Hybrid *ab initio* Kohn-Sham density functional theory/frozen-density orbital-free density functional theory simulation method suitable for biological systems

Miroslav Hodak^{a)}

Center for High Performance Simulation and Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-7518, USA

Wenchang Lu and J. Bernholc

Center for High Performance Simulation and Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-7518, USA and CSMD and CNMS, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6367, USA

(Received 13 September 2007; accepted 25 October 2007; published online 2 January 2008)

A hybrid computational method intended for simulations of biomolecules in solution is described. The *ab initio* Kohn-Sham (KS) density functional theory (DFT) method is used to describe the chemically active part of the system and its first solvation shells, while a frozen-density orbital-free (FDOF) DFT method is used to treat the rest of the solvent. The molecules in the FDOF method have fixed internal structures and frozen electron densities. The hybrid method provides a seamless description of the boundary between the subsystems and allows for the flow of molecules across the boundary. Tests on a liquid water system show that the total energy is conserved well during molecular dynamics and that the effect of the solvent environment on the KS subsystem is well described. An initial application to copper ion binding to the prion protein is also presented. © 2008 American Institute of Physics. [DOI: [10.1063/1.2814165](https://doi.org/10.1063/1.2814165)]

I. INTRODUCTION

Ab initio density functional theory¹ (DFT) calculations, originally formulated by Kohn and Sham,² are widely used for computational investigations of a variety of systems. With the recent improvements in algorithms and computing power, the Kohn-Sham (KS) DFT can treat systems containing several hundreds of atoms. This, however, is still not enough for studies on biological systems consisting of many thousands of atoms. Such systems are usually treated using classical molecular dynamics with force fields, but such simulations cannot accurately describe chemical processes such as bond breaking, bond creation, or charge transfer.

Of particular interest for biological systems is the treatment of solvent molecules. The solvent molecules can either be treated explicitly or implicitly. In implicit methods, the solvent is represented as a polarizable continuum dielectric.^{3–5} The attractiveness of this approach stems from the fact that in many biosimulations, solvent molecules outnumber the solute and thus implicit modeling of the solvent can substantially improve the efficiency of simulations. This approach was successfully used^{6–9} in many instances, but evidence suggests that for a complete understanding of the behavior of many biological systems, an explicit treatment of the solvent is essential. For example, molecular mechanics simulations of proteins have shown that an implicit representation of the solvent leads to incorrect predictions of native structures and folding pathways, while explicit treatment of the solvent yields results in agreement with experimental

observation.^{10,11} Also, for many solvated biomolecules, water molecules are located near the active sites and are essential to their functions. To properly describe such water molecules, hydrogen bonds with other water molecules have to be included and this can only be done when the solvent is treated explicitly.

In *ab initio* simulations, it is not feasible to represent the solvent explicitly and thus implicit solvation is often used. The polarizable continuum model^{12,13} is particularly popular. Explicit description of the solvent can only be used when the solvent molecules are represented in an efficient manner, such as at a lower level of theory than the solute. The quantum molecular (QM)/molecular mechanics (MM) method,^{14–16} which combines *ab initio* and classical force field simulations, is often used to achieve full explicit solvation in studies on biomolecules. In this method, the chemically active part of the system is treated at the *ab initio* level, while the rest of the system, including the solvent and the remainder of the solute, is treated with classical force field methods. The weak point of QM/MM schemes is the treatment of the interaction between the QM and MM subsystems. A special energy term is required to describe this interaction. This term usually employs force field parameters, which were developed and tested in calculations within classical molecular mechanics and not for interactions between the QM and MM subsystems. Furthermore, to allow for the flow of solvent molecules across the QM/MM interface, resulting in exchange of molecules between the subsystems, buffer regions and force smoothing functions have to be used.¹⁷

In this paper, we explore a different approach to hybrid

^{a)}Electronic mail: hodak@chips.ncsu.edu.

simulation: combining two density functional methods. The chemically active part of the system is treated with the usual *ab initio* Kohn-Sham DFT, while a frozen-density orbital-free (FDOF) DFT method, which scales linearly with the number of atoms, is used for the rest of the system. In this method, the solvent molecules are assumed to have rigid geometries and frozen electron densities. The advantage of combining two DFT methods is that interactions between the subsystems can be treated at the quantum level rather than at the level of molecular mechanics, as is usually the case in QM/MM methods. In addition, the KS/FDOF hybrid method provides seamless integration between the subsystems and allows for the flow of the molecules across the KS/FDOF interface without the need for a buffer region or force smoothing. Within the KS/FDOF method, energy conserving molecular dynamics simulations can be run even when molecules are exchanged between the subsystems. Tests on liquid water at ambient conditions show that KS water molecules have appropriate dipole moments, proving that the effects of the solvent environment are well described. As an example of a realistic application, the KS/FDOF method is used to study copper ion binding to the prion protein. It is found that improved description of solvation effects leads to refinement of the binding site geometry.

This paper is organized as follows: Section II defines the frozen-density orbital-free DFT method and Sec. III describes its application and tests on a homogeneous system of liquid water. In Secs. IV and V, the hybrid KS/FDOF method is formulated and its implementation is discussed. Sections VI and VII present applications to liquid water and to copper ion binding to the prion protein. Finally, Sec. VIII gives the summary and conclusions.

II. FROZEN-DENSITY ORBITAL-FREE DFT SIMULATION METHOD

The KS energy functional is usually written as

$$E[\rho] = T[\rho] + E_H[\rho] + E_{XC}[\rho] + E_{Ac}[\rho] + E_{AA}, \quad (1)$$

where ρ is the charge density, $T[\rho]$ is the noninteracting kinetic energy (KE) functional, $E_H[\rho]$ is the Hartree energy, $E_{XC}[\rho]$ is the exchange-correlation energy, $E_{Ac}[\rho]$ is the atom-electron interaction energy, and E_{AA} is the energy of interaction between atoms. By introducing a set of orbitals ψ_i satisfying

$$\rho = \sum_i |\psi_i|^2 \quad (2)$$

and defining kinetic energy as

$$T[\rho] = -\frac{1}{2} \sum_i \langle \psi_i | \nabla^2 | \psi_i \rangle, \quad (3)$$

the ground state electron density can be obtained by solving a set of single-particle (Kohn-Sham) equations for orbitals ψ_i . The introduction of orbitals, while very successful, makes the KS DFT calculations computationally demanding: The orbitals have to be evaluated, stored, and kept orthogonal throughout the calculation. Because of this, there has been an interest in developing orbital-free (OF) DFT methods. The

orbital-free approaches eliminate orbitals and attempt to find density by directly minimizing the energy functional. These methods are significantly more efficient than the KS approach and scale linearly with the system size. They can thus be used to study much larger systems than orbital-based methods.

An OF DFT requires that the KE functional in Eq. (1) depends explicitly on the charge density ρ . The simplest such functional is the Thomas-Fermi^{18–20} (TF) KE functional

$$T_{TF} = C_{TF} \int_{\Omega} \rho^{5/3}(\mathbf{r}) d\mathbf{r}, \quad (4)$$

where C_{TF} is $\frac{3}{10}(3\pi^2)^{2/3}$ and Ω denotes the volume of the simulation cell. The von Weizsacker KE functional²¹ adds the dependence on the gradient of charge density to the TF formula:

$$T_W = T_{TF} + \frac{1}{\lambda} \frac{1}{8} \int_{\Omega} \frac{|\nabla \rho(\mathbf{r})|^2}{\rho(\mathbf{r})} d\mathbf{r}, \quad (5)$$

where λ is a constant. For $\lambda=9$, the T_W is the second-order gradient expansion of the Thomas-Fermi KE functional, but other values are also used.²² These functionals, however, yield poor results for the ground state charge density.²³ Various approaches for designing better KE functionals have been suggested, but a highly accurate form of the KE remains elusive. Following the success of the generalized-gradient approximation (GGA) approach in exchange-correlation energy functional, a lot of effort^{24–29} went into the development of GGA KE functionals. For example, Lee *et al.*²⁸ formulated a KE functional (LLP functional) which is an equivalent of the B88 (Ref. 30) exchange-correlation functional,

$$T_{LLP} = C_{TF} \int_{\Omega} \rho^{5/3} \left(1 + \frac{0.0045x^2}{1 + 0.0253x \sinh^{-1} x} \right) d\mathbf{r}, \quad (6)$$

where

$$x = 2^{1/3} \frac{|\nabla \rho(\mathbf{r})|}{\rho^{4/3}(\mathbf{r})}. \quad (7)$$

While the GGA approach improved the quality of OF KE functionals, the obtained ground state charge densities are still not satisfactory.²³ Using the nonlocal density approximation,^{31–35} even more accurate OF KEs have been obtained. These, however, break²³ linear scaling of the OF DFT, thus limiting the system sizes to which they can be applied.

Because of the difficulties in obtaining the correct ground state charge densities within OF DFT, the molecules treated by an OF method in this work have frozen charge densities and fixed internal geometries, which do not change throughout the simulation. The OF DFT is thus only used to calculate interactions within the system and not for obtaining ground state densities. Frozen-density approximations have been frequently utilized in DFT methods: Kim and Gordon used it to study interactions between noble gas atoms³⁶ and to investigate properties of solids,³⁷ Wesolowski and Warshel^{38,39} used it to calculate free energies of solvation,

Barker and Sprik⁴⁰ used it to study liquid water, and recently Huang and Carter⁴¹ used it to study Cu and Co atoms on the (111) surface of Cu.

To efficiently treat interactions between electrons and atoms, it is desirable to make use of pseudopotentials. The *ab initio* pseudopotentials^{42–44} used in KS DFT calculations contain terms dependent on angular momentum that are applied to the KS orbitals. Most of the currently used KS pseudopotentials, such as the Kleinman-Bylander⁴⁵ or Vanderbilt ultrasoft pseudopotentials,⁴⁶ are usually of separable type, i.e., they are written as a sum of a short-range angular-dependent nonlocal part and a long-range spherically symmetric local part. Since the nonlocal part acts on the orbitals of a system rather than on its density, these pseudopotentials cannot be directly used in OF methods. To enable use of the pseudopotentials in OF methods, one of two approaches is typically used: (i) Special local pseudopotentials that are directly applicable to charge density are generated^{47,48} or (ii) existing orbital-based pseudopotentials are used by defining orbitals corresponding to the charge density^{49–51} (usually a single orbital $\psi = \rho^{1/2}$ is used). In our model, the standard orbital-based pseudopotentials are utilized, but only their local parts are applied. Neglecting the nonlocal contributions is well justified because the present model only deals with intermolecular interactions and the short-range nonlocal effects are negligible at typical intermolecular distances.

The total electronic energy functional for the present FDOF DFT model is thus

$$E_{\text{FDOF}}[\rho] = T_{\text{OF}}[\rho] + E_H[\rho] + E_{\text{XC}}[\rho] + E_{\text{PP}_{\text{local}}}[\rho] + E_{\text{AA}}, \quad (8)$$

where T_{OF} stands for any of the OF KE functionals mentioned above and $E_{\text{PP}_{\text{local}}}$ evaluates the energy of interaction between the electrons and the local part of the pseudopotentials. This expression is formally very similar to the KS energy functional, Eq. (1). The only differences are the type of the KE functional and the fact that the FDOF method only uses the local part of the pseudopotentials.

The implementation of the FDOF method was based on previously existing KS DFT real-space code^{52–54} with multigrid acceleration. This code is well tested and exhibits excellent performance on massively parallel supercomputers. In particular, the multigrid Hartree solver and the functions that calculate exchange-correlation and local pseudopotential energies were utilized directly in our implementation. The RATTLE algorithm⁵⁵ was used to keep intramolecular geometries fixed.

III. APPLICATION OF FROZEN-DENSITY ORBITAL-FREE DFT METHOD TO LIQUID WATER

Water is the most ubiquitous solvent. In particular, its effects must be accounted for when studying biological systems. In fact, in many biosimulations, the number of atoms in the solvent is much higher than the number of atoms in the solute. It is therefore attractive to formulate a FDOF DFT calculation that would allow for an efficient and accurate treatment of liquid water.

The geometry of water molecules used in the present FDOF method was obtained from KS DFT simulation of liquid water at 300 K. The distance between the oxygen and hydrogen atoms in a water molecule was found to be 0.98 Å and the HOH angle was 105.1°. These values are very close to previously reported KS DFT results:⁵⁶ $r_{\text{OH}} = 0.991$ Å, $\angle \text{HOH} = 105.5^\circ$.

For parametrization of the charge density of a water molecule, a functional form proposed by Barker and Sprik⁴⁰ is used—Gaussians centered on atoms in the molecule. For simplicity, one Gaussian per atom is used. The charge density of a water molecule is then

$$\rho_{\text{H}_2\text{O}}(\mathbf{r}) = G(q_{\text{O}}, \alpha_{\text{O}}, \mathbf{R}_{\text{O}}, \mathbf{r}) + G(q_{\text{O}}, \alpha_{\text{O}}, \mathbf{R}_{\text{H1}}, \mathbf{r}) + G(q_{\text{O}}, \alpha_{\text{O}}, \mathbf{R}_{\text{H2}}, \mathbf{r}), \quad (9)$$

where G is a Gaussian function centered at the atomic position $\mathbf{R}$,

$$G(q, \alpha, \mathbf{R}, \mathbf{r}) = q \left(\frac{\alpha}{\pi} \right)^{3/2} \exp(-\alpha |\mathbf{r} - \mathbf{R}|^2). \quad (10)$$

This leaves four parameters that need to be determined for the water molecule: the charges on the oxygen and the hydrogen (q_{O} and q_{H}) and the Gaussian widths (α_{O} and α_{H}). The atomic charges can be determined from the total molecular charge and dipole moment. When using pseudopotentials, the total charge of a water molecule is 8 and, therefore,

$$q_{\text{O}} + 2q_{\text{H}} = 8. \quad (11)$$

An important characteristic of a water molecule is its dipole moment. In gas phase, its value⁵⁷ is 1.85 D, but in liquid phase, it increases significantly. The correct value is still under discussion and the estimates range from 2.6 to 3.15 D.^{58–61} In this work, a value of 3.0 D is used, which is close to the value of 2.95 D obtained from KS DFT studies.^{56,62} Because the electron density is given by Gaussians, the dipole moment only depends on the oxygen and hydrogen charges q_{O} and q_{H} and not on the Gaussian widths. This gives the second condition for the charges and using the chosen value of 3.0 D for the dipole moment, q_{O} becomes 7.05 electrons and q_{H} becomes 0.475 electron.

To adjust the Gaussian widths, molecular dynamics simulations were run on a system of 432 water molecules at 300 K. The radial distribution functions (RDFs) were calculated for various values of parameters α_{O} and α_{H} and compared to the experimental RDFs.⁶³ Using the PBE exchange-correlation potential^{64,65} and Vanderbilt ultrasoft pseudopotentials,⁴⁶ various types of KE functionals were also tested. It was found that the LLP functional and values of $\alpha_{\text{O}} = 0.75$ and $\alpha_{\text{H}} = 0.8$ reproduce the experimental RDFs best. The obtained RDFs and their comparison to experimental results are shown in Fig. 1. The agreement between the calculated and experimental RDFs is not perfect, but it is comparable to the agreement between experiments and *ab initio* DFT simulations.^{56,62,66–68} Furthermore, the goal of the FDOF method is not to achieve a perfect description of liquid water. Instead, this procedure was designed to provide a reasonably good description of water *and* to be easily embeddable in a hybrid calculation method. Obviously, more

FIG. 1. (Color online) Radial distribution functions calculated using the Thomas-Fermi-like method (solid lines) for a system of 432 water molecules at 300 K. Experimental results (Ref. 63) are also shown (dashed lines). For better comparison, intramolecular features in H-H and O-H plots were removed from experimental data.

complex models of water can be formulated, which would further improve the agreement with experiment.

IV. KS DFT/FDOF DFT HYBRID METHOD

The energy functionals for KS and FDOF DFTs, Eqs. (1) and (8), are only different in the type of KE functional and the fact that the FDOF method does not use the nonlocal pseudopotential energy. Because of this similarity, it is relatively easy to write an expression for the total electronic energy of a hybrid KS DFT/FDOF DFT system.

To do this, consider a system that contains two types of atoms: those that are treated with the FDOF method (FDOF atoms) and those that are treated with the KS DFT method (KS atoms). The FDOF atoms have fixed charge density $\varrho_{\text{FDOF}}(\mathbf{r})$ attached to them, while the charge density corresponding to the KS atoms, $\varrho_{\text{KS}}(\mathbf{r})$, minimizes the energy functional for the hybrid system. The total charge density of the entire system is simply the sum of both densities,

$$\varrho_{\text{tot}}(\mathbf{r}) = \varrho_{\text{FDOF}}(\mathbf{r}) + \varrho_{\text{KS}}(\mathbf{r}). \quad (12)$$

Since the Hartree and exchange-correlation energies are explicit functions of the charge density, it is straightforward to calculate these quantities for the hybrid system using ϱ_{tot} . The atom-atom interaction energy is also straightforward, being the sum over all atoms present in the hybrid system, regardless of their type. As for the pseudopotential energy, the local and nonlocal parts are treated differently: Each atom has a local potential associated with it and it acts on the total density ϱ_{tot} , while only the KS atoms have nonlocal potentials associated with them and these only act on the KS orbitals.

Defining KE of the hybrid system is somewhat more

complicated. Taking the sum of the FDOF and KS KEs is not sufficient, since kinetic energies are nonadditive. The nonadditive part of KE can be estimated³⁸ as

$$T_{\text{nadd}}[\varrho_{\text{FDOF}}, \varrho_{\text{KS}}] = T_{\text{OF}}[\varrho_{\text{tot}}] - T_{\text{OF}}[\varrho_{\text{FDOF}}] - T_{\text{OF}}[\varrho_{\text{KS}}]. \quad (13)$$

In this and the following formulas, the spatial coordinates are omitted for clarity.

This gives the following functional for the total electron energy of a hybrid system:

$$\begin{aligned} E_{\text{hybrid}}[\varrho_{\text{FDOF}}, \varrho_{\text{KS}}, \{\mathbf{R}_i\}_{\text{FDOF}}, \{\mathbf{R}_j\}_{\text{KS}}] \\ = T_{\text{TF}}[\varrho_{\text{TF}}] + T_{\text{KS}}[\varrho_{\text{KS}}] + T_{\text{nadd}}[\varrho_{\text{FDOF}}, \varrho_{\text{KS}}] + E_{\text{H}}[\varrho_{\text{tot}}] \\ + E_{\text{XC}}[\varrho_{\text{tot}}] + E_{\text{PP}_{\text{nonloc}}}[\varrho_{\text{KS}}, \{\mathbf{R}_j\}_{\text{KS}}] + E_{\text{PP}_{\text{loc}}}[\varrho_{\text{tot}}, \{\mathbf{R}_i\}] \\ + E_{\text{AA}}[\{\mathbf{R}_i\}]. \end{aligned} \quad (14)$$

Here $\{\mathbf{R}_i\}$ denotes all atoms in the system and $\{\mathbf{R}_j\}_{\text{KS}}$ stands for atoms treated by the KS DFT. Note that in this functional, ϱ_{FDOF} is rigidly tied to the positions of FDOF atoms, while ϱ_{KS} is determined by minimizing the energy functional above.

To find the KS ground state charge density ϱ_{KS} , a system of Kohn-Sham-like equations for orbitals has to be solved,

$$\left(-\frac{1}{2}\nabla^2 + v_{\text{eff}}\right)\psi_i = \epsilon_i\psi_i, \quad (15)$$

where

$$\begin{aligned} v_{\text{eff}} = \frac{\delta}{\delta\tilde{\rho}_{\text{KS}}} \{ & T_{\text{nadd}}[\varrho_{\text{FDOF}}, \tilde{\varrho}_{\text{KS}}] + E_{\text{H}}[\tilde{\varrho}_{\text{KS}} + \varrho_{\text{FDOF}}] \\ & + E_{\text{XC}}[\tilde{\varrho}_{\text{KS}} + \varrho_{\text{FDOF}}] + E_{\text{PP}_{\text{nonloc}}}[\varrho_{\text{KS}}, \{\mathbf{R}_j\}_{\text{KS}}] \\ & + E_{\text{PP}_{\text{loc}}}[\varrho_{\text{tot}}, \{\mathbf{R}_i\}] \}_{\tilde{\rho}_{\text{KS}}=\rho_{\text{KS}}}. \end{aligned} \quad (16)$$

The effective potential entering Eq. (15) is thus

$$\begin{aligned} v_{\text{eff}} = & v_{\text{nadd}}[\varrho_{\text{FDOF}}, \varrho_{\text{KS}}] + v_{\text{H}}[\varrho_{\text{tot}}] + v_{\text{XC}}[\varrho_{\text{tot}}] \\ & + v_{\text{PP}}[\{\mathbf{R}_j\}_{\text{KS}}] + v_{\text{PP}_{\text{loc}}}[\{\mathbf{R}_i\}_{\text{FDOF}}] = v_{\text{OF}}[\varrho_{\text{tot}}] \\ & - v_{\text{OF}}[\varrho_{\text{KS}}] + \int \frac{\varrho_{\text{tot}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{\text{XC}}[\varrho_{\text{tot}}] \\ & + v_{\text{PP}}[\{\mathbf{R}_j\}_{\text{KS}}] + v_{\text{PP}_{\text{loc}}}[\{\mathbf{R}_i\}_{\text{FDOF}}], \end{aligned} \quad (17)$$

where

$$v_{\text{OF}}[\varrho] = \left. \frac{\delta T_{\text{OF}}[\tilde{\varrho}]}{\delta\tilde{\varrho}} \right|_{\tilde{\varrho}=\varrho} \quad (18)$$

and v_{PP} without an additional index denotes both local and nonlocal parts of the pseudopotential. This effective potential is similar to that in the KS DFT theory, the differences being the kinetic energy potential terms, the local pseudopotential due to FDOF atoms, and the fact that the Hartree and exchange-correlation potentials are evaluated using the total charge density rather than the KS density.

V. IMPLEMENTATION OF THE HYBRID METHOD

Hybrid calculation schemes take advantage of the fact that chemical processes in biological systems are usually

FIG. 2. (Color online) Schematic layout of simulation cells in a hybrid calculation. The KS atomic cell holds atoms treated by the KS DFT, while the KS charge density and orbitals are restricted to the KS electron cell. FDOF molecules can be anywhere in the total simulation cell. Molecules treated by the KS DFT are shown as solid circles; solvent molecules treated with the FDOF method are shown as open circles.

well localized. It should therefore be sufficient to treat only a small region of the whole system at a high level of theory. The hybrid KS/FDOF method, as defined in the previous section, has no partitioning between regions and KS or FDOF atoms can coexist anywhere in the simulation cell. However, for this method to be useful for biological simulations, the KS atoms have to be restricted to a part of the whole simulation cell.

To accomplish this, a region at the center of the total simulation cell, called the KS atomic cell, is defined. This cell holds all KS atoms. The KS density is restricted to the KS electron cell; this cell is typically larger by 3–4 Å in each direction than the KS atomic cell to allow for the KS charge density to decrease smoothly to zero at the boundary of the electron cell. The total simulation cell is usually periodic. There is no restriction on its size other than it has to be large enough so that biologically relevant molecules located in the KS atomic cell do not interact with their periodic images. We typically use a total cell that is at least 7 Å larger in each direction than the KS electron cell. The layout of the subcells is shown in Fig. 2. Note that OF solvent molecules can appear anywhere in the simulation cell, although usually only the first few shells of the solvent molecules closest to the active site are treated by KS DFT and the rest of the solvent molecules is treated by FDOF DFT. In this case, the KS atomic cell can contain both KS and FDOF atoms. Another possibility, used typically for test simulations on liquid water, is that all solvent molecules completely inside the KS atomic cell are treated by KS DFT and all the rest by FDOF DFT.

The present hybrid method allows for varying the number of solvent molecules treated by the KS method during the course of simulation. Being able to deal with a change in the number of KS atoms is important, since in dynamical calculations, the number of solvent molecules in first solvation shells or within the KS atomic cell can change. It is also possible that a solvent molecule originally treated by the KS method moves out of the KS atomic cell. Such molecule

cannot be treated by the KS DFT anymore, since it is required that all KS atoms are inside the KS atomic cell, and the number of KS atoms has to be decreased. When such an event occurs, the KS/FDOF simulation can readily continue, since the total electron energy in the KS/FDOF system [Eq. (14)] can be calculated for any number of KS and FDOF molecules. However, the exchange causes discontinuities in the total electron energy due to the fact that the electron energy of a solvent molecule is different when calculated with KS or FDOF methods. These discontinuities and the way to restore energy conservation in the simulation will be discussed in the next section.

The hybrid method was implemented to work with both norm-conserving and ultrasoft pseudopotentials (PPs). Using ultrasoft PP in real-space orbital-based DFT methods⁵⁴ require significant changes compared to using the norm-conserving ones. In particular, ultrasoft PP require different grids for the charge density and the KS orbitals, while only one global grid is sufficient when norm-conserving PP are utilized. Accordingly, in the hybrid KS/FDOF method, the KS part uses one or two global grids depending on the type of pseudopotential used. The FDOF part, which deals with charge density only, uses one grid regardless of the type of pseudopotential. To allow for simple transfers between the KS and FDOF grids, it is important that these grids are related. This relationship depends on the PP type: for norm-conserving PP, the KS and FDOF grids are the same, while for ultrasoft PP, the FDOF grid is the same as the KS grid used for representing KS orbitals. This means that when the ultrasoft PP are used, KS and FDOF densities are represented on different grids. The reason for this is efficiency: In ultrasoft KS real-space calculations, the charge density is typically represented on a very fine grid. Using such a grid in the FDOF part would be inefficient and, therefore, the KS orbital grid, which is less dense than the charge density grid, is used. The KS ultrasoft code⁵⁴ contains functions that interpolate between the two grids and, therefore, even when the grids for the charge densities in KS and FDOF parts are different, the quantities represented on one of the grids can be easily transferred to the other.

It should be noted that the present formulation of the hybrid method cannot treat biomolecules that do not fit inside the KS atomic box, since the FDOF method is currently limited to solvent molecules. This limitation is practical rather than principal, since at present, only parameters for water molecules are determined. It is possible and even likely that a frozen-density parametrization will be developed for biomolecular fragments enabling the treatment of large systems. Although at present only fragments of biomolecules with no more than a few hundred atoms can be used in the KS/FDOF DFT simulations, most chemical processes involving creating or breaking of bonds are usually well localized and representing biomolecules by such fragments is often justified. Furthermore, in a typical biosimulation, water molecules constitute a large fraction of the entire simulation system. Therefore, using the FDOF method to treat only the solvent molecules saves substantial computational effort and enables studies of much larger systems.

Finally, it should be mentioned that our embedding

FIG. 3. (Color online) The dipole moment of a KS water molecule surrounded by 63 FDOF water molecules. The average dipole moment is 2.85 D.

scheme is formally similar to that previously proposed by Wesolowski *et al.*^{38,69,70} While they also used frozen density for a part of the system and formulated a calculation scheme combining two DFT methods, the details of the methods are substantially different: Wesolowski *et al.* employed the so-called “freeze and thaw” procedure in which the two subsystems in the calculation were periodically frozen and thawed and until convergence was reached. Our frozen-density orbital-free approach allows for more efficient treatment of the molecules in the fixed region while still treating all the interactions at the DFT level.

VI. APPLICATION TO LIQUID WATER

An important question for the hybrid method is whether an environment of FDOF water molecules can emulate that of real water molecules. To test this, we run a hybrid simulation on a system of 64 water molecules at liquid density at ambient conditions, in which one water molecule was treated by KS DFT and the remaining 63 water molecules were treated by OF DFT. As mentioned above, in the liquid phase, the dipole moment of a water molecule increases from 1.85 D in the gas phase to a value close to 3 D. The value of the dipole moment of the KS molecule during the simulation is plotted in Fig. 3. It varies from 2.23 to 3.54 D and its average value is 2.85 D, which is within the expected range of 2.6–3.15 D mentioned above. This value is also close to the KS DFT value^{56,62} of 2.95 D.

The ability of the KS/FDOF DFT hybrid method to perform molecular dynamics simulation is also tested. A system of 432 water molecules at the liquid density at ambient conditions is used. The simulation cell and the internal subcells were all cubes with sides of 23.8 Å for the whole simulation cell, 11.9 Å for the KS electron box, and 6.0 Å for the KS atomic box. All the molecules completely inside the KS atomic box are treated by the KS DFT. During the course of the simulation, the number of such molecules charges between 4 and 7. The time step is 0.4 fs and the equations of motions are integrated using second-order Verlet velocity scheme. Ultrasoft pseudopotentials⁴⁶ are employed and five Nosé-Hoover thermostats are used to maintain the tempera-

FIG. 4. (Color online) Energies and temperature for a KS/FDOF molecular dynamics simulation of a system of 432 water molecules. (a) Total and electron energies without adjusting for energy due to exchange of solvent molecules. (b) Adjusted total and electron energies as well as Nosé-Hoover and kinetic energies. Nosé-Hoover and kinetic energies were shifted by -6841.91 and -6839.40 Hartree, respectively, in order to appear in the same plot. (c) Temperature variation during the simulation.

ture of the system at 300 K. This test simulation confirms that FDOF DFT is indeed much more efficient than the KS DFT calculation: KS DFT took 95% of the overall computational time, even though only a small fraction of water molecules is treated by it. The data from this test simulation run are shown in Fig. 4.

The top plot shows the electronic and total energies during the course of the simulation. The latter is a sum of the electronic energy, the KE of atoms, and the energy of Nosé-Hoover thermostats. Both these energies exhibit rather large steps, which are due to exchanges of molecules between the KS and FDOF subsystems. The exchange causes discontinuities in energy, since the energy of a solvent molecule is different when calculated with the KS or FDOF methods. The steps up occur when FDOF solvent molecules become KS solvent molecules, while the steps down are due to the opposite process. Because the steps are about 1.6 Hartree high, the fine details of energy curves cannot be properly displayed on such a scale. In order to remove the steps and recover smooth energy curves, we compare the total energy, which should be conserved during a molecular dynamics run, just before and after each exchange and subtract the difference from both the total and electron energies. As shown in the middle plot in Fig. 4, this procedure indeed leads to smooth energy curves. Note that no adjustment is needed for Nosé-Hoover and ion kinetic energies as well as temperature shown in the bottom plot.

It should be noted that switching of solvent molecules affects the kinetic energy of the system. When a water molecule changes from KS to FDOF description, its vibrational motion is frozen and its vibrational energy is immediately reduced to zero. This causes a change in the total kinetic energy and the lost vibrational energy is added back into the system through the Nosé-Hoover thermostats. However, since the lowest vibrational frequency of the water molecule is 1595 cm^{-1} , much higher than the average thermal energy per degree of freedom at room temperature, the lost vibra-

FIG. 5. (Color) The relaxed structure of Cu^{2+} bound to a fragment of PrP protein. The copper ion is shown in gold, carbon atoms are colored green, nitrogen atoms blue, oxygen atoms red, and hydrogens gray. For clarity, water molecules are not shown. The copper ion bond distances are (Cu–N1) 1.97 Å, (Cu–N2) 2.01 Å, (Cu–N3) 1.92 Å, and (Cu–O) 2.35 Å.

tional energy is tiny and much smaller than normal fluctuations in the kinetic energy or the Nosé-Hoover energy. Furthermore, since the FDOF parametrization was developed to reproduce results for real liquid water, it effectively includes vibrational effects. Therefore, the effects of lost vibrational energy can be expected to be negligible and we find this in our calculations, see Fig. 4, where kinetic and other energies appear as smooth and continuous lines.

The total energy, which appears as a straight line in Fig. 4(b), has oscillations of about 0.75 meV around its average value. This degree of energy conservation is similar to that in real-space KS DFT (Ref. 54) with ultrasoft pseudopotentials.

It is also important to evaluate how the switching of solvent molecules affects hybrid simulations. Usually, this is done by calculating forces on a molecule that is about to be switched by both methods and comparing them. Unfortunately, direct comparison of FDOF and KS forces in the present hybrid method is not meaningful. Due to the use of fixed geometry in the FDOF method, the FDOF forces include intramolecular forces, which are removed when the RATTLE procedure is applied to enforce the fixed geometry. The effective forces after RATTLE can be compared, but even this comparison is misleading, again due to the use of the fixed geometry. This is because the enforced geometry may not be fully appropriate for instantaneous state and position of a molecule under consideration, which would add a significant contribution to the calculated forces and distort the comparison.

Instead of comparing forces at a few discrete steps, it is better to compare effects of such transition on the entire simulation. To this end, we compare a regular hybrid calculation on the system of 432 water molecules mentioned above with the calculation in which the switching was disabled. During the course of the regular simulation, one switching event occurred and the calculations continued for another 200 molecular dynamics steps after the event. We have compared the displacements of the switched molecule in the two calculations and found that the difference is 0.05 bohr or 5.6% of the total displacement of the molecule. The small size of this difference indicates that switching of solvent molecules between the two subsystems does not significantly affect the results of the hybrid simulation.

VII. APPLICATION TO BINDING OF Cu(II) TO PRION PROTEIN

As first application to a biologically relevant system, the binding of copper ion to prion protein,⁷¹ PrP, is studied. The PrP is responsible for infectious neurodegenerative diseases such as the mad cow disease (BSE) and the Creutzfeldt-Jakob disease. The physiological function of the PrP is unknown, but recent investigations^{72–77} suggest that it may be linked to its ability to bind copper.

The fundamental copper binding site consists of amino acids HGGGW.⁷⁵ The binding of copper to HGGGW was previously investigated experimentally.⁷⁵ Theoretical studies^{78,79} also exist, although these did not consider all five amino acids. These studies indicate that the copper ion is five-coordinated in the PrP. It binds to a nitrogen in the imidazole ring, two deprotonated nitrogens, and an oxygen in the protein backbone.

In our simulation, the PrP is modeled as a fragment consisting of the five amino acids mentioned above. As usual in atomistic simulations of proteins, the ends of the fragment backbone are passivated with stable CH_3 groups. The total number of atoms in the fragment is 72. The protein fragment is first molded into a shape resembling the binding side geometry obtained from experiment.⁷⁵ The system is then solvated with 3101 water molecules. The KS electron and the full simulation cells are cubic with sides of 22.49 and 44.98 Å, respectively. The water molecules closer to the copper ion than 6 Å are treated by KS DFT, which typically adds 12 water molecules to the KS subsystem, increasing the total number of KS atoms to 108. This system is equilibrated by first running molecular dynamics at 300 K for FDOF molecules only for 5000 steps, with the KS part frozen, and then 750 steps of molecular dynamics is performed for the entire system. Following the simulation, the structure of the hybrid system was relaxed. The resulting structure is shown in Fig. 5.

The calculated binding site geometry is quite similar to that obtained in Ref. 78. The Cu–N bonds differ only by a few hundreds of angstroms, while the Cu–O bond is 2.35 Å instead of 2.09 Å. A significant part of this difference appears to be due to the different treatment of the solvent in Ref. 78, where only the first solvation shell was explicit and

the rest of the solvent was modeled implicitly. In our calculation, two water molecules are found in contact with the copper ion, which are stabilized by hydrogen bonding with other water molecules present in the system. No water molecules close to the copper are reported in Ref. 78. The presence of these two water molecules is the major cause of the difference between the two calculations: When the calculation is performed without the water molecules in contact with copper, the difference in Cu–O bond distance decreases by more than half to only 0.11 Å, while other bonds involving copper stay almost unchanged.

The binding energy of the copper ion to the PrP was also calculated. To obtain it, the energy of a free protein fragment with protonated backbone nitrogens (E_{fragH_2}) was calculated when solvated with the same number of water molecules as the copper-bound fragment. We also computed the energy of a copper atom (E_{Cu}) and the energy of a hydrogen molecule H_2 (E_{H_2}). The copper ion binding energy can then be obtained as

$$E_{\text{bind,Cu}} = E_{\text{fragCu}} - E_{\text{fragH}_2} - E_{\text{Cu}} + E_{\text{H}_2}, \quad (19)$$

where E_{fragCu} is the energy of the relaxed solvated HGGGW–Cu complex shown in Fig. 5. By calculating all four energies, we obtain 1.76 eV as the binding energy of the copper ion. We are not aware of any calculated or measured binding energy of copper ions in biological systems, but our calculated value is similar to the binding energy of a copper atom in copper clusters Cu_4 and Cu_{13} (1.26 and 2.19 eV respectively).⁸⁰

Additional calculations on copper binding to the PrP are in progress, which address different copper binding modes and the resulting structural changes in the protein. The results of these calculations will be reported elsewhere.

VIII. SUMMARY AND CONCLUSIONS

In simulations involving proteins or other biomolecules, the solvent plays an important role and its effects cannot be neglected. In a typical protein simulation, the number of atoms in the solvent is by far greater than the number of atoms in the solvated protein. The hybrid method presented in this paper is designed for such biosimulations. The biomolecule, or its fragment, and solvent molecules in the first solvation shells are treated by full Kohn-Sham DFT, while the remaining solvent molecules are efficiently described using a frozen-density orbital-free DFT. We have shown that this approach allows for an efficient and accurate description of the solvent effects and that it provides seamless integration between the regions treated by different DFT methods.

At present, the hybrid method requires that the biomolecules are fully treated by the KS DFT, which means that it can only treat biomolecular fragments with no more than a few hundred atoms. Hybrid simulations for large biomolecules with the current method are possible but require developing of charge density parametrization for important parts of biomolecules, such as amino acids in the case of proteins.

In the near future, the need for treating biomolecules by two different methods may decrease due to the steady ad-

vances in computer power and computational methods. Current state-of-the-art supercomputers contain tens of thousands of processors and this number is expected to increase rapidly over the next few years. This will allow quantum calculations on systems with thousands of atoms especially when linear-scaling methods are used. When such calculations become feasible, the present hybrid method will allow for simulation of fully solvated, full-length biomolecules at the quantum level, which at present can only be performed using classical potentials.

ACKNOWLEDGMENTS

We gratefully acknowledge support by NSF, DOE, DOD, and ORNL-LDRD and grants of supercomputer time provided by DOE and the DOD Challenge Program.

- ¹ P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
- ² W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
- ³ J. Tomasi and M. Persico, Chem. Rev. (Washington, D.C.) **94**, 2027 (1994).
- ⁴ J. L. Rivail and D. Rinald, *Liquid State Quantum Chemistry: Computational Applications of The Polarizable Continuum Models* (World Scientific, Singapore, 1995).
- ⁵ C. J. Cramer and D. G. Truhlar, *Continuum Solvation Models* (Kluwer, Dordrecht, 1996).
- ⁶ J. Fattbert and F. Gygi, J. Comput. Chem. **23**, 662 (2002).
- ⁷ D. Begue, P. Carbonniere, V. Barone, and C. Pouchan, Chem. Phys. Lett. **416**, 206 (2005).
- ⁸ L. Guimaraes, H. Avelino de Abreu, and H. Duarte, Chem. Phys. **333**, 10 (2007).
- ⁹ P. Bonifassi, P. Ray, and J. Leszczynski, Chem. Phys. Lett. **431**, 321 (2006).
- ¹⁰ R. Zhou and B. Berne, Proc. Natl. Acad. Sci. U.S.A. **99**, 12777 (2002).
- ¹¹ Y. M. Rhee, E. J. Sorin, G. Jayachandran, E. Lindahl, and V. S. Pande, Proc. Natl. Acad. Sci. U.S.A. **101**, 6456 (2004).
- ¹² S. Miertus, E. Scrocco, and J. Tomasi, Chem. Phys. **55**, 117 (1981).
- ¹³ R. Cammi and J. Tomasi, J. Comput. Chem. **16**, 1449 (1995).
- ¹⁴ A. Warshel and M. Levitt, J. Mol. Biol. **102**, 227 (1976).
- ¹⁵ U. C. Singh and P. A. Kollmann, J. Comput. Chem. **7**, 718 (1986).
- ¹⁶ M. J. Field, P. A. Bash, and M. Karplus, J. Comput. Chem. **11**, 700 (1990).
- ¹⁷ T. Kerdcharoen and K. Morokuma, Chem. Phys. Lett. **355**, 257 (2002).
- ¹⁸ L. Thomas, Proc. Cambridge Philos. Soc. **23**, 542 (1927).
- ¹⁹ E. Fermi, Rend. Accad. Naz. Lincei **6**, 602 (1927).
- ²⁰ E. Fermi, Z. Phys. **48**, 73 (1928).
- ²¹ C. von Weizsacker, Z. Phys. **96**, 431 (1934).
- ²² R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).
- ²³ Y. Wang and E. Carter, *Orbital Free Kinetic Energy Density Functional Theory* (Kluwer, Dordrecht, 2000), Chap. 5, pp. 117–184.
- ²⁴ A. E. DePristo and J. D. Kress, Phys. Rev. A **35**, 438 (1987).
- ²⁵ A. J. Thakkar, Phys. Rev. A **46**, 6920 (1992).
- ²⁶ P. I. Plindov and S. K. Pogrebnya, Chem. Phys. Lett. **143**, 535 (1988).
- ²⁷ J. P. Perdew, Phys. Lett. A **165**, 79 (1992).
- ²⁸ H. Lee, C. Lee, and R. Parr, Phys. Rev. A **44**, 768 (1991).
- ²⁹ D. Lacks and R. Gordon, J. Chem. Phys. **100**, 4446 (1994).
- ³⁰ A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
- ³¹ J. A. Alonso and L. A. Girifalco, Phys. Rev. B **17**, 3735 (1978).
- ³² E. Chacón, J. E. Alvarellos, and P. Tarazona, Phys. Rev. B **32**, 7868 (1985).
- ³³ E. Smargiassi and P. A. Madden, Phys. Rev. B **49**, 5220 (1994).
- ³⁴ Y. A. Wang, N. Govind, and E. A. Carter, Phys. Rev. B **58**, 13465 (1998).
- ³⁵ Y. A. Wang, N. Govind, and E. A. Carter, Phys. Rev. B **60**, 16350 (1999).
- ³⁶ R. Gordon and Y. Kim, J. Chem. Phys. **61**, 3122 (1972).
- ³⁷ Y. Kim and R. Gordon, Phys. Rev. B **9**, 3548 (1974).
- ³⁸ T. Wesolowski and A. Warshel, J. Phys.: Condens. Matter **97**, 8050 (1993).
- ³⁹ T. Wesolowski and A. Warshel, J. Phys.: Condens. Matter **98**, 5183 (1994).

- ⁴⁰D. Barker and M. Sprik, *Mol. Phys.* **101**, 1183 (2003).
- ⁴¹P. Huang and E. A. Carter, *J. Chem. Phys.* **125**, 084102 (2006).
- ⁴²D. R. Hamann, M. Schlüter, and C. Chiang, *Phys. Rev. Lett.* **43**, 1494 (1979).
- ⁴³G. P. Kerker, *J. Phys. C* **13**, L189 (1980).
- ⁴⁴A. Zunger and M. L. Cohen, *Phys. Rev. B* **18**, 5449 (1978).
- ⁴⁵L. Kleinman and D. M. Bylander, *Phys. Rev. Lett.* **48**, 1425 (1982).
- ⁴⁶D. Vanderbilt, *Phys. Rev. B* **41**, 7892 (1990).
- ⁴⁷T. Starkloff and J. D. Joannopoulos, *Phys. Rev. B* **16**, 5212 (1977).
- ⁴⁸S. C. Watson, B. J. Jesson, E. A. Carter, and P. A. Madden, *Europhys. Lett.* **41**, 37 (1998).
- ⁴⁹V. Shah, D. Nehete, and D. G. Kanhere, *Z. Phys.* **48**, 73 (1928).
- ⁵⁰D. Nehete, V. Shah, and D. G. Kanhere, *Phys. Rev. B* **53**, 2126 (1996).
- ⁵¹A. Dhavale, V. Shah, and D. G. Kanhere, *Phys. Rev. A* **57**, 4522 (1998).
- ⁵²E. L. Briggs, D. J. Sullivan, and J. Bernholc, *Phys. Rev. B* **52**, 5471 (1995).
- ⁵³E. L. Briggs, D. J. Sullivan, and J. Bernholc, *Phys. Rev. B* **54**, 14362 (1996).
- ⁵⁴M. Hodak, S. Wang, W. Lu, and J. Bernholc, *Phys. Rev. B* **76**, 085108 (2007).
- ⁵⁵J. Ryckaert, G. Ciccotti, and H. Berendsen, *J. Comp. Phys.* **23**, 327 (1977).
- ⁵⁶P. L. Silvestrelli and M. Parrinello, *J. Chem. Phys.* **111**, 3572 (1999).
- ⁵⁷A. Shepard, G. Beers, G. Klein, and L. Rothmann, *J. Chem. Phys.* **59**, 2254 (1973).
- ⁵⁸C. Coulson and D. Eisenbero, *Proc. R. Soc. London, Ser. A* **291**, 445 (1966).
- ⁵⁹E. Batista, S. Xantheas, and H. Jonsson, *J. Chem. Phys.* **109**, 4546 (1998).
- ⁶⁰M. Heggie, C. L. Latham, and S. Maynard, and R. Jones, *Chem. Phys. Lett.* **249**, 485 (1996).
- ⁶¹Y. Badyal, M. Saboungi, D. Price, S. Shastri, D. Haefner, and A. Soper, *J. Chem. Phys.* **112**, 9206 (2000).
- ⁶²P. L. Silvestrelli and M. Parrinello, *Phys. Rev. Lett.* **82**, 3308 (1999).
- ⁶³A. K. Soper, *Chem. Phys.* **258**, 121 (2000).
- ⁶⁴J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- ⁶⁵J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **78**, 1396 (1997).
- ⁶⁶J. Grossman, E. Schwegler, E. Draeger, F. Gygi, and G. Galli, *J. Chem. Phys.* **120**, 300 (2004).
- ⁶⁷E. Schwegler, J. Grossman, F. Gygi, and G. Galli, *J. Chem. Phys.* **121**, 5400 (2004).
- ⁶⁸M. Sprik, J. Hutter, and M. Parrinello, *J. Chem. Phys.* **105**, 1142 (1996).
- ⁶⁹T. Wesolowski, R. Muller, and A. Warshel, *J. Phys.: Condens. Matter* **100**, 15444 (1996).
- ⁷⁰T. Wesolowski and J. Weber, *Chem. Phys. Lett.* **248**, 71 (1996).
- ⁷¹S. Prusiner, *Science* **278**, 245 (1997).
- ⁷²N. Vassallo and J. Herms, *J. Neurochem.* **86**, 538 (2003).
- ⁷³N. Inestrosa, W. Cerpa, and L. Varela-Nallar, *IUBMB Life* **57**, 645 (2005).
- ⁷⁴D. Brown, K. Qin, J. Herms, A. Madlung, J. Manson, R. Strome, P. Fraser, T. Kruck, A. von Bohlen, W. Schulz-Schaeffer, A. Giese, D. Westaway, and H. Kretschmar, *Nature (London)* **390**, 684 (1997).
- ⁷⁵G. Millhauser, *Acc. Chem. Res.* **37**, 79 (2004).
- ⁷⁶R. Whittall, H. Ball, F. Cohen, A. Prusiner, and M. Baldwin, *Protein Sci.* **9**, 332 (2000).
- ⁷⁷P. Pauly and D. Harris, *J. Biol. Chem.* **273**, 33107 (1998).
- ⁷⁸M. Pushie and A. Rauk, *JBIC, J. Biol. Inorg. Chem.* **8**, 53 (2003).
- ⁷⁹H.-F. Ji and H.-Y. Zhang, *Chem. Res. Toxicol.* **17**, 471 (2004).
- ⁸⁰B. Delley, D. E. Ellis, A. J. Freeman, E. J. Baerends, and D. Post, *Phys. Rev. B* **27**, 2132 (1983).