Title
Self-consistent implementation of meta-GGA functionals for the ONETEP linear-scaling electronic structure package

Journal
The Journal of Chemical Physics, 145(20)

ISSN
0021-9606

Authors
Womack, James C
Mardirossian, Narbe
Head-Gordon, Martin
et al.

Publication Date
2016-11-28

DOI
10.1063/1.4967960

Peer reviewed

Self-consistent implementation of meta-GGA functionals for the ONETEP linear-scaling electronic structure package

James C. Womack,[1] Narbe Mardirossian,[2] Martin Head-Gordon,[2,3] and Chris-Kriton Skylaris[1,*]

[1] *School of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, UK*
[2] *Kenneth S. Pitzer Center for Theoretical Chemistry, Department of Chemistry, University of California, Berkeley, California 94720, USA*
[3] *Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

Accurate and computationally efficient exchange-correlation functionals are critical to the successful application of linear-scaling density functional theory (DFT). Local and semi-local functionals of the density are naturally compatible with linear-scaling approaches, having a general form which assumes the locality of electronic interactions and which can be efficiently evaluated by numerical quadrature. Presently, the most sophisticated and flexible semi-local functionals are members of the meta-generalized-gradient approximation (meta-GGA) family, and depend upon the kinetic energy density, τ, in addition to the charge density and its gradient. In order to extend the theoretical and computational advantages of τ-dependent meta-GGA functionals to large-scale DFT calculations on thousands of atoms, we have implemented support for τ-dependent meta-GGA functionals in the ONETEP program. In this paper we lay out the theoretical innovations necessary to implement τ-dependent meta-GGA functionals within ONETEP's linear-scaling formalism. We present expressions for the gradient of the τ-dependent exchange-correlation energy, necessary for direct energy minimization. We also derive the forms of the τ-dependent exchange-correlation potential and kinetic energy density in terms of the strictly localized, self-consistently optimized orbitals used by ONETEP. To validate the numerical accuracy of our self-consistent meta-GGA implementation, we performed calculations using the B97M-V and PKZB meta-GGAs on a variety of small molecules. Using only a minimal basis set of self-consistently optimized local orbitals, we obtain energies in excellent agreement with large basis set calculations performed using other codes. Finally, to establish the linear-scaling computational cost and applicability of our approach to large-scale calculations, we present the outcome of self-consistent meta-GGA calculations on amyloid fibrils of increasing size, up to tens of thousands of atoms.

I. INTRODUCTION

Kohn-Sham density functional theory (KS-DFT) [1] is founded on the premise that the exact charge density of a system of interacting electrons can be represented by the density of an auxiliary system of independent particles. This ansatz allows the total energy functional for a many-electron system to be expressed in terms of large contributions from known functionals for the non-interacting system, and a relatively smaller correction for many-body exchange and correlation effects. Unfortunately, at present no computationally practicable exact form for this exchange-correlation functional, $E_{\mathrm{xc}}[n]$, is known.

A key challenge presented by KS-DFT is the development of accurate and computationally tractable forms for $E_{\mathrm{xc}}[n]$. Since the inception of KS-DFT, numerous approximate forms have been proposed, varying significantly in their construction. The "Jacob's ladder" of density functional approximations [2, 3] provides a useful metaphor for categorizing the different types of approximations to $E_{\mathrm{xc}}[n]$, arranging them into "rungs" based on the density-dependent ingredients with which they are built. Each rung on the ladder ascending towards the "heaven of chemical accuracy" introduces additional density-dependent ingredients, which can be used to construct density functionals of increasing sophistication and accuracy.

The first three rungs of the Jacob's ladder are occupied by the local and semi-local functional families: the local density approximations (LDAs), generalized gradient approximations (GGAs) and meta-generalized-gradient approximations (meta-GGAs). These share the general form

$$E_{\mathrm{xc}}[n] = \int \mathrm{d}\mathbf{r}\, \epsilon_{\mathrm{xc}}(\mathbf{r}) \tag{1}$$

where $\epsilon_{\mathrm{xc}}(\mathbf{r})$ is the exchange-correlation energy per unit volume, which depends on the values of the density, and other density-derived variables only at point $\mathbf{r}$. Higher rungs on the ladder introduce non-local dependence on the Kohn-Sham orbitals, $\{\psi_i\}$, and eigenvalues, $\{\varepsilon_i\}$, and have more complicated and computationally demanding forms than Eq. 1.

The meta-GGA family of functionals offers the most sophisticated and flexible semi-local forms, in which $\epsilon_{\mathrm{xc}}(\mathbf{r})$ depends on the density, $n(\mathbf{r})$, its gradient, $\nabla n(\mathbf{r})$, and the kinetic energy density [4],

$$\tau(\mathbf{r}) = \frac{1}{2}\sum_{i}^{N_{\mathrm{occ}}} |\nabla\psi_i(\mathbf{r})|^2, \tag{2}$$

where the summation is over all N_{occ} occupied orbitals. The use of τ to construct exchange-correlation functionals offers significant theoretical benefits without abandoning the computationally simple form of Eq. 1. Nevertheless, τ-dependence does present additional challenges, particularly with respect to self-consistent implementation. We discuss these issues in sections II C and II D.

The addition of new density-dependent ingredients as the Jacob's ladder is ascended allows for the construction of more

* C.Skylaris@soton.ac.uk

sophisticated functionals, which are typically more accurate than those on the lower rungs. The metaphor of the Jacob's ladder is a useful shorthand for this general trend, however it should be noted that this is not a strict rule. The accuracy of any given functional depends on other factors not considered in the ladder metaphor, such as the design of the functional form and the context in which it is to be applied. When selecting a functional to study a particular chemical system, it is important to consider these other factors, rather than simply assuming that a higher-rung functional is the best tool for the job—a less computationally demanding lower-rung functional may be more suitable.

The computational demands of any approximate exchange-correlation functional must be considered in the context of the overall KS-DFT procedure. Conventional DFT calculations have a computational cost which scales cubically with the size of the system. This compares favorably to other quantum chemical methods but still constrains the type of systems which can practically be studied.

The cubic scaling cost of conventional DFT calculations arises from the need to orthogonalize Kohn-Sham orbitals which extend over the entire system, and may be overcome by exploiting the locality of electronic interactions (see for example Refs. 5 and 6). Using this "nearsightedness" [7], DFT calculations may be performed with a cost that scales linearly with the size of the system. These "linear scaling", or $O(N)$, methods have enabled DFT to be routinely applied to systems containing thousands of atoms.

ONETEP [8] is one of several software packages which implement linear-scaling DFT approaches. Other $O(N)$ DFT packages include BigDFT [9], CONQUEST [10], OpenMX [11], Quickstep [12] and SIESTA [13]. The combination of readily available linear-scaling DFT software and highly parallel modern supercomputing hardware has extended the scope of DFT to systems such as complex biological macromolecules and nanomaterials (see Refs. 14–17 for some recent applications of ONETEP to large systems of this type).

As mentioned above, a key challenge in the development of quantum chemical approaches and their application is the tension between computational cost and accuracy. This tension is particularly keenly felt in the development and extension of linear-scaling methods, where maintaining $O(N)$ scaling in all components is of paramount importance. Ideally, we would like to make use of the most accurate exchange-correlation approximations, but, as has already been mentioned, higher rungs of the Jacob's ladder of density functional approximations introduce non-local ingredients which pose a serious challenge for maintaining $O(N)$ scaling. High accuracy semi-local exchange-correlation functionals are of particular interest for linear-scaling DFT because they naturally incorporate the locality of electronic interactions and are computationally simple to evaluate using numerical integration. We suggest that the meta-GGAs, as the most sophisticated semi-local functionals, exist in a "Goldilocks zone" where accuracy and computational expense can be particularly favorably balanced for linear-scaling DFT approaches (Fig. 1).

Recently, linear-scaling exact exchange evaluation has been demonstrated within ONETEP [18] using a density-fitting-based approach. This opens up access the fourth rung of the Jacob's ladder, albeit at the cost of introducing an auxiliary basis set. Nevertheless, the comparative simplicity of the semi-local meta-GGAs remains highly desirable. Furthermore, recently developed meta-GGAs offer impressive accuracy. For example, the accuracy of the combinatorially designed semiempirical B97M-V functional [19] is comparable to popular hybrid functionals across a broad range of datasets. In some circumstances, notably for non-covalent interactions, B97M-V can outperform popular hybrid functionals. Promising results have also been reported for the recently developed nonempirical SCAN functional [20, 21]. In light of the performance of modern meta-GGAs such as B97M-V, the "Goldilocks zone" of Fig. 1 becomes very appealing for linear-scaling DFT. Indeed, this balance of computational simplicity and accuracy was the key motivator in our implementing support for meta-GGAs in ONETEP.

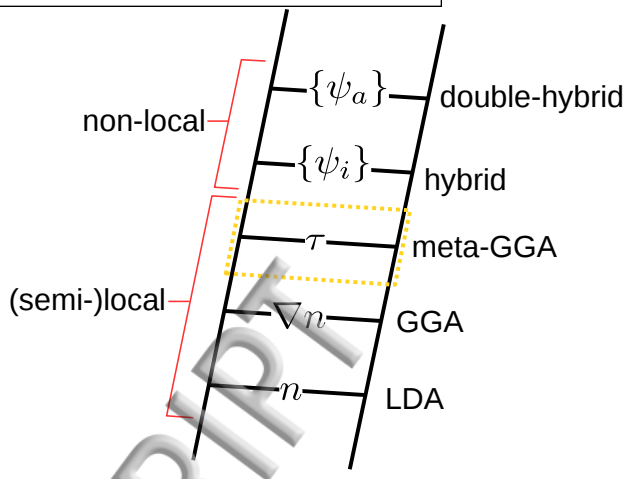

FIG. 1. The Jacob's ladder of density functional approximations ascending towards the "heaven of chemical accuracy" [2, 3], each rung introducing a new density-dependent ingredient. The region inside the yellow box is our suggested "Goldilocks zone", where the computational cost and accuracy are particularly well-balanced for linear-scaling DFT approaches.

In the remainder of this paper, the theoretical and computational details of our work to implement self-consistent meta-GGA support within ONETEP are presented alongside results demonstrating its numerical accuracy and computational efficiency. In section II, we describe the theoretical foundations of this work, with particular emphasis on the linear-scaling formalism used in ONETEP and the difficulties associated with self-consistent implementation of τ-dependent functionals. The theoretical innovations necessary to implement meta-GGAs within ONETEP's linear-scaling framework are described in section III. The meta-GGA functionals we implemented in ONETEP to validate our approach, PKZB [22] and B97M-V [19], are described in section III F. We present results concerning the numerical and computational performance of self-consistent meta-GGA calculations in ONETEP in section IV. Specifically, numerical compar-

isons with other codes are presented in section IV A, while the linear-scaling computational cost of our approach is verified in section IV B. Finally, in section V, we draw conclusions based on these results and make some suggestions for further work and potential future applications for meta-GGA functionals in ONETEP.

II. THEORY

A. Density-matrix DFT

KS-DFT represents the exact ground state electronic density via the density of an auxiliary system of independent particles [1], which interact via a mean-field. The principal advantage of using this auxiliary density is that the states of the independent electrons can be described by a set of one-electron orbitals, $\{\psi_i\}$. These Kohn-Sham (KS) orbitals are solutions to one-particle Schrödinger equations with eigenvalues, $\{\varepsilon_i\}$,

$$\left(-\frac{1}{2}\nabla^2 + V_{\mathrm{KS}}(\mathbf{r})\right)\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}). \tag{3}$$

where $V_{\mathrm{KS}}(\mathbf{r})$ is an effective local potential. This allows the obstacle of expressing the kinetic energy as an explicit functional of the density to be sidestepped, since the kinetic energy of a system of independent electrons has a simple form in terms of the KS orbitals:

$$T_{\mathrm{s}}[n] = -\frac{1}{2}\sum_{i=1}^{N}\int \mathrm{d}\mathbf{r}\psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r}). \tag{4}$$

The state of the independent electron system may be fully described by the one-particle density matrix

$$\rho(\mathbf{r},\mathbf{r}') = \sum_{i}^{N} f_i\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}') \tag{5}$$

with occupation numbers $\{f_i\}$ in the interval $[0,1]$, and from which the ground-state density is obtained from the density matrix by setting $\mathbf{r} = \mathbf{r}'$,

$$n(\mathbf{r}) = \rho(\mathbf{r},\mathbf{r}) = \sum_{i}^{N} f_i\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}). \tag{6}$$

The total energy functional of KS-DFT can be expressed in terms of the one-particle density matrix, by re-casting Eq. 4 in terms of $\rho(\mathbf{r},\mathbf{r}')$,

$$T_{\mathrm{s}}[n] = \int \mathrm{d}\mathbf{r}\left[-\frac{1}{2}\nabla_{\mathbf{r}}^2\rho(\mathbf{r},\mathbf{r}')\right]_{\mathbf{r}=\mathbf{r}'}, \tag{7}$$

where we have implicitly introduced a generalized definition of $T_{\mathrm{s}}[n]$, which allows for fractional occupancies, $\{f_i\}$ (as described in Ref. 23). The remaining external potential, Hartree and exchange-correlation contributions are explicit functionals of the density, and can be expressed in terms of $\rho(\mathbf{r},\mathbf{r}')$ using Eq. 6.

Formulating KS-DFT with the density matrix as the central quantity facilitates the development of methods which exploit the locality, or "nearsightedness" [7], of $\rho(\mathbf{r},\mathbf{r}')$ and opens the door to electronic structure methods in which computational cost scales linearly with system size, N. It is well-known that the density matrix for an insulating system (i.e. a system with a band-gap) decays exponentially with inter-electronic distance, $|\mathbf{r}-\mathbf{r}'|$ (see for example Ref. 24, and references therein). In practice, this locality can be enforced to yield $O(N)$ computational cost by truncating the density matrix such that $\rho(\mathbf{r},\mathbf{r}') = 0$ when $|\mathbf{r}-\mathbf{r}'| > r_{\mathrm{cut}}$, where r_{cut} is a predetermined cutoff distance. This truncation of the density matrix is the fundamental premise upon which ONETEP is designed.

B. ONETEP formalism

In ONETEP the density matrix is expressed in a separable form,

$$\rho(\mathbf{r},\mathbf{r}') = \varphi_\alpha(\mathbf{r})K^{\alpha\beta}\varphi_\beta^*(\mathbf{r}') \tag{8}$$

where $\{\varphi_\alpha\}$ are a set of strictly localized non-orthogonal orbitals, related to the KS orbitals by

$$\psi_i(\mathbf{r}) = \varphi_\alpha(\mathbf{r})M^\alpha{}_i\,. \tag{9}$$

and $\mathbf{K}$ is the density kernel,

$$K^{\alpha\beta} = \sum_{i=1}^{N} M^\alpha{}_i\, f_i M^\dagger{}_i{}^\beta, \tag{10}$$

constructed from elements of the transformation matrix $\mathbf{M}$ and occupation numbers $\{f_i\}$. Expressing the density matrix in separable form (Eq. 8) allows its intrinsic locality to be exploited by truncating the density kernel such that

$$R_{\alpha\beta} > r_{\mathrm{cut}} \implies K^{\alpha\beta} = 0 \tag{11}$$

where $R_{\alpha\beta}$ is the distance between the localized orbitals, $\varphi_\alpha(\mathbf{r})$ and $\varphi_\beta(\mathbf{r})$, and r_{cut} is a predetermined cutoff distance (the "density kernel cutoff") [8].

Note that in Eqs. 8 to 10 we have used the Einstein summation convention (implicit summation over repeated indices). This convention will be used for the remainder of this paper, though explicit summation signs will be used where this assists clarity. We have also distinguished between contravariant and covariant quantities by using upper and lower indices, as is conventional. The necessity of this distinction is a consequence of the non-orthogonality of ONETEP's local orbitals—care must be taken to respect the transformation properties of these orbitals and related objects.

The non-orthogonal orbitals in Eq. 8, called "non-orthogonal generalized Wannier functions" (NGWFs), are restricted to spherical, atom-centered localization regions, so $R_{\alpha\beta}$ in Eq. 11 is simply the distance between atoms. The NGWFs are constructed from an orthogonal basis of psinc

(periodic cardinal sine) functions positioned on a regular real-space grid [25–27], i.e.

$$\varphi_\alpha(\mathbf{r}) = \sum_{m}^{m\in L_\alpha} D(\mathbf{r}-\mathbf{r}_m)c_{m\alpha} \tag{12}$$

where each psinc, $D(\mathbf{r}-\mathbf{r}_m)$ is centered on a point of the real-space Cartesian grid, $\mathbf{r}_m$. Strict locality is enforced by only allowing psincs that fall within the localization region, L_α, to have non-zero $c_{m\alpha}$.

The psinc functions which form ONETEP's underlying basis can be thought of as a regular real-space grid of bandwidth-limited, periodic delta functions, and are constructed as a finite sum over plane waves—see appendix A of Ref. 25 for a detailed description of the psinc basis set. This relationship between the psinc basis and plane waves allows ONETEP to benefit from the useful properties of a plane wave basis, in particular the use of a single parameter, the kinetic energy cutoff, to systematically improve basis set quality. In addition, as in conventional plane wave DFT, it is simple to switch between real and reciprocal space representations of the basis by Fourier transform, enabling the most advantageous representation to be used for any given operation. To perform reciprocal-space operations while maintaining overall linear-scaling computational cost, ONETEP uses the "FFTbox" approximation, in which Fourier transforms are applied to a sub-region of the simulation cell, rather than the entire simulation cell—see Refs. 26, 28 for details.

A particular feature of the linear-scaling approach used in ONETEP is that the total energy is variationally minimized with respect to density kernel elements *and* the NGWFs [25, 29], using a conjugate gradients [30] scheme. This *in situ* optimization of the NGWFs allows ONETEP to achieve high-accuracy with a minimal set of NGWFs and has been shown to eliminate basis set superposition error [31].

The energy minimization takes place via two nested loops: in the inner loop, $\mathbf{K}$ is optimized with fixed NGWFs while in the outer loop, the NGWFs are varied, i.e.

$$E_{\min} = \min_{\{\varphi_\alpha\}} \left(\min_{\mathbf{K}} E[\mathbf{K},\{\varphi_\alpha\}] \right). \tag{13}$$

The procedure is repeated until a self-consistency threshold is satisfied, subject to the constraints that (i) the number of electrons in the system is unchanged, and (ii) the density kernel is idempotent.

Constraining the density kernel to be idempotent is equivalent to ensuring the orthogonality of the KS orbitals. This constraint is applied in ONETEP using a modified Li-Nunes-Vanderbilt (LNV) scheme [32, 33] (described in Ref. 34), where the total energy is minimized with respect to an auxiliary density kernel, $\mathbf{L}$, which is related to the true density kernel by the purifying transform

$$\mathbf{K} = 3\mathbf{LSL} - 2\mathbf{LSLSL}. \tag{14}$$

Provided that $\mathbf{L}$ is nearly idempotent, the minimization procedure naturally drives $\mathbf{K}$ towards idempotency. The requirement that the total number of electrons, N, does not change may be imposed by applying a constraint to the energy functional, for which several methods are available—see section 4 of Ref. 34 for details.

There are two key computational issues associated with the efficient implementation of this theoretical formalism in software:

1. Taking advantage of the sparsity of the density kernel, Hamiltonian and overlap matrices that naturally arises as a consequence of the strict locality of the NGWFs.

2. Efficient utilization of the available computational resources by distribution of computational work across parallel processing units.

From the beginning, ONETEP was designed to effectively address these practical issues, with efficient sparse matrix algebra routines [35] and parallel algorithms built around the Message Passing Interface (MPI) library [29]. More recently, shared-memory parallelism has been added using the OpenMP API [36] while support for graphics processing unit (GPU) acceleration is under development [37]. The new meta-GGA functionality described in this paper was developed to take full advantage of ONETEP's existing framework for sparse algebra and parallel computation.

C. Meta-GGA functionals

For local and semi-local exchange-correlation functionals, the exchange-correlation energy density, $\epsilon_{\mathrm{xc}}(\mathbf{r})$, at point $\mathbf{r}$ is determined entirely by the density and density-derived variables in the locality of $\mathbf{r}$ (Eq. 1). The distinction between the different classes of local/semi-local functionals lies in which variables the exchange-correlation energy density is constructed using, broadly categorized by rungs on the Jacob's ladder of density functional approximations (Fig. 1).

With each rung on the ladder, the introduction of new density-dependent ingredients allows the construction of more flexible and sophisticated functional forms, which correct for deficiencies in functionals on lower rungs. For example, a particular limitation of the LDAs is their tendency to overbind molecules, resulting in the overestimation of atomization energies. With the introduction of dependence on the gradient of the density, GGAs can dramatically improve upon LDA atomization energies, as well as many other properties (see, for example, Refs. 2, 38–40).

The meta-GGAs attempt to improve upon the LDAs and GGAs by introducing dependence on the kinetic energy density, τ (Eq. 2), i.e.

$$E_{\mathrm{xc}}^{\mathrm{MGGA}}[n] = \int \mathrm{d}\mathbf{r}\, \epsilon_{\mathrm{xc}}(n(\mathbf{r}), \nabla n(\mathbf{r}), \tau(\mathbf{r})). \tag{15}$$

The physical relevance of τ to the description of exchange and correlation is demonstrated by its appearance in expansions of both the exchange and correlation holes, as described in Refs. 41, 42. In practice, the addition of another density-derived ingredient offers greater flexibility, which may be used

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

to better satisfy formal constraints or fit empirical data (depending on the predilections of the designer).

The improvements offered by meta-GGAs over GGAs are typically less dramatic than those seen for GGAs over LDAs, but still significant. For example, the TPSS and SCAN meta-GGAs have been shown to consistently produce smaller errors than the PBE GGA [40, 43] for the calculation of a variety of properties [20, 44]. Since PBE, TPSS and SCAN are nonempirical functionals, the improvements seen for the meta-GGAs cannot be attributed to the quality of training data or fitting method, and must be a consequence of improvements in the functional form.

The addition of τ-dependence facilitates the construction of functional forms which satisfy more formal constraints. In particular, τ-dependence allows meta-GGAs to satisfy the constraint that the correlation energy of any one-electron density, n_1, is zero,

$$E_c[n_1, 0] = 0, \tag{16}$$

where E_c is a functional of two spin densities, i.e. $E_c[n_\uparrow, n_\downarrow]$. This can be achieved using expressions constructed from τ and the Weizsäcker kinetic energy density,

$$\tau^W = \frac{1}{8}\frac{|\nabla n|^2}{n}. \tag{17}$$

Since $\tau^W = \tau$ for a one-electron density and it can be shown that $\tau^W \leq \tau$ [45], the relationship between τ^W and τ can be used to construct correlation energy functionals which vanish for one-electron densities. For example, that $(\tau^W/\tau)^n = 1$ for a one-electron density is used in the PKZB [22] meta-GGA (section III F 1) to eliminate self-correlation (this form is also used in PKZB's successor, TPSS [44]). Similar self-correlation correcting expressions have been employed in the construction of many density functionals, including examples by Becke [42, 46, 47] and Van Voorhis and Scuseria [48] which precede PKZB's 1999 publication. In addition, related τ-containing expressions such as the electron localization function (ELF) [49] and localized orbital locator (LOL) [50] have been described which distinguish regions of electron localization, and can be used to produce striking visualizations of atomic and molecular electronic structure.

Many variants of the meta-GGA form have been developed since Becke's early work in the 1980s [42, 51, 52]. Some notable variants have already been mentioned, such as PKZB [22] and its nonempirical successor TPSS [44]. Another notable set of meta-GGAs are the semiempirical "Minnesota" functionals, which include local M06-L [53], M11-L [54] and the recent MN15-L [55], as well as hybrid variants which add exact exchange to the meta-GGA form (for example, M06 [56], M11 [57] and MN15 [58]). The development and refinement of meta-GGA functionals is an ongoing effort—recently published examples include the semiempirical B97M-V (which we implemented in this work, see section III F 2), the related range-separated hybrid meta-GGA, ωB97M-V [59], and nonempirical SCAN [20] functionals. For further examples and historical context, see the well-cited accounts in Refs 19, 60.

D. Self-consistent meta-GGA evaluation

In order to self-consistently solve the Kohn-Sham equations (Eq. 3), it is necessary to evaluate the effective potential, $V_{KS}(\mathbf{r})$, which includes an exchange-correlation contribution, defined as the functional derivative of the exchange-correlation energy with respect to the density,

$$V_{xc}(\mathbf{r}) = \frac{\delta E_{xc}}{\delta n(\mathbf{r})}. \tag{18}$$

For LDA and GGA functionals, the exchange-correlation energy density is an explicit functional of the charge density, n and its gradient, ∇n, and obtaining the functional form of V_{xc} is a simple matter of evaluating the functional derivative.

The τ-dependence of meta-GGA functionals (Eq. 15) presents an additional challenge because τ is itself an implicit functional of the density. Since the explicit dependence of τ on n is unknown, the functional derivative of Eq. 18 cannot be straightforwardly evaluated in the case of meta-GGAs.

One approach to this issue is to simply avoid ever needing to evaluate the meta-GGA exchange-correlation potential. This can be achieved by self-consistently determining the KS orbitals using a non-τ-dependent functional, and then evaluating the meta-GGA energy expression non-self-consistently using n and τ constructed from these orbitals. This method was used in Refs. 22, 61, and reportedly yields results close to those produced by self-consistent meta-GGA calculations. However, the method has limited utility beyond total energy evaluation, since the orbitals and density produced are completely determined by the non-τ-dependent functional used.

A second approach is to use the optimized effective potential (OEP) method [62]. The OEP method produces local, multiplicative exchange-correlation potentials for orbital-dependent functionals and is based on the premise of finding the local potential for which the total energy is stationary, i.e. $\delta E/\delta V_{KS}(\mathbf{r}) = 0$. The local potential which satisfies this requirement can be obtained by solution of an integral equation. This is a non-trivial task and approximate methods are generally used (as in the Krieger-Li-Iafrate (KLI) approach [63]). Although the OEP method is more often applied to functionals incorporating exact exchange, it is also applicable to meta-GGAs, as described in Refs. 64, 65. For a well-cited review of the OEP method and approximations to this, see Ref. 60.

A third approach, which avoids the theoretical and computational complexity of solving the OEP integral equation is to express the exchange-correlation potential in terms of functional derivatives with respect to the orbitals, rather than the density. Originally described by Neumann, Nobes and Handy [66] in the context of self-consistently evaluating the τ-dependent Becke-Roussel exchange functional [52], this approach yields a non-multiplicative, orbital-specific potential. Following Ref. 64, we refer to this approach as the "functional derivatives of τ-dependent functionals with respect to the orbitals", or FDO, approach.

For this work, we adopted the FDO approach, motivated by its relative simplicity compared to the OEP method. As demonstrated in section III, the FDO approach allows self-consistent calculations to be performed using τ-dependent

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

exchange-correlation functionals with only minor extensions to the existing density-matrix DFT framework used by ONETEP.

The FDO approach has previously been described as an "approximation", since its orbital-specific potential represents a step outside of the conventional Kohn-Sham method, which assumes a local effective potential (as in Ref. 64). While the FDO method could be considered an approximation to KS-DFT *with a local potential*, this terminology could be misleading since the results obtained using this method are not necessarily any less accurate or rigorous than those obtained using the OEP method. The FDO approach falls within the wider framework of the "generalized Kohn-Sham" (gKS) formalism [67], in which alternate realizations of DFT may be constructed based on auxiliary systems other than the usual non-interacting system of conventional KS-DFT. Such theories permit non-local effective potentials, as in the FDO method.

Since the exchange-correlation part of the effective potential in the FDO approach is non-multiplicative, we would expect the solutions to the corresponding eigenvalue equations (Eq. 3) to differ from those obtained using a local potential. It has been observed that the differences between the FDO and OEP methods for meta-GGAs are very small for total energies, though these can be more significant for quantities sensitive to electronic structure, such as nuclear shielding constants [64]. Additionally, recent work by Yang and coworkers has shown that band gaps for the SCAN meta-GGA [20] are improved using an FDO-type approach compared to OEP [65].

The FDO approach is perhaps the most widely used method of evaluating meta-GGA exchange-correlation potentials (see for example Refs. 44, 48, 68, 69). The widespread use of this approach is a further practical advantage, since in validating our work against other implementations which use the FDO method, we do not need to consider any error arising from the difference between the FDO and OEP approaches. In section IV A, we compare self-consistent meta-GGA energies calculated using ONETEP with those calculated in Q-Chem [70], both using the FDO approach [71].

In the FDO method, the functional derivative of the exchange-correlation energy with respect to the orbitals may be arrived at using the functional derivative chain rule [72]

$$\frac{\delta E_{\mathrm{xc}}^{\mathrm{mGGA}}}{\delta \psi_i(\mathbf{r})} = \int \mathrm{d}\mathbf{r}' \, \frac{\delta E_{\mathrm{xc}}^{\mathrm{mGGA}}}{\delta n(\mathbf{r}')} \frac{\delta n(\mathbf{r}')}{\delta \psi_i(\mathbf{r})} \tag{19}$$

and inserting the functional derivative

$$\frac{\delta n(\mathbf{r}')}{\delta \psi_i(\mathbf{r})} = 2\psi_i(\mathbf{r})\delta(\mathbf{r}' - \mathbf{r}), \tag{20}$$

to give

$$\frac{\delta E_{\mathrm{xc}}^{\mathrm{mGGA}}}{\delta \psi_i(\mathbf{r})} = 2\frac{\delta E_{\mathrm{xc}}^{\mathrm{mGGA}}}{\delta n(\mathbf{r})}\psi_i(\mathbf{r}). \tag{21}$$

The relationship in Eq. 21 implies the following form for the product of the exchange-correlation potential and an orbital:

$$\frac{1}{2}\frac{\delta E_{\mathrm{xc}}^{\mathrm{mGGA}}}{\delta \psi_i(\mathbf{r})} = \frac{\delta E_{\mathrm{xc}}^{\mathrm{mGGA}}}{\delta n(\mathbf{r})}\psi_i(\mathbf{r}) = \hat{V}_{\mathrm{xc}}^{\mathrm{FDO}}\psi_i(\mathbf{r}), \tag{22}$$

where we have used operator notation to emphasize the non-multiplicative nature of the FDO exchange-correlation potential, $\hat{V}_{\mathrm{xc}}^{\mathrm{FDO}}$. Evaluating the functional derivative on the left of Eq. 22 for a general meta-GGA dependent on n, ∇n and τ (Eq. 15), yields the FDO exchange-correlation potential [64, 68]

$$\begin{aligned}\frac{1}{2}\frac{\delta E_{\mathrm{xc}}^{\mathrm{mGGA}}}{\delta \psi_i(\mathbf{r})} = &\left\{\frac{\partial \epsilon_{\mathrm{xc}}}{\partial n} - \nabla\cdot\left(\frac{\partial \epsilon_{\mathrm{xc}}}{\partial \nabla n}\right)\right\}\psi_i(\mathbf{r}) \\ &- \frac{1}{2}\left(\nabla\frac{\partial \epsilon_{\mathrm{xc}}}{\partial \tau}\right)\cdot\nabla\psi_i(\mathbf{r}) - \frac{1}{2}\frac{\partial \epsilon_{\mathrm{xc}}}{\partial \tau}\nabla^2\psi_i(\mathbf{r}).\end{aligned} \tag{23}$$

A more compact expression of the FDO exchange-correlation potential is

$$\hat{V}_{\mathrm{xc}}^{\mathrm{FDO}}(\mathbf{r})\psi_i(\mathbf{r}) = V_{\mathrm{xc}}^{\mathrm{GGA}}\psi_i(\mathbf{r}) + \left\{\hat{V}_{\mathrm{xc}}^{\tau}\psi_i(\mathbf{r})\right\} \tag{24}$$

which consists of a GGA-like part,

$$V_{\mathrm{xc}}^{\mathrm{GGA}} = \frac{\partial \epsilon_{\mathrm{xc}}}{\partial n} - \nabla\cdot\left(\frac{\partial \epsilon_{\mathrm{xc}}}{\partial \nabla n}\right), \tag{25}$$

and a non-multiplicative, τ-dependent part,

$$\hat{V}_{\mathrm{xc}}^{\tau}\psi_i(\mathbf{r}) = -\frac{1}{2}\left(\nabla\frac{\partial \epsilon_{\mathrm{xc}}}{\partial \tau}\right)\cdot\nabla\psi_i(\mathbf{r}) - \frac{1}{2}\frac{\partial \epsilon_{\mathrm{xc}}}{\partial \tau}\nabla^2\psi_i(\mathbf{r}). \tag{26}$$

Using the relationship in Eq. 22 to derive an exchange-correlation potential means that the FDO exchange-correlation potential is only defined where $\delta E_{\mathrm{xc}}/\delta n$ appears as a product with a KS orbital. We have emphasized this relationship with the KS orbital that appears in $\delta E_{\mathrm{xc}}/\delta\psi_i$ by enclosing the non-multiplicative part of Eq. 24 in curly braces. Fortunately, in conventional self-consistent DFT calculations the exchange-correlation potential only arises as a product with a KS orbital in the evaluation of Hamiltonian matrix elements. As we shall see, this is also the case with the direct energy minimization approach used in ONETEP, though in this case a per-NGWF form of the exchange-correlation potential is used (section III B) and this additionally appears in expressions for the gradient of the energy (sections III D and III E).

Note that in the above, we have referred to the orbitals which solve the one-electron eigenvalue equations (Eq. 3) in both conventional and generalized Kohn-Sham (i.e. FDO) approaches as "KS orbitals". We will continue with this convention for the remainder of the paper, since this clearly distinguishes the orbitals which solve the one-electron equations from the localized orbitals (NGWFs) used in ONETEP.

III. IMPLEMENTATION

A. Kinetic energy density

To evaluate $\tau(\mathbf{r})$ within ONETEP's linear-scaling DFT formalism (section II B), it is necessary to recast Eq. 2 in terms of local quantities. This is easily achieved by expanding the KS orbitals in Eq. 2 in terms of the NGWFs (Eq. 9),

$$\tau(\mathbf{r}) = \frac{1}{2}\sum_{i}^{N_{\mathrm{occ}}} f_i \left(\nabla\varphi_\alpha(\mathbf{r})M^{\alpha}{}_{i}\right)\cdot\left(\nabla M^{\dagger}{}_{i}^{\beta}\varphi^*_\beta(\mathbf{r})\right), \tag{27}$$

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

and inserting the definition of the density kernel in terms of the orbital coefficients (Eq. 10)

$$\tau(\mathbf{r}) = \frac{1}{2}\sum_{\alpha\beta}(\nabla\varphi_\alpha(\mathbf{r}))\cdot\left(K^{\alpha\beta}\nabla\varphi_\beta(\mathbf{r})\right). \tag{28}$$

The introduction of fractional occupation numbers, $\{f_i\}$, in Eq. 27, follows from the generalized definition of the kinetic energy in terms of the density matrix (Eq. 7).

In practice, $\tau(\mathbf{r})$ is evaluated on the simulation cell grid in a manner analogous to $n(\mathbf{r})$,

$$n(\mathbf{r}) = \varphi_\alpha(\mathbf{r})K^{\alpha\beta}\varphi^*_\beta(\mathbf{r}), \tag{29}$$

described in detail in Ref. 29. This starts with the evaluation of "per-NGWF" contributions

$$\tau(\mathbf{r};\alpha) = (\nabla\varphi_\alpha(\mathbf{r}))\cdot\left(\nabla\sum_\beta K^{\alpha\beta}\varphi_\beta(\mathbf{r})\right). \tag{30}$$

where the summation over β is for all NGWFs which overlap with φ_α. It is convenient to perform the summation over β before applying the gradient operator, since this is easily applied to the summed quantity in reciprocal space, avoiding the need to individually calculate the gradient of each β NGWF. The per-NGWF contributions are then summed to yield the quantity on the full simulation cell grid,

$$\tau(\mathbf{r}) = \frac{1}{2}\sum_\alpha \tau(\mathbf{r};\alpha). \tag{31}$$

For completeness, we note that the evaluation of $\tau(\mathbf{r})$ by this method introduces a subtle approximation, not present for the same method applied to $n(\mathbf{r})$ [29]. For $n(\mathbf{r})$, summing over only NGWFs which overlap with φ_α to form $n(\mathbf{r};\alpha)$ (as in Eq. 30) is exact, since the NGWFs are zero-valued outside their localization regions. When the gradient operator is applied to a NGWF, however, it becomes delocalized. This is a consequence of the form of the psinc basis functions from which the NGWFs are constructed (Eq. 12)—while the psincs are zero-valued at the grid points they are not centered upon, they do not in general have zero gradient at these points. We have not explored the effect of this approximation in this work, but note that a similar approximation is assumed in ONETEP for the evaluation of the kinetic energy integrals over NGWFs (where the operator is the Laplacian, rather than the gradient) and is expected to have a negligible effect in that context [28].

B. Exchange-correlation potential

As with the kinetic energy density (section III A), the exchange-correlation potential must be expressed in a form containing only local quantities in order to be compatible with ONETEP's linear-scaling formalism (section II B). For LDA and GGA-type functionals, this is trivial, since the form of the exchange-correlation potential, as determined by the functional derivative $\delta E_{\mathrm{xc}}[n]/\delta n(\mathbf{r})$, is already a local potential. For meta-GGAs using the FDO method, the situation is more complicated, with a non-multiplicative "orbital-specific" potential, which is defined with respect to a KS orbital (Eqs. 22 and 23)—this presents a potential issue for ONETEP, since we must avoid explicit reference to the KS orbitals. Fortunately, it is simple to derive a "per-NGWF" potential, which avoids the need to deal directly with the KS orbitals.

Expanding the KS orbital in Eq. 23 in terms of NGWFs (Eq. 9) and using Eq. 22, we obtain

$$\begin{aligned}\sum_\alpha \frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})}\varphi_\alpha(\mathbf{r})M^\alpha{}_i \\ = \sum_\alpha V_{\mathrm{xc}}^{\mathrm{GGA}}\varphi_\alpha(\mathbf{r})M^\alpha{}_i + \hat{V}^\tau_{\mathrm{xc}}\varphi_\alpha(\mathbf{r})M^\alpha{}_i,\end{aligned} \tag{32}$$

where the GGA-like and τ-dependent parts of the potential are defined as in Eqs. 25 and 26 (but with ψ_i replaced by φ_α). Recognizing that the orbital coefficients, $M^\alpha{}_i$, commute with the non-multiplicative part of the potential, Eq. 32 can be rearranged to give

$$\sum_\alpha\left[\frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})}\varphi_\alpha(\mathbf{r}) - V_{\mathrm{xc}}^{\mathrm{GGA}}\varphi_\alpha(\mathbf{r}) - \left\{\hat{V}^\tau_{\mathrm{xc}}\varphi_\alpha(\mathbf{r})\right\}\right]M^\alpha{}_i = 0. \tag{33}$$

which implies a per-NGWF exchange-correlation potential

$$\frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})}\varphi_\alpha(\mathbf{r}) = \hat{V}_{\mathrm{xc}}^{\mathrm{FDO}}\varphi_\alpha(\mathbf{r}) = V_{\mathrm{xc}}^{\mathrm{GGA}}\varphi_\alpha(\mathbf{r}) + \left\{\hat{V}^\tau_{\mathrm{xc}}\varphi_\alpha(\mathbf{r})\right\} \tag{34}$$

with the same local GGA-like, and non-multiplicative τ-dependent parts seen in Eqs. 25 and 26 (but with ψ_i replaced by φ_α, as before).

Applying the product rule, the τ-dependent part can be expressed in a more compact form,

$$\left\{\hat{V}^\tau_{\mathrm{xc}}\varphi_\alpha(\mathbf{r})\right\} = -\frac{1}{2}\nabla\cdot\left(\frac{\partial\epsilon_{\mathrm{xc}}}{\partial\tau}\nabla\varphi_\alpha(\mathbf{r})\right), \tag{35}$$

which proves useful in evaluating integrals over the $\hat{V}^\tau_{\mathrm{xc}}$ (section III C).

C. Integrals

The self-consistent implementation of meta-GGAs under the FDO method requires integrals over the τ-dependent part of the exchange-correlation potential (Eq. 35), in addition to the usual integrals over the local, GGA-like part of the exchange-correlation potential (Eq. 25).

The form of the τ-dependent matrix element used in Neumann, Nobes and Handy's paper [66] and other later publications (e.g. Ref. 48) is

$$\langle\psi_i|\hat{V}^\tau_{\mathrm{xc}}|\psi_j\rangle = \frac{1}{2}\int \mathrm{d}\mathbf{r}\,\nabla\psi_i(\mathbf{r})\cdot\left(\frac{\partial\epsilon_{\mathrm{xc}}}{\partial\tau}\nabla\psi_j(\mathbf{r})\right) \tag{36}$$

which may be arrived at from Eq. 35 via integration-by-parts. This form of the integral has the apparent advantage that

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

the derivative of the potential $\partial\epsilon_{\rm xc}/\partial\tau$ is not required—only derivatives of the orbitals are needed. This is an attractive property for a Gaussian basis set code, since analytic derivatives of the basis functions (and hence orbitals) are easily accessible and the need to evaluate the numerical derivative of the potential is avoided.

The same result may be arrived at for NGWFs, when integrated over a unit cell with periodic boundary conditions (via the divergence theorem), i.e.

$$\langle\varphi_\alpha|\hat{V}^\tau_{\rm xc}|\varphi_\beta\rangle = \frac{1}{2}\int_V {\rm d}\mathbf{r}\,\nabla\varphi_\alpha(\mathbf{r})\cdot\left(\frac{\partial\epsilon_{\rm xc}}{\partial\tau}\nabla\varphi_\beta(\mathbf{r})\right), \quad (37)$$

where V is the volume of the simulation cell. However, for ONETEP, this form of the integral is problematic because the gradients of psinc functions (and by extension, NGWFs) are delocalized (as described in section III A).

For integrals over the local potential and kinetic energy operator,

$$V_{\alpha\beta} = \int_V {\rm d}\mathbf{r}\,\varphi_\alpha(\mathbf{r})V_{\rm loc}(\mathbf{r})\varphi_\beta(\mathbf{r}), \quad (38)$$

$$T_{\alpha\beta} = -\frac{1}{2}\int_V {\rm d}\mathbf{r}\,\varphi_\alpha(\mathbf{r})\nabla^2\varphi_\beta(\mathbf{r}), \quad (39)$$

the locality of the NGWFs naturally restricts the integration region. In $V_{\alpha\beta}$, the integrand is zero where the α and β NGWFs do not overlap, and thus it can be assumed that $V_{\alpha\beta} = 0$ where the localization regions of φ_α and φ_β do not overlap. Similarly for $T_{\alpha\beta}$, the integrand is always zero outside the localization region of the α NGWF, and the approximation is made that the $T_{\alpha\beta} = 0$ where there is no overlap between the localization regions of φ_α and φ_β (as described in Ref. 28). For Eq. 37 the delocalization of the NGWF gradients prevents the restriction of the integration to the localization sphere of either NGWF, necessitating a further approximation in addition to the assumption that the integral is zero where the localization regions of φ_α and φ_β are non-overlapping.

To make better use of the locality of the NGWFs, we use an alternative form of the integral over the τ-dependent part of the exchange-correlation potential, based on Eq. 35,

$$\langle\varphi_\alpha|\hat{V}^\tau_{\rm xc}|\varphi_\beta\rangle = -\frac{1}{2}\int_V {\rm d}\mathbf{r}\,\varphi_\alpha(\mathbf{r})\nabla\cdot\left(\frac{\partial\epsilon_{\rm xc}}{\partial\tau}\nabla\varphi_\beta(\mathbf{r})\right), \quad (40)$$

which has the advantage that the integrand is non-zero only in the localization region of φ_α. The need to evaluate the divergence of $(\partial\epsilon_{\rm xc}/\partial\tau)\nabla\varphi_\beta$ is not a significant issue in ONETEP, since the gradient operator can be trivially applied to each Cartesian component of the vector field in reciprocal space.

The form of Eq. 40 closely resembles the form of the kinetic energy integrals (Eq. 39). Indeed, for $\partial\epsilon_{\rm xc}/\partial\tau = 1$, the two integral types are equivalent—a property we used in testing the code to evaluate $\langle\varphi_\alpha|\hat{V}^\tau_{\rm xc}|\varphi_\beta\rangle$. As for the kinetic energy integrals [28] (and kinetic energy density, section III A), the evaluation of $\langle\varphi_\alpha|\hat{V}^\tau_{\rm xc}|\varphi_\beta\rangle$ in ONETEP makes the approximation that integrals over NGWFs with non-overlapping localization spheres are zero, despite the delocalizing effect of applying the gradient operator to psinc functions.

D. Density kernel optimization

Direct minimization of the energy under the LNV scheme used in ONETEP (section II B) requires the evaluation of the gradient of the energy with respect to the elements of the auxiliary density matrix, $\mathbf{L}$. This gradient is given by

$$\frac{\partial E}{\partial L^{\alpha\beta}} = \frac{\partial E}{\partial K^{\eta\theta}}\frac{\partial K^{\eta\theta}}{\partial L^{\alpha\beta}}. \quad (41)$$

where $\partial K^{\eta\theta}/\partial L^{\alpha\beta}$ can be derived from Eq. 14 (see Ref. 73). For the purposes of this work, it is sufficient to note that $\partial K^{\eta\theta}/\partial L^{\alpha\beta}$ does not depend on the form of $E_{\rm xc}[n]$ and we can therefore focus on the form of the $\partial E/\partial K^{\eta\theta}$.

The exchange-correlation component $\partial E/\partial K^{\eta\theta}$ can be expressed in terms of the functional derivative of $E_{\rm xc}[n]$ with respect to the density,

$$\frac{\partial E_{\rm xc}}{\partial K^{\eta\theta}} = \int {\rm d}\mathbf{r}\,\frac{\delta E_{\rm xc}}{\delta n(\mathbf{r})}\frac{\partial n(\mathbf{r})}{\partial K^{\eta\theta}}, \quad (42)$$

and thus the partial derivative with respect to the density kernel elements depends on the form of the exchange-correlation potential, $V_{\rm xc}(\mathbf{r}) \equiv \delta E_{\rm xc}/\delta n(\mathbf{r})$. For reference, the derivation of Eq. 42 is reproduced in appendix A.

Using Eq. 29, the partial derivative on the right of Eq. 42 is

$$\frac{\partial n(\mathbf{r})}{\partial K^{\eta\theta}} = \delta^\alpha_\eta\delta^\beta_\theta\varphi_\alpha(\mathbf{r})\varphi^*_\beta(\mathbf{r}) \quad (43)$$

and thus Eq. 42 becomes

$$\frac{\partial E_{\rm xc}}{\partial K^{\eta\theta}} = \int {\rm d}\mathbf{r}\,\varphi^*_\theta(\mathbf{r})\frac{\delta E_{\rm xc}}{\delta n(\mathbf{r})}\varphi_\eta(\mathbf{r}). \quad (44)$$

For LDA or GGA functionals the functional derivative is equivalent to the local exchange-correlation potential, and Eq. 44 is simply a matrix element over the local potential:

$$\frac{\partial E_{\rm xc}}{\partial K^{\eta\theta}} = \langle\varphi_\theta|V_{\rm xc}|\varphi_\eta\rangle. \quad (45)$$

Under the FDO approach (section II D), the meta-GGA exchange-correlation potential consists of a local GGA-like part (Eq. 25) and non-multiplicative, τ-dependent part (Eq. 35). Inserting the per-NGWF exchange-correlation potential (section III B) into Eq. 44 yields

$$\begin{aligned}\frac{\partial E_{\rm xc}}{\partial K^{\eta\theta}} &= \int {\rm d}\mathbf{r}\,\varphi^*_\theta(\mathbf{r})V^{\rm GGA}_{\rm xc}(\mathbf{r})\varphi_\eta(\mathbf{r}) \\ &\quad + \int {\rm d}\mathbf{r}\,\varphi^*_\theta(\mathbf{r})\{\hat{V}^\tau_{\rm xc}\varphi_\eta(\mathbf{r})\} \\ &= \langle\varphi_\theta|V^{\rm GGA}_{\rm xc}|\varphi_\eta\rangle + \langle\varphi_\theta|\hat{V}^\tau_{\rm xc}|\varphi_\eta\rangle,\end{aligned} \quad (46)$$

the form of the exchange-correlation component of $\partial E/\partial K^{\eta\theta}$ for meta-GGAs under the FDO approach.

The implementation of Eq. 46 in ONETEP was a simple matter of adding contributions from the additional matrix element, $\langle\varphi_\theta|\hat{V}^\tau_{\rm xc}|\varphi_\eta\rangle$, to the existing gradient expression for GGA-type functionals.

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

E. NGWF optimization

In addition to the gradient of the energy with respect to density kernel elements described in section III D, direct minimization of the energy in ONETEP also requires the evaluation of the gradient of the energy with respect to the NGWFs.

Since the total energy is a functional of the NGWFs (via the density, Eq. 29), the gradient of the energy with respect to the NGWFs is a functional derivative. This functional derivative has the form

$$\frac{\delta E_{\mathrm{xc}}}{\delta \varphi_{\alpha}^{*}(\mathbf{r})} = \varphi_{\kappa}(\mathbf{r})[\mathbf{Q}^{V}]^{\kappa\alpha} + \frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})}\varphi_{\kappa}(\mathbf{r})K^{\kappa\alpha} \tag{47}$$

where

$$\begin{aligned}[\mathbf{Q}^{V}]^{\kappa\alpha} &= 3L^{\kappa\nu}V_{\nu\mu}L^{\mu\alpha} \\ &\quad - 2L^{\kappa\gamma}S_{\gamma\delta}L^{\delta\nu}V_{\nu\mu}L^{\mu\alpha} \\ &\quad - 2L^{\kappa\nu}V_{\nu\mu}L^{\mu\gamma}S_{\gamma\delta}L^{\delta\alpha}\end{aligned} \tag{48}$$

and

$$V_{\nu\mu} = \int \mathrm{d}\mathbf{r}\, \varphi_{\nu}^{*}(\mathbf{r})\frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})}\varphi_{\mu}(\mathbf{r}). \tag{49}$$

For reference, a derivation of Eq. 47 is presented in appendix B (based on work presented in Ref. 73). Note that the result differs by a factor of two depending on whether the NGWFs are real or complex, and this is covered in appendix B. In this section, we only consider the case of complex NGWFs.

Inserting the per-NGWF FDO exchange-correlation potential (Eq. 34) into Eq. 47 yields

$$\begin{aligned}\frac{\delta E_{\mathrm{xc}}}{\delta \varphi_{\alpha}^{*}(\mathbf{r})} &= \varphi_{\kappa}(\mathbf{r})[\mathbf{Q}^{V,\mathrm{FDO}}]^{\kappa\alpha} \\ &\quad + V_{\mathrm{xc}}^{\mathrm{GGA}}(\mathbf{r})\varphi_{\kappa}(\mathbf{r})K^{\kappa\alpha} + \{\hat{V}_{\mathrm{xc}}^{\tau}\varphi_{\kappa}(\mathbf{r})\}K^{\kappa\alpha}.\end{aligned} \tag{50}$$

where $\mathbf{Q}^{V,\mathrm{FDO}}$ is defined as in Eq. 48, but with $V_{\nu\mu}$ replaced by

$$\begin{aligned}V_{\nu\mu}^{\mathrm{FDO}} &= \int \mathrm{d}\mathbf{r}\, \varphi_{\nu}^{*}(\mathbf{r})V_{\mathrm{xc}}^{\mathrm{GGA}}(\mathbf{r})\varphi_{\mu}(\mathbf{r}) \\ &\quad + \int \mathrm{d}\mathbf{r}\, \varphi_{\nu}^{*}(\mathbf{r})\left\{\hat{V}_{\mathrm{xc}}^{\tau}(\mathbf{r})\varphi_{\mu}(\mathbf{r})\right\}.\end{aligned} \tag{51}$$

As with the gradient of the energy with respect to density kernel elements, the implementation of Eq. 50 in ONETEP was achieved by the modification of the existing gradient expression for GGA-type functionals. This involved adapting the routines used to construct $\mathbf{Q}^{V}$ to use $V_{\nu\mu}^{\mathrm{FDO}}$ in place of the usual local potential integral, and adding an extra term to the GGA gradient expression, $\{\hat{V}_{\mathrm{xc}}^{\tau}\varphi_{\kappa}(\mathbf{r})\}K^{\kappa\alpha}$.

F. Functionals

1. PKZB

The PKZB meta-GGA was first described in 1999 [22, 74] and, following the design principles of the PBE GGA [40, 43], was constructed to satisfy a number of known exact constraints on $E_{\mathrm{xc}}[n]$ with only a single empirically fitted parameter (see Ref. 3 for further details of this constraint-satisfaction approach). Although PKZB performed well compared to contemporary semi-local functionals (see e.g. Ref. 45), it has since been superseded by more modern meta-GGAs such as the TPSS meta-GGA [44, 75, 76], a non-empirical refinement of PKZB.

While PKZB has been surpassed by more modern meta-GGAs, it has the advantage of theoretical simplicity. Our implementation of PKZB to test the self-consistent meta-GGA framework described above was primarily motivated by this simplicity. In particular, the forms of the exchange and correlation parts of the functional are closely related to the PBE GGA, which was already available and well-tested in ONETEP. This allowed PKZB to be relatively simply implemented by reuse and modification of existing code.

A further advantage of using PKZB as a testbed for meta-GGAs in ONETEP was its relatively wide availability in other codes. Since the publication of the original PKZB paper in 1999, there has been ample time for the creation of mature, stable implementations of the functional. We used this to our advantage in the implementation of this functional and the general framework for meta-GGAs in ONETEP, performing preliminary tests of our work against independent implementations in Libxc [77], NWChem [78] and Q-Chem [70]. A selection of results, validating our self-consistent implementation of PKZB against Q-Chem are reported in section IV A.

While the key equations required to evaluate energies using the PKZB functional are well described in the original paper [22], equations for the derivatives of the exchange and correlation energy density, $\epsilon_{\mathrm{xc}}(\mathbf{r})$, necessary to evaluate the exchange-correlation potential (sections II C and III B) are not provided. For convenience and ease of reproducing our work, the equations we derived for the derivatives of $\epsilon_{\mathrm{xc}}(\mathbf{r})$ are presented in appendix C, alongside reproductions of the original energy equations.

2. B97M-V

B97M-V [19] is a semi-local meta-GGA density functional with VV10 non-local correlation [79] designed by Mardirossian and Head-Gordon via a combinatorial approach [80]. In the spirit of the original B97 density functional [81], the semi-local part of B97M-V is partitioned into three components: exchange, same-spin correlation, and opposite-spin correlation. Consequently, B97M-V has three separate inhomogeneity correction factors (ICF) that enhance the three associated uniform electron gas energy densities. Each ICF is a two-dimensional power series utilizing two variables: u, a finite-domain dimensionless variable that depends on the density and its gradient [81], and w, a finite-domain dimensionless variable that depends on the density and the kinetic energy density [82]. The final functional form of B97M-V was selected by searching through more than 10^{10} of the over 10^{40} possible options and choosing the most transferable and well-behaved fit. B97M-V has a total of 12 linear parameters—4

for exchange, 4 for same-spin correlation, and 4 for opposite-spin correlation. For non-covalent interactions and isomerization energies that are not significantly affected by self-interaction error, B97M-V is much more accurate than other semi-local functionals such as B97-D3(BJ) [83–85], M06-L [53], and MN15-L [55], and competitive with the best hybrid functionals such as ωB97X-V [86] and ωB97M-V [59]. As an example of its performance, across a database of 1744 closed-shell non-covalent dimer binding energies, B97M-V affords an RMSD of 0.24 kcal mol^{-1}, while the three aforementioned semi-local functionals manage RMSDs of 0.47, 0.55, and 1.38 kcal mol^{-1}, respectively. Furthermore, across a database of 755 isomerization energies, B97M-V has an RMSD of 0.27 kcal mol^{-1}, compared to 0.78, 0.71, and 1.55 kcal mol^{-1}, respectively, for B97-D3(BJ), M06-L, and MN15-L [59].

As mentioned in section I, linear-scaling DFT methods are particularly sensitive to the balance of computational cost and accuracy. To produce meaningful results for very large chemical systems, exchange and correlation must be treated with accurate methods, but without compromising the computational scaling of the method such that calculations become intractable. Our motivation for implementing B97M-V was its potential to satisfy both of these requirements, sitting firmly in the "Goldilocks zone" of Fig. 1.

A domain in which we are particularly interested in applying B97M-V is the description of biomolecular association. ONETEP's linear-scaling DFT approach enables a full quantum-mechanical description of a protein-ligand system [14, 15], avoiding the need to partition the system into quantum and classical regions, as in competing QM/MM approaches. We hope that the combination of this capability with B97M-V's excellent accuracy in describing non-covalent interactions will yield a powerful tool for the study of protein-ligand binding, applicable to the challenging problem of computational drug optimization. For non-covalent interaction energies, B97M-V has been shown to significantly outperform several dispersion corrected GGAs (see Refs. 19, 59). For biological systems, in which non-covalent interactions often play an important role, we anticipate the use of B97M-V will yield improved accuracy over the dispersion-corrected GGA functionals that have been employed in previous studies using ONETEP (see for example, Refs. 14, 15). In practice, the calculation of free energies of binding using B97M-V would combine some kind of configurational sampling with the evaluation of the energies of configurations using DFT, as in the QM-PBSA [14, 87, 88] method or the "stepping stone" approach of Sampson and coworkers [89].

It could be argued that B97M-V is not a "pure" meta-GGA, because it incorporates a contribution from the VV10 non-local correlation functional [79], in order to account for dispersion interactions. VV10 and similar non-local correlation functionals in the vdW-DF family [90, 91] are not easily classified within the traditional Jacob's ladder hierarchy because their non-locality is a consequence of their functional form, rather than dependence on KS orbitals. These functionals take the form of a double integral

$$E_{\mathrm{c}}^{\mathrm{nl}}[n] = \int \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 n(\mathbf{r}_1)\Phi(\mathbf{r}_1,\mathbf{r}_2)n(\mathbf{r}_2) \qquad (52)$$

where $\Phi(\mathbf{r}_1,\mathbf{r}_2)$ is a kernel depending on the interelectronic distance, $\mathbf{r}_1 - \mathbf{r}_2$. Klimeš and Michaelides have proposed a "stairway to heaven" of DFT-based dispersion corrections, analogous to the Jacob's ladder (Fig. 1) with the vdW-DF family (including VV10) on the third step [92].

For clarity, we note that "VV10" is often used to refer to the combination of specific semi-local exchange and correlation functionals with the VV10 non-local correlation functional, as described in Ref. 79. We will distinguish between the non-local correlation component and the combined exchange-correlation functional by referring to the non-local part as "VV10 NLC".

The implementation of VV10 NLC in ONETEP follows the approach described in Ref. 93, which slightly modifies the original VV10 NLC kernel in order to make use of the efficient integration scheme for vdW-DF functionals described by Román-Pérez and Soler [94] (for details of the implementation of the vdW-DF functionals in ONETEP, see Ref. 95). To obtain good agreement between this "revised VV10" (rVV10), and the original VV10 [79] functional, Sabatini and coworkers reoptimized one of VV10's two empirical parameters, b and C [93]. ONETEP implements rVV10 NLC as described in Ref. 93, and thus uses the reoptimized b value ($b = 6.3$).

The implementation of B97M-V in ONETEP uses rVV10 NLC and we will refer to this combination of the local part of the B97M-V functional with revised VV10 NLC as "B97M-rV", as distinct from "B97M-V", which incorporates the original VV10 NLC functional. Since the local part of B97M-V [19] was optimized in combination with the original VV10 NLC functional and we have not reoptimized the empirical parameters for either part of B97M-rV to account for use of rVV10 NLC, we might expect B97M-V and B97M-rV to behave slightly differently. However, as described in section IV A, we find the differences in relative energies obtained using the two variants to be very small in most cases.

The impact of rVV10 NLC on the computational cost of B97M-rV compared to a "pure" meta-GGA, such as PKZB, is an interesting question, and particularly relevant to our intended application of B97M-rV to large scale calculations. This question is considered in section IV B.

To avoid the time-consuming process of implementing the local component of B97M-rV from scratch, this was implemented in ONETEP by adapting the existing Q-Chem routines for evaluating the B97M-V exchange-correlation energy density, $\epsilon_{\mathrm{xc}}(\mathbf{r})$, and its derivatives with respect to n, ∇n and τ (as described in Ref. 19). The initial testing and debugging of these modified routines within the context of ONETEP was done by comparison with an independent implementation of the local part of B97M-V in the Libxc library [77], which could be called directly within ONETEP. Comparisons of full self-consistent B97M-rV energies obtained using ONETEP and Q-Chem are presented in section IV A.

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

IV. RESULTS

A. Testing and validation

In order to test the numerical correctness of the theoretical framework outlined in section III, we compared meta-GGA energies calculated self-consistently using the ONETEP and Q-Chem packages for a number of small molecular systems. Q-Chem is an all-electron Gaussian basis set code and therefore differs substantially from ONETEP in its approach to self-consistent solution of the Kohn-Sham equations (see Refs. 70, 96, 97 for details). This difference presents both a challenge and a opportunity. The challenge arises in setting up calculations with comparable conditions in each package. For example, it is not immediately clear what kinetic energy cutoff and pseudopotential in ONETEP is appropriate to compare with a given atomic orbital basis set and integration grid in Q-Chem. The opportunity is in the strong independent validation provided by comparison against an independent electronic structure code built upon an alternative theoretical foundation.

For all Q-Chem results reported in this section, a development version of Q-Chem 4 was used [70]. In these calculations, the def2-QZVPPD basis set was used for all atoms [98, 99] and the (99,590) integration grid (99 radial shells per atom, 590 angular grid points per shell) was used to evaluate exchange-correlation functionals. The SG-1 grid [100] was used to evaluate the rVV10 NLC component of B97M-rV [101].

The ONETEP calculations were performed using norm-conserving pseudopotentials from the Rappe-Bennett pseudopotential library (GGA-optimized) [102] for H, C, N, O and F nuclei [103]. The Opium pseudopotential generator [104–106] was used to generate the pseudopotential for sulfur, used in calculations on cysteine [107]. The pseudopotential files used in the calculations presented in this section are included in the supplementary material for ease of reproduction. The number of NGWFs used to represent the valence electrons of each atom was as follows: H, 1; C, 4; N, 4; O, 4; S, 9. This corresponds to 1 NGWF for each valence-shell electron pair.

Through a set of calibration calculations, two combinations of kinetic energy cutoff and NGWF localization sphere radii were identified for comparison with the Q-Chem calculations, corresponding roughly to settings that would be used in "standard" production and "high-accuracy" reference calculations. The "standard" calculations used a kinetic energy cutoff of 900 eV and NGWF localization sphere radii of 8.0 a_0 for all nuclei, while the "high-accuracy" calculations used a kinetic energy cutoff of 1200 eV and NGWF localization sphere radii of 12.0 a_0. Note that these kinetic energy cutoff values are the values requested from ONETEP and in practice the kinetic energy cutoff used in the calculation deviates slightly from this in order that an integer number of grid points is positioned along each simulation cell dimension.

The simulation cell sizes used for each test case were derived from the molecular geometry such that there was at least a 5 a_0 vacuum gap between the edge of the cell and the furthest extent of the NGWF localization spheres (using 12 a_0 radii) in each Cartesian direction. This vacuum buffer was introduced to avoid interactions between periodic images of the molecular systems. The simulation cell dimensions are provided with molecular geometries in the supplementary material.

No correction for basis set superposition error (BSSE) was applied in either the Q-Chem or ONETEP calculations. As mentioned earlier, *in situ* optimization of NGWFs eliminates the need for such corrections in ONETEP [31]. In the case of Q-Chem, it was judged that BSSE correction was unnecessary for the large def2-QZVPPD atomic orbital basis set as the monomer basis sets would already be highly saturated.

The binding energies of four non-covalently bound complexes, shown in Fig. 2, and a set of relative energies for conformers of cysteine and melatonin, shown in Fig. 3, were used to test the implementation of meta-GGAs in ONETEP against Q-Chem. Geometry specifications for all the test cases are available in the supplementary material. The comparison of relative energies was necessary, since absolute energies from all-electron codes are not equivalent to absolute energies from pseudopotential codes, where the core orbitals are represented by an effective ionic potential. The results of Q-Chem and ONETEP calculations using the PKZB and B97M-rV meta-GGAs are presented in Tables I and II. Results are also reported for the PBE GGA, as a reference. Ideally, the difference between the codes for self-consistently evaluated meta-GGA energies should be similar to the difference for GGA energies, where the additional energy and gradient terms introduced by τ-dependence (section III) are not required.

As mentioned in section III F 2, "B97M-rV" denotes the B97M-V functional, used in conjunction with the revised VV10 NLC functional of Ref. 93, rather than the original VV10 NLC functional [79]. The results reported in the B97M-rV rows of Tables I and II were obtained using the B97M-rV functional in both Q-Chem and ONETEP, to avoid introducing an additional error due to the difference between original and revised VV10 NLC. The difference between B97M-V and B97M-rV is typically small, on the order of 0.1 kcal mol^{-1} (see Table S1 in the supplementary material for examples).

For all the test cases studied, the difference between Q-Chem and ONETEP meta-GGA energies for both the 900 eV/8 a_0 and 1200 eV/12 a_0 settings was of a similar magnitude to the differences seen for PBE, indicating that the implementation of self-consistent meta-GGA evaluation is well-behaved. For the binding energies, the RMSD for ONETEP and Q-Chem PBE results is $\sim$ 0.1 kcal mol^{-1}. Similar RMSDs are obtained for PKZB and B97M-rV, though B97M-rV actually produces better agreement than PBE, with both RMSDs slightly smaller than those obtained for PBE. For the conformational energies, the RMSDs for PBE are slightly larger, nearer to 0.2 kcal mol^{-1}. The meta-GGAs produce RMSDs close to 0.3 kcal mol^{-1} in this case. Differences between Q-Chem and ONETEP for the melatonin conformer energies are somewhat larger for PKZB and B97M-rV than for PBE, with some absolute differences exceeding 0.4 kcal mol^{-1}. For B97M-rV, the larger Q-Chem/ONETEP difference seen for the melatonin conformer energies is accompanied by a comparatively large difference of 0.7 kcal mol^{-1} between

ONETEP results with the two cutoff energy/NGWF radius settings. This may indicate that the melatonin conformer energies calculated with the lower quality 900 eV/8 a_0 basis set using B97M-rV are not well-converged with respect to the basis set.

Given the limited size of our dataset and the substantial differences between Q-Chem and ONETEP outlined earlier, it is difficult to identify specific reasons why the RMSDs for the conformational energies (Table II) are generally somewhat larger than those for the binding energies (Table I). It is plausible that the different magnitudes of the RMSDs observed for the binding and conformational energies are simply an artifact of the systems selected for this comparison—with a larger set of data points, the differences between the conformational and binding energies could evaporate. To draw meaningful comparisons between the differences in conformational energies and binding energies calculated using ONETEP and Q-Chem would require extensive benchmarking and comparison on a much larger number of data points. In the present work, we are primarily concerned with demonstrating the numerical validity of the implementation of meta-GGAs in ONETEP, using the behavior of the PBE GGA as a reference. A more extensive comparison of the differences in energies calculated using ONETEP and Q-Chem is therefore beyond the scope of this work.

In spite of the slightly larger deviations from the Q-Chem results seen for the conformer energies in Table II, the ONETEP results retain the same energetic ordering as the Q-Chem results for all functionals and basis set settings: aa < be < dw for melatonin, and I < II < III < IV for cysteine with PBE and PKZB and I < III < II < IV for B97M-rV. In addition, there is only one binding energy in which ONETEP predicts binding where Q-Chem does not, $C_6H_6\cdots H_2O$ (I) with PKZB. In this case, Q-Chem produces a very small positive binding energy of 0.026 kcal mol^{-1} while ONETEP gives small negative binding energies of -0.011 kcal mol^{-1} and -0.034 kcal mol^{-1} for the 900 eV/8 a_0 and 1200 eV/12 a_0 settings, respectively.

The comparisons presented in Tables I and II demonstrate that the theoretical framework outlined in this paper, as implemented in ONETEP, yields self-consistent meta-GGA energies in good qualitative and quantitative agreement with a well-established implementation in Q-Chem. In spite of the significant theoretical and algorithmic differences between the two codes, it is possible to produce results which agree to within a fraction of a kcal mol^{-1}, which is consistent with the extent of agreement achieved with the well-tested PBE GGA.

B. Linear-scaling performance testing

A key motivation in implementing meta-GGA functionals within ONETEP was their application to large systems with many thousands of atoms. When studying systems of this size, low scaling computational cost with respect to system size is vital. To evaluate the computational scaling of our implementation of meta-GGAs in ONETEP, we performed a series of single point DFT energy calculations on segments of a 13696 atom amyloid fibril. The results are plotted in Fig. 4, and clearly demonstrate the computational cost for self-consistent single-point energy calculations using meta-GGAs, as measured in terms of overall execution time, increases linearly with the number of atoms, N.

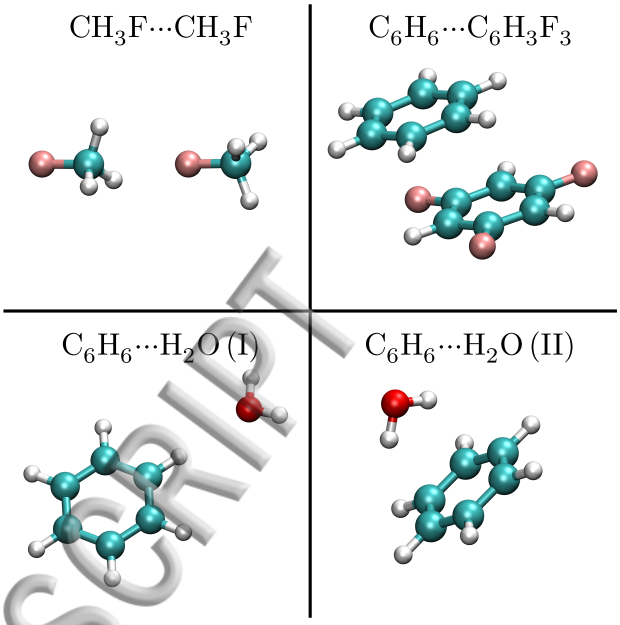

FIG. 2. The four systems for which binding energies are presented in Table I. The geometries for $CH_3F\cdots CH_3F$ and $C_6H_6\cdots C_6H_3F_3$ were obtained from the X40 set of non-covalent interactions of halogenated molecules [108] while the geometries for the two $C_6H_6\cdots H_2O$ structures were from the set of hydrocarbon-water interactions presented in Ref. 109 (this is referred to as the HW30 set in Refs. 19, 59).

All calculations presented in Fig. 4 were performed using a development version of ONETEP, running on the Iridis 4 supercomputer at the University of Southampton with 128 MPI processes and 4 OpenMP threads per process. ONETEP was compiled using the Intel Fortran compiler 16.0 and linked to the Intel Math Kernel Library 11.3. The kinetic energy cutoff and NGWF cutoff radii (for all atoms) were 600 eV and 7.0 a_0, respectively. These settings correspond to a medium quality basis set and while they are sufficient for testing the computational scaling of our implementation, we would recommend production calculations are run with a higher quality basis set (higher kinetic energy cutoff, larger NGWF radii). The calculation was converged until the RMS NGWF gradient was less than 2.0×10^{-6} a.u. and a density kernel cutoff (Eq. 11) of 20.0 a_0 was used. This is a relatively small kernel cutoff, suitable for testing purposes, but unlikely to be sufficient for production calculations, where larger values would be necessary. The structure of the 13696 atom fibril was kindly provided by the authors of Ref. 113 and smaller segments were derived from this structure. The atomic coordinates used in all the amyloid fibril calculations are provided in the supplementary material.

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

The larger prefactor associated with the meta-GGAs observed in Fig. 4 is unsurprising, given the additional operations required to evaluate the meta-GGA energy and gradient terms. The meta-GGA code is still a work-in-progress and it is likely that further optimization will improve the prefactor.

Though it may be possible to decrease the overall computational cost of meta-GGA calculations, we expect that meta-GGA calculations will remain somewhat more expensive than corresponding GGA and LDA calculations, for the simple reason that the evaluation of $\tau(\mathbf{r})$ over the simulation cell is inherently more costly than the evaluation of $n(\mathbf{r})$. As outlined in section III A, the evaluation of $\tau(\mathbf{r})$ follows the same multi-step procedure used to evaluate $n(\mathbf{r})$ (described in Ref. 29). However, the evaluation of $\tau(\mathbf{r})$ additionally involves application of the gradient operator and a scalar product over Cartesian components (Eq. 30). Since the evaluation of the $n(\mathbf{r})$ is one of the more costly steps in a typical ONETEP calculation, we would expect the additional cost of evaluating $\tau(\mathbf{r})$ to be a major contributor to the overall greater cost of calculations using meta-GGAs.

As a rough example of the extent of the additional cost associated performing a meta-GGA calculation, we can examine timings reported by ONETEP's internal timing routines. For the full 13696 atom fibril, the PKZB single-point energy calculation converged after 11 outer loop (NGWF optimization) iterations in 7.3 h. The amount of time spent in the key routines for evaluating the charge density and kinetic energy density on the simulation cell grid was 0.9 h and 2.0 h, respectively. For comparison, the PBE calculation took 3.4 h to complete 11 outer loop iterations, with 0.9 h spent inside the same key charge density evaluation routine. For the PKZB calculation, the time spent inside routines for evaluating integrals over the τ-dependent part of the exchange-correlation potential (Eq. 40) was 1.5 h, which accounts for most of the remaining 1.9 h difference between for PKZB and PBE calculations. The last 0.4 h of this difference is difficult to attribute, because the modifications made to ONETEP to support meta-GGAs were spread throughout the codebase and thus affect many individual routines to some extent. Nevertheless, these timings suggest that the additional cost of evaluating the kinetic energy density and new exchange-correlation integrals (Eq. 40) are significant contributors to the increased overall cost of calculating meta-GGA energies over GGA energies in our current implementation.

The data in Fig. 4 suggest that PKZB calculations have a higher prefactor than B97M-rV, since the gradient of the fitted

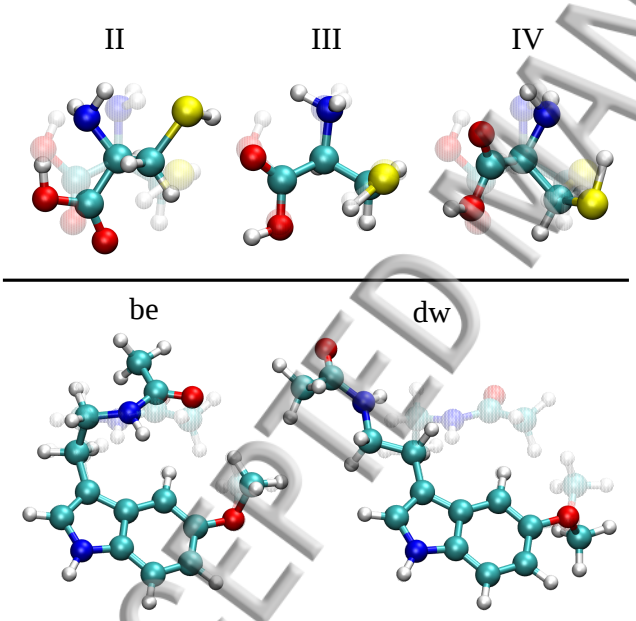

FIG. 3. Conformations of melatonin and cysteine, for which relative energies are presented in Table II. The structures of cysteine (I, II, III, and IV) were obtained from Ref. 110, forming part of the CYCONF set, which itself forms part of the GMTKN30 database [111]. The structures of melatonin ("aa", "be" and "dw") were obtained from Ref. 112 (this is referred to as the Melatonin52 set in Ref. 59). The conformational energies in Table II are relative to the lowest energy conformers, "I" and "aa" for cysteine and melatonin, respectively. In this figure, the higher-energy conformations of each molecule are superimposed on the lowest energy conformations, to illustrate the structural differences.

TABLE I. Comparison of binding energies, in kcal mol^{-1}, for the four non-covalently bound systems presented in Fig. 2 calculated using Q-Chem and ONETEP. Binding energies are presented for the PBE, and PKZB and B97M-rV exchange-correlation functionals. Monomer geometries were derived from the complex geometries (without optimization). The differences between energies calculated using Q-Chem and ONETEP are listed in the columns labeled Δ. The def2-QZVPPD [98, 99] basis set was used for all atoms in the Q-Chem calculations, while two levels of basis set quality were used in ONETEP, denoted X/Y, where X is the kinetic energy cutoff in eV and Y is the radius of the NGWF localization sphere in a_0. The RMSD of the Q-Chem and ONETEP energies is reported for each functional at both the ONETEP basis set levels.

		ONETEP			
	Q-Chem	900 eV/8 a_0	Δ	1200 eV/12 a_0	Δ
PBE					
$CH_3F\cdots CH_3F$	−0.925	−0.963	−0.038	−0.957	−0.033
$C_6H_6\cdots C_6H_3F_3$	0.650	0.856	0.206	0.794	0.144
$C_6H_6\cdots H_2O$ (I)	−0.922	−0.920	0.002	−0.925	−0.003
$C_6H_6\cdots H_2O$ (II)	−1.790	−1.706	0.084	−1.775	0.014
RMSD			0.113		0.074
PKZB					
$CH_3F\cdots CH_3F$	−0.528	−0.607	−0.079	−0.619	−0.091
$C_6H_6\cdots C_6H_3F_3$	1.221	1.014	−0.207	0.956	−0.265
$C_6H_6\cdots H_2O$ (I)	0.026	−0.011	−0.038	−0.034	−0.060
$C_6H_6\cdots H_2O$ (II)	−0.572	−0.716	−0.144	−0.790	−0.217
RMSD			0.133		0.180
B97M-rV					
$CH_3F\cdots CH_3F$	−1.435	−1.496	−0.061	−1.497	−0.062
$C_6H_6\cdots C_6H_3F_3$	−4.508	−4.358	0.149	−4.443	0.065
$C_6H_6\cdots H_2O$ (I)	−1.214	−1.246	−0.032	−1.249	−0.034
$C_6H_6\cdots H_2O$ (II)	−3.283	−3.249	0.034	−3.329	−0.046
RMSD			0.084		0.053

PKZB line appears larger than the gradient for the B97M-rV line. At first glance, this is an unexpected result, since one would expect B97M-rV to be more costly because of the additional work associated with rVV10 NLC, not present in PKZB. In fact, the explanation for this outcome is the slower convergence of PKZB calculations relative to B97M-rV. For all amyloid fibril segments, the PKZB calculations required 11 outer loop (NGWF optimization) iterations to achieve convergence (based on the RMS NGWF gradient), while B97M-rV required 8 or 9 iterations for most segments, as detailed in Table III. For the 5136 and 8560 atom segments, B97M-rV required 11 and 12 outer loop iterations to converge, respectively, and these individual calculations were more costly than the corresponding PKZB calculations. For comparison, the PBE calculations all required 11 iterations to achieve convergence. The results presented in Table III demonstrate that the average time per outer loop iteration was consistently lower for PKZB than for B97M-rV across all segment sizes, supporting the suggestion that the unexpectedly good relative performance of B97M-rV is a consequence of more rapid convergence.

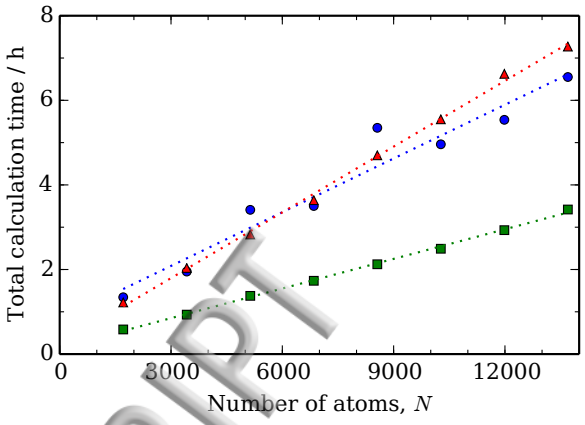

FIG. 4. Execution times for single point DFT calculations on amyloid fibril segments of increasing size, using the PBE (green squares), PKZB (red triangles) and B97M-rV (blue circles) exchange-correlation functionals. Linear least squares fits of the data for each functional are plotted as dotted lines. The computational details of these calculations are described in section IV B.

TABLE II. Comparison of conformational energies, in kcal mol^{-1}, for the conformers of cysteine and melatonin presented in Fig. 3, calculated using Q-Chem and ONETEP. The conformational energies are relative to the lowest energy conformer, labeled "I" and "aa" for cysteine and melatonin, respectively. As in Table I, differences between energies calculated using Q-Chem and ONETEP are listed in the columns labeled Δ and the RMSDs of the energies calculated using Q-Chem and ONETEP are provided for each exchange-correlation functional. The basis sets used for the Q-Chem and ONETEP calculations are as described for Table I.

			ONETEP			
		Q-Chem	900 eV/8 a_0	Δ	1200 eV/12 a_0	Δ
PBE						
Cysteine	II	0.894	0.910	0.016	1.013	0.119
Cysteine	III	3.033	3.261	0.228	3.292	0.259
Cysteine	IV	3.161	3.385	0.224	3.339	0.178
Melatonin	be	4.487	4.489	0.001	4.521	0.033
Melatonin	dw	6.319	6.125	−0.194	6.469	0.151
		RMSD		0.167		0.165
PKZB						
Cysteine	II	0.744	0.745	0.001	0.867	0.123
Cysteine	III	1.194	1.565	0.370	1.573	0.379
Cysteine	IV	2.102	2.484	0.382	2.212	0.110
Melatonin	be	4.853	5.316	0.463	4.972	0.119
Melatonin	dw	5.939	6.163	0.224	6.358	0.419
		RMSD		0.331		0.268
B97M-rV						
Cysteine	II	1.481	1.528	0.048	1.699	0.219
Cysteine	III	1.313	1.427	0.114	1.664	0.351
Cysteine	IV	1.481	1.683	0.202	1.863	0.382
Melatonin	be	5.582	5.193	−0.389	5.909	0.327
Melatonin	dw	9.424	9.011	−0.414	9.731	0.306
		RMSD		0.275		0.322

	PBE		PKZB		B97M-rV	
N_{atoms}	N_{iter}	t/N_{iter}	N_{iter}	t/N_{iter}	N_{iter}	t/N_{iter}
1712	11	191	11	400	9	538
3424	11	305	11	666	8	878
5136	11	451	11	925	11	1116
6848	11	567	11	1191	9	1401
8560	11	695	11	1538	12	1605
10272	11	815	11	1817	9	1984
11984	11	959	11	2168	9	2216
13696	11	1119	11	2379	9	2620

TABLE III. Outer loop iteration execution times, to the nearest second, for single point DFT calculations on amyloid fibril segments of increasing size. For each amyloid fibril segment, the number of outer loop iterations, N_{iter}, required to achieve convergence and the total time taken, t, divided by the number of outer loop iterations are given for the PBE, PKZB and B97M-rV functionals. The number of atoms in each amyloid fibril segment, N_{atoms}, is listed in the first column.

V. CONCLUSION

In this paper, we described a theoretical framework for self-consistent evaluation of meta-GGA exchange-correlation energies within ONETEP's linear-scaling DFT formalism. In developing this framework, we used the FDO method to evaluate functional derivatives (section II D) with respect to the density, yielding an orbital-specific exchange-correlation potential. To ensure that our approach maintained linear-scaling computational cost, we derived expressions for the kinetic energy density (section III A), τ, and exchange-correlation potential (section III B) in terms of ONETEP's strictly localized orbitals (NGWFs, Eq. 12) and density kernel (Eq. 10). We also derived expressions for the gradient of the τ-dependent exchange-correlation energy with respect to these localized

quantities in order that the meta-GGA energy could be self-consistently optimized using ONETEP's modified LNV scheme (sections III D and III E).

The PKZB and B97M-rV meta-GGAs were implemented in ONETEP, and their behavior within the self-consistent framework was validated against existing implementations in the Q-Chem package for a set of binding energies and relative conformational energies (Tables I and II). In all cases, the numerical agreement between ONETEP and Q-Chem was commensurate with the agreement obtained using the well-tested PBE GGA functional, indicating that the additional gradient and energy terms necessary to evaluate self-consistent meta-GGA energies (section III) are well-behaved. In addition, agreement in terms of the ordering of conformer energies was also observed for both the meta-GGAs tested (Table II). Given the significant algorithmic and theoretical differences between the Q-Chem and ONETEP packages, the quantitative and qualitative agreement in meta-GGA energies calculated using both codes is a strong validation of the theoretical framework presented in the preceding sections.

As outlined in the section I, a key motivation for this work was the proposed "Goldilocks zone" occupied by the meta-GGAs, balancing computational efficiency and numerical accuracy (Fig. 1). Meta-GGAs are particular attractive for linear-scaling DFT calculations because they have the potential to improve upon the accuracy of LDA and GGA results, while retaining a computationally efficient semi-local form (Eq. 15).

In section IV B we confirmed that our implementation of self-consistent meta-GGA evaluation in ONETEP exhibited $O(N)$ scaling, performing calculations on amyloid fibrils consisting of thousands of atoms. The computational cost of meta-GGA calculations was somewhat greater than for a corresponding GGA calculations, as would be expected due to the additional complexities associated with τ-dependence. In general, the meta-GGA calculations required approximately double the amount of time to complete compared to the corresponding GGA calculations, though even for the largest 13696 atom amyloid fibril, the total time was only around 7 h for both meta-GGA functionals tested. Given the excellent accuracy possible using modern meta-GGA functionals such as B97M-V (section III F 2), we would suggest that this factor of two increase in cost may be a price worth paying.

We anticipate that the implementation of self-consistent meta-GGA evaluation in ONETEP described in this paper will form the foundation for future applied studies in a variety of contexts. One of our motivations in implementing support for meta-GGAs in ONETEP was for the study of biomolecular association. In particular, the excellent performance of B97M-V [19, 59] for non-covalent interactions (section III F 2) suggested that we might be able to achieve significant improvements in accuracy over existing work with dispersion-corrected GGAs [14, 15] without the need for costly hybrid functionals. More broadly, meta-GGAs are capable of improving upon GGAs across a diverse range of systems and types of bonding (see for example, recent examples using the nonempirical SCAN functional in Ref. 21). This suggests that ONETEP users will find utility for meta-GGAs in a variety of applications where accurate description of exchange-correlation effects is important.

Of the linear-scaling DFT software packages listed in section I, only Quickstep (part of the CP2K suite [114]) currently supports meta-GGA exchange-correlation functionals, as mentioned in Ref. 12. Although both ONETEP and Quickstep use linear-scaling methods in order to apply DFT in large-scale calculations, the codes differ substantially in their implementation. Perhaps the starkest difference is in the types of basis function used. Quickstep is based on the Gaussian and plane wave (GPW) method [115], utilizing a combination of Gaussian functions and plane waves to construct the Kohn-Sham orbitals and electron density, while ONETEP employs self-consistently optimized NGWFs constructed from an underlying psinc basis set (section II). The NGWF/psinc basis and its *in situ* optimization presents unique challenges for the implementation of meta-GGA exchange-correlation functionals which have been addressed in this paper.

Although linear-scaling DFT packages can differ substantially in design and implementation, they often share common theoretical foundations. The essential ideas of direct energy minimization with respect to a separable density matrix and the use of strictly localized basis functions, as employed in ONETEP, are a common starting point for linear-scaling DFT methods (as described in Ref. 6). We therefore hope that the detailed account of the implementation of meta-GGAs in ONETEP presented in this paper will facilitate the implementation of meta-GGAs in other linear-scaling DFT codes which share this foundation.

All the calculations in this paper were performed using norm-conserving pseudopotentials, since the current implementation of self-consistent meta-GGA functionality in ONETEP is restricted to pseudopotentials of this type. An alternative approach to the problem of representing atomic cores is the projector augmented wave (PAW) method [116], in which all-electron orbitals are retained, but are expressed in terms of smooth pseudofunctions and partial waves representing the orbital in the region of the atomic core. The PAW method has some significant advantages over norm-conserving pseudopotentials for practical DFT calculations, notably improved transferability and better convergence with respect to the kinetic energy cutoff. Support for PAW has recently been implemented in ONETEP [117], allowing LDA and GGA calculations to be performed with the method. The next step for the work described in this paper is to extend this functionality to allow self-consistent meta-GGA evaluation using PAW. This will require the kinetic energy density to be decomposed into contributions from the smooth pseudofunctions and partial waves, and will also involve additional terms in the Hamiltonian, as described in Ref. 69.

SUPPLEMENTARY MATERIAL

See the supplementary material for: comparisons of binding and relative conformer energies calculated using B97M-V and B97M-rV; structures for all molecules used for calculations in this paper; and the pseudopotentials used in calcula-

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

tions reported in section IV A.

ACKNOWLEDGMENTS

JCW acknowledges the Engineering and Physical Sciences Research Council (EPSRC) UK for postdoctoral funding (EPSRC grant numbers EP/K039156/1 and EP/J015059/1). NM and MHG were supported by the Director, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences division of the U.S. Department of Energy under Contract DE-AC02-05CH11231. The authors would like to acknowledge the IRIDIS High Performance Computing Facility at the University of Southampton, which provided the resources used to perform the large scale calculations presented in this work. JCW would like to thank Dr Jacek Dziedzic for helpful discussions during the software development and testing phases of this work. JCW and CKS would also like to thank the UKCP consortium (EPSRC grant number EP/K013556/1) which has provided us with ongoing access to the UK's ARCHER national supercomputer.

Appendix A: Gradient of $E_{\rm xc}$ with respect to the density kernel elements

Eq. 42 may be derived by considering the definition of the partial derivative in terms of a limit

$$\frac{\partial E_{\rm xc}}{\partial K^{\eta\theta}} = \lim_{h\to 0} \frac{E_{\rm xc}[n(\mathbf{r}, K^{\eta\theta}+h)] - E_{\rm xc}[n(\mathbf{r})]}{h} \tag{A1}$$

where the change in $K^{\eta\theta}$ affects the density, n, of which $E_{\rm xc}$ is a functional. The density with $K^{\eta\theta}$ varied by h is

$$\begin{aligned} &n(\mathbf{r}, K^{\eta\theta}+h) \\ &\quad = \sum_{\alpha\beta} \varphi_\alpha(\mathbf{r}) K^{\alpha\beta} \varphi^*_\beta(\mathbf{r}) + h\varphi_\eta(\mathbf{r})\varphi^*_\theta(\mathbf{r}) \\ &\quad = n(\mathbf{r}) + h\frac{\partial n(\mathbf{r})}{\partial K^{\eta\theta}}. \end{aligned} \tag{A2}$$

The functional derivative of $E_{\rm xc}[n]$ with respect to the density may be written

$$\begin{aligned} &\int \mathrm{d}\mathbf{r} \frac{\delta E_{\rm xc}[n]}{\delta n(\mathbf{r})}\zeta(\mathbf{r}) \\ &\quad = \frac{\mathrm{d}}{\mathrm{d}\epsilon} E_{\rm xc}[n+\epsilon\zeta]\Big|_{\epsilon=0} \\ &\quad = \lim_{\epsilon\to 0} \frac{E_{\rm xc}[n(\mathbf{r})+\epsilon\zeta(\mathbf{r})] - E_{\rm xc}[n(\mathbf{r})]}{\epsilon} \end{aligned} \tag{A3}$$

where $\zeta(\mathbf{r})$ is an arbitrary test function (see for example, appendix A of Ref. 118). Using $\partial n(\mathbf{r})/\partial K^{\eta\theta}$ as test function in Eq. A3 produces a limit identical to the limit in Eq. A1, i.e.

$$\begin{aligned} &\int \mathrm{d}\mathbf{r} \frac{\delta E_{\rm xc}[n]}{\delta n(\mathbf{r})}\frac{\partial n(\mathbf{r})}{\partial K^{\eta\theta}} \\ &\quad = \lim_{\epsilon\to 0} \frac{E_{\rm xc}[n(\mathbf{r}) + \epsilon\frac{\partial n(\mathbf{r})}{\partial K^{\eta\theta}}] - E_{\rm xc}[n(\mathbf{r})]}{\epsilon} \end{aligned} \tag{A4}$$

where we have used Eq. A2 to express $n(\mathbf{r}, K^{\eta\theta}+h)$. The expression for the gradient of $E_{\rm xc}$ with respect to the density kernel elements may therefore be written in terms of the exchange-correlation potential (as in Eq. 42), i.e.

$$\frac{\partial E_{\rm xc}}{\partial K^{\eta\theta}} = \int \mathrm{d}\mathbf{r} \frac{\delta E_{\rm xc}}{\delta n(\mathbf{r})}\frac{\partial n(\mathbf{r})}{\partial K^{\eta\theta}}. \tag{A5}$$

Appendix B: Gradient of $E_{\rm xc}$ with respect to the NGWFs

The following derivation of the exchange-correlation component of the gradient of the energy with respect to the NGWFs is based on work presented in Ref. 73 (ch. 7), and is reproduced here for the convenience of interested readers.

Eq. 47 can be derived starting with the chain rule for functional derivatives (see appendix A of Ref. 119):

$$\begin{aligned} \frac{\delta E_{\rm xc}}{\delta \varphi^*_\alpha(\mathbf{r})} &= \int \mathrm{d}\mathbf{r}' \frac{\delta E_{\rm xc}}{\delta n(\mathbf{r}')}\frac{\delta n(\mathbf{r}')}{\delta \varphi^*_\alpha(\mathbf{r})} \\ &= \int \mathrm{d}\mathbf{r}' V_{\rm xc}(\mathbf{r}')\frac{\delta n(\mathbf{r}')}{\delta \varphi^*_\alpha(\mathbf{r})}. \end{aligned} \tag{B1}$$

The functional derivative of $n(\mathbf{r})$ with respect to the complex-conjugate of a NGWF can be expanded using the product rule for functional derivatives (appendix A, Ref. 119):

$$\begin{aligned} \frac{\delta n(\mathbf{r}')}{\delta \varphi^*_\alpha(\mathbf{r})} &= \varphi_\mu(\mathbf{r}')\frac{\delta K^{\mu\nu}}{\delta \varphi^*_\alpha(\mathbf{r})}\varphi^*_\nu(\mathbf{r}') \\ &\quad + \varphi_\mu(\mathbf{r}')K^{\mu\nu}\frac{\delta \varphi^*_\nu(\mathbf{r}')}{\delta \varphi^*_\alpha(\mathbf{r})}. \end{aligned} \tag{B2}$$

The functional derivative in the second term of Eq. B2 is simply

$$\frac{\delta \varphi^*_\nu(\mathbf{r}')}{\delta \varphi^*_\alpha(\mathbf{r})} = \delta^\alpha_\nu \delta(\mathbf{r}-\mathbf{r}'). \tag{B3}$$

To evaluate the functional derivative in the first term, the chain rule can be used again to yield

$$\frac{\delta K^{\mu\nu}}{\delta \varphi^*_\alpha(\mathbf{r})} = \int \mathrm{d}\mathbf{r}' \frac{\delta K^{\mu\nu}}{\delta S_{\iota\kappa}}\frac{\delta S_{\iota\kappa}}{\delta \varphi^*_\alpha(\mathbf{r})}. \tag{B4}$$

Since $\mathbf{K}$ is a function of $\mathbf{S}$, rather than a functional, $\delta K^{\mu\nu}/\delta S^{\iota\kappa}$ reduces to partial derivative, i.e.

$$\frac{\delta K^{\mu\nu}}{\delta S_{\iota\kappa}} = \frac{\partial K^{\mu\nu}}{\partial S_{\iota\kappa}}\delta(\mathbf{r}-\mathbf{r}') \tag{B5}$$

where, using the expression for the density kernel in terms of the auxiliary density kernel (Eq. 14),

$$\begin{aligned} \frac{\partial K^{\mu\nu}}{\partial S_{\iota\kappa}} &= 3L^{\mu\iota}L^{\kappa\nu} \\ &\quad - 2(L^{\mu\iota}L^{\kappa\gamma}S_{\gamma\delta}L^{\delta\nu} + L^{\mu\gamma}S_{\gamma\delta}L^{\delta\iota}L^{\kappa\nu}). \end{aligned} \tag{B6}$$

The other functional derivative in Eq. B4 is

$$\frac{\delta S_{\iota\kappa}}{\delta \varphi^*_\alpha(\mathbf{r})} = \delta^\alpha_\iota \varphi_\kappa(\mathbf{r}) \tag{B7}$$

and thus,

$$\frac{\delta K^{\mu\nu}}{\delta\varphi^*_\alpha(\mathbf{r})} = \frac{\partial K^{\mu\nu}}{\partial S_{\alpha\kappa}}\varphi_\kappa(\mathbf{r}). \tag{B8}$$

Using Eqs. B2, B3 and B8, Eq. B1 becomes

$$\frac{\delta E_{\rm xc}}{\delta\varphi^*_\alpha(\mathbf{r})} = \varphi_\kappa(\mathbf{r})V_{\nu\mu}\frac{\partial K^{\mu\nu}}{\partial S_{\alpha\kappa}} + \frac{\delta E_{\rm xc}}{\delta n(\mathbf{r})}\varphi_\kappa(\mathbf{r})K^{\kappa\alpha} \tag{B9}$$

where $V_{\nu\mu}$ is as defined in Eq. 49.

Finally, recognizing that

$$[\mathbf{Q}^V]^{\kappa\alpha} = V_{\nu\mu}\frac{\partial K^{\mu\nu}}{\partial S_{\alpha\kappa}}, \tag{B10}$$

where $[\mathbf{Q}^V]^{\kappa\alpha}$ is defined in Eq. 48, Eq. B9 becomes

$$\frac{\delta E_{\rm xc}}{\delta\varphi^*_\alpha(\mathbf{r})} = \varphi_\kappa(\mathbf{r})[\mathbf{Q}^V]^{\kappa\alpha} + \frac{\delta E_{\rm xc}}{\delta n(\mathbf{r})}\varphi_\kappa(\mathbf{r})K^{\kappa\alpha} \tag{B11}$$

which is the expression provided earlier (Eq. 47).

Note that the preceding derivation was for complex NGWFs (where each NGWF and its complex conjugate may be varied independently). For real NGWFs, Eq. B11 is simply multiplied by a factor of two, i.e.

$$\frac{\delta E_{\rm xc}}{\delta\varphi_\alpha(\mathbf{r})} = 2\varphi_\kappa(\mathbf{r})[\mathbf{Q}^V]^{\kappa\alpha} + 2\frac{\delta E_{\rm xc}}{\delta n(\mathbf{r})}\varphi_\kappa(\mathbf{r})K^{\kappa\alpha}. \tag{B12}$$

Appendix C: PKZB energy derivatives

To evaluate the PKZB exchange-correlation potential under the FDO method, we require derivatives of the exchange-correlation energy density, $\epsilon_{\rm xc}$, with respect to the charge density, n, its gradient, ∇n, and the kinetic energy density, τ. Equations for the PKZB exchange and correlation energy were provided in Ref. 22, but the derivatives necessary to evaluate the exchange-correlation potential were not, and do not appear to have been published elsewhere. To facilitate reproduction of our work, we present here the equations used to evaluate the PKZB energy and potential.

1. Exchange

The PKZB exchange energy functional [22, 74] has the form,

$$E_{\rm x}^{\rm PKZB}[n] = \int {\rm d}\mathbf{r}\, \epsilon_{\rm x}^{\rm unif}(n)F_{\rm x}(n,\nabla n,\tau) \tag{C1}$$

where

$$\epsilon_{\rm x}^{\rm unif} = -\frac{3}{4\pi}(3\pi^2)^{1/3}n^{4/3} \tag{C2}$$

is the exchange energy per unit volume of the uniform electron gas and

$$F_{\rm x}(p,\widetilde{q}) = 1 + \kappa - \frac{\kappa}{1 + x/\kappa} \tag{C3}$$

is the enhancement factor, which has a similar form to the factor used for PBE [40, 43]. PKZB differs from PBE in the definition of x:

$$x = c_1 p + c_2\widetilde{q}^2 + c_3\widetilde{q}p + c_4 p^2 \tag{C4}$$

where

$$p = \frac{|\nabla n|^2}{4(3\pi^2)^{2/3}n^{8/3}} \tag{C5}$$

$$\widetilde{q} = \frac{3\tau}{2(3\pi^2)^{2/3}n^{5/3}} - \frac{9}{20} - \frac{p}{12} \tag{C6}$$

with $c_1 = 10/81$, $c_2 = 146/2025$, $c_3 = -73/405$ and

$$c_4 = \left[D + \frac{1}{\kappa}\left(\frac{10}{81}\right)^2\right]. \tag{C7}$$

In the preceding equations, κ and D are constants, with values 0.804 and 0.113, as set out in Ref. 22. The exchange energy for a spin-polarized system may easily be obtained using the well-known spin-scaling relation [120]:

$$E_{\rm x}[n_\uparrow, n_\downarrow] = \frac{1}{2}E_{\rm x}[2n_\uparrow] + \frac{1}{2}E_{\rm x}[2n_\downarrow]. \tag{C8}$$

The derivative of $\epsilon_{\rm x}^{\rm PKZB}$ with respect to n is

$$\frac{\partial\epsilon_{\rm x}^{\rm PKZB}}{\partial n} = \frac{\partial\epsilon_{\rm x}^{\rm unif}}{\partial n}F_{\rm x} + \epsilon_{\rm x}^{\rm unif}\frac{\partial F_{\rm x}}{\partial n} \tag{C9}$$

where the derivative of $\epsilon_{\rm x}^{\rm unif}$ is trivial

$$\frac{\partial\epsilon_{\rm x}^{\rm unif}}{\partial n} = -\frac{(3\pi^2)^{1/3}}{\pi}n^{1/3}. \tag{C10}$$

The derivative of the enhancement factor is

$$\frac{\partial F_{\rm x}}{\partial n} = \frac{\partial F_{\rm x}}{\partial x}\frac{\partial x}{\partial n} \tag{C11}$$

where

$$\frac{\partial F_{\rm x}}{\partial x} = \left(1 + \frac{x}{\kappa}\right)^{-2} \tag{C12}$$

and

$$\frac{\partial x}{\partial n} = c_1\frac{\partial p}{\partial n} + 2c_2\widetilde{q}\frac{\partial\widetilde{q}}{\partial n} + c_3\left(\frac{\partial\widetilde{q}}{\partial n}p + \widetilde{q}\frac{\partial p}{\partial n}\right) + 2c_4 p\frac{\partial p}{\partial n}. \tag{C13}$$

The derivatives of p and $\widetilde{q}$ are:

$$\frac{\partial p}{\partial n} = -\frac{8}{3}\frac{|\nabla n|^2}{4(3\pi^2)^{2/3}n^{11/3}}, \tag{C14}$$

$$\frac{\partial\widetilde{q}}{\partial n} = -\frac{5}{3}\frac{3\tau}{2(3\pi^2)^{2/3}n^{8/3}} - \frac{1}{12}\frac{\partial p}{\partial n}. \tag{C15}$$

The derivative of $\epsilon_{\rm x}^{\rm PKZB}$ with respect to ∇n can be obtained via the derivative with respect to $|\nabla n|$, using

$$\frac{\partial\epsilon_{\rm x}^{\rm PKZB}}{\partial\nabla n} = \frac{\partial\epsilon_{\rm x}^{\rm PKZB}}{\partial|\nabla n|}\frac{\nabla n}{|\nabla n|} \tag{C16}$$

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

where

$$\frac{\partial \epsilon_{\mathrm{x}}^{\mathrm{PKZB}}}{\partial |\nabla n|} = \epsilon_{\mathrm{x}}^{\mathrm{unif}} \frac{\partial F_{\mathrm{x}}}{\partial x} \frac{\partial x}{\partial |\nabla n|}. \tag{C17}$$

The form of $\partial x/\partial|\nabla n|$ is identical to Eq. C13, but with derivatives with respect to n replaced with derivatives with respect to $|\nabla n|$. These derivatives with respect to $|\nabla n|$ are

$$\frac{\partial p}{\partial |\nabla n|} = 2\frac{|\nabla n|}{4(3\pi^2)^{2/3} n^{8/3}}, \tag{C18}$$

$$\frac{\partial \widetilde{q}}{\partial |\nabla n|} = -\frac{1}{12}\frac{\partial p}{\partial |\nabla n|}. \tag{C19}$$

Finally, the derivative of $\epsilon_{\mathrm{x}}^{\mathrm{PKZB}}$ with respect to τ is

$$\frac{\partial \epsilon_{\mathrm{x}}^{\mathrm{PKZB}}}{\partial \tau} = \epsilon_{\mathrm{x}}^{\mathrm{unif}}(n) \frac{\partial F_{\mathrm{x}}}{\partial x} \frac{\partial x}{\partial \tau} \tag{C20}$$

where, since p does not depend on τ,

$$\frac{\partial x}{\partial \tau} = 2c_2 \widetilde{q} \frac{\partial \widetilde{q}}{\partial \tau} + c_3 \frac{\partial \widetilde{q}}{\partial \tau} p \tag{C21}$$

and

$$\frac{\partial \widetilde{q}}{\partial \tau} = \frac{3}{2(3\pi^2)^{2/3} n^{5/3}}. \tag{C22}$$

2. Correlation

The PKZB correlation energy functional [22, 74] has the form

$$E_{\mathrm{c}}^{\mathrm{PKZB}}[n_\uparrow, n_\downarrow] = \int \mathrm{d}\mathbf{r} \left\{ \epsilon_{\mathrm{c},\uparrow\downarrow}^{\mathrm{PBE}} f_1 - \sum_{\sigma=\uparrow,\downarrow} f_{2,\sigma} \epsilon_{\mathrm{c},\sigma}^{\mathrm{PBE}} \right\} \tag{C23}$$

where the PBE correlation energy densities [40, 43] are

$$\epsilon_{\mathrm{c},\uparrow\downarrow}^{\mathrm{PBE}} \equiv \epsilon_{\mathrm{c}}^{\mathrm{PBE}}(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow) \tag{C24}$$

$$\epsilon_{\mathrm{c},\sigma}^{\mathrm{PBE}} \equiv \epsilon_{\mathrm{c}}^{\mathrm{PBE}}(n_\sigma, 0, \nabla n_\sigma, 0) \tag{C25}$$

and

$$f_1 = 1 + C\left(\frac{\tau_\uparrow^{\mathrm{W}} + \tau_\downarrow^{\mathrm{W}}}{\tau_\uparrow + \tau_\downarrow}\right)^2 \tag{C26}$$

$$f_{2,\sigma} = (1 + C)\left(\frac{\tau_\sigma^{\mathrm{W}}}{\tau_\sigma}\right)^2, \tag{C27}$$

with $\tau_\sigma^W = (1/8)|\nabla n_\sigma|^2/n_\sigma$, the Weizsäcker kinetic energy density, and $\uparrow$, $\downarrow$ denoting the two spin components of the charge density, n, and kinetic energy density, τ.

Clearly, the derivatives of the PKZB correlation energy density, $\epsilon_{\mathrm{c}}^{\mathrm{PKZB}}$ (the terms inside the curly braces of Eq. C23), with respect to n and ∇n will involve the corresponding derivatives of the PBE correlation energy densities (Eqs. C24 and C25). However, since these derivatives can be evaluated using unmodified code for the PBE functional (as in our implementation of PKZB), we will not present explicit forms for these here.

Since the spin-unpolarized PKZB derivatives can be obtained from the spin-polarized forms via

$$\frac{\partial \epsilon_{\mathrm{c}}^{\mathrm{PKZB}}}{\partial n} = \sum_{\sigma=\uparrow,\downarrow} \frac{\partial \epsilon_{\mathrm{c}}^{\mathrm{PKZB}}}{\partial n_\sigma} \frac{\partial n_\sigma}{\partial n} \tag{C28}$$

with identical relationships for derivatives with respect to ∇n and τ, it is only necessary to consider the spin-polarized forms.

The derivative of $\epsilon_{\mathrm{c}}^{\mathrm{PKZB}}$ with respect to spin-density n_σ is

$$\begin{aligned}\frac{\partial \epsilon_{\mathrm{c}}^{\mathrm{PKZB}}}{\partial n_\sigma} = & \frac{\partial \epsilon_{\mathrm{c},\uparrow\downarrow}^{\mathrm{PBE}}}{\partial n_\sigma} f_1 + \epsilon_{\mathrm{c},\uparrow\downarrow}^{\mathrm{PBE}} \frac{\partial f_1}{\partial n_\sigma} \\ & - \left(\frac{\partial \epsilon_{\mathrm{c},\sigma}^{\mathrm{PBE}}}{\partial n_\sigma} f_{2,\sigma} + \epsilon_{\mathrm{c},\sigma}^{\mathrm{PBE}} \frac{\partial f_{2,\sigma}}{\partial n_\sigma} \right).\end{aligned} \tag{C29}$$

In addition to the derivatives of the PBE correlation energy density, derivatives of f_1 and $f_{2,\sigma}$ are required. These are

$$\frac{\partial f_1}{\partial n_\sigma} = -2C\left(\frac{\tau_\uparrow^{\mathrm{W}} + \tau_\downarrow^{\mathrm{W}}}{\tau_\uparrow + \tau_\downarrow}\right)\left(\frac{\tau_\sigma^{\mathrm{W}}}{\tau_\uparrow + \tau_\downarrow}\right)\frac{1}{n_\sigma} \tag{C30}$$

$$\frac{\partial f_{2,\sigma}}{\partial n_\sigma} = -2(1 + C)\left(\frac{\tau_\sigma^{\mathrm{W}}}{\tau_\sigma}\right)^2 \frac{1}{n_\sigma}. \tag{C31}$$

For the correlation part of the PKZB functional we implemented the derivative of $\epsilon_{\mathrm{c}}^{\mathrm{PKZB}}$ with respect to ∇n_σ directly, rather than via $\partial \epsilon_{\mathrm{c}}^{\mathrm{PKZB}}/\partial|\nabla n_\sigma|$. This was done in order to make use of the existing code for the spin-polarized PBE functional in ONETEP, which outputs derivatives with respect to $|\nabla n|$, rather than the spin-components of this, $|\nabla n_\sigma|$. Derivatives with respect to the spin components of the gradient of the density, ∇n_σ, can be obtained from derivatives with respect to the magnitude of the gradient of the total density, $|\nabla n|$, using

$$\frac{\partial \epsilon_{\mathrm{c}}^{\mathrm{PBE}}}{\partial \nabla n_\sigma} = \frac{\partial \epsilon_{\mathrm{c}}^{\mathrm{PBE}}}{\partial |\nabla n|}\frac{\partial |\nabla n|}{\partial \nabla n_\sigma} = \frac{\partial \epsilon_{\mathrm{c}}^{\mathrm{PBE}}}{\partial |\nabla n|}\frac{\nabla n}{|\nabla n|}. \tag{C32}$$

The derivative of $\epsilon_{\mathrm{c}}^{\mathrm{PKZB}}$ with respect to ∇n_σ is identical to Eq. C29, but with ∂n_σ replaced with $\partial \nabla n_\sigma$. The required derivatives of f_1 and $f_{2,\sigma}$ with respect to ∇n_σ are

$$\frac{\partial f_1}{\partial \nabla n_\sigma} = 2C\left(\frac{\tau_\uparrow^{\mathrm{W}} + \tau_\downarrow^{\mathrm{W}}}{\tau_\uparrow + \tau_\downarrow}\right)\left(\frac{1}{\tau_\uparrow + \tau_\downarrow}\right)\frac{\partial \tau_\sigma^{\mathrm{W}}}{\partial \nabla n_\sigma}, \tag{C33}$$

$$\frac{\partial f_{2,\sigma}}{\partial \nabla n_\sigma} = 2(1 + C)\left(\frac{\tau_\sigma^{\mathrm{W}}}{\tau_\sigma}\right)\left(\frac{1}{\tau_\sigma}\right)\frac{\partial \tau_\sigma^{\mathrm{W}}}{\partial \nabla n_\sigma}, \tag{C34}$$

where

$$\frac{\partial \tau_\sigma^{\mathrm{W}}}{\partial \nabla n_\sigma} = \frac{1}{4}\frac{\nabla n_\sigma}{n_\sigma}. \tag{C35}$$

Finally, the derivative of $\epsilon_{\mathrm{c}}^{\mathrm{PKZB}}$ with respect to τ_σ is

$$\frac{\partial \epsilon_{\mathrm{c}}^{\mathrm{PKZB}}}{\partial \tau_\sigma} = \epsilon_{\mathrm{c},\uparrow\downarrow}^{\mathrm{PBE}} \frac{\partial f_1}{\partial \tau_\sigma} - \epsilon_{\mathrm{c},\sigma}^{\mathrm{PBE}} \frac{\partial f_{2,\sigma}}{\partial \tau_\sigma} \tag{C36}$$

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

which has fewer terms than Eq. C29 because $\epsilon_c^{\mathrm{PBE}}$ is independent of τ_σ. The required derivatives of f_1 and $f_{2,\sigma}$ are

$$\frac{\partial f_1}{\partial \tau_\sigma} = -2C\left(\frac{\tau_\uparrow^{\mathrm{W}} + \tau_\downarrow^{\mathrm{W}}}{\tau_\uparrow + \tau_\downarrow}\right)^2\left(\frac{1}{\tau_\uparrow + \tau_\downarrow}\right), \tag{C37}$$

$$\frac{\partial f_{2,\sigma}}{\partial \tau_\sigma} = -2(1+C)\left(\frac{\tau_\sigma^{\mathrm{W}}}{\tau_\sigma}\right)^2\left(\frac{1}{\tau_\sigma}\right). \tag{C38}$$

It is noteworthy that Eq. C37 is independent of σ, and thus need only be evaluated once for both spin-components.

[1] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[2] J. P. Perdew and K. Schmidt, in *AIP Conference Proceedings*, edited by V. Van Doren, C. Van Alsenoy, and P. Geerlings, volume 577, pages 1–20 (Antwerp (Belgium), 2001, AIP Publishing).
[3] J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria, and G. I. Csonka, J. Chem. Phys. **123**, 062201 (2005).
[4] The term "meta-GGA" has historically been applied to exchange-correlation functionals which depend upon the Laplacian of the density, $\nabla^2 n$, and/or the kinetic energy density, τ (see, for example, Refs. 2, 3). In the context of this paper we follow the more recent convention of considering τ-dependence as the defining feature of the meta-GGA family (as in Refs. 19, 20).
[5] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).
[6] D. R. Bowler and T. Miyazaki, Rep. Prog. Phys. **75**, 036503 (2012).
[7] W. Kohn, Phys. Rev. Lett. **76**, 3168 (1996).
[8] C.-K. Skylaris, P. D. Haynes, A. A. Mostofi, and M. C. Payne, J. Chem. Phys. **122**, 084119 (2005).
[9] S. Mohr, L. E. Ratcliff, L. Genovese, D. Caliste, P. Boulanger, S. Goedecker, and T. Deutsch, Phys. Chem. Chem. Phys. **17**, 31360 (2015).
[10] M. J. Gillan, D. R. Bowler, A. S. Torralba, and T. Miyazaki, Comput. Phys. Commun. **177**, 14 (2007).
[11] T. Ozaki and H. Kino, Phys. Rev. B **72**, 045121 (2005).
[12] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, Comput. Phys. Commun. **167**, 103 (2005).
[13] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and Daniel Sánchez-Portal, J. Phys.: Condens. Matter **14**, 2745 (2002).
[14] S. J. Fox, J. Dziedzic, T. Fox, C. S. Tautermann, and C.-K. Skylaris, Proteins **82**, 3335 (2014).
[15] G. Lever, D. J. Cole, R. Lonsdale, K. E. Ranaghan, D. J. Wales, A. J. Mulholland, C.-K. Skylaris, and M. C. Payne, J. Phys. Chem. Lett. **5**, 3614 (2014).
[16] G. C. Constantinescu and N. D. M. Hine, Nano Lett. **16**, 2586 (2016).
[17] E. Poli, J. D. Elliott, L. E. Ratcliff, L. Andrinopoulos, J. Dziedzic, N. D. M. Hine, A. A. Mostofi, C.-K. Skylaris, P. D. Haynes, and G. Teobaldi, J. Phys.: Condens. Matter **28**, 074003 (2016).
[18] J. Dziedzic, Q. Hill, and C.-K. Skylaris, J. Chem. Phys. **139**, 214103 (2013).
[19] N. Mardirossian and M. Head-Gordon, J. Chem. Phys. **142**, 074111 (2015).
[20] J. Sun, A. Ruzsinszky, and J. P. Perdew, Phys. Rev. Lett. **115**, 036402 (2015).
[21] J. Sun, R. C. Remsing, Y. Zhang, Z. Sun, A. Ruzsinszky, H. Peng, Z. Yang, A. Paul, U. Waghmare, X. Wu, M. L. Klein, and J. P. Perdew, Nat. Chem. **advance online publication** (2016), doi: 10.1038/nchem.2535.
[22] J. P. Perdew, S. Kurth, A. Zupan, and P. Blaha, Phys. Rev. Lett. **82**, 2544 (1999).
[23] J. F. Janak, Phys. Rev. B **18**, 7165 (1978).
[24] S. Ismail-Beigi and T. A. Arias, Phys. Rev. Lett. **82**, 2127 (1999).
[25] C.-K. Skylaris, A. A. Mostofi, P. D. Haynes, O. Diéguez, and M. C. Payne, Phys. Rev. B **66**, 035119 (2002).
[26] A. A. Mostofi, C.-K. Skylaris, P. D. Haynes, and M. C. Payne, Comput. Phys. Commun. **147**, 788 (2002).
[27] A. A. Mostofi, P. D. Haynes, C.-K. Skylaris, and M. C. Payne, J. Chem. Phys. **119**, 8842 (2003).
[28] C.-K. Skylaris, A. A. Mostofi, P. D. Haynes, C. J. Pickard, and M. C. Payne, Comput. Phys. Commun. **140**, 315 (2001).
[29] C.-K. Skylaris, P. D. Haynes, A. A. Mostofi, and M. C. Payne, Phys. Status Solidi B **243**, 973 (2006).
[30] M. R. Hestenes and E. Stiefel, J. Res. Nat. Bur. Stand. **49**, 409 (1952).
[31] P. D. Haynes, C. K. Skylaris, A. A. Mostofi, and M. C. Payne, Chem. Phys. Lett. **422**, 345 (2006).
[32] X.-P. Li, R. W. Nunes, and D. Vanderbilt, Phys. Rev. B **47**, 10891 (1993).
[33] R. W. Nunes and D. Vanderbilt, Phys. Rev. B **50**, 17611 (1994).
[34] P. D. Haynes, C.-K. Skylaris, A. A. Mostofi, and M. C. Payne, J. Phys.: Condens. Matter **20**, 294207 (2008).
[35] N. D. M. Hine, P. D. Haynes, A. A. Mostofi, C. K. Skylaris, and M. C. Payne, Comput. Phys. Commun. **180**, 1041 (2009).
[36] K. A. Wilkinson, N. D. M. Hine, and C.-K. Skylaris, J. Chem. Theory Comput. **10**, 4782 (2014).
[37] K. Wilkinson and C.-K. Skylaris, J. Comput. Chem. **34**, 2446 (2013).
[38] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).
[39] A. D. Becke, J. Chem. Phys. **96**, 2155 (1992).
[40] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[41] A. D. Becke, Int. J. Quantum Chem. **23**, 1915 (1983).
[42] A. D. Becke, J. Chem. Phys. **88**, 1053 (1988).
[43] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **78**, 1396 (1997).
[44] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys.

Rev. Lett. **91**, 146401 (2003).
[45] S. Kurth, J. P. Perdew, and P. Blaha, Int. J. Quantum Chem. **75**, 889 (1999).
[46] A. D. Becke, J. Chem. Phys. **104**, 1040 (1996).
[47] A. D. Becke, J. Chem. Phys. **109**, 2092 (1998).
[48] T. Van Voorhis and G. E. Scuseria, J. Chem. Phys. **109**, 400 (1998).
[49] A. D. Becke and K. E. Edgecombe, J. Chem. Phys. **92**, 5397 (1990).
[50] H. L. Schmider and A. D. Becke, J. Mol. Struct.: THEOCHEM **527**, 51 (2000).
[51] A. D. Becke, Int. J. Quantum Chem. **27**, 585 (1985).
[52] A. D. Becke and M. R. Roussel, Phys. Rev. A **39**, 3761 (1989).
[53] Y. Zhao and D. G. Truhlar, J. Chem. Phys. **125**, 194101 (2006).
[54] R. Peverati and D. G. Truhlar, J. Phys. Chem. Lett. **3**, 117 (2012).
[55] H. S. Yu, X. He, and D. G. Truhlar, J. Chem. Theory Comput. (2016).
[56] Y. Zhao and D. G. Truhlar, Theor. Chem. Acc. **120**, 215 (2007).
[57] R. Peverati and D. G. Truhlar, J. Phys. Chem. Lett. **2**, 2810 (2011).
[58] H. S. Yu, X. He, S. L. Li, and D. G. Truhlar, Chem. Sci. **7**, 5032 (2016).
[59] N. Mardirossian and M. Head-Gordon, J. Chem. Phys. **144**, 214110 (2016).
[60] S. Kümmel and L. Kronik, Rev. Mod. Phys. **80**, 3 (2008).
[61] M. Ernzerhof and G. E. Scuseria, J. Chem. Phys. **111**, 911 (1999).
[62] R. T. Sharp and G. K. Horton, Phys. Rev. **90**, 317 (1953).
[63] J. B. Krieger, Y. Li, and G. J. Iafrate, Phys. Rev. A **45**, 101 (1992).
[64] A. V. Arbuznikov and M. Kaupp, Chem. Phys. Lett. **381**, 495 (2003).
[65] Z.-h. Yang, H. Peng, J. Sun, and J. P. Perdew, Phys. Rev. B **93**, 205205 (2016).
[66] R. Neumann, R. H. Nobes, and N. C. Handy, Mol. Phys. **87**, 1 (1996).
[67] A. Seidl, A. Görling, P. Vogl, J. A. Majewski, and M. Levy, Phys. Rev. B **53**, 3764 (1996).
[68] A. V. Arbuznikov, M. Kaupp, V. G. Malkin, R. Reviakine, and O. L. Malkina, Phys. Chem. Chem. Phys. **4**, 5467 (2002).
[69] J. Sun, M. Marsman, G. I. Csonka, A. Ruzsinszky, P. Hao, Y.-S. Kim, G. Kresse, and J. P. Perdew, Phys. Rev. B **84**, 035117 (2011).
[70] Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuś, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. W. III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. D. Jr, H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. W. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. G. III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. S. III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon, Mol. Phys. **113**, 184 (2015).
[71] In Q-Chem, the FDO approach is derived in terms of matrix elements over V_{xc} (as in Neumann, Nobes and Handy's original paper [66]), rather than directly in terms of functional derivatives (as in this work). These approaches are equivalent and lead to the same orbital-specific exchange-correlation potential.
[72] F. Zahariev, S. S. Leang, and M. S. Gordon, J. Chem. Phys. **138**, 244108 (2013).
[73] A. A. Mostofi, Ph. D., University of Cambridge, Christ's College, Cambridge, 2003.
[74] J. P. Perdew, S. Kurth, A. Zupan, and P. Blaha, Phys. Rev. Lett. **82**, 5179 (1999).
[75] V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew, J. Chem. Phys. **119**, 12129 (2003).
[76] J. P. Perdew, J. Tao, V. N. Staroverov, and G. E. Scuseria, J. Chem. Phys. **120**, 6898 (2004).
[77] M. A. L. Marques, M. J. T. Oliveira, and T. Burnus, Comput. Phys. Commun. **183**, 2272 (2012).
[78] M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, and W. A. de Jong, Comput. Phys. Commun. **181**, 1477 (2010).
[79] O. A. Vydrov and T. Van Voorhis, J. Chem. Phys. **133**, 244103 (2010).
[80] N. Mardirossian and M. Head-Gordon, J. Chem. Phys. **140**, 18A527 (2014).
[81] A. D. Becke, J. Chem. Phys. **107**, 8554 (1997).
[82] A. D. Becke, J. Chem. Phys. **112**, 4020 (2000).
[83] S. Grimme, J. Comput. Chem. **27**, 1787 (2006).
[84] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).
[85] S. Grimme, S. Ehrlich, and L. Goerigk, J. Comput. Chem. **32**, 1456 (2011).
[86] N. Mardirossian and M. Head-Gordon, Phys. Chem. Chem. Phys. **16**, 9904 (2014).
[87] D. J. Cole, C.-K. Skylaris, E. Rajendra, A. R. Venkitaraman, and M. C. Payne, EPL **91**, 37004 (2010).
[88] S. Fox, H. G. Wallnoefer, T. Fox, C. S. Tautermann, and C.-K. Skylaris, J. Chem. Theory Comput. **7**, 1102 (2011).
[89] C. Sampson, T. Fox, C. S. Tautermann, C. Woods, and C.-K. Skylaris, J. Phys. Chem. B **119**, 7030 (2015).
[90] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **92**, 246401 (2004).
[91] K. Lee, É. D. Murray, L. Kong, B. I. Lundqvist, and D. C. Langreth, Phys. Rev. B **82**, 081101 (2010).
[92] J. Klimeš and A. Michaelides, J. Chem. Phys. **137**, 120901

(2012).
[93] R. Sabatini, T. Gorni, and S. de Gironcoli, Phys. Rev. B **87**, 041108 (2013).
[94] G. Román-Pérez and J. M. Soler, Phys. Rev. Lett. **103**, 096102 (2009).
[95] L. Andrinopoulos, Ph. D., Imperial College London, 2013.
[96] J. Kong, C. A. White, A. I. Krylov, D. Sherrill, R. D. Adamson, T. R. Furlani, M. S. Lee, A. M. Lee, S. R. Gwaltney, T. R. Adams, C. Ochsenfeld, A. T. B. Gilbert, G. S. Kedziora, V. A. Rassolov, D. R. Maurice, N. Nair, Y. Shao, N. A. Besley, P. E. Maslen, J. P. Dombroski, H. Daschel, W. Zhang, P. P. Korambath, J. Baker, E. F. C. Byrd, T. Van Voorhis, M. Oumi, S. Hirata, C.-P. Hsu, N. Ishikawa, J. Florian, A. Warshel, B. G. Johnson, P. M. W. Gill, M. Head-Gordon, and J. A. Pople, J. Comput. Chem. **21**, 1532 (2000).
[97] Y. Shao, L. F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. D. Jr, R. C. Lochan, T. Wang, G. J. O. Beran, N. A. Besley, J. M. Herbert, C. Y. Lin, T. Van Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. W. Iii, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. S. Iii, J. Kong, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon, Phys. Chem. Chem. Phys. **8**, 3172 (2006).
[98] F. Weigend and R. Ahlrichs, Phys. Chem. Chem. Phys. **7**, 3297 (2005).
[99] D. Rappoport and F. Furche, J. Chem. Phys. **133**, 134105 (2010).
[100] P. M. W. Gill, B. G. Johnson, and J. A. Pople, Chem. Phys. Lett. **209**, 506 (1993).
[101] For details on the construction of integration grids in Q-Chem, see the user manual, available at `http://www.q-chem.com/`.
[102] D. P. Landau, H. B. Schuttler, S. Lewis, M. Bachmann, and J. W. Bennett, Phys. Procedia **34**, 14 (2012).
[103] A ONETEP-compatible version of the Rappe-Bennett library is available to download from the CASTEP project page on CCPForge: `https://ccpforge.cse.rl.ac.uk/gf/project/castep/`.
[104] Opium pseudopotential generation project, `http://opium.sourceforge.net/`.
[105] A. M. Rappe, K. M. Rabe, E. Kaxiras, and J. D. Joannopoulos, Phys. Rev. B **41**, 1227 (1990).
[106] A. M. Rappe, K. M. Rabe, E. Kaxiras, and J. D. Joannopoulos, Phys. Rev. B **44**, 13175 (1991).
[107] This pseudopotential is part of a suite of norm-conserving pseudopotentials generated by K. Refson to supplement the Rappe-Bennett library, available to download from the CASTEP project page on CCPForge: `https://ccpforge.cse.rl.ac.uk/gf/project/castep/`.
[108] J. Řezáč, K. E. Riley, and P. Hobza, J. Chem. Theory Comput. **8**, 4285 (2012).
[109] K. L. Copeland and G. S. Tschumper, J. Chem. Theory Comput. **8**, 1646 (2012).
[110] J. J. Wilke, M. C. Lind, H. F. Schaefer, A. G. Császár, and W. D. Allen, J. Chem. Theory Comput. **5**, 1511 (2009).
[111] L. Goerigk and S. Grimme, J. Chem. Theory Comput. **7**, 291 (2011).
[112] U. R. Fogueri, S. Kozuch, A. Karton, and J. M. Martin, J. Phys. Chem. A **117**, 2269 (2013).
[113] J. T. Berryman, S. E. Radford, and S. A. Harris, Biophys. J. **97**, 1 (2009).
[114] J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, WIREs Comput Mol Sci **4**, 15 (2014).
[115] G. Lippert, J. Hutter, and M. Parrinello, Mol. Phys. **92**, 477 (1997).
[116] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[117] N. D. M. Hine, J. Phys. Condens.: Matter (2016), Under review.
[118] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, Oxford, 1994).
[119] E. Engel and R. M. Dreizler, *Density Functional Theory*, Theoretical and Mathematical Physics (Springer Berlin Heidelberg, Berlin, Heidelberg, 2011).
[120] G. L. Oliver and J. P. Perdew, Phys. Rev. A **20**, 397 (1979).