

The R^{-8} Dispersion Interaction:

Derivation and Application to the Effective Fragment Potential (EFP) Method

Peng Xu^{a*}, Samuel L. Leonard^a, William O'Brien^b, Mark S. Gordon^{a*}

a. Department of Chemistry, Iowa State University and Ames National Laboratory,
Ames, IA 50014

b. Science Undergraduate Research Internship (SULI): Department of Energy, Ames
National Laboratory, and Iowa State University

*Corresponding Authors:

mark@si.msg.chem.iastate.edu

peng@si.msg.chem.iastate.edu

Abstract

The anisotropic and isotropic R^{-8} dispersion contributions (disp8) are derived and implemented within the framework of the effective fragment potential (EFP) method, formulated with imaginary-frequency Cartesian polarizability tensors distributed at the centroids of the localized molecular orbitals (LMOs). Two forms of damping functions, intermolecular overlap-based and Tang-Toennies, are extended for disp8. To obtain LMO polarizability tensors centered at LMO centroids, an origin-shifting transformation is derived and implemented for the dipole-octopole polarizability tensor and the quadrupole-quadrupole polarizability tensor. The analytic gradient is derived and implemented for the isotropic disp8 contribution.

Relative to the previously implemented empirical EFP disp8 energy, the isotropic disp8 component of the interaction energy improves the overall agreement of the EFP dispersion energies with the symmetry adapted perturbation theory (SAPT) benchmarks, reducing the mean absolute errors (MAEs) and mean absolute percentage errors for most of the databases examined in this work. While the anisotropic disp8 can further enhance the accuracy of the EFP dispersion energy and yield smaller MAEs, significantly over-bound dispersion energies are predicted by the anisotropic disp8 when the maximum element in the intermolecular overlap matrix is greater than 0.1, possibly due to the breakdown of the approximations made in the EFP dispersion derivation at short range. For potential-energy-scan databases, the newly developed EFP dispersion model with isotropic disp8 yields overall correct curvature and good agreement with SAPT benchmarks around equilibrium and longer but overestimate the dispersion interactions at short range. While the overlap-based dispersion damping functions produce better MAEs than Tang-Toennies damping functions, further improvement is needed to better screen the large attractive dispersion energies at short range (overlap > 0.1).

I. Introduction

Dispersion, the weak interaction due to the correlated movement of electrons, impacts systems across chemistry, biology and material science. Correlated *ab initio* methods, such as Möller-Plesset 2nd order perturbation theory (MP2) and coupled cluster (CC) methods, can capture the dynamic correlation, in which dispersion is an integral part, at the cost of steep scaling, which severely limits their applicability to condensed phase, nanomaterials or bio-relevant systems. Introducing the resolution-of-identity (RI) approximation or local orbital approaches to these correlated methods and applying them within a fragmentation framework can mitigate the issue of high computational cost. However, these novel implementations may still be insufficient because of the increasing desire to use electronic structure theory to study larger and larger systems and to provide more detailed molecular modelling (e.g., explicit solvent molecules rather than a dielectric continuum approximation).

In the seminal paper of London¹, the long-range R^{-6} dispersion interaction was successfully accounted for by applying 2nd order perturbation theory to a pair of neutral atoms in S states approximated as two fluctuating dipoles. Extending the approach to higher multipoles, the fluctuating dipole-quadrupole interaction results in the R^{-8} term in the dispersion energy.^{2,3} Removing the limit to spherical systems such as atoms, the foundational work by Buckingham formulated the dispersion forces within the framework of intermolecular perturbation theory.^{4,5} The equations were made practical by introducing the Unsöld approximation⁶ and expressing the equations in terms of static molecular polarizabilities. The formulation and importance of higher order dispersion terms has been discussed for some special cases (e.g., linear molecule, tetrahedral molecule).^{4,5} In the late 1970s, Lekkerkerker *et al* extended an alternative approximate approach for dispersion forces applying the Kirkwood variational method to higher order dispersion terms.⁷⁻¹⁰ Instead of the Unsöld approximation, the Casimir-Polder formula¹¹ was employed, in which the dispersion interaction is expressed in terms of dynamic polarizabilities over the imaginary frequency range. Using spherical tensor formalism, Wormer *et. al.* developed a closed expression for the long-range interaction (including dispersion) in which the orientational dependence is simplified.¹²⁻¹⁶ Later, Wormer and coworkers also computed the dynamic polarizabilities using many-body perturbation theory.^{17,18} Due to the limited computational power at the time, early developments of dispersion interactions, especially higher order contributions, focused on noble gas atoms, diatomics and simple molecules like methane.^{4,5,7-10} Molecular polarizabilities were typically used, and certain higher-order multipolar contributions could be omitted by symmetry arguments.

With the advent of modern computers and consequently increased computational power for larger molecular systems, it became apparent that a single-site multipole expansion representation of molecules faces convergence difficulties. A distributed multipole expansion, in which the molecule is divided into regions, each described by its own multipoles, can better reflect the molecular complexity. Such distributed treatments naturally extend to the polarizabilities as well.

Several groups^{19–30} have developed dispersion models based on distributed polarizabilities. The underlying theoretical framework is perturbation theory, and the methods mainly differ in how the partition is done and how the so-called non-local polarizabilities are treated. Often, some form of Hirshfeld-type scheme^{31–33} is used to partition the molecular charge distribution into atomic contributions. Consequently, since polarizabilities involve two multipolar operators, the expansion centers of the two multipoles can coincide or differ, thereby giving rise to local and non-local polarizabilities, respectively. Some of these approaches^{22–30} were developed as dispersion corrections to density functional theory; i.e., a DFT-D type correction; hence, some parameters are fitted for different functionals.

The effective fragment potential (EFP) method was originally developed as an explicit solvation method³⁴, just for hydrated systems, and later extended to any closed-shell systems that are bound by intermolecular forces. EFP is an *ab initio* force field method, in which all of the parameters in an EFP potential are generated from preparatory *ab initio* calculations on the isolated monomer (called a fragment in EFP terminology). Once generated, the EFP fragments can interact with each other (EFP-EFP interactions) or with an *ab initio* wave function (QM-EFP interactions). EFP-EFP and QM-EFP interactions share the same theoretical foundation but differ in some key approximations made throughout the derivations. The EFP interaction energy between a pair of molecules consists of five terms: Coulombic interaction, polarization, dispersion, exchange repulsion and charge transfer. Unlike many other *ab initio* force field methods that distribute the molecular properties over atoms, localized molecular orbitals (LMOs) play a central role in the EFP method. The polarization and dispersion components are formulated using static and dynamic LMO polarizability tensors, respectively. The exchange repulsion component is derived from a truncated power expansion of the intermolecular overlap between LMOs. While mathematically equivalent to the canonical molecular orbitals (CMOs), LMOs provide a more intuitive and transferable picture for chemists.

The EFP-EFP dispersion interaction as currently implemented contains three contributions:

$$E_{EFP}^{disp} = E^{disp}(R^{-6}) + E^{disp}(R^{-7}) + \frac{1}{3}E^{disp}(R^{-6}) \quad (1)$$

where

$$E^{disp}(R^{-6}) = -\frac{3}{\pi} \sum_{k \in A}^{LMO} \sum_{j \in B}^{LMO} \frac{1}{R_{kj}^6} \sum_{n=1}^{12} W(n) \frac{2\omega_0}{(1-t_n)^2} \bar{\alpha}^k(i\omega_n) \bar{\alpha}^j(i\omega_n)$$

$$E^{disp}(R^{-7}) =$$

$$-\frac{1}{3} \frac{\hbar}{\pi} \sum_{k \in A}^{LMO} \sum_{j \in B}^{LMO} \sum_{\alpha\beta\gamma\sigma\kappa}^{x,y,z} T_{\alpha\beta}^{kj} T_{\gamma\sigma\kappa}^{kj} \sum_{n=1}^{12} W(n) \frac{2\omega_0}{(1-t_n)^2} \left[\alpha_{\alpha\gamma}^k(i\omega_n) A_{\beta,\sigma\kappa}^j(i\omega_n) - \alpha_{\beta\kappa}^j(i\omega_n) A_{\alpha,\gamma\sigma}^k(i\omega_n) \right]$$

R measures the intermolecular separation. The leading term with R⁻⁶ dependence is the contribution due to the isotropic dipole polarizabilities.³⁵ The disp7 term is an anisotropic

contribution formulated with dipole polarizabilities and dipole-quadrupole polarizabilities.³⁶ The third term in Eq. (1) can be considered to represent the higher order contributions and has been approximated as one-third of the isotropic disp6 term, an empirical approximation based on a limited number of test cases.³⁵

In order to maintain a method that is free of fitted parameters, the present work is devoted to the derivation and implementation of the analytic expressions for the disp8 contribution. Hereinafter, “old disp” refers to the dispersion energy given by Eq. (1), whereas “new disp” includes the analytic disp8 developed in this work.

In addition, note that EFP dispersion is a pairwise additive model based on intermolecular perturbation theory and does not account for many-body dispersion. In an earlier study³⁷, the many-body dispersion effects in molecular clusters bound by weak intermolecular forces were explored and were shown to be negligible. Dispersion methods such as MBD by Tkatchenko *et al.*³⁸ are based on a formulation of the random phase approximation called the adiabatic-connection fluctuation-dissipation theory, which formally gives an exact expression for the total electron correlation, including both intramolecular and intermolecular correlation.

This paper is organized as follows: Section II provides a brief overview of the EFP theory, followed by a detailed derivation of the analytic energy and gradient (in the Appendix) expressions for the R⁻⁸ dispersion contribution. Section III presents the computational details. Section IV presents and discusses the results. The conclusions are drawn, and future work is considered in Section V.

II. Theory

(A) Effective Fragment Potential Method

The effective fragment potential (EFP) method is an *ab initio* force field method that is developed to describe intermolecular interactions accurately and efficiently. As mentioned in Section I, parameters for an EFP (usually an individual molecule) are generated from a preparatory single point *ab initio* calculation, usually with Hartree-Fock (HF), coupled perturbed HF (CPHF) and time-dependent HF (TDHF) at the 6-311++G(3df,2p) basis set for a given geometry. This is accomplished at the so-called *MAKEFP* step. Detailed descriptions of each EFP term can be found elsewhere^{34-36,39-42}. Here, only a brief account of the EFP terms will be given.

The total EFP interaction energy consists of five components: Coulomb, polarization, dispersion, exchange repulsion (ExRep) and charge transfer (CT). The first three terms are electrostatic in origin (derived from a truncated multipole representation of the interacting potential operator), hence have R⁻ⁿ dependence, whereas the last two terms have an exponential dependence on the intermolecular separation R.

The Coulomb interaction is calculated as multipole-multipole interactions using the Stone distributed multipole analysis (DMA)⁴³ approach, carried out through octopole moments. These distributed multipoles are located at the atom centers and bond midpoints of an EFP fragment. The

polarization interaction, as the only EFP many-body term, is computed from the iteratively converged induced dipole moment located on each localized molecular orbital (LMO). The dispersion interaction, which is the focus of this work, is computed via the dynamic (imaginary frequency dependent) LMO polarizability tensors. The ExRep interaction energy is derived as a power series expansion of the intermolecular overlap, truncated at the quadratic term. The CT term arises from the interaction between occupied molecular orbitals on one fragment with the virtual molecular orbitals on another fragment. The virtual orbitals used in the CT term can be either the converged canonical virtual orbitals obtained after an SCF calculation of the fragment or the valence virtual orbitals (VVOs⁴⁴). The VVOs are the unoccupied complement to the valence occupied orbital space. Therefore, the number of occupied MOs + VVOs is the same as the number of minimal basis functions, thereby dramatically reducing the computational cost of the CT term relative to the full canonical orbital space.

At short range, the classical multipole expansion breaks down, and the Pauli exclusion principle becomes important; then, the use of the Hartree product instead of the antisymmetrized product for the total wave function is no longer a sound approximation. Therefore, the three R^{-n} dependent EFP components must be augmented with damping functions to obtain the correct asymptotic short-range behavior. For each of these three terms, two options for the damping functions are available. The Coulomb term can be screened by an exponential damping or an overlap-based damping; the latter utilizes the intermolecular overlap between the LMOs in the spherical Gaussian approximation.⁴⁵ For the polarization term one has a choice of exponential or Gaussian damping functions.⁴⁶ For the dispersion term, the damping functions can take an overlap-based form or a Tang-Toennies style.^{47,48}

(B) Theoretical Derivation of R^{-8} dispersion contribution

- *Molecular Model: Disp8 using molecular polarizabilities*

In Rayleigh-Schrödinger (RS) intermolecular perturbation theory, the dispersion energy is part of the second-order perturbation energy

$$E^{disp} = - \sum_{m \neq 0, n \neq 0}^{states} \frac{\langle 0_A 0_B | \hat{V} | mn \rangle \langle mn | \hat{V} | 0_A 0_B \rangle}{(E_m^A - E_0^A) + (E_n^B - E_0^B)} \quad (2)$$

where $|0_A\rangle$ and $|0_B\rangle$ represent the ground states of fragments A and B, respectively. $|m\rangle$ and $|n\rangle$ represent the excited states of fragments A and B, respectively. The E s in the denominator are the energies of the corresponding states. The summation is over all of the excited states of fragments A and B. The perturbation operator \hat{V} represents the long-range electrostatic interaction. To obtain all of the contributions for the R^{-8} order of interaction, \hat{V} must be expanded through the octopole moment,

$$\begin{aligned}
\hat{V} &= T^{AB} q^A q^B + \sum_{\alpha}^{x,y,z} T_{\alpha}^{AB} (q^A \mu_{\alpha}^B - \mu_{\alpha}^A q^B) \\
&- \sum_{\alpha\beta}^{x,y,z} T_{\alpha\beta}^{AB} \mu_{\alpha}^A \mu_{\beta}^B - \sum_{\alpha\beta\gamma}^{x,y,z} \frac{1}{3} T_{\alpha\beta\gamma}^{AB} (\mu_{\alpha}^A \theta_{\beta\gamma}^B - \theta_{\alpha\beta}^A \mu_{\gamma}^B) \\
&- \sum_{\alpha\beta\gamma\kappa}^{x,y,z} T_{\alpha\beta\gamma\kappa}^{AB} \left(\frac{1}{15} \mu_{\alpha}^A \Omega_{\beta\gamma\kappa}^B - \frac{1}{9} \theta_{\alpha\beta}^A \theta_{\gamma\kappa}^B + \frac{1}{15} \Omega_{\alpha\beta\gamma}^A \mu_{\kappa}^B \right)
\end{aligned} \tag{3}$$

where q, μ, θ and Ω are the monopole, dipole, quadrupole and octopole moments of an EFP fragment, and $\alpha, \beta, \gamma, \kappa$ run over the Cartesian coordinates x, y, z . The electrostatic T tensors in Eq. (3) are defined as follows:

$$T = \frac{1}{R} \quad O(R^{-1}) \tag{4a}$$

$$T_{\alpha} = \nabla_{\alpha} \frac{1}{R} = -\frac{R_{\alpha}}{R^3} \quad O(R^{-2}) \tag{4b}$$

$$T_{\alpha\beta} = \nabla_{\alpha} \nabla_{\beta} \frac{1}{R} = \frac{3R_{\alpha} R_{\beta} - R^2 \delta_{\alpha\beta}}{R^5} \quad O(R^{-3}) \tag{4c}$$

$$T_{\alpha\beta\gamma} = \nabla_{\alpha} \nabla_{\beta} \nabla_{\gamma} \frac{1}{R} = -\frac{15R_{\alpha} R_{\beta} R_{\gamma} - 3R^2 (R_{\alpha} \delta_{\beta\gamma} + R_{\beta} \delta_{\alpha\gamma} + R_{\gamma} \delta_{\alpha\beta})}{R^7} \quad O(R^{-4}) \tag{4d}$$

$$T_{\alpha\beta\gamma\kappa} = \nabla_{\alpha} \nabla_{\beta} \nabla_{\gamma} \nabla_{\kappa} \frac{1}{R} = \frac{\begin{bmatrix} 105R_{\alpha} R_{\beta} R_{\gamma} R_{\kappa} \\ -15R^2 \left(R_{\alpha} R_{\beta} \delta_{\gamma\kappa} + R_{\alpha} R_{\gamma} \delta_{\beta\kappa} + R_{\alpha} R_{\kappa} \delta_{\beta\gamma} \right. \\ \left. + R_{\beta} R_{\gamma} \delta_{\alpha\kappa} + R_{\beta} R_{\kappa} \delta_{\alpha\gamma} + R_{\gamma} R_{\kappa} \delta_{\alpha\beta} \right) \\ \left. + 3R^4 (\delta_{\alpha\beta} \delta_{\gamma\kappa} + \delta_{\alpha\gamma} \delta_{\beta\kappa} + \delta_{\alpha\kappa} \delta_{\beta\gamma}) \right]}{R^9} \quad O(R^{-5}) \tag{4e}$$

where R is the distance between the expansion centers (e.g., center of mass or centroids of LMOs) of the two charge distributions, and $\delta_{\alpha\beta}$ is a Kronecker delta.

Substituting Eq. (3) into Eq. (2), the dispersion interaction properly starts from the $T(R^{-3})T(R^{-3})$ combination, that is, the R^{-6} dipole-dipole term. The R^{-7} dispersion term comes from the $T(R^{-3})T(R^{-4})$ combination. The R^{-8} dispersion interaction comes from two combinations, $T(R^{-3})T(R^{-5})$ and $T(R^{-4})T(R^{-4})$. To simplify the notation, ΔE_{m0}^A is henceforth used in place of $E_m^A - E_0^A$, and ΔE_{n0}^B is similarly defined. These definitions lead to Eq. (5) below for E_8^{disp} :

$$E_8^{disp} = - \sum_{m \neq 0}^{states} \sum_{n \neq 0} \sum_{\alpha\beta\gamma\kappa\mu\nu}^{x,y,z} \frac{T_{\alpha\beta}^{AB} T_{\gamma\kappa\mu\nu}^{AB} \langle 0_A 0_B | \mu_{\alpha}^A \mu_{\beta}^B | mn \rangle \langle mn | \left(\begin{array}{c} \frac{1}{15} \mu_{\gamma}^A \Omega_{\kappa\mu\nu}^B + \frac{1}{15} \Omega_{\gamma\kappa\mu}^A \mu_{\nu}^B \\ -\frac{1}{9} \theta_{\gamma\kappa}^A \theta_{\mu\nu}^B \end{array} \right) | 0_A 0_B \rangle}{\Delta E_{m0}^A + \Delta E_{n0}^B}$$

$$\begin{aligned}
& - \sum_{m \neq 0} \sum_{n \neq 0} \sum_{\alpha\beta\gamma\kappa\mu\nu}^{states, x,y,z} \frac{T_{\alpha\beta\gamma\kappa}^{AB} T_{\mu\nu}^{AB} \langle 0_A 0_B | \left(\frac{1}{15} \mu_\alpha^A \Omega_{\beta\gamma\kappa}^B + \frac{1}{15} \Omega_{\alpha\beta\gamma}^A \mu_\kappa^B \right) | mn \rangle \langle mn | \mu_\mu^A \mu_\nu^B | 0_A 0_B \rangle}{\Delta E_{m0}^A + \Delta E_{n0}^B} \\
& - \sum_{m \neq 0} \sum_{n \neq 0} \sum_{\alpha\beta\gamma\kappa\mu\nu}^{states, x,y,z} \frac{\frac{1}{9} T_{\alpha\beta\gamma}^{AB} T_{\kappa\mu\nu}^{AB} \langle 0_A 0_B | (\mu_\alpha^A \theta_{\beta\gamma}^B - \theta_{\alpha\beta}^A \mu_\gamma^B) | mn \rangle \langle mn | (\mu_\kappa^A \theta_{\mu\nu}^B - \theta_{\kappa\mu}^A \mu_\nu^B) | 0_A 0_B \rangle}{\Delta E_{m0}^A + \Delta E_{n0}^B}
\end{aligned} \tag{5}$$

As for the derivation for disp6 and disp7, it is assumed that the fragments A and B are sufficiently separated so that short-range exchange effects can be neglected. Hence, the total wave function can be written as the Hartree product of the fragment wave functions:

$$|0_A 0_B \rangle = |0_A \rangle |0_B \rangle, |mn \rangle = |m \rangle |n \rangle \tag{6}$$

With the changes in Eq. (6), the numerator of Eq. (5) can be rewritten in terms of multipole integrals. To proceed further, note that the first and second terms in Eq. (5) are equivalent and can be combined:

$$\begin{aligned}
& E_8^{disp} \\
& = - \sum_{m \neq 0} \sum_{n \neq 0} \sum_{\alpha\beta\gamma\kappa\mu\nu}^{states, x,y,z} \left\{ \frac{2T_{\alpha\beta}^{AB} T_{\gamma\kappa\mu\nu}^{AB} \left[\begin{aligned} & \frac{1}{15} \langle 0_A | \mu_\alpha^A | m \rangle \langle 0_B | \mu_\beta^B | n \rangle \langle m | \mu_\gamma^A | 0_A \rangle \langle n | \Omega_{\kappa\mu\nu}^B | 0_B \rangle \\ & - \frac{1}{9} \langle 0_A | \mu_\alpha^A | m \rangle \langle 0_B | \mu_\beta^B | n \rangle \langle m | \theta_{\gamma\kappa}^A | 0_A \rangle \langle n | \theta_{\mu\nu}^B | 0_B \rangle \\ & + \frac{1}{15} \langle 0_A | \mu_\alpha^A | m \rangle \langle 0_B | \mu_\beta^B | n \rangle \langle m | \Omega_{\gamma\kappa\mu}^A | 0_A \rangle \langle n | \mu_\nu^B | 0_B \rangle \end{aligned} \right]}{\Delta E_{m0}^A + \Delta E_{n0}^B} \right\} \\
& - \sum_{m \neq 0} \sum_{n \neq 0} \sum_{\alpha\beta\gamma\kappa\mu\nu}^{states, x,y,z} \left\{ \frac{\frac{1}{9} T_{\alpha\beta\gamma}^{AB} T_{\kappa\mu\nu}^{AB} \left[\begin{aligned} & \langle 0_A | \mu_\alpha^A | m \rangle \langle 0_B | \theta_{\beta\gamma}^B | n \rangle \langle m | \mu_\kappa^A | 0_A \rangle \langle n | \theta_{\mu\nu}^B | 0_B \rangle \\ & - \langle 0_A | \theta_{\alpha\beta}^A | m \rangle \langle 0_B | \mu_\gamma^B | n \rangle \langle m | \mu_\kappa^A | 0_A \rangle \langle n | \theta_{\mu\nu}^B | 0_B \rangle \\ & - \langle 0_A | \mu_\alpha^A | m \rangle \langle 0_B | \theta_{\beta\gamma}^B | n \rangle \langle m | \theta_{\kappa\mu}^A | 0_A \rangle \langle n | \mu_\nu^B | 0_B \rangle \\ & + \langle 0_A | \theta_{\alpha\beta}^A | m \rangle \langle 0_B | \mu_\gamma^B | n \rangle \langle m | \theta_{\kappa\mu}^A | 0_A \rangle \langle n | \mu_\nu^B | 0_B \rangle \end{aligned} \right]}{\Delta E_{m0}^A + \Delta E_{n0}^B} \right\}
\end{aligned} \tag{7}$$

Next, the Casimir-Polder identity¹¹, $\frac{1}{A+B} = \frac{2}{\pi} \int_0^\infty \frac{AB}{(A^2 + \omega^2)(B^2 + \omega^2)} d\omega$, is introduced to rewrite the denominator as a product:

$$\frac{1}{\Delta E_{m0}^A + \Delta E_{n0}^B} = \frac{1}{\hbar} \frac{1}{\omega_{m0}^A + \omega_{n0}^B} = \frac{2}{\pi \hbar} \int_0^\infty \frac{\omega_{m0}^A \omega_{n0}^B}{[(\omega_{m0}^A)^2 + \omega^2][(\omega_{n0}^B)^2 + \omega^2]} d\omega \tag{8}$$

The ΔE_{m0}^A are expressed in terms of the transition frequency ω_{m0}^A in Eq. (8). This step allows the disp8 expression, Eq. (7), to be recast in terms of polarizabilities, which are defined as follows⁴:

$$\alpha_{\alpha\beta}(\omega) = 2 \sum_{m \neq 0}^{states} \frac{\omega_{m0} \langle m | \mu_\alpha | 0 \rangle \langle 0 | \mu_\beta | m \rangle}{\hbar(\omega_{m0}^2 - \omega^2)} \tag{9a}$$

$$A_{\alpha,\beta\gamma}(\omega) = 2 \sum_{m \neq 0}^{states} \frac{\omega_{m0} \langle m | \mu_\alpha | 0 \rangle \langle 0 | \theta_{\beta\gamma} | m \rangle}{\hbar(\omega_{m0}^2 - \omega^2)} \quad (9b)$$

$$C_{\alpha\beta,\gamma\kappa}(\omega) = \frac{2}{3} \sum_{m \neq 0}^{states} \frac{\omega_{m0} \langle m | \theta_{\alpha\beta} | 0 \rangle \langle 0 | \theta_{\gamma\kappa} | m \rangle}{\hbar(\omega_{m0}^2 - \omega^2)} \quad (9c)$$

$$D_{\alpha,\beta\gamma\kappa}(\omega) = 2 \sum_{m \neq 0}^{states} \frac{\omega_{m0} \langle m | \mu_\alpha | 0 \rangle \langle 0 | \Omega_{\beta\gamma\kappa} | m \rangle}{\hbar(\omega_{m0}^2 - \omega^2)} \quad (9d)$$

where $\alpha(\omega)$, $A(\omega)$, $C(\omega)$ and $D(\omega)$ are the dipole, dipole-quadrupole, quadrupole and dipole-octopole polarizability tensors at frequency ω , respectively. Note that the authors choose to use D for the dipole-octopole polarizability tensor, which is different from the notation in some other works. This choice of notation avoids confusion with the use of E for energy.

The disp8 interaction energy now takes the form,

$$E_8^{disp} = \sum_{\alpha\beta\gamma\kappa\mu\nu}^{x,y,z} -T_{\alpha\beta}^{AB} T_{\gamma\kappa\mu\nu}^{AB} \frac{\hbar}{\pi} \left[\begin{array}{l} \frac{1}{15} \int_0^\infty \alpha_{\alpha\gamma}^A(i\omega) D_{\beta,\kappa\mu\nu}^B(i\omega) d\omega \\ -\frac{1}{9} \int_0^\infty A_{\alpha,\gamma\kappa}^A(i\omega) A_{\beta,\mu\nu}^B(i\omega) d\omega \\ +\frac{1}{15} \int_0^\infty D_{\alpha,\gamma\kappa\mu}^A(i\omega) \alpha_{\beta\nu}^B(i\omega) d\omega \end{array} \right] - T_{\alpha\beta\gamma}^{AB} T_{\kappa\mu\nu}^{AB} \frac{\hbar}{\pi} \left[\begin{array}{l} \frac{1}{6} \int_0^\infty \alpha_{\alpha\kappa}^A(i\omega) C_{\beta\gamma,\mu\nu}^B(i\omega) d\omega \\ -\frac{1}{9} \int_0^\infty A_{\kappa,\alpha\beta}^A(i\omega) A_{\gamma,\mu\nu}^B(i\omega) d\omega \\ +\frac{1}{6} \int_0^\infty C_{\alpha\beta,\kappa\mu}^A(i\omega) \alpha_{\gamma\nu}^B(i\omega) d\omega \end{array} \right] \quad (10)$$

in which the superscripts A and B indicate that the expansion centers are located at the centers of mass of the corresponding molecules. In addition, in analogy with the R^{-6} and R^{-7} dispersion terms, the integrals in Eq. (10) are evaluated numerically via a 12-point Gauss-Legendre quadrature by a change of variable,

$$\omega = \omega_0 \frac{1+t}{1-t} \quad d\omega = \frac{2\omega_0}{(1-t)^2} dt = Z_{n0} dt \quad (11)$$

Now E_8^{disp} becomes,

$$\left[\frac{1}{\pi} \sum_{n=1}^{12} W(n) Z_{n0} \alpha_{\alpha\gamma}^A(i\omega_n) D_{\beta,\kappa\mu\nu}^B(i\omega_n) \right] \quad (12)$$

$$- \sum_{\alpha\beta\gamma\kappa\mu\nu}^{x,y,z} T_{\alpha\beta\gamma}^{AB} T_{\kappa\mu\nu}^{AB} \frac{\hbar}{\pi} \left[\begin{aligned} & \frac{1}{6} \sum_{n=1}^{12} W(n) Z_{n0} \alpha_{\alpha\kappa}^A(i\omega_n) C_{\beta\gamma,\mu\nu}^B(i\omega_n) \\ & - \frac{1}{9} \sum_{n=1}^{12} W(n) Z_{n0} A_{\kappa,\alpha\beta}^A(i\omega_n) A_{\gamma,\mu\nu}^B(i\omega_n) \\ & + \frac{1}{6} \sum_{n=1}^{12} W(n) Z_{n0} C_{\alpha\beta,\kappa\mu}^A(i\omega_n) \alpha_{\gamma\nu}^B(i\omega_n) \end{aligned} \right]$$

where $W(n)$ and t_n are the Gauss-Legendre weights and abscissas^{35,49} and the value for ω_0 is taken to be 0.3.⁵⁰ This E_8^{disp} is computed with the anisotropic polarizability tensors. Next, a formula using the isotropic forms of the polarizabilities will be introduced.

- Isotropic Approximation

The well-known C6 dispersion coefficient can be formally derived by neglecting the off-diagonal components of the dipole polarizability and averaging the diagonal components of the dipole polarizability.

$$\bar{\alpha} = \frac{\alpha_{xx} + \alpha_{yy} + \alpha_{zz}}{3} \quad (13)$$

It has been demonstrated that the isotropic averaging of the dipole-quadrupole polarizability is zero.⁴ It can be shown that the same is true for the dipole-octopole polarizability (see Appendix A). The only other surviving polarizability in Eq. (12) (within the isotropic approximation) is the quadrupole polarizability. So, the isotropic disp8 term contains only the terms involving the isotropic dipole polarizability, $\bar{\alpha}$, and the isotropic quadrupole polarizability, \bar{C} . (See Appendix B)

$$E_8^{disp,iso} = -\frac{15\hbar}{\pi} \frac{1}{(R_{AB})^8} \sum_{n=1}^{12} W(n) \frac{2\omega_0}{(1-t_n)^2} [\bar{\alpha}_A(i\omega_n) \bar{C}_B(i\omega_n) + \bar{C}_A(i\omega_n) \bar{\alpha}_B(i\omega_n)] \quad (14)$$

The isotropic dipole polarizability, $\bar{\alpha}$, is defined above and \bar{C} is the isotropic quadrupole polarizability⁵ defined as

$$\bar{C}(i\omega) = \frac{1}{5} C_{ijkl}(i\omega) \left[\frac{1}{2} (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) - \frac{1}{3} \delta_{ij} \delta_{kl} \right] \quad (15)$$

- Distributed Model: Disp8 using Localized Molecular Orbital (LMO) polarizabilities

Currently, the implemented R^{-6} and R^{-7} dispersion terms in the EFP method are computed using LMO polarizability tensors located at the centroid of each LMO. Such a distributed model allows a better description of the charge distribution than a single center model in a molecular environment. A detailed description of the distributed model has been given in previous studies^{35,36} for both the R^{-6} and R^{-7} dispersion interactions. The distributed disp8 can be derived in an analogous manner:

$$\begin{aligned}
E_8^{disp} = & - \sum_{k \in A}^{LMO} \sum_{l \in B}^{LMO} \sum_{\alpha\beta\gamma\kappa\mu\nu}^{x,y,z} T_{\alpha\beta}^{kl} T_{\gamma\kappa\mu\nu}^{kl} \frac{\hbar}{\pi} \left[\begin{aligned} & \frac{1}{15} \sum_{n=1}^{12} W(n) Z_{n0} \alpha_{\alpha\gamma}^k(i\omega_n) D_{\beta,\kappa\mu\nu}^l(i\omega_n) \\ & - \frac{1}{9} \sum_{n=1}^{12} W(n) Z_{n0} A_{\alpha,\gamma\kappa}^k(i\omega_n) A_{\beta,\mu\nu}^l(i\omega_n) \\ & + \frac{1}{15} \sum_{n=1}^{12} W(n) Z_{n0} D_{\alpha,\gamma\kappa\mu}^k(i\omega_n) \alpha_{\beta\nu}^l(i\omega_n) \end{aligned} \right] \\
& - \sum_{k \in A}^{LMO} \sum_{l \in B}^{LMO} \sum_{\alpha\beta\gamma\kappa\mu\nu}^{x,y,z} T_{\alpha\beta\gamma}^{kl} T_{\kappa\mu\nu}^{kl} \frac{\hbar}{\pi} \left[\begin{aligned} & \frac{1}{6} \sum_{n=1}^{12} W(n) Z_{n0} \alpha_{\alpha\kappa}^k(i\omega_n) C_{\beta\gamma,\mu\nu}^l(i\omega_n) \\ & - \frac{1}{9} \sum_{n=1}^{12} W(n) Z_{n0} A_{\kappa,\alpha\beta}^k(i\omega_n) A_{\gamma,\mu\nu}^l(i\omega_n) \\ & + \frac{1}{6} \sum_{n=1}^{12} W(n) Z_{n0} C_{\alpha\beta,\kappa\mu}^k(i\omega_n) \alpha_{\gamma\nu}^l(i\omega_n) \end{aligned} \right]
\end{aligned} \tag{16}$$

The isotropic version becomes,

$$E_8^{disp,iso} = - \frac{15\hbar}{\pi} \sum_{k \in A}^{LMO} \sum_{l \in B}^{LMO} \frac{1}{(R_{kl})^8} \sum_{n=1}^{12} W(n) Z_{n0} [\bar{\alpha}^k(i\omega_n) \bar{C}^l(i\omega_n) + \bar{C}^k(i\omega_n) \bar{\alpha}^l(i\omega_n)] \tag{17}$$

In Eq.(17), k and l represent the LMOs of fragments A and B, respectively. The T tensors of various ranks are defined in terms of the distance between the two LMO centroids (R_{kl}). As for the R^{-7} dispersion term, the LMO polarizability tensors are obtained via a two-step process: The first step is to compute the molecular polarizability located at the center of mass of the molecule and to partition the molecular polarizability into the LMO components. Then, an origin-shift transformation is carried out so that the expansion centers are shifted from the center of mass to the respective LMO centroids.

- Origin Dependence

For neutral molecules, the dipole polarizability is independent of the expansion center, but higher-rank polarizabilities will change with a shift of the expansion center. This is commonly referred to as origin dependence. A general form of a multipole-multipole polarizability can be defined as

$$P^{ik}(\omega) = \sum_{m \neq 0}^{states} \frac{\omega_{m0} Q^i Q^k}{[(\omega_{m0})^2 - \omega^2]} \tag{18}$$

where ω_{m0} is the transition frequency between the ground state and excited state m and Q^k symbolizes a multipole moment integral centered at origin k . For molecular polarizabilities, both i and k are the center of mass of the same molecule. However, after partitioning the molecular polarizability into LMO components, the origins i and k need to be shifted and they do not necessarily coincide at the same centroid. For the simplicity of the formulation and implementation,

it is desirable to have a local formalism ($i = k$). Currently, the ‘non-local’ polarizabilities ($i \neq k$) are neglected in the EFP distributed model. A brief discussion can be found in the Appendix.

Having finalized the form of the distributed disp8, the only missing piece is to have an origin-shifting transformation from the molecular center of mass to an LMO centroid for the quadrupole polarizabilities and the dipole-octopole polarizabilities. A quadrupole moment, θ' , with origin shifted by $-r'$, is related to the original quadrupole moment θ through,

$$\theta'_{\alpha\beta} = \theta_{\alpha\beta} - \left(\frac{3}{2} r'_\alpha \mu_\beta + \frac{3}{2} \mu_\alpha r'_\beta - \mu_\gamma r'_\gamma \delta_{\alpha\beta} \right) + \frac{3}{2} q r'_\alpha r'_\beta - \frac{1}{2} q r'^2 \delta_{\alpha\beta} \quad (19)$$

In Eq. (19), q , μ and θ are the charge, dipole and quadrupole moment at the original (un-shifted) center, and the subscripts run over x,y,z. Therefore, the origin-shifted quadrupole polarizability, C' , through some algebra, is obtained as

$$C'_{\alpha\beta,\mu\nu} = \left[\begin{aligned} & C_{\alpha\beta,\mu\nu} + \frac{3}{2} (r'_\alpha A_{\beta,\mu\nu} + r'_\beta A_{\alpha,\mu\nu}) - r'_\gamma \delta_{\alpha\beta} A_{\gamma,\mu\nu} - \frac{3}{2} (r'_\mu A_{\nu,\alpha\beta} + r'_\nu A_{\mu,\alpha\beta}) - r'_\kappa \delta_{\mu\nu} A_{\kappa,\alpha\beta} \\ & + \frac{9}{4} (r'_\alpha r'_\mu \alpha_{\beta\nu} + r'_\alpha r'_\nu \alpha_{\beta\mu} + r'_\beta r'_\mu \alpha_{\alpha\nu} + r'_\beta r'_\nu \alpha_{\alpha\mu}) \\ & - \frac{3}{2} (\delta_{\alpha\beta} r'_\gamma r'_\mu \alpha_{\gamma\nu} + \delta_{\alpha\beta} r'_\gamma r'_\nu \alpha_{\gamma\mu} - \delta_{\mu\nu} r'_\alpha r'_\kappa \alpha_{\beta\kappa} - \delta_{\mu\nu} r'_\beta r'_\kappa \alpha_{\alpha\kappa}) + r'_\gamma r'_\kappa \alpha_{\gamma\kappa} \delta_{\alpha\beta} \delta_{\mu\nu} \end{aligned} \right] \quad (20)$$

In Eq. (20) C , A and α without the superscript prime indicate the original polarizabilities with the origin at the center of mass. Note that the integrals involving the charge q vanish due to orthogonality of the ground and excited states, regardless of the charge.

The octopole moment origin shift transformation, although considerably more complicated algebraically, can be achieved as follows:

$$\Omega'_{\alpha\beta\gamma} = \Omega_{\alpha\beta\gamma} + \left\{ -\frac{5}{3} (r'_\alpha \theta_{\beta\gamma} + r'_\beta \theta_{\alpha\gamma} + r'_\gamma \theta_{\alpha\beta}) + \frac{2}{3} r'_\kappa (\theta_{\kappa\alpha} \delta_{\beta\gamma} + \theta_{\kappa\beta} \delta_{\alpha\gamma} + \theta_{\kappa\gamma} \delta_{\alpha\beta}) \right\} + \left\{ \begin{aligned} & \frac{5}{2} (r'_\alpha r'_\beta \mu_\gamma + r'_\alpha \mu_\beta r'_\gamma + \mu_\alpha r'_\beta r'_\gamma) - \mu_\kappa r'_\kappa (r'_\alpha \delta_{\beta\gamma} + r'_\beta \delta_{\alpha\gamma} + r'_\gamma \delta_{\alpha\beta}) \\ & + \frac{1}{2} q r'^2 (r'_\alpha \delta_{\beta\gamma} + r'_\beta \delta_{\alpha\gamma} + r'_\gamma \delta_{\alpha\beta}) - \frac{1}{2} r'^2 (\mu_\alpha \delta_{\beta\gamma} + \mu_\beta \delta_{\alpha\gamma} + \mu_\gamma \delta_{\alpha\beta}) - \frac{5}{2} q r'_\alpha r'_\beta r'_\gamma \end{aligned} \right\} \quad (21)$$

Then the shifted dipole-octopole polarizability, D' , becomes,

$$D'_{\sigma,\alpha\beta\gamma} = E_{\sigma,\alpha\beta\gamma} - \frac{5}{3} (r'_\alpha A_{\sigma,\beta\gamma} + r'_\beta A_{\sigma,\alpha\gamma} + r'_\gamma A_{\sigma,\alpha\beta}) + \frac{2}{3} r'_\kappa (A_{\sigma,\kappa\alpha} \delta_{\beta\gamma} + A_{\sigma,\kappa\beta} \delta_{\alpha\gamma} + A_{\sigma,\kappa\gamma} \delta_{\alpha\beta}) + \frac{5}{2} (r'_\beta r'_\gamma \alpha_{\sigma\alpha} + r'_\alpha r'_\gamma \alpha_{\sigma\beta} + r'_\alpha r'_\beta \alpha_{\sigma\gamma}) - r'_\kappa (r'_\alpha \delta_{\beta\gamma} + r'_\beta \delta_{\alpha\gamma} + r'_\gamma \delta_{\alpha\beta}) \alpha_{\sigma\kappa} - \frac{1}{2} r'^2 (\alpha_{\sigma\alpha} \delta_{\beta\gamma} + \alpha_{\sigma\beta} \delta_{\alpha\gamma} + \alpha_{\sigma\gamma} \delta_{\alpha\beta}) \quad (22)$$

The dipole-octopole polarizability tensor and quadrupole-quadrupole polarizability tensor are generated using the same procedure as that used for the dipole polarizability and dipole-quadrupole polarizability tensors⁵¹; that is, the dynamic analog of the coupled perturbed Hartree-Fock (CPHF) equations⁵²,

$$(H^{(2)}H^{(1)} - (iv)^2)Z = -H^{(2)}P \quad (23)$$

In Eq. (23), $H^{(1)}$ is the real orbital Hessian matrix

$$H_{aibj}^{(1)} = (\varepsilon_a - \varepsilon_i)\delta_{ab}\delta_{ij} + 4(ai|bj) - (ab|ij) - (aj|bi) \quad (24)$$

and $H^{(2)}$ is related to the magnetic field response of the system

$$H_{aibj}^{(2)} = (\varepsilon_a - \varepsilon_i)\delta_{ab}\delta_{ij} + (ab|ij) - (aj|bi) \quad (25)$$

in which i and j are the occupied molecular orbitals (MOs), a and b are the virtual MOs. ε_i and ε_a are the corresponding orbital energies. The matrix P in Eq. (23) provides the perturbation field. P is the dipole moment integral in the MO basis for the dipole-octopole polarizability and the quadrupole moment integral in the MO basis for the quadrupole-quadrupole polarizability. Then the response vectors Z obtained by solving Eq. (23) are used to form the desired polarizabilities:

$$C_{\alpha,\beta\gamma\kappa}(i\nu) = \sum_{ai} \frac{2}{3} \langle a | \hat{\theta}_{\alpha\beta} | i \rangle Z_{\gamma\kappa}^{ai}(i\nu) \quad (26)$$

$$D_{\alpha,\beta\gamma\kappa}(i\nu) = \sum_{ai} 2 \langle a | \hat{\Omega}_{\beta\gamma\kappa} | i \rangle Z_{\alpha}^{ai}(i\nu) \quad (27)$$

- Damping function

Both the anisotropic and isotropic disp8 in Eq. (16) and Eq. (17) are implemented, augmented with a damping function to ensure correct asymptotic behavior. Both disp6 and disp7 have their associated damping functions f_6 and f_7 . The damping functions for disp8 are natural extensions of the disp6 and disp7 contributions. Currently, two types of damping functions exist for the EFP dispersion interaction: overlap-based^{46,53} and Tang-Toennies^{46,47}, both of which are extended to account for the disp8 term.

Overlap-based:
$$f_8^S(i, j) = 1 - S_{ij}^2 \sum_{n=0}^8 \frac{(-2\ln|S_{ij}|)^{n/2}}{n!} \quad (28a)$$

where S_{ij} is the overlap between LMOs i and j .

Tang-Toennies:
$$f_8^{TT}(i, j) = 1 - \left(\sum_{n=0}^8 \frac{(bR)^n}{n!} \right) e^{-bR} \quad (28b)$$

where R is the distance between the centroids of LMOs i and j , and S_{ij} is an element of the interfragment overlap matrix. The parameter b is (somewhat arbitrarily) chosen to be 1.5 for consistency with disp6 and disp7, which have demonstrated reasonable performance^{36,46}.

- Analytic gradient of the isotropic disp8

The fully analytical gradient for the isotropic disp8 was derived and implemented to allow geometry optimizations and molecular dynamics simulations. (See derivations in Appendix).

All of the new developments are implemented in the GAMESS electronic structure package.⁵⁴⁻⁵⁶

III. Computational Details

The databases used as benchmarks in this work are listed below in Table 1. The first eight contain dimers at a single configuration, while the last four databases are potential energy scans. For all the systems, the EFP potentials were generated at the HF/6-311++G(3df,2p) level of theory using the monomer structure in the dimer as provided by the Biofragment Database (BFDdb).⁵⁷ The benchmarking dispersion energies are obtained from the same portal. The levels of symmetry adapted perturbation theory (SAPT) used for the benchmarks are summarized in Table 2.

Table 1. The databases examined in this work. The first eight sets contain dimers at a single configuration (around equilibrium) while the last four are interaction energy scans along certain reaction coordinates.

Database	Description
A24	A set of 24 small bimolecular complexes
BBI25	Peptide backbone-backbone complexes
HSG	Bimolecular complexes from protein-indinavir reaction site 1 HSG
JSCH	Nucleobase pairs
S22	Organic bimolecular complexes
S66	Organic bimolecular complexes
SSI100	Peptide sidechain-sidechain complexes
UBQ	Bimolecular complexes from native 1UBQ protein fold
Potential energy scan	
ACHC	Interaction energy curve for adenine-cytosine stacked nucleobases through 6 translations and rotations
HBC6	Dissociation curves of doubly H-bonded bimolecular complexes
NBC10	Dissociation curves of dispersion-bound bimolecular complexes
S22by7	Dissociation curves of organic bimolecular complexes

Table 2. The level of SAPT theory used for the databases.

Database	Dispersion Energies
A24	SAPT2+3(CCD)/aug-cc-pVTZ
BBI25	SAPT2+/aug-cc-pVDZ
HSG	SAPT2+3(CCD)/aug-cc-pVTZ
JSCH	sSAPT0/jun-cc-pVDZ
S22	SAPT2+3(CCD)/aug-cc-pVTZ
S66	sSAPT0/jun-cc-pVDZ
SSI100	SAPT2+/aug-cc-pVDZ
UBQ	sSAPT0/jun-cc-pVDZ
ACHC	sSAPT0/jun-cc-pVDZ
HBC6	SAPT2+3(CCD)/aug-cc-pVTZ
NBC10	SAPT2+3(CCD)/aug-cc-pVTZ

The mean absolute error (MAE) and mean absolute percentage error are computed for the database containing dimers at a single configuration, and are defined as:

$$\text{mean absolute error} = \frac{\sum_{i=1}^N |\text{Calculated}_i - \text{Benchmark}_i|}{N}$$

$$\text{mean absolute percentage error} = \frac{\sum_{i=1}^N \left| \frac{\text{Calculated}_i - \text{Benchmark}_i}{\text{Benchmark}_i} \right| \times 100}{N}$$

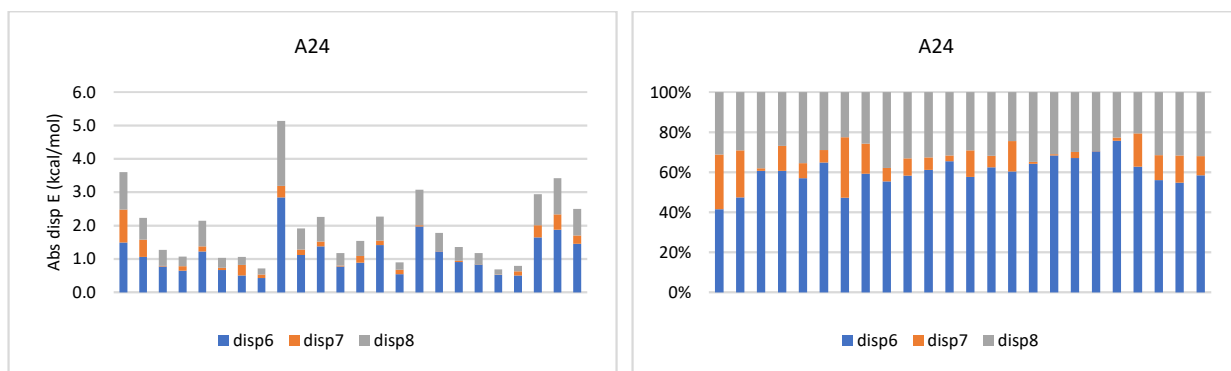
The signed error is defined as $\text{Calculated} - \text{Benchmark}$. Therefore, for the dispersion energy, underestimation by EFP is indicated by positive errors while negative errors mean overestimation by EFP.

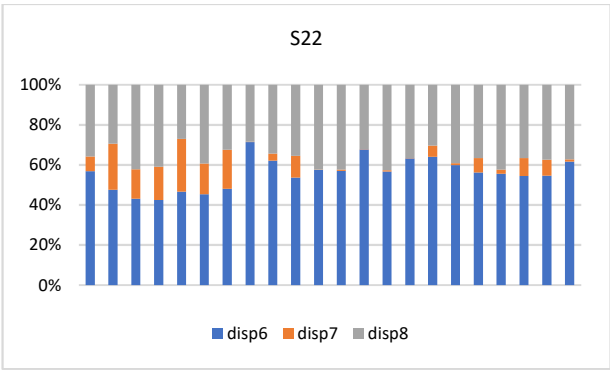
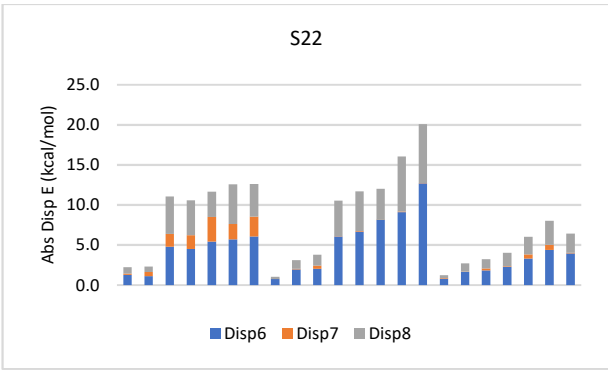
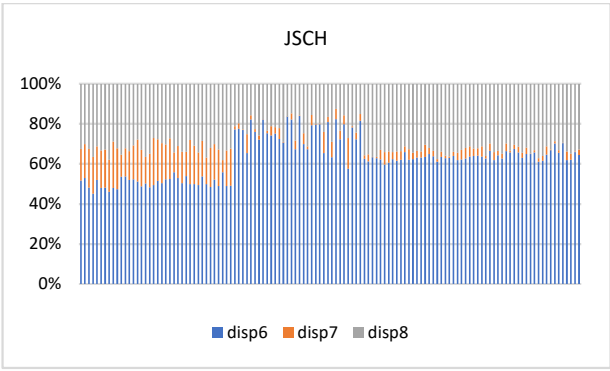
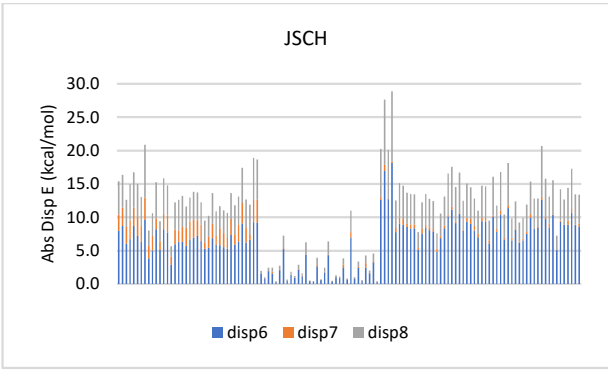
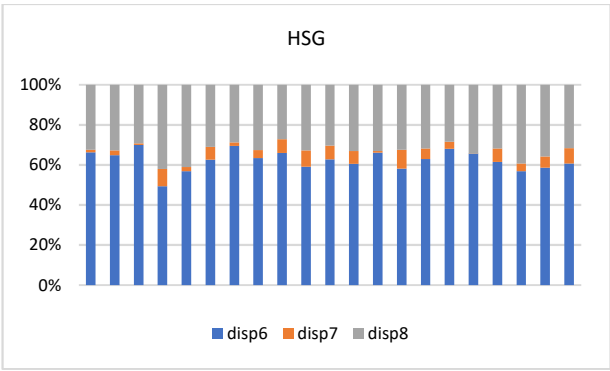
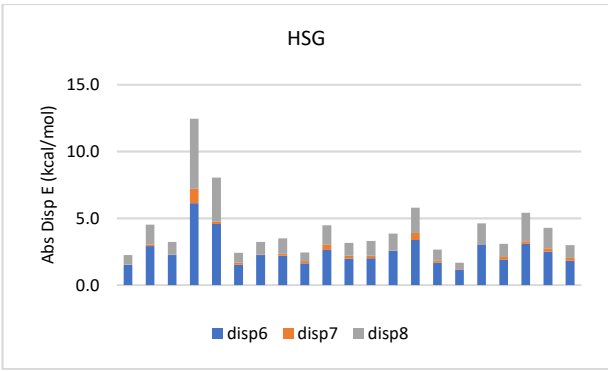
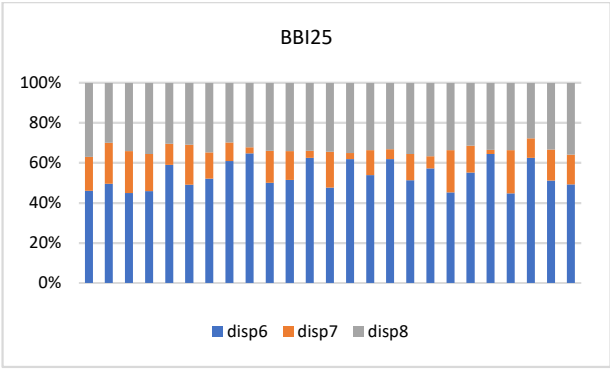
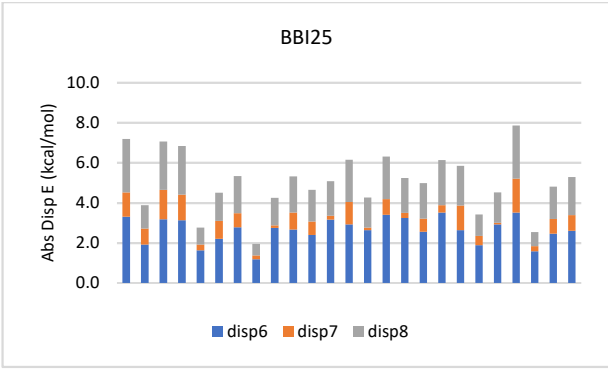
IV. Results and Discussion

A) Dimers around equilibrium configuration

The total dispersion energies predicted by EFP, “old disp” and “new disp” using either isotropic or anisotropic disp8, individual dispersion contributions (disp6, disp7 and disp8), as well as the three components in disp8 due to dipole-quadrupole polarizabilities, dipole-octopole polarizabilities and quadrupole-quadrupole polarizabilities for all the databases containing dimers at a single configuration are summarized in Supporting Information.

The magnitudes and percentages of disp6, disp7 and isotropic disp8 in the total dispersion energy for the single configuration databases are shown below in Figure 1. Disp6 is the leading term in the dispersion energy (Eq. 1), contributing 40% - 80% of the EFP total dispersion energy. Disp7 is negligible in many cases, but can be significant (~ 20%), particularly for hydrogen-bonded dimers. The isotropic disp8 contribution is smaller but still substantial (~ 20% - 40%).





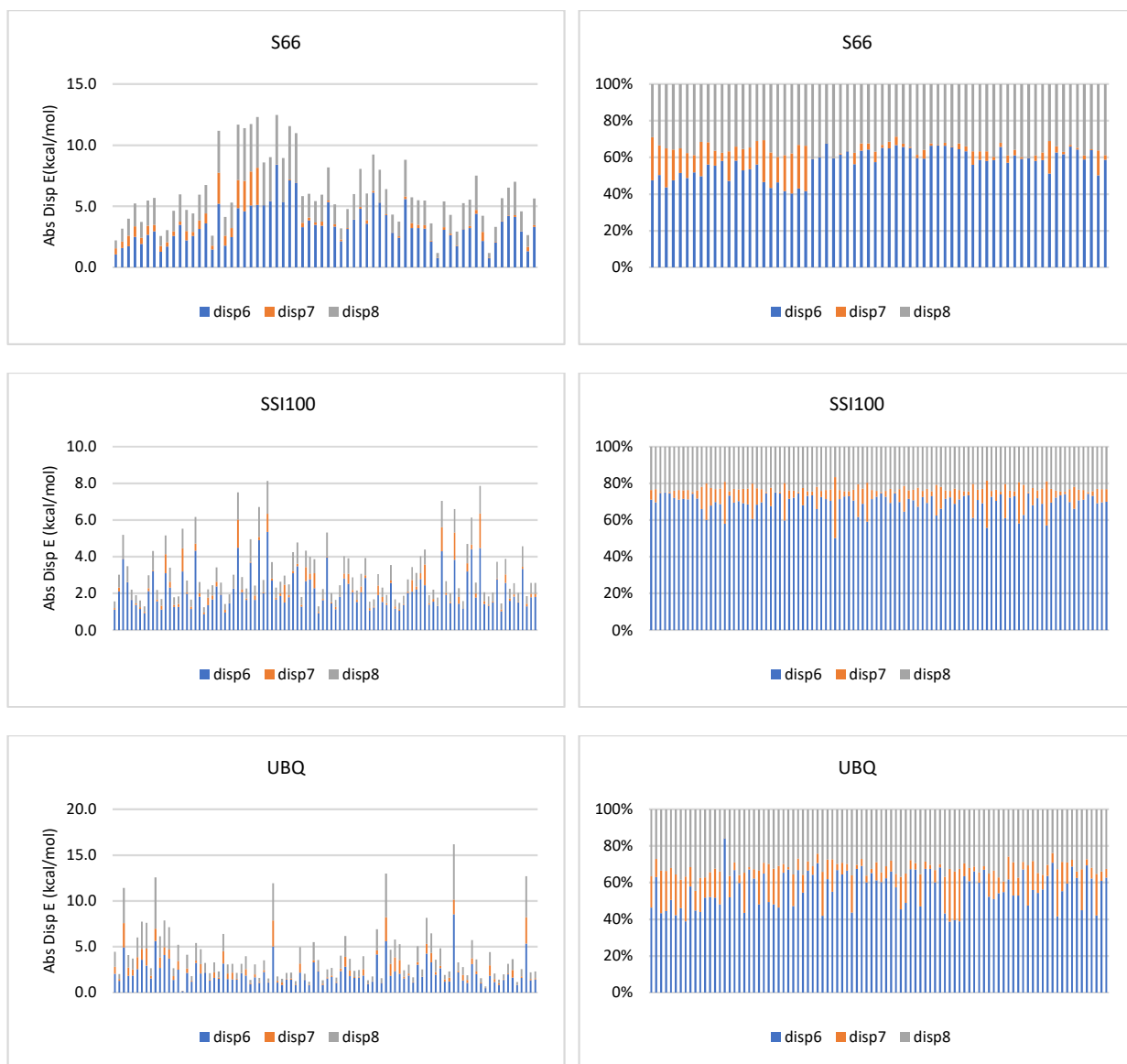


Figure 1. The magnitudes of the dispersion energy components (disp6, disp7 and isotropic disp8) (left) and the percentage of disp6, disp7 and disp8 in the total dispersion (right) for the A24, BBI25, HSG, JSCH, S22, S66, SSI100 and UBQ databases. Overlap-based damping functions are used. The bars are ordered from left to right according to the naming convention of each database.

The mean absolute errors (MAEs) and the mean absolute percentage errors of the EFP total dispersion energies for the eight databases containing dimers at a single configuration are summarized in Table 3. Of the eight databases, five yield MAEs less than 1.5 kcal/mol and show progressive improvement from the approximate disp8 (old disp) to the isotropic disp8 (new iso disp8) to the anisotropic disp8 (new aniso disp8). The three databases (JSCH, S66 and UBQ) that produce worse performance than the “old Disp8 all use sSAPT0/jun-cc-pVDZ for benchmarking

calculations, which consider only intermolecular perturbation, whereas other databases use at least the SAPT2+ level that contains intramolecular correlation including intramolecular dispersion.⁵⁸ Therefore, SAPT2+ and above provide better benchmarking values for the total dispersion interactions. The better agreement of the analytic disp8 (either isotropic or anisotropic) may be partially due to the over-estimation of disp8. As will be discussed in relation to the signed errors (Figure 2) and potential energy scans (for example, Figure 6), the overlap-based damping function, in its current form, may not provide sufficient screening at short range or for molecules with large intermolecular overlaps, thereby possibly contributing to an over-estimation of disp8.

Table 3. The mean absolute error (MAE) in kcal/mol and the mean absolute percentage error of EFP total dispersion energies for the eight databases. “Old” refers to approximating disp8 as one third of isotropic disp6. “New” dispersion computes the disp8 using either the isotropic or the anisotropic formulation.

Dataset	MAE (kcal/mol)			Mean Absolute Percentage Error (%)		
	Old	New (iso disp8)	New (aniso disp8)	Old	New (iso disp8)	New (aniso disp8)
A24	0.540	0.380	0.280	28.95	21.95	15.23
BBI25	1.694	0.864	0.436	36.31	18.83	9.22
HSG	0.993	0.509	0.497	23.06	13.42	12.89
JSCH	1.019	1.652	2.638	15.39	19.42	32.25
S22	1.696	1.320	1.215	25.47	18.02	15.42
S66	0.546	1.084	1.760	13.64	26.83	47.32
SSI100	0.664	0.269	0.395	19.44	8.16	10.48
UBQ	0.427	0.493	1.272	17.07	20.59	44.70

Table 4. Mean absolute error (MAE) of S22 dispersion energies compared against SAPT2+3(CCD)/aug-cc-pVTZ benchmarks. The EFP dispersion energies are screened with either overlap-based damping functions or Tang-Toennies damping functions. The dispersion expansion is truncated at R^{-6} (Disp6), R^{-7} (Disp6+7) and R^{-8} terms, where “Old Disp” includes all three contributions and uses an approximate disp8, and “New Disp” uses the isotropic disp8 developed in this work.

MAE (kcal/mol)	Disp6	Disp6+7	Old Disp	New Disp (iso disp8)
Overlap-based damping	2.46	3.07	1.69	1.31
Tang-Toennies damping	3.32	3.56	2.42	2.21

From Table 4, one can observe that (1) on top of disp6, adding only the disp7 correction is not sufficient, and, in fact, lowers the overall accuracy of the dispersion energies; (2) the even-powered higher order contribution (with disp8 being the leading term) is necessary; (3) the newly developed analytic disp8 provides the best accuracy and (4) the intermolecular overlap-based damping shows

better performance than Tang-Toennies style damping. So, for all subsequent discussion, the numerical results are obtained using the overlap-based damping functions.

The anisotropic disp8 (Eq. 16) consists of contributions from various types of polarizability tensors, namely, the dipole-quadrupole polarizability tensor (E8DQ), the dipole-octopole polarizability tensor (E8DO) and the quadrupole-quadrupole polarizability tensor (E8QQ). For the databases studied in this work (see Supplementary Information), it is found that the E8DQ contribution is always small, generally less than 5 % of the total dispersion energy and often even smaller than 1%. The E8QQ contribution is always attractive while the E8DO contribution can be attractive or repulsive, with a tendency to be attractive for hydrogen-bonded systems and repulsive for dispersion dominated systems.

Taking the S22 database as an example, the dispersion energies for individual dimers are listed in Table 5 and the statistics metrics (MAE and mean absolute percentage error) are summarized in Table 6. Several interesting observations can be made. (1) The old dispersion model significantly underestimates the dispersion energy for hydrogen-bonded dimers (3.95 kcal/mol MAE) but yields most likely fortuitously excellent agreement for dispersion dominated dimers (0.37 kcal/mol MAE). (2) The new dispersion model with isotropic disp8 predicts larger dispersion interactions than the old model, which means improved results for hydrogen-bonded dimers but larger errors for dispersion dominated dimers. (3) The dispersion model with anisotropic disp8 yields even larger dispersion energies for hydrogen-bonded dimers due to the attractive E8DO contribution. On the other hand, the repulsive E8DO contribution for the dispersion dominated dimers reduces the total dispersion interactions and provides much better agreement with available benchmarks.

Table 5. The total dispersion energies (kcal/mol) of the S22 database predicted by SAPT2+3(CCD)/aug-cc-pVTZ and EFP. The R^{-8} dispersion energy contribution to the total EFP dispersion is computed from the approximate disp8 (old), or the isotropic disp8 (iso disp8) or the anisotropic disp8 (aniso disp8).

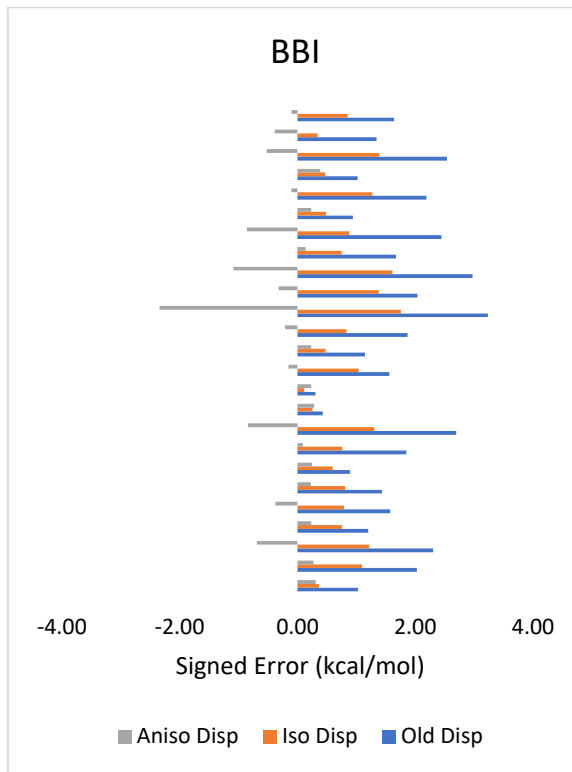
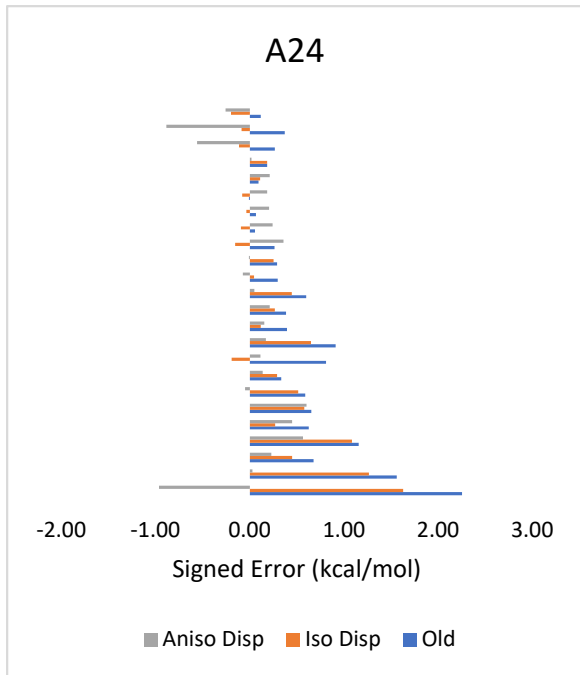
		SAPT ⁵⁸	Old Disp	New Disp (iso disp8)	New Disp (aniso disp8)
Hydrogen-bonded dimers					
Ammonia	ammonia	-2.1651	-1.5195	-1.8974	-1.7085
Water	water	-2.5274	-0.9208	-1.2324	-2.5379
formic acid	formic acid	-10.3917	-4.722	-7.8015	-16.4291
Formamide	formamide	-8.2317	-4.2207	-7.0517	-13.0303
Uracil	uracil	-9.9104	-4.156	-5.5077	-9.7334
Pyridoxine	aminopyridine	-10.3660	-5.6366	-8.6772	-14.4304
Adenine	thymine	-10.8311	-5.6227	-7.7021	-14.1919
Dispersion dominated dimers					
Methane	methane	-0.9482	-0.9605	-1.0091	-0.7881
Ethene	ethene	-2.5288	-2.6620	-3.0875	-2.4475
Benzene	methane	-2.8155	-2.2718	-2.9335	-3.2809
Benzene	benzene	-8.0158	-8.0494	-10.5018	-8.8422

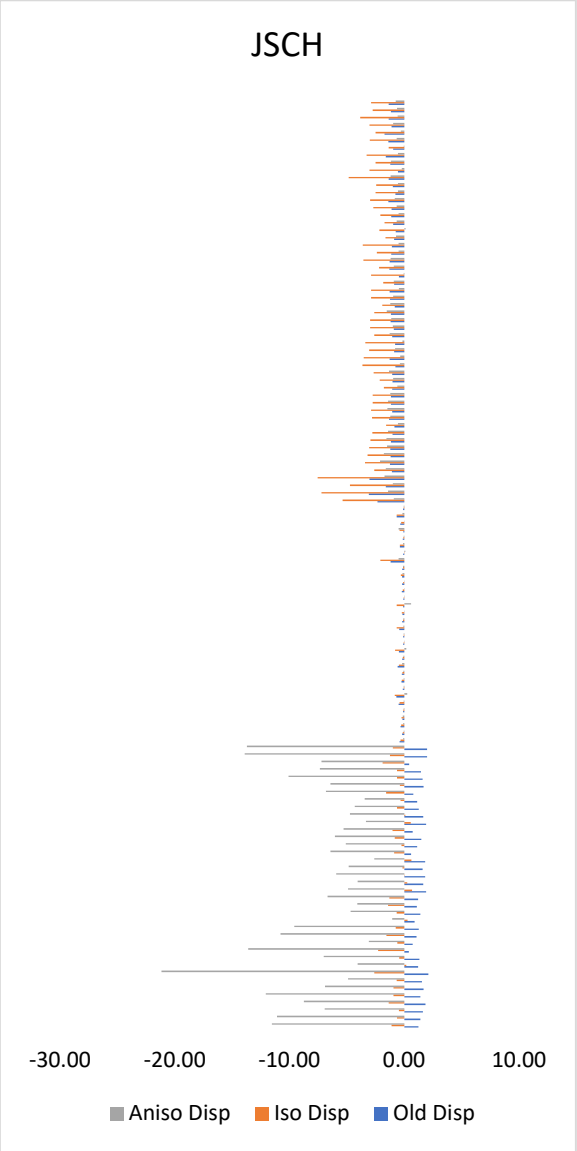
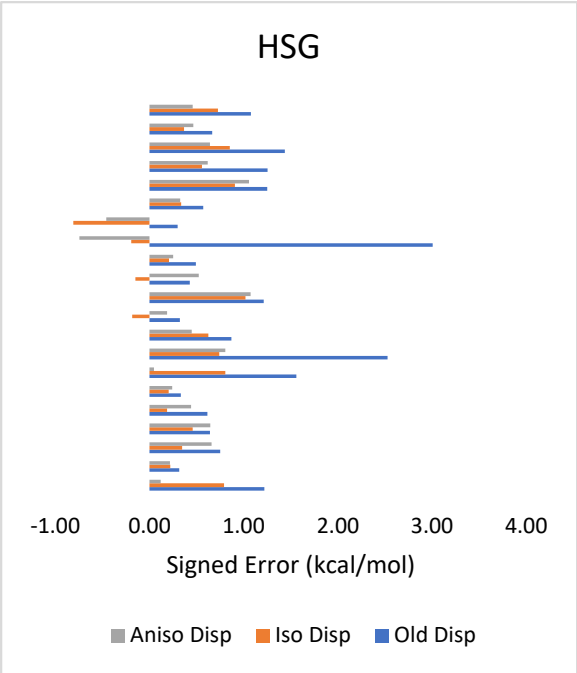
Pyrazine	pyrazine	-8.6512	-8.7689	-11.4931	-9.3063
Uracil	uracil	-11.9863	-10.8432	-12.0031	-10.6106
Indole	benzene	-11.8612	-12.1984	-16.0207	-12.8879
Adenine	thymine	-17.6415	-16.9627	-20.0753	-16.3613
Mixed dimers					
Ethene	ethyne	-1.4094	-0.9614	-1.0708	-1.2634
Benzene	water	-2.9557	-2.1166	-2.6331	-2.6609
Benzene	ammonia	-2.8436	-2.169	-2.7432	-2.9098
Benzene	HCN	-3.7630	-3.0642	-4.0268	-3.8127
Benzene	benzene	-4.7454	-3.8198	-4.9338	-5.1757
Indole	benzene	-6.6948	-5.2097	-6.7606	-6.8901
Phenol	phenol	-6.7825	-5.1683	-6.2454	-6.0133

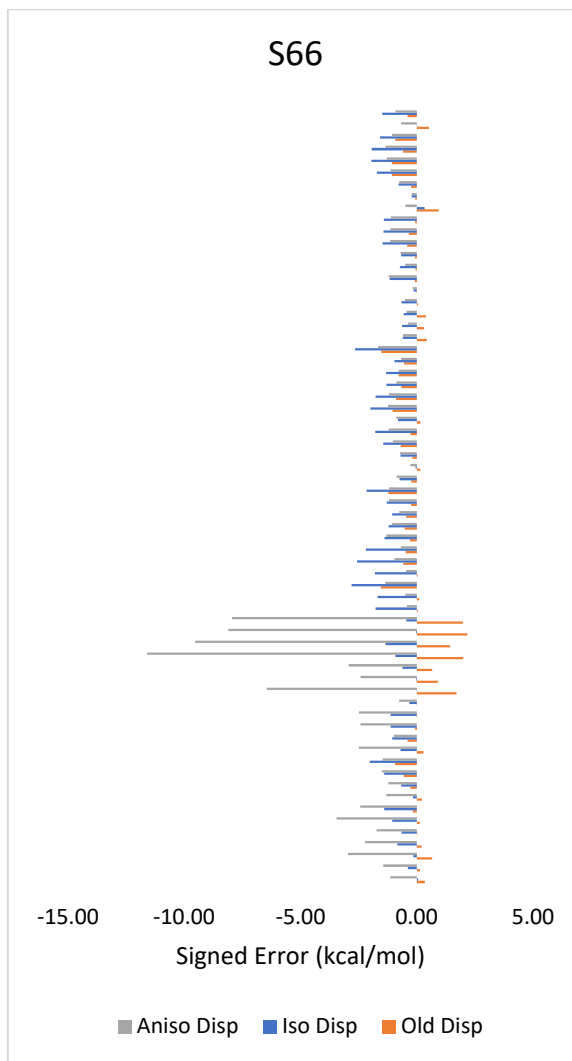
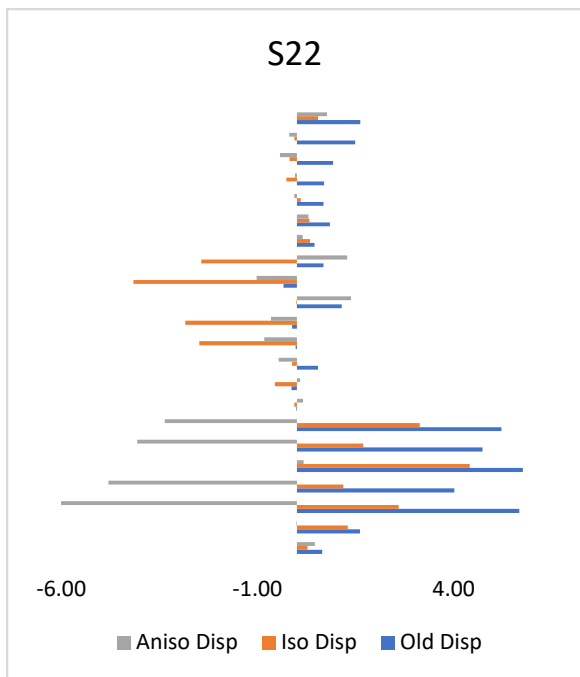
Table 6. The mean absolute error (kcal/mol) and mean absolute percentage error (%) for the full set, H-bonded subset, dispersion dominated subset and mixed subset of S22 database.

Dispersion	MAE (kcal/mol)			Mean absolute percentage error (%)		
	Old	New (iso)	New (aniso)	Old	New (iso)	New (aniso)
S22						
Full set	1.70	1.32	1.21	25.47	18.02	15.41
H-Bonded	3.95	2.08	2.70	49.78	27.50	29.99
Disp. bound	0.37	1.58	0.73	5.48	18.20	10.24
Mixed	0.96	0.26	0.28	23.99	8.34	6.76

The signed errors for all databases are plotted in Figure 2. For many dimers, there is an error reduction from the old dispersion (blue bar) to the new iso disp (orange bar) and then to the new aniso disp (grey bar). However, as discussed for the S22 set, the anisotropic disp8 is largely attractive for hydrogen-bonded dimers and yields substantial errors in some cases. Interestingly, if the dimer has a value greater than 0.1 for the maximum element in the intermolecular overlap matrix, it almost always results in a large over-bound dispersion energy. Such large errors could be the result of two approximations in the derivation of the EFP dispersion model. One is the use of a classical multipole expansion for the interaction operator. The second is the assumption that the total wave function can be represented by the Hartree product of monomer wave functions. An intermolecular overlap matrix element of 0.1 (increasingly likely at shorter inter-fragment distances) might suffice for the breakdown of one or both approximations. It is noted that at short range (overlap > 0.1), other dispersion terms (disp6 and disp7) should also be affected and the use of analytic disp8, either isotropic or anisotropic, should still reduce errors. However, fortuitous cancellation of errors when using the empirical (old) disp8 may appear to make the results worse, as exemplified in the S22 database.







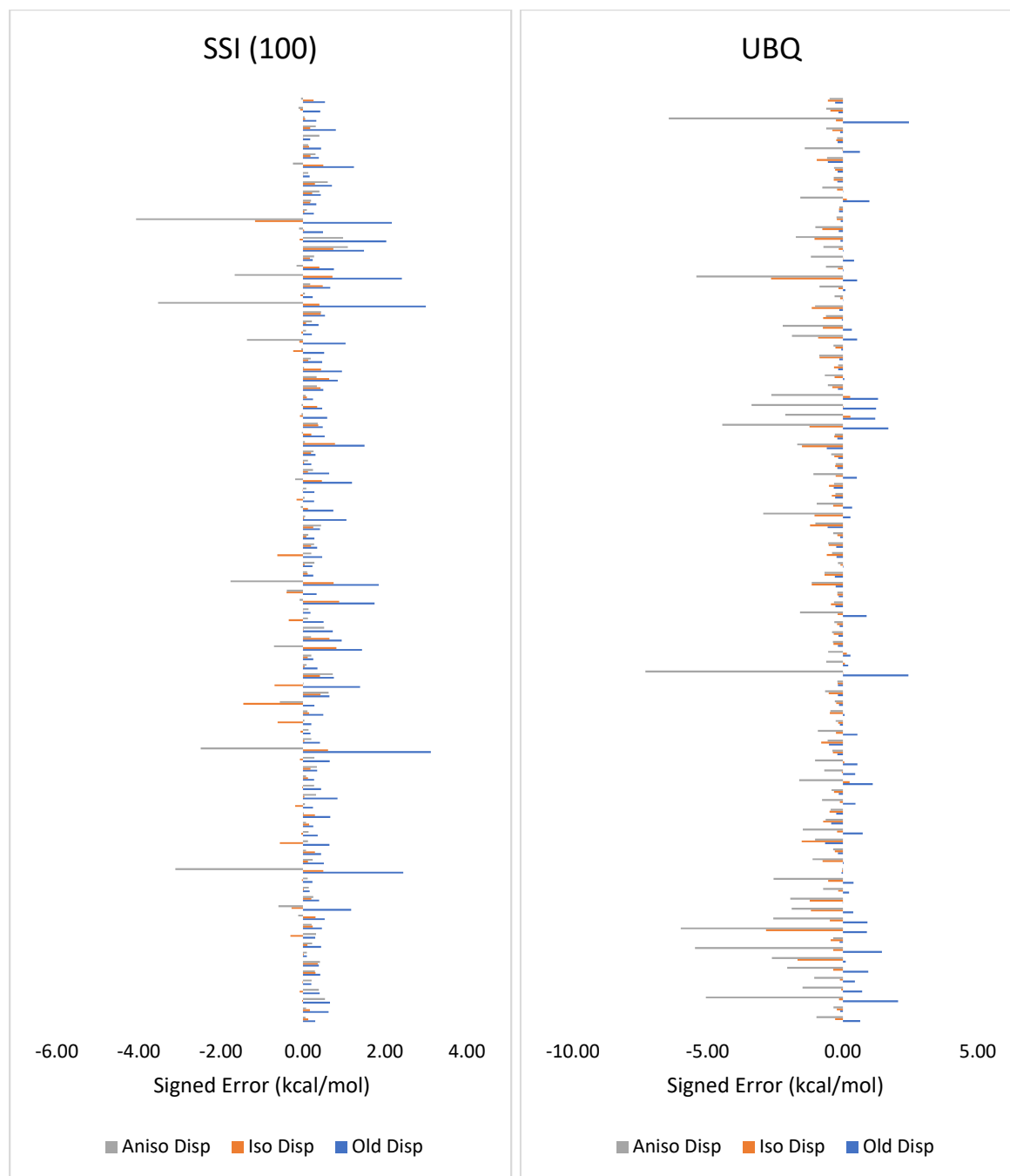


Figure 2. Signed errors (kcal/mol) for the total dispersion energies predicted by EFP relative to SAPT benchmarks. The disp8 contribution in the EFP total dispersion energy is calculated either as anisotropic disp8 (grey), isotropic disp8 (orange) or 1/3 of disp6 (blue). An underestimation by EFP is indicated by positive errors while negative errors mean overestimation by EFP. The bars are ordered from bottom to top according to the naming convention of each database.

B) Potential energy scans

Four potential energy scan databases are examined here (Figures 3-6). ACHC is composed of potential energy scans of adenine-thymine along six coordinates (Figure 3). Three translations are Rise (+Z), Slide (+Y), and Shift (+X) and three counterclockwise rotations includes Twist (+Z), Roll (+Y) and Tilt (+X). HBC6 consists of hydrogen-bonded bimolecular complexes of formic acid (FaOO), formamide (FaON) and formamidine (FaNN) monomers in homogeneous dimers or heterogenous dimers, along the hydrogen bonding dissociation coordinate (Figure 4). NBC10 contains a collection of dispersion bound bimolecular complexes including the sandwich (BzBz_S), T-shaped (BzBz_T), 3.2, 3.4 and 3.6 Å separated (BzBz_PD32, BzBz_PD34, BzBz_PD36) configurations of the benzene dimer, benzene with hydrogen sulfide (BzH2S), benzene with methane (BzMe), methane dimer (MeMe), the antiparallel sandwich (PyPy_S2) pyridine dimer and a T-shaped (PyPy_T3) pyridine dimer (Figure 5). S22by7 contains dimers in S22 database separated at 0.7, 0.8, 0.9, 1.0, 1.2, 1.5 and 2.0 X of the equilibrium separation (Figure 6).

The dispersion energies predicted by the new dispersion model with isotropic disp8 are compared with SAPT benchmarks and plotted in Figures 2-5 for the ACHC, HBC6, NBC10 and S22by7 databases, respectively. A general observation is that the new dispersion model is able to predict overall correct curvatures of the potential energy surfaces with good agreement with SAPT around equilibrium and longer separations, but tends to over-bind substantially at shorter range. For the significantly over-bound dimers, the maximum element in the intermolecular overlap matrix can be greater than 0.2 or even 0.3 whereas at equilibrium this value is generally smaller than 0.1. Also, the ACHC and S22by7 benchmarks are obtained at a lower level of accuracy, sSAPT0/jun-cc-pVDZ level while HBC6 and NBC10 dispersion benchmarks are obtained at SAPT2+3(CCD)/aug-cc-pVTZ. At shorter range, the importance of intramolecular correlation, including dispersion, is expected to increase.

ACHC

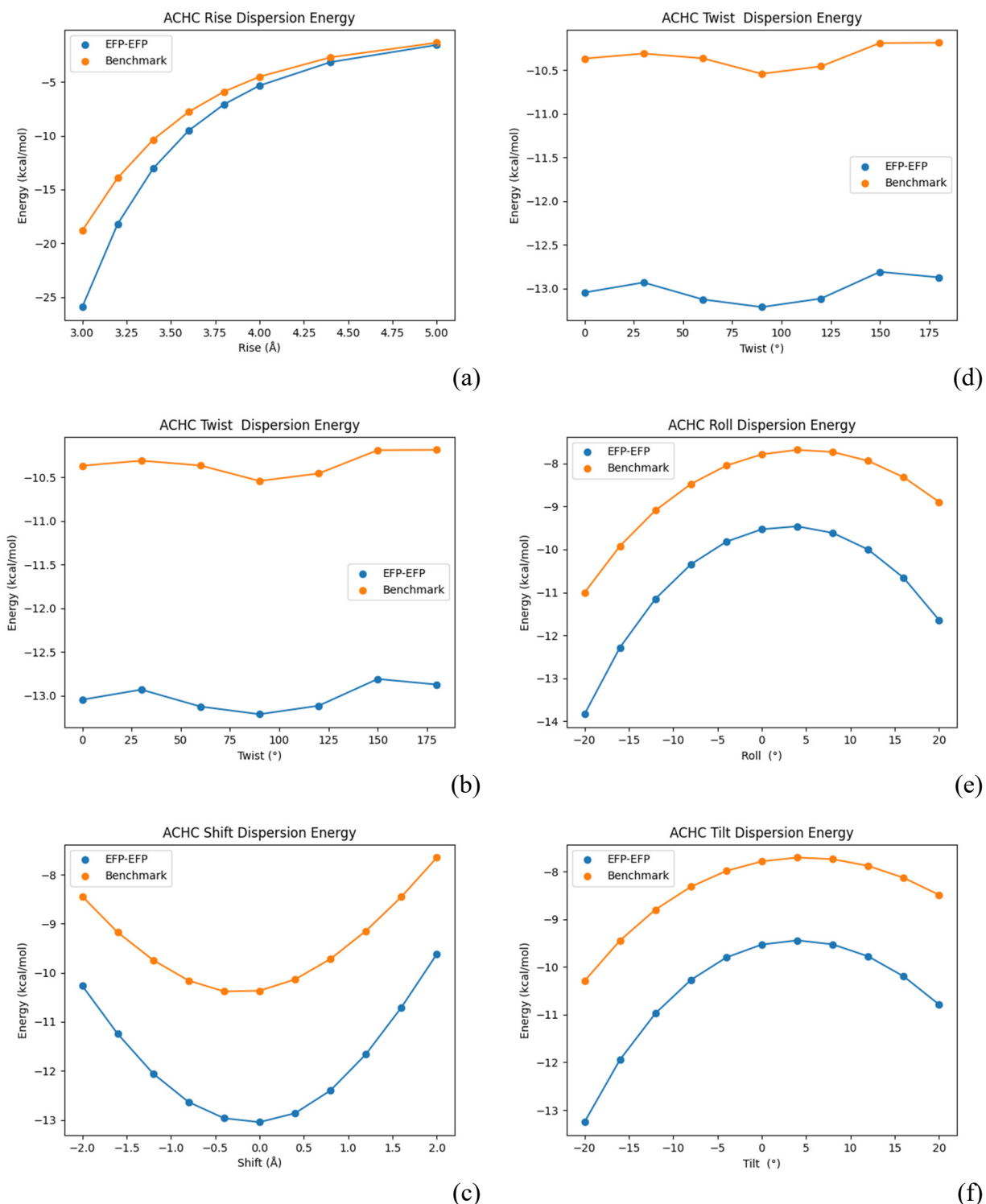
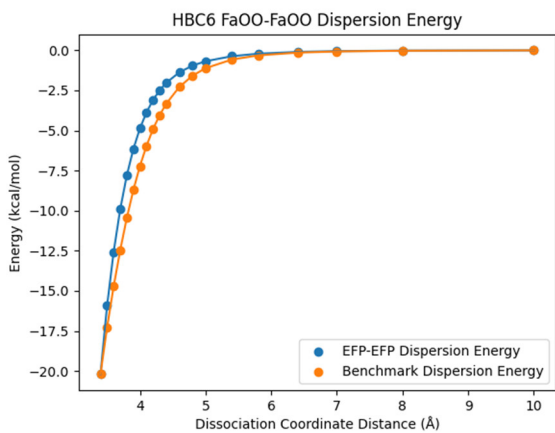
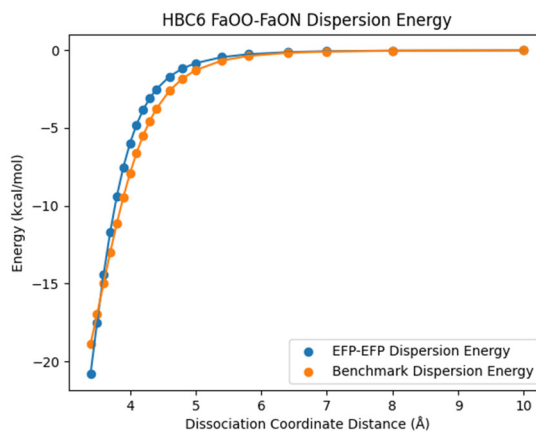


Figure 3. The dispersion energies of the ACHC database predicted by new EFP model with isotropic disp8 (blue) and SAPT (orange). Three translations are (a) Rise (+Z), (b) Slide (+Y), and (c) Shift (+X) and three counterclockwise rotations are (d) Twist (+Z), (e) Roll (+Y) and (f) Tilt (+X).

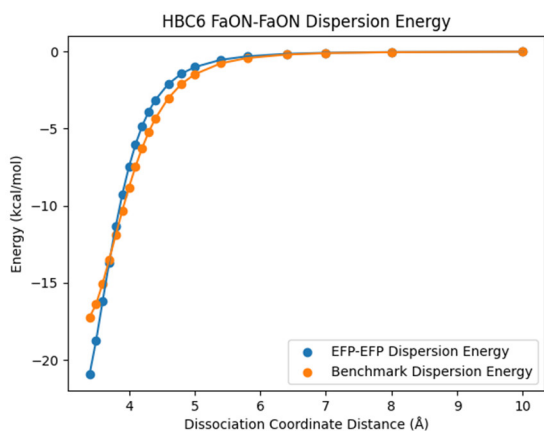
HBC6



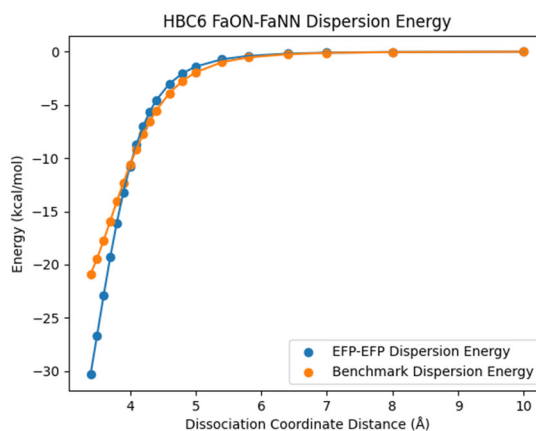
(a)



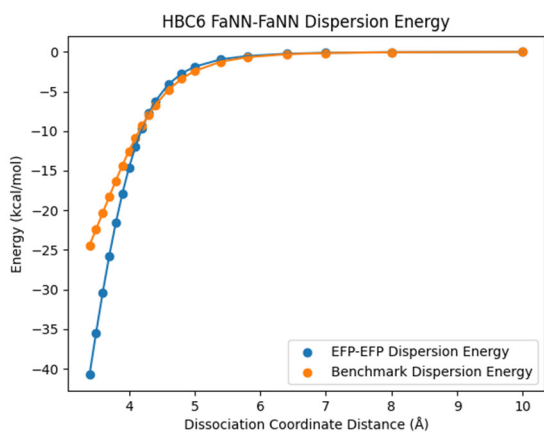
(d)



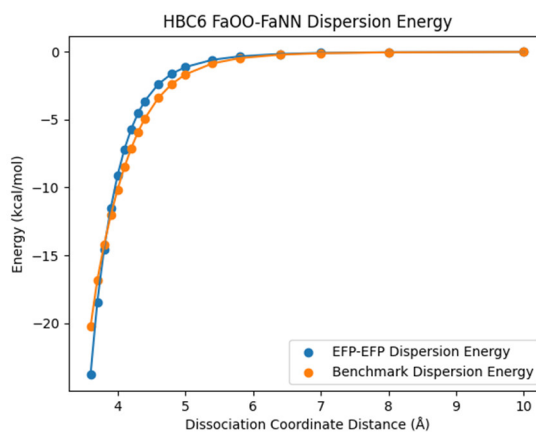
(b)



(e)



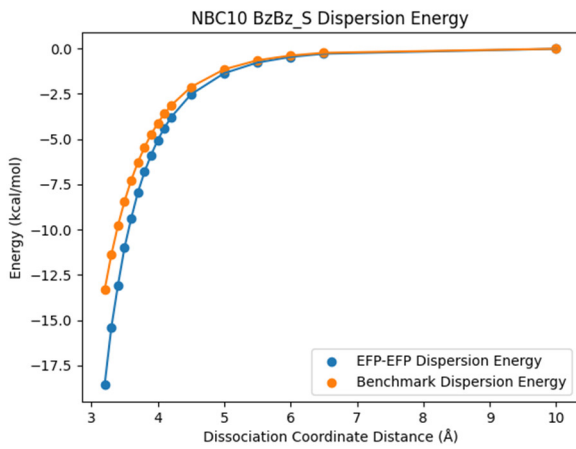
(c)



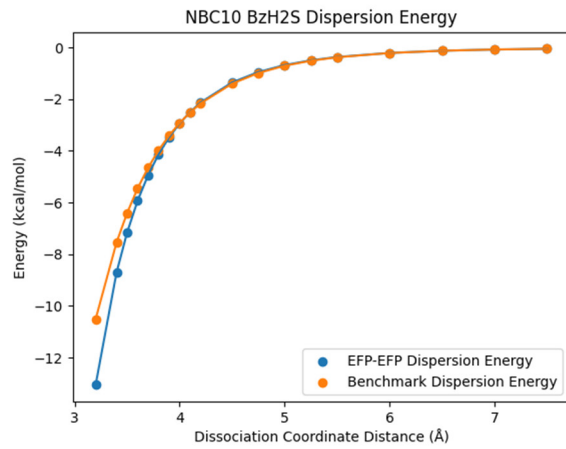
(f)

Figure 4. The dispersion energies of the HBC6 database predicted by the new EFP model with isotropic disp8 (blue) and SAPT (orange). Homogeneous dimers: (a) FaOO-FaOO, (b) FaON-FaON, (c) FaNN-FaNN. Heterogenous dimers: (d) FaOO-FaON, (e) FaON-FaNN, (f) FaOO-FaNN.

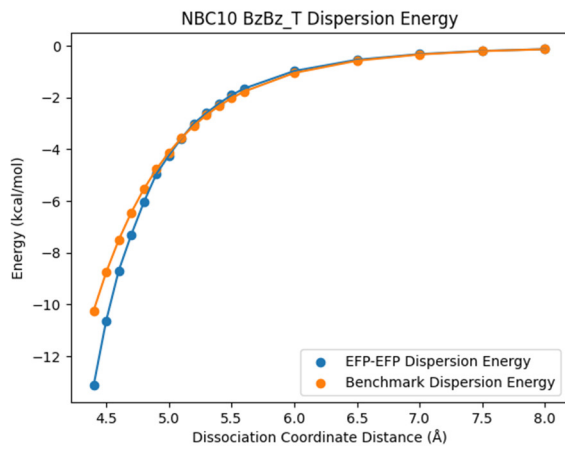
NBC10



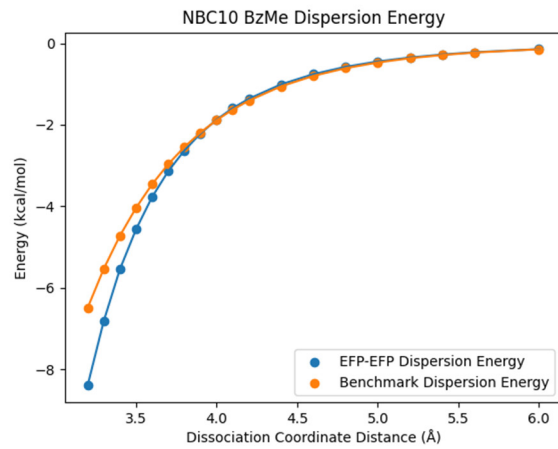
(a)



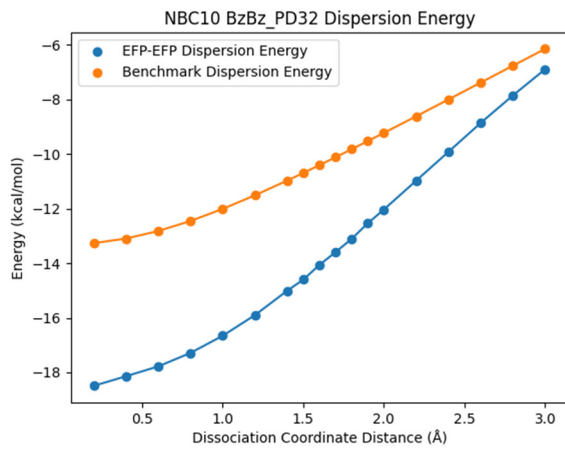
(f)



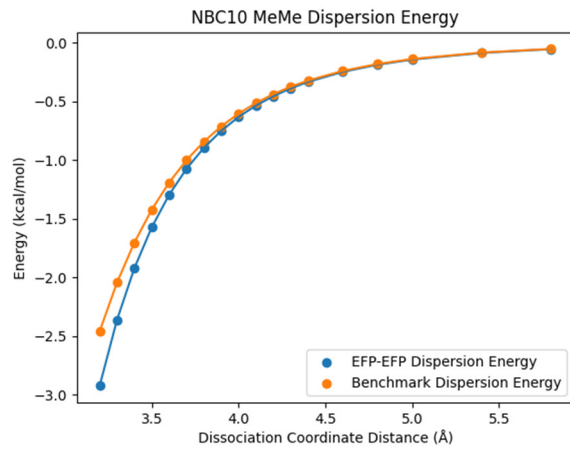
(b)



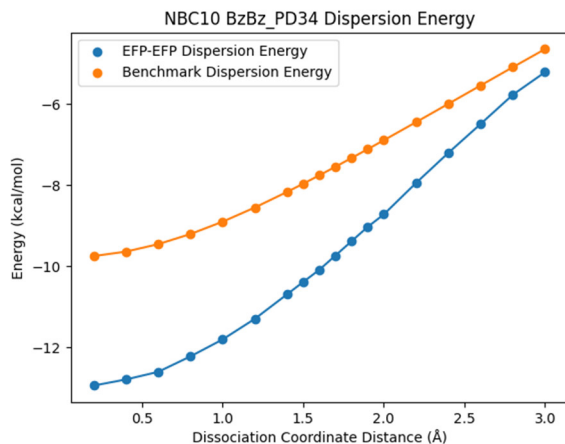
(g)



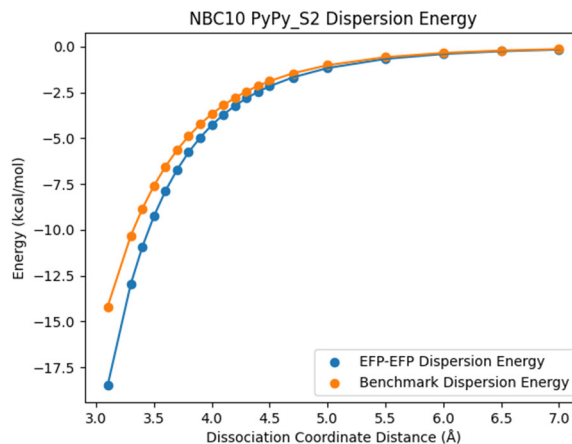
(c)



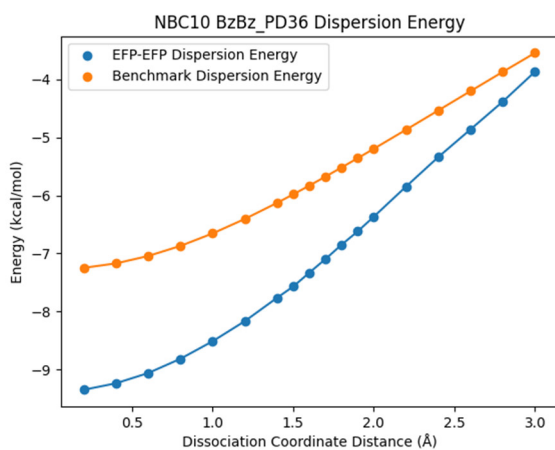
(h)



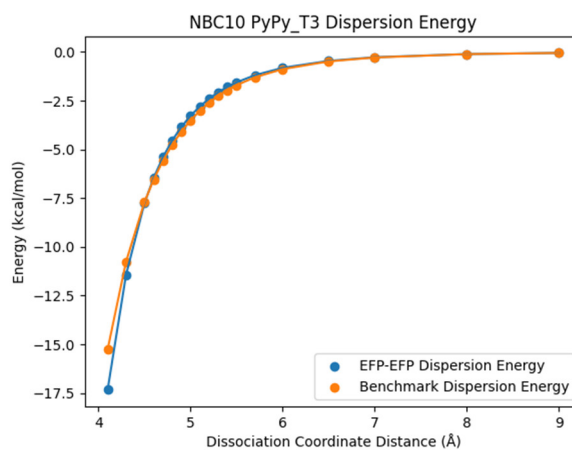
(d)



(i)



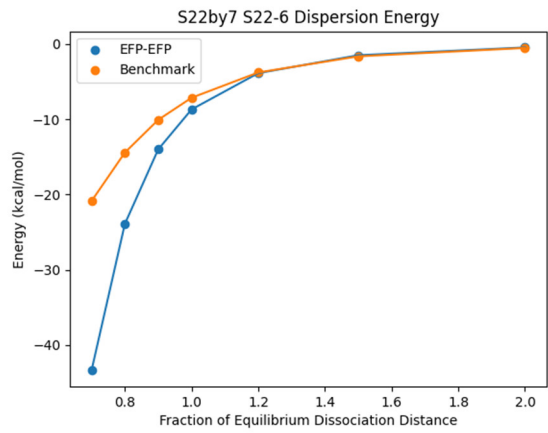
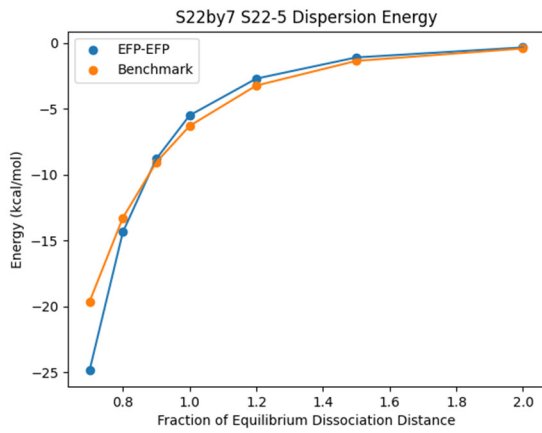
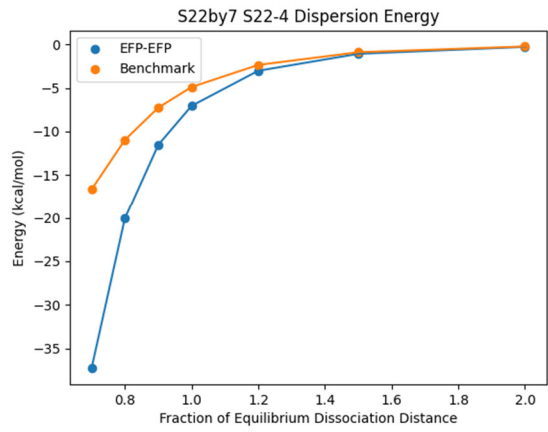
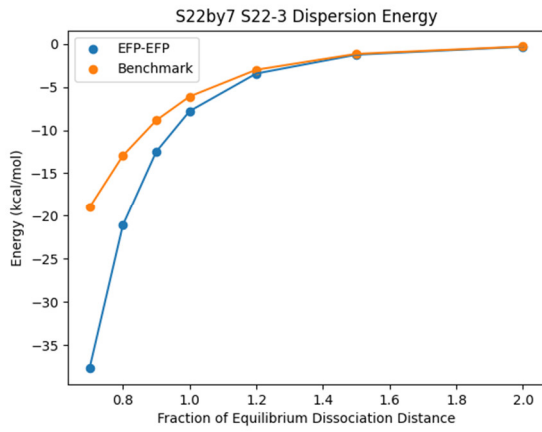
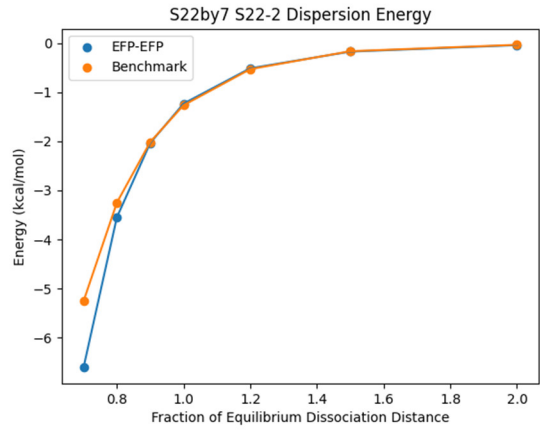
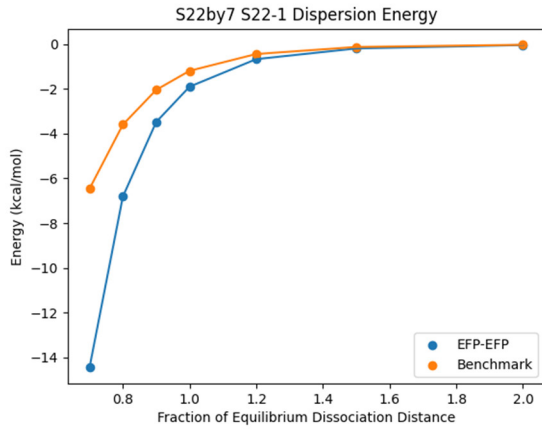
(e)

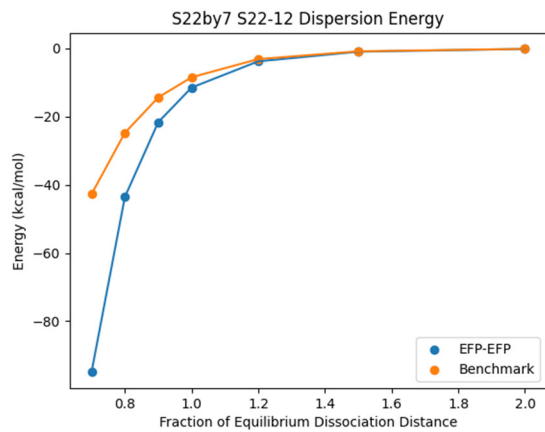
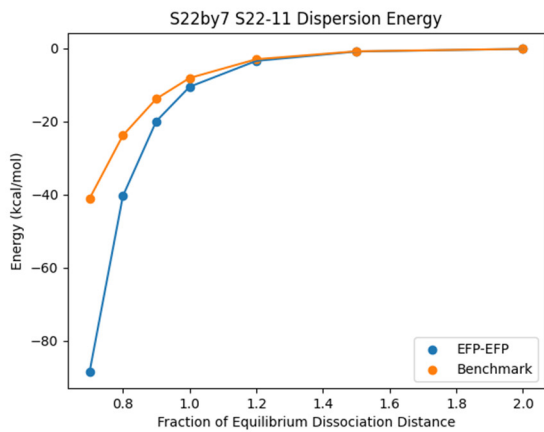
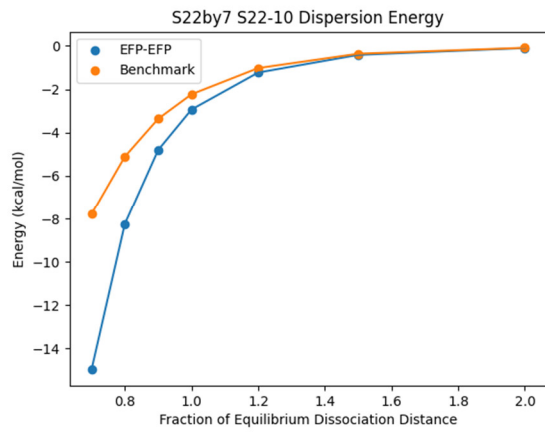
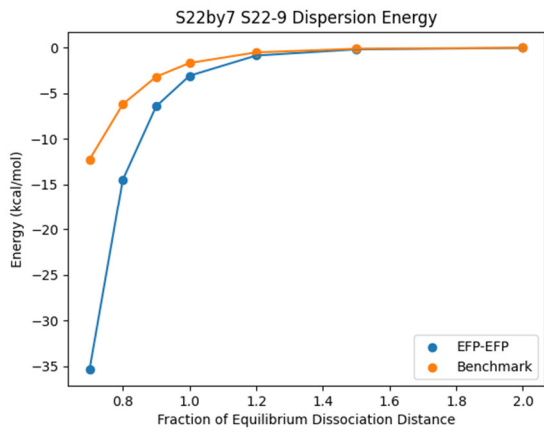
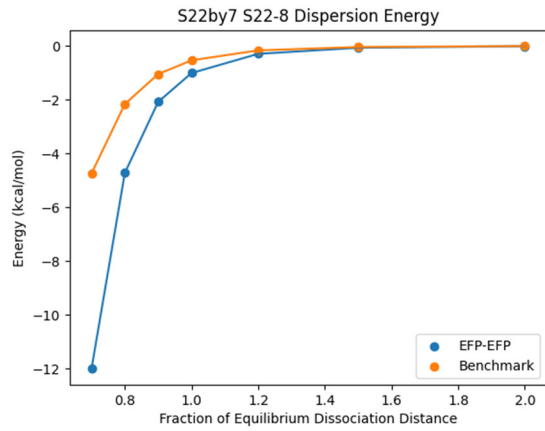
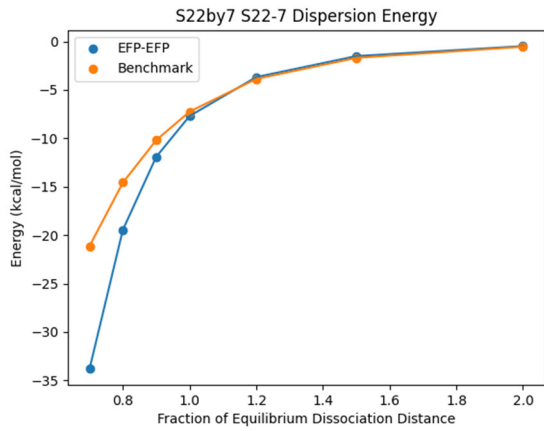


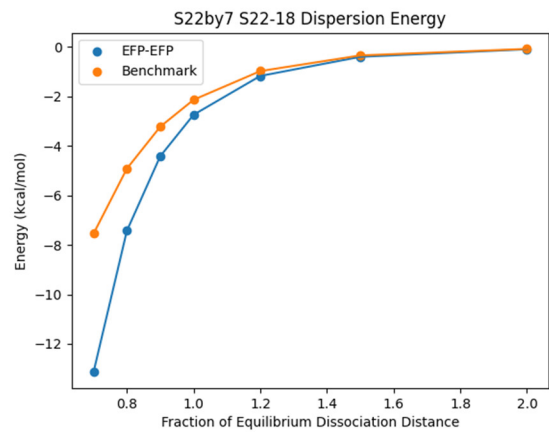
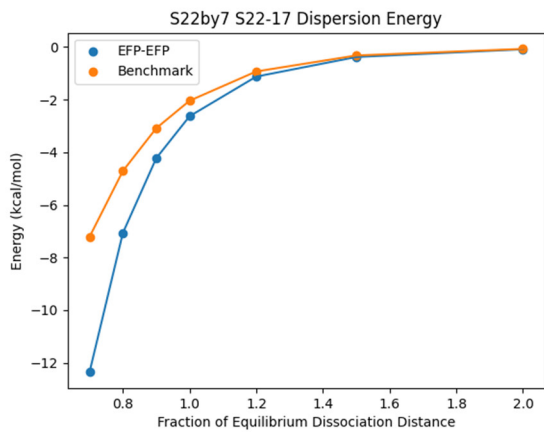
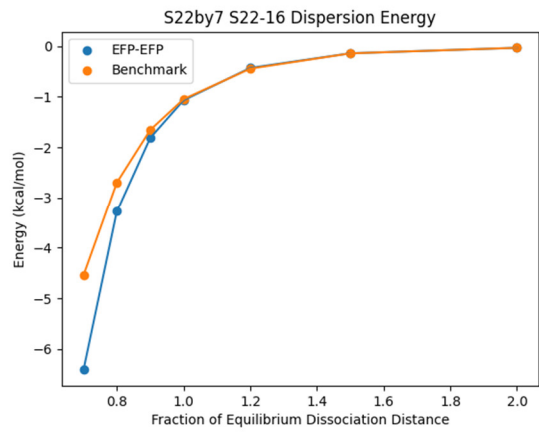
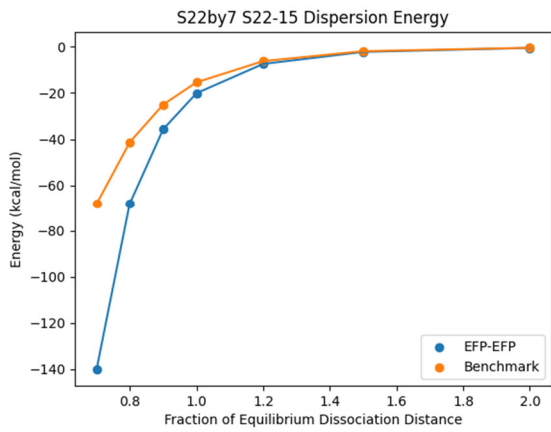
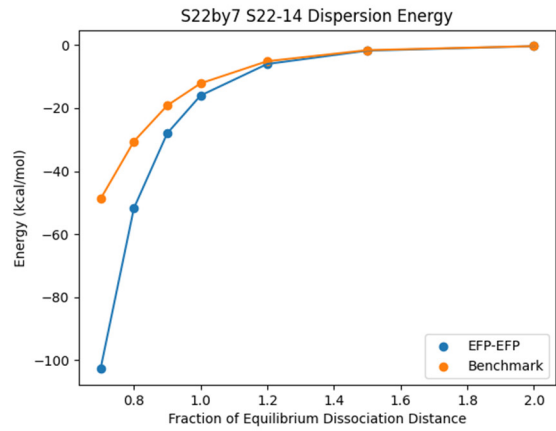
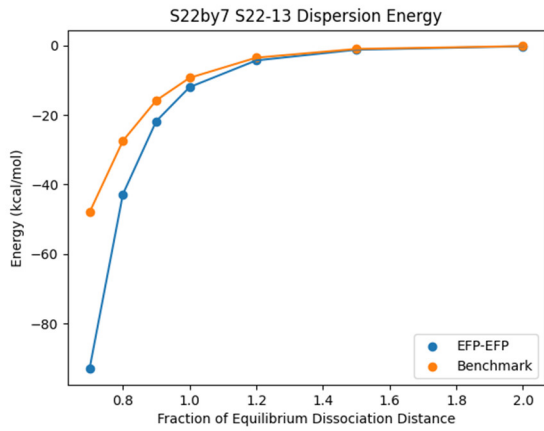
(j)

Figure 5. The dispersion energies of the NBC10 database predicted by the new EFP model with isotropic disp8 (blue) and SAPT (orange). (a) sandwich benzene dimer (BzBz_S), (b) T-shaped benzene dimer (BzBz_T), (c) 3.2, (d) 3.4 and (e) 3.6 Å separated (BzBz_PD32, BzBz_PD34, BzBz_PD36) configurations of the benzene dimer, (f) benzene with hydrogen sulfide (BzH2S), (g) benzene with methane (BzMe), (h) methane dimer (MeMe), (i) the antiparallel sandwich (PyPy_S2) pyridine dimer and (j) a T-shaped (PyPy_T3) pyridine dimer.

S22by7







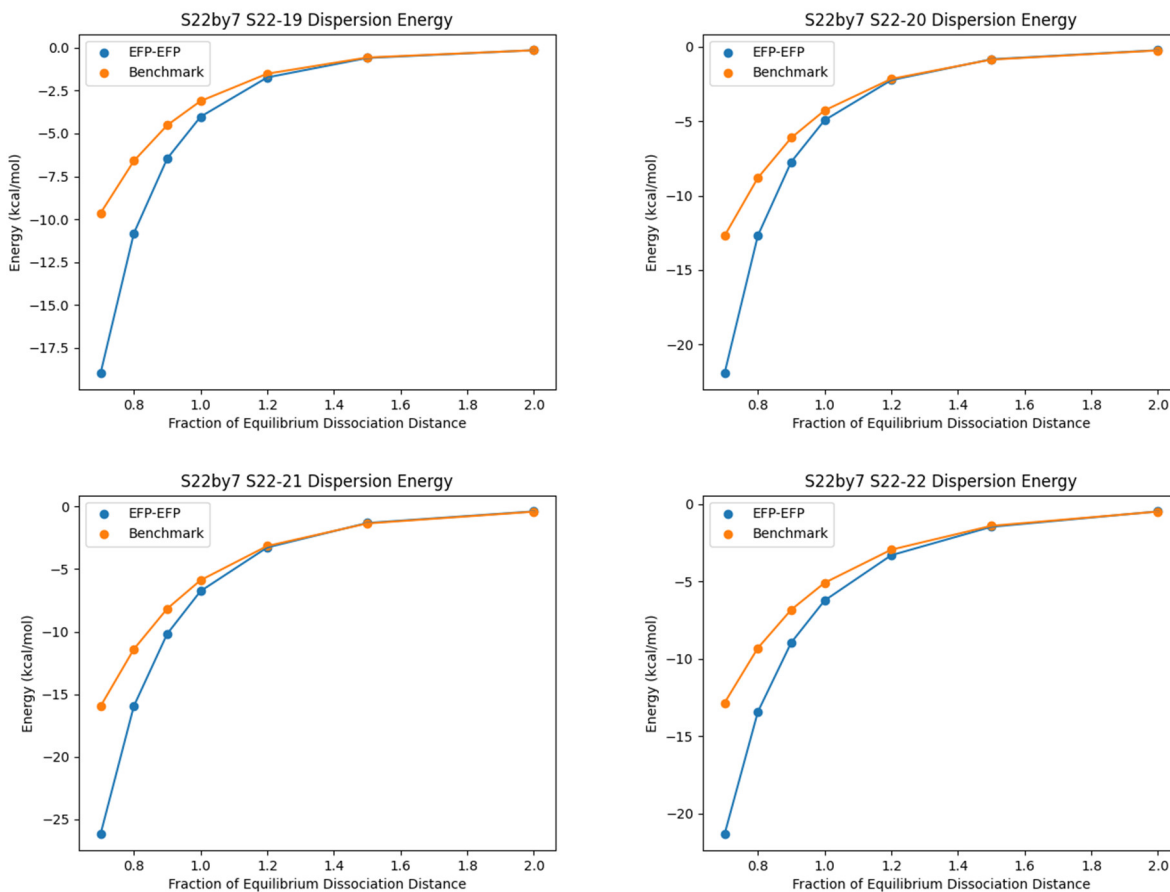
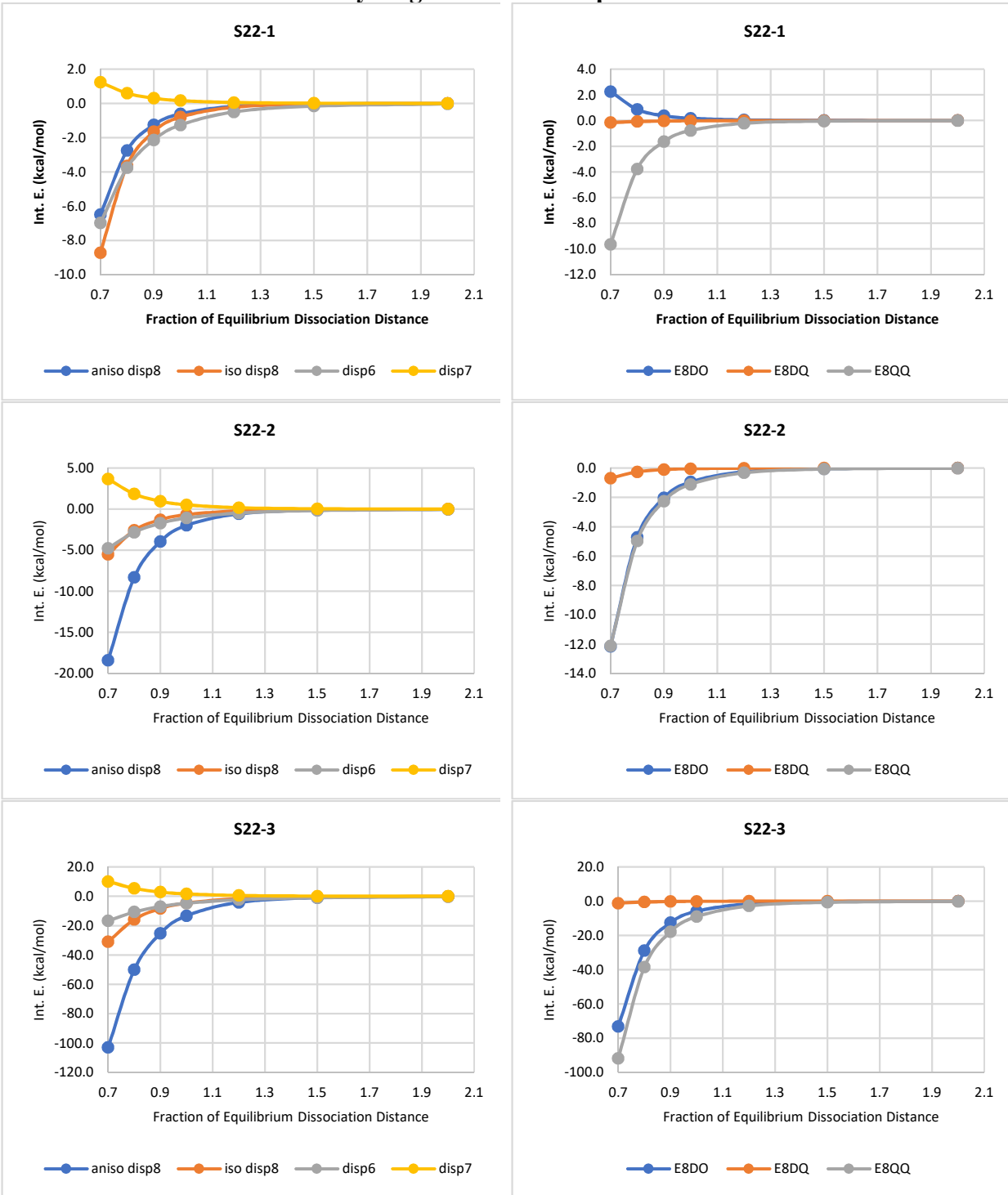
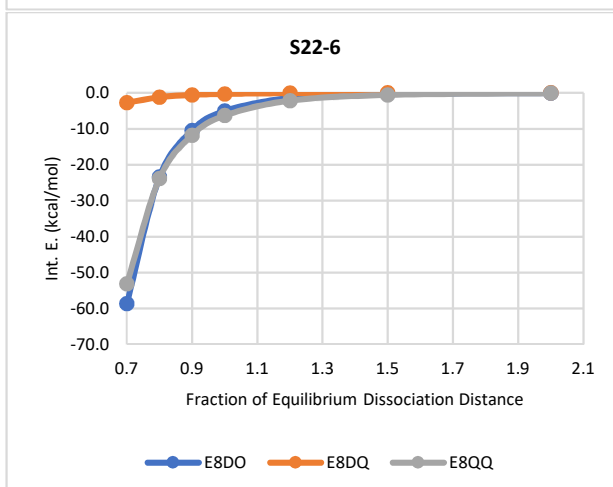
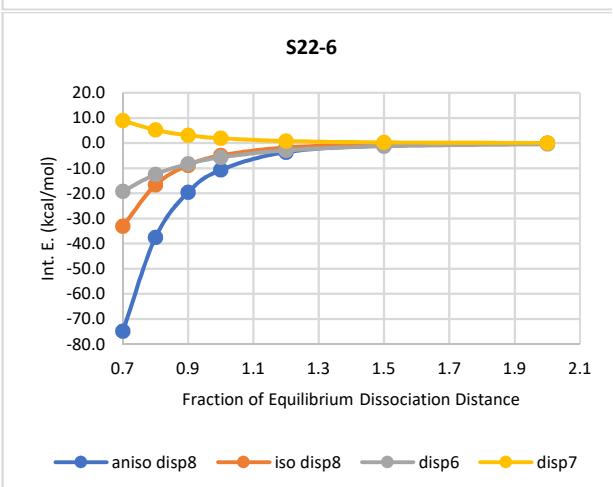
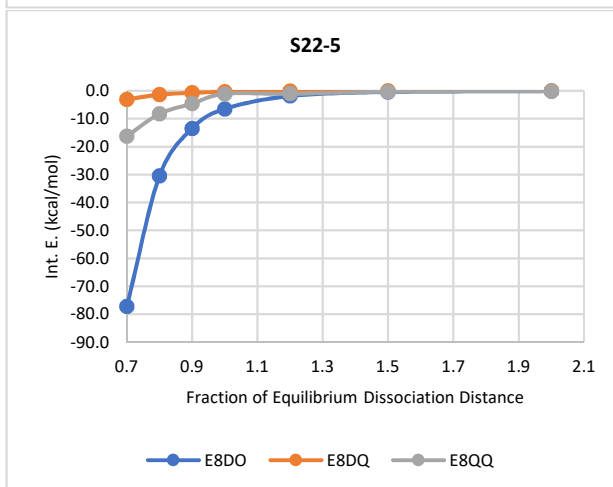
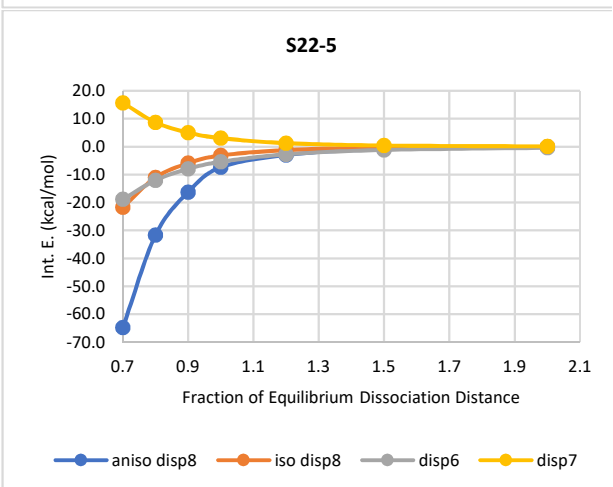
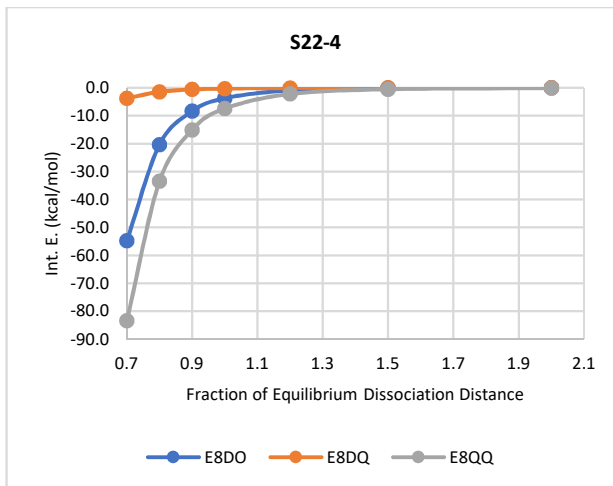
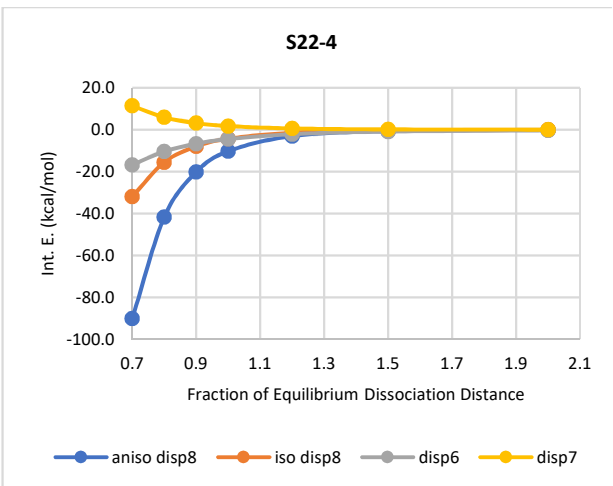


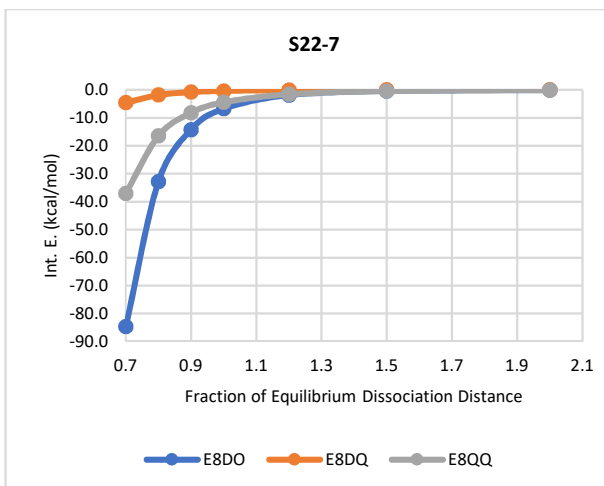
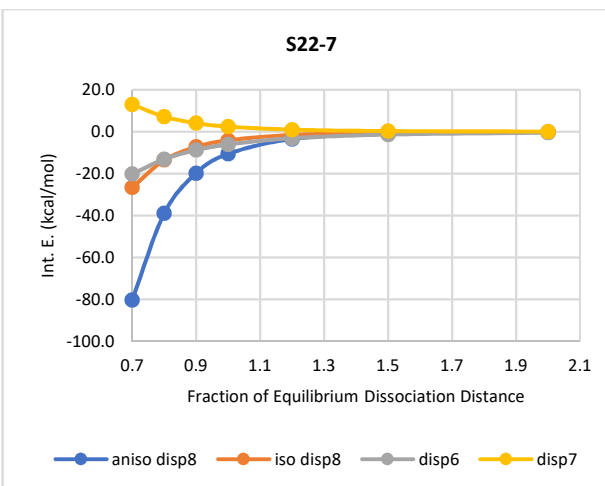
Figure 6. The dispersion energies predicted by EFP with isotropic disp8 and SAPT for S22by7 database. The list of 22 dimers is the same as Table 5. The separations are at 0.7, 0.8, 0.9, 1.0, 1.2, 1.5 and 2.0 x equilibrium separation.

The various components of the dispersion, R^{-6} dispersion (disp6), R^{-7} dispersion (disp7), isotropic R^{-8} dispersion (iso disp8) and the anisotropic R^{-8} dispersion (aniso disp8), as well as the three components in the anisotropic R^{-8} dispersion (E8DO, E8DQ and E8QQ), as a function of the fraction of the equilibrium separation, are plotted below for the S22by7 database (Figure 7).

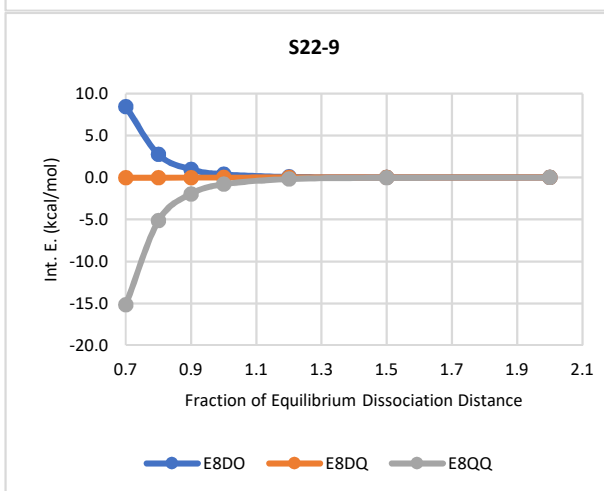
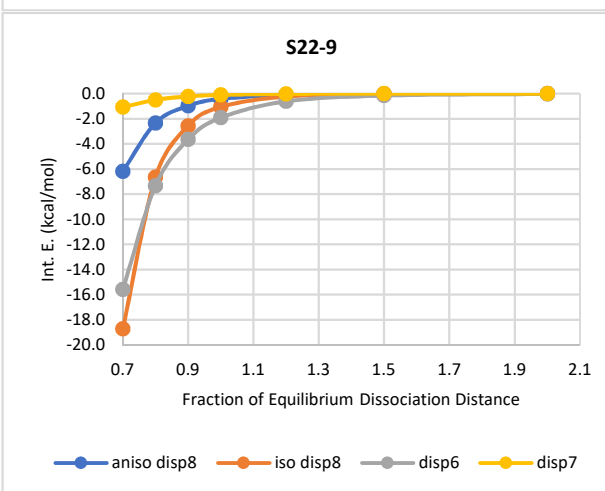
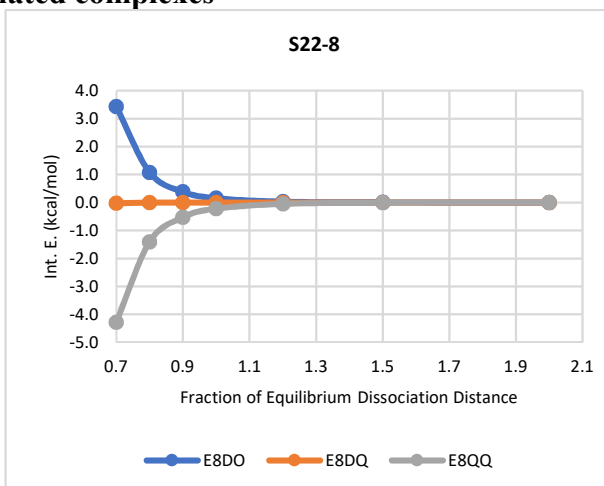
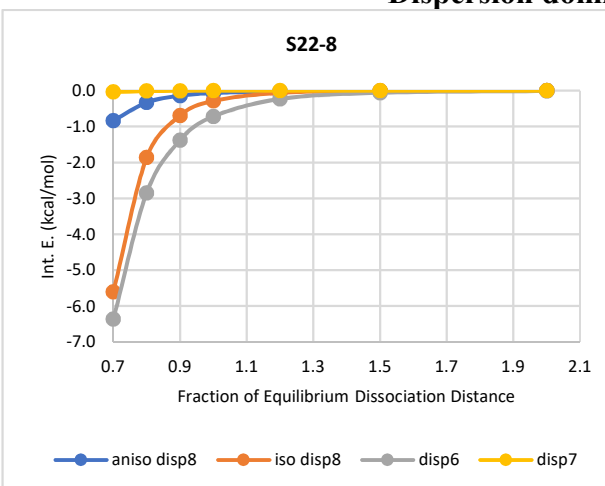
Hydrogen-Bonded complexes

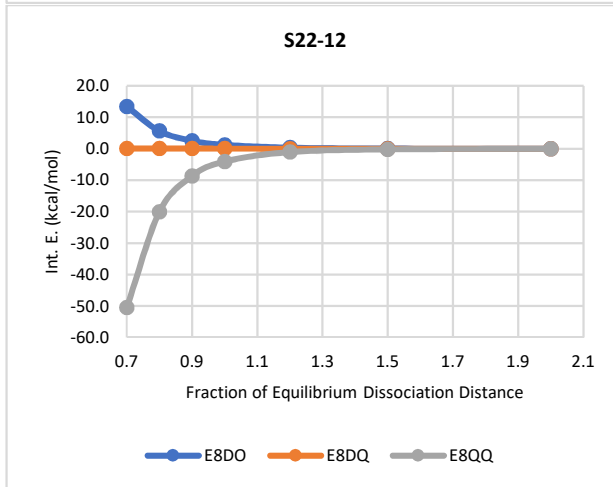
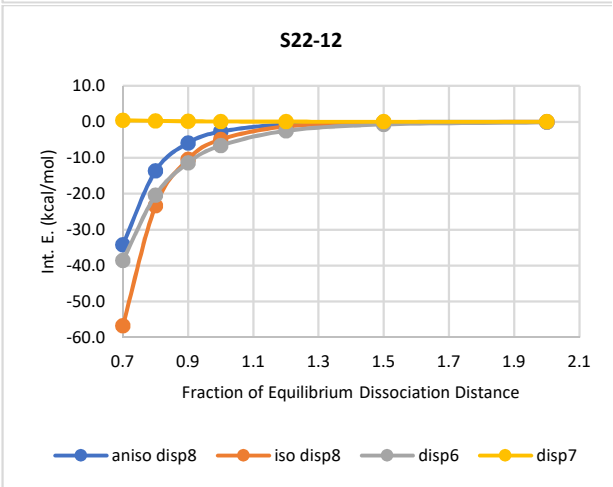
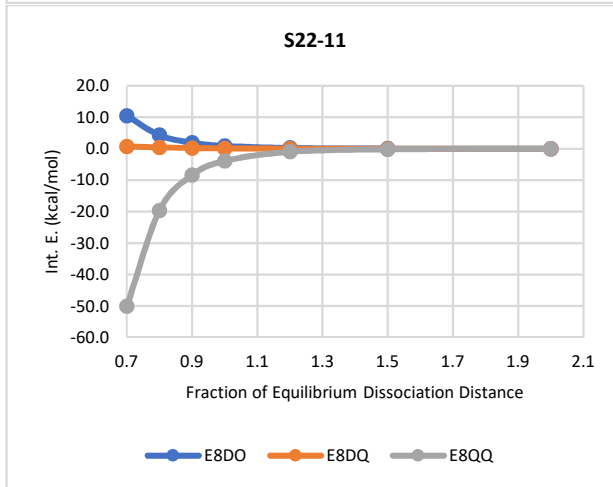
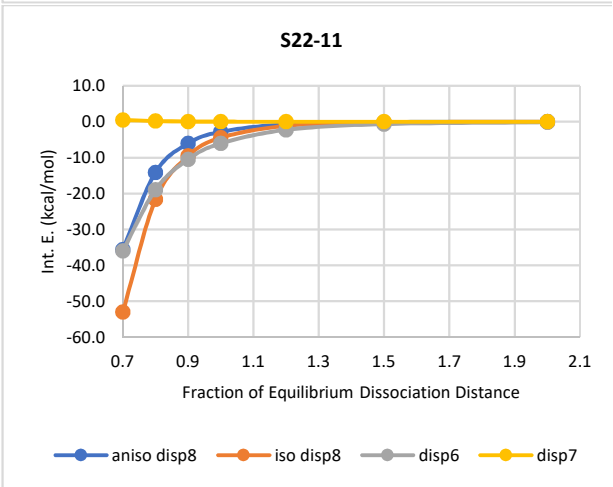
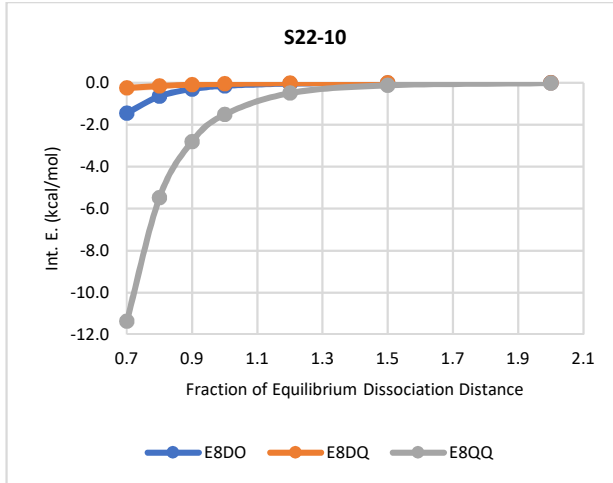
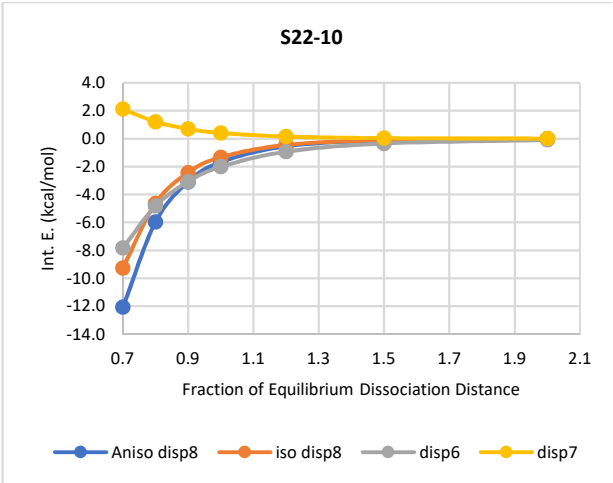


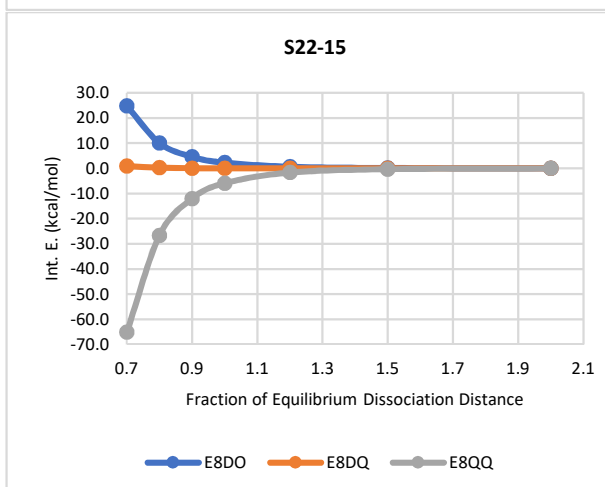
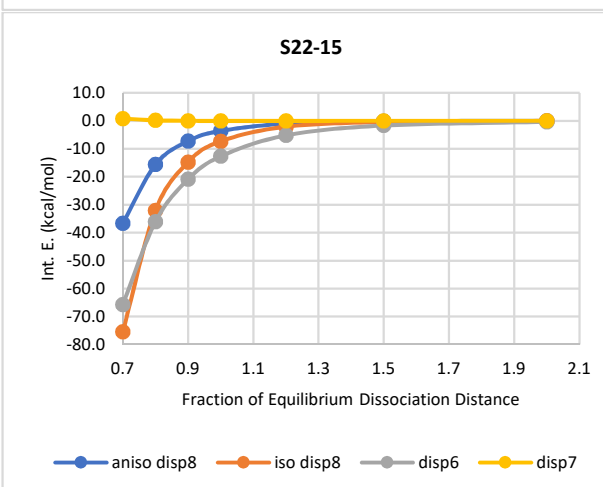
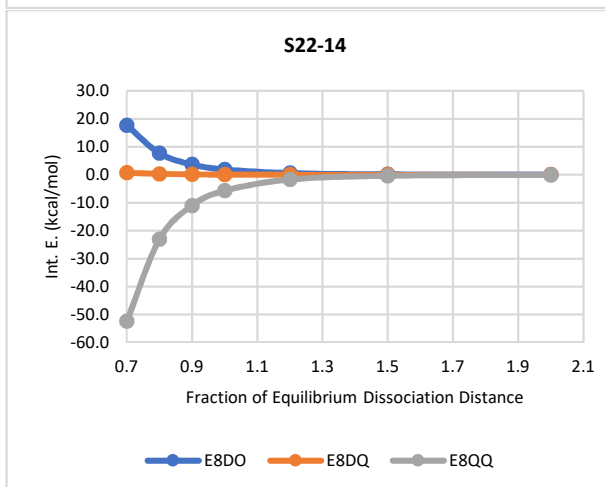
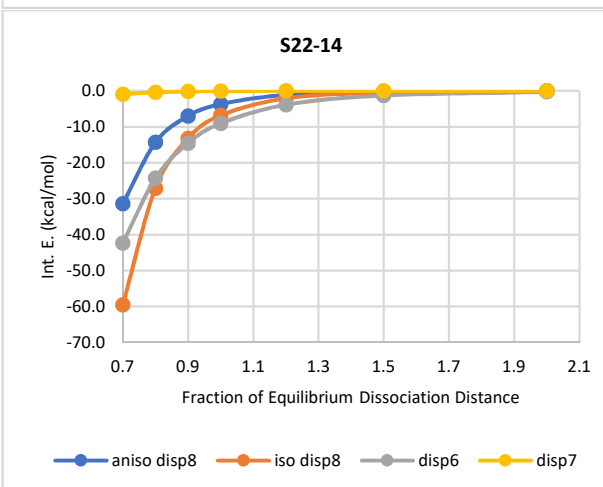
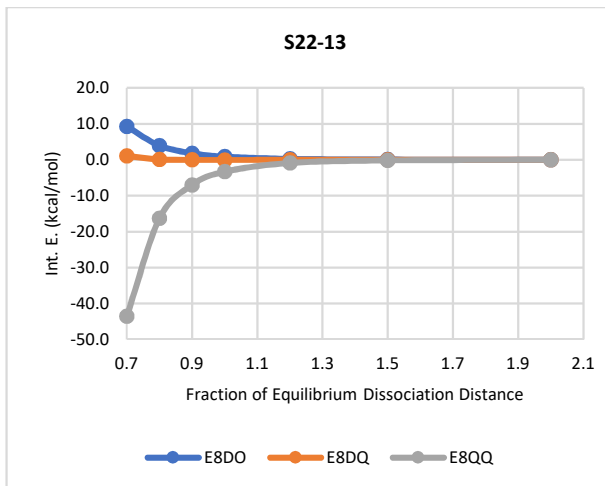
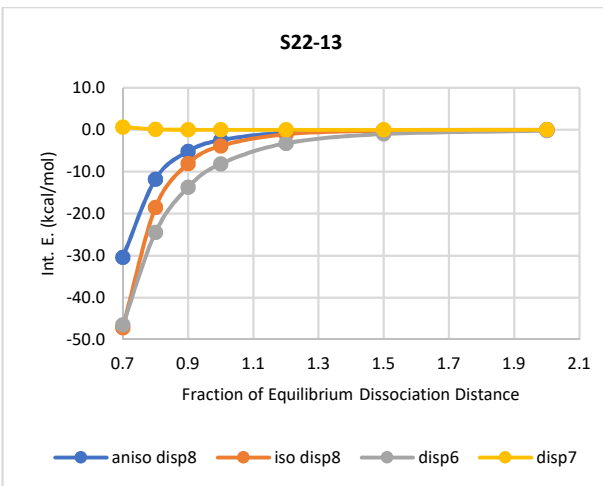




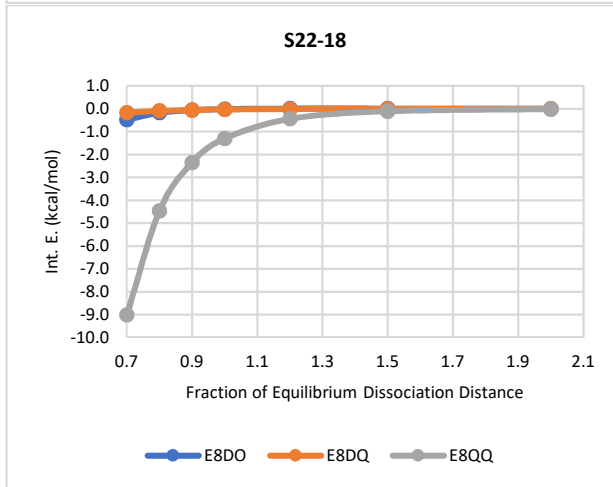
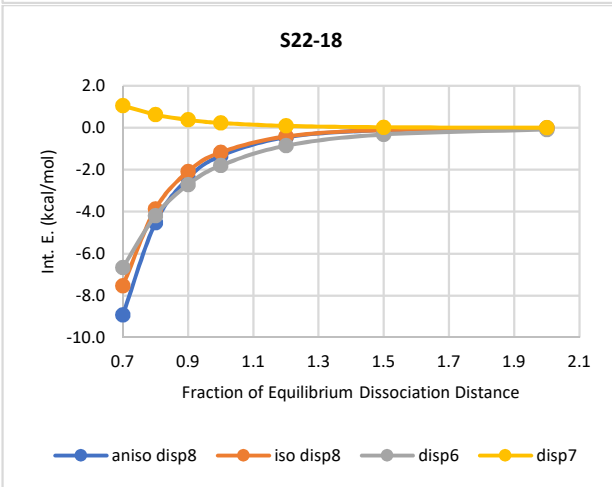
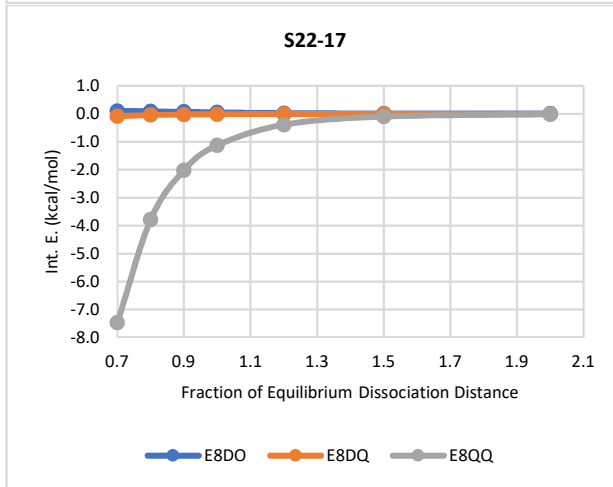
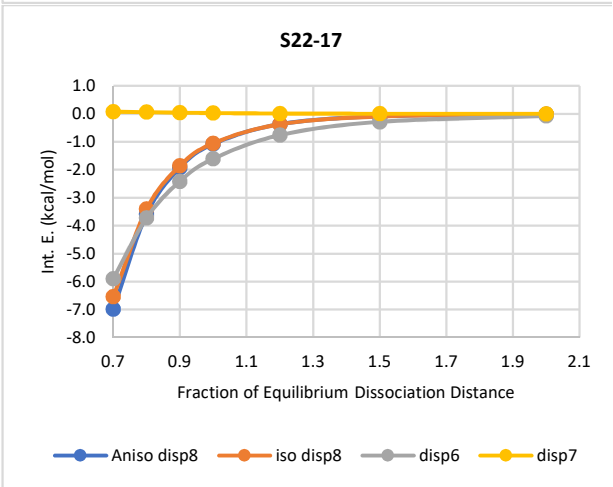
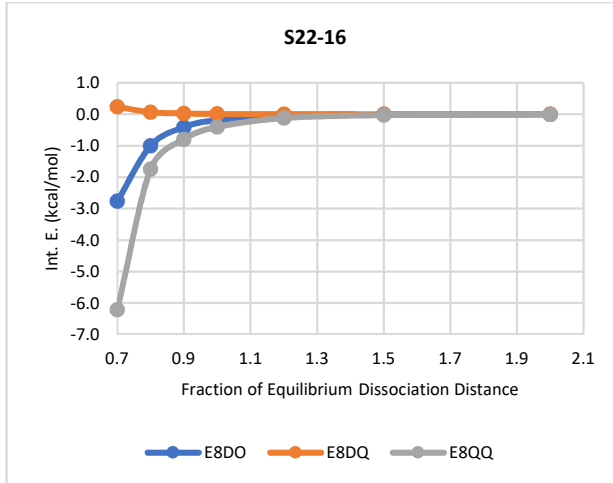
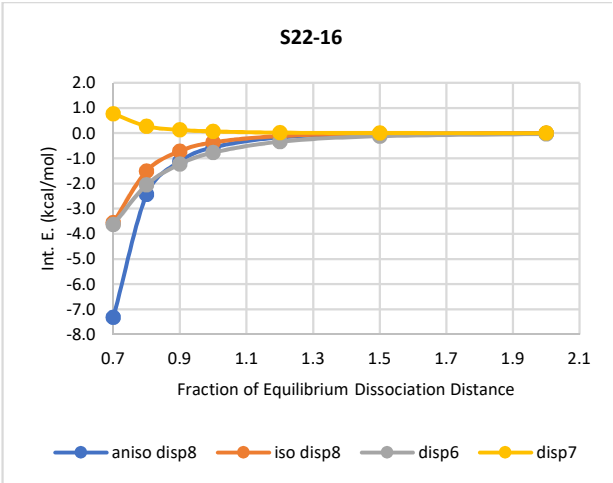
Dispersion dominated complexes

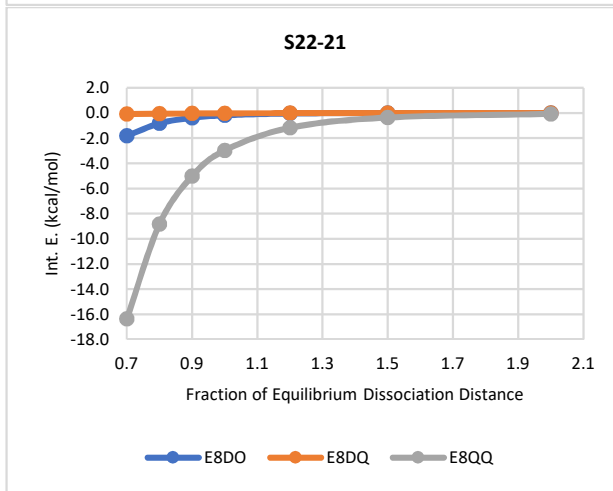
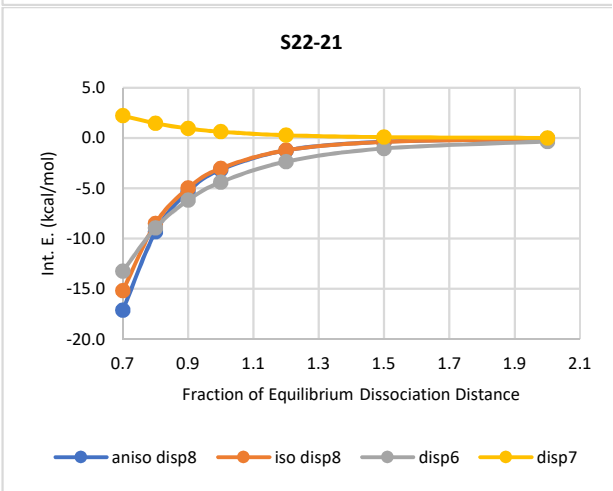
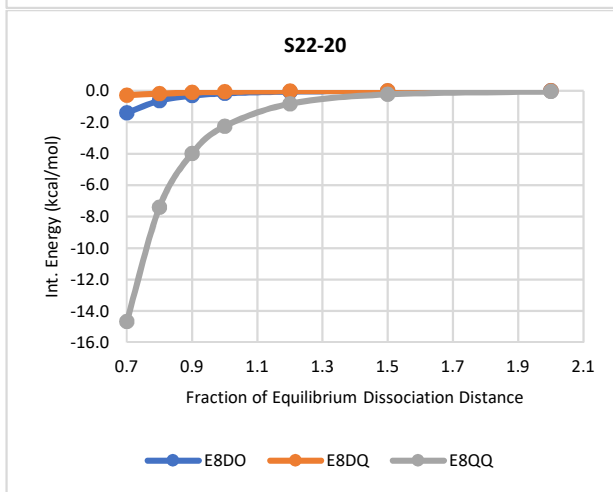
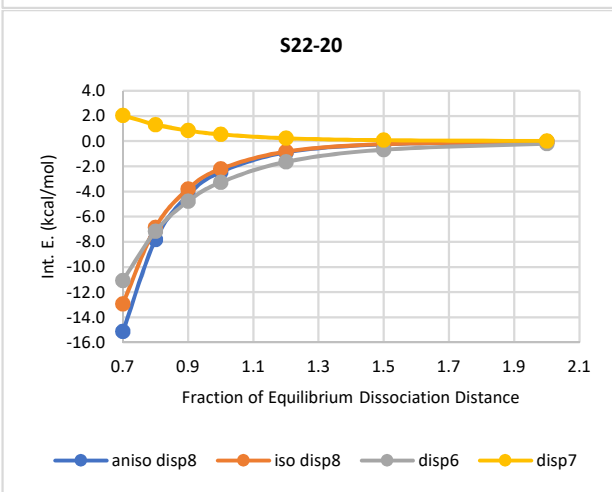
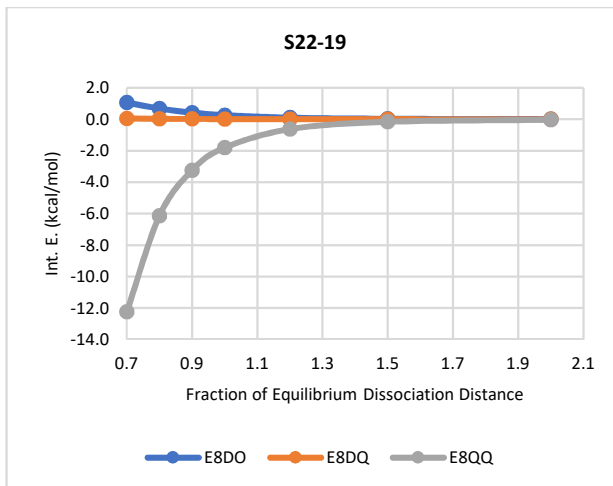
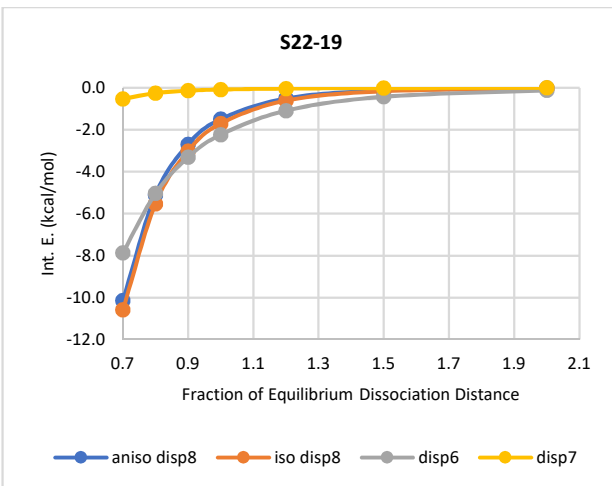






Mixed complexes





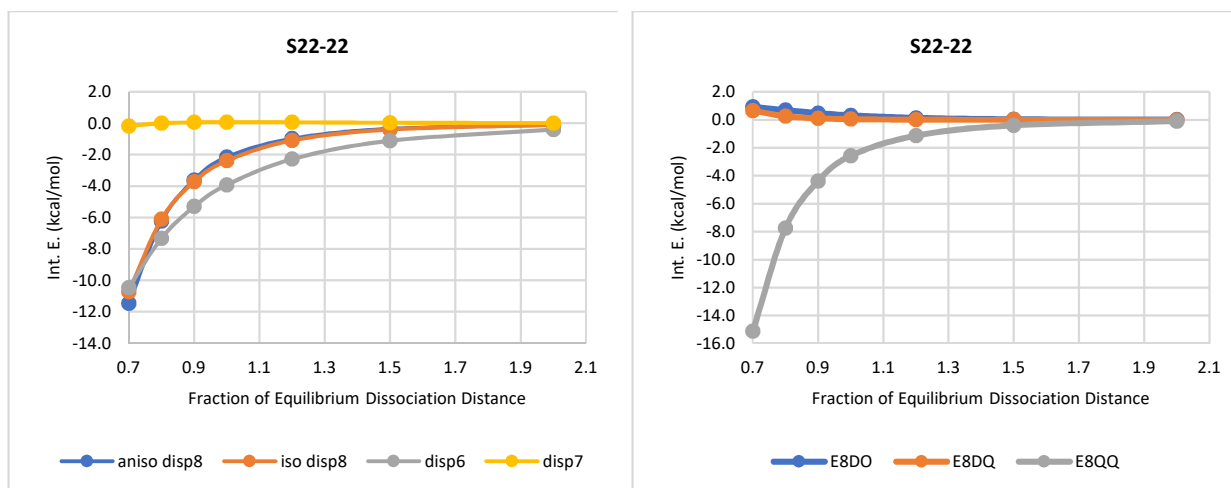


Figure 7. The various dispersion components predicted by EFP for the S22by7 database. The left panel shows disp6, disp7, isotropic disp8 and anisotropic disp8, and the right panel shows the three contributions to the anisotropic disp8, E8DO, E8DQ and E8QQ. The list of 22 dimers is the same as those in Table 5. The separations are at 0.7, 0.8, 0.9, 1.0, 1.2, 1.5 and 2.0 x equilibrium separation.

At equilibrium separations, disp8 can be a fraction of disp6. But, isotropic disp8 quickly increases relative to disp6 as the intermolecular separations shorten. The crossover point for many dimers in the S22 test set is at 0.8x equilibrium separation. The anisotropic disp8 can be much larger or smaller than disp6 at short range, depending on whether the E8DO component is attractive or repulsive. Disp7 is generally small and slowly changing. Most of the dimers have increasingly repulsive (positive) disp7 as the separation decreases.

As the distance between two molecules decreases, E8DQ remains small along the separation coordinate, whereas E8QQ and E8DO show more noticeable changes. E8QQ remains attractive (negative) at all separations for all 22 dimers in the set, while E8DO is attractive for most of the hydrogen-bonded dimers and repulsive for most of the dispersion dominated dimers. In terms of the magnitudes, E8QQ almost always undergoes steep changes as the intermolecular separation decreases. On the other hand, changes in E8DO do not demonstrate a discernable pattern.

V. Conclusions and Future work

The anisotropic and isotropic R^{-8} dispersion interactions (disp8) are derived in the framework of the effective fragment potential method by extending the interaction operator, which is expressed in terms of distributed multipoles, through octopoles. The final expressions for disp8 are a double summation of products of dynamic LMO polarizabilities. Two forms of damping functions for EFP dispersion terms, overlap-based and Tang-Toennies, have been extended to screen disp8. The analytic gradient has been derived and implemented for the isotropic disp8.

A total of twelve databases were examined by EFP, eight of which contain dimers at a single geometry. The other four contain potential energy scans along certain interacting coordinates (e.g., hydrogen bonding dissociation coordinate). Compared against the SAPT benchmarks, a

progressive improvement can be observed from the previous empirical disp8 model to the isotropic disp8 and the anisotropic disp8 formulation. For anisotropic disp8, the contributions due to dipole-quadrupole polarizabilities are always very small (< 5% of the total dispersion energy) while dipole-octopole polarizabilities or quadrupole-quadrupole polarizabilities make more prominent contributions. The contribution due to quadrupole-quadrupole polarizabilities is always attractive, whereas that due to dipole-octopole can be either attractive or repulsive. Interestingly, it was observed that the dipole-octopole term tends to be attractive for the hydrogen-bonded complexes and repulsive for the dispersion dominated complexes.

When the maximum element in the intermolecular overlap matrix is greater than 0.1, sizable errors in the dispersion energies are observed, likely due to the breakdown at short distances of the multipole expansion of the interacting operator and/or approximating the total wave function by the Hartree product of individual fragment wave functions.

The dispersion interaction energies for dimers along an interacting coordinate predicted by the new isotropic disp8 yield correct curvatures and good agreement with SAPT benchmarks around equilibrium and longer but overestimate the dispersion interactions at short range, due to the presence of large intermolecular overlap matrix elements.

To screen dispersion energies, the overlap-based damping functions provide better error statistics (smaller MAEs) than the Tang-Toennies damping functions and are free of empirical parameters. Future work will focus on refining the functional form of the overlap-based damping functions for all three terms in EFP dispersion interactions (disp6, disp7 and disp8).

Appendix

(A) Proof that the Dipole-Octopole Polarizability Tensor Vanishes upon Rotational Averaging

Throughout the derivation, Einstein notation is assumed, that is, the repeated suffix in a term is automatically summed over.

An n th rank tensor T with respect to different frames of reference can be related through the equation:

$$T_{i_1 \dots i_n} = l_{i_1 \lambda_1} \dots l_{i_n \lambda_n} T_{\lambda_1 \lambda_2 \dots \lambda_n} \quad (\text{A1})$$

where $T_{i_1 \dots i_n}$ and $T_{\lambda_1 \lambda_2 \dots \lambda_n}$ are n th rank tensors with respect to the space-fixed frame and the molecule-fixed frame, respectively, and $l_{i_p \lambda_p}$ is the cosine of the angle between the space-fixed axis and the molecule-fixed axis.

Denote the rotational average $\langle l_{i_1 \lambda_1} \dots l_{i_n \lambda_n} \rangle$ by $I^{(n)}$, which can be expressed in terms of a linear combination of isotropic tensors. Each member of this linear combination, called an *isomer*, is a product of two isotropic tensors, one referred to the space-fixed frame and the other to the molecule-fixed frame.

The fundamental isotropic tensors in three-dimensions are the Kronecker delta, δ_{ij} , and the Levi-Civita epsilon, ϵ_{ijk} , for which the usual quantum chemistry antisymmetrizer is a special case. For the dipole-octopole polarizability, which is a 4th rank tensor, the possible isomers are formed by permuting the indices from two Kronecker deltas:

$$f_1 = \delta_{i_1 i_2} \delta_{i_3 i_4}, \quad f_2 = \delta_{i_1 i_3} \delta_{i_2 i_4}, \quad f_3 = \delta_{i_1 i_4} \delta_{i_2 i_3} \quad (\text{A2})$$

According to Weyl's theorem⁵⁹,

$$\mathbf{I}^{(n)} = \sum_{r,s} f_r^{(n)} m_{rs}^{(n)} g_s^{(n)} \quad (\text{A3})$$

where $f_r^{(n)}$ and $g_s^{(n)}$ are the r th and s th members of the set of isomers in the space-fixed frame and the molecule-fixed frame, respectively. And $m_{rs}^{(n)}$ are numerical coefficients.

A given isomer $f_q^{(n)}$ and its corresponding isomer $g_q^{(n)}$ in the other frame are related by:

$$f_q^{(n)} = \mathbf{I}^{(n)} g_q^{(n)} \quad (\text{A4})$$

Substituting Eq. (A3) into Eq. (A4) and multiplying both sides by $f_t^{(n)}$ gives rise to

$$f_t^{(n)} f_q^{(n)} = f_t^{(n)} \left[\sum_{r,s} m_{rs}^{(n)} f_r^{(n)} g_s^{(n)} \right] g_q^{(n)} = \sum_{r,s} f_t^{(n)} f_r^{(n)} m_{rs}^{(n)} g_s^{(n)} g_q^{(n)} \quad (\text{A5})$$

Products of $f_t^{(n)} f_r^{(n)}$ upon contraction of indices (i_1, i_2 etc.) yield a matrix element $s_{tr}^{(n)}$. So, Eq. (A5) becomes

$$s_{tq}^{(n)} = \sum_{r,s} s_{tr}^{(n)} m_{rs}^{(n)} s_{sq}^{(n)} \quad (\text{A6})$$

That is,

$$\mathbf{S}^{(n)} = \mathbf{S}^{(n)} \mathbf{M}^{(n)} \mathbf{S}^{(n)} \quad (\text{A7})$$

Assuming that the inverse of $\mathbf{S}^{(n)}$ exists, it follows that,

$$\mathbf{M}^{(n)} = (\mathbf{S}^{(n)})^{-1} \quad (\text{A8})$$

In other words, solving for $\mathbf{S}^{(n)}$ yields $\mathbf{M}^{(n)}$ and therefore $\mathbf{I}^{(n)}$.

For $n=4$, one has a 4th rank tensor; e.g., the dipole-octopole polarizability. $\mathbf{S}^{(4)}$ is formed from the isomers in Eq. (A2).

$$\mathbf{S}^{(4)} = \begin{pmatrix} f_1^2 & f_1 f_2 & f_1 f_3 \\ f_2 f_1 & f_2^2 & f_2 f_3 \\ f_3 f_1 & f_3 f_2 & f_3^2 \end{pmatrix} = \begin{pmatrix} 9 & 3 & 3 \\ 3 & 9 & 3 \\ 3 & 3 & 9 \end{pmatrix} \quad (\text{A9})$$

Consider f_1^2 as an example for the diagonal terms. From Eq (A2),

$$f_1^2 = \delta_{i_1 i_2} \delta_{i_3 i_4} \delta_{i_1 i_2} \delta_{i_3 i_4} \quad (\text{A10})$$

By construction, the terms can survive only if $i_1 = i_2$ and $i_3 = i_4$.

Set $i_1 = i_2 = i, i_3 = i_4 = j$, with $i, j = x, y, z$. Then,

$$\begin{aligned} & \delta_{i_1 i_2} \delta_{i_3 i_4} \delta_{i_1 i_2} \delta_{i_3 i_4} \delta_{i_1 i_2} \delta_{i_3 i_4} = \delta_{ii} \delta_{jj} \delta_{ii} \delta_{jj} \\ & = \left(\delta_{xx} \delta_{xx} \delta_{xx} \delta_{xx} + \delta_{xx} \delta_{yy} \delta_{xx} \delta_{yy} + \delta_{xx} \delta_{zz} \delta_{xx} \delta_{zz} + \delta_{yy} \delta_{xx} \delta_{yy} \delta_{xx} + \delta_{yy} \delta_{zz} \delta_{yy} \delta_{zz} \right. \\ & \quad \left. + \delta_{yy} \delta_{yy} \delta_{yy} \delta_{yy} + \delta_{zz} \delta_{xx} \delta_{zz} \delta_{xx} + \delta_{zz} \delta_{yy} \delta_{zz} \delta_{yy} + \delta_{zz} \delta_{zz} \delta_{zz} \delta_{zz} \right) \\ & = (1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1) = 9 \end{aligned} \quad (\text{A11})$$

Similarly, for $f_2^2 = f_3^2 = 9$. Next, consider the off-diagonal term $f_1 f_2$.

$f_1 f_2 = f_2 f_1 = \delta_{i_1 i_2} \delta_{i_3 i_4} \delta_{i_1 i_3} \delta_{i_2 i_4}$ with terms surviving only if $i_1 = i_2 = i_3 = i_4$, and in Cartesian coordinates,

$$\delta_{i_1 i_2} \delta_{i_3 i_4} \delta_{i_1 i_3} \delta_{i_2 i_4} = \delta_{xx} \delta_{xx} \delta_{xx} \delta_{xx} + \delta_{yy} \delta_{yy} \delta_{yy} \delta_{yy} + \delta_{zz} \delta_{zz} \delta_{zz} \delta_{zz} = 1 + 1 + 1 = 3 \quad (\text{A12})$$

Similarly, $f_1 f_3 = f_3 f_1 = f_2 f_3 = f_3 f_2 = 3$

Thus,

$$M^{(4)} = (S^{(4)})^{-1} = \frac{1}{30} \begin{pmatrix} 4 & -1 & -1 \\ -1 & 4 & -1 \\ -1 & -1 & 4 \end{pmatrix} \quad (\text{A13})$$

From Eq. (A3)

$$\begin{aligned} \mathbf{I}^{(4)} &= \frac{1}{30} \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} \begin{pmatrix} 4 & -1 & -1 \\ -1 & 4 & -1 \\ -1 & -1 & 4 \end{pmatrix} \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} \\ &= \frac{1}{30} \begin{pmatrix} \delta_{i_1 i_2} \delta_{i_3 i_4} \\ \delta_{i_1 i_3} \delta_{i_2 i_4} \\ \delta_{i_1 i_4} \delta_{i_2 i_3} \end{pmatrix}^T \begin{pmatrix} 4 & -1 & -1 \\ -1 & 4 & -1 \\ -1 & -1 & 4 \end{pmatrix} \begin{pmatrix} \delta_{\lambda_1 \lambda_2} \delta_{\lambda_3 \lambda_4} \\ \delta_{\lambda_1 \lambda_3} \delta_{\lambda_2 \lambda_4} \\ \delta_{\lambda_1 \lambda_4} \delta_{\lambda_2 \lambda_3} \end{pmatrix} \\ &= \frac{1}{30} \begin{bmatrix} \delta_{i_1 i_2} \delta_{i_3 i_4} (4\delta_{\lambda_1 \lambda_2} \delta_{\lambda_3 \lambda_4} - \delta_{\lambda_1 \lambda_3} \delta_{\lambda_2 \lambda_4} - \delta_{\lambda_1 \lambda_4} \delta_{\lambda_2 \lambda_3}) \\ + \delta_{i_1 i_3} \delta_{i_2 i_4} (-\delta_{\lambda_1 \lambda_2} \delta_{\lambda_3 \lambda_4} + 4\delta_{\lambda_1 \lambda_3} \delta_{\lambda_2 \lambda_4} - \delta_{\lambda_1 \lambda_4} \delta_{\lambda_2 \lambda_3}) \\ + \delta_{i_1 i_4} \delta_{i_2 i_3} (-\delta_{\lambda_1 \lambda_2} \delta_{\lambda_3 \lambda_4} - \delta_{\lambda_1 \lambda_3} \delta_{\lambda_2 \lambda_4} + 4\delta_{\lambda_1 \lambda_4} \delta_{\lambda_2 \lambda_3}) \end{bmatrix} \end{aligned} \quad (\text{A14})$$

The rotationally averaged dipole-octopole polarizability tensor $\bar{D}_{i_1 i_2 i_3 i_4}$ can then be obtained from full dipole-octopole polarizability tensor by taking $\mathbf{I}^{(4)} D_{\lambda_1 \lambda_2 \lambda_3 \lambda_4}$, as follows:

$$\begin{aligned} \bar{D}_{i_1 i_2 i_3 i_4} &= \frac{1}{30} [\delta_{i_1 i_2} \delta_{i_3 i_4} (4\delta_{\lambda_1 \lambda_2} \delta_{\lambda_3 \lambda_4} - \delta_{\lambda_1 \lambda_3} \delta_{\lambda_2 \lambda_4} - \delta_{\lambda_1 \lambda_4} \delta_{\lambda_2 \lambda_3}) \\ & \quad + \delta_{i_1 i_3} \delta_{i_2 i_4} (-\delta_{\lambda_1 \lambda_2} \delta_{\lambda_3 \lambda_4} + 4\delta_{\lambda_1 \lambda_3} \delta_{\lambda_2 \lambda_4} - \delta_{\lambda_1 \lambda_4} \delta_{\lambda_2 \lambda_3}) \\ & \quad + \delta_{i_1 i_4} \delta_{i_2 i_3} (-\delta_{\lambda_1 \lambda_2} \delta_{\lambda_3 \lambda_4} - \delta_{\lambda_1 \lambda_3} \delta_{\lambda_2 \lambda_4} + 4\delta_{\lambda_1 \lambda_4} \delta_{\lambda_2 \lambda_3})] D_{\lambda_1 \lambda_2 \lambda_3 \lambda_4} \end{aligned} \quad (\text{A15})$$

$$\begin{aligned}
&= \frac{1}{30} \left[\delta_{i_1 i_2} \delta_{i_3 i_4} (4D_{\lambda_1 \lambda_1 \lambda_3 \lambda_3} - D_{\lambda_1 \lambda_2 \lambda_1 \lambda_2} - D_{\lambda_1 \lambda_2 \lambda_2 \lambda_1}) \right. \\
&\quad + \delta_{i_1 i_3} \delta_{i_2 i_4} (-D_{\lambda_1 \lambda_1 \lambda_3 \lambda_3} + 4D_{\lambda_1 \lambda_2 \lambda_1 \lambda_2} - D_{\lambda_1 \lambda_2 \lambda_2 \lambda_1}) \\
&\quad \left. + \delta_{i_1 i_4} \delta_{i_2 i_3} (-D_{\lambda_1 \lambda_1 \lambda_3 \lambda_3} - D_{\lambda_1 \lambda_2 \lambda_1 \lambda_2} + 4D_{\lambda_1 \lambda_2 \lambda_2 \lambda_1}) \right]
\end{aligned}$$

The terms that survive must satisfy either $i_1 = i_2 = i_3 = i_4$ or must have two pairs of equal indices that are not equal to each other; for example, $i_1 = i_2$ and $i_3 = i_4$ but $i_1 \neq i_3$. Explicitly expressing Eq. (A15) in terms of x, y and z,

$$\begin{aligned}
\bar{D}_{xxxx} &= \bar{D}_{yyyy} = \bar{D}_{zzzz} = \frac{1}{15} (D_{\lambda_1 \lambda_1 \lambda_3 \lambda_3} + D_{\lambda_1 \lambda_2 \lambda_1 \lambda_2} + D_{\lambda_1 \lambda_2 \lambda_2 \lambda_1}) \\
\bar{D}_{xxyy} &= \bar{D}_{xxzz} = \bar{D}_{yyzz} \dots = \frac{1}{30} (4D_{\lambda_1 \lambda_1 \lambda_3 \lambda_3} - D_{\lambda_1 \lambda_2 \lambda_1 \lambda_2} - D_{\lambda_1 \lambda_2 \lambda_2 \lambda_1}) \\
\bar{D}_{xyxy} &= \bar{D}_{xzzx} = \bar{D}_{yzyz} \dots = \frac{1}{30} (-D_{\lambda_1 \lambda_1 \lambda_3 \lambda_3} + 4D_{\lambda_1 \lambda_2 \lambda_1 \lambda_2} - D_{\lambda_1 \lambda_2 \lambda_2 \lambda_1}) \\
\bar{D}_{xyyx} &= \bar{D}_{xxzz} = \bar{D}_{yyzz} \dots = \frac{1}{30} (-D_{\lambda_1 \lambda_1 \lambda_3 \lambda_3} - D_{\lambda_1 \lambda_2 \lambda_1 \lambda_2} + 4D_{\lambda_1 \lambda_2 \lambda_2 \lambda_1})
\end{aligned} \tag{A16}$$

In Eq. (A16), explicitly expressing the λ_i as x, y and z gives rise to

$$\begin{aligned}
D_{\lambda_1 \lambda_1 \lambda_3 \lambda_3} &= D_{xxxx} + D_{xxyy} + D_{xxzz} + D_{yyxx} + D_{yyyy} + D_{yyzz} + D_{zzxx} + D_{zzyy} + D_{zzzz} \\
D_{\lambda_1 \lambda_2 \lambda_1 \lambda_2} &= D_{xxxx} + D_{xyxy} + D_{xzzx} + D_{yxyx} + D_{yyyy} + D_{yzyz} + D_{zxxz} + D_{zyzy} + D_{zzzz} \\
D_{\lambda_1 \lambda_2 \lambda_2 \lambda_1} &= D_{xxxx} + D_{xyyx} + D_{xzzx} + D_{yxxy} + D_{yyyy} + D_{yzzz} + D_{zxxz} + D_{zyyz} + D_{zzzz}
\end{aligned} \tag{A17}$$

Using the traceless property of octopole moments, $\Omega_{\alpha\beta\beta} = \Omega_{\beta\alpha\beta} = \Omega_{\beta\beta\alpha} = 0$, the sums in Eq. (A17), each group of three terms with the same field direction must vanish in order to preserve tracelessness. For example, $E_{xxxx} + E_{xxyy} + E_{xxzz} = 0$ and $E_{zxxz} + E_{zyzy} + E_{zzzz} = 0$. Therefore, $E_{\lambda_1 \lambda_1 \lambda_3 \lambda_3}$, $E_{\lambda_1 \lambda_2 \lambda_1 \lambda_2}$ and $E_{\lambda_1 \lambda_2 \lambda_2 \lambda_1}$ in Eq. (A17) are all zero, which means the rotational averaged $\bar{E}_{i_1 i_2 i_3 i_4}$ in Eq. (A15) is zero.

This completes the proof for the vanishing dipole-octopole polarizability upon rotational averaging. The derivation follows closely Appendix 2 of Craig and Thirunamachandran.⁶⁰

(B) Derivation of Isotropic R⁻⁸ Dispersion

Since the dipole-quadrupole polarizability and dipole-octopole polarizability are rotationally averaged to zero, the non-vanishing terms in the anisotropic R⁻⁸ dispersion energy (Eq. 12) upon rotation are the ones involving the quadrupole-quadrupole polarizability tensors:

$$E_8^{disp}(\text{non - vanishing upon rotational averaging}) = - \sum_{\alpha\beta\gamma\kappa\mu\nu}^{x,y,z} T_{\alpha\beta\gamma}^{AB} T_{\kappa\mu\nu}^{AB} \frac{\hbar}{\pi} \left[\frac{1}{6} \sum_{n=1}^{12} W(n) Z_{n0} \alpha_{\alpha\kappa}^A(i\omega_n) C_{\beta\gamma,\mu\nu}^B(i\omega_n) + \frac{1}{6} \sum_{n=1}^{12} W(n) Z_{n0} C_{\alpha\beta,\kappa\mu}^A(i\omega_n) \alpha_{\gamma\nu}^B(i\omega_n) \right] \quad (\text{B1})$$

Since the first term and the second term of Eq (B1) are equivalent for derivation purposes, we can focus on the first term.

$$-\frac{\hbar}{\pi} \frac{1}{6} \sum_{\alpha\beta\gamma\kappa\mu\nu}^{x,y,z} T_{\alpha\beta\gamma}^{AB} T_{\kappa\mu\nu}^{AB} \sum_{n=1}^{12} W(n) Z_{n0} \alpha_{\alpha\kappa}^A(i\omega_n) C_{\beta\gamma,\mu\nu}^B(i\omega_n) \quad (\text{B2})$$

The isotropic dipole polarizability is well known.

$$\bar{\alpha} = \frac{1}{3} \delta_{\alpha\kappa} \alpha_{\alpha\kappa} = \frac{\alpha_{xx} + \alpha_{yy} + \alpha_{zz}}{3} \quad (\text{B3})$$

All isotropic 4th-rank tensors take the form

$$C_{ijkl} = a \cdot P_{ijkl} + b \cdot Q_{ijkl} + c \cdot R_{ijkl} \quad (\text{B4})$$

where a, b, and c are constants and the elements of P , Q and R are given by

$$\begin{aligned} P_{ijkl} &= \delta_{ij} \delta_{kl} \\ Q_{ijkl} &= \delta_{ik} \delta_{jl} \\ R_{ijkl} &= \delta_{il} \delta_{jk} \end{aligned} \quad (\text{B5})$$

So Eq (B2) becomes

$$-\sum_{\alpha\beta\gamma\mu\nu}^{x,y,z} T_{\alpha\beta\gamma}^{AB} T_{\alpha\mu\nu}^{AB} \frac{\hbar}{\pi} \frac{1}{6} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) [a^B(i\omega_n) \cdot P_{\beta\gamma,\mu\nu} + b^B(i\omega_n) \cdot Q_{\beta\gamma,\mu\nu} + c^B(i\omega_n) \cdot R_{\beta\gamma,\mu\nu}] \quad (\text{B6})$$

Splitting Eq (B6) into components resulting from P , Q and R gives

$$\begin{aligned} &-\frac{\hbar}{6\pi} \sum_{\alpha\beta\gamma\mu\nu}^{x,y,z} T_{\alpha\beta\gamma}^{AB} T_{\alpha\mu\nu}^{AB} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) a^B(i\omega_n) \cdot P_{\beta\gamma,\mu\nu} \\ &-\frac{\hbar}{6\pi} \sum_{\alpha\beta\gamma\mu\nu}^{x,y,z} T_{\alpha\beta\gamma}^{AB} T_{\alpha\mu\nu}^{AB} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) b^B(i\omega_n) \cdot Q_{\beta\gamma,\mu\nu} \end{aligned} \quad (\text{B7})$$

$$-\frac{\hbar}{6\pi} \sum_{\alpha\beta\gamma\mu\nu}^{x,y,z} T_{\alpha\beta\gamma}^{AB} T_{\alpha\mu\nu}^{AB} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) c^B(i\omega_n) \cdot R_{\beta\gamma,\mu\nu}$$

Since $P_{\beta\gamma,\mu\nu} = \delta_{\beta\gamma}\delta_{\mu\nu}$, it is only necessary to consider terms with $\beta = \gamma$ and $\mu = \nu$ for the sum involving P:

$$\begin{aligned} & -\frac{\hbar}{6\pi} \sum_{\alpha\beta\gamma\mu\nu}^{x,y,z} T_{\alpha\beta\gamma}^{AB} T_{\alpha\mu\nu}^{AB} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) a^B(i\omega_n) \cdot P_{\beta\gamma,\mu\nu} \\ & = -\frac{\hbar}{6\pi} \sum_{\alpha\beta\mu}^{x,y,z} T_{\alpha\beta\beta}^{AB} T_{\alpha\mu\mu}^{AB} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) a^B(i\omega_n) \\ & = -\frac{\hbar}{6\pi} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) a^B(i\omega_n) \sum_{\alpha\beta\mu}^{x,y,z} T_{\alpha\beta\beta}^{AB} T_{\alpha\mu\mu}^{AB} \end{aligned} \quad (\text{B8})$$

However, note that

$$\begin{aligned} \sum_{\alpha\beta}^{x,y,z} T_{\alpha\beta\beta}^{AB} & = -\sum_{\alpha\beta}^{x,y,z} \frac{15R_\alpha R_\beta R_\beta - 3R^2(R_\alpha \delta_{\beta\beta} + R_\beta \delta_{\alpha\beta} + R_\beta \delta_{\alpha\beta})}{R^7} \\ & = \sum_{\alpha\beta}^{x,y,z} \frac{15R_\alpha R^2 - 3R^2(3R_\alpha + 2R_\beta \delta_{\alpha\beta})}{R^7} \\ & = \sum_{\alpha\beta}^{x,y,z} \frac{6R_\alpha R^2 - (6R^2 R_\beta \delta_{\alpha\beta})}{R^7} = 0 \end{aligned} \quad (\text{B9})$$

So, Eq (B8) vanishes and does not contribute to the energy.

Next, $Q_{\beta\gamma,\mu\nu} = \delta_{\beta\mu}\delta_{\gamma\nu}$, it is only necessary to consider terms with $\beta = \mu$ and $\gamma = \nu$ for the sum involving Q:

$$\begin{aligned} & -\frac{\hbar}{6\pi} \sum_{\alpha\beta\gamma}^{x,y,z} T_{\alpha\beta\gamma}^{AB} T_{\alpha\beta\gamma}^{AB} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) b^B(i\omega_n) \cdot Q_{\beta\gamma,\beta\gamma} \\ & = -\frac{\hbar}{6\pi} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) b^B(i\omega_n) \sum_{\alpha\beta\gamma}^{x,y,z} T_{\alpha\beta\gamma}^{AB} T_{\alpha\beta\gamma}^{AB} \\ & = -\frac{\hbar}{6\pi} \frac{90}{(4\pi\epsilon_0)^2 R_{AB}^3} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) b^B(i\omega_n) \end{aligned} \quad (\text{B10})$$

The third equality in Eq. (B10) arises from the following observation:

$$\begin{aligned} & \sum_{\alpha\beta\gamma}^{x,y,z} T_{\alpha\beta\gamma}^{AB} T_{\alpha\beta\gamma}^{AB} \\ & = \frac{1}{(4\pi\epsilon_0)^2} \sum_{\alpha\beta}^{x,y,z} \frac{15R_\alpha R_\beta R_\gamma - 3R^2(R_\alpha \delta_{\beta\gamma} + R_\beta \delta_{\alpha\gamma} + R_\gamma \delta_{\alpha\beta})}{R^7} \cdot \frac{15R_\alpha R_\beta R_\gamma - 3R^2(R_\alpha \delta_{\beta\gamma} + R_\beta \delta_{\alpha\gamma} + R_\gamma \delta_{\alpha\beta})}{R^7} \end{aligned} \quad (\text{B11})$$

$$\begin{aligned}
&= \frac{1}{(4\pi\epsilon_0)^2} \sum_{\alpha\beta}^{x,y,z} \frac{225R^6}{R^{14}} - \frac{90R^2 R_\alpha R_\beta R_\gamma (R_\alpha \delta_{\beta\gamma} + R_\beta \delta_{\alpha\gamma} + R_\gamma \delta_{\alpha\beta})}{R^{14}} + \frac{9R^4 (R_\alpha \delta_{\beta\gamma} + R_\beta \delta_{\alpha\gamma} + R_\gamma \delta_{\alpha\beta})^2}{R^{14}} \\
&= \frac{1}{(4\pi\epsilon_0)^2} \sum_{\alpha\beta}^{x,y,z} \frac{225R^6}{R^{14}} - \frac{90R^2(3R^4)}{R^{14}} + \frac{9R^4(3R^2 + 3R^2 + 3R^2 + 2R^2 + 2R^2 + 2R^2)}{R^{14}} \\
&= \frac{90}{(4\pi\epsilon_0)^2 R_{AB}^8}
\end{aligned}$$

The sum involving $R_{\beta\gamma,\mu\nu} = \delta_{\beta\nu}\delta_{\gamma\mu}$ can be simplified along similar lines:

$$\begin{aligned}
&-\frac{\hbar}{6\pi} \sum_{\alpha\beta\gamma\mu\nu}^{x,y,z} T_{\alpha\beta\gamma}^{AB} T_{\alpha\mu\nu}^{AB} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) c^B(i\omega_n) \cdot R_{\beta\gamma,\gamma\beta} \\
&= -\frac{\hbar}{6\pi} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) c^B(i\omega_n) \sum_{\alpha\beta\gamma\mu\nu}^{x,y,z} T_{\alpha\beta\gamma}^{AB} T_{\alpha\gamma\beta}^{AB} \\
&= -\frac{\hbar}{6\pi} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) c^B(i\omega_n) \sum_{\alpha\beta\gamma\mu\nu}^{x,y,z} T_{\alpha\beta\gamma}^{AB} T_{\alpha\beta\gamma}^{AB} \\
&= -\frac{\hbar}{6\pi} \frac{90}{(4\pi\epsilon_0)^2 R_{AB}^8} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) c^B(i\omega_n)
\end{aligned} \tag{B12}$$

Adding Eqs. B10 and B12 yields

$$-\frac{\hbar}{6\pi} \frac{90}{(4\pi\epsilon_0)^2 R_{AB}^8} \sum_{n=1}^{12} W(n) Z_{n0} \bar{\alpha}^A(i\omega_n) [b^B(i\omega_n) + c^B(i\omega_n)] \tag{B13}$$

The derivation for the second term of Eq. B1 can be done similarly and then combined with Eq. B13, resulting in

$$-\frac{\hbar}{6\pi} \frac{90}{(4\pi\epsilon_0)^2 R_{AB}^8} \sum_{n=1}^{12} W(n) Z_{n0} \left\{ \begin{aligned} &\bar{\alpha}^A(i\omega_n) \cdot [b^B(i\omega_n) + c^B(i\omega_n)] \\ &+ \bar{\alpha}^B(i\omega_n) \cdot [b^A(i\omega_n) + c^A(i\omega_n)] \end{aligned} \right\} \tag{B14}$$

Now, finding the rotational average of the quadrupole-quadrupole polarizability C_{ijkl} , a 4th rank tensor, is equivalent to finding the projection of C onto the space of the isotropic tensors spanned by P, Q and R (Eq. B5). However, P, Q and R are **not** orthogonal with respect to the tensor inner product. An orthogonal basis can be used instead:

$$\begin{aligned}
L_{ijkl} &= \frac{1}{3} \delta_{ij} \delta_{kl} = \frac{1}{3} P_{ijkl} \\
M_{ijkl} &= \frac{1}{2} (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) - \frac{1}{3} \delta_{ij} \delta_{kl} = \frac{1}{2} (Q_{ijkl} + R_{ijkl}) - \frac{1}{3} P_{ijkl} \\
N_{ijkl} &= \frac{1}{2} (\delta_{ik} \delta_{jl} - \delta_{il} \delta_{jk}) = \frac{1}{2} (Q_{ijkl} - R_{ijkl})
\end{aligned} \tag{B15}$$

The inner product of this basis is

$$\begin{aligned}
L \cdot L &= 1 \\
M \cdot M &= 5 \\
N \cdot N &= 3
\end{aligned} \tag{B16}$$

The projection of C onto the space of this set of isotropic tensors can be written as

$$\sum_{\beta\gamma\mu\nu}^{x,y,z} [C_{\beta\gamma\mu\nu}(i\omega)L_{\beta\gamma\mu\nu}] \cdot L_{\beta\gamma\mu\nu} + \frac{1}{5} [C_{\beta\gamma\mu\nu}(i\omega)M_{\beta\gamma\mu\nu}] \cdot M_{\beta\gamma\mu\nu} + \frac{1}{3} [C_{\beta\gamma\mu\nu}(i\omega)N_{\beta\gamma\mu\nu}] \cdot N_{\beta\gamma\mu\nu} \tag{B17}$$

The $b^B(i\omega_n)$ and $c^B(i\omega_n)$ in Eq. A13 can be obtained using Eq. B17 and the definitions given in Eq. B15.

$$\begin{aligned}
b(i\omega_n) &= \sum_{\beta\gamma\mu\nu}^{x,y,z} \frac{1}{2} \cdot \frac{1}{5} [C_{\beta\gamma\mu\nu}(i\omega)M_{\beta\gamma\mu\nu}] + \frac{1}{2} \cdot \frac{1}{3} [C_{\beta\gamma\mu\nu}(i\omega)N_{\beta\gamma\mu\nu}] \\
c(i\omega_n) &= \sum_{\beta\gamma\mu\nu}^{x,y,z} \frac{1}{2} \cdot \frac{1}{5} [C_{\beta\gamma\mu\nu}(i\omega)M_{\beta\gamma\mu\nu}] - \frac{1}{2} \cdot \frac{1}{3} [C_{\beta\gamma,\mu\nu}(i\omega)N_{\beta\gamma\mu\nu}]
\end{aligned} \tag{B18}$$

Substituting Eq. B18 into Eq. B14 yields

$$\begin{aligned}
& -\frac{\hbar}{6\pi(4\pi\epsilon_0)^2 R_{AB}^8} \sum_{n=1}^{12} W(n)Z_{n0} \left\{ \bar{\alpha}^A(i\omega_n) \cdot [b^B(i\omega_n) + c^B(i\omega_n)] \right. \\
& \left. + \bar{\alpha}^B(i\omega_n) \cdot [b^A(i\omega_n) + c^A(i\omega_n)] \right\} \\
& = -\frac{\hbar}{6\pi(4\pi\epsilon_0)^2 R_{AB}^8} \sum_{n=1}^{12} W(n)Z_{n0} \left\{ \bar{\alpha}^A(i\omega_n) \cdot \left[\frac{1}{5} \sum_{\beta\gamma\mu\nu}^{x,y,z} [C_{\beta\gamma\mu\nu}^B(i\omega)M_{\beta\gamma\mu\nu}] \right] \right. \\
& \left. + \bar{\alpha}^B(i\omega_n) \cdot \left[\frac{1}{5} \sum_{\beta\gamma\mu\nu}^{x,y,z} [C_{\beta\gamma\mu\nu}^A(i\omega)M_{\beta\gamma\mu\nu}] \right] \right\}
\end{aligned} \tag{B19}$$

and $\sum_{\beta\gamma\mu\nu}^{x,y,z} C_{\beta\gamma\mu\nu}^B(i\omega)M_{\beta\gamma\mu\nu}$ is a scalar. Note that $\frac{1}{5} \sum_{\beta\gamma\mu\nu}^{x,y,z} [C_{\beta\gamma\mu\nu}^B(i\omega)M_{\beta\gamma\mu\nu}]$ equivalent to the Buckingham definition (Eq. 15)

So, the isotropic R⁻⁸ dispersion is

$$-\frac{15\hbar}{\pi(4\pi\epsilon_0)^2 R_{AB}^8} \sum_{n=1}^{12} W(n)Z_{n0} \left\{ \bar{\alpha}^A(i\omega_n) \cdot \bar{C}^B(i\omega_n) \right. \\
\left. + \bar{\alpha}^B(i\omega_n) \cdot \bar{C}^A(i\omega_n) \right\} \tag{B20}$$

(C) Octopole origin-shift formulation

Recall the general forms of a quadrupole moment (θ) and an octopole moment (Ω) in cartesian coordinates, where \mathbf{r} is the vector relative to an expansion center and subscripts α, β and γ represent x, y and z, and

$$\theta_{\alpha\beta} = q \left(\frac{3}{2} r_\alpha r_\beta - \frac{1}{2} r^2 \delta_{\alpha\beta} \right) \quad (\text{C1})$$

$$\Omega_{\alpha\beta\gamma} = q \left[\frac{5}{2} r_\alpha r_\beta r_\gamma - \frac{1}{2} r^2 (r_\alpha \delta_{\beta\gamma} + r_\beta \delta_{\alpha\gamma} + r_\gamma \delta_{\alpha\beta}) \right] \quad (\text{C2})$$

Let the origin of the octopole moment be shifted by $-\mathbf{r}'$. Then,

$$\begin{aligned} \Omega'_{\alpha\beta\gamma} &= q \left[\frac{5}{2} (r_\alpha - r'_\alpha)(r_\beta - r'_\beta)(r_\gamma - r'_\gamma) \right. \\ &\quad \left. - \frac{1}{2} (r - r')^2 \left((r_\alpha - r'_\alpha) \delta_{\beta\gamma} + (r_\beta - r'_\beta) \delta_{\alpha\gamma} + (r_\gamma - r'_\gamma) \delta_{\alpha\beta} \right) \right] \\ &= q \left[\frac{5}{2} (r_\alpha - r'_\alpha)(r_\beta - r'_\beta)(r_\gamma - r'_\gamma) \right. \\ &\quad \left. - \frac{1}{2} (r^2 - 2r_\kappa r'_\kappa + r'^2) \left((r_\alpha - r'_\alpha) \delta_{\beta\gamma} + (r_\beta - r'_\beta) \delta_{\alpha\gamma} + (r_\gamma - r'_\gamma) \delta_{\alpha\beta} \right) \right] \\ &= q \left\{ \frac{5}{2} (\mathbf{r}_\alpha \mathbf{r}_\beta \mathbf{r}_\gamma - r'_\alpha r'_\beta r'_\gamma - r_\alpha r'_\beta r'_\gamma - r_\alpha r_\beta r'_\gamma + r'_\alpha r'_\beta r'_\gamma + r'_\alpha r_\beta r'_\gamma + r_\alpha r'_\beta r'_\gamma - r'_\alpha r'_\beta r'_\gamma) \right. \\ &\quad - \frac{1}{2} r^2 [(\mathbf{r}_\alpha) \delta_{\beta\gamma} + (\mathbf{r}_\beta) \delta_{\alpha\gamma} + (\mathbf{r}_\gamma) \delta_{\alpha\beta}] \\ &\quad - \frac{1}{2} r^2 [(-r'_\alpha) \delta_{\beta\gamma} + (-r'_\beta) \delta_{\alpha\gamma} + (-r'_\gamma) \delta_{\alpha\beta}] \\ &\quad + r_\kappa r'_\kappa [(r_\alpha - r'_\alpha) \delta_{\beta\gamma} + (r_\beta - r'_\beta) \delta_{\alpha\gamma} + (r_\gamma - r'_\gamma) \delta_{\alpha\beta}] \\ &\quad \left. - \frac{1}{2} r'^2 [(r_\alpha - r'_\alpha) \delta_{\beta\gamma} + (r_\beta - r'_\beta) \delta_{\alpha\gamma} + (r_\gamma - r'_\gamma) \delta_{\alpha\beta}] \right\} \end{aligned} \quad (\text{C3})$$

where the **bold** terms correspond to the original (un-shifted) octopole moment.

$$\begin{aligned} \Omega'_{\alpha\beta\gamma} &= \mathbf{\Omega}_{\alpha\beta\gamma} + q \left\{ \frac{5}{2} (-r'_\alpha r'_\beta r'_\gamma - r_\alpha r'_\beta r'_\gamma - r_\alpha r_\beta r'_\gamma + r'_\alpha r'_\beta r'_\gamma + r'_\alpha r_\beta r'_\gamma + r_\alpha r'_\beta r'_\gamma - r'_\alpha r'_\beta r'_\gamma) \right. \\ &\quad + \frac{1}{2} r^2 [(r'_\alpha) \delta_{\beta\gamma} + (r'_\beta) \delta_{\alpha\gamma} + (r'_\gamma) \delta_{\alpha\beta}] \\ &\quad + r_\kappa r'_\kappa (r_\alpha \delta_{\beta\gamma} - r'_\alpha \delta_{\beta\gamma} + r_\beta \delta_{\alpha\gamma} - r'_\beta \delta_{\alpha\gamma} + r_\gamma \delta_{\alpha\beta} - r'_\gamma \delta_{\alpha\beta}) \\ &\quad \left. - \frac{1}{2} r'^2 (r_\alpha \delta_{\beta\gamma} - r'_\alpha \delta_{\beta\gamma} + r_\beta \delta_{\alpha\gamma} - r'_\beta \delta_{\alpha\gamma} + r_\gamma \delta_{\alpha\beta} - r'_\gamma \delta_{\alpha\beta}) \right\} \end{aligned} \quad (\text{C4})$$

Now, separate the rest of the RHS of Eq. (C4) into two parts, which contain the one- and two-coordinate shifted vector component(s), respectively.

$$\Omega'_{\alpha\beta\gamma} = \mathbf{\Omega}_{\alpha\beta\gamma} + \text{term1} + \text{term2}$$

$$\begin{aligned} \text{term1} &= q \left\{ \frac{5}{2} (-r'_\alpha r'_\beta r'_\gamma - r_\alpha r'_\beta r'_\gamma - r_\alpha r_\beta r'_\gamma) + r_\kappa r'_\kappa (r_\alpha \delta_{\beta\gamma} + r_\beta \delta_{\alpha\gamma} + r_\gamma \delta_{\alpha\beta}) \right. \\ &\quad \left. + \frac{1}{2} r^2 [(r'_\alpha) \delta_{\beta\gamma} + (r'_\beta) \delta_{\alpha\gamma} + (r'_\gamma) \delta_{\alpha\beta}] \right\} \end{aligned} \quad (\text{C5})$$

$$\text{term2} = \left\{ \frac{5}{2} (r'_\alpha r'_\beta \mu_\gamma + r'_\alpha \mu_\beta r'_\gamma + \mu_\alpha r'_\beta r'_\gamma - q r'_\alpha r'_\beta r'_\gamma) - \mu_\kappa r'_\kappa (r'_\alpha \delta_{\beta\gamma} + r'_\beta \delta_{\alpha\gamma} + r'_\gamma \delta_{\alpha\beta}) - \frac{1}{2} r'^2 (\mu_\alpha \delta_{\beta\gamma} - q r'_\alpha \delta_{\beta\gamma} + \mu_\beta \delta_{\alpha\gamma} - q r'_\beta \delta_{\alpha\gamma} + \mu_\gamma \delta_{\alpha\beta} - q r'_\gamma \delta_{\alpha\beta}) \right\}$$

Now, recast term1 in Eq. (C4) in the form of the quadrupole moment,

$$\begin{aligned} \text{term1} &= -\frac{5}{2} r'_\alpha \frac{2}{3} q \left(\frac{3}{2} r_\beta r_\gamma - \frac{1}{2} r^2 \delta_{\beta\gamma} + \frac{1}{2} r^2 \delta_{\beta\gamma} \right) - \frac{5}{2} r'_\beta \frac{2}{3} q \left(\frac{3}{2} r_\alpha r_\gamma - \frac{1}{2} r^2 \delta_{\alpha\gamma} + \frac{1}{2} r^2 \delta_{\alpha\gamma} \right) \\ &\quad - \frac{5}{2} r'_\gamma \frac{2}{3} q \left(\frac{3}{2} r_\alpha r_\beta - \frac{1}{2} r^2 \delta_{\alpha\beta} + \frac{1}{2} r^2 \delta_{\alpha\beta} \right) \\ &\quad + r'_\kappa \delta_{\beta\gamma} \frac{2}{3} q \left(\frac{3}{2} r_\kappa r_\alpha - \frac{1}{2} r^2 \delta_{\kappa\alpha} + \frac{1}{2} r^2 \delta_{\kappa\alpha} \right) \\ &\quad + r'_\kappa \delta_{\alpha\gamma} \frac{2}{3} q \left(\frac{3}{2} r_\kappa r_\beta - \frac{1}{2} r^2 \delta_{\kappa\beta} + \frac{1}{2} r^2 \delta_{\kappa\beta} \right) \\ &\quad + r'_\kappa \delta_{\alpha\beta} \frac{2}{3} q \left(\frac{3}{2} r_\kappa r_\gamma - \frac{1}{2} r^2 \delta_{\kappa\gamma} + \frac{1}{2} r^2 \delta_{\kappa\gamma} \right) + \frac{1}{2} q r^2 [(r'_\alpha) \delta_{\beta\gamma} + (r'_\beta) \delta_{\alpha\gamma} + (r'_\gamma) \delta_{\alpha\beta}] \\ &= -\frac{5}{3} r'_\alpha (\theta_{\beta\gamma} + \frac{1}{2} q r^2 \delta_{\beta\gamma}) + \frac{1}{2} q r'_\alpha r^2 \delta_{\beta\gamma} - \frac{5}{3} r'_\beta (\theta_{\alpha\gamma} + \frac{1}{2} q r^2 \delta_{\alpha\gamma}) + \frac{1}{2} q r'_\beta r^2 \delta_{\alpha\gamma} \\ &\quad - \frac{5}{3} r'_\gamma (\theta_{\alpha\beta} + \frac{1}{2} q r^2 \delta_{\alpha\beta}) + \frac{1}{2} q r'_\gamma r^2 \delta_{\alpha\beta} + r'_\kappa \delta_{\beta\gamma} \frac{2}{3} (\theta_{\kappa\alpha} + \frac{1}{2} q r^2 \delta_{\kappa\alpha}) \\ &\quad + r'_\kappa \delta_{\alpha\gamma} \frac{2}{3} (\theta_{\kappa\beta} + \frac{1}{2} q r^2 \delta_{\kappa\beta}) + r'_\kappa \delta_{\alpha\beta} \frac{2}{3} (\theta_{\kappa\gamma} + \frac{1}{2} q r^2 \delta_{\kappa\gamma}) \\ &= -\frac{5}{3} r'_\alpha \theta_{\beta\gamma} - \frac{1}{3} q r'_\alpha r^2 \delta_{\beta\gamma} - \frac{5}{3} r'_\beta \theta_{\alpha\gamma} - \frac{1}{3} q r'_\beta r^2 \delta_{\alpha\gamma} - \frac{5}{3} r'_\gamma \theta_{\alpha\beta} - \frac{1}{3} q r'_\gamma r^2 \delta_{\alpha\beta} \\ &\quad + \left(\frac{2}{3} r'_\kappa \delta_{\beta\gamma} \theta_{\kappa\alpha} + \frac{1}{3} q r'_\kappa \delta_{\beta\gamma} r^2 \delta_{\kappa\alpha} \right) + \left(\frac{2}{3} r'_\kappa \delta_{\alpha\gamma} \theta_{\kappa\beta} + \frac{1}{3} q r'_\kappa \delta_{\alpha\gamma} r^2 \delta_{\kappa\beta} \right) \\ &\quad + \left(\frac{2}{3} r'_\kappa \delta_{\alpha\beta} \theta_{\kappa\gamma} + \frac{1}{3} q r'_\kappa \delta_{\alpha\beta} r^2 \delta_{\kappa\gamma} \right) \\ &= -\frac{5}{3} r'_\alpha \theta_{\beta\gamma} - \frac{5}{3} r'_\beta \theta_{\alpha\gamma} - \frac{5}{3} r'_\gamma \theta_{\alpha\beta} + \frac{2}{3} r'_\kappa \delta_{\beta\gamma} \theta_{\kappa\alpha} + \frac{2}{3} r'_\kappa \delta_{\alpha\gamma} \theta_{\kappa\beta} + \frac{2}{3} r'_\kappa \delta_{\alpha\beta} \theta_{\kappa\gamma} \end{aligned} \quad (\text{C6})$$

The **bold** terms in Eq (C6) cancel out because

$$\frac{1}{3} r'_\kappa \delta_{\beta\gamma} r^2 \delta_{\kappa\alpha} = \frac{1}{3} (r'_\kappa \delta_{\kappa\alpha}) \delta_{\beta\gamma} r^2 = \frac{1}{3} r'_\alpha r^2 \delta_{\beta\gamma} \quad (\text{C7})$$

So, the final form of the origin-shifted octopole moment is

$$\begin{aligned} \Omega'_{\alpha\beta\gamma} &= \Omega_{\alpha\beta\gamma} + \left\{ -\frac{5}{3} (r'_\alpha \theta_{\beta\gamma} + r'_\beta \theta_{\alpha\gamma} + r'_\gamma \theta_{\alpha\beta}) + \frac{2}{3} r'_\kappa (\theta_{\kappa\alpha} \delta_{\beta\gamma} + \theta_{\kappa\beta} \delta_{\alpha\gamma} + \theta_{\kappa\gamma} \delta_{\alpha\beta}) \right\} \\ &\quad + \left\{ \frac{5}{2} (r'_\alpha r'_\beta \mu_\gamma + r'_\alpha \mu_\beta r'_\gamma + \mu_\alpha r'_\beta r'_\gamma) - \mu_\kappa r'_\kappa (r'_\alpha \delta_{\beta\gamma} + r'_\beta \delta_{\alpha\gamma} + r'_\gamma \delta_{\alpha\beta}) \right. \\ &\quad \left. - \frac{1}{2} r'^2 (\mu_\alpha \delta_{\beta\gamma} + \mu_\beta \delta_{\alpha\gamma} + \mu_\gamma \delta_{\alpha\beta}) - \frac{5}{2} q r'_\alpha r'_\beta r'_\gamma + \frac{1}{2} q r'^2 (r'_\alpha \delta_{\beta\gamma} + r'_\beta \delta_{\alpha\gamma} + r'_\gamma \delta_{\alpha\beta}) \right\} \end{aligned} \quad (\text{C8})$$

(D) Analytic gradient of isotropic R^{-8} dispersion

The expression for the damped disp8 term is $E_8 = f_8 \cdot E_{8,0}$, where $E_{8,0}$ is the undamped disp8. So, the general expression for the disp8 gradient is

$$\frac{\partial E_8}{\partial q} = \frac{\partial(f_8 E_{8,0})}{\partial q} = \frac{\partial f_8}{\partial q} E_{8,0} + f_8 \frac{\partial E_{8,0}}{\partial q} \quad (D1)$$

where q is an EFP fragment degree of freedom. Since EFP fragments are rigid, the relevant degrees of freedom are the translation of the center of mass of a fragment and the rotation about the center of mass of a fragment.

First, focus on the derivative of the energy expression $E_{8,0}$.

$$\begin{aligned} \frac{dE_{8,0}^{disp,iso}}{dq_A} = & -\frac{15\hbar}{\pi} \sum_{k \in A}^{LMO} \sum_{l \in B}^{LMO} \frac{d}{dq_A} \left[\left(\frac{1}{(R_{kl})^8} \right) \sum_{n=1}^{12} W(n) \frac{2\omega_0}{(1-t_n)^2} [\bar{\alpha}^k(i\omega_n) \bar{C}^l(i\omega_n) \right. \\ & \left. + \bar{C}^k(i\omega_n) \bar{\alpha}^l(i\omega_n)] \right] \end{aligned} \quad (D2)$$

Since the isotropic dipole polarizability and the isotropic quadrupole polarizability are constants at a given imaginary frequency, their derivatives are zero. Only the derivative of R^{-8} is required, which can be straightforwardly derived, for example, for a specific case, dx_A

$$\frac{dR_{kl}^{-8}}{dx_A} = \frac{d[(x_k - x_l)^2 + (y_k - y_l)^2 + (z_k - z_l)^2]^{-4}}{dx_A} = -4R_{kl}^{-10} \cdot 2(x_k - x_l) \cdot \frac{dx_k}{dx_A} = -8 \frac{(x_k - x_l)}{R_{kl}^{10}} \quad (D3)$$

In Eq. (D3), the $\frac{dx_k}{dx_A}$ term is one, because the displacement of the centroid of LMO i is the same as that of the center of mass of fragment A (x_A). Therefore, Eq. (D2) becomes

$$\begin{aligned} \frac{dE_{8,0}^{disp,iso}}{dx_A} = & -\frac{15\hbar}{\pi} \sum_{i \in A}^{LMO} \sum_{j \in B}^{LMO} \sum_{n=1}^{12} -8 \frac{(x_k - x_l)}{R_{kl}^{10}} W(n) \frac{2\omega_0}{(1-t_n)^2} [\bar{\alpha}^k(i\omega_n) \bar{C}^l(i\omega_n) \\ & + \bar{C}^k(i\omega_n) \bar{\alpha}^l(i\omega_n)] \end{aligned} \quad (D4)$$

Next, consider the derivative of the damping functions.

Overlap-based damping function

$$\frac{\partial f_8^S(k, l)}{\partial x_A} = \frac{\partial}{\partial x} \left[1 - S_{kl}^2 \sum_{n=0}^8 \frac{(-2\ln|S_{kl}|)^n}{n!} \right] \quad (D5)$$

Let $RB = -2\ln|S_{kl}|$ for simplicity. The derivative of the powers of RB is

$$\frac{\partial}{\partial x} [-2\ln|S_{kl}|] = \frac{-2}{|S_{kl}|} \frac{\partial S_{kl}}{\partial x} \quad (D6)$$

$$\frac{\partial}{\partial x} [-2\ln|S_{kl}|]^n = -2n \frac{1}{|S_{kl}|} \frac{\partial S_{kl}}{\partial x} [-2\ln|S_{kl}|]^{n-1} = -2 \frac{1}{|S_{kl}|} \frac{\partial S_{kl}}{\partial x} nRB^{n-1} \quad (D7)$$

Then Eq. (D5) becomes

$$\begin{aligned}
\frac{\partial f_8^S(k, l)}{\partial x_A} &= \frac{\partial}{\partial x_A} \left\{ 1 - S_{kl}^2 \left[1 + RB^{1/2} + \frac{RB}{2} + \frac{RB^{3/2}}{6} + \frac{RB^2}{24} + \frac{RB^{5/2}}{120} + \frac{RB^3}{720} + \frac{RB^{7/2}}{5040} + \frac{RB^4}{40320} \right] \right\} \\
&= -\frac{\partial S_{kl}^2}{\partial x_A} \left[1 + RB^{1/2} + \frac{RB}{2} + \frac{RB^{3/2}}{6} + \frac{RB^2}{24} + \frac{RB^{5/2}}{120} + \frac{RB^3}{720} + \frac{RB^{7/2}}{5040} + \frac{RB^4}{40320} \right] \\
&\quad - S_{kl}^2 \left[\frac{\partial}{\partial x_A} RB^{1/2} + \frac{\partial}{\partial x_A} \frac{RB}{2} + \frac{\partial}{\partial x_A} \frac{RB^{3/2}}{6} + \frac{\partial}{\partial x_A} \frac{RB^2}{24} + \frac{\partial}{\partial x_A} \frac{RB^{5/2}}{120} + \frac{\partial}{\partial x_A} \frac{RB^3}{720} \right. \\
&\quad \left. + \frac{\partial}{\partial x_A} \frac{RB^{7/2}}{5040} + \frac{\partial}{\partial x_A} \frac{RB^4}{40320} \right] \\
&= -2 \frac{\partial S_{kl}}{\partial x_A} \cdot S_{kl} \cdot \left[1 + RB^{1/2} + \frac{RB}{2} + \frac{RB^{3/2}}{6} + \frac{RB^2}{24} + \frac{RB^{5/2}}{120} + \frac{RB^3}{720} + \frac{RB^{7/2}}{5040} + \frac{RB^4}{40320} \right] + 2S_{kl}^2 \cdot \frac{1}{S_{kl}} \\
&\quad \cdot \frac{\partial S_{kl}}{\partial x_A} \left[\frac{1}{2} RB^{-1/2} + \frac{1}{2} + \frac{13}{62} RB^{1/2} + \frac{1}{24} 2RB^1 + \frac{1}{120} \frac{5}{2} RB^{3/2} + \frac{1}{720} 3RB^2 + \frac{1}{5040} \frac{7}{2} RB^{5/2} \right. \\
&\quad \left. + \frac{1}{40320} 4RB^3 \right] \\
&= -\frac{\partial S_{kl}}{\partial x_A} \cdot S_{kl} \left[2 + 2RB^{1/2} + RB + \frac{RB^{3/2}}{3} + \frac{RB^2}{12} + \frac{RB^{5/2}}{60} + \frac{RB^3}{360} + \frac{RB^{7/2}}{2520} + \frac{RB^4}{20160} - RB^{-1/2} - 1 - \frac{1}{2} RB^{1/2} - \frac{1}{6} RB^1 \right. \\
&\quad \left. - \frac{1}{24} RB^{3/2} - \frac{1}{120} RB^2 - \frac{1}{720} RB^{5/2} - \frac{1}{5040} RB^3 \right] \\
&= -\frac{\partial S_{kl}}{\partial x_A} \cdot S_{kl} \left[1 - RB^{-1/2} + \frac{3}{2} RB^{1/2} + \frac{5}{6} RB + \frac{7}{24} RB^{3/2} + \frac{9}{120} RB^2 + \frac{11}{720} RB^{5/2} + \frac{13}{5040} RB^3 + \frac{RB^{7/2}}{2520} + \frac{RB^4}{20160} \right] \quad (D8)
\end{aligned}$$

Tang-Toennies damping function

Tang-Toennies damping function has a radial dependence on the intermolecular separation, in this case, the distance between two centroids of the LMOs of two EFP fragments. Hence, its derivative can be obtained straightforwardly from the derivatives of the intermolecular separation R .

$$\begin{aligned}
\frac{\partial}{\partial x_A} f_8^{TT}(k, l) &= \frac{\partial}{\partial x_A} \left[1 - \left(1 + bR + \frac{(bR)^2}{2} + \frac{(bR)^3}{6} + \frac{(bR)^4}{24} + \frac{(bR)^5}{120} + \frac{(bR)^6}{720} + \frac{(bR)^7}{5040} + \frac{(bR)^8}{40320} \right) e^{-bR} \right] \\
&= -\frac{\partial}{\partial x_A} \left(1 + bR + \frac{(bR)^2}{2} + \frac{(bR)^3}{6} + \frac{(bR)^4}{24} + \frac{(bR)^5}{120} + \frac{(bR)^6}{720} + \frac{(bR)^7}{5040} + \frac{(bR)^8}{40320} \right) e^{-bR} \\
&\quad + b \left(1 + bR + \frac{(bR)^2}{2} + \frac{(bR)^3}{6} + \frac{(bR)^4}{24} + \frac{(bR)^5}{120} + \frac{(bR)^6}{720} + \frac{(bR)^7}{5040} + \frac{(bR)^8}{40320} \right) e^{-bR} \frac{\partial R}{\partial x_A} \quad (D9)
\end{aligned}$$

Since

$$R = [(x_k - x_l)^2 + (y_k - y_l)^2 + (z_k - z_l)^2]^{\frac{1}{2}} \quad (D10)$$

where i and j are the LMOs of fragments A and B.

$$\frac{\partial R}{\partial x_A} = 2(x_k - x_l)[(x_k - x_l)^2 + (y_k - y_l)^2 + (z_k - z_l)^2]^{-\frac{1}{2}} = \frac{2(x_k - x_l)}{R} \quad (D11)$$

Then Eq (D9) becomes

$$\begin{aligned}
\frac{\partial}{\partial x_A} f_8^{TT}(k, l) &= -e^{-bR} \left(b \frac{\partial R}{\partial x_A} + b^2 R \frac{\partial R}{\partial x_A} + \frac{b^3 R^2}{2} \frac{\partial R}{\partial x_A} + \frac{b^4 R^3}{6} \frac{\partial R}{\partial x_A} + \frac{b^5 R^4}{24} \frac{\partial R}{\partial x_A} + \frac{b^6 R^5}{120} \frac{\partial R}{\partial x_A} + \frac{b^7 R^6}{720} \frac{\partial R}{\partial x_A} \right. \\
&\quad \left. + \frac{b^8 R^7}{5040} \frac{\partial R}{\partial x_A} \right) \\
&+ e^{-bR} \left(b \frac{\partial R}{\partial x_A} + b^2 R \frac{\partial R}{\partial x_A} + \frac{b^3 R^2}{2} \frac{\partial R}{\partial x_A} + \frac{b^4 R^3}{6} \frac{\partial R}{\partial x_A} + \frac{b^5 R^4}{24} \frac{\partial R}{\partial x_A} + \frac{b^6 R^5}{120} \frac{\partial R}{\partial x_A} + \frac{b^7 R^6}{720} \frac{\partial R}{\partial x_A} \right. \\
&\quad \left. + \frac{b^8 R^7}{5040} \frac{\partial R}{\partial x_A} + \frac{b^9 R^8}{40320} \frac{\partial R}{\partial x_A} \right) \\
&= e^{-bR} \frac{b^9 R^8}{40320} \frac{\partial R}{\partial x_A}
\end{aligned} \tag{D12}$$

(E) Non-local model of dispersion

When developing a distributed formulation of dispersion, the molecules A and B are divided into regions (atoms, LMOs or functional groups), each of which is described by its own multipole moments. Then the interaction operator V becomes

$$V = T^{ab} q^a q^b + T_{\alpha}^{ab} (q^a \mu_{\alpha}^b - \mu_{\alpha}^a q^b) \dots = \sum_a^A \sum_b^B \sum_{tu} Q_t^a T_{tu}^{ab} Q_u^b \tag{E1}$$

where Q_t^a symbolizes a multipole moment of rank t centered at point a within the molecule A. The T_{tu}^{ab} are the electrostatic T tensors in terms of the distance between the expansion centers a and b , defined analogously to Eq (4). Then, the dispersion interaction yields

$$E^{disp} = -\frac{\hbar}{\pi} \sum_a^A \sum_{a'}^A \sum_b^B \sum_{b'}^B T_{tu}^{ab} T_{t'u'}^{a'b'} \int_0^{\infty} P_{tt'}^{aa'} P_{uu'}^{bb'} \tag{E2}$$

where the distributed polarizability tensor, $P_{tt'}^{aa'}$, is defined as

$$P_{tt'}^{aa'} = \sum_{n \neq 0} \frac{\langle 0 | \hat{Q}_t^a | n \rangle \langle n | \hat{Q}_{t'}^{a'} | 0 \rangle + \langle 0 | \hat{Q}_{t'}^{a'} | n \rangle \langle n | \hat{Q}_t^a | 0 \rangle}{E_n - E_0} \tag{E3}$$

If $a \neq a'$, $P_{tt'}^{aa'}$ is a non-local polarizability, otherwise, $P_{tt'}^{aa'}$ is a local polarizability. The non-local polarizabilities are neglected in the local model of dispersion. Since only the local polarizabilities ($a = a'$ and $b = b'$) are involved, the conventional dispersion model involves only a double summation over the expansion centers instead of a quadruple sum. An early exploration of a non-local dispersion model was done by Stone and Tong (ST).¹⁹ Unlike the conventional model, the non-local model involves terms that contain so-called charge-flow terms such as charge-multipole and charge-charge polarizabilities since the charge of one region can flow into that of another and is no longer a constant. ST have shown that for small exemplar molecules such as N₂ dimer, Cl₂ dimer and C₂H₂ dimer, the contributions to the dispersion energy from non-local multipole-multipole polarizabilities and charge-multipole polarizabilities are rather small. But the charge-

charge polarizability contribution is non-negligible. ST introduced a localization scheme to transform the non-local polarizabilities into local ones, through origin shifting the multipole moment operators appearing in the polarizabilities and reducing the quadruple summation to a double summation. However, as they note, this origin shifting should not be large, otherwise, the convergence of the multipole expansion will be compromised and might not converge at all.¹⁹ Later, LeSueur and Stone proposed a scheme to localize all non-local dipole polarizabilities except for the charge flows between atoms separated by two or more atoms in order to achieve reasonably transferable polarizabilities for the alkanes and conjugated polyenes.²⁰ Lillestolen and Wheatley, around the same time period, introduced a scheme to determine localized atomic polarizabilities by calculating the polarization of the molecule in a finite field and localizing the resulting atomic multipoles.⁶¹ However, the test cases for these schemes are small molecules, e.g., diatomic or triatomic gas molecules, small alkanes, etc. On the other hand, the size of EFP fragments in some applications can be tens of atoms. Whether these localization schemes can achieve good results for larger systems remains uncertain. Furthermore, how effective these schemes are for LMO polarizabilities is also unknown, compounded by the fact that localizing the orbitals for such large fragments may not be possible. It is the authors' opinion that a careful examination of the non-local polarizability contribution must be carried out for at least moderate-sized molecules, as well as polymers, proteins and conjugated systems, for which large charge flow effects are expected, before implementation.

The current dispersion implementation in EFP assumes a local model and has been successful for predicting interaction energies for a wide range of systems.⁶²⁻⁶⁹ As the natural extension of the current EFP dispersion model, the conventional model (i.e., neglecting non-local polarizabilities) is adopted for disp8. If one were to implement the non-local model within the framework of EFP, all three contributions involved, disp6, disp7 and disp8, should all be considered in order to achieve a balanced approach, which is beyond the scope of the current study. A distributed non-local dispersion model will be considered in a future study.

Supporting Information. Dispersion energies at single configuration; Dispersion energies along potential energy surface; Analytic gradient and Numerical gradient comparison; OpenMP parallelization

Acknowledgements. The derivation and implementation, as well as the initial testing of the code in this work was supported by a Department of Energy Exascale Computing Project, under Project Number 17-SC-20-SC to Ames Laboratory. Ames Laboratory is operated by Iowa State University under Contract No. DE-AC02-07CH11338. Further work on databases was supported by the Department of Energy Science Undergraduate Research Internship (SULI) program. The SULI work performed by William O'Brien was in collaboration with the US Department of Energy, Ames National Laboratory, and Iowa State University. All of the computations were performed on a DOE Ames Laboratory computing cluster (CJ) (Intel(R) Xeon(R) CPU E5-2695 v2 @2.40

GHz 130 Gb of memory and Intel(R) Xeon(R) CPU E5-2699 v3 @ 2.30 GHz and 256 Gb of memory)

References

- (1) London, F. The General Theory of Molecular Forces. *Transactions of the Faraday Society* **1937**, 33, 8–26. <https://doi.org/10.1039/tf937330008b>.
- (2) Margenau, H. The Role of Quadrupole Forces in Van Der Waals Attractions. *Physical Review* **1931**, 38 (4), 747. <https://doi.org/10.1103/PhysRev.38.747>.
- (3) Hirschfelder, J. O.; Löwdin, P. O. Long-Range Interaction of Two Ls-Hydrogen Atoms Expressed in Terms of Natural Spin-Orbitals. *Mol Phys* **1959**, 2 (3), 229–258. <https://doi.org/10.1080/00268975900100231>.
- (4) Buckingham, A. D. Theory of Long-Range Dispersion Forces. *Discuss Faraday Soc* **1965**, No. 40, 232–238. <https://doi.org/10.1039/df9654000232>.
- (5) Buckingham, A. D. Permanent and Induced Molecular Moments and Long-Range Intermolecular Forces. *Advances in Chemical Physics*. January 1, 1967, pp 107–142. <https://doi.org/https://doi.org/10.1002/9780470143582.ch2>.
- (6) Unsöld, A. Quantentheorie Des Wasserstoffmolekülions Und Der Born-Landéschen Abstoßungskräfte. *Zeitschrift für Physik* **1927**, 43 (8), 563–574. <https://doi.org/10.1007/BF01397633>.
- (7) Lekkerkerker, H. N. W.; Coulon, Ph.; Luyckx, R. Dispersion Forces between Closed Shell Atoms. *Physica A: Statistical Mechanics and its Applications* **1977**, 88 (2), 375–384. [https://doi.org/https://doi.org/10.1016/0378-4371\(77\)90011-5](https://doi.org/https://doi.org/10.1016/0378-4371(77)90011-5).
- (8) Coulon, Ph.; Luyckx, R.; Lekkerkerker, H. N. W. Dispersion Forces between Linear Molecules. *J Chem Phys* **1979**, 71 (8), 3462–3466. <https://doi.org/10.1063/1.438735>.
- (9) Luyckx, R.; Coulon, Ph.; Lekkerkerker, H. N. W. Dispersion Forces between Noble Gas Atoms. *J Chem Phys* **1978**, 69 (6), 2424–2427. <https://doi.org/10.1063/1.436927>.
- (10) Coulon, P.; Luyckx, R.; Lekkerkerker, H. N. W. Approximate Calculation of the Dynamic Polarizabilities and Dispersion Interaction for Ethylene Molecules. *Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics* **1981**, 77 (1), 201–207. <https://doi.org/10.1039/F29817700201>.
- (11) Casimir, H. B. G.; Polder, D. The Influence of Retardation on the London-van Der Waals Forces. *Physical Review* **1948**, 73, 360–372. <https://doi.org/10.1103/PhysRev.73.360>.
- (12) Wormer, P. E. S.; Mulder, F.; Van Der Avoird, A. Quantum Theoretical Calculations of van Der Waals Interactions between Molecules. Anisotropic Long Range Interactions. *Int J Quantum Chem* **1977**, 11 (6), 959–970. <https://doi.org/10.1002/QUA.560110608>.
- (13) Visser, F.; Wormer, P. E. S. The Non-Empirical Calculation of Second-Order Molecular Properties by Means of Effective States. II. Effective TDCHF Spectra for NO+ CO, CO₂, and C₂H₂. *Chem Phys* **1985**, 92 (1), 129–140. [https://doi.org/10.1016/0301-0104\(85\)80012-4](https://doi.org/10.1016/0301-0104(85)80012-4).
- (14) Visser, F.; Wormer, P. E. S.; Stam, P. Time-dependent Coupled Hartree–Fock Calculations of Multipole Polarizabilities and Dispersion Interactions in van Der Waals Dimers Consisting of He, H₂, Ne, and N₂. *J Chem Phys* **1998**, 79 (10), 4973. <https://doi.org/10.1063/1.445591>.

- (15) Visser, F.; Wormer, P. E. S.; Stam, P. Erratum: Time-dependent Coupled Hartree–Fock Calculations of Multipole Polarizabilities and Dispersion Interactions in van Der Waals Dimers Consisting of He, H₂, Ne, and N₂ [*J. Chem. Phys.* 79, 4973 (1983)]. *J Chem Phys* **1998**, 81 (8), 3755. <https://doi.org/10.1063/1.448203>.
- (16) Mulder, F.; van der Avoird, A.; Wormer, P. E. S. Anisotropy of Long Range Interactions between Linear Molecules: H₂-H₂ and H₂-He. *Mol Phys* **1979**, 37 (1), 159–180. <https://doi.org/10.1080/00268977900100131>.
- (17) Wormer, P. E. S.; Hettema, H. Many-body Perturbation Theory of Frequency-dependent Polarizabilities and van Der Waals Coefficients: Application to H₂O–H₂O and Ar–NH₃. *J Chem Phys* **1992**, 97 (8), 5592. <https://doi.org/10.1063/1.463767>.
- (18) Thakkar, A. J.; Hettema, H.; Wormer, P. E. S. Ab Initio Dispersion Coefficients for Interactions Involving Rare-gas Atoms. *J Chem Phys* **1992**, 97 (5), 3252. <https://doi.org/10.1063/1.463012>.
- (19) Stone, A. J.; Tong, C. S. Local and Non-Local Dispersion Models. *Chem Phys* **1989**, 137 (1–3), 121–135. [https://doi.org/10.1016/0301-0104\(89\)87098-3](https://doi.org/10.1016/0301-0104(89)87098-3).
- (20) Misquitta, A. J.; Stone, A. J. Dispersion Energies for Small Organic Molecules: First Row Atoms. *Mol Phys* **2008**, 106 (12–13), 1631–1643. <https://doi.org/10.1080/00268970802258617>.
- (21) Misquitta, A. J.; Stone, A. J. ISA-Pol: Distributed Polarizabilities and Dispersion Models from a Basis-Space Implementation of the Iterated Stockholder Atoms Procedure. *Theor Chem Acc* **2018**, 137 (11), 153. <https://doi.org/10.1007/s00214-018-2371-4>.
- (22) Sato, T.; Nakai, H. Density Functional Method Including Weak Interactions: Dispersion Coefficients Based on the Local Response Approximation. *J Chem Phys* **2009**, 131 (22), 224104. <https://doi.org/10.1063/1.3269802>.
- (23) Sato, T.; Nakai, H. Local Response Dispersion Method. II. Generalized Multicenter Interactions. *J Chem Phys* **2010**, 133 (19), 194101. <https://doi.org/10.1063/1.3503040>.
- (24) Krishtal, A.; Geldof, D.; Vanommeslaeghe, K.; Alsenoy, C. Van; Geerlings, P. Evaluating London Dispersion Interactions in DFT: A Nonlocal Anisotropic Buckingham–Hirshfeld Model. *J Chem Theory Comput* **2012**, 8 (1), 125–134. <https://doi.org/10.1021/ct200718y>.
- (25) Krishtal, A.; Van Alsenoy, C.; Geerlings, P. Evaluating Interaction Energies of Weakly Bonded Systems Using the Buckingham–Hirshfeld Method. *J Chem Phys* **2014**, 140 (18), 184105. <https://doi.org/10.1063/1.4873133>.
- (26) Becke, A. D.; Johnson, E. R. Exchange-Hole Dipole Moment and the Dispersion Interaction. *J Chem Phys* **2005**, 122 (15), 154104. <https://doi.org/10.1063/1.1884601>.
- (27) Johnson, E. R.; Becke, A. D. A Post-Hartree–Fock Model of Intermolecular Interactions. *J Chem Phys* **2005**, 123 (2), 24101. <https://doi.org/10.1063/1.1949201>.
- (28) Johnson, E. R.; Becke, A. D. A Post-Hartree–Fock Model of Intermolecular Interactions: Inclusion of Higher-Order Corrections. *Journal of Chemical Physics* **2006**, 124 (17), 174104. <https://doi.org/10.1063/1.2190220>.
- (29) Becke, A. D.; Johnson, E. R. Exchange-Hole Dipole Moment and the Dispersion Interaction Revisited. *J Chem Phys* **2007**, 127 (15), 154108. <https://doi.org/10.1063/1.2795701>.
- (30) Otero-de-la-Roza, A.; Johnson, E. R. Van Der Waals Interactions in Solids Using the Exchange-Hole Dipole Moment Model. *J Chem Phys* **2012**, 136 (17), 174109. <https://doi.org/10.1063/1.4705760>.

- (31) Hirshfeld, F. L. Bonded-Atom Fragments for Describing Molecular Charge Densities. *Theor Chim Acta* **1977**, *44* (2), 129–138. <https://doi.org/10.1007/BF00549096>.
- (32) Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbó-Dorca, R. Critical Analysis and Extension of the Hirshfeld Atoms in Molecules. *J Chem Phys* **2007**, *126* (14), 144111. <https://doi.org/10.1063/1.2715563>.
- (33) Krishtal, A.; Senet, P.; Yang, M.; Van Alsenoy, C. A Hirshfeld Partitioning of Polarizabilities of Water Clusters. *J Chem Phys* **2006**, *125* (3), 34312. <https://doi.org/10.1063/1.2210937>.
- (34) Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. An Effective Fragment Method for Modeling Solvent Effects in Quantum Mechanical Calculations. *J Chem Phys* **1996**, *105* (5), 1968–1986. <https://doi.org/10.1063/1.472045>.
- (35) Adamovic, I.; Gordon, M. S. Dynamic Polarizability, Dispersion Coefficient C₆ and Dispersion Energy in the Effective Fragment Potential Method. *Mol Phys* **2005**, *103* (2–3), 379–387. <https://doi.org/10.1080/00268970512331317246>.
- (36) Xu, P.; Zahariev, F.; Gordon, M. S. The R–7 Dispersion Interaction in the General Effective Fragment Potential Method. *J Chem Theory Comput* **2014**, *10* (4), 1576–1587. <https://doi.org/10.1021/ct500017n>.
- (37) Alkan, M.; Xu, P.; Gordon, M. S. Many-Body Dispersion in Molecular Clusters. *J Phys Chem A* **2019**, *123* (39), 8406–8416. <https://doi.org/10.1021/acs.jpca.9b05977>.
- (38) Tkatchenko, A.; DiStasio, R. A.; Car, R.; Scheffler, M. Accurate and Efficient Method for Many-Body van Der Waals Interactions. *Phys Rev Lett* **2012**, *108* (23), 236402. <https://doi.org/10.1103/PhysRevLett.108.236402>.
- (39) Jensen, J. H.; Gordon, M. S. An Approximate Formula for the Intermolecular Pauli Repulsion between Closed Shell Molecules. *Mol Phys* **1996**, *89* (5), 1313–1325. <https://doi.org/10.1080/00268979609482543>.
- (40) Li, H.; Gordon, M. S. Gradients of the Exchange-Repulsion Energy in the General Effective Fragment Potential Method. *Theor Chem Acc* **2006**, *115* (5), 385–390. <https://doi.org/10.1007/s00214-006-0080-x>.
- (41) Li, H.; Gordon, M. S.; Jensen, J. H. Charge Transfer Interaction in the Effective Fragment Potential Method. *Journal of Chemical Physics* **2006**, *124* (21), 214108/1-214108/16. <https://doi.org/10.1063/1.2196884>.
- (42) Xu, P.; Gordon, M. S. Charge Transfer Interaction Using Quasiatomic Minimal-Basis Orbitals in the Effective Fragment Potential Method. *Journal of Chemical Physics* **2013**, *139* (19), 194104/1-194104/11. <https://doi.org/10.1063/1.4829509>.
- (43) Stone, A. J.; Alderton, M. Distributed Multipole Analysis Methods and Applications. *Mol Phys* **1985**, *56* (5), 1047–1064. <https://doi.org/10.1080/00268978500102891>.
- (44) Lu, W. C.; Wang, C. Z.; Schmidt, M. W.; Bytautas, L.; Ho, K. M.; Ruedenberg, K. Molecule Intrinsic Minimal Basis Sets. I. Exact Resolution of Ab Initio Optimized Molecular Orbitals in Terms of Deformed Atomic Minimal-Basis Orbitals. *J Chem Phys* **2004**, *120* (6), 2629–2637. <https://doi.org/10.1063/1.1638731>.
- (45) Jensen, J. H. Modeling Intermolecular Exchange Integrals between Nonorthogonal Molecular Orbitals. *Journal of Chemical Physics* **1996**, *104* (19), 7795–7796. <https://doi.org/10.1063/1.471485>.

- (46) Slipchenko†, L. V.; Gordon, M. S. Damping Functions in the Effective Fragment Potential Method. *Mol Phys* **2009**, *107* (8–12), 999–1016. <https://doi.org/10.1080/00268970802712449>.
- (47) Tang, K. T.; Toennies, J. Peter. An Improved Simple Model for the van Der Waals Potential Based on Universal Damping Functions for the Dispersion Coefficients. *Journal of Chemical Physics* **1984**, *80* (8), 3726–3741. <https://doi.org/10.1063/1.447150>.
- (48) Slipchenko, L. V.; Gordon, M. S. Damping Functions in the Effective Fragment Potential Method. *Mol Phys* **2009**, *107* (8–12), 999–1016. <https://doi.org/10.1080/00268970802712449>.
- (49) Smith, Q. A.; Ruedenberg, K.; Gordon, M. S.; Slipchenko, L. V. The Dispersion Interaction between Quantum Mechanics and Effective Fragment Potential Molecules. *Journal of Chemical Physics* **2012**, *136* (24), 244107/1-244107/12. <https://doi.org/10.1063/1.4729535>.
- (50) Gross, E. K. U.; Ullrich, C. A.; Gossmann, U. J. Density Functional Theory of Time-Dependent Systems. *NATO ASI Series, Series B: Physics* **1995**, *337* (Density Functional Theory), 149–171.
- (51) Quinet, O.; Liégeois, V.; Champagne, B. TDHF Evaluation of the Dipole-Quadrupole Polarizability and Its Geometrical Derivatives. *J Chem Theory Comput* **2005**, *1* (3), 444–452. <https://doi.org/10.1021/CT049888Y/ASSET/IMAGES/MEDIUM/CT049888YE00052.GIF>.
- (52) Yamaguchi, Y.; Goddard, J. D.; Osamura, Y.; Schaefer III., H. F. *A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory*; Oxford Univ. Press, 1994.
- (53) Guidez, E. B.; Xu, P.; Gordon, M. S. Derivation and Implementation of the Gradient of the R–7 Dispersion Interaction in the Effective Fragment Potential Method. *J Phys Chem A* **2016**, *120* (4), 639–647. <https://doi.org/10.1021/acs.jpca.5b11042>.
- (54) Schmidt, M. W.; Baldrige, 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. General Atomic and Molecular Electronic Structure System. *J Comput Chem* **1993**, *14* (11), 1347–1363. <https://doi.org/10.1002/jcc.540141112>.
- (55) Gordon, M. S.; Schmidt, M. W. Advances in Electronic Structure Theory: GAMESS a Decade Later. In *Theory Appl. Comput. Chem.: First Forty Years*; Elsevier B.V., 2005; pp 1167–1189.
- (56) Barca, G. M. J.; Bertoni, C.; Carrington, L.; Datta, D.; De Silva, N.; Deustua, J. E.; Fedorov, D. G.; Gour, J. R.; Gunina, A. O.; Guidez, E.; et al. Recent Developments in the General Atomic and Molecular Electronic Structure System. *J Chem Phys* **2020**, *152* (15), 154102. <https://doi.org/10.1063/5.0005188>.
- (57) Burns, L. A.; Faver, J. C.; Zheng, Z.; Marshall, M. S.; Smith, D. G. A.; Vanommeslaeghe, K.; MacKerell, A. D.; Merz, K. M.; Sherrill, C. D. The BioFragment Database (BFDb): An Open-Data Platform for Computational Chemistry Analysis of Noncovalent Interactions. *J Chem Phys* **2017**, *147* (16), 161727. <https://doi.org/10.1063/1.5001028>.
- (58) Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. Levels of Symmetry Adapted Perturbation Theory (SAPT). I. Efficiency and Performance for Interaction Energies. *J Chem Phys* **2014**, *140* (9), 094106. <https://doi.org/10.1063/1.4867135>.

- (59) Weyl, H. *The Classical Groups: Their Invariants and Representations*. Princeton University Press, **1997**; 320.
- (60) Craig, D. P.; Thirunamachandran, T. Appendix 2: Rotational Averaging of Tensors. *Molecular Quantum Electrodynamics: An Introduction to Radiation-Molecule Interactions*; Theoretical chemistry; Academic Press: London ; **1984**; 310.
- (61) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J Chem Phys* **2010**, *132* (15), 154104. <https://doi.org/10.1063/1.3382344>.
- (62) Smith, Q. A.; Gordon, M. S.; Slipchenko, L. V. Effective Fragment Potential Study of the Interaction of DNA Bases. *J Phys Chem A* **2011**, *115* (41), 11269–11276. <https://doi.org/10.1021/jp2047954>.
- (63) Smith, T.; Slipchenko, L. V.; Gordon*, M. S. Modeling π - π Interactions with the Effective Fragment Potential Method: The Benzene Dimer and Substituents. *J Phys Chem A* **2008**, *112* (23), 5286–5294. <https://doi.org/10.1021/jp800107z>.
- (64) Smith, Q. A.; Gordon, M. S.; Slipchenko, L. V. Benzene–Pyridine Interactions Predicted by the Effective Fragment Potential Method. *J Phys Chem A* **2011**, *115* (18), 4598–4609. <https://doi.org/10.1021/jp201039b>.
- (65) Hands, M. D.; Slipchenko, L. V. Intermolecular Interactions in Complex Liquids: Effective Fragment Potential Investigation of Water–Tert-Butanol Mixtures. *J Phys Chem B* **2012**, *116* (9), 2775–2786. <https://doi.org/10.1021/jp2077566>.
- (66) Slipchenko, L. V.; Gordon, M. S. Water–Benzene Interactions: An Effective Fragment Potential and Correlated Quantum Chemistry Study. *J Phys Chem A* **2009**, *113* (10), 2092–2102. <https://doi.org/10.1021/jp808845b>.
- (67) DeFusco, A.; Minezawa, N.; Slipchenko, L. V.; Zahariev, F.; Gordon, M. S. Modeling Solvent Effects on Electronic Excited States. *J Phys Chem Lett* **2011**, *2* (17), 2184–2192. <https://doi.org/10.1021/jz200947j>.
- (68) Rankin, B. M.; Hands, M. D.; Wilcox, D. S.; Fega, K. R.; Slipchenko, L. V.; Ben-Amotz, D. Interactions between Halide Anions and a Molecular Hydrophobic Interface. *Faraday Discuss.* **2013**, *160* (0), 255–270. <https://doi.org/10.1039/C2FD20082A>.
- (69) Adamovic, I.; Li, H.; Lamm, M. H.; Gordon, M. S. Modeling Styrene–Styrene Interactions. *J Phys Chem A* **2006**, *110* (2), 519–525. <https://doi.org/10.1021/jp058140o>.