CHEMICAL REVIEWS

REVIEW

pubs.acs.org/CR

Fragmentation Methods: A Route to Accurate Calculations on Large Systems

Mark S. Gordon,*[,†] Dmitri G. Fedorov,[‡] Spencer R. Pruitt,[†] and Lyudmila V. Slipchenko[§]

[†]Department of Chemistry and Ames Laboratory, Iowa State University, Ames Iowa 50011, United States

[‡]Nanosystem Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan

[§]Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, United States

CONTENTS

1. INTRODUCTION

Theoretical chemists have always strived to perform quantum mechanics (QM) calculations on larger and larger molecules and molecular systems, as well as condensed phase species, that are frequently much larger than the current state-of-the-art would suggest is possible. The desire to study species (with acceptable accuracy) that are larger than appears to be feasible has naturally led to the development of novel methods, including semiempirical approaches, reduced scaling methods, and fragmentation methods. The focus of the present review is on fragmentation methods, in which a large molecule or molecular system is made more computationally tractable by explicitly considering only one part (fragment) of the whole in any particular calculation. If one can divide a species of interest into fragments, employ some level of ab initio QM to calculate the wave function, energy, and properties of each fragment, and then combine the results from the fragment calculations to predict the same properties for the whole, the *possibility* exists that the accuracy of the outcome can approach that which would be obtained from a full (nonfragmented) calculation. It is this goal that drives the development of fragmentation methods.

An additional potential positive aspect of fragmentation methods is their ability to take advantage of massively parallel computers. If one can calculate the energy of each fragment in such a manner that the fragment calculation is essentially independent of the calculations for all of the other fragments, then each fragment calculation can be performed on a separate compute node. This is sometimes called coarse-grained parallel computing, and it can be very efficient. If, in addition, the QM algorithm one is using (e.g., for second order perturbation theory, MP2) has itself been developed for parallel hardware, one can implement a multilevel parallel approach, with fine-grained parallelism employed among the cores within each node. This further improves the parallel efficiency of the calculation.[1]

Aiming at achieving computational efficiency, a variety of methods based on localized orbitals was developed starting with

Special Issue: 2011 Quantum Chemistry

Received: March 27, 2011

Adams in 1961.[2] These methods are only briefly mentioned in this review, as they typically deal with orbitals in the full system rather than molecular fragments. Klessinger and McWeeny[3] in 1965 in their Group SCF (self-consistent field) method suggested defining molecular orbital groups to reduce the scaling of SCF. Christoffersen and co-workers[4−14] employed localized molecular orbitals (LMOs) and floating spherical Gaussian orbitals (FSGO)[15] to separate the subdensity matrices of fragments that are then summed to obtain the total density and (thereby) desired properties. The method has been applied with success to many species, including molecules of biological importance. As another important example, Stoll and Preuss[16] in 1977 used a subsystem-based approximation to define the total density, and suggested many-body corrections to the energy from the density computed at these levels. Stoll[17] in 1992 suggested many-body incremental corrections to the correlation energy, based on LMOs. The development of LMO methods for fragments is still an active area of research.[18,19] In a group of methods based on the divide-and-conquer (DC) idea,[20] an ad hoc total density from fragment calculations is constructed, followed by a single energy evaluation using this density.

The elongation method (usually abbreviated as ELG), based upon localized orbitals, was developed by Imamura[21] in 1991 (related to the earlier work by Demelo et al.[22]) to enable calculations on large polymeric chains at the computational cost of a much smaller system. A good example of an application for which the ELG method was developed is that of a polymerization/copolymerization reaction. Reactions of this type begin with an initiation step, followed by some number of propagation steps, and ultimately complete with a termination step. This is the series of steps that are employed in the elongation methodology. Initially, the applicability of the elongation method was limited to periodic polymer chains,[21] however the method has since been extended to aperiodic systems[23] and hydrogen bonded systems such as water clusters.[24] Since the original formulation, a new localization scheme[25] (a central aspect of the formalism) as well as extensions have been proposed.[26−44] The elongation method is fundamentally an LMO approach, rather than a fragment-based approach, so it is discussed only briefly here.

There have been many fragment-like approaches to electronic structure theory, in which fragments appear as groups of atoms and the electronic state of the full system is computed. Dreyfus and Pullman[45] in 1970 and Morokuma[46] in 1971 systematically considered both the Hartree product wave function and the fully antisymmetrized total wave functions constructed from wave functions of two weakly interacting fragments. This approach was extended to an arbitrary number of fragments by Ohno and Inokuchi[47] for an antisymmetrized product in 1972, and by Gao[48] for a Hartree product in 1997. Some theoretical foundations for fragment-based methods were given by Kutzelnigg and Maeder in 1978.[49]

The idea of performing fragment calculations in the Coulomb field of other fragments with or without the corresponding exchange has been reinvented and reformulated over the years in various forms, and given names such as mutual consistent field (MCF),[50] structural SCF,[51] self-consistent embedded ions (SCEI),[52,53] double self-consistent field (DSCF),[54] self-consistent charge (SCC)[55] and monomer SCF.[56] One of the early methods to accomplish this self-consistent cycle is attributed to Morokuma[46] in 1971, who introduced it to add the polarization term to the analysis of Dreyfus and Pullmann.[45] The Morokuma development was followed by the systematic work of Ohno[47] in 1972 and the energy decomposition (EDA) method by Kitaura and Morokuma[57] in 1976. The EDA method was later extended to more than two fragments by Chen and Gordon[58] in 1996. While these earlier approaches were not intended for large-scale calculations, the MCF method by Otto and Ladik[50] is the prototypical approach for self-consistent fragment calculations that are used today for large systems. Later it was applied to periodic systems by Bohm[59] in 1982, Kubota et al.[51] in 1994, and Pascual and Seijo[53] in 1995. Gao employed this MCF approach in his molecular orbital-derived empirical potential for liquid simulations (MODEL)[48] in 1997 (later renamed explicit polarization, X-Pol) and by Kitaura et al.[60] in the fragment molecular orbital (FMO) method in 1999.

An important advance in the development of fragment-based methods was made by the explicit division of the equations into blocks derived for individual fragments. While the earlier methods had to diagonalize the full Fock matrix,[46,47,57] the introduction of strong orthogonality between fragment wave functions made it possible to block the fragment Fock matrices,[50−52,54,60] although they are still coupled in many methods via the electrostatic field. Here again, note the paramount importance of the MCF approach,[50] which with modifications has been integrated as a part of modern methods such as X-Pol and FMO.

On the basis of the starting point of computing the electronic state of individual fragments, two main branches of methods emerged for the subsequent refinement of properties. The first group constitutes a perturbation treatment of fragment interactions, a typical example being the MCF method,[50,61−71] in which the mutually polarized fragment wave functions are used in a perturbative manner to obtain the exchange and charge transfer interactions between fragments. The other group of methods was inspired by the desire to decompose intermolecular interactions into conceptually familiar components. In this group, pairs or larger conglomerates of fragments are computed, and the important concept is the treatment of many-body effects other than the electrostatics. One of the early methods was developed by Hankins et al.,[72] who considered many-body corrections to intermolecular interactions in vacuum. By a clever introduction of the electrostatic potential (ESP) in the many-body expansion, Kitaura et al.[60] in 1999 were able to incorporate many-body effects in the framework of a two-body expansion, conceptually reminiscent of some earlier methods.[72,16,17] The systematic cancellation of the double counting of the Coulomb interaction in many-body calculations was shown diagrammatically by Fedorov and Kitaura in 2004.[56] Suárez et al.[73] in 2009 discussed some general aspects of many-body expansions. Gao et al.[74] suggested a means for including many-body effects using the total wave function. Many methods discussed in detail below partially consider many-body effects implicitly by computing large conglomerations of atoms (fragments).

In the original 1975 implementation of MCF by Otto and Ladik,[50] the wave function of a fragment was polarized by a partner fragment through modifying the Fock operator by the Coulomb potential of a partner. A simple monopole approximation was used originally to represent the electrostatic potential; HF equations for a pair of fragments were solved until self-consistency. This strategy automatically takes care of electrostatic and polarization effects. Exchange and charge-transfer energies were then obtained perturbatively based on the polarized wave functions of the fragments. Later on, the method was extended by including a local exchange potential (using the original Slater expression) into the Fock operator and

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

substituting the monopole approximation by distributed charges.[61] In 1985, implementation of the pseudopolarization tensor (PPT) MCF,[67,69] point charges and pseudopolarizabilities were calculated for each individual fragment, and the Coulomb and polarization interaction energies were computed using these charges and polarizabilities. So, PPT-MCF is related to universal force fields, like EFP,[75−77] SIBFA,[78] and AMOEBA.[79]

The idea of a perturbative treatment of fragment interactions has been further extended and developed in the symmetry-adapted perturbation theory (SAPT)[80,81] by Moszynski, Jeziorski, and Szalewicz in which the interaction energy is expressed in orders of an intermolecular interaction operator V and a many body perturbation theory (MBPT) operator W. The polarization energies are obtained from a regular Rayleigh−Schrödinger perturbation theory; for example, the Coulomb energy appears in the first order, the induction and dispersion energies in the second order, etc. Additionally, the exchange corrections arise from the use of a global antisymmetrizer to force the correct permutational symmetry of the dimer wave function in each order. In this way, the exchange-repulsion energy appears in the first order, and exchange-induction and exchange-dispersion contribute to the second order of the perturbation theory. The supermolecular HF energy in SAPT can be represented by corrections in zero order in W, namely, Coulomb, induction, exchange, and exchange-induction. The SAPT2 level is roughly equivalent to supermolecular second-order MBPT calculations. The new contributions to the interaction energy in the second-order theory are Coulomb, induction, and exchange-induction corrections that are second-order in W, in addition to first- and second-order corrections to W for exchange, and zeroth-order dispersion and exchange-dispersion corrections. The highest routinely used level of SAPT is equivalent to fourth order supermolecular MBPT and includes third-order corrections to W for the Coulomb energy and higher orders of the exchange and dispersion energies.

SAPT was extended to enable its use with density functional theory (DFT).[82] In this SAPT(DFT) scheme, Kohn−Sham (KS) orbitals and orbital energies are used to obtain Coulomb and exchange in zero order of W, similarly to how it is done in the wave function-based SAPT (essentially, by replacing HF orbitals with KS orbitals). Induction, dispersion, exchange-induction, and exchange-dispersion are calculated by using the frequency-dependent density susceptibility functions obtained from the time-dependent DFT theory at the coupled Kohn−Sham level of theory.[82−84] The scaling of SAPT(DFT) is $O(N^6)$ and becomes $O(N^5)$ if density fitting is used.[85,86] This is significantly better than the $O(N^7)$ scaling of the wave function-based SAPT.

In variational (i.e., HF) energy decomposition schemes, the total energy is typically represented as a sum of a frozen density interaction energy, a polarization energy, and a charge-transfer energy. The frozen density term is calculated as the interaction of the unrelaxed electron densities on the interacting molecules and consists of Coulomb and exchange contributions. The polarization term originates from the deformation of the electron clouds of the interacting molecules in the fields of each other. The charge transfer term arises because of the electron flow between the molecules in the system. Quantum mechanically, polarization and charge transfer terms can be described as energy lowerings due to the intramolecular and intermolecular relaxation of the molecular orbitals, respectively. The main differences in the variational EDA schemes come from the manner in which the intermediate self-consistent energies, corresponding to the variationally optimized antisymmetrized wave functions constructed from MOs localized on the individual molecules, are determined.

The original EDA method of Kitaura and Morokuma[57,87,88] (KM), which has become prototypical for many subsequent methods, lacks the antisymmetrization of the intermediate wave function, resulting in numerical instabilities of the polarization and charge-transfer components at short distances and with large basis sets. It is, however, not entirely artificial, but is connected to the intrinsic similarity between the intramolecular (polarization) and intermolecular charge transfers. The concept of the fragment (monomer) polarization is apparently not well-defined with large basis sets, especially those that contain diffuse functions and are therefore not obviously localized on a given fragment. In this sense, when the polarization of fragments as an important physical concept is of interest, one may have to employ smaller basis sets.

The restricted variational space (RVS) analysis[89,58] and the constrained space orbital variations (CSOV) method[90,91] improve on the KM scheme by employing fully antisymmetrized intermediate wave functions. The main deficiency of both methods is that they do not produce self-consistent polarization energies (which results, for example, in a dependence on the order in which the fragments are treated) and do not completely separate charge-transfer from polarization. In the natural energy decomposition analysis (NEDA)[92−94] the intermediate wave function is not variational and the resulting polarization and charge-transfer energies may be under- and overestimated, respectively. In the absolutely localized molecular orbital (ALMO) approach[95] by Head-Gordon, all energy terms are calculated variationally and the polarization and charge-transfer terms are naturally separated. Additionally, the charge-transfer energy can be decomposed into forward donation and back-bonding contributions. An earlier implementation of EDA based on the block localized wave function developed by Mo et al. is similar in spirit.[96] An energy decomposition analysis based on the divide and conquer approach was also formulated.[97]

Fedorov and Kitaura have extended the EDA method to covalently bound fragments by introducing the pair interaction energy decomposition analysis (PIEDA)[98] within the FMO framework[60] and showed the equivalency of the fragment SCF in FMO and the polarized state of monomers in EDA. In the FMO method, one performs fragment calculations in the electrostatic field of other fragments, mutually self-consistent with each other. This approach provides the polarized state of the fragments including many-body polarization. In the EDA method, the same thing is accomplished by the restriction of orbital rotations, which only allows many-body polarization to take place for the interfragment interactions. The exact agreement of the resultant many-body polarization, shown numerically with the aid of PIEDA for water clusters, clearly demonstrated the equivalence of the self-consistent fragment polarizations in FMO and EDA, in this regard.

In the EDA approach developed by Su and Li,[99] the interaction energy is separated into Coulomb, exchange, repulsion, polarization, and dispersion terms. In contrast with the KM and other EDA schemes, the exchange and repulsion energies in this scheme are separated according to the method of Hayes and Stone.[100] Dispersion is obtained using a supermolecular approach and size-consistent correlation methods (MP2 or CCSD(T)). Polarization is defined as the orbital relaxation energy that occurs on going from the monomer to the supermolecular orbitals. That is, the polarization interaction in the Su-Li EDA

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

method includes both the polarization and charge-transfer contributions of the KM scheme. A DFT version of the Su-Li EDA scheme was also developed.

Wu et al. developed a density-based EDA,[101] in which the energies of the intermediate states are calculated using the densities of the fragments, rather than their wave functions. Similarly to the KM EDA method, the total energy is separated into a frozen density Coulomb plus exchange repulsion term, polarization, and charge-transfer. The frozen density energy is obtained using the constrained search technique developed by Wu and Yang.[102] The polarization and charge-transfer terms are separated from each other by constructing an intermediate state in which the density is relaxed without charge transfer, using the constrained DFT method of Wu and Van Voorhis.[103] The EDA approach is a useful tool for the development of reliable force fields for condensed phase molecular simulations.[79,104−108,202,203,201]

Related to EDA schemes and ideas is the self-consistent-field method for molecular interactions (SCF-MI) first introduced by Gianinetti et al.[109,110] and further reformulated and extended by Nagata et al.[111] In the SCF-MI method one expands molecular orbitals of a given fragment in terms of only the atomic orbitals belonging to atoms of that fragment. This leads to absolutely localized MOs that are free from basis set superposition error (BSSE) but also prevents charge-transfer between the fragments. The charge-transfer interactions between the fragments can be added by means of single excitation second order perturbation theory, as suggested by Nagata and Iwata.[112] Khaliullin et al. showed that the SCF-MI method can result in significant computational savings (e.g., $(N/O)^2$ speedups for the diagonalization step compared to the conventional SCF method) and thus is applicable to systems containing hundreds of molecules.[113]

Murrell and co-workers developed a method called diatomics-in-molecules,[114] that relied on a power expansion in the intermolecular overlap integral, in order to obtain a general expression for the intermolecular exchange repulsion. Jensen and Gordon[115] built upon the Murrell approach using LMOs to derive a general expression for the intermolecular exchange repulsion interaction energy.[116,117]

A number of fragment methods have been proposed based on the so-called thermochemical[118] analogy, i.e., by capping the fragments of interest and eliminating or subtracting the effect of the caps. These include the molecular fractionation with conjugate caps (MFCC) method,[119] its later extension, the generalized energy-based fragmentation (GEBF) method,[120] the molecular tailoring approach (MTA),[121] the kernel energy method (KEM), and the fragment energy method (FEM).[73] The relation between these has been discussed by Suárez et al.[73] Collins and co-workers[122] have developed the systematic fragmentation method (SFM) in order to describe large molecular systems with accurate QM methods. Smaller subsystems are treated with a high level of accuracy. This accuracy is, in principle, retained by incorporating nonbonded interactions between the fragments using model potentials.[123]

A related, but somewhat different approach has been developed and implemented by several groups in order to expand the size of accessible species and active species in multiconfigurational self-consistent field (MCSCF) methods, in particular the complete active space (CAS) SCF approach. Among the most well-known of these methods are the restricted active space (RAS) SCF method, developed by Roos and co-workers,[124] the quasi (Q) CAS method of the Hirao group,[125] and the occupation restricted multiple active space (ORMAS) method developed by Ivanic.[126] The general philosophy of these methods is to divide a large (possibly computationally intractable) CAS active space into logically determined subspaces, so that each subspace is amenable to MCSCF calculations. ORMAS is the most general of these methods and subsumes the others. Because MCSCF calculations account only for static correlation, it is necessary to add dynamic correlation, either variationally via configuration interaction (multireference CI = MRCI) or perturbatively (MRPT). MRPT is the computationally more efficient approach. Therefore, the RASPT2 method[127] and ORMAS-PT2 method[128] have recently been introduced.

Another class of methods that is related to fragmentation is the use of LMOs to reduce the scaling of (most commonly) correlated electronic structure methods. This is accomplished by first performing a Hartree−Fock (HF) calculation on the entire system and localizing the orbitals. In this sense the "local orbital" approach is distinct from most fragmentation methods in that the latter first separate the system into separate physical collections of atoms, and subsequently perform explicit calculations on only one fragment at a time, not the entire system. Most local orbital methods[129] have built upon the beautiful early work of Pulay and co-workers,[130,131] whose work primarily addressed second-order perturbation theory (MP2). Many others have contributed to this field, most notably Werner and co-workers[132] (MP2 and coupled cluster (CC) methods), Head-Gordon et al.[133] (MP2 and CC methods), and Carter and co-workers[134] (multireference configuration interaction). A recent method developed by the Piecuch group[135] called cluster in molecule (CIM) is generally applicable and can be used as a multilevel method, in which different parts of a system can be treated with different levels of theory (Werner and co-workers[136] used MP2 for small contributions in their local CC method).

Another important example is the incremental method,[17,137] which is mainly used to estimate the correlation energy in periodic systems, similar to other local orbital methods, with the important distinction that a many-body expansion in terms of orbital contributions is used. This is conceptually similar to the expansion used in some fragment methods, although molecular applications have also been reported.[138−140] Manby et al.[141] suggested a hierachical method for the additive calculation of the properties of clusters using edge, corner, surface and bulk unit energies, and compared their method to the incremental method.[142]

A number of fragment-based methods represent an important class of linear scaling methods[143] (sometimes called order N, $O(N)$ methods), which are a thriving field of current research.[144,145] Strictly speaking, it is often not entirely clear that a particular method really scales linearly, and a description like "nearly linear scaling" may be more appropriate for practically relevant regimes of N (as discussed by Nagata et al. in Chapter 2 of ref 144), although one can in principle also define an asymptotic scaling for infinite N.

Another group of approaches, only briefly mentioned is related to the integrated MO and molecular mechanics (MM), IMOMM method,[146] its later extension our own n-layered integrated molecular orbital and molecular mechanics (ONIOM),[147] methods known as integrated MO and MO (IMOMO),[148] QM:QM,[149] and multiple area QM/MM.[150,151] Most of these methods compute the whole system at a low level and add higher level results for a selected part of the system. Some of the methods such as multiple area QM/MM[150,151] and extended ONIOM[152] can be considered fragment-based approaches.

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

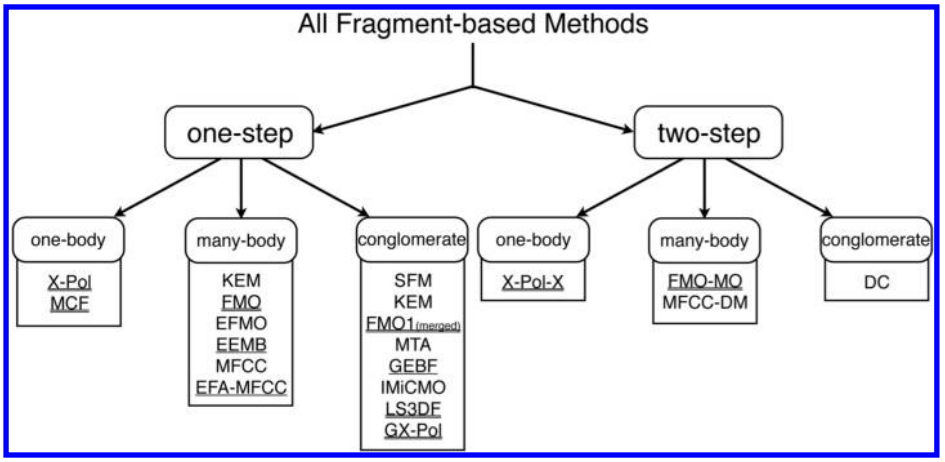


Figure 1. Classification of fragment-based methods. Singly underlined approaches include constant embedding potentials, while doubly underlined approaches include some form of self-consistent fragment embedding potential (SCC); other approaches have no embedding potential.

2. METHODOLOGIES

One can classify fragment-based methods in various ways. Fedorov and Kitaura[226] suggested three categories: divide-and-conquer (DC),[20] transferable approaches (SFM[122]), and methods based on many-body molecular interactions (FMO,[60] MFCC[119]). The distinction between the latter two groups is in the way in which fragment interactions are added: the latter category typically has a simple many-body expansion inspired by the theory of molecular interactions;[72] while the transferable group has a different recipe for taking into account these interactions. Fedorov and Kitaura also proposed the name *e pluribus unum* for those methods in which the total properties are obtained by fragment calculations under the influence (for example, in the electrostatic field) of the whole system.

Li et al.[120] grouped fragment methods into density matrix (DC,[20] ELG,[21] MFCC[119]) or energy based approaches (IMiCMO,[153] MFCC,[119] SFM,[122] MTA[121]). The density based methods compute the density of the whole system using fragments, followed by the calculation of the energy; the energy based methods compute the energy directly from fragment energies avoiding the expensive step of calculating the energy from the density of the whole system. For some methods, like MFCC, it is possible to either compute the energy directly or to first obtain the density and then the energy; thus MFCC belongs to both categories.

Suárez et al.[73] classified fragment-based methods into two main groups, those using (a) overlapping fragments (MTA,[121] MFCC,[119] FEM,[73] and SFM[122]) and (b) disjoint fragments (FMO, KEM). The main idea of this classification is in the fragment definition: whether an atom is assigned to a single fragment as in (a) or can be found in several, as in (b).

Figure 1 illustrates an elaboration and generalization of the classification introduced by Li et al.[120] The purpose of this modification is to introduce a more elaborate division than the two groups that were used in the original classification, and to extend it to apply when a property other than the energy is directly computed from the fragment calculations. The energy-based group is renamed the *one-step* group, which refers to the direct calculation of total properties from the fragment-derived values. This approach allows one to include properties, other than energy, such as the density and properties linear in the density (i.e., one-electron properties) to be included in the classification. The distinctive feature of the one-step group is that the total properties are obtainable directly from fragment calculations. That is, the one-electron properties from the density can be rewritten as an appropriate combination of fragment properties.

The *two-step* group, which corresponds to the density based method of Li et al.,[120] includes methods which consist of two steps: (1) computing some total property from piecewise values (usually, the electron density); (2) calculating a related total property dependent on the one determined in the first step in a nonlinear way (usually, the total energy). For methods in this group it is necessary to perform calculations for the whole system at once, which imposes various limitations and leads to complications.

Each of the one and two-step groups can be further divided to reflect the manner in which the QM calculations on the combinations of fragments are performed. In the 1-body methods, where "1-body" refers to fragments, there are no QM calculations of the conglomerates of fragments. Instead, often either force field derived terms are added (e.g., X-Pol[48]) or perturbation theory is used based on the wave functions of fragments as the starting point (MCF[50]). In the many-body subgroup pairs, triples or large unions of fragments are computed, whereas in the conglomerate subgroup different principles are used to combine fragments into unions, for example, based on the cardinality (MTA method) or buffer zone (DC method) to include a certain number of atoms around each fragment. The many-body group eventually can be said to be inspired by the theory of molecular interactions[72] with its many-body expansion and typically contains many-body energy corrections such as $E_{IJ}-E_I-E_J$, for dimer IJ corrections to monomers I and J. On the other hand, the conglomerate group uses different recipes for the construction of groups of atoms to be computed.

Several families of methods, namely, FMO, MFCC, and KEM have variants that fall into different categories. For instance, in the FMO-MO method (introduced below), an FMO calculation is performed in the first step, whereas the second step is performed as in a full ab initio QM calculation for the whole system, corresponding to one SCF iteration. In FMO-based NMR calculations, denoted FMO1 (merged), one first performs self-consistent calculations on fragments (FMO1). This is followed by the calculation of NMR shifts for a large "super-fragment" composed of a central fragment I and all fragments within a desired radius around it (I consequently goes over all

dx.doi.org/10.1021/cr200093j |*Chem. Rev.* XXXX, XXX, 000–000

fragments). This is the reason that some methods appear in several different groups.

2.1. QM-Based Force Fields

Quantum mechanics calculations are often used to assist the development and parametrization of modern polarizable force fields. However, while it is straightforward to extract the electrostatic point charges or multipoles from electronic structure calculations on small fragments or molecules, rigorous formulation of accurate but computationally inexpensive ways to model exchange-repulsion self-consistent induction, dispersion, and charge-transfer terms can be quite intricate, especially in the interfacial region between QM and the force field. Therefore, various approaches have been suggested and tried over the years, often combining quantum-mechanically based terms (typically Coulomb and polarization) and parametrizing the rest of the potential.

For example, in the AMOEBA (atomic multipole optimized energetics for biomolecular applications) force field, multipoles on atoms (up to quadrupoles)[79] are obtained from the distributed multipolar analysis (DMA) and (experimental) isotropic atomic polarizabilities with Thole's damping functions are used.[154] Parameters for Thole's damping functions are fitted using EDA computations[155,156] Also, in AMOEBA the van der Waals R^{-7}−R^{-14} term is parametrized to reproduce experimental gas phase and condense phase data. Parameters of bonded terms in AMOEBA are also obtained by fitting to experimental data. In the NEMO (nonempirical molecular orbital) force field for water,[157] multipoles up to quadrupoles are also used to calculate Coulomb interactions, while atomic polarizabilities are obtained by solving coupled HF equations. Dispersion is fitted to an analytic R^{-6} expression with exponential damping, while the exchange-repulsion is represented by a sum of exponential and R^{-14} terms.

The essence of the direct reaction field (DRF) approach[158] of van Duijnen is an explicit evaluation of the polarization (induction) term using atomic polarizabilities and Thole's damping functions. The Coulomb term is evaluated using screened charges; dispersion energy is calculated using the Slater−Kirkwood expression[159] with atomic polarizabilities, and the repulsion component is taken to be proportional to the dispersion scaled by van der Waals atomic radii.

In the SIBFA (sum of interactions between fragments ab initio) force field,[78,160−164] developed by Gresh and co-workers, the interaction energy between fragments is a sum of Coulomb, polarization, repulsion, dispersion, and charge-transfer components. The electrostatic term is computed using distributed multipoles up to quadrupoles centered at atoms and bond barycenters. The multipoles are obtained using the procedure developed by Vigné-Maeder and Claverie.[165] The Coulomb term is augmented by an explicit penetration contribution.[166−168] The repulsion term is formulated as a sum of bond−bond, bond-lone pair, and lone pair-lone pair interactions expressed as S^2/R terms. Here, S is an approximation of the overlap between LMOs of the interacting partners; R is the distance between the LMO centroids. The S^2/R^2 term was added to improve the accuracy of the repulsion term.[167] Polarization energies are obtained by using permanent multipoles (the same as those that appear in the Coulomb term) and distributed polarizability tensors computed at the bond barycenters and on the heteroatom lone pairs. Polarization interactions are screened by Gaussian functions that depend on the distance between the interaction centers. The dispersion term is described as a sum of R^{-6}, R^{-8}, and R^{-10} contributions,[169] calibrated on the basis of a SAPT analysis. The uniqueness of SIBFA is in the treatment of the charge-transfer interactions that are modeled following the Murrell formulation[170] using the ionization potential of the electron donor and the electron affinity and "self-potential" of the electron acceptor.[162,161] The charge transfer term is essential for describing polycoordinated complexes of cations.[171,164]

2.1.1. Effective Fragment Potential Method. The original idea of the effective fragment potential method[75,172,173] (EFP) was to describe aqueous solvent effects on molecules of biological interest, in which the solute was described by some QM level (usually HF) and the solvent molecules were represented by EFPs. The EFP potential was represented as series deduced from the long-range (in powers of $(1/R)$ Coulomb operator) and short-range (in powers of intermolecular overlap S) perturbation theory. In this formulation each water molecule is represented as a fragment of fixed geometry with a set of parameters deduced from ab initio calculations. In the original implementation,[172,173] now called EFP1, the interaction energy between water molecules consists of Coulomb, polarization, and repulsion terms

$$E^{\mathrm{EFP1}} = E^{\mathrm{Coul}} + E^{\mathrm{pol}} + E^{\mathrm{rep}} \tag{1}$$

In the presence of a quantum region, the total energy of the QM/EFP1 system can be written as

$$E^{\mathrm{QM-EFP1}} = \langle \Psi | H^{\mathrm{QM}} + V^{\mathrm{Coul}} + V^{\mathrm{pol}} + V^{\mathrm{rep}} | \Psi \rangle + E^{\mathrm{Coul}} + E^{\mathrm{pol}} + E^{\mathrm{rep}} \tag{2}$$

V^{Coul}, V^{pol}, and V^{rep} are one-electron contributions to the (unperturbed) quantum Hamiltonian H^{QM} due to Coulomb, polarization, and repulsion terms of EFP1 water fragments. E^{Coul}, E^{pol}, and E^{rep} are fragment-fragment Coulomb, polarization, and exchange-repulsion energies, respectively.

The Coulomb E^{Coul} term is evaluated using classical multipoles up to octopoles centered at each atom and bond midpoint. The distributed multipole moments are obtained using the Stone distributed multipole analysis;[174,175] each water fragment has five points with distributed multipoles. The fragment-fragment Coulomb interactions consist of charge−charge, charge-dipole, charge-quadrupole, charge-octopole, dipole−dipole, dipole−quadrupole, and quadrupole−quadrupole terms. The Coulomb contribution V^{Coul} from a multipole point k to the ab initio Hamiltonian consists of four terms due to EFP charges q, dipoles μ, quadrupoles Θ, and octopoles Ω

$$V_k^{\mathrm{Coul}}(x) = q_k T(r_{kx}) - \sum_{\alpha}^{x,y,z} \mu_\alpha^k T_\alpha(r_{kx}) + \frac{1}{3}\sum_{\alpha,\beta}^{x,y,z} \Theta_{\alpha\beta}^k T_{\alpha\beta}(r_{kx}) - \frac{1}{15}\sum_{\alpha,\beta,\gamma}^{x,y,z} \Omega_{\alpha\beta\gamma}^k T_{\alpha\beta\gamma}(r_{kx}) \tag{3}$$

T, T_α, $T_{\alpha\beta}$, and $T_{\alpha\beta\gamma}$ are electrostatic tensors of zero, first, second, and third rank, respectively; x is the electron coordinate; r_{kx} is the distance between the position of electron x and multipole point k.

To account for short-range charge-penetration effects, the charge−charge fragment-fragment terms and the charge-based term in the Hamiltonian are augmented by Gaussian like damping functions of the form $1 - \beta \exp(-\alpha R^2)$. The α and β parameters were determined from a fit of the damped multipole potential to the quantum Hartree−Fock potential on a set of

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

points around the fragment. R is the distance between two multipole points or between a multipole point and an electron.

The fragment polarization energy E^{pol} is evaluated as an interaction of induced dipoles of each fragment with the static field due to the Coulomb multipoles and the induced field due to the induced dipoles of the other fragments. The induced dipoles originate at the centroids of localized molecular orbitals (LMO), where (anisotropic) distributed polarizability tensors are placed. Each water fragment has five distributed polarizability points: at the oxygen (inner shell), the centers of the two O−H bonds, and at the centroids of the two lone pairs. The polarization energy of the fragments is calculated self-consistently using an iterative procedure. Thus, polarization accounts for some many-body effects that are important in aqueous systems.

When a quantum region is present in the system, the induced dipoles of the fragments also interact with the electron density and nuclei of the quantum part, by means of V^{pol} a one-electron contribution to the quantum Hamiltonian:

$$V_k^{pol}(x) = \frac{1}{2}\sum_{\alpha}^{x,y,z}(\mu_\alpha^{ind,k} + \tilde{\mu}_\alpha^{ind,k})T_\alpha(r_{kx}) \quad (4)$$

where μ^{ind} and $\tilde{\mu}^{ind}$ are the induced and the conjugate induced dipoles at the polarization point k. The total electric field acting on each fragment now consists of static and induced fields (due to the multipoles and induced dipoles of the other fragments), as well as fields due to the electron density and nuclei of the ab initio part. The total polarization energy of the QM/EFP1 system is

$$E^{pol} = -\frac{1}{2}\sum_k\sum_{\alpha}^{x,y,z}\mu_\alpha^k(F_\alpha^{mult,k} + F_\alpha^{nuc,k}) + \frac{1}{2}\sum_k\sum_{\alpha}^{x,y,z}\tilde{\mu}_\alpha^k F_\alpha^{ai,k} \quad (5)$$

where $\mathbf{F}^{mult}$ is the field due to the static multipoles on the EFP fragments, and $\mathbf{F}^{ai}$ and $\mathbf{F}^{nuc}$ are fields due to the density and the nuclei of the quantum region. Polarization in a QM/EFP system is computed self-consistently using a two-level iterative procedure. The lower level, which takes care of the convergence of the induced dipoles, is identical to the fragment-only system. The wave function is kept frozen at this level. The lower-level iterative procedure exits when the induced dipoles are self-consistent and are consistent with the frozen ab initio wave function. At the higher level, the wave function is updated based on the converged values of the induced dipoles from the lower level. Convergence of the upper level is determined by convergence of wave function parameters (molecular orbital coefficients). Convergence of the two-level iterative procedure yields self-consistent induced dipoles and the ab initio wave function.

The remaining contribution to the ab initio-EFP interactions accounts for exchange-repulsion and charge transfer effects. This V^{rep} term is modeled in the form of Gaussian functions

$$V_k^{rep}(x) = \sum_{i=1}^{2}\beta_{i,k}\exp(-\alpha_{i,k}r_{kx}^2) \quad (6)$$

where k represents repulsion centers (atoms and center of mass), r_{kx} is the distance between a repulsion center and an electron. Parameters α and β are obtained by fitting the repulsion term to the difference between the HF energies of water dimers and a sum of electrostatic and polarization terms. A total of 192 water dimers were used in the fitting procedure. The fragment-fragment repulsion energy E^{rep} is modeled similarly to the ab initio-EFP1 repulsion. However, a single exponential function has been used instead of Gaussians.

The EFP1 water model has been combined with various wave functions in the quantum region, including HF, DFT, MCSCF and MP2. Excited state methods such as configuration interaction with single excitations (CIS[176]), TDDFT,[177] and MRPT[178] have been interfaced with EFP1 as well. The general EFP2 model (see below) is interfaced with CIS(D)[179] (CIS with perturbative double excitations) and equations of motion (EOM)-CCSD.[180,181] Additionally, the EFP1 model has been combined with the FMO method.[182,183] The interface between the EFP1 method and the polarizable continuum model (PCM) has also been developed.[184−187] To describe heterogeneous catalytic systems in the presence of a solvent (the liquid-surface interface), an interface between the EFP method and the universal force field has been developed.[188] The implementations of the EFP1 and QM/EFP1 methods have been efficiently parallelized.[189]

Analytic gradients have been developed for the EFP1 and QM/EFP1 models[172,173] and the gradients have been used extensively in studies of water clusters and bulk,[190−193] aqueous reactions,[194,195] amino acid neutral-zwitterion equilibria,[184,196] and photochemistry in water.[176,177]

To summarize, the EFP1 water potential described above has proven to be a robust and useful model for treating interactions in water. However, the main drawback of the original version of the EFP1 model is that it is fitted to HF or DFT energies of water dimers. Therefore, this potential does not include long-range correlation effects such as dispersion. In order to improve the accuracy of the water model, two more variants of EFP1 have been developed. In one, referred to as EFP1/DFT,[197] the B3LYP functional was used to obtain EFP Coulomb and polarization parameters, and the repulsion term was fitted to reproduce the B3LYP/DH(d,p) energies in dimers. In the EFP1/MP2 variant,[198] the potential is based on the MP2 method, that is, the multipoles and polarizability tensors are obtained from MP2 calculations on a water monomer, and the repulsion is fitted to HF dimer energies. To incorporate correlation effects, an additional dispersion term of the form $E^{disp} = C_6/R^6 + C_8/R^8$ is added to fragment-fragment interaction energies. The C_6 and C_8 coefficients were obtained from a fit to MP2 correlation energies of the dimers. The original HF-based potential is now referred to as EFP1/HF.

To extend the EFP idea to a general solvent, one needs to design a way to accurately model exchange-repulsion and charge-transfer interactions without extensive fitting. Moreover, the dispersion term must be added to the model for a balanced description of π-stacking, van der Waals complexes, and other weak interactions. This has been realized in a general formulation of the EFP method, called EFP2.[77,199,200] The fragment-fragment energy in the EFP2 method consists of the following terms:

$$E^{EFP2} = E^{Coul} + E^{pol} + E^{exrep} + E^{disp} + E^{ct} \quad (7)$$

where E^{exrep} is a general (not-fitted) exchange-repulsion energy, E^{disp} and E^{ct} are dispersion and charge-transfer energies, respectively. The EFP2 potential can be built for any solvent molecule in a MAKEFP run in GAMESS. However, solvents with flexible degrees of freedom (for example, long alcohols or chain hydrocarbons) should be treated with care since EFP is inherently a rigid-body model.

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

The Coulomb E^{Coul} and polarization E^{pol} terms in the EFP2 method are obtained similarly to those in the EFP1 models. That is, the multipoles at the atoms and bond-midpoints are generated by the DMA, and distributed polarizability tensors are calculated at the LMO centroids by using the coupled perturbed HF (CPHF) approach. Since the general EFP2 potential is intended to be used for various species, including biologically relevant highly polar and charged fragments, such as NH_4^+, H_3O^+, OH^-, F^-, Cl^-, metal cations, etc., special care should be taken for short-range and quantum effects. Therefore, several damping formulas have been developed and investigated in combination with the Coulomb and polarization energies.[201−203] In particular, Coulomb interactions at short distances can be damped either by an exponential damping function applied to charge−charge term only[201]

$$f_{kl}^{\mathrm{ch-ch}} = 1 - \frac{\alpha_l^2}{\alpha_l^2 - \alpha_k^2}\exp(-\alpha_k R_{kl}) - \frac{\alpha_k^2}{\alpha_k^2 - \alpha_l^2}\exp(-\alpha_l R_{kl}) \tag{8}$$

or by using a set of exponential-based damping functions applied to higher-order Coulomb terms as well (charge-dipole, charge-quadrupole, dipole−dipole, dipole−quadrupole).[202,203] Damping parameters α at each multipole distributed point (k or l in the above equation) are obtained in a MAKEFP run by fitting the damped classical Coulomb potential to the quantum (Hartree−Fock) potential on a set of points around the fragment. Another option is to estimate the short-range charge penetration energy using the spherical Gaussian overlap (SGO) approximation:[204,203]

$$E_{kl}^{\mathrm{pen}} = -2\left(\frac{1}{-2\ln|S_{kl}|}\right)^{1/2}\frac{S_{kl}^2}{R_{kl}} \tag{9}$$

Damping of the polarization energy is extremely important to avoid the "polarization catastrophe" that may happen due to breaking the multipole approximation at short separations between the fragments. Exponential and Gaussian damping functions have been investigated;[203] damping parameters for these functions (one parameter per polarization point, i.e., a LMO centroid) have been set to predefined values that were picked to be the same for all fragment types and all polarization points.

The dispersion energy between fragments is calculated using the leading induced dipole−induced dipole $1/R^6$ term.[205] An empirical correction for the $1/R^8$ term is added as one-third of the $1/R^6$ term. The resulting formula for the fragment−fragment dispersion energy is

$$E^{\mathrm{disp}} = \frac{4}{3}\sum_{k,l}\frac{C_{6,kl}}{R_{kl}^6} \tag{10}$$

where k and l are distributed (LMO) dispersion points on fragments A and B, respectively (A ≠ B), R_{kl} is the distance between these points, and C_6 is obtained using the following integration:[205,84]

$$C_{6,kl} = \int_0^\infty \overline{\alpha}_k(i\nu)\overline{\alpha}_l(i\nu)\mathrm{d}\nu \tag{11}$$

where $\overline{\alpha}_{k(l)}$ is the 1/3 of the trace of the dynamic (frequency dependent) polarizability tensor at the point k (or l), respectively. The integration is performed on the fly between all pairs of dispersion points of all fragments, using a 12-point Gauss-Legendre quadrature. Distributed dynamic polarizability tensors on each fragment are obtained from time-dependent HF calculations within the MAKEFP run in GAMESS.

Dispersion interactions are corrected for charge penetration effects. Two variants of damping functions have been employed within the EFP2 method. One is the Tang-Toennies damping formula[206] with damping parameter set to $\beta = 1.5$:[205]

$$f_{kl}^{\mathrm{TT}} = \left(1 - \exp(-\beta R_{kl})\sum_{n=0}^{6}\frac{(\beta R_{kl})^n}{n!}\right) \tag{12}$$

Another damping formula[203] uses the intermolecular overlap integrals S_{kl} between LMOs k and l on fragments A and B:

$$f_{kl}^{S} = 1 - S_{kl}^2(1 - 2\ln|S_{kl}| + 2\ln^2|S_{kl}|) \tag{13}$$

The latter formulation is parameter free and was shown to provide accurate dispersion energies up to very short (almost valence) separations between the fragments.

The exchange-repulsion energy in EFP2 is derived from the exact HF expression for the exchange-repulsion energy of two closed-shell molecules. Truncating the sequence at the quadratic term in the intermolecular overlap and applying an infinite basis set and the spherical Gaussian overlap approximation[207] lead to the following expression for the exchange-repulsion energy:[115,208]

$$E^{\mathrm{exrep}} = \sum_{k,l}\left[-4\sqrt{\frac{-2\ln|S_{kl}|}{\pi}}\frac{S_{kl}^2}{R_{kl}} - 2S_{kl}\Big(\sum_{i\in\mathrm{A}}F_{ki}^{\mathrm{A}}S_{il} + \sum_{j\in\mathrm{B}}F_{lj}^{\mathrm{B}}S_{jk} - 2T_{kl}\Big) + 2S_{kl}^2\left(-\sum_{j\in\mathrm{B}}\frac{Z_J}{R_{kJ}} - \sum_{I\in\mathrm{A}}\frac{Z_I}{R_{Il}} + 2\sum_{j\in\mathrm{B}}\frac{1}{R_{kj}} + 2\sum_{i\in\mathrm{A}}\frac{1}{R_{il}} - \frac{1}{R_{kl}}\right)\right] \tag{14}$$

where i, j, k, and l are LMOs; I and J are nuclei, S and T are the intermolecular overlap and kinetic energy integrals, respectively; and F is the intramolecular Fock matrix. The overlap and kinetic energy integrals are calculated on the fly between each pair of LMOs on different fragments A and B. The intramolecular Fock matrix elements are precalculated as part of the MAKEFP run.

The charge transfer fragment-fragment energy E^{ct} is derived by considering the interactions between the occupied valence molecular orbitals on one fragment with the virtual orbitals on another fragment.[209,210] The charge transfer term results in significant energy-lowering in polar or ionic species. An approximate formula, based on a second-order perturbative treatment of the intermolecular interactions, uses canonical HF orbitals of individual fragments and a multipolar expansion of the electrostatic potential V of the fragment. The charge transfer energy of fragment A induced by fragment B is approximated as

$$E_{\mathrm{A(B)}}^{\mathrm{ct}} = 2\sum_{i}^{\mathrm{occA}}\sum_{n}^{\mathrm{virB}}\frac{V_{in}^{\mathrm{B}} - \sum_{m}^{\mathrm{allA}}S_{nm}V_{im}^{\mathrm{B}}}{(1 - \sum_{m}^{\mathrm{allA}}S_{nm}^2)(F_{ii}^{\mathrm{A}} - T_{nn})}\Big[V_{in}^{\mathrm{B}} - \sum_{m}^{\mathrm{allA}}S_{nm}V_{im}^{\mathrm{B}} + \sum_{j}^{\mathrm{occB}}S_{ij}(T_{nj} - \sum_{m}^{\mathrm{allA}}S_{nm}T_{mj})\Big] \tag{15}$$

where V^{B} is the electrostatic potential on fragment B; i, j, n, and m are canonical orbitals. A similar expression for the energy $E_{\mathrm{B(A)}}^{\mathrm{ct}}$ is obtained when fragment A induces charge transfer in fragment B. Charge transfer is treated as an additive pairwise interaction, so the total charge transfer energy is a sum of the $E_{\mathrm{B(A)}}^{\mathrm{ct}}$ and $E_{\mathrm{B(A)}}^{\mathrm{ct}}$ terms

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

of all pairs of fragments A and B. The charge transfer term is the most computationally expensive term in EFP2; therefore, since it is relatively small in nonpolar or weakly polar systems, the charge transfer interaction is often omitted in EFP2 calculations.

Exact analytic gradients are available for all EFP2 fragment-fragment terms.[211,212] Similar to the EFP1 models, the general EFP2 potential can be used in Monte Carlo and molecular dynamics simulations. It was also shown that in Monte Carlo simulations, EFP can serve as an accurate importance function for a QM/EFP potential.[213] The general EFP2 model is also used in combination with the SFM, as discussed below. A hybrid EFP-FMO method, called EFMO, is also described below.

A general EFP2 ab initio-fragment interface is under development. The Coulomb and polarization quantum-EFP terms are treated similarly to the QM-EFP1 scheme, through inclusion of one-electron EFP terms into the QM Hamiltonian. Recently, a formula for the exchange-repulsion energy between quantum and EFP regions has been derived and implemented.[214] Implementation of the exchange-repulsion gradients is in progress, as is the formulation of the ab initio-EFP dispersion interaction.

To extend the general EFP model to covalently bound systems, a covalently bound ab initio/EFP interface has been developed.[215] The method is similar in spirit to that of Assfeld and Rivail[216] and is based on defining a buffer region that separates the QM and EFP parts. The buffer region consists of several LMOs that are kept frozen in the EFP calculations. In order to avoid variational collapse of the ab initio wave function into the buffer region, MOs of the ab initio part are kept orthogonal to the buffer LMOs. The buffer region creates a necessary separation between the ab initio and EFP regions such that the QM−EFP interactions can be considered to be nonbonded interactions and can therefore be treated in the same way as other EFP−QM interactions. Using the buffer region does not increase the cost of the QM/EFP calculations.

A variant of a flexible EFP potential has been developed by Nemukhin and co-workers[217,218] in which a covalently bound system (such as a polypeptide) is divided into small EFP fragments. Ab initio−EFP interactions are described in a similar manner to those in a QM−EFP1 model (with Coulomb, polarization, and repulsion terms), while fragment−fragment interactions are described at the level of a customary force field. This allows flexibility in a covalently bound molecule and optimization of internal degrees of freedom.

2.1.2. Explicit Polarization Potential. The original formulation of X-Pol by Gao[48,54] was called MODEL and formulated for homogeneous clusters representing liquids. In 2007, MODEL was extended by Xie and Gao[219] to covalently connected fragments with the use of the generalized hybrid orbital method (GHO).[220] From the beginning, X-Pol was envisioned as the next generation force field, based on QM calculations of fragments, and it was suggested that some parameters are needed for the implementation. X-Pol is typically used with semiempirical methods that describe fragment wave functions. The distinct feature of the X-Pol method is the incorporation of many-body polarization through performing a double SCF calculation; that is, self-consistent calculations of fragments in the electrostatic field of the other fragments (similar to the EFP approach). In the X-Pol method, the electrostatic field is usually represented by the field of Mulliken charges (often scaled by an empirical factor).

The basic equations of X-Pol can be summarized as follows. The total energy of the system, E_{tot}, divided into N fragments, is defined relative to the sum of the energies of noninteracting fragments E_I^0

$$E_{\text{tot}} = \langle \Phi | \hat{H} | \Phi \rangle - \sum_{I=1}^{N} E_I^0 \tag{16}$$

where the total Hamiltonian is given by

$$\hat{H} = \sum_{I=1}^{N} \hat{H}_I^0 + \frac{1}{2} \sum_{I=1}^{N} \sum_{J \neq I} \hat{H}_{IJ} \tag{17}$$

and the total wave function Φ is the Hartree product of fragment wave functions Ψ_I.

$$\Phi = \prod_{I=1}^{N} \Psi_I \tag{18}$$

$\hat{H}_I^0$ is the electronic Hamiltonian of isolated fragment I, and the interaction potential is expressed as

$$\hat{H}_{IJ} = -\sum_{i=1}^{2M} \sum_{\beta=1}^{B} \frac{q_\beta^J}{|\mathbf{r}_i^I - \mathbf{r}_\beta^J|} + \sum_{\alpha=1}^{A} \sum_{\beta=1}^{B} \frac{Z_\alpha^I q_\beta^J}{|\mathbf{r}_a^I - \mathbf{r}_\beta^J|} + E_{IJ}^{\text{vdW}} \tag{19}$$

The first term in eq 19 describes the interaction of electrons i in fragment I with partial charges q on atoms β in fragment J. The second term in eq 19 describes the interaction of nuclear charges Z on atoms α in fragment I with partial charges on atoms β in fragment J. The third term describes the Van der Waals (vdW) interaction energy usually represented by the Lennard-Jones potential between atoms in fragments I and J.

In applying the X-Pol method, one performs SCF calculations on the individual fragments in the electrostatic field of the point charges due to the other fragments, determined using the electron density of the other fragments, until full self-consistency is reached. The fragment boundaries are treated with frozen orbitals, whereby the four bonds on the typical boundary carbon atom are divided equally (meaning two bonds to each fragment) between two fragments. The first group of orbitals, which is inside the fragment, is fully SCF-optimized, while the other group of orbitals that belong to other fragments is frozen and contributes to the total Fock matrix and density, while remaining frozen. Polypeptides are usually divided at Cα atoms, which are the boundary atoms in X-Pol. Thus, the fragments correspond to peptide units as defined by IUPAC[221] rather than residues.

In 2008, Xie et al.[222] modified the formulation of X-Pol to obtain rigorously analytic gradients. The modification was necessitated by the use of the Mulliken charges to represent the external field. In this method, individual fragment Fock operators are modified with additional terms derived from the variation of the total electronic energy with respect to the electronic states of all fragments. In 2009, Song et al.[223] extended the RHF version of X-Pol to DFT.

In 2010, Cembran et al.[224] proposed the X-Pol-X method, which adds interfragment exchange interactions. Different from the earlier X-Pol, this new formulation appears to belong to the two-step category in the classification of fragment-based methods. In the X-Pol-X formalism, the fragments in X-Pol are replaced by blocks, which may or may not be associated with specific groups of atoms, although the practical implementation still uses fragments as blocks. The blocks are described by block-localized wave functions (BLW). The total X-Pol-X

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

energy is given by

$$E^{\text{X-Pol-X}} = \sum_{\mu\nu}^{K} D_{\mu\nu}(\mathbf{H}_{\mu\nu} + \mathbf{F}_{\mu\nu}) \tag{20}$$

where K is the total number of basis functions in the whole system and $\mathbf{H}$ and $\mathbf{F}$ are the one-electron and Fock matrices, respectively, as in an ab initio calculation for the whole system. The density matrix is defined as

$$\mathbf{D} = \tilde{\mathbf{C}}(\tilde{\mathbf{C}}^T\mathbf{R}\tilde{\mathbf{C}})^{-1}\tilde{\mathbf{C}}^T \tag{21}$$

In eq 21, $\tilde{\mathbf{C}}$ is the matrix of occupied MOs, with nonzero blocks on the diagonal. Each block in the $\mathbf{D}$ matrix corresponds to the respective block in X-Pol-X. $\mathbf{R}$ is the matrix of the interblock AO-based overlap integrals (a and b denote blocks)

$$R_{\mu\nu}^{ab} = \langle\chi_\mu^a|\chi_\nu^b\rangle \tag{22}$$

The interfragment exchange is apparently accounted for through the block-orthogonalization constraint $(\tilde{C}^T R\tilde{C})^{-1}$, although the interblock potential is described with Coulomb terms.

Practically, one still has to solve Hartree−Fock equations for blocks, with the appropriately constructed Fock matrix.[224] In X-Pol-X, in contrast to X-Pol, the free block energies are not subtracted from the QM energy as in eq 16. This suggests that X-Pol-X has become more sophisticated than a "next generation force field".

In 2010, Gao et al.[74] proposed the generalized X-Pol method, GX-Pol. This approach has two formulations, based on (a) consistent diabatic configurations (CDC) and (b) variational diabatic configurations (VDC). The two approaches have some similarity with (a) the valence bond (VB) (or MCSCF) and (b) CI theories, respectively. The analogy is not complete, since the VB, MCSCF, and CI methods all account for the occupation of virtual orbitals and describe electron correlation, whereas GX-Pol effectively accounts for quantum effects (such as charge transfer) between fragments in the RHF fashion. This GX-Pol approach is conceptually analogous to the MP2 and CI in vibrational SCF (VSCF),[225] which accounts for the anharmonicity rather than the electron correlation.

The CDC approach is based on pairs of blocks X2; the energy is obtained by

$$E_{\text{X2}} = \langle\Theta_{\text{X2}}|\hat{H}|\Theta_{\text{X2}}\rangle \tag{23}$$

where the wave function for M blocks is given by

$$\Theta_{\text{X2}} = \sum_{a=1}^{M}\sum_{b=a+1}^{M} c_{ab}\Psi_{ab}^{\text{A}} \tag{24}$$

The pair wave functions in eq 24 are

$$\Psi_{ab}^{\text{A}} = R_{ab}^{\text{A}}\hat{A}(\,\Phi_1 \quad \ldots \quad \Phi_{ab} \quad \ldots \quad \Phi_M\,) \tag{25}$$

where $\hat{A}$ is the antisymmetrizer and R_{ab}^{A} is the normalization constant. Φ_{ab} is the wave function of the pair of blocks (ab). The coefficients c_{ab} and other variables in the block wave functions are determined self-consistently. An extension to an arbitrary combination of blocks (i.e., more than pairs) was also proposed.

In the VDC approach, the energy E is obtained by diagonalizing a CI-like secular equation with each matrix element, such as $H_{I,J}$ corresponding to the combination of pairs $I = (ab)$ and $J = (rs)$

$$\begin{vmatrix} H_{11}\text{-}ES_{11} & \ldots & H_{1,M2}\text{-}ES_{1,M2} \\ \ldots & \ldots & \ldots \\ H_{M2,1}\text{-}ES_{M2,1} & \ldots & H_{M2,M2}\text{-}ES_{M2,M2} \end{vmatrix} = 0 \tag{26}$$

In eq 26, $M2 = M(M-1)/2$ is the total number of dimer configurations Ψ_{ab}^{A}, and the corresponding overlaps and Hamiltonian matrix elements are given, respectively, by

$$S_{ab,st} = \langle\Psi_{ab}^{\text{A}}|\Psi_{st}^{\text{A}}\rangle \tag{27}$$

$$H_{ab,st} = \langle\Psi_{ab}^{\text{A}}|\hat{H}|\Psi_{st}^{\text{A}}\rangle \tag{28}$$

It should be noted that CDC and VDC give different energies, like CI and MCSCF do for electron correlation. This new formulation of GX-Pol can describe charge transfer between blocks, previously not considered in X-Pol. As for X-Pol-X, GX-Pol is a more general method than the original X-Pol.

2.2. Fragment Molecular Orbital Methods

Since the original formulation[60] during the last century, the FMO method[226,227] has undergone considerable development extending the applicability of the method to a wide variety of types of systems. A review of the FMO method in 2007 by Fedorov and Kitaura[228] was followed by a book[229,230] summarizing the activity of a number of research groups developing and applying FMO.

The gist of the FMO method is to incorporate high order interactions into low order expansions in terms of fragments, inspired by molecular interaction models.[72] In particular, the electrostatic interaction is treated at the full N-body order ($M = N$), where N is the total number of fragments; this is accomplished by incorporating the Coulomb field of all N fragments into the self-consistent monomer SCF cycle (see introduction for a more detailed discussion), while nonelectrostatic interactions (exchange-repulsion, charge transfer, and dispersion) are treated at a lower order, typically $M = 2$ or 3. The fragments, their pairs ($M = 2$, FMO2) and triples ($M = 3$, FMO3) are explicitly treated quantum-mechanically in the presence of the Coulomb (i.e., electrostatic) field of the whole system. In this manner, the low-order nonelectrostatic interactions are coupled to the electrostatics. In most systems, this is a reasonable approximation, because the nonelectrostatic interactions are typically shorter-range.

Beginning with this assumption, it becomes acceptable to break a system into many smaller, localized pieces, treating the long-range effects of the full system using only a Coulomb operator. The driving motivation for this approximation and subsequent fragmentation is to find a route to drastically reduce the computational cost required for calculations. Even when the most efficient algorithms are used, for example, with MP2 or CCSD(T), the system size in terms of basis functions is the limiting factor. It is this roadblock in current ab initio methods that the FMO method attempts to overcome, first by reducing computational cost through fragmentation as mentioned above, and second by employing multilevel parallelization through the use of the generalized distributed data interface (GDDI).[1] It is this combination of theoretical approximations and exploitation of modern computational resources that has allowed the FMO method to perform all electron calculations on over 20 000 atoms.[231]

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

Fragment creation in the FMO method involves the detaching of bonds electrostatically, assigning two electrons from the broken bond to one fragment and none to the other. Where these bond detachments occur is left up to the user, relying on their own chemical intuition and knowledge of the system being investigated. A general guideline is to avoid detaching bonds involved in considerable electron delocalization, for example, within benzene rings. There are well established standard fragments for polypeptides (fragmented at Cα atoms), nucleotides, and saccharides. Metal ions have very considerable charge transfer to solvent[232] and represent an example of the need to define large fragments. When fragmenting a chemical system in this manner, the charge on fragments is not affected by bond detachment, so a negatively charged residue retains this same negative charge after fragmentation. The technical details of this are well illustrated elsewhere.[229,199] Note that the FMO, ELG, and X-Pol methods appear to be among the few fragment methods that do not add caps (hydrogen or more involved) to detached bonds.

The detachment of bonds in the FMO method can be done in two ways. In the original method, hybrid orbital projection (HOP) operators[233] are used to properly divide the variational space for the atom from which a bond is detached, because this atom has to be present in two fragments: This atom is present in one fragment just to describe the detached bond; the same atom is present in the second fragment since it contains the rest of the electrons that "belong" to that atom. This procedure is readily accomplished with projection operators, based on precomputed hybridized orbitals (e.g., sp^3) for the atom at the detached ends of the fragmented bond. In this method, often referred to as the HOP method, there are no restrictions on the variational optimization of the occupied orbitals of the fragments, in contrast to the second approach described in the next paragraph. This HOP method is considered to be a good approach for very polar systems such as proteins with charged residues.

The second detachment method is based on adaptive frozen orbitals (AFO),[234] in which the electron density distribution is precomputed for an appropriate MO in a model system that represents the system of interest. This precomputed electron density is subsequently frozen in the FMO calculation. The treatment of the variational space division for each bond-detached atom in a fragmentation process is very similar to the HOP procedure, although the AFO procedure is achieved in practice in a different manner by Fock matrix transformations. However, an important distinction in the AFO procedure is that there is a restriction on the variation of the occupied fragment orbitals on the atoms that are involved in the fragmentation (the detached bond orbitals). This is particularly important for systems in which the detached bonds are inevitably close to one another. This happens in covalent crystals and related systems such as nanowires, for which the AFO procedure is preferred. The HOP method is based on projection operators in the general form $|\phi\rangle\langle\phi|$. Such projection operators are used in many methods, for instance, in the core operator of model core potentials.[235] The FMO AFO approach was adapted from the EFP method,[75,215] which is in turn related to other methods, such as the GHO approach[220] of Assfeld and Rivail.[216] The AFO method is similar to these other methods in that the density of some occupied orbitals is frozen, but differs in the variational space division, as well as in other technical details such as the method to obtain the bond orbitals, and their orthogonality.

HOP-based approaches require a separate preliminary generation of hybrid orbitals. Although this is straightforward, it does add an extra step. To date, bonds have been detached at carbon and silicon[236] atoms. On the other hand, in the AFO approach the frozen orbitals are generated on the fly. Therefore, in principle any system can be computed, although a practical implementation may require an elaborate algorithm to properly build a model system in complicated cases. The AFO-based FMO approach has been applied to bonds detached at carbon and silicon.

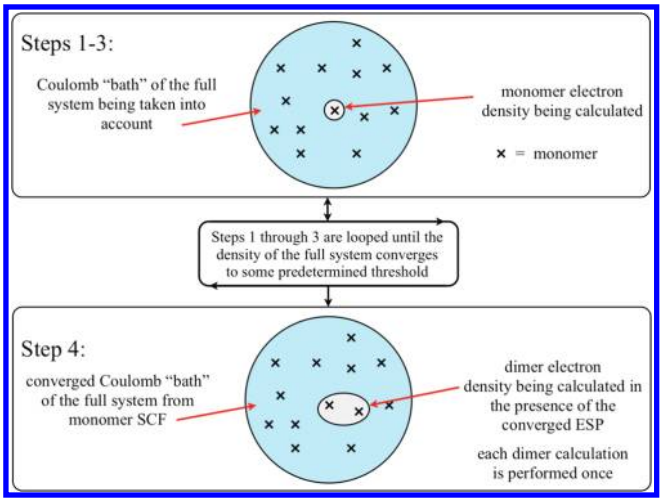


Figure 2. General outline of a standard FMO calculation including up to dimer interactions.

Now, consider a summary of the basic FMO formulation. For a more detailed mathematical description the reader is referred to chapter 2 of ref 144. The basic algorithm for calculating the energy of a system using the FMO method is as follows (see Figure 2):

(1) The initial electron density distribution is calculated for each monomer.
(2) The monomer Fock operators are then constructed using these densities and the energy of each monomer is calculated in the Coulomb bath of the rest of the system.
(3) Each of the monomer energies is iterated to self-consistency, leading to a converged ESP. This step in FMO development is usually called the monomer SCF or SCC.
(4) Fragment dimer calculations (FMO2) are then performed in the converged ESP of the rest of the system. Each dimer calculation is only performed once (not self-consistently).
(5) Optionally, fragment trimer calculations (FMO3) are performed next in the converged ESP of the rest of the system. Each trimer calculation is only performed once.

Within the two- and three-body FMO methods (FMO2 and FMO3, respectively), the total energy of the system can be written as

$$E^{\mathrm{FMO2}} = \sum_{I}^{N} E_I + \sum_{I>J}^{N} (E_{IJ} - E_I - E_J) \tag{29}$$

$$E^{\mathrm{FMO3}} = E^{\mathrm{FMO2}} + \sum_{I>J>K}^{N} \{(E_{IJK} - E_I - E_J - E_K) - (E_{IJ} - E_I - E_J) - (E_{IK} - E_I - E_K) - (E_{JK} - E_J - E_K)\} \tag{30}$$

with monomer (I), dimer (IJ), and trimer (IJK) energies being obtained as described above. The beauty of the seemingly

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

simplistic form of eq 29 is its ability to encapsulate the concepts of properly handling many-body effects, shown in the diagrammatic treatment[56] and the Green's function formalism.[237]

The Fock equation used in the FMO method

$$\tilde{\mathbf{F}}^x\mathbf{C}^x = \mathbf{S}^x\mathbf{C}^x\,\tilde{\boldsymbol{\varepsilon}}^x,\ x = I, IJ, IJK \tag{31}$$

$$\tilde{\mathbf{F}}^x = \tilde{\mathbf{H}}^x + \mathbf{G}^x \tag{32}$$

is a modified version of the standard form. The modification adds a term, $V^x_{\mu\nu}$, to the one-electron Hamiltonian, $\tilde{\mathbf{H}}^x$, that represents the ESP calculated during the monomer SCF

$$\tilde{H}^x_{\mu\nu} = H^x_{\mu\nu} + V^x_{\mu\nu} + B\sum_i \langle\mu|\phi^h_i\rangle\langle\phi^h_i|\nu\rangle \tag{33}$$

To properly divide basis functions across fractioned bonds, a projection operator (HOP), $B\Sigma_i\langle\mu|\phi^h_i\rangle\langle\phi^h_i|\nu\rangle$, is also added to the Hamiltonian, where i runs over hybrid orbitals φ^h_i on all bond detached atoms in fragment x. The constant B is chosen to be sufficiently large to remove the corresponding orbitals out of the variational space. Typically, $B = 10^6$ a.u.

One of the most important components of the FMO approach is the ESP, calculated during the monomer SCF process. The ESP of the full system takes the form

$$V^x_{\mu\nu} = \sum_{K(\neq x)} (u^K_{\mu\nu} + v^K_{\mu\nu}) \tag{34}$$

$$u^K_{\mu\nu} = \sum_{A\,\in\,K} \langle\mu|(-Z_A/|\mathbf{r}-\mathbf{r}_A|)|\nu\rangle \tag{35}$$

$$v^K_{\mu\nu} = \sum_{\lambda\sigma\,\in\,K} D^K_{\lambda\sigma}(\mu\nu|\lambda\sigma) \tag{36}$$

The term $u^K_{\mu\nu}$ represents the nuclear attraction contribution to the energy, with the two-electron contribution represented by the second term $v^K_{\mu\nu}$, expressed in terms of AOs μ and ν. Both of these terms are calculated for each of the surrounding monomers K with electron density D^K.

One might wonder if the addition of exchange integrals to the ESP in eq 36 could improve the results. Following an initial analysis[237] in 2006, it was later concluded[238] that because of the lack of orthogonality between fragment wave functions, the addition of exchange to ESPs in the SCC procedure actually decreases the accuracy. It is also possible to use point charges for all ESPs. This was shown to work particularly well for large basis sets by Fedorov et al.[239] These authors also proposed to screen point charges by introducing an exponent prefactor to the point charge potentials in eq 35. Such screening was found to be useful in particular for the AFO scheme when ESP are represented by point charges, and for molecular clusters the effects of screening were found to be small.

Even using the formulation described thus far, the computational cost can still become excessive due to the increasing cost of the two-electron term, $v^K_{\mu\nu}$, contained in the ESP. Additional approximations can be used to reduce this cost by creating a cutoff value R_{app}. Two-electron terms outside of this boundary can be treated in a more approximate manner; however, this additional layer of complexity can decrease the accuracy of the method. This is due to the delicate balance among the approximations in different FMO terms.

To understand this, consider a dimer IJ separated from a fragment L, which exerts an ESP upon I, J, or IJ, schematically represented as $IJ...L$. If the distance between J and L is short (no ESP approximation), but that between I and L is long enough to apply the ESP approximation, there is a loss of balance, because in dimer IJ and monomer J the ESP due to L is computed without approximations, but in monomer I it is approximated. Consequently, pair corrections of the form $E_{IJ} - E_I - E_J$ include potentials due to L both with and without approximations. Because of the rapidly increasing number of dimers with system size, the error accumulates, creating a significant loss of accuracy in the total energy of the system.

To avoid the problems created by using distance based approximations to the ESP, a reformulation of the energy expression is used[240] that is equivalent to eq 29 but more accurate when using approximations to the ESP

$$E^{\text{FMO2}} = \sum_I^N E_I' + \sum_{I>J}^N \Delta E_{IJ}' + \sum_{I>J}^N Tr(\Delta\mathbf{D}^{IJ}\mathbf{V}^{IJ}) \tag{37}$$

with the analogous expression[241] for FMO3 being

$$\begin{aligned} E^{\text{FMO3}} = {} & \sum_I^N E_I' + \sum_{I>J}^N \Delta E_{IJ}' \\ & + \sum_{I>J>K}^N (E_{IJK}' - E_I' - E_J' - E_K' - \Delta E_{IJ}' - \Delta E_{IK}' - \Delta E_{JK}') \\ & + \sum_{I>J}^N Tr(\Delta\mathbf{D}^{IJ}\mathbf{V}^{IJ}) + \sum_{I>J>K}^N [Tr(\Delta\mathbf{D}^{IJK}\mathbf{V}^{IJK}) \\ & - Tr(\Delta\mathbf{D}^{IJ}\mathbf{V}^{IJ}) - Tr(\Delta\mathbf{D}^{IK}\mathbf{V}^{IK}) - Tr(\Delta\mathbf{D}^{JK}\mathbf{V}^{JK})] \end{aligned} \tag{38}$$

where $\Delta E'_{IJ} = E'_{IJ} - E'_I - E'_J$ and the new energy terms E'_x are defined as the internal n-mer energies with the ESP contributions subtracted out

$$E_x' = E_x - Tr(\mathbf{D}^x\mathbf{V}^x), x = I, J, IJ \tag{39}$$

and $\Delta\mathbf{D}^x$ is the difference density matrix, defined as

$$\Delta\mathbf{D}^{IJ} = \mathbf{D}^{IJ} - \mathbf{D}^I \oplus \mathbf{D}^J = \mathbf{D}^{IJ} - \begin{bmatrix} \mathbf{D}^I & 0 \\ 0 & \mathbf{D}^J \end{bmatrix} \tag{40}$$

The above formulation in eq 37 allows the total energy to be calculated explicitly from only the dimer ESP $\mathbf{V}^{IJ}$. Equation 38 contains both dimer and trimer ESPs, resulting in a more considerable effect of approximations on FMO3. Approximations can be applied to monomers and dimers separately, with the dimer ESP directly contributing to the total energy. The monomer ESP determines the monomer electron densities, and therefore only contributes to the total energy indirectly.

The total density can be calculated in the same manner as the total energy (cf., eq 29), and one-electron properties such as multipole moments are straightforward to calculate once the total density is obtained. For example, Sekino et al. computed moments up to quadrupoles.[242]

Equation 38 includes both the two-body $\Delta\mathbf{D}^{IJ}$ and three-body $\Delta\mathbf{D}^{IJK} = \mathbf{D}^{IJK} - \mathbf{D}^I \oplus \mathbf{D}^J \oplus \mathbf{D}^K$ difference density matrices. This form of the energy equation, when applying three-body approximations to the ESP, partially retains the problem of

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

ESP imbalance that the reformulation attempts to avoid. This retention of error is due to the uniform application of the distance definition R_{app}. To avoid this the distance based approximation was also reformulated[241] to treat the matrix elements of $V^{x}_{\mu\nu}$ block-wise so that

$$\tilde{R}_{IJ,L}(\mu,\nu) = \left\{ \begin{array}{lll} R_{I,L} & \text{for} & \mu,\nu \in I \\ R_{J,L} & \text{for} & \mu,\nu \in J \\ R_{IJ,L} & \text{for} & \mu \in I, \nu \in J \text{ or } \mu \in J, \nu \in I \end{array} \right\} \quad (41)$$

This allows for the application of approximations, for example in the I block of $V^{x}_{\mu\nu}$, to use the distance $R_{I,L}$, while using the uniform distance definition $R_{IJ,L}$ for the off-diagonal elements corresponding to both fragments I and J.

It is now appropriate to discuss the types of approximations applied to the ESP using the previously described energy expressions. Two levels of approximation can be used, both of which are applied to the two-electron integral term in eq 34. The first level of approximation, applied at intermediate distances, is the Mulliken atomic orbital population to the two-electron integrals.[243,244] Eq 36 can be rewritten as

$$v^{K}_{\mu\nu} \cong \sum_{\lambda \in K} (\mathbf{D}^{K}\mathbf{S}^{K})_{\lambda\lambda}(\mu\nu|\lambda\lambda) \quad (42)$$

The application of this approximation can reduce the computational cost of the two-electron integrals by a factor of N_B (number of basis functions).

The second level of approximation uses fractional atomic charges Q_A derived from the Mulliken atomic populations P^A of the monomers. This approximation is applied at long distances and simplifies eq 36 as

$$v^{K}_{\mu\nu} \cong \sum_{A \in K} \langle\mu|(P_A/|\mathbf{r} - \mathbf{r}_A|)|\nu\rangle \quad (43)$$

effectively reducing the computational cost of the two-electron integrals by another factor of N_B.

For far separated monomers the corresponding dimers do not need to be computed with the SCF procedure, as the monomer states already polarized by the ESP are very little perturbed by an explicit dimer SCF. However, there is a Coulomb interaction between such monomers that should be computed and added to the total energy

$$E'_{IJ} \cong E'_I + E'_J + Tr(D^I u^{1,I(J)}) + Tr(D^J u^{1,J(I)}) + \sum_{\mu\nu \in I} \sum_{\rho\sigma \in J} D^I_{\mu\nu} D^J_{\rho\sigma} (\mu\nu|\rho\sigma) + \Delta E^{NR}_{IJ} \quad (44)$$

where $u^{1,I(J)}$ and $u^{1,J(I)}$ (cf. eq 35) are the one-electron Coulomb potentials of the force exerted by fragment J on fragment I, and vice versa. The electron−electron interaction (cf. eq 36) and the nuclear repulsion (NR) are also added in eq 44.

The final equations for the basic energy evaluation in FMO2 and FMO3 for RHF wave functions are eqs 37 and 38, respectively. These equations can be readily extended to other wave functions. A presentation of the application to other methods can be found in chapter 2 of ref 144.

The original energy gradient[245] based on the HOP approach was proposed for FMO2 in 2001 and contained several approximations. The derivative of Mulliken point charges in ESP was added by Nagata et al.[246] in 2009, and the derivative of the HOP terms was developed by Nagata et al.[247] in 2010. The FMO3 gradient was introduced in 2010 by Komeiji et al.[248] The FMO2 gradient for the AFO approach was proposed in 2009 by Fedorov et al.[249] A fully analytic FMO gradient without approximation was developed by Nagata et al.[250] in 2011 with the introduction of the self-consistent Z-vector (SCZV) procedure to obtain the exact derivatives of the dimer terms coupled to the electrostatic potential.

In 2002, Inadomi et al.[251] proposed the FMO-MO method, which according to the classification of fragment methods as discussed above, belongs to a different category from the rest of the FMO methods. While most FMO approaches are one-step approaches, the FMO-MO method is two-step, in which the density is obtained from the FMO calculation and used to construct the Fock matrix of the whole system in the exact ab initio fashion. Watanabe et al.[252] benchmarked the FMO-MO method for DNA, and Umeda et al.[253] proposed an efficient parallel algorithm for this method. It should be noted that the FMO-MO method (so far limited to RHF) can be used as a traditional ab initio method, because it produces the energy, MOs and other properties from the total Fock matrix.

Sekino et al.[242] in 2003 suggested taking the union of fragment MOs as a representation of the MOs of the whole system. Tsuneyuki et al.[254] in 2009 developed the FMO-LCMO method, which has the limited but important aim of producing MOs near the HOMO and LUMO. This is accomplished by constructing a small MO-based Fock matrix collecting monomer and dimer contributions for a few orbitals of interest. Nishioka and Ando[255] in 2011 took advantage of the FMO-LCMO scheme and used it for the theoretical modeling of the electronic coupling. Fedorov and Kitaura[238] in 2009 proposed to build the total AO-based Fock matrix for the whole system (FMO/F) from monomer and dimer Fock matrices. This is in contrast to the FMO-MO method, in which a similar matrix is computed in the ab initio fashion using the total density. Adding the exchange to this total Fock matrix (FMO/FX) considerably improves the description of the virtual MOs. All of these methods are FMO extensions for generating a specific property (the MOs), rather than a density-based method like FMO-MO, aimed at delivering the energy, MOs and other properties.

Sugiki et al.[55d] in 2003 began the process of extending FMO to wave functions other than RHF, by developing FMO2-based DFT, extended to FMO3 by Fedorov and Kitaura[256] in 2004; Shimodo et al.[257] in 2006 reported some accuracy benchmarks for FMO2-DFT. In 2004, Møller−Plesset perturbation theory was interfaced with FMO by Fedorov and Kitaura,[258] and an alternative implementation was reported by Mochizuki et al.[259,260] In 2009, RI-based FMO2-MP2 was proposed by Ishikawa and Kuwata[261] and Cholesky decomposition based FMO2-MP2 by Okiyama et al.[262] Mochizuki et al. developed an efficient parallel algorithm on the Earth Simulator for FMO-based MP2 in 2008[263] and MP3 in 2010.[264] The FMO3-based MP2 method was developed for the energy by Fedorov and Kitaura[265] in 2009 and the gradient by Mochizuki et al.[266] in 2011.

The first MD simulations using FMO were reported by Komeiji et al.[267] in 2003, indicating that the gradient accuracy problem (which has subsequently been solved by Nagata et al.[250]) requires larger fragments. Ishimoto et al.[268,269] in 2004−2005

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

used their Hamiltonian algorithm for FMO-MD. Dynamic fragmentation was proposed by Komeiji et al.[270] in 2009 to cope with the problems of bond breaking in FMO-MD. A review of FMO-MD methods was published by Komeiji et al.[271] in 2009. A path integral based FMO-MD was developed by Fujita et al.[272] in 2009. A FMO-MD method with periodic boundary conditions was developed by Fujita et al.[273] in 2011.

Fedorov and Kitaura[274] in 2005 developed FMO-based MCSCF and coupled cluster[275] methods. As an important step to improve the efficiency, Fedorov et al.[276] introduced the multilayer FMO (MFMO) method, in which the system can be divided into several layers, and each layer is allowed to have its own basis set and level of theory. The original MFMO formulation[276] suggested that layers can have a different many-body expansion, which was later realized by Mochizuki et al.[266] under the name of FMO(3)-MP2, which should also be called FMO3-RHF:FMO2-MP2 for the case when the two layers include all fragments. In a quest for the exact solution, Maezono et al.[277,278] in 2006−2007 proposed the FMO-based quantum Monte Carlo method.

One of the main targets of the FMO method is biochemical simulations, for which it is very important to properly consider solvent effects. The FMO-based PCM was developed for both the energy by Fedorov et al.[279] in 2006 and the gradient by Li et al.[280] in 2010. A Poisson−Boltzmann model was incorporated into FMO by Watanabe et al.[281] in 2010. For discrete models of explicit solvation, water can be described either as FMO or EFP fragments. The latter is due to the work of Nagata et al. for the FMO/EFP energy[182] in 2009 and gradient[183] in 2011.

Complementing the ground state methods, FMO-based MCSCF was the first excited state method, reporting accuracy for singlet and triplet calculations of solvated phenol.[274] Mochizuki et al. developed FMO-based CI with singles (CIS)[282] in 2005 and added the perturbative treatment of doubles in 2007.[283] The FMO2-based TDDFT method was proposed by Chiba et al.[284,285] in the gas phase in 2007 and solution[286] in 2008; the gas-phase gradient was developed[287] in 2009 and the FMO3-based TDDFT method was proposed[288] in 2010. For open shells, the ROHF-based MP2 and CC methods were developed by Pruitt et al.[289] in 2010.

All of these excited state methods are built around the central excited state fragment, which naturally contains the chromophore. In the FMO2 model, dimers that include the central fragment add explicit quantum effects, while other fragments only exert their electrostatic potential (ESP). In contrast to the ground state FMO, FMO1 for excited states is quite useful as an approximation to full FMO2, if one is interested in the excitation energies. The working equation is very similar to eq 29, with the difference that the energies of the central monomer and dimers that contain the central monomer correspond to the excited state, while other energies are for the ground state.

There is an important difference between MCSCF on the one hand, and CI or TDDFT on the other. In the former, ESPs are computed for the central excited state (described by MCSCF), while in the latter the ESPs correspond to the ground state (RHF or DFT, respectively). The FMO1-based CI method has been described as a multilayer approach with a single chromophore fragment in the second layer,[282] but it can also be thought of as a single layer FMO1-CI, because these two descriptions are identical. On the other hand, for TDDFT, one can define FMO1-RHF:TDDFT, which corresponds to performing TDDFT calculations for the chromophore in the field of the ESP computed at the RHF level. This is different from FMO1-TDDFT, in which the ESPs are computed with DFT.

To analyze the effect of isotopic substitutions by using wave functions for nuclei, Ishimoto et al.[290] in 2006 developed a multicomponent method. Auer et al.[291] in 2009 proposed the FMO-based nuclear-electronic orbital (NEO) method. Mochizuki et al.[292] in 2006 developed a method for calculating the dynamic polarizability in the FMO framework. An interface to model core potentials (MCP) for the treatment of heavy atoms was added in 2006 by Ishikawa et al.[293] The total electrostatic potential was used to obtain atomic charges fitted to it by Okiyama et al. in 2007[294] and 2009.[295] Sekino et al.[296] in 2007 implemented the calculation of NMR chemical shifts with FMO2. Gao et al.[297] in 2007 proposed an approach for evaluating NMR shifts from dimers, by using two fragmentations shifted with respect to each other, extended[298] later in 2010 to larger conglomerates of fragments but with a single fragmentation in the FMO1(merged) method.

The FMO method is suited to various analyses, as it provides information on fragments and their interactions that are naturally built into the method. In 2006, Amari et al.[299] developed the visualized cluster analysis of protein−ligand interactions based on the FMO method. Du and Sakurai[300] in 2010 proposed a multivariate analysis of properties of amino acid residues.

An important goal is to obtain more information than simple pair interaction energies (PIEs), also called interfragment interaction energies, IFIEs. Mochizuki et al.[301] in 2005 developed the configuration analysis for fragment interaction (CAFI), providing a means to extract for each orbital the stabilization component of the polarization and the charge transfer for IFIEs. Fedorov and Kitaura built EDA into the FMO method, developing the pair interaction energy decomposition analysis (PIEDA),[98] which decomposes PIEs into their electrostatic, exchange-repulsion, charge transfer, and dispersion components. A fragment interaction analysis based on local MP2 (FILM) was proposed by Ishikawa et al.[302] in 2007. FILM allows extracting orbital-based contributions to the electron correlation component of the interaction energy.

The counterpoise correction[303] for the basis set superposition error (BSSE) conceptually is a two-fragment method, so it is natural to attempt to include it into fragment methods. After the initial probing,[293,302] two methods[304,305] have been developed for BSSE corrections of the FMO energy in eq 29. The validity of these corrections remains to be seen, as it has not been shown that they indeed bring the results closer to the complete basis set limit. One should also be aware that careful tests[306] indicate that in particular for MP2 the CP correction does not improve the results in this sense, especially for small basis sets.

Geometry optimizations for the FMO method were reported by Fedorov et al.[307] in 2007. Subsequently, Ishikawa et al.[308] in 2010 suggested the partial energy gradient (PEG) method and Fedorov et al.[309] in 2011 introduced the frozen domain (FD) concept in their FMO/FD method. The focus of both of these methods is to perform an efficient geometry optimization of a part of the system. In the FMO/FD approach (Figure 3), the system is divided into several domains: frozen, polarizable and active. The electronic state of the frozen domain is computed only at the initial geometry, and the polarizable domain is recomputed for new geometries during geometry optimizations of atoms in the active domain. The main purpose of FMO/FD and its faster derivative FMO/FDD (in which the number of dimer calculations

in the polarizable domain is reduced by omitting pairs of fragments except those in the active domain) is to efficiently perform geometry optimizations of the active sites in large molecular systems. To demonstrate the efficiency of the FMO/FDD approach, Fedorov et al.[309] optimized the geometry of a large protein−ligand complex with 19471 atoms at the B3LYP/6-31G(d) level of theory for the polarizable domain. This calculation took 32 h on six dual-CPU quad-core 2.83 GHz Xeon nodes.

There are several fragment methods that are closely related to the FMO method. The electrostatically embedded many-body expansion (EEMB) method[310−313] corresponds to FMO with constant ESP, which is not updated in the SCC fashion, and is often taken to be represented by point charges from molecular mechanics. Hirata et al.[314] in 2005 proposed representing the ESP by dipoles and in 2008[315] by point charges fitted to reproduce the fragment electrostatic potential.

In the effective fragment molecular orbital (EFMO) method Steinmann et al.[316] merged the ideas of EFP and FMO (Scheme 1). In their method, no SCC calculation is performed, and all fragments are computed in vacuum (i.e., in the absence of the ESP). The many-body polarization is added, computed from fragment polarizabilities (as in the EFP method). The long-range Coulomb interaction in eq 44 is computed using multipoles, also following the EFP method. Short range dimers (for which the interfragment distance R_{IJ} is smaller than a threshold R_{ESDIM}) are computed self-consistently, in vacuum.

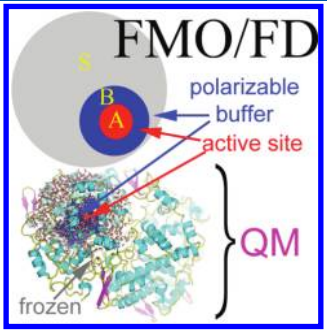


Figure 3. Schematic representation of FMO/FD and FMO/FDD. Fragments are divided into three domains, frozen (gray), polarizable (blue), and active (red). All three domains are computed with QM (using multilayer FMO). An actual geometry optimization of the partially solvated prostaglandin H(2) synthase-1 in complex with the reversible competitive inhibitor ibuprofen (PDB 1EQG) is shown at the lower part. Reproduced with permission from the TOC graphic of *J. Phys. Chem. Lett.* **2011**, *2*, 282−288. Copyright 2011 American Chemical Society.

The EFMO energy expression is given by

$$E = \sum_{I=1}^{N} E_I^0 + \sum_{\substack{I>J \\ R_{IJ} \le R_{\mathrm{ESdim}}}}^{N} (E_{IJ}^0 - E_I^0 - E_J^0 - E_{IJ}^{\mathrm{ind}}) + \sum_{\substack{I>J \\ R_{IJ} > R_{\mathrm{ESdim}}}}^{N} \Delta E_{IJ}^{\prime \mathrm{ESM}} + E_{\mathrm{total}}^{\mathrm{ind}} \quad (45)$$

The superscript 0 indicates that the corresponding calculations are performed in vacuum. The total induction energy $E_{\mathrm{total}}^{\mathrm{ind}}$ is corrected for double counting by subtracting SCF dimer contributions E_{IJ}^{ind} (because SCF dimers have implicit polarization). The long-range electrostatic term $\Delta E_{IJ}^{\prime\,\mathrm{ESM}}$ of dimers, for which the interfragment separation is larger than R_{ESDIM}, is computed with multipoles, consistent with the approximate method of computing the energy of far separated dimers in FMO, eq 44.

In the EFP method, multipoles up to octopoles are used. In the EFMO method, the multipole expansion is truncated at the level of quadrupoles since higher multipoles become less important at large interfragment separations. Distributed multipoles and polarizability tensors for computing polarization terms are calculated on-the-fly for each individual fragment. The main differences between EFMO, FMO, and EFP are summarized in Scheme 1. The computational cost of the EFMO method is significantly less than the cost of a corresponding FMO calculation. The EFMO method employs explicit ab initio calculations on the closely lying dimers where quantum effects such as charge penetration and charge-transfer are important and the semiclassical EFP terms could become less accurate.

Several methods, which appeared earlier than EFMO, are related to it and differ in the details of computing EFP-related terms in EFMO. These methods include the polarizable multipole interaction with supermolecular pairs (PMISP)[317−319] and the approach by Beran.[320,321] PMISP involves capping fragments so it may be closer to methods like KEM rather than FMO, modified with the addition of the explicit polarization.

2.3. Molecular Tailoring Approach

Since the original formulation of the molecular tailoring approach (MTA)[121,322] the fragmentation scheme has undergone significant changes and improvements.[323,324] The general idea of the MTA is similar in nature to both MFCC and SFM, with aspects of the divide and conquer method

Scheme 1. Comparison of the Effective Fragment Molecular Orbital Method (EFMO) with EFP and FMO

Differences between EFMO and EFP	Differences between EFMO and FMO
1. The monomer energies are explicitly included in EFMO to give an estimate of the total energy of the system and allow internal geometries of the fragments to change.	1. In EFMO, monomer calculations are performed in the absence of the external field and done once.
2. In EFMO, short-range interactions between the fragments are computed *ab initio* rather than with EFP expression: $E^{EFP} = E^{Coul} + E^{pol} + E^{disp} + E^{exrep} + E^{ct}$	2. In EFMO, all long-range and many-body interactions are included through classical Coulomb and polarization terms.
3. In EFMO, long-range interactions are modeled by electrostatic Coulomb and polarization terms only; dispersion, exchange-repulsion and charge-transfer are considered to be negligible at large separations.	

dx.doi.org/10.1021/cr200093j |*Chem. Rev.* XXXX, XXX, 000–000

incorporated as well. The MTA has been tested on a number of systems[325−328] including large π-conjugated systems, for example, graphene sheets.[329] Further refinements of the method allow for geometry optimizations based on the adjustment of the cardinality of the fragments and their overlapping sections. This newest implementation of the MTA is appropriately named the cardinality guided molecular tailoring approach (CG-MTA).[323]

The current implementation of the MTA is highly automated, with the user only needing to specify two values in addition to providing the coordinates of the target system. The first value specified is the maximum fragment size in terms of atoms per fragment. The second value, *R*-goodness (R_g), is central to the MTA fragmentation scheme and will be discussed in greater detail below.

The general fragmentation process for the MTA begins with the specification of the maximum fragment size and R_g and proceeds as follows. (1) Create an initial set of fragments by centering a sphere of radius R_g at each atom and assigning all atoms falling within the sphere to the fragment. During this process care is taken not to break aromatic rings or double bonds. Additional atoms are included or excluded based on the maximum fragment size. (2) Fragments created in step one are merged according to their proximity, while staying within the maximum fragment size set initially. (3) Merging of fragments is performed recursively depending on the maximum overlap of nearest neighbor fragments. This recursive merging stops once all fragments reach the maximum fragment size. (4) The final set of fragments is checked for the respective R_g value of the included atoms. (5) Broken bonds are capped using hydrogen atoms positioned along the appropriate bond vectors. (6) All intersecting portions of the merged fragments are computed, with the sign of each contribution being set according to $(-1)^{K-1}$ where K = number of fragments involved in the intersection. (7) The energy expression for the fragmentation scheme is created and the total energy of the system is calculated.

The value R_g used for fragment definition defines the radius of the sphere centered at atom i in the molecule of interest. All atoms that fall within this sphere are chosen to be part of fragment F_i. Since a sphere of radius R_g is created around all atoms in the system of interest at the beginning of the fragmentation process, each atom is bound to appear in more than one fragment even after the merging process in step 2. When this occurs, such atoms are evaluated for how accurately the local environment is represented by measuring the R_g value of the atom in each fragment. The largest of these values is chosen to represent the R_g of the atom in that particular fragmentation scheme. This is important since the value of R_g measures the quality of the fragmentation scheme based on how well the chemical environment around atom i is represented. A larger value of R_g will give a better representation of the chemical environment, providing more accurate results. It has been shown that values of R_g in the range 3−4 Å is necessary to achieve accurate results. While the minimum value of R_g ensures that each atom will have a certain representation of the surrounding chemical environment, the maximum fragment size assures that fragments do not become too large based on the computational resources available. The use of spheres centered around atoms also allows for an inherent treatment of nonbonded effects. For example, in a system such as a protein with a three-dimensional structure, one can imagine atoms in parts of the system that are not connected to the atom center i falling within the sphere created by R_g. In such cases such, it is possible for only hydrogen atoms in the unconnected region to fall into the sphere. When this occurs, atoms bonded to these hydrogens are added to the fragment as well. In general, the fragmentation process is automated to account for such systems.

Consider a general molecule M that, after the complete fragmentation procedure, is broken into two fragments F_1 and F_2. The overlap of these two fragments, $F_1|F_2$, must be subtracted from the energy, giving the total energy expression for the system as

$$E_M = E_{F_1} + E_{F_2} - E_{F_1 \cap F_2} \tag{46}$$

This expression for the total energy holds true for any system with only single overlaps between fragments. In the case of nonlinear systems, where more than two fragments overlap in the same region, additional terms must be added or subtracted from the total energy. Taking these cases into consideration, a general expression for the total energy of a chemical system containing K fragments can be derived as

$$E_M = \sum_i^K E_{F_i} - \sum_{i>j}^K E_{F_i|F_j} + \ldots + (-1)^{K-1} \sum_{i>j>k}^K E_{F_i|F_j|\ldots|F_K} \tag{47}$$

During the course of energy calculations for all fragments, the individual fragment densities are obtained during the standard SCF procedure. The complete density matrix for the full system can be constructed using the density matrix from each fragment. To obtain the most accurate density matrix for the system, individual density matrix elements are chosen from fragments based upon the quality of the R_g value calculated for each atom during the fragmentation process.

The derivative of the energy expression eq 47 with respect to coordinates X^μ can be easily derived as

$$\frac{\partial E_M}{\partial X^\mu} = \sum_i^K \frac{\partial E_{F_i}}{\partial X_{F_i}{}^\mu} - \sum_i^K \frac{\partial E_{F_i|F_j}}{\partial X^\mu_{F_i|F_j}} + \ldots + (-1)^{K-1} \sum_i^K \frac{\partial E_{F_i|F_j|\ldots|F_K}}{\partial X^\mu_{F_i|F_j|\ldots|F_K}} \tag{48}$$

where $X_{F_i{}^\mu}$ refers to the nuclear coordinates of atom μ in fragment F_i and $X_{F_{i|}F_j|^\mu}$ refers to the coordinates of the overlapping section of fragments F_i and F_j. An analogous expression for the Hessian can be written as

$$\mathbf{H} = \sum_i^K \mathbf{H}_{F_i} - \sum_i^K \mathbf{H}_{F_i|F_j} + \ldots + (-1)^{K-1} \sum_i^K \mathbf{H}_{F_i|F_j|\ldots|F_K} \tag{49}$$

Additional capabilities have been added to the CG-MTA including correlated methods such as MP2 and RI-MP2,[330] course grained parallelization[323] and the calculation of vibrational frequencies.[331]

2.4. Kernel Energy Method

The original formulation of the kernel energy method (KEM)[332] was tested on a number of systems.[333−336] Further improvement to the method[337] allowed the KEM to be applied to

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

systems beyond its original limitation to peptides and polymers. Based loosely on the energy decomposition analysis, the kernel energy method takes into account fragment or "kernel" interactions up to fourth order terms.[338] More recent improvements have enabled the KEM to accurately model π-conjugated systems, most notably graphene.[339]

The purpose for developing the KEM was to perform quantum calculations on systems with biological importance, such as proteins. By considering a system to be composed of individual "kernels" with known atomic coordinates, the contributions of all kernels in a system can be combined to provide properties for the full system. In this way the energy and properties of large systems can be obtained even if calculations on the full system are impossible. A few simple rules govern a calculation using the KEM, the most important being the requirement that each atom must be present in some kernel once and only once.

Once a system has been divided into separate kernels, care must be taken to remove any dangling bonds at the periphery of the kernels through the use of hydrogen caps. After all dangling bonds are capped, each kernel energy is evaluated, followed by all "double kernel" calculations of nearest neighbor kernels. In the original method, only those kernels covalently bonded to one another were considered during double kernel calculations. However, during subsequent development, the method was modified to include separated kernels during the double kernel evaluation, providing a more accurate description of the systems tested.

The sum of kernel contributions to the total energy of a system can be written mathematically as

$$E_{\text{total}} = \sum_{i=1;j=i+1}^{n-1} E_{ij} - \sum_{i=2}^{n-1} E_i \tag{50}$$

E_{ij} represents the energy of two covalently bonded kernels and E_i is the energy of a single kernel. During further development, the definition of double kernel calculations was extended to include all double kernels, not only those connected by a covalent bond. The modified formula for the total energy is then written as

$$E_{\text{total}} = \sum_{i=1;j=i+1}^{n-1} \left(\sum_{\substack{i=1 \\ j=i+m}}^{n-m} E_{ij} \right) - (n-2) \sum_{i=1}^{n} E_i \tag{51}$$

or written more intuitively

$$E_{\text{total}} = \sum_{1 \le i \le n} E_i + \sum_{1 \le i < j \le n} \Delta E_{ij} \tag{52}$$

ΔE_{ij} is defined as the interaction energy of two kernels ($\Delta E_{ij} = E_{ij} - (E_i + E_j)$. To improve the accuracy of the method beyond the inclusion of disconnected double kernel calculations, both triple and quadruple kernel calculations were implemented. The total energy including triple kernel contributions is

$$E_{\text{total}} = \sum_{1 \le i \le n} E_i + \sum_{1 \le i < j \le n} \Delta E_{ij} + \sum_{1 \le i < j < k \le n} \Delta E_{ijk} \tag{53}$$

with the expression for ΔE_{ijk} following that of ΔE_{ij} ($\Delta E_{ijk} = E_{ijk} - (E_i - E_j - E_k) - (\Delta E_{ij} - \Delta E_{ik} - \Delta E_{jk})$). Similarly, inclusion of quadruple kernel energies gives the expression

$$E_{\text{total}} = \sum_{1 \le i \le n} E_i + \sum_{1 \le i < j \le n} \Delta E_{ij} + \sum_{1 \le i < j < k \le n} \Delta E_{ijk} + \sum_{1 \le i < j < k < l \le n} \Delta E_{ijkl} \tag{54}$$

and ΔE_{ijkl} follows from the expressions for ΔE_{ij} and ΔE_{ijk}. Although the use of quadruple kernels is not typically required, addition of these terms can be advantageous for large systems or if a computer with many nodes and cores is available. In most cases, inclusion of double and triple kernel interactions is sufficient to achieve a high level of accuracy compared to fully ab initio calculations. Equations 52 and 53 of the KEM can be compared with those in the FMO method, eqs 37 and 38. The difference is the inclusion of the ESP in FMO.

Applications of the KEM have included proteins,[332] as well as other biologically relevant systems such as DNA.[333] More recently the KEM has been applied successfully to systems containing extended aromatic character[339] and nitrate ester.[340] Application to systems with such diffuse electrons is typically a failing point for most fragmentation methods, however the KEM overcomes this deficiency through the use of a new bond fractioning scheme. Instead of fractioning single bonds perpendicular to the direction of bonding, the KEM method employs a "fissioning" process where the aromatic bonds are divided in half parallel to the direction of bonding. This creates two aromatic bonds, one in each of the two kernels created. Breaking a conjugated system such as graphene into kernels using this process has produced results accurate to within 1 kcal/mol of full ab initio HF and MP2 calculations. A recent development of the KEM includes a generalized fragmentation scheme[341] based on the approach of Deev and Collins[122] that aims to increase the computational efficiency through elimination of extraneous double, triple, and quadruple kernel calculations.

2.5. Molecular Fractionation with Conjugate Caps and Related Fragmentation Methods

2.5.1. MFCC. As with many fragmentation methods, the molecular fractionation with conjugate caps (MFCC) approach[119,342−349] attempts to reduce computational costs and provide a means to calculate interaction energies, but specifically for protein−ligand systems. The original formulation of the MFCC approach[342] fractioned only peptide bonds to enable the calculation of protein−ligand binding energies. The fractioned bonds are then capped with so-called "concaps" that resemble the local environment of the fragment. By adding together the individual contributions of the fragments and subtracting the contributions from the merged concaps, the total interaction energy of the protein−ligand system can be calculated. An important difference between the MFCC method and other fragmentation methods that employ capping of fractioned bonds is the nature of the caps used. Instead of simple hydrogen caps, the caps in the MFCC approach are formed using portions of the neighboring sections of the molecule. This provides both an efficient method for choosing caps as well as including a representation of the local environment during individual fragment calculations. Further developments have provided the ability to fraction disulfide bonds and allowed for the inclusion of nonbonded interactions in globular proteins.

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

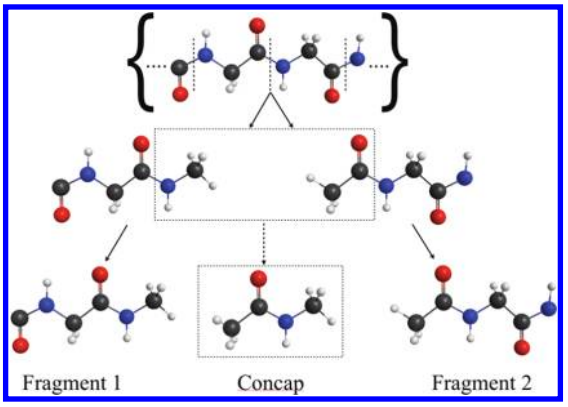


Figure 4. Example of a typical MFCC fractionation scheme. The parent system is divided at the peptide bond into two fragments that are then capped. The caps are then combined into a single concap to be subtracted from the sum of fragment contributions.

As mentioned previously, the focus of the original formulation of the MFCC approach was to break a protein into its constituent amino acids and calculate the interaction energies between individual protein fragments and the ligand of interest. The simplest example is a protein P composed of N amino acids

$$P = nA_1A_2A_3A_4...A_N \tag{55}$$

where A_i represent the individual amino acids and n distinguishes the N-terminus of the protein

$$n = NH_3^+(NH_2) \tag{56}$$

with the opposite end, A_N, representing the C-terminus

$$A_N = R_NCHCOO^-(R_NCHCOOH) \tag{57}$$

To calculate the interaction between protein P and an arbitrary molecule M, the protein is divided into single amino acid fragments across the homolytically broken N−C peptide bonds (Figure 4). This creates N fragments with either one or two unpaired electrons located where the N−C bond(s) used to be. To avoid this unnatural electronic state, each fragment is assigned either one or two concaps, C_{ap}^1 and $C_{ap}^{1^*}$. The main purpose of the caps[350] is to complete the valency requirements of the "dangling" bonds left over after fractionation.

Consider the fractioning of the bond in a simple two amino acid (dipeptide) system. The bond fractioning and subsequent capping gives

$$A_1A_2 = A_1C_{ap}^1 + C_{ap}^{1*}A_2 - C_{ap}^{1*}C_{ap}^1 \tag{58}$$

The energy contained in A_1A_2 can be represented by the sum of the individual amino acid fragments minus the artificial molecule created by joining the two concaps. One can similarly break a larger tripeptide system of three amino acids to give three fragments and the corresponding concaps

$$A_1A_2A_3 = A_1C_{ap}^1 + C_{ap}^{1*}A_2C_{ap}^2 + C_{ap}^{2*}A_3 - C_{ap}^{1*}C_{ap}^1 - C_{ap}^{2*}C_{ap}^2 \tag{59}$$

These two simple examples illustrate the two cases of singly and doubly capped fragments, with the general systems of interest consisting of both the protein and the general ligand M. Note that the original MFCC formulation was only concerned with the interaction energy between a rigid protein and ligand, foregoing the subsequent intramolecular energy calculation for the full system. The total interaction energy can be represented by

$$E(M-P) = \sum_i^N E(M - C_{ap}^{i-1^*}A_iC_{ap}^i) - \sum_i^{N-1} E(M - C_{ap}^{i*}C_{ap}^i) \tag{60}$$

The term $E(M - C_{ap}^{i-1^*}A_iC_{ap}^i)$ represents the interaction energy between molecule M and the capped protein fragment $C_{ap}^{i-1^*}A_iC_{ap}^i$. The second sum of terms represents the interaction of molecule M with the artificial molecule formed by connecting concaps, $C_{ap}^{i^*}C_{ap}^i$. The regularity of amino acid bonding provides a simple choice for the concaps. Since only peptide bonds are being broken, the N-terminus side of each amino acid can be capped with an NH_2 group and the C-terminus side can be capped with $R_{i+1}C_\alpha H_2$ (see Figure 4). This choice of caps ensures that the valence requirements of the broken bonds are complete and the approximate chemical environment around the fragment is being properly represented. Another benefit of choosing caps in this way is the nature of the artificial molecules formed when joining caps, creating reasonable molecular species such as $H_2NR_{i+1}C_\alpha H_2$. The most recent implementation of the MFCC method employs two different sizes of concaps,[351] a "small" and a "large" version. The small concap only extends across the nearest neighbor amino acid, while the large concap extends across the two nearest neighbor amino acids. Use of the large concap provides the obvious benefit of increased accuracy, with the trade-off being an increase in computational cost.

In the special case where disulfide bonds are fractioned,[352] for example a bond between two cystines, an additional term must be added to eq 60:

$$E(M-P) = \sum_i^N E(M - C_{ap}^{i-1*}A_iC_{ap}^i) - \sum_i^{N-1} E(M - C_{ap}^{i*}C_{ap}^i) - \sum_i^{N-1} E(M - DC_{ap}^{i*}DC_{ap}^i) \tag{61}$$

DC_{ap}^* and DC_{ap}^i can be either MeS or HS caps, with the MeS caps shown to give slightly more accurate results.

Further development of the MFCC approach[353,354] addressed the limitation of calculating only the interaction energy by allowing for the calculation of the total electron density, electrostatic potential and dipole moment of proteins. Using the basic dipeptide example in eq 58, the total electron density of the system, ρ, can be expressed as

$$\rho = \rho_{A_1} + \rho_{A_2} - \rho^{cc} \tag{62}$$

where ρ_{A_1} and ρ_{A_2} represent the densities of the capped amino acid fragments after fractionation and ρ^{cc} represents the density of the merged concap. By calculating the densities of the separate fragments, the total density of the system can be determined. One advantage of this formulation lies in the choice of concaps after fractionation. If the cap chosen for fragment A_1 includes the entirety of fragment A_2, then eq 62 is exact, giving the correct limiting behavior if the caps are chosen to be sufficiently large. For a system of N amino acids, eq 62 can be generalized as

$$\rho = \sum_{i=1}^N \rho_i - \sum_{i=1}^{N-1} \rho_i^{cc} - \sum_{i=1}^{N_d} \rho_i^{dc} \tag{63}$$

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

where the third summation over ρ_i^{dc} is only included when disulfide bonds are fractioned. Using this representation of the total density of the system, the electrostatic potential can easily be obtained as

$$\varphi(\mathbf{r}) = -\int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}' \tag{64}$$

with the total electrostatic potential derived through the combination of individual electrostatic potentials of the fragments:

$$\varphi = \sum_{i=1}^{N}\varphi_i - \sum_{i=1}^{N-1}\varphi_i^{cc} - \sum_{i=1}^{N_d}\varphi_i^{dc} \tag{65}$$

In a similar fashion, the dipole moment, μ, of the protein can be calculated as

$$\mu = \sum_{i=1}^{N}\mu_i - \sum_{i=1}^{N-1}\mu_i^{cc} - \sum_{i=1}^{N_d}\mu_i^{dc} \tag{66}$$

Obtaining the total electron density of the system provides a route to calculating the total energy of the system,[351] either through the standard SCF method or by using Kohn−Sham orbitals and the DFT formalism. The implementation of the MFCC approach that allows for the calculation of the total energy using the density matrix (DM) has been termed the MFCC-DM approach.

Two versions of the MFCC-DM approach,[354] the "simple" approach (MFCC-SDM) and the "ghost" approach (MFCC-GDM), are available. These two variants are the result of using extra hydrogens as part of the conjugate caps. If the MFCC approach were exact, then the density matrix elements associated with the hydrogen atomic orbitals (AOs) would be effectively zero. Since the MFCC approach is not exact in practice, these matrix elements add small contributions to the density matrix. The two implementations, MFCC-SDM and MFCC-GDM, differ in how these extraneous hydrogen AOs are handled. In the MFCC-SDM approach the matrix elements are simply neglected, whereas in the MFCC-GDM approach the extra hydrogens are treated as ghost atoms and the matrix elements are accounted for explicitly.

Other than how the extraneous hydrogen atoms are handled, the most important difference between the two MFCC-DM implementations is the structure of the density matrix. In the MFCC-SDM approach the number of electrons is not exactly conserved since the contributions from the extra hydrogen atoms are simply ignored. In the MFCC-GDM approach this deficiency is addressed, since the extraneous hydrogen atoms are treated explicitly as ghost atoms in the system. The trade-off for the exact treatment of the density matrix in the MFCC-GDM approach is an increase in computational cost compared to the MFCC-SDM approach. In practice it has been shown that both approaches produce sufficiently accurate results when compared to full ab initio HF calculations.[354]

To improve the description of globular macromolecules with two- and three-dimensional structures, the addition of nonbonded or "through-space" interactions was implemented.[355] Currently this includes a description of two-body interactions only due to the negligible contribution of higher order interactions in proteins. In general, if two fragments are separated by more than one fractioned bond, they may still be adjacent to one another in the structure of the macromolecule. Following the EDA, the interaction of these two fragments can be calculated by

$$\Delta E^{(2)} = \sum_i \sum_j [E(A_iA_j) - E(A_i) - E(A_j)] \tag{67}$$

Instead of capping these fragments in the same fashion as described earlier, these separated fragments are capped only with hydrogen atoms. This avoids overcounting the two-body correction as well as simplifying the expression needed to calculate the "through-space" two-body correction. This correction can be added to the previous representation for the full energy of the system, giving

$$E = \sum_{i=1}^{N} E(A_i) - \sum_{i=1}^{N-1} E(A_i^{cc}) - \sum_{i=1}^{N_d} E(A_i^{dc}) + \Delta E^{(2)} \tag{68}$$

This new expression for the total energy of the system is termed the energy-corrected MFCC (EC-MFCC) approach.

In addition to the MFCC formalism described above, a number of other capabilities have been added recently. These capabilities include gradients for geometry optimizations[356,355] and the hybrid generalized molecular fractionation with conjugate caps/molecular mechanics (GMFCC/MM) approach[357] that uses MM for long-range interactions. The addition of the conductorlike polarizable continuum model (MFCC-CPCM)[358] and the pairwise interaction correction to the density matrix formulation of the MFCC approach (MFCC-DM-PIC),[359] allow for the calculation of the electrostatic solvation energy of macromolecules and the treatment of short-range polarization interactions, such as hydrogen bonding, respectively. Additionally, the electrostatic field-adapted molecular fractionation with conjugate caps (EFA-MFCC)[360] approach improves the ability to treat charged systems by adding a description of the surrounding environment using point charges. All of these improvements to the original formulation provide a means to perform calculations on a variety of large macromolecular systems, while including a number of important intermolecular interactions.

2.5.2. Generalized Energy-Based Fragmentation Approach. The generalized energy-based fragmentation (GEBF) approach has been proposed[120] as a reformulation of the EFA-MFCC approach. This method builds upon the fragmentation scheme of the MFCC approach, but improves upon the description of the environmental electrostatic field added in the EFA-MFCC approach. The atoms far separated from the fragment of interest are represented as point charges and included in the ab initio calculation on the fragment. The main improvement over the EFA-MFCC approach is the inclusion of electrostatic interactions with polar groups, improving the ability of the method to calculate dipole moments and static polarizabilities.

A new fragmentation scheme is also introduced for the GEBF approach, distinguishing the method from previous MFCC implementations. The general fragmentation scheme consists of the same ideas, dividing a large system into smaller fragments and capping each fragment with the neighboring fragment to conserve the valence requirements of the broken bonds. Any dangling bonds left, such as those on the capping fragment, are capped with hydrogen atoms. However, the inclusion of neighboring fragments is not restricted to those exclusively covalently bonded to the fragment; nonbonded fragments are included as well. By choosing the number of fragments to be included in each quantum calculation using a distance based cutoff, the size of the

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

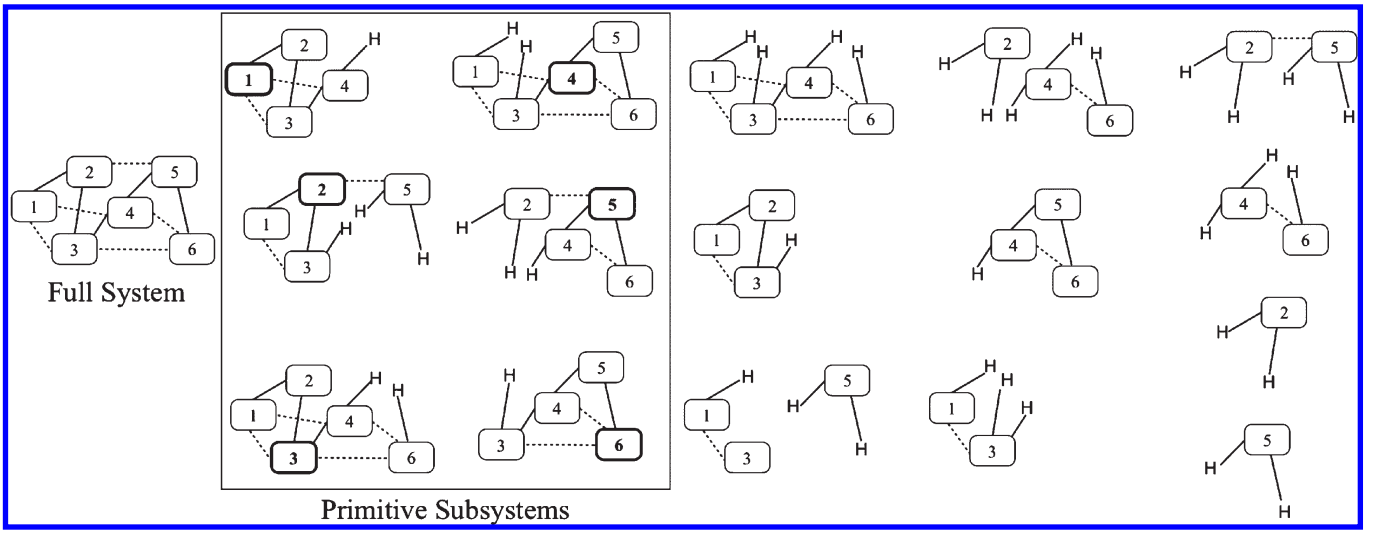


Figure 5. Depiction of a fragmentation scheme created by the GEBF method containing primitive subsystems derived from each central fragment. Solid lines depict covalent bonds while dashed lines represent hydrogen bonds. All possible subsystems are shown, including those that would be removed (i.e., 1234) because of double counting.

fragments and therefore the accuracy of the method can be easily varied depending on computational resources.

Consider a linear system, M, of six fragments

$$M = m_1 m_2 m_3 m_4 m_5 m_6 \tag{69}$$

with connectivity represented in Figure 5. Starting with m_1, derivative subsystems are formed including the central fragment (m_1) and all fragments covalently and, in this case, hydrogen bonded to m_1. This process is then repeated for all m_i (i = total number of fragments) fragments in the full system M, producing i subsystems in the general case. Subsystems that may be present in a larger subsystem are eliminated to avoid double counting. For example, the subsystem $m_1m_2m_3m_4$ appears in the subsystem $m_1m_2m_3m_4m_6$ and should therefore be eliminated from consideration. In this case, two subsystems are eliminated, leaving four unique subsystems.

The remaining subsystems are then checked for double counting of fragment interactions. For example, if the four-body interaction $m_1m_3m_4m_6$ occurs more than once in the four derivative subsystems, a complementary subsystem must be built and subtracted from the sum of the subsystems. This process is followed through for all n-body interactions down to single fragment terms, with each subsystem being assigned a coefficient C_i of either 1 or −1. The total energy of the system M can then be represented as

$$E = \sum_i^N C_i E_i \tag{70}$$

where N is the total number of subsystems. This fragmentation scheme provides nearly all of the three- and four-body interactions, with a small number being neglected. However, these neglected terms were shown to be a small source of error.[120] The final requirement of the fragmentation process in the GEBF approach is for the net number of hydrogen atoms used for capping to be zero.

After the fragmentation and construction of derivative subsystems is complete, each subsystem calculation is performed in the field of point charges on all other atoms. The partial charges Q_A used in the GEBF approach are derived from the natural population analysis (NPA)[361,362] and computed for the central fragment only during an initial HF or DFT calculation on each subsystem. These point charges are then incorporated into a second HF or DFT calculation for all subsystems. Including the point charges in the energy expression, the total energy of the system is

$$E = \sum_i^N C_i \tilde{E}_i - \left(\sum_i^N C_i - 1\right) \sum_A \sum_{B>A} \frac{Q_A Q_B}{R_{AB}} \tag{71}$$

$\tilde{E}_i$ is the total energy of the i-th subsystem that includes the effects of the electrostatic field. The derivatives of this equation have also been derived,[363] allowing for geometry optimizations and vibrational frequency calculations. The most recent improvements to the method include an algorithm for automatic fragmentation and derivative subsystem construction for general molecules.[364]

2.5.3. Other MFCC-Related Methods. A number of more recent approaches have been proposed based on the general fractionation scheme suggested in the MFCC method. An extension of the frozen-density embedding (FDE)[365] scheme to the MFCC approach uses overlapping electron densities of different subsystems to provide a more accurate representation of the surrounding environment, as well as a better representation of the total electron density. Another approach[366] rigorously derives the additivity rule used in the MFCC approach within the semilocal DFT formalism. Finally, the multilevel fragment-based approach (MFBA)[367] employs a fragmentation scheme that is similar to the MFCC formulation (using hydrogen for caps instead of nearest neighbor fragments), while using different levels of ab initio theory for nonbonded fragment calculations depending on the fragment separation.

The polarized protein-specific charge (PPC) method proposed by Ji et al.[368] in 2008 is based on fitting atomic charges to MFCC-derived electrostatic potentials, analogously to the FMO-derived charges[294,295] developed in 2007. The PPC method was applied[369−376] to a number of studies showing the importance of the polarization missing in commonly used force fields.

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

2.6. Systematic Fragmentation Method

A more recent approach[122,377] to fragmenting large chemical systems is the systematic fragmentation method (SFM). Originally designed to treat large proteins, polymers and surfaces, the SFM takes a unique approach to fragmenting systems. In the same vein as the other fragmentation methods, SFM employs highly accurate methods, such as MP2 or CC, to perform calculations on the smaller fragments, obtaining the total energy of the system. However, instead of having uniquely defined fragments, the SFM uses overlapping sections of the molecule to account for interactions between fragments. By performing the fragmentation in this way, each atom "feels" the presence of the full system without the use of an externally applied field. Nonbonded interactions are also accounted for between far separated regions of the system, originally using simple electrostatics, but more recently by using the EFP method to obtain more accurate results.[199] The SFM has been used to describe the isomerization of DNA helices[199] and retinal,[378] as well as to give a proper description of potential energy surfaces of chemical reactions.[379]

The underlying premise behind the SFM is in thinking of a chemical system as a collection of single-bonded functional groups. The definition of a functional group is central to the fragmentation, allowing for different levels of fragmentation (larger individual fragments) to increase the accuracy of the individual calculations. This also allows for a better description of many-body effects, as any atom in the fragment would be influenced by more of the system during the full ab initio calculations. The general fragmentation scheme is best described using a linear molecule, M, of arbitrary length K

$$M = G_1G_2G_3...G_K \tag{72}$$

The "super" molecule can be broken into two distinct molecules by stretching the bond between G_{n-1} and G_n to infinity. The bond is broken homolytically, assigning one electron from the broken bond to G_{n-1} and the other to G_n. To avoid charged fragments created by such a scheme, hydrogen atoms are used as "caps" on each of the fragments. The overall fragmentation creates two molecules

$$M \rightarrow M_1 + M_2 \tag{73}$$

which are composed of

$$M_1 = G_1G_2G_3...G_{n-1}H^{(n-1)} \tag{74}$$

$$M_2 = H^{(n)}G_nG_{n+1}...G_K \tag{75}$$

where $H^{(n-1)}$ and $H^{(n)}$ are the hydrogen caps for their respective fragments, located along the direction of the broken bond at a chemically sensible distance for the specific GH bond.[122]

The energies of these fragments can then be calculated and are related by

$$E(M) = E(M_1) + E(M_2) + \mathrm{dE}_1 \tag{76}$$

where dE_1 represents the energy change created by the bond breakage. To allow for overlapping fragments, it is acknowledged that the fragmentation choice just described is not the only possible fragmentation scheme. The complete molecule M could also be broken at some other single bond, creating different fragments, M_3 and M_4, giving the following energy expression

$$E(M) = E(M_3) + E(M_4) + \mathrm{dE}_2 \tag{77}$$

where

$$M_3 = G_1G_2G_3...G_{i-1}H^{(i-1)} \tag{78}$$

$$M_4 = H^{(i)}G_iG_{i+1}...G_K \tag{79}$$

By performing both fragmentations at the same time, one can represent the total molecule M as

$$M \rightarrow G_1G_2G_3...G_{n-1}H^{(n-1)} + H^{(n)}G_nG_{n+1}...G_{i-1}H^{(i-1)} + H^{(i)}G_iG_{i+1}...G_K \tag{80}$$

This double fragmentation creates a new energy expression

$$E(M) = E(M_1) + E(M_5) + E(M_4) + \mathrm{dE}_3 \tag{81}$$

The new fragment M_5 is the result of fragment M_2 being broken into two pieces by the second fragmentation. Specifically,

$$M_5 = H^{(n)}G_nG_{n+1}...G_{i-1}H^{(i-1)} \tag{82}$$

The new term dE_3 is the result of an approximation; namely, if the $G_{n-1}G_n$ bond is separated from the $G_{i-1}G_i$ bond by a great enough distance, the energy difference from simultaneous fragmentation will equal the sum of the energy changes of each fragmentation performed separately

$$\mathrm{dE}_3 \approx \mathrm{dE}_1 + \mathrm{dE}_2 \tag{83}$$

As the distance between the two fragmentation sites increases, the approximation in eq 83 becomes more and more reliable. Rearranging eqs 77 and 81, while using the equality in eq 83, it can be shown that

$$\begin{aligned}\mathrm{dE}_3 &= E(M) - E(M_1) - E(M_5) - E(M_4) \\ &\approx E(M) - E(M_1) - E(M_2) + E(M) - E(M_3) - E(M_4)\end{aligned} \tag{84}$$

or more simply

$$E(M) \approx E(M_2) + E(M_3) - E(M_5) \tag{85}$$

Using the molecular definitions, it becomes apparent that M_5 is simply the overlapping or "double counted" region common to both M_2 and M_3.

Within the foregoing formulation, there are different "levels" of fragmentation. Fragmentation level 1 consists of fragmentation sites separated by one functional group, level 2 has fragmentation sites separated by two functional groups and so on. The current implementation of the SFM allows for up to fragmentation level 3, with the obvious possibility for extension to higher levels. As the fragmentation sites become farther separated, the approximation in eq 83 becomes more reliable and the total energy of the system approaches that of the full ab initio calculation.

To illustrate the exhaustive fragmentation of a system using the different fragmentation levels, consider the acyclic molecule M again

$$M = G_1G_2G_3G_4G_5G_6G_7G_8 \tag{86}$$

In this case, the fixed length $K = 8$. Under the level 1 fragmentation scheme the two broken bonds are separated by only one functional group. The first fragmentation site is chosen between

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

G_1 and G_2 and the second site between G_2 and G_3. The first fragmentation creates

$$M = G_1 + G_2G_3G_4G_5G_6G_7G_8 \quad (87)$$

while the second creates

$$M = G_1G_2 + G_3G_4G_5G_6G_7G_8 \quad (88)$$

Subtracting group G_2 to avoid double counting, we get the following representation of molecule M:

$$M = G_1G_2 + G_2G_3G_4G_5G_6G_7G_8 - G_2 \quad (89)$$

Following this scheme through to all possible fragmentation sites allowed by level 1, until no fragment larger than two functional groups remains, the bonded energy of molecule M can be represented by the sum of fragment energies as follows:

$$\begin{aligned} E_{\text{level1}}^{\text{bonded}}(M) = {} & E(G_1G_2) + E(G_2G_3) \\ & + E(G_3G_4) + E(G_4G_5) + E(G_5G_6) \\ & + E(G_6G_7) + E(G_7G_8) - E(G_2) - E(G_3) \\ & - E(G_4) - E(G_5) - E(G_6) - E(G_7) \end{aligned} \quad (90)$$

Using the same methodology for level 2, with fragmentation sites instead separated by two functional groups, the bonded energy of M is represented by

$$\begin{aligned} E_{\text{level1}}^{\text{bonded}}(M) = {} & E(G_1G_2) + E(G_2G_3) \\ & + E(G_3G_4) + E(G_4G_5) + E(G_5G_6) \\ & + E(G_6G_7) + E(G_7G_8) - E(G_2) - E(G_3) \\ & - E(G_4) - E(G_5) - E(G_6) - E(G_7) \end{aligned} \quad (91)$$

and using level 3, with fragmentation sites separated by three functional groups, gives a representation of the bonded energy of M

$$\begin{aligned} E_{\text{level2}}^{\text{bonded}}(M) = {} & E(G_1G_2G_3) + E(G_2G_3G_4) \\ & + E(G_3G_4G_5) + E(G_4G_5G_6) + E(G_5G_6G_7) \\ & + E(G_6G_7G_8) - E(G_2G_3) - E(G_3G_4) \\ & - E(G_4G_5) - E(G_5G_6) - E(G_6G_7) \end{aligned} \quad (92)$$

Carrying this fragmentation scheme out to level n, where n is the number of groups in the system, one would be left with the unfragmented system. By using higher levels of SFM, and consequently larger fragments, the total energy will approach that of the exact system.

So far, only the bonded energy of the example system has been discussed, leaving out a great deal of important contributions to the total energy contained in the nonbonded interactions of separated functional groups. Recently, the calculation of these interactions has moved from a simple electrostatic model in the original formulation, to a more sophisticated approach combining ab initio and EFP method calculations.[378] The choice of performing either full ab initio calculations or EFP calculations is based on the shortest atom−atom distance between the interacting fragments. At short ranges (<2.7 Å) the nonbonded interactions are calculated with the full ab initio method being employed. For intermediate distances (2.7−4.5 Å) the EFP method is used, effectively reducing the number of ab initio calculations required and increasing the computational efficiency.

The simplest case of nonbonded interactions occurs between just two separated functional groups, for example G_1 and G_4. The nonbonded energy contained in the "super-group" G_1G_4 can be given by

$$E_{\text{nb}}^{(1,1)}[G_1;G_4] = E(G_1G_4) - E(G_1) - E(G_4) \quad (93)$$

$E(G_1G_4)$ is the supermolecular energy of the two separated functional groups and $E(G_1)$ and $E(G_4)$ are the one-body fragment energies. During the calculation of the supermolecular energy $E(G_1G_4)$, the two functional groups retain their original positions from the complete system M. This procedure can be carried out to its eventual conclusion of all possible pairs of separated functional groups, providing the total two-body nonbonded interaction energy of the entire system M.

In some instances, three-body interactions can play a very important role in chemical systems.[380−384] To account for these interactions a similar approach can be taken by considering the interaction of three functional groups, G_1, G_2, and G_3. The interactions of these three groups is ignored unless two of the three are bonded in the full system M. To give a specific example, consider the case in which G_2 is directly bonded to G_3, giving the three body interaction energy as

$$\begin{aligned} E_{\text{nb}}^{(1,2)}[G_1;G_2G_3] = {} & E(G_1G_2G_3) - E(G_1) - E(G_2G_3) \\ & - E_{\text{nb}}^{(1,1)}[G_1;G_2] - E_{\text{nb}}^{(1,1)}[G_1;G_3] \end{aligned} \quad (94)$$

The three-body energy is simply the supermolecular energy $E(G_1G_2G_3)$ minus the one-body energy $E(G_1)$, the bonded energy of $E(G_2G_3)$ and the two-body nonbonded energies $E_{nb}^{(1,1)}[G_1;G_2]$ and $E_{nb}^{(1,1)}[G_1;G_3]$.

Using all three of these expressions for the bonded and nonbonded energies, the total energy of the entire system, E_{SFM}, can be expressed as a sum of bonded and non-bonded energies

$$E_{\text{SFM}} = E^{\text{bonded}} + E^{\text{non-bonded}} \quad (95)$$

The term $E^{\text{nonbonded}}$ includes all of the terms up to n-th order. In practice it has been shown that the inclusion of nonbonded terms in the level 3 fragmentation scheme provides the best combination of accuracy and computational efficiency. The choice of including three-body nonbonded interactions depends on the system of interest, with molecular clusters such as water necessitating the inclusion of three-body interactions.

The SFM has some limitations, the most important of which is the inability to break bonds more complex than single bonds. There is also a limitation in the fragmentation of cyclic molecules such as cyclohexane. During fragmentation the capping hydrogens may become too close, causing a steric interaction not present in the complete system and violating the approximation that fragmentation sites are energetically independent. This problem can be alleviated through the use of the so-called ring repair rule[122] to effectively avoid such nonphysical interactions. Active development of the SFM currently includes a reformulation of the polarization interactions to give a more accurate description of highly polar molecular clusters.

2.7. Divide-and-Conquer Methods

2.7.1. Original Divide and Conquer Approach. The original formulation of the divide and conquer (DC) approach was proposed by Yang[20] in 1991. Based on the Kohn−Sham (KS) formalism, the DC approach aims to avoid the use of the N/2 KS orbitals and instead divide the density of the system of interest into the sum of the densities of the subsystems. The DC formalism was generalized to a more efficient one-electron

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

density matrix approach that is generally applicable to ab initio[385−388] and semiempirical methods.[389,390] As with other fragmentation methods, the local environment of each subsystem is taken into account, in this case through the use of buffer regions surrounding each fragment. Since the method was initially proposed, it has been expanded to electron correlation methods as well.[391] Nakai's group[392−401] and others[402,403] made many contributions to further development. The main contributors to these enhancements of the DC approach will be discussed, beginning with both the original and improved density matrix formulations proposed by Yang.

In the KS formalism, the total energy of a system of N electrons in an external field $v(\mathbf{r})$ can be represented using the electron density $\rho(\mathbf{r})$ as

$$E[\rho] = T_s[\rho] + \int v(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r} + E_{\mathrm{xc}}[\rho] + \frac{1}{2}\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' + \sum_{a,b}\frac{Z_aZ_b}{R_{ab}} \quad (96)$$

$T_s[\rho]$ represents the kinetic energy of the noninteracting electron gas with density ρ in the ground state. The term $E_{\mathrm{xc}}[\rho]$ is the exchange-correlation energy. The final term in eq 96 is the nuclear repulsion. In the conventional KS formalism, the energy functional $E[\rho]$ is minimized with respect to the density by satisfying the KS equation

$$\hat{H}\psi_i(\mathbf{r}) = \left[-\frac{1}{2}\nabla^2 + V_{\mathrm{eff}}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}) \quad (97)$$

$V_{\mathrm{eff}}(\mathbf{r})$ is the KS effective local potential, and $\hat{H}$ is the KS Hamiltonian. As in the KS formalism, the DC approach uses the electron density as the main variable. However, the DC approach represents the total electron density of the full system as a sum of subsystem contributions. This is accomplished through the use of normalized "partition functions"

$$\sum_{\alpha} p^{\alpha}(\mathbf{r}) = 1 \quad (98)$$

$p^{\alpha}(\mathbf{r})$ is a positive weighting function for subsystem α. More specifically, $p^{\alpha}(\mathbf{r})$ is large for the subspace of subsystem α and small otherwise. This leads to an expression for the total electron density of the system

$$\rho(\mathbf{r}) = \sum_{\alpha} p^{\alpha}(\mathbf{r})\rho(\mathbf{r}) = \sum_{\alpha}\rho^{\alpha}(\mathbf{r}) \quad (99)$$

The partition functions have been determined previously,[20] and apparently the density and energy do not depend significantly upon the specific form of the partition function. The density of an individual subsystem can now be defined using the Fermi function[404] $f_\beta(x)$ as

$$\rho^{\alpha}(\mathbf{r}) = 2p^{\alpha}(\mathbf{r})\sum_m f_\beta(\varepsilon_{\mathrm{F}} - \varepsilon_m^{\alpha})|\psi_m^{\alpha}(\mathbf{r})|^2 \quad (100)$$

where

$$f_\beta(x) = [1 + \exp(-\beta x)]^{-1} \quad (101)$$

ψ_m^{α} represents an eigenfunction that is localized on the particular subsystem α, and ε_m^{α} is the corresponding eigenvalue of the subsystem α. The eigenfunctions ψ_m^{α} are linear combinations of a set of local basis functions.

$$\psi_m{}^{\alpha}(\mathbf{r}) = \sum_j C_{jm}^{\alpha}\varphi_j^{\alpha}(\mathbf{r}) \quad (102)$$

It is the use of localized basis functions for a particular subsystem that allows the DC method to scale nearly linearly with system size. The coefficients C_{jm}^{α} are solutions to the generalized eigenvalue equation, derived from the Rayleigh−Ritz variational principle

$$(\mathbf{H}^{\alpha} - \varepsilon_m^{\alpha}\mathbf{S}^{\alpha})\mathbf{C}_m^{\alpha} = 0 \quad (103)$$

The Fermi energy[404] ε_{F} from eq 98 is determined from

$$N = \int \rho(\mathbf{r})\mathrm{d}\mathbf{r} = 2\sum_{\alpha}\sum_m f_\beta(\varepsilon_{\mathrm{F}} - \varepsilon_m^{\alpha})\langle\psi_m^{\alpha}|\rho^{\alpha}(\mathbf{r})|\psi_m^{\alpha}\rangle \quad (104)$$

N is simply the normalization condition for the electron density. After solving the preceding equations self-consistently, the energy expression from eq 96 can be expressed in terms of the eigenvalues

$$\tilde{E}[\rho] = \tilde{\varepsilon} + \int \rho[-\varphi(\mathbf{r})/2 - V_{\mathrm{xc}}(\mathbf{r})]\mathrm{d}\mathbf{r} + E_{\mathrm{xc}}[\rho] + \sum_{a,b}\frac{Z_aZ_b}{R_{ab}} \quad (105)$$

$\tilde{\varepsilon}$ is an approximation to the KS eigenvalues, represented by

$$\tilde{\varepsilon} = 2\sum_{\alpha}\sum_m f_\beta(\varepsilon_{\mathrm{F}} - \varepsilon_m^{\alpha})\langle\psi_m^{\alpha}|\rho^{\alpha}(\mathbf{r})|\psi_m^{\alpha}\rangle \quad (106)$$

In an effort to improve upon the original formulation, Yang proposed a reformulation of the DC formalism based on the one-electron density matrix.[385] Defined in terms of the KS orbitals, the one-electron density matrix can be expressed as

$$\rho(\mathbf{r},\mathbf{r}') = 2\sum_m^{N/2}\psi_m(\mathbf{r})\psi_m(\mathbf{r}') = \sum_{ij}\rho_{ij}\varphi_j(\mathbf{r})\varphi_j(\mathbf{r}') \quad (107)$$

The density matrix ρ_{ij} is given by the linear coefficients in the expansion of the KS orbitals

$$\rho_{ij} = 2\sum_m^{N/2} C_{im}C_{jm} \quad (108)$$

The partition matrix for each subsystem can now be defined in the atomic orbital space, with a normalization condition that is similar to the original formulation in eq 98

$$\sum_{\alpha}\mathbf{p}_{ij}^{\alpha} = 1 \quad (109)$$

and constructed using the following rules:

$$\mathbf{p}_{ij}^{\alpha} = \begin{cases} 1 & \text{if} \quad i \in \alpha \text{ and } j \in \alpha \\ \frac{1}{2} & \text{if} \quad i \in \alpha \text{ and } j \notin \alpha \\ 0 & \text{if} \quad i \notin \alpha \text{ and } j \notin \alpha \end{cases} \quad (110)$$

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

Subsystem contributions to the density matrix can now be written as

$$\rho_{ij} = \sum_{\alpha} \mathbf{p}_{ij}^{\alpha} \rho_{ij} = \sum_{\alpha} \rho_{ij}^{\alpha} \quad (111)$$

The expression in eq 111 is equivalent to the expression from the original formulation in eq 99. A corresponding approximation to the original formulation, using a set of local eigenvectors to approximate the density matrix of a subsystem, can be applied to give

$$\rho_{ij}{}^{\alpha} = 2\mathbf{p}_{ij}^{\alpha} \sum_{m} f_{\beta}(\varepsilon_{\mathrm{F}} - \varepsilon_{m}^{\alpha}) C_{im}^{\alpha} C_{jm}^{\alpha} \quad (112)$$

An analogous expression for the Fermi energy, determined by the normalization, can then be expressed as

$$N = \sum_{ij} \rho_{ij} \mathbf{S}_{ij} = \sum_{ij} (2 \sum_{\alpha} \mathbf{p}_{ij}^{\alpha} \sum_{m} f_{\beta}(\varepsilon_{\mathrm{F}} - \varepsilon_{m}^{\alpha}) C_{im}^{\alpha} C_{jm}^{\alpha}) \mathbf{S}_{ij} \quad (113)$$

Finally, the analogous expression for the sum of the eigenvalues is written as

$$\begin{aligned} \varepsilon &= 2 \sum\sum f_{\beta}(\varepsilon_{\mathrm{F}} - \varepsilon_{m}^{\alpha}) \varepsilon_{m}^{\alpha} \sum \mathbf{p}_{ij}^{\alpha} C_{im}^{\alpha} C_{jm}^{\alpha} \mathbf{S}_{ij} \\ &= \sum_{ij} (2 \sum_{\alpha} \mathbf{P}_{ij}{}^{\alpha} \sum_{m} f_{\beta}(\varepsilon_{\mathrm{F}} - \varepsilon_{m}^{\alpha}) C_{im}^{\alpha} C_{jm}^{\alpha}) \mathbf{H}_{ij} \end{aligned} \quad (114)$$

The main advantages of the new DC formulation include the removal of the time-consuming partition function integrals, as well as the general applicability the approach now has to ab initio methods such as Hartree−Fock. Analogous expressions for the gradients have been derived from the original formulation as well.

The implementation of the DC method described divides a molecular system into various subsystems, each described by a set of local orbitals. To aid in the description of each subsystem, the basis functions of neighboring subsystems, or "buffer" regions, are included in each subsystem calculation. The inclusion of buffer regions is determined by a distance based cutoff, R_b. Any atom that falls inside the sphere created by R_b is included as a buffer atom in the subsystem calculation. By including the atomic basis functions from both the subsystem and the buffer region, computational requirements scale by a factor of N_{α}^3, where N_{α} is the number of basis functions contained in subsystem α and the corresponding buffer region. Yang determined that the buffer region required to achieve a certain level of accuracy remains constant regardless of the system size.[386] Linear scaling is then possible by fixing the size of the buffer region for each subsystem.

In addition to linear scaling in terms of the computational time required, Yang also proposed a reduction in memory requirements by storage of a sparse matrix for the total system.[386] This is accomplished using a distance based cutoff for matrix elements between atom pairs with an interatomic distance less than R_h. By applying this cutoff the storage requirements of the density matrix become proportional to the size of the molecule. The amount of CPU time required also decreases because of the reduction in matrix element evaluations. In addition to the density matrix, both the one-electron core Hamiltonian and Fock matrices are treated in this way.

Following the two formulations of the DC method in 1995 by Yang, a number of other groups proposed extensions to the method to enhance the functionality. The first to propose such an extension was Merz in 1996 with his semiempirical MO implementation.[389,390] A notable improvement pointed out by Merz was the need for overlap between adjacent subsystems. To illustrate why this is necessary, consider the closed shell Fock matrix

$$F_{\mu\nu} = H_{\mu\nu} + \sum_{\lambda=1}^{M} \sum_{\sigma=1}^{M} \left[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\sigma|\lambda\nu) \right] P_{\lambda\sigma} \quad (115)$$

where the two electron integrals are represented as

$$(\mu\nu|\lambda\sigma) = \int\int \chi_{\mu}^{*}(\mathbf{r}_1)\chi_{\nu}(\mathbf{r}_1) \times \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \chi_{\lambda}^{*}(\mathbf{r}_2)\chi_{\sigma}(\mathbf{r}_2) \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \quad (116)$$

The density matrix for a particular subsystem, $\mathbf{P}^{\alpha}$, depends on how the subsystem is defined. Buffer regions are included to reduce the truncation effects of fragmentation, however their basis functions are only included during the solution of the SCF equations. This leads to their exclusion during construction of the density matrix, since only basis functions that are part of the subsystem directly, and not contributed by a buffer region, are included. Since the density matrix elements contributed exclusively from buffer region basis functions are zero, overlap between subsystems is required to provide these contributions to the density matrix.

A number of improvements to the DC formulation of Yang were implemented by the Nakai group, including the ability to perform MP2 calculations,[392,395] application to delocalized systems,[393] the addition of Hartree−Fock exchange,[394] CC calculations[396] and a two level hierarchical scheme.[399] The problem of overlapping subsystems pointed out by Merz was also overcome in the Nakai implementation by treating each atom as a subsystem. Detailed test calculations were performed and showed that the use of a sufficiently large buffer region is adequate to overcome such a small subsystem division. Even in the case of delocalized systems, subsystems consisting of a single atom were capable of producing reasonably accurate results when a large buffer region (greater than 10 Å) was considered.[393] It was noted, however, that the accuracy of any DC calculation can obviously be improved through the use of larger subsystems if desired.

Correlated calculations were improved through the use of a "dual buffer" scheme.[395] The first buffer region surrounding each subsystem includes the calculation of the correlation energy, while the second layer buffer only performs a Hartree−Fock calculation. This scheme effectively exploits the local nature of electron correlation and provides additional reductions in the computational effort for correlated calculations.

A dual-level hierarchical scheme[399] is built upon the foundation of a dual buffer scheme. Following the same procedure for calculating the correlation energy as in the dual buffer scheme, a second level of approximation is introduced by using a small and large basis set for HF and correlated calculations respectively. The use of a larger basis set for the correlated buffer calculations provides more accurate energies, while the use of a smaller basis set for the HF buffer regions reduces the computational cost for each subsystem calculation.

2.7.2. Adjustable Density Matrix Assembler Approach (ADMA). Developed contemporarily with the divide and conquer method, the adjustable density matrix assembler approach divides the density of a molecular system using a density matrix approach. Originating from a similar approach called the molecular electron density lego approach[405,406] (MEDLA), and inspired by earlier work by Michl[407] and Stoddart,[408] the ADMA[409−414] method uses the idea of a fuzzy-set to remove any discontinuities

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

between fragments, circumventing the need for explicit capping or other such bond conserving approaches. Using a density matrix formalism allows the ADMA approach to be applied to many modern ab initio methods. The most recent formulation of the ADMA approach will be discussed, as well as recent improvements such as the inclusion of point charges during individual fragment calculations.

The ADMA approach expresses the electron density of a molecular system in terms of a sum of contributions from individual fragments. Consider the exact total density of a molecule at a given nuclear configuration

$$\rho(\mathbf{r}) = \sum_{i=1}^{n}\sum_{j=1}^{n} P_{ij}\phi_i(\mathbf{r})\phi_j(\mathbf{r}) \tag{117}$$

where $\phi_i(\mathbf{r})$ and $\phi_j(\mathbf{r})$ are atomic orbitals (AOs). To rewrite this expression in terms of a summation of smaller contributions, the molecule of interest must be divided into a set of mutually exclusive "families" f^k where $k = 1, ..., m$ and m is the total number of families. Division of AOs among families is accomplished through the use of a membership function $m^k(i)$ such that

$$m^k(i) = \begin{cases} 1 & \text{if AO}\phi(\mathbf{r}) \text{ is centered on one nuclei of set } f^k \\ 0 & \text{otherwise} \end{cases} \tag{118}$$

Each fragment k then has a density matrix defined by

$$P_{ij}^k = \begin{cases} P_{ij} & \text{if both } \phi_i \text{ and } \phi_j \text{ are centered on a nucleus of } f_k \\ \frac{1}{2} & \text{if one of } \phi_i \text{ and } \phi_j \text{ are centered on a nucleus of } f_k \\ 0 & \text{otherwise} \end{cases} \tag{119}$$

This scheme allows for an exactly additive decomposition of the total density matrix. To illustrate this, consider a molecular system M containing nine families

$$M = f_1 f_2 f_3 \ldots f_9 \tag{120}$$

The interactions of these nine families can be artificially represented in matrix form

$$\begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{matrix} \tag{121}$$

Any nearest neighbor families are taken to have a strong interaction, for example, family 4 has a strong interaction with family 5 but not family 6. Building the total interactions from family 4, a reduced form of the interaction matrix can be written as

$$\begin{matrix} 1 & 2 \\ \mathbf{4} & 5 \\ 7 & 8 \end{matrix} \tag{122}$$

which includes all families with a strong interaction with the central family 4. Of the other families considered in this reduced system, families 2, 5, and 8 contain dangling bonds that were previously connected to families 3, 6, and 9 respectively. These bonds are capped with hydrogen atoms, however, the size of each family (in this case the family represented by eq 122) is chosen to be large enough that the corresponding orbitals from the capping hydrogen have negligible contributions to the orbitals on the central family 4. This allows the capping hydrogen orbitals to be excluded from construction of the density matrix. The contribution of the reduced system in eq 122 to the total density matrix can then be written as

$$\begin{matrix} 1 & & & * & & & & & \\ & 2 & & * & & & & & \\ & & 3 & 0 & & & & & \\ * & * & 0 & 4 & * & 0 & * & * & 0 \\ & & & * & 5 & & & & \\ & & & 0 & & 6 & & & \\ & & & * & & & 7 & & \\ & & & * & & & & 8 & \\ & & & 0 & & & & & 9 \end{matrix} \tag{123}$$

Each asterisk represents a contribution to the total density matrix created by interactions contained in the reduced subsystem of eq 122. Every interaction is scaled by 0.5, and all other off diagonal contributions are zero. This procedure can then be repeated, assigning each family as the central family and building the corresponding reduced systems. Following this procedure through to completion, a matrix of fragment contributions to the total density matrix is obtained.

$$\begin{matrix} 1 & * & 0 & * & * & 0 & 0 & 0 & 0 \\ * & 2 & * & * & * & * & 0 & 0 & 0 \\ 0 & * & 3 & 0 & * & * & 0 & 0 & 0 \\ * & * & 0 & 4 & * & 0 & * & * & 0 \\ * & * & * & * & 5 & * & * & * & * \\ 0 & * & * & 0 & * & 6 & 0 & * & * \\ 0 & 0 & 0 & * & * & 0 & 7 & * & 0 \\ 0 & 0 & 0 & * & * & * & * & 8 & * \\ 0 & 0 & 0 & 0 & * & * & 0 & * & 9 \end{matrix} \tag{124}$$

Each asterisk now represents a sum of two scaled contributions. This relatively sparse fragment interaction matrix is then used to construct a good approximation to the total density matrix. The example shown is overly simplified; however, it provides a reasonable illustration of how the ADMA method can reduce the computational cost required to compute the total density matrix.

In general, the fragment choice in the ADMA approach is unimportant, however in practice chemically reasonable fragments are chosen. For example, when fragmenting a protein, a set of chemically sensible rules are followed: (i) all heavy atoms double bonded to an oxygen atom are considered to be fragments; (ii) if a heavy atom is part of an aromatic moiety, the entire aromatic region defines a fragment; (iii) specific chemically important substituents such as the $CONH_2$ groups of asparagine and gluatamine are defined as a fragment. After the initial fragment choice, in a manner similar to the MTA method, a sphere of a user defined radius is used to include neighboring families. If a single atom belonging to another nuclear family falls within the radius around the fragment of interest, then all atoms in that family are included. This includes any atoms that are not directly bonded to the fragment of interest. All dangling bonds are considered at this point and capped using hydrogen atoms.

Once the fragment of interest and the included surroundings are defined, the density matrix contribution can be calculated for

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

all defined fragments. The accuracy of the ADMA approach may be improved either by choosing larger fragments or by increasing the radius of inclusion for neighboring nuclear families. The composite fragment density matrix can then be used to calculate a number of properties including the total energy, dipole moment, electrostatic potential, and total density matrix for the molecular system.

The accuracy of the ADMA approach depends on how well the local environment, determined by the radius of inclusion, is described during each fragment calculation. The most obvious way to improve the description of the environment is to increase the radius of inclusion of neighboring atoms. It was shown that a radius of up to 10 Å is needed to obtain sufficiently accurate results (errors of less than 1 kcal/mol).[415] Unfortunately, this decreases the efficiency of the approach, particularly for large molecules. The solution was implemented by Mezey and Exner in 2006, termed the field-adapted ADMA (FA-ADMA)[416] approach. By using a smaller radius of inclusion coupled with a more approximate description of far separated atoms, the FA-ADMA approach effectively increases the accuracy of the original ADMA approach without any significant increase in computational cost.

The most effective approximation to the surroundings of fragments in the FA-ADMA was determined to be partial charges based on Mulliken or Löwdin charges. The computational scheme employed begins with a standard ADMA calculation which is used to obtain the partial charges of all the atoms present in the fragment calculation. During subsequent fragment calculations, these partial charges are included, gradually increasing the number of partial charges until the full system is represented. The entire process is then started anew, using all of the partial charges obtained during the previous step, and continued in an iterative process until the charges and densities for all fragments do not change up to a certain threshold. Atoms present in the point charge description of the environment that have been replaced by capping hydrogens in the fragment (so-called junction atoms) are represented by a 60% scaling of the total Mulliken charge or an 80% scaling of the total Löwdin charge, aiding in the reduction of error incurred by overcounting these atoms during the ab initio portion of the FA-ADMA calculations.

Additional improvements to the ADMA approach have been implemented in recent years. These include the use of an alternative, more general fragmentation scheme,[417] the use of the ADMA density matrix as an initial guess for ab initio SCF calculations on large molecules[418] and an improved description of the junction atoms at the edges of the surroundings encompassed by the radius of inclusion.[419]

2.8. Other Methods

Das et al.[420] in 2003 proposed the ab initio fragment orbital-based theory (AFOT), in which the total wave function is constructed from orbitals of fragments, which are formed and computed without capping.

The integrated multicenter molecular orbital method (IMiCMO) formulated by Sakai and Morita[153,421] for molecular clusters relies on the addition of the properties of target molecules computed in the buffer region of adjacent molecules, while the effect of the rest of the system is computed with point charge interactions. In addition to the total energies, forces and second derivatives have also been developed.

Mata et al.[422] in 2009 suggested a model to estimate the excitation energies of large systems by diagonalizing the total Hamiltonian H_{ij}, $i,j = 1, ..., N$, where N is the number of fragments. In this Hamiltonian only the diagonal elements H_{ii} are computed quantum-mechanically (for each fragment, electrostatically embedded in the field of others) while off-diagonal elements are estimated semiclassically with the dipole−dipole interaction model. This model takes into account the excitonic coupling in molecular clusters and it has been used to compute the first electronic absorption band of water.

Following the same methodology as the SFM, Bettens et al.,[423−425] have developed an energy based fragmentation method based on the idea of isodesmic reactions.[426] Molecules are divided into groups in a similar fashion to the SFM and capped with hydrogens. The main difference between the two methods is how the groups are recombined into fragments. In many cases the two fragmentation schemes produce the exact same fragments. However, for higher levels of fragmentation such as SFM level 3, the Bettens method can produce a different fragmentation scheme. The fragments produced in these cases are typically smaller in size than those produced by the SFM, but many more fragments are formed. In a specialized case provided by Bettens,[425] fragments formed are one group size smaller than the SFM fragments. In this instance SFM produces three fragments, while the Bettens method produces eleven fragments.

In the linear-scaling three-dimensional fragment method for large-scale electronic structure calculations (LS3DF) developed by Wang et al.[427] in 2008, the energy is constructed as a sum of additive contributions of small conglomerates of fragments in the self-consistent Coulomb potential determined at each iteration by solving the Poisson equation for the total density; the conglomerates are used to reduce the error caused by the caps added to dangling bonds. This method has been applied to CdSe quantum dots that contain up to 2616 atoms.

3. SOFTWARE AND PARALLEL COMPUTING

Developing efficient software is a very important aspect of computational and theoretical chemistry. Many research groups concentrate on method development using some "in-house" programs or locally modified versions of commonly used software. Although this often results in important advances in method development, such results are often not reproducible by other scientists without reproducing the code, an arduous process. Moreover, there is a growing need by researchers in adjacent fields, such as biochemistry material science and engineering, computational biology and others to use ab initio methods for practical calculations. It is clear that there is a demand for easy-to-use software that can be run by users on their own. This need is at present only partially filled by software development of fragment-based methods.

There are two types of software: programs to perform calculations, and software to aid in input file preparation and to analyze and visualize results. For the first group, the most fragment-friendly program at present is the general atomic and molecular electronic structure system (GAMESS),[428,429] which has EFP, FMO, ELG, and DC in the production version. Q-Chem supports EFP, while ABINIT-MP[430,431] and PAICS[432] are FMO programs. PEACH[430] is an MD program used in conjunction with ABINIT-MP to run FMO-MD simulations. MTA is based on locally modified versions of GAMESS, and it has a Web interface.[433,434] Some recent X-Pol developments use a locally modified version of GAMESS.

GAMESS has MacMolPlt[435] as its graphical interface, although it has limited capability specifically for fragment methods.

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

Facio[436] has a very elaborate Windows GAMESS/FMO interface for input file preparation and result visualization; it has an automatic fragmentation engine for dividing polypeptides (including proteins), nucleotides (including DNA), saccharides or any combination thereof into standard fragments. Biostation[431] is a graphical interface developed in conjunction with ABINIT-MP for FMO calculations.

One of the important advantages of fragmentation methods is that they are frequently inherently scalable to many compute nodes and cores. The FMO method in GAMESS presents a nice example of this, since it has been implemented with the associated distributed data interface (DDI)[437,438] and its generalized, multilevel partner GDDI.[1] DDI facilitates the distribution of large arrays across many compute nodes, while GDDI makes it possible to use both coarse graining and fine graining. The coarse-grained parallelism allows the calculation (for example) of the contribution from each fragment to the total energy on a different node, while the fine-grained parallelism occurs among the cores within each node. This approach enables FMO calculations to take advantage of virtually perfect scaling on tens of thousands of cores. While the example presented here is for FMO in GAMESS, most of the fragmentation methods discussed here have the potential for similar scaling.

4. APPLICATIONS

The applications below are described in subsections according to the application field.

4.1. Homogeneous Clusters and Explicit Solvent Treatments

Fragment methods are a powerful alternative to force fields in running molecular dynamics simulations (in fact, some ab initio based methods are thought of as new generation force fields[48,75,316]). Of course, MD simulations require very accurate gradients, and in some cases the gradient used in the initial MD simulations had questionable accuracy.[267] Also, during QM MD simulations, there is no enforced bond definition, and bonds can be broken along the trajectory. If this happens, the pieces of the broken fragment can move away from each other and attach themselves to other fragments. If this happens, the fragment definition becomes inappropriate. The dynamic fragmentation[270] method is designed to address this issue: When a fragment breaks into pieces, the fragment definition is changed on the fly; this can, however, result in a discontinuity on the potential energy surface. Some fragment methods like EFP have frozen fragment geometries, and do not suffer from this problem. On the other hand, rigid fragments cannot adjust their structure to the environment, which limits their accuracy.

Many methods assume that fragments do not interact very strongly. But a small interfragment distance can cause large charge transfer and other strong interactions. In the methods that use multipole models to describe the electrostatics, the short-range multipole interaction can break down because of the neglect of charge penetration; multipole screening[202,239] can be used as a remedy.

Nevertheless, fragment based methods can take into account many-body polarization and charge transfer, and thus are new generation force fields. In charged solutes these effects of the polarization and charge transfer between the solute and solvent are large and difficult to neglect.

One important application of fragment methods is the investigation of chemical reactions of small molecules in solution with MD, with solvent molecules represented as fragments, on a par with the solute. Sato et al.[439] applied the FMO method to study a S_N2 reaction of the hydrolysis of the methyl diazonium ion, and found two types of reaction pathway. Solute−solvent charge transfer plays a very important role in this reaction and the real trajectories are quite different from the textbook images of the S_N2 reaction. Sato et al.[440] used the FMO method to study the amination of formaldehyde in water and found that this reaction proceeds via a stepwise mechanism through a zwitterionic intermediate, not by a concerted mechanism.

Pomogaev et al.[441] employed ELG to compute the absorption spectra for explicitly solvated estradiol and tryptophan in the Trp-cage protein, using snapshots from classical MD simulations. Kistler and Matsika[442] applied the multilayer FMO-MCSCF approach to study solvatochromic shifts of uracil and cytosine, using snapshots of classical MD simulations. They showed by comparison of various methods that explicit solvent methods including the use of QM or MM to describe the solvent produce similar results, whereas the continuum-based PCM for some excited states gives significantly different and parameter dependent results.

Fujiwara et al.[443] applied the FMO method to MD simulations of Zn^{2+} solvated by a droplet with 64 water molecules and reproduced the experimental value of the first peak in the radial distribution function to within 0.01 Å.

Several EFP1 and QM/EFP1 studies have been devoted to the investigation of structures and energetics of small to medium water clusters.[190,192,444] Structural properties of bulk water[193] and origins of enhancement of the dipole moment in bulk water[445] were also investigated. Another important area of application of QM/EFP1 methodology is hydration of various ions, including halogen[446] and molecular anions,[447,448] as well as metal[449,450] and molecular[451] cations.

Hydration and dissociation of ionic species were investigated by Peterson and Gordon[452] for NaCl and by Yoshikawa and Morales[453] for LiOH. Solvation effects on the electronic properties of a solute, such as electronegativity, hardness, HOMO−LUMO gaps, were investigated for ammonia[454] and a set of polyatomic molecules and ions.[455]

Solvent effects on neutral-zwitterionic equilibrium of aminoacids were also studied with QM/EFP1. Day et al.[456] investigated small (up to ten waters) clusters of glutamic acid and water. Solvation of glycine was studied by Bandyopadhyay and Gordon.[196] Solvation of alanine by increasing number of water molecules was studied by Mullin and Gordon.[457,458] Song et al. investigated solvent effects on the conformational potential energy surfaces (gauche versus trans conformations) of acetylcholine and acetylthiocholine ACh and ATCh.[459]

The EFP2 model has been used to investigate intermolecular interactions in complex molecular clusters. Adamovic and Gordon showed that water−methanol mixtures are heterogeneic at the microscopic level.[460] The same authors investigated interactions in styrene clusters,[461] where both H-bonding and π-bonding structures take place. π-stacking interactions were investigated in benzene dimer,[202] substituted benzene dimers,[462] and benzene-pyridine dimers;[463] an intriguing competition between H-bonding and $\pi-\pi$ bonding was observed in water-benzene clusters.[464] Interactions in π-stacks of DNA base-pairs were studied by Ghosh et al.[200] The shortening of the B−N bond in H_3BNH_3 on going from the gaseous to the solid state was explained using the EFP2 method.[465]

The QM/EFP method has also been used to investigate the electronic excited states of chromophores in solution. The spectroscopy of enzyme active sites in the presence of several

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

EFP1 waters was studied by Krauss.[466,467] Aqueous solvation of the lowest singlet excited state of Coumarin 151 was investigated with the CIS/EFP1 method.[176] The absorption spectrum of acetone in water was modeled with TD-DFT/EFP1 by Yoo et al.[177] Ab initio molecular dynamics with CASSCF/EFP1 was used to investigate the dynamics of solvated excited states in coumarin 151[468] and excited state hydrogen transfer in 7-azaindole.[469] The EFP2 model combined with CIS(D) was used to investigate solvatochromic shifts of singlet and triplet electronic states of para-nitroaniline in water, dioxane, and cyclohexane by Kosenkov and Slipchenko.[179] A combination of the EOM-CCSD for ionization potentials (IP) method with EFP allowed accurate description of vertical ionization energy (VIE) of thymine in bulk water.[181] Intrestingly, the first hydration shell increases VIE by ~0.1 eV, while subsequent solvation lowers the ionization energy and the bulk value of the solvent-induced shift of thymine's VIE is approximately −0.9 eV.

Influence of the water environment on energetics and pathways of various reactions were investigated with QM/EFP1 methods. The energetics of the Menshutkin reaction between ammonia and methyl bromide in the presence of an increasing number of water molecules were studied by HF/EFP1.[194] Adamovic and Gordon showed that MP2/EFP1 provides an accurate description of hydration effects on structures and barriers in S_N2 reaction between Cl^- and CH_3Br.[195] The EFP solvent model was also used to study the kinetics of the hydrogen abstraction from H_2O_2 by $OHO^{\bullet}$.[470] The reaction mechanism of phosphate monoester aminolysis in aqueous solution was investigated with DFT/EFP1 by Ferreira et al.[471] The role of solvation on the relative thermodynamic stabilities of cis- and trans-platinum dichloride in aqueous solution was investigated by Hush et al.[472] The energy profiles for the reaction $OH^- + CO_2 \rightarrow HCO_3^-$ in the presence of 30 water molecules were investigated by Nemukhin et al.[473] In another study by the same authors,[474] the QM/EFP method was used to model the reactions of hydroxymethyl radical (CH_2OH)-C with glutathione tripeptide (GSH) and with methylthiol(CH_3SH) in water.

Jose and Gadre[475] applied the MTA method to investigate the properties of Li clusters and CO_2 clusters.[476] Mahadevi et al.[477] studied benzene clusters using the MTA method. Yang et al.[478] employed GEBF to study water clusters.

4.2. Biochemical Systems

Biochemical systems are a major application field for fragment methods, because the molecules in biology are large, often containing thousands of atoms. Fragment methods have several competitors, one is force fields; another is electronic structure approaches that rely either on semiempirical approximations[479,480] or multilayer approaches, like ONIOM[147] and QM/MM.[481] In addition, purely ab initio methods are also becoming increasingly efficient.[482,483]

Force fields are highly tuned for biochemical calculations, and are ubiquitously used. Their drawback is that frequently they are not polarizable and do not account for charge transfer (EFP and SIBFA are examples in which both are considered), and in addition, they are often difficult to use in practice by nonexperts if the system contains nonstandard parts for which no precomputed parameters are available. Force fields do have a major advantage of their speed and the ability to use them for molecular dynamics simulations, often essential to describe biological processes. On the other hand, ab initio based methods, including fragment-based approaches can incorporate full many-body polarization and charge transfer, and they do not rely on the need for parameters (aside from basis sets); of course the associated cost is an increased amount of computations.

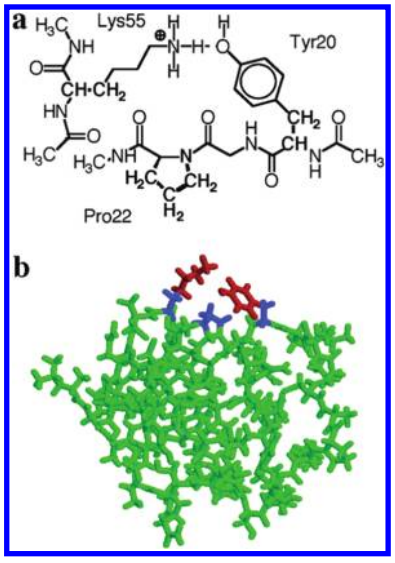


Figure 6. Subsystem of OMTKY3 (protein turkey ovomucoid third domain) used to obtain the buffer region (bold) used for (b) ab initio/buffer/EFP regions (red/blue/green) used for the computation of the pK_a of Lys55. Reproduced with permission from ref 485. Copyright 2005 American Chemical Society.

4.2.1. Polypeptides, Proteins, Saccharides, and Oligoamides. Jensen and co-workers[484−490] have successfully applied the QM/EFP approach with a buffer region to predicting pK_a values (Figure 6) of ionizable residues in proteins. Xie et al.[491] employed the X-Pol method at the AM1 level to perform a 50 ps MD simulation of bovine pancreatic trypsin inhibitor in water with periodic boundary conditions.

Komeiji et al.[492] applied the FMO/MP2/6-31G* method to analyze the dependence of the change in the electronic structure of ubiquitin (PDB 1UBQ) on the thickness of solvating water layers. Five configurations corresponding to local minima were averaged. He et al.[493] applied the FMO/MP2/PCM/6-31G* method to study the ability of FMO and empirical dispersion to discriminate between the native and decoy structures for the Pin1 WW domain (PDB: 1I6C) and the Co repressor protein (PDB 1ORC). Sawada et al.[494] used the FMO/RHF/PCM/6-31G* method to optimize the structure of helical heparin oligosaccharides (PDB 1HPN). They found that the optimized structure is in good agreement with the NMR experiment. In addition, a comparison of the FMO predictions with those of the force field revealed the differences pointing to possible deficiencies in the force field model.

Huang et al.[495] applied the KEM with MP2/6-31G** to analyze the interactions in vesicular stomatitis virus nucleoprotein (PDB 2QVJ) containing 33,175 atoms. Duan et al.[496] used the MFCC method at the HF, B3LYP, and MP2 levels with the 6-31+G* basis set to investigate the interaction of HIV-1 protease with the water molecule that bridges the flaps of the protease with the inhibitors. Dong et al.[497] applied the GEBF method at the DFT/6-311+G** level to study the formation of single and double helices of aromatic oligoamides. Deshmukh et al. applied the MTA method to study intramolecular hydrogen bonding in

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

sugars[498] at the MP2/6-311++G(2d,2p) level and polypeptides[499] with the DFT/6-311++G** level.

4.2.2. Protein−Ligand Binding. A number of applications of the FMO method to various protein−ligand complexes have been performed. Fukuzawa et al.[500−502] studied human estrogen receptor and showed good correlation with the experimental binding energies. Sawada et al.[503−506] performed a number of FMO applications to influenza virus hemagglutinin and showed the importance of considering the full-sized hemagglutinin upon the binding energies. A very detailed analysis[507] of the recognition patterns of sialosides by avian and human influenza hemagglutinins was performed using MP2/PCM/6-31G*, resulting in good agreement with the experimental binding energies In another study[508] at the same level it was shown that binding of influenza A virus hemagglutinin to the sialoside receptor is not controlled by the homotropic allosteric effect (i.e., hemagglutinin trimerization does not not increase the binding energy per ligand). Future mutations[509] of the influenza virus were predicted by a combination of hemadsorption experiment and quantum chemical calculations for antibody binding, and details of the ligand binding were investigated.[264,510,511]

Yamagishi et al.[512−514] performed an analysis of the functions of key residues in the ligand-binding pocket of vitamin D. Ito et al.[515−517] analyzed the role of the functional groups in retinoid X receptor and studied the influence of mutations upon the transcriptional activation. Nakanishi et al.[518] applied FMO/MP2/6-31G* to elucidate the molecular recognition mechanism in the FK506 binding protein. A number of other protein−ligand complexes,[519−530] as well as systems involving nucleic acids,[531−536] were also studied with the FMO method.

Taking advantage of the FMO method in the GAMESS and ABINIT-MP packages, several pharmaceutical companies have used it for drug-design related research. Ozawa et al. applied the FMO MP2/6-31G* approach[537] to demonstrate that CH/π hydrogen bonds determine the selectivity of the Src homology 2 domain to tyrosine phosphotyrosyl peptides. This was followed by a study[538] that showed the importance of CH/π hydrogen bonds in rational drug design as exemplified by leukocyte-specific protein tyrosine kinase. Fujimura and Sasabuchi[539] applied FMO/MP2/6-31G* to elucidate the role of fluorine atoms in a fluorinated prostaglandin agonist. Ohno et al.[540] employed FMO MP2/6-31G and discussed the strong correlation of pair interaction energies (PIEs) with the drug's potency.

The MFCC method has been applied to a number of ligand binding studies.[541−545] Huang et al.[546] employed KEM to study the interaction of aminoglycoside drugs and ribosomal A site RNA targets. Using the ELG method, Orimoto et al.[547] investigated the electronic structure of B-type poly(dG)·poly(dC) DNA.

4.2.3. Quantitative Structure−Activity Relationship (QSAR). It is very difficult to evaluate the free energies of binding from first principles with the accuracy of 1 kcal/mol, which is often the difference between several ligands in protein−ligand binding. In practical ab initio calculations, there are not only basis set and wave function limitations and the difficulty in describing solvent and counterions but also the entropic contribution at room temperature requires proper configurational sampling, which at present is usually done with MD and can require long-time trajectories.

In QSAR studies, one introduces empirical factors by taking computed physical quantities, $\{X_i\}$ called descriptors as arguments of a function $f(\{X_i\})$ (often, f is the binding energy), the coefficients in which are optimized for some training set of systems with $f(\{X_i\})$ known experimentally. The obtained relation is used to predict properties of compounds outside the training set, for which $\{X_i\}$ are computed. The underlying condition for successful QSAR studies is that descriptors should correlate with the desired property. Because of the fitting nature of QSAR, various completely unrelated methods can be used to compute descriptors.[548] For example, some descriptors can be computed from gas phase ab initio based calculations, the solvation effects can be estimated with the Poisson−Boltzmann model, and the entropic factor descriptor can be obtained from other simple models such as the number of rotatable bonds.[549]

Fragment-based methods seem to be a perfect match for fragment-based drug discovery (FBDD)[550,551] and they can be utilized to provide descriptors in QSAR. In particular, a very useful set of descriptors is given by the pair interactions between ligands and constituent parts of the protein (residues or residue fragments). Because the FMO method provides PIEs as a byproduct of the calculations, it is straightforward to apply FMO to QSAR studies, as reviewed by Yoshida et al.[552] In other one-step methods, such as KEM, MFCC, or PMISP methods, one can have similar pair interaction energies, whereas two-step methods operate with the properties of the whole system. Ishikawa et al.[553] discussed the basis sets effects upon the PIEs.

Yoshida et al.[554] employed the FMO RHF/6-31G method in QSAR studies of cyclic urea type HIV-1 PR inhibitors, using the sum of the residue fragment−ligand PIEs and the charge transfer between the ligand and protein from FMO calculations as descriptors. They found a strong correlation between the FMO binding energy in vacuum and the sum of protein−ligand PIEs, thus only the former was used.

Because the foregoing strategy was also used in further studies, it is instructive to comment on the relation between the sum of the protein−ligand PIEs ΔE_{PL} in the protein−ligand complex PL and the binding energy ΔE_{b} computed as the difference between the energies of the complex PL and the isolated protein (P) and ligand (L). Both ΔE_{PL} and ΔE_{b} can be used as descriptors; the former is often used as an approximation and replacement of the latter. For simplicity, the deformation energy, that is, the change in the geometry of the protein and the ligand, when comparing their isolated and complexed minima, is implicitly included in the following derivations.[518]

$$\begin{aligned}\Delta E_{\mathrm{b}} = E_{\mathrm{PL}} - E_{\mathrm{P}} - E_{\mathrm{L}} = [&\sum_{I\in P} E_I'^{\mathrm{P(PL)}} + \sum_{I\in L} E_I'^{\mathrm{L(PL)}} \\ &+ \sum_{\substack{I>J\\ I,J\in P}} \Delta E_{IJ}^{\mathrm{P(PL)}} + \sum_{\substack{I>J\\ I,J\in L}} \Delta E_{IJ}^{\mathrm{L(PL)}} + \sum_{I\in P, J\in L} \Delta E_{IJ}^{\mathrm{PL(PL)}}] \\ &- [\sum_{I\in P} E_I'^{\mathrm{P}} + \sum_{\substack{I>J\\ I,J\in P}} \Delta E_{IJ}^{\mathrm{P}}] - [\sum_{I\in L} E_I'^{\mathrm{L}} + \sum_{\substack{I>J\\ I,J\in L}} \Delta E_{IJ}^{\mathrm{L}}] \\ \Delta E_{\mathrm{PL}} = &\sum_{I\in P, J\in L} \Delta E_{IJ}^{\mathrm{PL(PL)}} \end{aligned} \quad (125)$$

P(PL) indicates the partial properties of P in the complex PL (and, similarly, for L). The difference is then

$$\begin{aligned}\Delta E_{\mathrm{b}} - \Delta E_{\mathrm{PL}} = &\sum_{I\in P} \Delta E_I'^{\mathrm{P(PL)}} + \sum_{I\in L} \Delta E_I'^{\mathrm{L(PL)}} \\ &+ \sum_{\substack{I>J\\ I,J\in P}} \Delta\Delta E_{IJ}^{\mathrm{P(PL)}} + \sum_{\substack{I>J\\ I,J\in L}} \Delta\Delta E_{IJ}^{\mathrm{L(PL)}} \end{aligned} \quad (126)$$

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

where (A = P or L)

$$\begin{aligned} \Delta E_I^{\prime A(\mathrm{PL})} &= E_I^{\prime A(\mathrm{PL})} - E_I^{\prime A} \\ \Delta\Delta E_{IJ}^{\prime A(\mathrm{PL})} &= \Delta E_{IJ}^{\prime A(\mathrm{PL})} - \Delta E_{IJ}^{\prime A} \end{aligned} \tag{127}$$

$\Delta E_I^{A(\mathrm{PL})}$ and $\Delta\Delta E_{IJ}^{A(\mathrm{PL})}$ are the internal fragment and pair interaction energies, respectively, and are affected by the polarization and deformation in the complex formation. In other words

$$\Delta E_\mathrm{b} - \Delta E_\mathrm{PL} = \Delta E_\mathrm{P}^{\mathrm{pold+def}} + \Delta E_\mathrm{L}^{\mathrm{pold+def}} \tag{128}$$

$$\Delta E_A^{\mathrm{pol+def}} = \sum_{I \in A} \Delta E_I^{\prime A(\mathrm{PL})} + \sum_{\substack{I>J \\ I,J \in A}} \Delta\Delta E_{IJ}^{A(\mathrm{PL})} \tag{129}$$

The difference between the binding energy ΔE_b and the sum of the pair interaction energies ΔE_PL is simply the destabilization polarization (denoted as "pold") plus the deformation energy of the protein and ligand. $\Delta E_A^{\mathrm{pold+def}}$ does not include all terms in the polarization. The polarization process (see PIEDA[300]) between isolated systems (P and L here) is divided into two contributions: first, each of the two interacting systems is destabilized relative to its own lowest energy state due to the polarization by the other (this gives the destabilization polarization). Consequently, the polarized systems interact, and this interaction energy includes the electrostatic energy, part of which is the stabilizing polarization energy, exchange-repulsion, charge transfer, dispersion and higher order terms. These terms are separable in the EDA[57] or its FMO extension, PIEDA[300] when A are individual fragments.

When the polarization and deformation energies vary considerably from ligand to ligand (for instance, when comparing ligands with different charges), the use of ΔE_PL in place of ΔE_b may lead to errors in consequent QSAR and care should be taken to use a proper training set. Finally, for completeness, it is useful to separate[518] the sum of the polarization and deformations energies given by the difference in the energy of the protein or ligand in the complexed and isolated states.

$$\begin{aligned} \Delta E_A^{\mathrm{pold+def}} &= E^{A(\mathrm{PL})} - E^A = E^{A(\mathrm{PL})} - E^A - \tilde{E}^A + \tilde{E}^A \\ &= \Delta E_A^{\mathrm{pold}} + \Delta E_A^{\mathrm{def}} \\ \Delta E_A^{\mathrm{pold}} &= E^{A(\mathrm{PL})} - \tilde{E}^A \\ \Delta E_A^{\mathrm{def}} &= \tilde{E}^A - E^A \end{aligned} \tag{130}$$

$\tilde{E}^A$ is the energy of isolated protein or ligand at the geometry in the complex, whereas E^A is computed at the geometry of A in its minimum. $E^{A(PL)}$ is the internal energy of A in the complex PL (defined as the FMO sum of monomer and dimers terms for fragments in A, cf., eq 37). This analysis can be used for any type of interactions, not just those between a protein and a ligand, and more subsystems can be defined. In addition, the FMO method provides all details of individual fragment contributions in the total sums, for further insight regarding drug design and other applications. Moreover, this analysis defines the protein and ligand destabilization polarization energy $\Delta E_A^{\mathrm{pold}}$ (often neglected in MM studies), as well as the deformation energy $\Delta E_A^{\mathrm{def}}$ (typically included). The stabilizing polarization (pols) and charge transfer, as well as other types of interactions are included in the sum of the protein−ligand PIEs ΔE_PL. The former can be estimated[518] using the simple response relation $\Delta E_A^{\mathrm{pols}} \approx -2\Delta E_A^{\mathrm{pold}}$ (see actual calculations[300] for numerical justification), which leads to the total polarization

$$E_A^{\mathrm{pol}} = \Delta E_A^{\mathrm{pold}} + \Delta E_A^{\mathrm{pols}} = -\Delta E_A^{\mathrm{pold}} \tag{131}$$

This way, the total polarization energy of the protein and ligand in their complex can be estimated, which requires 5 calculations (complex, plus protein and ligand, both at the geometry in the complexed and isolated states).

Fischer et al.[555] improved the scoring functions representing the binding energy for human estrogen receptor subtype α and human retinoic acid receptor of isotype γ, using atomic charges from FMO calculations at the RHF/STO-3G level and concluded that such quantum scoring functions (QSF) describe the electrostatics accurately, and that QSF performs better than force field analogues. Yoshida et al.[556] applied FMO/RHF/6-31G to QSAR studies of the binding affinity of substituted benzenesulfonamides with carbonic anhydrase, noting the difficulties in modeling Zn^{2+}-containing systems. Hitaoka et al.[557] used FMO/MP2/6-31G in QSAR studies of the binding affinity of sialic acid analogues with influenza virus neuraminidase-1, where they divided the protein into three binding pockets and used the sums of PIEs for these subsystems as separate descriptors.

Munei et al.[558] used FMO/RHF/6-31G in QSAR studies of the binding affinity of substituted benzenesulfonamides with carbonic anhydrase. Mazanetz et al.[548] used FMO/MP2/6-31G* in QSAR studies of cyclin-dependent kinase 2 inhibitor potency and did a careful comparison of FMO and MM derived QSAR models popular in drug design industry. They found that FMO outperformed all three MM based QSAR models.

4.2.4. Excited States and Chemical Reactions. Excited states and chemical reactions are an attractive application field of ab initio based methods. Ishida et al.[559] using the FMO method at the RHF and MP2 levels with the 6-31G(d) basis set analyzed the mechanism of the chorismate mutase reaction. Nakamura et al.[560] applied the FMO/MP2/6-31G method to clarify the role of K151 and D180 in L-2-haloacid dehalogenase from Pseudomonas sp. YL (PDB 1ZRM). Pruitt et al.[289] using the multilayer open shell (MP2/6-311G**: CR-CC(2,3)/3-21G) FMO method evaluated the enthalpy of the reversible addition−fragmentation chain transfer (RAFT) reaction involving radical species.

In treating excited states with the FMO method, a frequently used approach is to apply an excited state calculation such as CI (or TDDFT) to a single fragment (chromophore), computed in the electrostatic field due to the other fragments, evaluated using RHF (or DFT). In this method (called FMO1, indicating that only monomers are computed) the electrostatic effect upon excited states is considered, which is often the major effect of the environment (i.e., the rest of the system excluding the chromophore), and higher order effects such as charge transfer can be considered including dimer calculations of excited states (FMO2).

Mochizuki et al. did several FMO1 studies. They computed[561] red fluorescent protein with CIS(D)/6-31G(d), they simulated emission spectra of bioluminescent luciferases[562] with CIS(D)/6-31G, and they also performed CIS(D)/6-31G(d) calculations on the family of red[563] as well as blue and yellow[564] fluorescent proteins. Chiba et al. computed the yellow photoactive protein (PDB 2PHY) both in gas phase[284] and solution[284] using FMO/TDDFT/6-31G(d), considering both FMO1 and FMO2 types of excitations. Ikegami et al.[565] analyzed the asymmetric excitations in the left and right branches of the photosynthetic reaction center of *Blastochloris viridis* (PDB 1PRC) using the FMO method, where the asymmetry was studied for monomers and selected dimers at the CIS/6-31G(d) level.

dx.doi.org/10.1021/cr200093j |*Chem. Rev.* XXXX, XXX, 000–000

Milne et al.[566] applied FMO at the FMO1-TDDFT/6-311G(d,p) and FMO2-MP2/6-311G(d,p) levels to investigate the role of AMP protonation in firefly luciferase pH-sensitivity.

4.3. Solid-State Applications

For systems with perfect periodicity, cluster based approaches compete with the methods employing periodic boundary conditions, which often use plane waves as a natural basis set for these systems. The latter group of methods is ultimately better, however, given the practical limitations of the existing theories and computer programs, and the difficulty in deriving and implementing high level of calculations for them often drives users to use cluster based approaches.

An important question is how relevant is the electron delocalization for a property one is interested in? Some properties, such as band gaps and the density of states, appear to require a consideration of the whole system, and fragment-based methods are at a serious disadvantage. However, there are ways to account for the delocalization at the final step by building the Fock matrices for the whole system, as is done in FMO-MO, FMO-LCMO, and FMO/F, or in a different way in ELG. On the other hand, one is often interested in some local properties in periodic systems, such as the interactions determining their global properties, and for this, pair interactions in the FMO and KEM methods were found to be useful as described below. Applications of fragment-based methods to π-conjugated systems[329] and graphene[339] suggest their usefulness even for such delocalized systems.

4.3.1. Crystals, Surfaces, and Nanomaterials. Huang et al.[567] applied the KEM/MP2/6-31G(d,p) method to investigate the interactions in the crystal of two molecules TDA1 and RangDP52. Fukunaga et al.[568] applied the FMO/TDDFT/6-31G(d,p) method to investigate the role of intermolecular interactions upon the excitations energies in three isomers of quinacridone crystals. To facilitate calculations and make them more realistic, an embedding model was used, in which a cluster of quinacridone molecules was immersed in the field of a large number of atomic charges, computed with the BLYP functional and periodic boundary conditions, with the 6-31G(d,p) basis set.

Faujasite zeolites were modeled by Fedorov et al.[234] using the FMO/RHF/6-31G(d) method. It was found that the adsorption energies can be quite accurately modeled with the FMO method, despite a large number of detached bonds. The fragmentation scheme is shown in Figure 7. The phenol molecule occupies a place in the zeolite pore near the aluminum-containing fragment with acidic hydrogens.

Zhang et al.[569] applied the ELG RHF/STO-3G method to study the adsorption of Si and C chains onto unfaulted and faulted Si(111) surfaces. Chen et al.[570] calculated the electronic structure of the single-wall pristine boron nitride (BN) and boron nitride-carbon (BN/C) heterostructured nanotubes using ELG at the RHF/6-31G level.

Fedorov et al.[249] applied the FMO method at the B3LYP/3-21G(d) level to optimize the structure of silicon nanowires of diameter 1.2 and length 4.8 (nm), and showed that FMO optimized structures closely agree with those by PBC methods and experiment, suggesting that the geometry can be optimized with reasonable results using fragment methods. Otaki and Ando[571] investigated the dielectric properties of 5-bromo-9-hydroxyphenalenone and found that the induced polarization is enhanced by the weak hydrogen bonding.

4.3.2. Polymers. Many applications of fragment-based methods are performed with the ELG method, which is particularly suited for (although not limited to) linear polymer calculations. Orimoto et al.[572] applied the ELG/PM3 level to optimize structures of polysilane derivatives, poly[bis(4-propoxybutyl)silylene].

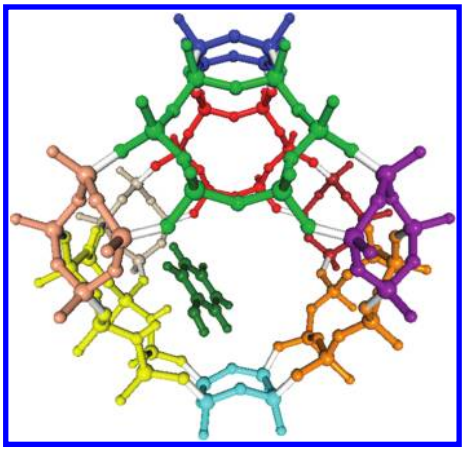


Figure 7. Fragmentation of the complex of phenol and faujasite zeolite, computed with FMO.[234] The phenol molecule is shown in dark green, and the active catalytic site containing three aluminum atoms and three acidic hydrogens is shown in yellow.

The ELG method was used extensively to calculate polarizabilities and hyperpolarizabilities of important nonlinear optics materials. These two properties were studied using ELG by a number of researchers. Ohnishi et al.[573] computed donor/acceptor substituted polydiacetylenes at the RHF/6-31G level. Yu et al. using RHF/6-31G calculated polyimides[574] and $[Li^+[calix[4]pyrrole]Li^-]_n$, up to 15 units.[575] Pomogaeva et al.[576] at the RHF/6-311G level with ECP/VDZ for chalcogen atoms (S, Se and Te) studied series of benzo-2,1,3-chalcogendiasoles ribbon oligomers (up to 15 units). At the unrestricted PM3 level, Orimoto et al.[577] calculated a pyrrole-based spin-polarized molecular wire containing 1-pyrrolylphenyl nitronyl nitroxide with oligothiophene units under the influence of an applied electric field. Yan et al.[578] computed meso−meso-linked metalloporphyrin oligo-mers up to 22 units at RHF/6-31G (ECP/VDZ for metals Mg, Zn and Ni).

Pomogaev et al. using ELG at the CIS/INDO level computed absorption spectra for aromatic molecules (benzene, anthracene, 4-dicyanomethylene-4H-pyran, tryptophan, and estradiol) bound to polyethylene[579] and explicitly solvated estradiol and tryptophan.

5. CONCLUSIONS AND PROGNOSIS

During several decades, fragment-based methods have come a long way from the initial stage of method development to large scale applications which span many types of systems: molecular clusters, proteins, DNA, oligosaccharides, zeolites, quantum dots, nanowires, and others. Despite the very considerable progress, they remain underused; there may be several reasons for this. Some of the software developments are only locally implemented, making it difficult for most interested users to utilize the methods. Many, perhaps most methods are specific to one specific program; users who are unfamiliar with that program may have an inertial barrier to using it. Second, many applications so far have been performed in what should be considered demonstrative fashion, with low level wave functions and basis sets. Third, in some cases the applications are not conducted without properly considering all necessary effects and factors; the most conspicuous example is the need to incorporate solvent effects and entropy in biochemical applications.

dx.doi.org/10.1021/cr200093j |*Chem. Rev.* XXXX, XXX, 000–000

Another concern which many potential users of fragment-based methods face is the uncertainty of the accuracy. An automatic error control with an estimate of possible errors would be a welcome addition to these methods, whose accuracy is known to depend upon the physical nature of the systems, in particular, the electron density localization, as well as on the wave function and basis set. However, in the case of "standard" systems such as proteins, the error behavior can be studied in advance and reasonable automatic ways to fragment can be suggested; on the other hand, nonstandard systems may require some preliminary accuracy tests which can be seen as a barrier to users who are interested in a completely automatic, black-box way of doing calculations. It can be expected that some properties that result from fragment-based calculations, for example, the interfragment charge transfer in FMO,[98] can be utilized to estimate the error in practical calculations. Methods that have frozen fragments, such as EFP, can result in errors that are caused by the inability to relax the internal structures of the fragments.

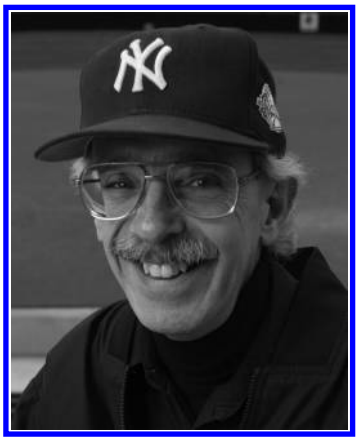
Nevertheless, fragment-based methods also offer many advantages. One is the efficiency and the ability to compute realistic systems. Another is the additional information which they can deliver, such as the intrinsic details of the physical picture of the interactions in the system. As shown above, the application field is very broad encompassing most systems of finite size, in which chemists and physicists are interested. An important example of the match between fragmentation models and their applications is given by fragment-based drug discovery, where some progress involving the use of the FMO method has already been achieved.[551] In the near future with the revolutionary progress in computer technology, the advent of multicore CPUs and GPUs, increasing levels of calculations and the ease of their execution, fragment-based methods will grow more popular in the computational community.

AUTHOR INFORMATION

Corresponding Author

*E-mail: mark@si.msg.chem.iastate.edu.

BIOGRAPHIES



Mark Gordon, Frances M. Craig Distinguished Professor of Chemistry at Iowa State University and Director of the Ames Laboratory Applied Mathematical and Computational Sciences program, was born and raised in New York City. After completing his B.S. in Chemistry in 1963, Professor Gordon entered the graduate program at Carnegie Institute of Technology, where he received his PhD in 1967 under the guidance of Professor John Pople, 1998 Chemistry Nobel Laureate. Following a postdoctoral research appointment with Professor Klaus Ruedenberg at Iowa State University, Professor Gordon accepted a faculty appointment at North Dakota State University in 1970, where he rose through the ranks, eventually becoming distinguished professor and department chair. He moved to Iowa State University and Ames Laboratory in 1992. Professor Gordon's research interests are very broadly based in electronic structure theory and related fields, including solvent effects, the theory of liquids, surface science, the design of new materials, and chemical reaction mechanisms. He has authored more than 500 research papers and is a member (and Treasurer) of the International Academy of Quantum Molecular Science. He received the 2009 American Chemical Society Award for Computers in Chemical and Pharmaceutical Research, and he is a Fellow of the American Physical Society and the American Association for the Advancement of Science. In 2009, he was elected to the inaugural class of American Chemical Society Fellows.

Dmitri G. Fedorov received his M.S. in quantum chemistry from Saint Petersburg State University in Russia in 1993. He was awarded Ph. D. in physical chemistry at Iowa State University in 1999, working under the supervision of Prof Mark S. Gordon. He spent two years at the University of Tokyo as a JSPS postdoctoral research fellow, working with Prof Kimihiko Hirao and then moved to the National Institute of Advanced Industrial Science and Technology in Japan in 2002, where he is currently employed as a senior research scientist at the Nanosystem Research Institute. His research interests include spin−orbit coupling and other relativistic effects, as well as development of quantum-chemical methods for describing large molecular systems.

Spencer R. Pruitt received his B.S. degree in chemistry from the University of Minnesota Duluth in 2006. Afterwards he

dx.doi.org/10.1021/cr200093j |Chem. Rev. XXXX, XXX, 000–000

entered the graduate program at Iowa State University, joining the research group of Prof. Mark S. Gordon, where he is currently working towards his PhD in chemistry. His research interests include the study of intermolecular interactions, reactivity of ionic liquids and the development of the Fragment Molecular Orbital method within GAMESS.

Lyudmila V. Slipchenko is an Assistant Professor in the chemistry department at Purdue University. She received her B.S. and M.S. in Applied Mathematics and Physics from Moscow Institute of Physics and Technology and a Ph.D. in Chemistry from the University of Southern California. As a postdoctoral research associate, she worked with Prof. Mark S. Gordon at ISU. Her research focuses on the study of electronic structure, electronic excited states, and intermolecular interactions in the condensed phase.

ACKNOWLEDGMENT

D.G.F. thanks Professor Kazuo Kitaura for numerous fruitful and inspiring discussions. S.R.P. and M.S.G. thank Professor Michael Collins and Dr. Anuja Rahalkar for helpful discussions regarding the SFM and MTA methods. D.G.F. was partially supported by the Next Generation SuperComputing Project, Nanoscience Program (MEXT, Japan). This work was supported in part by grants from the National Science Foundation (Petascale Applications S.R.P., D.G.F., M.S.G.; CAREER award L.V.S.), by the Air Force Office of Scientific Research (S.R.P., M.S.G.), by a grant from the Department of Energy ASCR and BES programs to the Ames Laboratory (M.S.G.), and by ACS Petroleum Research Foundation (L.V.S.).

REFERENCES

(1) Fedorov, D. G.; Olson, R. M.; Kitaura, K.; Gordon, M. S.; Koseki, S. *J. Comput. Chem.* **2004**, *25*, 872.
(2) Adams, W. H. *J. Chem. Phys.* **1961**, *34*, 89.
(3) Klessinger, M.; McWeeny, R. *J. Chem. Phys.* **1965**, *42*, 3343.
(4) Christoffersen., R. E.; Maggiora, G. M. *Chem. Phys. Lett.* **1969**, *3*, 419.
(5) Christoffersen, R. E.; Cheney, V.; Maggiora, G. M. *J. Chem. Phys.* **1971**, *54*, 239.
(6) Shipman, L. L.; Christoffersen, R. E. *Chem. Phys. Lett.* **1972**, *15*, 469.
(7) Shipman, L. L.; Christoffersen, R. E. *Proc. Nat. Acad. Sci U.S.A.* **1972**, *69*, 3301.
(8) Cheney, B. V.; Christoffersen, R. E. *J. Chem. Phys.* **1972**, *56*, 3503.
(9) Christoffersen, R. E. *Int. J. Quantum Chem.* **1973**, *7*, 169.
(10) Shipman, L. L.; Christoffersen, R. E. *Theor. Chim. Acta* **1973**, *31*, 75.
(11) Shipman, L. L.; Christoffersen, R. E.; Cheney, V. *J. Med. Chem.* **1974**, *17*, 583.
(12) Oie, T.; Maggiora, G. M.; Christoffersen, R. E. *Int. J. Quantum Chem.* **1976**, *10*, 119.
(13) Petke, J. D.; Christoffersen, R. E.; Maggiora, G. M.; Shipman, L. L. *Int. J. Quantum Chem.* **1977**, *12*, 343.
(14) Murk, D.; Nietzsche, L. E.; Christoffersen, R. E. *J. Am. Chem. Soc.* **1978**, *100*, 1371.
(15) Frost, A. A. *J. Chem. Phys.* **1967**, *47*, 3707.
(16) Stoll, H.; Preuss, H. *Theor. Chim. Acta* **1977**, *46*, 11.
(17) Stoll, H. *Phys. Rev. B* **1992**, *46*, 6700.
(18) Szekeres, Z.; Surján, P. R. *Chem. Phys. Lett.* **2003**, *369*, 125.
(19) Reinhardt, P.; Piquemal, J.-P.; Savin, A. *J. Chem. Theory Comput.* **2008**, *4*, 2020 and references therein.
(20) Yang, W. *Phys. Rev. Lett.* **1991**, *66*, 1438.
(21) Imamura, A.; Aoki, Y.; Maekawa, K. *J. Chem. Phys.* **1991**, *95*, 5419.
(22) Demelo, C. P.; Dossantos, M. C.; Matos, M.; Kirtman, B. *Phys. Rev. B* **1987**, *35*, 7847.
(23) Aoki, Y.; Imamura, A. *J. Chem. Phys.* **1992**, *97*, 8432.
(24) Korchowiec, J.; Lewandowski, J.; Makowski, M.; Gu, F. L.; Aoki, Y. *J. Comput. Chem.* **2009**, *30*, 2515.
(25) Gu, F. L.; Aoki, Y.; Korchowiec, J.; Imamura, A.; Kirtman, B. *J. Chem. Phys.* **2004**, *121*, 10385.
(26) Mitani, M.; Aoki, Y.; Imamura, A. *J. Chem. Phys.* **1994**, *100*, 2346.
(27) Aoki, Y.; Suhai, S.; Imamura, A. *J. Chem. Phys.* **1994**, *101*, 10808.
(28) Aoki, Y.; Suhai, S.; Imamura, A. *Int. J. Quantum Chem.* **1994**, *52*, 267.
(29) Imamura, A.; Aoki, Y.; Nishimoto, K.; Kurihara, Y.; Nagao, A. *Int. J. Quantum Chem.* **1994**, *52*, 309.
(30) Mitani, M.; Aoki, Y.; Imamura, A. *Int. J. Quantum Chem.* **1995**, *54*, 167.
(31) Kurihara, Y.; Aoki, Y.; Imamura, A. *J. Chem. Phys.* **1997**, *107*, 3569.
(32) Mitani, M.; Aoki, Y.; Imamura, A. *Int. J. Quantum Chem.* **1997**, *64*, 301.
(33) Imamura, A.; Aoki, Y. *Adv. Colloid Interface Sci.* **1997**, *71*−*2*, 147.
(34) Kurihara, Y.; Aoki, Y.; Imamura, A. *J. Chem. Phys.* **1998**, *108*, 10303.
(35) Ladik, J.; Imamura, A.; Aoki, Y.; Ruiz, M.; Otto, P. *J. Mol. Struct. (THEOCHEM)* **1999**, *491*, 49.
(36) Rather, G.; Aoki, Y.; Imamura, A. *Int. J. Quantum Chem.* **1999**, *74*, 35.
(37) Kim, J. T.; Lee, M. J.; Kim, U. R.; Kimura, M.; Aoki, Y.; Imamura, A. *J. Polym. Sci., Part A: Polym. Chem.* **2001**, *39*, 2677.
(38) Gu, F. L.; Aoki, Y.; Imamura, A.; Bishop, D. M.; Kirtman, B. *Mol. Phys.* **2003**, *101*, 1487.
(39) Korchowiec, J.; Gu, F. L.; Imamura, A.; Kirtman, B.; Aoki, Y. *Int. J. Quantum Chem.* **2005**, *102*, 785.
(40) Korchowiec, J.; Gu, F. L.; Aoki, Y. *Int. J. Quantum Chem.* **2005**, *105*, 875.
(41) Makowski, M.; Korchowiec, J.; Gu, F. L.; Aoki, Y. *J. Comput. Chem.* **2006**, *27*, 1603.
(42) Pomogaeva, A.; Springborg, M.; Kirtman, B.; Gu, F. L.; Aoki, Y. *J. Chem. Phys.* **2009**, *130*, 194106.
(43) Korchowiec, J.; De Silva, P.; Makowski, M.; Gu, F. L.; Aoki, Y. *Int. J. Quantum Chem.* **2010**, *110*, 2130.
(44) Makowski, M.; Korchowiec, J.; Gu, F. L.; Aoki, Y. *J. Comput. Chem.* **2010**, *31*, 1733.
(45) Dreyfus, M.; Pullman, A. *Theor. Chim. Acta* **1970**, *19*, 20.
(46) Morokuma, K. *J. Chem. Phys.* **1971**, *55*, 1236.
(47) Ohno, K.; Inokuchi, H. *Theor. Chim. Acta* **1972**, *26*, 331.
(48) Gao, J. *J. Phys. Chem. B* **1997**, *101*, 657.
(49) Kutzelnigg, W.; Maeder, F. *Chem. Phys.* **1978**, *32*, 451.
(50) Otto, P.; Ladik, J. *Chem. Phys.* **1975**, *8*, 192.
(51) Kubota, H.; Aoki, Y.; Imamura, A. *Bull. Chem. Jpn.* **1994**, *67*, 13.
(52) Barandiaran, Z.; Seijo, L. *J. Chem. Phys.* **1988**, *89*, 5739.

(53) Pascual, J. L.; Seijo, L. *J. Chem. Phys.* **1995**, *102*, 5368.
(54) Gao, J. *J. Chem. Phys.* **1998**, *109*, 2346.
(55) The term "self-consistent-charge" was used in various contexts with quite different exact mathematical meaning but alluding to the same physical concept. Some examples of the usages are (only the last one is for a fragment-based method): (a) Bardeen, J. *Phys. Rev.* **1936**, *49*, 653. (b) Rosén, A.; Ellis, D. E.; Adachi, H.; Averill, F. W. *J. Chem. Phys.* **1976**, *65*, 3629. (c) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569. (d) Sugiki, S.-I.; Kurita, N.; Sengoku, Y.; Sekino, H. *Chem. Phys. Lett.* **2003**, *382*, 611.
(56) Fedorov, D. G.; Kitaura, K. *J. Chem. Phys.* **2004**, *120*, 6832.
(57) Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325.
(58) Chen, W.; Gordon, M. S. *J. Chem. Phys.* **1996**, *100*, 14316.
(59) Bohm, M. C. *Chem. Phys. Lett.* **1982**, *89*, 126.
(60) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. *Chem. Phys. Lett.* **1999**, *313*, 701.
(61) Otto, P.; Ladik, J. *Chem. Phys.* **1977**, *19*, 209.
(62) Otto, P. *Chem. Phys.* **1978**, *33*, 407.
(63) Otto, P. *Chem. Phys. Lett.* **1979**, *62*, 538.
(64) Otto, P.; Ladik, J. *Int. J. Quantum Chem.* **1980**, *18*, 1143.
(65) Förner, W.; Otto, P.; Bernhardt, J.; Ladik, J. *Theor. Chim. Acta* **1981**, *60*, 269.
(66) Zyss, J.; Berthier, G. *J. Chem. Phys.* **1982**, *77*, 3635.
(67) Otto, P. *Int. J. Quantum Chem.* **1985**, *28*, 895.
(68) Otto, P.; Ladik, J.; Liu, S. C. *J. Mol. Struct. (THEOCHEM)* **1985**, *123*, 129.
(69) Otto, P. *Int. J. Quantum Chem.* **1986**, *30*, 275.
(70) Ladik, J. *Prog. Surf. Sci.* **1987**, *26*, 135.
(71) Otto, P. *J. Mol. Struct. (THEOCHEM)* **1989**, *188*, 277.
(72) Hankins, D.; Moskowitz, J. W.; Stillinger, F. H. *J. Chem. Phys.* **1970**, *53*, 4544.
(73) Suárez, E.; Díaz, N.; Suárez, D. *J. Chem. Theory Comput.* **2009**, *5*, 1667.
(74) Gao, J.; Cembran, A.; Mo, Y. *J. Chem. Theory Comput.* **2010**, *6*, 2402.
(75) Jensen, J. H.; Day, P. N.; Gordon, M. S.; Basch, H.; Cohen, D.; Garmer, D. R.; Kraus, M.; Stevens, W. J. *Model. Hydrogen Bond* **1994**, *569*, 139.
(76) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *J. Phys. Chem. A* **2001**, *105*, 293.
(77) Gordon, M. S.; Slipchenko, L.; Li, H.; Jensen, J. H. *Annu. Rep. Comput. Chem.* **2007**, *3*, 177.
(78) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J. P. *J. Chem. Theory Comput.* **2007**, *3*, 1960.
(79) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549.
(80) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887.
(81) Moszynski, R. *Mol. Phys.* **1996**, *88*, 741.
(82) Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. *J. Chem. Phys.* **2005**, *123*, 214103.
(83) Bukowski, R.; Podeszwa, R.; Szalewicz, K. *Chem. Phys. Lett.* **2005**, *414*, 111.
(84) Amos, R. D.; Handy, N. C.; Knowles, P. J.; Rice, J. E.; Stone, A. J. *J. Phys. Chem.* **1985**, *89*, 2186.
(85) Hesselmann, A.; Jansen, G.; Schutz, M. *J. Chem. Phys.* **2005**, *122*, 014103.
(86) Podeszwa, R.; Bukowski, R.; Szalewicz, K. *J. Chem. Theory Comput.* **2006**, *2*, 400.
(87) Kitaura, K.; Morokuma, K. Variational approach (SCF ab-initio calculations) to the study of molecular interactions: the origin of molecular interactions. In *Molecular Interactions*, Vol. 1; Ratajczak, H., Orville-Thomas, W. J., Eds.; John Wiley and Sons: Chichester, U.K., 1980; p 21.
(88) Morokuma, K.; Kitaura, K. *Chemical Applications of Atomic and Molecular Electrostatic Potentials*; Politzer, P., Truhlar, D. G., Eds.; Plenum: New York, 1981; p 215.
(89) Stevens, W. J.; Fink, W. H. *Chem. Phys. Lett.* **1987**, *139*, 15.
(90) Bagus, P. S.; Hermann, K.; Bauschlicher, J. C. W. *J. Chem. Phys.* **1984**, *80*, 4378.
(91) Bagus, P. S.; Illas, F. *J. Chem. Phys.* **1992**, *96*, 8962.
(92) Glendening, E. D.; Streitwieser, A. *J. Chem. Phys.* **1994**, *100*, 2900.
(93) Schenter, G. K.; Glendening, E. D. *J. Phys. Chem.* **1996**, *100*, 17152.
(94) Glendening, E. D. *J. Phys. Chem. A* **2005**, *109*, 11936.
(95) Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. *J. Phys. Chem. A* **2007**, *111*, 8753.
(96) Mo, Y. R.; Gao, J. L.; Peyerimhoff, S. D. *J. Chem. Phys.* **2000**, *112*, 5530.
(97) van der Vaart, A.; Merz, K. M. *J. Phys. Chem. A* **1999**, *103*, 3321.
(98) Fedorov, D. G.; Kitaura, K. *J. Comput. Chem.* **2007**, *28*, 222.
(99) Su, P. F.; Li, H. *J. Chem. Phys.* **2009**, *131*, 014102.
(100) Hayes, I. C.; Stone, A. J. *Mol. Phys.* **1984**, *53*, 83.
(101) Wu, Q.; Ayers, P. W.; Zhang, Y. K. *J. Chem. Phys.* **2009**, *131*, 164112.
(102) Wu, Q.; Yang, W. *J. Chem. Phys.* **2003**, *118*, 2498.
(103) Wu, Q.; Van Voorhis, T. *Phys. Rev. A* **2005**, *72*, 024502.
(104) Curutchet, C.; Bofill, J. M.; Hernández, B.; Orozco, M.; Luque, F. J. *J. Comput. Chem.* **2003**, *24*, 1263.
(105) Cisneros, G. A.; Darden, T. A.; Gresh, N.; Pilmé, J.; Reinhardt, P.; Parisel, O.; Piquemal, J. P.,Design Of Next Generation Force Fields From AB Initio Computations: Beyond Point Charges Electrostatics. In *Multi-scale Quantum Models for Biocatalysis*; York, D. M., Lee, T.-S., Eds.; Springer: Dordrecth, the Netherlands, 2009; p 137.
(106) Piquemal, J. P.; Cisneros, G. A.; Reinhardt, P.; Gresh, N.; Darden, T. A. *J. Chem. Phys.* **2006**, *124*, 104101.
(107) Piquemal, J. P.; Perera, L.; Cisneros, G. A.; Ren, P. Y.; Pedersen, L. G.; Darden, T. A. *J. Chem. Phys.* **2006**, *125*, 054511.
(108) Halgren, T. A.; Damm, W. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236.
(109) Gianinetti, E.; Raimondi, M.; Tornaghi, E. *Int. J. Quantum Chem.* **1996**, *60*, 157.
(110) Gianinetti, E.; Vandoni, I.; Famulari, A.; Raimondi, M. Extension of the SCF-MI Method to the Case of K Fragments one of which is an Open-Shell System. In *Advanced Quantum Chemistry*; Academic Press: 1998; p 251.
(111) Nagata, T.; Takahashi, O.; Saito, K.; Iwata, S. *J. Chem. Phys.* **2001**, *115*, 3553.
(112) Nagata, T.; Iwata, S. *J. Chem. Phys.* **2004**, *120*, 3555.
(113) Khaliullin, R. Z.; Head-Gordon, M.; Bell, A. T. *J. Chem. Phys.* **2006**, *124*, 204105.
(114) Nemukhin, F. Y.; Knowles, P. J.; Murrell, J. N. *Chem. Phys.* **1995**, *193*, 27.
(115) Jensen, J. H.; Gordon, M. S. *Mol. Phys.* **1996**, *89*, 1313.
(116) Huxley, P.; Knowles, D. B.; Murrell, J. N.; Watts, J. D. *J. Chem. Soc., Faraday Trans. 2* **1984**, *80*, 1349.
(117) Brobjera, J. T.; Murrell, J. N. *Chem. Phys. Lett.* **1981**, *77*, 601.
(118) Benson, S. W.; Cruickshank, F. R.; Golden, D. M.; Haugen, G. R.; O'Neal, H. E.; Rodgers, A. S.; Shaw, R.; Walsh, R. *Chem. Rev.* **1969**, *69*, 279.
(119) Zhang, D. W.; Xiang, Y.; Zhang, J. Z. H. *J. Phys. Chem. B* **2003**, *107*, 12039.
(120) Li, W.; Li, S. H.; Jiang, Y. S. *J. Phys. Chem. A* **2007**, *111*, 2193.
(121) Babu, K.; Gadre, S. R. *J. Comput. Chem.* **2003**, *24*, 484.
(122) Collins, M. A.; Deev, V. *J. Chem. Phys.* **2006**, *125*, 104104.
(123) Deev, V.; Collins, M. A. *J. Chem. Phys.* **2005**, *122*, 154102.
(124) Olsen, J.; Roos, B. O.; Jøgensen, P.; Jensen, H. J. Aa *J. Chem. Phys.* **1988**, *89*, 2185.
(125) Nakano, H.; Hirao, K. *Chem. Phys. Lett.* **2000**, *317*, 90.
(126) Ivanic, J. *J. Chem. Phys.* **2003**, *119*, 9364.

(127) Sauri, V.; Serrano-Andres, L.; Shahi, A. R. M.; Gagliardi, L.; Vancoilie, S.; Pierloot, K. *J. Chem. Theory Comput.* **2010**, *7*, 153.
(128) Roskop, L.; Gordon, M. S. *J. Chem. Phys.* **2011**, *135*, 044101.
(129) Werner, H.-J.; Pfluger, K. *Annu. Rep. Comput. Chem.* **2006**, *2*, 53.
(130) Pulay, P. *Chem. Phys. Lett.* **1983**, *100*, 151.
(131) Saebo, S.; Pulay, P. *Chem. Phys. Lett.* **1985**, *113*, 13.
(132) Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2009**, *130*, 241101 and references cited therein.
(133) Subotnik, J. E.; Sodt, A.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 034103.
(134) Venkatnathan, A.; Szilva, A. B.; Walter, D.; Gdanitz, R. J.; Carter, E. A. *J. Chem. Phys.* **2004**, *120*, 1693.
(135) Li, W.; Piecuch, P. *J. Phys. Chem. A* **2010**, *114*, 6721.
(136) Schütz, M.; Werner, H.-J. *J. Chem. Phys.* **2000**, *114*, 661.
(137) Paulus, B. *Phys. Rep.* **2006**, *428*, 1.
(138) Friedrich, J.; Hanrath, M.; Dolg, M. *J. Phys. Chem. A* **2007**, *111*, 9830.
(139) Friedrich, J.; Hanrath, M.; Dolg, M. *Chem. Phys.* **2008**, *346*, 266.
(140) Mata, R. A.; Stoll, H. *Chem. Phys. Lett.* **2008**, *465*, 136.
(141) Manby, F. R.; Alfè, D.; Gillan, M. J. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5178.
(142) Nolan, S. J.; Bygrave, P. J.; Allan, N. L.; Manby, F. R. *J. Phys.: Condens. Matter* **2010**, *22*, 074201.
(143) Goedecker, S. *Rev. Mod. Phys.* **1999**, *71*, 1085.
(144) Zalesny, R., Papadopoulos, M. G., Mezey, P. G., Leszczynski, J., Eds. *Linear-Scaling Techniques in Computational Chemistry and Physics*; Springer: Berlin, 2011.
(145) Reimers, J. R., Ed. *Computational Methods for Large Systems: Electronic Structure Approaches for Biotechnology and Nanotechnology*; Wiley: New York, 2011.
(146) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170.
(147) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357.
(148) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959.
(149) Elsohly, A. M.; Shaw, C. L.; Guice, M. E.; Smith, B. D.; Tschumper, G. S. *Mol. Phys.* **2007**, *105*, 2777.
(150) Li, H.; Li, W.; Li, S. H.; Ma, J. *J. Phys. Chem. B* **2008**, *112*, 7061.
(151) He, X.; Wang, B.; Merz, K. M. *J. Phys. Chem. B* **2009**, *113*, 10380.
(152) Guo, W.; Wu, A.; Xu, X. *Chem. Phys. Lett.* **2010**, *498*, 203.
(153) Morita, S.; Sakai, S. *J. Comput. Chem.* **2001**, *22*, 1107.
(154) Thole, B. T. *Chem. Phys.* **1981**, *59*, 341.
(155) Piquemal, J.-P.; Perera, L.; Cisneros, G. A.; Ren, P.; Pedersen, L. G.; Darden, T. A. *J. Chem. Phys.* **2006**, *125*, 054511.
(156) Wu., J.; Piquemal, J.-P.; Chaudret, R.; Reinhardt, P.; Ren, P. *J. Chem. Theory Comput.* **2010**, *6*, 2059.
(157) Åstrand, P. O.; Linse, P.; Karlström, G. *Chem. Phys.* **1995**, *191*, 195.
(158) Swart, M.; van Duijnen, P. T. *Mol. Sim.* **2006**, *32*, 471.
(159) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(160) Gresh, N.; Claverie, P.; Pullman, A. *Theor. Chim. Acta* **1984**, *66*, 1.
(161) Gresh, N.; Claverie, P.; Pullman, A. *Int. J. Quantum Chem.* **1986**, *29*, 101.
(162) Gresh, N. *J. Comput. Chem.* **1995**, *16*, 856.
(163) Gresh, N.; Guo, H.; Kafafi, S. A.; Salahub, D. R.; Roques, B. P. *J. Am. Chem. Soc.* **1999**, *121*, 7885.
(164) Gresh, N.; Piquemal, J. P.; Krauss, M. *J. Comput. Chem.* **2005**, *26*, 1113.
(165) Vigné-Maeder, F.; Claverie, P. *J. Chem. Phys.* **1988**, *88*, 4934.
(166) Piquemal, J. P.; Gresh, N.; Giessner-Prettre, C. *J. Phys. Chem. A* **2003**, *107*, 10353.
(167) Piquemal, J. P.; Chevreau, H.; Gresh, N. *J. Chem. Theory Comput.* **2007**, *3*, 824.
(168) Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A. *J. Chem. Phys.* **2006**, *125*, 184101.
(169) Creuzet, S.; Langlet, J.; Gresh, N. *J. Chim. Phys. Phys.-Chim. Biol.* **1991**, *88*, 2399.
(170) Murrell, J. N.; Randic, M.; Williams, D. R. *Proc. R. Soc. London, Ser. A* **1966**, *284*, 566.
(171) Chaudret, R.; Ulmer, S.; van Severen, M. C.; Gresh, N.; Parisel, O.; Cisneros, G. A.; Darden, T. A.; Piquemal, J. P. *Theory. Appl. Comput. Chem.: 2008* **2009**, *1102*, 185.
(172) Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. *J. Chem. Phys.* **1996**, *105*, 1968.
(173) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *J. Phys. Chem. A* **2001**, *105*, 293.
(174) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233.
(175) Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: Oxford, 1996.
(176) Arora, P.; Slipchenko, L. V.; Webb, S. P.; Defusco, A.; Gordon, M. S. *J. Phys. Chem. A* **2010**, *114*, 6742.
(177) Yoo, S.; Zahariev, F.; Sok, S.; Gordon, M. S. *J. Chem. Phys.* **2008**, *129*, 144112.
(178) Defusco, A.; Gordon, M. S. Unpublished work.
(179) Kosenkov, D.; Slipchenko, L. V. *J. Phys. Chem. A* **2010**, *115*, 392.
(180) Slipchenko, L. V. *J. Phys. Chem. A* **2010**, *114*, 8824.
(181) Ghosh, D.; Isayev, O.; Slipchenko, L. V.; Krylov, A. I. *J. Phys. Chem. A* **2011**, *115*, 6028.
(182) Nagata, T.; Fedorov, D. G.; Kitaura, K.; Gordon, M. S. *J. Chem. Phys.* **2009**, *131*, 024101.
(183) Nagata, T.; Fedorov, D. G.; Sawada, T.; Kitaura, K.; Gordon, M. S. *J. Chem. Phys.* **2011**, *134*, 034110.
(184) Li, H.; Pomelli, C. S.; Jensen, J. H. *Theo. Chem. Acc.* **2003**, *109*, 71.
(185) Bandyopadhyay, P.; Gordon, M. S.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **2002**, *116*, 5023.
(186) Li, H.; Gordon, M. S. *J. Chem. Phys.* **2007**, *126*, 124112.
(187) Li, H. *J. Chem. Phys.* **2009**, *131*, 184103.
(188) Zorn, D.; Lin, V. S.-Y.; Pruski, M.; Gordon, M. S. *J. Phys. Chem. B* **2008**, *112*, 12753.
(189) Netzloff, H. M.; Gordon, M. S. *J. Comput. Chem.* **2004**, *25*, 1926.
(190) Day, P. N.; Pachter, R.; Gordon, M. S.; Merrill, G. N. *J. Chem. Phys.* **2000**, *112*, 2063.
(191) Chen, W.; Gordon, M. S. *J. Chem. Phys.* **1996**, *105*, 11081.
(192) Merrill, G. N.; Gordon, M. S. *J. Phys. Chem. A* **1998**, *102*, 2650.
(193) Netzloff, H. M.; Gordon, M. S. *J. Chem. Phys.* **2004**, *121*, 2711.
(194) Webb, S. P.; Gordon, M. S. *J. Phys. Chem. A* **1999**, *103*, 1265.
(195) Adamovic, I.; Gordon, M. S. *J. Phys. Chem. A* **2005**, *109*, 1629.
(196) Bandyopadhyay, P.; Gordon, M. S. *J. Chem. Phys.* **2000**, *113*, 1104.
(197) Adamovic, I.; Freitag, M. A.; Gordon, M. S. *J. Chem. Phys.* **2003**, *118*, 6725.
(198) Song, J.; Gordon, M. S. Unpublished work.
(199) Gordon, M. S.; Mullin, J. M.; Pruitt, S. R.; Roskop, L. B.; Slipchenko, L. V.; Boatz, J. A. *J. Phys. Chem. B* **2009**, *113*, 9646.
(200) Ghosh, D.; Kosenkov, D.; Vanovschi, V.; Williams, C. F.; Herbert, J. M.; Gordon, M. S.; Schmidt, M. W.; Slipchenko, L. V.; Krylov, A. I. *J. Phys. Chem. A* **2010**, *114*, 12739.
(201) Freitag, M. A.; Gordon, M. S.; Jensen, J. H.; Stevens, W. J. *J. Chem. Phys.* **2000**, *112*, 7300.
(202) Slipchenko, L. V.; Gordon, M. S. *J. Comput. Chem.* **2007**, *28*, 276.
(203) Slipchenko, L. V.; Gordon, M. S. *Mol. Phys.* **2009**, *107*, 999.
(204) Kairys, V.; Jensen, J. H. *Chem. Phys. Lett.* **1999**, *315*, 140.
(205) Adamovic, I.; Gordon, M. S. *Mol. Phys.* **2005**, *103*, 379.
(206) Tang, K. T.; Toennies, J. P. *J. Chem. Phys.* **1984**, *80*, 3726.
(207) Jensen, J. H. *J. Chem. Phys.* **1996**, *104*, 7795.
(208) Jensen, J. H.; Gordon, M. S. *J. Chem. Phys.* **1998**, *108*, 4772.

(209) Jensen, J. H. *J. Chem. Phys.* **2001**, *114*, 8775.
(210) Li, H.; Gordon, M. S.; Jensen, J. H. *J. Chem. Phys.* **2006**, *124*, 214108.
(211) Li, H.; Netzloff, H. M.; Gordon, M. S. *J. Chem. Phys.* **2006**, *125*, 194103.
(212) Li, H.; Gordon, M. S. *Theor. Chem. Acc.* **2006**, *115*, 385.
(213) Bandyopadhyay, P. *J. Chem. Phys.* **2005**, *122*, 091102.
(214) Kemp, D.; Rintelman, J.; Gordon, M. S.; Jensen, J. H. *Theor. Chem. Acc.* **2010**, *125*, 481.
(215) Kairys, V.; Jensen, J. H. *J. Phys. Chem. A* **2000**, *104*, 6656.
(216) Assfeld, X.; Rivail, J.-L. *Chem. Phys. Lett.* **1996**, *263*, 100.
(217) Nemukhin, A. V.; Grigorenko, B. L.; Topol, I. A.; Burt, S. K. *J. Comput. Chem.* **2003**, *24*, 1410.
(218) Grigorenko, B. L.; Nemukhin, A. V.; Topol, I. A.; Burt, S. K. *J. Phys. Chem. A* **2002**, *106*, 10663.
(219) Xie, W.; Gao, J. *J. Chem. Theory Comput.* **2007**, *3*, 1890.
(220) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. *J. Phys. Chem. A* **1998**, *102*, 4714.
(221) *Pure Appl. Chem.* **1974**, *40*, 291.
(222) Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. *J. Chem. Phys.* **2008**, *128*, 234108.
(223) Song, L.; Han, J.; Lin, Y.; Xie, W.; Gao, J. *J. Phys. Chem. A* **2009**, *113*, 11656.
(224) Cembran, A.; Bao, P.; Wang, Y.; Song, L.; Truhlar, D. G.; Gao, J. *J. Chem. Theory Comput.* **2010**, *6*, 2469.
(225) Matsunaga, N.; Chaban, G. M.; Gerber, R. B. *J. Chem. Phys.* **2002**, *117*, 3541.
(226) Fedorov, D. G.; Kitaura, K. Theoretical Development of the Fragment Molecular Orbital (FMO) Method. In *Modern Methods for Theoretical Physical Chemistry of Biopolymers*; Starikov, E. B., Lewis, J. P., Tanaka, S., Eds.; Elsevier: Amsterdam, 2006; pp 3−38.
(227) Nakano, T.; Mochizuki, Y.; Fukuzawa, K.; Amari, S.; Tanaka, S. Developments and Applications of ABINIT-MP Software Based on the Fragment Molecular Orbital Method. In *Modern Methods for Theoretical Physical Chemistry of Biopolymers*; Starikov, E. B., Lewis, J. P., Tanaka, S., Eds.; Elsevier: Amsterdam, 2006, p 39.
(228) Fedorov, D. G.; Kitaura, K. *J. Phys. Chem. A* **2007**, *111*, 6904.
(229) Fedorov, D. G., Kitaura, K., Eds. *The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems*; CRC Press: Boca Raton, FL, 2009.
(230) Massa, L. *Int. J. Quantum Chem.* **2011**, *111*, 3251.
(231) Ikegami, T.; Ishida, T.; Fedorov, D. G.; Kitaura, K.; Inadomi, Y.; Umeda, H.; Yokokawa, M.; Sekiguchi, S. *Proceedimgs of Supercomputing 2005*; IEEE Computer Society: Seattle, WA, 2005.
(232) Fujita, T.; Fukuzawa, K.; Mochizuki, Y.; Nakano, T.; Tanaka, S. *Chem. Phys. Lett.* **2009**, *478*, 295.
(233) Nakano, T.; Kaminuma, T.; Sato, T.; Akiyama, Y.; Uebayasi, M.; Kitaura, K. *Chem. Phys. Lett.* **2000**, *318*, 614.
(234) Fedorov, D. G.; Jensen, J. H.; Deka, R. C.; Kitaura, K. *J. Phys. Chem. A* **2008**, *112*, 11808.
(235) Klobukowski, M.; Huzinaga, S.; Sakai, Y. Model core potentials: Theory and applications, in *Computational Chemistry: Reviews of Current Trends*; Leszczynski, J., Ed.; World Scientific: Singapore, 1999; Vol. 3.
(236) Ishikawa, T.; Mochizuki, Y.; Imamura, K.; Nakano, T.; Mori, H.; Tokiwa, H.; Tanaka, K.; Miyoshi, E.; Tanaka, S. *Chem. Phys. Lett.* **2006**, *430*, 361.
(237) Yasuda, K.; Yamaki, D. *J. Chem. Phys.* **2006**, *125*, 154101.
(238) Fedorov, D. G.; Kitaura, K. *J. Chem. Phys.* **2009**, *131*, 171106.
(239) Fedorov, D. G.; Slipchenko, L. V.; Kitaura, K. *J. Phys. Chem. A* **2010**, *114*, 8742.
(240) Nakano, T.; Kaminuma, T.; Sato, T.; Fukuzawa, K.; Akiyama, Y.; Uebayasi, M.; Kitaura, K. *Chem. Phys. Lett.* **2002**, *351*, 475.
(241) Fedorov, D. G.; Kitaura, K. *Chem. Phys. Lett.* **2006**, *433*, 182.
(242) Sekino, H.; Sengoku, Y.; Sugiki, S.-I.; Kurita, N. *Chem. Phys. Lett.* **2003**, *378*, 589.
(243) Ruedenberg, K. *J. Chem. Phys.* **1951**, *19*, 1433.
(244) Companion, A. L.; Parr, R. G. *J. Chem. Phys.* **1961**, *35*, 2268.
(245) Kitaura, K.; Sugiki, S.-I.; Nakano, T.; Komeiji, Y.; Uebayasi, M. *Chem. Phys. Lett.* **2001**, *336*, 163.
(246) Nagata, T.; Fedorov, D. G.; Kitaura, K. *Chem. Phys. Lett.* **2009**, *475*, 124.
(247) Nagata, T.; Fedorov, D. G.; Kitaura, K. *Chem. Phys. Lett.* **2010**, *492*, 302.
(248) Komeiji, Y.; Mochizuki, Y.; Nakano, T. *Chem. Phys. Lett.* **2010**, *484*, 380.
(249) Fedorov, D. G.; Avramov, P. V.; Jensen, J. H.; Kitaura, K. *Chem. Phys. Lett.* **2009**, *477*, 169.
(250) Nagata, T.; Brorsen, K.; Fedorov, D. G.; Kitaura, K.; Gordon, M. S. *J. Chem. Phys.* **2011**, *134*, 124115.
(251) Inadomi, Y.; Nakano, T.; Kitaura, K.; Nagashima, U. *Chem. Phys. Lett.* **2002**, *364*, 139.
(252) Watanabe, T.; Inadomi, Y.; Umeda, H.; Fukuzawa, K.; Tanaka, S.; Nakano, T.; Nagashima, U. *J. Comput. Theor. Nanosci.* **2009**, *6*, 1328.
(253) Umeda, H.; Inadomi, Y.; Watanabe, T.; Yagi, T.; Ishimoto, T.; Ikegami, T.; Tadano, H.; Sakurai, T.; Nagashima, U. *J. Comput. Chem.* **2010**, *31*, 2381.
(254) Tsuneyuki, S.; Kobori, T.; Akagi, K.; Sodeyama, K.; Terakura, K.; Fukuyama, H. *Chem. Phys. Lett.* **2009**, *476*, 104.
(255) Nishioka, H.; Ando, K. *J. Chem. Phys.* **2011**, *134*, 204109.
(256) Fedorov, D. G.; Kitaura, K. *Chem. Phys. Lett.* **2004**, *389*, 129.
(257) Shimodo, Y.; Morihashi, K.; Nakano, T. *J. Mol. Struct. (THEOCHEM)* **2006**, *770*, 163.
(258) Fedorov, D. G.; Kitaura, K. *J. Chem. Phys.* **2004**, *121*, 2483.
(259) Mochizuki, Y.; Koikegami, S.; Nakano, T.; Amari, S.; Kitaura, K. *Chem. Phys. Lett.* **2004**, *396*, 473.
(260) Mochizuki, Y.; Nakano, T.; Koikegami, S.; Tanimori, S.; Abe, Y.; Nagashima, U.; Kitaura, K. *Theor. Chem. Acc.* **2004**, *112*, 442.
(261) Ishikawa, T.; Kuwata, K. *Chem. Phys. Lett.* **2009**, *474*, 195.
(262) Okiyama, Y.; Nakano, T.; Yamashita, K.; Mochizuki, Y.; Taguchi, N.; Tanaka, S. *Chem. Phys. Lett.* **2010**, *490*, 84.
(263) Mochizuki, Y.; Yamashita, K.; Murase, T.; Nakano, T.; Fukuzawa, K.; Takematsu, K.; Watanabe, H.; Tanaka, S. *Chem. Phys. Lett.* **2008**, *457*, 396.
(264) Mochizuki, Y.; Yamashita, K.; Fukuzawa, K.; Takematsu, K.; Watanabe, H.; Taguchi, N.; Okiyama, Y.; Tsuboi, M.; Nakano, T.; Tanaka, S. *Chem. Phys. Lett.* **2010**, *493*, 346.
(265) Fedorov, D. G.; Ishimura, K.; Ishida, T.; Kitaura, K.; Pulay, P.; Nagase, S. *J. Comput. Chem.* **2007**, *28*, 1476.
(266) Mochizuki, Y.; Nakano, T.; Komeiji, Y.; Yamashita, K.; Okiyama, Y.; Yoshikawa, H.; Yamataka, H. *Chem. Phys. Lett.* **2011**, *504*, 95.
(267) Komeiji, Y.; Nakano, T.; Fukuzawa, K.; Ueno, Y.; Inadomi, Y.; Nemoto, T.; Uebayasi, M.; Fedorov, D. G.; Kitaura, K. *Chem. Phys. Lett.* **2003**, *372*, 342.
(268) Ishimoto, T.; Tokiwa, H.; Teramae, H.; Nagashima, U. *Chem. Phys. Lett.* **2004**, *387*, 460.
(269) Ishimoto, T.; Tokiwa, H.; Teramae, H.; Nagashima, U. *J. Chem. Phys.* **2005**, *122*, 094905.
(270) Komeiji, Y.; Ishikawa, T.; Mochizuki, Y.; Yamataka, H.; Nakano, T. *J. Comput. Chem.* **2009**, *30*, 40.
(271) Komeiji, Y.; Mochizuki, Y.; Nakano, T.; Fedorov, D. G. *J. Mol. Struct. (THEOCHEM)* **2009**, *898*, 2.
(272) Fujita, T.; Watanabe, H.; Tanaka, S. *J. Phys. Soc. Jpn.* **2009**, *78*, 104723.
(273) Fujita, T.; Nakano, T.; Tanaka, S. *Chem. Phys. Lett.* **2011**, *506*, 112.
(274) Fedorov, D. G.; Kitaura, K. *J. Chem. Phys.* **2005**, *122*, 054108.
(275) Fedorov, D. G.; Kitaura, K. *J. Chem. Phys.* **2005**, *123*, 134103.
(276) Fedorov, D. G.; Ishida, T.; Kitaura, K. *J. Phys. Chem. A* **2005**, *109*, 2638.
(277) Maezono, R.; Watanabe, H.; Tanaka, S. Ab initio biomolecular calculations using quantum Monte Carlo combined with the fragment molecular orbital method. In *Advances in Quantum Monte Carlo*; Anderson, J. B., Rothstein, S. M., Eds.; ACS Symposium Series 953; American Chemical Society: Washington, DC, 2006; pp 141−146.

(278) Maezono, R.; Watanabe, H.; Tanaka, S.; Towler, M. D.; Needs, R. J. *J. Phys. Soc. Jpn.* **2007**, *76*, 064301.
(279) Fedorov, D. G.; Kitaura, K.; Li, H.; Jensen, J. H.; Gordon, M. S. *J. Comput. Chem.* **2006**, *27*, 976.
(280) Li, H.; Fedorov, D. G.; Nagata, T.; Kitaura, K.; Jensen, J. H.; Gordon, M. S. *J. Comput. Chem.* **2010**, *31*, 778.
(281) Watanabe, H.; Okiyama, Y.; Nakano, T.; Tanaka, S. *Chem. Phys. Lett.* **2010**, *500*, 116.
(282) Mochizuki, Y.; Koikegami, S.; Amari, S.; Segawa, K.; Kitaura, K.; Nakano, T. *Chem. Phys. Lett.* **2005**, *406*, 283.
(283) Mochizuki, Y.; Tanaka, K.; Yamashita, K.; Ishikawa, T.; Nakano, T.; Amari, S.; Segawa, K.; Murase, T.; Tokiwa, H.; Sakurai, M. *Theor. Chem. Acc.* **2007**, *117*, 541.
(284) Chiba, M.; Fedorov, D. G.; Kitaura, K. *Chem. Phys. Lett.* **2007**, *444*, 346.
(285) Chiba, M.; Fedorov, D. G.; Kitaura, K. *J. Chem. Phys.* **2007**, *127*, 104108.
(286) Chiba, M.; Fedorov, D. G.; Kitaura, K. *J. Comput. Chem.* **2008**, *29*, 2667.
(287) Chiba, M.; Fedorov, D. G.; Nagata, T.; Kitaura, K. *Chem. Phys. Lett.* **2009**, *474*, 227.
(288) Chiba, M.; Koido, T. *J. Chem. Phys.* **2010**, *133*, 044113.
(289) Pruitt, S. R.; Fedorov, D. G.; Kitaura, K.; Gordon, M. S. *J. Chem. Theory Comput.* **2010**, *6*, 1.
(290) Ishimoto, T.; Tachikawa, M.; Nagashima, U. *J. Chem. Phys.* **2006**, *124*, 014112.
(291) Auer, B.; Pak, M. V.; Hammes-Schiffer, S. *J. Phys. Chem. C* **2010**, *114*, 5582.
(292) Mochizuki, Y.; Ishikawa, T.; Tanaka, K.; Tokiwa, H.; Nakano, T.; Tanaka, S. *Chem. Phys. Lett.* **2006**, *418*, 418.
(293) Ishikawa, T.; Mochizuki, Y.; Nakano, T.; Amari, S.; Mori, H.; Honda, H.; Fujita, T.; Tokiwa, H.; Tanaka, S.; Komeiji, Y.; Fukuzawa, K.; Tanaka, K.; Miyoshi, E. *Chem. Phys. Lett.* **2006**, *427*, 159.
(294) Okiyama, Y.; Watanabe, H.; Fukuzawa, K.; Nakano, T.; Mochizuki, Y.; Ishikawa, T.; Tanaka, S.; Ebina, K. *Chem. Phys. Lett.* **2007**, *449*, 329.
(295) Okiyama, Y.; Watanabe, H.; Fukuzawa, K.; Nakano, T.; Mochizuki, Y.; Ishikawa, T.; Ebina, K.; Tanaka, S. *Chem. Phys. Lett.* **2009**, *467*, 417.
(296) Sekino, H.; Matsumura, N.; Sengoku, Y. *Comput. Lett.* **2007**, *3*, 423.
(297) Gao, Q.; Yokojima, S.; Kohno, T.; Ishida, T.; Fedorov, D. G.; Kitaura, K.; Fujihira, M.; Nakamura, S. *Chem. Phys. Lett.* **2007**, *445*, 331.
(298) Gao, Q.; Yokojima, S.; Fedorov, D. G.; Kitaura, K.; Sakurai, M.; Nakamura, S. *J. Chem. Theory Comput.* **2010**, *6*, 1428.
(299) Amari, S.; Aizawa, M.; Zhang, J.; Fukuzawa, K.; Mochizuki, Y.; Iwasawa, Y.; Nakata, K.; Chuman, H.; Nakano, T. *J. Chem. Inf. Comput. Sci.* **2006**, *46*, 221.
(300) Du, S.; Sakurai, M. *Chem. Phys. Lett.* **2010**, *488*, 81.
(301) Mochizuki, Y.; Fukuzawa, K.; Kato, A.; Tanaka, S.; Kitaura, K.; Nakano, T. *Chem. Phys. Lett.* **2005**, *410*, 247.
(302) Ishikawa, T.; Mochizuki, Y.; Amari, S.; Nakano, T.; Tokiwa, H.; Tanaka, S.; Tanaka, K. *Theor. Chem. Acc.* **2007**, *118*, 937.
(303) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 533.
(304) Ishikawa, T.; Ishikura, T.; Kuwata, K. *J. Comput. Chem.* **2009**, *30*, 2594.
(305) Okiyama, Y.; Fukuzawa, K.; Yamada, H.; Mochizuki, Y.; Nakano, Y.; Tanaka, S. *Chem. Phys. Lett.* **2011**, *509*, 67.
(306) Schwenke, D. W.; Truhlar, D. G. *J. Chem. Phys.* **1985**, *82*, 2418.
(307) Fedorov, D. G.; Ishida, T.; Uebayasi, M.; Kitaura, K. *J. Phys. Chem. A* **2007**, *111*, 2722.
(308) Ishikawa, T.; Yamamoto, N.; Kuwata, K. *Chem. Phys. Lett.* **2010**, *500*, 149.
(309) Fedorov, D. G.; Alexeev, Y.; Kitaura, K. *J. Phys. Chem. Lett.* **2011**, *2*, 282.
(310) Dahlke, E. E.; Truhlar, D. G. *J. Chem. Theor. Comput.* **2008**, *4*, 1.
(311) Dahlke, E. E.; Leverentz, H. R.; Truhlar, D. G. *J. Chem. Theor. Comput.* **2008**, *4*, 33.
(312) Sorkin, A.; Dahlke, E. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 683.
(313) Leverentz, H. R.; Truhlar, D. G. *J. Chem. Theory Comput.* **2009**, *5*, 1573.
(314) Hirata, S.; Valiev, M.; Dupuis, M.; Xantheas, S. S.; Sugiki, S.; Sekino, H. *Mol. Phys.* **2005**, *103*, 2255.
(315) Kamiya, M.; Hirata, S.; Valiev, M. J. *Chem. Phys.* **2008**, *128*, 074103.
(316) Steinmann, C.; Fedorov, D. G.; Jensen, J. H. *J. Phys. Chem. A* **2010**, *114*, 8705.
(317) Söderhjelm, P.; Ryde, U. *J. Phys. Chem. A* **2009**, *113*, 617.
(318) Söderhjelm, P.; Aquilante, F.; Ryde, U. *J. Phys. Chem. B* **2009**, *113*, 11085.
(319) Söderhjelm, P.; Kongsted, J.; Ryde, U. *J. Chem. Theory. Comput.* **2010**, *6*, 1726.
(320) Beran, G. J. O. *J. Chem. Phys.* **2009**, *130*, 164115.
(321) Sebetci, A.; Beran, G. J. O. *J. Chem. Theory Comput.* **2010**, *6*, 155.
(322) Babu, K.; Ganesh, V.; Gadre, S. R.; Ghermani, N. E. *Theor. Chem. Acc.* **2004**, *111*, 255.
(323) Ganesh, V.; Dongare, R. K.; Balanarayan, P.; Gadre, S. R. *J. Chem. Phys.* **2006**, *125*, 104109.
(324) Gadre, S. R.; Ganesh, V. *J. Theor. Comput. Chem.* **2006**, *5*, 835.
(325) Deshmukh, M. M.; Gadre, S. R.; Bartolotti, L. J. *J. Phys. Chem. A* **2006**, *110*, 12519.
(326) Deshmukh, M. M.; Suresh, C. H.; Gadre, S. R. *J. Phys. Chem. A* **2007**, *111*, 6472.
(327) Elango, M.; Subramanian, V.; Rahalkar, A. P.; Gadre, S. R.; Sathyamurthy, N. *J. Phys. Chem. A* **2008**, *112*, 7699.
(328) Gadre, S. R.; Jose, K. V. J.; Rahalkar, A. P. *J. Chem. Sci.* **2010**, *122*, 47.
(329) Yeole, S. D.; Gadre, S. R. *J. Chem. Phys.* **2010**, *132*, 094102.
(330) Rahalkar, A. P.; Katouda, M.; Gadre, S. R.; Nagase, S. *J. Comput. Chem.* **2010**, *31*, 2405.
(331) Rahalkar, A. P.; Ganesh, V.; Gadre, S. R. *J. Chem. Phys.* **2008**, *129*, 234101.
(332) Huang, L. L.; Massa, L.; Karle, J. *Int. J. Quantum Chem.* **2005**, *103*, 808.
(333) Huang, L. L.; Massa, L.; Karle, J. *Biochemistry* **2005**, *44*, 16747.
(334) Huang, L.; Massa, L.; Karle, J. *Proc. Natl. Acad. Sci. U.S.A* **2005**, *102*, 12690.
(335) Huang, L.; Massa, L.; Karle, J. *Proc. Natl. Acad. Sci. U.S.A* **2006**, *103*, 1233.
(336) Huang, L.; Massa, L.; Karle, J. *J. Chem. Theory Comput.* **2007**, *3*, 1337.
(337) Huang, L. L.; Massa, L.; Karle, J. *Int. J. Quantum Chem.* **2006**, *106*, 447.
(338) Huang, L.; Massa, L.; Karle, J. *Proc. Natl. Acad. Sci. U.S.A* **2008**, *105*, 1849.
(339) Huang, L.; Bohorquez, H. J.; Matta, C. F.; Massa, L. *Int. J. Quantum Chem.* **2011**, DOI: 10.1002/qua.22975.
(340) Huang, L.; Massa, L. *Int. J. Quantum Chem.* **2011**, *111*, 2180.
(341) Weiss, S. N.; Huang, L.; Massa, L. *J. Comput. Chem.* **2010**, *31*, 2889.
(342) Zhang, D. W.; Zhang, J. Z. H. *J. Chem. Phys.* **2003**, *119*, 3599.
(343) Zhang, D. W.; Chen, X. H.; Zhang, J. Z. H. *J. Comput. Chem.* **2003**, *24*, 1846.
(344) Zhang, D. W.; Xiang, Y.; Gao, A. M.; Zhang, J. Z. H. *J. Chem. Phys.* **2004**, *120*, 1145.
(345) Zhang, D. W.; Zhang, J. Z. H. *J. Theor. Comput. Chem.* **2004**, *3*, 43.
(346) Chen, X. H.; Zhang, J. Z. H. *J. Theor. Comput. Chem.* **2004**, *3*, 277.
(347) Mei, Y.; Zhang, D. W.; Zhang, J. Z. H. *J. Phys. Chem. A* **2005**, *109*, 2.
(348) Li, W.; Fang, T.; Li, S. *J. Chem. Phys.* **2006**, *124*, 154102.
(349) Mei, Y.; Wu, E. L.; Han, K. L.; Zhang, J. Z. H. *Int. J. Quantum Chem.* **2006**, *106*, 1267.

(350) Chen, X. H.; Zhang, J. Z. H. *J. Chem. Phys.* **2004**, *120*, 11386.
(351) He, X.; Zhang, J. Z. H. *J. Chem. Phys.* **2005**, *122*, 031103.
(352) Chen, X. H.; Zhang, D. W.; Zhang, J. Z. H. *J. Chem. Phys.* **2004**, *120*, 839.
(353) Gao, A. M.; Zhang, D. W.; Zhang, J. Z. H.; Zhang, Y. K. *Chem. Phys. Lett.* **2004**, *394*, 293.
(354) Chen, X. H.; Zhang, Y. K.; Zhang, J. Z. H. *J. Chem. Phys.* **2005**, *122*, 184105.
(355) Li, S. H.; Li, W.; Fang, T. *J. Am. Chem. Soc.* **2005**, *127*, 7215.
(356) Xiang, Y.; Zhang, D. W.; Zhang, J. Z. H. *J. Comput. Chem.* **2004**, *25*, 1431.
(357) He, X.; Zhang, J. Z. H. *J. Chem. Phys.* **2006**, *124*, 184703.
(358) Mei, Y.; Ji, C. G.; Zhang, J. Z. H. *J. Chem. Phys.* **2006**, *125*, 094906.
(359) Chen, X. H.; Zhang, J. Z. H. *J. Chem. Phys.* **2006**, *125*, 044903.
(360) Jiang, N.; Ma, J.; Jiang, Y. *J. Chem. Phys.* **2006**, *124*, 114112.
(361) Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211.
(362) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735.
(363) Hua, W. J.; Fang, T.; Li, W.; Yu, J. G.; Li, S. H. *J. Phys. Chem. A* **2008**, *112*, 10864.
(364) Hua, S. G.; Hua, W. J.; Li, S. H. *J. Phys. Chem. A* **2010**, *114*, 8126.
(365) Jacob, C. R.; Visscher, L. *J. Chem. Phys.* **2008**, *128*, 155102.
(366) He, J.; Di Paola, C.; Kantorovich, L. *J. Chem. Phys.* **2009**, *130*, 144104.
(367) Řezáč, J; Salahub, D. R. *J. Chem. Theory Comput.* **2010**, *6*, 91.
(368) Ji, C. G.; Mei, Y.; Zhang, J. Z. H. *Biophys. J.* **2008**, *95*, 1080.
(369) Ji, C. G.; Zhang, J. Z. H. *J. Am. Chem. Soc.* **2008**, *130*, 17129.
(370) Duan, L. L.; Mei, Y.; Zhang, Q. G.; Zhang, J. Z. H. *J. Chem. Phys.* **2009**, *130*, 115102.
(371) Ji, C. G.; Zhang, J. Z. H. *J. Phys. Chem. B* **2009**, *113*, 13898.
(372) Tong, Y.; Ji, C. G.; Mei, Y.; Zhang, J. Z. H. *J. Am. Chem. Soc.* **2009**, *131*, 8636.
(373) Ji, C. G.; Zhang, J. Z. H. *J. Phys. Chem. B* **2009**, *113*, 16059.
(374) Lu, Y.; Mei, Y.; Zhang, J. Z. H.; Zhang, D. *J. Chem. Phys.* **2010**, *132*, 131101.
(375) Tong, Y.; Mei, Y.; Li, Y. L.; Ji, C. G.; Zhang John, Z. H. *J. Am. Chem. Soc.* **2010**, *132*, 5137.
(376) Duan, L. L.; Mei, Y.; Zhang, D.; Zhang, Q. G.; Zhang, J. Z. H. *J. Am. Chem. Soc.* **2010**, *132*, 11159.
(377) Netzloff, H. M.; Collins, M. A. *J. Chem. Phys.* **2007**, *127*, 134113.
(378) Mullin, J. M.; Roskop, L. B.; Pruitt, S. R.; Collins, M. A.; Gordon, M. S. *J. Phys. Chem. A* **2009**, *113*, 10040.
(379) Collins, M. A. *J. Chem. Phys.* **2007**, *127*, 024104.
(380) Xantheas, S. S.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 8037.
(381) Xantheas, S. S. *J. Chem. Phys.* **1994**, *100*, 7523.
(382) Hodges, M. P.; Stone, A. J.; Xantheas, S. S. *J. Phys. Chem. A* **1997**, *101*, 9163.
(383) Xantheas, S. S. *Chem. Phys.* **2000**, *258*, 225.
(384) Xantheas, S. S. *Struct. Bonding (Berlin)* **2005**, *116*, 119.
(385) Yang, W.; Lee, T. S. *J. Chem. Phys.* **1995**, *103*, 5674.
(386) Lee, T. S.; York, D. M.; Yang, W. *J. Chem. Phys.* **1996**, *105*, 2744.
(387) Li, W.; Li, S. *J. Chem. Phys.* **2005**, *122*, 194109.
(388) He, X.; Merz, K. M., Jr. *J. Chem. Theory Comput* **2010**, *6*, 405.
(389) Dixon, S. L.; Merz., K. M. *J. Chem. Phys.* **1996**, *104*, 6643.
(390) Dixon, S. L.; Merz, K. M. *J. Chem. Phys.* **1997**, *107*, 879.
(391) Li, W.; Li, S. *J. Chem. Phys.* **2004**, *121*, 6649.
(392) Kobayashi, M.; Akama, T.; Nakai, H. *J. Chem. Phys.* **2006**, *125*, 204106.
(393) Akama, T.; Fujii, A.; Kobayashi, M.; Nakai, H. *Mol. Phys.* **2007**, *105*, 2799.
(394) Akama, T.; Kobayashi, M.; Nakai, H. *J. Comput. Chem.* **2007**, *28*, 2003.
(395) Kobayashi, M.; Imamura, Y.; Nakai, H. *J. Chem. Phys.* **2007**, *127*, 074103.
(396) Kobayashi, M.; Nakai, H. *J. Chem. Phys.* **2008**, *129*, 044103.
(397) Akama, T.; Kobayashi, M.; Nakai, H. *Int. J. Quantum Chem.* **2009**, *109*, 2706.
(398) Kobayashi, M.; Nakai, H. *J. Chem. Phys.* **2009**, *131*, 114108.
(399) Kobayashi, M.; Nakai, H. *Int. J. Quantum Chem.* **2009**, *109*, 2227.
(400) Touma, T.; Kobayashi, M.; Nakai, H. *Chem. Phys. Lett.* **2010**, *485*, 247.
(401) Kobayashi, M.; Kunisada, T.; Akama, T.; Sakura, D.; Nakai, H. *J. Chem. Phys.* **2011**, *134*, 034105.
(402) Song, G.; Li, Z; Liu, Z.; Cao, X.; Wang, W.; Fan, K.; Xie, Y.; Schaefer, H. F., III *J. Chem. Theory Comput.* **2008**, *12*, 2049.
(403) Elliott, P.; Burke, K.; Cohen, M. H.; Wasserman, A. *Phys. Rev. A* **2010**, *82*, 024501.
(404) Fermi, E. *Rend. Lincei* **1930**, *11*, 980.
(405) Walker, P. D.; Mezey, P. G. *J. Am. Chem. Soc.* **1993**, *115*, 12423.
(406) Walker, P. D.; Mezey, P. G. *J. Am. Chem. Soc.* **1994**, *116*, 12022.
(407) Michl, J.; Kaszynski, K.; Friedli, A. C.; McMurdi, N. D.; Kim, T. *NATO ASI Ser., Ser. C* **1989**, *273*, 469.
(408) Mathias, J. P.; Stoddart, J. F. *Chem. Soc. Rev.* **1992**, 215.
(409) Mezey, P. G. *J. Math. Chem.* **1995**, *18*, 141.
(410) Mezey, P. G. *Computational Chemistry: Reviews and Current Trends*; Leszczynski, J., Ed.; World Scientific: Singapore, 1996; Vol. 1, p 109.
(411) Exner, T. E.; Mezey, P. G. *J. Phys. Chem. A* **2002**, *106*, 11791.
(412) Exner, T. E.; Mezey, P. G. *Phys. Chem. Chem. Phys.* **2005**, *7*, 4061.
(413) Szekeres, Z.; Mezey, P. G. *Mol. Phys.* **2005**, *103*, 1013.
(414) Eckard, S.; Exner, T. E. *Z. Phys. Chem.* **2006**, *220*, 927.
(415) Exner, T. E.; Mezey, P. G. *J. Comput. Chem.* **2003**, *24*, 1980.
(416) Exner, T. E.; Mezey, P. G. *J. Phys. Chem. A* **2004**, *108*, 4301.
(417) Szekeres, Z.; Exner, T.; Mezey, P. G. *Int. J. Quantum Chem.* **2005**, *104*, 847.
(418) Szekeres, Z.; Mezey, P. G.; Surjan, P. R. *Chem. Phys. Lett.* **2006**, *424*, 420.
(419) Eckard, S.; Exner, T. E. *Int. J. Quantum Chem.* **2009**, *109*, 1451.
(420) Das, G. P.; Yeates, A. T.; Dudis, D. S. *Int. J. Quantum Chem.* **2003**, *92*, 22.
(421) Sakai, S.; Morita, S. *J. Phys. Chem. A* **2005**, *109*, 8424.
(422) Mata, R. A.; Stoll, H.; Cabral, B. J. C. *J. Chem. Theory Comput.* **2009**, *5*, 1829.
(423) Bettens, R. P. A; Lee, A. M. *J. Phys. Chem. A* **2006**, *110*, 8777.
(424) Lee, A. M.; Bettens, R. P. A. *J. Phys. Chem. A* **2007**, *111*, 5111.
(425) Le, H.-A.; Lee, A. M.; Bettens, R. P. A. *J. Phys. Chem. A* **2009**, *113*, 10527.
(426) Hehre, W. J.; Ditchfield, R.; Radom, L.; Pople, J. A. *J. Am. Chem. Soc.* **1970**, *92*, 4796.
(427) Wang, L.; Zhao, Z.; Meza, J. *Phys. Rev. B* **2008**, *77*, 165113.
(428) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; et al. *J. Comput. Chem.* **1993**, *14*, 1347, http://www.msg.ameslab.gov/gamess/index.html.
(429) Gordon, M. S.; Schmidt, M. W. Advances in Electronic Structure Theory: GAMESS a Decade Later. In *Theory and Applications of Computational Chemistry, The First Forty Years*; Elsevier: Amsterdam, 2005.
(430) Komeiji, Y.; Inadomi, Y.; Nakano, T. *Comput. Biol. Chem.* **2004**, *28*, 155.
(431) RSS21 Project. http://www.ciss.iis.u-tokyo.ac.jp/rss21/result/download/index.php#download_2 (accessed August 2011) (in Japanese).
(432) PAICS. http://www.paics.net/index_e.html.
(433) Ganesh, V.; Kavathekar, R.; Rahalkar, A.; Gadre, S. R. *J. Comput. Chem.* **2008**, *29*, 488.
(434) Kavathekar, R.; Khire, S.; Ganesh, V.; Rahalkar, A. P.; Gadre, S. R. *J. Comput. Chem.* **2009**, *30*, 1167.
(435) Bode, B. M.; Gordon, M. S. *J. Mol. Graphics Mod* **1998**, *16*, 133. http://www.scl.ameslab.gov/~brett/MacMolPlt/.

(436) Suenaga, M. *J. Comput. Chem. Jpn.* **2008**, *7*, 33. http://www1.bbiq.jp/zzzfelis/Facio.html .
(437) Schmidt, M. W.; Fletcher, G. D.; Bode, B. M.; Gordon, M. S. *Comput. Phys. Commun.* **2000**, *128*, 190.
(438) Olson, R. M.; Schmidt, M. W.; Gordon, M. S.; Rendell, A. P. *Proc. Supercomput.* **2003**.
(439) Sato, M.; Yamataka, H.; Komeiji, Y.; Mochizuki, Y.; Ishikawa, T.; Nakano, T. *J. Am. Chem. Soc.* **2008**, *130*, 2396.
(440) Sato, M.; Yamataka, H.; Komeiji, Y.; Mochizuki, Y.; Nakano, T. *Chem.—Eur. J.* **2010**, *16*, 6430.
(441) Pomogaev, V.; Pomogaeva, A.; Aoki, Y. *J. Phys. Chem. A* **2009**, *113*, 1429.
(442) Kistler, K. A.; Matsika, S. *J. Phys. Chem. A* **2009**, *113*, 12396.
(443) Fujiwara, T.; Mochizuki, Y.; Komeiji, Y.; Okiyama, Y.; Mori, H.; Nakano, T.; Miyoshi, E. *Chem. Phys. Lett.* **2010**, *490*, 41.
(444) Bandyopadhyay, P. *Theor. Chem. Acc.* **2008**, *120*, 307.
(445) Kemp, D. A.; Gordon, M. S. *J. Phys. Chem. A* **2008**, *112*, 4885.
(446) Kemp, D. A.; Gordon, M. S. *J. Phys. Chem. A* **2005**, *109*, 7688.
(447) Merrill, G. N.; Webb, S. P. *J. Phys. Chem. A* **2003**, *107*, 7852.
(448) Merrill, G. N.; Webb, S. P. *J. Phys. Chem. A* **2004**, *108*, 833.
(449) Merrill, G. N.; Webb, S. P.; Bivin, D. B. *J. Phys. Chem. A* **2003**, *107*, 386.
(450) Chandrakumar, K. R. S.; Ghanty, T. K.; Ghosh, S. K.; Mukherjee, T. *J. Mol. Struct. (THEOCHEM)* **2007**, *807*, 93.
(451) Merrill, G. N.; Fletcher, G. D. *Theor. Chem. Acc.* **2008**, *120*, 5.
(452) Petersen, C. P.; Gordon, M. S. *J. Phys. Chem. A* **1999**, *103*, 4162.
(453) Yoshikawa, A.; Morales, J. A. *J. Mol. Struct. (THEOCHEM)* **2004**, *681*, 27.
(454) Balawender, R.; Safi, B.; Geerlings, P. *J. Phys. Chem. A* **2001**, *105*, 6703.
(455) Safi, B.; Balawender, R.; Geerlings, P. *J. Phys. Chem. A* **2001**, *105*, 11102.
(456) Day, P. N.; Pachter, R. *J. Chem. Phys.* **1997**, *107*, 2990.
(457) Mullin, J. M.; Gordon, M. S. *J. Phys. Chem. B* **2009**, *113*, 8657.
(458) Mullin, J. M.; Gordon, M. S. *J. Phys. Chem. B* **2009**, *113*, 14413.
(459) Song, J.; Gordon, M. S.; Deakyne, C. A.; Zheng, W. C. *J. Phys. Chem. A* **2004**, *108*, 11419.
(460) Adamovic, I.; Gordon, M. S. *J. Phys. Chem. A* **2006**, *110*, 10267.
(461) Adamovic, I.; Li, H.; Lamm, M. H.; Gordon, M. S. *J. Phys. Chem. A* **2006**, *110*, 519.
(462) Smith, T.; Slipchenko, L. V.; Gordon, M. S. *J. Phys. Chem. A* **2008**, *112*, 5286.
(463) Smith, Q. A.; Gordon, M. S.; Slipchenko, L. V. *J. Phys. Chem. A* **2011**, *115*, 4598.
(464) Slipchenko, L. V.; Gordon, M. S. *J. Phys. Chem. A* **2009**, *113*, 2092.
(465) Mohri, F.; Granovsky, A. A. *Int. J. Quantum Chem.* **2008**, *108*, 544.
(466) Krauss, M. *Comput. Chem.* **1995**, *19*, 33.
(467) Krauss, M. *Comput. Chem.* **1995**, *19*, 199.
(468) Kina, D.; Arora, P.; Nakayama, A.; Noro, T.; Gordon, M. S.; Taketsugu, T. *Int. J. Quantum Chem.* **2009**, *109*, 2308.
(469) Kina, D.; Nakayama, A.; Noro, T.; Taketsugu, T.; Gordon, M. S. *J. Phys. Chem. A* **2008**, *112*, 9675.
(470) Atadinc, F.; Gunaydin, H.; Ozen, A. S.; Aviyente, V. *Int. J. Chem. Kinetics* **2005**, *37*, 502.
(471) Ferreira, D. E. C.; Florentino, B. P. D.; Rocha, W. R.; Nome, F. *J. Phys. Chem. B* **2009**, *113*, 14831.
(472) Hush, N. S.; Schamberger, J.; Bacskay, G. B. *Coord. Chem. Rev.* **2005**, *249*, 299.
(473) Nemukhin, A. V.; Topol, I. A.; Grigorenko, B. L.; Burt, S. K. *J. Phys. Chem. B* **2002**, *106*, 1734.
(474) Nemukhin, A. V.; Grigorenko, B. L.; Topol, I. A.; Burt, S. K. *Phys. Chem. Chem. Phys.* **2004**, *6*, 1031.
(475) Jose, K. V. J.; Gadre, S. R. *J. Chem. Phys.* **2008**, *129*, 164314.
(476) Jose, K. V. J.; Gadre, S. R. *Int. J. Quantum Chem.* **2009**, *109*, 2238.
(477) Mahadevi, A. S.; Rahalkar, A. P.; Gadre, S. R.; Sastry., G. N. *J. Chem. Phys.* **2010**, *133*, 164308.
(478) Yang, Z.; Hua, S. G.; Hua, W. J.; Li, S. H. *J. Phys. Chem. A* **2010**, *114*, 9253.
(479) Nikitina, E.; Sulimov, V.; Zayets, V.; Zaitseva, N. *Int. J. Quantum Chem.* **2004**, *97*, 747.
(480) Stewart, J. J. P. *J. Mol. Mod.* **2009**, *15*, 765.
(481) *Combined Quantum Mechanical and Molecular Mechanical Methods*; Gao, J., Thompson, M. A., Eds.; ACS Symposium Series 712; Oxford University Press: New York, 1998.
(482) Scuseria, G. E. *J. Phys. Chem. A* **1999**, *103*, 4782.
(483) Ufimtsev, I. S.; Martinez, T. J. *J. Chem. Theory Comput.* **2009**, *5*, 2628.
(484) Li, H.; Hains, A. W.; Everts, J. E.; Robertson, A. D.; Jensen, J. H. *J. Phys. Chem. B* **2002**, *106*, 3486.
(485) Jensen, J. H.; Li, H.; Robertson, A. D.; Molina, P. A. *J. Phys. Chem. A* **2005**, *109*, 6634.
(486) Minikis, R. M.; Kairys, V.; Jensen, J. H. *J. Phys. Chem. A* **2001**, *105*, 3829.
(487) Porter, M. A.; Hall, J. R.; Locke, J. C.; Jensen, J. H.; Molina, P. A. *Proteins: Struct., Funct., Bioinf.* **2006**, *63*, 621.
(488) Naor, M. M.; Jensen, J. H. *Proteins: Struct., Funct., Bioinf.* **2004**, *57*, 799.
(489) Wang, P. F.; Flynn, A. J.; Naor, M. M.; Jensen, J. H.; Cui, G. L.; Merz, K. M.; Kenyon, G. L.; McLeish, M. J. *Biochemistry* **2006**, *45*, 11464.
(490) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 689.
(491) Xie, W.; Orozco, M.; Truhlar, D. G.; Gao, J. *J. Chem. Theory Comput* **2009**, *5*, 459.
(492) Komeiji, Y.; Ishida, T.; Fedorov, D. G.; Kitaura, K. *J. Comput. Chem.* **2007**, *28*, 1750.
(493) He, X.; Fusti-Molnar, L.; Cui, G.; Merz, K. M., Jr. *J. Phys. Chem. B* **2009**, *113*, 5290.
(494) Sawada, T.; Fedorov, D. G.; Kitaura, K. *Int. J. Quantum Chem.* **2009**, *109*, 2033.
(495) Huang, L.; Massa, L.; Karle, J. *Proc. Natl. Acad. Sciences U.S.A* **2009**, *106*, 1731.
(496) Duan, L. L.; Tong, Y.; Mei, Y.; Zhang, Q. G.; Zhang, J. Z. H. *J. Chem. Phys.* **2007**, *127*, 145101.
(497) Dong, H.; Hua, S. G.; Li, S. H. *J. Phys. Chem. A* **2009**, *113*, 1335.
(498) Deshmukh, M. M.; Bartolotti, L. J.; Gadre, S. R. *J. Phys. Chem. A* **2008**, *112*, 312.
(499) Deshmukh, M. M.; Gadre, S. R. *J. Phys. Chem. A* **2009**, *113*, 7927.
(500) Fukuzawa, K.; Kitaura, K.; Nakata, K.; Kaminuma, T.; Nakano, T. *Pure Appl. Chem.* **2003**, *75*, 2405.
(501) Fukuzawa, K.; Kitaura, K.; Uebayasi, M.; Nakata, K.; Kaminuma, T.; Nakano, T. *J. Comput. Chem.* **2005**, *26*, 1.
(502) Fukuzawa, K.; Mochizuki, Y.; Tanaka, S.; Kitaura, K.; Nakano, T. *J. Phys. Chem. B* **2006**, *110*, 16102.
(503) Sawada, T.; Hashimoto, T.; Nakano, H.; Suzuki, T.; Ishida, H.; Kiso, M. *Biochem. Biophys. Res. Commun.* **2006**, *351*, 40.
(504) Sawada, T.; Hashimoto, T.; Nakano, H.; Suzuki, T.; Suzuki, Y.; Kawaoka, Y.; Ishida, H.; Kiso, M. *Biochem. Biophys. Res. Commun.* **2007**, *355*, 6.
(505) Sawada, T.; Hashimoto, T.; Tokiwa, H.; Suzuki, T.; Nakano, H.; Ishida, H.; Kiso, M.; Suzuki, Y. *Glycoconj. J.* **2008**, *25*, 805.
(506) Sawada, T.; Hashimoto, T.; Tokiwa, H.; Suzuki, T.; Nakano, H.; Ishida, H.; Kiso, M.; Suzuki, Y. *J. Mol. Genet. Med.* **2009**, *3*, 133.
(507) Sawada, T.; Fedorov, D. G.; Kitaura, K. *J. Am. Chem. Soc.* **2010**, *132*, 16862.
(508) Sawada, T.; Fedorov, D. G.; Kitaura, K. *J. Phys. Chem. B* **2010**, *114*, 15700.
(509) Takematsu, K.; Fukuzawa, K.; Omagari, K.; Nakajima, S.; Nakajima, K.; Mochizuki, Y.; Nakano, T.; Watanabe, H.; Tanaka, S. *J. Phys. Chem. B* **2009**, *113*, 4991.

(510) Iwata, T.; Fukuzawa, K.; Nakajima, K.; Aida-Hyugaji, S.; Mochizuki, Y.; Watanabe, H.; Tanaka, S. *Comput. Biol. Chem.* **2008**, *32*, 198.
(511) Fukuzawa, K.; Omagari, K.; Nakajima, K.; Nobusawa, E.; Tanaka, S. *Prot. Pept. Lett.* **2011**, *18*, 530.
(512) Yamagishi, K.; Yamamoto, K.; Yamada, S.; Tokiwa, H. *Chem. Phys. Lett.* **2006**, *420*, 465.
(513) Motoyoshi, S.; Yamagishi, K.; Yamada, S.; Tokiwa, H. *J. Steroid Biochem. Mol. Biol.* **2010**, *121*, 56.
(514) Yamagishi, K.; Tokiwa, H.; Makishima, M.; Yamada, S. *J. Steroid Biochem. Mol. Biol.* **2010**, *121*, 63.
(515) Ito, M.; Fukuzawa, K.; Mochizuki, Y.; Nakano, T.; Tanaka, S. *J. Phys. Chem. B* **2007**, *111*, 3525.
(516) Ito, M.; Fukuzawa, K.; Mochizuki, Y.; Nakano, T.; Tanaka, S. *J. Phys. Chem. A* **2008**, *112*, 1986.
(517) Ito, M.; Fukuzawa, K.; Ishikawa, T.; Mochizuki, Y.; Nakano, T.; Tanaka, S. *J. Phys. Chem. B.* **2008**, *112*, 12081.
(518) Nakanishi, I.; Fedorov, D. G.; Kitaura, K. *Proteins: Struct., Funct., Bioinf.* **2007**, *68*, 145.
(519) Sugiki, S.-I.; Matsuoka, M.; Usuki, R.; Sengoku, Y.; Kurita, N.; Sekino, H.; Tanaka, S. *J. Theor. Comput. Chem.* **2005**, *4*, 183.
(520) Nemoto, T.; Fedorov, D. G.; Uebayasi, M.; Kanazawa, K.; Kitaura, K.; Komeiji, Y. *Comput. Biol. Chem.* **2005**, *29*, 434.
(521) Watanabe, H.; Enomoto, T.; Tanaka, S. *Biochem. Biophys. Res. Commun.* **2007**, *361*, 367.
(522) Harada, T.; Yamagishi, K.; Nakano, T.; Kitaura, K.; Tokiwa, H. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2008**, *377*, 607.
(523) Tada, M.; Nagasima, T.; Udagawa, T.; Tachikawa, M.; Sugawara, H. *J. Mol. Struct. (THEOCHEM)* **2009**, *897*, 149.
(524) Dedachi, K.; Khan, M. T. H.; Sylte, I.; Kurita, N. *Chem. Phys. Lett.* **2009**, *479*, 290.
(525) Van Schouwen, B. M. B.; Nakano, M.; Watanabe, H.; Tanaka, S.; Gordon, H. L.; Rothstein, S. M. *J. Mol. Struct. (THEOCHEM)* **2010**, *944*, 12.
(526) Yamagishi, K.; Yamamoto, K.; Mochizuki, Y.; Nakano, T.; Yamada, S.; Tokiwa, H. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3344.
(527) Yoshikawa, E.; Miyagi, S.; Dedachi, K.; Ishihara-Sugano, M.; Itoh, S.; Kurita, N. *J. Mol. Graphics Modell.* **2010**, *29*, 197.
(528) Ishikawa, T.; Kuwata, K. *J. Chem. Theory Comput.* **2010**, *6*, 538.
(529) Hasegawa, K.; Mohri, S.; Yokoyama, T. *Prion* **2010**, *4*, 38.
(530) Tsuji, S.; Kasumi, T.; Nagase, K.; Yoshikawa, E.; Kobayashi, H.; Kurita, N. *J. Mol. Graph. Modell.* **2011**, *29*, 975.
(531) Fukuzawa, K.; Komeiji, Y.; Mochizuki, Y.; Kato, A.; Nakano, T.; Tanaka, S. *J. Comput. Chem.* **2006**, *27*, 948.
(532) Kurisaki, I.; Fukuzawa, K.; Komeiji, Y.; Mochizuki, Y.; Nakano, T.; Imada, J.; Chmielewski, A.; Rothstein, S. M.; Watanabe, H.; Tanaka, S. *Biophys. Chem.* **2007**, *130*, 1.
(533) Watanabe, T.; Inadomi, Y.; Fukuzawa, K.; Nakano, T.; Tanaka, S.; Nilsson, L.; Nagashima, U. *J. Phys. Chem. B.* **2007**, *111*, 9621.
(534) Kurisaki, I.; Fukuzawa, K.; Nakano, T.; Mochizuki, Y.; Watanabe, H.; Tanaka, S. *J. Mol. Struct. (THEOCHEM)* **2010**, *962*, 45.
(535) Ohyama, T.; Hayakawa, M.; Nishikawa, S.; Kurita, N. *J. Comput. Chem.* **2011**, *32*, 1661.
(536) Mori, H.; Ueno-Noto, K. *J. Phys. Chem. B* **2011**, *115*, 4774.
(537) Ozawa, T.; Okazaki, K. *J. Comput. Chem.* **2008**, *29*, 2656.
(538) Ozawa, T.; Tsuji, E.; Ozawa, M.; Handa, C.; Mukaiyama, H.; Nishimura, T.; Kobayashi, S.; Okazaki, K. *Bioorg. Med. Chem.* **2008**, *16*, 10311.
(539) Fujimura, K.; Sasabuchi, Y. *Chem. Med. Chem.* **2010**, *5*, 1254.
(540) Ohno, K.; Mori, K.; Orita, M.; Takeuchi, M. *Curr. Med. Chem.* **2011**, *18*, 220.
(541) Zhang, D. W.; Zhang, J. Z. H. *Int. J. Quantum Chem.* **2005**, *103*, 246.
(542) Mei, Y.; He, X.; Xiang, Y.; Zhang, D. W.; Zhang, J. Z. H. *Proteins: Struct., Funct., Bioinf.* **2005**, *59*, 489.
(543) He, X.; Mei, Y.; Xiang, Y.; Zhang, D. W.; Zhang, J. Z. H. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 423.
(544) Wu, E. L.; Han, K. L.; Zhang, J. Z. H. *J. Theor. Comput. Chem.* **2009**, *8*, 551.
(545) Tong, Y.; Mei, Y.; Zhang, J. Z. H.; Duan, L. L.; Zhang, Q. G. *J. Theor. Comput. Chem.* **2009**, *8*, 1265.
(546) Huang, L. L.; Massa, L.; Karle, J. *Proc. Natl. Acad. Sciences U.S.A* **2007**, *104*, 4261.
(547) Orimoto, Y.; Gu, F. L.; Imamura, A.; Aoki, Y. *J. Chem. Phys.* **2007**, *126*, 215104.
(548) Mazanetz, M. P.; Ichihara, O.; Law, R. J.; Whittaker, M. *J. Cheminf.* **2011**, *3*, 2.
(549) Hayik, S. A.; Dunbrack, R., Jr.; Merz, K. M., Jr. *J. Chem. Theory Comput.* **2010**, *6*, 3079.
(550) Whittaker, M.; Law, R. J.; Ichihara, O.; Hesterkamp, T.; Hallett, D. *Drug Discovery Today: Technol.* **2010**, *7*, No. e163 and references therein.
(551) Ichihara, O.; Barker, J.; Law, R. J.; Whittaker, M. *Mol. Inf.* **2011**, *30*, 298.
(552) Yoshida, T.; Fujita, T.; Chuman, H. *Curr. Comput.-Aided Drug Des.* **2009**, *5*, 38.
(553) Ishikawa, T.; Mochizuki, Y.; Amari, S.; Nakano, T.; Tanaka, S.; Tanaka, K. *Chem. Phys. Lett.* **2008**, *463*, 189.
(554) Yoshida, T.; Yamagishi, K.; Chuman, H. *QSAR Comb. Sci.* **2008**, *27*, 694.
(555) Fischer, B.; Fukuzawa, K.; Wenzel, W. *Proteins: Struct., Funct., Bioinf.* **2008**, *70*, 1264.
(556) Yoshida, T.; Munei, Y.; Hitaoka, S.; Chuman, H. *J. Chem. Inf. Model* **2010**, *50*, 850.
(557) Hitaoka, S.; Harada, M.; Yoshida, T.; Chuman, H. *J. Chem. Inf. Model* **2010**, *50*, 1796.
(558) Munei, Y.; Shimamoto, K.; Harada, M.; Yoshida, T.; Chuman, H. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 141.
(559) Ishida, T.; Fedorov, D. G.; Kitaura, K. *J. Phys. Chem. B* **2006**, *110*, 1457.
(560) Nakamura, T.; Yamaguchi, A.; Kondo, H.; Watanabe, H.; Kurihara, T.; Esaki, N.; Hirono, S.; Tanaka, S. *J. Comput. Chem.* **2009**, *30*, 2625.
(561) Mochizuki, Y.; Nakano, T.; Amari, S.; Ishikawa, T.; Tanaka, K.; Sakurai, M.; Tanaka, S. *Chem. Phys. Lett.* **2007**, *433*, 360.
(562) Tagami, A.; Ishibashi, N.; Kato, D.; Taguchi, N.; Mochizuki, Y.; Watanabe, H.; Ito, M.; Tanaka, S. *Chem. Phys. Lett.* **2009**, *472*, 118.
(563) Taguchi, N.; Mochizuki, Y.; Nakano, T.; Amari, S.; Fukuzawa, K.; Ishikawa, T.; Sakurai, M.; Tanaka, S. *J. Phys. Chem. B* **2009**, *113*, 1153.
(564) Taguchi, N.; Mochizuki, Y.; Nakano, T. *Chem. Phys. Lett.* **2011**, *504*, 76.
(565) Ikegami, T.; Ishida, T.; Fedorov, D. G.; Kitaura, K.; Inadomi, Y.; Umeda, H.; Yokokawa, M.; Sekiguchi, S. *J. Comput. Chem.* **2010**, *31*, 447.
(566) Milne, B. F.; Marques, M. A. L.; Nogueira, F. *Phys. Chem. Chem. Phys.* **2010**, *12*, 14285.
(567) Huang, L.; Massa, L.; Karle, I.; Karle, J. *Proc. Natl. Acad. Sci. U.S.A* **2009**, *106*, 3664.
(568) Fukunaga, H.; Fedorov, D. G.; Chiba, M.; Nii, K.; Kitaura, K. *J. Phys. Chem. A* **2008**, *112*, 10887.
(569) Zhang, R. J.; Tian, W. Q.; Gu, F. L.; Aoki, Y. *J. Phys. Chem. C* **2007**, *111*, 6350.
(570) Chen, W.; Yu, G. T.; Gu, F. L.; Aoki, Y. *J. Phys. Chem. C* **2009**, *113*, 8447; ibid. 11424.
(571) Otaki, H.; Ando, K. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10719.
(572) Orimoto, Y.; Aoki, Y. *J. Polymer Science Part B-Polymer Phys.* **2006**, *44*, 119.
(573) Ohnishi, S. I.; Orimoto, Y.; Gu, F. L.; Aoki, Y. *J. Chem. Phys.* **2007**, *127*, 084702.
(574) Yu, G. T.; Chen, W.; Gu, F. L.; Orimoto, Y.; Aoki, Y. *Mol. Phys.* **2009**, *107*, 81.
(575) Yu, G. T.; Chen, W.; Gu, F. L.; Aoki, Y. *J. Comput. Chem.* **2010**, *31*, 863.
(576) Pomogaeva, A.; Gu, F. L.; Imamura, A.; Aoki, Y. *Theor. Chem. Acc.* **2010**, *125*, 453.

(577) Orimoto, Y.; Gu, F. L.; Korchowiec, J.; Imamura, A.; Aoki, Y. *Theor. Chem. Acc.* **2010**, *125*, 493.
(578) Yan, L. K.; Pomogaeva, A.; Gu, F. L.; Aoki, Y. *Theor. Chem. Acc.* **2010**, *125*, 511.
(579) Pomogaev, V.; Gu, F. L.; Pomogaeva, A.; Aoki, Y. *Int. J. Quantum Chem.* **2009**, *109*, 1328.