

Orbital Analysis and Excited-State Calculations in an Energy-Based Fragmentation Method

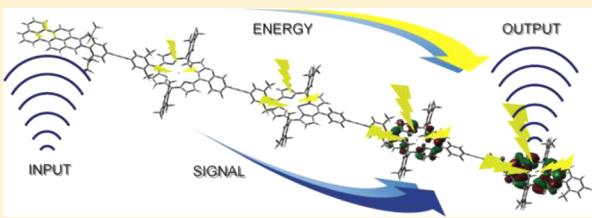
Takashi Tsuchiya, Kushal Shrestha, and Elena Jakubikova*

Department of Chemistry, North Carolina State University, Raleigh, North Carolina 27695, United States



S Supporting Information

ABSTRACT: Covalently bound molecular arrays composed of porphyrins or related pigments have gained a lot of interest as components of artificial light-harvesting systems and molecular photonic devices. The large size of these arrays, however, makes their theoretical investigation employing the *ab initio* or density functional methodologies difficult. Energy-based fragmentation methods (EBF) represent a set of conceptually simple approaches to theoretical investigation of large systems and were therefore chosen as a tool to study these systems. Here a new approach to EBF, EBF-MO, is introduced that enables one to obtain orbitals and orbital energies and to perform population analysis and excited-state calculations of large systems composed of hundreds of atoms. This approach was implemented into a parallel program, JETT, and the benchmark calculations have shown its accuracy and applicability to the ground- and excited-state calculations of systems containing transition metals and extended π -conjugation. EBF-MO was then applied to the density functional theory (DFT) and the time-dependent density functional theory (TDDFT) calculations of ground- and excited-state properties of a porphyrin-based molecular photonic wire composed of 472 atoms and 4265 basis functions at the B3LYP/LANL08,6-31G* level. The TDDFT calculations have revealed the character of the excited states, and the unidirectionality of the excitation energy transfer across the array relevant to its signal transfer function. The computational approaches introduced here have widened the applicability of the *ab initio* and density functional methodologies to calculations of extended systems such as natural and artificial light-harvesting systems and molecular photonic devices.



1. INTRODUCTION

Porphyrins and related complexes (chlorins, bacteriochlorins) play an important role as light-harvesting pigments in plants and bacteria,^{1,2} as well as in artificial systems mimicking the photosynthetic events that occur in nature.^{3–5} Covalently linked arrays of porphyrins have been studied both experimentally^{6–11} and theoretically,^{12–17} gaining interest as simplified models of natural photosynthetic antennas. Porphyrins and porphyrin-based arrays are also key building blocks for molecular photonic devices,¹⁸ and components in various nanoscale electronic devices such as molecular switches^{19,20} and memory devices.^{21–23}

Today, thanks to the advances in the experimental synthetic techniques, extended arrays with 10 or even 100 covalently bound porphyrin subunits can be synthesized.^{7,9,24,25} The large size of these arrays, that can easily reach hundreds or thousands of atoms, represents a significant challenge to electronic structure calculations. While ground- and excited-state properties of a single porphyrin molecule or porphyrin oligomers composed of 100–300 atoms are accessible through conventional *ab initio* and density functional theory (DFT) calculations, calculating properties of a larger collection of pigments is not straightforward. First-principle calculations of their excited-state properties are particularly challenging.

Here we introduce a new approach to ground- and excited-state calculations of porphyrin-based molecular arrays that

stem from energy-based fragmentation (EBF) methodologies. The paper is organized as follows. First, a short overview of various computational approaches to calculations of large systems is given along with our rationale for employing an EBF method in calculations of porphyrin-based arrays. The overview is followed by a detailed description of a computational approach that introduces a new definition of molecular orbitals within the EBF methods (EBF-MO). The calculations on benchmark systems (pentaphenyl, porphyrin triad) are presented next. Finally, an application of the EBF-MO approach to ground- and excited-state calculations of a porphyrin-based molecular photonic wire is described.

Note that, while the development of the EBF-MO method was motivated by our desire to study extended arrays of porphyrins and related pigments, its applicability is more general.

1.1. Overview of Computational Approaches for Large Systems. Several theoretical approaches can be used for calculations of large molecular systems. For example, a number of linear scaling methodologies have been proposed to tackle calculations on large molecules.^{26–29} Those based on Hartree–Fock (HF) or DFT algorithms exploit the “nearsightedness” of an electron in a many-electron system.³⁰ On the

Received: January 10, 2013

Published: June 13, 2013

other hand, linear (or at least reduced) scaling in post HF methodologies is often achieved via the localization of atomic³¹ or molecular orbitals,³² or via the Laplace transform techniques.³³

Multilevel approaches that apply different levels of theory in different regions of a system of interest represent an additional group of methodologies capable of dealing with large molecular systems. For example, quantum methods combined with molecular mechanics^{34,35} (QM/MM) or with quantum mechanics³⁶ (QM/QM') are widely used multilevel methods. These techniques are suitable for problems that lend themselves to treatment of a small region of interest (e.g., an active site of an enzyme) at a high level of accuracy, while including the rest of the system in a more approximate way.³⁷

Methodologies based on frozen-density embedding theory (FDE)³⁸ are other examples of multilevel approaches to calculations of large systems. FDE based methods calculate the electron density for a subsystem of interest using accurate DFT or wave function theory based approaches, while treating the rest of the system in a more approximate way.³⁹

Subsystem DFT based on FDE theory tackles large systems by dividing them into smaller subsystems and by partitioning the total electron density of a system into subsystem densities.^{40,41} The interaction of a subsystem density with the environment (represented by the frozen electron density of all other subsystems) is taken into account by an effective embedding potential, which is in principle exact.^{38,42} Mutual polarization of all subsystem densities can be accounted for through an iterative “freeze-and-thaw” procedure,⁴¹ and the resulting subsystem densities also allow for accurate calculations of excitation energies that are local to individual subsystems.^{39,43,44} Additionally, time dependent DFT (TDDFT) has been reformulated in the framework of FDE and has been used to determine the coupling of local excitations on different subsystems.^{45–47} The method has also been successfully extended to calculations of charge-transfer excitation energies.⁴⁸ FDE methods are especially well suited for treating large aggregates of subsystems that are not covalently linked to each other.

Molecular fragmentation methods provide an additional set of methodologies applicable to calculations on large systems. They can be roughly divided into density matrix (DM)-based and energy-based. DM-based approaches (such as the divide-and-conquer approach (DC),⁴⁹ the elongation method (ELG),^{50,51} the density-fragment interaction method (DFI),⁵² and the adjustable density matrix assembler approach (ADMA)^{53,54}) partition a large system into several subsystems, calculate electron densities of these subsystems separately, and finally reconstruct the total electronic structure by considering interactions between the electron densities of the individual subsystems.

While energy-based fragmentation (EBF) approaches (e.g., molecular tailoring approach (MTA),⁵⁵ fragment molecular orbital method (FMO),⁵⁶ kernel energy method (KEM),⁵⁷ generalized energy-based fragment method (GEBF),⁵⁸ systematic molecular fragmentation (SMF),⁵⁹ and many others⁶⁰) also divide large molecules into a number of subsystems based on a fragmentation scheme, the total energy in this case is determined from the energies of individual subsystems. However, a drawback of the energy-based fragmentation methods is that the orbitals and, thus, the wave function of the parent system are normally not obtained. Therefore,

characteristics of the electronic states that require direct inspection of the molecular orbitals are difficult to study.

Some methodologies, such as FMO or MFCC (molecular fractionation with conjugate caps),⁶¹ have been extended to density based versions.^{62–64} This allows them to access molecular orbitals of the entire system as well as other wavefunction-based properties.⁶⁴ MTA, originally introduced as a DM-based approach, is also capable of obtaining various one-electron properties such as the molecular electrostatic potential or dipole moments.⁵⁵ FMO has also been successfully employed to calculations of excited-state properties of various systems in the framework of the multiconfiguration self-consistent field,⁶⁵ configuration interaction singles (CIS),⁶⁶ CIS with perturbative doubles,⁶⁷ and TDDFT methodologies.^{68,69} While these excited-state calculations are very successful at describing electronic excitations within the same fragment, they are not suitable to calculations of excitations with charge-transfer character.

1.2. EBF as a Tool for Calculations of Porphyrin-Based Arrays.

Our approach to calculations on covalently bound conjugated molecular systems, such as porphyrin arrays, stems from the EBF approaches. The methodology employed in the present work is based on the GEBF, SMF, and MTA methods in particular.^{55,59,70} These methods first divide the parent system into unique fragments from which a set of overlapping super-fragments is created. The total energy of the system is then calculated from the super-fragment energies.

We further propose a new definition of molecular orbitals (MOs) in the framework of the energy based fragmentation method that allows us to obtain molecular orbitals and orbital energies and to perform population analysis for the parent system. In our energy-based fragmentation with molecular orbitals (EBF-MO) approach, molecular orbitals of a parent system are obtained from the Fock matrix constructed from Fock matrices of individual super-fragments. Note that our approach is somewhat related to but distinct from the definition of MOs in the FMO-MO method that proceeds by constructing the density matrix of the parent system out of the density matrices of the individual fragments.⁶⁴ The size of the Fock matrix constructed in the EBF-MO approach can be further reduced by a transformation in the basis of select molecular orbitals calculated for individual super-fragments, and this allows one to obtain a relevant subset of the orbitals for the parent system. This is especially useful in case of very large systems, in which the size of the Fock matrix exceeds the computational ability. The orbitals obtained from the EBF-MO calculations can then be used to construct the wave function or electron density, based on which one can perform subsequent correlation and excited-state calculations, such as TDDFT.

As our target systems contain subunits with extended π -conjugation, the ability of EBF, and EBF-MO in particular, to properly describe these systems is important. EBF approaches have been widely used to obtain various properties of large biological molecules, such as proteins,⁷¹ and their validity for these types of systems is well established. Until recently,^{72,73} these methods have not been applied to extended π -conjugated systems, since the fragments considered are typically quite small (e.g., the size of an amino acid residue), and breaking up the π -conjugation is not advisable. Lately, a new method based on Kekulé structures was proposed for partitioning of aromatic molecules, in which each double bond in a chosen Kekulé structure is selected as a fragment.^{58,73} This approach is,

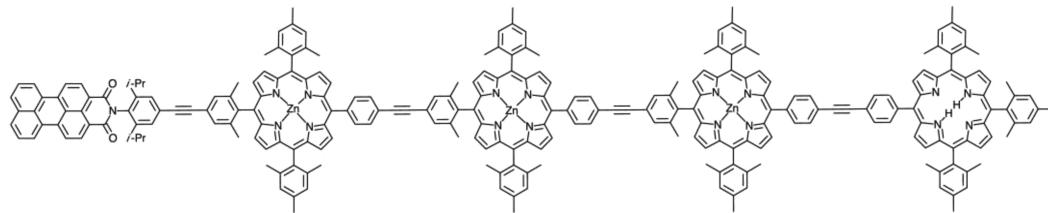


Figure 1. Porphyrin-based molecular wire (**1**) considered in this study.

however, unable to describe compounds containing transition metal atoms.⁵⁸

With the currently available computational power and parallel quantum codes, one does not need to restrict the size of molecular fragments to a small number of atoms. *Ab initio* or DFT calculations on molecules composed of tens of atoms or even a hundred atoms are easily affordable, so the individual fragments and super-fragments for EBF methods can in principle consist of up to a hundred atoms. By keeping the fragments suitably large, the fragmentation of aromatic rings in the molecular-based assemblies can be avoided. This also provides a way to treat large molecular arrays containing transition metal atoms—one only needs to make sure that the transition metal is not separated from its ligands in the fragmentation scheme used to partition the molecule. Therefore, while our implementation of EBF-MO draws on already existing energy-based fragmentation methodologies, the size of fragments considered is significantly larger than in most other applications.

The EBF-MO method developed here was implemented into a parallel program, JETT. The total energy, gradient, and Hessian are made available in JETT; therefore, a geometry optimization and a vibrational frequency analysis can be performed for a very large system using the energy-based fragmentation method. In addition, the ability to obtain molecular orbitals of the parent system is available in JETT. The implementation was tested employing DFT and TDDFT on two benchmark systems: a pentaphenyl molecule and a zinc-porphyrin triad. The EBF-MO approach was further applied to calculations of ground- and excited-state properties of a porphyrin-based molecular photonic wire (**1**) that was previously synthesized and characterized by Lindsey and co-workers (see Figure 1).^{11,74} **1** absorbs visible light and transmits a signal from an input unit (perylene) to the output unit (free-base porphyrin) through excited-state energy transfer with 86% efficiency. The system consists of 472 atoms, and the number of basis functions in our calculations has reached as many as 4265. Our calculations have elucidated the nature of the ground and excited states of **1**, as well as the mechanism of excited-state energy transfer along the wire. DFT and TDDFT calculations performed within the framework of the EBF-MO method presented here demonstrate the applicability of the energy-based fragmentation methods to calculations of molecular architectures relevant to solar energy conversion and molecular electronic applications.

2. ENERGY-BASED FRAGMENTATION METHOD WITH MOLECULAR ORBITALS (EBF-MO)

2.1. Energy, Gradient, and Hessian. Figure 2 illustrates the fragmentation scheme employed in our implementation of the energy-based fragment method using the pentaphenyl molecule as an example. The system is first divided into five fragments, numbered from one through five. A single phenyl

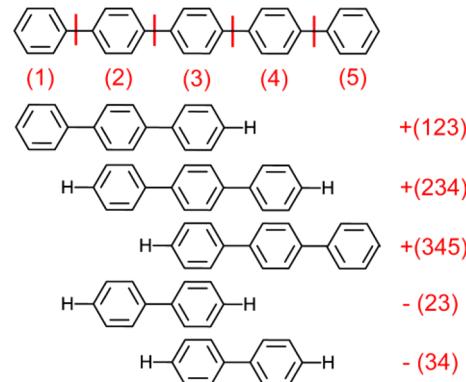


Figure 2. The fragmentation scheme using pentaphenyl molecule as an example.

group constitutes a fragment in the current example. Using three consecutive fragments as constituents, large super-fragments, (123), (234), and (345), are created. To keep the number of fragments in the system consistent, small super-fragments, (23) and (34), are created from overlapping pairs of large super-fragments and subtracted from the system. The unsaturated bonds in each super-fragment are capped with hydrogen atoms. In the calculations presented here, capping hydrogens are placed along the axes of the original C–C bonds and are unconstrained during the structure optimizations. The initial C–H bond length is either set equal to the original C–C bond length (pentaphenyl and zinc-porphyrin triad) or to 1.07 Å (porphyrin array **1**).

The fragmentation scheme described above incorporates interactions between the neighboring fragments into the quantum mechanical calculations explicitly by including nearest neighbors of each fragment in the super-fragment. For example, the interactions between fragments (1), (2), and (3) are included in the quantum mechanical calculation of super-fragment (123). This fragmentation scheme is equivalent with the level-2 systematic molecular fragmentation by Deev and Collins.⁵⁹ However, it is still possible that the effect of neglected interactions, such as the interaction between fragments (1) and (4), is significant. We solved this problem by making a single fragment sufficiently large. For example, in later sections, we treat a porphyrin unit in a molecular array as one fragment. The number of atoms in the resulting super-fragments reaches as many as 176. With such fragment sizes, the effect of neglected interaction can be made negligible.

In this fragmentation scheme, the energy of a parent system, E^{parent} , is calculated as

$$\begin{aligned} E^{\text{parent}} = & E_{(123)}^{\text{super-fragment}} + E_{(234)}^{\text{super-fragment}} + E_{(345)}^{\text{super-fragment}} \\ & - E_{(23)}^{\text{super-fragment}} - E_{(34)}^{\text{super-fragment}} \end{aligned} \quad (1)$$

where $E_{(123)}^{\text{super-fragment}}$, $E_{(234)}^{\text{super-fragment}}$, ... represent the total energies of super-fragments (123), (234), In general, the energy of the parent system is written as a sum of super-fragment energies, $E_i^{\text{super-fragment}}$

$$E^{\text{parent}} = \sum_i C_i E_i^{\text{super-fragment}} \quad (2)$$

where i runs over all the super-fragments and C_i stands for the coefficient for the super-fragment i , which is typically either 1 or -1.

Because the gradient and Hessian are energy derivatives with respect to the atomic displacement, the gradient, $\mathbf{g}^{\text{parent}}$, and Hessian, $\mathbf{H}^{\text{parent}}$, of the parent system can be derived rigorously from the energy of the system, eq 2, to obtain

$$\mathbf{g}^{\text{parent}} = \sum_i C_i \mathbf{g}_i^{\text{super-fragment}} \quad (3)$$

and

$$\mathbf{H}^{\text{parent}} = \sum_i C_i \mathbf{H}_i^{\text{super-fragment}} \quad (4)$$

where $\mathbf{g}_i^{\text{super-fragment}}$ and $\mathbf{H}_i^{\text{super-fragment}}$ are the gradient and Hessian, respectively, of the super-fragment, i , while C_i is the same as in eq 2.

2.2. Molecular Orbitals, Orbital Energies, and Population Analysis. In the Hartree–Fock and Kohn–Sham procedures, orbitals and orbital energies are obtained by solving the Roothaan equation, that is,

$$\mathbf{FC} = \mathbf{SC}\epsilon \quad (5)$$

where \mathbf{F} represents the Fock matrix, \mathbf{S} represents the overlap matrix, \mathbf{C} is the orbital coefficient matrix, and ϵ is the diagonal matrix consisting of orbital energies. Therefore, having the Fock and overlap matrices available makes it possible to obtain orbitals and orbital energies of the parent system. In the framework of the standard energy-based fragmentation methods, orbitals of the parent system are generally not available. Therefore, understanding the electronic nature of the system can be quite difficult. To overcome this deficiency, we simply assume that the Fock and overlap matrices of the parent system, $\mathbf{F}_{\text{parent}}$ and $\mathbf{S}_{\text{parent}}$, respectively, can be constructed from the Fock and overlap matrices of the super-fragments according to eqs 6 and 7, in the same way the energy, gradient, and Hessian are constructed.

$$\mathbf{F}^{\text{parent}} = \sum_i C_i \mathbf{F}_i^{\text{super-fragment}} \quad (6)$$

and

$$\mathbf{S}^{\text{parent}} = \sum_i C_i \mathbf{S}_i^{\text{super-fragment}} \quad (7)$$

where $\mathbf{F}_i^{\text{super-fragment}}$ and $\mathbf{S}_i^{\text{super-fragment}}$ are the Fock matrix and the overlap matrix, respectively, of the super-fragment, i , and the coefficients C_i are the same as in eqs 2, 3, and 4. In terms of the matrix elements of the Fock and overlap matrices, F_{ij} and S_{ij} , respectively, this can be written as

$$F_{ij}^{\text{parent}} = \sum_k C_k F_{ij}^{\text{super-fragment} k} \quad (8)$$

$$F_{ij}^{\text{super-fragment} k} = \begin{cases} \langle \chi_i | \hat{F}_k | \chi_j \rangle & \text{if both } \chi_i \text{ and } \chi_j \text{ belong to} \\ & \text{super-fragment } k \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

$$S_{ij}^{\text{parent}} = \sum_k C_k S_{ij}^{\text{super-fragment} k} \quad (10)$$

$$S_{ij}^{\text{super-fragment} k} = \begin{cases} \langle \chi_i | \chi_j \rangle & \text{if both } \chi_i \text{ and } \chi_j \text{ belong to} \\ & \text{super-fragment } k \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

in which \hat{F}_k is the Fock operator of super-fragment k and χ_i and χ_j are atomic orbital basis functions.

Solution to the Roothaan equation

$$\mathbf{F}^{\text{parent}} \mathbf{C}^{\text{parent}} = \mathbf{S}^{\text{parent}} \mathbf{C}^{\text{parent}} \epsilon^{\text{parent}} \quad (12)$$

defines molecular orbitals and orbital energies in the EBF method (EBF-MO).

The reasoning behind the construction of the overlap matrix (eq 7) is quite simple—the matrix constructed using eqs 10 and 11 includes exact overlap integrals between the basis functions localized on the same super-fragment, while the overlap integrals between the basis functions localized on two different super-fragments are set to zero. This should provide a very good approximation to the overlap matrix of the parent system, provided that the fragments are chosen to be sufficiently large, as one would expect a negligible overlap between basis functions localized on distant centers.

It is less straightforward to evaluate the suitability of our approximation to the Fock matrix of the parent system, since eqs 8 and 9 cannot be derived rigorously. The approximations to the overlap and Fock matrices resulting from the application of eqs 8–11 make it, however, possible to obtain molecular orbitals and orbital energies of the parent system. As shown in later sections, despite the assumptions made, orbitals and orbital energies are reproduced with very good accuracy.

Systems targeted by the fragmentation methods are normally quite large (composed of hundreds of atoms) and require the use of thousands or tens of thousands of basis functions. Consequently, the dimension of the Fock and overlap matrices may become too large to solve eq 12, which includes a matrix diagonalization. Since matrix diagonalization is an order (N^3) procedure, this creates a bottleneck in the methodology. Fortunately, in most cases, a few orbitals are important for the chemical properties or reactivity of the system. Typically, only the highest occupied molecular orbital, the lowest unoccupied molecular orbital (HOMO and LUMO, respectively), and orbitals lying close to them in energy are of particular interest. Therefore, we reduce the dimension of the Fock and overlap matrices by means of a transformation in the basis of selected molecular orbitals calculated for individual super-fragments. The transforming matrix is constructed from the HOMO, LUMO, and other orbitals close to them in energy in each super-fragment. Normally, it is sufficient to include only a small number of orbitals from each super-fragment. This greatly reduces the dimension of matrices to be diagonalized and makes the relevant orbitals of the parent system obtainable. Particularly, for the system composed of N super-fragments, we construct the reduced Fock and overlap matrices using eqs 13–16 and solve the reduced Roothaan eq 17:

Table 1. Total Energies of Pentaphenyl and Zinc-Porphyrin Triad Calculated by the Conventional and Fragmentation Methods^a

compound	calculation	energy (hartree)		ΔE (kcal/mol)
		conventional method	fragmentation method	
pentaphenyl	single point ^b	-1156.433430	-1156.433466	-0.023
	optimization ^c	-1156.455111	-1156.455106	0.003
zinc-porphyrin triad	optimization	-3621.776241	-3621.776228	0.008

^aThe energy differences (ΔE) between the conventional and fragmentation method calculations are also shown. ^bSingle point energies were calculated at nonoptimized structure common for both conventional and EBF method calculations. ^cGeometry optimizations were done independently starting from the same initial structure.

$$\mathbf{F}^{\text{reduced}} = \mathbf{U}^T \mathbf{F}^{\text{parent}} \mathbf{U} \quad (13)$$

and

$$\mathbf{S}^{\text{reduced}} = \mathbf{U}^T \mathbf{S}^{\text{parent}} \mathbf{U} \quad (14)$$

where $\mathbf{F}^{\text{parent}}$ and $\mathbf{S}^{\text{parent}}$ are from eqs 6 and 7, respectively, and

$$\mathbf{U} = \{\Phi_1^{\text{super-fragment}} \Phi_2^{\text{super-fragment}} \dots \Phi_N^{\text{super-fragment}}\} \quad (15)$$

In eq 15,

$$\Phi_i^{\text{super-fragment}} = \{\phi_{\text{HOMO}-k}^i \dots \phi_{\text{HOMO}-1}^i \phi_{\text{HOMO}}^i \phi_{\text{LUMO}}^i \phi_{\text{LUMO}+1}^i \dots \phi_{\text{LUMO}+l}^i\} \quad (16)$$

where ϕ_j^i represents the j th orbital in super-fragment i . As described above, $\Phi_i^{\text{super-fragment}}$ is constructed from a small number of orbitals including the HOMO and LUMO and orbitals energetically close to them in super-fragment i ; i.e., k and l are small integers.

After obtaining the reduced Fock and overlap matrices,

$$\mathbf{F}^{\text{reduced}} \mathbf{C}^{\text{reduced}} = \mathbf{S}^{\text{reduced}} \mathbf{C}^{\text{reduced}} \epsilon^{\text{reduced}} \quad (17)$$

is solved to obtain the reduced set of orbitals and orbital energies. The orbitals obtained, $\mathbf{C}^{\text{reduced}}$, are written in terms of an alternative basis set given by eqs 15 and 16. It is always possible to back-transform the reduced orbital set into the original basis set representation. A similar method was previously applied to a conjugated polymer system combined with a localized molecular orbital method and density functional tight binding (DFTB) theory.⁷⁵

The reduced Fock matrix can also be used to perform electron population analysis. To account for the correct number of electrons, all the occupied orbitals are considered as constituents of the alternative basis set. That is, eq 16 is now substituted by

$$\Phi_i^{\text{super-fragment}} = \{\phi_1^i \phi_2^i \dots \phi_{\text{HOMO}-1}^i \phi_{\text{HOMO}}^i\} \quad (18)$$

From eqs 15, 17, and 18, a set of orbitals including all occupied orbitals is obtained for the system. Using these occupied molecular orbitals, electron populations are calculated.⁷⁶

2.3. Excited-State Calculations. The time-dependent density functional theory (TDDFT) is employed in subsequent excited-state calculations, using Casida's formulation as implemented in Gaussian 09.^{77–80} The excited-state calculations are done for the parent system in its entirety, using the MOs obtained with the EBF-MO approach described above. In the standard implementation, all occupied valence and virtual MOs are used in constructing Casida's equation, which is a matrix equation with the dimension of $2N \times 2N$ (N = the number of occupied MOs times the number of virtual MOs employed in TDDFT calculations). Therefore, as the size of the

parent system increases, the computational cost needed to obtain TDDFT excitation energies becomes prohibitive. To solve this problem, we limit the subspace of excitations to include MOs obtained from eqs 15–17, so only a subset of MOs in a defined energy window near the HOMO–LUMO region is employed in the calculations. To obtain the UV-vis spectrum with sufficient accuracy, a relatively large reduced Fock matrix needs to be constructed. Note that similar approaches (albeit employing different criteria for selection of the relevant MO subset) have been used previously in calculations of electronic spectra of large systems.^{81–86}

3. COMPUTATIONAL METHOD

The EBF-MO methodology described above was implemented in JETT and used in all the calculations reported here. JETT is capable of producing the energy, gradient, Hessian, as well as molecular orbitals and orbital energies of the system using the energy-based fragmentation method. Our implementation contains two layers of parallelism—first, the calculations on individual super-fragments are done in parallel independently of each other; second, the energies, gradients, and Hessians of each super-fragment are obtained using an electronic structure code that also runs in parallel.

Gaussian 09⁸⁷ (G09) was used for all quantum mechanical calculations performed on super-fragments. Total energy, gradient, Hessian, Fock matrix, and overlap matrix for each super-fragment were calculated by G09 and passed to JETT to stitch together and construct the energy, gradient, Hessian, Fock, and overlap matrices of the parent system. The geometry optimization was done using the stitched energy and gradient by means of the energy minimization algorithm of the quasi-Newton–Raphson method.⁸⁸ The Hessian was updated by the Broyden–Fletcher–Goldfarb–Shanno (BFGS) procedure.^{89–93}

Excited-state calculations were performed in G09, using orbitals obtained from the EBF-MO calculations. The orbitals were passed to G09 by the use of checkpoint files. Modifications to the G09 source code were made to bypass the initial SCF calculation, and to create the excitation space for the TDDFT calculation from the reduced orbital set.

The calculations were performed at the DFT and TDDFT levels, using the B3LYP^{94,95} exchange-correlation functional. The LANL08⁹⁶ relativistic effective core potential (ECP) and the associated basis set were used for zinc atom, while the 6-31G* basis set^{97,98} was used for all other atoms.

4. RESULTS AND DISCUSSION

4.1. Benchmark Systems. Geometry Optimization and Vibrational Frequency Analysis. First, we performed single-point calculations for the pentaphenyl molecule by the conventional, all-as-whole, method and by the EBF method,

to confirm the capability of the fragmentation method to reproduce the total energy obtained from the conventional calculation. The molecule was fragmented as shown in Figure 2. The structure of pentaphenyl was not optimized by either method but was kept identical in both calculations. The fragmentation method calculated the total energy to be -1156.433466 hartree, while the conventional method, -1156.433430 hartree, with the error being only -0.023 kcal/mol (Table 1).

Next, we performed geometry optimization using both conventional and fragmentation methods for pentaphenyl and zinc-porphyrin triad. The fragmentation scheme for zinc-porphyrin triad is shown in Figure 3. Initial geometries were

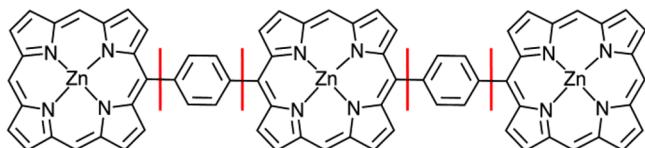


Figure 3. Structure and fragmentation scheme of the zinc-porphyrin triad. The full fragmentation scheme showing all super-fragments can be found in the Supporting Information.

chosen to be common for both methods. The optimized geometries for the pentaphenyl molecule are depicted in Figures S3a and S3b (Supporting Information), while optimized geometries for zinc-porphyrin triad are displayed in Figures S3c and S3d (Supporting Information). Geometry optimizations resulted in virtually identical structures for both methods. The mean absolute error (MAE) calculated for the carbon–carbon distances in pentaphenyl is 0.00004 Å, and the MAE for zinc–nitrogen distances in zinc-porphyrin triad is 0.00004 Å. As shown in Table 1, total energies at the optimized geometries for pentaphenyl are -1156.455106 and -1156.455111 hartree for the fragmentation and conventional methods, respectively, with the error being 0.003 kcal/mol. For zinc-porphyrin triad, total energies are -3621.776228 and -3621.776241 hartree, respectively, with an error of 0.008 kcal/mol. In both systems, errors are negligibly small.

The vibrational frequencies of pentaphenyl calculated by the conventional and EBF methods are shown in Figure S4 (Supporting Information). In the fragmentation approach, the Hessian matrix of the parent system is constructed using eq 4 and subsequently diagonalized in the mass-weighted coordinate system. The EBF method reproduced frequencies obtained from the conventional calculation with good accuracy. For example, the lowest vibrational frequencies are determined to be 14.25 and 13.85 cm $^{-1}$ by the EBF and conventional

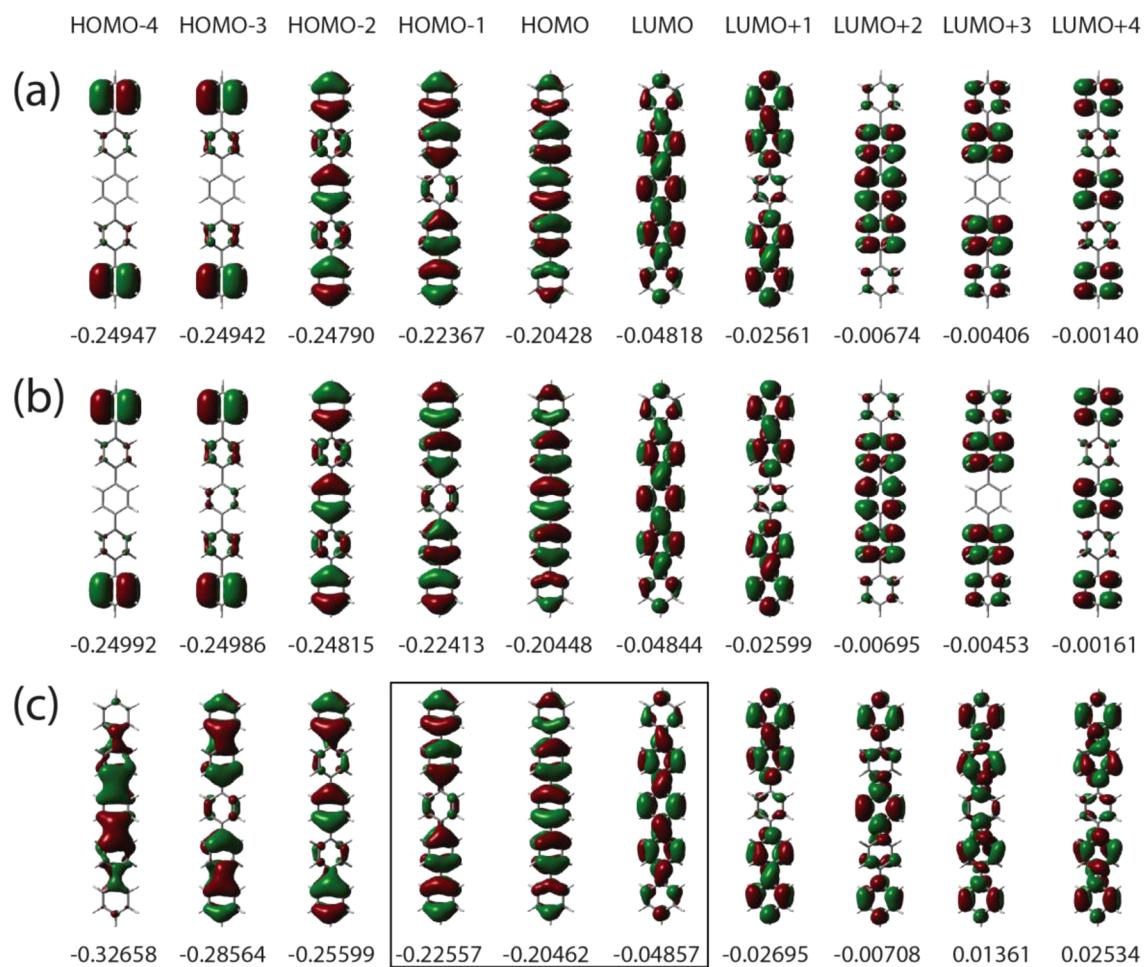


Figure 4. Pentaphenyl orbitals calculated by the conventional method (a), the EBF-MO method with the full Fock matrix (b), and the EBF-MO method with a reduced Fock matrix (c). The box around the MOs obtained from the reduced Fock matrix denotes MOs within the energy window of the chosen super-fragment orbitals (-0.2256 to -0.0317 hartree). The numbers shown below the orbitals represent orbital energies (in hartree).

methods, respectively. The MAE and the maximum error between the frequencies calculated via the fragmentation and conventional methods are 0.46 and 2.86 cm^{-1} , respectively. The IR spectra for zinc-porphyrin triad obtained as a result of the vibrational frequency analysis by the conventional and fragment methods, respectively, are shown in the Supporting Information (Figure S5). Intensities are calculated from the dipole moment and dipole moment derivative of the parent system,⁹⁹ created by stitching together the dipole moments and dipole moment derivatives of the super-fragments, respectively. The two IR spectra are almost identical. The MAE for vibrational frequencies is 0.08 cm^{-1} , and the MAE for intensities is 2.40 km mol^{-1} .

Orbital Analysis. Figure 4 depicts orbitals of the pentaphenyl molecule calculated from Fock matrices constructed by the conventional (Figure 4a) and EBF-MO (Figure 4b) methods. Orbitals and orbital energies are very similar between the two methods. The MAE in orbital energy is 0.0017 hartree. While the validity of stitching the Fock matrices of super-fragments to construct the Fock matrix of the parent system is not obvious, the stitched Fock and overlap matrices produce orbitals and orbital energies of the parent system that are in very good agreement with those calculated by the conventional method.

In Figure 4c, pentaphenyl orbitals calculated from the reduced Fock matrix are shown. We constructed a new, alternative basis set using the HOMO and LUMO of each super-fragment. With this choice, we reduced the dimension of the final Fock matrix from 464×464 to 10×10 . As can be seen in Figure 4c, the reduced Fock matrix reproduced the shapes and orbital energies of the HOMO, LUMO, HOMO-1, and LUMO+1 with very good accuracy, while the other orbitals differ significantly from those calculated using the full Fock matrix (see Figure 4b). Since only the HOMO and LUMO from each super-fragment were included in the alternative basis set, only orbitals with energies close to the orbital energies of HOMOs and LUMOs of the super-fragments can be reproduced accurately. Orbital energies of HOMOs are -0.2134 , -0.2149 , -0.2134 , -0.2256 , and -0.2256 hartree, and those of LUMOs are -0.0400 , -0.0424 , -0.0400 , -0.0317 , and -0.0317 hartree for super-fragments 1–5, respectively. Therefore, only orbitals lying in the energy window between -0.2256 and -0.0317 hartree are meaningful. In the present case, HOMO-1 (the orbital energy is -0.2256 hartree), HOMO (-0.2046 hartree), and LUMO (-0.0486 hartree) fall within this energy window (see Figure S6a, Supporting Information). By using additional orbitals from each super-fragment, the reliable energy range can be widened.

Orbitals of the zinc-porphyrin triad calculated by the conventional and EBF-MO methods are shown in the Supporting Information (Figure S7). Unlike pentaphenyl, this system contains near-degenerate orbitals in both HOMO and LUMO energy regions (see Figure S6b, Supporting Information). We have included all these near-degenerate orbitals in the alternative basis set used to create the reduced Fock matrix.

It is important to note that we experienced several cases in which spurious, unphysical orbitals appeared (particularly near the HOMO and LUMO energy regions), when all of the near-degenerate orbitals were not included in the alternative basis set. We call these unphysical orbitals “ghost orbitals”. To prevent the occurrence of ghost orbitals, all near-degenerate orbitals must be included in the alternative basis. As a result, the number of orbitals used to reduce the Fock matrix is, using the notation of (occupied, virtual), (4, 4), (2, 2), (4, 4), (2, 2), and

(2, 2) for super-fragments 1–5, respectively. There are 12 orbitals out of the resulting 28 orbitals that are within the valid energy range (from around -0.2 to -0.05 eV, see Figure S6b, Supporting Information). These orbitals accurately reproduced the shape and energy of the orbitals calculated by both the conventional method and the EBF-MO method using the full Fock matrix. The MAE in the orbital energies between the reduced and full sets of orbitals is 0.000 hartree (see Figure S7, Supporting Information).

We frequently observed near-zero eigenvalues of the stitched overlap matrix, $\mathbf{S}^{\text{parent}}$, in eq 7. The eigenvalue of the overlap matrix is used in the denominator of the relationship

$$\mathbf{X}^T \mathbf{F} \mathbf{X} = \mathbf{F}' \quad (19)$$

$$\mathbf{X} = \mathbf{S}^{-1/2} \quad (20)$$

where \mathbf{F} and \mathbf{S} are the same as in eq 5 and \mathbf{F}' is the final Fock matrix to be diagonalized. Therefore, a near-zero eigenvalue can incorrectly lower the energy of one of the virtual orbitals. We call these virtual orbitals “invader orbitals”. Ideally, near-zero eigenvalues should not occur because the original basis set is linearly independent. Near-zero eigenvalues observed in this case might be a consequence of approximating the overlap matrix by stitching the super-fragment overlap matrices, or a consequence of a drastic reduction of the dimension when creating the reduced overlap matrix. To prevent the occurrence of the invader orbitals, we add a small constant (0.001 hartree) to the near-zero eigenvalues (i.e., eigenvalues with a magnitude smaller than 5.0×10^{-5}) of the final overlap matrix whenever they appear. After this level-shift correction, HOMO and LUMO are correctly determined as the 14th and 15th orbitals, respectively.

Population Analysis. Table 2 shows selected atomic charges for pentaphenyl calculated from Mulliken electron populations. Atomic charges of the complete set of atoms in pentaphenyl are listed in Table S1 in the Supporting Information. Charges calculated by the EBF-MO method using the full Fock matrix agree well with the results of the conventional calculation. The MAE is 0.0116, which is sufficiently small. Although the signs of the atomic charges on 12C (and 33C, see Table S1, Supporting Information) are different, this error is not significant, as these charges are almost zero.

Charges were also calculated from the reduced Fock matrix. To account for all the electrons in the system correctly, all occupied orbitals of each super-fragment were used to construct an alternative basis set, while excluding the virtual orbitals. In this way, the dimension of the final Fock matrix was reduced from 520×520 to 265×265 , 8 times smaller in the computational cost. Tables 2b and S1b (Supporting Information) show the atomic charges calculated using the reduced Fock matrix. Errors seen in the atomic charges calculated from the full Fock matrix (i.e., opposite signs on 12C and 33C) are not observed in this case. The MAE is also smaller when the reduced Fock matrix is used (0.00693) than for its full Fock matrix counterpart (0.0116).

Further reduction in the size of the Fock matrix is possible by using an alternative basis set employing the occupied orbitals of the large super-fragments only. In this way, the dimension of the final Fock matrix was reduced to 183×183 . The charges calculated with this smaller alternative basis set are shown in Tables 2c and S1c (Supporting Information). The MAE is 0.00543. Interestingly, the smaller size of an alternative basis set consistently results in more accurate calculated charges. Overall,

Table 2. Mulliken Atomic Charges of Pentaphenyl Calculated by the EBF-MO Method with Full Fock Matrix (a), by the EBF-MO Method with Reduced Fock Matrix Using All Super-Fragments (b) and Only Large Super-Fragments (c) as Well as the Conventional Method (d)^a

position	atom	a	b	c	d
1	C	0.05894	0.09118	0.08973	0.08716
2	C	-0.12518	-0.12493	-0.12480	-0.12508
3	C	-0.17163	-0.17048	-0.17067	-0.17137
4	C	-0.17163	-0.17048	-0.17067	-0.17137
5	C	-0.13221	-0.13261	-0.13264	-0.13265
6	C	-0.13225	-0.13260	-0.13263	-0.13265
12	C	-0.02451	0.07305	0.07779	0.09412
13	C	0.10839	0.12291	0.11655	0.09298
14	C	-0.18154	-0.16799	-0.16998	-0.17942
15	C	-0.18155	-0.16801	-0.17000	-0.17942
16	C	-0.17123	-0.20453	-0.19929	-0.17946
17	C	-0.17131	-0.20452	-0.19929	-0.17946
22	C	0.17784	0.11870	0.11219	0.09453
23	C	0.17813	0.11871	0.11219	0.09453
24	C	-0.16724	-0.18703	-0.18546	-0.17967
25	C	-0.16732	-0.18705	-0.18547	-0.17967
26	C	-0.16763	-0.18706	-0.18548	-0.17967
27	C	-0.16761	-0.18706	-0.18548	-0.17967
total ^b		0.00000	0.00017	0.00008	0.00000
MAE ^c		0.01169	0.00693	0.00543	

^aSee Scheme 1 for the atom position labeling scheme. ^bTotal of all atoms. ^cMean absolute error of all atoms.

charges obtained from calculations employing the EBF-MO with a reduced Fock matrix are qualitatively and quantitatively correct, and the results are even better than those obtained from a calculation using the full Fock matrix.

For the zinc-porphyrin triad, the MAEs in atomic charges calculated from the full Fock matrix, reduced Fock matrix, and reduced Fock matrix using large-super fragments only are 0.00315, 0.00224, and 0.00188, respectively, showing the same trend in the accuracy.

Excited States. In the TDDFT calculation, ground-state reference orbitals are of the utmost importance. Therefore, we have first compared calculations using two different sets of orbitals as the reference, namely, a full set of orbitals calculated by the conventional method and a full set of orbitals calculated by the EBF-MO method. The latter is obtained from the stitched full Fock matrix. Table 3 lists the excitation energies and oscillator strengths of the first 10 excitations for the pentaphenyl system. Comparing with the conventional TDDFT calculation, both excitation energies and oscillator strengths were calculated accurately using the orbital set obtained from the EBF-MO. The MAE in the excitation energy is 0.0013 eV. In addition, the character of the excited states was also correctly reproduced. For example, the main component of

Table 3. Excitation Energies and Oscillator Strengths for the First 10 Excited States Calculated by TDDFT for Pentaphenyl^a

excited state	TDDFT with EBF-MO orbitals			
	full ^b	reduced ^{c,d}	reduced ^{e,f}	conventional TDDFT
1	3.8345	3.8993	3.8854	3.8366
2	1.7402	1.6841	1.5166	1.7448
3	4.3940	4.4003	4.4000	4.3912
4	0.0000	0.0000	0.0000	0.0000
5	4.5700	4.6055	4.6055	4.5723
6	0.0000	0.0002	0.0002	0.0000
7	4.6674	4.6889	4.6881	4.6687
8	0.0000	0.0000	0.0000	0.0000
9	4.7076	4.7285	4.7278	4.7089
10	4.8630	4.8775	4.8767	4.8627
	0.0004	0.0006	0.0006	0.0003
	0.0000	0.0000	0.0000	0.0000
	4.8638	4.8786	4.8778	4.8636
	0.0007	0.0008	0.0009	0.0008
	4.9935	5.0036	5.0033	4.9936
	0.0105	0.0136	0.0113	0.0154
	5.0727	5.1281	5.1242	5.0708
	0.4415	0.5116	0.4867	0.4314

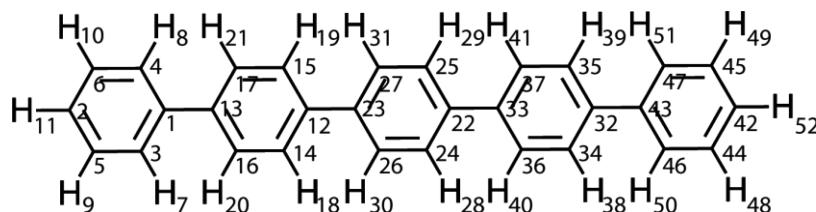
^aThe first entry is the excitation energy (in eV), and the second entry is the oscillator strength. ^bOrbitals obtained from the full Fock matrix are used as a reference. ^cOrbitals obtained from the reduced Fock matrix are used as a reference. ^dAll ghost orbitals are eliminated. ^eGhost orbitals are eliminated using the volume criteria (see text).

the first excited state was calculated to be the HOMO to LUMO excitation by both methods, and the major component of the second excitation was a mixture of the HOMO-1 to LUMO and the HOMO to LUMO+1 excitations. Reference orbitals are shown in Figure 4.

Although the TDDFT combined with the EBF-MO method can calculate excitations correctly and accurately, the full set of reference orbitals is not always available because of the large size of the target system. However, a reduced set of orbitals obtained from the reduced Fock matrix can provide orbitals crucial to the description of relevant excitations. Table 3 lists results of a TDDFT calculation employing a reduced set of reference orbitals (third column in Table 3). Agreement with the conventional method is very satisfactory, with the MAE in the excitation energy equal to 0.0249 eV. Although the order of states 3 and 4 was switched in the reduced calculation, the deviation is insignificant, since those states are almost degenerate.

To obtain a good description of important electronic excitations, reference orbitals were created from a relatively large reduced Fock matrix. 275 super-fragment orbitals with

Scheme 1. Atom Position Labels Used in Table 2



orbital energies between -0.5 and 0.4 hartree were employed to form the reduced Fock matrix. Those values were chosen because reasonable orbital energy gaps were observed around those energy values. Unfortunately, inclusion of all of these orbitals in the alternative basis set led to the presence of a large number of ghost orbital states, which in turn resulted in the presence of spurious (“ghost”) excited states originating from the excitations involving the ghost orbitals. To remedy this problem, one needs to eliminate the ghost orbitals from the reference orbital set. Table 3 (column three) contains the results of the TDDFT calculation in which all ghost orbitals were removed from the reference orbital set, showing an excellent agreement with the conventional calculations as described above. The reference set with ghost orbitals eliminated consists of 43 occupied and 63 virtual orbitals.

Clearly, elimination of ghost orbitals is crucial for reliable TDDFT calculations with a reduced orbital set. Unfortunately, their complete deletion is only possible by comparison of the reduced orbital set with the true orbital set obtained by the conventional calculation, which is impractical. Figure 5 shows

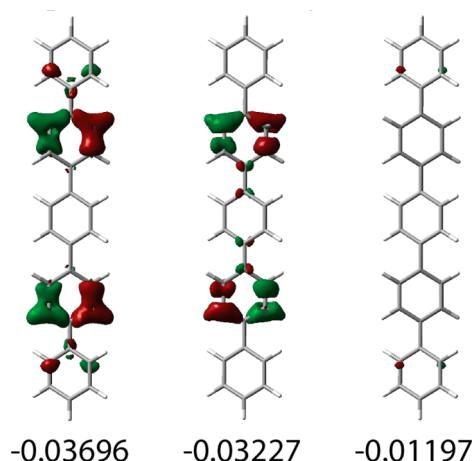


Figure 5. Example ghost orbitals calculated from the reduced Fock matrix for pentaphenyl. Orbitals are shown at the 0.02 isosurface value. The numbers shown below the orbitals represent orbital energies (in hartree).

representative ghost orbitals. Note that the pentaphenyl orbitals shown in Figure 4 were visualized at the same isosurface value. As can be seen in Figures 4 and 5, in comparison with the genuine orbitals of the system under consideration, ghost orbitals are characterized by a significantly smaller volume inside an isosurface with a given value of the electron density. Therefore, we can use the size of the volume with electron

density greater or equal to a selected value as a criterion for ghost orbital elimination. We define a parameter r as

$$r = \frac{V^{\text{orbital}}}{V^{\text{grid}}} \quad (21)$$

where V^{orbital} is the volume within an isosurface with a fixed isovalue and V^{grid} is the total volume spanned by the cube grid. Orbitals with r below a threshold value are regarded as ghost orbitals. In the practical calculation, V^{orbital} is approximated by the number of volume elements within a particular isosurface, while V^{grid} is approximated by the total number of volume elements in the grid. We set the isovalue to $0.02 \text{ e}/\text{\AA}^3$ and chose the threshold value of r ($r = 0.02$) that eliminated all the ghost orbitals in the chemically relevant HOMO and LUMO energy region, while several ghost orbitals were left in the region far from the HOMO and LUMO. By this procedure, 113 ghost orbitals out of a total 117 ghost orbitals were deleted. Finally, 44 occupied and 66 virtual orbitals were employed in the TDDFT calculation. Even with a small number of ghost orbitals remaining in the reference orbital set (in the region far from HOMO and LUMO orbitals), excitation energies and oscillator strengths are in a very good agreement with the results obtained with all the ghost orbitals eliminated (see Table 3, compare columns 3 and 4). Note that the threshold value of r will be different depending on the isovalue used as well as the molecular system under consideration. Therefore, the determination of the proper value of r requires a case-by-case inspection.

4.2. Extended Porphyrin Array. The EBF-MO approach described in previous sections was applied to calculations of a porphyrin-based molecular wire (**1**) shown in Figure 1. **1** consists of 472 atoms. Employing the 6-31G* basis set and LANL08 ECP in the calculations of this system resulted in the use of 4265 basis functions. Conventional calculations for a system of this size would be difficult. The methodologies developed, however, enable us to perform both ground- and excited-state calculations on this system with relative ease.

Ground State. **1** was fragmented into 10 fragments, as shown in Figure 6. Fifteen super-fragments were then constructed, as shown in Figure S2 (Supporting Information), and the total energy, gradient, Hessian, Fock, and overlap matrices were obtained as described earlier. Single point energies, gradients, and Hessian matrices were calculated in parallel for each super-fragment. Both geometry optimization and vibrational frequency analysis were performed. The entire optimization of the geometry took 272 steps. On average, a single optimization step took 38 min of wall time on 2.3 GHz Opteron CPU using 16 cores per super-fragment. The full geometry optimization was completed in 172 h, in about 7 days. Calculation of the analytical Hessian for the largest super-

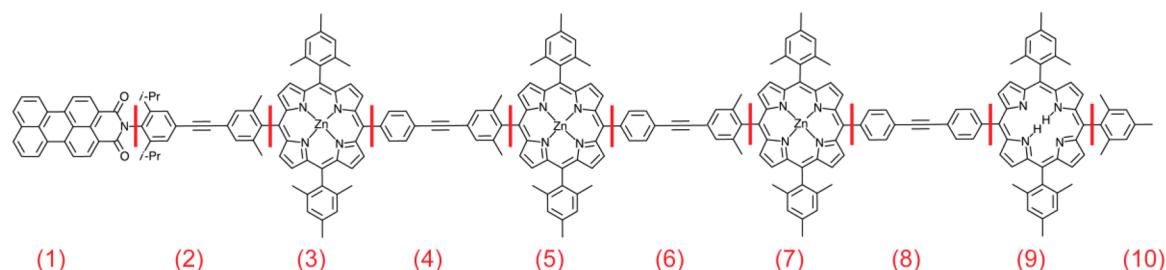


Figure 6. Extended porphyrin array **1**. Red lines indicate the fragmentation scheme. Numbers in parentheses are the fragment numbers.

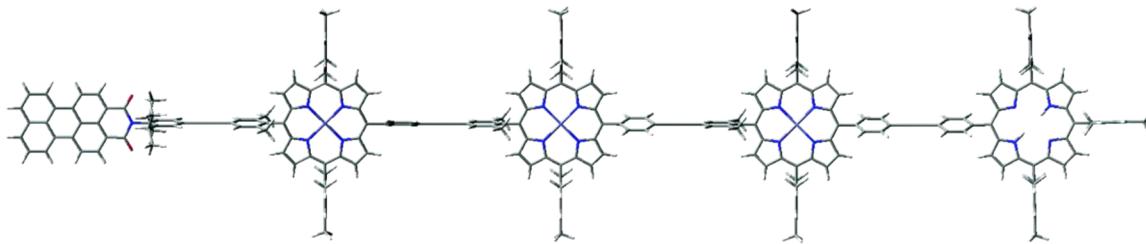


Figure 7. Optimized structure of porphyrin wire 1.

fragment took 30 h to complete on a 0.8 GHz Opteron CPU, using 14 cores per fragment. Note that, once the Hessian matrices are calculated for all the super-fragments, vibrational analysis of the parent system takes an insignificant amount of time.

Figure 7 shows the optimized geometry of **1**. The structure was confirmed to be a minimum by the vibrational frequency analysis. There were two imaginary frequencies found, $0.46i$ and $0.22i\text{ cm}^{-1}$. A close inspection of the normal mode vectors revealed that these imaginary frequencies correspond to very slight and floppy bending of the entire molecule. Considering the large size of the molecular wire and the fact that the method used to obtain the Hessian of the parent system is not exact, these values are insignificant and one can assume that the optimized structure is at a minimum. Figure 8 depicts the IR

site, while the LUMO is localized on the perylene site. Remaining orbitals, with the exception of HOMO-8, are localized on one of the four porphyrin sites. Molecular orbitals in this energy range play a major role in low-energy electronic excitations.

Excited States. TDDFT calculation was performed next, using the ground-state orbitals calculated from the reduced Fock matrix (eq 17). In order to include all relevant molecular orbitals in the excited-state calculation, the reduced Fock matrix was created on the basis of 582 super-fragment orbitals selected from an energy range between -0.25 and 0.1 hartree. Out of the 582 parent orbitals obtained, 483 orbitals lie within the same energy range. Ghost orbitals were eliminated next by the procedure described previously, using the value of $0.02\text{ e}/\text{\AA}^3$ for the isosurface, and setting $r = 0.00125$. The TDDFT calculation was performed on the resulting 122 orbitals (43 occupied and 79 virtual orbitals).

It is well-known that TDDFT tends to calculate the charge transfer states incorrectly; these states are too low in energy.¹² In our calculations employing the B3LYP functional, the charge transfer states appear in the energy range from 570 to 620 nm, although their oscillator strengths are zero. The same charge transfer states are calculated in the energy range from 310 to 340 nm when using the long-range corrected B3LYP functional,¹⁰⁰ which is expected to remedy this failure of ordinary B3LYP. Details of the long-range corrected calculation can be found in the Supporting Information. As can be seen from this example, because our excited-state calculation employs the TDDFT routine as it is implemented in the Gaussian 09 computational package, improvement of the TDDFT in general leads to the improvement of the results of our calculation.

Figure 10 shows the Q-band region of the UV-vis absorption spectrum obtained from the TDDFT calculation. The peaks at 491 nm (A), 530 nm (B-D), 559 nm (E), and 608 nm (F) agree well with the experimental spectrum.¹¹ Table 4 summarizes the excitation energies and character of the electronic excitations for peaks A-F. These electronic excitations correspond to the transitions localized within the perylene subunit (A), three Zn-porphyrin subunits (B-D), and the free-base porphyrin subunit (E, F). The small peak at 468 nm corresponds to a transition delocalized over the neighboring Zn-porphyrin and Fb-porphyrin subunits involving HOMO, HOMO-3, LUMO+1, and LUMO+3 orbitals.

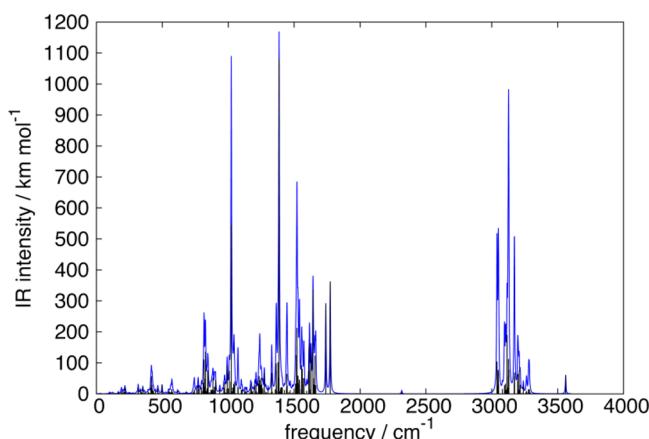


Figure 8. Calculated IR spectrum of the extended porphyrin array (**1**).

spectrum of **1**. The highest peak at 1404 cm^{-1} corresponds to a phenyl ring distortion in the perylene subunit, while the second highest peak at 1023 cm^{-1} corresponds to a pyrrole distortion in the three zinc porphyrin subunits. A pyrrole distortion in the free-base porphyrin subunit was also found at around 1000 cm^{-1} .

Orbitals of the parent system were calculated by means of the EBF-MO method employing the reduced Fock matrix. All super-fragment orbitals within the energy range from -0.22 to -0.07 hartree were used. The remaining orbitals are energetically well separated (see Figure S8 in the Supporting Information) and were therefore not included. The final dimension of the Fock matrix was reduced from 4265×4265 to 92×92 . A total of 22 orbitals of the parent system were found to lie within this energy range (13 occupied and 9 virtual orbitals). Figure 9 displays HOMO-9 through HOMO and the LUMO through LUMO+8 of **1** along with their orbital energies. The HOMO is localized on the free-base porphyrin

TDDFT results provide an insight into the mechanism of the light-induced excited-state energy transfer process in **1**. First, an intense visible-light excitation occurs at 491 nm (2.52 eV) in the perylene subunit. This excitation is then transmitted along three Zn-porphyrin subunits with a slight loss of energy, ending at the free-base porphyrin site at 559 nm (2.22 eV) for the excited state E, or at 608 nm (2.04 eV) for F. The transferred energy can then be released in the process of radiation decay to

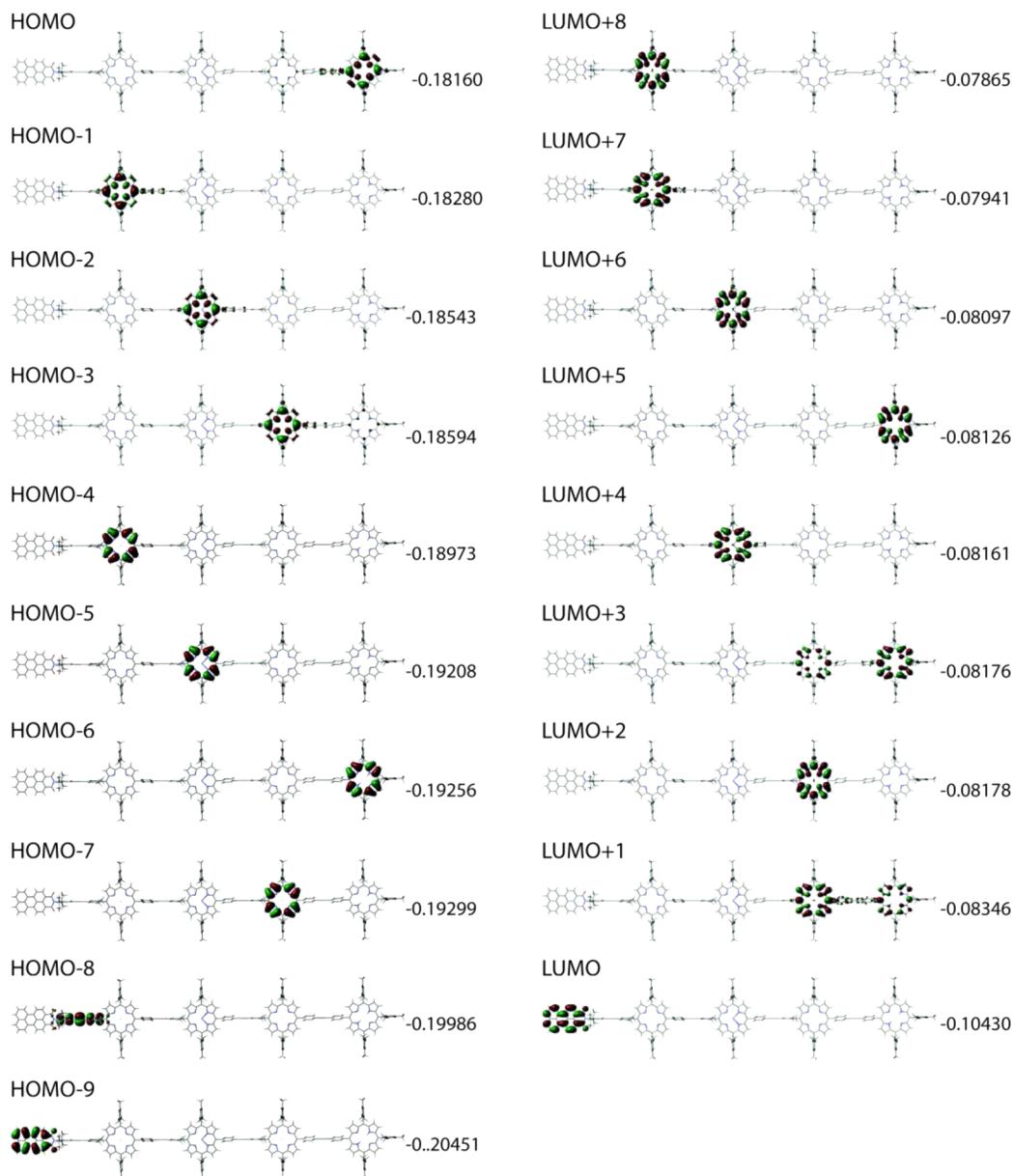


Figure 9. Orbitals of the extended porphyrin array (**1**) calculated by the EBF-MO method with a reduced Fock matrix. The numbers shown are the orbital energies (in hartree).

ground state or passed onto a different component inside a molecular photonic device or an artificial photosynthetic system. The energy transfer occurs with an approximately 12% loss in the excitation energy if the excitation energy transfer stops at E and 19% loss if the transfer reaches the excited state F. By this mechanism, an input signal (491 nm) is sent across the wire to obtain the output signal (559 or 608 nm). Therefore, the excited-state energy transfer is unidirectional. These results agree well with the experimental observations by Ambroise et al.¹¹ and confirm the suitability of **1** as a transmission component of molecular photonic devices.

5. CONCLUSIONS

A new approach to the energy-based fragmentation method, EBF-MO, has been introduced that allows one to obtain molecular orbitals and orbital energies and to perform population analysis and excited-state calculations of extended

systems. Orbitals are obtained from the full or reduced Fock matrices constructed from the Fock matrices of individual super-fragments. The EBF-MO approach was implemented into a parallel program, JETT, that interfaces with Gaussian 09 to perform the electronic structure calculations.

The approach was tested on two benchmark systems, pentaphenyl and a zinc porphyrin trimer, demonstrating its capability to obtain qualitatively and quantitatively accurate single point energies, optimized structures, IR spectra, Mulliken electron populations, and molecular orbitals of the parent systems. Molecular orbitals obtained from the EBF-MO calculations were then successfully used in the excited-state calculations employing TDDFT. This methodology has widened the applicability of conceptually simple and easily implementable energy-based fragmentation methods to orbital analysis that has been thus far the domain of more complex density-based approaches.

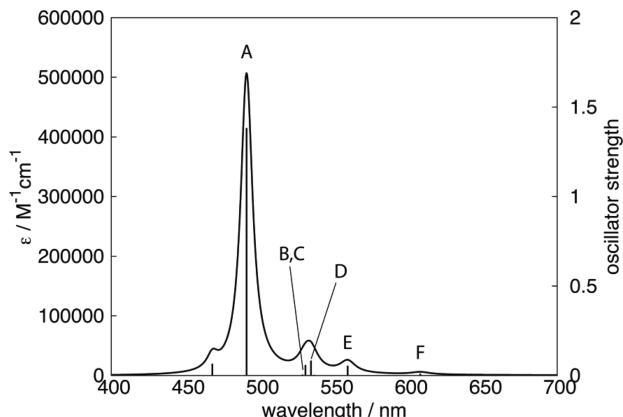


Figure 10. The Q-band region of the UV-vis spectrum of the extended porphyrin array (1) calculated by TDDFT. See Figure 1 in ref 11 for the experimental spectra.

Table 4. Energies, Wavelengths, and Significant Components of Excitations for the Extended Porphyrin Array (1)

excited state ^a	energy (eV)	wavelength (nm)	excitation ^{b,c}
A	2.52	491	HOMO-9 to LUMO
B	2.35	529	HOMO-2 to LUMO+4 HOMO-5 to LUMO+6
C	2.34	531	HOMO-1 to LUMO+7 HOMO-4 to LUMO+8
D	2.32	534	HOMO-7 to LUMO+2 HOMO-3 to LUMO+1
E	2.22	559	HOMO-6 to LUMO+5 HOMO to LUMO+5 HOMO to LUMO+3
F	2.04	608	HOMO to LUMO+5

^aRefer to Figure 10. ^bRefer to Figure 9 for the orbitals. ^cOnly components with the absolute value of the coefficient larger than 0.3 are shown.

The EBF-MO method was also applied to ground- and excited-state calculations of a covalently linked porphyrin array composed of perylene and four porphyrin subunits, designed to serve as a molecular photonic wire.¹¹ The system consists of 472 atoms. The calculations were performed at the B3LYP/LANL8,6-31G* level of theory, leading to the use of 4265 basis functions. For the ground state, geometry optimization and vibrational analysis were performed and chemically important molecular orbitals were obtained. HOMO, LUMO, and orbitals close to them in energy were found to be localized on individual subunits (i.e., perylene, Zn porphyrins, free-base porphyrin) of this system, confirming a weak nature of the coupling among the individual pigments. The TDDFT calculations have found a number of excited states in the visible region of the absorption spectrum, again localized on the individual subunits. The most intense and highest energy excitation in the 450–700 nm region was found to be localized on the perylene subunit, while the lowest energy excitation arises from the free-base porphyrin subunit. These findings have revealed a mechanism for signal transfer across the molecular wire by means of an excited-state energy transfer from the input end (perylene) to the output end (free-base porphyrin) via the transmission elements (Zn porphyrins).

The results demonstrate applicability of the EBF-MO to the systems with extended π -conjugation as well as to the systems containing transition metal atoms. The approaches introduced open up new possibilities for the *ab initio* and DFT calculations of large molecular arrays important for artificial photosynthetic assemblies and molecular photonic devices.

ASSOCIATED CONTENT

S Supporting Information

Table of Mulliken atomic charges for the pentaphenyl array, optimized structures of pentaphenyl and zinc-porphyrin triad, results of vibrational frequency analyses for pentaphenyl and zinc-porphyrin triad, orbital energy level diagrams for pentaphenyl, zinc-porphyrin triad, and porphyrin array (1). Full fragmentation scheme for zinc-porphyrin triad and porphyrin arrays (1). Table of wall times comparing the conventional and EBF-MO calculations for calculations of benchmark systems. Cartesian coordinates (in angstroms) of the optimized structures of pentaphenyl, zinc-porphyrin triad, and porphyrin array (1). This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: ejakubi@ncsu.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was supported by the North Carolina State University (Department of Chemistry). This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number TG-CHE120019.

REFERENCES

- (1) Hu, X.; Ritz, T.; Damjanovic, A.; Schulten, K. *J. Phys. Chem. B* **1997**, *101*, 3854.
- (2) Blankenship, R. E. *Molecular Mechanisms of Photosynthesis*; Blackwell Science Ltd: Oxford, U.K., 2002.
- (3) Gust, D.; Moore, T. A. *Science* **1989**, *244*, 35.
- (4) Imahori, H. *J. Phys. Chem. B* **2004**, *108*, 6130.
- (5) Nakamura, Y.; Aratani, N.; Osuka, A. *Chem. Soc. Rev.* **2007**, *36*, 831.
- (6) Song, H.-e.; Taniguchi, M.; Diers, J. R.; Kirmaier, C.; Bocian, D. F.; Lindsey, J. S.; Holten, D. *J. Phys. Chem. B* **2009**, *113*, 16483.
- (7) Kozaki, M.; Uetomo, A.; Suzuki, S.; Okada, K. *Org. Lett.* **2008**, *10*, 4477.
- (8) Sprafke, J. K.; Kondratuk, D. V.; Wykes, M.; Thompson, A. L.; Hoffmann, M.; Drevinskas, R.; Chen, W. H.; Yong, C. K.; Karnbratt, J.; Bullock, J. E.; Malfois, M.; Wasielewski, M. R.; Albinsson, B.; Herz, L. M.; Zigmantas, D.; Beljonne, D.; Anderson, H. L. *J. Am. Chem. Soc.* **2011**, *133*, 17262.
- (9) del Rosario Benites, M.; Johnson, T. E.; Weghorn, S.; Yu, L.; Rao, P. D.; Diers, J. R.; Yang, S. I.; Kirmaier, C.; Bocian, D. F.; Holten, D.; Lindsey, J. S. *J. Mater. Chem.* **2002**, *12*, 65.
- (10) Tsuda, A.; Osuka, A. *Science* **2001**, *293*, 79.
- (11) Ambroise, A.; Kirmaier, C.; Wagner, R. W.; Loewe, R. S.; Bocian, D. F.; Holten, D.; Lindsey, J. S. *J. Org. Chem.* **2002**, *67*, 3811.
- (12) Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126*, 4007.
- (13) Duchemin, I.; Deutsch, T.; Blase, X. *Phys. Rev. Lett.* **2012**, *109*, 167801.
- (14) Kobayashi, R.; Amos, R. D. *Chem. Phys. Lett.* **2006**, *420*, 106.
- (15) Nakai, K.; Sahnoun, R.; Kato, T.; Kono, H.; Fujimura, Y. *J. Phys. Chem. B* **2005**, *109*, 13921.

- (16) Wilson, G. J.; Arnold, D. P. *J. Phys. Chem. A* **2005**, *109*, 6104.
- (17) Yamaguchi, Y.; Yokoyama, S.; Mashiko, S. *J. Chem. Phys.* **2002**, *116*, 6541.
- (18) Holten, D.; Bocian, D. F.; Lindsey, J. S. *Acc. Chem. Res.* **2001**, *35*, 57.
- (19) Duan, X.; Huang, Y.; Lieber, C. M. *Nano Lett.* **2002**, *2*, 487.
- (20) Winkelmann, C. B.; Ionica, I.; Chevalier, X.; Royal, G.; Bucher, C.; Bouchiat, V. *Nano Lett.* **2007**, *7*, 1454.
- (21) S. Kwok, K. *Mater. Today* **2003**, *6*, 20.
- (22) Li, C.; Ly, J.; Lei, B.; Fan, W.; Zhang, D.; Han, J.; Meyyappan, M.; Thompson, M.; Zhou, C. *J. Phys. Chem. B* **2004**, *108*, 9646.
- (23) Marinado, T.; Nonomura, K.; Nissfolk, J.; Karlsson, M. K.; Hagberg, D. P.; Sun, L.; Mori, S.; Hagfeldt, A. *Langmuir* **2009**, *26*, 2592.
- (24) Tour, J. M. *Chem. Rev.* **1996**, *96*, 537.
- (25) Aratani, N.; Osuka, A.; Kim, Y. H.; Jeong, D. H.; Kim, D. *Angew. Chem., Int. Ed.* **2000**, *39*, 1458.
- (26) Saebi, S.; Pulay, P. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213.
- (27) Wu, S. Y.; Jayanthi, C. S. *Phys. Rep.* **2002**, *358*, 1.
- (28) Friesner, R. A.; Murphy, R. B.; Beachy, M. D.; Ringnalda, M. N.; Pollard, W. T.; Dunietz, B. D.; Cao, Y. *J. Phys. Chem. A* **1999**, *103*, 1913.
- (29) Yang, W. *J. Chem. Phys.* **1991**, *94*, 1208.
- (30) Kohn, W. *Phys. Rev. Lett.* **1996**, *76*, 3168.
- (31) Scuseria, G. E.; Ayala, P. Y. *J. Chem. Phys.* **1999**, *111*, 8330.
- (32) Li, S.; Ma, J.; Jiang, Y. *J. Comput. Chem.* **2002**, *23*, 237.
- (33) Haser, M.; Almlöf, J. *J. Chem. Phys.* **1992**, *96*, 489.
- (34) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198.
- (35) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357.
- (36) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959.
- (37) Gao, J.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467.
- (38) Wesolowski, T. A.; Warshel, A. *J. Phys. Chem.* **1993**, *97*, 8050.
- (39) Hofener, S.; Gomes, A. S.; Visscher, L. *J. Chem. Phys.* **2012**, *136*, 044104.
- (40) Cortona, P. *Phys. Rev. B* **1991**, *44*, 8454.
- (41) Wesolowski, T. A.; Weber, J. *Chem. Phys. Lett.* **1996**, *248*, 71.
- (42) Wesolowski, T. A.; Weber, J. *Int. J. Quantum Chem.* **1997**, *61*, 303.
- (43) Khait, Y. G.; Hoffmann, M. R. *J. Chem. Phys.* **2010**, *133*, 044107.
- (44) Casida, M. E.; Wesolowski, T. A. *Int. J. Quantum Chem.* **2004**, *96*, 577.
- (45) Konig, C.; Neugebauer, J. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10475.
- (46) Neugebauer, J. *J. Chem. Phys.* **2007**, *126*, 134116.
- (47) Neugebauer, J.; Curutchet, C.; Muñoz-Losa, A.; Mennucci, B. *J. Chem. Theory Comput.* **2010**, *6*, 1843.
- (48) Pavanello, M.; Van Voorhis, T.; Visscher, L.; Neugebauer, J. *J. Chem. Phys.* **2013**, *138*, 054101.
- (49) Yang, W.; Lee, T.-S. *J. Chem. Phys.* **1995**, *103*, 5674.
- (50) Imamura, A.; Aoki, Y.; Maekawa, K. *J. Chem. Phys.* **1991**, *95*, 5419.
- (51) Korchowiec, J.; Gu, F. L.; Imamura, A.; Kirtman, B.; Aoki, Y. *Int. J. Quantum Chem.* **2005**, *102*, 785.
- (52) Fujimoto, K.; Yang, W. *J. Chem. Phys.* **2008**, *129*, 054102.
- (53) Mezey, P. G. *J. Math. Chem.* **1995**, *18*, 141.
- (54) Exner, T. E.; Mezey, P. G. *J. Phys. Chem. A* **2002**, *106*, 11791.
- (55) Gadre, S. R.; Ganesh, V. *J. Theor. Comput. Chem.* **2006**, *5*, 835.
- (56) Fedorov, D. G.; Nagata, T.; Kitaura, K. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7562.
- (57) Huang, L.; Bohorquez, H. J.; Matta, C. F.; Massa, L. *Int. J. Quantum Chem.* **2011**, *111*, 4150.
- (58) Hua, S.; Hua, W.; Li, S. *J. Phys. Chem. A* **2010**, *114*, 8126.
- (59) Deev, V.; Collins, M. A. *J. Chem. Phys.* **2005**, *122*, 154102.
- (60) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. *Chem. Rev.* **2011**, *112*, 632.
- (61) Zhang, D. W.; Zhang, J. Z. *H. J. Chem. Phys.* **2003**, *119*, 3599.
- (62) Chen, X.; Zhang, Y.; Zhang, J. Z. *J. Chem. Phys.* **2005**, *122*, 184105.
- (63) He, X.; Zhang, J. Z. *J. Chem. Phys.* **2005**, *122*, 31103.
- (64) Inadomi, Y.; Nakano, T.; Kitaura, K.; Nagashima, U. *Chem. Phys. Lett.* **2002**, *364*, 139.
- (65) Fedorov, D. G.; Kitaura, K. *J. Chem. Phys.* **2005**, *122*, 054108.
- (66) Mochizuki, Y.; Koikegami, S.; Amari, S.; Segawa, K.; Kitaura, K.; Nakano, T. *Chem. Phys. Lett.* **2005**, *406*, 283.
- (67) Mochizuki, Y.; Tanaka, K.; Yamashita, K.; Ishikawa, T.; Nakano, T.; Amari, S.; Segawa, K.; Murase, T.; Tokiwa, H.; Sakurai, M. *Theor. Chem. Acc.* **2006**, *117*, 541.
- (68) Chiba, M.; Koido, T. *J. Chem. Phys.* **2010**, *133*, 044113.
- (69) Chiba, M.; Fedorov, D. G.; Kitaura, K. *Chem. Phys. Lett.* **2007**, *444*, 346.
- (70) Li, W.; Li, S.; Jiang, Y. *J. Phys. Chem. A* **2007**, *111*, 2193.
- (71) Fedorov, D. G.; Kitaura, K. *J. Phys. Chem. A* **2007**, *111*, 6904.
- (72) Yeole, S. D.; Gadre, S. R. *J. Chem. Phys.* **2010**, *132*, 094102.
- (73) Bettens, R. P. A.; Lee, A. M. *J. Phys. Chem. A* **2006**, *110*, 8777.
- (74) Hindin, E.; Forties, R. A.; Loewe, R. S.; Ambroise, A.; Kirmaier, C.; Bocian, D. F.; Lindsey, J. S.; Holten, D.; Knox, R. S. *J. Phys. Chem. B* **2004**, *108*, 12821.
- (75) McMahon, D. P.; Troisi, A. *Chem. Phys. Lett.* **2009**, *480*, 210.
- (76) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.
- (77) Casida, M. E.; Jamorski, C.; Casida, K. C.; Salahub, D. R. *J. Chem. Phys.* **1998**, *108*, 4439.
- (78) Scalmani, G.; Frisch, M. J.; Mennucci, B.; Tomasi, J.; Cammi, R.; Barone, V. *J. Chem. Phys.* **2006**, *124*, 094107.
- (79) Furche, F.; Ahlrichs, R. *J. Chem. Phys.* **2002**, *117*, 7433.
- (80) Stratmann, R. E.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **1998**, *109*, 8218.
- (81) Besley, N. A. *Chem. Phys. Lett.* **2004**, *390*, 124.
- (82) DeBeer George, S.; Petrenko, T.; Neese, F. *Inorg. Chim. Acta* **2008**, *361*, 965.
- (83) Robinson, D.; Besley, N. A.; Lunt, E. A.; O'Shea, P.; Hirst, J. D. *J. Phys. Chem. B* **2009**, *113*, 2535.
- (84) Robinson, D.; Besley, N. A.; O'Shea, P.; Hirst, J. D. *J. Phys. Chem. B* **2009**, *113*, 14521.
- (85) Robinson, D.; Besley, N. A.; O'Shea, P.; Hirst, J. D. *J. Phys. Chem. B* **2011**, *115*, 4160.
- (86) Stener, M.; Fronzoni, G.; de Simone, M. *Chem. Phys. Lett.* **2003**, *373*, 115.
- (87) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaraillo, J.; Gomperts, R.; Stratmann, R. E.; Yazayev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Willingford, CT, 2009.
- (88) Schlegel, H. B. *J. Comput. Chem.* **2003**, *24*, 1514.
- (89) Goldfarb, D. *Math. Comp.* **1970**, *24*, 23.
- (90) Shanno, D. F. *Math. Comp.* **1970**, *24*, 647.
- (91) Broyden, G. C. *J. Inst. Math. Its Appl.* **1976**, *6*, 76.
- (92) Fletcher, N. C.; Robinson, T. C.; Behrendt, A.; Jeffery, J. C.; Reeves, Z. R.; Ward, M. D. *J. Chem. Soc., Dalton Trans.* **1999**, 2999.
- (93) Hratchian, H. P.; Schlegel, H. B. In *Theory and Applications of Computational Chemistry: The First Forty Years*; Dykstra, C., Ed.; Elsevier B. V.: Amsterdam, the Netherlands, 2005; p 195.
- (94) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- (95) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

- (96) Roy, L. A.; Hay, P. J.; Martin, R. L. *J. Chem. Theory Comput.* **2008**, *4*, 1029.
- (97) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.
- (98) Harihara, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.
- (99) Hess, B. A.; Schaad, L. J.; Carsky, P.; Zahradnik, R. *Chem. Rev.* **1986**, *86*, 709.
- (100) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51.