COMPARATIVE STUDY OF POPULATION ANALYSIS WITH QUASI-ATOMIC ORBITALS

COMPARATIVE STUDY OF POPULATION ANALYSIS WITH QUASI-ATOMIC ORBITALS

By

AYSE KUMRU DIKMENLI, B.Sc., M.Sc.

A Thesis Submitted to the School of Graduate Studies In Partial Fulfillment of the Requirements for the Degree Master of Science

MASTER OF SCIENCE (2019)

McMaster University Chemistry Department

Hamilton, Ontario

TITLE Comparative Study of Population Analysis Methods With Quasi Atomic-Orbitals

AUTHOR Ayse Kumru Dikmenli, M.Sc., B.Sc. (Bogazici University)

SUPERVISOR Prof. Paul W. Ayers (McMaster University)

PAGES xi, 147

Abstract

Atom is not an observable of the molecular wavefunction, in quantum chemistry there are myriad ways of defining an atom in a molecule. Partitioning a molecule's electrons between its atomic constituents (population analysis) remains a challenge.

A popular approach is based on Mulliken's overlap-based population analysis, which exploits the fact that molecular orbitals can be expressed as linear combinations of user-defined functions: atomic orbitals. In turn, this creates a dependency on the selection of the predetermined atomic orbitals that are used to expand the molecular orbitals. Chemically intuitive atomic orbitals like Minimal Atomic Orbitals (minAO) produces chemically intuitive atomic charges but a non-accurate wave function. Accurate wave functions can be obtained from large atomic basis sets like def2-QZVPd at the cost of chemically unintuitive atomic charges. With this problem in sight, Quasi-Atomic Orbitals (QAO) are constructed from the accurate wave function to resemble the minAO and maximally span the molecular orbital space. The key idea is Mulliken population analysis can be carried out for wave functions with the chemical intuitive power of minAO, without sacrificing the wave function's accuracy by using QAO. To ensure that overlaps of QAO are divided between different atoms without bias. Zero-Bond Dipole (ZBD) orthogonalization is proposed as a novel way to orthogonalize QAO.

Common population analysis from literature: Charge Model 5 (CM5), QH, Hu-Lu-Yang (ESP), Mulliken, NPA atomic charges will be compared to QAO (Mulliken with QAO) and ZBD-QAO (Mulliken after ZBD orthogonalization of QAO) and tested for mathematical accuracy and expected chemical trends.

Acknowledgments

I would like to express my deepest regards for Prof. Ayers. Anyone working with him is blessed with an environment encouraging self-improvement. I feel grateful to have supportive lab mates. Thank you, Jen, for helping me build and structure the Vetee database code and naming it with the initials of “database” in Turkish (for of course “veritabani” is too long for a name). It has not been easy to code, and I appreciate your contribution and listening me at tea breaks or code reviews. I would like to thank David for writing Quasi-Atomic Orbital code, explaining the theory and helping me use his code. Thank you, Xiao and Will, for helping me with the math and the coding hell of Zero-Bond Dipole code. Also, Israel has been a great non-wavering support for me in this journey as well as my parents. I would like to express my deepest gratitude for my friends who where here to morally support me Ayşegül and Melih, and my friends who are in Turkey namely Idil, Neşe and Deniz.

Table of Contents

Table of Figures

Table of Tables

Acronyms

CM5 : Charge Model 5 Population Analysis

QH : Original Hirshfeld Population Analysis

MBS : Mulliken with Minimal Basis Projection

NPA : Natural Population Analysis

AO : Atomic Orbitals

minAO : Minimal Atomic Orbitals

MO : Molecular Orbitals

oMO : Occupied Moleculart Oribtals

vMO : Virtual (Unoccupied) Molecular Orbitals

QAO : Quasi-Atomic Orbitals

ZBD : Zero- Bond Dipole (orthogonalization)

ZBD-QAO : Zero-Bond Dipole orthogonalized Quasi-Atomic Orbitals

Q1 : Quasi-Atomic Orbital QUAMBO

Q2 : Quasi-Atomic Orbital QUAO

Q3 : Quasi-Atomic Orbital IAO

Z-Q1 : Zero -Bond Dipole orthogonalized QUAMBO

Z-Q2 : Zero -Bond Dipole orthogonalized QUAO

Z-Q3 : Zero -Bond Dipole orthogonalized IAO

Chapter 1

1 Introduction

1.1 Background

The periodic table of the elements lies at the heart of chemistry, based on the precept that molecules are built from atoms in this table. For example, in the Lewis model, lines between two atoms symbolize the sharing of two electrons, and the number of lines corresponds to bond order.[1] As straightforward as it is to think about atoms as the building blocks of molecules in chemistry, in quantum mechanics there is no strict definition for the atom within a system containing multiple atoms.[2,3] One of the many postulates of quantum mechanics is that a wave function contains all the information about a molecular system.[4] It is accepted that for every physical observable there is a Hermitian operator, and that the possible values of the measured property are the eigenvalues of that operator. For example, the energy, *E*, is the eigenvalue of the Hamiltonian operator, $\hat{H}$, and the wavefunction, Ψ, is its eigenvector. This relationship is encapsulated in the time-independent Schrödinger equation,

$$\hat{H}\Psi\left(\mathbf{R}_1,\mathbf{R}_2,\ldots,\mathbf{R}_P,\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N\right)=E\Psi\left(\mathbf{R}_1,\mathbf{R}_2,\ldots,\mathbf{R}_P,\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N\right) \quad (1)$$

where $\mathbf{R}_1,\mathbf{R}_2,\ldots,\mathbf{R}_P$ denote the locations of the atomic nuclei and $\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N$ denote the positions of the electrons. Solving the Schrödinger equation is intractable for systems with more

than two particles, which motivates approximations that simplify the solution, e.g., the Born-Oppenheimer approximation,[5] which allows the molecular wavefunction to be approximated as the product of nuclear and electronic pieces. In this thesis we will consider the nuclei to be fixed (thereby ignoring nuclear quantum effects) and treat only the electronic Schrödinger equation. Even determining the electronic wave function is challenging and so, in this thesis, only two methods of approximation will be used; Hartree-Fock[4,6] (HF) and density-functional theory[7–9] (DFT). In both models, the electronic wavefunction is approximated as an anti-symmetrized product of molecular orbitals (i.e., a single Slater determinant). We will use atom-centered Gaussian basis sets, so each molecular orbital is expressed as a linear combination of atom-centered Gaussian basis functions. The approximations we are making are common among chemists, and we made them for computational expedience, as our primary focus is not obtaining accurate electronic wavefunctions for molecules, but assessing different ways of defining atoms in a molecule.

Most methods for partitioning molecules into atoms fall into two main families. In the first family, pioneered by Hirshfeld and Bader,[10–12] the molecule's electron density is divided into atomic contributions. The atomic densities are then used to define atomic properties, including the number of electrons (by integrating the atomic density), multipole moments, etc.. In this thesis, we will consider only Hirshfeld's original definition (referred to as QH),[13] and one revision thereof that is built into the Gaussian program[14] (CM5 short for Charge-Model 5).[15] Hirshfeld partitioning is based on the idea that the utility of the periodic table is maximized if the atoms in a molecule resemble the isolated items atoms in the periodic table to the maximum possible extent. If one defines the density of the atoms-in-a-molecule so that their divergence from the isolated atoms in the periodic table is minimized, then the Hirshfeld partitioning results.[3,16–20] (Many possible

definitions of divergence are possible, but it's traditional to use information theory because the electron density is a probability distribution function.[3,16] The Hirshfeld partitioning is not obviously appropriate for charged molecules or for ionic bonding (where forcing atoms in an ionic molecule to resemble neutral atoms is chemically misinformed), which has led to the development of many other Hirshfeld-based schemes, which we will not consider in this thesis.[21–23] We will, however, consider the Charge Model 5 (CM5) revision of Hirshfeld. CM5 charges are an empirical revision of Hirshfeld charges designed to give good dipole moments. The basic idea is to revise the Hirshfeld charge of an atom, *A*, based on nearby atoms, *B*, where the definition of "nearby" is determined by the empirical exponentially-decaying bond-length bond-order relationship proposed by Pauling. Specifically, the CM5 atomic charges are defined as:

$$q_{\mathrm{A}}^{(\mathrm{CM5})} = q_{\mathrm{A}}^{\mathrm{Hirshfeld}} + \sum_{\substack{B=1 \\ B \neq A}}^{N_{\mathrm{atoms}}} T_{Z_{\mathrm{A}} Z_{\mathrm{B}}} e^{-\alpha\left(|\mathbf{R}_{\mathrm{A}} - \mathbf{R}_{\mathrm{B}}| - \left(r_{\mathrm{A}}^{\mathrm{cov}} + r_{\mathrm{B}}^{\mathrm{cov}}\right)\right)} \tag{2}$$

where $T_{Z_{\mathrm{A}} Z_{\mathrm{B}}}$ is an antisymmetric matrix (otherwise charge would not be conserved) based on the atom types (e.g., the atomic numbers of atoms *A* and *B*). For element-pairs that occur frequently in the training set (CH, NH, OH, CN, CO, NO), there is a specific value of $T_{Z_{\mathrm{A}} Z_{\mathrm{B}}}$; for other elements for which there is adequate data, one chooses to define only atomic parameters $T_{Z_{\mathrm{A}} Z_{\mathrm{B}}} = t_{Z_A} - t_{Z_{\mathrm{B}}}$; otherwise one takes the value of t_{Z_A} from the top parameterized element in a column of the periodic table and multiplies it by a diminishing factor, *C*, for each row of the periodic table one descends before one reaches the element of interest. Clearly this approach is limited by its empirical fitting and, moreover, by the assumption that the Hirshfeld atomic charges can be improved using only local changes (which is not true for zwitterions or ion pairs). It is also true

that, unlike most population analysis methods, CM5 defines only atomic charges, and not atomic properties more generally.

The approach to partitioning molecules into atomic contributions is based on orbitals or, more generally, reduced density matrices. This strategy was pioneered by Mulliken in the formative years of quantum chemistry.[24] The molecular orbitals that enter into the single Slater determinant approximation for the electronic wavefunction are typically expanded in terms of atom-centered basis functions. Intuitively, these basis functions represent the atomic orbitals that constructively and deconstructively interfere to form bonding and antibonding molecular orbitals, respectively. Thus, molecular orbitals can be mathematically expressed as a linear combination of atomic orbitals, [25]

$$\left|\mathrm{MO}_i\right\rangle = \sum_i C_{il} \left|\mathrm{AO}_l\right\rangle \tag{3}$$

This expansion is exact if all possible atomic orbitals (including the continuum) were used, but always approximate in practice. It can be quite accurate, however, when large atomic basis sets are used.

Molecular orbitals can be further classified as occupied molecular orbitals (oMO), including both core (cMO) and valence-occupied (voMO) molecular orbitals. Unoccupied, or virtual, molecular orbitals (vMO) can be similarly decomposed into two groups, the valence virtual (vvMO) and external virtual (exMO) molecular orbitals. The relativistic contraction of cMOs is, of course, important for the chemical and physical properties of molecuels containing heavy elements.[26] To overcome this issue, pseudopotential basis sets have been developed to include relativistic effects, and we will use pseudopotentials for elements beyond Kr in our calculations.[27] Our chemical intuition is based on the idea that cMO and exMO have little qualitative chemical

effect, though these MOs can be quantitatively important. Conversely, the valence occupied and virtual MOs (the voMO and vvMO) are very important for bonding in chemistry.

While the number of oMO is determined by the number of electrons in a molecule, the number of vMO changes with the choice of atomic basis set. Increasing the number of basis functions increases the mathematical accuracy of the molecule's wavefunction and properties, and also increase the number of vMOs. However, the high-energy vMOs are difficult to interpret and, indeed, are usually considered merely mathematical artifacts required by the approach by the complete basis set limit. Indeed, some atomic basis functions are added without any consideration for chemical interpretation. For example, diffuse functions are added to treat the long-range portions of molecular orbitals, which is especially important for electronic excited states and anions, where the probability of observing an electron far from the molecule is relatively large. Polarization functions are added to model the way atomic orbitals deform in the presence of other atoms and/or external electric fields. The "deformed" polarized atomic orbitals can lead to chemically unintuitive atomic charges.[28]

1.2 Mullliken and Löwdin Population Analysis

The main focus of this thesis is Mulliken's (overlap) population analysis, which is an orbital based approach to directly distribute molecular electronic charge to an atom pair.[29] Assuming that molecular orbitals are normalized, the total number of electrons in a system can be written as the sum of the overlap of each oMO with itself, multiplied by the occupation numbers.

$$N = \sum_i n_i \left\langle \mathrm{MO}_i \middle| \mathrm{MO}_i \right\rangle \tag{4}$$

Where N is the total number of electrons in the molecule, n_i is the number of electrons (occupation number) of the molecular orbital $\left|\mathrm{MO}_i\right\rangle$. Notice that this equation is valid even for exact calculations, where the occupation numbers are not just zero and one (as they are for a single Slater determinant). Expanding the MOs in terms of AOs using equation (3),

$$\begin{aligned} N &= \sum_i n_i \sum_{lk} C_{il} C_{ik}^{\dagger} \left\langle \mathrm{AO}_k \middle| \mathrm{AO}_l \right\rangle \\ &= \sum_{lk} \left\langle \mathrm{AO}_k \middle| \mathrm{AO}_l \right\rangle \sum_i C_{il} n_i C_{ik}^{\dagger} \\ &= \sum_{lk} N_{lk} \end{aligned} \tag{5}$$

Therefore, the total number of electrons in the molecule can decomposed into its contribution from different atomic basis functions. The AO decomposition can be re-expressed as an atom-based composition as long as the number of electrons associated with each atom N_A sums up to the total number of electrons in the molecule,

$$N = \sum_A N_A \tag{6}$$

Electrons of atom *A* can be further decomposed into mono and di-atomic terms

$$N_A = \sum_{k,l \in A} N_{lk}^{AA} + \sum_{\substack{k \in A \\ l \in B \\ A \neq B}} N_{lk}^{AB} = \sum_{\substack{k \in A \\ l \in B}} N_{lk}^{AB} \tag{7}$$

where atomic basis functions *k* and *l* either belong to the same atom A or different atoms A and B. Notice that the Mulliken partitioning makes the *ad hoc* assumption that the diatomic terms in Eq. (7), $N^{A \neq B}$, are divided equally between the contributing atoms. Different orbital-based partitioning methods differ based on (a) the choice of atomic basis functions and (b) how they deal with the diatomic terms in Eq. (7). These two decisions are not unrelated. In typical Mulliken analysis, the nonorthogonal atom-centered basis functions are used as atomic orbitals in Eq. (5);

this is done even though these atomic basis functions may not resemble the atomic orbitals and, indeed, even though the atomic basis functions may have significant amplitude on atoms that are far from the atom on which they are centered. This delocalization is especially acute for diffuse basis functions. This motivates the strategy of using Löwdin, or symmetric, orthogonalization to define a new set of atomic basis functions before performing the partitioning.

We will denote transformed orbitals with a tilde, so the old and new (orthogonalized) orbitals are denoted as $\left|\mathrm{AO}_k\right\rangle$ and $\left|\mathrm{AO}_k\right\rangle$, respectively. The new orbitals will be orthogonalized,

$$\left\langle \mathrm{AO}_l^A \middle| \mathrm{AO}_k^B \right\rangle = \delta_{AB}\delta_{lk} \tag{8}$$

subject to the constraint that the new basis set is as close to the old basis set as possible (in the L2-norm). This Löwdin orthogonalized orbitals are then computed as

$$\left|\mathrm{AO}_l\right\rangle = \sum_k S_{lk}^{-1/2} \left|\mathrm{AO}_k\right\rangle \tag{9}$$

Projection onto the subspace of orbitals associated with atom *A* is then defined in the usual way,

$$\hat{p}_A = \sum_{l\in A} \left|\mathrm{AO}_l\right\rangle\left\langle \mathrm{AO}_l\right| = \sum_{l\in A}\sum_{k,m} \left|\mathrm{AO}_m\right\rangle S_{ml}^{-1/2} S_{lk}^{-1/2} \left\langle \mathrm{AO}_k\right| \tag{10}$$

The advantage of using Löwdin-orthogonalized AOs from Eq. (9): there are no longer any contributions at all from the di-atomic terms, so

$$N_A = \sum_{l\in A} N_{ll}^{AA} \tag{11}$$

Therefore, the problem of dividing di-atomic contributions between atoms has been removed. However, notice from Eq. (11) that the AOs assigned to atom *A* have contributions from all of the

other atoms in the molecule. The Löwdin-orthogonalized AOs are, therefore, not highly localized on the atoms either. In the next two sections we will discuss other ways of defining effective atomic orbitals and dividing the di-atomic contributions between atoms.

1.3 Mulliken with Quasi-Atomic Orbitals

The aforementioned problems with Mulliken and Löwdin population analysis are not present if only the core and valence atomic orbitals—that is, functions from a minimal atomic basis set—are used. This motivates various techniques that are construct an appropriate atomic-orbital basis set. Two of the more popular options are natural population analysis (NPA) and the minimal basis set Mulliken analysis (MBS).

In NPA, one expands the occupied molecular orbitals (equivalently, the one-electron density matrix) using the atomic orbitals of the isolated neutral atoms,

$$w_{kl}^{AB} = \sum_{k,l,A,B} \int \left\langle \mathrm{isoAO}_k^A \left| \gamma\left(\mathbf{r},\mathbf{r}'\right) \right| \mathrm{isoAO}_l^B \right\rangle \tag{12}$$

One then uses these weights to perform an occupation-weighted version of the Löwdin orthogonalization, i.e., one chooses the orthogonal AOs that it is as close as possible to the isolated AO basis in a weighted sense,

$$\min_{\text{transformations}} \sum_{Ak} w_{kk}^{AA} \left\langle \left| \mathrm{isoAO}_k^A - \mathrm{AO}_k^A \right|^2 \right\rangle \tag{13}$$

This is only the essence of the NPA algorithm, which is a complicated method requiring orthogonalization wherein various orbital sets (core, Rydberg, valence, etc.) are orthogonalized separately. However, this reveals that NPA is, in essence, a cleverly-weighted Löwdin partitioning. NPA is therefore more robust than the standard (unweighted) Löwdin partitioning, but can have

problems when the isolated AO basis sets are unbalanced (e.g., if different basis sets are used for different atoms) or, more generally, when an isolated AO that should not have a large weight acquires an anomalously high weight via Eq. (12).

In MBS, the occupied molecular orbitals (equivalently, the one-electron reduced density matrix) are expanded in a minimal basis set, which is traditionally chosen to be the STO-3G* basis. The minimal basis set is then Lowdin orthogonalized. The resulting molecular orbitals are no longer normalized because the accurate molecular orbitals cannot be fully represented with a minimal basis set, so a normalization correction is required. The resulting scheme is not easily represented purely as an effective atomic orbital basis, but the key idea is that the wavefunction is expanded in terms of the Löwdin-orthogonalized minimal atomic orbital basis set. This clearly fails when the wavefunction has features that cannot be adequately represented with the minimal atomic basis set selected, as can be the case for highly-correlated systems, and states (e.g., excited states and anions) with diffuse electrons.

MBS can be improved by selecting an minimal atomic orbital basis set that is adapted to the system of interest in such a way that accuracy is not compromised. This is the strategy used in quasi-atomic orbital (QAO) methods. Unlike the Mulliken, Löwdin, and MBS approaches, but similar to NPA, the AOs are determined by the molecular orbitals, and are thereby adapted to the system. Unlike NPA and perhaps similar to Hirshfeld analysis, an auxiliary minimal basis set of accurate AOs are used so that chemists' intuition about the AOs that are used to construct MOs are retained. (Unfortunately, also similar to Hirshfeld analysis, only neutral AOs are used, so these methods are less obviously appropriate for charged molecules and ionic bonding.)

We consider three different types of QAO: QUAMBO, QUAO and IAO. All three of these methods start with a reference minimal basis set of accurate atomic orbitals called minAOs. The

goal is to find a set of quasi-atomic orbitals that (a) is as close to the minAOs as possible and (b) can exactly reconstruct the occupied molecular orbitals. Mathematically, this means that the quasi-atomic orbitals are a linear combination of the molecular orbitals and a subset of the virtual molecular orbitals called the valence virtual molecular orbitals,

$$\left|\mathrm{QAO}_k\right\rangle = \sum_{i=1}^{N_{\mathrm{occupied}}} x_{ki}\left|\mathrm{oMO}_i\right\rangle + \sum_{a=N_{\mathrm{occupied}}+1}^{N_{\mathrm{minAO}}} x_{ka}\left|\mathrm{vvMO}_a\right\rangle \tag{14}$$

We have chosen the number of vvMOs plus the number of oMOs and the number of QAOs to both equal the number of minAOs,[30] but there are times when including fewer, or more, vvMOs is helpful.

Finding the QAOs that resemble the minAOs most strongly amounts to choosing the best possible vvMOs, which means that one wishes to find the vvMOs that have the highest overlap with the minAOs. (These are the vMOs that have the greatest contributions from the minAOs. Consequently, they define the QAOs that have the greatest overlap with the minAOs.) In QUAMBO, the optimal set of vvMOs are those that maximize, [3132]

$$\max_{\mathrm{vvMOs}} \sum_{a=N_{\mathrm{occupied}}+1}^{N_{\mathrm{minAO}}} \sum_{k=1}^{N_{\mathrm{minAO}}} \left\langle \mathrm{vvMO}_a \middle| \mathrm{min\ AO}_k \right\rangle \left\langle \mathrm{min\ AO}_k \middle| \mathrm{vvMO}_a \right\rangle \tag{15}$$

The vvMOs are therefore given by the left singular vectors corresponding to the $N_{\mathrm{min\,AOs}} - N_{\mathrm{occupied}}$ largest singular values of the overlap matrix between the virtual molecular orbitals and the minAOs, $\left\langle \mathrm{vMO}_a \middle| \mathrm{minAO}_k \right\rangle$.[33,34] When singular vectors with large singular values are thus discarded, it is reasonable to increase the space of vvMOs to include them. Similarly, in the (less likely) case where singular vectors with tiny singular values are thus included, it is reasonable to decrease the

space of vvMOs. In cases where the occupied MOs cannot be well-represented with the minAO basis set, it is reasonable to increase the minAO basis.

Eq. (15) resembles a projection of the vvMOs onto the space of minAOs, but because the minAOs are not orthonormal it is not a true projection operator. In quasi-atomic orbitals (QUAO), [35] one tries to find QAOs that (a) maximally resemble the Löwdin-orthogonalized minAOs and (b) span the space of occupied molecular orbitals. Equation (15) is obtained again but because orthogonalized minAOs are used instead of the original minAOs,

$$\begin{aligned} \left|\min \mathrm{AO}_k\right\rangle &= \sum_{l=1}^{N_{\mathrm{minAO}}} \left(S_{kl}^{\mathrm{minAO}}\right)^{-1/2} \left|\mathrm{minAO}_l\right\rangle \\ S_{kl}^{\mathrm{minAO}} &= \left\langle \mathrm{minAO}_k \middle| \mathrm{minAO}_l \right\rangle \\ &\max_{\mathrm{vvMOs}\ a=N_{\mathrm{occupied}}+1} \sum_{k=1}^{N_{\mathrm{minAO}}} \sum_{k=1}^{N_{\mathrm{minAO}}} \left\langle \mathrm{vvMO}_a \middle| \min \mathrm{AO}_k \right\rangle \left\langle \min \mathrm{AO}_k \middle| \mathrm{vvMO}_a \right\rangle \end{aligned} \tag{16}$$

As before, the QUAO vvMOs are obtained by performing a singular value decomposition on the overlap matrix between the virtual MOs and the orthogonalized minAOs, $\left\langle \mathrm{vMO}_a \middle| \mathrm{minAO}_k \right\rangle$.

Once the vvMOs have been determined using either the QUAMBO or QUAO procedure, the QAOs are obtained by projecting the minAOs onto the basis of occupied molecular orbitals and valence virtual molecular orbitals,

$$\left|\mathrm{QAO}_k\right\rangle = \sum_{i=1}^{N_{\mathrm{occ}}} \left|\mathrm{oMO}_i\right\rangle \left\langle \mathrm{oMO}_i \middle| \mathrm{minAO}_k \right\rangle + \sum_{a=1}^{N_{\mathrm{vvMOs}}} \left|\mathrm{vvMO}_a\right\rangle \left\langle \mathrm{vvMO}_a \middle| \mathrm{minAO}_k \right\rangle \tag{17}$$

Intrinsic atomic orbitals (IAOs) were proposed by Knizia based on similar intuition, but without specifically optimizing an objective function as in the case of QUAMBO and QUAO.[36] While QUAMBO and QUAO focus on the vMOs, in IAO the focus is on the oMOs. Like MBS, in IAO a key step is to project the occupied molecular orbitals onto the minAO basis, though in

IAO the accurate atomic orbitals are used for the minAO basis. Knizia calls the oMOs projected onto the minAOs *depolarized*,

$$\begin{aligned} \left|\mathrm{doMOs}_k\right\rangle &= \sum_{l,m=1}^{N_{\mathrm{minAO}}} \left|\mathrm{minAO}_l\right\rangle \left(S_{lm}^{\mathrm{minAO}}\right)^{-1} \left\langle \mathrm{min\,AO}_m \middle| \mathrm{oMO}_k \right\rangle \\ S_{lm}^{\mathrm{minAO}} &= \left\langle \mathrm{minAO}_l \middle| \mathrm{minAO}_m \right\rangle \end{aligned} \tag{18}$$

The doMOs are the portion of the oMOs that can be expressed in the minimal basis set, and we could analyze them directly (which would be similar to MBS) or project an accurate wavefunction onto this space and then analyze it (which inspires the IAO approach). The projection onto doMOs can be written as

$$\begin{aligned} \hat{P}_{\mathrm{doMO}} &= \sum_{l,m=1}^{N_{\mathrm{minAO}}} \left|\mathrm{doMO}_l\right\rangle \left(S_{lm}^{\mathrm{doMO}}\right)^{-1} \left\langle \mathrm{doMO}_m \right| \\ S_{lm}^{\mathrm{doMO}} &= \left\langle \mathrm{doMO}_l \middle| \mathrm{doMO}_m \right\rangle \end{aligned} \tag{19}$$

However, unlike MBS, the IAO method does not ignore the portion of the oMOs that cannot be expressed using the minAOs and, indeed, the IAOs span the space of oMOs. We mentioned earlier that the IAOs are occupied-orbital-centric, as opposed to the virtual-orbital-centric approach in QUAMBO and QUAO. One way this is achieved is by using the resolution of the identity instead. I.e., one uses identities like

$$\sum_{a=1}^{N_{\mathrm{vMOs}}} \left|\mathrm{vMO}_a\right\rangle \left\langle \mathrm{vMO}_a \right| \approx 1 - \sum_{i=1}^{N_{\mathrm{oMOs}}} \left|\mathrm{oMO}_i\right\rangle \left\langle \mathrm{oMO}_i \right| \tag{20}$$

which are exact only in the basis-set limit. With this background, IAOs are given by the following formula,

$$\left|\mathrm{IAO}_k\right\rangle = \left[\left(\sum_{i=1}^{N_{\mathrm{oMOs}}} \left|\mathrm{oMO}_i\right\rangle \left\langle \mathrm{oMO}_i \right| \right) \hat{P}_{\mathrm{doMO}} + \left(1 - \sum_{i=1}^{N_{\mathrm{oMOs}}} \left|\mathrm{oMO}_i\right\rangle \left\langle \mathrm{oMO}_i \right| \right) \left(1 - \hat{P}_{\mathrm{doMO}} \right) \right] \left|\mathrm{minAO}_k\right\rangle \tag{21}$$

The interpretation of this formula is that IAOs are obtained by

(a) projecting the minAOs onto the depolarized occupied MOs, and then the expansion of the minAOs in the doMO basis is projected onto the original occupied MOs.

(b) projecting the minAOs onto the depolarized virtual MOs (using the resolution of the identity), and then projecting the minAOs in the depolarized virtual basis onto the original virtual MOs (again using the resolution of the identity).

1.4 Zero-Bond Dipole Orthogonalization (ZBD)

After a set of AOs has been selected, population analysis can be performed either using the Löwdin approach or the Mulliken approach. If the Mulliken approach is chosen, then it is necessary to contemplate how the di-atomic contributions should be partitioned between their composing atoms. Most generally, one considers a weighted division of the di-atomic contributions,

$$\begin{aligned} N &= \sum_{lk} N_{lk} \\ &= \sum_{A} \sum_{lk} w_{lk}^{A} N_{lk} \\ &= \sum_{A} N_{A} \end{aligned} \tag{22}$$

weight w_{jk}^{A} of atomic basis functions j and k belonging to atom A is normalized for every given atom in the molecule so that number of electrons N_A belonging to atom A can be written as

$$N_A = \sum_{lk} w_{lk}^{A} N_{lk} \tag{23}$$

In traditional Mulliken population analysis N_{jk} is split equally between the atoms that atomic basis j and k belong to. I.e.,

$$w_{lk}^{A}\begin{cases} l \in A \text{ and } k \in A = 1 \\ l \in A \text{ and } k \notin A = 0.5 \\ l \notin A \text{ and } k \in A = 0.5 \\ l \notin A \text{ and } k \notin A = 0 \end{cases} \tag{24}$$

Which means that if j and k both are associated with atom A the weight is equal to one and all the considered electrons belong to one atom; this corresponds to the atomic term, N_{jk}^{AA}, in equation (7). If either j or k basis functions belong to another atom $B \neq A$ then the weight is half, and atom A only gets half of the electrons. The last case is where neither AO is associated with atom A, in which case this no contribution is made to this atom's population. These simple half-and-half partitioning of the di-atomic populations ignores basic chemical concepts like the way electronegativity leads to polar covalent bonds and the fact that distant atoms do not bond (even if their diffuse functions overlap significantly).

There are many possible refinements to the conventional Mulliken partitioning of di-atomic contributions, but we focus on the zero-bond dipole (ZBD) method proposed by Laikov.[37] The general framework is that one to transform the QAOs (whether QUAMBO, or QUAO, or IAO) into an orthogonal set of quasi-atomic orbitals, oQAOs, that are in some sense optimal. For example, most common orthogonalization methods can be expressed as the minimization of a quadratic functional,

$$f = \sum_{k,l} w_{kl} \left\langle \mathrm{QAO}_k - \mathrm{oQAO}_k \left| \hat{F} \right| \mathrm{QAO}_l - \mathrm{oQAO}_l \right\rangle \tag{25}$$

For example, the Löwdin corresponds to $w_{kl} = \delta_{kl}$ and $\hat{F} = 1$.[38] The zero-bond-dipole orthogonalization is not expressed exactly in this form (though it is inspired by such an objective function). The idea is that the problem of dividing atomic charges in Mulliken analysis is caused

by the fact the product of two QAOs, $\mathrm{QAO}_k^A(\mathbf{r})\cdot\mathrm{QAO}_l^B(\mathbf{r})$ is polarized along the bond, and therefore an equal division of this contribution between atoms *A* and *B* is ill-advised. This suggests orthogonalizing the orbitals in such a way that the orthogonal orbitals have zero dipole moment along the bond,

$$\mathbf{R}_{AB}\cdot\left\langle \mathrm{oQAO}_k^A\left|\mathbf{r}\right|\mathrm{oQAO}_l^B\right\rangle = 0 \tag{26}$$

where $\mathbf{R}_{AB}=\left(\mathbf{R}_A-\mathbf{R}_B\right)\Big/\left[\left(\mathbf{R}_A-\mathbf{R}_B\right)\cdot\left(\mathbf{R}_A-\mathbf{R}_B\right)\right]$ is the (normalized) internuclear distance vector. The new orthogonal QAOs will be called ZBD-QAO and will satisfy, by construction,

$$\left\langle \text{ZBD-QAO}_k^A\middle|\text{ZBD-QAO}_l^B\right\rangle = \delta_{kl}\delta_{AB} \tag{27}$$

The QAOs we have defined are already orthogonalized within an atom, but not between atoms. We start by performing a Löwdin orthogonalization of the QAOs. ZBD-QAOs can then be obtained as a unitary transformation of the Lowdin-orthogonalized QAOs,

$$\left|\text{ZBD-QAO}_k^A\right\rangle = \sum_{l,B} T_{kl}^{AB}\left|\mathrm{oQAO}_l^B\right\rangle \tag{28}$$

$$T_{ab}^{AB} = \exp\left(\mathbf{L}-\mathbf{L}^{\dagger}\right) \tag{29}$$

The lower triangular matrix **L** is chosen so that the dipole moment overlap, minimized over all diatomic terms, becomes as close to zero as possible

$$\mathbf{f}(\mathbf{L}) \equiv \sum_{\substack{A'k'B'l' \\ A\neq B}} \mathbf{R}_{AB}\cdot\left[\exp\left(\mathbf{L}-\mathbf{L}^{\dagger}\right)\right]_{kk'}^{AA'}\left\langle \mathrm{QAO}_{k'}^{A'}\left|\mathbf{r}\right|\mathrm{QAO}_{l'}^{B'}\right\rangle\left[\exp\left(\mathbf{L}-\mathbf{L}^{\dagger}\right)\right]_{l'l}^{B'B} = 0 \tag{30}$$

This function is minimized using Scipy with the threshold for the function minimization set to 10^{-7}. Equation (30) can be minimized directly but the efficiency can be improved by providing the Jacobian,

$$\left.\frac{\partial f_{Ak;Bl}}{\partial \ell_{Cm;Dn}}\right|_{\mathbf{L}=\mathbf{0}} \tag{31}$$

The transformation matrix is antisymmetric, for each iteration during the optimization only lower triangular part is updated. The new transformation is constructed as the product of lower triangular matrix and transpose of itself.

$$\left[\mathbf{L}-\mathbf{L}^{\dagger}\right] = \ell_{\mu\nu}\left(\delta_{Cm;Dn}-\delta_{Cm;Dn}\right) = \ell_{\mu\nu}\left(\delta_{\mu\nu}-\delta_{\nu\mu}\right) \tag{32}$$

to ensure lower triangular shape $\mu < \nu$.

$$\left.\frac{\partial f_{Ak;Bl}}{\partial \ell_{Cm;Dn}}\right|_{\mathbf{L}=\mathbf{0}} \equiv \mathbf{R}_{AB}\cdot\begin{pmatrix} +\delta_{Cm;Ak}\left\langle \chi_{Dn}^{(i)}\left|\mathbf{r}\right|\chi_{Bl}^{(i)}\right\rangle - \delta_{Ak;Dn}\left\langle \chi_{Cm}^{(i)}\left|\mathbf{r}\right|\chi_{Bl}^{(i)}\right\rangle \\ +\delta_{Cm;Bl}\left\langle \chi_{Ak}^{(i)}\left|\mathbf{r}\right|\chi_{Dn}^{(i)}\right\rangle - \delta_{Bl;Dn}\left\langle \chi_{Ak}^{(i)}\left|\mathbf{r}\right|\chi_{Cm}^{(i)}\right\rangle \end{pmatrix} \tag{33}$$

Jacobian in equation (33) is minimized alongside the objective function until converge criteria (10^{-7}) is reached. The target minAO orbitals are maximally resembled by defining an overlap between minAO and converged orthogonal orbitals.

$$S_{kl}^{A} = \left\langle \mathrm{minAO}_{k}^{A}\middle|\mathrm{oQAO}_{l}^{A}\right\rangle \tag{34}$$

By using polar decomposition[1] converged orbitals are transformed into ZBD-QAO

[1] https://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.linalg.polar.html

$$\left|\text{ZBD-QAO}_k^A\right\rangle = \sum_{Ak;A'k'} U_{kk'}^{AA'} \left|\text{oQAO}_{k'}^A\right\rangle \tag{35}$$

where the transformation matrix within a given atom is $U_{kk'}^{AA'}$

1.5 Conclusion

Three different QAO populations can be carried out by using QUAMBO, QUAO and IAO, these will be referred to as to Q1, Q2, and Q3 for the comparison of charges in chapter four for better table fitting. The reason why we are suggesting these QAO and ZBD-QAO instead of the canonical ones is:

1. Derived from the MO so expected to be less dependent to basis set
2. Easy to calculate
3. Qualitatively intuitive like minAO
4. Mulliken populations should improve with QAO
5. ZBD orthogonalization for chemical trends

Chapter 2

2 Aim of the Study

2.1 Problems with Mulliken

Mulliken population ensures that partitioning of the electrons between two basis functions belonging to two different atoms is systematically solved. Even though Mulliken is mathematically unambiguous and elegant, dividing electrons equally between two different atoms has no base on chemical intuition. Using a minimal basis set gives good results with Mulliken, but as the basis set gets bigger atomic charges become unrealistic. This problem can be remedied by Löwdin by orthogonalization of orbitals to atoms. The downside of this technique is undesirable mixing of orbitals. During the orthogonalization procedure (the projection onto atomic space of atom *A*) occupied and virtual orbitals are mixed and this is expected in a molecular system. However, the

core electrons rarely contribute to bonding and Löwdin orthogonalization mixes core and valence spaces. Having higher core-valence coefficients in the overlap matrix can lead to chemically unintuitive atomic charges.

Mulliken's population may be systematic, but it fails to follow elementary chemical rules such as electronegativity, since the bond electrons are divided to atoms equally in equation (10). Even though the electrons are divided equally we still get chemically relevant results such as the chlorine of HCl is less negative than HF. But this trend is only insured by the overlap matrix, fluorine basis functions overlap more with hydrogens than chlorines. Many studies on concludes that Mulliken population depends heavily on the atomic orbitals chosen for the calculation. There are many options for choosing basis set for quantum mechanics calculations, considering that there is no computational limitation for the quality of the wavefunction bigger basis sets are preferred. On the other hand, if there are computational limitations (which is rarely the case) a minimal basis set is preferred. Mulliken can benefit from the chemically intuitive basis sets such as a minimal basis set at the cost of mathematical accuracy.

Table 2-1. Mulliken population for methane atomic charges (C and H) for different basis sets with varying number of total atomic functions (Nbasis) B3LYP/(STO-3G, Def2-SVPP, Def2-SVPD, Def2-TVPP, Def2-TVPD, Def2-QZVPP, Def2-QZVPD).

N_{basis}	Basis set	C	H
9	STO-3G	-0.309	0.077
22	Def2-SVPP	-0.443	0.111
52	Def2-SVPD	0.447	-0.112
73	Def2-TVPP	-0.382	0.096
87	Def2-TVPD	-0.889	0.222
177	Def2-QZVPP	-0.156	0.039
195	Def2-QZVPD	-0.579	0.145

Different basis sets and resulting atomic charges for methane molecule is provided in Table 2-1. The variation between different basis set cannot be just attributed to the number of basis functions

but we can confidently say that Mulliken is a basis set dependent population. For Def2-SVPD diffuse basis set we see a negative hydrogen charge on methane, while the chemical expectation is that the hydrogen donates electrons to carbon which results in a positive atomic charge.

Populations that to make up for the chemical intuition with specialized series of orthogonalizations with separation of chemically relevant spaces already exists.[39] NPA population, which will be compared to Mulliken in the later chapters, is very ambiguous and mathematically nonsensical. We propose a more elegant way of using non-canonical atomic orbitals that will be unique to the molecular wave function called Quasi-Atomic Orbital (QAO). These orbitals will be obtained by trying to resemble minimal atomic orbitals from the

2.2 Motivation

The aim of this work is twofold:

1. New orbital based population analysis QAO with Mulliken in comparison with the popular (CM5, QH, ESP, Mulliken, NPA) or referred to as "common population analysis" in this work. QAO populations will be assessed for mathematical and chemical traits and compared to common population analysis with a large database of molecules in chapter three.
2. New orthogonalization scheme for QAO orbitals. We expect QAO might fail to provide reliable chemical trends because they are non-orthogonal. ZBD orthogonalization is a more chemically intuitive way of orthogonalization than Löwdin. The resulting chemical trends with ZBD-QAO is expected to be more chemically intuitive than non-orthogonal QAO. Before the comparative assessment, ZBD-QAO charges will be tested with a small dataset from the reference[37] then chemical trends performance of all 11 population analysis will be assessed in chapter 4. The structures of molecular interactions are provided in appendix B.

2.2.1. Chemical Database

Molecules and molecular interactions chosen for comparison consists of 1866 structures ranging from well-known to unusual, organic to metal clusters. Most of these structures have been reported by the literature and some of them were generated and optimized with the specifications given in section 2.3. The list of the datasets from literature and the a list of the chemical formulas are included in appendix A. Structure are ordered by ID used in Vetee, program written for handling population analysis data. It should be noted that ID's don't always represent the same structure or state; conformers and different charges states of a molecule is grouped in the same ID. Different states are distinguished by the job ID which is not provided in this work.

Before comparing populations, a set of realistic expectations of behavior should be provided since there is no experimental data to compare to. Therefore, there is not a way to deduct one method is strictly better than the other. Instead, what we can do is define properties that we expect to observe from populations and compare their performances. Desired traits can be divided into two main topics of chemical and mathematical assessment. While both traits are desired in any given partitioning method, the end goal is to define a method that is chemically reasonable and mathematically reliable.

2.2.2. Chemical Intuition

The goal of every population is to gain a better understanding of chemical system. Therefore, the atomic charges should match the experimentally observed chemical reactivity. In retrospect, this possibly makes this trait the most difficult to assess because there are no target values to compare to. Instead, a viable option is comparing the atomic charge differences in intermolecular and intramolecular system. Atomic charges are expected to be in a reasonable range. Universally accepted intermolecular theory of atoms can include electronegativity and radius size. For example, going down the periodic table for halogens we expect increase in radius size and decrease in electronegativity. A population analysis can fail to satisfy chemical expected intermolecular behavior if a fluorine is more positive than a chlorine atom. While ideally each molecule should be distinct, the atomic constituents should be transferrable to different chemical environments. Populations should be applicable to all atoms and should give reasonable atomic charges for different total molecular charges. Some of the populations that will be discussed later are optimized for neutral molecules, but it is crucial that any population should be able to describe charged structures. Atomic charges should also be correlated with the oxidation state of the atom. While oxidation states are integers, atomic populations should never be rounded to integers. Ideally,

we expect atomic charges lower than oxidation states. While there is not a definite answer to what an atoms atomic charge should be, for a ground state molecule it should not exceed valence capacity of an atom. For example, the atomic charge of a hydrogen atom of in an organic system should not be close to (or greater than) one. When the electronic structure is delocalized like in a benzene molecule, the atomic charge should be evenly spread out to atoms. Minor changes to molecular system should not result in drastic or unpredictable changes, instead it should be a systematic change in numerical values. Conformational, or rotations around the free bonds, effect on charge partitioning methods have been studied over two decades.[40] Ideally, we expect small changes in the molecule to not greatly affect the atomic charges. Similarly, atomic charges of different conformations of a molecule should be similar and stable with rotations around bonds. The charges should be in accordance with the anionic and cationic makeup of the molecule such that the actives sites can be identified.

2.2.3. Mathematical Stability

Partitioning method should be bias-free and only dependent on the unique wavefunction of the molecule. Computational robustness is expected, as well as unique atomic charges. For canonical orbital based population methods (Mulliken, NPA, atomic basis functions are predetermined to reflect the chemists' expectations from orbitals. However, this expectation makes the atomic orbitals used to describe the molecular system not unique and open to bias resulting in basis set dependency. There are many different basis sets in literature and partitioning methods should be independent of this choice. Similarly, level of theory used to construct the wave function should not greatly affect the results.

2.3 Computational Details

Molecules from each dataset were selected to test the limitations of some common density-based partitioning populations(Hirshfeld, Charge-model 5 or CM5),[13,15] electrostatic potential fitting population (ESP) from Hu-Lu-Yang,[41] and three widely available orbital-based population analysis: Mulliken (M), Minimal Basis Set (MBS),[28] and Natural Population Analysis (NPA).[42] These results are compared to three methods based on quasi-atomic orbitals (QUAMBO, QUAO, IAO) in both their Mulliken and Löwdin formulations, with and without ZBD orthogonalization.

Structures without coordinates were generated using PubChemPy[43] package. These coordinates were optimized using *Gaussian16* (version B.01) software package.[14] (For datasets with optimized geometries, the optimization step was skipped and only the population analysis was carried out.) Geometry optimizations were performed with HF[44], B3LYP[45] and ωB97XD[46] levels of theory and Def2-SVP basis set, ultrafine integration grids, very tight linear equation solutions for SCF procedure and very tight optimization criteria. After the convergence, using stable=opt keyword the geometry was confirmed to be the local minimum. All other options in the Gaussian program were set to their default values, except occasionally additional keywords were needed to obtain convergence. NPA, Hirshfeld, Mulliken, ESP atomic charge calculations were generated with same integral and SCF options using basis sets[47,48] def2-SVPP, def2-SVPD, def2-TZVPP, def2-TZVPD, def2-QZVPP basis sets were used; these basis sets use pseudopotentials to estimate relativistic effects for atoms beyond Krypton (Z=36).[47,48] These property-optimized augmented basis sets were chosen because they cover the wide variety of atom types in our database

Chapter 3

3 Comparison of Population Analysis

3.1 Method

Approximation of the wave function method (or referred to as the level of theory) effect on the atomic charges have been assessed in this section.

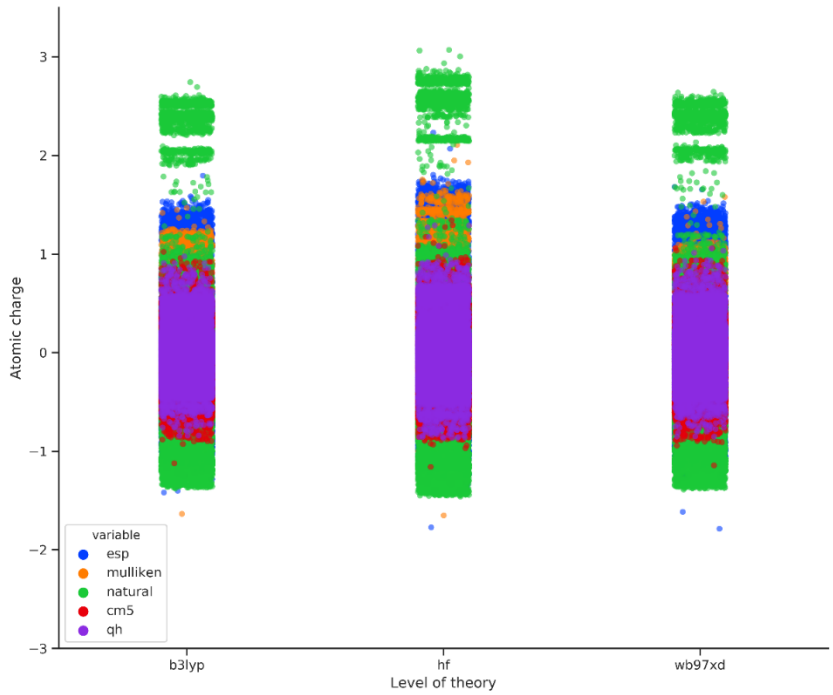

Figure 3-1. Scatter plot of atomic charges with three methods (B3LYP, HF, ωB97XD)/Def2-SVPP of common populations (CM5, Mulliken, NPA, QH, ESP) charges below negative three are omitted.

Atomic charge spread with three different level of theory for common population methods has been plotted Figure 3-1 as a stripplot[2]. This plot is a scatterplot where one variable is a category, in this case is the wave function method: B3LYP, HF, ωB97X. Some of the low-lying atomic charges have been omitted due to scaling issues and will be discussed later in Outlier Structures section. This graph gives a very rough look at atomic charge spread with different methods, HF method gives the highest spread out of all three. While the lower boundaries for all three levels of theory are similar (except for some outliers from Mulliken and ESP) the spread of the higher boundary is considerably more significant. Between B3LYP and HF, NPA charges above two can shift up to ~0.5. High positive charges can be dependent on the level of theory, especially for the NPA population. The same magnitude of method dependency can be observed with Mulliken as well for positive charges around positive one. We can tentatively say that ESP population is also level of theory dependent, but it is not very clear from this Figure 3-1. Hirshfeld flavors CM5 and QH seem to be the most stable population with changing level of theory and are the least spread out. These atomic charges tend to stay in the range of positive to negative one with CM5 being slightly more spread out than QH. The spread of atomic charges for each partitioning method can be further studied numerically in Figure 3-2. Each histogram is the distribution of $|q_{max} - q_{min}|$ where q_{max} and q_{max} are the maximum and minimum atomic charge of an atom for three methods: HF, B3LYP, and ωB97XD. All these graphs are skewed right histograms with greatest densities around zero and two to three minor peaks varying between 0.05 to 0.3. Ideally, we should be looking for a shorter tail for a smaller range in atom charge difference with changing levels of

[2] https://seaborn.pydata.org/generated/seaborn.stripplot.html

theory. CM5 and QH (almost identical in density) are still the least method dependent with sharp and the highest peak around zero compared to the other population methods. Also, these two graphs have the shortest tails not exceeding 0.5 and three minor peaks around 0.04, 0.06 and 0.11. The major peak around zero is shorter and wider for the ESP method. The tail is extending almost to 0.8 pointing towards a much higher level of theory difference. The minor peaks from the Hirshfeld techniques, CM5 and QH, are shifted by 0.02, 0.18 and 0.2. If a method dependency occurs with Hirshfeld techniques, then this dependency may be higher with ESP. NPA is comparable but a bit worse than ESP in dependency; the major peak is lower in density and shifted slightly away from zero. Also, the densities of the minor peaks are less prominent. Since the minor peaks are common to all population analysis methods, their presence may be caused by the molecular structures. The Mulliken population analysis histogram has the widest peaks among all tested methods, and its highest density peak is not even close to zero, instead lying above 0.05. There are only two minor peaks and they lie around 0.11 and 0.3 making them the highest out of all the considered methods. Although the tail of the graph is slightly above 0.5, looking at the frequency densities we can summarize that Mulliken is the most level of theory dependent population analysis technique. Out of all the flavors of QAO, IAO seems to be the least level of theory dependent with a shorter tail that is comparable to CM5 and QH methods.

The tails of CM5 and QH graphs are from dataset MB08 (ID 1354, 1402, 1273) and all different atoms (F, C and H respectively). Higher atomic charge differences between different methods cannot be attributed to an atom type with the MB08 dataset, but the dissociating bond effect on robustness of CM5 and QH should be be considered. Also, description of long distance interactions are very different with ωB97XD compared to HF. NPA tail (ID 1339, 1419, 1344) and ESP tail are both from the MB08 dataset (ID 1339, 1402, 1344) at the HF level. Unlike other

population analyses, the QUAO tail is mostly the X40 dataset (ID 1460 1460, 1326) for both HF and B3LYP. Molecule ID 1460 seems to be high in variance for carbon and bromine atoms with QUAO; bromine is the end of the tail with variance of one.

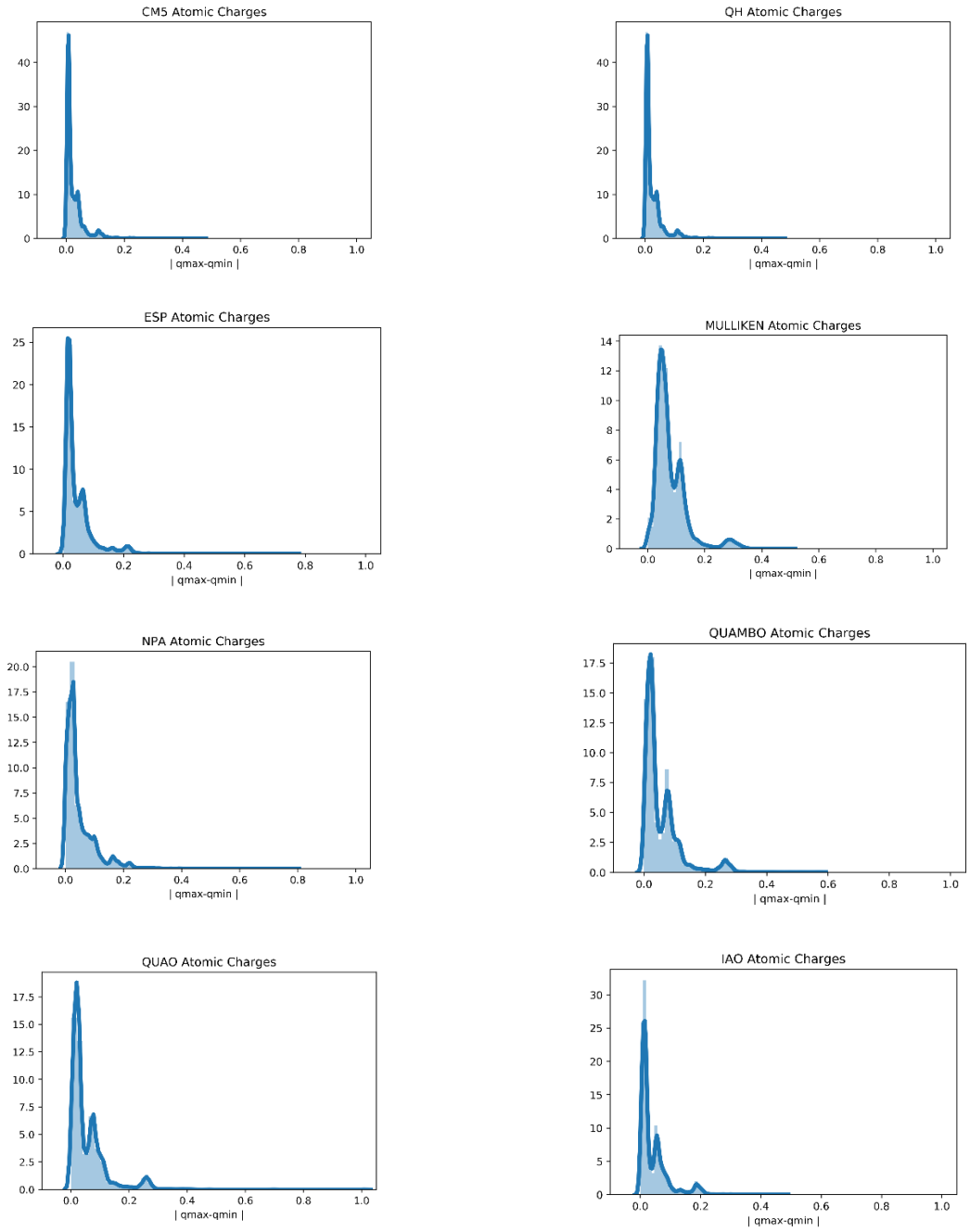

Figure 3-2 Distribution of max. min. atomic charge difference for three different levels of theory (B3LYP, HF, ωB97XD) for CM5, QH, ESP, Mulliken, NPA, QUAMBO, QUAO, IAO populations. One unit on the y axis is a thousand atomic charges.

The spread for the level of theory dependency is further studied with the variance σ^2 which is a measure of how far each atomic population is from the mean for studied methods.

$$\sigma_A^2 = \frac{\sum_{m\in\text{molecules}} \sum_{\text{atom } a\in m} \sum_{q=\{\text{HF,B3LYP,}\omega\text{B97XD}\}} \left(x_{maq} - \overline{x_{ma}}\right)^2}{n_A}$$
$$\sigma^2 = \frac{\sum_{\text{all molecules m}} n_A \sigma_A^2}{n_{\text{total number atoms in all molecules}}} \quad (36)$$

where average variance σ^2 is calculated as the variance of each atom as molecules for three different methods. Atomic charge x_{maq} and $\overline{x_{ma}}$ is the average of the atomic charge over all methods: HF, B3LYP, ωB97XD. To omit the effect of basis set dependency only population calculations carried out with one basis set Def2-SVPP is considered:

Table 3-1. Average and maximum variance of each population for same atom with three different methods (HF, B3LYP, ωB97XD) with one basis set Def2-SVPP.

	Avg(σ2)	Max(σ2)
CM5	7.00E-05	6.90E-02
QH	7.00E-05	6.90E-02
ESP	1.08E-03	1.68E-01
Mulliken	6.54E-03	7.56E-02
NPA	7.54E-04	1.89E-01
QUAMBO	7.80E-04	9.29E-01
QUAO	6.51E-04	2.21E-01
IAO	3.15E-04	8.72E-02

Average values of variance of each atom with different methods are presented in Table 3-1. On average Mulliken population analysis, followed by ESP and NPA, is the most method dependent population analysis. The least method dependent population is Hirshfeld flavors CM5 and QH followed by QAO flavors. The most method dependent QAO flavor is QUAMBO and the least dependent is IAO. In accordance with the previous findings, the HF method increase the variance

for all methods, if HF calculations are omitted all average variance values go down. Maximum values of the variance values are also provided in the second column of Table 3-1, while the highest average variance is seen for Mulliken the maximum variance is seen for QUAMBO. This is mainly because of some atom types are unstable with QUAMBO and QUAO. Maximum variance of IAO is less than other QAO flavors. It is interesting to observe that maximum variance of Mulliken is comparable to that of CM5 and QH. Meaning, this value is not an outlier (atoms from MB08 dataset), but Mulliken population has high variance for most atoms with different methods.

To summarize the findings of method dependency, with or without the mindless DFT benchmarking dataset MB08, is given by:

	CM5	QH	ESP	Mulliken	NPA	QUAMBO	QUAO	IAO
Method independence (with MB08)	✓	✓	×	×	×	×	×	✓
Method independence (without MB08)	✓	✓	✓	×	✓	✓*	✓*	✓

Mulliken is the most method dependent population followed by ESP and NPA. Method dependency can be improved for ESP and NPA by taking out the MB08 dataset. This cannot be said for Mulliken, this population stays method dependent for all datasets. QUAMBO and QUAO in general improves the method dependency of Mulliken except for Br atoms.

3.2 Basis Set

In this section the effect of increasing the number of basis functions, or using a bigger basis set, on population analysis methods is studied.

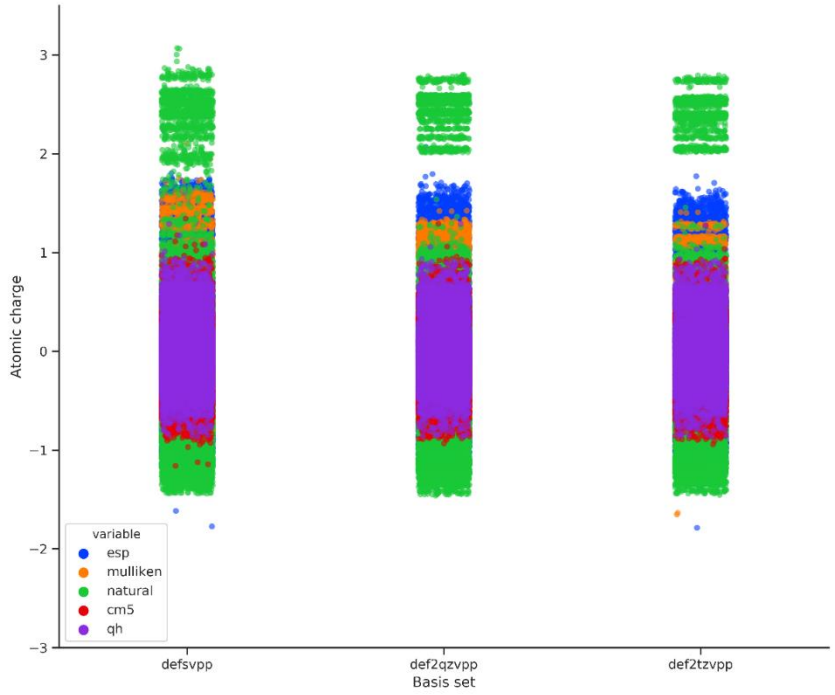

Figure 3-3 Scatter plot of atomic charges with three different level of theory basis sets (Def2-SVPP, Def2- TZVPP, Def2-QZVPP) for population analysis methods (CM5, Mulliken, NPA, QH, ESP) omitted atomic charges below negative three.

The most spread out atomic charges can be seen with Def2-SVPP basis set from Figure 3-3. Mulliken charges are considerably higher and the band structures seen with atomic charges above positive two for NPA with Def2-TZVPP and Def2-QZVPP basis sets are fuzzier for this basis set. Comparing these band like structures seen in higher charges for NPA with Figure 3-1, for basis set these band structures are more prominent. This can be attributed to NPA qualitatively being more basis dependent than method dependent. The preferable basis set in atomic charge quality framework is Def2-QVPP. Use of this basis set results in the least spread out atomic charges.

Furthermore, the benefit of increasing the number of basis functions is clear with the decreasing number of outlier atomic charges. Like the previously seen with method dependency, basis set dependency is the least for CM5 and QH.

Like level of theory dependence histograms in in Figure 3-2 basis set dependency can be assessed quantitatively Figure 3-4. A higher dependency (on average) can be observed with basis set compared to method. This can be proven by looking at the maximum tail of a histogram; for method dependency it is one (with QUAO) while in Figure 3-4 it is above two (with Mulliken). CM5 and QH graphs show a very narrow peak around 0.03 which makes them the most basis set stable two population methods. These methods also have the shortest tail extending to a smaller range of 0.44. Opposite to this case, Mulliken and ESP are more basis set dependent than method. Both have wider peaks resulting in a more spread out distribution with the change of basis. Primary density peak of Mulliken is located below 0.1, for basis set it is approximately located between 0-0.5 almost the five times more spread out than the former. Furthermore, both methods have longer tails extending past 1.75 and 2.0 respectively. Considering both the peak width and tail length we can summarize that Mulliken followed by ESP is the most basis set dependent population analysis methods. There is higher density at the tail ends for CM5, QH and QUAMBO. for basis sets than methods. The tails of the CM5 and QH like the method graphs, are composed of molecules from MB08 dataset (ID 1383, 1343, 1393) and all are oxygen type atoms. The longest tail seen for Mulliken consists of carbon atomic charges from zwitterionic molecules of JOC dataset (ID 812). The same molecule ID 812 seems to be a problematic molecule for population analysis methods because we can find it also at the end of ESP tail. Instead of the carbon atomic charges like Mulliken, ESP is unstable with nitrogen atoms. Molecular structures with greater basis set dependency for QUAMBO, QUAO and IAO are not the same. Starting with QUAMBO, the tail

is almost the same carbon atomic charges as Mulliken tail in a lower magnitude. QUAO has tail is mostly from Br charges intermolecular interaction structures of X40 dataset (ID 1460,1479,1466). Bromine atom disrupts QUAO computational robustness. Last QAO flavor IAO tail has molecules from (ID 1383, 1343, 1393) from MB08 mindless dataset and for only HF method and only for oxygen atoms. The same oxygen atomic charges also destabilize NPA method (ID 1343, 1383, 1303). From basis set dependency perspective QUAMBO and Mulliken, IAO and NPA have the same challenges. QUAO seems to be the worst out of all three at explaining halogen bonding.

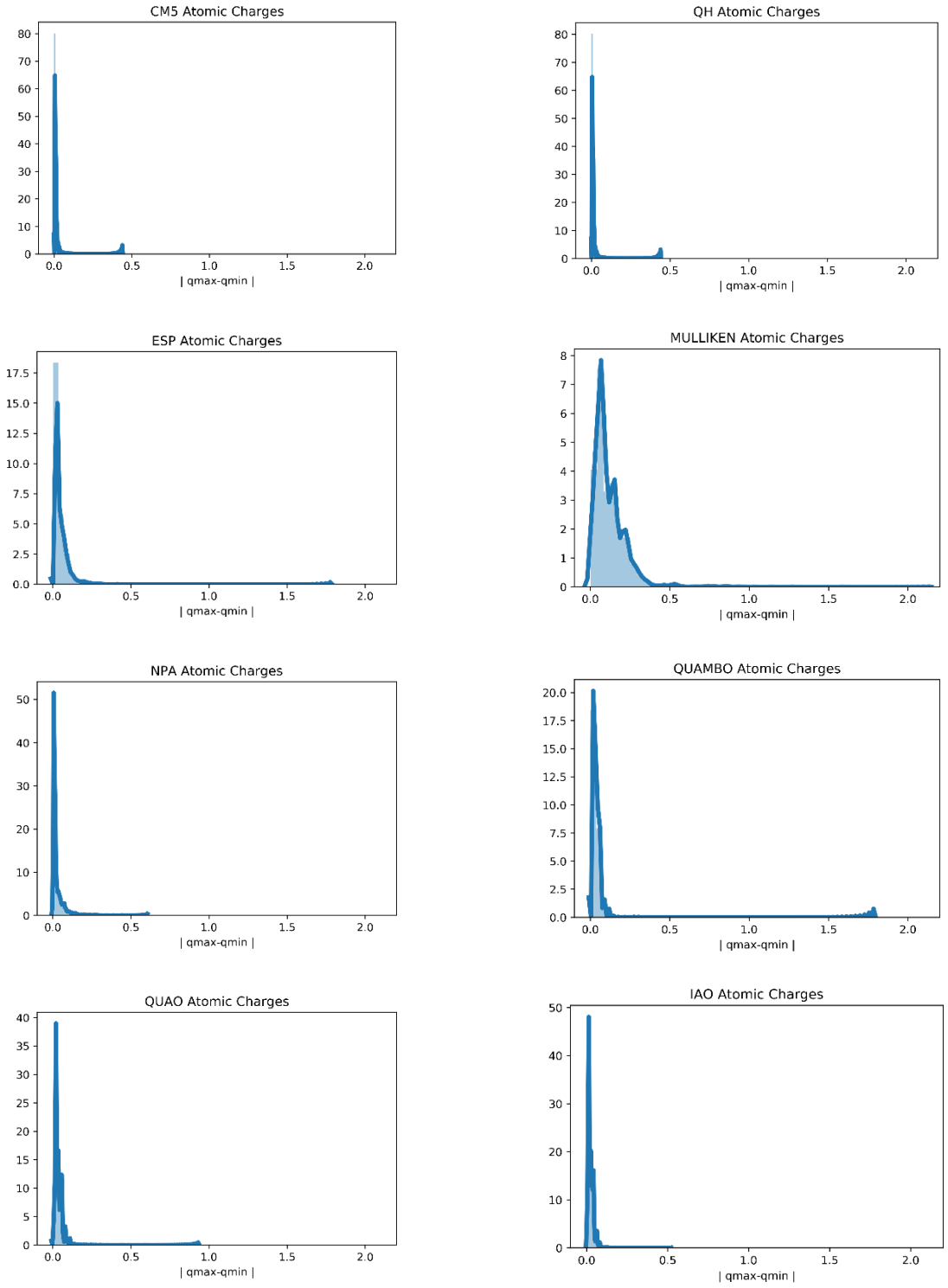

Figure 3-4. Distribution of max. min. atomic charge difference for three different basis sets (Def2-SVPP, Def2-TZVPP, Def2-QZVPP) for CM5, QH, ESP, Mulliken, NPA, QUAMBO, QUAO, IAO populations. One unit on the y axis is a thousand atomic charges.

Next, we will consider how basis set dependency by with looking at the average and maximum variance σ^2 for three different basis sets. Comparing this basis set table with the previous method variance in Table 3-1, higher numbers are observed. This means that all populations are more basis set dependent than method dependent.

Table 3-2. Average variance of each population for same atom with three different basis sets (Def2- SVPP, Def2- TZVPP, Def2-QZVPP) with B3LYP method.

	Avg(σ2)	Max(σ2)
CM5	6.37E-04	6.42E-02
QH	6.37E-04	6.42E-02
ESP	1.40E-03	8.78E-01
Mulliken	2.60E-03	1.12E+00
NPA	1.60E-03	1.05E-01
QUAMBO	2.08E-03	9.29E-01
QUAO	2.02E-03	2.21E-01
IAO	3.15E-04	8.72E-02

The highest value is seen with Mulliken population from Table 3-2 followed by NPA, ESP. The lowest average variance is with Hirshfeld CM5 and QH, so these are the least basis set dependent populations. QUAMBO and QUAO are only slightly less basis set dependent than Mulliken population. IAO is an improvement over Mulliken and other QAO flavors, there is a clear decline in maximum variance from Mulliken-QUAMBO-QUAO-IAO.

Using QAO as orbitals instead of canonical ones decreases basis set and method dependency for Mulliken population analysis. Dependency for QUAMBO, QUAO, and ESP can be decreased by omitting some datasets like MB08. Some atom types increase the dependency with QAO flavors so in the next section we will be talking about effect of atom type.

	CM5	QH	ESP	Mulliken	NPA	QUAMBO	QUAO	IAO
Basis set independence (with MB08)	✓	✓	×	×	✓	×	×	✓
Basis set independence (without MB08)	✓	✓	✓	×	✓	✓*	✓*	✓

3.3 Atom Type

Atoms are the fundament of every population analysis, so any population analysis method should be applicable to all atom types. Atomic charges in the database (for total number of 143883 atoms) are pair plotted for clear visualization of correlation between different population analysis. Due to the number of studied methods we have split the pair plots into three sections; common-common, QAO-QAO and finally QAO-common correlation graphs. To start with, correlation of common studied population analysis methods is presented Figure 3-5. For each atom type a different marker is used and provided in the legend on the right side of the graph. It should be emphasized that charges that were omitted in previous graphs (extremely negative sodium charges) is shown in this graph. However, this is not applied in the next two correlation graphs (QAO-QAO and QAO-common) since the outlier charges for QAO flavors are on the positive side it causes scaling issues. These omitted charges will be revisited in outlier charges section.

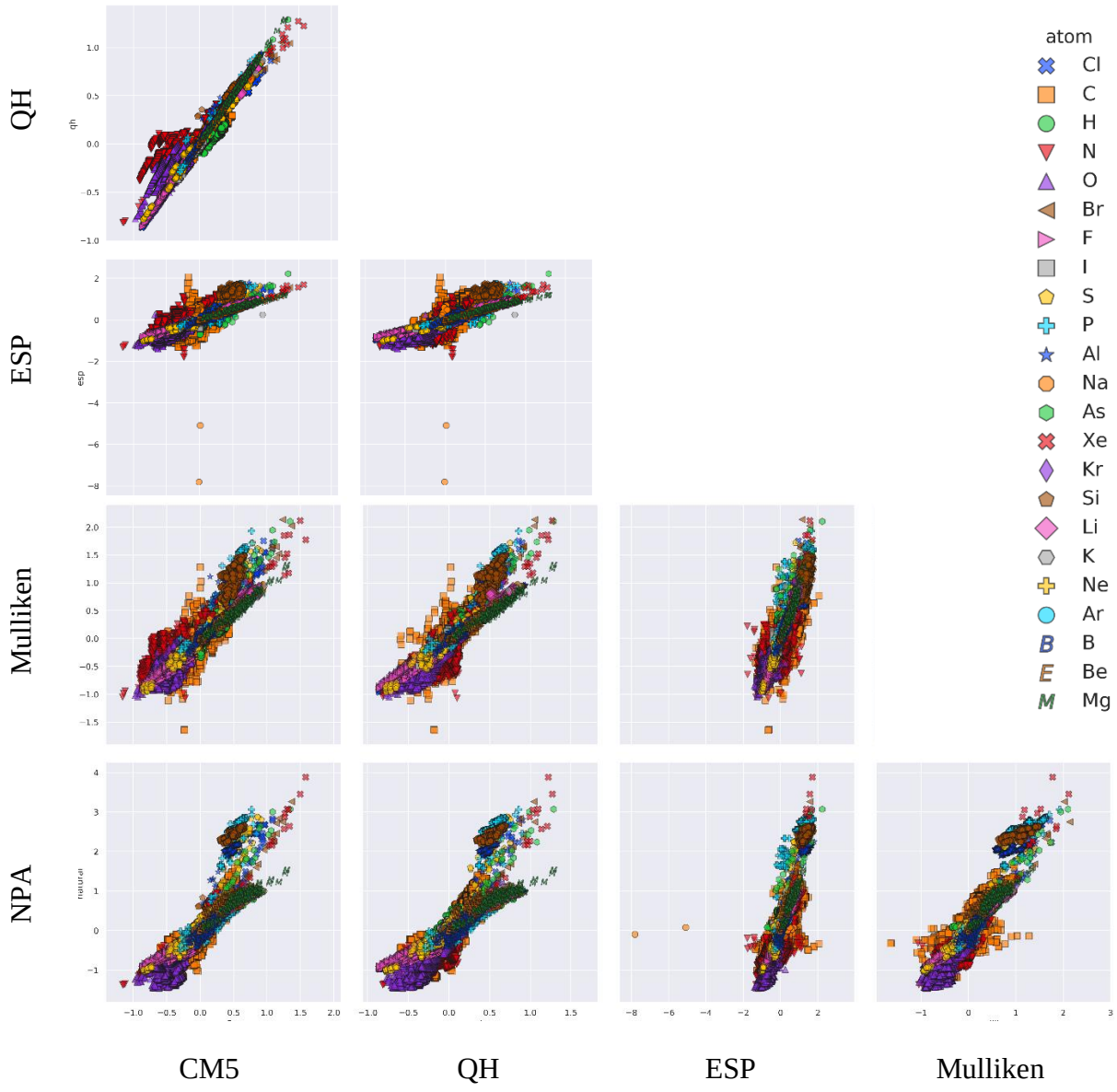

Figure 3-5 Correlation of atomic charges from different (CM5, Mulliken, NPA, QH, ESP) population analysis for B3LYP/Def2-SVPP separated by atom type.

The highest (almost linear) correlation is seen between CM5 and QH. This is to be expected because both are Hirshfeld population analysis flavors. From the scatter plot the most different charges seem to be for nitrogen atom. If this atom type is taken out of the database the regression line slope increases from 0.58 to 0.64, hence increasing the correlation. Considering that database at hand has some charged species it is unusual to see atomic charges for different atom types are

all in a small range for these methods and should be studied with high expected charges. Oxygen charges, concealed by the nitrogen markers, are also slightly different QH charges are on average higher (-0.08) than CM5 charges (-0.42). Attributing this difference between two methods to atom types might be an oversimplification of the problem, since most datasets with nitrogen and oxygen have charged molecules. This behavior might be due to the existence of charged species especially considering both atom types are spread out for other methods. Nitrogen and oxygen charges for CM5 and QH will be further studied in the next sections.

The most notable difference in the correlation plots is for ESP with outlier Na charges from ID 46, which were omitted for other figures better scaling in most figures. Compared to other populations, not only ESP has the highest carbon charges (higher than positive two) but also it has the most spread out charges. Further investigation on atom type effect on population analysis is required. NPA seems to produce the highest range for atomic charges. NPA correlation graphs are the most unique out of all the others, there is a clear deviation form linear dependence with atom types silica, potassium, aluminum. Atomic charges of xenon extend to positive four, which for other populations extend up to positive two.

Localization of several atom types can be seen among most population analysis. An example is silica which is most delocalized with Mulliken ranging between approximately 0.5 to 1.5, while for others it is very localized to a 0.5 atomic charge range. A higher range is not necessarily a defect, but it requires closer inspection to the reason why.

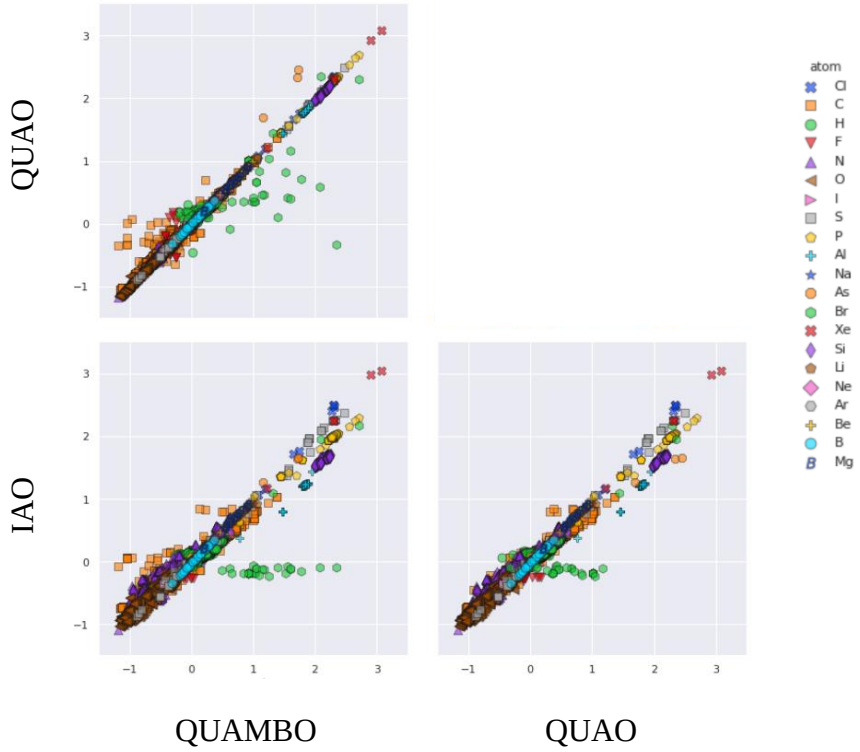

Figure 3-6. Correlation of atomic charges between different QAO flavors (QUAMBO, QUAO, IAO) population analysis for B3LYP/Def2-SVPP separated by atom type.

Atomic charge difference between three types of QAO methods can be seen from Figure 3-6. Bromine and carbon charges are the most prominently difference. The best regression line fit is for QUAMBO and QUAO, which seems to be unstable with bromine atoms ranging between 0 – 2.5 and 0 – 1.1 respectively. Bromine is not the only type of atomic charge to be localized with IAO, carbon charges also are in a smaller range. There are still omitted charges above sodium atom, these molecules will be discussed later. Furthermore, when compared to Figure 3-5 there are less data points. Chromium atoms are missing from the data and the reason should be investigated. Correlation graphs show clear localization of charges (except for the unstable atoms) such as xenon into three distinct categories. This should be further investigated if this pattern makes sense or not.

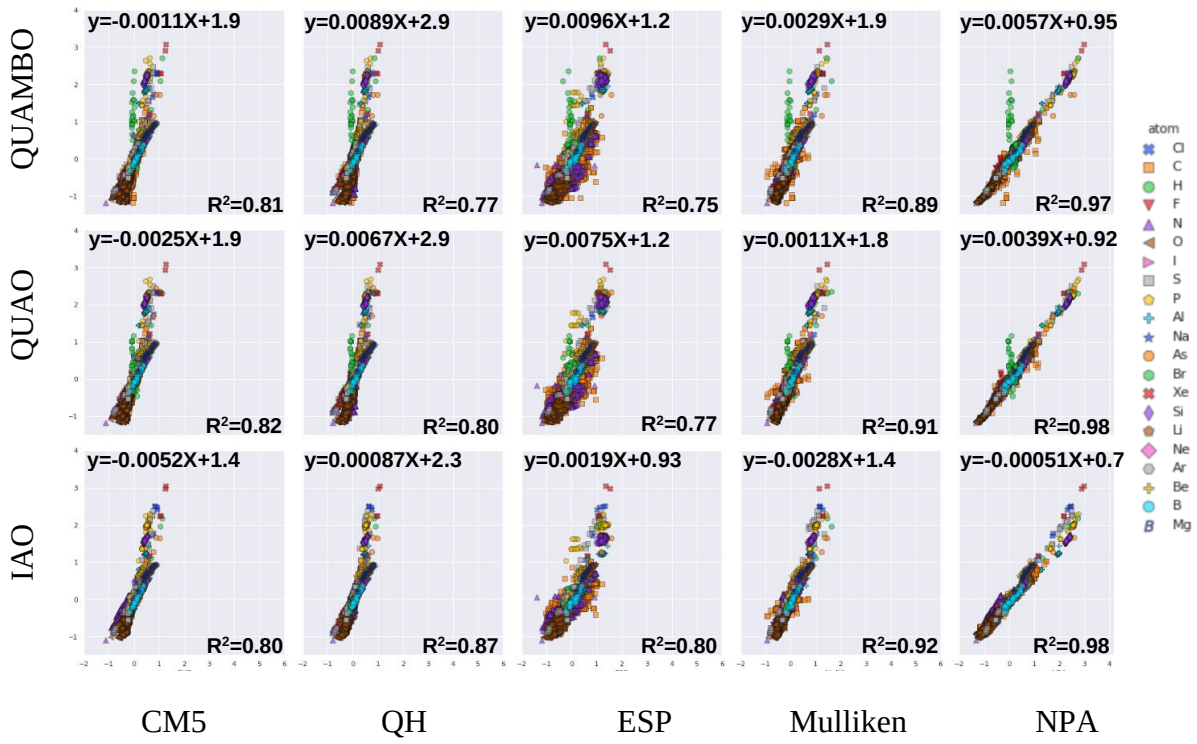

Figure 3-7. Correlation of atomic charges of QAO population methods (QUAMBO, QAO, IAO) with common population methods (CM5, QH, ESP, Mulliken, NPA) linear regression line equation and R^2 values given on each graph. Data is filtered for B3LYP/Def2-SVPP.

QAO methods have a clear different slope for each common population methods. The highest QAO correlation is with NPA, signified by a highest regression line fit of ~0.98. The fit is improved by omitting outlier charges (bromine and sodium) from QAO calculations. For well-known molecular structures QAO can be suggested instead of NPA populations. IAO is the QAO flavor that has the highest correlation with NPA population.

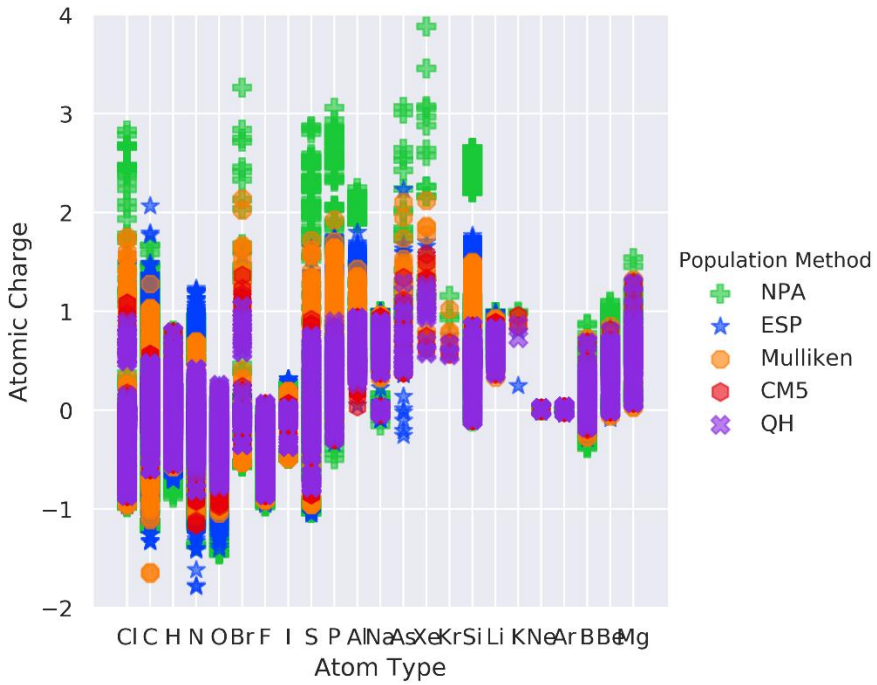

Figure 3-8 Spread of atomic charges categorized by atom type for common population methods (CM5, QH, ESP, Mulliken, NPA) with B3LYP/Def2-SVPP. Sodium outlier charges are omitted.

To be able to show atomic charge data categorized by atom type in Figure 3-8 Spread of atomic charges categorized by atom type for common population methods (CM5, QH, ESP, Mulliken, NPA) with B3LYP/Def2-SVPP. Sodium outlier charges are omitted., sodium charges had to be neglected because this is specific case, an “extreme” molecule so to say, which is considered a “breaking point” for some population analysis methods. These cases will be visited in later sections. From this graph it can be clearly seen that the variance for different atom types can be different for population analysis methods. NPA appears at the upper boundary for many atoms except for a few cases. Atomic charges of C, H, N, O, F, I with NPA is not the highest observed value. Granted, some of these are not chemically intuitive, such as carbon charges above two. The highest atomic charge boundary is seen with NPA for xenon atom, close to positive four (some extreme charges are still not shown). We can also see that the bromine charges go up to an unexpected positive

number. The close to one hydrogen charges should be studied as well, HF level of theory is expected to fail, but other methods should be giving a reasonable charge.

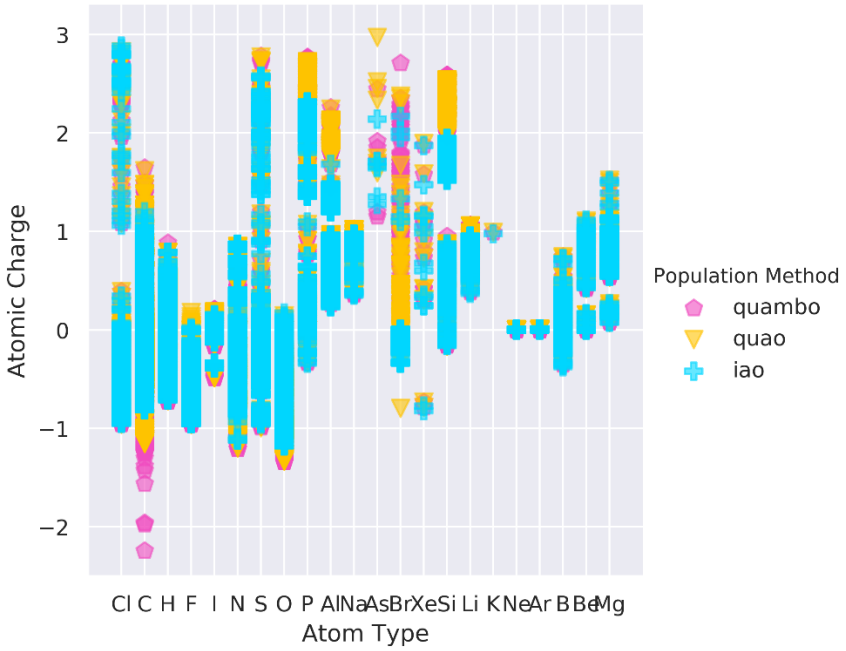

Figure 3-9. Spread of atomic charges categorized by atom type for QAO population methods with B3LYP/Def2-SVPP. Na outlier charges are omitted.

When the atomic charges for QAO methods are separated by atoms, we can see discrete levels of atomic charges like valence theory[49,50] Like the case for common populations in Figure 3-9, QAO methods break with sodium charges that were omitted for scaling sake. The most spread out flavor seems to be QUAO. There are some chemically unintuitive charges, like the chlorine and bromine charge going up to three. For the lower limit QUAMBO gives lowest atomic charges, below negative two for carbon. This is lower than the common population analysis methods limit, but there are improvements such as IAO method being stable with carbon atom in acceptable charges range. Clear clustering of the charges can be seen for many atoms (Mg, Be, Si, Xe) unlike common population methods.

	CM5	QH	ESP	Mulliken	NPA	QUAMBO	QUAO	IAO
Atom type	✓	✓	×	×	×	×	×	✓

3.4 Conformation

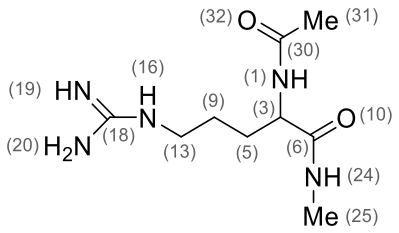

Figure 3-10. Structure of Arginine with indices of heavy atoms given in parenthesis ID 2.

Conformational stability is asses with amino acids. These molecules have similar backbones of amine, alpha carbon bonded to side chain specific for the amino acid and a carboxyl group. Four neutral amino acids: Alanine, Arginine, Aspartic Acid with 11, 58, 12, and 21 conformers are studied. Since Arginine has the highest number of conformations it will be the example for conformation assessment.

The variance of heavy atom indexed as in Figure 3-10 is given in Table 3-3. Overall, the conformer variance values are lower than the average variances seen for method and basis set. This might be because the populations are less dependent on conformational changes, but we cannot be certain until these assessments have the same sample size. CM5 and QH are the least conformation dependent population methods, rarely deviating from zero. ESP turns out to be the most conformation dependent method out of all studied population analysis. As the number of considered conformations increase, NPA and Mulliken atomic charges destabilize. For the amino acid with the least number of conformations Alanine, almost all population methods are stable except for ESP.

Table 3-3. Variances of heavy atom charges for 58 conformers of Arginine for common (CM5, QH, ESP, Mulliken, NPA) and QAO (QUAMBO, QUAO, IAO) populations with B3LYP/Def2-SVPP ID 2. For clarity, atomic charges are conditional formatted on color scale; background darkens as the charge increases.

Atom	CM5	QH	ESP	M	NPA	QUAMBO	QUAO	IAO
N (1)	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
C (3)	0.00	0.00	0.07	0.00	0.01	0.01	0.01	0.01
C (5)	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00
C (6)	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00
C (9)	0.01	0.00	0.05	0.01	0.04	0.04	0.03	0.03
O (10)	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01
C (13)	0.01	0.00	0.04	0.00	0.01	0.01	0.01	0.01
N (16)	0.02	0.00	0.04	0.01	0.04	0.03	0.03	0.03
C (18)	0.00	0.00	0.02	0.00	0.01	0.00	0.00	0.00
N (19)	0.02	0.01	0.06	0.01	0.06	0.05	0.04	0.04
N (20)	0.01	0.00	0.03	0.01	0.02	0.02	0.02	0.02
N (24)	0.01	0.00	0.02	0.01	0.03	0.03	0.02	0.02
C (25)	0.00	0.00	0.04	0.01	0.02	0.03	0.03	0.02
C (30)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C (31)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
O (32)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Highest variance is seen for Alanine for two carbon atoms with index 3, 5, and 19. Since these atoms are not buried in the center of the molecule (index 19 is at the end of the molecule) a possible reason might be that ESP being not stable for carbon charges. Higher variance persists with ESP for Arginine and Aspartic Acid. For Arginine the highest variation is for carbon index 2 and like in Alanine's case this is the carbon between the amine and carbonyl group, referred to as the alpha carbon.

	CM5	QH	ESP	Mulliken	NPA	QUAMBO	QUAO	IAO
Conformational stability	✓	✓	×	✓	✓	✓	✓	✓

3.5 Outlier Structures

3.5.1. N-methyl-[2,2,2]-azabicyclooctane

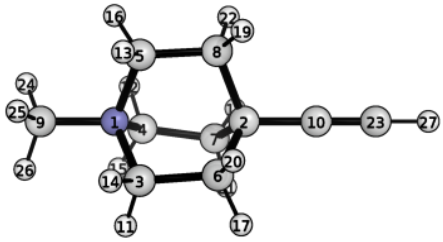

Figure 3-11. Structure of N-methyl-[2,2,2]-azabicyclooctane given indices on the atoms ID 812

So far, we have seen some prominent molecules form the database that causes population analysis methods to behave computationally less robust. Molecule ID 812 from JOC[51] (17-H structure on reference) zwitterion dataset has proven to cause a high basis set dependency and it requires closer inspection. The atomic charges of the heavy atoms for this molecule can be seen in Figure 3-11. Design of this anion sheds some light on the data, Figure 19 in JOC[51] and refers the unprotonated molecules (without hydrogen indexed 27) as the ligand. Positive to negative shift in atomic charge of neighboring atoms makes this molecule a strong nucleophile with a high dipole moment. As debated over by Weinstein, this is the proof of the non-classical cation due to the ionic carbon indexed two in Figure 3-3. Difference of atomic charge between neighboring carbon atoms causes the protonated structure enables stabilization through homoaromaticity[52] creating a σ -delocalized carbocation. Electron delocalization between carbon indices 4, 5, 6, 7, 8 is considered a chemically debatable in a neutral molecular environment.

Table 3-4. Heavy atomic (and hydrogen 27) charges and method/basis set variance in parenthesis for molecule ID 812 for common (CM5, QH, ESP, Mulliken, NPA) and QAO (QUAMBO, QUAO, IAO) populations. For clarity, atomic charges are conditional formatted on color scale; background darkens as the charge increases.

Atom	CM5	QH	ESP	Mulliken	NPA	QUAMBO	QUAO	IAO
N(1)	-0.23(0.01)	0.10(0.01)	-1.40(0.66)	-0.32(0.53)	-0.40(0.10)	-0.17(0.00)	-0.16(0.00)	0.05(0.00)
C(2)	0.00(0.00)	0.02(0.00)	-0.09(0.12)	-0.20(0.82)	-0.14(0.02)	-0.02(0.01)	-0.03(0.00)	0.02(0.00)
C(3)	-0.10(0.01)	-0.04(0.01)	0.87(0.14)	0.21(0.17)	-0.20(0.05)	-0.36(0.00)	-0.32(0.00)	-0.25(0.00)
C(4)	-0.10(0.01)	-0.04(0.01)	0.88(0.13)	0.21(0.17)	-0.20(0.05)	-0.36(0.00)	-0.32(0.00)	-0.25(0.00)
C(5)	-0.10(0.01)	-0.04(0.01)	0.87(0.13)	0.21(0.17)	-0.20(0.05)	-0.36(0.00)	-0.32(0.00)	-0.25(0.00)
C(6)	-0.18(0.01)	-0.08(0.01)	0.12(0.13)	0.04(0.14)	-0.40(0.03)	-0.45(0.00)	-0.40(0.00)	-0.28(0.00)
C(7)	-0.18(0.01)	-0.08(0.01)	0.09(0.16)	0.04(0.14)	-0.40(0.03)	-0.45(0.00)	-0.40(0.00)	-0.28(0.00)
C(8)	-0.18(0.01)	-0.08(0.01)	0.14(0.12)	0.04(0.14)	-0.40(0.03)	-0.45(0.00)	-0.40(0.00)	-0.28(0.00)
C(9)	-0.18(0.01)	-0.08(0.01)	0.95(0.42)	0.03(0.32)	-0.40(0.06)	-0.61(0.00)	-0.56(0.00)	-0.41(0.00)
C(10)	-0.05(0.01)	-0.04(0.01)	0.28(0.07)	0.68(0.61)	0.00(0.02)	0.01(0.38)	-0.01(0.00)	0.00(0.00)
C(23)	-0.22(0.01)	-0.17(0.01)	-0.57(0.04)	-0.81(0.96)	-0.27(0.03)	-0.40(0.65)	-0.36(0.00)	-0.29(0.00)
H(27)	0.11	0.06	0.24	0.01	0.2	0.26	0.24	0.15

Overall, the charges are mostly neutral except for a few outliers and high variances values cause by method and/or basis set dependency. The darkest colors (highest charges) are seen mostly for ESP and Mulliken, and the lightest color (lowest charges) with the first atom for ESP. The most negative charge Table 3-6 is a nitrogen indexed one, and the most positive is the carbon indexed nine with ESP.

N(1) : The carbon indexed 9 will not be discussed because atomic charges are directly corelated with charge of this nitrogen. Nitrogen indexed one is bonded to four carbon atoms and expected to be on the positive side. However, there is a negative charge delocalized in the bicyclooctane so the charge should be less than a regular nitrogen bonded to four carbons. Greatest difference is seen with ESP, not only the charge is wrong sign, the variance is quite high as well. Only positive charge is seen with QH and IAO.

C(3-5) : For negative N(1) charges the neighboring carbons (carbons with index 3, 4, 5, 9) are also positive (almost close to one with ESP). We expect electron delocalization in bicyclooctane , so

these carbons should be negatively charged. Mulliken and ESP fail in this regard, they also have a high variance, CM5 and QH have the correct right sign but the values are too low to indicate a delocalization. NPA and QAO populations are reasonable, if not for NPA's high variance.

C (2) : Expected to be negative, CM5 and QH fail to satisfy this, while other populations are all negative. Mulliken and ESP continue to have a high variance and not considered stable.

C (23) : In the reference this carbon atom is reported to be the negative probe of the molecule. Their findings are supported with ESP charges and this is the same case in Table 3-6. This atomic charge is negative for all populations, but too neutral for CM5 and QH populations.

H (27) : Reported as the acidic hydrogen, expected to have positive charge. All populations are positive, but QH and Mulliken charges are too neutral.

To conclude this outlier, the results of the assessment is presented in Table 2-1 IAO seems to be the best choice out of all populations. The presence of a large dipole moment and curried delocalization in the bicyclooctane cage results in ESP and Mulliken charges to be unreasonable. This is an example of neutral homoaromatic molecule where the scaling difference with CM5 can be seen clearly when the charges are compared to QH.[53]

Table 3-5. Population recommendation for N-methyl-[2,2,2]-azabicyclooctane ID 812 summary with CM5, QH, ESP, Mulliken, NPA, QUAMBO, QUAO, IAO

	CM5	QH	ESP	Mulliken	NPA	QUAMBO	QUAO	IAO
N (1)	×	✓	×	×	×	×	×	✓
C (3-5)	×	✓	×	×	×	×	×	✓
C (2)	×	×	×	×	✓	✓	✓	✓
C (23)	×	×	✓	✓	✓	✓	✓	✓
H (27)	✓	×	✓	×	✓	✓	✓	✓
Total	×	×	×	×	✓	✓	✓	✓✓

3.5.2. Trifluorobromomethane

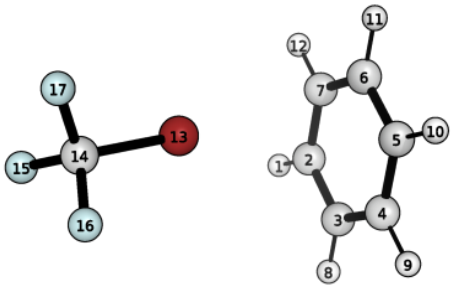

Figure 3-12. Structure of trifluorobromomethane (F_3CBr) interaction with benzene (C_6H_6) indices given on the atoms ID 1460.

The second outlier is from non-covalent halogenated compounds interaction dataset X40. The two interacting molecules are benzene and trifluorobromomethane ID 1460, labelled 29 in reference[54] the structure of the interaction with atom indices are given in Figure 3-12. It has been noted that there is as slight variance for the atomic charges for the trifluorobromomethane. Charges and variances in parenthesis are given in Figure 3-12.

Table 3-6. Trifluorobromomethane atomic charges B3LYP/Def2-SVPP and method/basis set variance in parenthesis ID 1460. For clarity, atomic charges are conditional formatted on color scale; background darkens as the charge increases

Atom	CM5	QH	ESP	Mulliken	NPA	QUAMBO	QUAO	IAO
Br (13)	0.01(0.00)	-0.01(0.00)	-0.13(0.00)	-0.05(0.00)	0.04(0.00)	2.35(0.03)	-0.34(0.09)	-0.09(0.00)
C (14)	0.35(0.00)	0.31(0.00)	0.47(0.01)	0.46(0.02)	1.03(0.01)	0.13(0.03)	0.37(0.09)	0.84(0.00)
F (15)	-0.12(0.00)	-0.10(0.00)	-0.13(0.00)	-0.14(0.00)	-0.36(0.00)	-0.30(0.00)	0.18(0.04)	-0.25(0.00)
F (16)	-0.12(0.00)	-0.10(0.00)	-0.13(0.00)	-0.14(0.00)	-0.36(0.00)	-0.35(0.00)	0.11(0.03)	-0.25(0.00)
F (17)	-0.12(0.00)	-0.10(0.00)	-0.13(0.00)	-0.14(0.00)	-0.36(0.00)	-0.30(0.00)	0.14(0.04)	-0.25(0.00)

Br(1) : Bromine is more electronegative than carbon, and less than fluorine. Therefore, we expect the charge of this atom to be more negative than carbon and more positive than fluorine atoms. CM5, QH, NPA, and QUAMBO populations all predict bromine as positively charged. QUAMBO

is the worst population for describing bromine charges and should not be used if the system has bromine atoms.

C(14) : Central carbon atom is expected to be positively charged. This is satisfied for all populations. NPA predicts this structure as a carbocation with the carbon charge above positive one.

F(15-17) : Fluorine is the most electronegative atom in the molecule. We expect it to be the lowest and a negative charge. QUAO charges are positive and the variances are slightly higher than zero. QUAO is the only population that fails.

One of the consistent atomic charge expectation failures for this molecule is seen for bromine atom. This atom has a very large size which can get the overlap coefficients higher than expected. Also, these atomic charges are from an intermolecular structure. The existence of highly delocalized electrons should be affecting the charges. General assessment breakdown is presented in Table 3-7. We advise ESP, Mulliken and IAO for compounds containing bromine atom.

Table 3-7. Population recommendation for trifluorobromomethane ID 1460 summary with CM5, QH, ESP, Mulliken, NPA, QUAMBO, QUAO, IAO

	CM5	QH	ESP	Mulliken	NPA	QUAMBO	QUAO	IAO
Br (1)	×	×	✓	✓	×	×	✓	✓
C (14)	✓	✓	✓	✓	✓	✓	✓	✓
F(15-17)	✓	✓	✓	✓	✓	✓	×	✓
Total	×	×	✓	✓	×	×	×	✓

3.5.3. 1,1,2-Trichlorotrifluoroethane

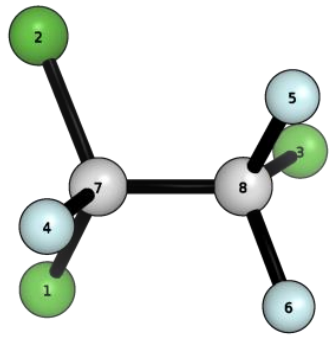

Figure 3-13. Structure of 1,1,2-Trichlorotrifluoroethane (FCl_2CCF_2Cl) ID 193.

We have seen that halogens can be positive from previous example ID 1460. The database is filtered for positive halogen charges several molecules from Mobley dataset stands out. The next example is neutral substituted ethane molecule.

Table 3-8. Atomic charges for 1,1,2-Trichlorotrifluoroethane with B3LYP/Def2-SVPP ID 193. For clarity, atomic charges are conditional formatted on color scale; background darkens as the charge increases

Atom	CM5	QH	ESP	Mulliken	NPA	QUAMBO	QUAO	IAO
Cl(1)	-0.029	-0.031	0.070	-0.036	0.010	-0.017	-0.023	-0.015
Cl(2)	-0.029	-0.031	0.070	-0.036	0.010	-0.017	-0.023	-0.013
Cl(3)	-0.035	-0.039	0.017	-0.066	-0.023	-0.037	-0.043	-0.031
F(4)	-0.117	-0.093	0.055	-0.148	-0.335	-0.314	-0.313	-0.241
F(5)	-0.118	-0.095	-0.021	-0.146	-0.347	-0.327	-0.325	-0.248
F(6)	-0.118	-0.095	-0.019	-0.146	-0.347	-0.327	-0.325	-0.247
C(7)	0.184	0.162	-0.240	0.174	0.292	0.331	0.343	0.260
C(8)	0.262	0.223	0.069	0.404	0.741	0.710	0.709	0.535

Cl(1-3) : Chlorine should be negative in sign and lower than carbon, higher than the fluorine charges. ESP and NPA charges are positive unlike expected. For other populations all conditions are satisfied.

F(4-6) : Fluorine is the most electronegative element in this molecule. It is expected to be negative in sign and the lowest out of all the other atomic charges. ESP charge for fluorine index 4 is positive. The other populations are all negative. QH charges are too neutral.

C(7,8) : Carbons should be positive and the highest atomic charges in the molecule. Only ESP for carbon index 7 negative, the other populations give positive charges. Carbon 8 has a higher number of fluorine substitutions so it should be more positive than carbon index 7. There is not much difference between the two carbons for CM5 and QH so they are not recommended.

Overall, Hirshfeld flavors satisfy the right trend but the charges are overall too neutral. These populations are not advised for halogen atoms over QAO or Mulliken. QAO is an improvement over Mulliken because the fluorine charges are more negative.

Table 3-9. Population recommendation for 1,1,2-Trichlorotrifluoroethane ID 1460 summary with CM5, QH, ESP, Mulliken, NPA, QUAMBO, QUAO, IAO

	CM5	QH	ESP	Mulliken	NPA	QUAMBO	QUAO	IAO
Cl (1-3)	✓	✓	×	✓	×	✓	✓	✓
F (4-6)	✓	×	×	✓	✓	✓	✓	✓
C (7,8)	×	×	×	✓	✓	✓	✓	✓
Total	×	×	×	✓	×	✓	✓	✓

3.5.4. Halotane

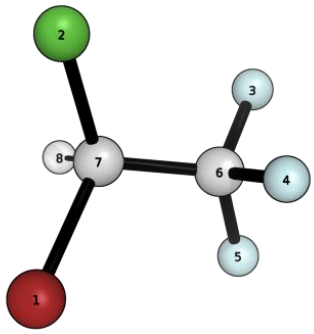

Figure 3-14. Structure of 2-bromo-2-chloro-1,1,1-trifluoroethane ($BrClHCCF_3$) ID 411.

The structure of 2-bromo-2-chloro-1,1,1-trifluoroethane is another molecule from Mobley dataset and the structure is given in Figure 3-14. In this example ethane is substituted with fluorines on one carbon, and one bromine and one chlorine on the other.

Table 3-10. Atomic charges for $BrClHCCF_3$ with B3LYP/Def2-SVPP ID 411. For clarity, atomic charges are conditional formatted on color scale; background darkens as the charge increases

Atom	CM5	QH	ESP	Mulliken	NPA	QUAMBO	QUAO	IAO
Br(1)	-0.007	-0.021	0.108	-0.010	0.071	0.013	0.003	-0.089
Cl(2)	-0.040	-0.037	0.110	-0.055	0.003	0.143	0.154	-0.019
F(3)	-0.128	-0.106	-0.124	-0.160	-0.367	-0.259	-0.301	-0.266
F(4)	-0.124	-0.102	-0.105	-0.150	-0.363	-0.082	-0.113	-0.260
F(5)	-0.129	-0.106	-0.135	-0.160	-0.367	-0.001	-0.082	-0.267
C(6)	0.338	0.286	0.532	0.555	1.151	0.534	0.428	0.804
C(7)	-0.027	0.028	-0.787	-0.199	-0.392	-0.519	-0.290	-0.086
H(8)	0.116	0.057	0.400	0.179	0.264	0.169	0.201	0.183

Br(1) : Bromine is slightly more electronegative than carbon, so we expect the bromine charge to be lower than carbon and hydrogen Preferably, expected sign should be slightly negative. ESP, NPA, QUAMBO and QUAO charges are all positive.

Cl(2) : Chlorine charges is expected to be definitely negative and lower than bromine charges. QH and IAO fail to distinguish between bromine and chlorine atoms. Populations with positive bromines also have positive chlorines. This might suggest that the reason behind these charges can be the same.

F(3-5) : As the most electronegative element in the molecule, is expected to be the lowest and negative charge. All populations satisfy these two criteria, but QUAMBO and QUAO populations are not recommended because of a significant charge variance for the chemically equivalent fluorine atoms.

C(6,7) : We expect the carbon atoms attached to halogens to be positively charged. This condition is satisfied with all populations for index 6 carbon. NPA charge is unreasonably positive, above positive one. We expect the carbon 7 to have lower charge than the one with three fluorine atoms attached. With ESP, NPA, QUAMBO and QUAO charges are too negative for carbon 7.

Some populations have difficulty differentiating between halogens. CM5 and Mulliken seems to be the most sensitive to halogen differences.

	CM5	QH	ESP	Mulliken	NPA	QUAMBO	QUAO	IAO
Br (1)	✓	✓	×	✓	×	×	×	✓
Cl (2)	✓	×	×	✓	×	×	×	×
F (3-5)	✓	✓	✓	✓	✓	×	×	✓
C (6,7)	✓	✓	×	✓	×	×	×	✓
Total	✓	×	×	✓	×	×	×	×

3.5.5. Sodium cluster

Sodium charges that were omitted in most figures are from 13 sodium atoms arranged in a icosahedron cluster. This molecule is one of the "extreme" cases in the database to test the populations for Coulomb interaction dominant structures. [55] The geometry optimized with Def2-SVP in the neutral charges with a symmetry of I_h . After the optimization is confirmed as a stable minimum, a series of single point calculations have been carried out with -4 to +4 total charges.

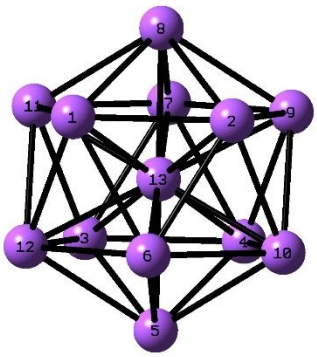

Figure 3-15. Na_{13} icosahedron structure molecule ID 46

Molecular structure given in Figure 3-15 symmetry of the molecule charges can be classified as outer and central sodium atom charges. Central sodium charges are presented in Table 3-11 for different total molecular charge. the first noticeable trend is that QAO flavors are all close to positive six with ESP with -4 molecular charge.

Table 3-11. Atomic charges for central sodium atom (index 13) for range of total molecular charge -4 to +4 with common (CM5, QH, ESP, Mulliken, MBS, NPA) and QAO (QUAMBO, QUAO, IAO) populations B3LYP/Def2-SVPP ID 46. For clarity, atomic charges are conditional formatted on color scale; background darkens as the charge increases

Total Charge	CM5	QH	ESP	Mulliken	MBS	NPA	QUAMBO	QUAO	IAO
-4	0.018	0.018	5.917	-4.631	-3.058	0.000	5.721	5.728	5.732
-3	0.025	0.025	0.483	-4.935	-2.834	0.019	5.650	5.662	5.666
-2	0.024	0.024	-3.297	-5.370	-2.612	-0.007	5.662	5.678	5.682
-1	0.015	0.015	-6.217	-5.946	-2.383	-0.006	5.612	5.639	5.644
0	-0.008	-0.008	-7.812	-6.576	-2.117	-0.097	5.588	5.627	5.634
1	-0.035	-0.035	-7.869	-6.788	-1.789	-0.186	5.509	5.568	5.576
2	-0.067	-0.067	-8.867	-6.784	-1.424	-0.336	5.456	5.541	5.551
3	-0.103	-0.103	-6.796	-6.700	-1.006	-0.463	5.363	5.453	5.465
4	-0.148	-0.148	-7.861	-6.639	-0.622	-0.663	5.293	5.394	5.410

The solid crystal cluster of thirteen sodium atom charges have been explored in total molecular charge range of negative to positive four. Our expectation for sodium charges is to be stable and reasonably distributed. For all population analysis in question as the total charge increases the charge on central sodium atom decreases in charge for most methods, but the rate of decrease is different for some methods. This a clear example of ESP charges failing with buried atoms. Mulliken and ESP charges have the highest spread and most unintuitive atomic charges; charges around negative seven is a clear indication that these population methods are not stable. MBS performs better than Mulliken especially in higher total molecular charges. Still, both methods yield negative charges for a sodium atom.

Sodium cluster breaks QAO methods and the charges are the same with ZBD-QAO. It is also possible that there is a mistake in the code. If it is a constant error the charges can be low and robust.

	CM5	QH	ESP	Mulliken	MBS	NPA	QUAMBO	QUAO	IAO
Na13	✓	✓	×	×	×	✓	×	×	×

3.6 Correlation

To gain a broader understanding of how the atomic charges from each dataset resemble each other, this correlation is presented in Figure 3-16. Each value from the figure represents the coefficient of determination (R^2) that was calculated by comparing two different population methods for approximately 1 million data points. From the heatmap, the most correlated methods ($R^2 \approx 1$) are QUAMBO and QUAO, whereas the least correlated methods ($R^2 \approx 0.9$) are IAO and CM5. Natural Population Analysis (NPA) results resemble those from Quasi-based populations (QUAMBO, QUAO, IAO). Mathematically speaking, the IAO method is derived differently from QUAMBO and QUAO, which may account for the lower correlation.

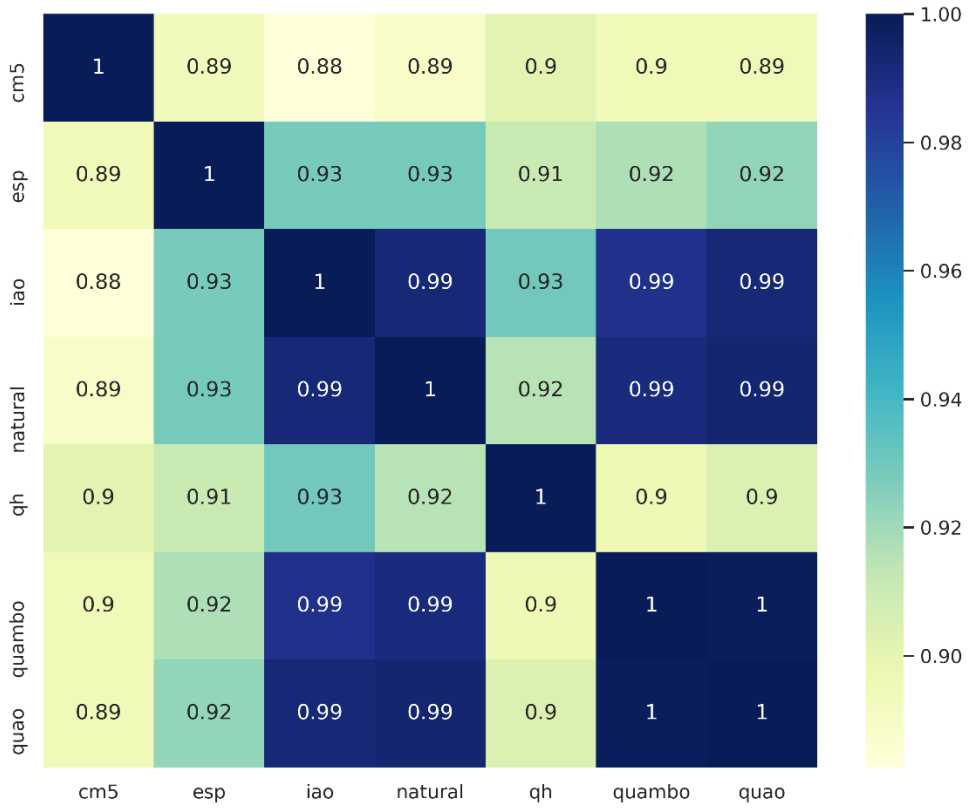

Figure 3-16. Heatmap of common (CM5, QH, ESP, Mulliken, NPA) and QAO (QUAMBO, QUAO, IAO) populations.

In the next we will be looking at ZBD orthogonalization. The correlation of QAO and ZBD-QAO populations are given in Figure 3-17.

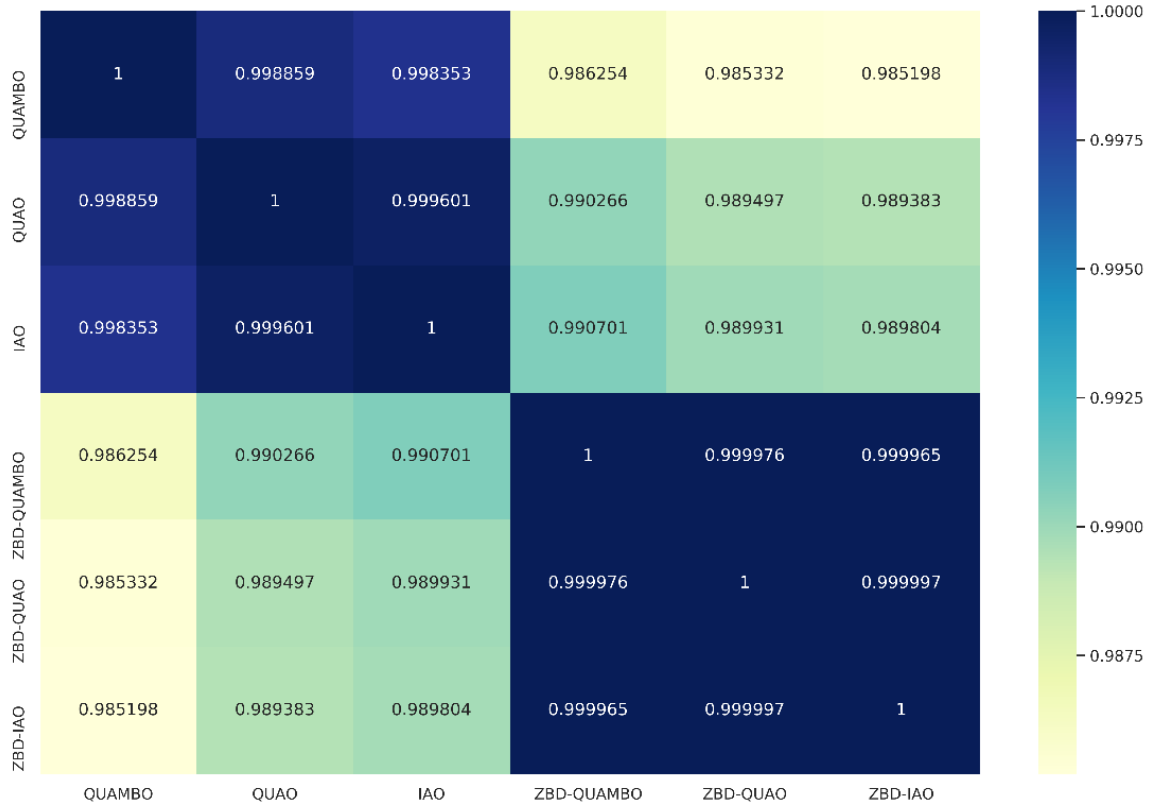

Figure 3-17. Heatmap of common QAO (QUAMBO, QUAO, IAO) and ZBD-QAO (ZBD-QUAMBO, ZBD-QUAO, ZBD-IAO) populations.

Out of all the QAO populations IAO is resembles the ZBD-QAO the most. QUAMBO is the least correlated method to ZBD-QAO (specifically ZBD-IAO). Still, compared to the common populations in Figure 3-16 the scale of Figure 3-17 is significantly smaller.

In the next section we will be identifying small differences between QAO and ZBD-QAO after comparing ZBD-QAO numbers from the reference. This comparison will also extend to effect on chemical trends.

Chapter 4

4 Zero-Bond Dipole effect on Chemical Trends

In this section atomic charges form different molecules will be compared using the population analysis after testing the Zero-Bond Dipole (ZBD) orthogonalization scheme. So far, we have compared population analysis numerically for stability, testing the similarities and outlier atomic charges. The focus of the previous sections has been solely on single molecule assessment. To have more rounded understating of the performance of studied methods, especially from a chemists' perspective, we will be comparing trends from multiple molecules. This is assessment will undoubtedly biased due to the limited number of molecules included in the database, which can be improved upon with additional data, but nevertheless essential with consideration of chemically well understood systems.

4.1 Zero-Bond Dipole Assessment

The dipole sensitive new orthogonalization ZBD will be applied to three different QAO flavors: QUAMBO, QUAO, and IAO (Q1, Q2, Q3). The total number of QAO based population is six with the addition of ZBD-QUAMBO, ZBD-QUAO, and ZBD-IAO (Z-Q1, Z-Q2, Z-Q3). In the reference Laikov[37] utilizes a different QAO procedure which is closer to IAO. This might mean that the atomic charges should not vary greatly compared to IAO population. Each molecule in Table 4-1 has two rows; the first row is with the QAO populations and the second row is the ZBD orthogonalized version. Orb column is the index to QAO and ZBD that represent this indexing. The last column that is the reference values from Laikov's paper.[37]

Overall when we expect that the atomic charges from Q1, Q2, and Q3 to be comparable to Ref column. The most notable difference is for with higher atomic charges on right side of Table 4-1. There reference values are slightly higher for both the reference QAO and ZBD versions. This shows that reference QAO behaves differently for positive charges (at least). ZBD and QAO values should be assessed for higher atomic charged systems. The highest charges have been seen with QAO populations for xenon atoms in Figure 3-6.

Table 4-1. Comparison of bolded atomic charges with non-orthogonal quasi-atomic populations QAO and Zero-bond orthogonalized versions (ZBD) to the reference values (Ref). For clarity, atomic charges are conditional formatted on color scale; background darkens as the charge increases.

Mol	Orb	Q1	Q2	Q3	Ref	Mol	Orb	Q1	Q2	Q3	Ref
OH_2	QAO	-0.94	-0.91	-0.9	-0.93	**Be**H_2	QAO	0.82	0.82	0.84	0.29
	ZBD	-0.59	-0.58	-0.58	-0.42		ZBD	0.84	0.84	0.87	0.47
O_2 H_2	QAO	-0.5	-0.48	-0.48	-0.44	**Al**H_3	QAO	0.29	0.29	0.29	0.74
	ZBD	-0.32	-0.31	-0.31	-0.22		ZBD	0.26	0.26	0.26	1.06
SiH_4	QAO	0.64	0.59	0.59	0.58	**Mg**H_2	QAO	0.82	0.82	0.84	0.55
	ZBD	0.91	0.89	0.89	1.1		ZBD	0.84	0.84	0.86	0.72
NH_3	QAO	0.64	0.59	0.59	0.58	**P**H_3	QAO	-0.03	-0.02	-0.01	-0.19
	ZBD	0.91	0.89	0.89	1.1		ZBD	0.3	0.31	0.31	0.59
C**O**	QAO	-0.48	-0.41	-0.41	-0.36	**P**F_6-	QAO	1.63	1.63	1.63	2.87
	ZBD	-0.22	-0.2	-0.2	-0.22		ZBD	1.51	1.51	1.51	2.7
C**O_2**	QAO	-0.52	-0.5	-0.5	-0.52	**As**F_6-	QAO	1.19	1.74	1.74	2.78
	ZBD	-0.28	-0.27	-0.27	-0.31		ZBD	1.44	1.46	1.48	2.57
H_2**C**CH_2	QAO	-0.49	-0.43	-0.43	-0.27	**As**F_5	QAO	1.83	2.43	2.35	2.7
	ZBD	-0.17	-0.16	-0.15	-0.01		ZBD	1.85	1.88	1.93	2.53
H_3**C**CH_3	QAO	-0.15	-0.15	-0.15	-0.29	**S**O_2	QAO	1.59	1.56	1.57	1.57
	ZBD	-0.12	-0.12	-0.12	0.05		ZBD	1.44	1.43	1.44	1.67
CH_4	QAO	-0.89	-0.8	-0.79	-0.54	**P**F_5	QAO	2.31	2.31	2.31	2.83
	ZBD	-0.38	-0.35	-0.35	0.1		ZBD	1.99	1.99	1.99	2.71
NaH	QAO	0.5	0.5	0.49	0.38	**Cl**FO_3	QAO	2.41	2.44	2.46	2.83
	ZBD	0.51	0.51	0.51	0.47		ZBD	2.59	2.61	2.62	2.98
BH_3	QAO	0.01	0.01	0.01	0.01	**Br**FO_3	QAO	2.82	2.56	2.8	2.79
	ZBD	0.02	0.02	0.02	0.41		ZBD	2.45	2.44	2.57	2.92
LiH	QAO	0.47	0.47	0.47	0.31						
	ZBD	0.48	0.48	0.49	0.43						

4.2 Hydrocarbons

Hydrocarbons are among the most well know and studied chemical structures.[56] The first trend presented Table 4-2 is for assessment of chain length effect with alkane structures.

Table 4-2. Bolded hydrogen atomic charges for hydrocarbons with B3LYP/Def2-SVPP. ID 221, 97, 209, 746, 572. Q1, Q2, Q3 are QUAMBO, QUAO, IAO and Zero-bond orthogonalized versions are prefixed with letter "Z". For clarity, atomic charges are conditional formatted on color scale; background darkens as the charge increases.

Molecule	CM5	QH	ESP	M	NPA	Q1	Q2	Q3	Z-Q1	Z-Q2	Z-Q3
H$_3$C–CH$_3$	0.075	0.026	-0.004	0.093	0.199	0.213	0.190	0.189	0.082	0.075	0.074
C**H**$_3$–CH$_2$–CH$_3$	0.074	0.024	0.052	0.087	0.198	0.208	0.184	0.183	0.077	0.070	0.069
CH$_3$–(CH$_2$)$_2$–CH$_3$	0.076	0.026	0.067	0.091	0.207	0.215	0.191	0.190	0.082	0.076	0.075
CH$_3$–(CH$_2$)$_3$–CH$_3$	0.075	0.025	0.039	0.086	0.199	0.208	0.185	0.184	0.077	0.070	0.069
CH$_3$–(CH$_2$)$_4$–CH$_3$	0.076	0.026	0.053	0.090	0.207	0.215	0.191	0.189	0.082	0.076	0.074

As the hydrocarbon chain grows down the Table 4-2 we expect the atomic charge on the hydrogen of the terminal carbon to be (a) impervious to drastic fluctuations (b) relatively neutral. Population analysis that fail to satisfy these conditions NPA and non-orthogonalized QAO flavors since these both produce higher than expected overall hydrogen charges. With ZBD orthogonalization atomic charges become more neutral, instead of NPA they resemble Mullliken. CM5 and QH hydrogen charges for different hydrocarbons are very neutral and constant with increasing number of carbons. The highest difference between two hydrogen charges are seen with ESP from methane to ethane. Considering that the carbon atom is more electronegative than the hydrogen, methane hydrogen with ESP being slightly negative is acceptable. When more carbons are added to the alkane chain, we see a slight fluctuation for atomic charges.

Table 4-3. Bolded hydrogen atomic charges for alkanes and alkenes with B3LYP/Def2-SVPP. ID Q1, Q2, Q3 are QUAMBO, QUAO, IAO and Zero-bond orthogonalized versions are prefixed with letter "Z". For clarity, atomic charges are conditional formatted on color scale; background darkens as the charge increases.

Molecule	CM5	QH	ESP	M	NPA	Q1	Q2	Q3	Z-Q1	Z-Q2	Z-Q3
H$_3$C–CH$_3$	0.075	0.026	-0.003	0.093	0.199	0.213	0.190	0.119	0.082	0.075	0.074
H$_2$C=CH$_2$	0.088	0.037	0.139	0.103	0.190	0.236	0.208	0.131	0.082	0.073	0.072
C**H**$_3$–CH$_2$–CH$_3$	0.074	0.024	0.052	0.087	0.198	0.208	0.184	0.183	0.077	0.070	0.069
C**H**$_2$=CH–CH$_3$	0.085	0.033	0.160	0.093	0.195	0.233	0.205	0.204	0.078	0.070	0.068
CH$_3$–CH$_2$–CH$_2$–CH$_3$	0.076	0.026	0.067	0.091	0.207	0.215	0.191	0.190	0.082	0.076	0.075
C**H**$_2$=CH–CH$_2$–CH$_3$	0.083	0.031	0.155	0.088	0.187	0.222	0.196	0.123	0.075	0.067	0.066
C**H**$_2$=CH–CH=CH$_2$	0.089	0.038	0.152	0.101	0.198	0.242	0.213	0.212	0.089	0.080	0.079

To assess the atomic charge on hydrogen of the terminal vinylic carbon, structures from the previous table are presented followed by their double bonded versions. In the presence of a double bond we expect the hydrogen to be more positive. For ESP the atomic charge difference between the alkane and the akene is ~0.1 greatest, whereas for Hirshfeld populations CM5 and QH produce the difference is ~0.01 almost one tenth of ESP. While Mulliken, NPA and non-orthogonal QAO flavors have high overall charges for all hydrogens ZBD-QAO have Hirshfeld-like low charges. Propane and propene are slightly more positive than butane and butene trend is only present with IAO. CM5 is fitted for these neutral molecules and there is a clear pattern in stability of the atomic charges as the alkane compared to an alkene. This pattern is not seen in any of the orbital populations in the scope (Mulliken, NPA, QAO, ZBD-QAO). The last molecule in Table 4-3 is for comparison for growing the hydrocarbon chair, instead is to assess the effect of charge distribution; the butadiene should have resonance with a slightly lower hydrogen charge. This trend is not observed with any charge except slight with ESP population.

Table 4-4. Bolded hydrogen atomic charges ketones and aldehyde with B3LYP/Def2-SVPP. ID 681, 201, 200. Q1, Q2, Q3 are QUAMBO, QUAO, IAO and Zero-bond orthogonalized versions are prefixed with letter "Z". For clarity, atomic charges are conditional formatted on color scale; background darkens as the charge increases.

Molecule	CM5	QH	ESP	M	NPA	Q1	Q2	Q3	Z-Q1	Z-Q2	Z-Q3
C**H**$_3$–(CH2)$_3$–HCO	0.077	0.027	0.068	0.093	0.208	0.217	0.193	0.191	0.086	0.080	0.078
C**H**$_3$–(CH2)$_2$–CO–CH$_3$	0.079	0.029	0.060	0.096	0.210	0.221	0.196	0.195	0.090	0.084	0.083
C**H**$_3$–CH$_2$–CO–CH$_2$–CH$_3$	0.077	0.027	0.048	0.089	0.207	0.214	0.192	0.190	0.085	0.080	0.078
C**H**$_3$–CO–(CH2)$_2$–CH$_3$	0.094	0.040	0.122	0.112	0.227	0.254	0.225	0.224	0.112	0.104	0.103
HCO–(CH2)$_3$–CH$_3$	0.085	0.026	-0.054	0.057	0.114	0.200	0.170	0.169	0.044	0.036	0.034

Effect of carbonyl distance to the hydrogen is presented in Table 4-4 with fixed hydrocarbon chain length. Starting with the furthest methyl hydrogen, carbonyl gets closer the methyl group until the hydrogen in question is from the aldehyde. As expected, hydrogen atomic charge is almost the

same as the ethane charge when the carbonyl group is the furthest. CM5 and QH charges are neutral and unchanging, there is no difference the hydrogens of methyl and aldehyde. We expect the hydrogen of alpha carbon (row four) to be the most positive, but this trend is barely noticeable with Hirshfeld populations. ESP charges are stable for the first three rows and highest for the alpha carbon. The only negative charge for hydrogens in this table is seen with ESP for the aldehyde hydrogen. For the same molecule, the charge separation of carbonyl group is the greatest (difference between atomic charge of oxygen and carbon is close to one) with ESP and NPA. While the negative charge is localized to electronegative atom oxygen with NPA, this charge is delocalized to the hydrogen with ESP, causing it to be negative. Like the previous tables, QAO methods resemble NPA charges while the ZBD-QAO resemble regular Mulliken.

	CM5	QH	ESP	Mulliken	NPA	Q1	Q2	Q3	Z-Q1	Z-Q2	Z-Q3
Hydrocarbons	✓	✓	×	✓	✓	×	×	×	✓	✓	✓

4.3 Halogens

In this section we will be looking at small molecules with halogen bonding. Halogen compounds are regularly used in supramolecular and solvent chemistry and have been a challenge to describe with quantum mechanics.[57,58] Most populations can fail to presented expected trends because of electronegativity difference, large multipole moments and “fuzzy” atoms. The latter is especially challenging for orbital based populations studied in this thesis.

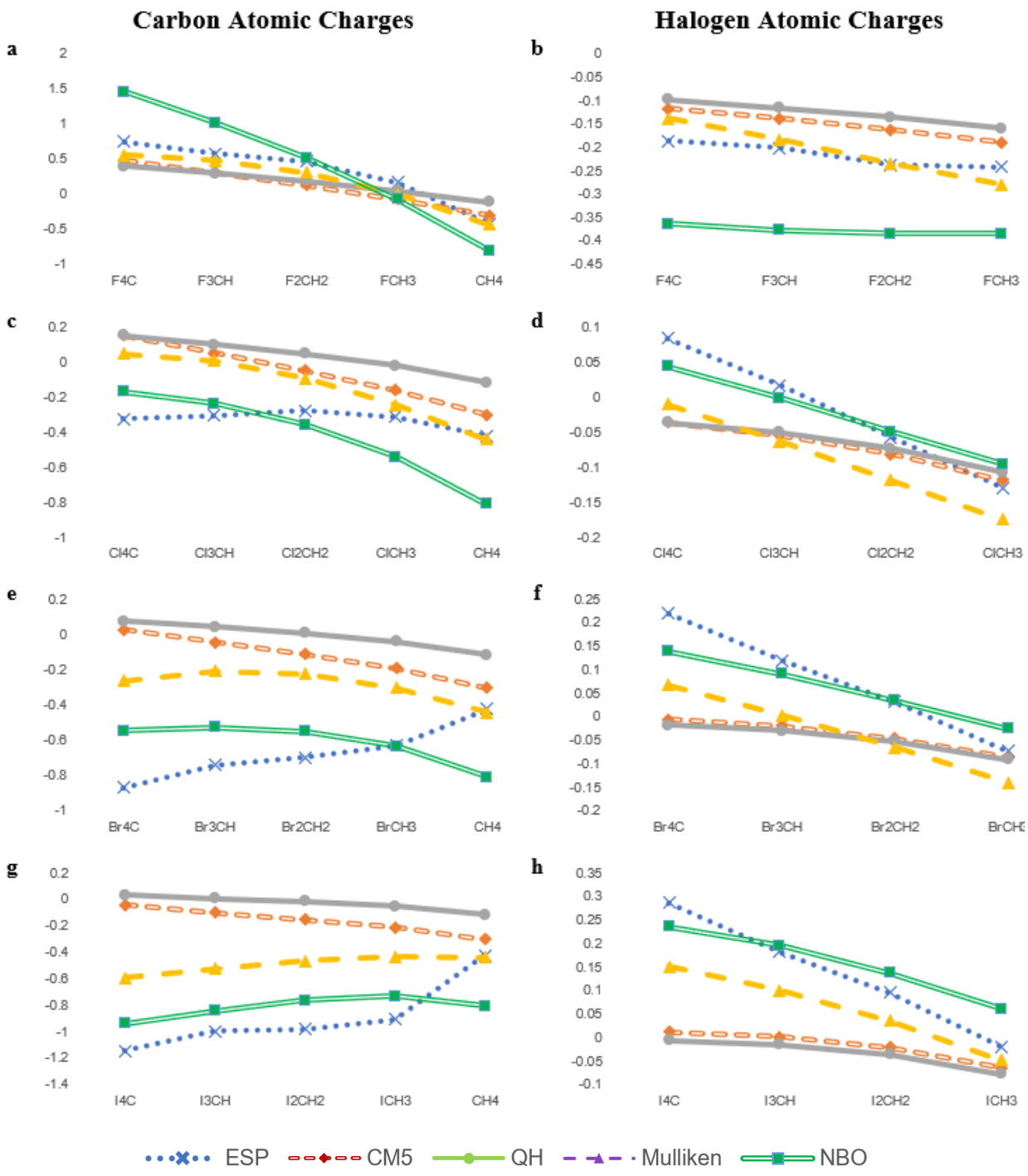

Figure 4-1 Carbon and halogen atomic charges (y axis) plotted against chemical formula (x axis) $X_nCH_{(4-n)}$ where n=[0,4] X=(F,Cl,Br) with common populations (CM5, QH, ESP, Mulliken, NPA).

The first trend we will be looking at is simple methane substitution case; for halogen atoms F, Cl, Br, I we will tests for (1) number of halogen substitution (2) different halogen effects. Both carbon and halogen charges shall be subject to assessment. To be able to see clear trends atomic charges results are presented in graphs Figure 3-1 where charges are split into two columns by atom type; on the left column carbon and on the right is the halogen atomic charges. Our first expectation is that a carbon atom substituted with four halogen atoms should be positively charged. However, carbon atomic charges on the left column (Figure 3-1 **a**, **c**, **e**, **g**) are not positive like we expected. Carbons for CCl_4 (ESP, NPA), CBr_4 (ESP, NPA, Mulliken) and CI_4 (ESP, NPA, Mulliken, CM5) are all negative. QH and CM5 charges for carbon atom are the most positive, but neither halogen number effect nor halogen atom type effect is observed, the carbon atomic charge difference is too small. Another expected trend is that as more halogens are substituted to the central carbon atom, we expect carbon charges to increase. Down the periodic table, ESP is the first population to show the opposite trend (signified by a positive slope) with Cl substitutions (Figure 3-1 **c**); carbon charge for Cl_3CH is smaller than Cl_2CH_2 . Regarding bromine substitutions in Figure 3-1 **b,** positive slope increases with ESP while the expected result is a negative slope. For other populations still produce a positive slope. As we go down the periodic table, we expect atomic charges for halogens to decrease so that I > Br > Cl > F (for same number of halogen substituents). Even though the difference between halogen charges are very small ~0.01 the trend is as expected with CM5 and QH while all the orbital based populations fail. We can summarize that Hirshfeld flavors (QH and CM5) are the most "reliable" population analysis method for halogen substitutions.

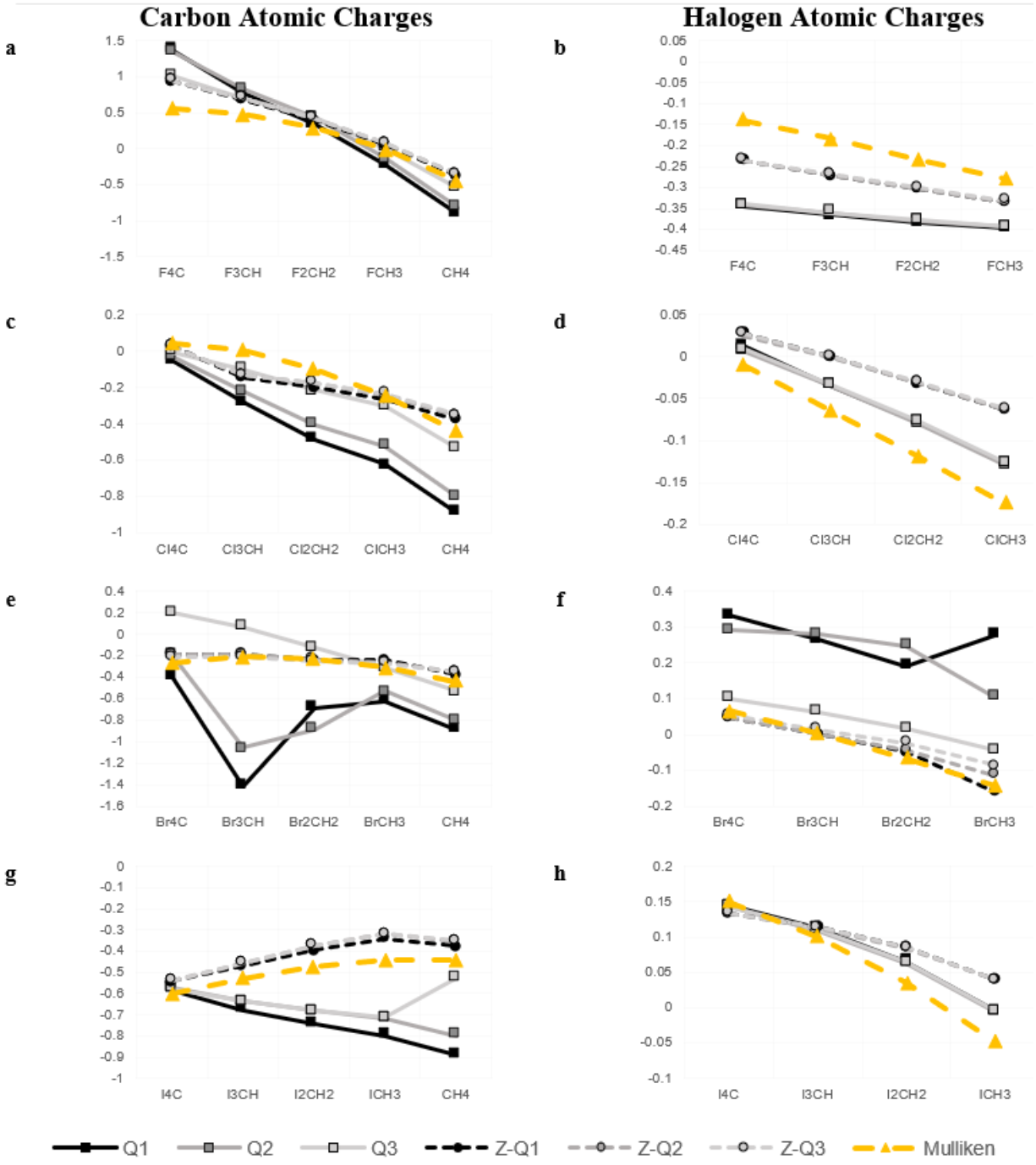

Figure 4-2. Carbon and halogen atomic charges (y axis) ploted for tetra-halogen structures $X_nCH_{(4-n)}$ for n=[0,4] X=(F,Cl,Br). Comparison of Q1, Q2, Q3 are QUAMBO, QUAO, IAO and Zero-bond orthogonalized versions (prefixed with letter "Z") with classical Mulliken population.

The same trend in Figure 3-1 is investigated with QAO and ZBD-QAO populations. We can see that neither QAO or ZBD-QAO fix the expectation of carbon atoms to be positively charged. Figure 4-2**a** fluorine substitution atomic charges for the central carbon is not improved with QUAMBO and QUAO flavors. For these methods, one improvement is that the central carbon of tetrafluoromethane has a higher positive charge. However, methane carbon is even more negative compared to Mulliken. ZBD orthogonalization lowers the tetrafluoromethane carbon atomic charge between Mulliken and QUAMBO/QUAO and increases the atomic charge of methane carbon to Mulliken level. There is a clear difference in the fluorine charges for different populations in Figure 4-2**b**. This time both QAO and ZBD-QAO populations offer better charges than Mulliken. Overall, ZBD fluorine charges are higher than QAO charges while both populations are not as negative as NPA in Figure 3-1**b**. A more significant difference can bee seen with chlorine charges than fluorine. Central carbon for chlorine substitutions is too negative with QUAMBO and QUAO. The curve in Mulliken atomic charges seen with this carbon is fixed with ZBD-QUAMBO. For the chlorine charges both QAO and ZBD-QAO is higher than Mulliken, failing to satisfy the expectation that halogens should be negatively charged. Looking back at the first half of Figure 4-2, QUAMBO-QUAO populations and IAO-ZBD-QAO populations produced similar results. However, QUAMBO and QUAO start behaving unexpectedly for bromine. Carbon atomic charges drop to below negative one for bromoform (Br_3CH) with these methods. IAO is not affected by this localization of charge on carbon instead of the bromines and the only population to yield positive charge for the carbon atom of bromoform. For all bromine substituted molecules ZBD orthogonalization corrects QUAMBO and QUAO charges to Mulliken level. Iodine charges with QAO or ZBD-QAO are not improved compared to Mulliken, the trends are still chemically unintuitive.

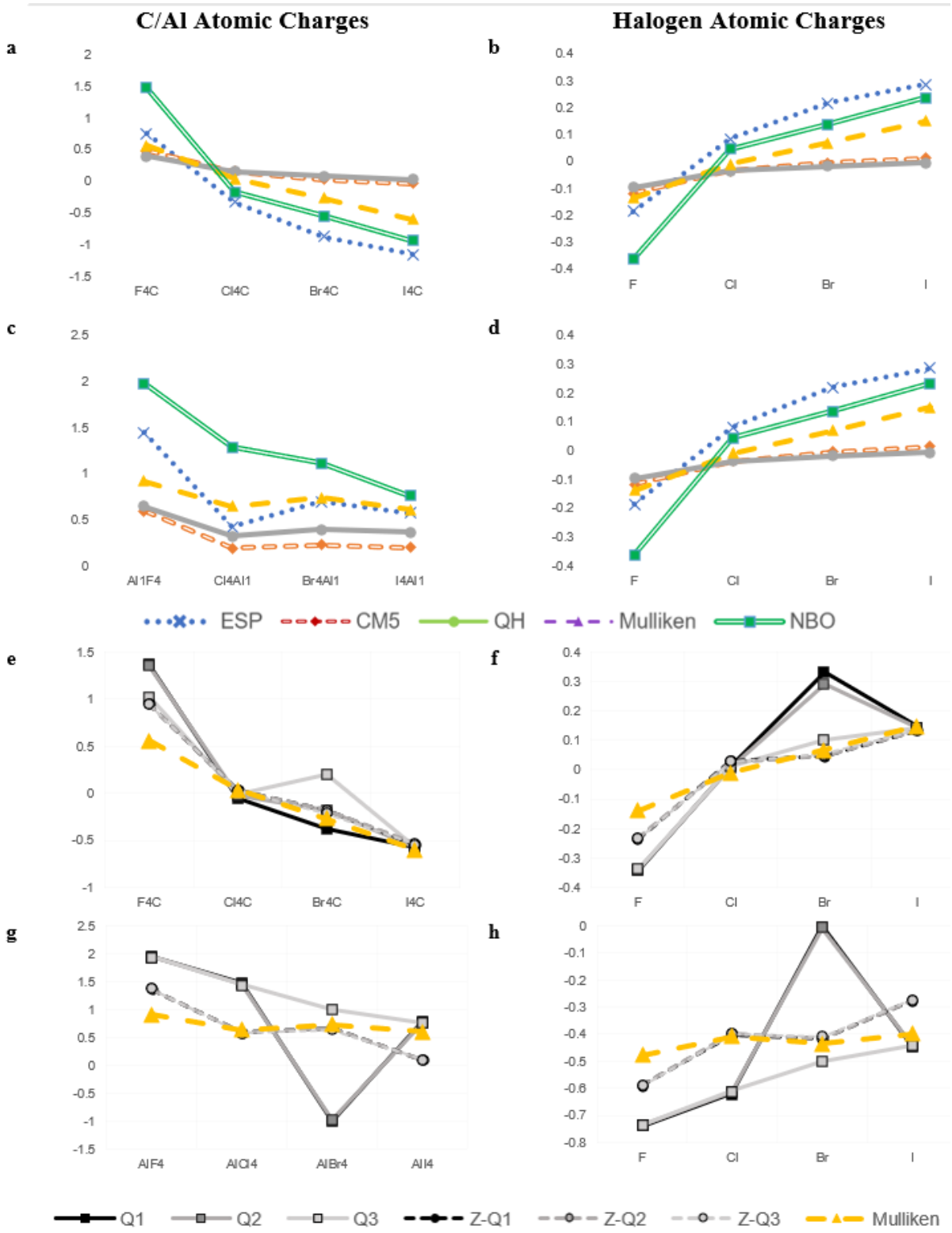

Figure 4-3. Central atom effect on atomic charges (y axis) ploted for tetra-halogen structures AX_4 for A=(C, Al) X=(F, Cl, Br, I). Common population methods (CM5, QH, ESP, Mulliken, NPA) and Q1, Q2, Q3 (QUAMBO, QUAO, IAO) and Zero-bond orthogonalized versions (prefixed with letter "Z").

Switching from the central atom of tetra-halogen structure to aluminum which is a less electronegative atom. Common population methods for carbon center Figure 4-3.**a** give the expected trend (slope is negative) even though we see some negative charges. For aluminum center this trend is broken with ESP because AlCl4 is lower than AlBr4. Figure 4-3.**c** most populations (CM5, QH, Mulliken) have no aluminum charge difference between chlorine or bromine substitutions. QAO and ZBD-QAO populations carbon charges are all in trend except for the bromine charge problem with QUAMBO and QUAO. IAO charges Figure 4-3.**e,g** are very close to NPA charges in Figure 4-3.**a,c**. For bromine error with QUAMBO and QUAO changes sign between carbon and aluminum central atom. This is one of the reasons why QUAMBO and QUAO have unusually high variance. Atomic charge range for the central atom with CM5 and QH type is too low. We expect that if a carbon (or an aluminum) is bonded to four fluorine atoms, the positive charge on the central atom should be greater than 0.5. We see a slight difference in results for aluminum central atom between carbon and aluminum centers between CM5 and QH; CM5 charges are more neutral than QH. There is almost no difference between atomic charges of carbon and aluminum for CM5 and QH either. This is a clear example why these methods have very low variance. On the other hand, NPA population has the highest variance for different halogens. ESP and Mulliken break when the carbon is replaced with the aluminum atom. ZBD-QAO performs slightly better than Mulliken for fluorine and iodine substitutions but offer no differentiation between bromine and chlorine. Overall, IAO seems to perform the best out of all the QAO populations because it can differentiate between bromine and chlorine substitutions.

	CM5	QH	ESP	Mulliken	NPA	Q1	Q2	Q3	Z-Q1	Z-Q2	Z-Q3
Halogen trends	×	×	×	×	×	×	×	✓	✓	✓	✓

4.4 Hydrogen bond

After the discovery that the DNA bases stabilized by intermolecular hydrogen bonds[59,60] (H-bond), it has been excepted as one of the most important non-covalent interactions. Studied extensively for its applications in bioorganic chemistry, [61,62] a H-bond can be shown as $X\text{-}H\cdots Y$ where X is the proton donor and Y is the acceptor.[63] Most commonly Y is an electronegative atom such as N, O or F but it has been shown that electron rich regions such as delocalized π system could also participate in H-bonding. The strength of the H-bond is dependent on the molecules, distance, angle and it can b a weak or strong as a covalent bond. Hydrogen bonds are very complicated and have contributions not just from electrostatic but also polarization, van der Waals and even charge transfer.[64] In the framework of atomic charges, in the presence of H-bond a charge transfer is expected from proton acceptor to donors sigma antibonding orbital. This causes the proton donator to have elongated X-H bond.[65] There different types of hydrogen bonding where this is not the case (blue shifting or anti H-bond).[66]

Total anion charges for anion-neutral intermolecular hydrogen bonded interactions are presented in Table 4-5. All molecules in this table are from dataset AHB21, a dataset specifically designed to test quantum chemical methods for binding energies of H-bonds for structures can be found in appendix B.[67] The reference found the strongest H-bonds in this dataset are seen for F^-(HF) > Cl^-(HCl) > OH^-(H2O) interactions. Within the same anionic species, these molecules also produce the highest atomic charges

Table 4-5. Bolded anion charges H-bonded to a neutral molecule with B3LYP/Def2-SVPP. ID 818-834. Q1, Q2, Q3 are QUAMBO, QUAO, IAO and Zero-bond orthogonalized versions are prefixed with letter "Z". For clarity, atomic charges are conditional formatted on color scale; background darkens as the charge increases.

interaction	CM5	QH	ESP	M	NPA	Q1	Q2	Q3	Z-Q1	Z-Q2	Z-Q3
F$^-$ . . . H_3N	-0.76	-0.75	-0.87	-0.78	-0.86	-0.83	-0.84	-0.84	-0.83	-0.83	-0.83
F$^-$ . . . H_2O	-0.69	-0.67	-0.82	-0.76	-0.83	-0.79	-0.79	-0.79	-0.76	-0.76	-0.76
F$^-$ . . . HF	-0.59	-0.55	-0.73	-0.69	-0.76	-0.74	-0.74	-0.74	-0.65	-0.65	-0.65
Cl$^-$ . . . H_3N	-0.84	-0.84	-0.90	-0.84	-0.92	-0.91	-0.91	-0.91	-0.90	-0.91	-0.90
Cl$^-$ . . . H_2O	-0.80	-0.79	-0.86	-0.82	-0.89	-0.87	-0.87	-0.87	-0.85	-0.85	-0.85
Cl$^-$ . . . HF	-0.74	-0.73	-0.82	-0.79	-0.84	-0.82	-0.82	-0.82	-0.78	-0.78	-0.78
Cl$^-$ . . . H_2S	-0.75	-0.74	-0.82	-0.75	-0.81	-0.80	-0.80	-0.80	-0.80	-0.80	-0.80
Cl$^-$ . . . HCl	-0.55	-0.52	-0.65	-0.62	-0.62	-0.60	-0.60	-0.60	-0.53	-0.53	-0.53
HO$^-$ . . . H_3N	-0.76	-0.74	-0.88	-0.75	-0.87	-0.81	-0.82	-0.82	-0.80	-0.80	-0.80
HO$^-$ . . . H_2O	-0.67	-0.61	-0.81	-0.69	-0.79	-0.73	-0.73	-0.72	-0.67	-0.67	-0.67
N_3 $^-$ . . . H_3N	-0.90	-0.88	-0.98	-0.90	-0.97	-0.94	-0.95	-0.95	-0.93	-0.93	-0.93
N_3 $^-$ . . . H_2O	-0.86	-0.83	-0.94	-0.88	-0.95	-0.91	-0.91	-0.91	-0.89	-0.89	-0.89
N_3 $^-$ . . . HF	-0.79	-0.73	-0.86	-0.83	-0.88	-0.84	-0.84	-0.84	-0.78	-0.78	-0.78
N_3 $^-$ . . . H_2S	-0.85	-0.82	-0.91	-0.85	-0.92	-0.90	-0.90	-0.90	-0.89	-0.89	-0.89
HS$^-$ . . . H_3N	-0.83	-0.82	-0.89	-0.84	-0.92	-0.90	-0.90	-0.90	-0.89	-0.89	-0.89
HS$^-$ . . . H_2O	-0.77	-0.76	-0.84	-0.81	-0.88	-0.86	-0.86	-0.86	-0.83	-0.83	-0.83
HS$^-$ . . . HF	-0.71	-0.69	-0.81	-0.78	-0.82	-0.80	-0.80	-0.80	-0.75	-0.75	-0.75

From charges in Table 4-5 we expect reasonable numbers for both anion and neutral molecule. All anions should be negatively charged; the neutral molecules should be more positive than the anions. While chemical intuition dictates that the total charge on the anion should be as close to -1 as possible, this is not a desirable trend in quantum mechanics. We need unique and non-integer charges especially in the case of two interacting molecules to account for charge transfer.

One of the reasons why H-bond interactions are complicated is also the strength cannot be attributed to a single chemically intuitive expectation. If the electronegativity was the main driving force for the H-bond, as the anion becomes more electronegative the charge should become more

negative. However, the trend for Cl^- and F^- is the opposite, where the former is more negative than the. Instead considering the effect charge transfer in the of Pearson hardness-softness or Lewis acid-base framework.[68] With softer hydrogen donors the negative charge on the anion can be relocate. As the strength of the bond increases the anions gets more positively charged as expected. For N3 molecule the anion charge is the most negative because instead of participating in intermolecular, intramolecular charge transfer is observed. It is not recommended to use ESP population for hydrides because atomic charges of the anion interacting with ammonia and water is the same.

	CM5	QH	ESP	Mulliken	NPA	Q1	Q2	Q3	Z-Q1	Z-Q2	Z-Q3
H-bond	✓	✓	×	✓	✓	✓	✓	✓	✓	✓	✓

Table 4-6. Bolded atomic charges for non-interacting water and cation-water interaction of with B3LYP/Def2-SVPP. ID 839-841&101. Q1, Q2, Q3 are QUAMBO, QUAO, IAO and Zero-bond orthogonalized versions are prefixed with letter "Z". For clarity, atomic charges are conditional formatted on color scale; background darkens as the charge increases.

interaction	CM5	QH	ESP	M	NPA	Q1	Q2	Q3	Z-Q1	Z-Q2	Z-Q3
Li^+... H_2O	0.87	0.77	0.96	0.85	0.98	0.98	0.98	0.98	0.95	0.95	0.95
Na^+... H_2O	0.93	0.83	0.98	0.86	0.98	0.97	0.97	0.97	0.95	0.95	0.95
K^+... H_2O	0.96	0.87	0.97	0.89	0.98	0.99	0.99	0.99	0.98	0.98	0.98
H_2**O**	-0.65	-0.32	-0.77	-0.59	-0.86	-0.88	-0.86	-0.85	-0.57	-0.57	-0.56
Li^+... H_2**O**	-0.65	-0.23	-1.03	-0.64	-1.00	-1.02	-0.99	-0.99	-0.65	-0.64	-0.64
Na^+... H_2**O**	-0.67	-0.25	-0.99	-0.61	-0.97	-0.98	-0.95	-0.95	-0.62	-0.61	-0.61
K^+... H_2**O**	-0.70	-0.27	-0.91	-0.44	-0.94	-0.98	-0.95	-0.95	-0.63	-0.63	-0.63
H_2O	0.32	0.16	0.38	0.30	0.43	0.44	0.43	0.43	0.29	0.28	0.28
Li^+... **H_2**O	0.39	0.23	0.53	0.39	0.51	0.52	0.51	0.50	0.35	0.35	0.35
Na^+... **H_2**O	0.37	0.21	0.51	0.38	0.49	0.50	0.49	0.49	0.33	0.33	0.33
K^+... **H_2**O	0.37	0.20	0.47	0.36	0.48	0.50	0.48	0.48	0.33	0.32	0.32

In the presence of a cation the oxygen should be a bit more positive

M^+ : Going down the periodic table the anion becomes more softer and positive so that $K^+ > Na^+ > Li^+$. For ESP and Mulliken population this trend is not observed.

O : In the presence of a cation interaction, the oxygen atomic charge of water molecule is expected to be slightly more positive than the non-interacting water oxygen. QH is the only population this trend is seen for all three different interactions. For Mulliken only K interaction has a more positive oxygen.

H : Water hydrogen interacting with the cation should be become less positive as the cation becomes a softer Lewis acid. While every population shows this trend, the hydrogen charge of Li^+ interaction is significantly higher than

Cation and hydrogen interactions (alkali-water interactions) from CHB6 dataset can be seen in Table 4-6. These alkali-water interactions CM5 and QH is lower than any other population for

cation M^+ (Li, Na, K) atomic charges. We expect the metal to be close to positive one, but darker background is only seen with ESP, NPA, QAO, ZBD-QAO. There is a clear difference between CM5 and QH, oxygen charges with QH is ~0.3 higher than CM5 charges. For QH population charges for alkali metals are the lowest. Lowest oxygen charges are seen with ESP, NPA, QUAMBO, QUAO and IAO populations. ZBD-QAO oxygen atomic charges are higher than the QAO populations, close to CM5 charges. Muliken population predicts the oxygen charge of the water interacting with potassium slightly more positive than the non-interaction water molecule. For ZBD-QAO populations oxygen charges in the presence of an anion are barely different than non-interacting water case. CM5 the existence of the cation has no effect on the charges, but for ESP, NPA, and QAO oxygen charge decreases. Expected trend is in the presence of a cation, the oxygen of the hydrogen donor should be slightly more positive. So we can summarize the results as:

	CM5	QH	ESP	M	NPA	Q1	Q2	Q3	Z-Q1	Z-Q2	Z-Q3
M^+	✓	✓	×	×	✓	✓	✓	✓	✓	✓	✓
O	×	✓	×	×	×	×	×	×	✓	✓	✓
H	✓	✓	✓	✓	✓	×	×	×	✓	✓	✓
Total	✓	✓	×	✓	×	×	×	×	✓	✓	✓

4.5 Silica

Another chemical trend we can look at is the different silica dimer dataset from MBID paper.[23] This dataset consists of permutations of intramolecular interactions of 12 silica containing molecules with 20 different molecules. In the previous chapter we have seen that silica is one of the atoms that have low atomic charge spread. In reference paper standard deviations of 15 different population methods are provided. There populations CM5, QH, and ESP are among these methods

Table 4-7. Standard deviations (STD) for atomic charges of silica atoms in ZG237 dataset compared with reference STD.[23] ID 1034-1271.

Population	STD	REF STD
CM5	0.053	0.003
QH	0.050	0.004
ESP	0.087	0.092
Mulliken	0.062	-
NPA	0.063	-
QUAMBO	0.075	-
QUAO	0.071	-
IAO	0.070	-
ZBD-QUAMBO	0.053	-
ZBD-QUAO	0.053	-
ZBD-IAO	0.052	-

Starting with the populations also provided in the REF values, standard deviance (STD) is higher for Hirshfeld populations CM5 and QH. There are two reasons for this difference (a) the reference STD values are grouped together for same type of silica atoms, then divided by the mean value of the whole dataset (b) the values include three different methods and three different basis sets. ESP population surprisingly has lower STD compared to the reference values. The highest variance is

seen with QAO populations Q1, Q2, Q3 (QUAMBO, QUAO, IAO). ZBD orthogonalization decreases the STD of QAO populations. From the common populations NPA is the highest STD.

Silica charges from ZG237 dataset has been plotted against the molecule ID (mid) with different populations are given in Figure 4-4. As shown in the legend at the bottom right corner there are six different silica structures. A single data point on the graph is colored by the silica structure non-covalently interacting with another molecule differentiated by a molecule ID. For silica atomic charges two trends are expected:

1. $H_3O_3Si–O–SiO_3H_3 < H_3O_3Si–O–AlO_3H_3 < H_3O_3Si–O–PO_3H$ and

with distinct atomic charge clusters. Only with ESP population clustering of same silica molecules is not seen. While other populations present distinct clusters, Mulliken there is no difference between $H_3O_3Si–O–AlO_3H3$ and $H_3O_3Si–O–PO_3H$

2. $SiO_4H_3 < SiO_4H_3CH_3 < SiO_4H_4$

For CM5, QH and Mulliken there is no difference between SiO_4H_4 and SiO_4H_3.

Overall, QAO improves the clustering of silica charges of Mulliken and ZBD orthogonalization lowers the charges. Mulliken should not be used for silica charges.

	CM5	QH	ESP	Mulliken	NPA	Q1	Q2	Q3	Z-Q1	Z-Q2	Z-Q3
Trend (1)	✓	✓	×	×	✓	×	×	×	✓	✓	✓
Trend (2)	×	×	✓	×	✓	✓	✓	✓	✓	✓	✓
Clustering	✓	✓	×	✓	✓	✓	✓	✓	✓	✓	✓
Total	×	×	×	×	✓	×	×	×	✓	✓	✓

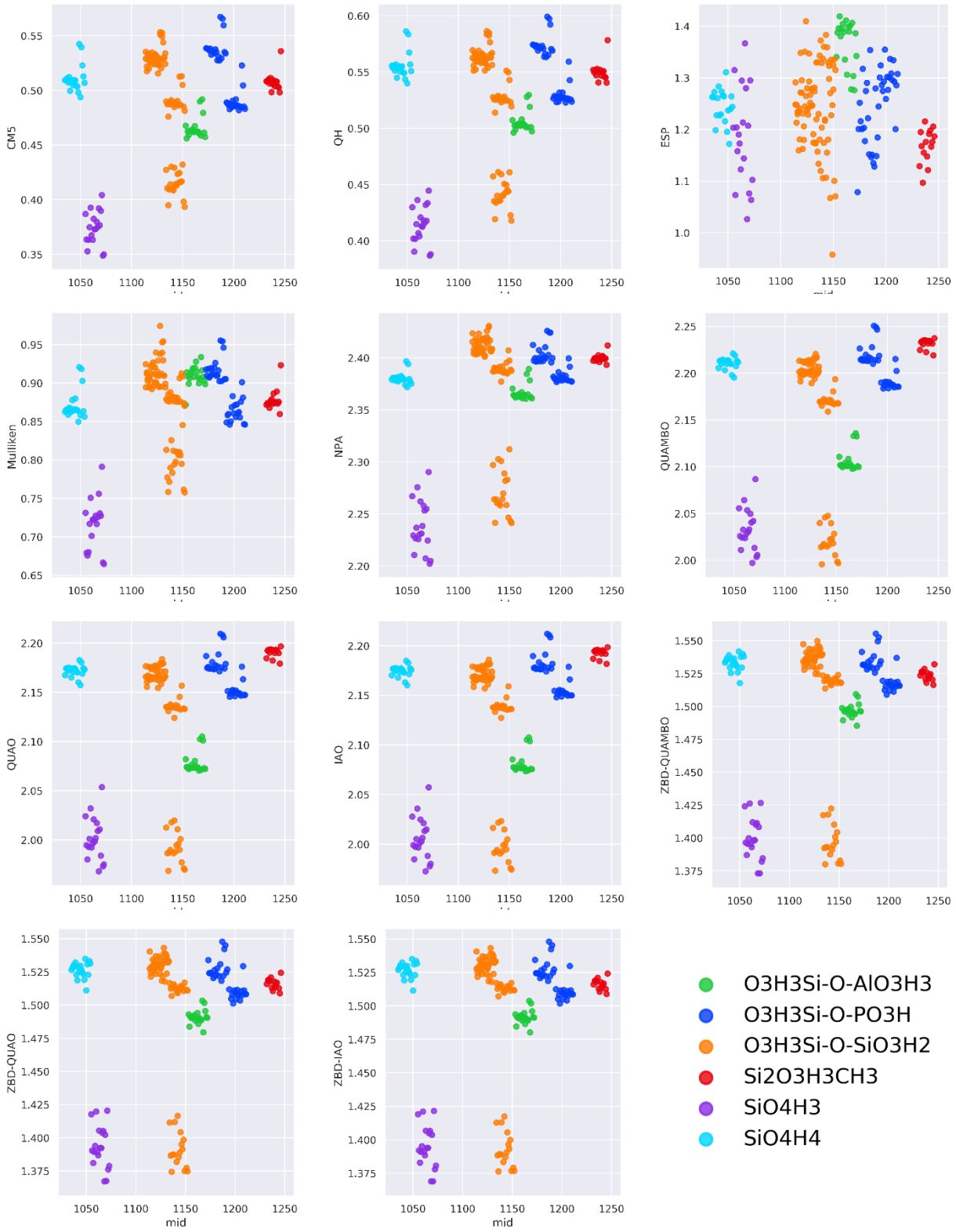

Figure 4-4. Silicon charges (y axis) from ZG237 dataset with molecule ID (x axis) 1035-1271 for populations CM5, QH, ESP, Mulliken, NPA, QAO and ZBD-QAO flavors for B3LYP/Def2-SVPP separated by molecular structure presented in legend.

4.6 Xenon

After looking at silica charges that appeared in very localized in chapter two graphs, we will consider the band like atomic charges seen in NPA and QAO populations. All these atomic charges are from stable structures of the heavy noble gas xenon from G18 dataset. Thes molecules have been studied since the 60's[69–71] and received great chemical interest because they are can break the octet rule (XeF_4 which is planar).[72] Noble gas halides have been considered as ionic in character[73] where halide is negatively charged while the noble gas is positively charged. Xenon atom have been found in many oxidation states ranging from zero to eight. Earlier explanations of the bonding patterns with Molecular Orbital Theory states that xenon bonds are mainly p_σ atomic orbitals.[74]

Table 4-8. Bolded atomic charges for xenon containing compounds with B3LYP/Def2-SVPP. ID 839-841&101. Q1, Q2, Q3 are QUAMBO, QUAO, IAO and Zero-bond orthogonalized versions are prefixed with letter "Z". For clarity, atomic charges are conditional formatted on color scale; background darkens as the charge increases.

Molecule	CM5	QH	ESP	M	NPA	Q1	Q2	Q3	Z-Q1	Z-Q2	Z-Q3
XeF_2	0.62	0.57	0.61	0.75	1.14	1.22	1.22	1.22	1.39	1.39	1.39
XeF_4	1.03	0.93	1.05	1.29	2.17	2.31	2.32	2.32	2.52	2.52	2.52
XeF_6	1.07	0.97	1.09	1.29	2.16	2.30	2.30	2.30	2.51	2.51	2.51
XeOF_4	1.27	1.08	1.35	1.45	2.97	3.07	3.08	3.09	3.32	3.33	3.33
XeO_4	1.26	0.99	1.53	1.14	2.88	2.91	2.93	2.93	3.18	3.18	3.19
Xe**F_2**	-0.24	-0.21	-0.26	-0.30	-0.55	-0.61	-0.61	-0.61	-0.70	-0.70	-0.70
Xe**F_4**	-0.26	-0.23	-0.26	-0.32	-0.54	-0.58	-0.58	-0.58	-0.63	-0.63	-0.63
Xe**F_6**	-0.18	-0.16	-0.18	-0.21	-0.36	-0.38	-0.38	-0.38	-0.42	-0.42	-0.42
Xe**O**F_4	-0.24	-0.21	-0.26	-0.30	-0.55	-0.57	-0.58	-0.58	-0.62	-0.62	-0.62
Xe**O_4**	-0.32	-0.25	-0.38	-0.28	-0.72	-0.73	-0.73	-0.73	-0.79	-0.80	-0.80

Previously studies on compounds containing xenon with NPA population concluded that Xe-C and Xe-O bonds are electrostatic in character because of Coulombic attraction between highly

positive xenon and negatively charged carbon or oxygen.[69] In Table 4-8 we will look at some highly stable xenon compounds from literature to see the bonding trends. For xenon charges there is a clear difference between CM5, QH, ESP, Mulliken and NPA, QAO, ZD-QAO populations. Former population methods produce significantly lower xenon charges, range seen for different structures is very small. While the latter populations have a wide range of atomic charges proving that QAO, much like NPA, considers these structures electrostatic in nature. For all studied populations, xenon charges in tetra- and hexa-fluoride are indistinguishable. Furthermore, QH and Mulliken fail to distinguish between xenon charges of tetrafluoride and tetraoxide. XeF_4 is square planar molecule while XeO_4 is tetrahedral with four double bonded oxygens. Unlike QH and Mulliken populations suggest, there should be difference between the atomic charges of oxygen and fluorine atoms. Overall, highest (or lowest) charges are observed with ZBD-QAO populations. This is expected because ZBD is localizes the electrons to atom centers, while density partitioning methods like CM5 and QH allow the electrons to be more delocalized. As one would expect, these xenon structures are the band like atomic charges seen for NPA and QAO populations.

	CM5	QH	ESP	Mulliken	NPA	Q1	Q2	Q3	Z-Q1	Z-Q2	Z-Q3
Xe	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓

5 Conclusion

In this work CM5, QH, ESP, Mulliken, NPA and the new QAO populations (QUAMBO, QUAO, IAO) and Zero-Bond Dipole orthogonalized versions (ZBD-QUAMBO , ZBD-QUAO, ZBD-IAO) were compared with three different methods (HF, B3LYP, ωB97XD) and basis sets (Def2-SVPP, Def2-TZVPP, Def2-QZVPP) over 1894 different molecular structures. For mathematical accuracy method, basis set, conformation, atom type dependency have been considered.

	CM5	QH	ESP	Mulliken	NPA	QUAMBO	QUAO	IAO
Method independence (with MB08)	✓	✓	×	×	×	×	×	✓
Basis set independence (with MB08)	✓	✓	×	×	✓	×	×	✓
Atom type	✓	✓	×	×	×	×	×	✓
Conformational stability	✓	✓	×	✓	✓	✓	✓	✓
Outliers	×	×	×	×	×	×	×	×

The resulting suggestion is CM5, QH and IAO for covering desired mathematical traits. Nevertheless, extreme cases or outliers to break the expected charges exist for all methods.

	CM5	QH	ESP	Mulliken	NPA	Q1	Q2	Q3	Z-Q1	Z-Q2	Z-Q3
Hydrocarbons	✓	✓	×	✓	✓	×	×	×	✓	✓	✓
Halogen trends	×	×	×	×	×	×	×	✓	✓	✓	✓
H-bond	✓	✓	×	✓	✓	✓	✓	✓	✓	✓	✓
Silica	✓	✓	×	✓	×	×	×	×	✓	✓	✓
Xenon	×	×	×	×	✓	✓	✓	✓	✓	✓	✓

For chemical trend expectations, non-orthogonalized QUAMBO and QUAO are not an improvement over Mulliken population. We recommend IAO to improve chemical trends or using ZBD orthogonalization with any of the QAO flavors. ESP population should not be used if reliable chemical trends or mathematical accuracy is desired.

References

(1) Ruedenberg, K. The Physical Nature of the Chemical Bond. *Rev. Mod. Phys.* **1962**, *34* (2), 326–376.

(2) Matta, C. F.; Bader, R. F. W. An Experimentalist's Reply to "What Is an Atom in a Molecule?" *J. Phys. Chem. A* **2006**, *110* (19), 6365–6371.

(3) Parr, R. G.; Ayers, P. W.; Nalewajski, R. F. What Is an Atom in a Molecule? *J. Phys. Chem. A* **2005**, *109* (17), 3957–3959.

(4) Szabo. *Modern Quantum Chemistry*; 1967; Vol. 35.

(5) Born, M.; Oppenheimer, R. Zur Quantentheorie Der Molekeln. *Ann. Phys.* **1927**, *389* (20), 457–484.

(6) McLachlan, A. D.; Ball, M. A. Time-Dependent Hartree Fock Theory for Molecules. *Rev. Mod. Phys.* **1964**, *36* (3), 844–855.

(7) Burke, K.; Perdew, J. P.; Wang, Y. *Electronic Density Functional Theory: Recent Progress and New Directions*; 1998.

(8) Runge, E.; Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **1984**, *52* (12), 997–1000.

(9) Parr, R. G. Density Functional Theory of Atoms and Molecules. In *Horizons of Quantum Chemistry*; Springer Netherlands, 1980; pp 5–15.

(10) Hirshfeld, F. L. Bonded-Atom Fragments for Describing Molecular Charge Densities. *Theor. Claim. Acta* **1977**, *44* (2), 129–138.

(11) Bader, R. F. W.; Anderson, S. G.; Duke, A. J. Quantum Topology of Molecular Charge Distributions. *J. Am. Chem. Soc.* **1979**, *101* (6), 1389–1395.

(12) Bader, R. F. W. Atoms in Molecules. *Acc. Chem. Res.* **1985**, *18* (1), 9–15.

(13) Spackman, M. A.; Jayatilaka, D. Hirshfeld Surface Analysis. *CrystEngComm* **2009**, *11* (1), 19–32.

(14) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V., and D. J. F. Gaussian 16. Gaussian, Inc., Wallingford CT 2016.

(15) Marenich, A. V.; Jerome, S. V.; Cramer, C. J.; Truhlar, D. G. Charge Model 5: An Extension of Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in Gaseous and Condensed Phases. *J. Chem. Theory Comput.* **2012**, *8* (2), 527–541.

(16) Nalewajski, R. F.; Parr, R. G.; Guminski, K. *Information Theory, Atoms in Molecules, and Molecular Similarity*; 2000.

(17) Heidar-Zadeh, F.; Ayers, P. W. How Pervasive Is the Hirshfeld Partitioning? *J. Chem. Phys.* **2015**, *142* (4).

(18) Heidar-Zadeh, F.; Ayers, P. W.; Bultinck, P. Deriving the Hirshfeld Partitioning Using Distance Metrics. *J. Chem. Phys.* **2014**, *141* (9).

(19) Heidar-Zadeh, F.; Vinogradov, I.; Ayers, P. W. Hirshfeld Partitioning from Non-Extensive Entropies. *Theor. Chem. Acc.* **2017**, *136* (4).

(20) Heidar-Zadeh, F.; Ayers, P. W.; Verstraelen, T.; Vinogradov, I.; Vöhringer-Martinez, E.; Bultinck, P. Information-Theoretic Approaches to Atoms-in-Molecules: Hirshfeld Family of Partitioning Schemes. *J. Phys. Chem. A* **2018**, *122* (17), 4219–4245.

(21) Heidar-Zadeh, F.; Ayers, P. W.; Verstraelen, T.; Vinogradov, I.; Vöhringer-Martinez, E.; Bultinck, P. Information-Theoretic Approaches to Atoms-in-Molecules: Hirshfeld Family of Partitioning Schemes. *J. Phys. Chem. A* **2018**, *122* (17), 4219–4245.

(22) Bultinck, P.; Alsenoy, C. Van; Ayers, P. W.; Carbó-Dorca, R. Critical Analysis and Extension of the Hirshfeld Atoms in Molecules. **2007**.

(23) Verstraelen, T.; Vandenbrande, S.; Heidar-Zadeh, F.; Vanduyfhuys, L.; Van Speybroeck, V.; Waroquier, M.; Ayers, P. W. Minimal Basis Iterative Stockholder: Atoms in Molecules for Force-Field Development. *J. Chem. Theory Comput.* **2016**, *12* (8), 3894–3912.

(24) Mulliken, R. S. Electronic Population Analysis on LCAO–MO Molecular Wave Functions. II. Overlap Populations, Bond Orders, and Covalent Bond Energies. *J. Chem. Phys.* **1955**, *23* (10), 1841–1846.

(25) Slater, J. C.; Koster, G. F. Simplified LCAO Method for the Periodic Potential Problem. *Phys. Rev.* **1954**, *94* (6), 1498–1524.

(26) Ziegler, T.; Snijders, J. G.; Baerends, E. J. Relativistic Effects on Bonding. *J. Chem. Phys.* **1981**, *74* (2), 1271–1284.

(27) Delley, B. A Scattering Theoretic Approach to Scalar Relativistic Corrections on Bonding. *Int. J. Quantum Chem.* **1998**, *69* (3), 423–433.

(28) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. A Complete Basis Set Model Chemistry. VII. Use of the Minimum Population Localization Method. *J. Chem. Phys.* **2000**, *112* (15), 6532–6542.

(29) Vedene H. Smith, J. Theoretical Determination and Analysis of Electronic Charge Distributions. *Phys. Scripta.* **1977**, *15*, 147–162.

(30) Derricotte, W. D.; Evangelista, F. A. Localized Intrinsic Valence Virtual Orbitals as a Tool for the Automatic Classification of Core Excited States. *J. Chem. Theory Comput.* **2017**, *13* (12), 5984–5999.

(31) Kotos, W. Possible Improvements of the Interaction Energy Calculated Using Minimal Basis Sets. *Theor. Chim. Acta* **1979**, *51* (3), 219–240.

(32) West, A. C.; Schmidt, M. W.; Gordon, M. S.; Ruedenberg, K. A Comprehensive Analysis in Terms of Molecule-Intrinsic, Quasi-Atomic Orbitals. III. the Covalent Bonding Structure of Urea. *J. Phys. Chem. A* **2015**, *119* (41), 10368–10375.

(33) 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.

(34) Schmidt, M. W.; Hull, E. A.; Windus, T. L. Valence Virtual Orbitals: An Unambiguous Ab Initio Quantification of the LUMO Concept. *J. Phys. Chem. A* **2015**, *119* (41), 10408–10427.

(35) West, A. C.; Schmidt, M. W.; Gordon, M. S.; Ruedenberg, K. A Comprehensive Analysis in Terms of Molecule-Intrinsic Quasi-Atomic Orbitals. IV. Bond Breaking and Bond Forming along the Dissociative Reaction Path of Dioxetane. *J. Phys. Chem. A* **2015**, *119* (41), 10376–10389.

(36) Knizia, G. Intrinsic Atomic Orbitals: An Unbiased Bridge between Quantum Theory and Chemical Concepts. *J. Chem. Theory Comput.* **2013**, *9* (11), 4834–4843.

(37) Laikov, D. N. Intrinsic Minimal Atomic Basis Representation of Molecular Electronic Wavefunctions. *Int. J. Quantum Chem.* **2010**, *111* (12), 2851–2867.

(38) Heßelmann, A. Local Molecular Orbitals from a Projection onto Localized Centers. *J. Chem. Theory Comput.* **2016**, acs.jctc.6b00321.

(39) Quandt, A. *Valency and Bonding. A Natural Bond Orbital Donor–Acceptor Perspective. By Frank Weinhold and Clark Landis.*; 2006; Vol. 7.

(40) Distribution, S. E. Computer Simulation of the Conformational Properties of Retro-Inverso Peptides. 11. *Biopolymers* **1983**, *22*, 1901–1917.

(41) Hu, H.; Lu, Z.; Yang, W. Fitting Molecular Electrostatic Potentials from Quantum Mechanical Calculations. *J. Chem. Theory Comput.* **2007**, *3* (3), 1004–1013.

(42) Reed, A.; Weinstock, R.; Weinhold, F. Natural Population Analysis. *J. Chem. Phys.* **1985**, *83*, 735.

(43) Swain, M. PubChemPy https://github.com/mcs07/PubChemPy.

(44) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785–789.

(45) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38* (6), 3098–3100.

(46) Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom-Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10* (44), 6615–6620.

(47) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297–3305.

(48) Weigend, F. Accurate Coulomb-Fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8* (9), 1057–1065.

(49) Mayer, I. Charge, Bond Order and Valence in the Ab Initio SCF Theory. *Chem. Phys. Lett.* **1983**, *97* (3).

(50) Glezakou, V. A.; Elbert, S. T.; Xantheas, S. S.; Ruedenberg, K. Analysis of Bonding

Patterns in the Valence Isoelectronic Series O 3, S3, SO2, and OS2 in Terms of Oriented Quasi-Atomic Molecular Orbitals. *J. Phys. Chem. A* **2010**, *114* (33), 8923–8931.

(51) Wannere, C. S.; Schleyer, P. V. R.; Schaefer, H. F. The Design of "Neutral" Carbanions with Intramolecular Charge Compensation. *J. Org. Chem.* **2016**, *81* (5), 1885–1898.

(52) Winstein, S.; Sonnenberg, J.; DeVries, L. The Tris-Homocyclopropenyl Cation. *J. Am. Chem. Soc.* **1959**, *81* (24), 6523–6524.

(53) Dodda, L. S.; Vilseck, J. Z.; Cutrona, K. J.; Jorgensen, W. L. Evaluation of CM5 Charges for Nonaqueous Condensed-Phase Modeling. *J. Chem. Theory Comput.* **2015**, *11* (9), 4273–4282.

(54) Kesharwani, M. K.; Manna, D.; Sylvetsky, N.; Martin, J. M. L. The X40×10 Halogen Bonding Benchmark Revisited: Surprising Importance of (n-1)d Subvalence Correlation. *J. Phys. Chem. A* **2018**, *122* (8), 2184–2197.

(55) Calvo, F.; Spiegelmann, F. Mechanisms of Phase Transitions in Sodium Clusters: From Molecular to Bulk Behavior. *J. Chem. Phys.* **2000**, *112* (6), 2888–2908.

(56) Bader, R. F. W.; Tang, T. H.; Tal, Y.; Biegler-Koenig, F. W. Properties of Atoms and Bonds in Hydrocarbon Molecules. *J. Am. Chem. Soc.* **1982**, *104* (4), 946–952.

(57) Torii, H.; Yoshida, M. Properties of Halogen Atoms for Representing Intermolecular Electrostatic Interactions Related to Halogen Bonding and Their Substituent Effects. *J. Comput. Chem.* **2010**, *31* (1), 107–116.

(58) Valerio, G.; Raos, G.; Meille, S. V.; Metrangolo, P.; Resnati, G. Halogen Bonding in Fluoroalkylhalides: A Quantum Chemical Study of Increasing Fluorine Substitution. *J. Phys. Chem. A* **2000**, *104* (8), 1617–1620.

(59) Hobza, P.; Sandorfy, C. Nonempirical Calculations on All the 29 Possible DNA Base Pairs. *J. Am. Chem. Soc.* **1987**, *109* (5), 1302–1307.

(60) Nagata, C.; Aida, M. Ab Initio Molecular Orbital Study on the Pairing and Stacking Interactions between Nucleic Acid Bases in Relation to the Biological Activities. *J. Mol. Struct. THEOCHEM* **1988**, *179* (1), 451–466.

(61) Aida, M. An Ab Initio Molecular Orbital Study on the Sequence-Dependency of DNA Conformation: An Evaluation of Intra- and Inter-Strand Stacking Interaction Energy. *J. Theor. Biol.* **1988**, *130* (3), 327–335.

(62) Aida, M. Characteristics of the Watson-Crick Type Hydrogen-Bonded DNA Base Pairs: Anab Initio Molecular Orbital Study. *J. Comput. Chem.* **1988**, *9* (4), 362–368.

(63) Hornby, D. Hydrogen Bonding in Biological Structures. *FEBS Lett.* **1993**, *323* (3), 295–295.

(64) Zhou, G.; Zhang, J. L.; Wong, N. B.; Tian, A. Theoretical Study of the Blue-Shifting Intramolecular Hydrogen Bonds of Nitro Derivatives of Cubane. *J. Mol. Struct. THEOCHEM* **2003**, *639* (1–3), 43–51.

(65) Weinhold, F.; Reed', A. E.; Curtiss, L. A. Intermolecular Interactions from a Natural Bond Orbital, Donor-Acceptor Viewpoint.

(66) Hobza, P.; Havlas, Z. Blue-Shifting Hydrogen Bonds. *Chem. Rev.* **2000**, *100* (11), 4253–

4264.

(67) Lao, K. U.; Schäffer, R.; Jansen, G.; Herbert, J. M. Accurate Description of Intermolecular Interactions Involving Ions Using Symmetry-Adapted Perturbation Theory. *J. Chem. Theory Comput.* **2015**, *11* (6), 2473–2486.

(68) Pearson, R. G. The Principle of Maximum Hardness. *Acc. Chem. Res.* **1993**, *26* (5), 250–255.

(69) Haiduke, R. L. A.; de Paiva Martins Filho, H.; da Silva, A. B. F. A Theoretical Study on the XeF2 Molecule. *Chem. Phys.* **2008**, *348* (1–3), 89–96.

(70) Pysh, E. S.; Jortner, J.; Rice, S. A. Forbidden Electronic Transitions in XeF2and XeF4. *J. Chem. Phys.* **1964**, *40* (7), 2018–2032.

(71) Jortner, J.; Rice, S. A.; Guy Wilson, E. Speculation Concerning the Nature of Binding in Xenon Fluorine Compounds [1]. *The Journal of Chemical Physics*. 1963, pp 2302–2303.

(72) Rundle, R. E. On the Probable Structure of XeF4 and XeF2. *Journal of the American Chemical Society*. January 1, 1963, pp 112–113.

(73) Coulson, C. A. Force Fields in KrF2and XeF2. *J. Chem. Phys.* **1966**, *44* (2), 468–469.

(74) Liao, M. S.; Zhang, Q. E. Chemical Bonding in XeF2, XeF4, KrF2, KrF4, RnF2, XeCl2, and XeBr2: From the Gas Phase to the Solid State. *J. Phys. Chem. A* **1998**, *102* (52), 10647–10654.

(75) Mobley, D. L.; Guthrie, J. P. FreeSolv: A Database of Experimental and Calculated Hydration Free Energies, with Input Files. *J. Comput. Aided. Mol. Des.* **2014**, *28* (7), 711–720.

(76) Zahn, S.; Macfarlane, D. R.; Izgorodina, E. I. Assessment of Kohn-Sham Density Functional Theory and Møller-Plesset Perturbation Theory for Ionic Liquids. *Phys. Chem. Chem. Phys.* **2013**, *15* (32), 13664–13675.

(77) Verstraelen, T.; Vandenbrande, S.; Heidar-Zadeh, F.; Vanduyfhuys, L.; Speybroeck, V. Van; Waroquier, M.; Ayers, P. W. *Supporting Information for "Minimal Basis Iterative Stockholder: Atoms in Molecules for Force-Field Development."*

(78) Korth, M.; Grimme, S. Mindless DFT Benchmarking. *J. Chem. Theory Comput.* **2009**, *5* (4), 993–1003.

(79) Řezáč, J.; Riley, K. E.; Hobza, P. Benchmark Calculations of Noncovalent Interactions of Halogenated Molecules. *J. Chem. Theory Comput.* **2012**, *8* (11), 4285–4292.

(80) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. Benchmark Database of Accurate (MP2 and CCSD(T) Complete Basis Set Limit) Interaction Energies of Small Model Complexes, DNA Base Pairs, and Amino Acid Pairs. *Phys. Chem. Chem. Phys.* **2006**, *8* (17), 1985–1993.

(1) Ruedenberg, K. The Physical Nature of the Chemical Bond. *Rev. Mod. Phys.* **1962**, *34* (2), 326–376.

(2) Matta, C. F.; Bader, R. F. W. An Experimentalist's Reply to "What Is an Atom in a Molecule?" *J. Phys. Chem. A* **2006**, *110* (19), 6365–6371.

(3) Parr, R. G.; Ayers, P. W.; Nalewajski, R. F. What Is an Atom in a Molecule? *J. Phys. Chem. A* **2005**, *109* (17), 3957–3959.

(4) Szabo. *Modern Quantum Chemistry*; 1967; Vol. 35.

(5) Born, M.; Oppenheimer, R. Zur Quantentheorie Der Molekeln. *Ann. Phys.* **1927**, *389* (20), 457–484.

(6) McLachlan, A. D.; Ball, M. A. Time-Dependent Hartree Fock Theory for Molecules. *Rev. Mod. Phys.* **1964**, *36* (3), 844–855.

(7) Burke, K.; Perdew, J. P.; Wang, Y. *Electronic Density Functional Theory: Recent Progress and New Directions*; 1998.

(8) Runge, E.; Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **1984**, *52* (12), 997–1000.

(9) Parr, R. G. Density Functional Theory of Atoms and Molecules. In *Horizons of Quantum Chemistry*; Springer Netherlands, 1980; pp 5–15.

(10) Hirshfeld, F. L. Bonded-Atom Fragments for Describing Molecular Charge Densities. *Theor. Claim. Acta* **1977**, *44* (2), 129–138.

(11) Bader, R. F. W.; Anderson, S. G.; Duke, A. J. Quantum Topology of Molecular Charge Distributions. *J. Am. Chem. Soc.* **1979**, *101* (6), 1389–1395.

(12) Bader, R. F. W. Atoms in Molecules. *Acc. Chem. Res.* **1985**, *18* (1), 9–15.

(13) Spackman, M. A.; Jayatilaka, D. Hirshfeld Surface Analysis. *CrystEngComm* **2009**, *11* (1), 19–32.

(14) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V., and D. J. F. Gaussian 16. Gaussian, Inc., Wallingford CT 2016.

(15) Marenich, A. V.; Jerome, S. V.; Cramer, C. J.; Truhlar, D. G. Charge Model 5: An Extension of Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in Gaseous and Condensed Phases. *J. Chem. Theory Comput.* **2012**, *8* (2), 527–541.

(16) Nalewajski, R. F.; Parr, R. G.; Guminski, K. *Information Theory, Atoms in Molecules, and Molecular Similarity*; 2000.

(17) Heidar-Zadeh, F.; Ayers, P. W. How Pervasive Is the Hirshfeld Partitioning? *J. Chem. Phys.* **2015**, *142* (4).

(18) Heidar-Zadeh, F.; Ayers, P. W.; Bultinck, P. Deriving the Hirshfeld Partitioning Using Distance Metrics. *J. Chem. Phys.* **2014**, *141* (9).

(19) Heidar-Zadeh, F.; Vinogradov, I.; Ayers, P. W. Hirshfeld Partitioning from Non-Extensive Entropies. *Theor. Chem. Acc.* **2017**, *136* (4).

(20) Heidar-Zadeh, F.; Ayers, P. W.; Verstraelen, T.; Vinogradov, I.; Vöhringer-Martinez, E.; Bultinck, P. Information-Theoretic Approaches to Atoms-in-Molecules: Hirshfeld Family of Partitioning Schemes. *J. Phys. Chem. A* **2018**, *122* (17), 4219–4245.

(21) Heidar-Zadeh, F.; Ayers, P. W.; Verstraelen, T.; Vinogradov, I.; Vöhringer-Martinez, E.; Bultinck, P. Information-Theoretic Approaches to Atoms-in-Molecules: Hirshfeld Family of Partitioning Schemes. *J. Phys. Chem. A* **2018**, *122* (17), 4219–4245.

(22) Bultinck, P.; Alsenoy, C. Van; Ayers, P. W.; Carbó-Dorca, R. Critical Analysis and Extension of the Hirshfeld Atoms in Molecules. **2007**.

(23) Verstraelen, T.; Vandenbrande, S.; Heidar-Zadeh, F.; Vanduyfhuys, L.; Van Speybroeck, V.; Waroquier, M.; Ayers, P. W. Minimal Basis Iterative Stockholder: Atoms in Molecules for Force-Field Development. *J. Chem. Theory Comput.* **2016**, *12* (8), 3894–3912.

(24) Mulliken, R. S. Electronic Population Analysis on LCAO–MO Molecular Wave Functions. II. Overlap Populations, Bond Orders, and Covalent Bond Energies. *J. Chem. Phys.* **1955**, *23* (10), 1841–1846.

(25) Slater, J. C.; Koster, G. F. Simplified LCAO Method for the Periodic Potential Problem. *Phys. Rev.* **1954**, *94* (6), 1498–1524.

(26) Ziegler, T.; Snijders, J. G.; Baerends, E. J. Relativistic Effects on Bonding. *J. Chem. Phys.* **1981**, *74* (2), 1271–1284.

(27) Delley, B. A Scattering Theoretic Approach to Scalar Relativistic Corrections on Bonding. *Int. J. Quantum Chem.* **1998**, *69* (3), 423–433.

(28) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. A Complete Basis Set Model Chemistry. VII. Use of the Minimum Population Localization Method. *J. Chem. Phys.* **2000**, *112* (15), 6532–6542.

(29) Vedene H. Smith, J. Theoretical Determination and Analysis of Electronic Charge Distributions. *Phys. Scripta.* **1977**, *15*, 147–162.

(30) Derricotte, W. D.; Evangelista, F. A. Localized Intrinsic Valence Virtual Orbitals as a Tool for the Automatic Classification of Core Excited States. *J. Chem. Theory Comput.* **2017**, *13* (12), 5984–5999.

(31) Kotos, W. Possible Improvements of the Interaction Energy Calculated Using Minimal Basis Sets. *Theor. Chim. Acta* **1979**, *51* (3), 219–240.

(32) West, A. C.; Schmidt, M. W.; Gordon, M. S.; Ruedenberg, K. A Comprehensive Analysis in Terms of Molecule-Intrinsic, Quasi-Atomic Orbitals. III. the Covalent Bonding Structure of Urea. *J. Phys. Chem. A* **2015**, *119* (41), 10368–10375.

(33) 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.

(34) Schmidt, M. W.; Hull, E. A.; Windus, T. L. Valence Virtual Orbitals: An Unambiguous Ab Initio Quantification of the LUMO Concept. *J. Phys. Chem. A* **2015**, *119* (41), 10408–10427.

(35) West, A. C.; Schmidt, M. W.; Gordon, M. S.; Ruedenberg, K. A Comprehensive Analysis in Terms of Molecule-Intrinsic Quasi-Atomic Orbitals. IV. Bond Breaking and Bond

Forming along the Dissociative Reaction Path of Dioxetane. *J. Phys. Chem. A* **2015**, *119* (41), 10376–10389.

(36) Knizia, G. Intrinsic Atomic Orbitals: An Unbiased Bridge between Quantum Theory and Chemical Concepts. *J. Chem. Theory Comput.* **2013**, *9* (11), 4834–4843.

(37) Laikov, D. N. Intrinsic Minimal Atomic Basis Representation of Molecular Electronic Wavefunctions. *Int. J. Quantum Chem.* **2010**, *111* (12), 2851–2867.

(38) Heßelmann, A. Local Molecular Orbitals from a Projection onto Localized Centers. *J. Chem. Theory Comput.* **2016**, acs.jctc.6b00321.

(39) Quandt, A. *Valency and Bonding. A Natural Bond Orbital Donor–Acceptor Perspective. By Frank Weinhold and Clark Landis.*; 2006; Vol. 7.

(40) Distribution, S. E. Computer Simulation of the Conformational Properties of Retro-Inverso Peptides. 11. *Biopolymers* **1983**, *22*, 1901–1917.

(41) Hu, H.; Lu, Z.; Yang, W. Fitting Molecular Electrostatic Potentials from Quantum Mechanical Calculations. *J. Chem. Theory Comput.* **2007**, *3* (3), 1004–1013.

(42) Reed, A.; Weinstock, R.; Weinhold, F. Natural Population Analysis. *J. Chem. Phys.* **1985**, *83*, 735.

(43) Swain, M. PubChemPy https://github.com/mcs07/PubChemPy.

(44) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785–789.

(45) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38* (6), 3098–3100.

(46) Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom-Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10* (44), 6615–6620.

(47) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297–3305.

(48) Weigend, F. Accurate Coulomb-Fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8* (9), 1057–1065.

(49) Mayer, I. Charge, Bond Order and Valence in the Ab Initio SCF Theory. *Chem. Phys. Lett.* **1983**, *97* (3).

(50) Glezakou, V. A.; Elbert, S. T.; Xantheas, S. S.; Ruedenberg, K. Analysis of Bonding Patterns in the Valence Isoelectronic Series O 3, S3, SO2, and OS2 in Terms of Oriented Quasi-Atomic Molecular Orbitals. *J. Phys. Chem. A* **2010**, *114* (33), 8923–8931.

(51) Wannere, C. S.; Schleyer, P. V. R.; Schaefer, H. F. The Design of "Neutral" Carbanions with Intramolecular Charge Compensation. *J. Org. Chem.* **2016**, *81* (5), 1885–1898.

(52) Winstein, S.; Sonnenberg, J.; DeVries, L. The Tris-Homocyclopropenyl Cation. *J. Am. Chem. Soc.* **1959**, *81* (24), 6523–6524.

(53) Dodda, L. S.; Vilseck, J. Z.; Cutrona, K. J.; Jorgensen, W. L. Evaluation of CM5 Charges

for Nonaqueous Condensed-Phase Modeling. *J. Chem. Theory Comput.* **2015**, *11* (9), 4273–4282.

(54) Kesharwani, M. K.; Manna, D.; Sylvetsky, N.; Martin, J. M. L. The X40×10 Halogen Bonding Benchmark Revisited: Surprising Importance of (n-1)d Subvalence Correlation. *J. Phys. Chem. A* **2018**, *122* (8), 2184–2197.

(55) Calvo, F.; Spiegelmann, F. Mechanisms of Phase Transitions in Sodium Clusters: From Molecular to Bulk Behavior. *J. Chem. Phys.* **2000**, *112* (6), 2888–2908.

(56) Bader, R. F. W.; Tang, T. H.; Tal, Y.; Biegler-Koenig, F. W. Properties of Atoms and Bonds in Hydrocarbon Molecules. *J. Am. Chem. Soc.* **1982**, *104* (4), 946–952.

(57) Torii, H.; Yoshida, M. Properties of Halogen Atoms for Representing Intermolecular Electrostatic Interactions Related to Halogen Bonding and Their Substituent Effects. *J. Comput. Chem.* **2010**, *31* (1), 107–116.

(58) Valerio, G.; Raos, G.; Meille, S. V.; Metrangolo, P.; Resnati, G. Halogen Bonding in Fluoroalkylhalides: A Quantum Chemical Study of Increasing Fluorine Substitution. *J. Phys. Chem. A* **2000**, *104* (8), 1617–1620.

(59) Hobza, P.; Sandorfy, C. Nonempirical Calculations on All the 29 Possible DNA Base Pairs. *J. Am. Chem. Soc.* **1987**, *109* (5), 1302–1307.

(60) Nagata, C.; Aida, M. Ab Initio Molecular Orbital Study on the Pairing and Stacking Interactions between Nucleic Acid Bases in Relation to the Biological Activities. *J. Mol. Struct. THEOCHEM* **1988**, *179* (1), 451–466.

(61) Aida, M. An Ab Initio Molecular Orbital Study on the Sequence-Dependency of DNA Conformation: An Evaluation of Intra- and Inter-Strand Stacking Interaction Energy. *J. Theor. Biol.* **1988**, *130* (3), 327–335.

(62) Aida, M. Characteristics of the Watson-Crick Type Hydrogen-Bonded DNA Base Pairs: Anab Initio Molecular Orbital Study. *J. Comput. Chem.* **1988**, *9* (4), 362–368.

(63) Hornby, D. Hydrogen Bonding in Biological Structures. *FEBS Lett.* **1993**, *323* (3), 295–295.

(64) Zhou, G.; Zhang, J. L.; Wong, N. B.; Tian, A. Theoretical Study of the Blue-Shifting Intramolecular Hydrogen Bonds of Nitro Derivatives of Cubane. *J. Mol. Struct. THEOCHEM* **2003**, *639* (1–3), 43–51.

(65) Weinhold, F.; Reed', A. E.; Curtiss, L. A. Intermolecular Interactions from a Natural Bond Orbital, Donor-Acceptor Viewpoint.

(66) Hobza, P.; Havlas, Z. Blue-Shifting Hydrogen Bonds. *Chem. Rev.* **2000**, *100* (11), 4253–4264.

(67) Lao, K. U.; Schäffer, R.; Jansen, G.; Herbert, J. M. Accurate Description of Intermolecular Interactions Involving Ions Using Symmetry-Adapted Perturbation Theory. *J. Chem. Theory Comput.* **2015**, *11* (6), 2473–2486.

(68) Pearson, R. G. The Principle of Maximum Hardness. *Acc. Chem. Res.* **1993**, *26* (5), 250–255.

(69) Haiduke, R. L. A.; de Paiva Martins Filho, H.; da Silva, A. B. F. A Theoretical Study on

the XeF2 Molecule. *Chem. Phys.* **2008**, *348* (1–3), 89–96.

(70) Pysh, E. S.; Jortner, J.; Rice, S. A. Forbidden Electronic Transitions in XeF2and XeF4. *J. Chem. Phys.* **1964**, *40* (7), 2018–2032.

(71) Jortner, J.; Rice, S. A.; Guy Wilson, E. Speculation Concerning the Nature of Binding in Xenon Fluorine Compounds [1]. *The Journal of Chemical Physics*. 1963, pp 2302–2303.

(72) Rundle, R. E. On the Probable Structure of XeF4 and XeF2. *Journal of the American Chemical Society*. January 1, 1963, pp 112–113.

(73) Coulson, C. A. Force Fields in KrF2and XeF2. *J. Chem. Phys.* **1966**, *44* (2), 468–469.

(74) Liao, M. S.; Zhang, Q. E. Chemical Bonding in XeF2, XeF4, KrF2, KrF4, RnF2, XeCl2, and XeBr2: From the Gas Phase to the Solid State. *J. Phys. Chem. A* **1998**, *102* (52), 10647–10654.

(75) Mobley, D. L.; Guthrie, J. P. FreeSolv: A Database of Experimental and Calculated Hydration Free Energies, with Input Files. *J. Comput. Aided. Mol. Des.* **2014**, *28* (7), 711–720.

(76) Zahn, S.; Macfarlane, D. R.; Izgorodina, E. I. Assessment of Kohn-Sham Density Functional Theory and Møller-Plesset Perturbation Theory for Ionic Liquids. *Phys. Chem. Chem. Phys.* **2013**, *15* (32), 13664–13675.

(77) Verstraelen, T.; Vandenbrande, S.; Heidar-Zadeh, F.; Vanduyfhuys, L.; Speybroeck, V. Van; Waroquier, M.; Ayers, P. W. *Supporting Information for "Minimal Basis Iterative Stockholder: Atoms in Molecules for Force-Field Development."*

(78) Korth, M.; Grimme, S. Mindless DFT Benchmarking. *J. Chem. Theory Comput.* **2009**, *5* (4), 993–1003.

(79) Řezáč, J.; Riley, K. E.; Hobza, P. Benchmark Calculations of Noncovalent Interactions of Halogenated Molecules. *J. Chem. Theory Comput.* **2012**, *8* (11), 4285–4292.

(80) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. Benchmark Database of Accurate (MP2 and CCSD(T) Complete Basis Set Limit) Interaction Energies of Small Model Complexes, DNA Base Pairs, and Amino Acid Pairs. *Phys. Chem. Chem. Phys.* **2006**, *8* (17), 1985–1993.

A. APPENDIX : Chemical Database

JPCA (ID 860-133) quantum chemistry benchmarking dataset for polarizability of DFT functionals. This is a detailed study on wave function methods and basis sets. Basis sets used in this thesis are not included in this list. However, five different basis sets and ten different methods have been put to test. Their results on dipole moments show that B3LYP is better than all other methods.

Mobley (ID 132–775) this dataset is from Mobley's hydration free energies database FreeSolv[75] and consists of small neutral molecules. This research has some interesting molecules which they refer to as "extremas".

JOC (ID 774-818) intermolecular charge transfer compounds from zwitterions and neutral carbocations.[51]

AHB21 (ID 817–839) anion-neutral interactions[67]

CHB6 (ID 838–845) cation-neutral interactions[67]

IL16 (ID 844–861) anion-cation interactions, a subset of IL-2013[76]

ZG237 (ID 1035–1271) non-covalent interactions of silica clusters with different types of molecules including noble gases, cationic and neutral small organic molecules.[23] There are 12 different silica clusters interacting with one of the 20 different molecules.[77]

MB08 (ID 1272–1451) unusual bonding dataset for "mindless" DFT benchmarking.[78] This dataset is first introduced as one of the sets to benchmark Becke and Minnesota density functionals. This dataset consists of 165 molecules which are artificially created to strip chemical bias from methods. While it provides a way to test for electronically challenging systems it is not completely random ansatz. All the geometries were optimized by PBE/TZVP level and benchmarked with CCSD(T)/cc-pVTZ.

X40 (ID 1452–1491) is a dataset of 40 non-covalent interactions of halogens and small organic molecules.[54,79] It covers a variety of different interaction types but possibly the most notable are halogen bonding and π–halogen interactions. All geometries are optimized using CCSD(T)/CBS method.

S66 (ID 1492–1558) well-balanced non-covalent interactions dataset for benchmarking interaction energies. It is an extension of their bioorganic chemistry dataset S22[80] aimed for parametrization of various computational techniques. Like S22, S66 dataset is considered well-balanced because it includes almost equal amounts of electrostatic, dispersions and mixed interaction types. These interactions are further categorized into H-bonding, dispersion dominated and other type of interactions. The geometries of these molecules were optimized using in several series of steps ending with MP2/cc-pVTZ and reported in the supporting information.

Figure A-1. Molecule ID and chemical formulas in database.

0	Cl4C1	31	N1H2	62	As1Cl3O1	93	S1O1
1	Cl3C1H1	32	S1O4H2	63	S1F2	94	O2C1
2	Cl2C1H2	33	S1O4H1	64	S1F4	95	S1O2
3	Cl1C1H3	34	S1O4	65	S1F6	96	C2H4
4	C1H4	35	S1O3H2	66	S1F2O1	97	C3H8
5	Br4C1	36	S1O3H1	67	Cl1F1O1	98	C4H6
6	Br3C1H1	37	P1O4H3	68	Cl1F1O2	99	C4H8
7	Br2C1H2	38	P1O3H1	69	Cl1F3O1	100	C5H10
8	Br1C1H3	39	Cl1O4H1	70	Cl1F1O3	101	O1H2
9	C1H4	40	Cl1O3H1	71	Cl1F3O2	102	O1C1H4
10	F4C1	41	C60	72	Br1F1O2	103	O1C2H6
11	F3C1H1	42	Al1F4	73	Br1F3O1	104	O2C2H4
12	F2C1H2	43	Cl4Al1	74	Br1F1O3	105	O2C3H6
13	F1C1H3	44	Br4Al1	75	Cl1F1	106	O1C3H6
14	C1H4	45	I4Al1	76	Br1F1	107	O1C2H4
15	I4C1	46	Na13	77	Cl1F3	108	O2C2H4
16	I3C1H1	47	P1F3	78	Br1F3	109	O1C2H6
17	I2C1H2	48	Cl3P1	79	Cl1F5	110	O1N1C3H7
18	I1C1H3	49	P1F5	80	Br1F5	111	S1C2H6
19	C1H4	50	Cl5P1	81	Xe1F2	112	S1O2C2H6
20	N1C3H9	51	P1F6	82	Xe1F4	113	F1C1H3
21	N1C2H7	52	Cl6P1	83	Xe1F6	114	P1H3
22	N1C1H5	53	P1F3O1	84	Xe1O1F4	115	S1H2
23	N1H3	54	Cl3P1O1	85	Xe1O4	116	Si1H4
24	N1C4H12	55	As1F3	86	Kr1F2	117	N1H3
25	N1C3H10	56	As1Cl3	87	Br2	118	N1C2H7
26	N1C2H8	57	As1F5	88	Cl2	119	N1C3H9
27	N1C1H6	58	As1Cl5	89	N2	120	N1C2H3
28	N1H4	59	As1F6	90	O2	121	N2C3H4
29	N1C2H6	60	As1Cl6	91	O1C1	122	N1C4H5
30	N1C1H4	61	As1F3O1	92	O1N1	123	N2C3H4

124	O1C4H4	158	Cl5O2N1C6	192	O1C6H12	224	S1C3H8
125	S1C4H4	159	Cl3S1O2N1C9H8	193	Cl3F3C2	225	O1C6H6
126	N1C5H5	160	C8H10	194	Cl1S3P1O2C11H16	226	C3H8
127	O1C6H6	161	C10H14			227	O2C13H18
128	Cl1C6H5	162	Cl2O2C12H6	195	Cl4C6H2	228	O3C16H14
129	F1C6H5	163	O1C9H18	196	Cl2C2H4	229	I1C3H7
130	C7H8	164	N1C6H7	197	O1C8H10	230	O1N1C7H5
131	C6H6	165	F1O2C15H13	198	S1C3H8	231	O1C8H10
132	O1N3C4H5	166	S1C4H10	199	C7H8	232	C5H10
133	O2C7H14	167	O1C8H10	200	O1C5H10	233	O1C1H2
134	O1C4H10	168	O2C8H16	201	O1C5H10	234	Cl1C5H11
135	Cl6C12H4	169	O1C6H14	202	Cl4C2H2	235	C10H14
136	N1C6H13	170	O1C1H4	203	Cl1C6H13	236	C10H22
137	O2C7H6	171	O1C4H10	204	O5C5H10	237	Cl1C2H5
138	O2C4H8	172	O1N1C4H9	205	C6H12	238	C6H14
139	Cl2C6H4	173	C4H8	206	O1C4H10	239	O1C4H10
140	I1C2H5	174	C6H10	207	Br1C7H7	240	C12H10
141	C8H18	175	S1O2C2H6	208	S1P1O5N1C8H10	241	O3C14H14
142	O1C4H10	176	O2C2H4	209	C4H10	242	O3C4H10
143	O1C10H16	177	O2C6H12	210	S1H2	243	O2C3H6
144	Cl1O1C6H5	178	F8C4	211	Br1O2N2C9H13	244	O3C8H8
145	S1C7H8	179	O1N1C7H7	212	O1N1C17H21	245	O1C2H6
146	I1C3H7	180	O2C12H8	213	O2N1C1H3	246	O3N1C16H13
147	Cl1C3H7	181	Br1O1C6H5	214	O1N1C3H7	247	C10H20
148	O2N1C5H11	182	O2C6H12	215	O1C2H4	248	O2C6H12
149	O1C7H14	183	O4C8H14	216	O1C5H8	249	O1C8H10
150	C5H12	184	O1C8H10	217	O1C9H18	250	O2C6H12
151	Cl6C2	185	Cl1C7H7	218	Cl1C3H7	251	N1C7H9
152	O1C8H18	186	S2P1O3N3C10H12	219	C9H12	252	O1C7H14
153	P1O4C6H15	187	O2C6H12	220	C12H10	253	O4C7H12
154	Cl3C2H3	188	Br2C2H4	221	C2H6	254	N1C2H7
155	O4N1C8H7	189	S1C2H6	222	Cl1F5O1C3H2	255	O1C4H10
156	S1O1N1C10H21	190	C5H12	223	O1C8H14	256	O2N2C14H10
157	N1C7H9	191	N1C9H9			257	N1C7H5

258	O2C7H8	290	Br1C4H9	324	C9H12	358	Cl1C4H9
259	Cl3O2N1C1	291	C10H8	325	N1C7H9	359	O1C10H8
260	Cl1C7H7	292	N1C4H5	326	O2N4C11H18	360	O1C6H10
261	O1C6H12	293	Cl4O2C12H4	327	O6N2C3H6	361	O2C6H14
262	Cl2C6H4	294	Cl4O2C12H4	328	C6H10	362	O2C7H14
263	Cl5C2H1	295	O2C7H6	329	C8H10	363	C9H18
264	Cl2C2H2	296	C8H8	330	O3C3H8	364	C12H12
265	F3O4N3C14H16	297	O1C5H12	331	N1C10H9	365	O2C6H4
266	Br1C7H15	298	O3N1C4H9	332	Cl1O4C9H9	366	Br1C5H11
267	Cl1S2P1O4N1C14H17	299	O4C9H8	333	F3N1C9H10	367	C9H10
		300	O1C8H10	334	O1C4H10	368	Cl3C2H3
268	Cl1O1N3C10H8	301	O1C7H8	335	Cl7C12H3	369	Cl2C1H2
269	C7H16	302	O2C5H10	336	O5N2C10H12	370	N1C5H7
270	C16H10	303	N1C6H15	337	O1C9H18	371	O2N2C6H6
271	O2C8H16	304	Cl6O2C12H2	338	S1O2C4H8	372	O3N1C6H13
272	C7H16	305	N1C5H5	339	O3C10H14	373	N1C10H9
273	O2C8H10	306	C14H12	340	O2C10H14	374	O1C3H6
274	O6N2C3H6	307	S1C4H4	341	F2C2H4	375	O1C5H8
275	O1N1C10H13	308	C4H8	342	F1O2N2C4H3	376	O2C8H8
276	O1N1C7H5	309	Cl3C1H1	343	O1C9H20	377	O1C7H6
277	Cl3C6H3	310	F3O1C2H3	344	O1C5H10	378	O2C9H10
278	C6H12	311	C3H6	345	I1C6H5	379	Cl5O2C12H3
279	N2C8H12	312	O2C2H4	346	O1C6H14	380	Cl3C6H3
280	S1O6N3C13H19	313	C9H12	347	Cl1F1C1H2	381	O2C3H6
281	O1N1C6H11	314	S1N5C9H17	348	C9H12	382	Cl2C2H4
282	I1O2N2C4H3	315	C6H6	349	O1C9H10	383	O1C7H14
283	O3N1C6H5	316	C9H20	350	O2C7H8	384	O1C7H8
284	O1C8H10	317	O2C3H8	351	O4C6H10	385	Br1O2N2C4H3
285	N2C5H12	318	Cl1O2N2C9H13	352	C11H16	386	C5H10
286	O2N1C8H5	319	O1C3H8	353	O1C6H14	387	C7H16
287	C3H6	320	O1C8H10	354	O2C8H16	388	N2C6H14
288	Cl1N1C5H4	321	O1C10H22	355	O1C7H14	389	Cl2O2C12H6
289	O1C6H10	322	C8H18	356	Cl1O3C5H11	390	C8H18
		323	Cl1N1C6H6	357	O1C8H18	391	O1N1C9H11

392	O2N1C6H5	426	O2C6H14	460	S1O2N2C7H14	494	C12H12
393	O1N1C7H9	427	O1C8H16	461	N2C6H8	495	C6H10
394	Cl4C6H2	428	Cl2C3H6	462	S1C1H4	496	N1C2H7
395	C10H14	429	C8H16	463	C12H12	497	Cl2O3C8H6
396	N1C6H7	430	I1C7H15	464	O6N2C2H4	498	N2C4H6
397	Cl1N1C5H4	431	C7H14	465	N1C20H23	499	O2N1C7H7
398	C5H8	432	O1C8H10	466	Cl2C4H8	500	O4C6H14
399	N1C3H5	433	O2C5H10	467	C6H12	501	O1C5H12
400	Cl1O2N2C4H3	434	Cl6C6	468	O1C7H8	502	S1O4C2H6
401	I1C1H3	435	Cl1S1O2C1H3	469	N1C8H11	503	O2C3H8
402	O3N1C14H9	436	O3C8H10	470	O1C6H12	504	C6H12
403	C9H20	437	N1C6H7	471	Cl2F2O1C3H4	505	N1C9H7
404	Cl1O1C7H7	438	N1C7H9	472	Cl1O2C12H7	506	O2C6H14
405	Cl1C1H3	439	N1C5H9	473	O2C5H10	507	C12H18
406	O4C6H10	440	O2C6H12	474	S1P1O3N2C12H21	508	O2C6H12
407	Cl2C2H2	441	C7H12	475	C9H12	509	S2C4H10
408	C14H12	442	O3C11H16	476	C5H12	510	O2C8H10
409	F1C6H5	443	O2N1C15H15	477	Cl10C12	511	C6H14
410	Br1C7H7	444	O4N1C3H5	478	O1C10H14	512	O2N2C6H6
411	Br1Cl1F3C2H1	445	S1C4H10	479	O2C7H14	513	Cl4C2
412	Cl2C6H4	446	O2N1C14H9	480	C12H12	514	O1C5H12
413	C6H14	447	O1C10H14	481	F6C3	515	N1C6H7
414	N1C7H9	448	N2C6H4	482	N2C4H6	516	O2N1C4H9
415	O6C6H14	449	S1C6H14	483	O3C5H12	517	F2O3C13H8
416	Cl2O1C4H8	450	O1C5H10	484	Cl4O2C12H4	518	N1C6H15
417	O3C9H10	451	O3N1C6H5	485	N1C7H9	519	O1C8H14
418	N2C10H10	452	O1N1C7H9	486	S1C6H6	520	O2N2C6H6
419	O2C5H8	453	F3O4N4C11H13	487	Cl3C12H7	521	C2H4
420	O2C4H10	454	O1C7H16	488	Cl5C12H5	522	O1C12H10
421	O2C2H6	455	C9H10	489	N1C8H11	523	P1O4C3H9
422	F6C8H4	456	C12H12	490	O2N1C10H13	524	O1C8H8
423	O1C9H12	457	C4H6	491	O1C10H18	525	Cl4C2H2
424	O2C5H12	458	O1C6H14	492	O1C4H8	526	Cl2O4C9H8
425	Cl7C10H5	459	C4H10	493	N1H3	527	O6C9H14

528 O3C10H12
529 O1N1C6H5
530 Cl1C4H9
531 O3N3C3H3
532 N1C8H19
533 I1C6H13
534 Cl1N1C6H6
535 N1C3H7
536 Br1C4H9
537 F3O1C3H5
538 Cl4C12H6
539 N1C7H9
540 N1C5H11
541 O2C9H10
542 Cl3P1O4C4H8
543 S1N5C10H19
544 O1N1C7H7
545 F6C8H4
546 Cl4C1
547 Cl2C12H8
548 N1C6H13
549 C14H10
550 C8H16
551 C9H12
552 O2C8H16
553 N1C4H9
554 Cl1C2H3
555 O1C4H6
556 Cl3P1O4C12H14
557 C7H12
558 O2C7H14
559 O3N1C12H15
560 O1C3H6
561 I2C1H2
562 Cl7C12H3
563 Cl2O3C8H8
564 N1C1H5
565 O4N1C2H5
566 Cl8C10H6
567 O2C4H8
568 O1C8H10
569 F4C1
570 O1N1C7H7
571 Cl4C6H2
572 C6H14
573 F3O1C8H7
574 O2C4H8
575 O1N2C9H12
576 C8H18
577 O1C3H8
578 O1N1C3H9
579 Cl1O2C3H5
580 O2C12H20
581 O2C4H8
582 O2C8H14
583 O2C5H10
584 N1C7H9
585 O1C8H10
586 O1C6H12
587 O1C2H6
588 C14H10
589 O1C10H8
590 Cl1C4H9
591 Cl1O1C6H5
592 Cl1C3H5
593 C6H14
594 N1C5H13
595 Br2C6H4
596 Cl1C7H15
597 F3O4N3C13H16
598 O3N1C6H5
599 N1C7H17
600 F1O1C6H5
601 O1N1C6H13
602 O1C5H12
603 Cl1O2C12H7
604 N2H4
605 O1C7H8
606 O2N1C7H7
607 Cl3S1P1O3N1C9H11
608 O1C3H8
609 F3O3C5H9
610 O1C9H14
611 O1C11H22
612 O2N4C8H10
613 O3C9H10
614 O3N2C9H10
615 O3C6H12
616 O2C5H10
617 Cl2O1N1C9H9
618 O1C9H12
619 O1C8H8
620 N1C2H3
621 O1C6H14
622 O1N1C7H9
623 O2C5H10
624 Br3C1H1
625 Br1C6H5
626 O2C9H18
627 Cl1C6H5
628 O1C6H14
629 Cl3C2H1
630 O1C10H16
631 O2C7H6
632 S2C2H6
633 Br1C3H7
634 Br1F3C1
635 N2C3H4
636 S4P2O4C9H22
637 O2C3H6
638 Br1Cl1C2H4
639 Br1C4H9
640 N1C7H9
641 C5H10
642 F3O2N2C5H3
643 Cl3C6H3
644 F3O4N3C13H16
645 O1C7H14
646 N1C8H19
647 O3N1C4H9
648 C10H14
649 O3C11H14
650 O2C7H16
651 O2C5H10
652 O1C6H14
653 O1C5H12
654 O2C7H6
655 C5H10
656 O1N1C3H7
657 O1N1C2H5
658 O2N1C14H9
659 Cl1O1C6H5
660 Cl6O1C12H8
661 O1C3H6

662	Cl1O2N1C14H20	696	Cl1N1C6H6	730	C4H6	764	O6N2C4H8
663	C8H10	697	O1C5H10	731	C8H16	765	O1C5H12
664	O1C10H18	698	N1C7H9	732	N1C3H9	766	Cl8O2C12
665	Cl1F3C2H2	699	O2C4H8	733	C8H16	767	O1C5H12
666	Cl1N5C7H12	700	N2C6H4	734	S1O3C2H6	768	C7H16
667	O2C4H8	701	C5H8	735	N1C3H9	769	O1C10H18
668	Cl8C12H2	702	I1C4H9	736	C5H8	770	Cl2C2H2
669	N1C4H7	703	O1C6H12	737	Br2C1H2	771	C9H12
670	C3H4	704	I1O1C6H5	738	Cl2O1C6H4	772	Br1C2H5
671	Cl6C12H4	705	O3N1C6H11	739	S1P1O5N1C10H14	773	S1C6H14
672	Cl1F2C1H1	706	O1N1C6H5	740	Br1C6H13	774	F3O1C3H5
673	O2N1C2H5	707	C10H14	741	O1C8H16	776	C8H6
674	O1C10H20	708	F3C7H5	742	C11H10	777	C6N2O2H6
675	O1C6H12	709	Cl5C12H5	743	Br1C3H7	778	C16H8
676	N1C4H11	710	Cl3O2C12H5	744	S3P1O2C7H17	779	C14H12
677	O1N1C3H7	711	N1C7H9	745	C8H10	780	C15H12
678	C7H16	712	O1C4H8	746	N1C7H9	781	C11H8
679	N1C6H15	713	O1C8H8	747	Cl4C12H6	782	C16H16
680	C7H14	714	C5H8	748	N2C4H10	783	C16H16
681	O1C5H10	715	O1C5H10	749	O6C6H12	784	C16H16
682	O2C7H14	716	S2P1O6C10H19	750	Cl1O2N2C4H3	785	C16H16
683	C8H14	717	N2C5H6	751	C13H10	786	N1C5H9
684	I1C5H11	718	Br1C1H3	752	Cl6S1O3C9H6	787	N1C5H10
685	F1O1C6H5	719	C6H10	753	O1C10H20	788	C2H2
686	O1N1C5H11	720	O2N1C12H11	754	Cl2C3H6	789	C2H1
687	O1C4H8	721	Br1C8H17	755	O2C4H10	790	N1C3H7
688	S1O1C2H6	722	F3O2C3H3	756	O1C10H16	791	C3H7N1
689	O2C6H12	723	O3N1C4H9	757	O2N1C3H7	792	H5C3
690	C7H16	724	F1C1H3	758	O2N1C3H7	793	C8N1H13
691	C7H16	725	C1H4	759	S1C5H6	794	C5H5
692	Br1F4C2H1	726	S1C2H6	760	Cl1C12H9	795	C8N1H13
693	O1C7H8	727	C7H14	761	C7H8	796	C8N1H13
694	N1C6H15	728	C5H10	762	N1C4H11	797	C10N1H9
695	S1O1N1C10H21	729	O2N1C4H5	763	Cl2N1C7H3	798	N1C7H5

799	N1C7H6	833	S1H3O1	867	C8O1N1H9	902	N1C6H6Cl1
800	C14N2H22	834	S1H2F1	868	H7C3O1N1	903	C3H6
801	C8H6	835	O2C2H6N1	869	H7C3N1O2	904	C7N1O2H7
802	C14N2H24	836	O3C2H5	870	H9C5Cl1O1	905	O2C1H2
803	C14N2H24	837	O3C1H3	871	C1O1H2	906	C2O1N1H5
804	C8H8	838	O2C1H2F1	872	C4H9F1	907	H3F1C4N2
805	C9N1H13	839	Li1O1H2	873	H5N1C2O2	908	C3H8
806	C9N1H13	840	Na1O1H2	874	H11C4N1	910	C2H3F3
807	C8N1H15	841	K1O1H2	875	C3H5F3	911	H9C5F1O1
808	N1C8H15	842	Li1C6H6	876	O1H5C6Cl1	912	F1C1H4N1
809	N1C8H16	843	Na1C6H6	877	H5C3O1Cl1	913	C2H3N1
810	C9H15	844	K1C6H6	878	O1H5C6Cl1	914	C3H4
811	N1C10H15	845	C5N2H9Cl1	879	N1C6H7	915	Cl1C6H11
812	N1C10H16	846	C4N1H9O2F2	880	C6H6	916	H3N1C1O2
813	C8H9	847	C4N1H8O2F3	881	O1H2	917	C3H8N1Cl1
814	C7N1H9	848	C4N2H7Cl1	882	H3C4Cl1O1	918	C5H12
815	N1C7H9	849	C4N3H7O3	883	H5F1C6	919	H9C4O1N1
816	C7N1H9	850	C4N2H7O4Cl1	884	C1H4	920	C4N2H4
817	C7N1H9	851	C5N1H12Cl1	885	C3H8N1F1	921	C3O1H6
818	F1N1H3	852	C4N1H12Cl1	886	N1C6H6F1	922	C1O1H4
819	F1O1H2	853	C3N1H10Cl1	888	H10C5N1F1	923	N1C6H6F1
820	F2H1	854	C3N2H10O3	889	O1H5C6F1	924	C2O1H4
821	Cl1N1H3	855	C3N1H10O4Cl1	890	O1Cl1C6H5	925	N1C7H9
822	Cl1O1H2	856	C2N1H8Cl1	891	C3H6F2	926	C4O1H10
823	Cl1F1H1	857	C2N2H8O3	892	H8C5	927	C7O4N2H6
824	Cl1S1H2	858	C2N1H8O4Cl1	893	C4H8F2	928	H4C6F2
825	Cl2H1	859	C3N1H11O3S1	894	H4F2C6	929	H3C4F1N2
826	O1H4N1	860	C3N1H8O3S1F3	895	H6C4O1	930	H9C5O1F1
827	O2H3	861	Cl1C5H10N1	896	C5N1H5	931	H11C5N1
828	N4H3	862	C5N2H12	897	N2C5O1H12	932	C4H6
829	N3O1H2	863	C3H7F1	898	Cl1C5H9O1	933	N2C7O1H6
830	N3F1H1	864	N1C6H6F1	899	H4C5Cl1N1	934	H8N1C4Cl1O1
831	N3S1H2	865	H4C3N2	900	C8O5N4H12	935	N5C11H11
832	S1H4N1	866	C2H6	901	H8N1C4F1O1	936	C4H7F3

937	H4C4O1	975	H4C5F1N1	1010	H3C4Cl1N2	1045	Si1O4H12C3
938	H3C4Cl1O1	976	O2C2H4	1011	C4N2H6	1046	Si1O4H10C3
939	F1O2H3C2	977	H6N2C7	1012	H8N1F1C4O1	1047	Si1O4H8C3
940	C3O1H8	978	H2C2	1013	C3H6F2	1048	Si1O4H7N1
941	C4H10	979	N1H3	1014	O1C2H6	1049	Si1O4H8N1
943	H4C4F1N1	980	H8C3O1	1015	H7C7O1N1	1050	Si1O4H9N1C1
944	H4F1C5N1	981	N1C8H11	1016	H4C4F1N1	1051	Si1O4H10N1C1
945	C15N2O2H12	982	H5C2N1O2	1017	H10C5N1Cl1	1052	Si1O4H16N1C4
946	H10C5Cl1N1	984	H11C9O1N1	1018	C2O1H6	1053	Si1O4H4Ne1
947	Cl1C1H4N1	985	C5O1N1H11	1019	C7F3H5	1054	Si1O4H4Ar1
949	H4C5F1N1	986	O1C6H6	1020	F1H3C4N2	1055	Si1O5H5
950	C7H11F3	987	Cl1H3C4N2	1021	H9C4N1O2	1056	Si1O4H7C1
951	H4C4Cl1N1	988	O1F1C6H5	1022	N1C6H6Cl1	1057	Si1O4H3N2
952	H13C5N1	989	H4Cl1C5N1	1023	C2N1H7	1058	Si1O4H5
953	H5C3O1F1	990	O3N2C6H6	1024	C4H9F1	1059	Si1O4H9C6
954	C4N2H4	991	H11C9N1O1	1025	H5C2N1	1060	Si1O5H7C1
955	H9C4O1N1	992	H3C4F1O1	1026	H3Cl1C4N2	1061	Si1O4H9C2
956	C5N3O1H5	993	C6N1O2H5	1027	C2H4	1062	Si1O4H7C2
957	H6C3O2	994	C3N1H7O2	1029	H10C5F1N1	1063	Si1O4H5C2
958	C2H4F2	995	H9C5O1Cl1	1030	H6C2N1Cl1	1064	Si1O4H11C3
959	N2C3O1H8	996	H4C4Cl1N1	1031	C3N1O4H5	1065	Si1O4H9C3
960	O2C4N2H4	997	H4F2C6	1032	F1C6H11	1066	Si1O4H7C3
961	H3C4F1O1	998	F1C2H6N1	1033	C1N1H5	1067	Si1O4H6N1
962	C6H12	999	H9C8O1N1	1034	C4N1H5	1068	Si1O4H7N1
963	C5H7N1	1000	C2O2H4	1035	Si1O5H6	1069	Si1O4H8N1C1
964	Cl1O2H3C2	1001	N1C6H6Cl1	1036	Si1O6H4C1	1070	Si1O4H9N1C1
966	C4O1H10	1002	C6N1H13	1037	Si1O4H8C1	1071	Si1O4H15N1C4
967	C6N2O2H6	1003	C2H5F1	1038	Si1O4H4N2	1072	Si1O4H3Ne1
968	F1C5H9O1	1004	H4C5Cl1N1	1039	Si1O4H6	1073	Si1O4H3Ar1
969	O1H5C6F1	1005	H4C2O1	1040	Si1O4H10C6	1074	Al1O5H6
971	C3H7F1	1006	C4H10	1041	Si1O5H8C1	1075	Al1O6H4C1
972	F1C5H10N1	1007	C2N1H7	1042	Si1O4H10C2	1076	Al1O4H8C1
973	C3N1H9	1008	H8N1Cl1C4O1	1043	Si1O4H8C2	1077	Al1O4H4N2
974	H10C5O1	1009	O6C6N1H9	1044	Si1O4H6C2	1078	Al1O4H6

1079	Al1O4H10C6	1113	P1O4H3Ar1	1147	Si2O7H9N1	1176	Si1P1O7H5N2
1080	Al1O5H8C1	1114	Si2O8H8	1148	Si2O7H10N1C1	1177	Si1P1O7H7
1081	Al1O4H10C2	1115	Si2O9H6C1	1149	Si2O7H11N1C1	1178	Si1P1O7H11C6
1082	Al1O4H8C2	1116	Si2O7H10C1	1150	Si2O7H17N1C4	1179	Si1P1O8H9C1
1083	Al1O4H6C2	1117	Si2O7H6N2	1151	Si2O7H5Ne1	1180	Si1P1O7H11C2
1084	Al1O4H12C3	1118	Si2O7H8	1152	Si2O7H5Ar1	1181	Si1P1O7H9C2
1085	Al1O4H10C3	1119	Si2O7H12C6	1153	Si1Al1O8H8	1182	Si1P1O7H7C2
1086	Al1O4H8C3	1120	Si2O8H10C1	1154	Si1Al1O9H6C1	1183	Si1P1O7H13C3
1087	Al1O4H7N1	1121	Si2O7H12C2	1155	Si1Al1O7H10C1	1184	Si1P1O7H11C3
1088	Al1O4H8N1	1122	Si2O7H10C2	1156	Si1Al1O7H6N2	1185	Si1P1O7H9C3
1089	Al1O4H9N1C1	1123	Si2O7H8C2	1157	Si1Al1O7H8	1186	Si1P1O7H8N1
1090	Al1O4H10N1C1	1124	Si2O7H14C3	1158	Si1Al1O7H12C6	1187	Si1P1O7H9N1
1091	Al1O4H16N1C4	1125	Si2O7H12C3	1159	Si1Al1O8H10C1	1188	Si1P1O7H10N1C1
1092	Al1O4H4Ne1	1126	Si2O7H10C3	1160	Si1Al1O7H12C2	1189	Si1P1O7H11N1C1
1093	Al1O4H4Ar1	1127	Si2O7H9N1	1161	Si1Al1O7H10C2	1190	Si1P1O7H17N1C4
1094	P1O5H5	1128	Si2O7H10N1	1162	Si1Al1O7H8C2	1191	Si1P1O7H5Ne1
1095	P1O6H3C1	1129	Si2O7H11N1C1	1163	Si1Al1O7H14C3	1192	Si1P1O7H5Ar1
1096	P1O4H7C1	1130	Si2O7H12N1C1	1164	Si1Al1O7H12C3	1193	Si1P1O8H6
1097	P1O4H3N2	1131	Si2O7H18N1C4	1165	Si1Al1O7H10C3	1194	Si1P1O9H4C1
1098	P1O4H5	1132	Si2O7H6Ne1	1166	Si1Al1O7H9N1	1195	Si1P1O7H8C1
1099	P1O4H9C6	1133	Si2O7H6Ar1	1167	Si1Al1O7H10N1	1196	Si1P1O7H4N2
1100	P1O5H7C1	1134	Si2O8H7	1168	Si1Al1O7H11N1C1	1197	Si1P1O7H6
1101	P1O4H9C2	1135	Si2O7H9C1			1198	Si1P1O7H10C6
1102	P1O4H7C2	1136	Si2O7H5N2	1169	Si1Al1O7H12N1C1	1199	Si1P1O8H8C1
1103	P1O4H5C2	1137	Si2O7H7			1200	Si1P1O7H10C2
1104	P1O4H11C3	1138	Si2O7H11C6	1170	Si1Al1O7H18N1C4	1201	Si1P1O7H8C2
1105	P1O4H9C3	1139	Si2O8H9C1			1202	Si1P1O7H6C2
1106	P1O4H7C3	1140	Si2O7H11C2			1203	Si1P1O7H12C3
1107	P1O4H6N1	1141	Si2O7H9C2	1171	Si1Al1O7H6Ne1	1204	Si1P1O7H10C3
1108	P1O4H7N1	1142	Si2O7H7C2	1172	Si1Al1O7H6Ar1	1205	Si1P1O7H8C3
1109	P1O4H8N1C1	1143	Si2O7H13C3	1173	Si1P1O8H7	1206	Si1P1O7H7N1
1110	P1O4H9N1C1	1144	Si2O7H11C3	1174	Si1P1O9H5C1	1207	Si1P1O7H9N1C1
1111	P1O4H15N1C4	1145	Si2O7H9C3	1175	Si1P1O7H9C1	1208	Si1P1O7H10N1C1
1112	P1O4H3Ne1	1146	Si2O7H8N1			1209	Si1P1O7H16N1C4

1210	Si1P1O7H4Ne1	1244	Si1O4C4H10	1278	h7si1	1312	p1h4f2al1
1211	Si1P1O7H4Ar1	1245	Si1O4C1H9N1	1279	h5li1c2	1313	h5b1f1li1
1212	P1Al1O8H7	1246	Si1O4C1H10N1	1280	h4be1al1b1f1	1314	s1c1h4cl1al1
1213	P1Al1O9H5C1	1247	Si1O4C2H11N1	1281	h5o2na1	1315	h7al1
1214	P1Al1O7H9C1	1248	Si1O4C2H12N1	1282	h7be1	1316	h6na1b1
1215	P1Al1O7H5N2	1249	Si1O4C5H18N1	1283	h7b1	1317	h4b1na1be1si1
1216	P1Al1O7H7	1250	Si1O4C1H6Ne1	1284	h6be1p1	1318	h6li1p1
1217	P1Al1O7H11C6	1251	Si1O4C1H6Ar1	1285	li2h5o1	1319	o1h5c1al1
1218	P1Al1O8H9C1	1252	Si1O4C1H8	1286	h6o1be1	1320	h6be1s1
1219	P1Al1O7H11C2	1253	Si1O5C2H6	1287	h6li1b1	1321	h6cl1al1
1220	P1Al1O7H9C2	1254	Si1O3C2H10	1288	h5s1cl1be1	1322	h5n1f1li1
1221	P1Al1O7H7C2	1255	Si1O3C1H6N2	1289	h6mg1si1	1323	s1h5b1c1
1222	P1Al1O7H13C3	1256	Si1O3C1H8	1290	h5mg1p1b1	1324	h5c1o1b1
1223	P1Al1O7H11C3	1257	Si1O3C7H12	1291	h5li1f1b1	1325	h5n1be2
1224	P1Al1O7H9C3	1258	Si1O4C2H10	1292	h3c2s1f1be1	1326	cl1h4f1mg1s1
1225	P1Al1O7H8N1	1259	Si1O3C3H12	1293	h6p1o1	1327	li1h6f1
1226	P1Al1O7H9N1	1260	Si1O3C3H10	1294	li2h5b1	1328	b1h6na1
1227	P1Al1O7H10N1C1	1261	Si1O3C3H8	1295	b1h5s1be1	1329	h5li1n2
1228	P1Al1O7H11N1C1	1262	Si1O3C4H14	1296	h6s1n1	1330	h7cl1
1229	P1Al1O7H17N1C4	1263	Si1O3C4H12	1297	h5b1n2	1331	h5b1n1f1
1230	P1Al1O7H5Ne1	1264	Si1O3C4H10	1298	h6c1si1	1332	h6be1s1
1231	P1Al1O7H5Ar1	1265	Si1O3C1H9N1	1299	h5b1o1f1	1333	h7na1
1232	Si1O5C1H8	1266	Si1O3C1H10N1	1300	c1mg1h4f1o1	1334	h4s1mg1f1al1
1233	Si1O6C2H6	1267	Si1O3C2H11N1	1301	n1be1h6	1335	h5c1si1n1
1234	Si1O4C2H10	1268	Si1O3C2H12N1	1302	h6f1b1	1336	h6o1mg1
1235	Si1O4C1H6N2	1269	Si1O3C5H18N1	1303	h2si1li2n1al1o1	1337	h6al1n1
1236	Si1O4C1H8	1270	Si1O3C1H6Ne1	1304	h5b1f1cl1	1338	n1h5b1al1
1237	Si1O4C7H12	1271	Si1O3C1H6Ar1	1305	h5li1f1c1	1339	h6b1n1
1238	Si1O5C2H10	1272	h6c1o1	1306	h4b1c1mg1li1	1340	h6n1f1
1239	Si1O4C3H12	1273	h5o1p1n1	1307	h5li1na1c1	1341	h5be1f1c1
1240	Si1O4C3H10	1274	n1li1b2h3c1	1308	si1h5c1mg1	1342	h5al1c1o1
1241	Si1O4C3H8	1275	h6s1be1	1309	h4c1o1be1n1	1343	h5na1o1si1
1242	Si1O4C4H14	1276	be1h5n1cl1	1310	h4p1b1c1cl1	1344	c1h5b1be1
1243	Si1O4C4H12	1277	h4al1b1be1n1	1311	be2li1n2h2cl1	1345	li1si1be1b2h3

1346	h4n1f1c1o1	1380	h6n2	1414	h5c2n1	1448	si1h4
1347	h6li1b1	1381	c3h4mg1	1415	h5al1cl1p1	1449	p2
1348	h4f1n2b1	1382	n1c3h4	1416	b2h3c2p1	1450	s2
1349	o1al1h5be1	1383	h3c3o1p1	1417	h4c2na1al1	1451	cl2
1350	h7f1	1384	h3mg1c2o1li1	1418	h5c1si1f1	1452	C9Br1H14N1
1351	h4be1c1b1li1	1385	al1n1h5c1	1419	c2s1h3b1si1	1453	Cl1H5O1C1
1352	h6li1p1	1386	o1h5n1c1	1420	c1h5n1s1	1454	H6C7I1F3
1353	n1h5c2	1387	p1h4o1al1f1	1421	n1al1c2h4	1455	C2F3H5
1354	c3f1h3li1	1388	c2f1h5	1422	c2h5n1	1456	C9H11Cl1O1
1355	h6c2	1389	h5al1o2	1423	c1h5n1o1	1457	C2I1F3O1H2
1356	h5c2o1	1390	h5f1o1p1	1424	si1cl2o1h4	1458	C2Br1F3O1H2
1357	h7b1	1391	h6s1li1	1425	h3c3f1si1	1459	Cl1H6N1C1
1358	h4mg1c1b1al1	1392	h5c2al1	1426	al2li1h3c2	1460	H6C7Br1F3
1359	c3h3n1s1	1393	p1h3c1o2si1	1427	h5si2cl1	1461	C1Cl2H4
1360	h4n1c2b1	1394	h6c1n1	1428	h4c4	1462	C2Cl1H7
1361	h7c1	1395	c2si1p1h4	1429	c2h4si1o1	1463	C1I2H4
1362	c1h4b1o2	1396	li1h4c1cl1s1	1430	h5mg1p1o1	1464	F6C12H6
1363	h6c1mg1	1397	h5o1b1c1	1431	h5cl2p1	1465	C1H4F2
1364	h5p1si1n1	1398	si1h6c1	1432	h5n2cl1	1466	C7Br1H9S1
1365	c1s1h5o1	1399	p1c1h5mg1	1433	h6c1si1	1467	O2H3C1F3
1366	p1h3c3mg1	1400	c5h2o1	1434	o1c2s1h4	1468	C9H11Br1O1
1367	h5c2p1	1401	h4s1c2si1	1435	c1h5o1n1	1469	C9Cl1H14N1
1368	h5s1c2	1402	h5c2p1	1436	c3h3o1li1	1470	C2Cl2H6
1369	s1h6o1	1403	c2f1h4b1	1437	h2	1471	H7O1C2Cl1
1370	h6c1o1	1404	h6p1f1	1438	li1h1	1472	C2F2H6
1371	h5c2b1	1405	n2c1si1h4	1439	be1h2	1473	F3H9C12
1372	h7o1	1406	p1cl1n2h4	1440	b1h3	1474	C2Br1H5O1
1373	h5c1mg1o1	1407	h5o1c1b1	1441	c1h4	1475	H9C7I1
1374	h4o1c2n1	1408	h6c1si1	1442	n2	1476	C2Cl1H5O1
1375	c2h4si1li1	1409	h6na1n1	1443	o2	1477	H7O1C2F1
1376	c2h5n1	1410	h5n1c2	1444	f2	1478	C1H4Br2
1377	h4p1c2f1	1411	h5o1b1c1	1445	na1h1	1479	H9C7Br1
1378	h5c1n2	1412	h6s1c1	1446	mg1h2	1480	C2Cl1F3O1H2
1379	h4s1p1c1n1	1413	h4o1s1c2	1447	al1h3	1481	C7I1H9S1

1482	F1H6N1C1	1516	N2C10H10	1550	C2H4O1	1584	C2N1H8O1
1483	F1H5O1C1	1517	N4H8C8O4	1551	C4H6O2	1585	C2N1H8O1
1484	Br1H5O1C1	1518	C11H11N1	1552	C7H16O2	1586	C2N1H8O1
1485	C9I1H14N1	1519	C10H10N2O2	1553	C7H17O1N1	1587	C2N1H8O1
1486	I1H5O1C1	1520	N3C9H9O2	1554	C8H10O2	1588	C2N1H8O1
1487	C2I1H5O1	1521	C8H10	1555	C5H11O1N1	1589	C2N1H8O1
1488	C2F1H7	1522	N2H8C6O2	1556	N1C7H7	1590	C2N2H11
1489	C2Cl3H5	1523	N2H6C6O2	1557	N2H10C6	1591	C2N2H11
1490	O2H3C1Cl3	1524	N1C7H9	1558	C3O3H7	1592	C2N2H11
1491	C9H11I1O1	1525	C10H24	1559	C3O3H7	1593	C2N2H11
1492	O2H4	1526	C10H24	1560	C3O3H7	1594	C2N2H11
1493	O2H6C1	1527	C10H24	1561	C3O3H7	1595	C2N2H11
1494	O1H7N1C1	1528	C10H22	1562	C3O3H7	1596	C2N2H11
1495	O2H9C3N1	1529	C10H20	1563	C3O3H7	1597	C2N2H11
1496	O2H8C2	1530	C11H16	1564	C3O3H7	1598	C2N1H10O1
1497	O1H9C2N1	1531	C11H18	1565	C3O3H7	1599	C2N1H10O1
1498	O2H11C4N1	1532	N2H16C9O2	1566	C2O3H5	1600	C2N1H10O1
1499	O2H6C1	1533	N2H14C9O2	1567	C2O3H5	1601	C2N1H10O1
1500	N1H9C2O1	1534	N2H16C9O2	1568	C2O3H5	1602	C2N1H10O1
1501	N2H10C2	1535	C7H16	1569	C2O3H5	1603	C2N1H10O1
1502	N2H12C4O1	1536	C7H14	1570	C2O3H5	1604	C2N1H10O1
1503	N1H7C1O1	1537	C8H19O1N1	1571	C2O3H5	1605	C2N1H10O1
1504	C4H11O2N1	1538	C12H12	1572	C2O3H5	1606	C1N1H8O1
1505	C4H12O1N2	1539	N2C10H10	1573	C2O3H5	1607	C1N1H8O1
1506	C6H14O2N2	1540	C11H11N1	1574	C3O2H8N1	1608	C1N1H8O1
1507	C3H9O2N1	1541	C8H8	1575	C3O2H8N1	1609	C1N1H8O1
1508	N4H8C8O4	1542	C4H4	1576	C3O2H8N1	1610	C1N1H8O1
1509	O1H7N1C5	1543	C8H10O2	1577	C3O2H8N1	1611	C1N1H8O1
1510	O1H9C6N1	1544	C8H11O1N1	1578	C3O2H8N1	1612	C1N1H8O1
1511	C4O4H8	1545	C6H8O1	1579	C3O2H8N1	1613	C1N1H8O1
1512	C4O2N2H10	1546	C7H10O1	1580	C3O2H8N1	1614	C2N3H8O1
1513	C6O4H8N2	1547	C7H11N1	1581	C3O2H8N1	1615	C2N3H8O1
1514	C6O3N3H9	1548	C9H13O1N1	1582	C2N1H8O1	1616	C2N3H8O1
1515	C12H12	1549	N2C10H10	1583	C2N1H8O1	1617	C2N3H8O1

1618	C2N3H8O1	1652	C4N2H7O1	1686	N1H3	1720	C5H12O1
1619	C2N3H8O1	1653	C4N2H7O1	1687	C5H9CL1	1721	C4H7N1
1620	C2N3H8O1	1654	C4N3H10	1688	C4H9N1O2	1722	C5O1H6
1621	C2N3H8O1	1655	C4N3H10	1689	C4H8O2	1723	C4O1H8
1622	C2N4H11	1656	C4N3H10	1690	C1H4O1	1724	C7H8
1623	C2N4H11	1657	C4N3H10	1691	C6H12O2	1725	C6H14
1624	C2N4H11	1658	C4N3H10	1692	C6N1H7	1726	C7O1N2H16
1625	C2N4H11	1659	C4N3H10	1693	C7H14	1727	C6H10
1626	C2N4H11	1660	C4N3H10	1694	C6N1H9	1728	C6H14
1627	C2N4H11	1661	C4N3H10	1695	C7H10O1	1729	C5H12O2
1628	C2N4H11	1662	C4N2H9O1	1696	C7H14	1730	C5H10
1629	C2N4H11	1663	C4N2H9O1	1697	C6H12	1731	C7H12
1630	C2N3H10O1	1664	C4N2H9O1	1698	C5H9N1O1	1732	C3H8S1
1631	C2N3H10O1	1665	C4N2H9O1	1699	C7S1H10	1733	C3H9N1
1632	C2N3H10O1	1666	C4N2H9O1	1700	N1C5H11O1	1734	C4H10S1O2
1633	C2N3H10O1	1667	C4N2H9O1	1701	C4H10O2	1735	C5N1H7
1634	C2N3H10O1	1668	C4N2H9O1	1702	C4N2H6	1736	C5H10
1635	C2N3H10O1	1669	C4N2H9O1	1703	C5H8	1737	C4H10O2
1636	C2N3H10O1	1670	C3N2H7O1	1704	C6H10O2	1738	C7H10
1637	C2N3H10O1	1671	C3N2H7O1	1705	C1H2O1	1739	C4H10S1O1
1638	C1N3H8O1	1672	C3N2H7O1	1706	C5H12N2O1	1740	C6O2H10
1639	C1N3H8O1	1673	C3N2H7O1	1707	CL1H1	1741	C6H12
1640	C1N3H8O1	1674	C3N2H7O1	1708	C6H14	1742	C1O1
1641	C1N3H8O1	1675	C3N2H7O1	1709	C4O1H9N1	1743	C7H4
1642	C1N3H8O1	1676	C3N2H7O1	1710	C6H10O2	1744	C6H14S2
1643	C1N3H8O1	1677	C3N2H7O1	1711	C6H10O2	1745	C5H10O2
1644	C1N3H8O1	1678	C6O1H12	1712	C4H10S1	1746	H2
1645	C1N3H8O1	1679	C4H11N1	1713	C5H8O1	1747	C1CL2O1
1646	C4N2H7O1	1680	N1H11C4	1714	O2H2	1748	C4N2H6
1647	C4N2H7O1	1681	C7H10	1715	C6H9N1	1749	C4H10O3S1
1648	C4N2H7O1	1682	S1H2	1716	C1H1N1	1750	C4O2H8
1649	C4N2H7O1	1683	C4H9O2N1	1717	C1O2	1751	C5H10O1
1650	C4N2H7O1	1684	CL2	1718	C3O1H7N1	1752	C7H4
1651	C4N2H7O1	1685	C5H8	1719	C6O1H12	1753	C4H7N1O1

1754	C5H10	1783	C17N8B1	1812	B2C1N3O6H1	1841	C9B1N8O16H1
1755	C9H20	1784	C9N8O16B1	1813	B2C1N3O6H1	1842	C9B1N8O16H1
1756	C4H10O2	1785	C1B2H4	1814	B2C1N3O6H1	1843	C9B1N8O16H1
1757	O1H2	1786	C1B2F3H1	1815	C5F5H1	1844	C9B1N8O16H1
1758	C9H20	1787	B2C4N3H1	1816	C10N5H1	1845	C9B1N8O16H1
1759	C3O2H6	1788	B2C1N3O6H1	1817	C5B1F6H1	1846	C9B1N8O16H1
1760	C6H14	1789	C5H6	1818	C5B1F6H1	1847	C9H9B1
1761	C7H12	1790	C5F5H1	1819	C11B1N6H1	1848	C9H9B1
1762	C5H8	1791	C10N5H1	1820	C11B1N6H1	1849	C9H9B1
1763	C5O1H10	1792	C5N5O10H1	1821	C11B1N6H1	1850	C9F8B1H1
1764	C2H4	1793	C5H7B1	1822	C11B1N6H1	1851	C9F8B1H1
1765	C1B2H3	1794	C5B1F6H1	1823	C5B1N6O12H1	1852	C9F8B1H1
1766	C1B2F3	1795	C11B1N6H1	1824	C5B1N6O12H1	1853	C9F8B1H1
1767	B2C4N3	1796	C5B1N6O12H1	1825	C5B1N6O12H1	1854	C9F8B1H1
1768	B2C1N3O6	1797	C9H9B1	1826	C5B1N6O12H1	1855	C17N8B1H1
1769	C5H5	1798	C9B1F8H1	1827	C9H9B1	1856	C17N8B1H1
1770	C5F5	1799	C17B1N8H1	1828	C9H9B1	1857	C17N8B1H1
1771	C10N5	1800	C9B1N8O16H1	1829	C9H9B1	1858	C17N8B1H1
1772	C5N5O10	1801	C9H9B1	1830	C9B1F8H1	1859	C17N8B1H1
1773	C5H6B1	1802	C9F8B1H1	1831	C9B1F8H1	1860	C17N8B1H1
1774	C5B1F6	1803	C17N8B1H1	1832	C9B1F8H1	1861	C17N8B1H1
1775	C11B1N6	1804	C9N8O16B1H1	1833	C9B1F8H1	1862	C9N8O16B1H1
1776	C5B1N6O12	1805	C1B2H4	1834	C9B1F8H1	1863	C9N8O16B1H1
1777	C9H8B1	1806	C1B2F3H1	1835	C9B1F8H1	1864	C9N8O16B1H1
1778	C8B1F8	1807	C1B2F3H1	1836	C17B1N8H1	1865	C9N8O16B1H1
1779	C17B1N8	1808	B2C4N3H1	1837	C17B1N8H1	1866	C9N8O16B1H1
1780	C9B1N8O16	1809	B2C4N3H1	1838	C17B1N8H1		
1781	C9H8B1	1810	B2C4N3H1	1839	C17B1N8H1		
1782	C9F8B1	1811	B2C4N3H1	1840	C17B1N8H1		

B. APPENDIX : Structure Figures

Molecular formulas in Appendix A is not meaningful at all for intermolecular interactions. The structures used in Chapter 4 generated with PyMol are provided in this section. The color legend for atoms can be found at https://pymolwiki.org/index.php/Color_Values

ABH21

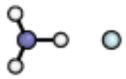

Figure 5: m818

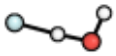

Figure 6: m819

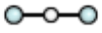

Figure 7: m820

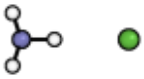

Figure 8: m821

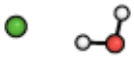

Figure 9: m822

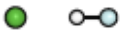

Figure 10: m823

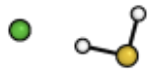

Figure 11: m824

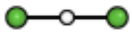

Figure 12: m825

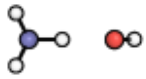

Figure 13: m826

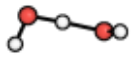

Figure 14: m827

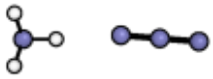

Figure 15: m828

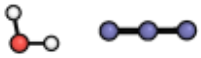

Figure 16: m829

Figure 17: m830

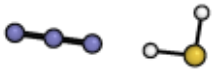

Figure 18: m831

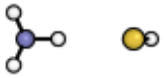

Figure 19: m832

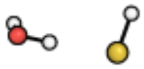

Figure 20: m833

Figure 21: m834

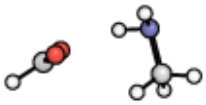

Figure 22: m835

Figure 23: m836

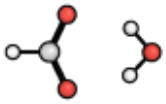

Figure 24: m837

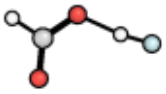

Figure 25: m838

IL16

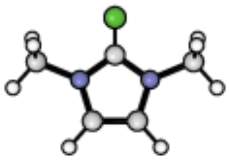

Figure 26: m845

Figure 27: m846

Figure 28: m847

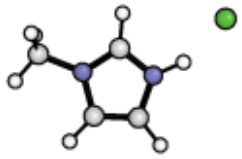

Figure 29: m848

Figure 30: m849

Figure 31: m850

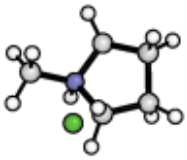

Figure 32: m851

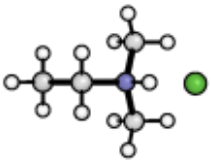

Figure 33: m852

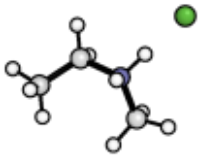

Figure 34: m853

Figure 35: m854

Figure 36: m855

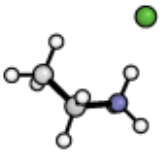

Figure 37: m856

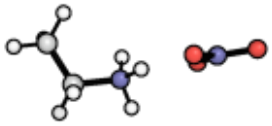

Figure 38: m857

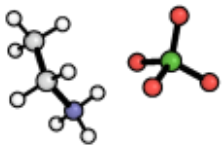

Figure 39: m858

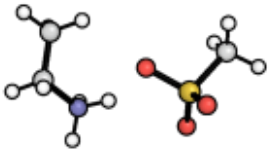

Figure 40: m859

Figure 41: m860

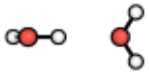

Figure 42: m1492

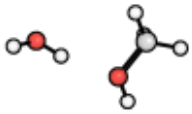

Figure 43: m1493

Figure 44: m1494

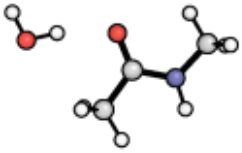

Figure 45: m1495

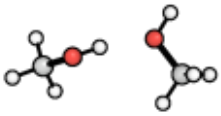

Figure 46: m1496

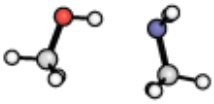

Figure 47: m1497

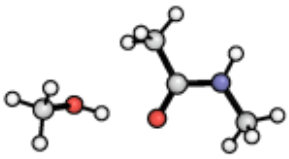

Figure 48: m1498

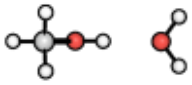

Figure 49: m1499

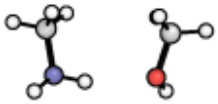

Figure 50: m1500

Figure 51: m1501

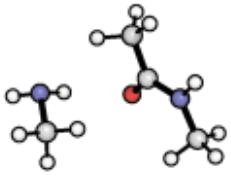

Figure 52: m1502

Figure 53: m1503

Figure 54: m1504

Figure 55: m1505

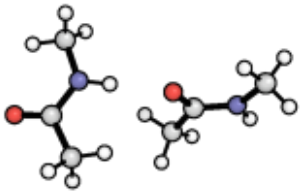

Figure 56: m1506

Figure 57: m1507

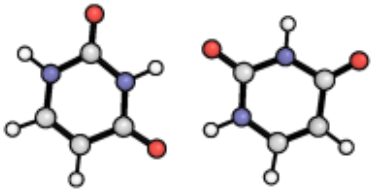

Figure 58: m1508

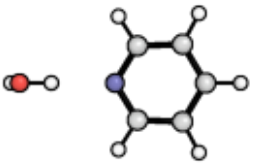

Figure 59: m1509

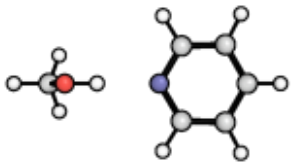

Figure 60: m1510

Figure 61: m1511

Figure 62: m1512

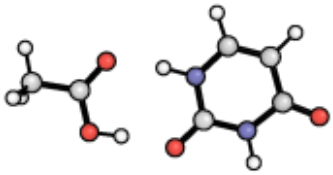

Figure 63: m1513

Figure 64: m1514

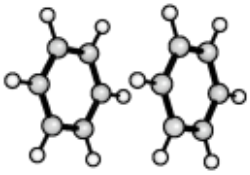

Figure 65: m1515

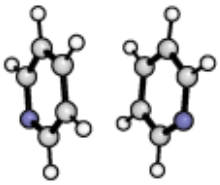

Figure 66: m1516

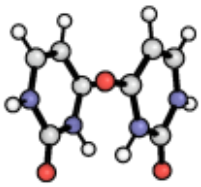

Figure 67: m1517

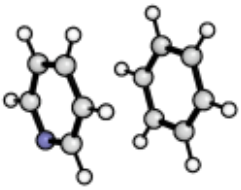

Figure 68: m1518

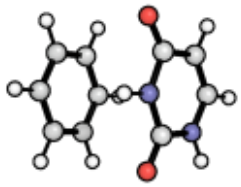

Figure 69: m1519

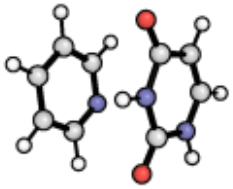

Figure 70: m1520

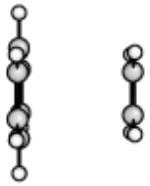

Figure 71: m1521

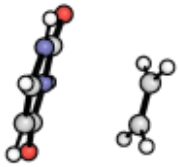

Figure 72: m1522

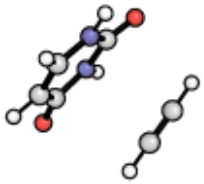

Figure 73: m1523

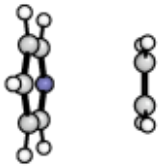

Figure 74: m1524

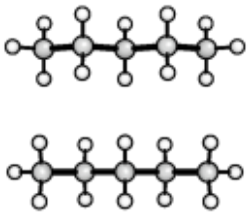

Figure 75: m1525

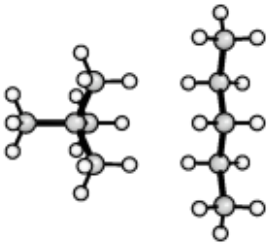

Figure 76: m1526

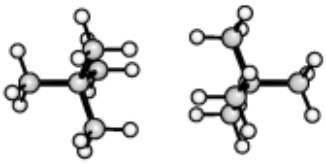

Figure 77: m1527

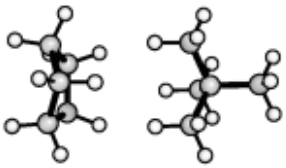

Figure 78: m1528

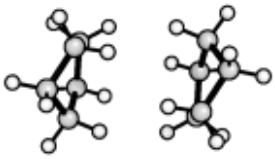

Figure 79: m1529

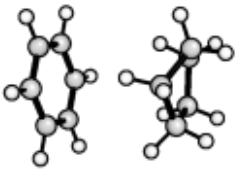

Figure 80: m1530

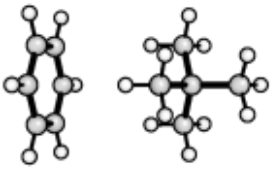

Figure 81: m1531

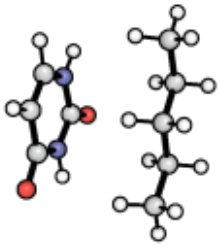

Figure 82: m1532

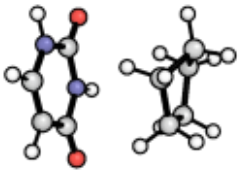

Figure 83: m1533

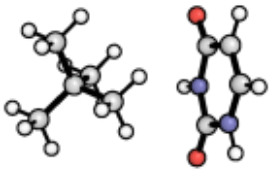

Figure 84: m1534

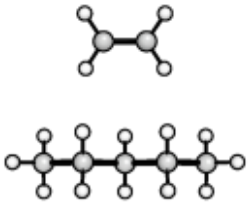

Figure 85: m1535

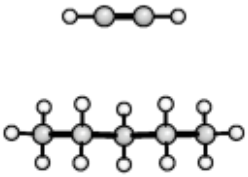

Figure 86: m1536

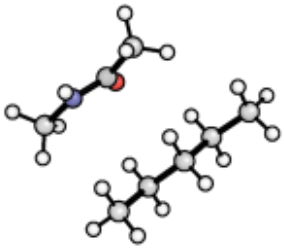

Figure 87: m1537

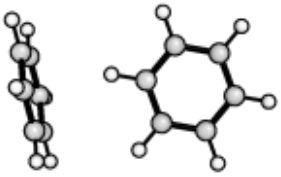

Figure 88: m1538

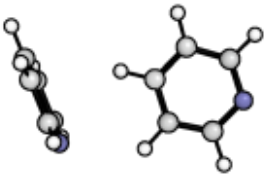

Figure 89: m1539

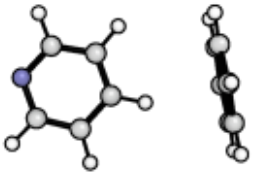

Figure 90: m1540

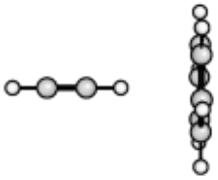

Figure 91: m1541

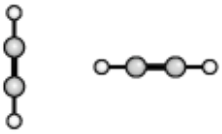

Figure 92: m1542

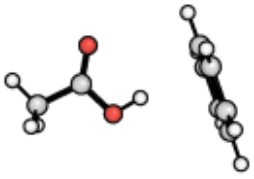

Figure 93: m1543

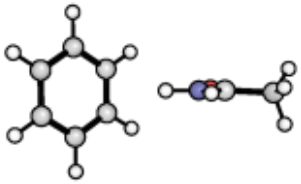

Figure 94: m1544

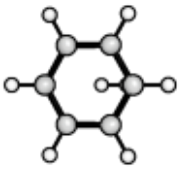

Figure 95: m1545

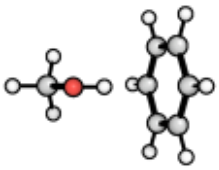

Figure 96: m1546

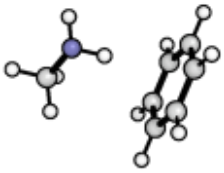

Figure 97: m1547

Figure 98: m1548

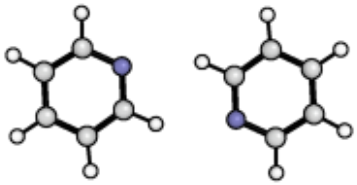

Figure 99: m1549

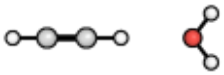

Figure 100: m1550

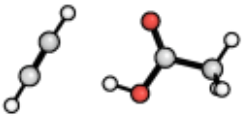

Figure 101: m1551

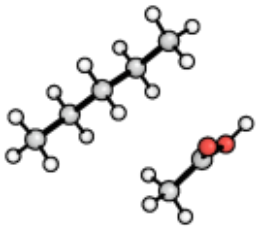

Figure 102: m1552

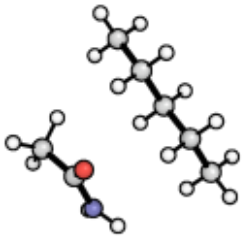

Figure 103: m1553

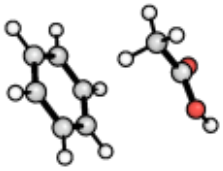

Figure 104: m1554

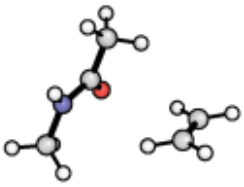

Figure 105: m1555

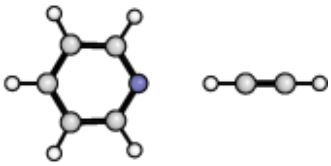

Figure 106: m1556

Figure 107: m1557

ZG237

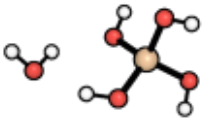

Figure 108: m1035

Figure 109: m1036

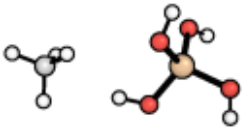

Figure 110: m1037

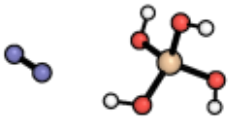

Figure 111: m1038

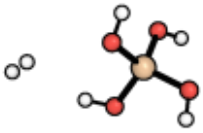

Figure 112: m1039

Figure 113: m1040

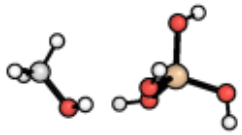

Figure 114: m1041

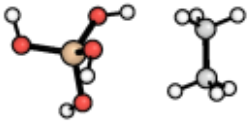

Figure 115: m1042

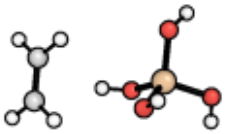

Figure 116: m1043

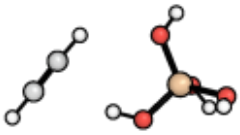

Figure 117: m1044

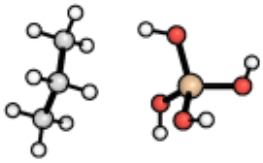

Figure 118: m1045

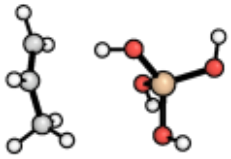

Figure 119: m1046

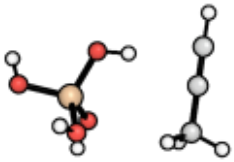

Figure 120: m1047

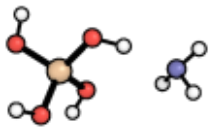

Figure 121: m1048

Figure 122: m1049

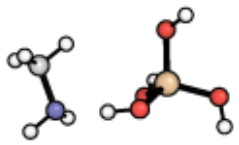

Figure 123: m1050

Figure 124: m1051

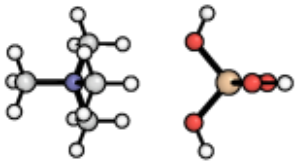

Figure 125: m1052

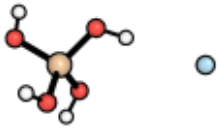

Figure 126: m1053

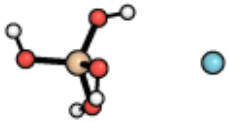

Figure 127: m1054

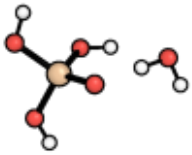

Figure 128: m1055

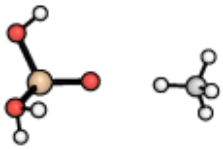

Figure 129: m1056

Figure 130: m1057

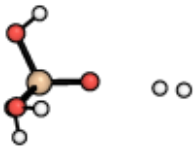

Figure 131: m1058

Figure 132: m1059

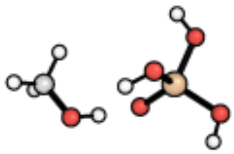

Figure 133: m1060

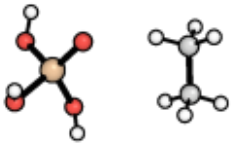

Figure 134: m1061

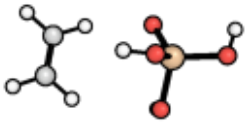

Figure 135: m1062

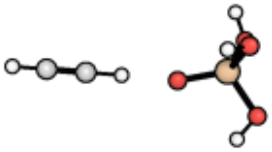

Figure 136: m1063

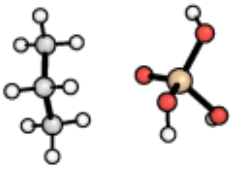

Figure 137: m1064

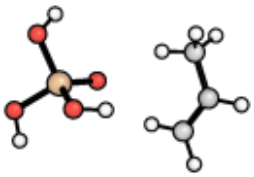

Figure 138: m1065

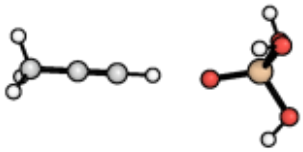

Figure 139: m1066

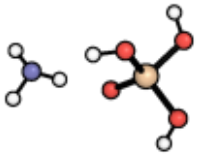

Figure 140: m1067

Figure 141: m1068

Figure 142: m1069

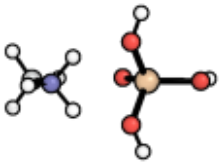

Figure 143: m1070

Figure 144: m1071

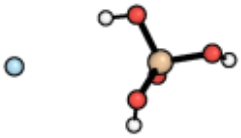

Figure 145: m1072

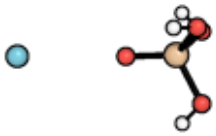

Figure 146: m1073

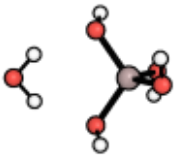

Figure 147: m1074

Figure 148: m1075

Figure 149: m1076

Figure 150: m1077

Figure 151: m1078

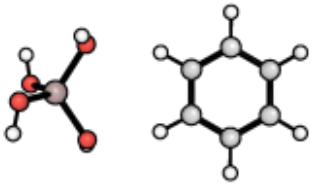

Figure 152: m1079

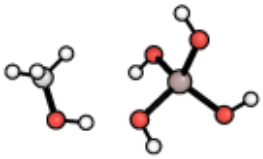

Figure 153: m1080

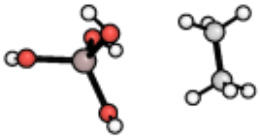

Figure 154: m1081

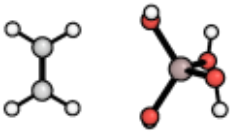

Figure 155: m1082

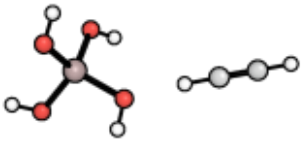

Figure 156: m1083

Figure 157: m1084

Figure 158: m1085

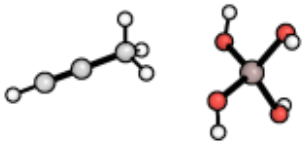

Figure 159: m1086

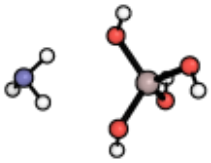

Figure 160: m1087

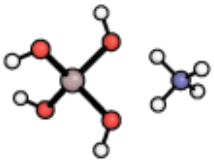

Figure 161: m1088

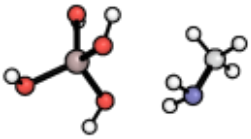

Figure 162: m1089

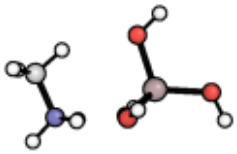

Figure 163: m1090

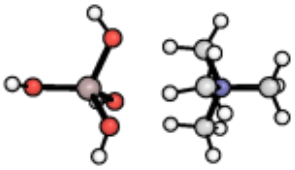

Figure 164: m1091

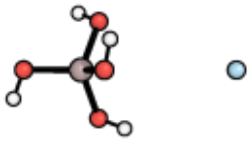

Figure 165: m1092

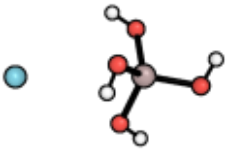

Figure 166: m1093

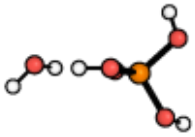

Figure 167: m1094

Figure 168: m1095

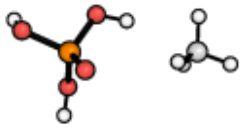

Figure 169: m1096

Figure 170: m1097

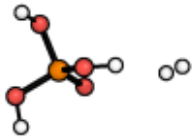

Figure 171: m1098

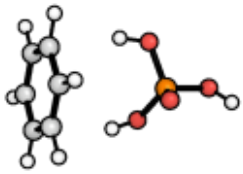

Figure 172: m1099

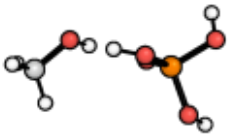

Figure 173: m1100

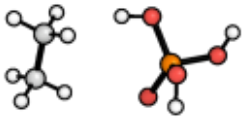

Figure 174: m1101

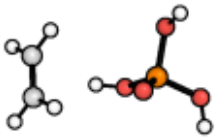

Figure 175: m1102

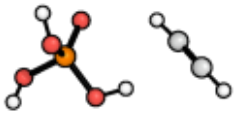

Figure 176: m1103

Figure 177: m1104

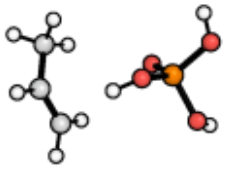

Figure 178: m1105

Figure 179: m1106

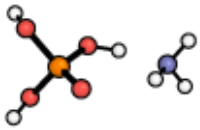

Figure 180: m1107

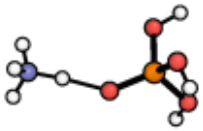

Figure 181: m1108

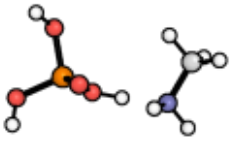

Figure 182: m1109

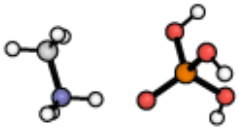

Figure 183: m1110

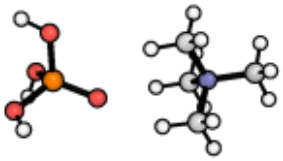

Figure 184: m1111

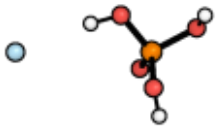

Figure 185: m1112

Figure 186: m1113

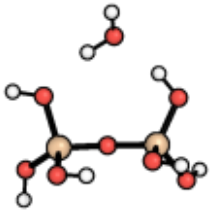

Figure 187: m1114

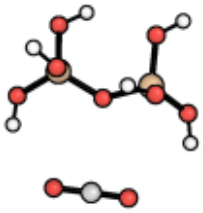

Figure 188: m1115

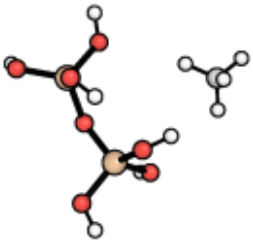

Figure 189: m1116

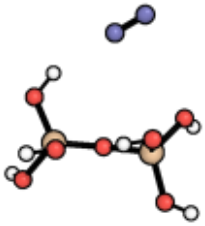

Figure 190: m1117

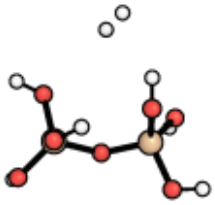

Figure 191: m1118

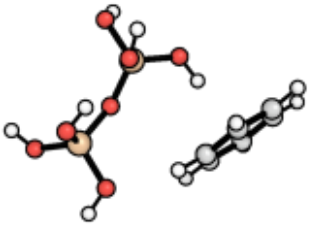

Figure 192: m1119

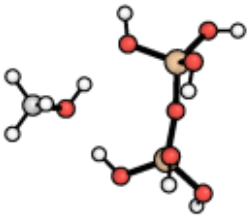

Figure 193: m1120

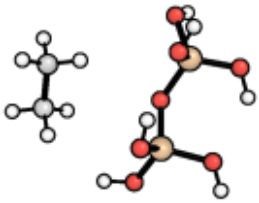

Figure 194: m1121

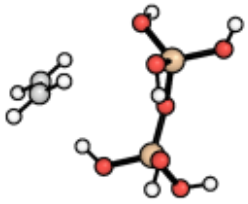

Figure 195: m1122

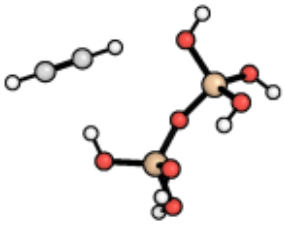

Figure 196: m1123

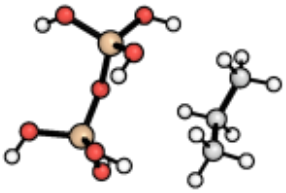

Figure 197: m1124

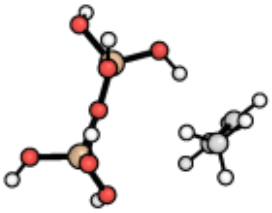

Figure 198: m1125

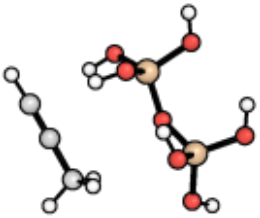

Figure 199: m1126

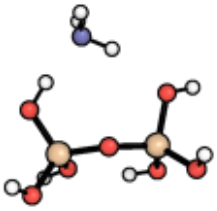

Figure 200: m1127

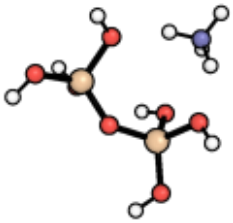

Figure 201: m1128

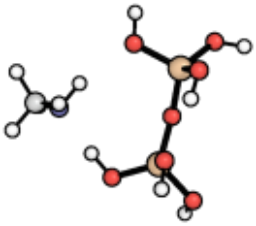

Figure 202: m1129

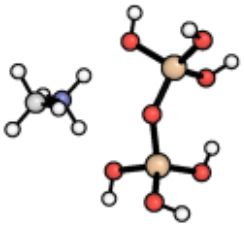

Figure 203: m1130

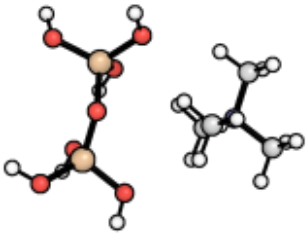

Figure 204: m1131

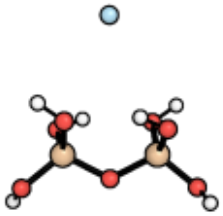

Figure 205: m1132

Figure 206: m1133

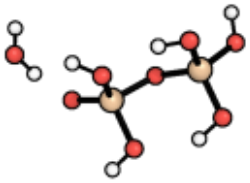

Figure 207: m1134

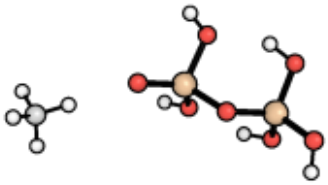

Figure 208: m1135

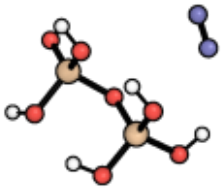

Figure 209: m1136

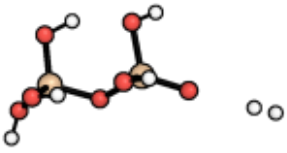

Figure 210: m1137

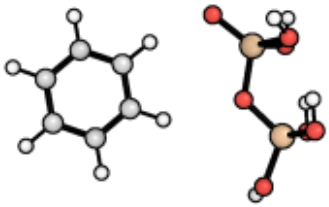

Figure 211: m1138

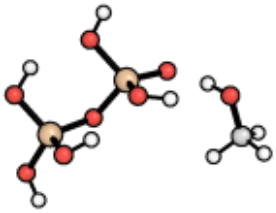

Figure 212: m1139

Figure 213: m1140

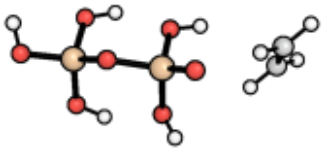

Figure 214: m1141

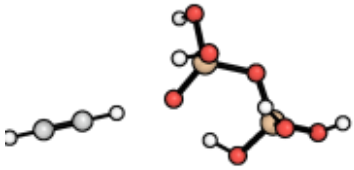

Figure 215: m1142

Figure 216: m1143

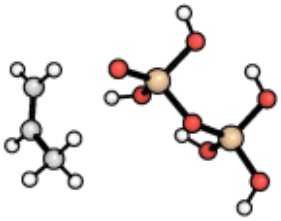

Figure 217: m1144

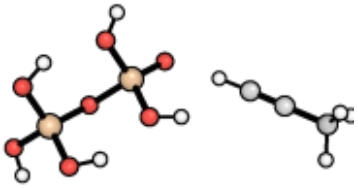

Figure 218: m1145

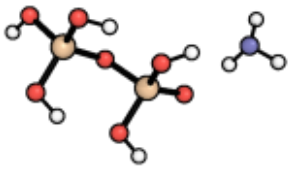

Figure 219: m1146

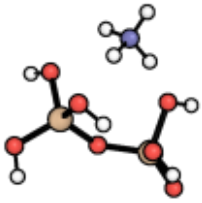

Figure 220: m1147

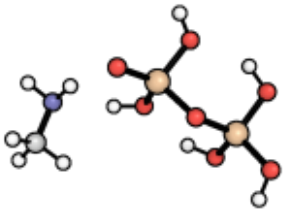

Figure 221: m1148

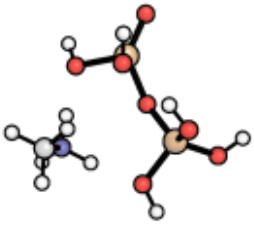

Figure 222: m1149

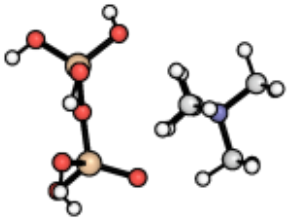

Figure 223: m1150

Figure 224: m1151

Figure 225: m1152

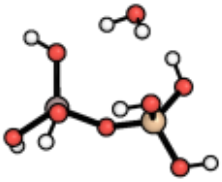

Figure 226: m1153

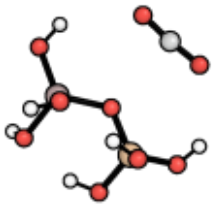

Figure 227: m1154

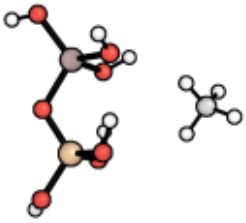

Figure 228: m1155

Figure 229: m1156

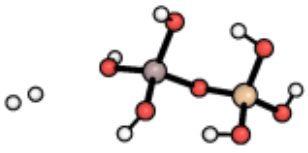

Figure 230: m1157

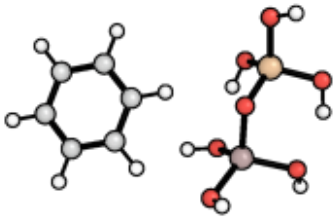

Figure 231: m1158

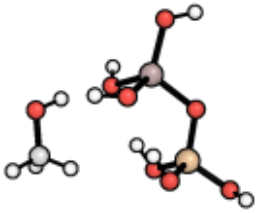

Figure 232: m1159

Figure 233: m1160

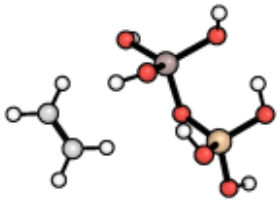

Figure 234: m1161

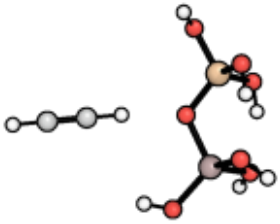

Figure 235: m1162

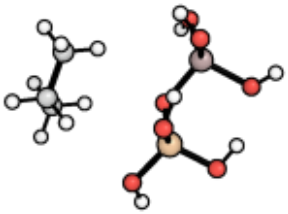

Figure 236: m1163

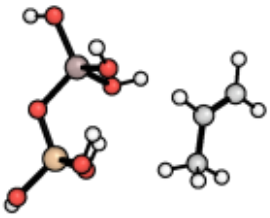

Figure 237: m1164

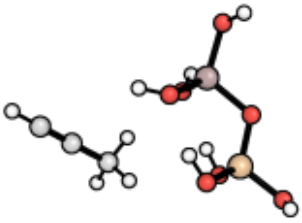

Figure 238: m1165

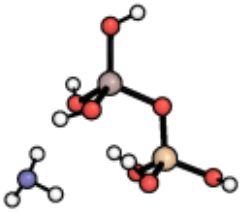

Figure 239: m1166

Figure 240: m1167

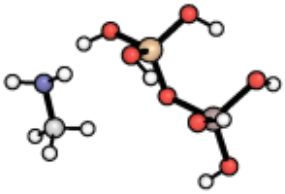

Figure 241: m1168

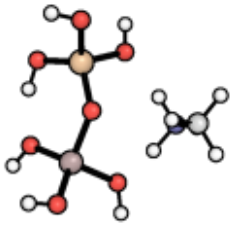

Figure 242: m1169

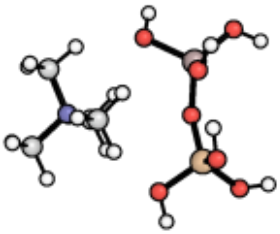

Figure 243: m1170

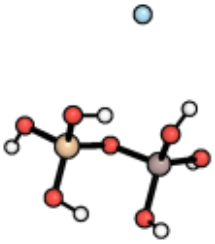

Figure 244: m1171

Figure 245: m1172

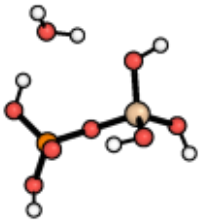

Figure 246: m1173

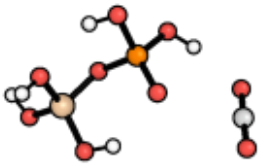

Figure 247: m1174

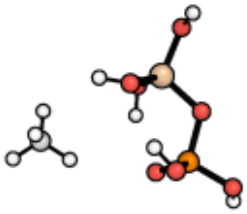

Figure 248: m1175

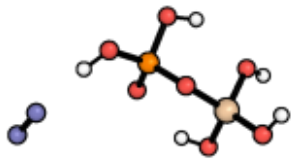

Figure 249: m1176

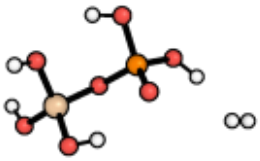

Figure 250: m1177

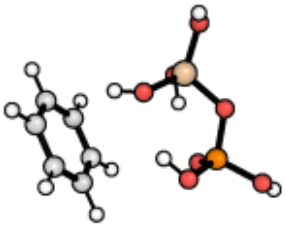

Figure 251: m1178

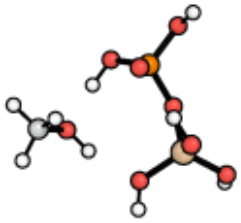

Figure 252: m1179

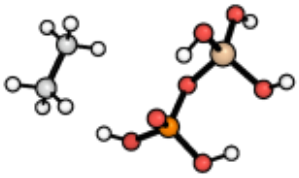

Figure 253: m1180

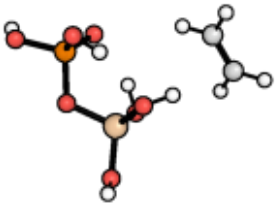

Figure 254: m1181

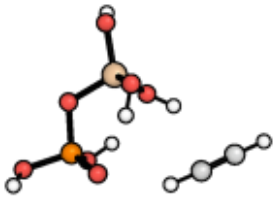

Figure 255: m1182

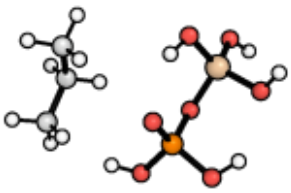

Figure 256: m1183

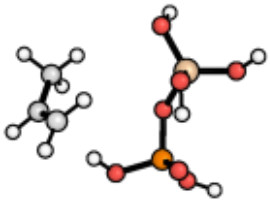

Figure 257: m1184

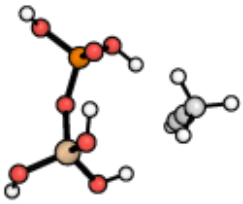

Figure 258: m1185

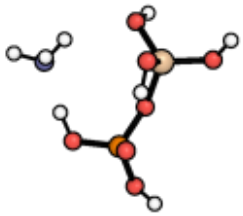

Figure 259: m1186

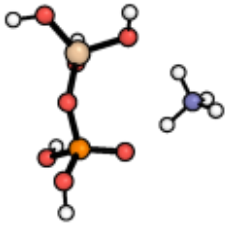

Figure 260: m1187

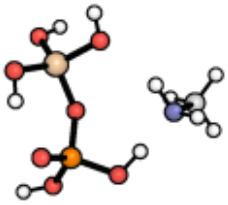

Figure 261: m1188

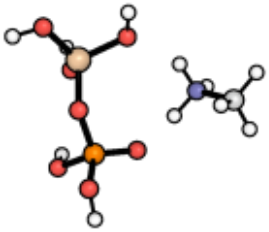

Figure 262: m1189

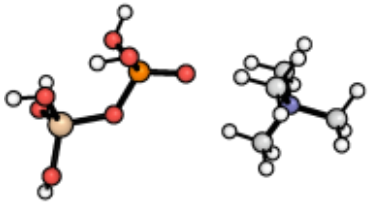

Figure 263: m1190

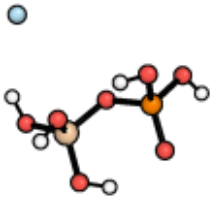

Figure 264: m1191

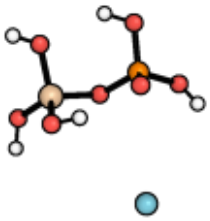

Figure 265: m1192

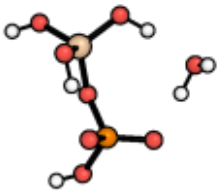

Figure 266: m1193

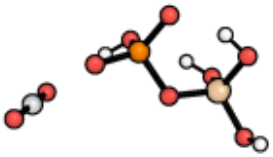

Figure 267: m1194

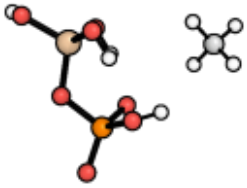

Figure 268: m1195

Figure 269: m1196

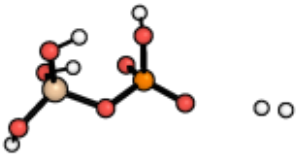

Figure 270: m1197

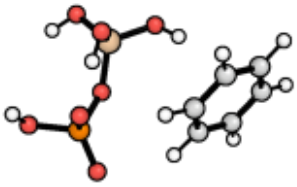

Figure 271: m1198

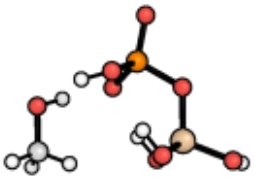

Figure 272: m1199

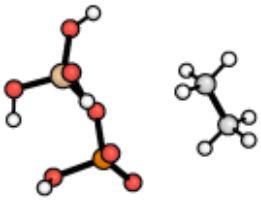

Figure 273: m1200

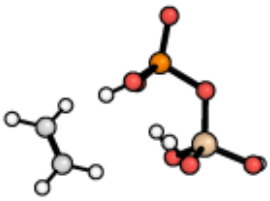

Figure 274: m1201

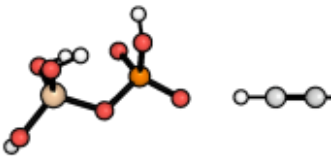

Figure 275: m1202

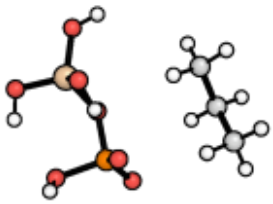

Figure 276: m1203

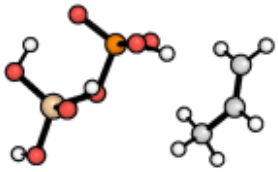

Figure 277: m1204

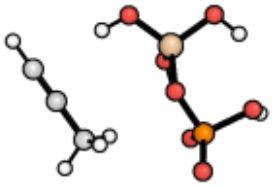

Figure 278: m1205

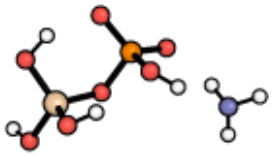

Figure 279: m1206

Figure 280: m1207

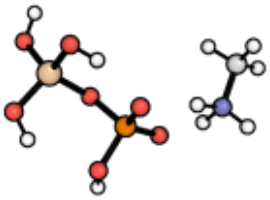

Figure 281: m1208

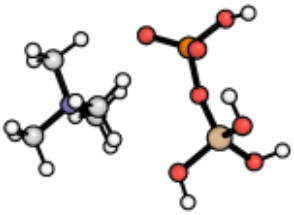

Figure 282: m1209

Figure 283: m1210

Figure 284: m1211

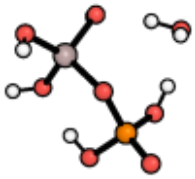

Figure 285: m1212

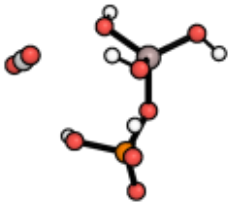

Figure 286: m1213

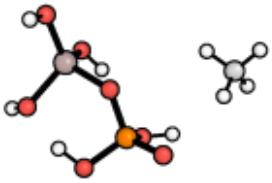

Figure 287: m1214

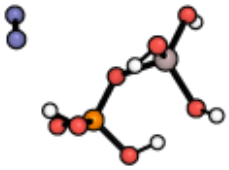

Figure 288: m1215

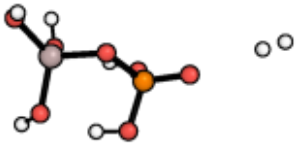

Figure 289: m1216

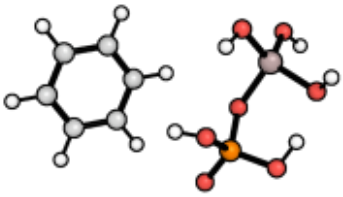

Figure 290: m1217

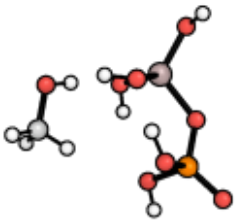

Figure 291: m1218

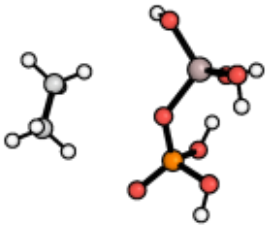

Figure 292: m1219

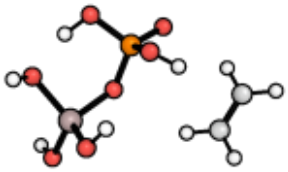

Figure 293: m1220

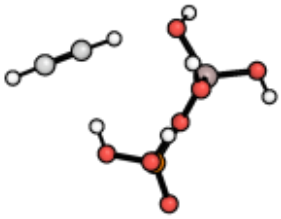

Figure 294: m1221

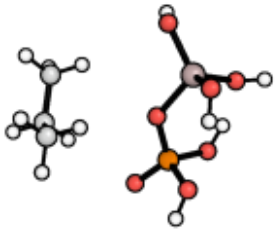

Figure 295: m1222

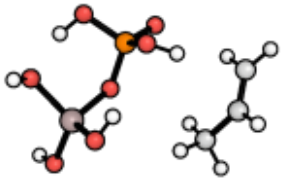

Figure 296: m1223

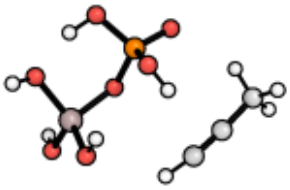

Figure 297: m1224

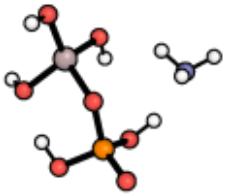

Figure 298: m1225

Figure 299: m1226

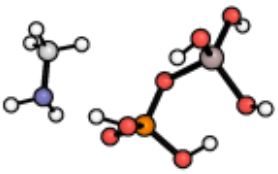

Figure 300: m1227

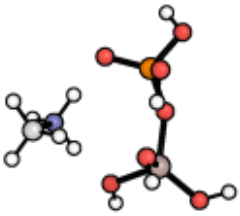

Figure 301: m1228

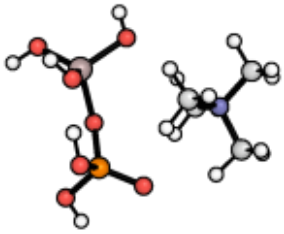

Figure 302: m1229

Figure 303: m1230

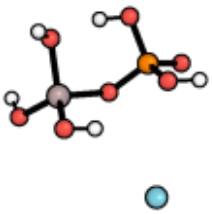

Figure 304: m1231

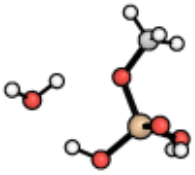

Figure 305: m1232

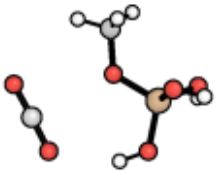

Figure 306: m1233

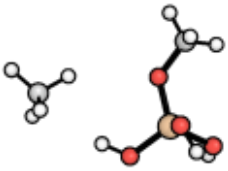

Figure 307: m1234

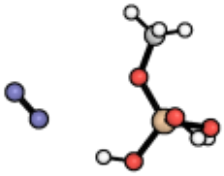

Figure 308: m1235

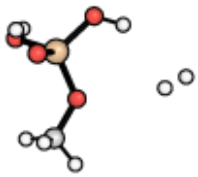

Figure 309: m1236

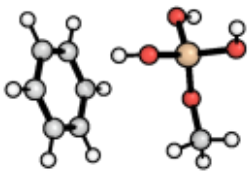

Figure 310: m1237

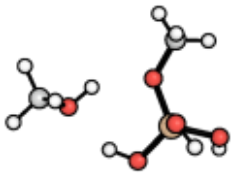

Figure 311: m1238

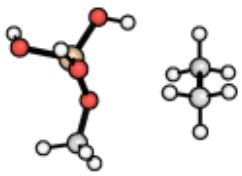

Figure 312: m1239

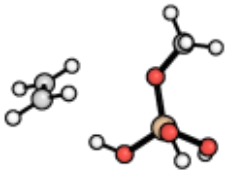

Figure 313: m1240

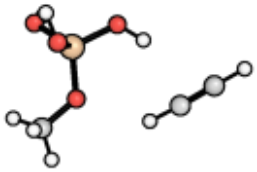

Figure 314: m1241

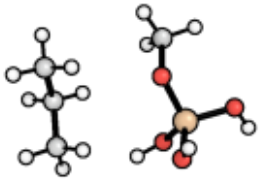

Figure 315: m1242

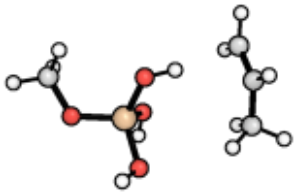

Figure 316: m1243

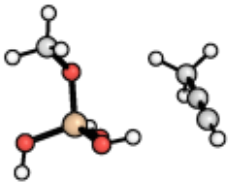

Figure 317: m1244

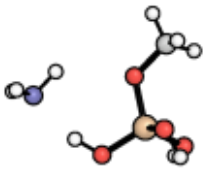

Figure 318: m1245

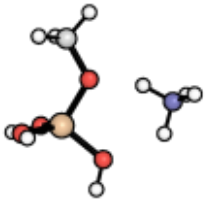

Figure 319: m1246

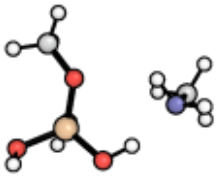

Figure 320: m1247

Figure 321: m1248

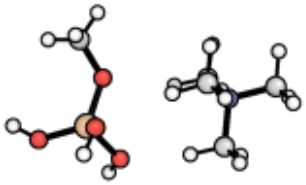

Figure 322: m1249

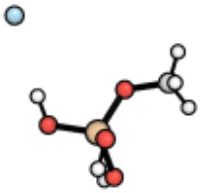

Figure 323: m1250

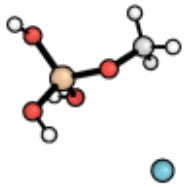

Figure 324: m1251

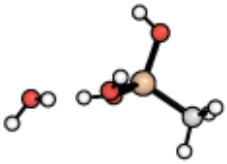

Figure 325: m1252

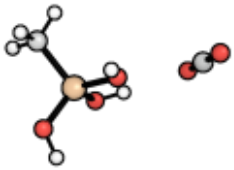

Figure 326: m1253

Figure 327: m1254

Figure 328: m1255

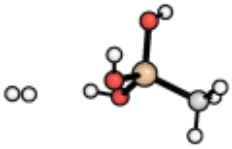

Figure 329: m1256

Figure 330: m1257

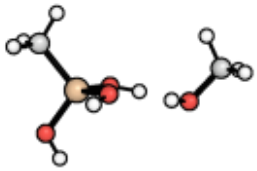

Figure 331: m1258

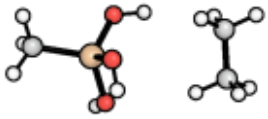

Figure 332: m1259

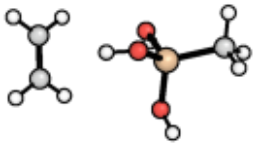

Figure 333: m1260

Figure 334: m1261

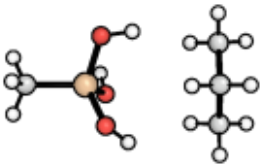

Figure 335: m1262

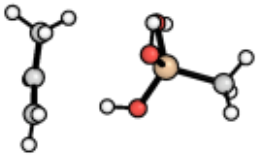

Figure 336: m1263

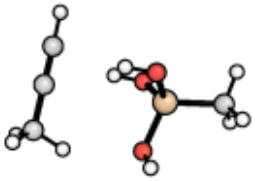

Figure 337: m1264

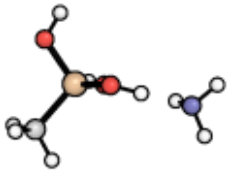

Figure 338: m1265

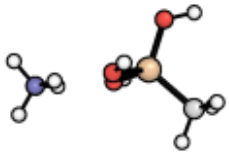

Figure 339: m1266

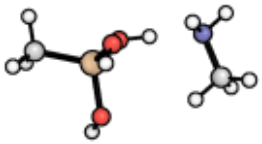

Figure 340: m1267

Figure 341: m1268

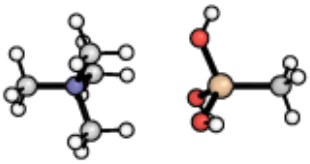

Figure 342: m1269

Figure 343: m1270

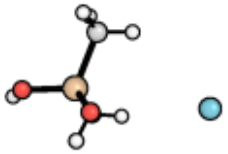

Figure 344: m1271