PCCP

Dynamic Article Links

Cite this: *Phys. Chem. Chem. Phys.*, 2011, **13**, 11731–11738

www.rsc.org/pccp

PAPER

A combined quantum mechanical and molecular mechanical method using modified generalized hybrid orbitals: implementation for electronic excited states

Yukio Kawashima,*[ab] Haruyuki Nakano,[a] Jaewoon Jung[c] and Seiichiro Ten-no[d]

Received 19th February 2011, Accepted 18th April 2011
DOI: 10.1039/c1cp20438f

The generalized hybrid orbital (GHO) method is implemented at the second-order approximate coupled cluster singles and doubles (CC2) level for quantum mechanical (QM)/molecular mechanical (MM) electronic excited state calculations. The linear response function of CC2 in the GHO scheme is derived and implemented. The new implementation is applied to the first singlet excited states of three aromatic amino acids, phenylalanine, tyrosine, and tryptophan, and also bacteriorhodopsin for assessment. The results obtained for aromatic amino acids agreed well with the full QM CC2 calculations, while the calculated excitation energies of bacteriorhodopsin and its chromophore, all-*trans* retinal, reproduced the environmental shift of the experimental data. For the bacteriorhodopsin case, the environmental shift of GHO also showed good agreements with the experimental data. The contribution of the quantum effect of certain moieties in the excited states is elucidated by changing the partitioning of QM and MM regions.

1. Introduction

Combined quantum mechanical and molecular mechanical (QM/MM)[1] methods have been successful in tackling electronic structure calculations in large-scale systems. Particularly in the past decade, QM/MM methods have been applied to various problems, such as chemical reactions in enzymes or solvent environments.[2,3] In QM/MM methods, the calculated molecular system is divided into at least two different subsystems: a subsystem where the quantum effect must be included is treated by QM methods, and the other is treated by MM methods to include environmental effects. If the molecular system consists only of small molecules, partitioning into QM and MM regions will be straightforward; the QM and MM regions can be separated between molecules. However, applying QM/MM methods to enzymes, the boundary between QM and MM regions will be within the molecule, *i.e.*, the boundary of the QM and MM regions will be an atom or a covalent bond. Many studies have been published on this topic.[4–6]

The scheme most widely used to partition QM and MM regions is the "link atom" approach.[7–15] This method partitions the QM and MM regions between covalent bonds. Extra atoms are added to the dangling atoms of the QM system to saturate the valencies. The conventional link atom scheme, which adds a hydrogen atom to the QM system, has two drawbacks. First, the interaction between these link atoms and atoms in the MM region is not physical. Second, the electronegativity of hydrogen atoms is different from the covalently bonded atoms before partitioning. To overcome the first problem, the scaled-position link atom method has been proposed to transfer forces on link atoms to other "physical" atoms.[16] Zhang *et al.* added a fluorine-like atom using the pseudobond scheme,[17–19] Antes and Thiel added an atom with a modified basis set so that the link atom behaves as a methyl group,[20] and DiLabio *et al.* employed conventional effective core potentials to the boundary atom[21] to reduce errors caused by the second problem.

The other well-known partitioning schemes, the local self-consistent field (LSCF) method[22–25] developed by Philipp and Friesner[26,27] and the generalized hybrid orbital method (GHO),[28–31] set the boundaries between QM and MM regions on atoms by constructing frozen hybrid orbitals. In other words, the partitioning is done between hybrid orbitals. The LSCF method developed by Rivail *et al.* saturates the valency of the QM boundary atom by utilizing strictly localized orbitals with a predetermined density obtained from calculations on a smaller model system. In the LSCF method, the boundary atom is a QM atom with a frozen density for the hybrid orbital pointing toward the MM fragment. Philipp and Friesner employed Boys-localized orbitals obtained from model molecules in place of bonds. In the GHO method,

[a] *Department of Chemistry, Graduate School of Sciences, Kyushu University, Fukuoka 812-8581, Japan. E-mail: snow@ccl.scc.kyushu-u.ac.jp*
[b] *Institute of Advanced Research, Kyushu University, Fukuoka 812-8581, Japan*
[c] *RIKEN Advanced Institute for Computational Science, Kobe, 650-0047, Japan*
[d] *Graduate School of System Informatics, Kobe University, Kobe 657-8501, Japan*

which was initially developed by Gao *et al.*, the boundary atom is treated as both a QM and a MM atom. The hybridization coefficients of the hybrid orbitals on the boundary atom in the GHO method are determined by the molecular geometry. Thus, the GHO method does not require a model system, which is necessary in the method using localized orbitals. Moreover, the charge density of the hybrid orbital that is bonded to the QM region is optimized self-consistently with all other atomic orbitals in the SCF calculation.

Recently, Jung *et al.*[32] have modified the GHO scheme for the *ab initio* Hartree–Fock calculations introduced by Pu *et al.*[30] Two new techniques have been implemented. The first is the determination of the occupation number linked to the boundary atom. This technique takes into account inhomogeneity in the occupation numbers of the hybrid orbitals of the boundary atoms from different types of MM atoms in such a way that the formal charge condition is fulfilled. This allows the scheme to include the effect of introducing different atoms next to the boundary atom, while the previous scheme sets the occupation to be homogeneous regardless of the atom species. The second technique is a rigorous orthogonalization procedure of auxiliary orbitals for more than two boundary atoms. This technique widens the applicability of this scheme. Furthermore, Jung *et al.* improved the transformation matrix for constructing the hybrid orbitals[33] and also formulated and implemented MP2 gradients for GHO.[34] For the GHO scheme, implementation has only been achieved for the electronic ground state. To the best of our knowledge, the GHO scheme has not been applied to electronic excited states. Treatment of the polarization effect is essential to treat the excited states. Thus, the seamless electrostatic interaction of the GHO scheme that results from the introduction of the hybrid orbitals is expected to be able to treat the excited states.

In this article, we present the implementation and application of the GHO scheme at the second-order approximate coupled cluster singles and doubles (CC2) and CC2 response[35–38] level of theory to treat electronic excited states of biomolecules. Here, all single excitations are included, but double excitations are treated in an approximate way. This lowers the computational cost because the CC2 model scales as N^5, whereas the coupled cluster with singles and doubles (CCSD) model scales as N^6, where N is the number of basis functions. This reduction of computational cost in CC2 allows us to compute larger systems than CCSD. The QM/MM model has already been implemented at this level[39,40] to obtain energies,[41] electric dipoles,[42] and quadrupole moments,[43] together with linear and higher order response properties;[44,45] however, the applications have been limited to molecules in solvents.[46–49] Studies of excited states of biomolecules using EOM-CCSD, such as the work of Valiev and Kowalski,[50] have been published as well.

This article is organized as follows. In the next section, we describe the QM/MM scheme, along with the computational details. We report the calculation results and analysis of the excited state calculation of aromatic amino acids and bacteriorhodopsin (bR) in Sections 3 and 4, respectively, and then conclude the article.

2. Method

In the GHO method, a set of hybridized orbitals $\{\eta_B, \varphi_{3(B-1)+1}, \varphi_{3(B-1)+2}, \varphi_{3(B-1)+3}\}$ is constructed from the *s*- and *p*-valence orbitals at the boundary atom B (see Fig. 1), where η_B denotes the active hybrid orbital, and the other three are the auxiliary orbitals. The transformation matrix $\mathbf{t}_B$ relates the hybrid orbitals with the atomic orbitals according to eqn (1).

$$\begin{pmatrix} \eta_B \\ \varphi_{3(B-1)+1} \\ \varphi_{3(B-1)+2} \\ \varphi_{3(B-1)+3} \end{pmatrix} = \mathbf{t}_B^{\mathrm{T}} \begin{pmatrix} s_B \\ p_{Bx} \\ p_{By} \\ p_{Bz} \end{pmatrix} \tag{1}$$

For details of the transformation matrix, refer to ref. 33. The GHO Hamiltonian (eqn (2)) is expressed as the sum of Hamiltonians for the QM region and QM–MM interactions.

$$\hat{H}_{\mathrm{GHO}} = \hat{H}_{\mathrm{QM}} + \hat{H}_{\mathrm{QM-MM}} \tag{2}$$

The QM–MM interaction,

$$\hat{H}_{\mathrm{QM-MM}} = \sum_{pq} (h_{pq}^{(\mathrm{EE})} + h_{pq}^{(\mathrm{aux})}) \hat{E}_{pq}, \tag{3}$$

consists of the electrostatic embedding term of MM partial charges,

$$h_{pq}^{(\mathrm{EE})} = -\sum_{a=1}^{N_{\mathrm{MM}}} q_a \langle p| \frac{1}{|\mathbf{r} - \mathbf{r}_a|} |q\rangle, \tag{4}$$

and the contribution from auxiliary orbitals (aux),

$$h_{pq}^{(\mathrm{aux})} = \frac{1}{2} \sum_{rs} P_{rs}^{(\mathrm{aux})} [2(pq|rs) - (ps|rq)], \tag{5}$$

where $\hat{E}_{pq}$ is the unitary group generator, N_{MM} and q are the number and partial charges of MM atoms, and $P_{pq}^{(\mathrm{aux})} = \sum_{\alpha}^{\mathrm{aux}} n'_\alpha A_{p\alpha} A_{q\alpha}$ is the density matrix of the auxiliary orbitals with the occupation numbers n'_α (see details in ref. 32 and 33). The Hamiltonian matrix is spanned by the union of the active hybrid orbitals and the basis functions in the pure QM region. Similarly, the Fock operator of the GHO scheme is

$$\hat{F}_{\mathrm{GHO}} = \hat{F}_{\mathrm{QM}} + \hat{H}_{\mathrm{QM-MM}}. \tag{6}$$

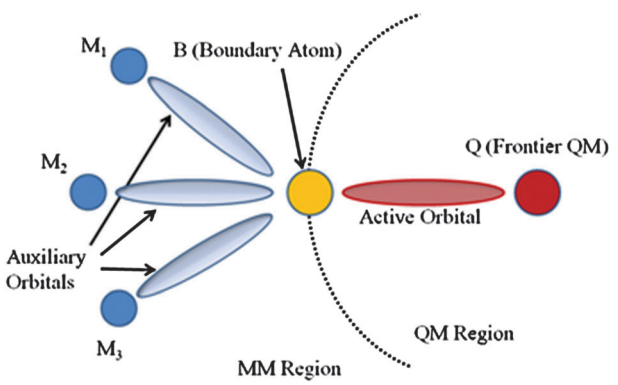

Fig. 1 Schematic representation of the active and auxiliary orbitals in the GHO method.

Note that the present expressions differ from those of our previous papers, in which the contribution of the auxiliary orbitals is included in the density matrix for the Fock assembly. The RHF energy in the GHO scheme is given by

$$E_{\mathrm{GHO}}^{(\mathrm{HF})} = \tfrac{1}{2}\mathrm{tr}[\mathbf{P}_{\mathrm{QM}}(\mathbf{h}_{\mathrm{GHO}}+\mathbf{f}_{\mathrm{GHO}})] + E_{\mathrm{MM}}, \quad (7)$$

where $\mathbf{P}_{\mathrm{QM}}$ is the density matrix of the pure QM region, $\mathbf{h}_{\mathrm{GHO}}$ is the one electronic part of $\hat{H}_{\mathrm{GHO}}$, and E_{MM} is the sum of the classical bonded interaction and nonbonded van der Waals energies including at least one MM atom.

We apply the GHO scheme to CC2 and CC2 linear response calculations, and we briefly summarize our implementation of CC2 and CC2 linear response in the GHO scheme. For details of CC2 and CC2 linear response, refer to ref. 35–38.

First, we show our implementation using the ground state CC2 method. By substituting the Fock operator and the Hamiltonian in the equations for the conventional gas phase CC2 originally derived by Christiansen *et al.*,[36] we express the CC2 energy,

$$E_{\mathrm{GHO}}^{(\mathrm{CC2})} = \langle \Psi_{\mathrm{HF}}|\hat{\bar{H}}_{\mathrm{GHO}}(\hat{T}_1+\hat{T}_2)|\Psi_{\mathrm{HF}}\rangle, \quad (8)$$

and single and double amplitude equations in the GHO scheme as

$$\langle \mu_1|\hat{\bar{H}}_{\mathrm{GHO}} + [\hat{\bar{H}}_{\mathrm{GHO}}, \hat{T}_2]|\Psi_{\mathrm{HF}}\rangle = 0 \quad (9)$$

and

$$\langle \mu_2|\hat{\bar{H}}_{\mathrm{GHO}} + [\hat{F}_{\mathrm{GHO}}, \hat{T}_2]|\Psi_{\mathrm{HF}}\rangle = 0, \quad (10)$$

respectively, where T is the cluster operator, and $\hat{\bar{H}}_{\mathrm{GHO}}$ is the $\hat{T}_1$ transformed Hamiltonian,

$$\hat{\bar{H}}_{\mathrm{GHO}} = \exp(-\hat{T}_1)\hat{H}_{\mathrm{GHO}}\exp(\hat{T}_1). \quad (11)$$

The cluster amplitude equation for the single excitation is identical to the CCSD case, while $\hat{\bar{H}}$ in the commutator is replaced by $\hat{F}$. This leads to a large reduction in computation in CC2.

Next, we show our implementation to calculate excitation energies and oscillator strengths using the CC2 model in the GHO scheme. For excited state calculations, linear response functions of the GHO scheme are necessary. The response theories for exact states and CC models are obtained employing the quasienergy approach of Sasagane *et al.*[51] For details and derivation of the exact state and CC response functions, refer to ref. 35. Substituting the Hamiltonian and the Fock operator of eqn (2) and (4) in the quasienergy Lagrangian of the conventional gas phase CC2, which was derived by Christiansen *et al.*,[35] the CC2 Lagrangian in the GHO scheme can be obtained. The response function is derived as the derivatives of the quasienergy.

The poles and residues of the linear response function determine the excitation energies and transition moments. Thus, the electronic excitation energies ω_f are obtained as eigenvalues of the asymmetric CC2 Jacobian $\mathbf{A}$ as

$$\mathbf{A}R^f = \omega_f R^f, \quad (12)$$

and

$$L^f\mathbf{A} = L^f\omega_f, \quad (13)$$

where R and L correspond to the right- and left-hand side eigenvectors. These eigenvectors are chosen to be biorthogonal such that

$$L^f R^g = \delta_{fg}. \quad (14)$$

From the residues of the linear response function, the transition strengths can be obtained from

$$S_{XY}^{0f} = \tfrac{1}{2}(T_{0f}^X T_{f0}^Y + (T_{0f}^X T_{f0}^Y)^*), \quad (15)$$

where

$$T_{f0}^Y = L^f\xi^Y, \quad (16)$$

and

$$T_{0f}^Y = \eta^Y R^f + \bar{M}^f(\omega_f)\xi^Y. \quad (17)$$

The vector $\bar{M}^f(\omega_f)$ introduced in eqn (15) is defined as

$$\bar{M}^f(\omega_f)(\omega_f\mathbf{1}+\mathbf{A}) + \mathbf{F}R^f = 0. \quad (18)$$

Matrices CC2 Jacobian $\mathbf{A}$ and $\mathbf{F}$, and vectors η^Y and ξ^Y, in the GHO scheme are described as

$$\mathbf{A} = \begin{bmatrix} \langle \mu_1|[\hat{\bar{H}}_{\mathrm{GHO}} + [\hat{\bar{H}}_{\mathrm{GHO}}, \hat{T}_2], \hat{\tau}_{\nu_1}]|\Psi_{\mathrm{HF}}\rangle & \langle \mu_1|[\hat{\bar{H}}_{\mathrm{GHO}}, \hat{\tau}_{\nu_2}]|\Psi_{\mathrm{HF}}\rangle \\ \langle \mu_2|[\hat{\bar{H}}_{\mathrm{GHO}}, \hat{\tau}_{\nu_1}]|\Psi_{\mathrm{HF}}\rangle & \delta_{\mu\nu}\omega_{\mu_2} \end{bmatrix}, \quad (19)$$

$$\mathbf{F} = \begin{bmatrix} (\langle \Psi_{\mathrm{HF}}| + \langle \bar{t}_1| + \langle \bar{t}_2|)[[\hat{\bar{H}}_{\mathrm{GHO}}, \hat{\tau}_{\mu_1}], \hat{\tau}_{\nu_1}]|\Psi_{\mathrm{HF}}\rangle & \langle \bar{t}_1|[[\hat{\bar{H}}_{\mathrm{GHO}}, \hat{\tau}_{\mu_1}], \hat{\tau}_{\nu_2}]|\Psi_{\mathrm{HF}}\rangle \\ \langle \bar{t}_1|[[\hat{\bar{H}}_{\mathrm{GHO}}, \hat{\tau}_{\mu_2}], \hat{\tau}_{\nu_1}]|\Psi_{\mathrm{HF}}\rangle & 0 \end{bmatrix}, \quad (20)$$

$$\eta^Y = ((\langle \Psi_{\mathrm{HF}}| + \langle \bar{t}_1|)[\hat{Y}, \hat{\tau}_{\nu_1}]|\Psi_{\mathrm{HF}}\rangle + \langle \bar{t}_2|[\hat{Y} + [\hat{Y}, \hat{T}_2], \hat{\tau}_{\nu_1}](\langle \bar{t}_1| + \langle \bar{t}_1|[\hat{Y}, \hat{\tau}_{\nu_2}])) \quad (21)$$

and

$$\xi^Y = \begin{pmatrix} \langle \mu_1|\hat{Y} + [\hat{Y}, \hat{T}_2]|\Psi_{\mathrm{HF}}\rangle \\ \langle \mu_2|\hat{Y} + [\hat{Y}, \hat{T}_2]|\Psi_{\mathrm{HF}}\rangle \end{pmatrix}, \quad (22)$$

respectively.

Operator $\hat{Y}$ corresponds to a perturbation oscillating with frequency ω_y, and $\hat{\tau}$ is the excitation operator. In eqn (21) and (22), the zero-order Lagrangian multipliers $\bar{t}$ were introduced. These multipliers,

$$\langle \bar{t}_i| = \sum_{\mu_i} \bar{t}_{\mu_i}\langle \mu_i|, \quad (23)$$

can be obtained from the equation

$$\bar{t}\mathbf{A} + \eta = 0. \quad (24)$$

All CC2 methods were newly implemented in the GELLAN program.[52]

3. Electronic excited states of aromatic amino acids

First, we performed calculations to assess the accuracy of the GHO scheme at the CC2 level. For critical assessments, we need to focus on the errors originating only from our model. To reduce the effect of errors from outside our implementation, test calculations using small molecules are effective because the molecules can be fully optimized at the QM level and our excited state calculations can be compared with full QM results. Furthermore, it is necessary to reduce errors in the QM/MM calculations that arise from classical force fields. Amino acids are the smallest units of proteins on which we can perform full QM calculations and upon which the classical force fields have been studied intensively. Rocha-Rinza *et al.* performed critical assessments of their calculation of the excited state of photoactive yellow protein by comparing full QM and QM/MM.[53] Among the amino acids, the excited states of aromatic amino acids have drawn a great deal of interest. In particular, the fluorescence spectra of tryptophan have been utilized to explore the secondary structures of proteins.[54] Thus, aromatic amino acids are the best systems for our assessment.

We calculated the vertical excitation energies of the first singlet excited states of the aromatic amino acids phenylalanine, tyrosine, and tryptophan in the gas phase. In the gas phase, the amino acids do not form zwitterions, so we created dipeptides by acetylating the *N*-terminus and *N*-methylating the *C*-terminus. The molecular formulae of the studied systems are depicted in Fig. 2. The geometries were optimized at the MP2 level. The vertical excitation energies and oscillator strengths were calculated at the CC2 linear response level of theory. All calculations were performed using the cc-pVDZ basis set.[55] CHARMM27 force fields[56] were used in the GHO calculations, and all of the excited state calculations were performed using the GELLAN program.[53]

For the full QM CC2 calculations on the three amino acids, the calculated excitation energies and oscillator strengths are listed in Table 1. The first singlet excitation energy of phenylalanine is the largest, and the excitation energy of tyrosine is the second largest. The oscillator strength of phenylalanine is small compared with other amino acids. The oscillator strength of tryptophan is the largest among the amino acids. The first singlet excited states of the three molecules have a $\pi \rightarrow \pi^*$ character with a dominant configuration of HOMO (highest occupied molecular orbital) $\rightarrow$ LUMO (lowest unoccupied molecular orbital) transition. The HOMO and LUMO of phenylalanine calculated at full QM CC2 are shown in Fig. 3. Both HOMO and LUMO orbitals are localized at the aromatic ring. Next, we tested the performance of the GHO scheme at the CC2 level for the calculation of excited states. Two different partitionings of the QM and MM regions were carried out. In the first partitioning, denoted by 1-A hereafter, we set the boundary atom at C_β and included the aromatic ring in the QM region. In the second partitioning, denoted by 1-B hereafter, we set the boundary atom at C_α and enlarged the QM region compared with the first partitioning. An example of the partitioning in phenylalanine is illustrated in Fig. 4.

Table 1 The first singlet excitation energies (eV) and oscillator strengths of aromatic peptides

	1-A	1-B	Full QM CC2
Phenylalanine	5.40 (0.0021)	5.28 (0.0001)	5.24 (0.0004)
Tyrosine	5.00 (0.0462)	5.01 (0.0282)	4.97 (0.0293)
Tryptophan	4.86 (0.0347)	4.91 (0.0332)	4.89 (0.0377)

Oscillator strengths are shown in parentheses.

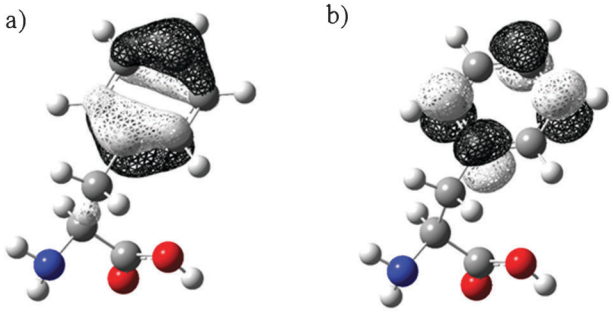

Fig. 3 Molecular orbitals of phenylalanine calculated using a full QM calculation: (a) HOMO and (b) LUMO.

The calculated results for 1-B agree well with the full QM calculation. The first singlet excitation energies using partitioning 1-B were 5.28, 5.01, and 4.91 eV for phenylalanine, tyrosine, and tryptophan, respectively. The excitation energy of phenylalanine was the largest among the amino acids, as in the full quantum CC2 calculation, while the excitation energy of tryptophan was the smallest. The oscillator strength in the calculation of phenylalanine with partitioning 1-B was small compared with other amino acids. The excitation energy agreed with the full QM CC2 results within 0.04 eV. The oscillator strengths showed the same trend as in the full CC2 calculations.

On the other hand, the agreement between our calculations using partitioning 1-A and the full quantum CC2 calculations deteriorated for the excitation energy of phenylalanine and the oscillator strengths of phenylalanine and tyrosine. The excitation energy of the first singlet excited state of phenylalanine was

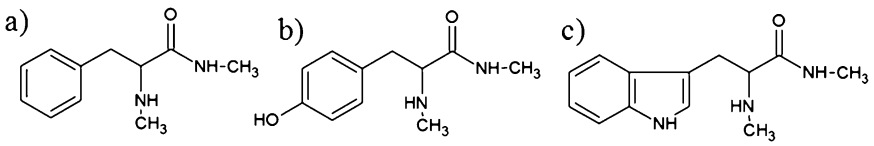

Fig. 2 Molecular formulas of aromatic amino acids: (a) phenylalanine, (b) tyrosine, and (c) tryptophan.

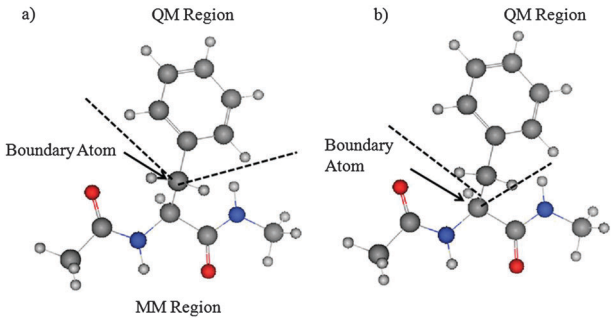

Fig. 4 Two different partitionings for the GHO calculation on phenylalanine: (a) 1-A and (b) 1-B.

overestimated by 0.16 eV using partitioning 1-A compared with the full quantum CC2 calculation. The oscillator strength in calculations using partitioning 1-A was overestimated for tyrosine. The calculation error in 1-A partitioning can be explained by the HOMO π orbitals obtained from the full QM calculation. Although the π electrons are mostly localized on the aromatic ring, there exist nonnegligible influences on the C_β atom. Thus, the $-CH_2-$ group including the C_β atom must be included in the QM region. The calculation using 1-B partitioning suggests that our model is accurate, but at the same time, the results using 1-A warn us to partition the QM and MM regions carefully.

4. Excited state of bacteriorhodopsin

The photoreaction of the retinal chromophore of bR found in *Halobacterium salinarum* establishes an electrochemical gradient across the membrane and activates the proton pump. The isomerization of the chromophore from all-*trans* retinal to the 13-*cis* form occurs in the excited state. Thus, numerous studies of the excited state of bR, both experimental and theoretical, have been reported.[57–65] This well-characterized photoreceptor is a good system for assessment of GHO at the CC2 level because of the extensive data available. Now, we test our scheme with calculations on a *real* biomolecule.

In bR, the chromophore all-*trans* retinal is covalently bonded to the opsin *via* Lys216 and forms a protonated Schiff base (PSB). The counterion of this Schiff base, the side chain of Asp85, is located near the nitrogen atom of the Schiff base. The water molecule, Wat402, plays a role as a bridge and forms hydrogen bonds with the PSB and Asp85. The importance of the quantum effects of the atoms in these regions has been suggested by Fujimoto *et al.*[62]

To elucidate the contribution of these regions to the excitation energy of the first singlet excited state, a QM/MM calculation was performed on bR employing three different partitionings. The QM region of each partitioning is depicted in Fig. 5. The calculation using the first partitioning, denoted by 2-A hereafter, includes the chromophore all-*trans* retinal PSB, the side chain of the Asp85 residue, and the water molecule Wat402, in the QM region. In this partitioning, there are two boundary atoms, the C_δ atom of Lys216 and the C_α atom of Asp85. The calculation using the second partitioning, denoted by 2-B hereafter, removes the side chain of the Asp85 residue from the QM region in the first partitioning. The calculation using the third partitioning, denoted by 2-C hereafter, removes Wat402 from the QM region in the second partitioning. Thus, only the all-*trans* retinal PSB moiety is included in the QM region. We follow the same approach to partitioning the QM region as Fujimoto *et al.*[62] Next, we removed all of the partial charges of the opsin from the third partitioning, denoted by 2-D hereafter, and employed the conventional CC2 method for the retinal Schiff base. Finally, we calculated the gas phase all-*trans* retinal PSB using conventional gas phase CC2.

We calculated the excitation energy and oscillator strength to the first singlet excited state of bR. For comparison, we calculated the isolated all-*trans* retinal PSB in the gas phase as well. Protein data bank (PDB) structure 1C3W[66] was selected for the structure of bR. Hydrogen atoms were added to the structure obtained from the PDB using the CHARMM[67] program. The system was solvated in a TIP3P water box ($70 \times 80 \times 90$ Å), and we added 150 mM NaCl to neutralize the total charge of the system. The hydrogen atoms were relaxed with other atoms fixed using a 50 ps classical molecular dynamics simulation of the NVT ensemble at 310 K preceded by a 1000 step minimization run. The NAMD[68] program was employed for these classical simulations. QM/MM minimization using the adopted basis Newton–Raphson method at the density functional theory B3LYP[69–71] level with the 6-31G**[72] basis set was performed to obtain a stable snapshot structure. To keep the structure as close as possible to the PDB structure, we performed geometry optimization for QM and MM atoms close to the QM region and fixed atoms located

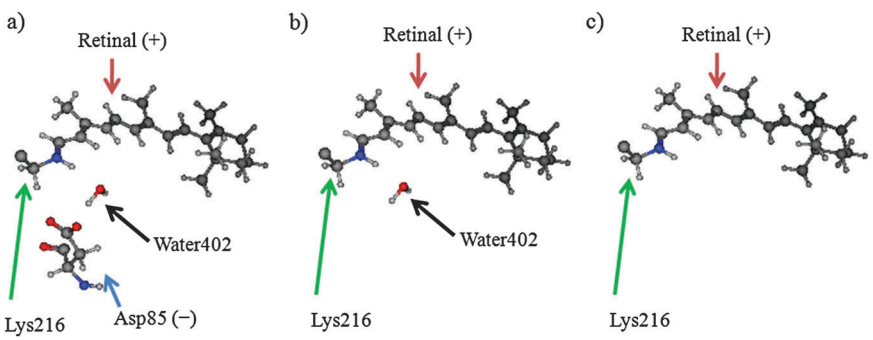

Fig. 5 Three different partitionings for the GHO calculation on bR: (a) partitioning 2-A including the all-*trans* protonated Schiff base, Wat402, and Asp85 in the QM region, (b) partitioning 2-B including the all-*trans* protonated Schiff base and Wat402 in the QM region, and (c) partitioning 2-C including the all-*trans* protonated Schiff base in the QM region.

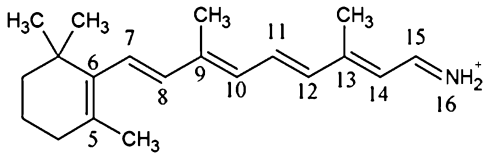

Fig. 6 Molecular formula of the all-*trans* retinal protonated Schiff base.

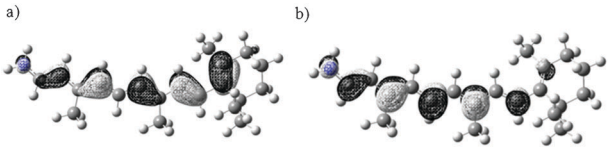

Fig. 8 Molecular orbitals of all-*trans* retinal: (a) HOMO and (b) LUMO.

more than 15 Å away from the all-*trans* retinal PSB center of mass. For the QM/MM simulation, the force fields from CHARMM27 and TIP3P[73] were used for the protein and water molecules, respectively. For retinal, the chromophore of bR, the parameters were taken from previous works on the rhodopsin family.[74–78] The CHARMM and Q-CHEM[79] programs were employed in this minimization. GHO at the CC2 level with the cc-pVDZ basis set was used to calculate the vertical excitation energy and oscillator strength. The excited state calculations were performed using the GELLAN program.

The molecular formula and atom indices are illustrated in Fig. 6. The C–C bond lengths of the all-*trans* retinal PSB were optimized in the gas phase, and the corresponding bond lengths in bR are illustrated in Fig. 7. The gas phase structure has a tendency toward strong bond alternation compared with the bR case, but otherwise the difference between the two structures was found to be small.

The results for the all-*trans* retinal PSB and bR are shown in Table 2. The first singlet excited state for these systems was found to have $\pi \rightarrow \pi^*$ character with a HOMO $\rightarrow$ LUMO dominant transition. The HOMO and LUMO of the all-*trans* retinal PSB are depicted in Fig. 8. First, we obtained the vertical excitation energy of the all-*trans* retinal PSB optimized in the gas phase. The CC2 excitation energy was 2.14 eV in the gas phase. Our theoretical value lies 0.14 eV above the experimental data observed in the gas phase.[80] For bR, the CC2 vertical excitation energy of 2.37 eV using the largest QM partitioning, 2-A, agrees well with the experimental data.[60] The GHO–CC2 excitation energy of bR was 0.19 eV above the experimental value of 2.18 eV. The difference in vertical excitation energies between the gas phase and bR with the 2-A partitioning is 0.23 eV, which agrees with the observed peak difference of 0.18 eV in the absorption spectra. The accuracy of the environmental shift using GHO–CC2 shows that the present GHO–CC2 is very promising.

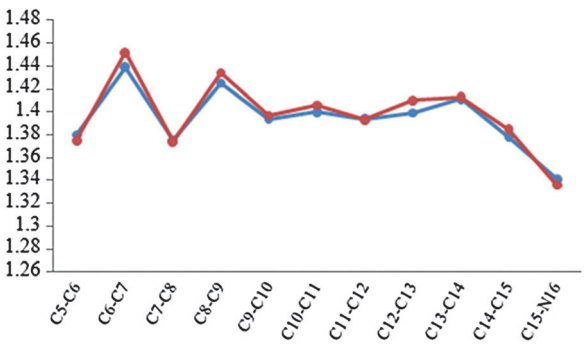

Fig. 7 Bond lengths of the all-*trans* retinal protonated Schiff base and bR in the polyene chain region. The red line indicates the gas-phase optimized all-*trans* retinal protonated Schiff base, and the blue line indicates bR.

Table 2 The first singlet $\pi \rightarrow \pi^*$ excitation energies (eV) and oscillator strengths of bR and the all-*trans* retinal protonated Schiff base

	Calculation	Experiment
2-A (GHO–CC2)	2.37 (2.157)	2.18[a]
2-B (GHO–CC2)	2.40 (2.163)	
2-C (GHO–CC2)	2.39 (2.184)	
2-D (Conventional CC2)	2.12 (2.189)	
All-*trans* retinal PSB (Conventional CC2)	2.14 (1.868)	2.00[b]

Oscillator strengths are shown in parentheses. [a] Ref. 55. [b] Ref. 75.

Finally, we briefly discuss the contribution to the environmental shift. The difference in the shift between 2-A and 2-B is the quantum effect of the side chain in residue Asp85, and that between 2-B and 2-C is the effect of Wat402. The difference between these calculations is very small, and the contribution to the excitation energy of the quantum effect among the counterions and water is negligible. The difference between 2-C and 2-D arises from the electrostatic interaction between the chromophore and the opsin. The difference of 0.25 eV between these partitionings represents a large blue shift in the excitation energy from the gas phase to bR. Finally, the difference between the 2-D and gas phase calculation originates from the geometry distortion in bR, yet the difference was very small and negligible. Thus, we conclude that the main contribution to the excitation energy is the quantum effect of the chromophore and the environmental effect of the opsin.

On the other hand, the oscillator strengths showed different tendencies. The differences in these values among different partitionings showed similar values. The largest difference was found between the gas-phase optimized all-*trans* retinal and partitioning 2-D, which is the distortion effect of the all-*trans* retinal structure. In our calculations, we found that the geometry distortion of the oscillator strength shift between the gas phase retinal PSB and bR had the largest contribution.

5. Conclusion

In this paper, we presented a new implementation of CC2 linear response theory for excited states in conjunction with the GHO QM/MM method, which was implemented into the GELLAN program. The method was applied to the excited states of three aromatic amino acids, phenylalanine, tyrosine, and tryptophan, and also to bR for assessment. Overall, our new implementation showed quantitatively reliable results.

We calculated the excitation energies of aromatic amino acids obtained from our implementation and compared them

with full QM CC2 results. The excitation energies of the first singlet excited states agreed with the full quantum calculation within 0.04 eV. The results also showed that partitioning of the QM and MM regions must be done carefully. The calculated results for bR agreed well with the experimental data. The environmental shifts in the excitation energy for the first singlet excited states are explained utilizing our implementation. The new implementation provides an accurate calculation of the environmental shift between the gas phase all-*trans* retinal PSB and bR. The main contribution to the shift was found to be the electrostatic contribution from the opsin.

In this article, we focused on QM/MM excited state calculations and assessments for our newly implemented method. For accurate QM/MM excited state calculation, we need further investigation of our new implementation. First, the basis set effect of excited state calculation using larger basis sets is necessary. Second, the effect of introducing of polarizable potential, which is introduced in many studies such as in ref. 65, may also be important. Moreover, we used one snapshot structure for excitation energy calculations. Next, we aim to combine this scheme with molecular simulation to introduce the stochastic nature of the excited states in biomolecules. Furthermore, the inclusion of solvation effects in terms of integral equation theories[81,82] is within the scope of this formalism. We plan to proceed to these subjects in the near future.

Acknowledgements

YK was supported by the Program for Improvement of the Research Environment for Young Researchers from the Special Coordination Funds for Promoting Science and Technology (SCF), Japan. HN and TS are grateful to CREST, Japan Science and Technology Agency (JST) for funding this research. YK thanks Dr Suyong Re of RIKEN for valuable discussion.

References

1 A. Warshel and M. J. Levitt, *J. Mol. Biol.*, 1976, **103**, 227.
2 J. Gao and D. G. Truhlar, *Annu. Rev. Phys. Chem.*, 2002, **53**, 467.
3 A. Warshel, *Annu. Rev. Biophys. Biomol. Struct.*, 2003, **32**, 425.
4 D. Das, K. P. Eurenius, E. M. Billings, P. Sherwood, D. C. Chatfield, M. Hodoŝĉek and B. R. Brooks, *J. Chem. Phys.*, 2002, **117**, 10534.
5 P. H. König, M. Hoffman, Th. Frauen and Q. Cui, *J. Phys. Chem. B*, 2005, **109**, 9082.
6 H. M. Senn and W. Thiel, *Angew. Chem., Int. Ed.*, 2009, **48**, 1198.
7 U. C. Singh and P. A. Kollman, *J. Comput. Chem.*, 1986, **7**, 718.
8 K. P. Eurenius, D. C. Chatfield, B. R. Brooks and M. Hodoscek, *Int. J. Quantum Chem.*, 1996, **60**, 1189.
9 M. J. Field, P. A. Bash and M. Karplus, *J. Comput. Chem.*, 1990, **11**, 700.
10 J. Bentzien, R. P. Muller, J. Florián and A. Warshel, *J. Phys. Chem. B*, 1998, **102**, 2293.
11 P. D. Lyne, M. Hodoscek and M. Karplus, *J. Phys. Chem. A*, 1999, **103**, 3462.
12 M. Eichinger, P. Tavan, J. Hutter and M. Parrinello, *J. Chem. Phys.*, 1999, **110**, 10452.
13 F. Maseras and K. Morokuma, *J. Comput. Chem.*, 1995, **16**, 1170.
14 M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber and K. Morokuma, *J. Phys. Chem.*, 1996, **100**, 19357.
15 T. Vreven, K. Morokuma, Ö. Farkas, H. B. Schlegel and M. J. Frisch, *J. Comput. Chem.*, 2003, **24**, 760.
16 M. Eichinger, P. Tavan, J. Hutter and M. Parrinello, *J. Chem. Phys.*, 1999, **110**, 10452.
17 Y. Zhang, T. Lee and W. Yang, *J. Chem. Phys.*, 1999, **110**, 46.
18 Y. Zhang, *J. Chem. Phys.*, 2005, **122**, 24114.
19 Y. Zhang, *Theor. Chem. Acc.*, 2006, **116**, 43.
20 I. Antes and W. Thiel, *J. Phys. Chem. A*, 1999, **103**, 9290.
21 G. A. DiLabio, M. M. Hurley and P. A. Christiansen, *J. Chem. Phys.*, 2002, **116**, 9578.
22 V. Théry, D. Rinaldi, J. Rivail, B. Maigret and G. G. Ferenczy, *J. Comput. Chem.*, 1994, **15**, 269.
23 G. Monard, M. Loos, V. Théry, K. Baka and J. Rivail, *Int. J. Quantum Chem.*, 1996, **58**, 153.
24 X. Assfeld and J. Rivail, *Chem. Phys. Lett.*, 1996, **263**, 100.
25 N. Ferre, X. Assfeld and J. Rivail, *J. Comput. Chem.*, 2002, **23**, 610.
26 D. M. Philipp and R. A. Friesner, *J. Comput. Chem.*, 1999, **20**, 1468.
27 R. B. Murphy, D. M. Philipp and R. A. Friesner, *J. Comput. Chem.*, 2000, **21**, 1442.
28 J. Gao, P. Amara, C. Alhambra and M. J. Field, *J. Phys. Chem. A*, 1998, **102**, 4714.
29 J. Gao, P. Amara, C. Alhambra and M. J. Field, *Theor. Chem. Acc.*, 2000, **104**, 336.
30 J. Pu, J. Gao and D. G. Truhlar, *J. Phys. Chem. A*, 2004, **108**, 632.
31 J. Pu, J. Gao and D. G. Truhlar, *ChemPhysChem*, 2005, **6**, 1853.
32 J. Jung, C. H. Choi, Y. Sugita and S. Ten-no, *J. Chem. Phys.*, 2007, **127**, 204102.
33 J. Jung and S. Ten-no, *Chem. Phys. Lett.*, 2010, **484**, 344.
34 J. Jung, Y. Sugita and S. Ten-no, *J. Chem. Phys.*, 2010, **132**, 084106.
35 O. Christiansen, P. Jorgensen and C. Hättig, *Int. J. Quantum Chem.*, 1998, **68**, 1.
36 O. Christiansen, H. Koch and P. Jørgensen, *Chem. Phys. Lett.*, 1995, **243**, 409.
37 O. Christiansen, A. Halkier, H. Koch, P. Jørgensen and T. Helgaker, *J. Chem. Phys.*, 1998, **108**, 2801.
38 O. Christiansen, H. Koch, P. Jørgensen and T. Helgaker, *Chem. Phys. Lett.*, 1998, **263**, 530.
39 J. Kongsted, A. Osted, K. V. Mikkelsen and O. Christiansen, *Mol. Phys.*, 2002, **100**, 1813.
40 A. Osted, J. Kongsted, K. V. Mikkelsen and O. Christiansen, *Mol. Phys.*, 2003, **101**, 2055.
41 J. Kongsted, A. Osted, K. V. Mikkelsen and O. Christiansen, *J. Phys. Chem. A*, 2003, **107**, 2578.
42 J. Kongsted, A. Osted, K. V. Mikkelsen and O. Christiansen, *Chem. Phys. Lett.*, 2002, **364**, 379.
43 J. Kongsted, A. Osted, K. V. Mikkelsen and O. Christiansen, *J. Chem. Phys.*, 2002, **118**, 1620.
44 J. Kongsted, A. Osted, K. V. Mikkelsen and O. Christiansen, *J. Chem. Phys.*, 2003, **119**, 10519.
45 J. Kongsted, A. Osted, K. V. Mikkelsen and O. Christiansen, *J. Chem. Phys.*, 2004, **120**, 3787.
46 J. Kongsted, A. Osted, K. V. Mikkelsen and O. Christiansen, *THEOCHEM*, 2003, **632**, 207.
47 J. Kongsted, T. B. Pedersen, A. Osted, A. E. Hansen, K. V. Mikkelsen and O. Christiansen, *J. Phys. Chem. A*, 2004, **108**, 3632.
48 J. Kongsted, A. Osted, T. B. Pedersen, K. V. Mikkelsen and O. Christiansen, *J. Phys. Chem. A*, 2004, **108**, 86242.
49 J. Kongsted, A. Osted, P. O. Astrand, K. V. Mikkelsen and O. Christiansen, *J. Chem. Phys.*, 2004, **121**, 8435.
50 M. Valiev and K. Kowalski, *J. Chem. Phys.*, 2006, **125**, 211101.
51 K. Sasagane, F. Aiga and R. Itoh, *J. Chem. Phys.*, 1993, **99**, 3738.
52 GELLAN, a hierarchical quantum chemistry program, Kobe University, 2010.
53 T. Rocha-Rinza, K. Sneskov, O. Christiansen, U. Ryde and J. Kongsted, *Phys. Chem. Chem. Phys.*, 2011, **13**, 1585.
54 M. R. Eftink, *Methods Biochem. Anal.*, 1991, **35**, 127.
55 T. H. Dunning, Jr., *J. Chem. Phys.*, 1989, **90**, 1007.
56 D. MacKerell, Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586.

57 S. Shim, J. Dasgupta and R. A. Mathies, *J. Am. Chem. Soc*., 2009, **131**, 7592.
58 S. Subramaniam and R. Henderson, *Nature*, 2000, **406**, 653.
59 R. A. Mathies, S. W. Lin, J. B. Ames and W. Pollard, *Annu. Rev. Biophys*., 1991, **20**, 491.
60 R. R. Birge and C. F. Zhang, *J. Chem. Phys*., 1990, **92**, 7178.
61 N. Hampp, *Chem. Rev*., 2000, **100**, 1755.
62 K. Fujimoto, S. Hayashi, J. Hasegawa and H. Nakatsuji, *J. Chem. Theory Comput*., 2007, **3**, 605.
63 S. Hayashi and I. Ohmine, *J. Phys. Chem. B*, 2000, **104**, 10678.
64 S. Hayashi, E. Tajkhorshid, E. Pebay-Peyroula, A. Royant, E. M. Landau, J. Navarro and K. Schulten, *J. Phys. Chem. B*, 2001, **105**, 10124.
65 M. Wanko, M. Hoffmann, P. Strodel, A. Koslowski, W. Thiel, F. Neese, T. Frauenheim and M. Elstner, *J. Phys. Chem. B*, 2005, **109**, 3606.
66 H. Luecke, B. Schobert, H. T. Richter, J. P. Cartailler and J. K. Lanyi, *J. Mol. Biol*., 1999, **291**, 899.
67 B. R. Brooks, C. L. Brooks III, A. D. MacKerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus, *J. Comput. Chem*., 2009, **30**, 1545.
68 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comput. Chem*., 2005, **26**, 1781.
69 A. D. Becke, *J. Chem. Phys*., 1993, **98**, 5648.
70 A. D. Becke, *Phys. Rev. A: At., Mol., Opt. Phys*., 1988, **38**, 3098.
71 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter*, 1988, **37**, 785.
72 W. J. Hehre, R. F. Stewart and J. A. Pople, *J. Chem. Phys*., 1969, **51**, 2657.
73 W. L. Jørgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc*., 1996, **118**, 11225.
74 E. Tajkhorshid, B. Paizs and S. Suhai, *J. Phys. Chem*., 1997, **101**, 8021.
75 E. Tajkhorshid and S. Suhai, *J. Phys. Chem*., 1999, **193**, 5581.
76 E. Tajkhorshid, J. Baudry, K. Schulten and S. Suhai, *Biophys. J*., 2000, **78**, 683.
77 M. Nina, B. Roux and J. C. Smith, *Biophys. J*., 1995, **68**, 25.
78 J. Baudry, S. Crouzy, B. Roux and J. C. Smith, *J. Chem. Inf. Comput. Sci*., 1997, **37**, 1018.
79 Y. Shao, L. Fusti-Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. Distasio Jr., R. C. Lochan, T. Wang, G. J. O. Beran, N. A. Besley, J. M. Herbert, C. Y. Lin, T. Van Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock III, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer III, J. Kong, A. I. Krylov, P. M. W. Gill and M. Head-Gordon, *Phys. Chem. Chem. Phys*., 2006, **8**, 3172.
80 I. B. Nielsen, L. Lammich and L. H. Andersen, *Phys. Rev. Lett*., 2006, **96**, 018304.
81 S. Ten-no, F. Hirata and S. Kato, *Chem. Phys. Lett*., 1993, **214**, 391.
82 S. Ten-no, F. Hirata and S. Kato, *J. Chem. Phys*., 1994, **100**, 7443.