On the Use of the rVV10 Nonlocal Correlation Functional in the B97M-V Density Functional: Defining B97M-rV and Related Functionals

Narbe Mardirossian,[1, a)] Luis Ruiz Pestana,[1, b)] James C. Womack,[2, c)] Chris-Kriton Skylaris,[2, d)] Teresa Head-Gordon,[1, 3, e)] and Martin Head-Gordon[1, 3, f)]

[1)] *Kenneth S. Pitzer Center for Theoretical Chemistry, Department of Chemistry, University of California, Berkeley, California 94720, USA*

[2)] *School of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, UK*

[3)] *Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

Abstract

The VV10 and rVV10 nonlocal correlation functionals are consistently implemented and assessed, with the goal of determining if the rVV10 nonlocal correlation functional can replace the VV10 nonlocal correlation functional in the recently-developed B97M-V density functional, to give the B97M-rV density functional. Along the way, four density functionals are simultaneously tested: VV10, rVV10, B97M-V, and B97M-rV. An initial assessment is carried out across the S22 dataset, and the short-range damping variable, b, is varied for all four density functionals in order to determine the sensitivity of the functionals to the empirical parameter. The results of this test indicate that a value of $b = 6$ (fortuitously the same as in B97M-V) is suitable for B97M-rV. The functionals are then compared across an extensive database of interaction energies, and it is demonstrated that B97M-rV either matches or outperforms B97M-V for all of the tests considered. Finally, the optimization of b across the S22 dataset is extended to two range-separated hybrid density functionals, ωB97X-V and ωB97M-V, and a value of $b = 6.2$ is recommended for both ωB97X-rV and ωB97M-rV.
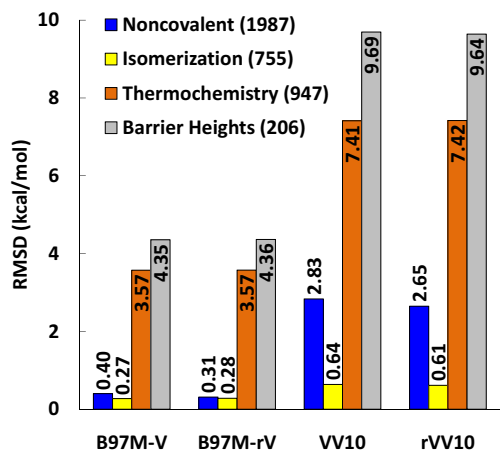
In the past decade, the focus of density functional development and benchmarking has expanded from thermochemical properties such as atomization energies to a wide variety of diverse energetics such as isomerization energies, barrier heights, excited states, and noncovalent interactions. The latter has certainly received the most attention[1], and the literature contains numerous prescriptions for correcting existing functionals[2,3] (e.g., B3LYP-D3 and B3LYP-NL) and parameterizing dispersion-corrected functionals[4,5] (e.g., B97-D and B97M-V), as well as an abundance of useful benchmark datasets[6,7] of hydrogen-bonded, dispersion-bound, and mixed systems (e.g., S22 and S66).

While damped atom-atom potentials[2,4,8], such as DFT-D2 and DFT-D3, are certainly more commonplace, nonlocal correlation functionals[9,10] (NLCF) such as vdW-DF2 and VV10 have been gaining popularity recently, because they have more physical content, fewer empirical parameters, and the potential of being significantly more accurate. The VV10 NLCF was originally paired with rPW86 exchange[11] and PBE correlation[12] to give the VV10 density functional[10]. Since the NLCF and the density functional both have the same name, the former will be referred to either as -V (analogous to Grimme's -D) or the VV10 NLCF, and the latter as VV10. The VV10 NLCF has since been parameterized onto semi-local (BLYP, revPBE, PBE, TPSS), hybrid (LC-ωPBE08, B3LYP, B3PW91, revPBE0, revPBE38, PBE0, TPSS0), and even double hybrid density functionals[3,13,14].

Vydrov and Van Voorhis accomplished the implementation of nonlocal correlation functionals in Gaussian basis set codes[15] for energy and nuclear gradient evaluations, and NLCFs are available in Q-Chem[16], Orca[17], and ERKALE[18]. More recently, the efficient implementation of NLCFs in plane-wave basis set codes has been made possible due to an interpolation scheme[19] introduced by Román-Pérez and Soler (RPS). While the RPS procedure was initially only applicable to the vdW-DF NLCFs, Sabatini and coworkers recently introduced a slight modification[20] to the VV10 NLCF (called rVV10), enabling its efficient implementation in periodic codes such as Quantum Espresso[21], CP2K[22], and ONETEP[23].

The VV10 NLCF is a critical component of three combinatorially optimized, semi-empirical density

---

[a)] Electronic mail: nmardirossian@berkeley.edu
[b)] Electronic mail: lar739@lbl.gov
[c)] Electronic mail: J.C.Womack@soton.ac.uk
[d)] Electronic mail: C.Skylaris@soton.ac.uk
[e)] Electronic mail: thg@berkeley.edu
[f)] Electronic mail: mhg@cchem.berkeley.edu

functionals[5,24,25]: ωB97X-V, ωB97M-V, and B97M-V. The lattermost has a semi-local parent functional, and can potentially be very useful for solid-state calculations. Since B97M-V is very accurate for non-covalent interactions compared to most existing density functionals, it is important to verify that replacing -V with -rV in B97M-V does not lead to a significant degradation in accuracy. This is of immediate relevance to recent work in the ONETEP linear-scaling electronic structure package[23], where an efficient implementation of B97M-rV has recently been accomplished, allowing the application of this functional to systems containing thousands of atoms[26]. The rest of this letter focuses on defining and assessing B97M-rV relative to B97M-V, along with VV10 and rVV10. In addition to the S22 dataset[6], the functionals are compared across the extensive database of interaction energies from Reference 25.

I. COMPUTATIONAL DETAILS

The def2-QZVPPD basis set is used for all calculations without counterpoise corrections. A (99,590) grid (99 radial shells with 590 angular grid points per shell) is used to calculate the contribution from local and semi-local exchange-correlation functionals, while the SG-1 grid[27] is used to calculate the contribution from the VV10 nonlocal correlation functional.

The database[25] used to assess the functionals in this benchmark is constructed from 84 existing datasets and contains 4986 data points, requiring 5931 single-point calculations. 82 of these 84 datasets (AE18 and RG10 are excluded from this study) are classified according to eight categories (or datatypes): NCED (non-covalent dimers (easy)), NCEC (non-covalent clusters (easy)), NCD (non-covalent dimers (difficult)), IE (isomerization energies (easy)), ID (isomerization energies (difficult)), TCE (thermochemistry (easy)), TCD (thermochemistry (difficult)), and BH (barrier heights). The number of data points (and datasets) in NCED, NCEC, NCD, IE, ID, TCE, TCD, and BH are 1744 (18), 243 (12), 91 (5), 755 (12), 155 (5), 947 (15), 258 (7), and 206 (8), respectively. "Difficult" interactions involve either self-interaction error or strong correlation, while "easy" interactions are not significantly characterized by either. Detailed information about the datasets can be found in Table 1 of Reference 25. All of the calculations are performed self-consistently with a development version of the Q-Chem 4 package[16].

II. THEORY

The VV10 nonlocal correlation functional energy expression is given in Equation 1, and the associated non-local correlation (NLC) kernel, Φ, is defined in Equations 2-5, where $R = |\mathbf{r} - \mathbf{r}'|$, and primed variables are a function of $\mathbf{r}'$ rather than $\mathbf{r}$.

$$E_c^{VV10} = \int d\mathbf{r}\rho(\mathbf{r})\left[\frac{1}{32}\left(\frac{3}{b^2}\right)^{\frac{3}{4}} + \frac{1}{2}\int d\mathbf{r}'\rho(\mathbf{r}')\Phi(\mathbf{r},\mathbf{r}')\right] \quad (1)$$

$$\Phi = -\frac{3}{2}\left(gg'(g+g')\right)^{-1} \quad (2)$$

$$g = \omega_0 R^2 + \kappa \quad (3)$$

$$\omega_0 = \sqrt{C\left|\frac{\nabla\rho}{\rho}\right|^4 + \frac{4\pi}{3}\rho} \quad (4)$$

$$\kappa = \frac{3\pi b}{(576\pi)^{\frac{1}{6}}}\rho^{\frac{1}{6}} \quad (5)$$

The VV10 NLCF depends on two empirical parameters, b and C. The former controls the short-range damping of the r^{-6} asymptote, and the latter controls the accuracy of the long-range C_6 coefficients. While it is clear from Equation 1 that E_c^{VV10} and b are inversely related, it is generally true that E_{bind} and b are inversely related, such that as $b \to \infty$, the binding energy of the parent functional is recovered.

The difference between the VV10 and rVV10 NLCFs originates in the nonlocal correlation kernel. To elucidate the difference, the NLC kernel in Equation 2 can be rewritten in the form given in Equation 6. Despite appearing more complicated, the approximation of setting $\sqrt{\frac{\kappa}{\kappa'}} = 1$ results in the simplified rVV10 kernel (Equation 7) and enables the use of the RPS procedure.

$$\Phi = -\frac{3}{2}\kappa^{-\frac{3}{2}}\kappa'^{-\frac{3}{2}}\left(\frac{g}{\kappa}\frac{g'}{\kappa'}\left(\sqrt{\frac{\kappa}{\kappa'}}\frac{g}{\kappa} + \sqrt{\frac{\kappa'}{\kappa}}\frac{g'}{\kappa'}\right)\right)^{-1} \quad (6)$$

$$\Phi^* = -\frac{3}{2}\kappa^{-\frac{3}{2}}\kappa'^{-\frac{3}{2}}\left(\frac{g}{\kappa}\frac{g'}{\kappa'}\left(\frac{g}{\kappa} + \frac{g'}{\kappa'}\right)\right)^{-1} \quad (7)$$

Vydrov and Van Voorhis originally paired the VV10 NLCF with rPW86 exchange and PBE correlation, and optimized the values of the two empirical parameters to $b = 5.9$ and $C = 0.0093$, giving the VV10 density functional. Sabatini and coworkers retained the value of C, as well as the exchange and correlation functionals, switched -V with -rV, and optimized $b = 6.3$ on the S22 dataset, to define the rVV10 density functional.

III. RESULTS AND DISCUSSION

Table I and Figure 1 contain the results for varying b across the S22 dataset for VV10, rVV10, B97M-V, and B97M-rV. The root-mean-square deviations (RMSD) are computed both with respect to the original functional (either VV10 or B97M-V), as well as the S22 reference values. Beginning with the comparison of VV10@b to VV10, the RMSD should be exactly 0 at the VV10 defined value of $b = 5.9$. The RMSD increases nearly linearly as one deviates from $b = 5.9$, at a rate of approximately 0.065 kcal/mol per 0.1 units of b. Moving on to the comparison of VV10@b to the S22 reference values, the minimum of VV10@b should be at 5.9, but it lies at about 6.05. There are two possible explanations for this discrepancy. First, the S22 reference values[28] (S22B) used here are slightly different from those employed in the original paper. However, the mean absolute deviation (MAD) between the

two sets of reference values is only 0.012 kcal/mol. The other, more significant source of error is the choice of basis set, which was counterpoise-corrected (CP) aug-cc-pVTZ in the original paper, but non-counterpoise-corrected (noCP) def2-QZVPPD in this work. The MAD and mean signed error (MSE) between the former and latter results is 0.044 and 0.045 kcal/mol, respectively, meaning that the CP aug-cc-pVTZ results are systematically underbound compared to the noCP def2-QZVPPD results. This explains the optimization of b toward a slightly larger value. Regardless, the difference between the VV10@b=5.9 RMSD of 0.50 kcal/mol and the VV10@b=6.05 RMSD of 0.49 kcal/mol is insignificant.

Comparing rVV10@b to VV10, the optimal value of b shifts from 5.9 to around 6.1, which is consistent with the finding of Sabatini and coworkers, namely that the rVV10 NLCF tends to systematically overbind compared to the VV10 NLCF. Therefore, a slightly larger value of b is warranted, as it helps reduce the overbinding. It is remarkable that with a value of $b = 6.1$ the RMSD of rVV10@b=6.1 vs. VV10 can be as small as 0.02 kcal/mol, while the RMSD for the rVV10 defined value of $b = 6.3$ is actually 0.14 kcal/mol relative to VV10. Nevertheless, this is because $b = 6.3$ was parameterized to minimize errors with respect to the S22 reference values, and not VV10. Considering the comparison of rVV10@b to the S22 reference values demonstrates that the value of $b = 6.3$ found by Sabatini and coworkers is indeed not far from the optimum value found in this work, which is 6.25.

Considering the B97M-V@b results, the minimum vs. B97M-V should be exactly 0 at $b = 6$, since that is the value used in the B97M-V density functional. Consistent with the trend witnessed with VV10@b, the RMSD increases linearly as one deviates from the defined value, at a rate of approximately 0.0625 kcal/mol per 0.1 units of b. Compared to the reference, however, the minimum for b is not at 6.0, but at 5.81. The explanation for this result is simply that B97M-V was parameterized using the aug-cc-pVTZ basis set without counterpoise corrections for non-covalent interactions, and closer to the basis set limit (e.g., noCP def2-QZVPPD) the binding energies will be smaller in magnitude, leading to the need for a smaller value of b. Nevertheless, the difference between the smallest possible RMSD (0.28 kcal/mol) and the $b = 6$ RMSD (0.31 kcal/mol), is only about 0.03 kcal/mol, so the basis set effect is not significant.

Moving on to the comparison of B97M-rV@b to B97M-V, the optimal value of b once again shifts to a larger value (6.2), further confirming that the rVV10 NLCF tends to bind systems more than the VV10 NLCF. The most interesting outcome appears in the comparison of the B97M-rV@b results to the S22 reference values. The optimal value of b is around 6 (6.02 to be more precise), which happens to be almost identical to the defined value of b for B97M-V. A fortuitous cancellation of errors accounts for this result. Closer to the basis set limit, B97M-V tends to be underbound, but the rVV10 NLCF overbinds relative to the VV10 NLCF.

Thus, the underbinding is offset by the overbinding, and a minimum near $b = 6$ is recovered. **As a result of this assessment, the B97M-rV density functional will be defined as having a value of $b = 6$.**

With the results of the S22 test in hand, seven functionals will be assessed on a database of 4399 data points (82 datasets). The first four functionals are based on the combination of the rPW86 exchange functional and the PBE correlation functional, and are rPW86PBE, rPW86PBE-V ($b = 5.9$), rPW86PBE-rV ($b = 5.9$), and rPW86PBE-rV ($b = 6.3$). The second and fourth functionals are already defined in the literature and are VV10 and rVV10, respectively. The third functional is rVV10@b=5.9, and it will be referred to as rVV10* for brevity. Although rVV10* uses a suboptimal value for b, it is included in the comparison in order to directly gauge the differences between the VV10 and rVV10 NLCFs, independent of the effect of different values of b. The other three functionals are based on the B97M-V functional, and are B97M (B97M-V with the VV10 NLCF correlation energy removed), B97M-V, and B97M-rV.

Table II contains RMSDs for the 8 energetic datatypes found in the assessment database, as well as RMSDs for interpolated equilibrium bond lengths (EBL) and equilibrium binding energies (EBE) of 81 non-covalent dimers. Considering first the results for the functionals whose parent is rPW86PBE, it is noticeable that across VV10, rVV10*, and rVV10, the RMSD for the thermochemistry categories (TCE and TCD) changes negligibly. In fact, the most pronounced changes are seen in the categories that pertain to conventional closed-shell non-covalent interactions (NCED, NCEC, EBL, EBE) and isomerization energies (IE). This is a positive (and perhaps expected) result, because it indicates that minor modifications to the NLCF (either in the value of b or the kernel) do not have much of an effect on bonded interactions.

Going from VV10 to rVV10*, the NCEC RMSD increases from 8.02 kcal/mol to 8.48 kcal/mol, and this is solely due to the fact that VV10 is already overbound for molecular clusters (e.g., water hexamers), and changing the NLCF from -V to -rV (without altering the value of b) should worsen the overbinding and lead to a larger RMSD. This effect is confirmed in Table III, which contains mean signed errors for several NCED and NCEC datasets. For example, considering the hexamers in H2O6Bind8, the MSE for VV10 is -5.42 kcal/mol, but -5.68 kcal/mol for rVV10*. To see an even more pronounced effect, the water 20-mers in H2O20Bind10 can be analyzed. For this dataset, the MSE for VV10 is -22.68 kcal/mol, but -24.32 kcal/mol for rVV10*.

If the b value of rVV10@b were fit to reproduce VV10 results, the optimal value would be at $b = 6.1$, which is smaller than the rVV10 defined value of $b = 6.3$. When rVV10 itself is considered, the NCEC result improves even over VV10. The MSEs in Table III confirm that this is because rVV10 binds less than VV10, with the former having an MSE of -20.83 kcal/mol for

RMSD (kcal/mol)	5.5	5.6	5.7	5.8	5.9	6.0	6.1	6.2	6.3	6.4	6.5	6.6	6.7	6.8	6.9	7.0
VV10@b vs. VV10	0.34	0.25	0.16	0.08	**0.00**	0.08	0.15	0.22	0.29	0.36	0.43	0.49	0.55	0.61	0.67	0.73
VV10@b vs. Ref	0.66	0.61	0.56	0.52	0.50	**0.49**	0.49	0.50	0.52	0.55	0.58	0.62	0.66	0.70	0.74	0.78
rVV10@b vs. VV10	0.53	0.43	0.34	0.25	0.17	0.09	**0.02**	0.07	0.14	0.22	0.28	0.35	0.42	0.48	0.54	0.60
rVV10@b vs. Ref	0.79	0.72	0.65	0.59	0.55	0.51	0.49	0.47	**0.47**	0.48	0.50	0.53	0.56	0.60	0.64	0.68
B97M-V@b vs. B97M-V	0.40	0.31	0.23	0.15	0.07	**0.00**	0.07	0.14	0.21	0.27	0.34	0.40	0.46	0.51	0.57	0.62
B97M-V@b vs. Ref	0.38	0.33	0.29	**0.28**	0.28	0.31	0.35	0.39	0.44	0.50	0.55	0.60	0.65	0.71	0.76	0.81
B97M-rV@b vs. B97M-V	0.58	0.49	0.40	0.32	0.24	0.16	0.08	**0.02**	0.06	0.13	0.20	0.26	0.33	0.39	0.44	0.50
B97M-rV@b vs. Ref	0.53	0.45	0.39	0.34	0.30	**0.29**	0.29	0.31	0.35	0.39	0.44	0.49	0.54	0.59	0.65	0.70

TABLE I. Root-mean-square deviations (RMSD) in kcal/mol for VV10, rVV10, B97M-V, and B97M-rV for different values of the empirical parameter, b. The RMSDs are computed both with respect to the original functional (either VV10 or B97M-V), as well as the S22 reference values.
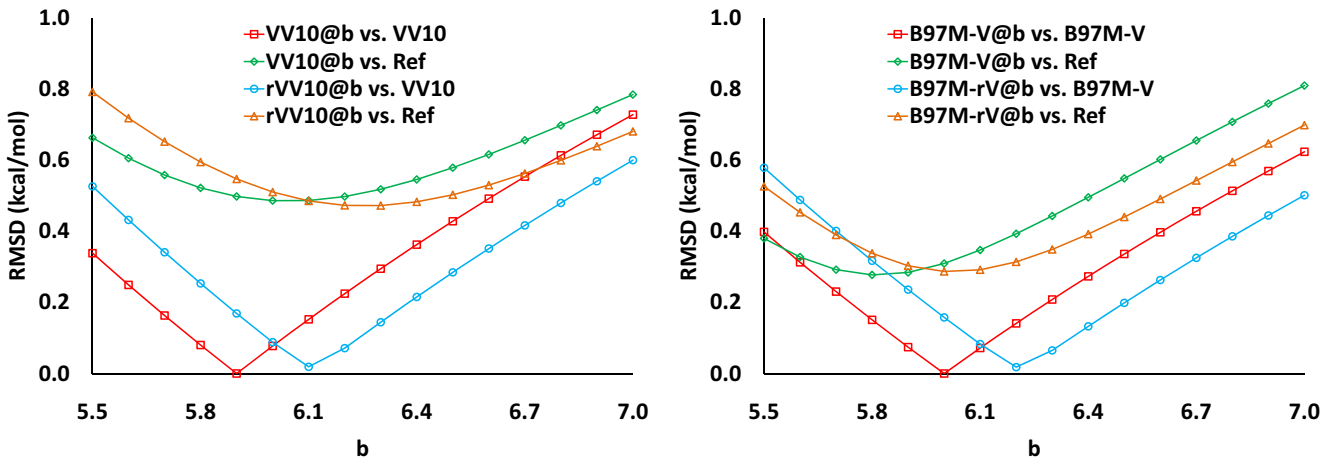

FIG. 1. Plots of root-mean-square deviations (RMSD) in kcal/mol for VV10, rVV10, B97M-V, and B97M-rV for different values of the empirical parameter, b. The RMSDs are computed both with respect to the original functional (either VV10 or B97M-V), as well as the S22 reference values.

Datatype	NCED	NCEC	NCD	IE	ID	TCE	TCD	BH	EBL (Å)	EBE
# Data Points	1744	243	91	755	155	947	258	206	81	81
rPW86PBE	2.10	4.88	2.80	1.13	12.05	8.39	12.74	9.31	0.323	2.36
VV10	0.44	8.02	3.65	0.64	8.42	7.41	10.86	9.69	0.028	0.47
rVV10*	0.48	8.48	3.70	0.66	8.38	7.41	10.87	9.73	0.027	0.53
rVV10	0.42	7.49	3.60	0.61	8.72	7.42	10.75	9.64	0.033	0.42
B97M	2.20	11.41	1.80	1.20	9.31	5.02	10.65	5.39	0.326	2.48
B97M-V	0.24	0.95	2.01	0.27	6.48	3.57	4.82	4.35	0.026	0.17
B97M-rV	0.22	0.67	2.04	0.28	6.45	3.57	4.78	4.36	0.025	0.17

TABLE II. Root-mean-square deviations in kcal/mol for 8 energetic datatypes (NCED through BH), as well as interpolated equilibrium bond lengths (EBL) and equilibrium binding energies (EBE) for a set of 81 dimers. NCED stands for non-covalent dimers (easy), NCEC stands for non-covalent clusters (easy), NCD stands for non-covalent dimers (difficult), IE stands for isomerization energies (easy), ID stands for isomerization energies (difficult), TCE stands for thermochemistry (easy), TCD stands for thermochemistry (difficult), and BH stands for barrier heights. "Difficult" interactions involve either self-interaction error or strong correlation, while "easy" interactions are not significantly characterized by either.

H2O20Bind10 and the latter having an MSE of -22.68 kcal/mol. Regardless, both functionals are severely overbound for water clusters. The reason for this is ultimately due to the choice of the parent functional. In order to effectively combine -V or -rV with a parent functional, the latter should underbind hydrogen bonds and dispersion interactions in a similar fashion. In general, rVV10 is the best performing functional out of the three rPW86PBE-based dispersion-corrected functionals, followed by VV10, and then rVV10*.

Table IV contains individual binding energies for the systems in the S22 dataset. Considering the results for the water dimer, it is clear that rPW86PBE is only slightly underbound (-4.87 kcal/mol vs. a reference of -4.99 kcal/mol), and that the three rPW86PBE-based dispersion-corrected functionals are overbound by about 10%. In fact, a value of b much larger than 10 is necessary to make the water dimer binding energies for these functionals match the reference. However, while a very large value of b may be suitable for strong hydrogen bonds, considering the binding energy of the methane dimer in Table IV indicates that rPW86PBE is not bound at all, and a value of $b = 5.8$ is necessary for rVV10 to recover the reference value of -0.53 kcal/mol. Furthermore, looking back to the S22 results in Table I, it is clear that the rVV10@b results would degrade by almost 50% if the value of b was increased from 6.3 to 7,

Dataset	S22	S66	HSG	H2O6Bind8	FmH2O10	H2O20Bind10
rPW86PBE	2.86	2.35	1.80	1.17	8.68	16.15
VV10	-0.21	-0.31	-0.31	-5.42	-11.51	-22.68
rVV10*	-0.34	-0.42	-0.40	-5.68	-12.33	-24.34
rVV10	-0.09	-0.21	-0.24	-5.06	-10.44	-20.83
B97M	3.11	2.60	2.03	6.24	18.72	38.49
B97M-V	0.20	0.08	0.05	0.19	0.07	2.75
B97M-rV	0.08	-0.03	-0.04	-0.05	-0.69	1.24

TABLE III. Mean signed errors in kcal/mol for 6 datasets of non-covalent interactions. S22 consists of 22 dispersion-bound, mixed, and hydrogen-bonded non-covalent complexes, S66 consists of 66 non-covalent interactions found in organic molecules and biomolecules, and HSG consists of 21 small ligands interacting with protein receptors. H2O6Bind8 consists of 8 isomers of the water hexamer, FmH2O10 consists of 10 isomers of $F^-(H_2O)_{10}$, and H2O20Bind10 consists of 10 isomers of $(H_2O)_{20}$.

so a value much larger than 10 will simply be unsuitable for different types of interactions.

Moving on to the results for the B97M-based functionals, it is once again evident that the performance for thermochemistry is unaffected by the choice of nonlocal correlation functional. For example, the TCE RMSDs for B97M-V and B97M-rV are both 3.57 kcal/mol, and for TCD, the values are 4.82 and 4.78 kcal/mol, respectively. In fact, the only substantial differences between B97M-V and B97M-rV come from standard non-covalent interaction binding energies (NCED and NCEC). Luckily, due to the cancellation of underbinding and overbinding identified earlier, NCED improves by about 10% when going from B97M-V to B97M-rV, and NCEC improves by nearly 30%. Looking at the values in Table III, the H2O20Bind10 MSE is 2.75 kcal/mol for B97M-V, yet only 1.24 kcal/mol for B97M-rV. Similarly, for the S66 dataset, the underbound MSE of 0.08 kcal/mol of B97M-V changes to a very slightly overbound MSE of -0.03 kcal/mol for B97M-rV. In fact, for all 6 datasets in Table III, the MSE decreases when going from B97M-V to B97M-rV.

For the rPW86PBE-based functionals, the hydrogen-bonded systems in S22 prefer a large value of b, while for the dispersion-bound systems, a value around $b = 6$ is appropriate. It is interesting to see if the same is true for the B97M-based dispersion-corrected functionals. Contrary to rVV10, the optimal b value for B97M-rV@b for the methane dimer is around 5.5, while the optimal b value for the water dimer is around 6.3. Thus, it is clear that a parent functional that underbinds different types of non-covalent interactions to the same degree is important, since a single value of b is more suitable for different types of interactions. Compared to the rPW86PBE binding energies of 0.07 and -4.87 kcal/mol for the methane dimer and water dimer, respectively, B97M predicts values of 0.09 and -4.39 kcal/mol, respectively. In comparison, the rVV10 binding energies are -0.47 and -5.50 kcal/mol, respectively, while the B97M-rV binding energies are -0.46 and -5.03 kcal/mol, respectively. Compared to the reference values of -0.53 and -4.99 kcal/mol, respectively, it is clear that the B97M-rV results are in better agreement.

During their assessment of the VV10 NLCF, Hujo and Grimme[3] fit -V to the BLYP and revPBE semi-local GGA density functionals. The combination of revPBE and -V, termed revPBE-NL (or revPBE-V using the present notation), turned out to be significantly more accurate than VV10 for non-covalent interactions, by more than a factor of 2 in the WTMAD (weighted total mean absolute deviation of relative energies). The success of revPBE-NL is related to the fact that revPBE severely underbinds all of the systems in S22, with the MAD and MSE both equal to 5.29 kcal/mol (vs. 2.86 kcal/mol for rPW86PBE). In fact, the S22 MSE of revPBE is even larger than that of Hartree–Fock, which is 4.99 kcal/mol. For the methane dimer and water dimer, revPBE gives binding energies of 0.46 and -3.39 kcal/mol, respectively, while revPBE-NL gives binding energies of -0.48 and -4.66 kcal/mol, respectively. Thus, it is indeed important to either pick a suitable parent functional (as in revPBE-NL) or to simultaneously train the parent functional onto the nonlocal correlation functional (as in B97M-V).

Finally, despite the heavy computational cost of evaluating exact exchange for extended systems, it is inevitable that computational and algorithmic advances will facilitate the use of hybrid density functionals in periodic codes. Thus, it is likely that hybrid density functionals that incorporate the rVV10 nonlocal correlation functional will be utilized in future applications. While the two aforementioned, combinatorially optimized range-separated hybrid density functionals, ωB97X-V and ωB97M-V, use the VV10 NLCF, it is straightforward to optimize b on the S22 dataset in order to recommend optimal parameters for ωB97X-rV and ωB97M-rV. Following the same procedure used to produce Table I and Figure 1, a value of $b = 6.2$ is recommended for both ωB97X-rV and ωB97M-rV.

IV. CONCLUSIONS

In summary, we recommend a value of $b = 6$ to define the B97M-rV density functional. B97M-rV is a small but significant modification to the B97M-V density functional, where the VV10 nonlocal correlation functional (NLCF) is replaced by the rVV10 NLCF. Since the rVV10 NLCF can be efficiently used in calculations on extended systems, it is important to establish that the change of NLCF does not adversely affect the performance of B97M-V. This work represents an important and necessary first step in validating B97M-rV on the types of systems that were used to train and test B97M-V. The results of this study indicate that since B97M-V is slightly underbound for dimers and clusters at the basis set limit, and since the rVV10 NLCF overbinds relative to the VV10 NLCF, B97M-rV actually improves upon B97M-V for non-covalent interactions. Moreover, comparisons of the VV10 and rVV10 NLCFs across the S22 dataset (as well as a comprehensive database of over 4000 data points) indicate that the rVV10 NLCF can seamlessly replace existing pa-

Complex (Symmetry)	Ref.	rPW86PBE	VV10	rVV10*	rVV10	B97M	B97M-V	B97M-rV
CH_4 dimer (D_{3d})	-0.53	0.07	-0.49	-0.52	-0.47	0.09	-0.44	-0.46
C_2H_4 dimer (D_{2d})	-1.47	-0.18	-1.45	-1.50	-1.39	-0.13	-1.33	-1.38
Benzene–CH_4 (C_3)	-1.45	0.07	-1.47	-1.53	-1.41	0.09	-1.37	-1.43
Benzene dimer (C_{2h})	-2.65	1.88	-2.76	-2.96	-2.59	1.85	-2.61	-2.80
Pyrazine dimer (C_s)	-4.26	0.78	-4.04	-4.25	-3.86	0.64	-3.95	-4.14
Uracil dimer (C_2)	-9.81	-2.61	-9.76	-10.10	-9.53	-3.09	-9.87	-10.19
Indole–benzene (C_1)	-4.52	2.18	-4.57	-4.86	-4.32	2.01	-4.47	-4.74
Adenine–thymine (C_1)	-11.73	-1.34	-11.47	-11.92	-11.11	-2.27	-11.93	-12.35
C_2H_4–C_2H_2 (C_{2v})	-1.50	-1.05	-1.69	-1.72	-1.66	-0.92	-1.53	-1.55
Benzene–H_2O (C_s)	-3.28	-1.85	-3.34	-3.41	-3.29	-1.74	-3.14	-3.21
Benzene–NH_3 (C_s)	-2.31	-0.80	-2.32	-2.40	-2.27	-0.75	-2.18	-2.25
Benzene–HCN (C_s)	-4.54	-2.53	-4.31	-4.38	-4.23	-2.57	-4.26	-4.33
Benzene dimer (C_{2v})	-2.72	0.05	-2.57	-2.68	-2.48	0.12	-2.38	-2.49
Indole–benzene (C_s)	-5.63	-1.69	-5.30	-5.46	-5.18	-1.65	-5.09	-5.24
Phenol dimer (C_1)	-7.10	-3.84	-7.05	-7.19	-6.94	-3.58	-6.62	-6.75
NH_3 dimer (C_{2h})	-3.13	-2.72	-3.45	-3.48	-3.41	-2.44	-3.11	-3.14
H_2O dimer (C_s)	-4.99	-4.87	-5.54	-5.57	-5.50	-4.39	-5.01	-5.03
Formic acid dimer (C_{2h})	-18.75	-17.68	-20.05	-20.13	-19.91	-16.50	-18.72	-18.79
Formamide acid dimer (C_{2h})	-16.06	-14.57	-16.80	-16.88	-16.68	-13.59	-15.67	-15.74
Uracil dimer (C_{2h})	-20.64	-18.25	-21.19	-21.29	-21.06	-17.42	-20.17	-20.26
2-pyridone–2-aminopyridine (C_1)	-16.93	-14.92	-18.13	-18.24	-17.99	-13.44	-16.46	-16.56
Adenine–thymine WC (C_1)	-16.66	-13.96	-17.51	-17.64	-17.36	-12.56	-15.90	-16.02

TABLE IV. Binding energies in kcal/mol for the S22 dataset. The S22B reference values are provided in the second column. The first set of interactions are dispersion-bound, the middle set of interactions are mixed, and the last set of interactions are hydrogen-bonded.

rameterizations involving the VV10 NLCF by increasing the value of b to compensate for the inherent tendency of the rVV10 NLCF to overbind relative to the VV10 NLCF. Although recent AIMD simulations of liquid water with B97M-rV have already yielded encouraging results across a large set of water properties[29], it remains to further test the B97M-rV density functional for solids, liquids, and interfaces in order to see if its demonstrated accuracy for various molecular interactions extends to properties that are significantly different from those used to train its parameters.

[1] J. Klimeš and A. Michaelides, "Perspective: Advances and Challenges in Treating van der Waals Dispersion Forces in Density Functional Theory," J. Chem. Phys. **137**, 120901 (2012).

[2] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu," J. Chem. Phys. **132**, 154104 (2010).

[3] W. Hujo and S. Grimme, "Performance of the van der Waals Density Functional VV10 and (hybrid)GGA Variants for Thermochemistry and Noncovalent Interactions," J. Chem. Theory Comput. **7**, 3866–3871 (2011).

[4] S. Grimme, "Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction," J. Comput. Chem. **27**, 1787–1799 (2006).

[5] N. Mardirossian and M. Head-Gordon, "Mapping the Genome of meta-Generalized Gradient Approximation Density Functionals: The Search for B97M-V," J. Chem. Phys. **142**, 074111 (2015).

[6] P. Jurečka, J. Šponer, J. Černý, and P. Hobza, "Benchmark Database of Accurate (MP2 and CCSD(T) Complete Basis Set Limit) Interaction Energies of Small Model Complexes, DNA Base Pairs, and Amino Acid Pairs," Phys. Chem. Chem. Phys. **8**, 1985–1993 (2006).

[7] J. Řezáč, K. E. Riley, and P. Hobza, "S66: A Well-Balanced Database of Benchmark Interaction Energies Relevant to Biomolecular Structures," J. Chem. Theory Comput. **7**, 2427–2438 (2011).

[8] S. Grimme, S. Ehrlich, and L. Goerigk, "Effect of the Damping Function in Dispersion Corrected Density Functional Theory," J. Comput. Chem. **32**, 1456–1465 (2011).

[9] K. Lee, E. D. Murray, L. Kong, B. I. Lundqvist, and D. C. Langreth, "Higher-Accuracy van der Waals Density Functional," Phys. Rev. B **82**, 081101 (2010).

[10] O. A. Vydrov and T. V. Voorhis, "Nonlocal van der Waals Density Functional: The Simpler the Better," J. Chem. Phys. **133**, 244103 (2010).

[11] E. D. Murray, K. Lee, and D. C. Langreth, "Investigation of Exchange Energy Density Functional Accuracy for Interacting Molecules," J. Chem. Theory Comput. **5**, 2754–2762 (2009).

[12] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple," Phys. Rev. Lett. **77**, 3865–3868 (1996).

[13] J. Aragó, E. Ortí, and J. C. Sancho-García, "Nonlocal van der Waals Approach Merged with Double-Hybrid Density Functionals: Toward the Accurate Treatment of Noncovalent Interactions," J. Chem. Theory Comput. **9**, 3437–3443 (2013).

[14] F. Yu, "Spin-Component-Scaled Double-Hybrid Density Functionals with Nonlocal van der Waals Correlations for Noncovalent Interactions," J. Chem. Theory Comput. **10**, 4400–4407 (2014).

[15] O. A. Vydrov, Q. Wu, and T. V. Voorhis, "Self-Consistent Implementation of a Nonlocal van der Waals Density Functional with a Gaussian Basis Set," J. Chem. Phys. **129**, 014106 (2008).

[16] Y. Shao *et al.*, "Advances in Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package," Mol. Phys. **113**, 184–215 (2015).

[17] F. Neese, "The ORCA Program System," WIREs Comput Mol Sci **2**, 73–78 (2012).

[18] J. Lehtola, M. Hakala, A. Sakko, and K. Hämäläinen, "ERKALE–A Flexible Program Package for X-Ray Properties of Atoms and Molecules," J. Comput. Chem. **33**, 1572–1585 (2012).

[19] G. Román-Pérez and J. M. Soler, "Efficient Implementation of a van der Waals Density Functional: Application to Double-Wall Carbon Nanotubes," Phys. Rev. Lett. **103**, 096102 (2009).

[20] R. Sabatini, T. Gorni, and S. de Gironcoli, "Nonlocal van der Waals Density Functional Made Simple and Efficient," Phys. Rev. B **87**, 041108 (2013).

[21] P. Giannozzi *et al.*, "QUANTUM ESPRESSO: A Modular and Open-Source Software Project for Quantum Simulations of Materials," J. Phys.: Condens. Matter **21**, 395502 (2009).

[22] J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, "cp2k: Atomistic Simulations of Condensed Matter Systems," WIREs Comput Mol Sci **4**, 15–25 (2014).

[23] C.-K. Skylaris, P. D. Haynes, A. A. Mostofi, and M. C. Payne, "Introducing ONETEP: Linear-Scaling Density Functional Simulations on Parallel Computers," J. Chem. Phys. **122**, 084119 (2005).

[24] N. Mardirossian and M. Head-Gordon, "ωB97X-V: A 10-Parameter, Range-Separated Hybrid, Generalized Gradient Approximation Density Functional with Nonlocal Correlation, Designed by a Survival-of-the-Fittest Strategy," Phys. Chem. Chem. Phys. **16**, 9904–9924 (2014).

[25] N. Mardirossian and M. Head-Gordon, "ωB97M-V: A Combinatorially Optimized, Range-Separated Hybrid, meta-GGA Density Functional with VV10 Nonlocal Correlation," J. Chem. Phys. **144**, 214110 (2016).

[26] J. C. Womack, N. Mardirossian, M. Head-Gordon, and C.-K. Skylaris, "Self-Consistent Implementation of meta-GGA Functionals for the ONETEP Linear-Scaling Electronic Structure Package," J. Chem. Phys. **145**, 204114 (2016).

[27] P. M. W. Gill, B. G. Johnson, and J. A. Pople, "A Standard Grid for Density Functional Calculations," Chem. Phys. Lett. **209**, 506 – 512 (1993).

[28] M. S. Marshall, L. A. Burns, and C. D. Sherrill, "Basis Set Convergence of the Coupled-Cluster Correction, $\delta_{MP2}^{CCSD(T)}$: Best Practices for Benchmarking Non-Covalent Interactions and the Attendant Revision of the S22, NBC10, HBC6, and HSG Databases," J. Chem. Phys. **135**, 194102 (2011).

[29] L. R. Pestana, N. Mardirossian, M. Head-Gordon, and T. Head-Gordon, "Ab Initio Simulations of Liquid Water using High Quality meta-GGA Functionals," Chem. Sci. (Submitted).