JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Energy Decomposition Analysis Based on Absolutely Localized Molecular Orbitals for Large-Scale Density Functional Theory Calculations in Drug Design

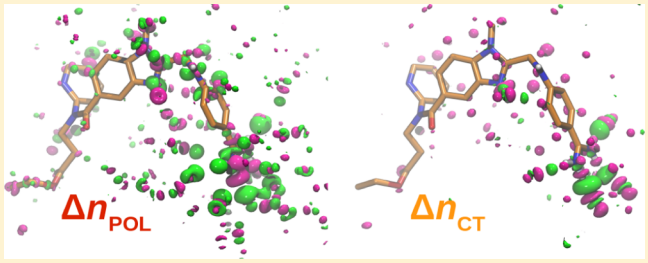
M. J. S. Phipps,[†] T. Fox,[‡] C. S. Tautermann,[‡] and C.-K. Skylaris*[,†]

[†]School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom

[‡]Lead Identification and Optimization Support, Boehringer Ingelheim Pharma GmbH & Co. KG, 88397 Biberach, Germany

Supporting Information

ABSTRACT: We report the development and implementation of an energy decomposition analysis (EDA) scheme in the ONETEP linear-scaling electronic structure package. Our approach is hybrid as it combines the localized molecular orbital EDA (Su, P.; Li, H. *J. Chem. Phys.*, **2009**, *131*, 014102) and the absolutely localized molecular orbital EDA (Khaliullin, R. Z.; et al. *J. Phys. Chem. A*, **2007**, *111*, 8753−8765) to partition the intermolecular interaction energy into chemically distinct components (electrostatic, exchange, correlation, Pauli repulsion, polarization, and charge transfer). Limitations shared in EDA approaches such as the issue of basis set dependence in polarization and charge transfer are discussed, and a remedy to this problem is proposed that exploits the strictly localized property of the ONETEP orbitals. Our method is validated on a range of complexes with interactions relevant to drug design. We demonstrate the capabilities for large-scale calculations with our approach on complexes of thrombin with an inhibitor comprised of up to 4975 atoms. Given the capability of ONETEP for large-scale calculations, such as on entire proteins, we expect that our EDA scheme can be applied in a large range of biomolecular problems, especially in the context of drug design.



1. INTRODUCTION

Intermolecular interactions are of key importance in determining the physical and chemical properties of molecular systems. For example, the intermolecular forces that govern a hydrogen bond are of great interest to chemists due to the role they play in determining polymeric structures and macroscopic properties of structures.[1,2] The effect of σ holes as observed in a number of halogen-bonding complexes[3] has implications for drug−host binding, and therefore, this interaction is of key pharmaceutical interest. Scientific fields that involve the study of systems containing large numbers of intermolecular interactions, such as the disciplines of biomolecular[4,5] and supramolecular chemistry[6,7] and condensed matter physics,[8,9] benefit from the insights gained from studies of intermolecular bonding.

The interaction energy is defined in the supermolecular approach as the difference between the ground-state energies of a number of monomers in isolation and the ground-state energy of the supermolecular complex that these monomers form, and the energy value obtained from this approach can provide for highly detailed and accurate studies of systems of interest. However, chemical effects such as electrostatics, Pauli repulsion, polarization, and charge transfer that each contribute to the overall intermolecular interaction are hidden within this value. Such bonding components are important from a chemical stance as they provide valuable insight into the synergy and nature of bonding. Additionally, these components are also valuable pedagogically as they provide a link between the world of *ab initio* quantum calculations and the familiar picture of intermolecular forces known to chemists.

In recent times, many developments have been made in gaining insight into these often complex interactions. As there is no unique decomposition, a large number of perturbative and variational energy decomposition analysis methods exist[10−20] that seek to decompose the interaction energy ΔE into its constituent chemical components. These include the symmetry-adapted perturbation theory (SAPT) scheme[19,20] based on the perturbative treatment of fragment wave functions in which chemically useful terms such as electrostatic, exchange repulsion, induction, and dispersion are obtained. Many variational schemes have been developed from the pioneering Kitaura−Morokuma (KM) scheme.[10,21,22] These include, for example, the reduced variational space (RVS)[11] approach, the constrained space orbital variations (CSOV)[12,23] energy decomposition, and the natural energy decomposition analysis (NEDA) scheme[13,14,24,25] based on natural bond orbital theory.[15,26] Additionally, schemes have been proposed and implemented for analyses of periodic extended systems including the EDA of Raupach and Tonner,[27] the CSOV

Received: March 15, 2016
Published: June 1, 2016

DOI: 10.1021/acs.jctc.6b00272
J. Chem. Theory Comput. 2016, 12, 3135−3148

EDA implementation in the CRYSTAL package,[28] and the EDA of Philipsen and Baerends[29] within the BAND code.[30]

These methods have demonstrated varying degrees of success in the study of protein−ligand interactions[31] and show that the problem of separating polarization and charge transfer effects is approached in a variety of ways. For example, schemes such as the localized molecular orbital (LMO)[16] and extended transition state (ETS)[32−35] EDAs simply do not seek to separate polarization and charge transfer effects at all. Other schemes that attempt this partitioning consider the problem as involving the optimization of molecular orbitals (MOs) subject to conditions that impose localization to particular fragments. Schemes that belong to this category include the KM, RVS, CSOV, absolutely localized molecular orbital (ALMO),[17] and block-localized wave function (BLW)[36,37] EDAs. The NEDA scheme that involves a highly localized picture of bonding also considers this to be an orbital localization problem; however, its solution does not involve variational optimization of a polarized state. The EDA of Wu et al.[18] considers charge transfer to be more appropriately partitioned by charge density using the constrained density functional theory (CDFT)[38] approach.

The use of these methods has generally been limited to studies involving no more than a few tens of atoms.[31] Exceptions to this include use of the fragment molecular orbital (FMO) framework[39−41] in schemes such as the PIEDA.[42] Recently, work has been published[43] investigating 18 class A GPCR-ligand crystal structures using the FMO-MP2 (second-order Møller−Plesset perturbation theory)[44] PIEDA method with the 6-31G* basis set, demonstrating the components of the interactions and hence the high suitability of EDA to the field of drug design especially in the case of large drug−protein systems. With the developments made in the field of linear-scaling density functional theory (DFT) methods[45,46] a new opportunity to remove this size limitation arises. The ALMO EDA scheme has been shown to give reasonable polarization and charge transfer energies[31] and benefits from the use of a properly antisymmetrized state in performing this partitioning. Notably, use of SCF-MI in separating polarization and charge transfer offers improvement upon the scaling of SCF calculations due to partitioning of the diagonalization bottleneck.[47] This has been demonstrated by Mo and Gao[48] in their quantum mechanical and molecular mechanical coupled BLW EDA (of relation to ALMO EDA) study of ionic system solvation with up to 32 water molecules being treated quantum mechanically. The frozen density energy component of the ALMO EDA decomposition provides insights into the interaction of fragment densities; however, also contained within this term are effects originating from electrostatic, exchange, Pauli repulsion, and correlation interactions. A rigorous treatment of the frozen density component is possible that further decomposes the term into its constituents via a modified LMO EDA[16] analysis.

Within this paper, we provide a brief summary of the ALMO EDA scheme of Head-Gordon et al.[17] and the LMO EDA scheme of Su and Li[16] and describe our implementation of a hybrid EDA scheme in the ONETEP[49] linear-scaling DFT package. The limitations of using the ALMO approach to partition polarization from charge transfer effects is explored and a solution proposed that exploits the strictly localized property of the ONETEP orbitals. An assessment of the method is provided through its application to a test set of small complexes that display biomolecule-like bonding characteristics. We then demonstrate the ability of our approach to be applied on entire protein−ligand systems by performing EDA calculations on a 4975 atom complex between thrombin and one of its inhibitors.

2. THEORY

In the following discussion of the theory, we have adopted the Einstein convention of implicit summation over repeated Greek indices only. We will use the word "fragment" to describe any molecule which is part of a complex held together by intermolecular interactions (i.e., no chemical bonds). All our fragments in the following examples are closed shell, but it is a straightforward task to extend the implementation to open-shell systems.

2.1. Localized and Absolutely Localized Energy Decomposition Analyses. In order to develop the equations used in the EDA approaches, it is first necessary to formally express the supermolecule interaction energy. This is given by

$$\Delta E = E[\Psi_X] - \sum_{A\in X}^{N_{frag}} [\Psi_A^{0,A}] + \Delta E_{nuc} + \Delta E_{BSSE}$$

$$\Delta E_{nuc} = E_{nuc}(X) - \sum_{A\in X}^{N_{frag}} E_{nuc}(A) \quad (1)$$

where Ψ represents an electronic wave function with the subscripts A and X used to represent the fragments A that compose the supermolecule X, with the superscript "0" used to denote states optimized in isolation from the field of the remaining fragments of the supermolecule and the superscript "A" used to indicate optimization only in the fragment A basis set, and where N_{frag} is the number of fragments comprising the supermolecule X. Here, $E[\Psi_Y]$ is the energy of system Y and $E_{nuc}(Y)$ represents its nuclei−nuclei electrostatic energy. Also, ΔE_{BSSE} is included as a correction for the basis set superposition error (BSSE).

2.1.1. Absolutely Localized Energy Decomposition Analysis. The ALMO EDA developed by Head-Gordon et al.[17] decomposes the interaction energy ΔE into the frozen density ΔE_{FRZ}, polarization ΔE_{POL}, and charge transfer ΔE_{CT} interaction terms as

$$\Delta E = \Delta E_{FRZ} + \Delta E_{POL} + \Delta E_{CT} \quad (2)$$

We define a number of key wave functions that are necessary to allow a clear description of the EDA theory in this paper. The electronic wave function $\Psi_A^{0,A}$ is the wave function of the arbitrary fragment A, and $\Psi_X^{0,orth}$ is the properly antisymmetrized wave function of the interacting X complex constructed from the nonorthogonal frozen occupied MOs of the fragments. Here, Ψ_X^{ALMO} is defined as the wave function of the polarized state produced by relaxation of the nonorthogonal frozen MOs while still ensuring localization of the MOs within their corresponding fragments. This localization is imposed through the constraint that the MOs are only expanded using basis functions on a particular fragment and therefore prevents charge delocalization between fragments. Also, Ψ_X is defined as the wave function of a fully antisymmetrized and fully relaxed state. Each energy component may be described in terms of energy functionals of these intermediate wave functions as

$$\Delta E_{FRZ} = E[\Psi_X^{0,orth}] - \sum_{A\in X}^{N_{frag}} E[\Psi_A^{0,A}] + \Delta E_{nuc} \quad (3)$$

DOI: 10.1021/acs.jctc.6b00272
J. Chem. Theory Comput. 2016, 12, 3135−3148

$$\Delta E_{\mathrm{POL}} = E[\Psi_{\mathrm{X}}^{\mathrm{ALMO}}] - E[\Psi_{\mathrm{X}}^{0,\mathrm{orth}}] \tag{4}$$

$$\Delta E_{\mathrm{CT}} = E[\Psi_{\mathrm{X}}] - E[\Psi_{\mathrm{X}}^{\mathrm{ALMO}}] + \Delta E_{\mathrm{BSSE}} \tag{5}$$

where N_{frag} is the number of fragments comprising the supermolecule X, and ΔE_{BSSE} is calculated using the Boys–Bernardi[50] counterpoise correction (CP) approach. The geometries of fragments A are defined as found in the (geometry optimized) X complex (i.e., only the electronic density is optimized in this calculation).

The ALMO EDA achieves the separation of polarization and charge transfer effects using the ALMO state, $\Psi_{\mathrm{X}}^{\mathrm{ALMO}}$. This state describes a system which is polarized but that has the restriction of no charge transfer between the fragments. The optimization of the ALMO state proceeds via a procedure termed locally projected SCF for molecular interactions (SCF-MI).[47,51] The SCF-MI equations we use are those of Stoll et al.[52] and are further described in Section SI.1 of the Supporting Information. We note that this approach of separating out polarization and charge transfer effects is the same as that described within the BLW[36,37] EDA procedure. Specifically, the BLW state that is identical to the ALMO state is optimized either by Jacobi transformation[53] or by the algorithm of Gianinetti and Raimondi et al.[54,55]

2.1.2. Frozen Density analysis. In this section, we describe a prescription for a frozen density analysis produced via the LMO EDA theory of Su and Li,[16] in which ΔE_{FRZ} is further decomposed into individual electrostatic (ΔE_{ES}), exchange (ΔE_{EX}), Pauli repulsion (ΔE_{REP}), and correlation (ΔE_{CORR}) terms as

$$\Delta E_{\mathrm{FRZ}} = \Delta E_{\mathrm{ES}} + \Delta E_{\mathrm{EX}} + \Delta E_{\mathrm{REP}} + \Delta E_{\mathrm{CORR}} \tag{6}$$

Combining the classical-like terms of the Kohn–Sham (KS) energy (by excluding the exchange and correlation contributions) as

$$E_{\mathrm{cl}} = 2\sum_{i,j}^{N_{\psi(\mathrm{occ})}} h_{ij}\sigma^{ij} + \sum_{i,j,k,l}^{N_{\psi(\mathrm{occ})}} \langle ij|kl\rangle \times \sigma^{ij}\sigma^{kl} + E_{\mathrm{nuc}} \tag{7}$$

where $N_{\psi(\mathrm{occ})}$ is the number of occupied MOs, kinetic energy effects are included by the presence of the full core Hamiltonian matrix h_{ij}, and where the contravariant metric tensor σ^{ij} is the inverse of the MO overlap matrix σ_{ij}

$$\sigma_{ij} = \int d\mathbf{r}\psi_i^*(\mathbf{r})\psi_j(\mathbf{r}) = \langle\psi_i|\psi_j\rangle \tag{8}$$

$$\sigma^{ij} \equiv (\sigma^{-1})_{ij} \tag{9}$$

the components of the frozen density interaction may be expressed using intermediate wave functions as

$$\Delta E_{\mathrm{ES}} = E_{\mathrm{cl}}[\Psi_{\mathrm{X}}^0] - \sum_{A\in\mathrm{X}}^{N_{\mathrm{frag}}} E_{\mathrm{cl}}[\Psi_A^{0,A}] \tag{10}$$

$$\Delta E_{\mathrm{EX}} = E_{\mathrm{X}}[\Psi_{\mathrm{X}}^0] - \sum_{A\in\mathrm{X}}^{N_{\mathrm{frag}}} E_{\mathrm{X}}[\Psi_A^{0,A}] \tag{11}$$

$$\Delta E_{\mathrm{REP}} = E_{\mathrm{cl}}[\Psi_{\mathrm{X}}^{0,\mathrm{orth}}] - E_{\mathrm{cl}}[\Psi_{\mathrm{X}}^0] + E_{\mathrm{X}}[\Psi_{\mathrm{X}}^{0,\mathrm{orth}}] - E_{\mathrm{X}}[\Psi_{\mathrm{X}}^0] \tag{12}$$

$$\Delta E_{\mathrm{CORR}} = E_{\mathrm{C}}[\Psi_{\mathrm{X}}^{0,\mathrm{orth}}] - \sum_{A\in\mathrm{X}}^{N_{\mathrm{frag}}} E_{\mathrm{C}}[\Psi_A^{0,A}] \equiv \Delta E_{\mathrm{FRZ}} - (\Delta E_{\mathrm{ES}} + \Delta E_{\mathrm{EX}} + \Delta E_{\mathrm{REP}}) \tag{13}$$

where $E_{\mathrm{X}}[n(\mathbf{r})]$ and $E_{\mathrm{C}}[n(\mathbf{r})]$ are the DFT exchange and correlation energy functionals, respectively, of the electronic density $n(\mathbf{r})$ associated with the wave function $\Psi(\mathbf{r})$ and where we note that the kinetic energy contribution to ΔE_{ES} is zero due to the kinetic energy of the fragment states $\Psi_A^{0,A}$ canceling with the kinetic energy of the frozen supermolecule state Ψ_{X}^0 as σ^{ij} is the identity matrix for $E_{\mathrm{cl}}[\Psi_{\mathrm{X}}^0]$, while it is as in eqs 8 and 9 for $E_{\mathrm{cl}}[\Psi_{\mathrm{X}}^{0,\mathrm{orth}}]$.

The components of this frozen energy decomposition are formally similar to the components described in the LMO EDA[16] but with inclusion of a new component ΔE_{CORR} that isolates the correlation energy change on orthogonalization contained within the LMO EDA "polarization" term. It is also noted that the electrostatic term of this decomposition is formally similar to that within the BLW[36,37] scheme. Additionally, it is important to note the development of a number of alternative approaches to the above theory such as the scheme of Horn et al.[56] for investigating (decomposing) the frozen interaction in KS DFT calculations and the density-based energy decomposition analysis (DEDA) of Wu[57] that calculates the frozen density interaction in a variational manner.

2.2. Combined ONETEP Approach. Our aim has been to implement within the ONETEP[49] package a combination of the ALMO EDA approach[17] and the frozen density component analysis based upon the LMO EDA theory of Su and Li[16] above. Our prescription includes the electrostatic, exchange, Pauli repulsion, and modified correlation terms of the LMO EDA of Su and Li[16] in the frozen density component of the ALMO EDA, with the ALMO EDA polarization and charge transfer components remaining as in their original form. By substitution of eq 6 into eq 2, we obtain the fully decomposed ΔE as

$$\Delta E = \Delta E_{\mathrm{ES}} + \Delta E_{\mathrm{EX}} + \Delta E_{\mathrm{REP}} + \Delta E_{\mathrm{CORR}} + \Delta E_{\mathrm{POL}} + \Delta E_{\mathrm{CT}} \tag{14}$$

The implementation of this EDA scheme within a linear-scaling DFT code such as ONETEP that aims to achieve large basis set accuracy in routine calculations is an important capability. By decomposing the frozen density component into its substituents and expressing the interaction energy in terms of six rigorously defined energy terms, our EDA scheme provides a detailed analysis of the interaction energy that allows understanding of the key chemical components of binding. In what follows, we describe the ONETEP EDA implementation and provide examples of calculations using this.

3. IMPLEMENTATION OF EDA IN ONETEP

3.1. ONETEP Formalism. In the ONETEP formalism, the set of KS orbitals $\{|\psi_i\rangle\}$ can be considered as linear combinations of a set of localized orbitals $\{|\phi_\alpha\rangle\}$ known as nonorthogonal generalized Wannier functions (NGWFs)[58] as

$$\psi_i(\mathbf{r}) = \phi_\alpha(\mathbf{r})M_i^\alpha \tag{15}$$

where $\{M_i^\alpha\}$ are the KS MO expansion coefficients. The NGWFs $\{|\phi_\alpha\rangle\}$ are expanded in a basis of N_p periodic cardinal sine (psinc) functions[59,60] $\{D_p\}$ as

DOI: 10.1021/acs.jctc.6b00272
J. Chem. Theory Comput. 2016, 12, 3135–3148

$$\phi_\alpha(\mathbf{r}) = \sum_{p=1}^{N_p} D_p(\mathbf{r}) c_{p\alpha} \tag{16}$$

where $c_{p\alpha}$ are the expansion coefficients. The psinc functions are centered on a regular grid that spans the whole simulation cell. The grid point each psinc function is centered on is given by the index p. The information on the KS states is contained within the density matrix given by

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_{i}^{N_{\psi(\mathrm{occ})}} \psi_i(\mathbf{r}) \psi_i^*(\mathbf{r}') \tag{17}$$

and the electronic density is obtained from $\rho(\mathbf{r},\mathbf{r}')$ as

$$n(\mathbf{r}) = 2\rho(\mathbf{r}, \mathbf{r}) \tag{18}$$

Orthonormality of $\{|\psi_i\rangle\}$ is a necessary physical constraint for the Pauli principle to be respected, which can be expressed as

$$\sigma_{ij} = \langle\psi_i|\psi_j\rangle = (M^\dagger)_i^{\ \alpha} S_{\alpha\beta} M^\beta_{\ j} = \delta_{ij} \tag{19}$$

where the NGWF overlap matrix $S_{\alpha\beta}$ is given by

$$S_{\alpha\beta} = \langle\phi_\alpha|\phi_\beta\rangle \tag{20}$$

By substitution of eq 15 into eq 17, we introduce the density kernel $K^{\alpha\beta}$ as the representation of the density operator in terms of NGWFs

$$\rho(\mathbf{r}, \mathbf{r}') = \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi_\beta^*(\mathbf{r}') \tag{21}$$

where the density kernel $K^{\alpha\beta}$ is given by

$$K^{\alpha\beta} = \sum_{i}^{N_\psi} M^\alpha_{\ i} f_i (M^\dagger)_i^{\ \beta} \tag{22}$$

where f_i are the MO occupancies which are equal to 1 for the doubly occupied MOs $\psi_i(\mathbf{r})$.

3.2. Construction of Frozen Density Kernels. In the frozen density analysis, we seek expressions of the kernel and set of NGWFs that represent the two intermediate wave functions Ψ_X^0 and $\Psi_X^{0,\mathrm{orth}}$. The NGWF set of these wave functions, $\{\phi_{A\alpha}(\mathbf{r})\}$, is constructed by superposition of the NGWFs of the isolated fragments A onto the supermolecule X. We now define the two density kernels $K_0^{A\alpha,B\beta}$ and $K_{\mathrm{orth}}^{A\alpha,B\beta}$ that are used to represent Ψ_X^0 and $\Psi_X^{0,\mathrm{orth}}$ in this "frozen" NGWF set.

The density kernel of the frozen state, $K_0^{A\alpha,B\beta}$, is constructed by superposition of the occupied states of the fragments onto the supermolecule. This is constructed from the fragment density kernels $(K_A)^{\alpha\beta}$ as

$$K_0^{A\alpha,A\beta} = (K_A)^{\alpha\beta}$$
$$K_0^{A\alpha,B\beta} = 0|_{A\neq B} \tag{23}$$

The density kernel $K_{\mathrm{orth}}^{A\alpha,B\beta}$ is the properly antisymmetrized form of $K_0^{A\alpha,B\beta}$. The construction of this kernel requires the coefficients $(M_A)_i^\alpha$ of the isolated fragments. From eq 22 for the density kernel, we can write a generalized eigenvalue problem as

$$K^{\alpha\beta} S_{\beta\gamma} M_i^\gamma = M^\alpha_{\ i} f_i \tag{24}$$

This eigenvalue equation may be expressed for the fragment problems as

$$(K_A)^{\alpha\beta} (S_A)_{\beta\gamma} (M_A)_i^\gamma = (M_A)^\alpha_{\ i} (f_A)_i \tag{25}$$

where the subscript A is used to denote a fragment quantity. The vectors of the fragment matrices M_A that are obtained for unoccupied MOs are discarded at this point. This is achieved by ordering the MOs by decreasing orbital occupancy and retaining only $(M_A)_{\beta i}|_{i\leq N_{\psi(\mathrm{occ,A})}}$, where $N_{\psi(\mathrm{occ,A})}$ is the number of occupied MOs on fragment A.

The block-diagonal MO coefficient matrix $M_{Bi}^{A\alpha}$ of the frozen supermolecule is constructed from the fragment coefficient matrices $(M_A)_i^\alpha$ as

$$M^{A\alpha}_{\ Ai} = (M_A)^\alpha_{\ i}$$
$$M^{A\alpha}_{\ Bi} = 0|_{A\neq B} \tag{26}$$

Finally, the density kernel $K_{\mathrm{orth}}^{A\alpha,B\beta}$ of the properly antisymmetrized state is expressed using these MO coefficients as

$$K_{\mathrm{orth}}^{A\alpha,B\beta} = \sum_{C,D\in X}^{N_{\mathrm{frag}}} \sum_{i}^{N_{\psi(\mathrm{occ,C})}} \sum_{j}^{N_{\psi(\mathrm{occ,D})}} M^{A\alpha}_{\ Ci} \sigma^{Ci,Dj} (M^\dagger)_{Dj}^{\ B\beta}$$
$$= \sum_{i}^{N_{\psi(\mathrm{occ,A})}} \sum_{j}^{N_{\psi(\mathrm{occ,B})}} M^{A\alpha}_{\ Ai} \sigma^{Ai,Bj} (M^\dagger)_{Bj}^{\ B\beta} \tag{27}$$

where we have exploited the block-diagonal structure of M described in eq 26, where i and j run over the occupied MOs of the fragments, and where the inverse MO overlap matrix $\sigma^{Ai,Bj}$ is constructed from the quantity $\sigma_{Ai,Bj}$

$$\sigma_{Ai,Bj} = \sum_{C,D\in X}^{N_{\mathrm{frag}}} (M^\dagger)_{Ai}^{\ C\alpha} S_{C\alpha,D\beta} M^{D\beta}_{\ Bj} \tag{28}$$

where $S_{C\alpha,D\beta}$ is the full NGWF set overlap matrix for the supermolecule.

For clarity, the structures of the quantities $M^{A\alpha}_{\ Bi}$ and $\sigma_{Ai,Bj}$ involved in constructing $K_{\mathrm{orth}}^{A\alpha,B\beta}$ are shown diagrammatically in Figure 1.

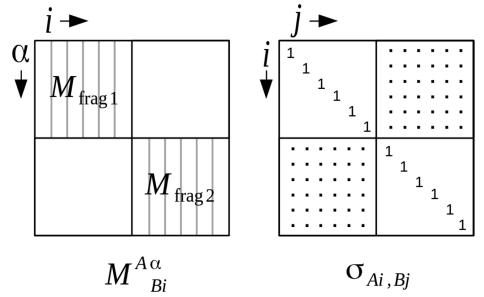


Figure 1. Diagrammatic representation of the quantities involved in the construction of the frozen density kernel $K_{\mathrm{orth}}^{A\alpha,B\beta}$ for a system comprising two fragments "frag1" and "frag2". Empty matrix sub-blocks represent zero blocks, and sub-blocks containing dots represent nonzero blocks. The diagonal intrafragmental MO overlap sub-blocks of $\sigma_{Ai,Bj}$ are identity due to orthonormality of the KS states within fragments and nonzero for the sub-blocks representing interfragment MO overlaps due to lack of orthonormality.

The energies of the two intermediate wave functions Ψ_X^0 and $\Psi_X^{0,\mathrm{orth}}$ are expressed in terms of energy functionals of the quantities $K_0^{A\alpha,B\beta}$, $K_{\mathrm{orth}}^{A\alpha,B\beta}$, and $\{\phi_{A\alpha}(\mathbf{r})\}$ as

$$E[\Psi_X^0] \equiv E[\{K_0^{A\alpha,B\beta}\}, \{\phi_{A\alpha}(\mathbf{r})\}] \tag{29}$$

DOI: 10.1021/acs.jctc.6b00272
J. Chem. Theory Comput. 2016, 12, 3135−3148

$$E[\Psi_X^{0,\mathrm{orth}}] \equiv E[\{K_{\mathrm{orth}}^{A\alpha,B\beta}\}, \{\phi_{A\alpha}(\mathbf{r})\}] \tag{30}$$

3.3. Calculation of Polarized Density Kernel. A brief overview of the ALMO polarization procedure as implemented within the ONETEP package is now described. As discussed in Section 3.1, the problem of obtaining the ground state of a system in ONETEP involves the optimization of the NGWFs and the optimization of the density kernel for each current set of NGWFs. In ONETEP, the Stoll eigenvalue problem is not solved directly by diagonalization, but the ground state is obtained by a conjugate gradient minimization of the energy. An overview of the Stoll equations on which this procedure is based in terms of the quantities used in ONETEP is provided in Section SI.2 of the Supporting Information.

To describe the optimization of the density kernel in the SCF-MI equations within ONETEP, we begin by introducing the density kernel corresponding to the nonantisymmetrized wave function at SCF-MI iteration m as $K_{0(m)}^{A\alpha,B\beta}$. At $m = 0$, this kernel is simply set to the kernel $K_0^{A\alpha,B\beta}$ constructed in the previous frozen density analysis by direct summation of the fragment density kernels. Having obtained this kernel, SCF-MI loops are performed in order to optimize the total energy with respect to the density kernel. The electronic energy is evaluated at each iteration using the antisymmetrized representation of this kernel, $K_{\mathrm{orth}}^{A\alpha,B\beta}$. This electronic energy is expressed in terms of the quantities used in ONETEP as

$$\begin{aligned} E_{\mathrm{orth}} &= \sum_{A,B\in X}^{N_{\mathrm{frag}}} \sum_{i}^{N_{\psi(\mathrm{occ,A})}} \sum_{j}^{N_{\psi(\mathrm{occ,B})}} \sigma^{Ai,Bj} \langle \psi_{Bj} | \hat{H} | \psi_{Ai} \rangle - E_{\mathrm{DC}} \\ &= \sum_{A,B\in X}^{N_{\mathrm{frag}}} K_{\mathrm{orth}}^{A\alpha,B\beta} H_{B\beta,A\alpha} - E_{\mathrm{DC}} \end{aligned} \tag{31}$$

where E_{DC} is the double-counting correction

$$E_{\mathrm{DC}} = E_{\mathrm{H}}[n(\mathbf{r})] - E_{\mathrm{XC}}[n(\mathbf{r})] + \int \frac{\delta E_{\mathrm{XC}}[n(\mathbf{r})]}{\delta n(\mathbf{r})} n(\mathbf{r}) d\mathbf{r} \tag{32}$$

where E_{H} and E_{xc} are the Hartree and exchange and correlation energies, respectively, and where

$$n(\mathbf{r}) = 2 \sum_{A,B\in X}^{N_{\mathrm{frag}}} \phi_{A\alpha}(\mathbf{r}) K_{\mathrm{orth}}^{A\alpha,B\beta} \phi_{B\beta}(\mathbf{r}) \tag{33}$$

A single iteration of the SCF-MI kernel loop contains the following steps:

1. The fragment eigenvalue problems described in eq 25 are solved to obtain the isolated fragment MO coefficients $(M_A)_i^{\alpha}$ and subsequently $M^{A\alpha}_{Bi}$.
2. The MO overlap matrix $\sigma_{Ai,Bj}$ is constructed using eq 28 and inverted to obtain $\sigma^{Ai,Bj}$.
3. The kernel $K_{\mathrm{orth}}^{A\alpha,B\beta}$ is constructed from $M^{A\alpha}_{Bi}$ using eq 27 and $\sigma^{Ai,Bj}$.
4. The unprojected Hamiltonian $H_{\alpha\beta}$ is calculated and the energy of the ALMOs evaluated as

$$E_{\mathrm{orth}} = \sum_{A,B\in X}^{N_{\mathrm{frag}}} K_{\mathrm{orth}}^{A\alpha,B\beta} H_{B\beta,A\alpha} - E_{\mathrm{DC}} \tag{34}$$

   If variations in the density kernel are less than a specified tolerance, then exit.
5. The locally projected Hamiltonian $(H_S^A)_{A\alpha,A\beta}$ is constructed from $H_{A\alpha,B\beta}$, $S_{A\alpha,B\beta}$, $M^{A\alpha}_{Bi}$, and $\sigma^{Ai,Bj}$ using eq S9 provided in the Supporting Information.
6. $(H_S^A)_{A\alpha,A\beta}$ and $K_{0(m)}^{A\alpha,B\beta}$ are used in a conjugate gradient minimization of the Stoll locally projected energy of the fragments E_S

$$E_S = \sum_{A\in X}^{N_{\mathrm{frag}}} K_{0(m)}^{A\alpha,A\beta} (H_S^A)_{A\beta,A\alpha} - E_{\mathrm{DC}} \tag{35}$$

   where E_{DC} is evaluated using the associated fragment densities of $K_{0(m)}^{A\alpha,B\beta}$ (eq 33) to obtain the new density kernel $K_{0(m+1)}^{A\alpha,B\beta}$. The minimization approach used[61] is the method of Li, Nunes, and Vanderbilt;[62,63] wherein, the density kernel $K_0^{A\alpha,B\beta}$ is defined in terms of an auxiliary density kernel $L_0^{A\alpha,B\beta}$ by the McWeeny purifying transform[64]

$$K_0^{A\alpha,B\beta} = 3(L_0S_0L_0)^{A\alpha,B\beta} - 2(L_0S_0L_0S_0L_0)^{A\alpha,B\beta} \tag{36}$$

   where S_0 is the fragment overlap matrix constructed from the fragment overlap matrices S_A as

$$\begin{aligned} (S_0)_{A\alpha,A\beta} &= (S_A)_{A\alpha,A\beta} \\ (S_0)_{A\alpha,B\beta} &= 0|_{A\neq B} \end{aligned} \tag{37}$$

   This purification transformation is iterated as required within the LNV subroutine to yield a fully idempotent $K_0^{A\alpha,B\beta}$, equivalent to orthonormalization of the KS states within the fragments. The gradient of the total energy with respect to the auxiliary density kernel $L_0^{A\alpha,A\beta}$ is calculated as

$$\begin{aligned} \frac{\partial E_S}{\partial L_0^{A\alpha,A\beta}} &= 6(S_0L_0H_S^A + H_S^AL_0S_0)_{A\beta,A\alpha} \\ &\quad - 4(S_0L_0S_0L_0H_S^A + S_0L_0H_S^AL_0S_0 + H_S^AL_0S_0L_0S_0)_{A\beta,A\alpha} \\ \frac{\partial E_S}{\partial L_0^{A\alpha,B\beta}} &= 0|_{A\neq B} \end{aligned} \tag{38}$$

   and minimization of E_S is equivalent to minimization of E_{orth} with the ALMO constraint.

Having obtained the SCF-MI optimized density kernel $K_{\mathrm{orth}}^{A\alpha,B\beta}$, the optimization of the NGWFs proceeds as described in Section SI.3 of the Supporting Information. Once the solution to the minimization of both the NGWFs and density kernel is found and the final polarized ALMO state energy obtained, the polarization energy is then calculated using eq 4. Charge transfer is calculated using eq 5 but with omission of the ΔE_{BSSE} contribution, i.e.,

$$\Delta E_{\mathrm{CT}} = E[\Psi_X] - E[\Psi_X^{\mathrm{ALMO}}] \tag{39}$$

where

$$E[\Psi_X] \equiv E[\{K^{A\alpha,B\beta}\}, \{\phi_\alpha(\mathbf{r})\}] \tag{40}$$

$$E[\Psi_X^{\mathrm{ALMO}}] \equiv E[\{K_{\mathrm{orth}}^{A\alpha,B\beta}\}, \{(\phi_{\mathrm{ALMO}})_{A\alpha}(\mathbf{r})\}] \tag{41}$$

where $\{K^{A\alpha,B\beta}\}$ and $\{\phi_\alpha(\mathbf{r})\}$ are, respectively, the fully optimized density kernel and NGWFs of the supermolecule without the ALMO constraint (i.e., the quantities obtained from a standard singlepoint calculation); $\{(\phi_{\mathrm{ALMO}})_{A\alpha}(\mathbf{r})\}$ are the NGWFs determined by the minimization of eq S10

DOI: 10.1021/acs.jctc.6b00272
J. Chem. Theory Comput. 2016, 12, 3135−3148

provided in the Supporting Information and where $\{K_{\mathrm{orth}}^{A\alpha,B\beta}\}$ is the optimized ALMO-constrained kernel (not equal to $\{K_{\mathrm{orth}}^{A\alpha,B\beta}\}$ of eq 27). The exclusion of ΔE_{BSSE} in the ONETEP implementation is justified as it is zero due to the fact that the psinc basis set is independent of atomic positions.

4. RESULTS AND DISCUSSION

4.1. Validation of the Method. We have applied our EDA to a number of small test systems which model interactions relevant within the field of drug design in order to validate the accuracy of the ONETEP EDA against its ALMO and LMO EDA counterparts in the Q-Chem[65] and GAMESS-US[66] *ab initio* packages, respectively. These systems have been selected as they provide examples of important interactions such as hydrogen bonding, $\pi-\pi$, and dispersion.[67,68] The test systems are derived from those of our earlier work[31] in which we present a benchmark study of a selection of EDA methodologies using six test sets of small interacting molecules of biomolecular relevance.

4.1.1. Calculation Setup. Geometry optimization was performed at the PBE-D2/aug-cc-pVTZ level of theory[69,70] on all structures using the Q-Chem *ab initio* package.[65] The D2 correction for dispersion of Grimme et al.[70] was used in order to properly model the dispersion interactions especially observed in the case of the $\pi-\pi$ interacting systems. The optimized geometries of the systems studied are shown within Figures 2 and 3.

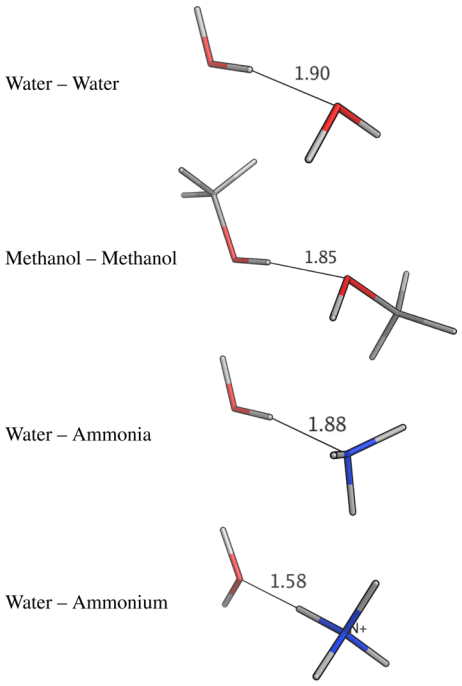


Figure 2. Test set 1 of PBE-D2/aug-cc-pVTZ geometry-optimized systems for EDA (intermolecular distances are given in Å).

EDA was performed on the geometry-optimized structures at the PBE-D2/aug-cc-pVTZ level of theory in the case of the LMO and ALMO EDAs and at the PBE-D2 level with a 1200 eV psinc basis set cutoff energy and 8 Bohr NGWF radii in the case of the ONETEP EDA. [For all calculations, excluding the benzene−ammonium system, the polarization stage NGWFs were initialized to the converged fragment NGWFs. For the benzene−ammonium system, reinitialization to guess NGWFs

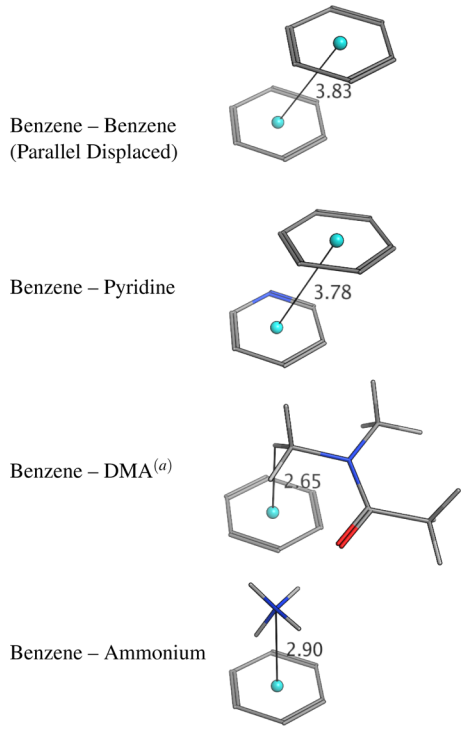


Figure 3. Test set 2 of PBE-D2/aug-cc-pVTZ geometry-optimized systems for EDA (intermolecular distances are given in Å).

at the polarization stage was required in order to prevent underconvergence of the polarization energy component.] [Grimme D3 correction results are also included for comparison in Section SI.4 of the Supporting Information. The D3 contributions were calculated using the DFT-D3 program as only the D2 correction is currently implemented in ONETEP.] Additionally, the Martyna−Tuckerman approach was used[71,72] in the ONETEP calculations to prevent interactions with periodic system images. Evidence of convergence of the calculations with respect to psinc basis set kinetic energy cutoff is provided in the Supporting Information.

4.1.2. Test Set 1: Hydrogen-Bonding Interactions. In this test set, we assess the bonding components of important hydrogen-bonded systems such as the water dimer, methanol−methanol, water−ammonia, and water−ammonium (geometries shown in Figure 2). Studies were published that investigate the covalency of hydrogen bonding in water,[73−80] and this covalency may be further analyzed by considering the strength of the charge transfer component through EDA.

The ALMO EDA frozen density component ΔE_{FRZ}, further partitioned into its constituent electrostatic, exchange, and Pauli repulsion parts, is shown in Figure 4. These constituent components follow the same definitions as in the LMO EDA and are in good agreement throughout the set. Additionally included in the ONETEP EDA is the correlation component that is required to form ΔE_{FRZ} as shown in eq 6.

We observe that the net effect of these frozen density components is almost negligible for all systems excluding the ammonium interacting case as shown by the small magnitudes of the ΔE_{FRZ} component values (+0.1 kcal/mol > ΔE_{FRZ} > −0.9 kcal/mol). Intuitively, a more favorable ΔE_{FRZ} component is expected for the water−ammonium system due to the strong electrostatic interaction of the charged ammonium molecule with the dipole of the water molecule. Our expectations are

DOI: 10.1021/acs.jctc.6b00272
J. Chem. Theory Comput. 2016, 12, 3135−3148

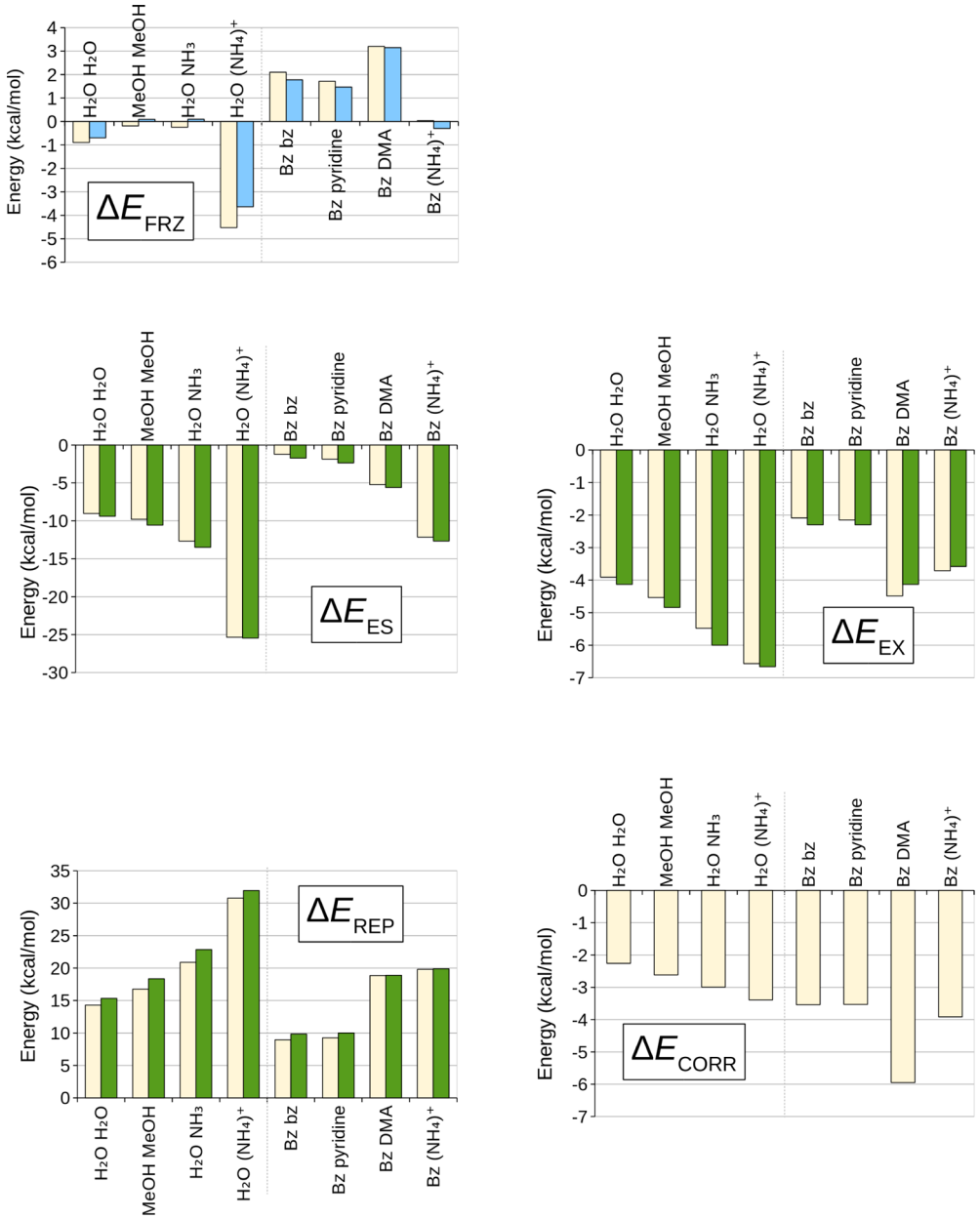


Figure 4. Frozen density analysis energy components for test sets 1 and 2. EDA component values are given in kcal/mol and were calculated with the PBE-D2 functional. ONETEP EDA calculations (yellow) were performed with a psinc basis set with a kinetic energy cutoff of 1200 eV. Q-Chem ALMO EDA calculations (blue) and GAMESS-US LMO EDA calculations (green) were performed with an aug-cc-pVTZ Gaussian basis set. The ALMO frozen density component is the sum of the electrostatic, exchange, Pauli repulsion, and correlation energy components as shown in eq 6.

confirmed by assessment of the electrostatic energy: For this system, we observe electrostatic contributions of −25.3 and −25.5 kcal/mol through the ONETEP and LMO EDAs, respectively, whereas in the case of the uncharged systems we observe $\Delta E_{ES} \leq -13.5$ kcal/mol. The increased effect of electrostatics in the water−ammonium system is also moderated significantly by a 9.9 kcal/mol (ONETEP) and 9.1 kcal/mol (LMO) increase in the Pauli repulsion term compared to the water−ammonia system.

Exchange in the ONETEP EDA is observed to increase through the test set from −3.9 kcal/mol for the water dimer to −6.6 kcal/mol for the ammonium interacting system, in good agreement with the LMO EDA results. The correlation term is also observed to increase in magnitude through the test set from −2.3 kcal/mol for the water dimer to −3.4 kcal/mol for the ammonium interacting system. This observation is rationalized through the increasing electron counts of the systems through the set and decreasing intermolecular separations.

Polarization and charge transfer energies obtained through the ALMO and ONETEP EDAs are shown in Figure 5. The LMO EDA does not attempt to distinguish polarization from charge transfer effects and instead simply describes both as an inter- and intramolecular polarization term. Again, the ALMO and ONETEP schemes show generally good agreement of the polarization term, with the strength of polarization increasing through the set to the ammonium interacting system [−10.2 kcal/mol (ONETEP) and −10.8 kcal/mol (ALMO)]. Polarization is shown to be similar in magnitude to the charge transfer contribution in the ONETEP EDA of the water dimer

DOI: 10.1021/acs.jctc.6b00272
J. Chem. Theory Comput. 2016, 12, 3135−3148

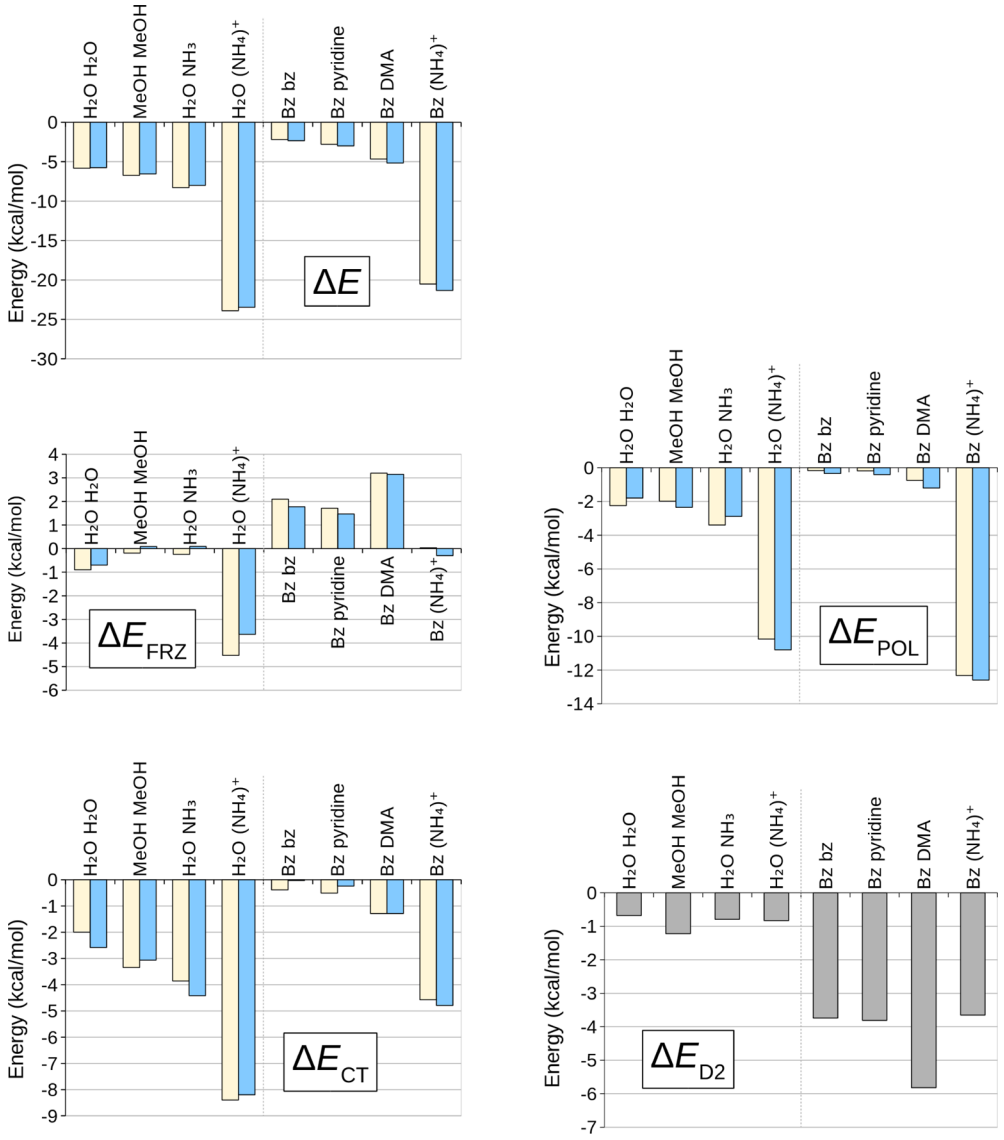


Figure 5. EDA energy components for test sets 1 and 2. EDA component values are given in kcal/mol and were calculated with the PBE-D2 functional. ONETEP EDA calculations (yellow) were performed with a psinc basis set with a kinetic energy cutoff of 1200 eV. Q-Chem ALMO EDA calculations (blue) were performed with an aug-cc-pVTZ Gaussian basis set. Values of Grimme's D2 correction for dispersion are also provided (gray).

system, with a polarization energy of −2.2 kcal/mol and a charge transfer energy of −2.0 kcal/mol. In the case of the ALMO EDA, charge transfer is observed to be much more significant (−2.6 kcal/mol) than polarization effects (−1.8 kcal/mol). Despite these differences which are expected due to the different basis sets used in the ONETEP and ALMO EDAs, the values observed and inferences drawn for these components are generally in good agreement between the schemes.

4.1.3. Test Set 2: π-Bonding Systems. In this test set, we assess the energy components of a number of molecules interacting with the π-electron cloud of a benzene molecule. Specifically, we have investigated the interactions of benzene, pyridine, dimethylacetamide (DMA), and ammonium with benzene (geometries shown in Figure 3).

We note overall good agreement between the ONETEP and LMO EDAs through the electrostatic, exchange, and Pauli repulsion energy components as displayed in Figure 4. We also observe good agreement for the ALMO EDA frozen density component that is partially constructed from these terms. Through the set, the contributions of electrostatics [−1.2 to −12.1 kcal/mol (ONETEP)], Pauli repulsion [9.0 to 19.8 kcal/mol (ONETEP)], polarization [−0.2 to −12.3 kcal/mol (ONETEP)], and charge transfer [−0.4 to −4.6 kcal/mol (ONETEP)] generally increase as expected with decreasing intermolecular separation and the presence of charged or larger molecules.

The largest correlation effects are observed for the benzene−DMA system (−5.9 kcal/mol). This compares to correlation energies of −3.5 kcal/mol for the benzene−benzene and benzene−pyridine systems and −3.9 kcal/mol for the benzene−ammonium interacting system. This difference likely arises due to the small intermolecular separation (2.65 Å) of the DMA molecule compared to the other systems within the set (2.90 to 3.83 Å), and the larger number of interacting atoms. A similar trend is observed in the contribution of exchange, and this can also be rationalized through similar considerations.

ALMO and ONETEP polarization and charge transfer contributions for test set 2 are shown in Figure 5 and also

DOI: 10.1021/acs.jctc.6b00272
J. Chem. Theory Comput. 2016, 12, 3135−3148

display good agreement. The ability of DMA to delocalize charge through its structure appears to enhance polarization to a degree, as shown by a polarization energy of −0.8 kcal/mol for the ONETEP EDA and −1.2 kcal/mol for the ALMO EDA. Interestingly, polarization is observed to increase markedly to −12.3 kcal/mol (ONETEP) and −12.6 kcal/mol (ALMO) in the case of the ammonium interacting system. This is likely the result of the placement of a positively charged molecule directly within the polarizable π-cloud of the partner benzene molecule.

Charge transfer is most significant in the DMA [−1.3 kcal/mol (ONETEP and ALMO)] and ammonium [−4.6 kcal/mol (ONETEP) and −4.8 kcal/mol (ALMO)] interacting systems. This is expected due to the ability of the DMA molecule to delocalize and stabilize incumbent charge and due to the high favorability of the positively charged ammonium molecule to accept electron density. Charge transfer is especially strong in the ammonium interacting system, reflecting the consequence of placement of a charged molecule directly within the π-cloud of the partner benzene molecule as was also observed in the case of polarization.

4.1.4. Basis Set Dependence. A notable weakness of using the SCF-MI approach to separate polarization and charge transfer effects is its significant basis set dependency. As the basis set size increases, the polarized state increasingly becomes described using basis functions that extend into the space of the partner fragment, and therefore, this component increasingly includes charge transfer contributions. This ill definition of the energy components is similar to the sensitivity[81,82] of Mulliken charges[83] to basis set size. Rigorous separation of the polarization and charge transfer terms is of key importance for obtaining an accurate chemical description, and the presence of this ill definition presents an issue. Plots demonstrating the basis set dependence of the frozen density, polarization, and charge transfer components are provided in Figure 6.

This effect may be countered, and chemically accurate contributions can be obtained by adopting a consistent basis set choice of carefully considered spatial extension. In previous work, the use of the aug-cc-pVTZ basis set for ALMO EDA calculations has been shown to offer a fair balance between accuracy and stability of terms.[80,84] Figure 6 shows that adopting between 7.5 to 8.5 Bohr NGWF radii is approximately equal to using the aug-cc-pVTZ basis set in the Gaussian code ALMO EDA. We suggest the use of 8.0 Bohr NGWF radii NGWFs in order to fairly balance the accuracy of ΔE with accuracy in partitioning polarization and charge transfer effects.

Ascribing polarization and charge transfer contributions using approaches based on the ALMO methodology is not without challenge. Methods that seek to avoid this ambiguity include the CDFT approach of Wu that exploit population analysis calculations to partition polarization and charge transfer by real space definitions[18] and also the recent fragment electric-field response functions (FERF) approach of Horn and Head−Gordon[85] that adopts the use of fragment polarization subspaces for variationally defining the polarization contribution. These approaches have shown success in isolating energy contributions with minimal basis set dependence, presenting the opportunity for the development of new EDA schemes with ever more chemically relatable energy components.

4.1.5. Exchange−Correlation Functional Dependence. It is interesting to also compare the dependence of the EDA with respect to exchange−correlation functional. We have compared the results of the water−ammonium and benzene−DMA

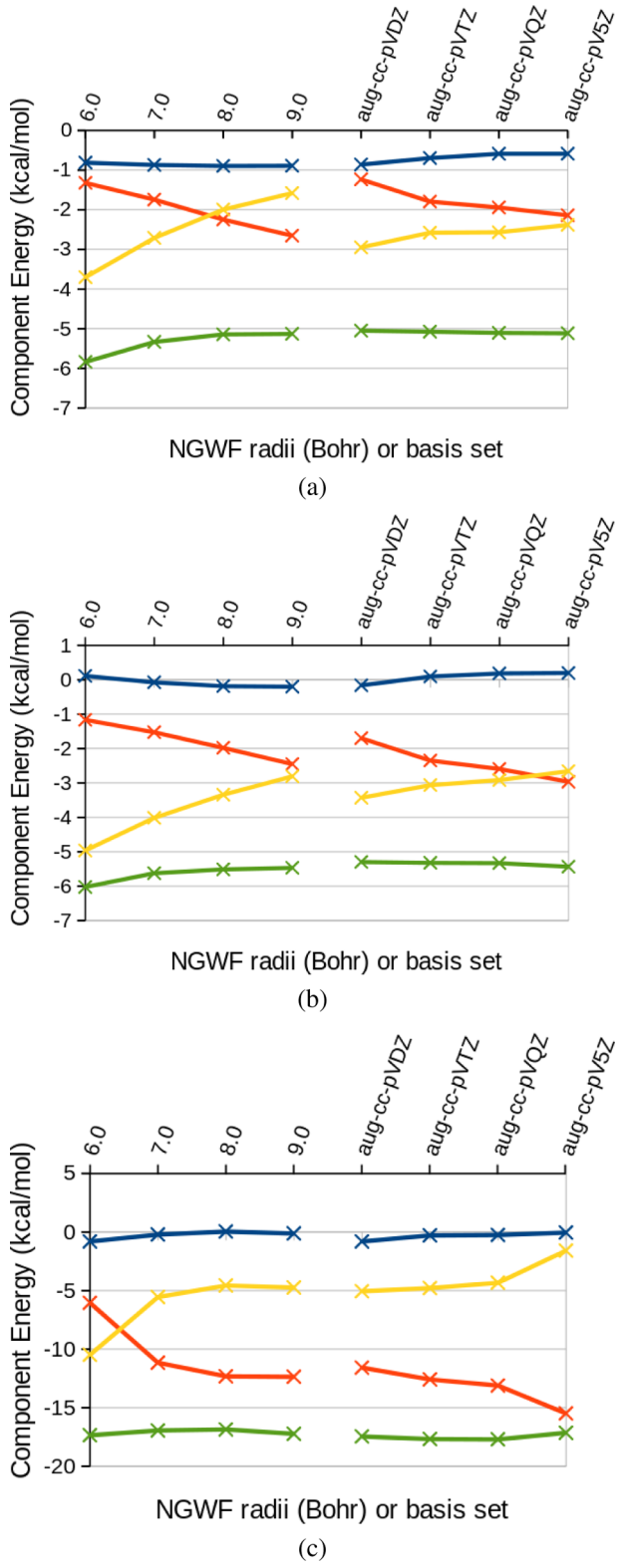


Figure 6. ONETEP (6 to 9 Bohr NGWF radii) and ALMO (aug-cc-pVDZ to aug-cc-pV5Z) EDA of PBE-D2/aug-cc-pVTZ geometry optimized (a) water dimer, (b) methanol−methanol, and (c) benzene−ammonium systems. The PBE/1200 eV level ΔE (green) is partitioned into the frozen density (blue), polarization (red), and charge transfer (yellow) energy components as described in eq 2. The D2 correction for dispersion of Grimme et al.[70] is independent of basis set size and so was not included for clarity.

DOI: 10.1021/acs.jctc.6b00272
J. Chem. Theory Comput. 2016, 12, 3135−3148

systems using the PBE[69] and BLYP[86,87] GGA exchange−correlation functionals and the LDA (with VWN correlation[88]) functional using ONETEP with a 1200 eV basis set kinetic energy cutoff. As Grimme's correction for dispersion depends upon the choice of exchange−correlation functional by a simple applied scaling factor, we have omitted this correction from our discussion. The results of this comparison are displayed in Tables 1 and 2.

Table 1. Water−Ammonium ONETEP EDA Energy Components Using a 1200 eV Basis Set Kinetic Energy Cutoff

energy term		exchange−correlation functional		
(kcal/mol)		PBE	BLYP	LDA (VWN)
FRZ		−4.5	−3.4	−8.3
	ES	−25.3	−25.7	−25.7
	EX	−6.6	−4.9	−12.9
	REP	30.8	31.4	32.4
	CORR	−3.4	−4.2	−2.1
POL		−10.2	−10.2	−10.4
CT		−8.4	−8.0	−8.6
ΔE		−23.1	−21.5	−27.4

Table 2. Benzene−Dimethylacetamide (DMA) ONETEP EDA Energy Components Using a 1200 eV Basis Set Kinetic Energy Cutoff

energy term		exchange−correlation functional		
(kcal/mol)		PBE	BLYP	LDA (VWN)
FRZ		3.2	6.8	−3.0
	ES	−5.2	−5.4	−5.5
	EX	−4.5	−0.6	−15.6
	REP	18.9	19.2	21.2
	CORR	−5.9	−6.5	−3.1
POL		−0.8	−0.9	−0.8
CT		−1.3	−1.3	−1.8
ΔE		1.2	4.7	−5.6

As we may expect, the LDA results display generally poor agreement of ΔE values with the GGA functionals. Analyzing the EDA components, we observe that this is primarily due to differences in the ΔE_{EX} components; for example, this component is 11.1 kcal/mol lower for the LDA functional than the PBE functional in the case of the benzene−DMA system. This disagreement is partially canceled by the ΔE_{REP} and ΔE_{CORR} components, giving a ΔE_{FRZ} component energy 6.2 kcal/mol lower for the LDA functional with this system. Polarization and charge transfer effects are captured with surprising accuracy by the LDA functional for the two systems studied, with ΔE_{POL} within 0.2 kcal/mol of the GGA functionals' values and ΔE_{CT} within 0.6 kcal/mol.

There is good agreement between the GGA results, indicating that energy components within this class of functionals could be expected to be comparable. As with the LDA functional, differences using the PBE and BLYP functionals are apparent within the ΔE_{FRZ} components, with the ΔE_{FRZ} term 1.1 kcal/mol lower for PBE than BLYP in the case of the water−ammonium system and 3.6 kcal/mol lower in the case of the benzene−DMA system. Polarization and charge transfer are in good agreement (0.4 kcal/mol) between the GGA functionals.

4.2. Demonstration of the Method on an Entire Protein with 4975 Atoms. *4.2.1. Calculation Setup.* Thrombin is a serine protease which plays a central role in the blood coagulation cascade.[89] Therefore, inhibition of thrombin is an approach toward the treatment and prevention of thrombosis. We have analyzed the bonding components of thrombin-inhibitor complex models with varying degrees of protein truncation using the ONETEP EDA. A brief description of these systems' preparation follows.

Four model systems were prepared from the structure of the ligand C24 (shown in Figure 7) bound to thrombin (pdb: 1KTS).[90,91] The initial thrombin-inhibitor crystal structure was capped using $COCH_3$ and $NHCH_3$ groups and protonated using the Protonate3d software.[92] [Caps were placed at the Glu1C, Tyr14J, Thr147, Gly150, and Leu264 residues.] The protein side chains, caps, and hydrogens were then optimized using the MMFF94x force field with the generalized Born implicit solvent (GBIS) model. The hirudin and its waters were subsequently removed from the system. This structure (referred to as the untruncated system) was then truncated to remove residues beyond various distances (3, 9, 15 Å) from the ligand. Caps ($COCH_3$ and $NHCH_3$ groups) were then added to the resulting protein fragments. The structures were then reprotonated, and any bond clashes introduced by capping were removed by further geometry optimization (MMFF94x, GBIS) of the caps and the hydrogens. ONETEP EDA was performed on the resulting thrombin−C24 systems using a 800 eV psinc basis set cutoff energy and 8 Bohr NGWF radii at the PBE-D2 level of theory. A summary of the protein models and EDA fragment definitions is given in Table 3, and visualization of the protein binding pocket is in Figure 8.

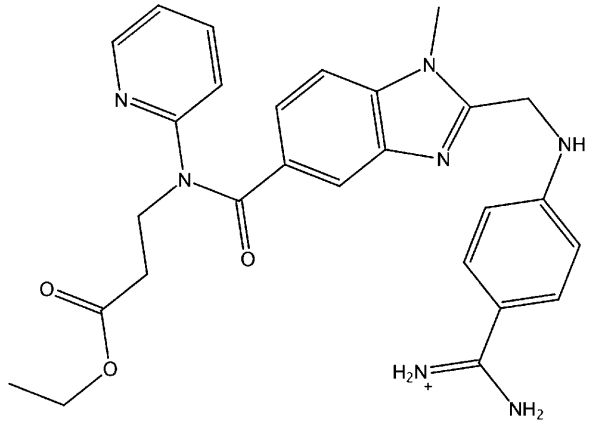


Figure 7. C24 ligand (pdb: 1KTS).

Table 3. Summary of Fragments of Thrombin−C24 Complexes Used in ONETEP EDA[a]

	fragment atom count		
protein truncation (Å)	protein and solvent	ligand	supermolecule
3	374 (−1)	67 (+1)	441 (0)
9	1328 (−1)	67 (+1)	1395 (0)
15	2621 (+1)	67 (+1)	2688 (+2)
∞	4908 (+2)	67 (+1)	4975 (+3)

[a]"∞" refers to the full, untruncated protein (charges given in parentheses).

DOI: 10.1021/acs.jctc.6b00272
J. Chem. Theory Comput. 2016, 12, 3135−3148

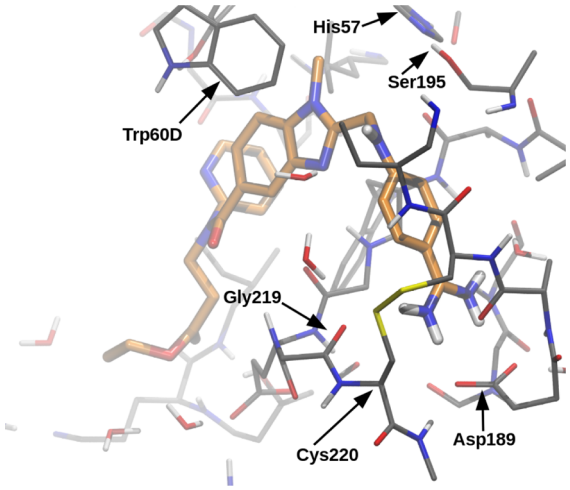


Figure 8. Binding pocket of the untruncated (4975 atom) protein system. The C24 ligand has been highlighted in orange.

4.2.2. Results and Discussion. Values of the ONETEP EDA components and their convergence with respect to the untruncated protein system are shown in Figures 9 and 10.

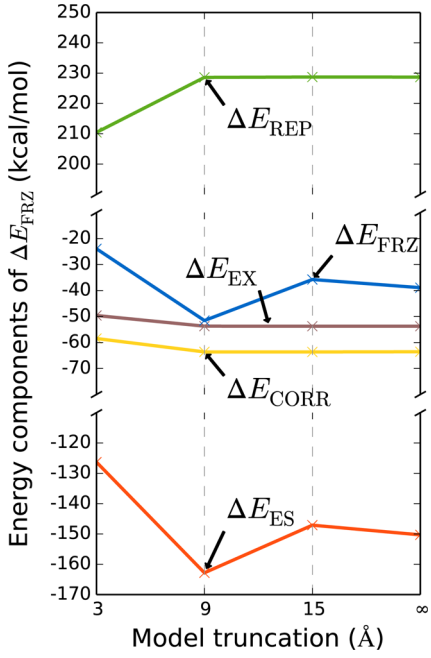


Figure 9. Frozen density analysis of the thrombin−C24 complexes calculated at the PBE-D2/800 eV level of theory (component values are given in kcal/mol). Error of the energy components is shown with respect to the full, untruncated (4975 atom) protein system. The frozen density component is formally equivalent to adding the electrostatic, exchange, Pauli repulsion, and correlation energy components as shown in eq 6.

The EDA components show a significant electrostatic contribution (−150.3 kcal/mol) to ΔE, reflecting the strong influence of the charged Asp189 residue interacting with the positively charged amidine of the ligand. The total contribution of exchange (−53.7 kcal/mol) and correlation (−63.6 kcal/mol) is shown to provide almost as much stability to binding as electrostatics. Stabilization from electrostatics, exchange, and correlation is however largely countered by steric Pauli

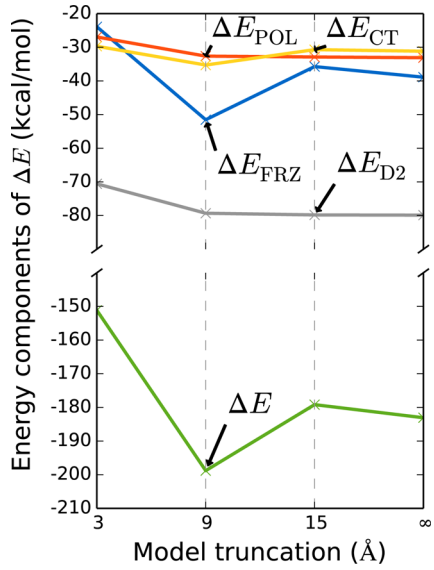


Figure 10. EDA components of the thrombin−C24 complexes calculated at the PBE-D2/800 eV level of theory (component values are given in kcal/mol). Error of the energy components is shown with respect to the full, untruncated (4975 atom) protein system. The full interaction energy ΔE is equivalent to adding the frozen density component, polarization, charge transfer, and Grimme D2 dispersion energy components.

repulsion effects (+228.7 kcal/mol). Polarization (−33.1 kcal/mol) is of similar magnitude to charge transfer (−31.2 kcal/mol), reflecting the balanced contribution of these effects to binding. A significant contribution to binding is provided by Grimme's D2 correction for dispersion (−79.9 kcal/mol), which arises due to the many long-range contributions over the large number of atoms in the protein.

Convergence of the exchange, Pauli repulsion, and correlation components is reached between 3 and 9 Å truncation, with full convergence of electrostatics observed between 15 Å truncation and the full protein. The large truncation radius required for convergence of electrostatics is likely due to differences in the charge distributions within the truncated proteins and the long-range nature of electrostatics and is interesting to observe as molecular mechanics descriptions electrostatics are often truncated at around 9−10 Å.[93,94] Polarization and dispersion are shown to converge smoothly with reduced protein truncation, with charge transfer displaying a similar convergence profile to the frozen density component and electrostatics. Again, the similar convergence profile of charge transfer may be an indication of differences in the protein charge distributions at the different truncation levels. Overall, our results indicate heavy protein truncation to have the potential to adversely affect the outcome of the EDA values. Minimal truncation of protein models is therefore shown to be necessary to ensure convergence of the EDA components.

Qualitative descriptions of intra- and interfragment delocalizations are provided by electron density difference (EDD) plots[37] (Figure 11). EDD plots are constructed as electron density differences between intermediate states of the EDA procedure. The polarization EDD is calculated by subtracting the electron density corresponding to the frozen density state from the polarized state electron density. Similarly, the charge transfer EDD is calculated by subtracting the electron density

DOI: 10.1021/acs.jctc.6b00272
J. Chem. Theory Comput. 2016, 12, 3135−3148

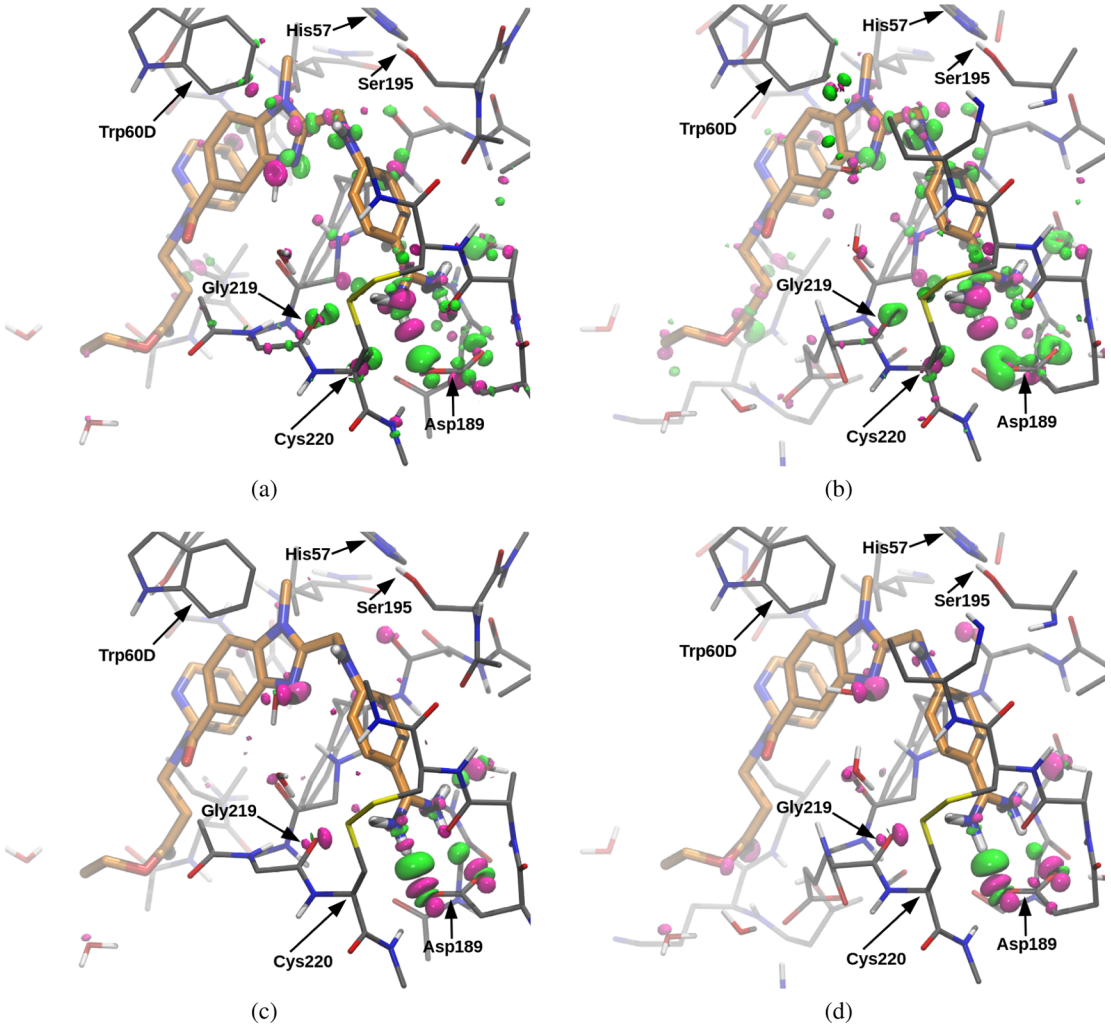


Figure 11. EDD plots of polarization for the (a) 3 Å system and (b) untruncated system and of charge transfer for the (c) 3 Å system and (d) untruncated system. The isosurface contour levels are displayed at 0.0175 electrons per Å^3 with green surfaces representing electron gain and magenta surfaces representing electron loss. The C24 ligand has been highlighted in orange. The EDD plots show the contributions of functional groups to polarization and charge transfer.

corresponding to the polarized state from that of the fully (electronically) relaxed state. Earlier work has demonstrated the high value of these plots in drug design, for example, by application to HIV-1 protease inhibitors by Hensen et al.[95] and to δ-opioid receptor binding models by Mo et al.[96] using the BLW EDA.

The EDD plots of the 3 Å truncated and untruncated systems both display the key polarization and charge transfer interactions. Overall, polarization is demonstrated as an important effect with charge transfer interactions involved in the formation of the many hydrogen bonds observed. Electron density is observed to gather in the carboxylate of Asp189 through polarization effects with bond formation to the amidine group of the C24 ligand being observed in the charge transfer EDD plot. The Gly219 and Cys220 residues of the protein are also shown to interact with the amidine group of the ligand. For these two residues, minor differences are observed between the 3 Å truncated and untruncated systems in the nature of the charge redistributions through polarization effects. In addition to interactions of the protein with the ligand, notable interactions involving water molecules are also observed, specifically with the ligand's amidine and 1-methyl-benzimidazole functional groups. These interactions highlight the importance of solvent participation in ligand−protein binding. Overall, the 3 Å and infinite cutoff EDD plots (for both polarization and charge transfer) appear qualitatively very similar. Thus, the 3 Å description is sufficient for qualitative and visual interpretations. For accurate energy components, however, the 15 Å description is needed as we saw in the previous section.

EDD plots are shown to be an important tool for visualizing the key interactions that are responsible for the polarization and charge transfer energies and allow us to associate these energies with particular functional groups. For example, it can be seen that during the polarization stage charge is redistributed in Asp189. Then, during charge transfer, charge delocalizes to the amidine group of the ligand, clearly showing the formation of bonding in this ion pair. Thus, the value of such plots can be very high for assisting in the fine-tuning of drug−protein interactions.

5. CONCLUSIONS

In this paper, we have presented a new energy decomposition analysis (EDA) scheme based on the ALMO[17] and LMO[16] approaches that decomposes ΔE into chemically meaningful

DOI: 10.1021/acs.jctc.6b00272
J. Chem. Theory Comput. 2016, 12, 3135−3148

electrostatic, exchange, correlation, Pauli repulsion, polarization, and charge transfer energy terms. This combined method, implemented in the ONETEP linear-scaling electronic structure package,[49] is a fully additive scheme with the advantage of decomposing the ALMO EDA frozen density component into chemically distinct constituents and is the first implementation of an EDA scheme in a psinc basis set code with strictly localized orbitals and plane wave basis set accuracy. With this new capability, we were able to present EDA calculations performed on systems several thousands of atoms in size, such as an entire protein−ligand complex.

We have validated the method in tests with hydrogen-bonding and π-bonding systems and have shown good agreement with the equivalent Q-Chem and GAMESS-US energy components. The recognized sensitivity of partitioning polarization and charge transfer with respect to basis set size has been shown to be present in our EDA implementation within ONETEP due to use of ALMOs. We suggest use of 8.0 Bohr NGWF radii when using the EDA within ONETEP to mitigate this effect, which we show to be approximately equivalent to use of the aug-cc-pVTZ basis set within the ALMO EDA. At this radius, convergence of the interaction energy ΔE with respect to basis set size is achieved, while also ensuring reasonable separation of charge transfer from polarization effects.

The ONETEP EDA is capable of handling large systems many thousands of atoms in size, and we have demonstrated application of the scheme to a whole biomolecular protein system of thrombin with the C24 ligand with 4975 atoms. This work has also shown that in order to ensure proper convergence of the EDA values it is a necessary requirement to have only minimal truncation of protein−ligand systems. We found that truncation of the protein system has to be at a radius of 15 Å or larger for converged EDA values. Truncation at this cutoff distance leads to a system with 2688 atoms. Smaller truncations are only useful for visual qualitative studies as we show by EDD plots.

EDD plots are a powerful tool for identifying and visualizing the key interactions that are responsible for polarization and charge transfer energies and allow us to associate these energies with particular functional groups. The combined EDA and EDD analysis is able to provide accurate quantitative and qualitative insights into the intermolecular bonding driving forces within systems of very large scale, and therefore, the method would be highly suitable for drug design applications.

Future extensions of the code for intramolecular energy component partitioning may also be made possible through the appropriate manipulation of fragment electron densities. Isolating energy components in this manner has the potential to provide individual functional group contributions to energy components, in addition to the contributions of intermolecular origin as described in this paper. Such an approach would be particularly suited to fragment-based techniques for drug design.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.6b00272.

Overview of the Stoll SCF-MI equations (SI.1). Stoll SCF-MI equations in ONETEP (SI.2). Outer loop NGWF optimization (SI.3). Calculated EDA component values (SI.4): PBE-D2/aug-cc-pVTZ (ALMO and LMO), PBE-D2/800 eV, and PBE-D2/1200 eV (ONETEP) EDA values of test set 1 (Table S1); PBE-D2/aug-cc-pVTZ (ALMO and LMO), PBE-D2/800 eV, and PBE-D2/1200 eV (ONETEP) EDA values of test set 2 (Table S2); PBE-D2/800 eV (ONETEP) thrombin-C24 EDA values (Table S3).(PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: C.Skylaris@soton.ac.uk.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

M.P. thanks the BBSRC and Boehringer Ingelheim for an industrial CASE PhD studentship (BBSRC Grant reference: BB/I015922/1). The authors are grateful for helpful discussions with Prof. Martin Head-Gordon. The calculations in this work were carried out on the Iridis4 Supercomputer of the University of Southampton and the ARCHER supercomputer via the UKCP consortium (EPSRC Grant reference: EP/K013556/1).

■ REFERENCES

(1) Schwierz, N.; Frost, C. V.; Geissler, P. L.; Zacharias, M. *J. Am. Chem. Soc.* **2016**, *138*, 527−539.
(2) Acquah, C.; Cagnetta, M.; Achenie, L. E. K.; Suib, S. L.; Karunanithi, A. T. *Ind. Eng. Chem. Res.* **2015**, *54*, 12108−12113.
(3) Clark, T.; Hennemann, M.; Murray, J.; Politzer, P. *J. Mol. Model.* **2007**, *13*, 291−296.
(4) Barik, A.; C, N.; Pilla, S. P.; Bahadur, R. P. *J. Biomol. Struct. Dyn.* **2015**, *33*, 2738−2751.
(5) Wu, E. L.; Qi, Y.; Park, S.; Mallajosyula, S. S.; MacKerell, A. D., Jr.; Klauda, J. B.; Im, W. *Biophys. J.* **2015**, *109*, 2090−2100.
(6) Ye, R.-P.; Yang, J.-X.; Zhang, X.; Zhang, L.; Yao, Y.-G. *J. Mol. Struct.* **2016**, *1106*, 192−199.
(7) Liu, X.; Dong, H.; Hu, D.; Shen, H.; Zheng, B.; Ling, B.; Bi, J. *Mol. Cryst. Liq. Cryst.* **2015**, *623*, 319−332.
(8) Tafreshi, S. S.; Roldan, A.; de Leeuw, N. H. *Surf. Sci.* **2015**, *637−638*, 140−148.
(9) Jasen, P. V.; Brizuela, G.; Padin, Z.; Gonzalez, E. A.; Juan, A. *Appl. Surf. Sci.* **2004**, *236*, 394−405.
(10) Morokuma, K. *J. Chem. Phys.* **1971**, *55*, 1236−1244.
(11) Stevens, W. J.; Fink, W. H. *Chem. Phys. Lett.* **1987**, *139*, 15−22.
(12) Bagus, P. S.; Hermann, K.; Bauschlicher, C. W. *J. Chem. Phys.* **1984**, *80*, 4378−4386.
(13) Glendening, E. D.; Streitwieser, A. *J. Chem. Phys.* **1994**, *100*, 2900−2909.
(14) Glendening, E. D. *J. Phys. Chem. A* **2005**, *109*, 11936−11940.
(15) Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211−7218.
(16) Su, P.; Li, H. *J. Chem. Phys.* **2009**, *131*, 014102.
(17) Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. *J. Phys. Chem. A* **2007**, *111*, 8753−8765.
(18) Wu, Q.; Ayers, P. W.; Zhang, Y. *J. Chem. Phys.* **2009**, *131*, 164112.
(19) Jeziorski, B.; Moszynski, R.; Ratkiewicz, A.; Rybak, S.; Szalewicz, K.; Williams, H. L. In *Methods and Techniques in Computational Chemistry: METECC-94*; Clementi, E., Ed.; STEF: Cagliari, 1993; Vol. *B*, Chapter 13, pp 79−129.
(20) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887−1930.
(21) Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325−340.
(22) Morokuma, K. *Acc. Chem. Res.* **1977**, *10*, 294−300.

DOI: 10.1021/acs.jctc.6b00272
J. Chem. Theory Comput. 2016, 12, 3135−3148

(23) Bagus, P. S.; Illas, F. *J. Chem. Phys.* **1992**, *96*, 8962−8970.
(24) Glendening, E. D. *J. Am. Chem. Soc.* **1996**, *118*, 2473−2482.
(25) Schenter, G. K.; Glendening, E. D. *J. Phys. Chem.* **1996**, *100*, 17152−17156.
(26) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735−746.
(27) Raupach, M.; Tonner, R. *J. Chem. Phys.* **2015**, *142*, 194105.
(28) Cruz Hernández, N.; Zicovich-Wilson, C. M.; Fdez. Sanz, J. *J. Chem. Phys.* **2006**, *124*, 194105.
(29) Philipsen, P. H. T.; Baerends, E. J. *J. Phys. Chem. B* **2006**, *110*, 12470−12479.
(30) te Velde, G.; Baerends, E. J. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1991**, *44*, 7888−7903.
(31) Phipps, M. J. S.; Fox, T.; Tautermann, C. S.; Skylaris, C.-K. *Chem. Soc. Rev.* **2015**, *44*, 3177−3211.
(32) Ziegler, T.; Rauk, A. *Theor. Chem. Acc.* **1977**, *46*, 1−10.
(33) Ziegler, T.; Rauk, A. *Inorg. Chem.* **1979**, *18*, 1558−1565.
(34) Ziegler, T.; Rauk, A. *Inorg. Chem.* **1979**, *18*, 1755−1759.
(35) Mitoraj, M. P.; Michalak, A.; Ziegler, T. *J. Chem. Theory Comput.* **2009**, *5*, 962−975.
(36) Mo, Y.; Gao, J.; Peyerimhoff, S. D. *J. Chem. Phys.* **2000**, *112*, 5530−5538.
(37) Mo, Y.; Bao, P.; Gao, J. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6760−6775.
(38) Wu, Q.; Van Voorhis, T. *Phys. Rev. A: At., Mol., Opt. Phys.* **2005**, *72*, 024502.
(39) Fedorov, D. G.; Kitaura, K. *J. Phys. Chem. A* **2007**, *111*, 6904−6914.
(40) Fedorov, D. G.; Nagata, T.; Kitaura, K. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7562−7577.
(41) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. *Chem. Phys. Lett.* **1999**, *313*, 701−706.
(42) Fedorov, D. G.; Kitaura, K. *J. Comput. Chem.* **2007**, *28*, 222−237.
(43) Heifetz, A.; Chudyk, E. I.; Gleave, L.; Aldeghi, M.; Cherezov, V.; Fedorov, D. G.; Biggin, P. C.; Bodkin, M. J. *J. Chem. Inf. Model.* **2016**, *56*, 159−172.
(44) Fedorov, D. G.; Kitaura, K. *J. Chem. Phys.* **2004**, *121*, 2483−2490.
(45) Bowler, D. R.; Miyazaki, T. *Rep. Prog. Phys.* **2012**, *75*, 036503.
(46) Goedecker, S. *Rev. Mod. Phys.* **1999**, *71*, 1085−1123.
(47) Khaliullin, R. Z.; Head-Gordon, M.; Bell, A. T. *J. Chem. Phys.* **2006**, *124*, 204105.
(48) Mo, Y.; Gao, J. *J. Phys. Chem. B* **2006**, *110*, 2976−2980.
(49) Skylaris, C.-K.; Haynes, P. D.; Mostofi, A. A.; Payne, M. C. *J. Chem. Phys.* **2005**, *122*, 084119.
(50) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553−566.
(51) Nagata, T.; Takahashi, O.; Saito, K.; Iwata, S. *J. Chem. Phys.* **2001**, *115*, 3553−3560.
(52) Stoll, H.; Wagenblast, G.; Preuß, H. *Theor. Chim. Acta* **1980**, *57*, 169−178.
(53) Mo, Y.; Peyerimhoff, S. D. *J. Chem. Phys.* **1998**, *109*, 1687−1697.
(54) Gianinetti, E.; Raimondi, M.; Tornaghi, E. *Int. J. Quantum Chem.* **1996**, *60*, 157−166.
(55) Gianinetti, E.; Vandoni, I.; Famulari, A.; Raimondi, M. *Adv. Quantum Chem.* **1998**, *31*, 251−266.
(56) Horn, P. R.; Mao, Y.; Head-Gordon, M. *J. Chem. Phys.* **2016**, *144*, 114107.
(57) Wu, Q. *J. Chem. Phys.* **2014**, *140*, 244109.
(58) Skylaris, C.-K.; Mostofi, A. A.; Haynes, P. D.; Diéguez, O.; Payne, M. C. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2002**, *66*, 035119.
(59) Mostofi, A. A.; Skylaris, C.-K.; Haynes, P. D.; Payne, M. C. *Comput. Phys. Commun.* **2002**, *147*, 788−802.
(60) Mostofi, A. A.; Haynes, P. D.; Skylaris, C.-K.; Payne, M. C. *J. Chem. Phys.* **2003**, *119*, 8842−8848.
(61) Haynes, P. D.; Skylaris, C.-K.; Mostofi, A. A.; Payne, M. C. *J. Phys.: Condens. Matter* **2008**, *20*, 294207.
(62) Li, X.-P.; Nunes, R. W.; Vanderbilt, D. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1993**, *47*, 10891−10894.
(63) Nunes, R. W.; Vanderbilt, D. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1994**, *50*, 17611−17614.
(64) McWeeny, R. *Rev. Mod. Phys.* **1960**, *32*, 335−369.
(65) Shao, Y.; et al. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172−3191.
(66) 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.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347−1363.
(67) Gabius, H.-J. *Pharm. Res.* **1998**, *15*, 23−30.
(68) Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2012**, *8*, 141−151.
(69) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(70) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(71) Martyna, G. J.; Tuckerman, M. E. *J. Chem. Phys.* **1999**, *110*, 2810−2821.
(72) Hine, N. D. M.; Dziedzic, J.; Haynes, P. D.; Skylaris, C.-K. *J. Chem. Phys.* **2011**, *135*, 204103.
(73) Isaacs, E. D.; Shukla, A.; Platzman, P. M.; Hamann, D. R.; Barbiellini, B.; Tulk, C. A. *Phys. Rev. Lett.* **1999**, *82*, 600−603.
(74) Hellemans, A. *Science* **1999**, *283*, 614−615.
(75) Martin, T. W.; Derewenda, Z. S. *Nat. Struct. Biol.* **1999**, *6*, 403−406.
(76) Ghanty, T. K.; Staroverov, V. N.; Koren, P. R.; Davidson, E. R. *J. Am. Chem. Soc.* **2000**, *122*, 1210−1214.
(77) Rashin, A.; Topol, I.; Tawa, G.; Burt, S. *Chem. Phys. Lett.* **2001**, *335*, 327−333.
(78) Barbiellini, B.; Shukla, A. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2002**, *66*, 235101.
(79) Beck, J. F.; Mo, Y. *J. Comput. Chem.* **2007**, *28*, 455−466.
(80) Khaliullin, R.; Bell, A.; Head-Gordon, M. *Chem. - Eur. J.* **2009**, *15*, 851−855.
(81) Åstrand, P.-O.; Ruud, K.; Mikkelsen, K. V.; Helgaker, T. *J. Phys. Chem. A* **1998**, *102*, 7686−7691.
(82) Cioslowski, J. *J. Am. Chem. Soc.* **1989**, *111*, 8333−8336.
(83) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833−1840.
(84) Azar, R. J.; Horn, P. R.; Sundstrom, E. J.; Head-Gordon, M. *J. Chem. Phys.* **2013**, *138*, 084102.
(85) Horn, P. R.; Head-Gordon, M. *J. Chem. Phys.* **2015**, *143*, 114111.
(86) Becke, A. D. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098−3100.
(87) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.
(88) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200−1211.
(89) Stubbs, M. T.; Bode, W. *Thromb. Res.* **1993**, *69*, 1−58.
(90) Hauel, N. H.; Nar, H.; Priepke, H.; Ries, U.; Stassen, J.-M.; Wienen, W. *J. Med. Chem.* **2002**, *45*, 1757−1766.
(91) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(92) Labute, P. *Proteins: Struct., Funct., Genet.* **2009**, *75*, 187−205.
(93) Izadi, S.; Aguilar, B.; Onufriev, A. V. *J. Chem. Theory Comput.* **2015**, *11*, 4450−4459.
(94) Karkov, H. S.; Krogh, B. O.; Woo, J.; Parimal, S.; Ahmadian, H.; Cramer, S. M. *Biotechnol. Bioeng.* **2015**, *112*, 2305−2315.
(95) Hensen, C.; Hermann, J. C.; Nam, K.; Ma, S.; Gao, J.; Höltje, H.-D. *J. Med. Chem.* **2004**, *47*, 6673−6680.
(96) Mo, Y.; Subramanian, G.; Gao, J.; Ferguson, D. M. *J. Am. Chem. Soc.* **2002**, *124*, 4832−4837.

DOI: 10.1021/acs.jctc.6b00272
J. Chem. Theory Comput. 2016, 12, 3135−3148