A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields

M. Eichinger and P. Tavan[a)]
Institut für Medizinische Optik, Theoretische Biophysik, Ludwig Maximilians Universität München, Oettingenstr. 67, D-80538 München, Germany

J. Hutter and M. Parrinello
Max-Planck-Institut für Festkörperforschung, Heisenbergstr. 1, D-70569 Stuttgart, Germany

(Received 5 November 1998; accepted 11 March 1999)

We present a hybrid method for molecular dynamics simulations of solutes in complex solvents as represented, for example, by substrates within enzymes. The method combines a quantum mechanical (QM) description of the solute with a molecular mechanics (MM) approach for the solvent. The QM fragment of a simulation system is treated by *ab initio* density functional theory (DFT) based on plane-wave expansions. Long-range Coulomb interactions within the MM fragment and between the QM and the MM fragment are treated by a computationally efficient fast multipole method. For the description of covalent bonds between the two fragments, we introduce the scaled position link atom method (SPLAM), which removes the shortcomings of related procedures. The various aspects of the hybrid method are scrutinized through test calculations on liquid water, the water dimer, ethane and a small molecule related to the retinal Schiff base. In particular, the extent to which vibrational spectra obtained by DFT for the solute can be spoiled by the lower quality force field of the solvent is checked, including cases in which the two fragments are covalently joined. The results demonstrate that our QM/MM hybrid method is especially well suited for the vibrational analysis of molecules in condensed phase. © *1999 American Institute of Physics.* [S0021-9606(99)71521-7]

I. INTRODUCTION

Enzymatic catalysis is a most prominent example of how a solvent environment exhibiting a complex structure and dynamics can steer the properties and chemical reactions of a solute molecule. In this context, the solute molecule is called a substrate and the immediate solvent environment is made up of the binding pocket of the enzyme. Within the pocket the substrate experiences highly specific local electrostatic fields and steric restraints, altering its properties in such a way that an otherwise unlikely reaction proceeds at a high rate.

As an example consider the primary photochemical reaction of the chromophore of bacteriorhodopsin (BR), which acts as a light-driven proton pump in the purple membrane of *Halobacterium salinarium*.[1] Figure 1 shows a model of BR embedded in a membrane-water system. Here, the chromophore represents the solute, whereas all other atoms in the system belong to the solvent.

The BR chromophore, whose chemical structure is depicted in Fig. 2, is a protonated Schiff base of all-*trans* retinal (PRSB) which is covalently bound through a lysine residue to the protein backbone. In solution a PRSB molecule isomerizes upon photoexcitation around each of the C=C double bonds of its polyene chain. In BR, however, only the 13-*cis* isomer is obtained at a high quantum yield (for details, see Ref. 2). Thus, the protein environment specifically catalyzes that photoisomerization of the retinal chromophore. In addition, the protein environment changes other properties of the chromophore, most notably its optical and vibrational spectra (for discussions, see Refs. 3 and 4). Note that the work to be presented in this article has been strongly motivated by the specific aim of describing the vibrational spectra of the BR chromophore *in situ*; therefore, we have used this system for illustration.

The general conclusion is that proteins represent most peculiar solvents for embedded solute molecules, altering their properties in highly specific ways. Theoretical descriptions have to account for the fact that here the solvent structure and dynamics are too complex to be covered by continuum models. Instead, quantum mechanical (QM) treatments of the solute have to be combined with explicit representations of the solvent. Due to computational limitations such representations cannot be based on quantum mechanics but must resort to empirical procedures of molecular mechanics (MM). Hence, so-called QM/MM hybrid models are required.

In this article, we wish to introduce a computationally efficient QM/MM hybrid method which aims at accurate descriptions of *ground state* properties, dynamics, and chemical reactions of solute molecules, embedded in large solvent environments; its extension towards excited solute states is left to the future.

As is apparent also from consideration of the example given in Fig. 1, QM/MM hybrid models first have to partition a large system into a small solute part, the QM fragment,

a)Author to whom correspondence should be sent.

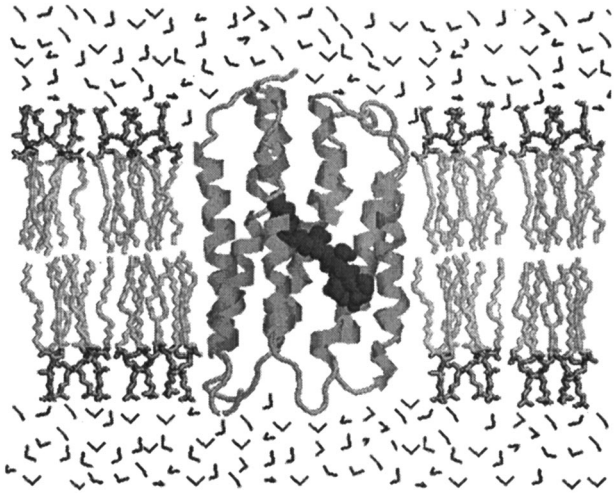

FIG. 1. Bacteriorhodopsin with its retinal chromophore (drawn in black) in a membrane-water model system (drawn in gray) as an example for a complex solvent-solute system.

and a larger solvent part, the MM fragment. In the figure, a possible partitioning is indicated by the choice of color. Furthermore, hybrid models have to specify (i) a particular QM method for the solute, (ii) an empirical force field for the solvent, and (iii) an interaction scheme for the two fragments.

In cases in which the QM and MM fragments are linked by covalent bonds, as illustrated in Fig. 3, the interaction scheme (iii) has to be complemented by (iv) an adequate treatment of these bonds. The various hybrid models that have been proposed previously and will be reviewed below differ among each other and from our proposal in these respects. Particularly concerning point (iv) we will suggest a new procedure which, as we intend to show, overcomes the problems that have hampered previous attempts.

In order to motivate our choices concerning the ingredients of the hybrid method [cf. points (i) and (ii) above], we will first shortly review the corresponding states of the art.

A. Quantum chemistry of isolated molecules

Methods for the quantum chemical description of small molecules and clusters have reached a stage at which accurate numerical computations of ground state molecular properties and reaction pathways are feasible.

As is well known, in quantum chemical calculations the computational effort depends on the molecular property to be described; it grows strongly with the size of the system and with the desired accuracy. For accurate descriptions advanced methods of *ab initio* quantum chemistry have to be employed, which are available through program packages such as Gaussian94.[5] Computational limitations currently restrict such calculations to systems with at most 100 atoms. If

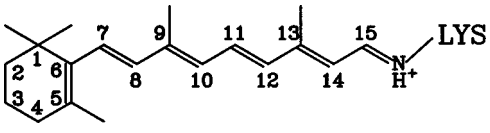

FIG. 2. Chemical structure of the BR chromophore.

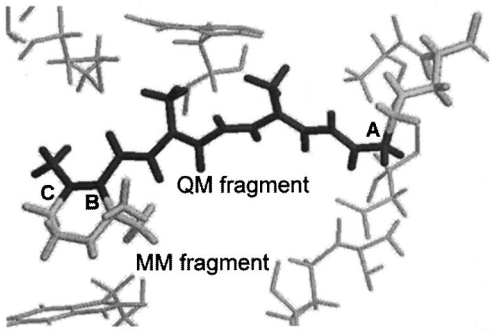

FIG. 3. Close look into the binding pocket of the BR chromophore. A possible partitioning of the system into an MM fragment drawn in gray and a QM fragment drawn in black is indicated; three covalent bonds at positions A–C join the MM and QM fragments.

instead semiquantitative results suffice for gaining qualitative insights one may also resort to the computationally less expensive methods of *semiempirical* quantum chemistry which are accessible, for example, through the MOPAC program package.[6]

Until recently, accurate *ab initio* descriptions of ground state properties typically had to first apply the Hartree–Fock (HF) approximation and subsequently take care of the most important effects of electron correlation.[7] But the huge computational cost restricted such post-HF methods to the study of molecules containing less than about 25 atoms. Larger molecules, like the retinal chromophore introduced above, could be treated only at the cost of the neglect of electron correlation. Here, particularly the prediction of *vibrational spectra*, which requires highly accurate intramolecular force fields and which, therefore, represents a most sensitive probe on the quality of the respective computational method, was severely hampered by the shortcomings of the HF approach.[4,8–10]

Fortunately, during the past decade a cost-effective alternative approach based on density-functional theory (DFT)[11,12] has been developed. Much effort has been devoted to refining the methodology. For the problem of *vibrational analysis*, for example, it has meanwhile become apparent that DFT with nonlocal exchange and correlation functionals like those suggested by Becke[13] and Lee–Yang–Parr[14] can very accurately predict vibrational spectra for sizeable molecules without any scaling (see, for example, Refs. 10,15–17). Thus, DFT calculations properly account for the most important effects of electron correlation at a computational cost of conventional *ab initio* HF treatments. In addition, DFT invites applications of algorithmic tricks like the use of plane-wave expansions together with fast Fourier transforms, allowing its computational efficiency to be enhanced further. Upon combination with parallelization strategies and the use of parallel computers DFT treatments of systems with up to 100 atoms have become possible.

As shown by Car and Parrinello[18] the DFT-based extension of high-quality quantum chemistry towards larger systems has also enabled descriptions of dynamical and, in particular, reactive processes within clusters of atoms or molecules through simulations of the molecular dynamics

(MD). The power and validity of this DFT method meanwhile has been demonstrated in a series of relevant applications (see, for example, Refs. 19–21). Thus, our choice of ingredient (i) of our method is the corresponding DFT program called CPMD.[22]

Previously, MD simulations had been restricted to the use of force fields obtained either from HF calculations or from empirical MM models. Although reactive processes are outside the scope of the latter and although MM force fields are much too crude to allow reasonable descriptions of vibrational spectra of molecules, for example, they are valuable if one wishes to calculate statistical properties and dynamical processes of *very large* systems comprising several 10^4 atoms. In particular, the theories of liquids and of biopolymers have strongly benefitted from the MM-MD approach. As a result of its ongoing development, reasonable microscopic descriptions of systems which we have termed *complex solvents* are feasible nowadays.

B. Molecular mechanics of complex solvents

Already the first MD simulations of simple liquids such as argon had shown that most important statistical properties like, in the case of argon, the radial distribution function can be obtained accurately.[23,24] Subsequently, the MD approach has been extended towards increasingly complicated molecular systems ranging from aqueous solutions of simple ions to those of lipid membranes or proteins (see Refs. 25–29). Intramolecular force fields for wide classes of macromolecules have been successfully developed, which decompose these molecules into local chemical bonding motives whose mechanical properties are then described in terms of simple semiempirical formulas.[30–33] But for a long time, the treatment of the long-range electrostatic forces between polar and charged molecules was less satisfactory.

In the MM-MD approach, electrostatic forces are considered to originate from partial or integral charges c_i centered within the atoms of the molecules. Common force fields[30,32,33] assign fixed values to these charges. Thus, polarization effects are partially neglected here, since fixed partial charges can only account for the average polarization of a molecule in a standard environment. But they cannot cover changes of the polarization within a dynamic and heterogeneous environment as provided by a protein. In view of the large polarizability of water,[34] for example, this approximation is definitely questionable.

Only a few MM-MD procedures have been proposed that allow polarization effects to be included. Because of the considerable computational effort associated with the required self-consistent evaluation of induced dipole moments (see Refs. 35,36), these methods are still restricted to systems of moderate size comprising at most 2000 atoms.

Until recently, computational limitations and algorithmic insufficiencies had enforced a further and even more questionable approximation. One tried to avoid the huge computational effort associated with the exact evaluation of the long-range electrostatic interactions between N charges c_i and c_j at distances r_{ij} from the Coulomb sum

$$E_{\mathrm{elstat}} = \sum_{i<j}^{N} c_i c_j / r_{ij}, \tag{1}$$

by simply neglecting these forces beyond a given cut-off distance R_c.[30,32] Such cut-off algorithms scale linearly with N in contrast to the N^2 behavior of Eq. (1). Values smaller than about 15 Å had to be chosen for R_c and these short-range cutoffs turned out to introduce serious artifacts into simulations of proteins and lipid membranes (see Refs. 25,29,37–40).

Considerations of accuracy would suggest (cf. Ref. 41) that the neglect of long-range contributions to the electrostatics in MD simulations of proteins seems safe only at cut-off distances larger by at least a factor of 5 than the quoted choice. Fortunately, the development of fast multipole methods,[41–44] which scale linearly with N, of multiple time step algorithms,[45–47] which can further enhance computational efficiency, and of combinations of these procedures,[48,49] notably of the FAMUSAMM algorithm,[50] has enabled computationally cost-effective and sufficiently accurate treatments of the long-range electrostatics. Similar considerations apply to recent improvements of the Ewald sum approach for simulation systems enclosed by periodic boundaries; here the computational complexity scales with $N\log N$.[51–55]

As a result, MM-MD simulations on parallel computers that account properly for the long-range electrostatics in protein-solvent systems comprising several 10^4 atoms and cover time spans in the nanosecond range have become feasible. Inspired by corresponding sample simulations[56] of an atomic force microscopy experiment concerning ligand-receptor binding,[57] it has recently been concluded that the MM-MD approach ''has come of age.''[28]

Despite its current deficiencies concerning the partial neglect of polarization effects we therefore chose the CHARMM/FAMUSAMM force field,[30,50] which is implemented in the parallelized MM-MD program EGO_VIII,[58] as the second ingredient (ii) for our QM/MM hybrid model.

C. QM/MM hybrid models

QM/MM hybrid methods have a long history and we have therefore partially adopted concepts developed in earlier work. A short review will allow comparison with our work.

The history of QM/MM hybrid models started two decades ago with the seminal paper by Warshel and Levitt.[59] The strict limitations of computer power available at that time forced the authors to resort to a simple semiempirical model for the QM fragment and its covalent linkages to the MM part. But for the MM fragment and its interactions with the QM fragment, the authors succeeded in creating an ingeniously consistent model. For instance, they included the atomic polarizabilities into the MM force field, whereas polarization effects have generally been neglected in subsequent models and the problem has been addressed again solely in Refs. 60 and 61. The van der Waals interactions between atoms of the MM and QM fragments were taken into account by a 6-12-Lennard-Jones potential and the elec-

trostatic interactions by partial charges and by evaluation of the Coulomb sum. The latter procedures have been common since then.

In subsequent models, either refined semiempirical[60–68] or computationally, more demanding *ab initio*[69–72] Hartree–Fock methods have been chosen for the description of the QM part. Recently, also DFT methods have been applied.[73–78] But here, in contrast to some of the former models, no procedures for the MD treatment of covalent linkages between the MM and QM fragments have been provided yet [cf. point (iv) above].

If a covalent bond joins the MM and QM fragments, as is the case in the sample system depicted in Fig. 3, a conceptual problem arises. For its solution, the bond can be replaced by a new bond to an artificially introduced hydrogen atom. This hydrogen atom is called a *link atom* and serves to ensure a saturated valence structure of the QM fragment. Evidently that is guaranteed only for simple σ bonds such as those bonds between CH_2 groups frequently occurring in proteins. Next, the mechanical properties of the original bond have to be restored and, finally, the link atom has to be removed again from the description.

A simple version of this link atom concept has been suggested by Singh and Kollman.[70] Here, the link atom remained decoupled from the MM fragment, but was kept as a dynamical object in the QM fragment, and the mechanics of the original bond was covered by an MM description. Later, Field *et al.*[63] developed a refined procedure by constraining the link atom to its equilibrium position (cf. also Bakowies and Thiel[66]). However, Bakowies[79] observed that serious artifacts show up whenever that equilibrium position deviates strongly from the line connecting the atoms of the original bond. Furthermore, he noticed that *vibrational spectra* derived by that hybrid procedure appear to be even less satisfactory than corresponding spectra obtained from the simple MM force field. As a result the link atom method acquired a rather bad image in the community.

The scaled position link atom method (SPLAM) which we will suggest in this article has been designed to solve these problems. To check its quality we will present sample calculations of vibrational spectra for selected hybrid models. Similarly, all other aspects of our hybrid method will be scrutinized by test calculations.

The article is organized as follows. In Sec. II we will first sketch the basic tasks that have to be tackled in any attempt to design a QM/MM hybrid method. We then describe the various procedures which combine the DFT description of the QM fragment with the CHARMM/FAMUSAMM force field for the MM fragment into a unified MD approach. Here we will emphasize the efficient computational treatment of the long-range solvent electrostatics, the SPLAM method, and accuracy issues. The quality of the hybrid method will be checked in Sec. III using a series of simple but relevant examples and comparisons both with experimental data and with results of previous computational studies. A short summary concludes the paper.

II. METHODS

We start the description of the physical concepts and computational methods entering our hybrid QM/MM approach with an outline of the basic tasks, thereby introducing some of the notation.

A. Basic tasks

Using the Born–Oppenheimer approximation, MD methods derive the forces acting on the nuclei at the positions $\mathbf{r}_i$, $i=1,\ldots,N$ from an effective potential energy function $E(\mathbf{R})$. For a fixed nuclear configuration $\mathbf{R}=(\mathbf{r}_1,\ldots,\mathbf{r}_N)$ that effective potential reflects the energy of the electron system in the electrostatic potential originating from the nuclei and electrons. Furthermore, the nuclei are treated as classical particles and the trajectory $\mathbf{R}(t)$ of the system within configuration space is obtained from a numerical solution of Newton's equations

$$m_i \frac{d^2}{dt^2}\mathbf{r}_i = -\nabla_i E(\mathbf{R}), \quad i=1,\ldots,N, \tag{2}$$

for a given set of initial positions and velocities of the atoms with the masses m_i. In pure MM models, the effective potential $E(\mathbf{R})$ is expressed by carefully parameterized semiempirical formulas. In contrast, the computationally much more demanding QM descriptions rely on numerical solutions of the stationary electronic Schroedinger equation in order to calculate the forces $-\nabla_i E(\mathbf{R})$.

Any QM/MM hybrid approach first has to partition the system into a large MM and a smaller QM fragment as exemplified in Figs. 1 and 3. Formally, such a partitioning may be expressed by collecting the coordinates of the atoms in the MM fragment into a configuration vector $\mathbf{R}_M=(\mathbf{r}_1,\ldots,\mathbf{r}_M)$ and those of the QM fragment into $\mathbf{R}_{M,Q}=(\mathbf{r}_{M+1},\ldots,\mathbf{r}_N)$. To simplify distinction, we use the index sets $\mathbf{M}=\{1,\ldots,M\}$ and $\mathbf{Q}=\{M+1,\ldots,N\}$ for the MM and QM atoms, respectively. Futhermore, the MM atoms are labeled by indices $l,m\in\mathbf{M}$ and the QM atoms by indices $q,s\in\mathbf{Q}$.

In addition to the partitioning, any hybrid method has to specify an energy function for the MM fragment, an electronic model Hamiltonian for the QM fragment, and prescriptions for the calculation of the interactions between the two fragments.

B. The hybrid model

An MM energy function E_M like the CHARMM/FAMUSAMM force field employed by us[30,50] is composed of short-range terms E_{bonded} accounting for chemical binding forces, and of more long-range contributions $E_{\text{nonbonded}}$ covering van der Waals and electrostatic interactions between nonbonded atoms within an MM fragment

$$E_M(\mathbf{R}_M)=E_{\text{bonded}}(\mathbf{R}_M)+E_{\text{nonbonded}}(\mathbf{R}_M). \tag{3}$$

Here, also atoms within a single molecule are classified as nonbonded whenever they are separated by more than three covalent bonds. This classification can be formally described by assigning a switching function $\epsilon(l,m)$ to each atom pair (l,m), which assumes the value zero if the atoms are sepa-

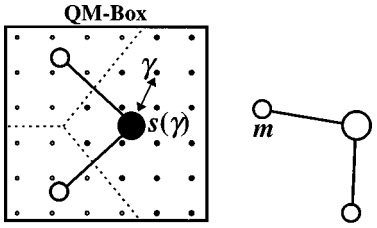

FIG. 4. Water molecule in a QM box discretized by a regular grid; also drawn is a water molecule of the MM fragment; the nearest neighbor classification $s(\gamma)$ of grid points γ relevant for the evaluation of the external potential is indicated by dashed boundaries.

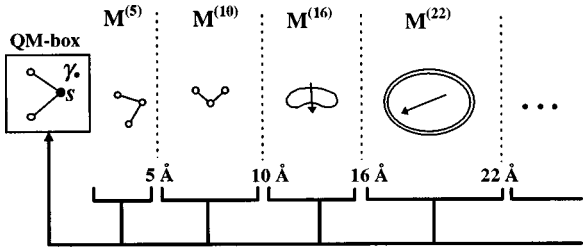

FIG. 5. FAMUSAMM distance classes; all MM atoms closer than 5 Å to a selected QM atom s (filled circle) are classified into the near zone $\mathbf{M}^{(5)}$; atoms in more distant zones are collected into classes $\mathbf{M}^{(10)}$, $\mathbf{M}^{(16)}$, ... and, starting at $\mathbf{M}^{(16)}$, are grouped into compact charge distributions; the arrows serve to indicate that the potential at a grid point γ (black dot) near the QM atom s is composed of contributions due to single charges in the two inner zones and to multipole moments of charge distributions in the outer zones.

rated by less than three covalent bonds, the value 0.4 if they are separated by exactly three bonds, and the value one otherwise. Upon multiplication with $\epsilon(l,m)$ the electrostatic and van der Waals pair interactions are gradually switched off within the thus defined bonded region around a given atom m.

Taking the decomposition in Eq. (3) as our guideline, we start the discussion of the various contributions to the total potential energy $E(\mathbf{R})$ of a hybrid system with the nonbonded interactions and here, in particular, with the electrostatics.

1. Electrostatics

The electronic Hamiltonian of a QM fragment describes the motion of the electrons within the electrostatic potential generated by the nuclei and electrons. The electrostatic influence of an MM fragment on a QM fragment can easily be included by adding the electrostatic potential $\Phi_M(\mathbf{r};\mathbf{R}_M)$ originating from the charges of the MM fragment as an external potential to the Hamiltonian. If, in particular, a DFT approach is used, then one simply has to complement the energy functional[11,12] $E_Q[\rho;\mathbf{R}]$ of the QM fragment at a given configuration $\mathbf{R}$ of the hybrid system by the additional term

$$E_Q^{\mathrm{ext}}[\rho;\mathbf{R}_M]=\int \rho(\mathbf{r})\Phi_M(\mathbf{r};\mathbf{R}_M)d\mathbf{r}, \tag{4}$$

where $\rho(\mathbf{r})$ is the electronic charge density in the QM fragment. For its evaluation, a suitable representation of the potential function $\Phi_M(\mathbf{r};\mathbf{R}_M)$ has to be supplied to a DFT program.

a. The external potential for the QM fragment. The DFT method of our choice is available through the program CPMD.[22] In that program, the electronic wave function is calculated in a rectangular box, the so-called QM box, using plane wave expansions. As illustrated in Fig. 4 for a water molecule, the box is chosen such that it just covers the QM fragment. The electrostatic potential Φ_M originating from the charges c_l of the MM fragment has to be provided on a regular grid. The grid discretizes the QM box and is also indicated in the figure. For sizeable QM fragments containing 10–50 atoms, the grid may comprise several 10 000 points γ. For large systems, the evaluation of Φ_M at all these grid points through the Coulomb sum

$$\Phi_M(\mathbf{r}_\gamma;\mathbf{R}_M)=\sum_{l\in\mathbf{M}}\frac{c_l}{|\mathbf{r}_\gamma-\mathbf{r}_l|},\quad \gamma=1,2,\ldots$$

would pose a considerable computational task.

Using the FAMUSAMM method[50] the electrostatic interactions among the MM atoms are efficiently calculated through a hierarchical combination of multipole and Taylor expansions. We will now show how an extension of that approach to the task specified above also enables a very efficient computation of $\Phi_M(\mathbf{r}_\gamma;\mathbf{R}_M)$ on the grid.

According to the scheme depicted in Fig. 5, the extended FAMUSAMM algorithm derives the potential at a grid point γ through summation of many contributions, which are generated by a hierarchy of charge distributions in the MM fragment. The sizes of these distributions are shown to increase with the distance from the atom $s(\gamma)$ closest to γ on a coarse scale (note that Fig. 4 illustrates with dashed lines the classification $s(\gamma)$ of grid points according to this nearest neighbor criterion). The computational prescription, according to which a given contribution to Φ_M is evaluated, depends on the distance class to which the source of this contribution belongs. Distant charge distributions beyond class $\mathbf{M}^{(10)}$ are represented by their multipole moments.

For the innermost distance class $\mathbf{M}^{(5)}$ associated with atom $s(\gamma)$, the contribution $\Phi^{(5)}(\mathbf{r}_\gamma)$ is calculated from the Coulomb potentials

$$\Phi^{(5)}(\mathbf{r}_\gamma)=\sum_{l\in\mathbf{M}^{(5)}}\tilde{\epsilon}(|\mathbf{r}_\gamma-\mathbf{r}_l|)\frac{c_l\mathrm{erf}(|\mathbf{r}_\gamma-\mathbf{r}_l|/\sigma)}{|\mathbf{r}_\gamma-\mathbf{r}_l|} \tag{5}$$

of Gaussian charge distributions centered at the positions of the atoms. Such extended distributions are advantageous because the potential of an MM point charge, which happens to be close to or even within the QM box, can cause difficulties concerning the convergence of the wave function. A proper choice for the width σ is 0.8 Å . The switching function $\tilde{\epsilon}(|\mathbf{r}_\gamma-\mathbf{r}_l|)$appearing in Eq. (5) assumes values different from one only for atoms in the link atom region of covalently connected QM and MM fragments. Therefore its discussion is postponed to Sec. II B 4.

The contribution of the next distance class $\mathbf{M}^{(10)}$ is computed in a two-stage process. First, the potential is evaluated at the atom $s(\gamma)$ using the usual Coulomb sum

$$\Phi^{(10)}(\mathbf{r}_s)=\sum_{l\in\mathbf{M}^{(10)}}\frac{c_l}{|\mathbf{r}_s-\mathbf{r}_l|},$$

but in addition also the first and second derivatives of $\Phi^{(10)}$ are calculated at $\mathbf{r}_s$, yielding the coefficient vector

$$\mathbf{k}^{(10)}=\nabla_{\mathbf{r}}\Phi^{(10)}(\mathbf{r})\big|_{\mathbf{r}=\mathbf{r}_s}$$

and a corresponding coefficient matrix $\mathbf{T}^{(10)}$ accounting for the second derivatives. With these coefficients the contribution of the charges in $\mathbf{M}^{(10)}$ to the potential Φ_M at $\mathbf{r}_\gamma$ can be approximated through a second-order Taylor expansion. But before that expansion is evaluated at $\mathbf{r}_\gamma$, the coefficients are modified by adding contributions accounting for the multipole moments of the charge distributions in the outer distance classes. These contributions are obtained in a highly efficient way through a hierarchical scheme of nested Taylor expansions centered at further reference points in the vicinity of $\mathbf{r}_s$ (for details, see Ref. 50). The resulting coefficients $\Phi^{(\infty)}(\mathbf{r}_s)$, $\mathbf{k}^{(\infty)}$, and $\mathbf{T}^{(\infty)}$ then enter the expansion

$$\Phi^{(\infty)}(\mathbf{r}_\gamma)=\Phi^{(\infty)}(\mathbf{r}_s)+(\mathbf{r}_\gamma-\mathbf{r}_s)^t\mathbf{k}^{(\infty)}+\tfrac{1}{2}(\mathbf{r}_\gamma-\mathbf{r}_s)^t\mathbf{T}^{(\infty)}(\mathbf{r}_\gamma-\mathbf{r}_s) \quad (6)$$

expressing the contributions of all outer distance classes. The external potential Φ_M at a grid point γ is then the sum

$$\Phi_M(\mathbf{r}_\gamma;\mathbf{R}_\mathrm{M})=\Phi^{(5)}(\mathbf{r}_\gamma)+\Phi^{(\infty)}(\mathbf{r}_\gamma). \quad (7)$$

Thus, at each grid point the sum in Eq. (5) and the expansion in Eq. (6) have to be evaluated. As a prerequisite, for each QM atom, an interaction list $\mathbf{M}^{(5)}$ with positions and charges of corresponding MM atoms and a set of ten expansion coefficients have to be provided. The interaction lists typically contain only about 100 elements and, therefore, the effort for the computation of Φ_M on the grid is limited. In the actual implementation this computational step is executed by a module of CPMD since then the properly extended MD program, which we will call from now on EGO_IX, has to transfer only a small set of data to CPMD. The implementation also provides some additional measures to enhance the accuracy of the computation of Φ_M (for these details, see Ref. 80).

b. The external potential for the MM fragment. For the treatment of the converse electrostatic influence of the QM fragment on the MM fragment, the electrostatic potential $\Phi_Q(\mathbf{r};\mathbf{R}_Q,\rho)$ originating from the electronic charge density ρ and the nuclear core charges C_s has to be evaluated. According to the Hellman–Feynman theorem the forces exerted on the MM atoms m with charges c_m are then obtained by calculating the gradient of that potential at the positions $\mathbf{r}_m$.

One may, however, simplify the calculation by approximating that electrostatic potential through the use of partial charges. Such charges c_s can be assigned to the QM atoms s by requiring that they should generate a potential $\tilde{\Phi}_Q$ which optimally represents the actual potential Φ_Q in the region outside the van der Waals boundary of the QM fragment.[81] When these partial charges together with the atomic positions $\mathbf{r}_s$ are fed into the FAMUSAMM scheme the contributions $\tilde{\Phi}_Q(\mathbf{r};\mathbf{R}_Q)$ of the QM fragment to the total electrostatic potential $\Phi(\mathbf{r};\mathbf{R})$ in the MM fragment can be obtained in exactly the same way as the contributions $\Phi_M(\mathbf{r};\mathbf{R}_M)$ generated by the charges of the MM fragment. Thus, the total potential within the MM fragment is approximately given by

$$\Phi(\mathbf{r};\mathbf{R})\approx\tilde{\Phi}_Q(\mathbf{r};\mathbf{R}_Q)+\Phi_M(\mathbf{r};\mathbf{R}_M) \quad (8)$$

and the electrostatic energy of the MM fragment is

$$E_{M,\mathrm{estat}}(\mathbf{R})=\sum_{m\in\mathbf{M}}c_m\,\Phi(\mathbf{r}_m;\mathbf{R}). \quad (9)$$

As sketched above, the use of external potentials allows the electrostatic action of either fragment on its respective counterpart to be accounted for in a straightforward manner and, therefore, has been applied in all previous hybrid methods. Note, however, that the use of the partial charge approximation for the action of the QM fragment on the MM fragment violates the principle that *actio* should be equal to *reactio*. Therefore, methods allowing corresponding dynamical artifacts to be corrected have to be provided and the quality of the treatment has to be scrutinized through appropriate test calculations (see Secs. II B 5 and III A below). Less critical in this respect is the treatment of the second type of nonbonded interaction, i.e., that of the van der Waals forces.

2. Van der Waals interactions

According to the CHARMM force field[30] the van der Waals interactions between nonbonded atoms l and m in an MM fragment are modeled by 6-12-Lennard-Jones potentials giving rise to energy contributions $E_{\mathrm{vdw}}(\mathbf{r}_l,\mathbf{r}_m)$. Within the QM fragment, these interactions are represented implicitly. But for the description of the van der Waals interactions between atoms q and m of the QM and MM fragments, respectively, one has to resort to the MM approach and has to include energy terms $E_{\mathrm{vdw}}(\mathbf{r}_q,\mathbf{r}_m)$. Upon proper summation of all van der Waals terms, one obtains a contribution $E_{\mathrm{vdw}}(\mathbf{R})$ to the total energy. Due to their rapid decay, the corresponding forces may be safely neglected at distances larger than about 10 Å throughout a given simulation system.

3. Bonded interactions

In CHARMM[30] the bonded contribution $E_{\mathrm{bonded}}(\mathbf{R}_M)$ to the potential energy of an MM fragment is specified by harmonic potentials E_{stretch} and E_{angle}, respectively, driving the bond lengths and bond angles towards default values for the given chemical bonding patterns. Additional nonharmonic potentials E_{stereo} describe forces originating from twisted bonds and other stereochemical effects. All these terms solely depend on the coordinates of at most four atoms covalently connected by three consecutive chemical bonds.

Since the bonded interactions within the QM fragment are already covered by $E_Q[\rho;\mathbf{R}]$, only a suitable energy expression remains to be given for simulation systems in which the two fragments are joined by covalent bonds.

4. Scaled position link atom method (SPLAM)

Unpolar C—C bonds joining CH_2 groups are ubiquitous in proteins. Therefore, in many cases, the partitioning of the simulation system into covalently linked QM and MM fragments can be chosen such that the joining covalent bonds are

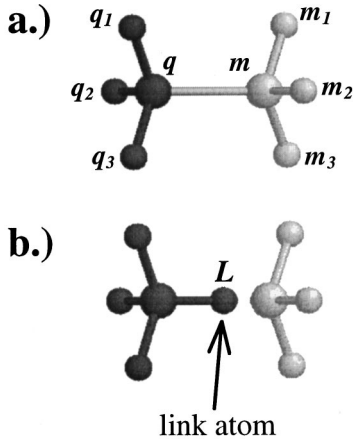

FIG. 6. Link atom concept exemplified for the ethane molecule: (a) the molecule is partitioned into two CH_3 fragments, a QM fragment drawn in black and an MM fragment drawn in gray; (b) the carbon atom *m* originally bound to the carbon atom *q* is replaced by the hydrogen atom *L* in order to recover a saturated pattern of covalent bonds in the QM fragment.

of that type. Note that the central C—C bond of the ethane molecule depicted in Fig. 6 provides a minimal model for such bonds. Therefore, we will initially use this molecule as a most simple example during the following presentation of our treatment of covalently joined fragments.

As outlined in the Introduction and illustrated in Fig. 6 for the ethane molecule, in such cases an artificial hydrogen atom, the so-called link atom, may be introduced into the QM fragment. The link atom replaces the missing bonding partner and serves to reconstruct a saturated valence structure. However, such a replacement perturbs the electronic wave function in the vicinity of the replaced atom and correspondingly changes the forces acting on the nearby nuclei in the QM fragment. The desired treatment should be designed to correct these errors, remove the added degrees of freedom, and restore the mechanical properties of the missing bond.

The CHARMM model of the ethane molecule describes the forces on the central carbon atoms *q* and *m* solely in terms of bonded interactions, because the switching functions $\epsilon(i,q)$ and $\epsilon(i,m)$ weighting the respective nonbonded interactions vanish at all of its atoms *i*. Thus, that molecule allows concentration on the bonded interactions in the derivation of the link atom method. We, therefore, start considering the forces originating from these dominant interactions.

a. Bonded interactions. We have stressed above that replacing a carbon atom in a C—C bond by a hydrogen atom *L* locally perturbs the wave function of the QM fragment. To keep that perturbation as small as possible we orient the artificial C—H bond relative to the remaining QM fragment in the same way as the original C—C bond would be oriented in a more extended QM fragment. This is achieved by constraining the position $\mathbf{r}_L$ of *L* to the line linking the position $\mathbf{r}_q$ of the C atom *q* in the QM fragment to the position $\mathbf{r}_m$ of the other C atom *m* of the original bond (the notation is that of Fig. 6). Thereby two of the three additional degrees of freedom connected with the artificial link atom have been removed from the description.

In order to remove also the remaining degree of freedom we adjust the extension of the artificial C—H bond to that of the original C—C bond by a simple scaling procedure. If, in a given configuration $\mathbf{R}$ of the system, the C—C distance $r_{\mathrm{CC}} \equiv |\mathbf{r}_q - \mathbf{r}_m|$ deviates from the equilibrium distance r^0_{CC} then we choose the C—H distance $r_{\mathrm{CH}} \equiv |\mathbf{r}_q - \mathbf{r}_L|$ according to the simple formula

$$r_{\mathrm{CH}} = r^0_{\mathrm{CH}} + \frac{k_{\mathrm{CC}}}{k_{\mathrm{CH}}}(r_{\mathrm{CC}} - r^0_{\mathrm{CC}}), \tag{10}$$

where r^0_{CH} is the equilibrium C—H bond length, and k_{CC}, k_{CH} are typical force constants of the corresponding bond stretches.

As can easily be deduced from a consideration of the respective bond stretching potentials in the harmonic approximation, the scaled position has been chosen such that the streching force of the C—H bond closely approximates that of a C—C bond linking the atoms *q* and *m*. Note, however, that the energy of a C—C bond extended to a distance r_{CC} is somewhat larger than the energy of a much stiffer C—H bond at the scaled distance. Therefore, a corresponding energy correction

$$\Delta E_{\mathrm{stretch}}(\mathbf{r}_q, \mathbf{r}_m) = k_{\mathrm{CC}}(1 - k_{\mathrm{CC}}/k_{\mathrm{CH}})(r_{\mathrm{CC}} - r^0_{\mathrm{CC}})^2 \tag{11}$$

is required. Because of the scaling procedure expressed by Eq. (10), which fixes the last undetermined parameter for the derivation of $\mathbf{r}_L$ from $\mathbf{r}_m$, we have termed our approach ''scaled position link atom method'' (SPLAM).

Technically speaking, the scaled position $\mathbf{r}_L$ of the link atom is calculated by the MD program EGO_IX according to the prescriptions given above from the positions $\mathbf{r}_q$ and $\mathbf{r}_m$ of the two atoms, whose bond is removed by the partitioning. Then, $\mathbf{r}_L(\mathbf{r}_q, \mathbf{r}_m)$is included in the configuration vector $\mathbf{R}_Q$ and passed to the DFT program CPMD. That program calculates E_Q, the partial charges c_s, and the forces $\mathbf{f}_s$ on all atoms *s* in the saturated QM fragment and returns these data to the MD program including a force $\mathbf{f}_L = -\nabla_L E_Q[\rho;\mathbf{R}]$ acting on the link atom. By construction, the stretching component $\mathbf{f}_{L,\mathrm{s}}$ of that force closely approximates the corresponding force component of $-\nabla_m\{E_Q[\rho;\mathbf{R}] + \Delta E_{\mathrm{stretch}}(q,m)\}$, where the dependence of $\mathbf{r}_L$ on $\mathbf{r}_m$ is used in the calculation of the gradient. Therefore, and this is a key point of SPLAM, the stretching component $\mathbf{f}_{L,\mathrm{s}}$ of the DFT force on *L* is simply interpreted as a force acting on the atom *m* of the MM fragment and is added to the MM forces acting on that atom.

Figure 7 demonstrates for the central C—C bond of the ethane molecule that the stretching forces obtained by the SPLAM procedure represent a very good approximation that also covers anharmonicities. The figure compares the absolute values of the SPLAM forces acting on the carbon atoms *q* and *m* at various C—C distances (dashed curve) with corresponding values obtained by a complete DFT treatment (solid curve) and with absolute values of harmonic MM stretching forces.

In the case of the ethane molecule the remaining force component $\mathbf{f}_{L,\mathrm{p}}$, which is oriented perpendicular to $\mathbf{f}_{L,\mathrm{s}}$, reflects the angle terms $E_{\mathrm{angle}}(q_\alpha, q, L)$, $\alpha \in \{1,2,3\}$ involving the link atom *L* and the QM atoms *q* and q_α (for the notation, cf. Fig. 6). But in contrast to the stretching component, this component of $\mathbf{f}_L$ may not be simply added to the forces act-

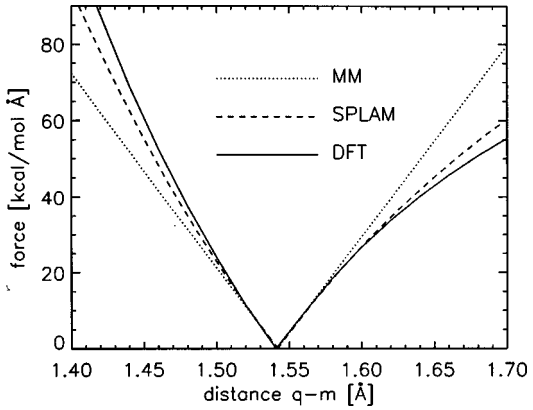

FIG. 7. Absolute values of the forces acting on the carbon atoms q and m, respectively, of the ethane molecule depicted in Fig. 6, as obtained by variation of the C—C distance; the solid curve results from a DFT treatment of the complete ethane molecule, the dashed curve represents the SPLAM result and the dotted lines are optimal harmonic MM forces; the force constants and equilibrium bond lengths required for the latter have been determined by DFT; the geometry has been optimized during stretching; the details of the DFT method applied are given in Secs. III A 1 and III B 1.

ing on atom m since such a transfer would induce a torque and, hence, violate the conservation of angular momentum. This artifact can be avoided by splitting the perpendicular component into fractions $(r_{\mathrm{CH}}/r_{\mathrm{CC}})\mathbf{f}_{L,\mathrm{p}}$ and $(1-r_{\mathrm{CH}}/r_{\mathrm{CC}})\mathbf{f}_{L,\mathrm{p}}$, of which the former is added to the force on m and the latter to $\mathbf{f}_q$.

The angular forces derived in this way from E_{Q} will deviate from forces expected according to CHARMM for an H—C—C bonding motive since they have been calculated for an H—C—H motive. An MM energy term effecting a corresponding correction is

$$\Delta E_{\mathrm{angle}}(\{q_\alpha\},q,m)=\sum_{\alpha\in\{1,2,3\}}(k_{\mathrm{HCC}}-k_{\mathrm{HCH}})[\theta(q_\alpha,q,m)-\theta^0(q_\alpha,q,m)]^2, \tag{12}$$

where the force constants have the usual CHARMM values and where the equilibrium angles θ^0 have been chosen to be identical for the two motives.

It remains to be stated that the replacement of the C—C bond by the artificial C—H bond completely removes a couple of bonded interactions which, in terms of the CHARMM force field, involve angular terms $E_{\mathrm{angle}}(q,m,m_\alpha)$ and stereochemical contributions $E_{\mathrm{stereo}}(q_\beta,q,m,m_\alpha)$, $\alpha,\beta\in\{1,2,3\}$, (cf. Fig. 6). Restoring these contributions the SPLAM energy correction for the bonded interactions of an ethane molecule is

$$\Delta E_b(\mathbf{R})=\Delta E_{\mathrm{stretch}}(\mathbf{r}_q,\mathbf{r}_m)+\Delta E_{\mathrm{angle}}(\{q_\alpha\},q,m)+\sum_{\alpha\in\{1,2,3\}}\left[E_{\mathrm{angle}}(q,m,m_\alpha)+\sum_{\beta\in\{1,2,3}E_{\mathrm{stereo}}(q_\beta,q,m,m_\alpha)\right], \tag{13}$$

where the terms $\Delta E_{\mathrm{stretch}}$ and $\Delta E_{\mathrm{angle}}$ are given by Eqs. 11 and 12, respectively.

Interestingly, a close inspection of the CHARMM force field reveals that this SPLAM energy function for the bonded interactions in ethane applies *identically* to C—C bonds between CH_2 groups in a protein. But in these more general cases, some further corrections concerning nonbonded interactions have to be considered. We will now present a couple of CHARMM-type energy corrections which serve to remove the corresponding influence of the link atom on the energetics and dynamics of the system as far as possible.

b. Non-bonded interactions. In contrast to the actual C—C bond the artificial C—H bond is polar. The energy of this dipole in the external potential Φ_M and the resulting contributions to the DFT forces $\mathbf{f}_L$ and $\mathbf{f}_q$ represent artifacts. A corresponding energy correction may be estimated from the partial charge c_L according to

$$\Delta E_{\mathrm{dip,ext}}(\mathbf{r}_L,\mathbf{r}_q;\mathbf{R}_M)=-c_L[\Phi_M(\mathbf{r}_L;\mathbf{R}_M)-\Phi_M(\mathbf{r}_q;\mathbf{R}_M)]. \tag{14}$$

Here, the external potential Φ_M is obtained through the FAMUSAMM scheme from the partial charges c_l of the MM atoms l. For the sake of consistency, in that calculation the switching function $\epsilon(l,L)\equiv\epsilon(l,m)$ introduced in connection with Eq. (3) has to be used. This function classifies the nonbonded interactions with the link atom L in terms of the actual bonding pattern of the atom m. As a result the electrostatic interactions with the MM atoms $l(m)$ located in the bonded region of atom m do not contribute to Φ_M. But then Eq. (14) can represent a good approximation to the DFT energy of the artificial dipole *only if* a corresponding classification scheme has also been applied to the calculation of the external potential Φ_M in the QM box. As indicated by the switching function $\tilde{\epsilon}(|\mathbf{r}_\gamma-\mathbf{r}_l|)$in Eq. (5) we therefore switch off smoothly at small distances the near zone contributions $\Phi^{(5)}(\mathbf{r}_\gamma)$ originating from the MM atoms $l(m)$. We use the function

$$\tilde{\epsilon}(r)=\begin{cases}0 & \text{if } r\leqslant d\\ 1-[1-(r-d)^2/\rho^2]^2 & \text{if } d<r<d+\rho\\ 1 & \text{else}\end{cases}$$

with the choices $d=0.5$ Å and $\rho=2.5$ Å for the parameters. Note that Eq. (14) yields the desired corrections to the forces $\mathbf{f}_s$, $s\in\{L,q\}$ in the usual way through the negative gradients $-\nabla_s\Delta E_{\mathrm{dip,ext}}(\mathbf{r}_L,\mathbf{r}_q;\mathbf{R}_{\mathrm{M}})$.

For the converse calculation of the electrostatic action of the QM fragment on its MM surroundings, the artificial C—H dipole has to be removed. This can be effected by adding c_L to c_q which ensures conservation of charge within the QM fragment. The electrostatic potential $\tilde{\Phi}_Q(\mathbf{r};\mathbf{R}_Q)$ is then calculated through FAMUSAMM using the corrected partial charges [cf. Eq. (8)].

In terms of the CHARMM force field, the DFT treatment of the QM fragment covers interactions of the artificial C—H dipole not only with the external potential but also with the partial charges of nonbonded atoms s within that fragment. An estimate for the corresponding energy correction is

$$\Delta E_{\mathrm{dip,int}}(\mathbf{R}_Q) = -\sum_{s\in\mathbf{Q}} \epsilon(s,L) c_s c_L \left(\frac{1}{|\mathbf{r}_s - \mathbf{r}_L|} - \frac{1}{|\mathbf{r}_s - \mathbf{r}_q|} \right), \tag{15}$$

where the switching function $\epsilon(s,L)$ is defined by the actual bonding pattern of atom m in the system (see above). This expression renders small corrections to nearly all DFT forces, particularly $\mathbf{f}_L$ and $\mathbf{f}_q$, and completes our CHARMM-inspired attempt to approximately remove the electrostatic influence of the link atom on the dynamics of the system.

Finally, the van der Waals interaction energy $\Delta E_{\mathrm{vdw}}(\mathbf{R}_Q)$ of the link atom with the other QM atoms, which can be approximated through the CHARMM force field, has to be subtracted in order to wipe out the last traces of the link atom.

Now the description of the SPLAM method is complete. Collecting Eqs. (13), (14), (15), together with $\Delta E_{\mathrm{vdw}}(\mathbf{R}_Q)$ yields the SPLAM energy function

$$E_{\mathrm{SPLAM}}(\mathbf{R}) = \Delta E_b(\mathbf{R}) + \Delta E_{\mathrm{dip,ext}}(\mathbf{r}_L, \mathbf{r}_q; \mathbf{R}_M) + \Delta E_{\mathrm{dip,int}}(\mathbf{R}_Q) + \Delta E_{\mathrm{vdw}}(\mathbf{R}_Q). \tag{16}$$

In the interpretation of that energy term the additional prescriptions concerning orientation and scaling of the C—C bond, splitting and transfer of the force $\mathbf{f}_L$, transfer of the partial charge c_L, and the appropriate use of the switching functions have to be kept in mind. Extensions of the SPLAM approach towards cutting and pasting other types of σ bonds than just CH_2—CH_2 bonds are straightforward.

Now all contributions to the potential energy of a hybrid system have been introduced and solely a few questions concerning the accuracy of the applied approximations remain to be addressed.

5. Dynamical corrections

We have noted above in connection with the partial charge approximation for the potential Φ_Q [cf. Eq. (8)] that this approximation weakly violates the reaction principle. Such violations are quite a common property of MD methods, which, for the sake of computational efficiency, apply approximations to the calculation of the electrostatics. They are also present in the FAMUSAMM algorithm. The resulting small deviations of the approximate forces from the exact ones can be considered as ''algorithmic noise.'' In FAMUSAMM, this noise is mainly due to the force extrapolations connected with the multiple time step procedure, whereas the second ingredient, the structure-adapted fast multipole method, is much more accurate.[50]

Algorithmic noise adds heat to the system, which can be removed through velocity scaling commonly used for simulations in the canonical ensemble. But noise can also affect dynamical properties as thoroughly discussed in Refs. 41,50,82. Here we merely mention that for FAMUSAMM, as opposed to other computational schemes, no changes of the dynamics could be identified.

For our hybrid model, in which the external potential Φ_M of the QM fragment is obtained through FAMUSAMM, the question is whether the fluctuations of the potential due to the force extrapolations can induce sizeable fluctuations of the electron density during a simulation. Test simulations of a QM water molecule embedded in a shell of MM water molecules (see also Sec. III A 1) have shown that a sufficiently smooth temporal behavior is obtained if the force extrapolations are applied to the outer distance classes starting at $\mathbf{M}^{(10)}$ instead of $\mathbf{M}^{(5)}$ used in the original MM version (cf. Fig. 5).

Concerning the simulation system as a whole, algorithmic noise can cause small total forces and torques. During a simulation they may add up and induce sizeable linear and angular momenta. One has to provide corrections for these two artifacts and, in fact, for pure MM simulations corresponding procedures[80] are implemented in the program EGO_IX. In the context of the hybrid method, they are applied to the electrostatic interactions within the MM fragment.

But here additional corrections have to be included since the total electrostatic reaction force

$$\mathbf{f}_{M,Q} = \sum_{m\in\mathbf{M}} \mathbf{f}_{m,Q},$$

on the MM fragment, where the local reaction forces $\mathbf{f}_{m,Q} = -c_m \nabla_m \tilde{\Phi}_Q(\mathbf{r}_m; \mathbf{R}_Q)$ are evaluated using the partial charge approximation, does not exactly cancel the total force on the QM fragment

$$\mathbf{f}_{Q,M} = \sum_{s\in\mathbf{Q}} \mathbf{f}_s,$$

which is caused by Φ_M and calculated by the CPMD program. Similar considerations apply to the respective total torques $\mathbf{t}_{Q,M}$ and $\mathbf{t}_{M,Q}$.

To ensure the conservation of the total linear momentum, we chose to apply corrections solely to the more inaccurate local reaction forces $-\mathbf{f}_{m,Q}$. For that purpose, we slightly rotate and scale the total force $\mathbf{f}_{M,Q}$ by multiplication with a matrix $\mathbf{S}$ such that $\mathbf{S}\mathbf{f}_{M,Q} = -\mathbf{f}_{Q,M}$ and observe that this condition completely determines $\mathbf{S}$ at each time step of an MD simulation. Multiplying $\mathbf{S}$ with the $\mathbf{f}_{m,Q}$ then yields the corrected local reaction forces. As to guarantee also the conservation of the total angular momentum, we apply corrections designed by a corresponding strategy (for details, see Ref. 80).

III. SIMPLE APPLICATIONS AND QUALITY CHECKS

In the remainder of this article, we will present the results of a series of test calculations which serve to check the validity of the hybrid method introduced above. As observables we will frequently use vibrational spectra since such spectra provide a most sensitive check of the force field quality. We start with tests concerning our treatment of the nonbonded interactions, as introduced in Secs. II B 1 and II B 2.

A. Non-bonded interactions

Liquid water is a highly polar fluid composed of very small molecules. Therefore, a hybrid simulation of pure water, in which the QM fragment is a single water molecule and the MM fragment is a large water environment, enables a first look at our description of the nonbonded interactions.

The properties of liquid water have been extensively studied, both experimentally and through computer simulations. Thus, comparisons of our results with those of similar hybrid simulations (see, e.g., Ref. 83) and of completely quantum mechanical CPMD simulations[84,21] can provide a quality check for our method.

1. A water molecule in aqueous solution

For a test simulation, we chose a water sphere composed of 534 TIP3P water molecules.[85] The sphere was enclosed by a soft boundary potential[86] with a radius of 15 Å . The O–H bond lengths of the TIP3P water molecules were kept fixed using the SHAKE algorithm.[87] But in contrast to the original TIP3P model, which is completely rigid, we chose the H–O–H angle to be flexible with a harmonic angular force constant of 55 kcal/mol deg^2. To keep the temperature constant at 300 K, all atoms were weakly coupled to a heat bath by velocity rescaling with a relaxation time $\tau = 0.1$ ps.[88]

Using the FAMUSAMM algorithm and an integration time step of 1.0 fs, the system was first equilibrated for 200 ps. In the resulting structure, a molecule at the center of the sphere was selected as the QM fragment for the subsequent hybrid simulations. For its CPMD description the Vanderbilt pseudopotential,[89] the local density approximation,[17] and a plane wave cutoff of 30 Rydberg were chosen, rendering a cost-effective but not necessarily most accurate approach denoted by local density approximation (LDA). Then the system was again equilibrated for 8 ps using a reduced time step of 0.25 fs. Finally, a 30 ps simulation was carried out for data collection; that time span is at least a factor of 5 longer than in the comparable studies quoted above.

a. Dipole moment. Using LDA, one obtains for the dipole moment of an *isolated* water molecule a value of 1.92 D, which is very close to the gas phase value of 1.85 D determined by measurements of the Stark effect.[90] The agreement with experiment may be further improved by application of generalized gradient corrections (GC), yielding a value of 1.86*D*.[84]

Due to its polarizability, the dipole moment of a water molecule increases greatly upon solvation in bulk liquid water. For water molecules in ice, a value of about 2.6 D has been estimated.[91] That analysis has been confirmed by a fully quantum mechanical CPMD simulation of 32 D_2O molecules in a 9.6 Å periodic box. Here an average dipole moment of 2.66 D has been calculated using LDA+GC, corresponding to an induced dipole moment of 0.80 D.[84]

The induced dipole moment of a solvated water molecule provides a clue to the average local electrostatic field generated by its structured environment. As a first check on the quality of our treatment of the nonbonded interactions, we have therefore extracted the dipole moment of the QM fragment from our hybrid simulation.

The results are collected in Fig. 8. The figure shows that the dipole moment of the QM fragment fluctuates along the C_2 symmetry axis of H_2O around an average value of 2.77 D, with a standard deviation of 0.18 D. The components of the induced dipole moment perpendicular to that axis vanish on the average and show slightly smaller fluctuations with standard deviations of about 0.15 D.

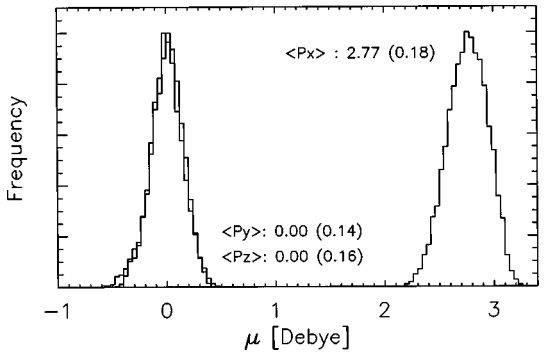

FIG. 8. Distributions of the components of the dipole moment of a QM water molecule solvated in TIP3P water; also given are the mean values and standard deviations along the symmetry axis (x axis) of H_2O and along the two perpendicular axes (the y axis is in the molecular plane).

The value of 0.85 D thus determined for the total induced dipole moment is in excellent agreement with the result of the fully quantum mechanical simulation.[84] The value indicates that the local electrostatic field in the hybrid simulation was about 6% stronger. Previous hybrid simulations[76,83] have obtained somewhat smaller induced dipole moments in the range between 0.38 and 0.66 D using the same set of approximations employed here for the QM part.

In Sec. II B 5, we stressed that the FAMUSAMM treatment of the electrostatics implies a certain amount of algorithmic noise. Within the QM fragment that noise should be small enough to ensure a smooth temporal behavior of physical observables, which, like the induced dipole moment, are directly proportional to the external field.

Figure 9 demonstrates for the dipole moment of the QM fragment that the fluctuations of that observable, when represented at a very high temporal resolution, are actually very smooth functions of time. Thus, fluctutations due to algorithmic noise are orders of magnitude smaller than the physical fluctuations that reflect the thermal motion and are depicted in the figure.

The strength of the local field as witnessed by the size of the induced dipole moment of the QM fragment depends, of course, on the parameterization of the TIP3P model. Concerning the latter, we first note that an isolated TIP3P molecule has a dipole moment of 2.35 D, which is closer to the dipole moment of a water molecule in the bulk liquid than to that in a non-polar environment. Due to our introduction of a flexible H–O–H angle, the TIP3P model acquires a slight

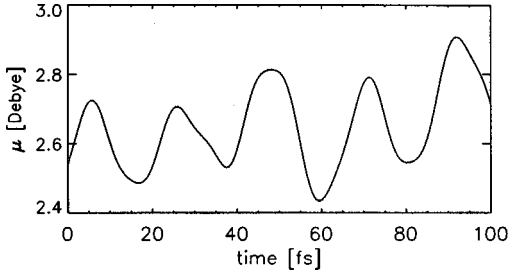

FIG. 9. Temporal development of the absolute value of the dipole moment of a solvated water molecule.

polarizability. Upon collection into bulk water, the TIP3P models elongate by 4% such that their dipole moments increase to an average value of 2.44 D.

b. Geometry and vibrational spectra. Actually, besides its dipole moment, also the geometry of a water molecule changes upon solvation. In our hybrid simulation, we found excellent agreement of these changes both with experimental data[92] and previous hybrid simulations.[83] But we will skip these details[80] here and turn to more sensitive checks of the force field by considering the vibrational spectra of H_2O and the associated effects of solvation.

A water molecule has three normal modes, the antisymmetric (*a*) and symmetric (*s*) stretches of the O–H bonds and the bend (*b*) of the O–H–O angle. In the gas phase,[93] these modes are observed at (*a*) 3756, (*s*) 3652, and (*b*) 1595 cm^{-1}; in the liquid,[94,95] the two stretching modes are strongly red-shifted by about 250 cm^{-1}, whereas the bending mode is shifted to the blue by about 75 cm^{-1}. Upon solvation the stretching bands experience a sizeable broadening from a few cm^{-1} in the gas phase to as much as about 250 cm^{-1} in the liquid.[94]

For the QM description of the gas phase spectrum, we first applied the harmonic approximation, i.e., using the CPMD program we calculated the Hessian matrix of an isolated water molecule at its equilibrium geometry, rendering the normal modes and frequencies in the usual way.[96] To estimate the influence of anharmonic effects, we also carried out a CPMD simulation of an isolated water molecule at 300 K over a time span of 6 ps. From the resulting trajectory, we calculated the autocorrelation functions for the three normal coordinates and obtained the vibrational spectrum by subsequent Fourier transformation (cf. Ref. 83). In the same way, the liquid spectrum was computed from the 30 ps trajectory of the QM fragment in the hybrid simulation.

According to our LDA calculations, the harmonic frequencies of the isolated water molecule have the values (*a*) 3752, (*s*) 3623, and (*b*) 1481 cm^{-1}, whereas the corresponding values derived from the CPMD simulation are 3757, 3620, and 1480 cm^{-1}, respectively. At the chosen temperature of 300 K, the anharmonic effects are thus negligibly small.

The QM/MM hybrid simulation for the liquid predicts redshifts of the stretching frequencies of (*a*) 123 cm^{-1} and (*s*) 66 cm^{-1} and a blueshift of the bending mode of 84 cm^{-1}. The sizes of the shifts agree quite well with the results of the comparable study of Alhambra *et al.*,[83] and, concerning the bending mode, also with the experimental findings. The shifts of the stretching modes are too small; these underestimates are probably due to quantum effects missing in the classical description of the hydrogens. For the widths of the bands at half maximum height, we have found the values (*a*) 180, (*s*) 230, and (*b*) 50 cm^{-1} matching closely the observed band widths.

In view of the well-known deficiencies of LDA concerning the description of hydrogen bonds (see Refs. 84 and 21 for discussions), the expert reader might wonder why our rather simple DFT treatment, when combined with molecular mechanics, can produce the reasonable results presented above. Therefore, we will now take a still closer look at the description of non-bonded interactions by the hybrid force field, using the water dimer as an example.

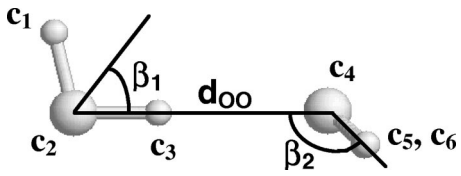

FIG. 10. Equilibrium structure of a water dimer; parameters for characterization of the intermolecular arrangement are given; also indicated are the locations of the six partial charges c_i required for the MM description.

2. Water dimer calculations

A water dimer is a strongly hydrogen bonded system whose equilibrium intermolecular arrangement can be characterized by the geometrical parameters indicated in Fig. 10. The structure is known experimentally and has been successfully described by DFT calculations.[21] In order to analyze our hybrid approach, we have calculated the water dimer in its equilibrium geometry with a set of different methods. As observables for quality assessment we use the structure, the association energy E_b, and the vibrational spectrum.

a. Structure and association energy. As shown in Fig. 10 the water dimer exhibits a broken symmetry. Therefore, we have looked at two different hybrid partitionings for that hydrogen bonded system: one in which the hydrogen donor is assigned to the QM fragment and the acceptor to the MM fragment (QM model) and the other with the converse partitioning (MQ). As references we take results of calculations, in which the complete dimer has been treated either by quantum mechanics (QQ) or by molecular mechanics (MM).

To check the influence of the DFT method on the hybrid description, we calculated the QM fragment not only by the simple LDA procedure given above. We also used the more elaborate Becke–Lee–Yang–Parr (BLYP) exchange-correlation functional together with the Martins–Troullier pseudopotential and a plane wave cutoff 70 Ry (for discussion and references, see Ref. 21). We denote this particular combination simply as BLYP.

The van der Waals parameters of the MM force field were taken from the TIP3P model of bulk water, but the partial charges were adjusted to the situation of an isolated dimer. Both DFT methods yield similar partial charges for the dimer. Therefore, we chose the values $c_1=0.36$, $c_2=-0.78$, $c_3=0.42$, $c_4=-0.74$, and $c_{5,6}=0.37$ with the numbering given in Fig. 10. The force constants for the O–H stretch and the H–O–H bend required for the calculation of vibrational spectra of MM fragments were determined by LDA and BLYP calculations on the H_2O monomer.

The optimized geometries and association energies obtained for the various models and DFT methods are summarized in Table I; the table also contains experimental data.

As had been shown previously[21] the results for the QQ model obtained with the elaborate BLYP procedures agree very well with the observations whereas the LDA treatment strongly overestimates the association energy E_b and underestimates the oxygen–oxygen distance d_{OO}.

TABLE I. Equilibrium geometries and association energies of water dimers at different levels of treatment; d_{OO} in Å , β_i in degrees, and E_b in kcal/mol.

	d_{OO}	β_1	β_2	E_b
Exp[97]	2.98	51	122	−5.3
BLYP QQ	2.98	56	123	−4.2
LDA QQ	2.73	53	136	−12.0
BLYP QM	2.78	52	163	−6.5
LDA QM	2.83	53	163	−6.1
BLYP MQ	2.85	50	117	−5.8
LDA MQ	2.80	52	120	−6.4
MM	2.77	47	162	−6.6

The MM results presented in the last line of the table reflect the choice of the van der Waals parameters and of the partial charges: the slight overestimate of E_b and underestimate of d_{OO} indicate that the van der Waals radii are a little too small; the angle β_2 is much too large since the stereochemical action of the oxygen lone pair electrons is neglected in the three-point charge model used here.

The results of the hybrid calculations are nearly identical for the two DFT approaches. The values of E_b and d_{OO} closely agree with the corresponding results for the MM model. In the case of the MQ hybrid model, the reduced value of the angle β_2 reveals the effects of the oxygen lone pairs, which are properly covered by both DFT treatments of the QM fragment. In the QM case, in which the electrostatics of the acceptor is modeled by three-point charges, this effect is lacking.

Thus, we may conclude that a four-point charge model for the MM fragment should be better suited for hybrid descriptions of structural details whenever water molecules are involved. Furthermore, no essential benefit is gained concerning structure and interaction energy by application of the more costly BLYP approach.

Note that results of hybrid calculations carried out by Tuñón *et al.*,[76] in which a Becke–Parr exchange-correlation functional had been used, support the above analysis. We would also like to add that the effects concerning the intramolecular structures of the H_2O molecules upon dimer formation turned out to be similar and small in all models.[80]

b. Vibrational spectra of the dimer. In contrast to the observables discussed so far, the vibrational spectra of the hybrid models should exhibit a distinct dependence on the choice of the DFT description, since the monomer shows such a dependence.[21] Using LDA and BLYP, we therefore calculated these spectra in the harmonic approximation by diagonalizing the mass-weighted Hessian matrices of second derivatives.

The high-frequency region of the dimer spectrum covers the O–H stretching vibrations and, therefore, accounts for the influence of the hydrogen bond on the intramolecular force fields. The low-frequency region comprises relative motions of the two monomers and, thus, directly monitors the properties of the hydrogen bond.

Figure 11 pertains to the LDA functional and compares the spectra of the QQ, QM, MQ and MM models in the high- and low-frequency regions with experimental data. The antisymmetric and symmetric O–H stretching modes localized in the donor (D) or acceptor (A) are denoted by a and s. Vibrational frequencies associated with modes localized in the respective QM fragments are drawn with thick lines.

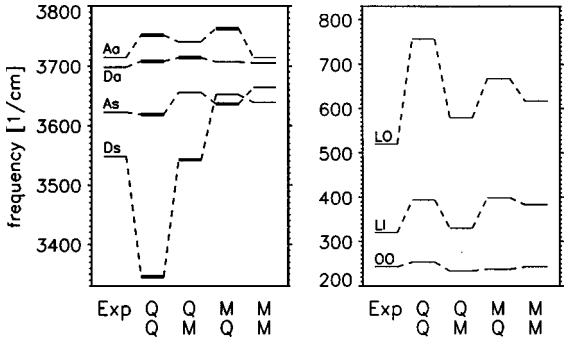

FIG. 11. Vibrational spectra of a water dimer in the high- and low-frequency regions; QQ: LDA for the dimer, QM and MQ: corresponding hybrid models; also given are MM results and experimental data (quoted from Ref. 21); the force constants of the MM fragments have been obtained from LDA results on the H_2O monomer.

c. High-frequency O–H stretches. According to the experimental data,[21] dimer formation leads to redshifts of the stretching frequencies measuring about 60–110 cm^{-1} for the donor and 30 cm^{-1} for the acceptor. In the QQ model, the BLYP functional (data not shown) correctly accounts for the redshifts of the donor modes, yielding shifts of 35–112 cm^{-1}. LDA, in contrast, predicts for the (Ds) mode (cf. Fig. 11) a redshift of 278 cm^{-1}, which is much too large. This deviation is in line with the LDA overestimate of the stability of the hydrogen bond and reflects a corresponding underestimate of the O–H force constant in the donor. Concerning the acceptor bands neither LDA nor BLYP reproduce the observed redshifts. The band pattern predicted by the MM model for the O–H stretches does not match the observations (cf. Fig. 11); an analysis has shown[80] that these deficiencies are partially inherited from the poor mechanical model of the monomer.

Figure 11 demonstrates that the *MM fragments* of the two hybrid models essentially retain the deficiencies of the empirical force field, whereas excellent descriptions are obtained for the modes localized in the respective *QM fragments*. In the MQ model, the splitting of the acceptor modes is as large as in the complete quantum mechanical treatment and their absolute frequencies are close (LDA) or very close (BLYP, data not shown) to those of the corresponding QQ treatment. In the LDA, QM model the interaction with the MM fragment induces redshifts of 40–80 cm^{-1} for the donor modes, which are much closer to the observed shifts than according to the LDA QQ model. Thus, the hybrid description repairs the LDA underestimate of the force constant belonging to the covalent O–H bond involved in the hydrogen bridge. Such repair is superfluous for the BLYP functional; here the QM hybrid description matches the QQ results *and* the experimental findings very well.[80]

d. Low-frequency relative motions. To gain further insight into the hybrid description of the intermolecular forces consider the low-frequency region of the dimer spectra shown in Fig. 11. In the QQ model, the LDA deficiencies render these forces and, consequently, also the frequencies

too large. Some of the forces calculated through the hybrid and purely mechanical models for the relative motions of the hydrogen-bonded pair are also too large, the least so for the QM model, which actually performs quite well. For the BLYP, functional comparably small overestimates are obtained both for the QQ and for the hybrid treatments.[80]

e. Discussion. The above results on the water dimer may be summarized as follows: (a) for LDA, the hybrid approach clearly outperforms the QQ model and provides an acceptable description of the experimental findings; that explains why the simple LDA hybrid simulation of a solvated water molecule was able to yield a reasonable description of the vibrational spectrum of liquid water; (b) for BLYP, the differences between the hybrid models and the QQ model are *extremely small*.

In particular, the high-frequency spectra have shown (i) that the electrostatic influence of the MM fragment on the wave function of the QM fragment is modeled with appreciable accuracy in our hybrid approach (if the partial charges in the MM fragment are properly chosen). The low-frequency spectra corroborate the hypothesis voiced further above (ii) that a choice of slightly larger van der Waals radii could further improve the hybrid descriptions of intermolecular forces, which currently exhibit small overestimates.

Finally, we would like to add that the various corrections for the unbalanced electrostatic reaction forces (see Sec. II B 5) contribute greatly to the accuracy of our hybrid approach, as shown by a series of further test calculations.[80]

B. Covalently linked QM-MM fragments

In Sec. II B 4, we introduced the SPLAM method for the treatment of covalently linked QM-MM fragments using the ethane molecule of Fig. 6 as a prototypical example, since that molecule allowed us to concentrate on the bonded interactions. Correspondingly we will first use the vibrational spectra of this molecule to check these aspects of the SPLAM force field.

1. Ethane

The vibrational spectrum of ethane is precisely known.[93] For its quantum mechanical and hybrid descriptions, we have used our standard BLYP approach and the harmonic approximation. Most of the parameters required for the MM fragments and the SPLAM force field [cf. Eqs. (10), (12), and (13)] have been determined through BLYP calculations of ethane and methane respectively. These include the force constants and equilibrium values for bond lengths and angles in H–C–H and H–C–C bonding motives. For the remaining MM parameter, we have retained the original CHARMM values.

Figure 12 compares observed band positions[93] with results of calculations, in which the QM fragment covered the whole ethane molecule (BLYP), one of the CH_3 groups (SPLAM, LAM), or was chosen empty (MM). The hybrid model labeled by LAM has been treated by the link atom approach of Field *et al.*[63] (cf. also Bakowies and Thiel[66]) using the same MM parameters as in SPLAM. We have chosen the 800–1700 cm^{-1} spectral range for comparisons since it covers a torsional mode around the central C–C bond at about 800 cm^{-1}, the stretching mode of that bond at about 1000 cm^{-1}, and the rocking and deformation modes of the methyl groups at higher frequencies. Therefore this range monitors not only the SPLAM description of the central bond but also the angular and torsional contributions collected in Eqs. (12) and (13).

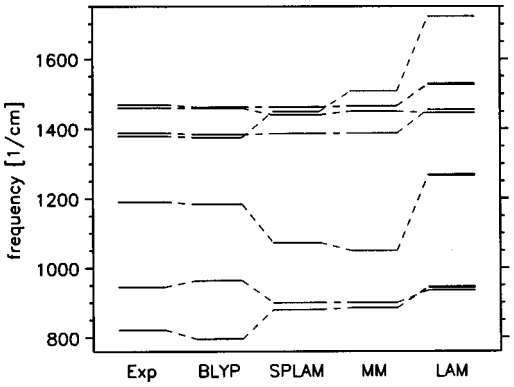

FIG. 12. Vibrational spectra of ethane.

From the first two columns of the figure, we observe the well-known fact that the harmonic BLYP force field yields a spectrum which closely agrees with the observed (anharmonic) spectrum in the given spectral range (see, for example, Ref. 15). Due to the partial optimization of force constants, the MM spectrum provides for some of the modes a satisfactory description. The sizeable deviations apparent for other modes testify to the inherent deficiencies of the MM approach. The SPLAM spectrum interpolates between the BLYP and the MM results as one would expect from a reasonable hybrid method. In contrast, the LAM spectrum depicted in the last column of the figure does not represent an interpolation. As noted earlier by Bakowies[79] vibrational spectra obtained by this conventional link atom method exhibit severe artifacts.

One can in fact explain every single detail of the hybrid spectra SPLAM and LAM by analyzing the compositions of the various modes and such analyses have been a mayor tool in the development of the details of the SPLAM method.[80] However, here we merely wish to emphasize the important result that the SPLAM description of bonded interactions transfers benefits of the more exact quantum mechanical description even to the region of the covalent bond linking the MM and QM fragments.

The question then arises as to whether the SPLAM treatment of the nonbonded interactions acting across the link region between more extended MM and QM fragments is of similar quality and whether the associated dynamical corrections actually work (cf. Secs II B 4 and II B 5).

2. A small protonated Schiff base

In order to answer the above question we have studied a larger molecule, which is charged, highly polar, and strongly polarizable. The model molecule of our choice is depicted in Fig. 13, is denoted by MPSB, and consists of a protonated Schiff base $CH{=}NH_2^+$ which is attached to the C=C double bond of a cyclohexene ring. MPSB is a small relative of the

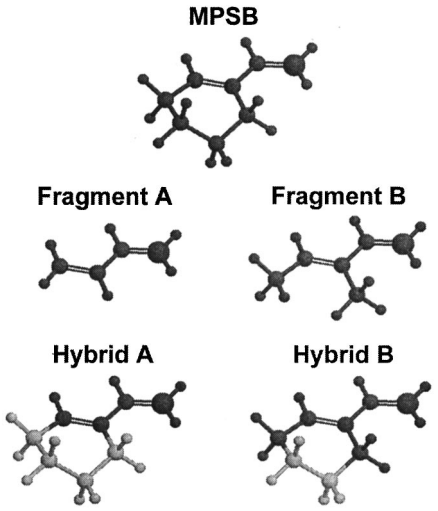

FIG. 13. Model protonated Schiff base MPSB; also shown are two different partitionings for SPLAM hybrid descriptions.

retinal chromophore of bacteriorhodopsin (cf. Fig. 2). Its positive charge delocalizes into the conjugated π-electron system, causing the strong polarizability and polarity. Therefore, the conjugated system of this molecule is highly sensitive towards the action of external electrostatic fields and is the source of a strong electrostatic field.

The QM fragments of two hybrid models **A** and **B** considered for MPSB are also depicted in Fig. 13. They both cover the conjugated π system, but QM fragment **B**, through the methyl groups at the C=C bond, can additionally cover the substituent effects of the cyclohexene ring on that system. Their respective MM parts are made up of dipolar CH_2 groups sensing and exerting electrostatic interactions with the π system. The vibrational modes localized in the C=C—C=NH_2 moiety of MPSB should therefore represent most sensitive monitors for the SPLAM treatment of the electrostatics in the hybrid models.

The C=C—C=NH_2 moiety of MPSB has three vibrational modes with frequencies in the 1500–1700 cm^{-1} spectral range. In the harmonic approximation a simple LDA treatment of MPSB (experimental data are lacking) predicts the NH_2 scissoring mode at 1511 cm^{-1}, the C=C stretch at 1600 cm^{-1}, and the C=N stretch at 1640 cm^{-1}. These results are collected in the central column of Fig. 14 and compared with the vibrational LDA spectra of the QM fragments **A** and **B** in the left and rightmost columns, respectively. Ring closures through SPLAM yield the spectra of the corresponding hybrid models which are also shown in the figure.

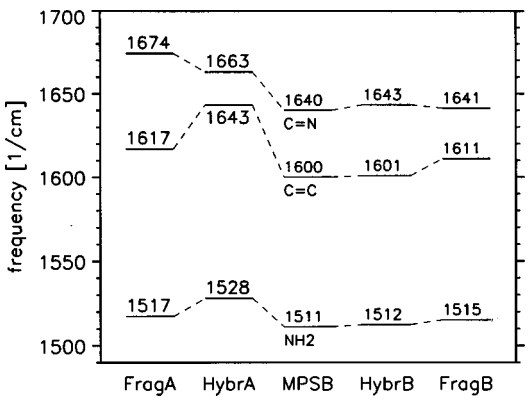

FIG. 14. Vibrational spectra of various models of MPSB.

Comparing first the frequencies obtained by LDA for MPSB and the small fragment **A**, we note that all fragment modes are sizably blueshifted. Much smaller blueshifts are observed for the larger fragment **B**, showing that the π system reacts sensitively to the methyl substituents. Solely the C=C stretching and the NH_2 scissoring modes still exhibit blueshifts of 11 and 4 cm^{-1}, respectively. A more detailed analysis of the LDA force fields has shown, for instance, that in fragment **A** the force constants of the C=C and C=N bonds are substantially larger than in MPSB and fragment **B**, for which closely matching values have been determined.[80]

As demonstrated by comparisons of the MPSB spectrum with those of hybrid **A** and fragment **A**, ring closure by SPLAM cannot account for the substituent effects missing in the LDA force field of fragment **A**. But such behavior was to be expected since SPLAM was designed for cutting CH_2—CH_2 linkages, whereas in hybrid **A** the two cuts are placed immediately adjacent to a conjugated system (cf. Fig. 13).

Hybrid **B**, however, in which the cuts have been placed at suitable positions, exhibits vibrational frequencies which are in nearly perfect agreement with those calculated through LDA for the whole molecule MPSB. Here, for instance, the increase of mass attached to the C=C bond, which results from the SPLAM ring closure in hybrid **B** as compared to fragment **B**, is the cause for the 10 cm^{-1} redshift of the C=C mode and repairs the corresponding deficiency of the fragment spectrum.

According to our experience, gathered during the development of the various details of SPLAM by monitoring MPSB hybrid spectra, slight changes of description leave sizeable traces in the spectra. We consider the close agreement of the spectra of MPSB and hybrid **B** revealed by Fig. 14 as strong evidence that particularly the electrostatic interactions across a SPLAM region between an MM fragment and a QM fragment are very well described by our approach.

IV. SUMMARY AND CONCLUSION

The test calculations presented and discussed above were designed for a close inspection of the various aspects of the QM/MM hybrid method introduced in Sec. II.

The MD simulation of a QM water molecule solvated in an aqueous MM solvent demonstrated that the algorithmic noise associated with the FAMUSAMM treatment of the solvent electrostatics is negligibly small. It also indicated that properties of a water molecule in the liquid phase, including its dipole moment and vibrational spectrum, are well described even if the simple LDA approximation is applied within the DFT description.

The subsequent discussion of the related water dimer, which focussed on the equilibrium structure and the vibrational spectrum in the harmonic approximation, has shown that complete QM and hybrid QM/MM descriptions are nearly equivalent as far as the intramolecular properties of

the QM fragment and the intermolecular forces are concerned. Regarding the latter, the hybrid model may even repair certain artifacts associated with lower-quality QM methods. Clearly, deficiencies of the empirical water model can deteriorate descriptions, but conversely, a hybrid method offering a high-quality QM method is well suited to the optimization of empirical force fields.

The computational results for the two examples summarized so far have led us to the conclusion that the hybrid method provides an essentially correct representation of the electrostatic interactions between the solvent and the solute, particularly with respect to the DFT description of the properties of the latter.

The subsequent examples addressed the question of whether the SPLAM force field for the description of covalently linked QM and MM fragments can remove those artifacts that had plagued previous link atom procedures. That this is actually the case, has been ascertained by analyzing vibrational spectra of ethane and of a small molecule related to the protonated retinal Schiff base that had been obtained for various partitionings into QM and MM fragments. In the immediate link region, the SPLAM force field turned out to interpolate between the QM and MM descriptions. For modes, which are localized within the QM fragment at a certain distance from that region the force field of a complete DFT treatment was recovered, so long as the SPLAM approach was restricted to covalent bonds between CH_2 groups.

Having now established a computationally efficient hybrid approach that combines *ab initio* DFT for a molecular site of interest with empirical MM for a large environment, we are curious to learn how it will perform in large-scale applications like, such as the MD simulation of enzymatic catalysis or the *in situ* calculation of vibrational spectra of protein chromophores.

ACKNOWLEDGMENTS

M. Eichinger and P. Tavan are grateful for a grant from the Deutsche Forschungsgemeinschaft (SFB-533/C3); use of the computer facilities at the Hochleistungs-Rechenzentrum Jülich (HLRZ) and of the Leibniz-Rechenzentrum München (LRZ) is also acknowledged.

[1] D. Oesterhelt and W. Stoeckenius, Nature New Biology **233**, 149 (1971).
[2] K. C. Hasson, F. Gai, and P. A. Afinrud, Proc. Natl. Acad. Sci. USA **93**, 15124 (1996).
[3] M. F. Großjean and P. Tavan, J. Chem. Phys. **88**, 4884 (1988).
[4] M. F. Großjean, P. Tavan, and K. Schulten, J. Phys. Chem. **94**, 8059 (1990).
[5] M. J. Frisch *et al.*, GAUSSIAN 94, Revision B.3, Gaussian, Inc., Pittsburgh PA, 1995.
[6] J. J. P. Stewart, Quantum Chemistry Program Exchange Bulletin **6**, 1986.
[7] K. B. Lipkowitz and D. B. Boyd, *Reviews in Computational Chemistry, Vol. IV* (VCH Publishers, Inc., New York, 1993).
[8] T. P. Hamilton and P. Pulay, J. Phys. Chem. **93**, 2341 (1989).
[9] S. Hirata, H. Yshida, H. Torii, and M. Tasumi, J. Chem. Phys. **103**, 8955 (1995).
[10] X. Zhou, S. J. Mole, and R. Liu, Spectroscopy **12**, 73 (1996).
[11] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[12] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[13] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[14] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[15] M. Nonella and P. Tavan, Chem. Phys. **199**, 19 (1995).
[16] P. Tavan, in *Spectroscopy of Biological Molecules, 6th European Conference on the Spectroscopy of Biological Molecules, 3–8 September, 1995, Villeneuve d'Ascq, France*, edited by J. C. Merlin, S. Turrell, and J. P. Huvenne (Kluwer Academic, Dordrecht, 1995), pp. 3–6.
[17] R. G. Parr and W. Yang, *Density-functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
[18] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[19] C. Lee *et al.*, Phys. Rev. B **47**, 4863 (1993).
[20] M. Tuckerman, K. Laasonen, M. Sprik, and M. Parrinello, J. Chem. Phys. **103**, 150 (1995).
[21] M. Sprik, J. Hutter, and M. Parrinello, J. Chem. Phys. **105**, 1142 (1996).
[22] J. Hutter *et al.*, *CPMD Version 3.0*, MPI für Festkörperforschung and IBM Zurich Research Laboratory, 1995–96.
[23] A. Rahman, Phys. Rev. **136**, A405 (1964).
[24] L. Verlet, Phys. Rev. **159**, 98 (1967).
[25] M. P. Allen and D. Tildesley, *Computer Simulations of Liquids* (Clarendon, Oxford, 1987).
[26] M. Karplus and A. McCammon, Annu. Rev. Biochem. **53**, 263 (1983).
[27] W. F. van Gunsteren and H. J. C. Berendsen, Angew. Chem. Int. Ed. Engl. **29**, 992 (1990).
[28] H. J. C. Berendsen, Science **271**, 954 (1996).
[29] K. M. Merz, Curr. Opin. Struct. Biol. **7**, 511 (1997).
[30] B. R. Brooks *et al.*, J. Comput. Chem. **4**, 187 (1983).
[31] S. J. Weiner *et al.*, J. Am. Chem. Soc. **106**, 765 (1984).
[32] A. Brünger, *X-PLOR Manual*, The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, 1992.
[33] W. van Gunsteren *et al.*, *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, Vdf Hochschulverlag AG an der ETH Zürich, Universitätstrasse 6, CH-8092 Zürich, 1996.
[34] J. K. Gregory *et al.*, Science **275**, 814 (1997).
[35] S. Kuwajima and A. Warshel, J. Phys. Chem. **94**, 460 (1990).
[36] L. X. Dang and T.-M. Chang, J. Chem. Phys. **106**, 8149 (1997).
[37] C. L. Brooks III, B. M. Pettitt, and M. Karplus, J. Chem. Phys. **83**, 5897 (1985).
[38] R. J. Loncharich and B. R. Brooks, Proteins **6**, 32 (1989).
[39] S. E. Feller, R. M. Venable, and R. W. Pastor, Langmuir **13**, 6555 (1997).
[40] D. J. Tobias, K. Tu, and M. L. Klein, Curr. Opin. Colloid Interface Sci. **2**, 15 (1997).
[41] C. Niedermeier and P. Tavan, J. Chem. Phys. **101**, 734 (1994).
[42] C. Niedermeier and P. Tavan, Mol. Simul. **17**, 57 (1996).
[43] L. Greengard and V. Rokhlin, Chem. Scr. **29A**, 139 (1989).
[44] J. A. Board, Jr. *et al.*, Chem. Phys. Lett. **198**, 89 (1992).
[45] W. B. Streett, D. J. Tildesley, and G. Saville, Mol. Phys. **35**, 639 (1978).
[46] H. Grubmüller, H. Hleer, A. Windemuth, and K. Schulten, Mol. Simul. **6**, 121 (1991).
[47] M. E. Tuckerman, B. J. Berne, and G. J. Martyna, J. Chem. Phys. **94**, 6811 (1991).
[48] R. Zhou and B. J. Berne, J. Phys. Chem. **103**, 9444 (1995).
[49] A. Windemuth, *Parallel Computing in Computational Chemistry* (ACS Books, Washington DC, 1995).
[50] M. Eichinger, H. Grubmüller, H. Heller, and P. Tavan, J. Comput. Chem. **18**, 1729 (1997).
[51] B. A. Luty, M. E. Davis, I. G. Tironi, and W. F. van Gunsteren, Mol. Simul. **14**, 11 (1994).
[52] U. Essmann *et al.*, J. Chem. Phys. **103**, 8577 (1995).
[53] A. Y. Toukmaji and J. J. A. Board, Comput. Phys. Commun. **95**, 73 (1996).
[54] E. L. Pollock and J. Glosli, Comput. Phys. Commun. **95**, 93 (1996).
[55] P. Procacci, T. Darden, and M. Marchi, J. Phys. Chem. **100**, 10464 (1996).
[56] H. Grubmüller, B. Heymann, and P. Tavan, Science **271**, 997 (1996).
[57] E.-L. Florin, V. T. Moy, and H. E. Gaub, Science **264**, 415 (1994).
[58] M. Eichinger, H. Grubmüller, and H. Heller, *User Manual for EGO_VIII, Release 2.0*, Theoretische Biophysik, Institut für Medizinische Optik, Ludwig-Maximilians-Universität, Oettingenstr. 67, D-80585 München, Germany, 1995, electronic access: http://www.imo.physik.uni-muenchen.de/ego.html.
[59] A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227 (1976).
[60] M. A. Thompson and G. K. Schenter, J. Phys. Chem. **99**, 6374 (1995).
[61] D. Bakowies and W. Thiel, J. Comput. Chem. **17**, 87 (1996).
[62] P. A. Bash, M. J. Field, and M. Karplus, J. Am. Chem. Soc. **109**, 8092 (1987).

[63] M. J. Field, P. A. Bash, and M. Karplus, J. Comput. Chem. **11**, 700 (1990).
[64] J. Gao, J. Phys. Chem. **96**, 437 (1992).
[65] M. A. Thompson, J. Phys. Chem. **99**, 4794 (1995).
[66] D. Bakowies and W. Thiel, J. Phys. Chem. **100**, 10580 (1996).
[67] I. B. Bersuker, M. K. Leong, J. E. Boggs, and R. S. Pearlman, Int. J. Quantum Chem. **63**, 1052 (1997).
[68] P. L. Cummins and J. E. Gready, J. Comput. Chem. **18**, 1496 (1997).
[69] G. Alagona, p. Desmeules, C. Ghia, and P. A. Kollman, J. Am. Chem. Soc. **106**, 3623 (1984).
[70] U. C. Singh and P. A. Kollman, J. Comput. Chem. **7**, 718 (1986).
[71] F. K. Brown *et al.*, J. Org. Chem. **57**, 4862 (1992).
[72] M. Freindorf and J. Gao, J. Comput. Chem. **17**, 386 (1996).
[73] R. V. Stanton, D. S. Hartsough, and K. M. Merz, J. Phys. Chem. **97**, 11868 (1993).
[74] T. Matsubara, F. Maseras, N. Koga, and K. Morokuma, J. Phys. Chem. **100**, 2573 (1996).
[75] M. Svenssons *et al.*, J. Phys. Chem. **100**, 19357 (1996).
[76] I. Tuñón *et al.*, J. Comput. Chem. **17**, 19 (1996).
[77] I. Tuñón, M. T. C. Martins-Costa, C. Millot, and M. F. Ruiz-López, J. Chem. Phys. **106**, 3633 (1997).
[78] M. Strnad *et al.*, J. Chem. Phys. **106**, 3643 (1997).
[79] D. Bakowies, Ph.D. thesis, Universität Zürich, 1994.
[80] M. Eichinger, Ph.D. thesis, Ludwig-Maximilians Universität München, Germany, 1999 (in preparation).
[81] U. C. Singh and P. A. Kollman, J. Comput. Chem. **5**, 129 (1984).
[82] H. Grubmüller and P. Tavan, J. Comput. Chem. **19**, 1534 (1998).
[83] C. Alhambra, K. Byun, and J. Gao, in *Combined QM and MM methods*, edited by J. Gao and M. A. Thompson (ACS Books, Washington DC, 1998).
[84] K. Laasonen, M. Sprik, M. Parrinello, and R. Car, J. Chem. Phys. **99**, 9080 (1993).
[85] W. L. Jorgensen *et al.*, J. Chem. Phys. **79**, 926 (1983).
[86] C. L. Brooks III and M. Karplus, J. Chem. Phys. **79**, 6312 (1983).
[87] W. F. van Gunsteren and H. J. C. Berendsen, Mol. Phys. **34**, 1311 (1977).
[88] H. J. C. Berendsen *et al.*, J. Chem. Phys. **81**, 3684 (1984).
[89] D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).
[90] S. A. Clough, Y. Beers, G. P. Klein, and L. S. Rothman, J. Chem. Phys. **59**, 2254 (1973).
[91] C. A. Coulson and D. Eisenberg, Proc. R. Soc. London, Ser. A **291**, 445 (1966).
[92] W. E. Thiessen and A. H. Narten, J. Chem. Phys. **77**, 2656 (1982).
[93] G. Herzberg, *Infrared and Raman Spectra of Polyatomic Molecules* (Van Nostrand Co., Inc., Princeton, NJ, 1966).
[94] T. A. Ford and M. Falk, Can. J. Chem. **46**, 3579 (1968).
[95] S. H. Chen *et al.*, Phys. Rev. Lett. **53**, 1360 (1984).
[96] H. Goldstein, *Classical Mechanics* (Addison–Wesley, Menlo Park, CA, 1970).
[97] T. R. Dyke, K. M. Mack, and J. S. Muenter, J. Chem. Phys. **66**, 498 (1977).