JCTC Journal of Chemical Theory and Computation

A Simple One-Body Approach to the Calculation of the First Electronic Absorption Band of Water

Ricardo A. Mata,*,† Hermann Stoll,‡ and B. J. Costa Cabral†,‡

Grupo de Física-Matemática da Universidade de Lisboa, Avenue Professor Gama Pinto 2, 1649-003 Lisboa, Portugal, Institut für Theoretische Chemie, Universität Stuttgart, Pfaffenwaldring 55, D-70569 Stuttgart, Germany, and Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade de Lisboa, 1749-016, Lisboa, Portugal

Received January 27, 2009

Abstract: A one-body decomposition approach for investigating the electronic absorption spectra of molecular systems was proposed and applied to water clusters $(H_2O)_N$ including up to $N =$ 80 water molecules. Two specific aspects of the present implementation are the inclusion of the coupling between excited states and a simplified representation for the *N*-body Coulombic effects. For smaller clusters, the results based on the one-body decomposition scheme are in good agreement with full EOM-CCSD calculations. Two different regimes can be identified in the electronic absorption profile of larger water clusters. The first low-energy regime is dominated by local excitonic states on the cluster surface, whereas the higher-energy excitations associated with the second one are of delocalized nature.

1. Introduction

Electronic properties of water are of fundamental importance for understanding chemical reaction mechanisms and kinetics in solution. However, they are not very well understood.[1,2] Specifically, the relationship between the electronic and topological properties of the hydrogen-bond network[3,4] and the influence of ionic solvation on the electronic properties of bulk or interfacial water[2] deserve further investigation. Therefore, several recent works on the electronic properties of water were reported.[3,5−8] Emphasis was placed on polarization effects in liquid phase,[5] electron binding energies,[5] dynamic polarizability,[8] and electronic absorption spectrum of water.[3,6−8] On the other hand, some relevant studies were also dedicated to water clusters with emphasis on cooperativity effects[9−11] and electronic properties.[12,13] The interest in clusters is motivated by the well-established fact that many of the cooperative effects characterizing the complex hydrogen-bond network of liquid water can also play a major role in determining the structure and electronic properties of clusters. Moreover, there is also theoretical evidence that even for relatively small clusters, the number of hydrogen bonds for the water molecules in the interior of the cluster is quite similar to what is found in condensed phases of water, whereas a more labile network can be observed for water molecules closer to the surface.[14] Consequently, it should be expected that the investigation of the electronic properties of water clusters can also provide relevant information on the relationship between the changing hydrogen-bond network associated with different molecular environments and the electronic properties of water.

In this work we will focus on the prediction of the first electronic absorption band of water clusters. Our main purpose is to adopt a simple, general, and accurate theoretical procedure for calculating the electronic absorption spectra of water clusters. In addition, we have investigated the dependence of the calculated spectra on the cluster size as well as the role played by water molecules at different regions of the aggregates in the absorption process. The present theoretical procedure is based on the many-body decomposition for the energy of a molecular aggregate.[15−29] In agreement with previous many-body decomposition schemes relying on the multilayer fragment molecular orbital

* Corresponding author e-mail: rmata@cii.fc.ul.pt.
† Grupo de Fisica-Matematica da Universidade de Lisboa.
‡ Universität Stuttgart.
‡ Departamento de Química e Bioquímica, Faculdade de Ciencias, Universidade de Lisboa.

10.1021/ct9001653 CCC: $40.75
Published on Web 05/29/2009

method[18,19,27] we are providing further evidence that reliable estimates of the excitation energies of water clusters can be carried out by using a one-body expansion of the cluster energy. Some specific aspects of the present implementation for calculating the electronic spectra of water clusters are the explicit consideration of the coupling between excited states and inclusion of different simplified representations for the *N*-body Coulombic effects. This approach is applied for the calculation of the electronic spectra for optimized structures of water clusters including up to 80 water molecules.

2. Theoretical Methods and Computational Details

2.1. One-Body Treatment of Excitations. Due to the high-order scaling of computational costs in ab initio methods, their use is prohibitive for more than a few water molecules. An alternative approach for the study of larger water clusters is to use a many-body expansion, which potentially reduces the scaling of the method to $\mathcal{O}(N^X)$, where X is the order to which the expansion is truncated. Recently, Chiba et al.[19] have proposed an extension of their FMO method for use in TD-DFT calculations. Here we present a many-body approach for investigating the electronic excitations of large water clusters with an explicit treatment of the coupling between excited states.

We follow the model of Harvey et al.[12] to treat the excitonic states. A matrix Hamiltonian is constructed in the representation of the basis of the uncoupled water molecules. Each of the basis functions represents the system where a single water molecule is excited

$$\Psi_i = \Phi_i^*(\mathbf{r}_i, \mathbf{R}_i) \prod_{j \neq i} \Phi_j(\mathbf{r}_j, \mathbf{R}_j) \tag{1}$$

where i stands for the index of the excited water molecule, Φ_i^* is the excited state of that molecule, Φ_j is the ground state of a neighboring water, $\mathbf{r}_i$ are the electronic coordinates, and $\mathbf{R}_i$ are the nuclear coordinates. Molecules in the excited state will be marked with an asterisk.

In contrast with the model of Harvey et al.,[12] which was based on a semiempirical model for both ground and excited electronic states, we will apply a one-body approximation to the calculation of the ab initio energy of each state. The diagonal elements of the Hamiltonian matrix H are given by

$$H_{ii} = E^{1B}(i^*) = E(i^*) + \sum_{j \neq i}^{N-1} E(j) \tag{2}$$

where $E(...)$ stands for the energy of the excited (or ground state) molecule at a given level of theory. In this work we will use the EOM-CCSD method. The terms $E(i^*)$ are taken from a EOM-CCSD calculation on the excited state of monomer i and the CCSD energies for ground-state energies $E(j)$.

A common approach to increase the convergence of the many-body expansion is to perform each monomer calculation in a point charge field representing the environment molecules. This leads to the approximate inclusion of *N*-body Coulombic effects. The diagonal terms in this case are given by

$$H_{ii} = \bar{E}^{1B}(i^*) = \bar{E}(i^*) + \sum_{j \neq i}^{N-1} \bar{E}(j) - C(i^*) \tag{3}$$

The $\bar{E}(i)$ terms are defined just as before, except that the calculations are performed with an operator added to the one-electron Hamiltonian

$$h_i^{PC} = \sum_{j \neq i} \sum_{\alpha \in j} \frac{q_\alpha}{r_{1\alpha}} \tag{4}$$

where j runs over surrounding monomers, α runs over all atoms in the monomer unit j (in case atom-centered point charges are used), q_α are the charges, and $r_{1\alpha}$ is the distance between an electron and a point charge. The $C(i^*)$ term in eq 3 is a correction energy added to avoid double counting. Since the energy sum in eq 3 runs over all monomers, and each single term already contains the interaction between the monomer and the other units, the interactions are double counted. This correction term will be later discussed.

The off-diagonal elements H_{ij} of the Hamiltonian, which give the coupling between two excited states, are approximated by the interaction of the transition dipoles $\mathbf{d}_{01}$ of the two excitations

$$H_{ij} = \frac{1}{R_{ij}^3}[\mathbf{d}_{01}^i \cdot \mathbf{d}_{01}^j - 3(\mathbf{d}_{01}^i \cdot \mathbf{R}_{ij})(\mathbf{d}_{01}^j \cdot \mathbf{R}_{ij})] \tag{5}$$

where $\mathbf{R}_{ij}$ is the distance vector between the two molecules i and j. The same approach was used by Harvey et al.,[12] where the amplitude of the transition dipole vectors was given by an analytical expression. In our case, the dipole moments are taken from the excited-state calculations performed on each monomer in eq 3 and placed at the center of mass. In the EOM-CCSD case, we use the geometric average of the right and left transition dipole moments

$$\mathbf{d}_{01}^i = \frac{\langle \Phi_i | \hat{\mathbf{d}} | \Phi_i^* \rangle + \langle \Phi_i^* | \hat{\mathbf{d}} | \Phi_i \rangle}{2} \tag{6}$$

where $\hat{\mathbf{d}}$ is the local dipole operator at center i.

Diagonalization of the Hamiltonian matrix gives N energies, one for each state. In order to obtain the excitation energies, one subtracts the ground-state result, which is also obtained under a one-body treatment

$$\bar{E}^{1B}(0) = \sum_i^N \bar{E}(i) - C(0) \tag{7}$$

where $C(0)$ is again a term to avoid double counting of particle interactions due to the point charge field.

Up to this point we have not discussed the point charge field to be used or the form of the correction terms, which are thereof dependent. In our calculations, the charges q_α in eq 4 are taken from the TIP3P model. This may seem at first a crude approximation, since one neglects polarization effects (the point charge field is fixed) and the fact that the density distribution of an excited water molecule is different

from its ground state. The reasons behind this choice are manifold. First, in the TIP3P model, which is an effective pair-potential for liquid water, the charges already include, at least partially, polarization effects in average way. Previous studies have also shown that these charges give reasonable results in expanding the total energy under a many-body approximation.[15,20] Including the full Hartree-Fock potential of neighboring molecules or simply replacing them by these atom-centered point charges has been shown to provide similar accuracy. We expect that this effect is also true when applied in excited-state calculations. Second, by using the same point charge field in all calculations, the correction term for all excited states is the same and equal to the one in the ground state. Under this approximation, including the correction in eq 3 amounts to adding a constant value to the diagonal terms and, therefore, has no effect on the eigenvectors of the Hamiltonian. It also cancels out when computing the difference between the ground and the excited state. In short, if one is only interested in excitation energies, this term can be neglected. It is unclear which approach leads to a higher accuracy, whether neglecting the correction term or including the excitation effect on the charge distribution. The first approximation, however, has the advantage of a nearly linear scaling computational cost. We consider it almost linear, since for larger structures the addition of point charges into the Hamiltonian will become a bottleneck (but only for extremely large structures). It also has a very small prefactor. If the changes in the charge distribution are considered, an extra calculation has to be performed for each monomer, which significantly affects the performance. However, as strongly indicated by some test calculations (see section 3.1), the first approximation already gives highly accurate results, with errors below the ones estimated for the level of theory used as a reference (EOM-CCSD/aug-cc-pVTZ). Also, as previous works on the absorption spectra of water show, the first absorption maximum of liquid water can be well reproduced by treating a single water quantum mechanically.[8] This is due to the localized character of the first excitation.

At this stage, we would like to point out the similarities and differences between our method and other related work. The diagonal terms H_{ii} are calculated as in the FMO method,[18,19,27,28] with the exception that the embedding is simplified, including only fixed point charges to mimic electrostatic effects. In the original method, Coulomb operators taken from converged monomer densities are used. In the work of Hirata et al.,[26] the embedding was based on self-consistent dipole moments. We find that the former approach significantly increases the computational effort. In the case of water, it has been seen to give only marginal improvements for ground-state energies.[15,20] The use of point dipoles per molecule, on the other hand, may be unsuitable to describe the electrostatics of hydrogen-bonded systems. Our main goal, however, is not to review the way the electrostatic embedding is performed but, instead, to expand the applicability of FMO-based formulations to cases where identical chromophores are present. A many-body expansion by itself cannot describe excitonic coupling effects.

The elongation method of Aoki and co-workers[23−25] does account for excitonic coupling. However, their formulation does not include the electrostatic environment in the intrafragment excitation and is also not easily extendable to an arbitrary correlation method. It is based on a CI-type formalism. Also, the coupling elements have to be limited to neighboring fragments due to the computational costs involved in their calculation. However, in a system with identical chromophores, excitonic coupling can lead to mixing of states throughout the system, regardless of its size. The symmetry-adapted-cluster/symmetry-adapted-cluster configuration interaction theory[29] is the most complete formulation, including long-range electrostatic embedding as well as state coupling. However, it is questionable whether it can be computationally efficient in the case of irregular clusters.

All excited-state calculations were performed with the equation-of-motion coupled cluster singles and doubles (EOM-CCSD) method.[30] The basis set chosen was the augmented correlation consistent triple-ζ valence basis set (aug-cc-pVTZ) of Kendall et al.[31] The one-body calculations were performed with in-house Python programs interfaced to the Molpro program package.[32]

2.2. Cluster Structures. One of the main purposes of the present work is to analyze the accuracy of the one-body energy decomposition scheme for predicting the absorption spectra of large water clusters. Since the optimization of very large clusters with ab initio methods is not affordable, we have performed calculations with the polarizable molecular mechanics potential AMOEBA.[33] The choice of this potential was driven by previous works on the properties of water clusters and liquid water that supports the reliability of the model. It should be observed that the force field has been mainly parametrized for reproducing the structure, energetics, and/or vibrational properties of the water monomer, dimer, and small water clusters. The successful applications of the AMOEBA potential to describe both the gas and condensed phases are basically related to the explicit inclusion of polarization effects, which favors transferability to different environments and thermodynamic conditions.

In order to discuss the accuracy of the AMOEBA potential for generating the structure of water clusters, comparison was made with ab initio structures. The use of this force field for the optimization of water clusters has been already tested for a few conformations of clusters ranging up to the hexamer. Further information can be found in ref 33. In this section we reevaluate the potential, taking special care to the evolution of its performance with increasing cluster size.

For $(H_2O)_N$ (with $N = 4-10$), three structures were chosen and optimized at the local second-order Møller−Plesset perturbation (LMP2) level of theory.[34] The basis set used was the Dunning cc-pVTZ basis set.[35] Previous calculations showed that the use of diffuse functions has little impact on the optimized structures. Up to the octamer, the starting cluster geometries were taken from ref 20. The clusters with 9 and 10 water molecules were optimized from starting structures taken from Monte Carlo simulations.[36] The LMP2 structures were then reoptimized with the AMOEBA potential, without any constraints, to an rms gradient per atom of

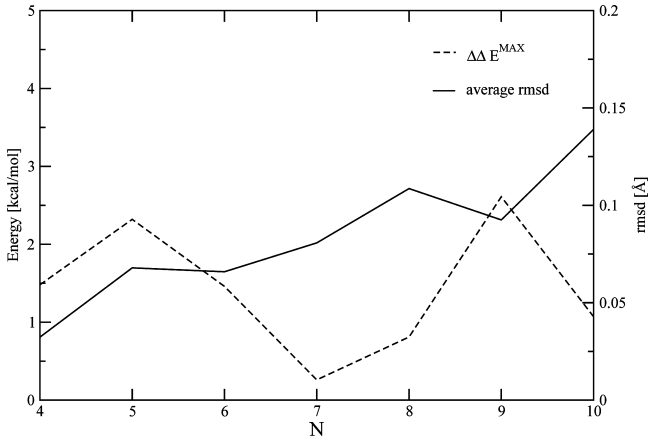

Figure 1. Largest deviations in energy differences of optimized cluster geometries with LMP2/cc-pVTZ and the AMOEBA potential ($\Delta\Delta E^{MAX}$ in kcal/mol) and the average root-mean-square distance (average rmsd in angstroms). Three conformers have been used for each point.

0.0001 kcal/mol/Å. All molecular mechanics calculations were carried out with the TINKER program.[37]

When the AMOEBA and LMP2/cc-pVTZ results are compared, the energetic ordering between different conformers is kept in all but two cases. The maximum deviation in the relative energies as a function of the cluster size is plotted in Figure 1. The results were also compared by superposing the optimized structures, based on mass-weighted coordinates, and computing the rms distance between all atoms. The average values are plotted in the same figure. Both analyses confirm the good performance of the potential in the optimization of small-sized water clusters. The rms distance obtained from superposing the two sets of structures is particularly small. The energy difference is somewhat harder to analyze since there are large fluctuations, but seem to point to a maximum error around 2−3 kcal/mol. All of these results seem to validate the use of the force field in sampling and optimization of large water cluster structures. In all of the following sections, and for consistency, only AMOEBA-optimized structures will be used.

3. Results and Discussion

3.1. Water Dimer. As a first test, calculations were performed on the water dimer, treating each monomer sequentially while representing the other molecule with a different embedding scheme. We neglect the coupling between the excited states since the transition dipole moments are close to perpendicular in the dimer orientation, and it would therefore be close to zero. The embedding schemes used were the same as in a previous work.[20] We compare results for calculations with no embedding potential ("no embedding"), with symmetrical orthogonalization of the monomer orbitals (with level shifting as in eq 7 in ref 20) and a Coulomb potential representing the other monomer ("J(0)"), with the full Hartree−Fock potential ("HF(0)") or the same potential iterated ("HF(1)"), and the calculation with TIP3P charges. The results are shown in Table 1.

The set of results which are in closest agreement to the full calculation are obtained with a simple TIP3P point charge

Table 1. Excitation Energies (in eV) for the AMOEBA-Optimized Water Dimer, Using Different Embedding Schemes

	monomer 1	monomer 2
no embedding	7.587	7.588
J(0)	7.817	8.079
HF(0)	7.734	8.100
HF(1)	7.669	8.152
TIP3P	7.613	8.015
full	7.571	8.018

embedding. This might seem at first surprising, but it should be noted that introduction of the accurate Coulomb or full Hartree−Fock potentials necessitates the use of level-shift operators which essentially freeze the surroundings of a given monomer.[20] The frozen surroundings, in turn, will lead to an upward shift of both ground and excited states, but more so for the excited state since it is less localized (cf. below). The TIP3P results seem to agree well with the full calculation by neglecting the Pauli repulsion of the surroundings and a rather fortunate error cancellation between the lack of exchange contributions and an approximate description of Coulomb effects. The only way to improve on the results with a Fock potential embedding would be to use different frozen surroundings for the ground and the excited state.

3.2. Water Tetramers and Pentamers. In order to validate our approach to the computation of the first absorption band of water, we tested our method in some smaller clusters, where it is still possible to apply the full quantum treatment. We first look at water tetramers, using AMOEBA structures taken from section 2.2.

In Table 2 we report the first four excitation energies, computed in a full EOM-CCSD calculation or with the use of a one-body approximation. In the latter, we considered two cases. In the first set we considered the off-diagonal elements of the excitonic states Hamiltonian to be zero. This amounts to neglecting the coupling between the excited states and simply taking the excitation energies from the one-body calculation. In the second set, we introduced the off-diagonal elements and obtained the energies by diagonalization of the matrix.

Both sets of results from the one-body approach replicate well the EOM-CCSD estimates. However, there is a relevant difference between the two. Neglecting the excited-states coupling, many are degenerate due to the structural symmetry of the cluster. We would like to point out that this error would not be corrected by including higher-order body terms in estimating H_{ii} (this is further discussed in the Conclusions section). It is an effect due to the decoupling of excitations. By the use of the nondiagonal matrix elements, we are able to correctly reproduce the mixing of states which leads to a lift of the degeneracy. The results with diagonalization of the Hamiltonian are on the average closer to the full EOM-CCSD values (the error is halved in cases where the simple one-body approach has a significant deviation). The largest deviation is around 0.07 eV, well below the error estimate of the full method.

In order to confirm that the diagonalization procedure does correctly mix the one-body excitonic states (as defined in eq 1), we have examined the eigenvectors of the solutions

Table 2. First Excitation Energies (in eV) for Three Selected Water Tetramer Clusters

ring 1			ring 2			cage		
full	1B ($H_{ij} = 0$)	1B ($H_{ij} \neq 0$)	full	1B ($H_{ij} = 0$)	1B ($H_{ij} \neq 0$)	full	1B ($H_{ij} = 0$)	1B ($H_{ij} \neq 0$)
8.094	8.103	8.081	7.937	8.040	7.990	7.908	7.933	7.908
8.094	8.103	8.081	7.942	8.040	8.019	7.939	7.933	7.947
8.102	8.103	8.114	8.108	8.057	8.083	8.203	8.264	8.270
8.117	8.103	8.137	8.130	8.057	8.103	8.381	8.353	8.358

Table 3. Relative Weights of Each Monomer for a Given Excitation (Given in Percent)[a]

		relative weight	
exc. energy	molecule	EOM-CCSD	1B ($H_{ij} \neq 0$)
7.908	1	0.0	0.0
	2	9.8	1.6
	3	45.1	49.2
	4	45.1	49.2
7.939	1	6.8	1.4
	2	0.0	0.0
	3	46.6	49.3
	4	46.6	49.3
8.203	1	0.0	0.0
	2	93.0	98.4
	3	3.5	0.8
	4	3.5	0.8
8.381	1	92.8	98.6
	2	2.5	0.0
	3	2.3	0.7
	4	2.3	0.7

[a] The excitation energy (in eV) is taken from the full EOM-CCSD calculation.

for the "cage" tetramer. These allow us to identify the molecules which mostly contribute to a given excited state. We compare these findings to a EOM-CCSD calculation using Pipek−Mezey localized orbitals[38] in the occupied space. By considering the largest coefficients (above 0.1), we calculated their weight in the excitation, labeling each coefficient according to the occupied orbital from which the excitation is made. Results are shown in Table 3.

Table 3 shows a strong agreement between the weights computed with localized orbitals and the ones derived from the transition dipole moments. The first two states are a combination of the individual excitations of molecules 3 and 4 (which in Table 2 are shown to give degenerate one-body excited states 7.933 eV above the ground state). The neighbors only contribute marginally to these states. The other higher-energy states are almost pure one-body states. The differences in the relative weights are relatively small; the largest difference is found in the two first excitations, where our analysis somewhat overestimates the contribution of molecules 3 and 4. Also, although the transition dipole interaction gives no contribution from molecule 2 to the highest-energy excitation, the EOM-CCSD coefficients give a balanced weight for the "spectrator" molecules. But even with these small discrepancies, there is a general agreement. The results for the other tetramers were similar and encourage us to make further use of this analysis to study the local (or delocalized) character of the excitations in larger clusters. Also, if the states are local (with a sparse eigenvector), we may be able to identify the region where the excitation takes place.

As a further test to our approach, we also computed the spectra for the water pentamers. The structures, optimized with the AMOEBA force field as detailed in section 2.2, are shown in Figure 2. One structure of particular interest is the one labeled "4coord", where a central water molecule is hydrogen-bonded to four neighbors (two times as acceptor and as a donor). This is the conformer which most closely resembles the liquid structure of water and therefore will become of greater relevance as one increases the cluster size.

The results for the three conformers of the water pentamer are shown in Table 4. Overall, the results are quite similar to the ones obtained for the tetramers. Except for the highest excitation energy of the "tetra + 1" conformer, which shows a 0.14 eV deviation from the full result, all other values have errors well below 0.10 eV (the average absolute deviation is 0.03 eV). Another fault should, however, be noted. Contrary to the tetramers, the energetic ordering of the excitations (comparing full results to the one-body approximation) is not preserved. In Table 4, we chose to order the excitations according to the weight analysis, and not by the energies. In some cases, where the excitations are near-lying, the ordering had to be inverted. This happens for the third and fourth excitations of the "ring" conformer and for the second and third excitations of "tetra + 1" and "4coord" conformers. Although undesirable, such a problem should be expected due to the small differences between the energies of these conformers and also by the simplified treatment we opted to use, by limiting the results to one-body terms. Inspecting the relative weights of each molecule in the excitations, one concludes that the estimates agree better in the case of localized excitations. In the delocalized cases, our results tend to underestimate delocalization effects. This is particularly so when three molecules have a significant weight on the state. The one-body approximation, although it may identify the most significant contributions, deviates

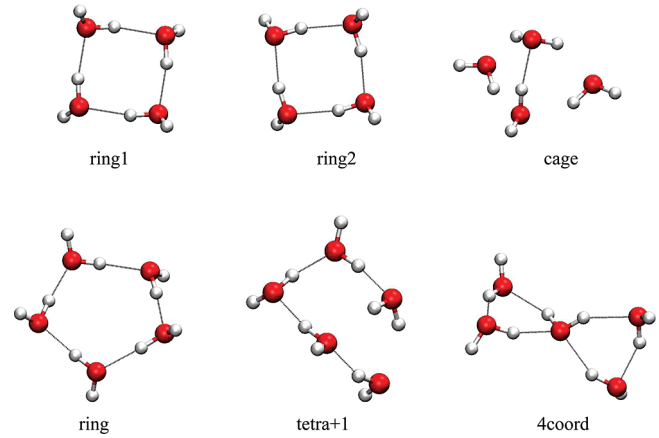

Figure 2. Water tetramer and pentamer structures.

Table 4. Excitation Energies (in eV) and Relative Weights of Each Monomer (Given in Percent), Calculated at the EOM-CCSD/aug-cc-pVTZ Level and Using a 1B Approximation with Coupling Terms

		ring		tetra + 1		4coord	
	molecule	EOM-CCSD	1B ($H_{ij} \neq 0$)	EOM-CCSD	1B ($H_{ij} \neq 0$)	EOM-CCSD	1B ($H_{ij} \neq 0$)
exc. energy(1)		7.915	7.950	7.870	7.925	7.990	8.044
	1	60.5	40.7	0.0	0.4	78.6	48.4
	2	0.0	0.7	0.0	0.4	0.0	1.0
relative weight	3	0.0	0.0	0.0	0.1	3.8	1.1
	4	0.0	0.3	1.5	1.8	0.0	13.8
	5	39.5	58.3	98.5	97.3	12.8	35.7
exc. energy(2)		8.078	8.030	8.089	8.130	8.032	8.065
	1	10.8	40.1	0.0	3.2	2.9	9.7
	2	18.0	1.2	22.7	18.6	1.5	7.4
relative weight	3	51.5	30.4	77.3	77.7	0.0	0.2
	4	0.0	0.1	0.0	0.1	27.7	28.8
	5	19.7	28.2	0.0	0.4	67.9	53.9
exc. energy(3)		8.090	8.078	8.115	8.098	8.054	8.061
	1	4.3	16.1	1.8	0.5	0.0	0.7
	2	19.4	1.0	74.1	76.6	66.1	79.0
relative weight	3	63.1	64.6	24.1	20.4	0.0	0.0
	4	0.0	6.2	0.0	2.4	33.9	20.2
	5	13.2	12.1	0.0	0.1	0.0	0.1
exc. energy(4)		8.096	8.063	8.281	8.288	8.081	8.078
	1	2.6	0.7	96.1	95.7	11.8	40.5
	2	18.1	52.8	2.0	2.4	34.8	12.6
relative weight	3	0.0	1.8	0.0	1.6	0.0	0.1
	4	75.9	44.4	1.9	0.1	40.1	37.2
	5	3.4	0.3	0.0	0.2	13.4	9.6
exc. energy(5)		8.110	8.097	8.437	8.294	8.630	8.591
	1	10.3	2.4	2.4	0.1	0.0	0.7
	2	38.2	44.3	0.0	2.0	0.0	0.0
relative weight	3	0.0	3.1	0.0	0.2	97.0	98.6
	4	37.0	49.1	97.6	95.7	0.0	0.0
	5	14.5	1.1	0.0	2.0	3.0	0.7

somewhat from the reference distribution. This problem could be solved by including higher-order terms in the coupling. For the case in study, and taking into account that the comparison is at best qualitative, the accuracy seems to be adequate, and suffices for an analysis of the excitonic states character.

We now look at the error in the highest-lying excitation of the "tetra + 1" structure. The state is localized, both in the full and one-body results. This indicates that the problem is not due to the coupling between states (the use of transition dipole moments interaction) but, instead, to the description of the one-body state. The excitation takes place at the monomer located in the tetramer ring acting as a donor to the external water molecule. The latter is also oriented toward another molecule in the ring, forming a trimer ring. However, comparison to a regular trimer ring shows that the extra molecule is oriented unfavorably relative to the first monomer. In this particular situation, the use of atom-centered point charges could be insufficient to describe the effect of this neighboring molecule.

3.3. Large Water Clusters ($N \geq 10$). In this section, we apply our approach to the spectra of larger clusters, looking into optimized structures of various sizes. We included water clusters $(H_2O)_N$ with sizes $N = 10$, 20, 30, 40, 60, and 80. In order to sample a significant conformational space, NVT molecular dynamics simulations were carried out for each cluster size at 250 K. The temperature was chosen so that no significant vaporization would occur, while giving enough energy for the system to explore different conformations. The time step used was 1 fs, using a modified Beeman algorithm for integration. Each system was thermalized for 50 ps, and afterward structures were sampled in 10 ps intervals during a 1 ns production run. This sums up to a total of 100 structures for each given size. Some of these simulations had to be repeated (with a different starting structure) since a water molecule would disattach from the cluster. The selected structures were then optimized, at the same level of theory as in the simulation, with use of the AMOEBA potential.[33] The optimization convergence criteria was set to an rms gradient per atom of 0.0001 kcal/mol/Å. We then performed a one-body EOM-CCSD calculation (with excitonic coupling) on each of the selected structures, taking the first N absorption values. In order to represent these in a suitable graphical form, we replaced each value (or peak) by a normalized Gaussian with a variance of 0.0025 eV^{-2}. The results are shown in Figure 3.

Before we discuss Figure 3 in more detail, there is some additional information that should be taken into account. The predicted excitation energy of the AMOEBA-optimized water monomer with EOM-CCSD/aug-cc-pVTZ is 7.6 eV. The absorption peak of gas-phase water is around 7.4 eV.[39,40] For liquid water, the first maximum of the one-photon absorption spectrum is around 8.2 eV,[41] which corresponds to a shift of 0.8 eV. In ice, this value is about 8.7 eV, a shift of 1.3 eV relative to the gas phase. If we consider that our combination of methods would lead, in the limit, to the same

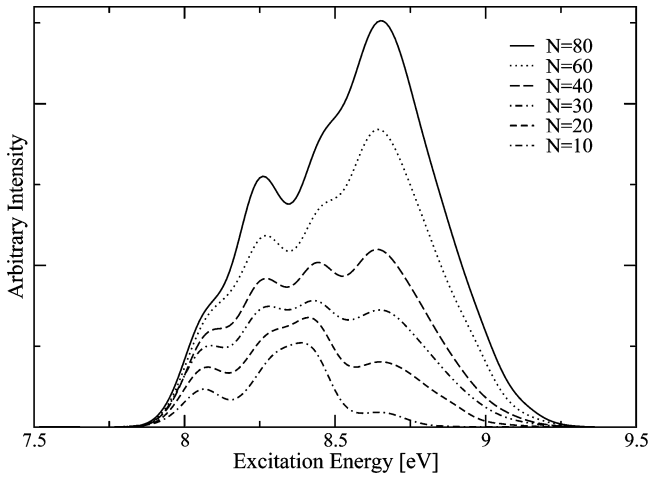

Figure 3. Absorption peaks for water clusters of varying sizes. Each peak has been replaced by a normalized Gaussian ($\sigma^2 = 0.0025$ eV^{-2}).

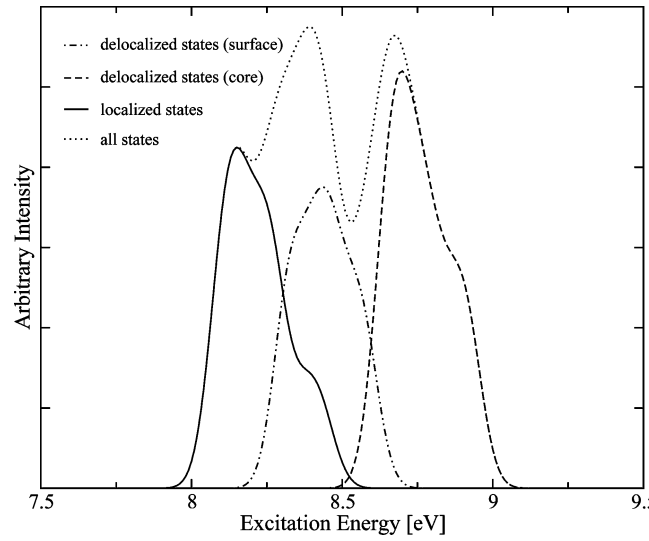

Figure 4. Absorption profile for a $(H_2O)_{40}$ cluster. The spectra is decomposed into localized states (found in the surface) and delocalized states, with or without significant elements from core water molecules. Each peak has been replaced by a normalized Gaussian ($\sigma^2 = 0.0025$ eV^{-2}).

shifts when describing the liquid and the solid water phase, our estimates would be 8.4 and 8.9 eV for the absorption maxima, respectively.

There are some identifiable trends as one increases the size of the cluster. With $N = 10$, one observes a broad range of excitations between 8.0 and 8.5 eV. There is also a small shoulder around 8.7 eV. All of the excitations are above 7.9 eV, which is expected since all of the waters will have at least one hydrogen-bond to a neighboring molecule. As the cluster size increases, the shoulder starts increasing, until it becomes the dominant band with $N = 80$. The value of 8.7 eV is between our estimates for the absorption maximum (in this level of theory) of liquid water and ice. This is also expected since the structures have been optimized, which leads to a rigid hydrogen-bond network (just as in the case of ice). The reason why the two values do not coincide is, however, uncertain. Such a small difference (0.2 eV) could be either due to the fact that EOM-CCSD predicts a different shift than experiment, to our model, to the sampling, or even the molecular mechanics potential used. In any case, the agreement seems reasonable, taking into account that we are studying clusters and the various approximations used.

The second maximum (to the red) seems to converge with increasing cluster size at a value around 8.2−8.3 eV. The first absorption band of water, both in the liquid and in the ice, is known to have no structure, so this splitting must be connected to surface effects. In order to better understand the structure of the band, we have analyzed the eigenvectors of the states computed for a structure of $(H_2O)_{40}$. Directly inspecting the vectors, we have found that the majority of states is significantly delocalized, with more than five significant contributions. We will consider significant eigenvector elements the ones above a 0.1 threshold (thus contributing more than 1%). The only localized states are found on the surface of the cluster. We confirmed this behavior in several other structures. In order to separate the contribution of surface molecules to the total spectra, we divided the excitations into classes. In the first group we include excitations localized exclusively on surface waters (looking only into the significant vector elements as defined above) and with few significant elements (≤ 5). The second group includes excitations which are delocalized, but still exclusively with surface water contributions. The third group is built from the remaining excitations (which are all delocalized and with significant core contributions). The decomposition of the spectra is shown in Figure 4.

The total spectra is divided into two distinguishable bands, one showing two maxima, the other with some structure on the higher excitation energies end. The band of lower excitation energies is the one extracted from the total graph when taking into account only local excitations found on the surface. These are, as expected, somewhat in between the value for the monomer and the one found in bulk water. The second maximum is mainly due to delocalized excitations over the surface. The maxima to the blue side of the spectra (at 8.7 eV) can be attributed to delocalized core excitons. The shoulder, around 8.9 eV, is hard to characterize but is also visible in the individual spectra of larger clusters. It is mainly dominated by the contribution of the water molecules closest to the center of mass, although it does delocalize significantly. The largest maximum represents the dominant character of the excitations present in the cluster. The classification is rather crude but does give some explanation for the clusters' elaborate UV spectra. Perhaps the most reliable (or even physical) classifications would be the localized surface states and delocalized core states. To classify the transitions in between, as done in Figure 4, is somewhat arbitrary but is an interesting way to look at the transition between the two regimes.

4. Conclusions

We have presented a new simple approach to the treatment of the first excitation band of water, which can be generally applied to very large systems. The method, by making use of only one-body embedded excited-state calculations, requires very little computational effort. For a system of N

water molecules, one only needs to perform N monomer calculations and diagonalize an $N \times N$ matrix. The errors are estimated to be small, even below the ones inherent to the high-end method we used as a benchmark in this work (EOM-CCSD/aug-cc-pVTZ). The method was used to compute and analyze the spectra of medium- to large-sized water clusters (up to 80 water molecules). The results show that for sizable structures, two dominant bands will be observed. The lower energy regime will be dominated by excitonic states of local character on the surface of the cluster. The remaining excitations show significant delocalization, with contributions from several different one-body states. However, the delocalization effect on the excitation energies, as estimated by our method, is rather small.

The approach, as formulated in this work, can and should be further improved before application to other systems. Here we give a short account of some of the developments we are currently pursuing. The diagonal elements of the Hamiltonian as well as the ground-state energy should be computed with at least two-body contributions. As previously remarked, the FMO method has been extended to the computation of excitation spectra. However, it is only applicable to excitations in a single molecule which is energetically well separated from the remaining possible excitations in the system. This is mainly due to difficulties in defining the excitations as belonging to a given molecule when computing two-body terms. At the one-body level, this problem is only partly avoided. Although all excitations are strictly localized, close energy-lying states may still swap, raising serious problems in the use of the many-body expansion. This was not a serious issue in this study since the second excitation is about 2 eV higher in energy.

The use of local occupied spaces can alleviate these problems. One possibility is to restrict the coefficients of a given excited state by grouping molecular orbitals according to the monomers where their main centers of charge are found. This would be equivalent to the use of local correlation domains.[30,42] The other possibility is to analyze the excited-state coefficients and derive monomer weights. Both alternatives could allow the inclusion of two-body terms. However, other problems remain to be solved at this level. A balanced virtual space has to be used in order to avoid inconsistencies in the computation of the one- and two-body terms. Preliminary calculations show that this is the main obstacle to a well-defined and convergent expansion. On the other hand, if the two-body terms are added, the correction terms introduced in eqs 3 and 7 are no longer needed. Another advantage lies in the fact that two-body contributions will account for polarization effects on neutral waters due to the excitation of a neighboring molecule. As previously discussed, this is absent in the one-body model.

Other improvements to be made are on the coupling between excited states. The use of the transition dipole seems to be adequate in the case studied, but the implementation of a more general scheme should be considered in future applications.

Acknowledgment. The authors thank Dr. Tatiana Korona for assistance in the use of the EOM-CCSD code, as well as Dr. Jeremy Harvey for some helpful comments on his work. R.A.M. gratefully acknowledges a Research Grant from Fundação para a Ciência e Tecnologia (reference SFRH/BPD/38447/2007).

References

(1) Winter, B.; Weber, R.; Widdra, W.; Dittmar, M.; Faubel, M.; Hertel, I. *J. Phys. Chem. A* **2004**, *108*, 2625.

(2) Winter, B.; Faubel, M. *Chem. Rev.* **2006**, *106*, 1176.

(3) Hermann, A.; Schmidt, W. G.; Schwerdtfeger, P. *Phys. Rev. Lett.* **2008**, *100*, 207403.

(4) Estácio, S. G.; Martiniano, H. F. M. C.; do Couto, P. C.; Cabral, B. J. C. In *Solvation Effects in Molecules and Biomolecules*; Canuto, S., Ed.; Elsevier: Heidelberg, 2008; Chapter 5, p 115.

(5) Millot, C.; Cabral, B. J. C. *Chem. Phys. Lett.* **2008**, *460*, 466.

(6) Brancato, G.; Rega, N.; Barone, V. *Phys. Rev. Lett.* **2008**, *100*, 107401.

(7) Lu, D.; Gygi, F.; Galli, G. *Phys. Rev. Lett.* **2008**, *100*, 147601.

(8) Mata, R. A.; Cabral, B.; Millot, C.; Coutinho, K.; Canuto, S. *J. Chem. Phys.* **2009**, *130*, 014505.

(9) Xantheas, S. S. *J. Chem. Phys.* **1994**, *100*, 7523.

(10) Liu, K.; Cruzan, J. D.; Saykally, R. J. *Science* **1996**, *271*, 929.

(11) Cruzan, J. D.; Braly, L. B.; Liu, K.; Brown, M. G.; Loeser, J. G.; Saykally, R. J. *Science* **1996**, *271*, 59.

(12) Harvey, J. N.; Jung, J. O.; Gerber, R. B. *J. Chem. Phys.* **1998**, *109*, 8747.

(13) Fredj, Y. M. E.; Harvey, J. N.; Gerber, R. B. *J. Phys. Chem. A* **2004**, *108*, 4405.

(14) Galamba, N.; Cabral, B. J. C. *J. Am. Chem. Soc.* **2008**, *130*, 17955.

(15) Dahlke, E. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 46.

(16) Sorkin, A.; Dahlke, E. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 683.

(17) Dahlke, E. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *3*, 46.

(18) Chiba, M.; Fedorov, D. G.; Kitaura, K. *Chem. Phys. Lett.* **2007**, *444*, 346.

(19) Chiba, M.; Fedorov, D. G.; Kitaura, K. *J. Chem. Phys.* **2007**, *127*, 104108.

(20) Mata, R. A.; Stoll, H. *Chem. Phys. Lett.* **2008**, *465*, 136.

(21) Fedorov, D. G.; Kitaura, K. *J. Phys. Chem. A* **2007**, *111*, 6904.

(22) Cui, J.; Liu, H.; Jordan, K. D. *J. Phys. Chem. B* **2006**, *110*, 18872.

(23) Kurihara, Y.; Aoki, Y.; Imamura, A. *J. Chem. Phys.* **1997**, *107*, 3569.

(24) Pomogaev, V.; Gu, F. L.; Pomogaeva, A.; Aoki, Y. *Int. J. Quantum Chem.* **2009**, *109*, 1328.

(25) Pomogaev, V.; Pomogaeva, A.; Aoki, Y. *J. Phys. Chem. A* **2009**, *113*, 1429.

(26) Hirata, S.; Valiev, M.; Dupuis, M.; Xantheas, S. S.; Sugiki, S.; Sekino, H. *Mol. Phys.* **2005**, *103*, 2255.

(27) Chiba, M.; Fedorov, D. G.; Kitaura, K. *J. Comput. Chem.* 200829, 2667.

(28) Fukunaga, H.; Fedorov, D. G.; Chiba, M.; Nii, K.; Kitaura, K. *J. Phys. Chem. A* **2008**, *112*, 10887.

(29) Nakatsuji, H.; Miyahara, T.; Fukuda, R. *J. Chem. Phys.* **2007**, *126*, 084104.

(30) Korona, T.; Werner, H.-J. *J. Chem. Phys.* **2003**, *118*, 3006–3019.

(31) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.

(32) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz, M. et al. MOLPRO, version 2008. 1, a package of ab initio programs, 2008. http://www.molpro.net.

(33) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933.

(34) Schütz, M.; Hetzer, G.; Werner, H.-J. *J. Chem. Phys.* **1999**, *111*, 5691–5705.

(35) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(36) do Couto, P. C.; Cabral, B. J. C.; Canuto, S. *Chem. Phys. Lett.* **2006**, *429*, 129.

(37) Ponder, J. W. In TINKER: Software Tools for Molecular Design, 4.2 ed.; Washington University School of Medicine: Saint Louis, MO, 2003.

(38) Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1989**, *90*, 4916–4926.

(39) Cheng, B.-M.; Chew, E. P.; Liu, C.-P.; Bahou, M.; Lee, Y.-P.; Yung, Y.; Grestell, M. F. *Geophys. Res. Lett.* **1999**, *26*, 3657.

(40) Cai, Z.-L.; Tozer, D. J.; Reimers, J. R. *J. Chem. Phys.* **2000**, *113*, 7084.

(41) Heller, J. M., Jr.; Hamm, R. N.; Birkhoff, R. D.; Painter, L. R. *J. Chem. Phys.* **1974**, *60*, 3483.

(42) Boughton, J. W.; Pulay, P. *J. Comput. Chem.* **1993**, *14*, 736–740.

CT9001653