4.97   |
| rVV10         | 3.39  | 4.77  | 4.69    | 2.29 | 21.03  | 1.93 | 3.33 | 2.85   |
| C09 (vdW-DF-c09) | 2.79  | 3.96  | 5.99    | 1.64 | 23.93  | 2.13 | 7.82 | 3.45   |
| OB86 (vdW-DF-optB86r) | 3.36  | 3.82  | 4.83    | 2.11 | 23.56  | 2.29 | 5.71 | 3.29   |
| DF3-opt1      | 3.96  | 3.65  | 4.43    | 1.54 | 23.96  | 1.79 | 8.05 | 3.58   |
| DF3-opt2      | 4.72  | 3.85  | 5.42    | 1.66 | 23.66  | 2.53 | 6.10 | 3.36   |
| OBK8 (vdW-DF-optB88) | 4.69  | 4.20  | 5.42    | 2.97 | 23.00  | 3.02 | 3.92 | 3.32   |
| CX (vdW-DF-cx) | 2.80  | 3.91  | 5.93    | 1.93 | 23.80  | 2.59 | 6.53 | 3.31   |
| B86R (rev-vdW-DF2) | 4.50  | 3.82  | 5.30    | 1.83 | 23.52  | 3.38 | 4.94 | 3.29   |
| CX0 (CX+0.25) | 2.36  | 3.47  | 3.72    | 2.98 | 15.20  | 2.01 | 4.62 | 2.32   |
| CX0P (vdW-DF-cx0p0.20) | 2.10  | 3.45  | 3.68    | 2.68 | 16.89  | 2.10 | 4.90 | 2.27   |
| DF2-BR0 (B86R+0.25) | 2.54  | 3.85  | 5.98    | 2.61 | 14.94  | 3.18 | 2.99 | 2.53   |
| AHCX (vdW-DF-ahcx0.20) | 2.17  | 3.51  | 3.92    | 2.74 | 17.00  | 2.02 | 4.85 | 2.30   |
| AHCX25 (0.25) | 2.45  | 3.52  | 3.89    | 3.05 | 15.34  | 1.88 | 4.54 | 2.34   |
| DF2-AH (vdW-DF2-ah0.20) | 6.52  | 7.63  | 15.48   | 5.02 | 15.26  | 6.39 | 4.80 | 4.97   |
| AHBR20 (0.20) | 2.48  | 3.66  | 5.04    | 2.32 | 16.75  | 3.11 | 3.17 | 2.42   |
| AHBR (vdW-DF2-ahbr0.25) | 2.32  | 3.75  | 5.62    | 2.67 | 15.09  | 3.05 | 2.94 | 2.55   |
TABLE S VII. Functional performance of regular vdW-DFs and associated unscreened hybrid and RSH vdW-DFs for individual benchmark sets of the GMTKN55 group 1: small-system properties. This table focuses on all-neutral systems. We abbreviate GMTKN55 benchmark ALKBDE10 as ‘ALKB’, YBDE18 as ‘YBDE’, HEAVYSB11 as ‘HSB’, and TAUT15 as ‘TAUT’. All entries are MAD values in kcal/mol.

| XC functional | W4-11 | ALKB | YBDE | AL2X6 | HSB | NBPRC | G2RC | FH51 | TAUT | DC13 |
|---------------|-------|------|------|-------|-----|-------|------|------|------|------|
| PBE           | 7.44  | 4.98 | 6.39 | 4.05  | 3.71| 2.88  | 5.85 | 3.34 | 2.05 | 10.67|
| PBE (OB)      | 14.96 | 6.21 | 5.91 | 4.26  | 4.58| 2.82  | 5.60 | 3.26 | 1.94 | 9.38 |
| revPBE+D3     | 5.88  | 5.15 | 5.91 | 1.91  | 3.40| 2.03  | 5.60 | 3.26 | 1.94 | 9.38 |
| revPBE+D3(OB)| 7.57  | 5.16 | 4.41 | 2.07  | 2.72| 1.98  | 6.16 | 3.34 | 1.55 | 8.87 |
| SCAN+D3(OB)   | 4.08  | 19.27| 3.12 | 2.13  | 6.64| 2.51  | 6.39 | 2.75 | 1.74 | 7.29 |
| HSE+D3        | 6.77  | 5.78 | 3.35 | 1.22  | 2.36| 2.68  | 6.48 | 2.63 | 1.36 | 8.24 |
| HSE+D3(OB)    | 3.56  | 5.57 | 1.41 | 0.93  | 1.14| 2.57  | 6.21 | 2.47 | 1.19 | 7.09 |
| vDW-DF1       | 12.98 | 4.15 | 13.28| 6.61  | 8.36| 6.84  | 7.10 | 5.57 | 1.60 | 17.06|
| vDW-DF2       | 18.69 | 4.02 | 16.04| 7.65  | 9.44| 8.74  | 9.43 | 6.63 | 1.78 | 24.21|
| rVV10         | 6.93  | 4.57 | 8.34 | 2.04  | 4.25| 2.96  | 4.14 | 3.59 | 1.94 | 13.13|
| C09 (vdW-DF-c09)| 11.14| 7.03 | 4.53 | 1.59  | 2.76| 2.26  | 7.09 | 3.50 | 1.93 | 9.41 |
| OB86 (vdW-DF-optB86r)| 6.57| 5.72 | 5.44 | 1.00  | 2.53| 1.57  | 6.16 | 3.32 | 1.80 | 6.87 |
| DF3-opt1      | 14.36 | 8.38 | 4.76 | 1.58  | 3.23| 2.22  | 6.93 | 3.50 | 2.03 | 8.51 |
| DF3-opt2      | 8.86  | 6.53 | 5.31 | 0.99  | 2.55| 1.28  | 5.91 | 3.29 | 1.86 | 6.65 |
| OBK8 (vdW-DF-optB88)| 4.40| 5.08 | 7.56 | 1.65  | 3.57| 2.36  | 5.41 | 3.74 | 1.71 | 9.10 |
| CX (vdW-DF-cx)| 8.55  | 6.29 | 5.15 | 1.47  | 2.80| 1.61  | 6.77 | 3.35 | 1.83 | 7.88 |
| B86R (rev-vdw-DF2)| 6.97| 5.88 | 5.96 | 2.35  | 2.81| 1.82  | 5.58 | 3.12 | 1.81 | 7.26 |
| CX0 (CX+0.25) | 7.56  | 5.60 | 3.90 | 0.76  | 2.66| 1.27  | 4.53 | 2.48 | 1.07 | 8.79 |
| CX0P (vdW-DF-cx0p,0.20)| 5.01| 5.27 | 3.80 | 0.81  | 2.27| 1.24  | 4.76 | 2.49 | 1.21 | 8.49 |
| DF2-BR0 (B86R+0,0.25)| 9.49| 5.44 | 5.71 | 1.61  | 4.05| 1.15  | 3.30 | 1.96 | 1.04 | 6.18 |
| AHCX (vdW-DF-ahcx,0.20)| 4.99| 5.16 | 3.92 | 0.76  | 2.28| 1.16  | 4.71 | 2.47 | 1.23 | 8.35 |
| AHCX25 (0.25)| 7.56  | 5.41 | 4.04 | 0.70  | 2.66| 1.15  | 4.44 | 2.44 | 1.09 | 8.63 |
| DF2-AH (vdW-DF2-ah,0.20)| 26.00| 7.92 | 14.97| 5.83  | 9.71| 7.17  | 5.94 | 4.40 | 1.39 | 18.17|
| AHB-R20 (0.20)| 6.55  | 4.76 | 5.52 | 1.86  | 3.50| 1.26  | 3.52 | 2.16 | 1.20 | 6.07 |
| AHB-R (vdW-DF2-ahbr,0.25)| 9.50| 5.29 | 5.81 | 1.73  | 4.02| 1.23  | 3.30 | 1.98 | 1.06 | 6.18 |
TABLE S VIII. Functional performance of regular vdW-DFs and associated unscreened hybrid and RSH vdW-DFs for individual benchmark sets of the GMTKN55 group 2: large-system isomerizations. All entries are MAD values in kcal/mol.

| XC functional       | MB16-43 | DARC | RSE43 | BSR36 | CDIE20 | ISO34 | ISOL24 | C60ISO | PArel |
|---------------------|---------|------|-------|-------|--------|-------|--------|--------|-------|
| PBE                 | 22.60   | 7.13 | 2.54  | 7.65  | 1.78   | 1.73  | 6.88   | 10.06  | 1.93  |
| PBE (OB)            | 22.78   | 6.94 | 3.10  | 7.67  | 1.81   | 1.80  | 6.83   | 11.06  | 1.81  |
| revPBE+D3           | 25.28   | 4.28 | 1.93  | 1.70  | 1.54   | 1.49  | 4.82   | 8.93   | 1.69  |
| revPBE+D3(OB)       | 27.71   | 3.71 | 2.31  | 1.80  | 1.50   | 1.50  | 4.56   | 9.82   | 1.53  |
| SCAN+D3(OB)         | 17.77   | 2.01 | 1.29  | 1.28  | 1.45   | 1.30  | 3.23   | 6.01   | 1.50  |
| HSE+D3              | 15.48   | 2.65 | 1.25  | 3.83  | 1.30   | 1.42  | 2.42   | 2.51   | 1.34  |
| HSE+D3(OB)          | 14.27   | 2.11 | 1.49  | 3.83  | 1.32   | 1.34  | 2.64   | 2.43   | 1.16  |
| vdw-DF1             | 57.84   | 15.99| 1.22  | 4.10  | 1.24   | 2.85  | 9.49   | 10.73  | 1.76  |
| vdw-DF2             | 75.22   | 22.24| 1.13  | 4.94  | 1.05   | 3.97  | 12.69  | 10.43  | 2.17  |
| rVV10               | 21.11   | 8.75 | 1.77  | 3.46  | 1.45   | 1.94  | 6.81   | 10.55  | 1.51  |
| C09 (vdW-DF-c09)    | 27.04   | 3.90 | 2.31  | 2.44  | 1.29   | 1.59  | 2.86   | 12.17  | 1.99  |
| OB86 (vdW-DF-optB86r)| 21.38  | 1.66 | 2.05  | 1.05  | 1.26   | 1.32  | 2.88   | 11.84  | 1.70  |
| DF3-opt1            | 25.00   | 2.53 | 2.54  | 0.53  | 1.30   | 1.46  | 2.81   | 11.59  | 2.04  |
| DF3-opt2            | 20.28   | 1.66 | 2.17  | 1.36  | 1.17   | 1.27  | 2.90   | 12.09  | 1.77  |
| OBK8 (vdW-DF-optB88)| 23.46   | 5.18 | 1.73  | 0.27  | 1.17   | 1.68  | 4.61   | 11.34  | 1.42  |
| CX (vdW-DF-cx)      | 24.27   | 1.71 | 2.21  | 1.27  | 1.34   | 1.49  | 2.65   | 12.01  | 1.86  |
| B86R (rev-vdW-DF2)  | 22.37   | 2.97 | 2.14  | 0.64  | 1.24   | 1.35  | 3.68   | 11.95  | 1.68  |
| CX0 (CX+0.25)       | 14.09   | 5.14 | 0.73  | 0.80  | 0.80   | 1.30  | 2.32   | 2.16   | 1.07  |
| CX0P (vdW-DF-cx0p,0.20) | 15.31 | 4.34 | 0.98  | 0.88  | 0.90   | 1.33  | 2.28   | 3.40   | 1.18  |
| DF2-BR0 (B86R+0.25) | 18.88   | 1.16 | 0.72  | 0.30  | 0.74   | 1.02  | 1.59   | 2.15   | 0.92  |
| AHCX (vdW-DF-ahcx,0.20) | 15.41 | 4.00 | 1.01  | 0.74  | 0.96   | 1.32  | 2.25   | 3.99   | 1.19  |
| AHCX25 (0.25)       | 14.15   | 4.71 | 0.75  | 0.62  | 0.87   | 1.30  | 2.28   | 2.83   | 1.08  |
| DF2-AH (vdW-DF2-ah,0.20) | 67.26 | 15.70| 0.82  | 4.38  | 0.76   | 3.10  | 9.34   | 2.80   | 1.74  |
| AHBR20 (0.20)       | 18.97   | 1.27 | 0.98  | 0.29  | 0.88   | 1.05  | 2.10   | 3.88   | 1.04  |
| AHBR (vdW-DF2-ahbr,0.25) | 18.95 | 1.06 | 0.74  | 0.35  | 0.80   | 1.02  | 1.82   | 2.72   | 0.94  |
TABLE S IX. Functional performance of regular vdW-DFs and associated unscreened hybrid and RSH vdW-DFs for individual benchmark sets of the GMTKN55 group 4: intermolecular noncovalent (NOC) interactions. We report MAD values in kcal/mol, abbreviating benchmark name HEAVY28 as ‘HEAVY’, PNICO23 as ‘PNICO’, CARBHB12 as ‘CARBH’, and WATER27 as ‘WATER’.

| XC functional | RG18 | ADIM6 | S22 | S66 | HEAVY | CARBH | PNICO | HAL59 | CHB6 | IL16 | AHB21 | WATER |
|---------------|------|-------|-----|-----|-------|-------|-------|-------|------|------|-------|-------|
| PBE           | 0.28 | 3.36  | 2.56| 2.13| 0.50  | 1.05  | 0.85  | 1.41  | 0.75 | 1.34 | 0.85  | 3.47  |
| PBE (OB)      | 0.28 | 3.38  | 2.55| 2.11| 0.47  | 1.09  | 0.82  | 1.30  | 0.79 | 1.49 | 0.83  | 2.83  |
| revPBE+D3     | 0.08 | 0.12  | 0.36| 0.25| 0.30  | 1.10  | 0.84  | 0.82  | 0.88 | 0.60 | 1.01  | 2.63  |
| revPBE+D3(OB) | 0.09 | 0.25  | 0.43| 0.28| 0.29  | 1.10  | 0.88  | 0.72  | 0.90 | 0.77 | 1.04  | 3.51  |
| SCAN+D3(OB)   | 0.18 | 0.12  | 0.47| 0.43| 0.27  | 1.38  | 1.08  | 1.03  | 0.45 | 0.98 | 1.67  | 10.15 |
| HSE+D3        | 0.11 | 0.25  | 0.54| 0.39| 0.38  | 1.41  | 0.86  | 0.64  | 1.24 | 0.38 | 1.32  | 5.73  |
| HSE+D3(OB)    | 0.13 | 0.14  | 0.52| 0.39| 0.44  | 1.47  | 0.97  | 0.73  | 1.29 | 0.31 | 1.32  | 6.29  |
| vdw-DF1       | 0.59 | 0.23  | 1.26| 0.69| 0.31  | 0.44  | 0.76  | 0.50  | 1.00 | 2.61 | 1.79  | 7.80  |
| vdw-DF2       | 0.35 | 0.51  | 0.69| 0.32| 0.14  | 0.46  | 0.39  | 0.69  | 1.13 | 1.82 | 1.05  | 1.75  |
| rVV10         | 0.16 | 0.28  | 0.42| 0.43| 0.40  | 1.59  | 0.91  | 1.47  | 0.36 | 0.82 | 1.28  | 11.23 |
| C09 (vdW-DF-c09) | 0.12 | 0.51  | 0.43| 0.38| 0.18  | 1.14  | 0.95  | 1.22  | 0.27 | 1.22 | 0.86  | 7.95  |
| OB86 (vdW-DF-optB86r) | 0.22 | 0.75  | 0.30| 0.35| 0.14  | 0.92  | 0.68  | 1.05  | 0.33 | 0.62 | 0.67  | 5.66  |
| DF3-opt1      | 0.11 | 0.42  | 0.49| 0.48| 0.22  | 1.62  | 1.31  | 1.50  | 0.13 | 1.82 | 1.58  | 13.66 |
| DF3-opt2      | 0.23 | 0.59  | 0.32| 0.41| 0.21  | 1.32  | 0.99  | 1.32  | 0.27 | 0.94 | 1.04  | 10.04 |
| OBK8 (vdW-DF-optB88) | 0.08 | 0.82  | 0.30| 0.36| 0.13  | 0.79  | 0.52  | 1.00  | 0.55 | 0.50 | 0.62  | 5.37  |
| CX (vdW-DF-cx) | 0.37 | 0.05  | 0.36| 0.28| 0.24  | 0.76  | 0.66  | 0.94  | 0.50 | 0.37 | 0.65  | 2.88  |
| B86R (rev-vdW-DF2) | 0.07 | 0.21  | 0.45| 0.36| 0.22  | 0.85  | 0.56  | 1.00  | 0.30 | 0.37 | 0.64  | 5.10  |
| CX0 (CX+0.025) | 0.47 | 0.78  | 0.41| 0.42| 0.26  | 0.58  | 0.41  | 0.58  | 0.71 | 0.32 | 0.60  | 2.94  |
| CX0P (vdW-DF-cx0p,0.20) | 0.45 | 0.61  | 0.30| 0.30| 0.25  | 0.59  | 0.44  | 0.62  | 0.56 | 0.32 | 0.50  | 2.82  |
| DF2-BR0 (B86R+0.025) | 0.06 | 0.04  | 0.27| 0.25| 0.35  | 0.48  | 0.24  | 0.57  | 0.52 | 0.49 | 0.46  | 2.50  |
| AHICX (vdW-DF-ahcx,0.20) | 0.43 | 0.56  | 0.26| 0.27| 0.25  | 0.60  | 0.44  | 0.63  | 0.53 | 0.33 | 0.50  | 2.71  |
| AHICX25 (0.25) | 0.47 | 0.71  | 0.36| 0.37| 0.26  | 0.58  | 0.40  | 0.59  | 0.67 | 0.32 | 0.61  | 2.79  |
| DF2-AH (vdW-DF2-ah,0.20) | 0.26 | 0.53  | 0.43| 0.19| 0.14  | 0.37  | 0.43  | 0.39  | 0.80 | 1.76 | 0.61  | 0.77  |
| AHB20 (0.20)  | 0.07 | 0.09  | 0.31| 0.28| 0.32  | 0.56  | 0.29  | 0.63  | 0.41 | 0.44 | 0.42  | 2.95  |
| AHB2 (0.25)   | 0.20 | 0.06  | 0.29| 0.26| 0.35  | 0.50  | 0.25  | 0.58  | 0.51 | 0.49 | 0.48  | 2.52  |
TABLE S X. Functional performance of regular vdW-DFs and associated unscreened hybrid and RSH vdW-DFs for individual benchmark sets of the GMTKN55 group 5: intramolecular NOC interactions. We report MAD values in kcal/mol, abbreviating benchmark name Amino20x4 as ‘Amino’, PCON21 as ‘PCONF’, and BUT14DIOL as ‘B14D’.

| XC functional | IDISP | ICONF | ACONF | Amino | PCONF | MCONF | SCONF | UPU23 | B14D |
|---------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| PBE           | 10.91 | 0.43  | 0.61  | 0.53  | 3.59  | 1.83  | 0.37  | 2.02  | 0.25  |
| PBE (OB)      | 10.78 | 0.43  | 0.61  | 0.51  | 3.48  | 1.80  | 0.35  | 1.99  | 0.26  |
| revPBE+D3     | 3.25  | 0.31  | 0.05  | 0.35  | 1.01  | 0.44  | 0.51  | 0.60  | 0.26  |
| revPBE+D3(OB) | 3.14  | 0.32  | 0.09  | 0.37  | 0.87  | 0.44  | 0.51  | 0.47  | 0.31  |
| SCAN+D3(OB)   | 2.15  | 0.31  | 0.13  | 0.22  | 0.47  | 0.46  | 0.66  | 0.39  | 0.40  |
| HSE+D3        | 2.96  | 0.29  | 0.19  | 0.29  | 1.34  | 0.31  | 0.21  | 0.70  | 0.20  |
| HSE+D3(OB)    | 2.55  | 0.30  | 0.14  | 0.27  | 1.05  | 0.27  | 0.26  | 0.55  | 0.26  |
| vdw-DF1       | 7.56  | 0.51  | 0.38  | 0.53  | 0.60  | 0.58  | 1.04  | 0.55  | 0.50  |
| vdw-DF2       | 7.89  | 0.52  | 0.12  | 0.38  | 0.40  | 0.48  | 0.52  | 0.53  | 0.13  |
| rVV10         | 3.91  | 0.33  | 0.12  | 0.33  | 0.73  | 0.41  | 1.08  | 0.43  | 0.75  |
| C09 (vdW-DF-c09) | 3.31 | 0.24  | 0.14  | 0.35  | 0.94  | 0.74  | 1.39  | 0.64  | 0.72  |
| OB86 (vdW-DF-optB86r) | 2.57 | 0.22  | 0.07  | 0.25  | 0.77  | 0.57  | 0.87  | 0.66  | 0.51  |
| DF3-opt1      | 2.79  | 0.26  | 0.19  | 0.42  | 0.90  | 0.57  | 1.58  | 0.37  | 0.94  |
| DF3-opt2      | 2.89  | 0.20  | 0.11  | 0.27  | 0.69  | 0.43  | 1.06  | 0.41  | 0.68  |
| OBK8 (vdW-DF-optB88) | 3.54 | 0.24  | 0.11  | 0.23  | 0.75  | 0.53  | 0.75  | 0.64  | 0.50  |
| CX (vdW-DF-cx) | 2.27 | 0.26  | 0.11  | 0.25  | 0.75  | 0.39  | 0.81  | 0.47  | 0.37  |
| B86R (rev-vdW-DF2) | 2.69 | 0.24  | 0.05  | 0.22  | 0.68  | 0.33  | 0.76  | 0.37  | 0.48  |
| CX0 (CX+0.025) | 1.94 | 0.21  | 0.05  | 0.23  | 0.40  | 0.55  | 0.26  | 0.65  | 0.20  |
| CX0P (vdW-DF-cx0p,0.20) | 1.79 | 0.20  | 0.06  | 0.22  | 0.42  | 0.50  | 0.35  | 0.60  | 0.23  |
| DF2-BR0 (B86R+0.0,25) | 1.26 | 0.22  | 0.09  | 0.18  | 0.22  | 0.17  | 0.13  | 0.36  | 0.20  |
| AHCX (vdW-DF-ahcx,0.20) | 1.61 | 0.21  | 0.08  | 0.22  | 0.40  | 0.46  | 0.32  | 0.59  | 0.22  |
| AHCX25 (0.25) | 1.72 | 0.21  | 0.07  | 0.22  | 0.36  | 0.50  | 0.22  | 0.63  | 0.19  |
| DF2-AH (vdW-DF2-ah,0.20) | 6.28 | 0.41  | 0.14  | 0.37  | 0.33  | 0.32  | 0.68  | 0.51  | 0.14  |
| AHBR20 (0.20) | 1.75 | 0.21  | 0.10  | 0.18  | 0.32  | 0.18  | 0.22  | 0.34  | 0.24  |
| AHBR (vdW-DF2-ahbr,0.25) | 1.50 | 0.23  | 0.11  | 0.19  | 0.24  | 0.16  | 0.11  | 0.36  | 0.19  |
2. Parameters for the analytical-exchange hole modeling and convergence tests

TABLE S XI. Parameters in the rational function defining $H(s)$ in the HJS AH model\textsuperscript{42,69} in its description for the exchange functionals PBEx, PBEsolx, rPW86 (exchange in vDW-DF2), cx13 (exchange in vDW-DF-cx), and B86R (exchange in vDW-DF2-b86r). The determination of the PBEx, PBEsolx, rPW86, and LV-rPW86 (or cx-13) parameters are given in Ref. 42 and repeated for the reader’s convenience; The last column gives parameters that we provide for the AH modeling of B86R and hence for the AHBR specification.

|         | PBEx      | PBEsolx   | rPW86     | LV-rPW86  | B86R      |
|---------|-----------|-----------|-----------|-----------|-----------|
| $a_2$   | 0.0154999 | 0.0045881 | 0.0000006 | 0.0024387 | 0.0045620 |
| $a_3$   | -0.0361006| -0.0085784| 0.0402647 | -0.0041526| -0.0087000|
| $a_4$   | 0.0379567 | 0.0072956 | -0.0353219| 0.0025826 | 0.0073696 |
| $a_5$   | -0.0186715| -0.0032019| 0.0402647 | -0.0041526| -0.0087000|
| $a_6$   | 0.0017426 | 0.006049  | -0.0353219| 0.0025826 | 0.0073696 |
| $a_7$   | -2.7062566| -1.449453 | -1.8779594| -2.2030319| -2.2089330|
| $b_1$   | 3.3316842 | 2.001104  | 1.5198811 | 2.1759315 | 2.1968353 |
| $b_2$   | -2.3871819| -1.1935421| -0.5383109| -1.2997841| -1.2662249|
| $b_3$   | 1.1197810 | 0.4476392 | 0.1352399 | 0.5347267 | 0.4689964 |
| $b_4$   | -0.3606638| -0.1172367| -0.0428465| -0.1588798| -0.1165714|
| $b_5$   | 0.0841990 | 0.0231625 | 0.0117903 | 0.0367329 | 0.0207188 |
| $b_6$   | -0.0114719| -0.0032578| 0.0033791 | -0.0073718| -0.0029772|
| $b_7$   | 0.0016928 | 0.0005399 | -0.0000493| 0.0012667 | 0.0005982 |
| $b_8$   | 0.0015054 | 0.0000158 | 0.0000071 | 0.0000008 | 0.0000047 |

TABLE S XII. Scope of impact by spurious intercellvdW-attraction on the GMTKN55 benchmarks sets on the MAD values (in kcal/mol) that characterize CX performance on individual-GMTKN55 benchmarks sets, given our choice of a 10 Å vacuum padding. The characterization rests on self-consistent electron-density variations (computed for the roughly 2450 different GMTKN55 system\textsuperscript{5} in a pilot study using the AbInit PPs\textsuperscript{109} at 80 Ry wavefunction-energy cutoff) and the extraction of effective per-unit-cell $C_{mol}^6$ asymptotic-vdW interaction coefficients,\textsuperscript{31,112,113} see Appendix A. We list assessments of impact (in kcal/mol) for the benchmarks that we find are most susceptible to offsets from such spurious vdW coupling in our planewave setup.

| Benchmark | Spurious-vdW impact |
|-----------|----------------------|
| G21EA     | 0.010                |
| ALK8      | 0.002                |
| ALKBDE10  | 0.002                |
| WATER27   | 0.001                |
| MB16-43   | 0.001                |
3. Bulk-system performance

**TABLE S XIII.** Comparison of computed bulk lattice constants $a$ and experiment values, back-corrected to zero point energy and thermal effects. All entries in Å.

|      | CX      | AHCX    | AHCX$_{0.25}$ | DF2-AH  | b86R    | AHB1$_{0.20}$ | AHB2   | Exper.* |
|------|---------|---------|---------------|---------|---------|---------------|---------|---------|
| Ag   | 4.065   | 4.078   | 4.082         | 4.287   | 4.104   | 4.115         | 4.118   | 4.070   |
| Au   | 4.101   | 4.098   | 4.097         | 4.305   | 4.134   | 4.127         | 4.126   | 4.067   |
| Al   | 4.041   | 4.040   | 4.039         | 4.044   | 4.030   | 4.032         | 4.033   | 4.022   |
| C    | 3.561   | 3.545   | 3.541         | 3.573   | 3.565   | 3.548         | 3.544   | 3.553   |
| Cu   | 3.576   | 3.587   | 3.592         | 3.736   | 3.602   | 3.613         | 3.617   | 3.599   |
| GaAs | 5.705   | 5.640   | 5.628         | 5.758   | 5.733   | 5.661         | 5.644   | 5.638   |
| LiF  | 4.052   | 4.012   | 4.002         | 4.026   | 4.036   | 4.004         | 3.996   | 3.972   |
| MgO  | 4.243   | 4.205   | 4.197         | 4.222   | 4.225   | 4.194         | 4.187   | 4.189   |
| NaCl | 5.661   | 5.623   | 5.612         | 5.635   | 5.626   | 5.603         | 5.597   | 5.569   |
| SiC  | 4.374   | 4.353   | 4.348         | 4.377   | 4.377   | 4.356         | 4.351   | 4.346   |
| Pt   | 3.929   | 3.910   | 3.906         | 4.057   | 3.952   | 3.929         | 3.925   | 3.917   |
| Rh   | 3.786   | 3.760   | 3.754         | 3.875   | 3.806   | 3.776         | 3.770   | 3.786   |
| Si   | 5.462   | 5.441   | 5.435         | 5.476   | 5.465   | 5.444         | 5.439   | 5.411   |

**TABLE S XIV.** Comparison of computed bulk cohesive energies $E_{coh}$ and experiment values, back-corrected to zero point energy and thermal effects. All entries in eV.

|      | CX      | AHCX    | AHCX$_{0.25}$ | DF2-AH  | b86R    | AHB1$_{0.20}$ | AHB2   | Exper.* |
|------|---------|---------|---------------|---------|---------|---------------|---------|---------|
| Ag   | 2.955   | 2.774   | 2.737         | 2.100   | 2.779   | 2.592         | 2.549   | 2.964   |
| Au   | 3.634   | 3.440   | 3.398         | 2.469   | 3.402   | 3.205         | 3.158   | 3.835   |
| Al   | 3.642   | 3.430   | 3.421         | 2.516   | 3.439   | 3.251         | 3.200   | 3.431   |
| C    | 7.891   | 7.565   | 7.553         | 6.570   | 7.777   | 7.414         | 7.424   | 7.452   |
| Cu   | 3.781   | 3.348   | 3.264         | 2.551   | 3.582   | 3.160         | 3.064   | 3.513   |
| GaAs | 3.358   | 3.317   | 3.321         | 2.739   | 3.242   | 3.190         | 3.179   | 3.337   |
| LiF  | 4.405   | 4.399   | 4.374         | 4.501   | 4.553   | 4.418         | 4.382   | 4.457   |
| MgO  | 5.110   | 5.057   | 5.105         | 4.889   | 5.247   | 5.065         | 5.020   | 5.203   |
| NaCl | 3.225   | 3.258   | 3.245         | 3.182   | 3.230   | 3.191         | 3.179   | 3.337   |
| SiC  | 6.590   | 6.406   | 6.393         | 5.698   | 6.514   | 6.301         | 6.296   | 6.478   |
| Pt   | 6.226   | 5.524   | 5.259         | 3.941   | 5.999   | 5.131         | 4.930   | 5.886   |
| Rh   | 6.367   | 5.244   | 4.972         | 3.956   | 6.389   | 5.164         | 4.856   | 5.783   |
| Si   | 4.758   | 4.664   | 4.624         | 4.155   | 4.679   | 4.563         | 4.531   | 4.685   |
TABLE S XV. Comparison of computed bulk-modulus constants $B_0$ and experiment values, back-corrected to zero point energy and thermal effects. All entries in GPa.

| Element | CX | AHCX | AHCX$_{0.25}$ | DF2-AH | b86R | AHB$_{0.20}$ | AHB | Exper.* |
|---------|----|------|---------------|--------|------|--------------|-----|---------|
| Ag      | 115.3 | 104.8 | 104.0 | 63.9 | 102.4 | 95.2 | 94.6 | 105.7 |
| Au      | 170.5 | 167.8 | 166.8 | 94.2 | 153.4 | 152.0 | 151.3 | 182.0 |
| Al      | 78.2 | 82.4 | 82.5 | 74.6 | 78.7 | 81.9 | 81.8 | 72.2 |
| C       | 439.8 | 466.2 | 472.6 | 434.8 | 434.1 | 461.9 | 468.5 | 454.7 |
| Cu      | 163.3 | 148.3 | 146.0 | 91.1 | 151.3 | 141.4 | 136.0 | 144.3 |
| GaAs    | 64.8 | 73.8 | 76.7 | 59.7 | 61.2 | 71.6 | 74.2 | 74.2 |
| LiF     | 68.3 | 74.5 | 75.9 | 75.7 | 69.7 | 74.9 | 76.0 | 76.3 |
| MgO     | 153.3 | 168.2 | 171.8 | 167.4 | 158.3 | 171.7 | 174.9 | 169.8 |
| NaCl    | 24.9 | 26.1 | 26.5 | 27.3 | 26.2 | 26.9 | 27.0 | 27.6 |
| SiC     | 215.0 | 228.5 | 231.8 | 214.5 | 212.6 | 226.7 | 230.1 | 229.1 |
| Pt      | 284.0 | 297.7 | 298.1 | 187.9 | 264.0 | 277.9 | 279.3 | 285.5 |
| Rh      | 295.8 | 312.7 | 317.3 | 225.2 | 275.7 | 298.0 | 303.3 | 277.1 |
| Si      | 90.1 | 96.5 | 98.3 | 90.7 | 89.1 | 95.9 | 97.5 | 101.3 |

4. Testing on DNA base-pair assembly

TABLE S XVI. Comparison of vdW-DF tool-chain performance for DNA assembly: Stepping energies defined by stacking two Watson-Crick (WC) base pairs, as discussed in the main text. A superscript ‘GBRV’ identifies more electron-sparse calculations (performed with ultrasoft PPs$^{108}$ at 50 Ry wavefunction-energy cutoff) than supplement our standard ONCV-SG15 PP/160 Ry characterizations. The coupled-cluster reference energies (as well as the atomic configurations) for 10 stacked base-pair combinations are taken from Ref.117. They are computed in DLPNO-CCSD(T), Ref.133. All entries are in kcal/mol.

|  | B3LYP+D3(BJ)* | CX$^{GBRV}$ | CX | AHCX | AHCX$_{0.25}$ | B86R | AHB$_{0.20}$ | AHB | DLPNO-CCSD(T) |
|---|----------------|------------|----|------|---------------|------|--------------|-----|---------------|
| ApA | -13.72 | -14.04 | -14.33 | -15.93 | -16.33 | -12.34 | -12.64 | -12.73 | -12.95 |
| ApC | -12.86 | -13.16 | -13.45 | -14.95 | -15.34 | -11.49 | -11.72 | -11.79 | -11.88 |
| ApG | -13.47 | -13.75 | -14.03 | -15.60 | -16.01 | -12.19 | -12.44 | -12.52 | -12.53 |
| ApT | -12.05 | -12.30 | -12.64 | -14.10 | -14.49 | -10.63 | -10.89 | -10.98 | -11.00 |
| CpC | -11.38 | -11.86 | -12.14 | -13.63 | -14.01 | -10.24 | -10.43 | -10.50 | -10.53 |
| GpG | -17.02 | -17.12 | -17.42 | -19.31 | -19.78 | -15.92 | -16.36 | -16.49 | -16.23 |
| GpC | -14.97 | -15.13 | -15.48 | -17.06 | -17.47 | -13.54 | -13.81 | -13.89 | -13.94 |
| TpA | -13.70 | -13.94 | -14.25 | -15.86 | -16.27 | -12.46 | -12.82 | -12.93 | -12.92 |
| TpC | -12.51 | -13.05 | -13.36 | -14.79 | -15.16 | -11.32 | -11.52 | -11.59 | -11.58 |
| TpG | -14.99 | -15.10 | -15.41 | -17.15 | -17.59 | -13.76 | -14.17 | -14.30 | -14.17 |
| MD | - | -1.17 | -1.48 | -3.06 | -3.47 | 0.38 | 0.09 | 0.00 | - |
| MAD | 0.89 | 1.17 | 1.48 | 3.06 | 3.47 | 0.38 | 0.12 | 0.08 | - |

*Ref.117
TABLE S XVII. Comparison of vdW-DF tool-chain performance for descriptions of DNA assembly by base-pair stacking, here focusing on molecular-pairing contributions, $\Delta E_{B-pair}$, for which there are also couple-cluster reference results. All entries are in kcal/mol.

|       | CX $^{GBRV}$ | CX | AHCX | AHCX$_{0.25}$ | B86R | AHBRO$_{0.20}$ | AHBRO | DLPNO-CCSD(T)$^a$ |
|-------|-------------|----|------|---------------|------|----------------|--------|------------------|
| ApA   | -14.75      | -15.02 | -16.51 | -16.89        | -12.29 | -12.50         | -12.57 | -13.16           |
| ApC   | -14.77      | -15.04 | -16.46 | -16.83        | -12.39 | -12.53         | -12.59 | -13.24           |
| ApG   | -15.37      | -15.63 | -17.14 | -17.52        | -13.02 | -13.19         | -13.26 | -13.67           |
| ApT   | -13.07      | -13.29 | -14.76 | -15.13        | -10.76 | -10.96         | -11.02 | -11.62           |
| CpC   | -15.41      | -15.67 | -17.06 | -17.41        | -12.92 | -13.00         | -13.04 | -13.50           |
| CpG   | -19.09      | -19.38 | -21.32 | -21.80        | -17.17 | -17.60         | -17.73 | -18.11           |
| GpC   | -17.34      | -17.64 | -19.20 | -19.59        | -14.99 | -15.17         | -15.24 | -15.77           |
| TpA   | -14.76      | -14.98 | -16.59 | -16.98        | -12.60 | -12.88         | -12.97 | -13.58           |
| TpC   | -14.74      | -14.99 | -16.35 | -16.70        | -12.18 | -12.30         | -12.34 | -12.90           |
| TpG   | -16.90      | -17.16 | -18.89 | -19.32        | -14.82 | -15.16         | -15.27 | -15.73           |
| MD    | -1.49       | -1.75  | -3.30  | -3.69         | 0.82   | 0.60           | 0.52   | -                |
| MAD   | 1.49        | 1.75   | 3.30   | 3.69          | 0.82   | 0.60           | 0.52   | -                |

$^a$Ref. 117.