Theor Chem Acc (2012) 131:1136
DOI 10.1007/s00214-012-1136-8

REGULAR ARTICLE

Analytic gradient and molecular dynamics simulations using the fragment molecular orbital method combined with effective potentials

Takeshi Nagata · Dmitri G. Fedorov · Kazuo Kitaura

Received: 9 May 2011 / Accepted: 29 August 2011 / Published online: 25 February 2012

Abstract The completely analytic energy gradients are derived and implemented for the two-body fragment molecular orbital (FMO2) method combined with the model core potentials (MCP) and effective fragment potentials (EFP). The many-body terms in EFP require solving coupled-perturbed Hartree-Fock equations, which are derived and implemented. The molecular dynamics (MD) simulations are performed using the FMO2/MCP method for the capped alanine decamer and with the FMO2/EFP method for the zwitterionic conformer of glycine tetramer immersed in the water layer of 6.0 Å (135 water molecules). The results of the MD simulations using the FMO2/EFP and FMO2/MCP gradients show that the total energy is conserved at the time steps less than 1 fs.

Keywords FMO · EFP · MCP · MD · FMO/EFP · FMO/MCP · Polypeptide · Zwitterion · NVT · Nose-Hoover · Large molecule · Protein · Solvent · QM/MM · Relativistic effect · Analytic gradient · SCZV

Published as part of the special collection of articles: From quantum mechanics to force fields: new methodologies for the classical simulation of complex systems.

T. Nagata (✉) · D. G. Fedorov · K. Kitaura
NRI, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan
e-mail: takeshi.nagata@aist.go.jp

T. Nagata · K. Kitaura
Graduate School of Pharmaceutical Sciences, Kyoto University, 46-29 Yoshidashimo Adachi, Sakyo-ku, Kyoto 606-8501, Japan

1 Introduction

Molecular dynamics (MD) of large systems such as proteins and DNA has become a widely used tool [1–7]. Ab initio MD [8, 9] is applicable to chemical reactions, but it is challenging for biological molecules, because the computational cost of ab initio methods scales at least as N^3 (N: system size) for a single step in MD. The reduction of scaling is a major problem in computational chemistry. For this purpose, a variety of quantum-mechanical methods has been developed based on fragmentation of large systems [10–25], whose relation has been introduced by Nagata et al. [26] and discussed in more detail in the recent review [27].

The fragment molecular orbital (FMO) method [28–31] is one such approach; it closely reproduces conventional ab initio properties while showing a nearly linear scaling. FMO has been extended to various kinds of wave functions, and its accuracy has been verified by comparing the FMO and conventional ab initio properties for Møller-Plesset perturbation theory (MP) [32–34], coupled cluster [35], density functional theory (DFT) [36, 37], multiconfigurational self-consistent field [38], configuration interaction [39, 40], time-dependent DFT [41–43], and open shell methods [44]. Continuum solvation models [45] can also be used with FMO [46–48].

The analytic energy gradient (the first derivative with respect to a nuclear coordinate) originally developed by Kitaura et al. [49] neglected the terms arising from the electrostatic potentials (ESP) coupled to the fragment electron densities, which can be shown to be related to the occupied-virtual orbital response terms, obtained by solving the coupled-perturbed Hartree-Fock (CPHF) equations of the entire system. Recently, Nagata et al. have formulated the CPHF equations necessary to obtain these terms in gas phase two-body FMO (FMO2) both at the restricted

Hartree-Fock (RHF) [26] and MP2 levels [50]. Their self-consistent Z-vector (SCZV) method [26] adds only a fraction to the total cost because it is successfully formulated for the CPHF equations of fragments, avoiding CPHF for the entire systems, although these fragment CPHF equations are coupled and require a self-consistent solution. It has been found that the response contribution due to the electrostatic potentials to the energy gradient is significant for polarizable systems and large basis sets. Using SCZV, the FMO2 energy gradients in gas phase are completely analytic, which ensures that MD simulations can be carried out without the accumulation of errors as the time evolves.

Komeiji et al. have implemented FMO/MD using the original incomplete energy gradient and shown that the loss of the energy conservation can take place [51]. In the following years, Komeiji et al. have suggested a number of improvements for FMO/MD [52–54], reviewed in 2009 [55]; Ishimoto et al. have developed alternative formulations [56, 57] and Fujita et al. have proposed path integral FMO/MD [58] and FMO/MD with periodic boundary conditions [59]. FMO/MD has been applied to a large number of systems, so far limited to small solute molecules in water, including the sampling of configurations for a consequent excited state calculations [60] and chemical reactions [61–64]. Recently, an important missing term in the gradient has been identified and implemented by Nagata et al. [65], which appears across the covalent bonds connecting fragments in proteins and other systems.

One well-established and efficient way to treat solvent explicitly is given by the effective fragment potential (EFP) method, which is derived from ab initio calculations [66, 67]. EFP can be thought of as a new generation of force fields in molecular mechanics (MM), which explicitly considers the Coulomb interaction and the polarization interaction of the solvent molecules represented in the Stone's multipole expansion and induced dipoles [68], respectively, while quantum-mechanical effects such as the electron exchange-repulsion and charge transfer are incorporated via effective one-electron operators, whose parameters are fitted to molecular orbital (MO) calculations. The first implementation of EFP using fitted parameters is denoted by EFP1 [66], which is typically used for water. Another version of EFP for the treatment of the electron-exchange and the charge-transfer contributions, which can consider any solvent, is denoted by EFP2 [69–72]. Development and applications of EFP are quite extensive [73–78]. Nagata et al. [79, 80] have interfaced FMO with EFP (FMO/EFP1) for the energy and an incomplete analytic gradient. In the application of FMO2/EFP1 to hydrated chignolin [81], the first structure in the experimental nuclear magnetic resonance (NMR) set of measured data showed the 0.819 Å root mean square deviation to the FMO optimized structure [80].

The model core potential (MCP) [82–84] is an extension to the effective core potentials [85, 86] with the proper consideration of the nodal structure of valence orbitals. MCP can be generated for heavy atoms using high level relativistic calculations [87]. A combined FMO and MCP method (FMO/MCP) has been developed by Ishikawa et al. and applied to hydrated Hg^{2+} and the Pt-containing cisplatin-DNA complex [88].

In this study, we develop completely analytic FMO2/EFP1 and FMO2/MCP gradients, of which the former presents particular difficulties because of its treatment of the many-body polarization. Secondly, we demonstrate the quality of the developed gradients by performing FMO/MD simulations for covalently connected fragments and thus extend the applicability of FMO/MD, previously limited mainly to molecular clusters.

2 Mathematical formulation

2.1 Energy gradient expression

FMO2 calculations [28, 31, 89] proceed as follows. At the first stage, individual fragment (=monomer) calculations are done using ab initio QM methods in the presence of the electrostatic potentials due to the remaining fragments. Because the monomer electron densities are mutually polarized via the electrostatic potentials, the calculations are repeated self-consistently. The procedure for determining the monomer densities and the electrostatic potentials at the monomer stage is referred to as the self-consistent charge (SCC). After convergence, dimers (=pairs of fragments) are calculated in the presence of the electrostatic potentials due to the monomer densities, which are frozen at this point. Finally, thus obtained monomer and dimer energies are summed up giving the total energy.

We briefly summarize the FMO2 energy expression [28, 30, 31, 90] and the corresponding gradient. The FMO2 total energy is given by

$$E = \sum_{I}^{N} E'_I + \sum_{I>J}^{N} (E'_{IJ} - E'_I - E'_J) + \sum_{I>J}^{N} \mathrm{Tr}(\Delta \mathbf{D}^{IJ} \mathbf{V}^{IJ}), \quad (1)$$

where E'_X is the internal fragment energy of fragment X ($X = I$ or IJ, for monomers and dimers, respectively). $\mathbf{V}^{IJ}$ is the matrix of the electrostatic potential (ESP) for dimer IJ due to the electron densities and nuclei of the remaining fragments, i.e.,

$$V^{IJ}_{\mu\nu} = \sum_{K \neq IJ}^{N} \left(u^K_{\mu\nu} + v^K_{\mu\nu} \right). \quad (2)$$

The one-electron and two-electron integrals in $V^{IJ}_{\mu\nu}$ are, respectively,

$$u_{\mu\nu}^{K} = \sum_{A\in K}\left\langle \mu\left|\frac{-Z_A}{|\mathbf{r}-\mathbf{R}_A|}\right|\nu\right\rangle, \tag{3}$$

$$v_{\mu\nu}^{K} = \sum_{\lambda\sigma\in K} D_{\lambda\sigma}^{K}(\mu\nu|\lambda\sigma), \tag{4}$$

where $D_{\lambda\sigma}^{K}$ is the density matrix element of fragment K and $(\mu\nu|\lambda\sigma)$ is the two-electron integral in the AO basis. The dimer density matrix difference $\Delta\mathbf{D}^{IJ}$ in Eq. 1 is defined by

$$\Delta\mathbf{D}^{IJ} = \mathbf{D}^{IJ} - (\mathbf{D}^{I} \oplus \mathbf{D}^{J}). \tag{5}$$

The internal fragment energy $E_X^{'}$ in Eq. 1 has the following form:

$$E'_X = \sum_{\mu\nu\in X} D_{\mu\nu}^{X} h_{\mu\nu}^{X} + \frac{1}{2}\sum_{\mu\nu\lambda\sigma\in X}\left[D_{\mu\nu}^{X}D_{\lambda\sigma}^{X} - \frac{1}{2}D_{\mu\lambda}^{X}D_{\nu\sigma}^{X}\right](\mu\nu|\lambda\sigma) + \sum_{\mu\nu\in X} D_{\mu\nu}^{X}P_{\mu\nu}^{X} + E_X^{\mathrm{NR}}, \tag{6}$$

where $h_{\mu\nu}^{X}$ is the X-mer one-electron Hamiltonian, $P_{\mu\nu}^{X}$ is the hybrid orbital projection matrix [65], and E_X^{NR} is the nuclear repulsion energy of X. In this study, the Roman induces $ijkl$ and Greek induces $\mu\nu\lambda\sigma$ denote the MO integrals and AO integrals, respectively.

The differentiation of $E_X^{'}$ with respect to a nuclear coordinate a leads to

$$\frac{\partial E'_X}{\partial a} = \sum_{\mu\nu\in X} D_{\mu\nu}^{X}\frac{\partial h_{\mu\nu}^{X}}{\partial a} + \frac{1}{2}\sum_{\mu\nu\lambda\sigma\in X}\left[D_{\mu\nu}^{X}D_{\lambda\sigma}^{X} - \frac{1}{2}D_{\mu\lambda}^{X}D_{\nu\sigma}^{X}\right] \times\frac{\partial(\mu\nu|\lambda\sigma)}{\partial a} + \sum_{\mu\nu\in X} D_{\mu\nu}^{X}\frac{\partial P_{\mu\nu}^{X}}{\partial a} - 2\sum_{i,j\in X}^{\mathrm{occ}} S_{ji}^{a,X}F_{ji}^{\prime X} - \overline{U}^{a,X,X} + \frac{\partial E_X^{\mathrm{NR}}}{\partial a}, \tag{7}$$

where the internal fragment Fock matrix element $F_{ij}^{\prime X}$ is given in terms of MOs:

$$F_{ij}^{\prime X} = h_{ij}^{X} + \sum_{k\in X}^{\mathrm{occ}}[2(ij|kk) - (ik|jk)] + P_{ij}^{X}, \tag{8}$$

and the overlap derivative is defined by

$$S_{ij}^{a,X} = \sum_{\mu\nu\in X} C_{\mu i}^{X*}\frac{\partial S_{\mu\nu}^{X}}{\partial a}C_{\nu j}^{X}. \tag{9}$$

The response term in Eq. 7 is defined as [91]

$$\overline{U}^{a,X,Y} = 4\sum_{i\in X}^{\mathrm{occ}}\sum_{r\in X}^{\mathrm{vir}} U_{ri}^{a,X}V_{ri}^{Y}, \tag{10}$$

where the MO coefficient derivative is

$$\frac{\partial C_{\mu i}^{X}}{\partial a} = \sum_{m\in X}^{\mathrm{occ+vir}} U_{mi}^{a,X}C_{\mu m}^{X}, \tag{11}$$

$U_{mi}^{a,X}$ is the orbital response of fragment X.

The differentiation of the ESP energy contribution in Eq. 1 with respect to a nuclear coordinate a leads to

$$\begin{aligned}\frac{\partial}{\partial a}\mathrm{Tr}(\Delta\mathbf{D}^{IJ}\mathbf{V}^{IJ}) = &\sum_{\mu\nu\in IJ}\Delta D_{\mu\nu}^{IJ}\sum_{K\neq IJ}^{N}\left[\frac{\partial u_{\mu\nu}^{K}}{\partial a} + \sum_{\lambda\sigma\in K} D_{\lambda\sigma}^{K}\frac{\partial(\mu\nu|\lambda\sigma)}{\partial a}\right]\\ &- 2\sum_{\mu\nu\in IJ} W_{\mu\nu}^{IJ}\frac{\partial S_{\mu\nu}^{IJ}}{\partial a} + 2\sum_{\mu\nu\in I} W_{\mu\nu}^{I}\frac{\partial S_{\mu\nu}^{I}}{\partial a}\\ &+ 2\sum_{\mu\nu\in J} W_{\mu\nu}^{J}\frac{\partial S_{\mu\nu}^{J}}{\partial a} - 2\sum_{K\neq IJ}^{N}\sum_{\mu\nu\in K}\Delta X_{\mu\nu}^{K(IJ)}S_{\mu\nu}^{a,K}\\ &+ \overline{U}^{a,IJ,IJ} - \overline{U}^{a,I,IJ} - \overline{U}^{a,J,IJ}\\ &+ 4\sum_{K\neq IJ}^{N}\sum_{\mu\nu\in IJ}\sum_{r\in K}^{\mathrm{vir}}\sum_{i\in K}^{\mathrm{occ}} U_{ri}^{a,K}\Delta D_{\mu\nu}^{IJ}(\mu\nu|ri),\end{aligned} \tag{12}$$

where

$$W_{\mu\nu}^{X} = \frac{1}{4}\sum_{\lambda\sigma\in X} D_{\mu\lambda}^{X}V_{\lambda\sigma}^{IJ}D_{\sigma\nu}^{X}, \tag{13}$$

and

$$\Delta X_{\mu\nu}^{K(IJ)} = \frac{1}{4}\sum_{\lambda\sigma\in K} D_{\mu\lambda}^{K}\left[\sum_{\zeta\eta\in IJ}\Delta D_{\zeta\eta}^{IJ}(\zeta\eta|\lambda\sigma)\right]D_{\sigma\nu}^{K}. \tag{14}$$

Collecting the response terms in Eqs. 7 and 12, the total contribution of response terms to the energy gradient can be given by the sum of $\overline{U}^{a}$ and $\Re^{a}$, where

$$\begin{aligned}\overline{U}^{a} = &-\sum_{I}^{N}\overline{U}^{a,I,I} - \sum_{I>J}^{N}\left(\overline{U}^{a,IJ,IJ} - \overline{U}^{a,I,I} - \overline{U}^{a,J,J}\right)\\ &+ \sum_{I>J}^{N}\left(\overline{U}^{a,IJ,IJ} - \overline{U}^{a,I,IJ} - \overline{U}^{a,J,IJ}\right),\end{aligned} \tag{15}$$

and

$$\Re^{a} = 4\sum_{K}^{N}\sum_{(I>J)\neq K}^{N}\sum_{\mu\nu\in IJ}\sum_{r\in K}^{\mathrm{vir}}\sum_{i\in K}^{\mathrm{occ}} U_{ri}^{a,K}\Delta D_{\mu\nu}^{IJ}(\mu\nu|ri). \tag{16}$$

Note that $\overline{U}^{a}$ cancels out unless some but not all ESP are computed with approximations (e.g., point charges) [26]. In this study, we do not employ them and use only the electrostatic dimer (ES-DIM) approximation, thereby one should compute just the $\Re^{a}$ term.

One purpose of this study is to derive the completely analytic energy gradient for the FMO2/EFP1 method and the FMO2/MCP method by considering the response contributions, Eq. 16. The FMO/EFP1 gradient has already been derived without the response contribution in Ref. [80]. To obtain the analytic FMO2/MCP gradient, one has to add the MCP derivative terms and the response terms. Since the fragment energies are variationally determined with those effective potentials in the fragment

Hamiltonians, there are no direct contributions to Eq. 16 for those methods. The additional contributions should appear only in the coupled-perturbed Hartree-Fock (CPHF) equations. We should note that Eq. 16 arises because the monomer densities are fixed in the dimer calculations.

2.2 CPHF equations for the FMO gradient

The FMO energy gradient expression is given in the previous subsection. Here, the set of CPHF equations is introduced for obtaining the orbital response contributions, Eq. 16. In the dimer RHF calculations, the dimer densities are computed without mutual self-consistency to monomer densities. Therefore, the response contribution, Eq. 16, emerges in the FMO energy gradient, and only the monomer responses $U_{ri}^{a,K}$ contribute to the gradient.

The CPHF equations in FMO are given by [26]

$$\mathbf{A}\mathbf{U}^a = \mathbf{B}_0^a, \tag{17}$$

where the fragment diagonal block of matrix $\mathbf{A}$ and the corresponding fragment off-diagonal block are, respectively,

$$A_{ij,kl}^{I,I} = \delta_{ik}\delta_{jl}\left(\epsilon_j^I - \epsilon_i^I\right) - [4(ij|kl) - (ik|jl) - (il|jk)], \tag{18}$$

and

$$A_{ij,kl}^{I,K} = -4(ij|kl). \tag{19}$$

Note that Eq. 17 has the dimension of the entire molecular system. In Eqs. 18 and 19, the former corresponds to the orbital Hessian of the target fragment I and the latter comes from the ESP acting upon I. The $ij \in I$ element of the derivative integrals $\mathbf{B}_0^a$ is

$$B_{0,ij}^{a,I} = F_{ij}^{a,I} - S_{ij}^{a,I}\epsilon_j^I - \sum_{kl\in I}^{\mathrm{occ}} S_{kl}^{a,I}[2(ij|kl) - (ik|jl)] - \sum_{K\neq I}\sum_{kl\in K}^{\mathrm{occ}} 2S_{kl}^{a,K}(ij|kl). \tag{20}$$

where the Fock matrix derivative $F_{ij}^{a,I} = F_{ij}^{\prime a,I} + V_{ij}^{a,I}$ includes the derivative of the electrostatic potential in the MO basis $V_{ij}^{a,I}$ (more detailed definitions can be found in [26]). The computation of $\mathbf{B}_0^a$ is straightforward [26]. We should note that the set of CPHF equations, Eq. 17, is derived from the perturbation of monomer RHF equations.

Substitution of Eq. 17 into the response contribution $\Re^a$ yields

$$\Re^a = \sum_K^N \sum_{r\in K}^{\mathrm{vir}} \sum_{i\in K}^{\mathrm{occ}} U_{ri}^{a,K} L_{ri}^K = \mathbf{L}^T\mathbf{U}^a = \mathbf{L}^T\mathbf{A}^{-1}\mathbf{B}_0^a = \mathbf{Z}^T\mathbf{B}_0^a \tag{21}$$

where the Lagrangian is defined by

$$L_{ri}^K = 4\sum_{(I>J)\neq K}^{N}\sum_{\mu\nu\in IJ} \Delta D_{\mu\nu}^{IJ}(\mu\nu|ri), \tag{22}$$

and the corresponding Z-vector equations with the dimension of the entire system should be solved:

$$\mathbf{A}\mathbf{Z} = \mathbf{L}. \tag{23}$$

In practice, Eq. 23 is solved by the SCZV procedure [26]. Next, the contributions due to the effective potentials are derived to be added into Eqs. 20, 18, and 19.

2.3 Effective potentials

This section considers the model core potential and the effective fragment potential (EFP1 is used throughout in this study). These effective potentials can be added into the Hamiltonian of fragment X. The particle treatment of some core electrons is replaced by MCP, and N_v represents the number of explicitly treated valence electrons. The FMO Hamiltonian of fragment X [89] is replaced by the corresponding MCP Hamiltonian [88] as follows:

$$\widehat{H}^{\mathrm{MCP},X} = \sum_{i\in X}^{N_\mathrm{v}}\left[\widehat{h}_i^{\mathrm{eff},X} + \widehat{h}_i^{\mathrm{MCP},X} + \widehat{V}^{\mathrm{eff},X}(\mathbf{r}_i) + \sum_{j(\in X)<i}^{N_\mathrm{v}}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}\right] + E_X^{\mathrm{NR,eff}}, \tag{24}$$

where the effective one-electron Hamiltonian of fragment X, the one-electron potential inherent in MCP, and the effective nuclear repulsion energy are, respectively,

$$\widehat{h}_i^{\mathrm{eff},X} = -\frac{1}{2}\nabla_i^2 - \sum_{L\in X}\frac{Z_L^{\mathrm{eff}}}{|\mathbf{r}_i - \mathbf{R}_L|} \tag{25}$$

$$\widehat{h}_i^{\mathrm{MCP},X} = -\sum_{L\in X}\frac{Z_L^{\mathrm{eff}}}{|\mathbf{r}_i - \mathbf{R}_L|}\left[\sum_k^{n_{L,\alpha}} A_{L,k}\exp(-\alpha_{L,k}|\mathbf{r}_i - \mathbf{R}_L|^2) + |\mathbf{r}_i - \mathbf{R}_L|\sum_k^{n_{L,\alpha}} B_{L,k}\exp(-\beta_{L,k}|\mathbf{r}_i - \mathbf{R}_L|^2)\right] + \sum_{L\in X}\sum_c^{N_{L,c}} B_{L,c}|\psi_{L,c}\rangle\langle\psi_{L,c}| \tag{26}$$

$$E_X^{\mathrm{NR,eff}} = \sum_{B\in X}\sum_{A(\in X)>B}\frac{Z_A^{\mathrm{eff}}Z_B^{\mathrm{eff}}}{R_{AB}}, \tag{27}$$

and $\widehat{V}_i^{\mathrm{eff},X}(\mathbf{r}_i)$ corresponds to the electrostatic potential in Eq. 2 in which the nuclear charge Z_A is replaced by the effective nuclear charge, $Z_A^{\mathrm{eff}} = Z_A - N_{A,\mathrm{c}}$ and $N_{A,\mathrm{c}}$ is the number of core electrons on atom A. L runs over atoms. The parameters $A_{L,k}$, $\alpha_{L,k}$, $B_{L,k}$, and $\beta_{L,k}$ are fitted from all-electron calculations when MCP is generated. The last term on the right hand side of Eq. 26 is the projection operator that shifts the core orbitals $|\psi_{L,c}\rangle$.

In the EFP method, the effective potential for electron i is composed of the Coulomb, polarization, and the remainder (the electron exchange-repulsion and charge transfer) terms:

$$\widehat{V}_i^{\mathrm{EFP},X} = \widehat{V}_i^{\mathrm{coul},X} + \widehat{V}_i^{\mathrm{pol},X} + \widehat{V}_i^{\mathrm{rem},X}. \tag{28}$$

The Coulomb potential acting on X is represented using the Stone's multipole expansion [68]:

$$\widehat{V}_i^{\mathrm{coul},X} = \sum_C^{N_{\mathrm{Coul}}} \left(-\frac{q_C}{r_{iC}} - \sum_a^{x,y,z} \mu_a^C \widehat{F}_a(\mathbf{r}_{iC}) - \frac{1}{3}\sum_{ab}^{x,y,z} \Theta_{ab}^C \widehat{F}'_{ab}(\mathbf{r}_{iC}) - \frac{1}{15}\sum_{abc}^{x,y,z} \Omega_{abc}^C \widehat{F}''_{abc}(\mathbf{r}_{iC}) \cdots \right), \tag{29}$$

where index C runs over the N_{Coul} expansion points of the whole solvent system (usually atoms and bond midpoints), and $r_{iC} = |\mathbf{r}_i - \mathbf{R}_C|$ is the distance between electron in X and point C. μ_a^C, Θ_{ab}^C and Ω_{abc}^C are the dipole, quadrupole, and octupole moments at C, respectively. $\widehat{F}_a(\mathbf{r}_{iC}) = -(a - a_C)/r_{iC}^3$, $\widehat{F}'_{ab}(\mathbf{r}_{iC})$, and $\widehat{F}''_{abc}(\mathbf{r}_{iC})$ give the electric field due to the QM charge, the field gradient, and the field second derivative operators, respectively. To improve the point multipole model accounting for the overlapping electron densities, the first term on the right hand side of Eq. 29 can be replaced by

$$-\frac{q_C}{r_{iC}} \rightarrow (1 - \beta_C \exp[-\alpha_C r_{iC}^2]) \left[-\frac{q_C}{r_{iC}} \right], \tag{30}$$

where α_C and β_C are fitted parameters [66].

The remainder potential for electron i due to EFP fragments is a one-electron operator,

$$\widehat{V}_i^{\mathrm{rem},X} = \sum_m^{N_{\mathrm{rem}}} \widehat{V}(m,i), \tag{31}$$

where m runs over the EFP remainder expansion points, and N_{rem} is the total number of them. For water acting as solvent, the ab initio electron exchange-repulsion plus charge-transfer contributions are fitted to Gaussian functions of $\widehat{V}^{\mathrm{rem}}(m,i)$ for water dimer [92].

The effective potentials have one-body terms, whose gradients are derived below. The polarization (induced dipole) potential contribution $\widehat{V}_i^{\mathrm{pol},X}$ in the EFP method needs a different treatment as a many-body contribution, introduced separately.

2.4 One-body terms

As mentioned above, the FMO/EFP1 gradient except the response contribution has been derived [80] and the FMO/MCP derivative contributions are straightforward to derive and implement. However, for these methods, an additional contribution should be introduced through the CPHF equations. Because they are derived from the HF equations, this subsection considers the contribution of a one-body potential to the Fock matrix.

Generally, an effective potential $\widehat{O}_1$ in the one-electron operator form combined with the electron density $D_{\mu\nu}^X$ gives rise to the following contribution to the total energy:

$$O^X(\mathbf{D}^X) = \sum_{\mu\nu\in X} D_{\mu\nu}^X \langle \mu | \widehat{O}_1 | \nu \rangle. \tag{32}$$

In FMO, the CPHF equations are obtained from monomer HF equations [26], thereby the contribution to the monomer Fock matrix in the MO basis is given by

$$O_{ij}^I = \langle i | \widehat{O}_1 | j \rangle. \tag{33}$$

The differentiation of O_{ij}^I with respect to a nuclear coordinate a yields

$$\frac{\partial O_{ij}^I}{\partial a} = \sum_{m\in I}^{\mathrm{occ+vir}} \left(U_{mi}^{a,I} \langle m | \widehat{O}_1 | j \rangle + U_{mj}^{a,I} \langle i | \widehat{O}_1 | m \rangle \right) + \sum_{\mu\nu\in X} C_{\mu i}^{I*} C_{\nu j}^I \frac{\partial \langle \mu | \widehat{O}_1 | \nu \rangle}{\partial a}. \tag{34}$$

The corresponding derivative terms contribute to the gradient directly (e.g., via the one-electron contribution to the Lagrangian and the AO-based integral derivatives similarly to the standard ab initio gradients; we do not repeat these well known formulas, see Ref. [93]). Here, we only concentrate on the modifications of the CPHF-related terms.

The first parenthesized term on the right hand side of Eq. 34 becomes a part of the Fock matrix element, leading to the orbital energies $\epsilon_j^I - \epsilon_i^I$ in Eq. 18, the second term is added into $F_{ij}^{a,I}$ in Eq. 20, and its computation is straightforward unless $\widehat{O}_1$ contains the electron density. The MCP potential $\widehat{O}_1 = \widehat{h}_1^{\mathrm{MCP},X}$, the Coulomb potential $\widehat{O}_1 = \widehat{V}_1^{\mathrm{coul},X}$, and the electron-exchange plus charge-transfer potential $\widehat{O}_1 = \widehat{V}_1^{\mathrm{rem},X}$ in EFP1 have this form (see Eqs. 26, 29–31), which is straightforward to implement. The many-body contribution contains the electron density in the operator, which is discussed in the following section.

2.5 Many-body terms

The addition of the many-body contributions to the response term in the CPHF equations is much more complicated than the one-body potentials to the Fock operator, because they usually contain the product of the electron densities.

The total EFP polarization energy contribution in the FMO2/EFP method is given by

$$E^{\mathrm{pol}} = \sum_{I}^{N} E_I^{\mathrm{pol}} + \sum_{I>J}^{N} (E_{IJ}^{\mathrm{pol}} - E_I^{\mathrm{pol}} - E_J^{\mathrm{pol}}) + E^{\mathrm{pol,efp}}, \quad (35)$$

where the X-mer polarization energy contribution E_X^{pol} and the EFP-EFP polarization energy $E^{\mathrm{pol,efp}}$ are, respectively,

$$E_X^{\mathrm{pol}} = -\frac{1}{2}\left[\mathbf{E}^X\right]^T \mathbf{p}^X - E^{\mathrm{pol,efp}} = -\frac{1}{2}\left[\mathbf{p}^X\right]^T \mathbf{E}^X - E^{\mathrm{pol,efp}} = -\frac{1}{2}\left[\mathbf{E}^X\right]^T \boldsymbol{\Xi}^{-1}\mathbf{E}^X - E^{\mathrm{pol,efp}}, \quad (36)$$

$$E^{\mathrm{pol,efp}} = -\frac{1}{2}\left[\mathbf{E}^{\mathrm{efp}}\right]^T \mathbf{p}^{\mathrm{efp}} = -\frac{1}{2}\left[\mathbf{E}^{\mathrm{efp}}\right]^T \boldsymbol{\Xi}^{-1}\mathbf{E}^{\mathrm{efp}}. \quad (37)$$

Equations 36 and 37 are defined in Ref. [80]. The superscript T stands for the transpose. The electrostatic field vector $\mathbf{E}^X$ and the induced dipole vector $\mathbf{p}^X$ are, respectively,

$$\mathbf{E}^X = \mathbf{E}^{X,\mathrm{el}}(\mathbf{D}^X) + \mathbf{E}^{X,\mathrm{nuc}} + \mathbf{E}^{\mathrm{efp}}, \quad (38)$$

$$\mathbf{p}^X = \boldsymbol{\Xi}^{-1}\mathbf{E}^X. \quad (39)$$

These vectors consist of $3N^{\mathrm{pol}}$ elements, where N^{pol} is the number of the polarizability points. In the case of a water molecule, the centroids of two OH bond orbitals and two lone-pair orbitals are chosen as the polarizability points. $\mathbf{E}^{X,\mathrm{nuc}}$ and $\mathbf{E}^{\mathrm{efp}}$ are the electrostatic fields due to a nuclei of X and the remaining EFPs. The ii diagonal block matrix of $\boldsymbol{\Xi}$ (i is a polarizability point) is defined by

$$\boldsymbol{\Xi}_{ii} = \begin{pmatrix} \alpha_{i,xx} & \alpha_{i,xy} & \alpha_{i,xz} \\ \alpha_{i,yx} & \alpha_{i,yy} & \alpha_{i,yz} \\ \alpha_{i,zx} & \alpha_{i,zy} & \alpha_{i,zz} \end{pmatrix}^{-1}, \quad (40)$$

$\boldsymbol{\Xi}_{ij} = \mathbf{0}$ for polarizability points i and j belonging to the same molecule. For different EFP molecules,

$$\boldsymbol{\Xi}_{ij} = -\begin{pmatrix} \frac{1}{r_{ij}^3} - \frac{3x_{ij}x_{ij}}{r_{ij}^5} & -\frac{3x_{ij}y_{ij}}{r_{ij}^5} & -\frac{3x_{ij}z_{ij}}{r_{ij}^5} \\ -\frac{3y_{ij}x_{ij}}{r_{ij}^5} & \frac{1}{r_{ij}^3} - \frac{3y_{ij}y_{ij}}{r_{ij}^5} & -\frac{3y_{ij}z_{ij}}{r_{ij}^5} \\ -\frac{3z_{ij}x_{ij}}{r_{ij}^5} & -\frac{3z_{ij}y_{ij}}{r_{ij}^5} & \frac{1}{r_{ij}^3} - \frac{3z_{ij}z_{ij}}{r_{ij}^5} \end{pmatrix}, \quad (41)$$

where the differences of the coordinates are defined by $x_{ij} = (x_i - x_j)$, $y_{ij} = (y_i - y_j)$, $z_{ij} = (z_i - z_j)$ and r_{ij} is the distance between polarizability points i and j.

Equation 40 represents the inverse of the polarizability tensor at a polarizability point i. The electronic contribution in Eq. 38 is described by

$$\mathbf{E}^{X,\mathrm{el}}(\mathbf{D}^X) = \sum_{\mu\nu\in X} D_{\mu\nu}^X \langle\mu|\widehat{\mathbf{E}}^{X,\mathrm{el}}|\nu\rangle, \quad (42)$$

where $\widehat{\mathbf{E}}^{X,\mathrm{el}}$ is the electrostatic field operator.

The EFP polarization contribution to the MO Fock matrix of monomer fragment I has a somewhat different form from the polarization energy E_I^{pol} because of its many-body feature:

$$W_{ij}^I = \langle i|\widehat{V}_1^{\mathrm{pol},I}|j\rangle, \quad (43)$$

where

$$\widehat{V}_1^{\mathrm{pol},I} = -\frac{1}{2}(\mathbf{p}^I + \widetilde{\mathbf{p}}^I)^T \widehat{\mathbf{E}}^{I,\mathrm{el}} \quad (44)$$

$$\widetilde{\mathbf{p}}^I = \left(\boldsymbol{\Xi}^{-1}\right)^T \mathbf{E}^I. \quad (45)$$

W_{ij}^I can be easily derived from the variation of $E_I^{\mathrm{pol}} = -\frac{1}{2}[\mathbf{E}^I]^T\boldsymbol{\Xi}^{-1}\mathbf{E}^I - E^{\mathrm{pol,efp}}$ with respect to the MO coefficients, and this expression is common both in QM/EFP (index I is not used) and FMO/EFP (for fragment I), because the variation of the EFP polarization contribution $E^{\mathrm{pol,efp}}$ is zero. Note that the induced dipoles in Eqs. 39 and 45 depend on the electron density $\mathbf{D}^X$ through the electric field $\mathbf{E}^X$, thus the operators in Eq. 43 are fundamentally different from that in Eq. 33.

The polarization contribution to the derivatives of the MO Fock matrix elements in the CPHF equations is obtained by the differentiation of W_{ij}^I with respect to a nuclear coordinate a:

$$\frac{\partial W_{ij}^I}{\partial a} = -\frac{1}{2}(\mathbf{p}^I + \widetilde{\mathbf{p}}^I)^T \left[\sum_{m\in I}^{\mathrm{occ+vir}} \left(U_{mi}^{a,I}\langle m|\widehat{\mathbf{E}}^{I,\mathrm{el}}|j\rangle + U_{mj}^{a,I}\langle i|\widehat{\mathbf{E}}^{I,\mathrm{el}}|m\rangle\right) + \sum_{\mu\nu\in I} C_{\mu i}^{I*} C_{\nu j}^I \frac{\partial\langle\mu|\widehat{\mathbf{E}}^{I,\mathrm{el}}|\nu\rangle}{\partial a}\right] - \frac{1}{2}\frac{\partial(\mathbf{p}^I + \widetilde{\mathbf{p}}^I)^T}{\partial a}\langle i|\widehat{\mathbf{E}}^{I,\mathrm{el}}|j\rangle. \quad (46)$$

Similarly to the one-body contributions (see Eq. 34 and its discussion below), these terms contribute to the gradient directly (for EFP, one can also see [80]), in addition to the CPHF terms. Below, we focus exclusively on the latter. In Eq. 46, the orbital response contributions are absorbed as a part of MO Fock matrix in the CPHF equation, leading to $\epsilon_j^I - \epsilon_i^I$ and the second term describes the atomic orbital (AO) derivative terms and the Hellman-Feynman term, corresponding to the contribution to the Fock derivative term, $F_{ij}^{a,I}$. Their implementation is straightforward.

The last term on the right hand side of Eq. 46 can be further transformed as follows:

$$\begin{aligned} &-\frac{1}{2}\frac{\partial(\mathbf{p}^I + \widetilde{\mathbf{p}}^I)^T}{\partial a}\langle i|\widehat{\mathbf{E}}^{I,\mathrm{el}}|j\rangle \\ &= -\frac{1}{2}\left(\frac{\partial\boldsymbol{\Xi}^{-1}}{\partial a}\mathbf{E}^I + \boldsymbol{\Xi}^{-1}\frac{\partial\mathbf{E}^I}{\partial a} + \frac{\partial\left(\boldsymbol{\Xi}^{-1}\right)^T}{\partial a}\mathbf{E}^I + \left(\boldsymbol{\Xi}^{-1}\right)^T\frac{\partial\mathbf{E}^I}{\partial a}\right)^T\langle i|\widehat{\mathbf{E}}^{I,\mathrm{el}}|j\rangle \\ &= -\frac{1}{2}\left(-\left(\mathbf{p}^I\right)^T\frac{\partial\boldsymbol{\Xi}^T}{\partial a}\left(\boldsymbol{\Xi}^{-1}\right)^T + \frac{\partial\left(\mathbf{E}^I\right)^T}{\partial a}\left(\boldsymbol{\Xi}^{-1}\right)^T - \left(\widetilde{\mathbf{p}}^I\right)^T\frac{\partial\boldsymbol{\Xi}}{\partial a}\boldsymbol{\Xi}^{-1} + \frac{\partial\left(\mathbf{E}^I\right)^T}{\partial a}\boldsymbol{\Xi}^{-1}\right)\langle i|\widehat{\mathbf{E}}^{I,\mathrm{el}}|j\rangle, \end{aligned} \quad (47)$$

where the relation

$$\frac{\partial \boldsymbol{\Xi}^{-1}}{\partial a} = -\boldsymbol{\Xi}^{-1}\frac{\partial \boldsymbol{\Xi}}{\partial a}\boldsymbol{\Xi}^{-1}, \tag{48}$$

is used. The substitution of Eq. 47 into Eq. 46 leads to

$$\begin{aligned}\frac{\partial W_{ij}^{I}}{\partial a} &= -\frac{1}{2}\left(\mathbf{p}^{I}+\widetilde{\mathbf{p}}^{I}\right)^{T}\left[\sum_{m\in I}^{\text{occ+vir}}\left(U_{mi}^{a,I}\langle m|\widehat{\mathbf{E}}^{I,\text{el}}|j\rangle\right.\right.\\ &\quad\left.\left.+U_{mj}^{a,I}\langle i|\widehat{\mathbf{E}}^{I,\text{el}}|m\rangle\right)+\sum_{\mu\nu\in I}C_{\mu i}^{I*}C_{\nu j}^{I}\frac{\partial\langle\mu|\widehat{\mathbf{E}}^{I,\text{el}}|\nu\rangle}{\partial a}\right]\\ &\quad+\frac{1}{2}\left(\left(\mathbf{p}^{I}\right)^{T}\frac{\partial\boldsymbol{\Xi}^{T}}{\partial a}\left(\boldsymbol{\Xi}^{-1}\right)^{T}-\frac{\partial\left(\mathbf{E}^{I}\right)^{T}}{\partial a}\left(\boldsymbol{\Xi}^{-1}\right)^{T}\right.\\ &\quad\left.+\left(\widetilde{\mathbf{p}}^{I}\right)^{T}\frac{\partial\boldsymbol{\Xi}}{\partial a}\boldsymbol{\Xi}^{-1}-\frac{\partial\left(\mathbf{E}^{I}\right)^{T}}{\partial a}\boldsymbol{\Xi}^{-1}\right)\langle i|\widehat{\mathbf{E}}^{I,\text{el}}|j\rangle\\ &= -\frac{1}{2}\left(\mathbf{p}^{I}+\widetilde{\mathbf{p}}^{I}\right)^{T}\left[\sum_{m\in I}^{\text{occ+vir}}\left(U_{mi}^{a,I}\langle m|\widehat{\mathbf{E}}^{I,\text{el}}|j\rangle\right.\right.\\ &\quad\left.\left.+U_{mj}^{a,I}\langle i|\widehat{\mathbf{E}}^{I,\text{el}}|m\rangle\right)+\sum_{\mu\nu\in X}C_{\mu i}^{I*}C_{\nu j}^{I}\frac{\partial\langle\mu|\widehat{\mathbf{E}}^{I,\text{el}}|\nu\rangle}{\partial a}\right]\\ &\quad-\frac{1}{2}\frac{\partial\left(\mathbf{E}^{I}\right)^{T}}{\partial a}\left(\mathbf{p}^{I,ij}+\widetilde{\mathbf{p}}^{I,ij}\right)\\ &\quad+\frac{1}{2}\left(\left(\widetilde{\mathbf{p}}^{I,ij}\right)^{T}\frac{\partial\boldsymbol{\Xi}}{\partial a}\mathbf{p}^{I}+\left(\widetilde{\mathbf{p}}^{I}\right)^{T}\frac{\partial\boldsymbol{\Xi}}{\partial a}\mathbf{p}^{I,ij}\right),\end{aligned} \tag{49}$$

where

$$\mathbf{p}^{I,ij} = \boldsymbol{\Xi}^{-1}\langle i|\widehat{\mathbf{E}}^{I,\text{el}}|j\rangle \tag{50}$$

$$\widetilde{\mathbf{p}}^{I,ij} = \left(\boldsymbol{\Xi}^{-1}\right)^{T}\langle i|\widehat{\mathbf{E}}^{I,\text{el}}|j\rangle. \tag{51}$$

In the second half of Eq. 49, the $\mathbf{E}^{I,\text{el}}(\mathbf{D}^{I})$ derivative term extracted from the second term leads to

$$\begin{aligned}&-\frac{1}{2}\frac{\partial\left[\mathbf{E}^{I,\text{el}}(\mathbf{D}^{I})\right]^{T}}{\partial a}\left(\mathbf{p}^{I,ij}+\widetilde{\mathbf{p}}^{I,ij}\right)\\ &= -\frac{1}{2}\left[\sum_{l\in I}^{\text{occ}}\sum_{m\in I}^{\text{occ+vir}}2\left(U_{ml}^{a,I}\langle m|\widehat{\mathbf{E}}^{I,\text{el}}|l\rangle+U_{ml}^{a,I}\langle l|\widehat{\mathbf{E}}^{I,\text{el}}|m\rangle\right)\right.\\ &\quad\left.+\sum_{\mu\nu\in I}D_{\mu\nu}^{I}\frac{\partial\langle\mu|\widehat{\mathbf{E}}^{I,\text{el}}|\nu\rangle}{\partial a}\right]^{T}\left(\mathbf{p}^{I,ij}+\widetilde{\mathbf{p}}^{I,ij}\right)\\ &= -\frac{1}{2}\sum_{l\in I}^{\text{occ}}\sum_{k\in I}^{\text{vir}}4U_{kl}^{a,I}\left(\langle k|\widehat{\mathbf{E}}^{I,\text{el}}|l\rangle\right)^{T}\left(\mathbf{p}^{I,ij}+\widetilde{\mathbf{p}}^{I,ij}\right)\\ &\quad-\frac{1}{2}\sum_{kl\in I}^{\text{occ}}\left(-2S_{kl}^{a,I}\right)\left(\langle k|\widehat{\mathbf{E}}^{I,\text{el}}|l\rangle\right)^{T}\left(\mathbf{p}^{I,ij}+\widetilde{\mathbf{p}}^{I,ij}\right)\\ &\quad-\frac{1}{2}\left[\sum_{\mu\nu\in X}D_{\mu\nu}^{I}\frac{\partial\langle\mu|\widehat{\mathbf{E}}^{I,\text{el}}|\nu\rangle}{\partial a}\right]^{T}\left(\mathbf{p}^{I,ij}+\widetilde{\mathbf{p}}^{I,ij}\right).\end{aligned} \tag{52}$$

Substitution of Eq. 52 into Eq. 49 yields

$$\begin{aligned}\frac{\partial W_{ij}^{I}}{\partial a} &= -\frac{1}{2}\left(\mathbf{p}^{I}+\widetilde{\mathbf{p}}^{I}\right)^{T}\left[\sum_{m\in I}^{\text{occ+vir}}\left(U_{mi}^{a,I}\langle m|\widehat{\mathbf{E}}^{I,\text{el}}|j\rangle\right.\right.\\ &\quad\left.\left.+U_{mj}^{a,I}\langle i|\widehat{\mathbf{E}}^{I,\text{el}}|m\rangle\right)+\sum_{\mu\nu\in I}C_{\mu i}^{I*}C_{\nu j}^{I}\frac{\partial\langle\mu|\widehat{\mathbf{E}}^{I,\text{el}}|\nu\rangle}{\partial a}\right]\\ &\quad-2\sum_{l\in I}^{\text{occ}}\sum_{k\in I}^{\text{vir}}U_{kl}^{a,I}\left(\langle k|\widehat{\mathbf{E}}^{I,\text{el}}|l\rangle\right)^{T}\left(\mathbf{p}^{I,ij}+\widetilde{\mathbf{p}}^{I,ij}\right)\\ &\quad+\sum_{kl\in I}^{\text{occ}}S_{kl}^{a,I}\left(\langle k|\widehat{\mathbf{E}}^{I,\text{el}}|l\rangle\right)^{T}\left(\mathbf{p}^{I,ij}+\widetilde{\mathbf{p}}^{I,ij}\right)\\ &\quad-\frac{1}{2}\left(\sum_{\mu\nu\in X}D_{\mu\nu}^{I}\frac{\partial\langle\mu|\widehat{\mathbf{E}}^{I,\text{el}}|\nu\rangle}{\partial a}+\frac{\partial\left[\mathbf{E}^{I,\text{nuc}}+\mathbf{E}^{\text{efp}}\right]}{\partial a}\right)^{T}\\ &\quad\times\left(\mathbf{p}^{I,ij}+\widetilde{\mathbf{p}}^{I,ij}\right)\\ &\quad+\frac{1}{2}\left(\left(\widetilde{\mathbf{p}}^{I,ij}\right)^{T}\frac{\partial\boldsymbol{\Xi}}{\partial a}\mathbf{p}^{I}+\left(\widetilde{\mathbf{p}}^{I}\right)^{T}\frac{\partial\boldsymbol{\Xi}}{\partial a}\mathbf{p}^{I,ij}\right).\end{aligned} \tag{53}$$

The terms in Eq. 53 can be classified into the contributions to $B_{0,ij}^{a,I}$ and $F_{ij}^{a,I}$:

$$\begin{aligned}B_{0,ij}^{a,I,\text{pol}} &= \sum_{kl\in I}^{\text{occ}}S_{kl}^{a,I}\left(\langle k|\widehat{\mathbf{E}}^{I,\text{el}}|l\rangle\right)^{T}\left(\mathbf{p}^{I,ij}+\widetilde{\mathbf{p}}^{I,ij}\right)\\ &\quad-\frac{1}{2}\left(\sum_{\mu\nu\in I}D_{\mu\nu}^{I}\frac{\partial\langle\mu|\widehat{\mathbf{E}}^{I,\text{el}}|\nu\rangle}{\partial a}\right)^{T}\left(\mathbf{p}^{I,ij}+\widetilde{\mathbf{p}}^{I,ij}\right)\\ &\quad-\frac{1}{2}\frac{\partial\left[\mathbf{E}^{I,\text{nuc}}+\mathbf{E}^{\text{efp}}\right]^{T}}{\partial a}\left(\mathbf{p}^{I,ij}+\widetilde{\mathbf{p}}^{I,ij}\right)\\ &\quad+\frac{1}{2}\left(\left(\widetilde{\mathbf{p}}^{I,ij}\right)^{T}\frac{\partial\boldsymbol{\Xi}}{\partial a}\mathbf{p}^{I}+\left(\widetilde{\mathbf{p}}^{I}\right)^{T}\frac{\partial\boldsymbol{\Xi}}{\partial a}\mathbf{p}^{I,ij}\right),\end{aligned} \tag{54}$$

$$F_{ij}^{a,I,\text{pol}} = -\frac{1}{2}\left(\mathbf{p}^{I}+\widetilde{\mathbf{p}}^{I}\right)^{T}\sum_{\mu\nu\in X}C_{\mu i}^{I*}C_{\nu j}^{I}\frac{\partial\langle\mu|\widehat{\mathbf{E}}^{I,\text{el}}|\nu\rangle}{\partial a}, \tag{55}$$

and the contribution to the matrix $\mathbf{A}$:

$$\begin{aligned}A_{ij,kl}^{I,I,\text{pol}} &= 2\left(\langle k|\widehat{\mathbf{E}}^{I,\text{el}}|l\rangle\right)^{T}\left(\mathbf{p}^{I,ij}+\widetilde{\mathbf{p}}^{I,ij}\right)\\ &= 2\left(\mathbf{p}^{I,kl}+\widetilde{\mathbf{p}}^{I,kl}\right)^{T}\langle i|\widehat{\mathbf{E}}^{I,\text{el}}|j\rangle.\end{aligned} \tag{56}$$

Note that the sign of the first term in Eq. 56 is opposite in the above contribution to the $\mathbf{A}$ matrix compared to the correction term in Eq. 53.

Practically, the product of $A_{ij,kl}^{I,I,\text{pol}}$ and the Z-vector Z_{kl}^{I} is computed in the SCZV procedure (also see Eq. 23) [26]:

$$\begin{aligned}\sum_{k\in I}^{\text{vir}}\sum_{l\in I}^{\text{occ}}A_{ij,kl}^{I,I,\text{pol}}Z_{kl}^{I} &= 2\sum_{k\in I}^{\text{vir}}\sum_{l\in I}^{\text{occ}}\left(\mathbf{p}^{I,kl}+\widetilde{\mathbf{p}}^{I,kl}\right)^{T}\langle i|\widehat{\mathbf{E}}^{I,\text{el}}|j\rangle Z_{kl}^{I}\\ &= 2\sum_{k\in I}^{\text{vir}}\sum_{l\in I}^{\text{occ}}\left(\left\{\boldsymbol{\Xi}^{-1}+\left[\boldsymbol{\Xi}^{-1}\right]^{T}\right\}\langle k|\widehat{\mathbf{E}}^{I,\text{el}}|l\rangle\right)^{T}\\ &\quad\times\langle i|\widehat{\mathbf{E}}^{I,\text{el}}|j\rangle Z_{kl}^{I}\end{aligned}$$

$$= 2\left(\left\{\boldsymbol{\Xi}^{-1} + \left[\boldsymbol{\Xi}^{-1}\right]^T\right\}\mathbf{E}^{I,\mathrm{el}}(\overline{\mathbf{Z}}^I)\right)^T \langle i|\widehat{\mathbf{E}}^{I,\mathrm{el}}|j\rangle$$
$$= 2\left(\mathbf{p}^{I,\mathbf{Z}^I} + \widetilde{\mathbf{p}}^{I,\mathbf{Z}^I}\right)^T \langle i|\widehat{\mathbf{E}}^{I,\mathrm{el}}|j\rangle, \quad (57)$$

where the following definitions

$$\mathbf{E}^{I,\mathrm{el}}(\overline{\mathbf{Z}}^I) = \sum_{\mu\nu\in I}\langle\mu|\widehat{\mathbf{E}}^{I,\mathrm{el}}|\nu\rangle\overline{Z}^I_{\mu\nu} \quad (58)$$

$$\overline{Z}^I_{\mu\nu} = \sum_{k\in I}^{\mathrm{vir}}\sum_{l\in I}^{\mathrm{occ}}\frac{C^{I*}_{\mu k}Z^I_{kl}C^I_{\nu l} + C^{I*}_{\nu k}Z^I_{kl}C^I_{\mu l}}{2}, \quad (59)$$

are used for the derivation of Eq. 57. The induced dipoles due to the symmetrized Z-vector, $\mathbf{p}^{I,\mathbf{Z}^I}$ and $\widetilde{\mathbf{p}}^{I,\mathbf{Z}^I}$ should be solved by the following equations on each SCZV iteration (because of the $\overline{\mathbf{Z}}^I$ dependence)

$$\boldsymbol{\Xi}\mathbf{p}^{I,\mathbf{Z}^I} = \mathbf{E}^{I,\mathrm{el}}(\overline{\mathbf{Z}}^I) \quad (60)$$

$$\boldsymbol{\Xi}^T\widetilde{\mathbf{p}}^{I,\mathbf{Z}^I} = \mathbf{E}^{I,\mathrm{el}}(\overline{\mathbf{Z}}^I). \quad (61)$$

The SCZV equations considering the effective potentials explicitly do not introduce a considerable extra cost. The calculation of the induced dipoles due to the Z-vector in Eqs. 60 and 61 corresponds to that of the induced dipoles due to the density matrix in the SCC calculation [80].

3 Computational details

The developed FMO2/EFP1 and FMO2/MCP gradients were implemented into GAMESS [94, 95]. To verify that the analytic FMO2/EFP1 and FMO2/MCP energy gradients are complete, short MD simulations of 50 fs were performed with the NVE ensemble. The accuracy of FMO2/MD based on the analytic gradient is investigated. The zwitterionic form of glycine tetramer is computed in gas phase and immersed in a water layer of 6.0 Å (135 water molecules) using VEGA [96]; alanine dacamer capped with acetyl and $-NHCH_3$ groups, $(ALA)_{10}$, is computed in gas phase. For all polypeptide molecules, the one residue per fragment partition is employed.

To determine the initial structure for the demonstrative short MD simulation, the geometry optimization [97] was performed for the trans-structure of hydrated tetraglycine at the FMO2-RHF/EFP1/cc-pVDZ level. The same optimization calculation was carried out at the conventional MO-RHF/EFP1/cc-pVDZ level as well. A solute part of the optimized structures were used for the respective short MD simulations in gas phase. Using the optimized structure of hydrated tetraglycine, the FMO2/EFP1 MD simulation proceeded at the same level with the time step of 0.2 fs in the NVT ensemble at 315 K (using Nose-Hoover thermostats [98]) until the temperature is stable, followed by the further MD simulation of 593 fs under the same conditions. The relaxed structure (Fig. 1a) from this MD simulation was used in the consequent shorter FMO2/EFP1 MD simulations. One can see from Fig. 1a that the EFP water molecules are still bound with each other and charged groups of the solute after the MD simulation. These structures were used for the short MD simulations in the conventional MO method and the FMO2/EFP1 method with the incomplete gradients for comparison. All MD simulations, whose results are compared to each other, always had the same initial structure and atomic velocities.

For $(ALA)_{10}$, the core 1s electrons of C, N, and O were treated with MCP. In the covalent bond detachment in FMO [89], the bond detached atom in each detached bond is put into two fragments (redundantly) and MCPs were used only in the main fragment containing this bond detached atom, while the other fragment, where this atom is only used to describe the detached bond, did not have the MCP on this atom. To obtain the initial structure for the MD simulation, the geometry was optimized at the FMO2-RHF/MCP level with the MCP-DZP basis set (see Fig. 1b) [83]. This is also used as the initial geometry of the MD simulation with the incomplete FMO2/MCP gradients. We should note that the MD simulations with the corresponding conventional MO gradients (using the geometry optimized at the corresponding level) and the incomplete FMO2 gradients were executed for comparison.

For all the FMO calculations, the electrostatic dimer (ES-DIM) approximation was applied with the default value, $L_{\mathrm{ES\text{-}DIM}} = 2.0$ and the other electrostatic approximations are switched off [90, 91]. In the ES-DIM approximation, only the response term of Eq. 16 can be considered [26]. There are two types of EFP1 parameter sets for the electron-exchange plus charge-transfer potential: one is the RHF parameter set and another is the DFT parameter set [99]. This study used the former throughout [92]. The velocity-Verlet algorithm was used for time integration throughout. All gradients developed in this work are parallelized with the generalized distributed data interface (GDDI) [100].

4 Results and discussion

4.1 Gradient accuracy

To verify the accuracy of the analytic FMO2/EFP1 and FMO2/MCP gradients, we compared them with the corresponding numeric gradients. Numeric FMO2/EFP1-RHF and FMO2/MCP-RHF energy gradients were computed

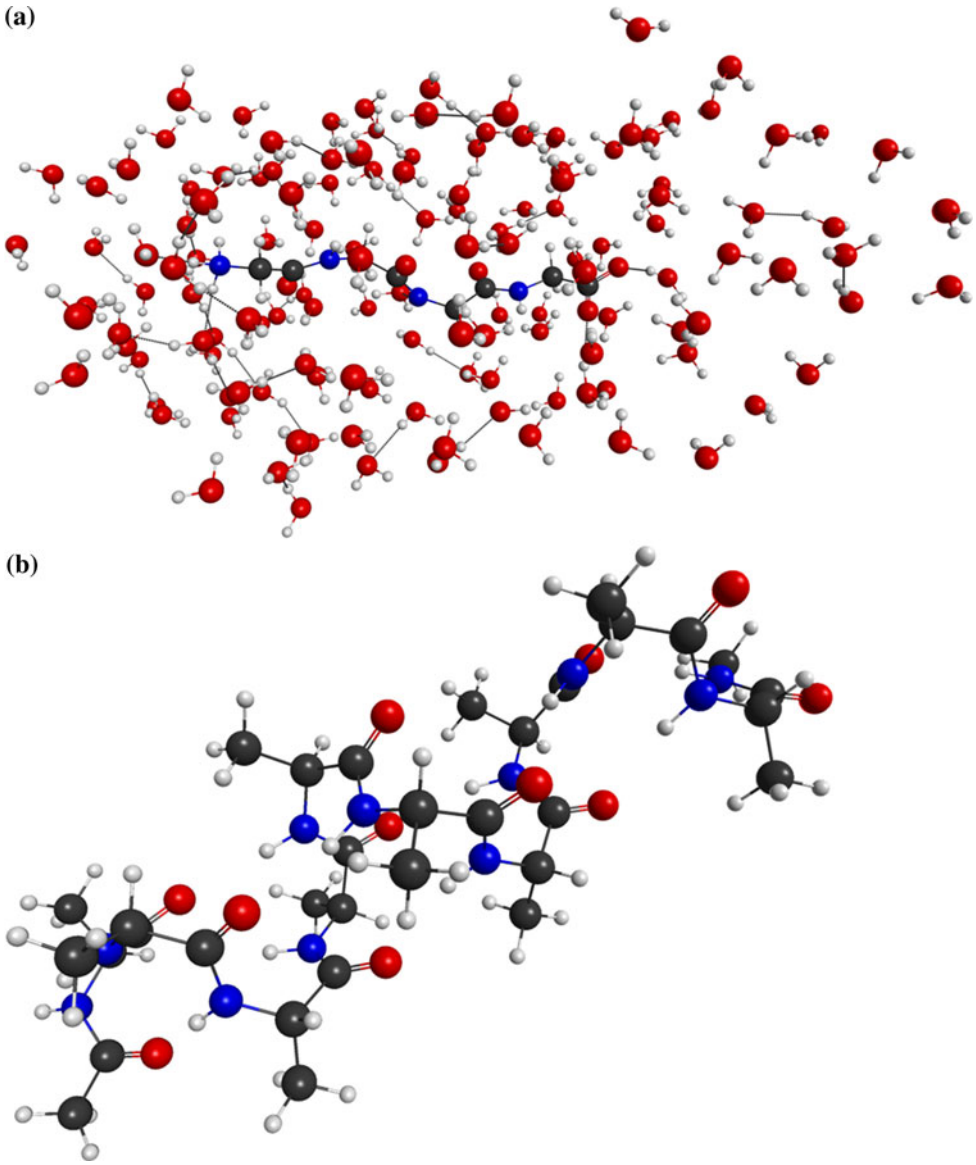

Fig. 1 **a** Structure of zwitterionic tetraglycine in the water layer of 6.0 Å relaxed with the MD simulation in the NVT ensemble (at 315 K) at the FMO-RHF/EFP/cc-pVDZ level and **b** structure of alanine decamer $(ALA)_{10}$ optimized at the FMO-RHF/MCP/MCP-DZP level

with double differencing and a coordinate step of 0.0005 Å. The initial structures of hydrated tetraglycine and $(ALA)_{10}$ for the short MD simulations were used for the respective gradient calculations.

Table 1 displays the errors of analytic FMO2/MCP gradients relative to the corresponding numeric gradients for $(ALA)_{10}$. The bond detached atoms (BDA) and bond attached atoms (BAA) at the two ends of detached bonds form a representative set for the accuracy tests. As shown in Table 1, all of the deviations are sufficiently small.

Next, the analytic and numeric FMO2/EFP1 gradient calculations were performed for hydrated tetraglycine with the geometry obtained after the NVT MD simulation of 593 fs. For the solute, the RMSD between the analytic and numeric gradients is 8.1×10^{-7} a.u. and the maximum absolute deviation is 5.9×10^{-6} a.u.

4.2 Zwitterionic tetraglycine

There are some arguments on the structural stability of peptides in water solvent both empirically and theoretically [101–105]. It is important to sample the structures and to check the free energy difference between the zwitterionic and neutral forms. When a 0 K geometry optimization is performed, the thermal movement of solvent is not accounted for, overestimating the hydrogen bonding. Also, MD sampling is capable of overcoming local minima and converging toward the global one.

Before discussing the results of the MD simulations, we first check the energy conservation of zwitterionic tetraglycine in gas phase. For systems with the fragmentation across the covalent bonds such as tetraglycine, the FMO2/MD simulation using the incomplete gradients does not

Table 1 Deviations of the analytic and numeric gradients (a.u.) for a representative set of bond detached atoms (BDA) and bond attached atoms (BAA) in $(ALA)_{10}$ at the FMO-RHF/MCP-DZP level

	x	*y*	*z*
C1(BDA)	−0.000007	0.000000	0.000000
C1(BAA)	−0.000004	0.000003	0.000004
C2(BDA)	−0.000001	0.000006	0.000007
C2(BAA)	0.000005	−0.000002	0.000008
C3(BDA)	0.000012	0.000005	0.000000
C3(BAA)	0.000000	0.000007	0.000000
C4(BDA)	−0.000002	0.000009	−0.000004
C4(BAA)	−0.000009	0.000011	−0.000002
C5(BDA)	−0.000002	0.000001	0.000007
C5(BAA)	−0.000007	0.000004	0.000008
C6(BDA)	0.000007	0.000001	0.000006
C6(BAA)	0.000002	−0.000001	0.000004
C7(BDA)	0.000003	0.000010	−0.000003
C7(BAA)	0.000009	0.000008	−0.000006
C8(BDA)	−0.000018	0.000005	0.000000
C8(BAA)	−0.000003	0.000002	0.000000
C9(BDA)	0.000003	0.000001	0.000011
C9(BAA)	0.000003	0.000003	0.000002

conserve the energy [51]. Figure 2a displays a double logarithmic plot of the root mean square deviation (RMSD) of the total energies relative to the energy of 10th step against time step Δt for zwitterionic tetraglycine in gas phase. For the FMO2/MD simulations with the completely analytic gradients, the slope of 2.0 indicates that the energy is sufficiently conserved even at the smaller time step such as 0.1 fs and the gradient is analytic because the RMSD value is proportional to Δt^2 in the velocity-Velret method [51]. This tendency is in good agreement with that in the conventional MO-MD simulations, while the FMO2/MD simulations with the incomplete gradients give a wrong description at the shorter time steps. The corresponding numeric values in Table 2 reinforce these findings. The results imply that FMO2/MD with the one residue per fragment partition is sufficiently accurate and shows perfect energy conservation. The energy conservation at 0.1 fs also means that it is possible to pursue the course of chemical reactions.

The initial structure of tetraglycine in gas phase for the MD simulations was obtained from the FMO2 geometry optimization for hydrated tetraglycine. Since the geometry optimization is done at 0 K, the structure is different from that determined at 315 K, which is typical for biological molecules in solution. Thus, the structure is more realistic when obtained by the MD simulation in the NVT ensemble. This study employed the Nose-Hoover thermostats in the NVT ensemble [98], and we performed the MD

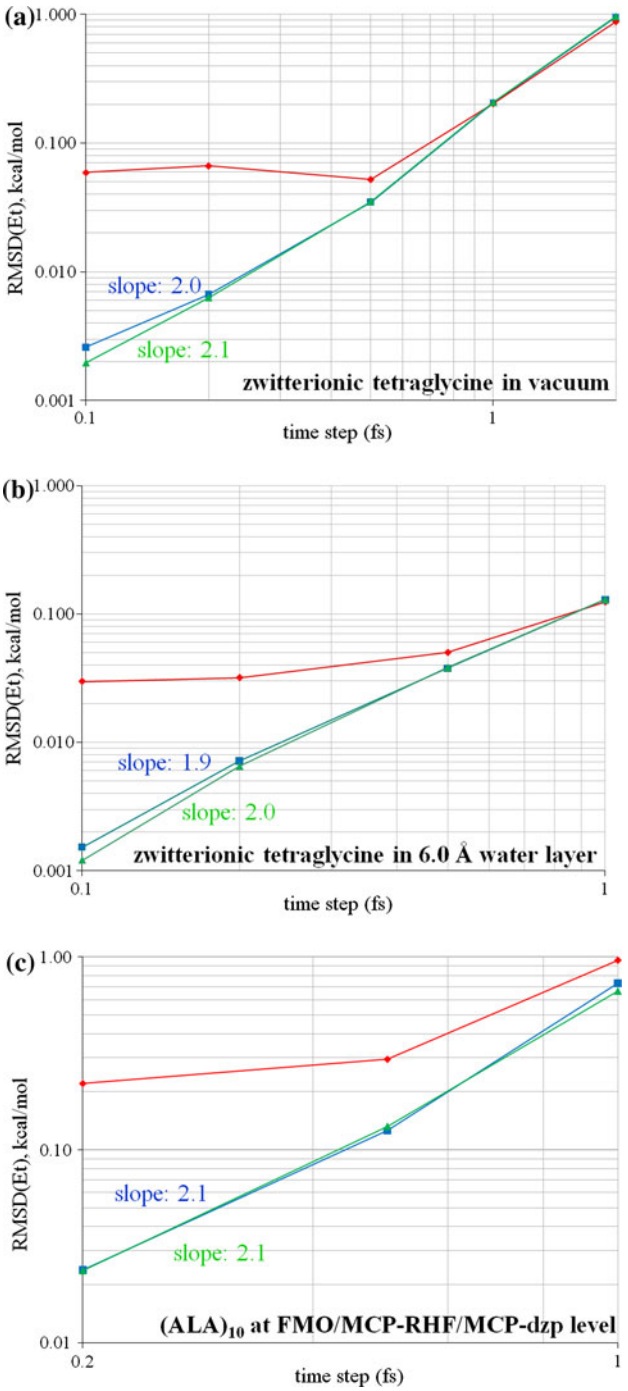

Fig. 2 Double logarithmic plots of the RMSD of the total energies relative to the energy of the 10th step against time step Δt in the short MD simulations of 50 fs in the NVE ensemble. *Triangle* markers are used for MO-MD, *square* for FMO/MD with the analytic gradients, and *diamond* for FMO/MD with the incompletely analytic gradients

simulation at 315 K for the hydrated tetraglycine until the total energy and temperature are stable and a stationary structure is taken for the short MD simulations of 50 fs for time steps of 0.1, 0.2, 0.5, and 1.0 fs. Figure 2b displays that the slopes in the MO- and FMO2-based MD/EFP1

Table 2 Root mean square deviations of the total energies relative to the 10th value (in kcal/mol) in the MD simulations of 50 fs at several time steps

Time step	Ab initio MO	FMO[a]	FMO[b]
Zwitterionic tetraglycine in gas phase			
0.1 fs	0.0020	0.0026	0.0592
0.2 fs	0.0063	0.0067	0.0667
0.5 fs	0.0352	0.0348	0.0524
1.0 fs	0.2068	0.2055	0.2030
2.0 fs	0.9608	0.9484	0.8769
Zwitterionic tetraglycine in 6.0 Å water layer			
0.1 fs	0.0012	0.0015	0.0298
0.2 fs	0.0065	0.0071	0.0319
0.5 fs	0.0384	0.0377	0.0501
1.0 fs	0.1293	0.1294	0.1243
Capped alanine decamer: $(ALA)_{10}$			
0.2 fs	0.0237	0.0239	0.2204
0.5 fs	0.1324	0.1256	0.2944
1.0 fs	0.6642	0.7295	0.9633

[a] Analytic gradient in this work

[b] Original incomplete gradient

Table 3 Time-Averages and RMSDs of the kinetic energy, potential energy, total energy (in kcal/mol) and temperature (in K)

	Kinetic energy	Potential energy	Total energy	Temperature
Hydrated tetraglycine at time step of 0.1 fs in the NVE ensemble				
MO/MD				
Average	276.737	-5.67877066×10^5	-5.67600324×10^5	
RMSD	7.058	7.058		
FMO/MD[a]				
Average	276.651	-5.67877335×10^5	-5.67600681×10^5	
RMSD	7.090	7.091		
FMO/MD[b]				
Average	276.637	-5.67877344×10^5	-5.67600704×10^5	
RMSD	7.092	7.105		
Hydrated tetraglycine at time step of 0.2 fs in the NTV ensemble (315K)				
FMO/MD[a]				
Average	281.536	-5.67695274×10^5	-5.67413738×10^5	313.891
RMSD	8.107	9.028	4.190	9.036
$(ALA)_{10}$ at time step of 0.2 fs in the NVE ensemble				
MO/MD				
Average	49.198	-3.11678589×10^5	-3.11629391×10^5	
RMSD	5.706	5.715		
FMO/MD[a]				
Average	48.984	-3.11672215×10^5	-3.11623232×10^5	
RMSD	6.029	6.039		
FMO/MD[b]				
Average	48.979	-3.11672195×10^5	-3.11623216×10^5	
RMSD	6.046	6.117		

[a] Analytic gradient in this work

[b] Original incomplete gradient

simulations are 1.9 and 2.0, showing that accurate long MD simulations are possible without accumulating the errors due to the incomplete gradients.

Table 3 shows the time-averaged kinetic energy, potential energy, and total energy calculated for the hydrated tetraglycine in the MD simulations of 50 fs for the time step of 0.1 fs. The FMO2/MD energies are in good agreement with MO-MD. The FMO2/MD simulation with the incomplete gradients also give reasonably accurate energies, because there is little accumulation of errors in

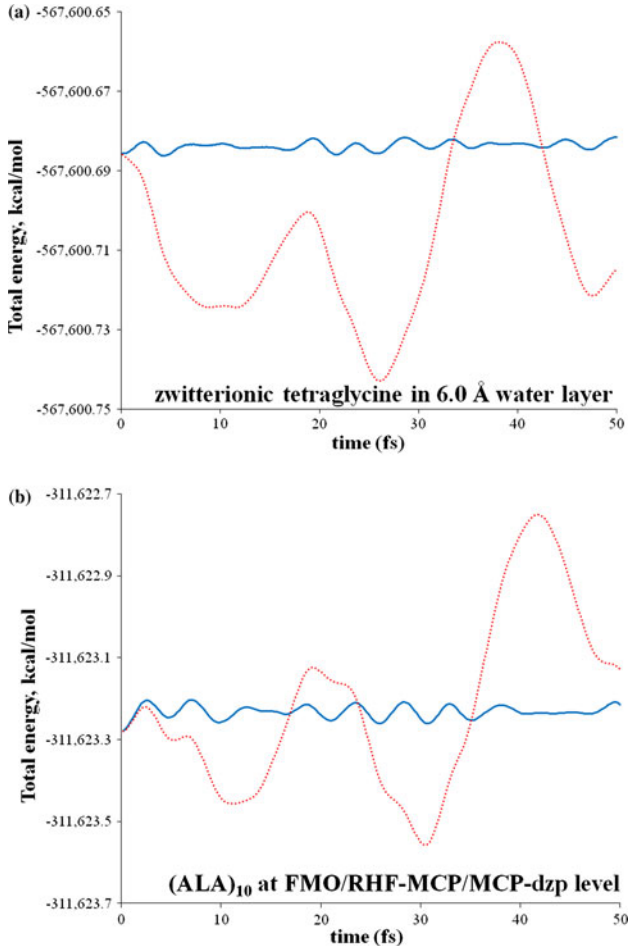

Fig. 3 Plots of the total energies against the time in the short FMO/MD simulations with the NVE ensemble: **a** for the zwitterionic tetraglycine in the water layer of 6.0 Å for the time step of 0.1 fs and **b** $(ALA)_{10}$ at the time step of 0.2 fs. *Solid line* is used for the analytic energy gradient. *Dashed line* is used for the incompletely analytic energy gradient

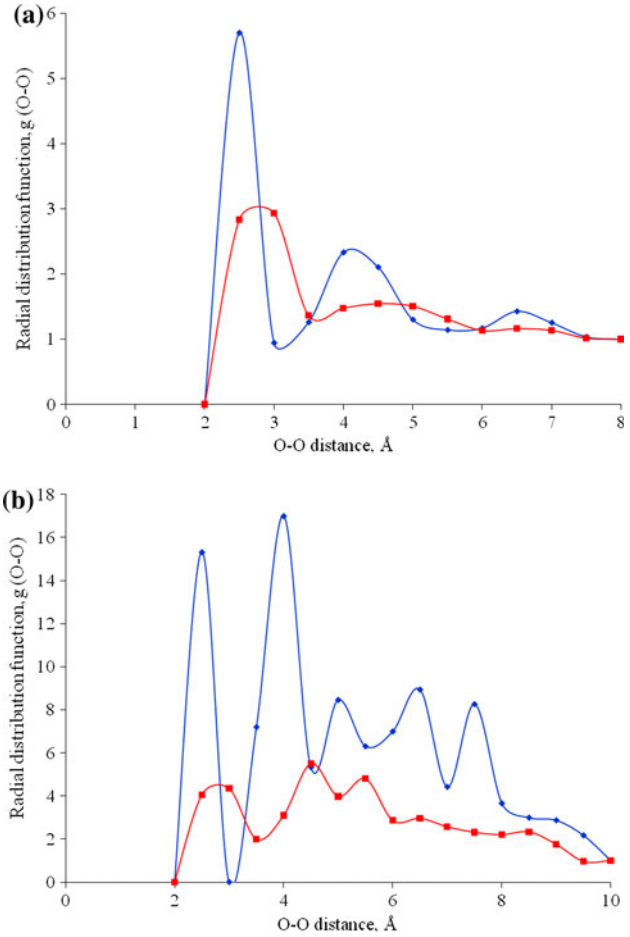

Fig. 4 Radial distribution functions of **a** O–O distances in the EFP water molecules of hydrated tetraglycine and **b** distances between oxygens of the COO^- group and of the EFP water molecules of hydrated tetraglycine. *Diamond* markers show the data for the structure optimized at the FMO-RHF/EFP/cc-pVDZ level. *Square* markers show the data for the time-averaged structures in the MD simulation of 592 fs

such a short MD simulation. The plot of the total energy against the time in the MD simulation is also another way of checking the energy conservation. In Fig. 3a, one can see that the total energies in the FMO2/MD simulation with the analytic gradients are conserved, while those with the incomplete gradients fluctuate. The preliminary MD simulation of 593 fs (see Table 3) is close to 315 K, meaning that the NVT MD simulations are properly conducted.

Figure 4a shows the radial distribution functions (RDF) against the O–O distances of the water molecules in the structures of hydrated tetraglycine optimized at the FMO2-RHF/EFP1/cc-pVDZ level versus those obtained in the NVT MD simulation of 593 fs for the time step of 0.2 fs. The fluctuations of O–O distances are larger for the latter, as one can easily expect, and the FMO2/EFP1-MD simulations properly describe the fluctuation of the water molecules. For the MD simulations, the peak in the first solvation shell reasonably corresponds to the typical O–O distance in water at 315 K, while the O–O distance optimized in the FMO method is somewhat shorter. Figure 4b plots the RDFs between the oxygen atoms of the COO^- group and oxygen atoms of the water molecules obtained at the same levels of theory as in Fig. 4a. For the RDF obtained in the MD simulation, the peak in the first solvation shell also shifts to the larger O–O distance compared with that for the FMO2 geometry optimization. In addition, for the latter, there is a vacancy between the first shell and the second shell, while for the former, the water molecules can be found there because of the fluctuation and the second shell shifts to the larger O–O distances.

4.3 Alanine decamer

The FMO/MCP method can applied to systems containing heavy metals such as transition metals, for which the relativistic effects are significant. Another use of the

FMO/MCP method is to reduce the basis set superposition errors (BSSE) [88], because the core orbitals, which is one of main causes of BSSE, are replaced by the potentials. The α-helix conformer of $(ALA)_{10}$ has several hydrogen bonds, so it is a good test system for the FMO2/MCP method. The intramolecular hydrogen bonds hold together the structures of polypeptides, proteins, and DNA.

In Fig. 2c, the slope of 2.1 in the FMO2/MD simulations shows that the FMO2/MD simulations at short time steps are accurate even with the one residue per fragment partition, and the slope is in good agreement with the conventional MO-MD simulations. Another energy conservation test in Fig. 3b shows that the total energy is well conserved for this system, when the analytic energy gradient is used in simulations. So far $(ALA)_{10}$ appears to be the largest system fragmented across covalent bonds as treated in the FMO2/MD simulations, and the applications to larger systems such as proteins and DNA are now possible.

We measured the total wall clock times for the FMO2/MCP and MO/MCP MD simulations of 50 fs at the time step of 0.2 fs using the Soroban cluster: 192 cores of Xeon and compared them and the number of GDDI groups was set to 5 for the former [100]. The total wall clock times for the former and the latter were 3763.9 min and 3721.7 min, respectively. The latter calculation is faster, because the system is too small to treat in the FMO method and the MO method in GAMESS is well parallelized. The computation time in the FMO gradient calculation is discussed in Ref. [26], which directly affects that in the MD simulation. It is concluded that FMO/MD is preferable for large systems with many fragments.

5 Conclusions

The analytic two-body FMO2/EFP1 and FMO2/MCP energy gradients have been successfully derived and implemented in this study. The results of FMO2/MD simulations using these gradients are in good agreement with those of the corresponding conventional MO-MD simulations. Our study has demonstrated that FMO2/MD can be performed for systems with covalently connected fragments, which very considerably expands its application field previously limited to molecular clusters.

By verifying the accuracy in MD simulations, we have been able to use an efficient partition of one residue per fragment, while earlier studies [51] suggested that even larger partitions of two residues per fragment do not have sufficient accuracy. It is clear that the ability to perform molecular dynamics using FMO will largely extend its usefulness, for instance, to evaluate the entropic contribution in the protein-ligand binding, so far often neglected in the applications of FMO or evaluated using force fields.

Acknowledgments This work has been supported by the Next Generation Super Computing Project, Nanoscience Program (MEXT, Japan), and Computational Materials Science Initiative (CMSI, Japan). We thank Dr. Yuto Komeiji, Prof. Mark Gordon and Kurt Brorsen for fruitful discussions.

References

1. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, Debolt S, Ferguson D, Seibel G, Kollman P (1995) Comput Phys Commun 91:1
2. Kollman P, Massova I, Reyes C, Kuhn B, Huo SH, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE (2000) Acc Chem Res 33:889
3. Bashford D, Case DA (2000) Annu Rev Phys Chem 51:129
4. Mackerell AD (2004) J Comput Chem 25:1584
5. Wang W, Donini O, Reyes CM, Kollman PA (2001) Annu Rev Biophys Biomol Struct 30:211
6. Warshel A (2003) Annu Rev Biophys Biomol Struct 32:425
7. Kollman P (1993) Chem Rev 93:2395
8. Car R, Parrinello M (1985) Phys Rev Lett 55:2471
9. Schlegel HB, Millam JM, Iyengar SS, Voth GA, Daniels AD, Scuseria GE, Frisch MJ (2001) J Chem Phys 114:9758
10. Morokuma K (1971) J Chem Phys 55:1236
11. Ohno K, Inokuchi H (1972) Theor Chim Acta (Berl.) 26:331
12. Otto P, Ladik J (1975) Chem Phys 8:192
13. Barandiaran Z, Seijo L (1988) J Chem Phys 89:5739
14. Gao JL (1997) J Phys Chem B 101:657
15. Xie W, Orozco M, Truhlar DG, Gao J (2009) J Chem Theory Comput 5:459
16. Leverentz HR, Truhlar DG (2009) J Chem Theory Comput 5:1573
17. Gordon MS, Mullin JM, Pruitt SR, Roskop LB, Slipchenko LV, Boatz JA (2009) J Phys Chem B 113:9646
18. Huang L, Massa L, Karle I, Karle J (2009) Proc Natl Acad Sci USA 106:3664
19. Tong Y, Mei Y, Zhang JZH, Duan LL, Zhang QG (2009) J Theor Comput Chem 8:1265
20. Söderhjelm P, Kongsted J, Ryde U (2010) J Chem Theory Comput 6:1726
21. Mata RA, Stoll H, Cabral BJC (2009) J Chem Theory Comput 5:1829
22. Yeole SD, Gadre S (2010) J Chem Phys 132:094102
23. Makowski M, Korchowiec J, Gu FL, Aoki Y (2010) J Comput Chem 31:1733
24. Kobayashi M, Kunisada T, Akama T, Sakura D, Nakai H (2011) J Chem Phys 134:034105
25. He X, Merz KM (2010) J Chem Theory Comput 6:405
26. Nagata T, Brorsen K, Fedorov DG, Kitaura K, Gordon MS (2011) J Chem Phys 134:124115
27. Gordon MS, Fedorov DG, Pruitt SR, Slipchenko LV (2012) Chem Rev 112:632
28. Kitaura K, Ikeo E, Asada T, Nakano T, Uebayasi M (1999) Chem Phys Lett 313:701
29. Fedorov DG, Kitaura K (2004) J Chem Phys 120:6832
30. Fedorov DG, Kitaura K (2007) J Phys Chem A 111:6904
31. Fedorov, DG, Kitaura, K (eds) (2009) The fragment molecular orbital method: practical applications to large molecular systems. CRC Press, Boca Raton
32. Fedorov DG, Kitaura K (2004) J Chem Phys 121:2483
33. Mochizuki Y, Nakano T, Koikegami S, Tanimori S, Abe Y, Nagashima U, Kitaura K (2004) Theor Chem Acc 112:442
34. Mochizuki Y, Yamashita K, Fukuzawa K, Takematsu K, Watanabe H, Taguchi N, Okiyama Y, Tsuboi M, Nakano T, Tanaka S (2010) Chem Phys Lett 493:346

35. Fedorov DG, Kitaura K (2005) J Chem Phys 123:134103/1
36. Sugiki SI, Kurita N, Sengoku Y, Sekino H (2003) Chem Phys Lett 382:611
37. Fedorov DG, Kitaura K (2004) Chem Phys Lett 389:129
38. Fedorov DG, Kitaura K (2005) J Chem Phys 122:054108/1
39. Mochizuki Y, Koikegami S, Amari S, Segawa K, Kitaura K, Nakano T (2005) Chem Phys Lett 406:283
40. Mochizuki Y, Tanaka K, Yamashita K, Ishikawa T, Nakano T, Amari S, Segawa K, Murase T, Tokiwa H, Sakurai M (2007) Theor Chem Acc 117:541
41. Chiba M, Fedorov DG, Kitaura K (2007) Chem Phys Lett 444:346
42. Chiba M, Fedorov DG, Kitaura K (2007) J Chem Phys 127:104108
43. Chiba M, Fedorov DG, Kitaura K (2008) J Comput Chem 29:2667
44. Pruitt SR, Fedorov DG, Kitaura K, Gordon MS (2010) J Chem Theory Comput 6:1
45. Tomasi J, Mennucci B, Cammi R (2005) Chem Rev 105:2999
46. Fedorov DG, Kitaura K, Li H, Jensen JH, Gordon MS (2006) J Comput Chem 27:976
47. Li H, Fedorov DG, Nagata T, Kitaura K, Jensen JH, Gordon MS (2010) J Comput Chem 31:778
48. Watanabe H, Okiyama Y, Nakano T, Tanaka S (2010) Chem Phys Lett 500:116
49. Kitaura K, Sugiki SI, Nakano T, Komeiji Y, Uebayasi M (2001) Chem Phys Lett 336:163
50. Nagata T, Fedorov DG, Ishimura K, Kitaura K (2011) J Chem Phys 135:044110
51. Komeiji Y, Nakano T, Fukuzawa K, Ueno Y, Inadomi Y, Nemoto T, Uebayasi M, Fedorov DG, Kitaura K (2003) Chem Phys Lett 372:342
52. Komeiji Y, Ishikawa T, Mochizuki Y, Yamataka H, Nakano T (2009) J Comput Chem 30:40
53. Komeiji Y, Mochizuki Y, Nakano T (2010) Chem Phys Lett 484:380
54. Mochizuki Y, Nakano T, Komeiji Y, Yamashita K, Okiyama Y, Yoshikawa H, Yamataka H (2011) Chem Phys Lett 504:95
55. Komeiji Y, Mochizuki Y, Nakano T, Fedorov DG (2009) J Mol Str (THEOCHEM) 898:2
56. Ishimoto T, Tokiwa H, Teramae H, Nagashima U (2004) Chem Phys Lett 387:460
57. Ishimoto T, Tokiwa H, Teramae H, Nagashima U (2005) J Chem Phys 122:094905
58. Fujita T, Watanabe H, Tanaka S (2009) J Phys Soc Jpn 78:104723
59. Fujita T, Nakano T, Tanaka S (2011) Chem Phys Lett 506:112
60. Mochizuki Y, Komeiji Y, Ishikawa T, Nakano T, Yamataka H (2007) Chem Phys Lett 437:66
61. Sato M, Yamataka H, Komeiji Y, Mochizuki Y, Ishikawa T, Nakano T (2008) J Am Chem Soc 130:2396
62. Fujiwara T, Mochizuki Y, Komeiji Y, Okiyama Y, Mori H, Nakano T, Miyoshi E (2010) Chem Phys Lett 490:41
63. Fujiwara T, Mori H, Mochizuki Y, Tatewaki H, Miyoshi E (2010) J Mol Str (THEOCHEM) 949:28
64. Sato M, Yamataka H, Komeiji Y, Mochizuki Y, Nakano T (2010) Chem Eur J 16:6430
65. Nagata T, Fedorov DG, Kitaura K (2010) Chem Phys Lett 492:302
66. Day NP, Jensen HJ, Gordon SM, Webb PS (1996) J Chem Phys 105:1968
67. Gordon MS, Freitag MA, Bandyopadhyay P, Jensen JH, Kairys V, Stevens WJ (2001) J Phys Chem A 105:293
68. Stone AJ (1996) The theory of intermolecular forces. Oxford University Press, New York
69. Jensen JH, Gordon MS (1998) J Chem Phys 108:4772
70. Jensen JH (2001) J Chem Phys 114:8775
71. Adamovic I, Gordon MS (2005) Mol Phys 103:379
72. Li H, Gordon MS (2006) J Chem Phys 124:214108
73. Aguilar CM, Rocha WR (2011) J Phys Chem B 115:2030
74. Minezawa N, De Silva N, Zahariev F, Gordon MS (2011) J Chem Phys 134:124115
75. Kosenkov D, Slipchenko LV (2011) J Phys Chem A 115:392
76. Ghosh D, Kosenkov D, Vanovschi V, Williams CF, Herbert JM, Gordon MS, Schmidt MW, Slipchenko LV, Krylov AI (2010) J Phys Chem A 114:12739
77. Slipchenko LV (2010) J Phys Chem A 114:8824
78. Arora P, Slipchenko LV, Webb SP, DeFusco A, Gordon MS (2010) J Phys Chem A 114:6742
79. Nagata T, Fedorov DG, Kitaura K, Gordon MS (2009) J Chem Phys 131:024101
80. Nagata T, Fedorov DG, Sawada T, Kitaura K, Gordon MS (2011) J Chem Phys 134:034110
81. Honda S, Yamasaki K, Sawada Y, Morii H (2004) Structure 12:1507
82. Huzinaga S (1995) Can J Chem 73:619
83. Miyoshi E, Mori H, Hirayama R, Osanai Y, Noro T, Honda H, Klobukowski M (2005) J Chem Phys 122:074104
84. Klobukowski M, Huzinaga S, Sakai Y (1999) In: Leszczynski J (eds) Computational chemistry: reviews of current trends, vol 3, World Scientific, Singapore, p 49
85. Frenking G, Antes I, Böhme M, Dapprich S, Ehlers AW, Jonas V, Neuhaus MOA, Stegmann R, Veldkamp A, Vyboishchikov SF (1996) In: Lipkowitz KB, Boyd DB (eds) Reviews in computational chemistry, vol 8, VCH Publishers, New York, p 63
86. Cundari TR, Benson MT, Lutz ML, Sommerer SO (1996) In: Lipkowitz KB, Boyd DB (eds) Reviews in computational chemistry, vol 8, VCH Publishers, New York, p 145
87. Zeng T, Fedorov DG, Klobukowski M (2010) J Chem Phys 133:114107
88. Ishikawa T, Mochizuki Y, Nakano T, Amari S, Mori H, Honda H, Fujita T, Tokiwa H, Tanaka S, Komeiji Y, Fukuzawa K, Tanaka K, Miyoshi E (2006) Chem Phys Lett 427:159
89. Nagata T, Fedorov DG, Kitaura K (2011) In: Zalesny R, Papadopoulos MG, Mezey PG, Leszczynski J (eds) Linear-scaling techniques in computational chemistry and physics, Springer, Berlin, pp 17–64
90. Nakano T, Kaminuma T, Sato T, Fukuzawa K, Akiyama Y, Uebayasi M, Kitaura K (2002) Chem Phys Lett 351:475
91. Nagata T, Fedorov DG, Kitaura K (2009) Chem Phys Lett 475:124
92. Chem W, Gordon MS (1996) J Chem Phys 105:11081
93. Yamaguchi Y, Schaefer HF III, Osamura Y, Goddard J (1994) A new dimension to quantum chemistry: analytical derivative methods in ab initio molecular electronic structure theory. Oxford University Press, New York CAN 123:66488 65-3 General Physical Chemistry USA Book written in English
94. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JJ, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA (1993) J Comput Chem 14:1347
95. Gordon MS, Schmidt MW (2005) Theory and applications of computational chemistry, the first forty years. Elsevier, Amsterdam
96. Pedretti A, Villa L, Vistoli G (2002) J Mol Graph Model 21:47
97. Fedorov DG, Ishida T, Uebayasi M, Kitaura K (2007) J Phys Chem A 111:2722
98. Nose S (1991) Prog Theor Phys Suppl 103:1
99. Adamovic I, Freitag MA, Gordon MS (2003) J Chem Phys 118:6725
100. Fedorov DG, Olson RM, Kitaura K, Gordon MS, Koseki S (2004) J Comput Chem 25:872

101. Jensen JH, Gordon MS (1995) J Am Chem Soc 117:8159
102. Aikens CM, Gordon MS (2006) J Am Chem Soc 128:12835
103. Wada G, Tamura E, Okina M, Nakamura M (1982) Bull Chem Soc Jpn 55:3064
104. Mullin JM, Gordon MS (2009) J Phys Chem B 113:8657
105. Mullin JM, Gordon MS (2009) J Phys Chem B 113:14413