# Generalized QM/MM Force Matching Approach Applied to the 11-cis Protonated Schiff Base Chromophore of Rhodopsin

Manuel Doemer, Patrick Maurer, Pablo Campomanes, Ivano Tavernelli, and Ursula Rothlisberger*

Laboratoire de Chimie et Biochimie Computationelle, Ecole Polytechnique Fédérale de Lausanne, Lausanne, CH-1025, Switzerland

S *Supporting Information*

ABSTRACT: We extended a previously developed force matching approach to systems with covalent QM/MM boundaries and describe its user-friendly implementation in the publicly available software package CPMD. We applied this approach to the challenging case of the retinal protonated Schiff base in dark state bovine rhodopsin. We were able to develop a highly accurate force field that is able to capture subtle structural changes within the chromophore that have a pronounced influence on the optical properties. The optical absorption spectrum calculated from configurations extracted from a MD trajectory using the new force field is in excellent agreement with QM/MM and experimental references.
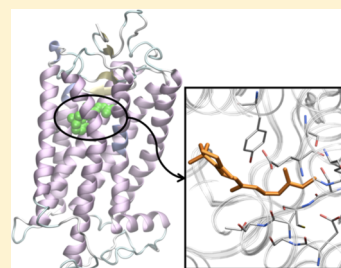
1. INTRODUCTION

Classical Molecular Dynamics (MD) simulations based on Force Fields (FFs) represent an important method in the investigation of large scale molecular systems of ~100,000 atoms and more in biology and materials science applications.[1−3] Typically, their functional form does not allow the explicit description of major rearrangements of the electronic structure, i.e. the bonding topology is retained over the entire course of a simulation. However, the energy and nuclear force evaluation is computationally so expedient that FFs can nowadays provide the underlying Potential Energy Surface (PES) for MD simulations up to the microsecond time scale. This exceeds by far the accessible time windows of first-principles MD based on electronic structure methods, such as Density Functional Theory (DFT). FFs therefore represent an indispensable tool for the computational modeling of e.g. large scale conformational motions of biomolecules at an atomistic resolution.

However, the functional form of classical FFs and their parametrization is largely empirical, and the determination of accurate and transferable parameter sets involves significant human effort. With a plethora of existing force fields and associated parameter sets a large variety of biological systems can be studied. For systems that contain molecules for which no parameters are available it can be, however, a difficult and time-consuming task to determine a reliable force field. Typical approaches employ electronic structure calculations of a suitable small model compound in the gas phase or in a continuum solvent to determine the missing parameters.[4] Such procedures, however, bear the risk, that the newly determined parameters are not necessarily transferable to the actual system under investigation in the condensed phase. Alternatively, in situ parametrization methods, such as the ″learn-on-the-fly″ approach[5] can be used to determine parameters from higher level reference calculations on the system at hand also in the condensed phase. The potential constructed in this way can be used for the long-time propagation of a specific system, while its transferability is expected to be limited. A similar approach is the force-matching technique[6] to parametrize classical or semiempirical potentials to reproduce the forces calculated by *ab inito* or DFT methods. The method has originally been applied to optimize a glue potential for aluminum and, later on, an embedded atom method (EAM) potential for magnesium.[7] This parametrization method is very appealing and can be used with a manifold of different classical and semiempirical potential energy functions.[8−11] For example, interaction potentials, based on a (nonself-consistent) tight-binding model[12] and a modified embedded atom method,[13] were fitted for silicon from, among other properties in the target function, DFT forces. In ref 14 a classical optimal potential (OP) was constructed to study the bulk properties of iron at earth's core conditions. In ref 15 the force-matching methodology was modified to directly parametrize a spline interpolation of interatomic forces. The method takes advantage of the linear dependence of the target function on the fitting parameters. Therefore, instead of directly minimizing an objective function the problem could be recast into solving an overdetermined system of linear equations. This modified parametrization scheme can be seen as more reliable and tractable for large numbers of parameters and was applied to studies of liquid water and hydrogen fluoride.[16] Recently, the force matching protocol has also been applied to parametrize a reactive FF.[17]

Our group has developed a force matching protocol for biomolecular FFs based on mixed Quantum Mechanics/ Molecular Mechanics (QM/MM) reference calculations.[18] A similar scheme was proposed in an adaptive formulation later on.[19] In our approach, the QM region is chosen in order to include all components of the system for which no parameters are available. Finite-temperature QM/MM MD simulations are

performed to generate a set of reference configurations. Note however, that such reference configurations could also be obtained from other sources. Next, nuclear forces on the QM atoms and electrostatic properties are extracted from this trajectory to serve as target properties for the subsequent parameter fitting scheme. The force field parameters are then determined in such a way as to optimally reproduce the electrostatic properties and the nuclear forces of the QM subsystem. The influence of the environment as well as finite temperature and pressure effects are taken into account automatically. The optimized FF parameters can then be used to perform simulations of a given system with essentially the accuracy of a QM/MM treatment at the computational cost of classical MD. This allows sampling times far beyond the limits of the QM/MM method and the calculation of properties with long correlation times, such as thermodynamic averages involving slow motions of large biomolecules. This QM/MM force matching method has been applied to various systems involving aqueous solutions of dihydrogenphosphate, a glycyl-alanine dipeptide, a nitrosyl-dicarbonyl complex of technetium-(I), an azole-bridged diplatinum anticancer drug interacting with DNA[20] and organometallic ruthenium complexes.[21] Note however, that these systems do not contain QM/MM boundaries across chemical bonds. The QM/MM interactions therefore involve only nonbonded contributions.

In this work we extended the previously developed force matching approach for systems with noncovalent QM/MM boundaries[18] to situations with covalent QM/MM boundaries using pseudopotential link atoms[22] and describe its user-friendly implementation in the publicly available software package CPMD v. 3.15.[23] The optimized FF parameters from the QM/MM force matching can be employed afterward in conjunction with any software package for classical simulations that offers the employed functional form.

We applied this protocol to the challenging case of deriving *in situ* FF parameters for the retinal protonated Schiff base (RPSB) of bovine rhodopsin in the dark state embedded in a lipid bilayer, as shown in Figure 1. Rhodopsin is a biological pigment in the photoreceptor cells of the retina and constitutes the first member in a signaling cascade responsible for the
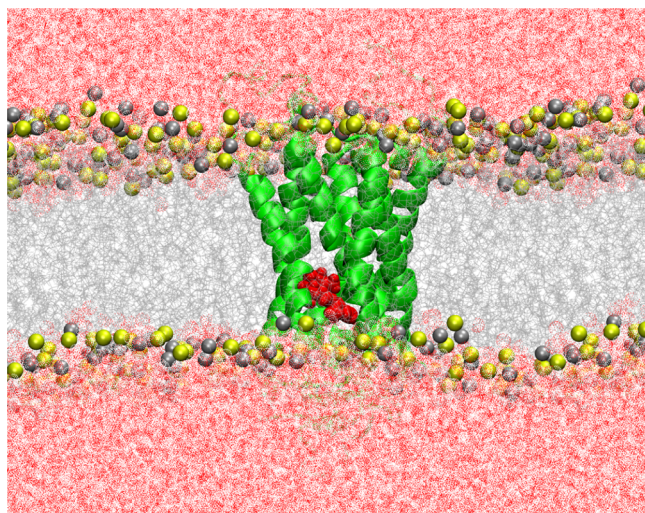

Figure 1. Rhodopsin (green) embedded in a lipid bilayer (light gray with yellow/dark gray phospholipid head groups) and water solvent (red/white). RPSB in red.

perception of light.[24,25] The initial event upon light absorption is the cis–trans photoisomerization of the retinal chromophore within the active site of the protein[26–29] (red in Figure 1). A more detailed representation of the retinal moiety is provided in Figure 2 (C_1 to C_{15}). It is covalently bound to the side chain
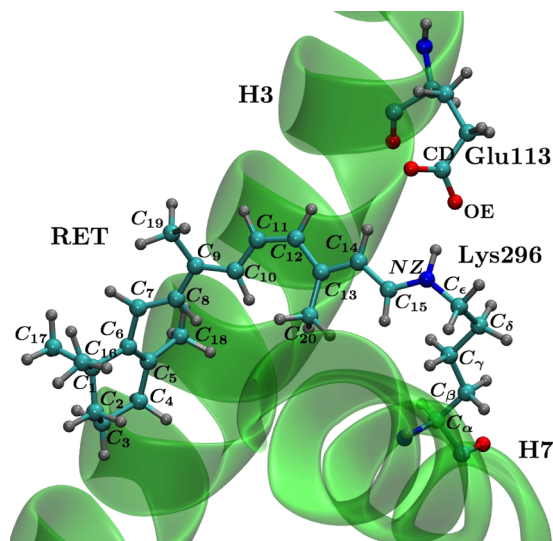

Figure 2. 11-cis retinal (RET) covalently bound by a protonated Schiff base (PSB) linkage to Lys296 (in helix H7) in ball and sticks representation. The counterion Glu113 (in helix H3) is also shown.

of Lys296 (C_β to NZ) via a protonated Schiff base. Another important residue is Glu113, which acts as a counterion for the positive charge of the protonated group. The light-induced isomerization occurs around the C_{11}=C_{12} double bond. The investigation of the structural changes after the absorption of light has been an active field of research both on the experimental[30–33] and computational sides.[34–37] Biomolecular force fields have been employed to illuminate equilibrium properties of dark state rhodopsin[38] and the large scale structural rearrangements of the protein after light absorption.[39,40] However, the parameter sets currently used for the retinal chromophore[35,41,42] do not account for the bond length alternation (*BLA*) among carbon–carbon single and double bonds in the conjugated π-system. This is a reasonable approximation if the focus is on the global structural properties of the protein. However, recent investigations have shown that optical properties calculated from configurations generated by such an approximate bonding topology do not agree well with experiments.[43] Currently, one has to rely on QM/MM methods in order to generate realistic structural models for a calculation of optical properties.[44] In order to overcome the time scale limitations associated with the QM/MM approach we apply the newly implemented QM/MM force matching protocol to derive a consistent set of FF parameters that reproduce the structural and dynamical properties at the QM/MM level. In particular, the new parameter set correctly captures thermal variations of the *BLA* in dark state RPSB. These subtle changes in the chromophore structure have a pronounced influence on the optical properties. The optical absorption spectrum calculated from configurations extracted from an MD trajectory using the new FF is in excellent agreement with the QM/MM based results and experimental references, while the original FF produces configurations that lead to a substantial red-shift in the calculated absorption spectrum.

2. METHODS

2.1. QM/MM Force Matching. In our force matching scheme QM/MM reference calculations on the system of interest are performed in order to derive FF parameters for all the atoms comprised in the QM fragment. The procedure is general for any functional form of the classical FF, but in this work we considered the Amber[41] form with

$$E^{nb} = \sum_{\alpha < \beta} \left[\frac{A_{\alpha\beta}}{R_{\alpha\beta}^{12}} - \frac{B_{\alpha\beta}}{R_{\alpha\beta}^{6}} + \frac{q_\alpha q_\beta}{\varepsilon R_{\alpha\beta}} \right] \tag{1}$$

the nonbonded interactions between atom pairs $\alpha-\beta$ involve van der Waals interactions, modeled by a Lennard-Jones potential with corresponding parameters $A_{\alpha\beta}$ and $B_{\alpha\beta}$, and the Coulomb potential for atomic point charges q_α in a medium with dielectric constant ε. The bonded interactions are calculated according to

$$E^b = \sum_{n=1}^{N^{\text{bonds}}} K_{b_n}(b_n - b_n^{\text{eq}})^2 + \sum_{n=1}^{N^{\text{angles}}} K_{\theta_n}(\theta_n - \theta_n^{\text{eq}})^2 \\ + \sum_{n=1}^{N^{\text{dihedrals}}} \frac{V_n}{2}[1 + \cos(m_n\phi_n - \varphi_n)] \tag{2}$$

where the first term represents the bond-stretching potential with force constant K_{b_n} and equilibrium value b_n^{eq}. The second term describes the angle bending potential with equilibrium value θ_n^{eq} and force constant K_{θ_n}. The third term represents the periodic dihedral potential with barrier height V_n, the phase shift φ_n, and the multiplicity m_n. Improper dihedral terms of the same form can be used to correct for the out-of-plane motions in ring structures.

Since in our force matching scheme we optimize only the FF parameters for the atoms in the QM region, we rely, as in any QM/MM scheme, on the assumption that the original FF parameters provide a good representation for the rest of the system, i.e. the MM region. The procedure involves four steps: 1) A set of L reference configurations is generated. 2) The nuclear forces on all the atoms of the QM subsystem as well as the electrostatic potential and field in the surrounding of the electronic charge density are stored. 3) An optimal set of atomic point charges $\{q_\alpha\}$ that reproduces the electrostatic potential and field in the surrounding of the QM region over all the reference configuration is calculated. Step 4) involves the actual force matching. The nonbonded contributions, computed with the charges obtained in the third step and a given set of Lennard-Jones parameters from the MM force field, are subtracted from the total reference forces on the QM atoms. The remaining forces are assumed to originate from bonded interactions. The parameters for bonded interactions (torsions, bending, and bonds) are subsequently optimized in order to reproduce the residual part of the reference forces.

The QM/MM force matching results in a new parameter set for the corresponding fully classical system. The subsequent simulations with the modified FF can then be performed in any software package with a classical MD implementation.

We have implemented our QM/MM force matching protocol, as described in the rest of this section, in the CPMD package starting with the release 3.15[23] interfaced with GROMOS96.[45] This procedure is general and can be combined with other QM/MM implementations. A description of the CPMD specific details for the implementation can be found in the corresponding manual and downloaded from our Web page.[46]

We extract the reference configurations from a finite temperature QM/MM Car−Parrinello MD trajectory at the same level of theory than the one that is used to compute the reference forces. At the typical time scale of DFT based QM/MM simulations, i.e. 10 to 100 ps, it is unlikely to observe dihedral transitions in the QM subsystem. Therefore, the procedure described in this work provides a parameter set that can be safely used only for a specific conformation of the system. However, it is possible to generate a force field that also reproduces torsional barriers with the same procedure if the QM/MM dynamics are performed under the action of a bias potential[47] or other enhanced sampling methods[48−53] that induce transitions within the available computational time. Alternatively, different configurations could be extracted from a classical trajectory or from a guided sampling along a given collective coordinate.

The target property to optimize the parameters of the classical FF are the nuclear forces on the QM atoms computed at the reference configurations. We compute references forces with the electronic structure quenched to the Born−Oppenheimer (BO) surface with tight convergence criteria (10^{-7} au for the orbital gradients). If the QM/MM reference trajectory is performed based on BO MD, the reference forces can be stored already during the reference MD run and no extra calculations have to be performed.

At difference to the conventional ESP[54,55] and RESP[4] procedures the partial atomic charges are fitted directly from the QM/MM electrostatic potential and field according to a modified D-RESP scheme[56,57] that takes also thermal effects into account. In this extended approach,[18] the effective point charges located on the QM atoms are determined in such a way as to optimally reproduce the electrostatic potential and field probed at the nuclear positions of a set of NN atoms surrounding the QM region. The NN atoms are defined by a set of MM atoms within a given cutoff distance from the QM atoms, that are explicitly coupled to the QM charge density. All reference configurations are included in the fitting. Furthermore, the effective point charges are harmonically restrained to their respective Hirshfeld values.[58] These additional restraints are necessary in order to exclude chemically unreasonable and strongly conformation dependent solutions to the over-determined problem.[4,56] Thus, the set of effective atomic point charges $\{q_\alpha\}$ is optimized by minimizing the penalty function:

$$\chi^2(\{q_\alpha\}) = \sum_{l=1}^{L} \Big[\sum_{\beta \in NN_l} (w^V(V_{\beta l}^{\text{MM}} - V_{\beta l}^{\rho})^2 + w^E \\ |\mathbf{E}_{\beta l}^{\text{MM}} - \mathbf{E}_{\beta l}^{\rho}|^2) \\ + \sum_{\alpha \in \text{QM}} w^H(q_\alpha - q_{\alpha l}^H)^2 \Big] + w^Q \\ (Q^{\text{tot}} - \sum_{\alpha \in \text{QM}} q_\alpha)^2 \tag{3}$$

The index l runs over all L reference configurations; β labels the classical atoms NN_l that are explicitly coupled to the quantum charge density in configuration l. α refers to atoms of the QM subsystem. The weighting factors w^V, w^E, w^H, and w^Q can be specifically tuned to each system in order to balance the individual contributions, as described in Section 3.1. $V_{\beta l}^{\rho}$ and $\mathbf{E}_{\beta l}^{\rho}$

414

dx.doi.org/10.1021/ct400697n | J. Chem. Theory Comput. 2014, 10, 412−422

are the electrostatic potential and field, respectively, on the classical atom β in configuration l due to the QM charge density, while $V_{\beta l}^{MM}$ and $\mathbf{E}_{\beta l}^{MM}$ are the potential and field due to the classical point charges $\{q_\alpha\}$ that are being optimized. The third term in eq 3 restrains the charges $\{q_\alpha\}$ to their respective Hirshfeld values $q_{\alpha l}^H$ and the last term assures that the total charge of the set $\{q_\alpha\}$ matches the reference total charge of the QM electron density Q^{tot}. It is straightforward to ensure that chemically equivalent atoms obtain the same charge values. Minimizing the penalty function χ^2 can be recast in the form of a least-squares problem and solved as an overdetermined system of linear equations in $\{q_\alpha\}$.

At the current state of the implementation the van der Waals parameters $A_{\alpha\beta}$ and $B_{\alpha\beta}$ are not optimized but are kept fixed to the original force field value. This choice is consistent with the QM/MM interaction Hamiltonian we use that retains the Lennard-Jones parameters from the classical force field.[58] Future developments will include a parametrization protocol for the van der Waals parameters based on reference values obtained with the recently developed Dispersion Corrected Atom-Centered Potentials method.[59−63]

Due to their different magnitudes, it is advantageous to determine the parameters for the bonded interactions in a separate step from the ones for nonbonded interactions. Once a new set of charges $\{q_\alpha\}$ has been determined, we therefore compute the classical nonbonded forces $\mathbf{F}_{l\alpha}^{MM_{nb}}$ (for the QM atom α in configuration l) due to the electrostatics employing the new charges and combine them with the van der Waals interactions according to eq 1 from the MM force field. In order to optimize the parameters in the bonded interactions we subtract the nonbonded forces from the QM/MM reference forces and minimize the remaining penalty function under variations of the bonded parameters $\{\tau_{b_n}\} = \{\{K_{b_n}\}, \{b_n^{eq}\}, \{K_{\theta_n}\}, \{\theta_n^{eq}\}, \{V_n\}\}$ (eq 2)

$$\sigma^2(\{\tau_{b_n}\}) = \sum_{l=1}^{L} \sum_{\alpha \in QM} \|\mathbf{F}_{l\alpha}^{MM_b} - (\mathbf{F}_{l\alpha}^{QM} - \mathbf{F}_{l\alpha}^{MM_{nb}})\|^2 \qquad (4)$$

For a numerical solution this minimization can be formulated as a nonlinear least-squares problem. Since the derivatives of $\sigma^2(\{\tau_{b_n}\})$ with respect to the individual parameters τ_{b_n} are readily available, gradient-based algorithms can be employed.[64,65]

As in any QM/MM scheme, special care has to be taken if the boundary between the QM and the MM parts cuts a chemical bond. In order to avoid strong perturbations of the electronic structure the valence of the QM region has to be saturated. This can be achieved by introducing a QM capping atom, typically a hydrogen atom. The capping atom is an artificial construction, and its unphysical interaction with the MM environment is typically removed via exclusion rules. However, in the QM/MM force matching scheme further complications arise from the unphysical nuclear forces on the capping atom, since it is included in the QM region and therefore contributes to the QM/MM reference forces. We therefore suggest to employ the alternative scheme implemented in CPMD, in which the MM boundary atom is incorporated into the QM region and replaced by a specially parametrized[22] monovalent pseudopotential. In case reference configurations were already generated using a capping atom, the monovalent pseudopotential can simply be introduced at the point when the reference forces are recalculated; however,

the capping atom would need to be removed at this point from the reference trajectory. This scheme has been applied successfully in QM/MM simulations over the past ten years. Note, however, that the electron density in the surrounding of the monovalent pseudopotential is slightly perturbed, and it is therefore expected that in our QM/MM force matching scheme the fitting of the charges leads to unphysical solutions in this region. We found it best practice to fit, in a first step, atomic point charges for all atoms included in the QM region as described above. The charges on the monovalent pseudopotential and its nearest QM neighbors adopt unphysical values, but they account for the perturbations of the electronic structure at the QM/MM boundary. The charges further than ∼4 Å away from the boundary represent well the electrostatic properties of the surrounding of the QM region. In a second step, i.e. before fitting the bonded interactions, the charges on the monovalent pseudopotential and its nearest QM neighbors are replaced by their original FF values, while preserving the total charge Q^{tot}. The final atomic point charges in this region are therefore not optimized but retained from the original FF, which should provide a more physically sound electrostatic representation here. Furthermore, it might be necessary to constrain the bond between the monovalent pseudopotential and the next QM atom during a QM/MM MD. A typical choice for the bond length is the equilibrium value in the original FF. After fitting the bonded interactions by force matching we recommend in such a case to retain all parameters involving one of these atoms at the original FF values.

This procedure to handle the parameters involved in interactions close to the QM/MM boundary works well if the QM region is chosen large enough and if the original classical point charges and bonded parameters in the vicinity of the monovalent pseudopotential are reasonably good. In the case of optimizing parameters for the RPSB moiety in rhodopsin, for example, we deliberately extended the QM region up to C_β of the LYS 296 side chain (see Figure 2). Naturally, in case the QM/MM boundary does not involve any bonded interactions, e.g. a molecule in the QM region with the surrounding solvent treated at the classical level, no such complications arise.[18]

2.2. Computational Details. As a test application for the newly implemented force matching option in CPMD, we applied this approach to develop an improved FF for the RPSB in rhodopsin.

Simulations were based on the crystal structure of dark-state rhodopsin (PDB code: 1U19).[66] All potentially charged amino acids were considered to be in their default protonation states at physiological pH (i.e., charged), except Asp83 and Glu122 that were assumed to be neutral in line with FTIR experiments.[67] Histidine residues were protonated either at the Nδ position (His100, His211) or the Nε position (His65, His152, His195, His278). The rhodopsin protein was embedded in an explicit membrane environment modeled by 1-palmitoyl-2-oleoyl-phosphatidylethanolamine (POPE) lipids, solvated with explicit water molecules (covering both the cytoplasmic and extracellular sides), and neutralized with Na+ counterions. The final size of the simulation box was approximately 96 × 99 × 125 Å³, containing about 25000 water molecules and 300 lipids, resulting in a total number of ∼118000 atoms. The all-atom AMBER/parm99SB force field[41] was used to model standard protein residues and counterions, whereas the TIP3P model[68] was employed for water molecules. The force field parameters for the palmitic acid residues bound to Cys322 and Cys323, for the RPSB, and for POPE lipids were

taken from previous studies.[34,69] Electrostatic interactions were taken into account using the Particle Mesh Ewald algorithm[70] with a real space cutoff of 10 Å. The same cutoff was employed for the treatment of the van der Waals interactions. Bonds involving hydrogen atoms were constrained using the SHAKE algorithm.[71] An integration time step of 2 fs was used. Constant temperature (310 K) and pressure (1 atm) were achieved by coupling the systems to a Langevin thermostat and a Nosé-Hoover Langevin barostat, respectively.[72,73] After insertion of the protein in a pre-equilibrated lipid bilayer, the system was minimized using the conjugate gradient algorithm and then heated up to 310 K in 500 ps while keeping positional restraints on the protein backbone atoms. A run of 3.2 ns in the NPT ensemble, slowly removing the restraints, was initially carried out. A subsequent simulation in the canonical ensemble was performed for another 120 ns. Data for analyses were collected over the last 40 ns. All the classical MD simulations were carried out using the NAMD package.[74]

The QM/MM MD simulations and force matching were performed within the CPMD package release 3.15[23] interfaced with GROMOS96.[45] A 30 ps QM/MM Car−Parrinello MD simulation was started using one of the equilibrated configurations from the aforementioned classical run as initial structure. The electronic structure of the QM subsystem, formed by the RPSB and the Lys296 side chain up to the C_γ atom, was described using the BLYP exchange-correlation functional,[75−77] while the remaining atoms, belonging to the MM part, were considered at the classical level using the force field parameters mentioned above. Valence electrons were described using a plane-wave expansion up to a kinetic energy cutoff of 80 Ry, and soft norm-conserving Martins-Troullier pseudopotentials[78] were employed to represent the interactions between the valence electrons and the ionic cores. A monovalent pseudopotential was included at the position of the C_β atom to saturate the electronic density of the QM region at the C_γ atom. The C_β−C_γ bond length was constrained to a value of 1.54 Å to preserve the proper electronic structure at the center of the QM subsystem. Long-range electrostatic effects between MM atoms were described using the P3M method[79] with a real space cutoff of 10 Å. Electrostatic interactions between QM and MM regions were taken into account by means of a fully Hamiltonian hierarchical coupling scheme,[56−58] whereas bonded and van der Waals interactions between both subsystems were described at the force field level. The inherent periodicity in the plane-wave calculations was circumvented solving Poisson's equation for nonperiodic boundary conditions,[80] while periodic boundary conditions were retained for the classical simulation box. The propagation of the equations of motion was performed within the Car−Parrinello scheme[81] using a fictitious electron mass of 500 au and a time step of 0.1 fs. Simulations were carried out in the canonical (NVT) ensemble using a Nosé-Hoover chain of thermostats[82] to maintain the temperature at 310 K. 1000 equally spaced snapshots were taken from the last 15 ps of trajectory to reparametrize the force field used for the RPSB via the force-matching protocol. On the extracted configurations, the electronic wave function was quenched to the Born−Oppenheimer surface using a convergence threshold of 10^{-7} au, from which the nuclear reference forces $\mathbf{F}_{l\alpha}^{QM}$ in eq 4 and the electrostatic potential and field on the NN atoms were computed.

To test the quality of the structures generated by the newly developed FF, vertical excitation energies were calculated for

these configurations using the ZINDO/S semiempirical method[83] implemented in Gaussian09.[84] This method has been calibrated on a large set of compounds, and its parameters have been optimized to give accurate excitation energies for the calculation of absorption spectra in the visible range. In particular, previous studies have shown that the application of the ZINDO/S methodology to the computation of vertical excitation energies and oscillator strengths on retinal and related polyenals provides results in very good agreement with experiments.[44,85]

All molecular images were generated with Visual Molecular Dynamics (VMD).[86]

3. RESULTS AND DISCUSSION

We applied the QM/MM force matching to derive parameters for the RPSB in rhodopsin. This represents a particularly challenging system, since the QM/MM boundary cuts a chemical bond.

We discuss first the fitting of the atomic charges and how the new parameters compare to the original force field. The complete list of the newly derived parameters can be found in the Supporting Information (SI). We then assess the performance of the FF to reproduce structural properties of the retinal binding pocket and the absorption spectrum from the QM/MM reference trajectory and experiment.

3.1. FIT OF ATOMIC POINT CHARGES

We derived individual atomic charges for the atoms in the QM region that best reproduce the electrostatic field and potential on the NN atoms (eq 3), with a sole constraint imposed on chemically equivalent hydrogens of the methyl groups to adopt the same charges.

As a measure for the quality of the charge set $\{q_\alpha\}$ we use the relative standard deviation (SD) of the electrostatic potential (SD_V) and field (SD_E) with respect to the QM reference over all L configurations on the NN_l probe sites:

$$\text{SD}_V = \sqrt{\frac{\sum_l^L \sum_{\alpha \in NN_l} (V_{\alpha l}^{MM} - V_{\alpha l}^\rho)^2}{\sum_l^L \sum_{\alpha \in NN_l} (V_{\alpha l}^\rho)^2}}$$

$$(5)$$

$$\text{SD}_E = \sqrt{\frac{\sum_l^L \sum_{\alpha \in NN_l} \|\mathbf{E}_{\alpha l}^{MM} - \mathbf{E}_{\alpha l}^\rho\|^2}{\sum_l^L \sum_{\alpha \in NN_l} \|\mathbf{E}_{\alpha l}^\rho\|^2}}$$

$$(6)$$

As described previously[18] we optimize the weighting parameters in eq 3 by deriving charge sets for different values of w^V, w^E, and w^H. Figure 3 shows SD_V and SD_E for all atoms comprised in the QM subsystem as a function of w^E with different values for w^H. Fixed values for $w^V = 0.1$ and $w^Q = 1000$ were used. Naturally, increasing w^E leads to a lower SD_E, however also to a worse description of the electrostatic potential. Furthermore, it was observed previously[18] that high values of w^E in combination with small w^H can lead to unphysical atomic charges. For RPSB in rhodopsin we find the best compromise at $\text{SD}_V \approx 0.05$ and $\text{SD}_E \approx 0.66$ without allowing large deviations in the potential. The corresponding weights are $w^V = 0.1$, $w^Q = 1000$, $w^H = 0.01$, and $w^E = 0.1$, which we use in order to derive the atomic charges for the RPSB.

A complication in deriving FF parameters for this particular system is the QM/MM boundary, which dissects a chemical bond, in our case C_β−C_γ in Lys296. In order to saturate the QM valence we have placed a monovalent pseudopotential on
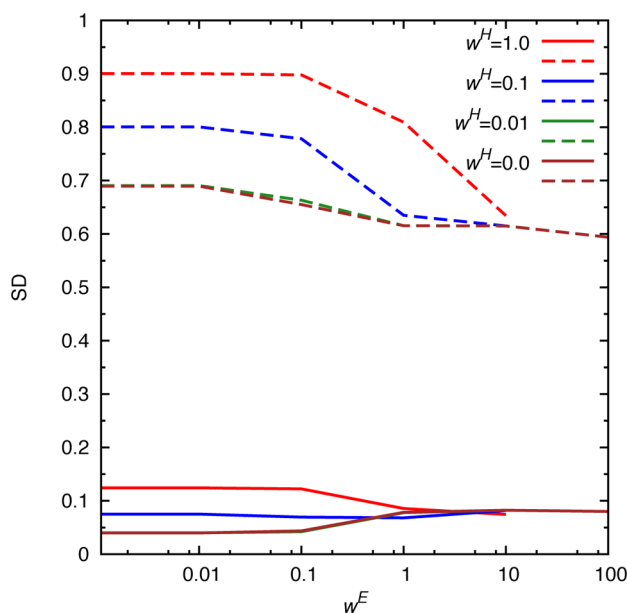
Figure 3. SD_V (solid lines) and SD_E (dashed lines) using different values for w^H. Fixed values for $w^V = 0.1$ and $w^Q = 1000$ were used.

C_β (Figure 2). Naturally, such a perturbation of the electron density in the surrounding of C_β affects the derivation of atomic charges in its vicinity. We have explored various possibilities to restrain the charges of the adjacent atoms to the values of the original force field during the charge fitting. However, this leads simply to a shift of the problematic region further into the quantum region. In each case the atomic charges closest to C_β that were not restrained during the fitting procedure adopted unphysical values in order to compensate for the boundary effects. The charges further inside the QM region, i.e. starting from roughly three bonds away from C_β, appear reasonable. Since we were mainly interested in optimized parameters for the retinal moiety we found it best practice to fit, in a first step, without any restraints atomic point charges for all atoms included in the QM region as described in section 2.1. The charges on C_β and its nearest QM neighbors are affected by the perturbations of the electronic structure at the QM/MM boundary, as can be seen in Figure 4. Clearly, the results from the unrestrained fitting (green) for C_β, C_γ, and C_δ are out of line compared to the original values (blue) and chemical intuition. The charges further away from the boundary, i.e. for the retinal moiety, are, however, chemically sound and represent the
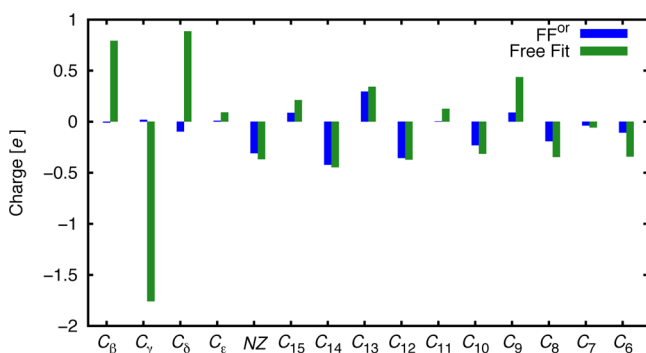
electrostatic properties of the surrounding of the QM region well. In a second step, i.e. before fitting the bonded interactions, the charges on the atoms $C_\beta - C_\varepsilon$ and the attached hydrogens were replaced by their original FF values, while preserving the total charge Q^{tot}. These charges were not optimized in this case but represented the best possible estimate. Note, however, that it is essential that the QM region is chosen large enough and the fragment for which new charges are to be derived is far enough from the QM/MM boundary.

Table 1 compares the charges of the newly derived force field FF^{FM} with the original FF^{or}. The most significant variations are highlighted in bold. In general the new charges are slightly more polar than the original ones. The average absolute change amounts to 0.06 e, while the largest change occurs for the charge on C_{11}, which increases by 0.12 e.

3.2. The Bonded Parameters. The bonded parameters are less sensitive to the perturbations of the electronic charge density at the QM/MM boundary. However, since C_β was replaced by a monovalent pseudopotential we retained all FF parameters of the interactions involving C_β from the original set FF^{or}. Here we discuss only the parameters with the most substantial changes, all values can be found in the SI.

Table 2 shows the newly derived parameters for the bonds among the heavy atoms, the parameters involving hydrogens changed even less. The parameters for $C_\beta - C_\gamma$ were retained from the original FF. In general, the new force constants are slightly smaller than the original ones. FF^{or} does not account for the single/double bond length alternation (BLA) along the conjugated chain, which is defined as the difference between the sum of the bond lengths of all the C—C single bonds and the sum of the bond lengths of all the C=C double bonds between C_5 and C_{15} of the RPSB chromophore (Figure 2)

$$BLA = [R_{C_6-C_7} + R_{C_8-C_9} + R_{C_{10}-C_{11}} + R_{C_{12}-C_{13}} + R_{C_{14}-C_{15}}]$$
$$- [R_{C_5=C_6} + R_{C_7=C_8} + R_{C_9=C_{10}} + R_{C_{11}=C_{12}} + R_{C_{13}=C_{14}}] \quad (7)$$

from C_{15} to C_5. It rather assigns the same equilibrium values and force constants to all these bonds. The latter are relatively high, typical for double bonds. In contrast, the new $\{b_n^{eq}\}$ describe the alternating bond lengths of the single and double bonds. Moreover, the new force constants for the single bonds are closer to typical values for single bonds.

Of the total of 108 new bending angle interactions only the ones with substantial changes are shown in Table 3. The force constants K_θ for angles involving a hydrogen atom changed on average by 9 kcal/mol/rad² to smaller values compared to K_θ FF^{or}. The equilibrium values θ_n^{eq} changed by 6° on average. A few stronger variations occurred mainly for the angles among the heavy atoms for which the average changes amount to 17 kcal/mol/rad² for K_θ and 9° for θ_n^{eq}, respectively.

The parameters for the dihedral angle interactions were not affected substantially; they can, however, be found in the SI.

3.3. Performance of the New Force Field. In order to assess the quality of the newly derived parameters we calculated several properties of the RPSB and its nearby environment. We compared the results from the simulations employing the original classical force field (FF^{or}), QM/MM, and the modified force field with the parameters for the RPSB optimized using the QM/MM force matching protocol (FF^{FM}) with experimental values.

The first two lines in Table 4 show the root-mean-square deviations (RMSD) of the atomic positions with respect to the

Figure 4. Atomic charges for the carbon atoms close to the monovalent pseudopotential on C_β. Blue: Values from the original FF, green: Results from unrestrained fitting.

Table 1. Comparison of the Newly Derived Atomic Point Charges (in Units of Electron Charge) for the RPSB with the Original Values[a]

atom	FF^{or}	FF^{FM}	atom	FF^{or}	FF^{FM}	atom	FF^{or}	FF^{FM}
C_β	−0.009	−0.009	H_{20i}	0.031	**0.064**	C_5	0.074	**0.329**
C_γ	0.019	0.019	C_{12}	−0.355	−0.372	C_{18}	−0.168	−0.290
$H_{\gamma i}$	0.010	0.010	H_{121}	0.157	0.208	H_{18i}	0.058	0.064
C_δ	−0.097	−0.097	C_{11}	0.003	**0.127**	C_4	−0.063	−0.048
$H_{\delta i}$	0.054	0.054	H_{111}	0.170	0.061	H_{4i}	0.041	0.002
C_ε	0.009	0.009	C_{10}	−0.232	−0.315	C_3	−0.063	**0.058**
$H_{\varepsilon i}$	0.081	0.081	H_{101}	0.179	0.059	H_{3i}	0.041	**-0.001**
NZ	−0.309	−0.368	C_9	0.090	**0.437**	C_2	−0.063	**-0.149**
H_{Z1}	0.353	0.455	C_{19}	−0.058	**-0.326**	H_{2i}	0.041	0.008
C_{15}	0.087	**0.212**	H_{19i}	0.051	**0.102**	C_1	0.125	**0.410**
H_{151}	0.231	0.206	C_8	−0.192	−0.346	C_{16}	−0.204	−0.198
C_{14}	−0.422	−0.447	H_{81}	0.096	0.191	H_{16i}	0.057	0.037
H_{141}	0.190	0.239	C_7	−0.038	−0.058	C_{17}	−0.204	−0.255
C_{13}	0.295	0.342	H_{71}	0.140	0.094	H_{17i}	0.057	0.061
C_{20}	0.019	**-0.126**	C_6	−0.107	**-0.342**			

[a] i labels the equivalent hydrogen atoms at a common carbon site. Charges up to C_ε and the attached hydrogens were kept at their FF^{or} values.

Table 2. Fitted Force Constants K_b and Equilibrium Values b^{eq} for the Bonds among Heavy Atoms in the RET and Lys296 Moiety[a]

bond	K_{b_n} FF^{or}	K_{b_n} FF^{FM}	b_n^{eq} FF^{or}	b_n^{eq} FF^{FM}
C_β−C_γ	310.00	310.00	1.53	1.53
C_γ−C_δ	310.00	168.97	1.53	1.56
C_δ−C_ε	310.00	152.74	1.53	1.56
C_ε−NZ	337.00	197.45	1.46	1.48
NZ＝C_{15}	481.00	460.63	1.34	1.33
C_{15}−C_{14}	469.00	339.29	1.40	1.41
C_{14}＝C_{13}	469.00	366.64	1.40	1.38
C_{13}−C_{12}	469.00	262.76	1.40	1.42
C_{12}＝C_{11}	469.00	386.88	1.40	1.37
C_{11}−C_{10}	469.00	292.17	1.40	1.42
C_{10}＝C_9	469.00	404.32	1.40	1.37
C_9−C_8	469.00	268.45	1.40	1.44
C_8＝C_7	469.00	458.11	1.40	1.36
C_7−C_6	469.00	233.65	1.40	1.45
C_6＝C_5	469.00	439.61	1.40	1.33
C_5−C_4	317.00	208.43	1.51	1.51
C_4−C_3	310.00	210.22	1.53	1.56
C_3−C_2	310.00	203.83	1.53	1.55
C_2−C_1	310.00	174.09	1.53	1.52
C_6−C_1	317.00	151.73	1.51	1.47
C_{16}−C_1	310.00	165.52	1.53	1.53
C_{17}−C_1	310.00	169.83	1.53	1.53
C_{18}−C_5	317.00	217.36	1.51	1.50
C_{19}−C_9	317.00	224.99	1.51	1.50
C_{20}−C_{13}	317.00	206.96	1.51	1.49

[a] Parameters involving C_β were retained from FF^{or}. Units force constants: [kcal/mol/Å^2], equilibrium values: [Å].

Table 3. Original and New Parameters for the Angles in the RET and Lys296 Moiety[a]

angles	K_{θ_n} FF^{or}	K_{θ_n} FF^{FM}	θ_n^{eq} FF^{or}	θ_n^{eq} FF^{FM}
H_{Z1}−NZ−C_ε	52.80	24.62	118.40	99.93
H_{42}−C_4−H_{41}	36.50	34.57	109.50	98.32
NZ−C_{15}−C_{14}	73.57	74.04	120.00	134.68
C_3−C_2−C_1	41.61	70.40	109.50	108.74
C_4−C_3−C_2	41.61	83.52	109.50	111.71
C_5−C_6−C_1	73.57	37.58	120.00	122.49
C_6−C_1−C_2	66.38	65.45	120.00	94.80
C_8−C_7−C_6	66.38	50.45	120.00	133.54
C_{10}−C_9−C_8	66.38	65.43	120.00	133.16
C_{16}−C_1−C_2	41.61	65.80	109.50	100.02
C_{16}−C_1−C_6	66.38	60.01	120.00	97.75
C_{17}−C_1−C_2	41.61	79.17	109.50	103.00
C_{17}−C_1−C_6	66.38	63.34	120.00	98.94
C_{17}−C_1−C_{16}	41.61	70.04	109.50	100.56
C_{19}−C_9−C_{10}	73.57	48.03	120.00	133.79
C_{19}−C_9−C_8	73.57	51.49	120.00	132.10

[a] Force constants K_θ for changes by more than 50% and equilibrium values θ^{eq} for changes by more than 10%. Parameters involving C_β were retained from FF^{or}. Units force constants: [kcal/mol/rad^2], equilibrium values: [°].

findings confirm the initial statement that the original FF^{or} already gives a good representation of the RPSB binding pocket if overall structural properties are concerned. Furthermore the values for the overall length of the conjugated chain R_{NZ-C6}, which accounts for the bent geometry of the chromophore in the binding pocket, are in very good agreement with the QM/MM data. The dihedral angle $\phi_{C_{10}-C_{11}=C_{12}-C_{13}}$ accounts for the interactions with the residues in the protein pocket that significantly distort the conjugated chain of the chromophore from planarity, which is well reproduced by all methods. $\phi_{C_5=C_6-C_7=C_8}$ measures the displacement of the ionone ring from the plane of the π-system.

The BLA in the experimental crystal structure amounts to 0.45 Å (Table 4). In contrast, the original FF does not account for the changes of the bond lengths between carbon−carbon single and double bonds in the RPSB chromophore and produces, as expected, a BLA close to 0. Since in the derivation

experimental crystal structure[66] for the heavy atoms of the RPSB only (RMSD$_{RET}$) and including all atoms within a radius of 5 Å around the RPSB (RMSD$_{RET+5}$). The values are all comparably small within 0.3 to 0.7 Å and do not differ much between the two FFs and the QM/MM results. These values support the robustness of FF^{FM}, at variance to a recent comparative study on squid rhodopsin in which it has been shown that different FFs for the RPSB can yield different configurational sampling of nearby residues.[88] Furthermore, our

Table 4. Comparison of Local Properties of the Retinal Binding Pocket (Figure 2) Averaged over Finite Temperature MD Trajectories Employing Different Levels of Theory: QM/MM (15 ps), the Original Force Field FF^{or}, and the Newly Optimized Force Field FF^{FM} (Both 40 ns)[a]

	QM/MM	FF^{or}	FF^{FM}	exp.
$RMSD_{RET}$ [Å]	0.32 ± 0.07	0.36 ± 0.08	0.34 ± 0.06	-
$RMSD_{RET+5}$ [Å]	0.57 ± 0.04	0.57 ± 0.05	0.70 ± 0.07	-
R_{NZ-C6} [Å]	11.12 ± 0.18	10.99 ± 0.17	$11.09 \pm$	11.21
$\phi_{C5-C6-C7-C8}$ [°]	-44.8 ± 9.1	-50.9 ± 9.8	-50.1 ± 10.6	-30.3
$\phi_{C10-C11-C12-C13}$ [°]	-16.1 ± 8.4	-15.7 ± 8.4	-16.6 ± 8.8	-40.8
BLA [Å]	0.31 ± 0.09	-0.02 ± 0.08	0.31 ± 0.10	0.45
$R_{NZ-CDGlu113}$ [Å]	3.68 ± 0.12	3.81 ± 0.15	3.75 ± 0.14	3.98
$R_{NZ-OEGlu113}$ [Å]	2.75 ± 0.08	2.85 ± 0.11	2.77 ± 0.09	3.45
E_{max} [eV]	2.44	2.10	2.41	2.49[87]

[a]Experimental values for structural properties were extracted from the crystal structure PDB code 1U19.[66] Atomic minimal RMSD values were calculated with respect to the crystal structure, considering the heavy atoms of the RPSB only ($RMSD_{RET}$) and including all heavy atoms within a radius of 5 Å around the RPSB ($RMSD_{RET+5}$). The absorption maximum E_{max} was obtained from a Gaussian fit to the ZINDO ($S_0 \rightarrow S_1$) spectrum, averaged over 3000 configurations.

of the parameters for the FF^{FM} all C−C bonds were treated individually, a BLA in line with the QM/MM reference value of 0.31 Å was obtained. The remaining discrepancy to the experimental represents the well-known underestimation due to the approximations in the exchange correlation functional employed for the QM part in the reference calculations. Naturally, these deficiencies of the QM/MM reference calculations cannot be overcome by the FF^{FM}. Furthermore the experimental values were obtained for a crystal structure, while our calculations were performed at pseudo *in vivo* conditions in the presence of the membrane and water.

As reported in previous studies,[44] we observe considerable differences in the counterion distances. In the crystal structure the NZ-H moiety appears to be involved in two equivalent hydrogen bonds with the carboxylate oxygens OE of the Glu113 counterion with a $R_{NZ-OEGlu113}$ distance of 3.45 Å. In contrast, all simulations prefer a stronger hydrogen bond to only one of the carboxylate oxygens. FF^{or} predicts a distance between NZ and the closer OE of 2.85 Å. An even shorter distance of 2.75 Å is observed during the QM/MM trajectory. As expected, the optimized parameters in FF^{FM} result in a distance closer to the QM/MM results around 2.77 Å. The FF^{or} predicts a distance between *NZ* and the *CD* of Glu113 ($R_{NZ-CDGlu113}$) of 3.81 Å, in good agreement with the crystal structure value of 3.98 Å . Again, the QM/MM reference calculations and FF^{FM} predict considerably shorter distances of 3.68 and 3.75 Å, respectively.

In order to investigate the effect of the modifications in the RPSB FF on the large scale structural properties of rhodopsin, we compared the local RMSDs of the individual trans-membrane helices employing FF^{or} and FF^{FM}, respectively. The results, summarized in the SI, did not reveal any significant differences between the two force fields. Furthermore, the ionic lock remained in its closed configuration for both FFs over the course of the 40 ns trajectories, showing very similar structural properties. Therefore, the newly derived FF^{FM} for the RPSB moiety does not alter the description of the global structural properties of rhodopsin significantly.

In order to validate further the optimized FF parameters for the RPSB we calculated absorption spectra from the MD trajectories at 310 K. From each trajectory we extracted 3000 equally spaced configurations over the respective total simulation times (15 ps for QM/MM, 40 ns for FF^{or} and FF^{FM}). For the calculation of the excitation energies we reduced

the total system of each configuration to a cluster model including the RPSB chromophore and a sufficiently large surrounding of about 250 atoms in total, which covers all residues with contribution of >0.01 eV to the excitation energies.[44] We calculated the ZINDO/S absorption spectra on the cluster models by weighting the excitation energy of the four lowest states with the corresponding oscillator strengths for the full set of frames. Figure 5 compares the resulting $S_0 \rightarrow$
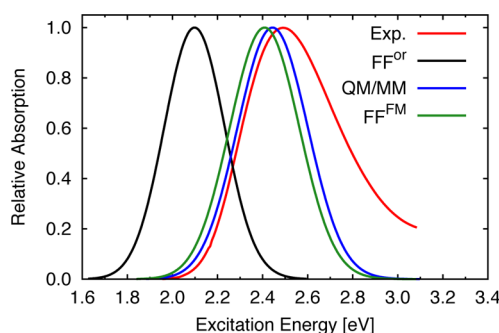

Figure 5. Relative absorption spectra with respect to the maximum as a function of the excitation energy. The calculated spectra were obtained from Gaussian fits to the ZINDO ($S_0 \rightarrow S_1$) spectra, which were averaged over 3000 configurations extracted from the respective MD trajectories at 310 K. The experimental spectrum was reconstructed from the numerical data of ref 87. Note, that the deviations in the high energy part are due to higher excitations that were not included in the calculations.

S_1 transition intensities, relative to the absorption maxima, calculated from the FF^{or}, QM/MM, and FF^{FM} trajectories. The maximum excitation energies are reported in Table 4. The absorption maximum in the spectrum obtained from the QM/MM configurations is in good agreement with the experimental result of 2.49 eV, while FF^{or} produces configurations that lead to a red-shift of 0.39 eV in the calculated absorption spectrum. As noted previously,[44] the red-shifted absorption maximum is due to the deficiency of FF^{or} to reproduce the correct BLA. This sensitive correlation between BLA and the absorption maximum of the RPSB in rhodopsin has already been reported in numerous studies.[89−93] Remarkably, in contrast to FF^{or} results, the flexibility and quality of the new parameter set FF^{FM} produces configurations that lead to absorption spectra in excellent agreement with the experimental reference.

4. CONCLUSIONS

Our group recently developed a QM/MM force matching protocol tailored to derive parameters for nonpolarizable biomolecular force fields from QM/MM reference calculations[18] with the QM fragment described at the DFT level of theory, typically described within the generalized gradient approximation. Here, we extend the original approach to systems with covalent QM/MM boundaries[18] and describe its user-friendly implementation in the publicly available software package CPMD version 3.15. The optimized FF parameters from the QM/MM force matching can be employed afterward in conjunction with any software package for classical simulations that offers the employed functional form.

We have applied the method to derive improved parameters for the RPSB in rhodopsin, which implies a covalent QM/MM boundary. The perturbation of the electron density at the boundary mainly influences the charge fitting procedure while bonded interactions are not sensitive. We have shown that within the monovalent pseudopotential approach to saturate the valence and with a sufficiently large QM region the derivation of atomic charges is unproblematic for a fragment extending up to a few bonds from the boundary. For the bonded parameters we only excluded the interactions involving the atom replaced by the monovalent pseudopotential from the fitting.

The original force field for the RPSB[35] is not flexible enough to account for the bond length alternation among the C−C single and double bonds of the RPSB chromophore. While this is a reasonable approximation to investigate large scale conformational properties of the protein it does not produce configurations of the RPSB appropriate for the calculation of absorption spectra. With our QM/MM force matching method we were able to derive a parameter set that accounts for the correct bonding properties consistent with the QM/MM reference calculations. The resulting absorption spectrum is in excellent agreement both with the results based on the QM/MM configurations and experimental values.

Overall we have shown that the MD simulations with optimized parameters perform well in reproducing properties from the QM/MM simulation. The agreement could further be improved by employing a more sophisticated functional form for the classical force field, for example by introducing a polarizable model. Our scheme can help at this point to determine whether it is sufficient to optimize *in situ* the atomic point charges or how much can be gained by adding polarization terms.

A limitation of our approach is the use of DFT to represent the electronic structure of the QM subsystem in the QM/MM reference calculations. Naturally, the drawbacks of the DFT method are inherited to the derived parameters, such as the failure in describing dispersion interactions of standard GGA functionals. However, future extensions will include parametrization schemes for the Lennard-Jones potentials based on the recently developed disperion-corrected atom-centered potentials.[59−63] A further concern is the limited time scale accessible in the QM/MM reference calculations which is typically insufficient to surpass torsional barriers. Enhanced sampling techniques or steered MD approaches could be exploited to help in such situations.[48−53]

ASSOCIATED CONTENT

⑤ Supporting Information

Complete list of the newly determined parameters for the RPSB FF^{FM} and a comparison of FF^{FM} and FF^{or} for describing global structural properties of rhodopsin and the ionic lock. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: ursula.roethlisberger@epfl.ch.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Hansson, T.; Oostenbrink, C.; van Gunsteren, W. *Curr. Opin. Struct. Biol.* **2002**, *12*, 190−196.
(2) Karplus, M.; McCammon, J. A. *Nat. Struct. Biol.* **2002**, *9*, 646−652.
(3) Reimers, J. R. *Computational Methods for Large Systems*; Electronic Structure Approaches for Biotechnology and Nanotechnology; Wiley: Weinheim, 2011.
(4) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(5) Csányi, G.; Albaret, T.; Payne, M.; De Vita, A. *Phys. Rev. Lett.* **2004**, *93*, 175503.
(6) Ercolessi, F.; Adams, J. *Europhys. Lett.* **1994**, *26*, 583−588.
(7) Liu, X.; Ercolessi, F.; Moriarty, J. *Modell. Simul. Mater. Sci. Eng.* **1996**, *4*, 293−303.
(8) Umeno, Y.; Kitamura, T.; Date, K.; Hayashi, M.; Iwasaki, T. *Comput. Mater. Sci.* **2002**, *25*, 447−456.
(9) Li, Y. H.; Siegel, D. J.; Adams, J. B.; Liu, X. Y. *Phys. Rev. B* **2003**, *67*, 125101.
(10) Aguado, A.; Bernasconi, L.; Madden, P. A. *J. Chem. Phys.* **2003**, *118*, 5704−5717.
(11) Aguado, A.; Madden, P. A. *Phys. Rev. B* **2004**, *70*, 245103.
(12) Lenosky, T.; Kress, J.; Kwon, I.; Voter, A.; Edwards, B.; Richards, D.; Yang, S.; Adams, J. *Phys. Rev. B* **1997**, *55*, 1528−1544.
(13) Lenosky, T. J.; Sadigh, B.; Alonso, E.; Bulatov, V. V.; de la Rubia, T. D.; Kim, J.; Voter, A. F.; Kress, J. D. *Modell. Simul. Mater. Sci. Eng.* **2000**, *8*, 825−841.
(14) Laio, A.; Bernard, S.; Chiarotti, G. L.; Scandolo, S.; Tosatti, E. *Science* **2000**, *287*, 1027−1030.
(15) Izvekov, S.; Parrinello, M.; Burnham, C. J.; Voth, G. A. *J. Chem. Phys.* **2004**, *120*, 10896−10913.
(16) Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2005**, *109*, 6573−6586.
(17) Knight, C.; Maupin, C. M.; Izvekov, S.; Voth, G. A. *J. Chem. Theory Comput.* **2010**, *6*, 3223−3232.
(18) Maurer, P.; Hugosson, H. W.; Colombo, M. C.; Rothlisberger, U. *J. Chem. Theory Comput.* **2007**, *3*, 628−639.
(19) Akin-Ojo, O.; Song, Y.; Wang, F. *J. Chem. Phys.* **2008**, *129*, 064108.
(20) Spiegel, K.; Magistrato, A.; Maurer, P.; Ruggerone, P.; Rothlisberger, U.; Reedijk, J.; Klein, M. L. *J. Comput. Chem.* **2008**, *29*, 39−49.
(21) Gossens, C.; Tavernelli, I.; Rothlisberger, U. *J. Am. Chem. Soc.* **2008**, *130*, 10921−10928.

(22) Sulpizi, M.; Laio, A.; VandeVondele, J.; Cattaneo, A.; Rothlisberger, U.; Carloni, P. *Proteins: Struct., Funct., Bioinf.* **2003**, *52*, 212−224.

(23) CPMD, version 3.15, Copyright IBM Corp 1990−2008, Copyright MPI für Festkörperforschung Stuttgart, 1997−2001. http://www.cpmd.org/ (accessed October 7, 2013).

(24) Palczewski, K. *Annu. Rev. Biochem.* **2006**, *75*, 743−767.

(25) Gomperts, B. D.; Kramer, I. M.; Tatham, P. E. R. *Signal Transduction*; Elsevier/Academic Press: Amsterdam, 2009.

(26) Hubbard, R.; Kropf, A. *Proc. Natl. Acad. Sci.* **1958**, *44*, 130−139.

(27) Green, B. H.; Monger, T. G.; Alfano, R. R.; Aton, B.; Callender, R. H. *Nature* **1977**, *269*, 179−180.

(28) Wang, Q.; Schoenlein, R. W.; Peteanu, L. A.; Mathies, R. A.; Shank, C. V. *Science* **1994**, *266*, 422−424.

(29) Hamm, P.; Zurek, M.; Röschinger, T.; Patzelt, H. *Chem. Phys. Lett.* **1996**, *263*, 613−621.

(30) Kukura, P.; McCamant, D. W.; Yoon, S.; Wandschneider, D. B.; Mathies, R. A. *Science* **2005**, *310*, 1006−1009.

(31) Nakamichi, H.; Okada, T. *Angew. Chem.* **2006**, *118*, 4376−4379.

(32) Salom, D.; Lodowski, D. T.; Stenkamp, R. E.; Trong, I. L.; Golczak, M.; Jastrzebska, B.; Harris, T.; Ballesteros, J. A.; Palczewski, K. *Proc. Natl. Acad. Sci.* **2006**, *103*, 16123−16128.

(33) Schwartz, T. W.; Frimurer, T. M.; Holst, B. *Annu. Rev. Pharmacol. Toxicol.* **2006**, *46*, 481−519.

(34) Röhrig, U. F.; Guidoni, L.; Rothlisberger, U. *Biochemistry* **2002**, *41*, 10799−10809.

(35) Saam, J.; Tajkhorshid, E.; Hayashi, S.; Schulten, K. *Biophys. J.* **2002**, *83*, 3097−3112.

(36) Röhrig, U. F.; Guidoni, L.; Laio, A.; Frank, I.; Rothlisberger, U. *J. Am. Chem. Soc.* **2004**, *126*, 15328−15329.

(37) Frutos, L. M.; Andruniow, T.; Santoro, F.; Ferre, N.; Olivucci, M. *Proc. Natl. Acad. Sci.* **2007**, *104*, 7764−7769.

(38) Sakmar, T. P.; Menon, S. T.; Marin, E. P.; Awad, E. S. *Annu. Rev. Biophys. Biomol. Struct.* **2002**, *31*, 443−484.

(39) Crozier, P. S.; Stevens, M. J.; Forrest, L. R.; Woolf, T. B. *J. Mol. Biol.* **2003**, *333*, 493−514.

(40) Neri, M.; Vanni, S.; Tavernelli, I.; Rothlisberger, U. *Biochemistry* **2010**, *49*, 4827−4832.

(41) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(42) Feldman, T.; Ostrovsky, M.; Kholmurodov, K.; Yasuoka, K. *Open Biochem. J.* **2012**, *6*, 94−102.

(43) Campomanes, P.; Horta, B.; Neri, M.; Röhrig, U. F.; Vanni, S.; Tavernelli, I.; Rothlisberger, U. Origin of the Color Shift among the Early Intermediates of the Rhodopsin Photocycle (to be submitted).

(44) Valsson, O.; Campomanes, P.; Tavernelli, I.; Rothlisberger, U.; Filippi, C. *J. Chem. Theory Comput.* **2013**, *9*, 2441−2454.

(45) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; Van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596−3607.

(46) http://lcbc.epfl.ch; section Research: QM/MM Force Matching (accessed July 11, 2013).

(47) VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2000**, *113*, 4863−4868.

(48) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.

(49) Tai, K. *Biophys. Chem.* **2004**, *107*, 213−220.

(50) Minary, P.; Tuckerman, M.; Martyna, G. *Phys. Rev. Lett.* **2004**, *93*, 150201.

(51) Christen, M.; Van Gunsteren, W. F. *J. Comput. Chem.* **2007**, *29*, 157−166.

(52) Laio, A.; Gervasio, F. L. *Rep. Prog. Phys.* **2008**, *71*, 126601.

(53) Liwo, A.; Czaplewski, C.; Ołdziej, S.; Scheraga, H. A. *Curr. Opin. Struct. Biol.* **2008**, *18*, 134−139.

(54) Momany, F. A. *J. Phys. Chem.* **1978**, *82*, 592−601.

(55) Cox, S. R.; Williams, D. E. *J. Comput. Chem.* **1981**, *2*, 304−323.

(56) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Phys. Chem. B* **2002**, *106*, 7300−7307.

(57) Laio, A.; Gervasio, F. L.; VandeVondele, J.; Sulpizi, M.; Rothlisberger, U. *J. Phys. Chem. B* **2004**, *108*, 7963−7968.

(58) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941−6947.

(59) Lilienfeld, O. A. v.; Tavernelli, I.; Rothlisberger, U. *Phys. Rev. Lett.* **2004**, *93*, 153004.

(60) Lin, I.; Coutinho-Neto, M.; Felsenheimer, C.; Lilienfeld, O. A. v. *Phys. Rev. B* **2007**, *75*, 1−5.

(61) Cascella, M.; Lin, I.-C.; Tavernelli, I.; Rothlisberger, U. *J. Chem. Theory Comput.* **2009**, *5*, 2930−2934.

(62) Aeberhard, P.; Arey, J.; Lin, I.; Rothlisberger, U. *J. Chem. Theory Comput.* **2009**, *5*, 23−28.

(63) Doemer, M.; Tavernelli, I.; Rothlisberger, U. *J. Chem. Theory Comput.* **2012**, *9*, 955−964.

(64) Garbow, B. S.; Hillstrom, K. E.; More, J. J. *Implementation guide for MINPACK-1*; 1980.

(65) Moré, J. J. *Numerical analysis*; Springer: Berlin, Heidelberg, 1978; pp 105−116.

(66) Okada, T.; Sugihara, M.; Bondar, A.-N.; Elstner, M.; Entel, P.; Buss, V. *J. Mol. Biol.* **2004**, *342*, 571−583.

(67) Fahmy, K.; Jäger, F.; Beck, M.; Zvyaga, T. A.; Sakmar, T. P.; Siebert, F. *Proc. Natl. Acad. Sci.* **1993**, *90*, 10206−10210.

(68) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Imprey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(69) Dickson, C. J.; Rosso, L.; Betz, R. M.; Walker, R. C.; Gould, I. R. *Soft Matter* **2012**, *8*, 9617−9627.

(70) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(71) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.

(72) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613−4621.

(73) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177−4189.

(74) Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comput. Phys.* **1999**, *151*, 283−312.

(75) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.

(76) Lee, C.; Yang, W.; Parr, R. *Phys. Rev. B* **1988**, *37*, 785−789.

(77) Colle, R.; Salvetti, O. *Theor. Chim. Acta* **1975**, *37*, 329−334.

(78) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993−2006.

(79) Hünenberger, P. H. *J. Chem. Phys.* **2000**, *113*, 10464−10476.

(80) Hockney, R. *Methods Comput. Phys.* **1970**, *9*, 135−211.

(81) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471−2474.

(82) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511−519.

(83) Ridley, J.; Zerner, M. *Theor. Chim. Acta* **1973**, *32*, 111−134.

(84) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Ö. Farkas, Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.

(85) López, C. S.; Faza, O. N.; Estévez, S. L.; de Lera, A. R. *J. Comput. Chem.* **2006**, *27*, 116−123.

(86) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.

(87) Partridge, J. C.; De Grip, W. J. *Vision Res.* **1991**, *31*, 619−630.

(88) Jardón-Valadez, E.; Bondar, A.-N.; Tobias, D. J. *Biophys. J.* **2010**, *99*, 2200−2207.

(89) Valsson, O.; Filippi, C. *J. Chem. Theory Comput.* **2010**, *6*, 1275−1292.

(90) Coccia, E.; Varsano, D.; Guidoni, L. *J. Chem. Theory Comput.* **2013**, *9*, 8−12.

(91) Bravaya, K.; Bochenkova, A.; Granovsky, A.; Nemulkhin, A. *J. Am. Chem. Soc.* **2007**, *129*, 13035−13042.

(92) Altun, A.; Yokoyama, S.; Morokuma, K. *J. Phys. Chem. B* **2008**, *112*, 6814−6827.

(93) Coto, P. B.; Strambi, A.; Ferre, N.; Olivucci, M. *Proc. Natl. Acad. Sci.* **2006**, *103*, 17154−17159.