Beyond molecular properties: computational study of TAPP cluster and bulk phase

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

(Dr. rer. nat.)

von der KIT-Fakultät für Chemie und Biowissenschaften

des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

M.Sc. Jing Liu

Tag der mündlichen Prüfung: 20. 07. 2022

Referent: Priv.-Doz. Dr. Sebastian Höfener

Korreferent: Prof. Dr. Marcus Elstner

Vorsitz: Prof. Dr. Mario Ruben

Weitere Prüfer: Prof. Dr. Silke Behrens

Priv.-Doz. Dr. Zbigniew Pianowski

CONTENTS

I. INTRODUCTION

Organic semiconductors (OSCs) are organic solids which the organic molecules formed the crystal structure [1–4], they can readily switch from conductor to insulator, whereby the process of doping or changes in the electric field control the state of the material [5–9]. Since 1950s when the discovery was made that polycyclic aromatic compounds formed semi-conducting charge-transfer complex salts with halogens [10, 11], there has been great developments in their applications. In 1977 the first highly conducting polymer, chemically doped polyacetylene was discovered [12], and despite the poor performance on semiconducting properties, these organic materials are gradually getting more and more attentions during the past years. Compare to their inorganic counterparts, OSCs have many desirable properties such as light weight, low-cost production, low-temperature processing, mechanical flexibility, and abundant availability based on the easy-to-tailor nature of the organic molecules [7, 13–16]. These properties encourage the varieties of applications based on the organic semiconductors in the last few decades. Many applications are emerging with organic semiconductors such as organic field effect transistors (OFETs), light-emitting diodes (OLEDs), organic photovoltaic cells (OPVs) [17–24].

Organic semiconductors can be broadly divided into two main categories: polymers and small-molecule materials [25–27]. From a chemical point of view there are differences between these two categories, which obviously affect their technological performance. Polymeric macromolecules are constituted by the repetition of a fundamental unit, the monomer, and are soluble in organic solvents [28, 29]. Small-molecule materials can be further divided into two subgroups based on their solubility, first one pigments, which are not soluble in organic solvents, and second one dyes, which are soluble [30]. As a result of their chemical properties, polymers can be solution processed (for example, being spin-coated using the appropriate organic solvents), whereas small-molecule materials must be in general thermally evaporated. Polycyclic compounds, which are constituted by the repetition of a fundamental unit, the monomer, or soluble in organic solvents are particularly promising and are used in various applications [31–33]. The π conjugated polymer systems are of especially importance because of their charge carrier mobility, which is heavily relied on the π-π overlaps between the stacks of the polymers.

Among π conjugated system, tetraazaperopyrenes (TAPPs) in particular draw increas-

ingly attention in the past few years [34–36]. TAPPs provide a solid approach on studying and synthesising the n-type semiconducting materials since this area receives less attention compare to the p-channel organic semiconductors during the last decades due to the reasons such as low mobility, instability in the air and large barrier to electron injection [34–42]. Building of the TAPPs is based on the use of nitrogen-containing oligoacenes, such as diphenylanthrazolines and tetraazaanthracene derivatives to form the backbone aromatic structure. Different TAPP compounds have been synthesised by halogenation or alkylation process [38]. TAPP compounds also show desirable electronic properties with different halogenation schemes [39–42].

In order to get better understanding the electronic properties of TAPPs both in vacuum and in bulk phase, varieties different theoretical methods are applied, from molecular dynamic (MD) simulation using force field to the highly accurate post-Hartree Fock (HF) and density functional theory (DFT) methods [43–45].

A fundamental precursor for understanding the macroscopic charge-transport behaviour is to study ionization potentials (IPs) and electron affinities (EAs) [2]. One approach to gain insight into these quantities is calculating frontier-orbital energies with density-functional theory (DFT) [46–50]. The advantage of conventional DFT is that it can be applied to ground-state properties of large systems due to its low scaling with system size [51–53]. However, DFT lacks a systematic way to improve the functional and it is often not sufficiently accurate for properties such as bandgaps, excited-state energies, or charge-transport processes [54–57]. One way to improve over DFT is given by the many-body Green's function GW formalism [58–60]. The G_0W_0 method only involves a one-step evaluation instead of an iterative process, thus greatly reducing the computational costs while still improving significantly over DFT [61–63]. However, although in more accurate GW methods, the results are rather insensitive to the choice of the functional, and it is not the case for G_0W_0 in which the results are influenced by the initial choice of the functional [64, 65]. Despite several advantages of the G_0W_0 method, e.g., the straightforward manner to obtain absolute IP/EA energies, it is more difficult to extract property information for the sought quasi-particle states.

An alternative to the GW approach is wavefunction methods as they can be systematically improved to the full configuration interaction (FCI) limit, and in general, they offer a straightforward way to investigate the target-state properties [66]. Two approaches that

have been successful in the last decades are the algebraic diagrammatic construction (ADC) scheme [67, 68] and coupled-cluster (CC) response theory [69–71]. These methods, particularly the second-order approximate schemes, which have costs that scale with N^5 , where N is a measure of the system size, have evolved into very efficient tools for studying the excited-state properties of medium-sized systems [72, 73]. Another approach is given by the CIS(D_∞) method, which integrates the second-order perturbation theory into the configuration interaction singles (CIS) framework to improve its performance on calculating the excited state properties [74]. The accuracy and efficiency of CIS(D_∞), ADC(2), and second-order approximate coupled-cluster (CC2) singles and doubles are similar because of close technical relations [75–77]. Using wavefunction methods, a simple scheme to investigate electron-detachment and electron-attachment energies is used to calculate the energy difference between the ionic and the neutral system, which is also denoted the ΔE approach [78, 79]. This scheme suffers from spin contamination when calculating the ions, and often, only the lowest ionic state can be investigated. In order to avoid spin contamination and to investigate higher ionic states, wavefunction methods can be combined with response theory or, in case of a coupled cluster, also with the equation-of-motion (EOM) ansatz to describe electron-detached or electron-attached states in which electrons are excited to the continuum or localized from the continuum to unoccupied MOs, respectively [80–84].

Wavefunction methods have shown their ability to yield highly accurate results, but due to a rather steep scaling of the computational costs with system size, even dimers and trimers of medium-sized molecules are out of reach, in particular if excited-state properties are sought. In case of weakly interacting molecules, embedding approaches offer the advantage of a priori partitioning, leading to conceptual decomposition and significantly reduced computational scaling. There are numerous embedding methods each with a different approach on how to treat the subsystems. To give an example, one difference between embedding schemes is whether quantum mechanics (QM) methods are applied for the target subsystem only while classical molecular mechancis (MM) are employed for the environment, also denoted QM/MM, or the embedding scheme allows us to treat all subsystems at QM levels, also denoted QM/QM. Among all the QM/QM embedding schemes, frozen density embedding (FDE) is one efficient approach to study environment effects beyond pure Coulomb contributions on target molecules without the need of parameterization. For example, FDE allows both an accurate subsystem mode, in which all subsystems are treated

with the same accuracy, and an embedding mode, in which more approximate methods are applied for the environment sub-systems [85–88]. Recently, an equation-of-motion coupled-cluster singles and doubles for ionization potential (EOM-IP-CCSD) method combined with FDE for vertical transitions has been reported in which DFT-in-DFT is used to obtain the embedding potential. EOM methods have also been developed recently in the framework of projector-based embedding [89–91].

In this work, the geometrical properties of the different TAPP derivatives are analysed. The charge transfer process of organic crystals is largely influenced by their geometrical configuration in the solid state. The molecular displacement, the angles between the molecules and the distances between the layers are particularly important and will be discussed in detail. On the other hand, the concept of short contact also provides a useful way to predict the possible interactions between atoms on neighboring molecules.

In organic materials, charge can be excited within molecules without migration, which is called Frenkel excitons (FE) [92, 93]; another excitation process, in which electrons are excited and then migrate to other molecules, is called charge transfer (CT) [94, 95]. These processes compete with each other, and the final outcome of the more favorable process is largely determined by the geometric arrangement of the organic molecules in the solid state. If the given bulk structure tends to favor the hopping process, which results in the exciton migrating to the neighboring molecules before the excitation relaxation process occurs, then the CT process will dominate instead of the FE process; otherwise, the FE process will dominate. The relationship between these configurations is strongly influenced by their electronic coupling. The relaxation of the molecular cluster geometry must be taken into account when analyzing the efficiency of the transfer process, which is related to the coupling between the CT and FE states. It is also known that the displacement between molecules also affects the overlap of the boundary orbitals between molecules by affecting the displacement between molecules. This would lead to different patterns of the charge transfer process. Therefore, the analysis of the geometric configuration of organic molecules and their relative positions in the solid state is of great importance.

II. THEORY

The methods involved in these works are post-HF methods such as CIS-D(∞) and ADC(2) methods combined with the continuum approach which are used to study the ionization potential (IP) and electron affinity (EA) properties of the vacuum TAPP molecule and clusters. Also the periodic density functional theory with plane wave basis function is also employed for the calculation of the geometry of the TAPPs in the bulk phase and PBE together with B3LYP functional are used for investigate the HOMO-LUMO gap of such bulk systems. In the present work, electron attachment energies are provided as $\Delta E_{\text{attachment}} = -E_{EA}$ in order to distinguish them more easily in the tables from the ionization potentials. We have, therefore, adapted also the sign of the experimental adiabatic electron affinity (AEA) value of azulene to -0.8 ± 0.1 eV accordingly [96, 97].

A. Hartree-Fock method

The fundamental of the electronic structure calculation is the Hartree-Fock method (HF), upon which the molecular electronic orbital can be described by a single particle wavefunction. In modern day computational chemistry the HF method serves mostly as the starting point for other more sophisticated and accurate methods such as configuration interaction and coupled cluster methods, which despite their capability on describing the electronic correlation interactions in great detail, are also extremely demanding both on computational resources and timing.

HF method is set to solve the Schrödinger equation (SEQ) of the system under the Born-Oppenheimer (BO) approximation. Consider the time independent version of the SEQ since this is the basic form for ground state property calculation:

$$\hat{H}\Psi = E\Psi \tag{1}$$

where $\hat{H}$ is the system Hamiltonian which can be expend as following:

$$\begin{aligned} \hat{H} &= \hat{T} + \hat{V}_{ee} + \hat{V}_{eN} \\ &= -\frac{1}{2}\sum_i \nabla_i^2 + \sum_{i>j} \frac{1}{r_{ij}} + \sum_{N,i} \frac{Z_N}{r_{Ni}} \end{aligned} \tag{2}$$

where $\hat{T}$ is the electronic kinetic energy, $\hat{V}_{ee}$ and $\hat{V}_{eN}$ are the electron-electron and electron-nuclei potential energy respectively. The wavefunction acquired by solving equation (1) is

the electron wavefunction while the corresponding energy is the electron energy. For the N-particle system, the corresponding N-particle wavefunction can be written as the product of the single electron wavefunction:

$$\Psi(x_1, x_2, \ldots, x_N) = \chi_1(x_1)\chi_2(x_2)\chi_3(x_3)\cdots\chi_N(x_N) \tag{3}$$

where $\chi_i(x_i)$ is the spin orbital of particle x_i. In general for the N-particle system, wavefunction of such system can be written as the Slater determinant:

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(x_1) & \chi_2(x_1) & \cdots & \chi_N(x_1) \\ \chi_1(x_2) & \chi_2(x_2) & \cdots & \chi_N(x_2) \\ \cdots & \cdots & \ddots & \cdots \\ \chi_1(x_N) & \chi_2(x_N) & \cdots & \chi_N(x_N) \end{vmatrix} \tag{4}$$

Hartree-Fock equation can be written as:

$$\hat{f}(x_i)\chi_i(x_i) = \epsilon_i\chi_i(x_i) \tag{5}$$

In order to solve the HF equation, Roothaan and Hall in 1951 independently proposed to expand the single electron spin orbital into linear combinations of series of basis functions.

$$\chi_i(x_i) = \sum_j c_{ji}\phi_j \tag{6}$$

During the minimization process the coefficient c_j will be optimized while the basis function ϕ_j will be kept unchanged. There are two types of basis sets in conventional molecular electronic structure calculations: Slater type orbital (STO) and Gaussian type orbital (GTO). Physically speaking STO is more favourable since it can correctly describe the cusp at the $r = 0$, i.e. the electronic behaviour on the site of nucleus while GTO will give wrong description. Yet due to the hardness of integration between two Slater type functions, in modern day computational chemistry Gaussian type of basis function is most commonly used although plane wave basis set is more popular in solid state calculation. Insert the basis function expression back to HF equation (5):

$$\hat{f}(x_i)\sum_j c_{ji}\phi_j(x_i) = \epsilon_i\sum_j c_{ji}\phi_j(x_i) \tag{7}$$

For N-particle system the Roothaan equation can be written as:

$$\mathbf{FC}_j = \varepsilon_j\mathbf{SC}_j \tag{8}$$

where **F** is Fock matrix and **S** is the overlap matrix of basis functions and **C** is the coefficient matrix.

Despite its relatively success on describe the ground state energy, there are still lot to be desired mostly due to the fact that basic HF formalism ignores the electronic correlation interaction, which is essential when dealing with the properties such as excited state energy, electronic spectra, transition state, etc.. Some of the post-HF methods included in this work would be explained in the following sections.

B. Post-HF methods

1. Second-order algebraic diagrammatic construction scheme

Among the various different many-body methods developed for calculating excited state properties of molecules, many-body Green's functions and equation-of-motion (EOM) method are the most popular choices while random phase approximation (RPA) method also serves as fundamental for the development of large potion of modern day many-body methods. Despite their popularity, these above mentioned methods also suffer some disadvantages, such as Hartree-Fock instability of RPA. In attempt to address these drawbacks, Jochen Schirmer proposed the new method called algebraic diagrammatic construction (ADC) [67, 68]. Algebraic diagrammatic construction (ADC) is a new approximation scheme for the polarization propagator, and to expand this scheme to the second order would results ADC2 method.

In order to construct the ADC scheme, we start with the polarization propagator in its Fourier transform:

$$\Pi_{jk,j'k'}(\omega) = \int_{-\infty}^{\infty} d(t-t') e^{i\omega(t-t')} \Pi_{jk,j'k'}(t-t') \tag{9}$$

with the spectral representation:

$$\begin{aligned} \Pi_{jk,j'k'}(\omega) = & \sum_{m\neq 0} \frac{\langle \Psi_0 | c_k^\dagger c_j | \Psi_m \rangle \langle \Psi_m | c_{j'}^\dagger c_{k'} | \Psi_0 \rangle}{\omega + E_0 - E_m + i\eta} + \\ & \sum_{m\neq 0} \frac{\langle \Psi_0 | c_{j'}^\dagger c_{k'} | \Psi_m \rangle \langle \Psi_m | c_k^\dagger c_j | \Psi_0 \rangle}{-\omega + E_0 - E_m + i\eta} \end{aligned} \tag{10}$$

with can be simplified as:

$$\underline{\Pi}(\omega) = \underline{\Pi}_+(\omega) + \underline{\Pi}_-(\omega) \tag{11}$$

The transition function can be build with transition operator $\hat{D}$ as following:

$$T(\omega) = \underline{D}^{\dagger}\underline{\Pi}_{+}(\omega)\underline{D} \tag{12}$$

which can be expended as following:

$$T(\omega) = \sum_{n=0}^{\infty} T^{(n)}(\omega) \tag{13}$$

with

$$T^{(n)}(\omega) = \underline{D}^{\dagger}\underline{\Pi}_{+}^{(n)}(\omega)\underline{D} \tag{14}$$

Now the transition function $T(\omega)$ can be written as:

$$T(\omega) = \underline{F}^{\dagger}\underline{\Gamma}_{+}(\omega)\underline{F} \tag{15}$$

$$\underline{\Gamma}_{+}(\omega) = [\omega\underline{1} - \underline{K} - \underline{C}]^{-1} \tag{16}$$

The vector $\underline{F}$ and matrices $\underline{\Gamma}$, $\underline{K}$ and $\underline{C}$ are defined within the familiar space of sing, double, triple excitations with respect to the unperturbed HF ground state. With the zero-th order term of $T(\omega)$ is given by:

$$T_0(\omega) = \underline{F}^{\dagger}(0)[\omega\underline{1} - \underline{K}]^{-1}\underline{F}(0) \tag{17}$$

where

$$K_{jk,j'k'} = \delta_{jj'}\delta_{kk'}(\epsilon_j - \epsilon_k)$$

$$F_{jk}(0) = D_{jk}$$

where j and k denote particle and hole states respectively.

The diagonalization within the p-h excitations space for the first-order ADC method yielding the excitation energies, which is equivalent to the TDA scheme. Even the first-order ADC method is fundamentally the same as the RPA method, however the second-order ADC method takes a decisive step away from the RPA method. The configuration space of the second-order ADC method contains the p-h and 2p-2h excitations so size-wise it is of the configuration interaction (CI) method. By construction, the excitation energies and the transition moments of the sing excited states are now exact up to second-order. It is also worth noticing that due to the similarity of construction, CIS(D_{∞}) always show results identical with ADC(2) scheme.

2. CIS(D_∞) method

In order to overcome some of the drawbacks of the single excitation configuration interaction method (CIS), second-order correction via perturbation theory is added to the CIS framework [74]. Using CIS wavefunction as reference, CIS(D) pproximately accounts for dynamical correlation effects on electronic transitions. CIS(D_∞) is a full diagonalization of the response matrix through second order in the space of singles and doubles and requires explicit iteration of double substitutions like CC2. The coupled cluster (CC) equation for the ground state amplitudes a_i where i goes for all the excitations:

$$\langle \phi_i | \hat{H} | exp \sum_i a_i \hat{t}_i \phi_0 \rangle_C = 0 \tag{18}$$

Where the coupled cluster response matrix $\mathbf{A}$ is given by:

$$A_{ij} = \langle \phi_i | \hat{H} | \hat{t}_j exp(\sum_k a_k \hat{t}_k) \phi_0 \rangle \tag{19}$$

with the eigenvalue as:

$$\mathbf{Ab} = \omega \mathbf{b} \tag{20}$$

Based on the different truncation of the response amplitudes, a series of size consistent approximation methods can be derived.

For CIS(D) approximation to expand the response matrix first the Hamiltonian is divided into mean-field Fock operator and the first-order fluctuation potential:

$$\hat{H} = \hat{F}^{(0)} + \hat{V}^{(1)} \tag{21}$$

CIS results are obtained at zero-order from the eigenvector of singlet-singlet block of the CIS response matrix and first-order terms are matrix elements of the fluctuation potential without the SS block. The expansion of the response matrix can be written as:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{SS}^{(0)} & 0 & 0 & \cdots \\ 0 & \mathbf{D}_{DD}^{(0)} & 0 & \cdots \\ 0 & 0 & \mathbf{D}_{TT}^{(0)} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} + \begin{bmatrix} 0 & \mathbf{A}_{SD}^{(1)} & \mathbf{A}_{ST}^{(1)} & \cdots \\ \mathbf{A}_{DS}^{(1)} & \mathbf{A}_{DD}^{(1)} & \mathbf{A}_{DT}^{(1)} & \cdots \\ 0 & \mathbf{A}_{TD}^{(1)} & \mathbf{A}_{TT}^{(1)} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} + \begin{bmatrix} \mathbf{A}_{SS}^{(2)} & \mathbf{A}_{SD}^{(2)} & 0 & \cdots \\ \mathbf{A}_{DS}^{(2)} & \mathbf{A}_{DD}^{(2)} & \mathbf{A}_{DT}^{(2)} & \cdots \\ \mathbf{A}_{TS}^{(1)} & \mathbf{A}_{TD}^{(2)} & \mathbf{A}_{TT}^{(2)} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} + \cdots \tag{22}$$

Due to the ill-behaved nature of the reference CIS states exhibit the near degeneracies, the entire manifold of CIS reference states are being treated as near degeneracies. Therefore

the entirety of the singles block of the response matrix are being rediagonalized, allow for remixing induced by including electron correlation effects through second order in the fluctuation potential.

$$\begin{bmatrix} \mathbf{A}_{SS}^{(0)} + \mathbf{A}_{SS}^{(2)} & \mathbf{A}_{SD}^{(1)} \\ \mathbf{A}_{DS}^{(1)} & \mathbf{D}_{DD}^{(0)} \end{bmatrix} = \begin{bmatrix} \langle \Phi_S | \hat{H} | \Phi_S \rangle + \langle \Phi_S | \hat{V} | T_2^{(1)} \Phi_S \rangle & \langle \Phi_S | \hat{V} | \Phi_D \rangle \\ \langle \Phi_D | \hat{V} | \Phi_S \rangle & \langle \Phi_D | \hat{F} | \Phi_D \rangle \end{bmatrix} \tag{23}$$

Diagonalization above equation is a quasi-degenerate generalization of CIS(D).
Define the following:

$$\Delta = \omega (\mathbf{D}_{DD}^{(0)})^{-1} \tag{24}$$

we have the following binomial expansion:

$$\begin{aligned} (\mathbf{D}_{DD}^{(0)} - \omega)^{-1} &= (\mathbf{D}_{DD}^{(0)})^{-1} (1 - \Delta)^{-1} \\ &= (\mathbf{D}_{DD}^{(0)})^{-1} (1 + \Delta + \Delta^2 + \cdots) \end{aligned} \tag{25}$$

The CIS(D_n) methods are defined based on the truncation of n-th Δ terms and CIS(D_∞) indicates solving the above binomial expansion exactly for the eigenvalues and eigenvectors.

3. CC2 method

Even as the most popular model for calculating the molecular properties and electron correlations, coupled cluster theory with singles and doubles (CCSD) already scale as N^6 and with triples (CCSDT) the scaling with respect to the system size would be up to N^8. This drastically increase of the scaling with respect to the system size would require a new approximation if by any chance the method of CCSD level of accuracy can be applied on any real chemical molecules. Ove Christiansen et. al. proposed a second-order coupled cluster (CC2) as approximation to the CCSD method in 1995.

CCSD energy can be expressed as:

$$E = \langle \phi_{HF} | \hat{H} exp(T_1 + T_2) | \phi_{HF} \rangle \tag{26}$$

with amplitudes from:

$$\begin{aligned} &\langle \mu_i | exp(-T_1 - T_2) \hat{H} \exp(T_1 + T_2) | \phi_{HF} \rangle = 0 \\ &i = 1, 2 \end{aligned} \tag{27}$$

with μ_i as the singles and doubles excitation manifold. CCSD equation can be further expend as following:

$$\langle \mu_2 | \hat{H} + [\hat{H}, T_2] + \frac{1}{2}[[\hat{H}, T_2], T_2] | \phi_{HF} \rangle = 0 \tag{28}$$

and Hamiltonian can be divided into Fock and fluctuation operator:

$$\hat{H} = \hat{F} + \hat{U} \tag{29}$$

The CCSD energy is correct through third order if using the converged HF wavefunction. To obtain an energy correct through second order it is sufficient to include the first order doubles amplitudes as in MP2. By assign the singles to be zeroth order in $\hat{U}$, CCSD equation can be rewritten as following approximation:

$$\langle \mu_2 | [\hat{F}, T_2] + \hat{H} | \phi_{HF} \rangle = 0 \tag{30}$$

Without external perturbation, CC2 model can be written as combination of above equation and the following one:

$$\langle \mu_1 | \hat{H} + [\hat{H}, T_2] | \phi_{HF} \rangle = 0 \tag{31}$$

C. IP and EA methods

1. Continuum orbital scheme

As shown about two decades ago, ionization potentials can be obtained numerically for conventional linear-response methods when using a continuum orbital [98]. While in the original work an exponent of $10-50$ was used, a diffuse coefficient of $10-10$ yielded satisfying accuracy for the present purpose. However, the continuum orbitals can not only be used to obtain IP and EA values numerically but they can also be used in a formal manner in the derivation of the methods, which then have eliminated the continuum orbital from the actual calculation. Continuum orbitals, denoted x in the following, have an orbital eigenvalue of exactly zero, i.e., $\hat{F}_{xx} = F_{xx} = \varepsilon_x = 0$, and they do not interact with the other orbitals located on the system so that all matrix elements of one-electron and two-electron integrals vanish, e.g., $F_{xq} = 0$, $\langle x|q\rangle = 0$ and $\langle xq|rs\rangle = 0$. These properties are used in the following for the derivation, but the continuum orbitals are finally omitted in the resulting implementation.

2. Ionization potential and analytical nuclear gradients

a. Vertical ionization potentials We relate the vertical ionization energies to vertical excitation energies where one electron is excited into an ultra-diffuse virtual orbital, denoted x in the following, with an energy of exactly zero and a vanishing overlap with the remaining basis functions centered on the investigated system. The excitation where one electron is excited into the vacuum orbital x is denoted as $\{\mu^h\} = \{xi, xibj, xibjck, \dots\}$. Alternatively, the index is sometimes dropped, leading to the excitation hole manifold $\{\mu^h\} = \{i, ibj, ibjck, \dots\}$, but in the present work, the index x is kept for a better readability. Since the electron in the orbital x cannot interact with the remaining system, $\mathbf{A}$ is the block diagonal with respect to the two excitation manifolds $\{\mu^h\}$ and $\{\mu\}\backslash\{\mu^h\}$. The ionization energies are, thus, obtained as the eigenvalues of the sub-block A_{μ^h,ν^h}.

To determine the eigenvalues of A_{μ^h,ν^h}, products with trial vectors of the following type are required:

$$\rho_{\mu^h} = \sum_{\nu^h} A_{\mu^h,\nu^h} R_{\nu^h} \tag{32}$$

Replacing $\rho_a^i \to \rho_x^i, R_a^i \to R_x^i$, and $R_{ab}^{ij} \to R_{xb}^{ij}$ or $R_{ab}^{ij} \to R_{ax}^{ij}$ and restricting the ground-state amplitudes to excitations into the usual virtual basis functions, one obtains the following for the CC2 model:

$$^{EA-CC2}\rho_i^x = -\sum_j E_{ji} R_x^j - \sum_{dlk} (ld\hat{|}ki)\tilde{R}_{xd}^{kl} + \sum_{dl} \tilde{R}_{xd}^{il} \hat{F}_{ld} \tag{33}$$

$$^{EA-CC2}\rho_{ij}^{xb} = (\varepsilon_b - \varepsilon_i - \varepsilon_j) R_{xb}^{ij} - \sum_k (ki\hat{|}aj) R_x^k \tag{34}$$

The expressions for CIS(D_∞) are obtained from those for CC2 by replacing the ground-state amplitudes with the first-order Møller–Plesset (MP1) perturbation theory amplitudes as follows:

$$^{EA-CIS(D_\infty)}\rho_x^i = -\sum_j E_{ji}^{MP1} R_x^j - \sum_{dlk} (ld|ki)\tilde{R}_{xd}^{kl} \tag{35}$$

$$^{EA-CIS(D_\infty)}\rho_{xb}^{ij} = (\varepsilon_b - \varepsilon_i - \varepsilon_j) R_{xb}^{ij} - \sum_k (ki|aj) R_x^k \tag{36}$$

The expressions for ADC(2) are obtained by symmetrizing in addition the contribution of what before has been the singles-singles block of the Jacobian matrix A_{μ^h,ν^h}. The effect of

this is now reduced to a symmetrization of the contributions from E^{MP1},

$$
\begin{aligned}
{}^{EA-ADC(2)}\rho_x^i &= -\frac{1}{2}\sum_j (E_{ji}^{MP1} + E_{ij}^{MP1}) R_x^j - \sum_{dlk} (ld|ki)\tilde{R}_{xd}^{kl} && (37)\\
{}^{EA-ADC(2)}\rho_{xb}^{ij} &= (\varepsilon_b - \varepsilon_i - \varepsilon_j) R_{xb}^{ij} - \sum_k (ki|aj) R_x^k && (38)
\end{aligned}
$$

b. Excited-state Lagrangian for CIS(D_∞) and ADC(2) Having obtained the eigenvalues and eigenvectors of the Jacobian, a Lagrange functional is constructed to facilitate the calculation of properties and analytical nuclear gradients of the ionized states,

$$
\begin{aligned}
\mathcal{L}^f = \quad & \langle HF|H + [H, T_2]|HF\rangle + \omega^f (1 - \sum_{\mu^h} \bar{L}_{\mu^h}^f R_{\mu^h}^f) \\
& + \sum_{\mu^h,\nu^h} \bar{L}_{\mu^h}^f A_{\mu^h\nu^h} R_{\mu^h}^f + \sum_\mu \bar{t}_\mu \Omega_\mu + \sum_{\mu_0} \bar{\kappa}_{\mu_0} F_{\mu_0} \qquad (39)
\end{aligned}
$$

where ω is the ionization energy and Ω are the constraints for the amplitudes. The quantities $\bar{\mathbf{t}}$, $\bar{\mathbf{L}}$, and $\bar{\bar{\kappa}}$ denote the amplitude multipliers, the excitation vector multipliers, and the orbital-rotation multipliers, respectively. Note that the Lagrange multipliers for the excited state, $\bar{\mathbf{L}}^f$, refer to the hole manifold μ^h, while the Lagrange multipliers for the ground-state amplitudes, $\bar{\mathbf{t}}$, refer to the conventional manifold μ. The index range of the Lagrange multipliers $\bar{\kappa}_{\mu 0}$ depends on whether, e.g., a canonical, non-canonical, or semi-canonical representation is chosen.

In the present work, the ground-state embedding potential is used for both vertical transitions and nuclear gradients, that is, geometry relaxation. Note that the use of the ground-state embedding potential for geometry relaxation is an approximation as the environment densities are lacking polarization of the electron change, which results in a too little back-polarization of the ionized compound. The study of back-polarization effects is, however, beyond the scope of the present work and shall be addressed in future work.

c. Excited-state Lagrange multipliers In order to make the Lagrangian stationary, the excited-state Lagrange multipliers $\bar{\mathbf{L}}^f$ have to be determined. They are constructed to enforce a vanishing derivative with respect to the excitation vectors,

$$
\frac{\partial}{\partial R_{\mu^h}} \mathcal{L}^f = \sum_{\nu^h} \bar{L}_{\nu^h}^f (A_{\nu^h\mu^h} - \omega^f) \stackrel{!}{=} 0 \qquad (40)
$$

which is expressed as an eigenvalue equation,

$$
\bar{\mathbf{L}}^f \mathbf{A} = \omega^f \bar{\mathbf{L}}^f \qquad (41)
$$

The Lagrange multipliers are, thus, nothing else than the left eigenvectors. However, the biorthogonality condition is taken care of by explicitly enforcing $|\mathbf{R}^f| = 1$ and $\langle \mathbf{L}^f | \mathbf{R}^f \rangle = 1$ for $\mathbf{L}^f$ and $\mathbf{R}^f$ respectively.

For a transformation of the Jacobian A_{μ^h,ν^h} from the left, one obtains in a doubles-direct formulation only a residual ρ for the singles while the doubles are constructed on-the-fly,

$$\rho_x^I = -\sum_j E_{ji} \bar{L}_x^j - \sum_{dlk} (ld|ik) \bar{L}_{dx}^{kl} \tag{42}$$

$$\bar{L}_{ax}^{ij} = -\frac{2\sum_k \bar{L}_x^k (j^p k^h | i^p a^h) - \sum_k \bar{L}_x^k (i^p k^h | j^p a^h)}{\varepsilon_a - \varepsilon_i - \varepsilon_j} \tag{43}$$

d. Ground-state amplitude multipliers The equation that determines the ground-state amplitude multipliers is obtained by the differentiation of the Lagrangian with respect to the ground-state amplitudes t_λ,

$$\begin{aligned} 0 \stackrel{!}{=} \frac{\partial}{\partial t_\lambda} \mathcal{L}^f = \quad & (\frac{\partial}{\partial t_\lambda} \langle HF | [H, T_2] | HF \rangle) + \sum_\mu \bar{t}_\mu^f (\frac{\partial}{\partial t_\lambda^f} \Omega_\mu) \\ & + \sum_{\mu^h \nu^h} L_{\mu^h}^f (\frac{\partial}{\partial t_\lambda} A_{\mu^h \nu^h}) R_{\nu^h}^f \end{aligned} \tag{44}$$

In the latter equation, the first and second term on the right-hand side (RHS) are simply the conventional CC2 equations. Often, for the third term on the RHS, the following matrix $\mathbf{B}$ is introduced:

$$B_{\mu^h,\nu^h;\lambda} = \frac{\partial}{\partial t_\lambda} A_{\mu^h \nu^h} \tag{45}$$

However, for the models CIS(D∞) and ADC(2) in the present work, only the following contribution is needed:

$$\begin{aligned} \sum_{ij} \bar{L}_x^i B_{i,j;manb} R_x^j = \quad & -\frac{\partial}{\partial t_{ab}^{mn}} \sum_{ij} \bar{L}_x^i E_{ji} R_x^j \\ = \quad & -2 \sum_j bar L_x^m (nb|ja) R_x^j + \sum_j \bar{L}_x^m (na|jb) R_x^j \end{aligned} \tag{46}$$

e. Z-vector equation and gradient In order to compute the orbital-relaxed excited-state properties, the Lagrangian also has to be stationary with respect to the ground-state orbital

rotations. The respective equations are obtained by differentiating:

$$\begin{aligned} 0 \stackrel{!}{=} \frac{\partial}{\partial \kappa_{\mu_0}} \mathcal{L}^f = \quad & (\frac{\partial}{\partial \kappa_{\mu_0}} \langle HF|H + [H, T_2]|HF\rangle) \\ & + \sum_{\mu\nu^h} L^f_{\mu^h} (\frac{\partial}{\partial \kappa_{\mu_0}} A_{\mu^h\nu^h}) R^f_{\nu^h} \\ & + \sum_{\mu} \bar{t}_\mu (\frac{\partial}{\partial \kappa_{\mu_0}} \Omega_\mu) + \sum_{\mu_0} \bar{\kappa}_{\mu_0} (\frac{\partial}{\partial \kappa_{\mu_0}} F_{\mu_0}) \end{aligned} \tag{47}$$

Re-arranging yields the following well-known Z-vector equation:

$$\begin{aligned} -\sum_{\mu_0} \bar{\kappa}_{\mu_0} (\frac{\partial}{\partial \kappa_{\mu_0}} F_{\mu_0}) = \quad & (\frac{\partial}{\partial \kappa_{\mu_0}} \langle HF|H + [H, T_2]|HF\rangle) \\ & + \sum_{\mu^h\nu^h} L^f_{\mu^h} (\frac{\partial}{\partial \kappa_{\mu_0}} A_{\mu^h\nu^h}) R^f_{\nu^h} \\ & + \sum_{\mu} \bar{t}^f_\mu (\frac{\partial}{\partial \kappa_{\mu_0}} \Omega_\mu) \end{aligned} \tag{48}$$

Using the effective densities, the occupied-virtual block on the right-hand side can be expressed as:

$$\begin{aligned} \eta_{AI} = \quad & \frac{1}{2}(D^{F,ex}_{AI} + D^{F,ex}_{IA})(\varepsilon_I - \varepsilon_A) - \frac{1}{4}(D^{F,ex}_{pq} + D^{F,ex}_{qp}) A_{pqAI} \\ & + ({}^{ex}H_{AI} - {}^{ex}\Omega_{AI} - {}^{ex}H_{IA} + {}^{ex}\Omega_{IA}) \end{aligned} \tag{49}$$

where the indices I, J and A, B denote the general, i.e., active and frozen, occupied and virtual indices, respectively. The only difference to Ref. [99] and [100] is that densities employing the particle/hole excitation vectors are used. For example, the one-electron density $\mathbf{D}^{F,A}$ is defined in the IP case as follows (for zero T_1 amplitudes):

$$D^{F,A}_{ij} = \quad -L^i_x R^j_x - \sum_{ak} L^{jk}_{ax} R^{ik}_{ax} \tag{50}$$

$$D^{F,A}_{ia} = \quad \bar{C}_{ai} = \sum_j \tilde{R}^{ij}_{ax} L^j_x \tag{51}$$

$$D^{F,A}_{ai} = \quad 0 \tag{52}$$

$$D^{F,A}_{aa} = \quad \sum_{ij} L^{ij}_{ax} R^{ij}_{ax} \tag{53}$$

$$\tag{54}$$

while the ground-state densities such as $\mathbf{D}^{\xi}$ remain unchanged. Similarly, the matrices ${}^{ex}\mathbf{H}$ and ${}^{ex}\mathbf{\Omega}$ are given as:

$$ {}^{ex}\mathbf{H} = \sum_{abc} {}^{\xi}\mathbf{H}^{(x)} + \sum_{ab} {}^{A}\mathbf{H}^{(x)} \tag{55} $$

$$ {}^{ex}\mathbf{\Omega} = \sum_{abc} {}^{\xi}\mathbf{\Omega}^{(x)} + \sum_{abc} {}^{A}\mathbf{\Omega}^{(x)} \tag{56} $$

for which the ground-state contributions, ${}^{\xi}\mathbf{H}^{(x)}$ and ${}^{\xi}\mathbf{\Omega}^{(x)}$, remain unaltered compared to the conventional excited-state case.

Finally, the contributions are computed from the matrices ${}^{A}\mathbf{H}$ and ${}^{A}\mathbf{\Omega}$, which are given in the case of IP-ADC(2), as follows,

$$ {}^{A}H^{(a)}_{pq} = \sum_{Qi} {}^{A}\tilde{\tilde{Y}}^{Q}_{pi} B^{Q}_{iq} \tag{57} $$

$$ {}^{A}H^{(b)}_{pq} = \sum_{a} \delta_{pa} \sum_{Qi} \breve{Y}^{Q}_{ai} \sum_{r} \delta^{h}_{ri} B^{Q}_{qr} + \bar{\lambda}^{p}_{px} \sum_{Qi} \breve{Y}^{Q}_{xi} \sum_{r} \delta^{h}_{ri} B^{Q}_{qr} \tag{58} $$

$$ {}^{A}\Omega^{(a)}_{pq} = \sum_{\alpha} C_{\alpha q} \sum_{j} \delta_{pj} \sum_{Q\beta} {}^{A}\Gamma^{(a),Q}_{j\beta}(\alpha\beta|Q) \tag{59} $$

$$ {}^{A}\Omega^{(b)}_{pq} = \sum_{\alpha} C_{\alpha q} \sum_{j} \delta^{h}_{pj} \sum_{Q\beta} {}^{A}\bar{\Gamma}^{(b),Q}_{j\beta}(\alpha\beta|Q) \tag{60} $$

$$ \tag{61} $$

3. Electron attachment and analytical nuclear gradients

a. Vertical electron attachment energies Similar to the IP case, vertical electron attachment energies are related to vertical excitation energies where one electron is excited from an ultra-diffuse occupied orbital, with an energy of exactly zero and a vanishing overlap with the remaining basis functions centered on the investigated system, into a virtual orbital of the system. The excitation where one electron is excited from x is denoted as $\{\mu^p\} = \{ax, axbj, axbjck, \ldots\}$. Similar to the IP case, the index can be dropped, resulting in $\{\mu^p\} = \{a, abj, abjck, \ldots\}$, but for a better readability, the index x is kept in the present work. Since the electron in the orbital x cannot interact with the remaining system, $\mathbf{A}$ is again a block diagonal with respect to the two excitation manifolds $\{\mu^p\}$ and $\{\mu\}\backslash\{\mu^p\}$. The electron attachment energies are, thus, obtained as the eigenvalues of the sub-block A_{μ^p,ν^p}.

To determine the eigenvalues of A_{μ^p,ν^p}, products with trial vectors of the following type

are required:

$$\rho_{\mu^p} = \sum_{\nu^p} A_{\mu^p,\nu^p} R_{\nu^p} \tag{62}$$

Replacing one of the occupied indices of the trial vector and one of the occupied indices of the result vector by x in the expressions for excitation energies, and restricting the ground-state amplitudes to excitation from the usual occupied orbitals with $i \neq x$, one obtains:

$$^{EA-CC2}\rho^x_a = \sum_b E_{ab} R^x_b + \sum_{dlc} \tilde{R}^{xl}_{cd} (ld\hat{|}ac) + \sum_{dl} R^{xl}_{ad} \hat{F}_{ld} \tag{63}$$

$$^{EA-CC2}\rho^{xj}_{ab} = (\varepsilon_a + \varepsilon_b + \varepsilon_j) R^{xj}_{ab} + \sum_c (ac\hat{|}bj) R^x_c \tag{64}$$

The expressions for CIS(D_∞) are obtained from the CC2 equations by replacing the ground-state amplitudes with the first-order Møller–Plesset (MP1) perturbation theory amplitudes as follows:

$$^{EA-CIS(D_\infty)}\rho^x_a = \sum_b E^{MP1}_{ab} R^x_b + \sum_{dlc} \tilde{R}^{xl}_{cd} (ld|ac) \tag{65}$$

$$^{EA-CIS(D_\infty)}\rho^{xj}_{ab} = (\varepsilon_a + \varepsilon_b + \varepsilon_j) R^{xj}_{ab} + \sum_c (ac|bj) R^x_c \tag{66}$$

The expressions for ADC(2) are obtained by symmetrizing in addition the contribution of what before has been the singles-singles block of the Jacobian matrix $\mathbf{A}_{\mu^p,\nu^p}$. The effect of this is now reduced to a symmetrization of the contributions from EMP1,

$$^{EA-ADC(2)}\rho^x_a = \frac{1}{2} \sum_b (E^{MP1}_{ab} + E^{MP1}_{ba}) R^x_b + \sum_{dlc} \tilde{R}^{xl}_{cd} (ld|ac) \tag{67}$$

$$^{EA-ADC(2)}\rho^{xj}_{ab} = (\varepsilon_a + \varepsilon_b + \varepsilon_j) R^{xj}_{ab} + \sum_c (ac|bj) R^x_c \tag{68}$$

b. Excited-state Lagrangian for CIS(D_∞) and ADC(2) Having obtained the eigenvalues and eigenvectors of the Jacobian, a Lagrange functional is constructed to facilitate the calculation of the properties and analytical nuclear gradients of the electron-attached states,

$$\begin{aligned} \mathcal{L}^f = \; & \langle HF|H + [H, T_2]|HF\rangle + \omega^f (1 - \sum_{\mu^p} \bar{L}^f_{\mu^p} R^f_{\mu^p}) \\ & + \sum_{\mu^p,\nu^p} \bar{L}^f_{\mu^p} A_{\mu^p\nu^p} R^f_{\mu^p} + \sum_\mu \bar{t}_\mu \Omega_\mu + \sum_{\mu_0} \bar{\kappa}_{\mu_0} F_{\mu_0} \end{aligned} \tag{69}$$

where ω^f is the electron-attachment energy. Note that the Lagrange multipliers for the excited state, $\bar{\mathbf{L}}^f$, refer to the particle manifold $\{\mu^p\}$ in the EA case.

In order to make the Lagrangian stationary, the excited-state Lagrange multipliers $\bar{\mathbf{L}}^f$ have to be determined, which are similar to the IP case, given as the left eigenvectors of the Jacobian $\mathbf{A}_{\mu^p,\nu^p}$. Therefore, follow the same approach as for IPs and just replace the μ^h manifold with the μ^p manifold, which leads to the following non-vanishing contributions to $\bar{\mathbf{L}}^{xj}_{ab}$:

$$\bar{L}^{xj}_{ab} = \frac{2\sum_c (ca|jb)\bar{L}^x_c - \sum_c (cb|ja)\bar{L}^x_c}{\varepsilon_b + \varepsilon_a + \varepsilon_j} \tag{70}$$

Similar to the IP case, the ground-state amplitudes are obtained by the differentiation of the Lagrangian with respect to the ground-state amplitudes t_λ. Following an analogous derivation as for IPs, for the models EA-CIS(D_∞) and EA-ADC(2), the doubles contribution is obtained as follows:

$$\begin{aligned}\sum_{cd} \bar{L}^x_c B_{c,d;manb} R^x_d = &\quad \frac{\partial}{\partial t^{mn}_{ab}} \sum_{cd} \bar{L}^x_c E_{cd} R^x_d \\ = &\quad -2\sum_d \bar{L}^x_a (nb|md) R^x_d + \sum_d \bar{L}^x_b (na|md) R^x_d \end{aligned} \tag{71}$$

c. Z-vector equation and gradient In case of EA methods, differentiating and rearranging yield a corresponding Z-vector equation and right-hand side. The difference is located in the effective densities, which are now constructed from the particle vectors,

$$D^{F,A}_{ij} = \quad -\sum_{ab} L^{xj}_{ab} R^{ix}_{ab} \tag{72}$$

$$D^{F,A}_{ia} = \quad \bar{C}_{ai} = \sum_{xb} \tilde{R}^{ix}_{ab} L^x_b \tag{73}$$

$$D^{F,A}_{ai} = \quad 0 \tag{74}$$

$$D^{F,A}_{ab} = \quad L^x_a R^x_b + \sum_{ic} L^{ix}_{ac} R^{ix}_{bc} \tag{75}$$

$$\tag{76}$$

The ground-state contributions, e.g., $\mathbf{D}^{xi}$, remain unchanged, while the matrices needed for the right-hand side vector η are given as:

$${}^A H^{(a)}_{pq} = \quad \sum_{Qi} {}^A\tilde{\tilde{\mathrm{Y}}}^Q_{pi} B^Q_{iq} \tag{77}$$

$${}^A H^{(b)}_{pq} = \quad \sum_a \lambda^p_{pa} \sum_{Qx} \breve{Y}^Q_{ax} \sum_r \bar{\lambda}^h_{rx} B^Q_{qr} + \sum_a \lambda^p_{pa} \sum_{Qi} \breve{\bar{Y}}^Q_{ai} \sum_r \bar{\lambda}^h_{ri} B^Q_{qr} \tag{78}$$

$${}^A H^{(b)}_{pq} = \quad \sum_{Qx} Y^{*Q}_{px} \sum_r \breve{\lambda}^p_{rx} B^Q_{rq} \tag{79}$$

and

$$
\begin{aligned}
{}^{A}\Omega^{(a)}_{pq} &= \sum_{\alpha} C_{\alpha q} \sum_{j} \lambda^{p}_{pj} \sum_{Q\beta} {}^{A}\Gamma^{(a),Q}_{j\beta}(\alpha\beta|Q) && (80)\\
{}^{A}\Omega^{(b)}_{pq} &= \sum_{\alpha} C_{\alpha q} \sum_{j} \lambda^{h}_{pj} \sum_{Q\beta} {}^{A}\bar{\Gamma}^{(b),Q}_{j\beta}(\alpha\beta|Q) + \sum_{\alpha} C_{\alpha q} \sum_{j} \bar{\lambda}^{h}_{px} \sum_{Q\beta} {}^{A}\Gamma^{(b),Q}_{x\beta}(\alpha\beta|Q) && (81)\\
{}^{A}\Omega^{(c)}_{pq} &= \sum_{\alpha} C_{\alpha q} \sum_{x} \breve{\lambda}^{p}_{px} \sum_{Q\beta} {}^{A}\Gamma^{(c),Q}_{x\beta}(\alpha\beta|Q) && (82)\\
& && (83)
\end{aligned}
$$

Some of these quantities are also used to compute the gradient itself. To give an example, the differentiated three-index integrals are contracted with an intermediate Δ^{ao},

$$
\frac{dL}{d\chi} \leftarrow \sum_{\alpha\beta p} \Delta^{P,ao}_{\alpha\beta} \frac{\partial(\alpha\beta|P)}{\partial\chi} \tag{84}
$$

For this contribution, the intermediate Δ^{ao} has to be contracted as follows for the excited-state contributions:

$$
\Delta^{P,ao}_{\alpha\beta} \leftarrow \sum_{j} (C_{\alpha j}\, {}^{A}\Gamma^{(a),Q}_{j\beta} + {}^{A}\bar{\Gamma}^{(b),Q}_{j\alpha} \Lambda^{h}_{\beta j}) + ({}^{A}\Gamma^{(b),Q}_{x\alpha} \bar{\Lambda}^{h}_{\beta x} + {}^{A}\breve{\Lambda}^{p}_{\alpha x} \Gamma^{(c),Q}_{x\beta}) \tag{85}
$$

d. Frozen-density embedding Employing the a priori partitioning of FDE, the total electron density is constructed as the sum of the subsystem electron densities,

$$
\rho(\mathbf{r}) = \sum_{z} \rho_z(\mathbf{r}) \tag{86}
$$

Each subsystem density ρ_z is obtained from individual subsystem calculations in which the polarization of the subsystems is obtained via an effective potential, vide infra. Due to such a partitioning, the energy can be expressed as follows for two subsystems, I and II:

$$
E_{tot}[\rho(\mathbf{r})] = E_I[\rho_I(\mathbf{r})] + E_{int}[\rho_I(\mathbf{r}), \rho_{II}(\mathbf{r})] + E_{II}[\rho_{II}(\mathbf{r})] \tag{87}
$$

where the interaction between the subsystems is given as:

$$
\begin{aligned}
E_{int}[\rho_I, \rho_{II}] = &\int \rho_I(\mathbf{r}) \upsilon^{II}_{nuc}(\mathbf{r}) d\mathbf{r} + \int \rho_{II}(\mathbf{r}) \upsilon^{I}_{nuc}(\mathbf{r}) d\mathbf{r} + E^{I,II}_{nuc} \\
&+ \iint \frac{\rho_I(\mathbf{r}) \rho_{II}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E^{nadd}_{emb:XCK}[\rho_I \rho_{II}]
\end{aligned} \tag{88}
$$

In the FDE approach, the non-additive exchange-correlation and the kinetic energy contributions are expressed using density functional theory (DFT), independently of the actual

subsystem methods, and read as follows:

$$E^{nadd}_{emb:xck}[\rho_I\rho_{II}] = \quad E_{XC}[\rho_I+\rho_{II}] - \sum_{z=I,II} E_{XC}[\rho_z] + T_s[\rho_I+\rho_{II}] \\ - \sum_{z=I,II} T_s[\rho_z] \tag{89}$$

The polarization of the subsystems is introduced by including an effective embedding potential in the individual subsystem calculations, which is obtained, due to the DFT formalism, as the functional derivative of the interaction energy,

$$\upsilon^{emb,I}(\mathbf{r}) = \frac{\partial R_{int}[\rho]}{\partial\rho(\mathbf{r})} \mid_{\rho=\rho_I} \tag{90}$$

In order to construct a supermolecular Lagrangian in the framework of FDE, the interaction energy E_{int} and the environment energy E_{II} are added to excited Lagrangian. As discussed earlier, however, we neglect the environment contributions of E_{II} in the spirit of uncoupled FDE (FDEu). The FDEu-ADC(2) Lagrange functional can then be written as:

$$\mathcal{L}^f_{ADC(2),exc,FDEu} = \quad \langle HF|H|HF\rangle + \sum_{pq}(D^{F,ex,f}_{pq} + D^{\kappa}_{pq})F_{pq} \\ + \frac{1}{2}\sum_{pqrs}(d^{nsep,\xi,f}_{pqrs} + d^{nsep,A,f}_{pqrs}) \times (pq|rs)_{DF} + E_{int} \tag{91}$$

where the Fock matrix $\mathbf{F}$ contains the embedding potential,

$$F_{pq} = F^{vac}_{pq} + \upsilon^{emb}_{pq} \tag{92}$$

Once the molecular orbital (MO) coefficients, coupled cluster amplitudes, multipliers, and excitation vectors, as well as the Lagrange multipliers for the orbital rotations, are computed, the gradient can be obtained as the partial derivative of the total Lagrangian, with respect to the displacement of the nuclear coordinates of the active subsystem $\vec{X}$,

$$\frac{d\mathcal{L}_{ADC(2),exc,FDEu}}{d\vec{\mathbf{X}}} = \quad \frac{\partial\mathcal{L}_{ADC(2),exc}}{\partial\vec{\mathbf{X}}} - \sum_{pq}\mathbf{F}^{eff,emb}_{pq}S^{[\vec{\mathbf{X}}]}_{pq} \\ + (\frac{\partial}{\partial\vec{\mathbf{X}}}E^{I,II}_{nuc}) + \int\rho_{II}(\mathbf{r})(\frac{\partial}{\partial\vec{\mathbf{X}}}\upsilon^I_{nuc}(\mathbf{r}))d\mathbf{r} \\ + \sum_{pq}D^{HF}_{pq}\left\langle p|\upsilon^I_{emb:coul}|q\right\rangle^{[\vec{\mathbf{X}}]} \\ + 2\sum_{pq}D^{HF}_{pq}\left\langle p^{[\vec{\mathbf{X}}]}|\upsilon^I_{emb:XCK}|q\right\rangle \\ + \sum_{pq}(D^{F,ex}_{pq} + D^{\kappa}_{pq})\left\langle p|\upsilon^I_{emb}|q\right\rangle^{[\vec{\mathbf{X}}]} \tag{93}$$

Here, $\mathbf{F}^{eff}$ is the effective Fock matrix, which accounts for the metric change upon geometric perturbations of the underlying Gaussian basis functions. The analytical nuclear excited-state gradients for conventional ADC(2), $\frac{\partial \mathcal{L}_{ADC(2),exc}}{\partial \vec{\mathbf{X}}}$, are given for the formalism used in the present work in [99] and [76].

D. Density functional theory

Density functional theory uses functional of the electron density to calculate the electronic structures of the system, which receives great amount of successes among the material and bio-chemical systems [46–50]. While the conventional electronic structure methods focus on solving Schrödinger equation of N electrons which interacting with each other in the background potential generate by the nuclei. The main challenge of this approach is although it provides a systematic way to improve the accuracy of the calculation, the computational cost grow rapidly even for medium sized system and for the high level methods such as coupled cluster. Density functional theory addresses these issues with the new approach, in which instead of dealing with N electrons with 3N coordinates, only the electron density $\rho(r)$ with 3 coordinates is used, and this fundamental difference gives the DFT methods better performances when calculate the medium even large size systems. The density of the system is defined as follow:

$$\rho(r_1) = N \int \cdots \int |\Psi(x_1, x_2, \ldots, x_N)|^2 dx_1, \ldots, dx_N \tag{94}$$

$$\tag{95}$$

with the constraint

$$\int \rho(r) dr = N \tag{96}$$

In order to put connect the electron density of the system to the energy of the system, Hohenberg-Kohn and Kohn-Sham theorems are introduced. The Hohenberg-Kohn theorems consist two theorems, which not only states that total energy of the system can be described by the electron density, but also provide a way to acquire such energy in practice. With the first Hohenberg-Kohn theorem the total energy can be written as functional of the electron density:

$$E[\rho(r)] = \int \rho(r) \upsilon_{ext}(r) dr + F[\rho(r)] \tag{97}$$

which $F[\rho(r)]$ is the universal functional with unknown form and $\upsilon_{ext}(r)$ is the external potential. Wavefunction of the system can be obtained by apply the variational principle as stated in second Hohenberg-Kohn theorem:

$$\frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \epsilon_0 \tag{98}$$

where $\epsilon_0 = E_0$ only when Ψ is the real wavefunction of the system and the Hamiltonian of the system can be written as following:

$$\hat{H} = \hat{V} + \hat{T} + \hat{V}_{ext} \tag{99}$$

$$\hat{V} = \hat{V}_{en} + \hat{V}_{nn} + \hat{V}_{ee} \tag{100}$$

where $\hat{V}_{en}$ is the electron-nuclei interaction, $\hat{V}_{nn}$ is the nuclei-nuclei interaction and $\hat{V}_{ee}$ is the electron-electron interaction, $\hat{T}$ is the kinetic energy. $\hat{V}_{ext}$ is the external potential in which all the particles are moving. Although the Hohenberg-Kohn theorems provides the fundamental of the density functional theory, there is no showing of how to calculate the properties related to the electronic structures of the system such as ground state energies from these principles. About one year after the seminal DFT paper by Hohenberg and Kohn [101], Kohn and Sham [102] proposed method to carry out the electronic calculation with density functional theory.

In Kohn-Sham (KS) formulation ground state energy can be acquired by building the following Lagrangian within a non-interacting system and minimize it:

$$0 = \int \rho(r)\upsilon_{ext}(r)dr + F[\rho(r)] - \lambda(\int \rho(r)dr - N) \tag{101}$$

By introduce the non-interacting system, KS formulation simplified the real problem into sum of independent particles, which can be solved with ease. In the non-interacting system the electronic energy can be expressed as following:

$$\begin{aligned} E[\rho(r)] &= T[\rho(r)] + V[\rho(r)] + F[\rho(r)] \\ &= T_s[\rho(r)] + T_c[\rho(r)] + V_{ee}[\rho(r)] + V_{ext}[\rho(r)] + E_{xc}[\rho(r)] \end{aligned} \tag{102}$$

where $T_s[\rho(r)]$ is the kinetic energy of non-interacting system and $T_c[\rho(r)]$ is the correlation correction since $T_s[\rho(r)]$ is defined as expectation value of Slate determinant with all the exchange interaction already described by the antisymmetrization.

$$T_s[\rho(r)] = \langle\Phi(\rho)|\hat{T}|\Phi(\rho)\rangle \tag{103}$$

$V_{ee}[\rho(r)]$ and $V_{ext}[\rho(r)]$ is classical electrostatic interaction between electrons and external potential generated by the nuclei respectively. All the non-classical electron-electron interactions are packed into one exchange-correlation term $E_{xc}[\rho(r)]$, however this term has no exact expression and remain to be the main focus on the development of the DFT method. By solving the Kohn-Sham equation (102), one can obtain the density of the non-interacting system and this density is exactly the same as the real interacting system. However by applying the KS formulation the unknown energy contribution of the system is reduced to the single exchange-correlation term, which is a successful simplification. Yet despite the fact that the exchange-correlation term is only an correction and comparably small to other terms in the KS equation, it is still an important contribution since the binding energy of many systems is about the same size as $E_{XC}[\rho(r)]$.

E. Periodic DFT

The conventional DFT made a great success calculating the electronic structures of different chemical system with different size [51–53], but for the periodic system, the application of Bloch's theorem to the Kohn-Sham formulation must be made in order to fully utilize their periodicity. The periodic system can be described mathematically by its smallest repeating unit and these unit cells are invariant under transformation.

$$\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3 \tag{104}$$

Where $\mathbf{R}$ is the lattice point, n_1, n_2 and n_3 are integers, a_1, a_2 and a_3 are lattice vectors.

1. Bloch's theorem

In crystallography the atoms within the unit cell are packed in such way that there is no overlap or empty space , and such unit cell is called Bravais lattices. For any symmetry operation applies on the wavefunction of particle in the Bravais lattice, following equation holds true in real space:

$$\mathbf{T}_{\mathbf{R}_n}\psi(r) = \psi(r + \mathbf{R}_n) = e^{i\phi_n}\psi(r) \tag{105}$$

Equation 105 shows that under symmetry operator electron at position $\mathbf{r}$ shifts by lattice vector $\mathbf{R}$. In another word wavefunction of particles in the periodic system only changes

nothing but a phase factor as following:

$$\psi_k(r) = e^{ikr}u_k(r) \tag{106}$$

with e^{ikr} is the plane wave and $u_k(r) = u_k(\mathbf{R}+r)$ is lattice periodic factor which is invariant under translations, rotations and reflections. Index $\mathbf{k}$ is a vector in reciprocal space which can be constructed for each point in real space:

$$\mathbf{k} = x_1 g_1 + x_2 g_2 + x_3 g_3 \tag{107}$$

$$g_i = 2\pi \frac{a_j \times a_j}{\Omega} \tag{108}$$

$$\Omega = |det(a_1 a_2 a_3)| \tag{109}$$

With g_i as lattice vectors in reciprocal space and Ω is the unit cell volume in real space. The relation between lattice vectors in real space a_1, a_2, a_3 and in reciprocal space g_1, g_2, g_3 is as following:

$$a_i \dot{g}_j = 2\pi\delta_{ij} \tag{110}$$

2. Kohn-Sham Equations in Plane-Wave Form

The Schrödinger equation for the periodic system can be written as following:

$$\hat{h}\psi_{nk} = \varepsilon_{nk}\psi_{nk} \tag{111}$$

For any reciprocal lattice vector $\mathbf{G}$ we have the following:

$$G = n_1 g_1 + n_2 g_2 + n_3 g_3 \tag{112}$$

where n_i is the band index.

Apply the above discussion on the conventional DFT formulation, Schrödinger equation of the system under influence of the periodic potential and with non-interacting electron can be written as following:

$$[-\frac{\hbar^2}{2m}\nabla^2 + V(r)]e^{ikr}u_{nk}(r) = E_n(k)e^{ikr}u_{nk}(r) \tag{113}$$

The periodic potential can be written as:

$$\upsilon(r) = \sum_G \bar{\upsilon}(G)e^{iGr} \tag{114}$$

$$\bar{\upsilon}(G) = \frac{1}{\Omega}\int V(r)e^{-iGr}dr \tag{115}$$

where $\bar{\upsilon}$ is the Fourier transformed quantities in reciprocal space.

Combine the equation (113), and (115) will result the Kohn-Sham equation in reciprocal space.

$$\sum_{G'}[\frac{1}{2}|k+G|^2\delta_{G,G'} \quad +\bar{\upsilon}_{ext}(G-G') + \bar{\upsilon}_{ee}(G-G') \\ +\bar{\upsilon}_{XC}(G-G')]c_{j,k+G'} = \varepsilon_i(k)c_{j,k+G} \tag{116}$$

where $\frac{1}{2}|k+G|^2\delta_{G,G'}$ is the diagonalized kinetic energy and rest of the left hand side terms are Fourier transformed external potential, electron-electron interaction and exchange-correlation interaction. It is possible that in the real case calculations only limited amount of plane wave basis set would be used despite for the fact that in theory infinite number of basis set would grantee the exact results. In practice this is done by truncating the basis set to the finite amount so only the most important part of the kinetic energy would be calculated while the computational cost would be kept at bay. The truncating of the basis set is defined by the kinetic energy cutoff E_{cut}:

$$\frac{1}{2}|k+G|^2 \leq E_{cut} \tag{117}$$

3. Brillouin Zone Sampling

In conventional DFT calculation, ground state energy of the system comes from summing all single electron contributions, however by combine with the Bloch's theorem, in periodic system the ground state energy can be acquired by only calculating electronic contributions from one unit cell in real space. Based on the discussion above, for a periodic system it is only necessary to calculate the electrons within the first Brillouin zone instead of calculating the infinity amount of the electrons. This conclusion also means that only finite amount of k-points are needed within the first Brillouin zone since wavefunctions at close k-points are nearly identical. Efficient k-point sampling schemes have been developed in the past few decades [103] and apply the symmetry of the lattice would results further reduce the k-point need for the calculation.

Thus one can write the function integration within the first Brillouin zone as following:

$$\begin{aligned} f(r) &= \frac{\Omega}{(2\pi)^3}\int_{BZ} F(k)dk \\ &= \sum_j w_j F(k_j) \end{aligned} \tag{118}$$

where $F(k)$ is the Fourier transform of $f(r)$, Ω is the cell volume and the w_j are weighting factors. By applying the symmetry of the Brillouin zone, above integral can be further simplified as following:

$$f(r) = \sum_{j=1}^{P(n_j)} w_j F(k_j) \tag{119}$$

where $P(n_j)$ is the symmetry-dependent number of points in the irreducible wedge of the Brillouin zone.

4. Pseudopotentials

Even with all the simplifications mentioned above, it is necessary to further reduce the computational cost since in the periodic calculations the system size would be rather large, combining with functional beyond local density approximation (LDA) would be impossible without them. One of the such basic simplification approach is the introduction of the pseudopotential [104–106]. Based on their chemical activity, electrons can be divided into two catalogue: valence electrons which are located far away from atomic core and able to participate most of the chemical reactions, core-electrons which are highly localized in the near core region and generally inactive during the chemical reactions.

However the main disadvantage of using plane wave basis set is the large number of basis set required for describe the system. This is due to the fact that electron wave function would form "cusp" in the near nucleus area, which requires plane waves with very short wavelengths and high cut off energy. For the most of the reciprocal space the electronic wave function is rather smooth which requires less plane wave basis function and most of the extra functions for describing the core electrons are gone waste. By introducing the pseudopotential the cusp problem of the core electron wavefunction can be avoided and computational cost will be greatly reduced.

The pseudopotential is constructed such that that the pseudo wave function has no radial nodes within the core region and the pseudo wave functions and potential agree with the real wave function and potential beyond the given cut-off radius. Except the above mentioned requirement, there are other criteria that pseudopotential must fulfil: the core charge produced by pseudo wavefunctions must be the same as the real ones; the pseudo wavefunctions must be continuous at the cut-off point; the eigenvalues of pseudo wavefunctions must be

the same as the all electron wavefunctions.

5. Auxiliary density matrix method

Despite the rapid development of varies density functional based method which made the *ab initio* calculation on the large scale gas or condensed phase and molecular dynamics (MD) possible, it also becomes increasingly clear that an improved description of the electronic structure, and thus more accurate results, can only be obtained by functionals that go beyond the GGA form. This also results the need for the incorporation of a non-local term such as Hartree-Fock exchange (HFX). The calculation of Hartree-Fock exchange (HFX) is computationally demanding for large systems with high quality basis sets and consequently, there is significant interests in finding efficient approaches to deal with these non-local terms. The efficiency of a HFX calculation depends strongly on the algorithm employed and there are different algorithms which can reduce the scaling with the system size to linear scaling, but HFX calculations still scale very poorly with basis set quality.

There are several reasons for why the HFX term is extremely costly despite the algorithms employed. First is that even in the linear scaling code, the cost scales able the N^4 which N is the number of the primitive basis function per atom. Therefore the basis set with the hight l-quantum number, i.e. polarization functions, are also costly since they scale quadratically with l. Basis function which are heavily contracted such as MOLOPT basis sets, are also expensive since for each quartet of basis functions a very larger number of primitive integrals needs to be considered. Moreover, very diffused basis sets are even more costly and they are essential in the calculation especially in the condensed phase, where periodic boundary conditions provide a potentially unlimited number of interacting atomic sites.

In order to address these problems, in 2010 Manuel Guidon Et al. proposed a new approach [107] which can greatly reduce the cost for the calculation with HFX term using high quality basis sets.

The energy expression in Kohn-Sham DFT can be written as:

$$E[\rho(r)] = T_s[\rho(r)] + J[\rho(r)] + E_{xc}[\rho(r)] + \int \nu(r)\rho(r)dr \tag{120}$$

where

$$E_{xc}[\rho(r)] = \alpha E_x^{HFX}[\Psi_i] + (1-\alpha)E_x^{DFT}[\rho] + E_c^{DFT}[\rho] \tag{121}$$

with α indicates the fraction of HFX being used in the hybrid functional, E_x and E_c are the exchange and correlation interaction respectively. Electron density of the system can be written as:

$$\rho(r) = \sum^{i} |\Psi(r)|^2 \tag{122}$$

and atomic centered basis set $\phi_\mu(r)$ can be used to express the wavefunction Ψ

$$\Psi_i(r) = \sum_{\mu} C^{\mu i} \phi_\mu(r) \tag{123}$$

Hartree-Fock exchange energy can be expressed in terms of a density matrix and two-electron integrals (ERIs) as following

$$E_x^{HFX}[P] = -\frac{1}{2} \sum_{\lambda\sigma\mu\nu} P^{\mu\sigma} P^{\nu\lambda} (\mu\nu|\lambda\sigma) \tag{124}$$

where the density matrix elements $P^{\mu\nu}$

$$P^{\mu\nu} = \sum_{i} C^{\mu i} C^{\nu i} \Leftrightarrow P = CC^T \tag{125}$$

and ERI can be written as

$$(\mu\nu|\lambda\sigma) = \iint \phi_\mu(r_1)\phi_\nu(r_1) g(|r_2 - r_1|) \phi_\lambda(r_2)\phi_\sigma(r_2) dr_1 dr_2 \tag{126}$$

with $g(r)$ as 1/r in standard Hartree-Fock theory.

Now introduce auxiliary density matrix $\hat{P} \approx P$ by choosing either smaller or more rapidly decaying basis function than the original one. New exchange interaction therefore can be written as

$$\begin{aligned} E_x^{HFX}[P] &= E_x^{HFX}[\hat{P}] + (E_x^{HFX}[P] - E_x^{HFX}[\hat{P}]) \\ &\approx E_x^{HFX}[\hat{P}] + (E_x^{DFT}[P] - E_x^{DFT}[\hat{P}]) \end{aligned} \tag{127}$$

In order for Equation (127) to stand, the assumption that the difference in the exchange energy between primary and auxiliary density matrix can be captured by a GGA functional, even in cases where GGA exchange and HFX might be qualitatively different. By introduce the correction which takes the difference between auxiliary and primary density matrix into account, this method only results very small error while the calculation efficiency has been greatly improved.

F. Frozen density embedding

1. *Partitioning the electron density*

In the frozen density embedding method, instead of dividing the system into subsystems, the electron density of the system is partitioned [108, 109].

$$\rho_{tot}(r) = \sum_i \rho_i(r) \tag{128}$$

The only constraint of partitioning the electron density is integrate into integral numbers and by fulfil this condition the charges of the system is also partitioned. The total energy of the system can be written as:

$$\begin{aligned} E[\rho_I, \rho_{II}] = E_{NN} \quad & + \int (\rho_I(r) + \rho_{II}(r))(v_I^{nuc}(r) + v_{II}^{nuc}(r))dr \\ & + \frac{1}{2}\int \frac{(\rho_I(r) + \rho_{II}(r))(\rho_I(r') + \rho_{II}(r'))}{|r - r'|} dr dr' \\ & + E_{XC}[\rho_I, \rho_{II}] + T_s[\rho_I] + T_s[\rho_{II}] + T_s^{nadd}[\rho_I, \rho_{II}] \end{aligned} \tag{129}$$

where E_{NN} is the nuclear repulsion energy, v_I^{nuc} and v_{II}^{nuc} are the electron-electron interaction between two subsystems, E_{XC} is the exchange-correlation energy, $T_s[\rho]$ is the kinetic energy of the non-interacting reference system and $T_s^{nadd}[\rho_I, \rho_{II}]$ is the non-additive kinetic energy which defined as following:

$$T_s^{nadd}[\rho_I, \rho_{II}] = T_s[\rho_I + \rho_{II}] - T_s[\rho_I] - T_s[\rho_{II}] \tag{130}$$

In order to calculate the $T_s^{nadd}[\rho_I, \rho_{II}]$ it is necessary to acquire $T_s[\rho_I + \rho_{II}]$, which can not be calculated conventionally due to the fact that electron density of the system has already been divided into ρ_I and ρ_{II}. Thus the approximate kinetic energy functional has to be applied and studies have shown that PW91k functional provide most accurate interaction energy. The non-additive exchange-correlation energy is defined as following:

$$E_{XC}^{nadd}[\rho_I, \rho_{II}] = E_{XC}[\rho_I + \rho_{II}] - E_{XC}[\rho_I] - E_{XC}[\rho_{II}] \tag{131}$$

2. *The embedding potential*

Based on the equation (129), electron density of subsystem I can be determined by minimizing the energy functional with respect to ρ_I while keep ρ_{II} fixed, which can be done

by minimizing the following Lagrangian:

$$0 = \frac{\delta}{\delta\rho_I}[E[\rho_I, \rho_{II}] + \lambda_I(\int \rho_I(r)dr - N_I)] \tag{132}$$

The electron density acquired from this equation can be expressed using the KS wavefunction of the same non-interacting system which can be calculated as following:

$$[-\frac{\nabla^2}{2} + \upsilon_{eff}[\rho_I, \rho_{II}](r)]\phi_i^I(r) = \varepsilon_i\phi_i^I(r) \tag{133}$$

and i goes up to $N_I/2$. In order to solve the equation (133), $\upsilon_{eff}[\rho_I, \rho_{II}](r)]$ has to be determined by the following:

$$E_s[\rho_I] = T_s[\rho_I] + \int \rho_I(r)\upsilon_{eff}[\rho_I, \rho_{II}](r)dr \tag{134}$$

Compare the equation (133) and (134), together with assumption that electron density of subsystem II is frozen, effective potential of whole system can be expressed as following:

$$\upsilon_{eff}[\rho_I, \rho_{II}](r) = \upsilon_{eff}[\rho_I](r) + \upsilon_{eff}^{emb}[\rho_I, \rho_{II}](r) \tag{135}$$

with $\upsilon_{eff}[\rho_I](r)$ is the effective potential of subsystem I with froze electron density of subsystem II:

$$\upsilon_{eff}[\rho_I](r) = \upsilon_I^{nuc}(r) + \int \frac{\rho_I(r')}{|r - r'|}dr' + \frac{E_{XC[\rho]}}{\delta\rho}|_{\rho=\rho_I(r)} \tag{136}$$

and $\upsilon_{eff}^{emb}[\rho_I, \rho_{II}](r)$ is the embedding potential in which contains all the inter-subsystem interactions:

$$\upsilon_{eff}^{emb}[\rho_I, \rho_{II}](r) = \upsilon_{II}^{nuc}(r) + \int \frac{\rho_{II}(r')}{|r - r'|}dr' + \frac{\delta E_{XC}^{nadd}[\rho_I, \rho_{II}]}{\delta\rho_I} + \upsilon_T[\rho_I, \rho_{II}](r) \tag{137}$$

The KS equation for the frozen embedding system can be acquired by insert equation (135) into equation (133):

$$[-\frac{\nabla^2}{2} + \upsilon_{eff}[\rho_I](r) + \upsilon_{eff}^{emb}[\rho_I, \rho_{II}](r)]\phi_i^I(r) = \varepsilon_i\phi_i^I(r) \tag{138}$$

3. Approximate the non additive kinetic energy

In conventional DFT method the largest contribution to the total energy of the system is the kinetic energy, which can be divided into non-interacting reference kinetic energy and the interacting kinetic energy that normally included in the exchange-correlation energy.

However in FDE method the non additive kinetic part can not be calculated exactly thus only approximation is possible. In principle any kinetic energy functional used in orbital free DFT can also be applied on approximating the non additive kinetic of FDE. With the following kinetic energy and potential expression:

$$T_s^{nadd}[\rho_I, \rho_{II}] = T_s[\rho_I + \rho_{II}] - T_s[\rho_I] - T_s[\rho_{II}] \tag{139}$$

$$\upsilon_T[\rho_I, \rho_{II}](r) = \frac{\delta T_s[\rho]}{\delta \rho}|_{\rho=\rho_{tot}(r)} - \frac{\delta T_s[\rho]}{\delta \rho}|_{\rho=\rho_I(r)} \tag{140}$$

It is worth noticing that accurate kinetic functional for $T_s[\rho]$ does not necessarily leads to improvement of the $\upsilon_T^{nadd}[\rho_I, \rho_{II}]$. Study shows that by compare results from supermolecular KS-DFT and FDE, PW91k functional proposed by Lembarki and Chermette gives good agreement with KS-DFT calculation and therefore widely used in different FDE implementations.

$$F_{LC94}(s) = \frac{1 + 0.093907s\ arcsinh(76.32s) + 0.26608 - 0.0809615e^{-100s^2}s^2}{1 + 0.093907s\ arcsinh(76.32s) + 0.57767 \times 10^{-4}s^4} \tag{141}$$

G. Short contact analysis

The concept of short contact is, from a theoretical viewpoint, rather weakly defined due to employing atomic van-der-Waals radii, but can serve as a qualitative discussion concerning the kinetic stability. The atoms in pair, which are taken from neighbouring molecules, can be divided into different types as shown in Fig. 1. These are denoted head-to-head and side-by-side in the present work for clarity. In the supplementary information additionally short contacts are listed occurring within a stack. In the present work, however, the short-contact distance γ is defined as:

$$\gamma_{pq} = d_{pq} - f\left(R_p^{\mathrm{vdW}} + R_q^{\mathrm{vdW}}\right), \tag{142}$$

where d_{pq} is the distance of the atoms p and q, f is the scaling factor and $R_p^{\mathrm{vdW}}, R_q^{\mathrm{vdW}}$ is the van-der-Waals radii of atoms p, q, respectively. Note that the two atoms p and q are located on different molecules so that d_{pq} corresponds always to an intermolecular distance. In the present work, for $\gamma < 0$ atoms are considered to be in short contact. By introducing the scaling factor, it is possible to set different criteria for different geometric configurations and only focus on the atoms which are most like to have interactions between each other. If not stated otherwise, the scaling factor 1.05 was used in the present work.

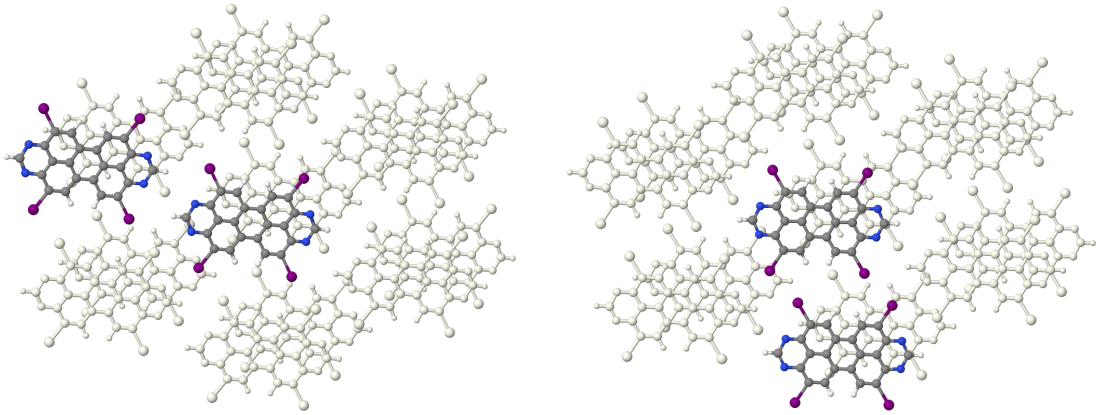


Fig. 1: *Ortho*-H-TAPP-I_4 clusters extracted from the optimized triclinic bulk phase. Highlighted dimer configurations are used to compute intermolecular distances in the present work. Left: head-to-head dimer, right: side-by-side dimer.

An additional measure for the short contact analysis is the gap Δ between short contact and non-short contact:

$$\Delta = \min_{ab} \gamma_{ab} - \max_{ij} \gamma_{ij} \,, \tag{143}$$

Where $\gamma_{ab} > 0$, i.e. pairs in long contact, and $\gamma_{ij} < 0$, i.e. pairs in short contact. Note that short contact atoms are not present in all cases.

III. COMPUTATIONAL DETAILS

In the following works, different computational settings are used for different property calculation. For the IP/EA calculation, unless stated otherwise, all calculations, in particular all FDE calculations, are carried out using KOALA [108, 109] which employs density fitting, also abbreviated as RI [110–112] using the self-consistent field (SCF) density in the embedding potential. The embedding exchange-correlation functional used was Perdew-Burke-Ernzerhof (PBE) [113] and the embedding kinetic energy functional used was PW91k [114]. Electron-detachment (IP) calculations were performed using the def2-TZVP basis set [115–117] and electron attachment (EA) calculations for neutral systems were performed with the def2-TZVPD basis set to account for accurate virtual orbitals [118]. In case of geometry-relaxed transitions, only the subsystem which is excited is relaxed, while all other subsystems' geometries are kept frozen.

The program CP2K [119] is used for the optimization of all the bulk structures of TAPP compounds. The optimization is performed using periodic density functional theory (DFT) with the PBE functional. Unless otherwise specified, the TZVP-MOLOPT-GTH basis set [120] with GTH-PBE pseudopotential [104–106] is used for all the periodic DFT calculations. DFT-D3 dispersion correction [121] is applied to account for intermolecular dispersion interaction. Both the unit cell and the molecular geometry are optimized. In the supplementary information, the results of all geometry optimizations are provided, including files in crystal information format (.cif) as well as extracted clusters.

Electronic-gap single point calculations using the B3LYP functional were carried out using the TZVP-GTH basis with the GTH-BLYP pseudopotential. Cutoff radii were chosen as half of the shortest cell dimension. Auxiliary density matrix methods (ADMM) [107] is used as MOLOPT basis sets are used, for which standard HFX becomes too expensive and non-MOLOPT basis sets are not available for fluorine. In ADMM, an auxiliary is introduced, which is used to create an auxiliary density matrix (ADM) by projection.

Due to different alkyl chains and crystal systems, the number of atoms differs in general when comparing different compounds or the same compound in different crystal systems. In order to be able to compare many different systems, in the present work we provide atomization energies (AE) with respect to unit cells ("AE per unit cell") and molecules ("AE per molecule"). To eliminate either system size effect, the atomization energy per unit cell

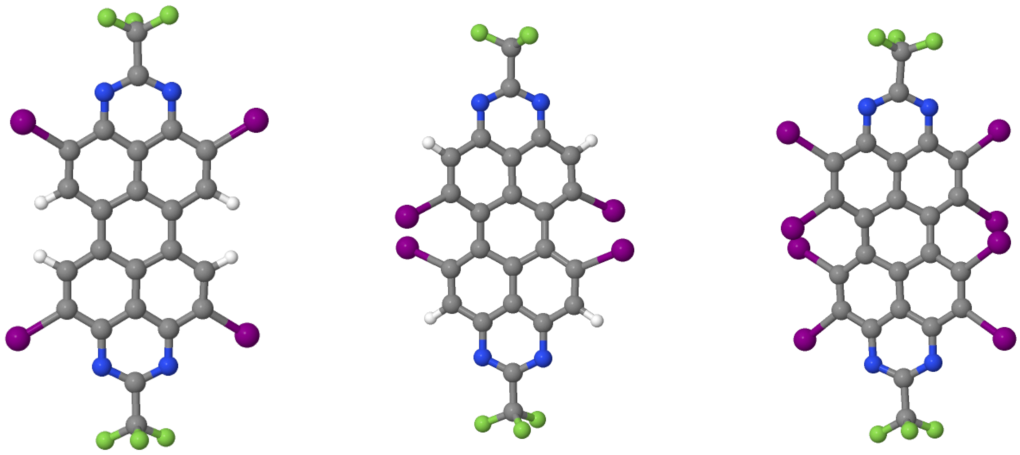


Fig. 2: The three basic TAPP configurations (left: *ortho*, middle: *bay*, right: *all*), displayed at the example of C_1F_3-TAPP-I as taken from bulk geometry optimizations. Hydrogen atoms are displayed white, carbon atoms gray, nitrogen atoms blue, fluorine atoms green, and iodine atoms purple.

is divided by the number of atoms in the cell, resulting in an average energy gain per atom ("AE average"). It should be emphasized, however, that crystal-structure prediction is a very active field in modern research and the present work only uses the tools available currently [122].

Vacuum calculations are carried out using the Turbomole program [123], employing the def2-TZVP basis set , combined with PBE and B3LYP [124–127] functional. Geometries were kept fixed to the bulk geometries and default values were used for the self-consistent field procedure.

A. Target systems

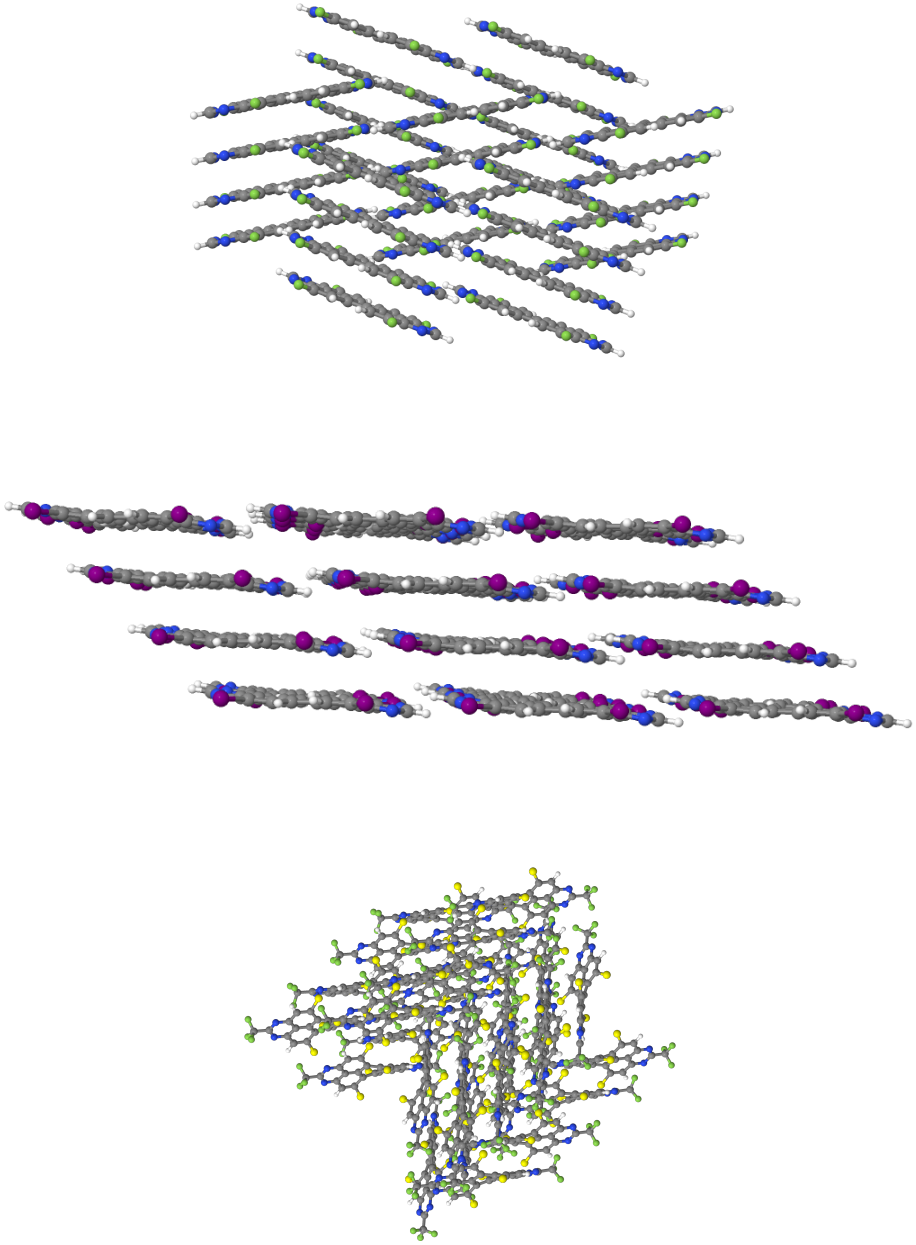


Fig. 3: From top to bottom: (a) side view of a monoclinic crystal (*ortho*-H-TAPP-F_4); (b) side view of a triclinic crystal (*ortho*-H-TAPP-I_4); (c) side view of an orthorhombic crystal (*bay*-C_1F_3-TAPP-Cl_4). Hydrogen atoms are displayed white, carbon atoms gray, nitrogen atoms blue, fluorine atoms green, chlorine atoms yellow, iodine atoms purple.

The TAPP derivatives are divided into the three categories *ortho*-, *bay*-, and *all*-TAPP, based on their substitution position, cf. Fig. 2. Reference experimental structures are taken from the Cambridge Crystallographic Data Centre (CCDC) [128] as crystal information file (cif). In this work most of the crystal system of TAPPs are either monoclinic or triclinic. Only *bay*-C_1F_3-TAPP-X compounds are observed with an orthorhombic crystal system. For TAPP derivatives that have not yet been synthesized, the cif file is manually modified from other TAPP compounds with similar structure. Fig. 3 shows these configurations using three TAPP compounds as example: *ortho*-H-TAPP-F_4, *ortho*-H-TAPP-I_4 and *bay*-C_1F_3-TAPP-Cl_4.

In order to allow for a full relaxation, the cell symmetry, *i.e.* monoclinic or triclinic or orthorhombic, were only employed in the first step but were not imposed during the optimization. The full optimizations, however, did not lead to a change of the crystal system. The largest deviation obtained is in the order of magnitude of about 0.5 degree in case of monoclinic C_1F_3-TAPP-H_8. While formally the angles deviate in case of monoclinic and orthorhomic from exact 90 degrees after the full relaxation, we nevertheless stick to the names monoclinic and orthorhombic as the deviations are significantly below 1 degree.

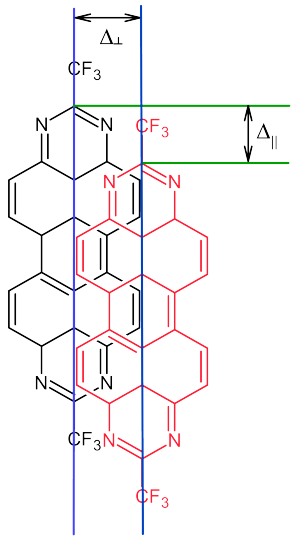


Fig. 4: Perpendicular $\Delta_\perp$ and parallel $\Delta_\parallel$ intermolecular displacement between TAPP molecules, using *ortho*-H-TAPP-H dimer as example.

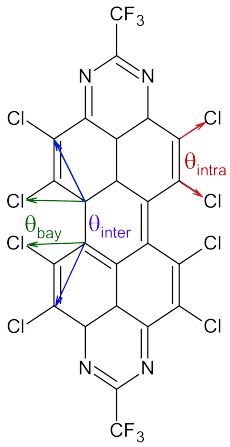


Fig. 5: Intramolecular inter-block torsion angle θ_{inter} and intra-block torsion angle θ_{intra}, using *all*-C_3F_7-TAPP-Cl_8 as example. For TAPPs with all substitution pattern, the angle between two neighbouring halogen atoms in the *bay*- position is denoted θ_{bay}.

The geometries after geometry optimization are compared with respect to intermolecular, cf. Fig. 4, and intramolecular, cf. Fig. 5, geometrical properties. The six-ring structure without any substitution is defined as the molecular plane and serves as the backbone of the analysis, and different TAPPs are defined based on the different substitutional groups on the molecular plane. Here the inter-layer distance is defined as the distance between two TAPP molecules surface, which is made of three nitrogen atoms in one TAPP molecule. The two displacements are defined based on their relations with the long molecule axis, one

perpendicular to the molecule axis is perpendicular displacement and one parallel to the molecule axis is the parallel displacement. There are three relevant intra-molecule torsion angles, as shown in Fig. 5. First, the angle between the two three-ring building blocks of same TAPP molecule, denoted inter block θ_{inter}, second, the angle between the closest halogen atoms in the neighbour building blocks, denoted θ_{bay}, and third, the angle between two neighbouring hydrogen atoms, or its substitution, denoted θ_{intra}.

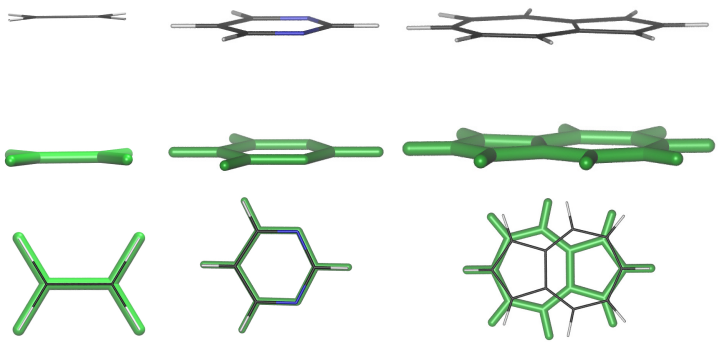
IV. ELECTRONIC PROPERTIES FOR THE CLUSTERS

In the section, the newly reported method for the calculation of vertical as well as relaxed IP and EA energies in the presence of explicit molecular environments using FDE is discussed. This approach allows for an efficient calculation of ionized systems with a large size such as TAPP com- pounds, and the results obtained this way are more suitable for analyzing molecular properties. The results of the small test systems will be discuss first in order to establish the reliability of the method, followed by the IP and EA results of the TAPP system. In this section only the results will be discussed and the method itself has already been discussed in the Section II.

A. Small reference systems

The small test systems which share the structural similarities of building blocks of TAPP molecules are tested with this approach in order to verify the reliability and the accuracy of the method.

Fig. 6: Reference dimer of ethylene, pyrimidine and azulene molecules used in validate the IP/EA method.



a. Reference values Reference results at the ΔCCSD(T) level of theory were obtained using TURBOMOLE. In order to avoid problems due to spin contamination, in addition, results using a continuum orbital at the coupled-cluster singles and doubles (CCSD) level of theory are reported. For these numerical calculations, a continuum orbital with an exponent of 10^{-10} was used and has been checked using the CFOUR program to confirm the use of the RI approximation. In case of IP calculations, the target state is the one which is dominated by the excitation from the HOMO to the continuum orbital. In case of EA calculations, two additional electrons were included in the calculation, leading to an occupation of the continuum orbital with two electrons. The EA value was then determined as the singlet excitation, which is dominated by the transition from the continuum orbital into the valence LUMO of the molecular system.

b. Ionization potentials In order to assess the accuracy of the present approach employing FDE, ethylene and pyrimide and selected dimers are computed, for which results are presented in Table I to II, including the results from the reference methods. Ethylene and pyrimide were chosen to investigate the performance on small structures that can be considered the basic building blocks of, e.g., TAPPs.

Starting with the monomer ethylene molecule, Table I shows that the methods IP-CIS(D_∞), IP-ADC(2), and IP-CC2 yield a sufficient agreement within 0.05 eV. The reference methods, i.e., CCSD and ΔCCSD(T), agree among each other within 0.02 eV, which points to a small influence of the triples. The reference values differ to the second-order approximate methods by about 0.55 eV, which is in agreement with a benchmark study. However, in case of supermolecular calculations on the dimers, the second-order methods are able

Tab. I: Vertical IP reference results: ethylene

Molecule	System	Distance	IP-ADC(2)	IP-CC2	CCSD	ΔCCSD(T)
Ethylene	Superm.	3.0 Å	8.91 (-1.36)	8.87 (-1.35)	9.30 (-1.32)	9.30 (-1.29)
		5.0 Å	10.09 (-0.18)	10.05 (-0.17)	10.45 (-0.17)	10.44 (-0.15)
		7.5 Å	10.24 (-0.03)	10.19 (-0.03)	10.59 (-0.03)	10.57 (-0.02)
	FDE	3.0 Å	10.01 (-0.26)	9.96 (-0.26)	-	-
		5.0 Å	10.19 (-0.08)	10.14 (-0.08)	-	-
		7.5 Å	10.24 (-0.03)	10.20 (-0.02)	-	-
	Monomer		10.27	10.22	10.62	10.59

Tab. II: Vertical IP reference results: pyrimidine*

Molecule	System	Distance	IP-ADC(2)	IP-CC2
Pyrimidine	Superm.	3.0 Å	8.23 (-0.45)	8.17 (-0.43)
		5.0 Å	8.59 (-0.09)	8.52 (-0.08)
		7.5 Å	8.65 (-0.03)	8.58 (-0.02)
	FDE	3.0 Å	8.57 (-0.11)	8.49 (-0.11)
		5.0 Å	8.63 (-0.05)	8.56 (-0.04)
		7.5 Å	8.66 (-0.02)	8.58 (-0.02)
	Monomer		8.68	8.60

* CCSD and ΔCCSD(T) results for monomer are 9.73 and 10.57 respectively

to reproduce the distance dependence of the reference methods. For example, in case of IP-ADC(2) at 3.0 Å distance, a shift of -1.36 eV with respect to the monomer is obtained, which shows a reduced deviation of about 0.04 eV to the shift obtained with CCSD, which is computed to be -1.32 eV.

In case of pyrimidine, as shown in Table II, the ΔCCSD(T) is not a sufficiently accurate reference method since the spin contamination at the SCF level of theory becomes significantly large: $\langle S^2 \rangle_{\mathrm{SCF}} = 0.917$; see Table II. However, the vertical IP of the monomer,

obtained from numerical CCSD, exhibits a deviation of about 1 eV to IP-ADC(2) and IP-CC2. Table II furthermore shows that the interaction in the pyrimidine dimer is significantly less pronounced compared to ethylene since the supermolecular IP at a distance of 3 Å is 8.23 eV, which differs from the monomer value of 8.68 eV by less than 0.5 eV.

Finally, Table I to II also contains values for dimers obtained by employing the FDE ansatz. Comparing these values to the supermolecular dimer calculations, it can be seen that with increasing distance, the supermolecular calculations and the FDE approach both converge smoothly to the monomer result. For the rather short distance of 3 Å , the supermolecular calculations and the FDE results differ by almost 1 eV in case of ethylene. The reason for this is rooted in the FDE ansatz itself, which localizes the charge on one fragment, while in case of the supermolecular calculations, the charge is evenly distributed among the two molecules.

Tab. III: Geometry-relaxed IP reference results: ethylene and pyrimidine dimer

Dimer	System	Distance	IP-CIS(D_∞)	IP-ADC(2)
Ethylene	Superm.	3.0 Å	8.84 (-1.19)	8.84 (-1.19)
		5.0 Å	9.92 (-0.11)	9.92 (-0.11)
		7.5 Å	10.00 (-0.03)	10.00(-0.03)
	FDE	3.0 Å	9.78 (-0.25)	9.78 (-0.25)
		5.0 Å	9.95 (-0.08)	9.95 (-0.08)
		7.5 Å	10.01 (-0.02)	10.01 (-0.02)
	Monomer		10.03	10.03
Pyrimidine	Superm.	3.0 Å	7.22 (-0.34)	7.22 (-0.34)
		5.0 Å	7.48 (-0.08)	7.48 (-0.08)
		7.5 Å	7.53 (-0.03)	7.53 (-0.03)
	FDE	3.0 Å	7.47 (-0.09)	7.47 (-0.09)
		5.0 Å	7.52 (-0.04)	7.52 (-0.04)
		7.5 Å	7.54 (-0.02)	7.54 (-0.02)
	Monomer		7.56	7.56

Geometry relaxation Having assessed the accuracy of the methods for vertical transitions, we now turn to adiabatic transitions for relaxed geometries, that is, the relaxed vertical ionization. The results in Table III are obtained from the optimization of one molecule, while the second one was frozen (in case of dimers). It should be noted that this decision is straightforward for FDE in which the charge is localized on one subsystem; it does not enable a one-to-one comparison to supermolecular calculations as in these the charge might be distributed evenly over the subsystems.

In case of ethylene, the monomer value relaxes by about 0.24–10.03 eV. This rather small change indicates that the geometry is not significantly perturbed in this case after relaxation in the IP state. Similar to the purely vertical excitations, in case of the supermolecular calculations, the relative influence of the second molecule is reducing the IPs even further to 8.78 eV of which the relative shift to the relaxed monomer is similar to the vertical transition.

In case of pyrimide, the monomer value relaxes significantly by -1.15 eV to 7.56 eV, and accordingly, geometry changes can be observed in the molecule. For example, the angle $\angle$(NCN) reduces from 127° to 113° , while the nitrogen–nitrogen distance reduces by about 0.2 Å from 2.39 Å to 2.21 Å. However, similar to ethylene and the purely vertical transitions, for pyrimidine, the dimer exhibits no strong interactions so that even in the supermolecular approach, the change in the IP is significantly less pronounced.

Finally, the active subsystems, at which the charge is localized, are optimized, while the environment is frozen using the FDE ansatz. Similar to the purely vertical transitions, the FDE ansatz reduces the interaction although the reduction goes even further for the relaxed systems. For example, in case of pyrimidine, the CIS(D_∞) vertical transition is reduced from -0.45 eV to -0.11 eV, while in case of the relaxed geometries, it is reduced from -0.34 eV to -0.09 eV.

Tab. IV: Vertical EA reference results: azulene

Molecule	System	Distance	EA-ADC(2)	EA-CC2	EA-CIS(D_∞)
Azulene	Superm.	3.0 Å	−1.42(−0.41)	−1.31(−0.43)	−1.42(−0.41)
		5.0 Å	−0.98(+0.03)	−0.86(+0.02)	−0.98(+0.03)
		7.5 Å	−0.97(+0.04)	−0.84(+0.04)	−0.97(+0.04)
	FDE	3.0 Å	−0.79(+0.22)	−0.66(+0.22)	−0.78(+0.23)
		5.0 Å	−0.89(+0.12)	−0.76(+0.12)	−0.89(+0.12)
		7.5 Å	−0.95(+0.06)	−0.88(+0.00)	−0.95(+0.06)
	Monomer		−1.01	−0.88	−1.01

c. Electron attachment energies In order to investigate the accuracy of the present approach for electron attachment energies, in the present work, the azulene molecule is discussed, since the ethylene anion is not stable. Azulene is well characterized and has a stable anion as well as shares the similarity of TAPP building block and also experimentally confirmed. Results for this molecule are collected in Table V, including monomer results, dimer results, see Fig 6, and values obtained using the reference methods.

The target methods are in qualitative agreement with the experimental adiabatic electron attachment energy of -0.8 ± 0.1 eV. The reference methods ΔCCSD(T) and CCSD yield a significantly smaller vertical transition energy of about -0.5 eV for the employed geometry, which might be rooted in comparing a computed vertical transition with a measured adiabatic transition. Altogether, the supermolecular calculations seem to differ strongly from the FDE results. For example, the dimer at 3.0 Å has a shift of about -0.4 in the supermolecular approach, while it is $\sim$+0.2 eV using FDE.

In order to analyze the azulene case more closely, we have computed atomic charges by fitting these to reproduce the electrostatic potential (ESP) of the orbital-relaxed EA-ADC(2) densities. Results are collected in Table VI. The atomic charges show that in case of the supermolecular dimer calculation, the charge is evenly distributed among the subsystems, while in case of FDE, the charge is forced on one subsystem. This charge separation is clearly reflected in the atomic charges. For example, while in case of supermolecular EA-ADC(2), the charges are "symmetric" and significantly differ from both the charged

and neutral monomer values; in case of EA-ADC(2) with FDE, the monomer charges are retained.

Tab. V: Geometry-relaxed EA reference results: azulene

Molecule	System	Distance	EA-ADC(2)	EA-CIS(D_∞)
Azulene	Superm.	3.0 Å	−1.53(−0.20)	−1.53(−0.20)
		5.0 Å	−1.28(+0.05)	−1.28(+0.05)
		7.5 Å	−1.29(+0.04)	−1.29(+0.04)
	FDE	3.0 Å	−1.11(+0.22)	−1.11(+0.22)
		5.0 Å	−1.21(+0.12)	−1.21(+0.12)
		7.5 Å	−1.27(+0.06)	−1.27(+0.06)
	Monomer		−1.33	−1.33

Geometry relaxation Having assessed the accuracy of the methods for vertical electron attachment energy, we again turn to transitions for relaxed geometries, that is, the relaxation energy due to electron attachment. Results collected in Table V are obtained from the optimization of one molecule, while the second one was frozen (in case of dimers). Similar to the IP case, this strategy is straightforward for FDE in which the charge is localized on one subsystem, but it does not enable a one-to-one comparison to supermolecular calculations as in these the charge is distributed evenly over the subsystems, as discussed in the context of Table VI.

The transition energy of azulene is reduced by about −0.25 eV to −1.33 eV. In case of a neutral system, the supermolecular results show a strong interaction at 3.0 Å and they converge to the monomer results with increasing dimer distance. However, it can be seen that in the supermolecular calculations, the sign of the shift is also inverted for 5.0 Å and 7.5 Å . The reason might be rooted in a localization of the electron on one molecule. Using FDE, the interaction is reduced, and similar to the purely vertical transitions, the sign is inverted even for the close distance of 3.0 Å.

Tab. VI: Atomic charges in e from an ESP fit based on the orbital-relaxed EA-ADC(2) density (using a neutral reference) for an azulene monomer and dimer (at 3.0 Å).

Atom	Monomer		Superm. MP2 [a]		MP2–in–MP2 [a]		Superm. EA-ADC(2)		EA-ADC(2)–in–MP2	
	I	II	I	II	I	II	I	II	I	II
C_1	–0.43	–0.18[a]	–0.15	–0.14	–0.17	–0.17	–0.12	–0.11	–0.44	–0.18[a]
C_4	+0.40	+0.31[a]	+0.32	+0.32	+0.31	+0.31	+0.40	+0.39	+0.42	+0.32[a]
C_5	–0.64	–0.53[a]	–0.57	–0.57	–0.53	–0.53	–0.70	–0.69	–0.63	–0.53[a]
H_1	+0.11	+0.14[a]	+0.14	+0.14	+0.13	+0.13	+0.12	+0.12	+0.10	+0.13[a]
H_2	+0.03	+0.10[a]	+0.12	+0.12	+0.08	+0.08	+0.12	+0.13	+0.01	+0.08[a]
H_5	+0.20	+0.21[a]	+0.23	+0.23	+0.20	+0.20	+0.25	+0.25	+0.19	+0.20[a]
$\sum$[b]	–1.0	0.0	0.0	0.0	0.0	0.0	–0.5	–0.5	–1.0	0.0

[a] ESP charges obtained from orbital-relaxed MP2.

[b] Sum of the system's atomic charges.

d. Charge distribution analysis of azulene dimer In order to further study the charge localization effect when apply the FDE method, here we use azulene dimer as example. Based on the Table VI, in the supermolecule approach the charge is evenly divided among two subsystems while in the FDE approach the charge is localized on one of the subsystems.

B. Tetraazaperopyrenes

In this section, we will study the properties and in particular their change due to molecular environments of tetraazaperopyrene (TAPP) monomer and small cluster derivatives using this IP/EA approach. TAPP compounds can be tuned using two different substituents X and R, with different substitution there are three different TAPP compounds; see Fig 11 and also Fig 2 in section III. The substituents R have been shown to be important for the packing of the crystal structure but rather unimportant for excited states since the alkyl chains do not participate in the lowest excited state. The substituents X tune the properties of the valence electrons, in particular excited states and changes such as the absorption spectra. In order to quantify the impact of the different substituents for IP and EA values,

we, therefore, investigate the properties of all monomer compounds R-TAPP-X and the selected clusters of H-TAPP-H. In this work we will focus on the TAPP monomer and small clusters from dimer up to 21-TAPP cluster with 756 atoms in total.

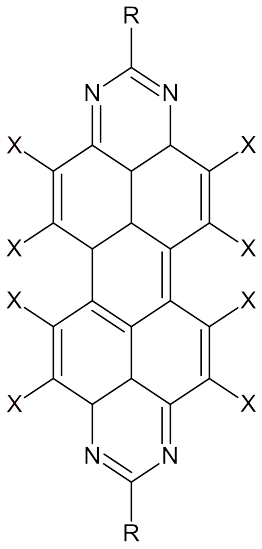


Fig. 7: TAPP molecule with R as different alkyl groups and X indicate the different halogen atoms. Only in all TAPP halogen substitution happens on all 8 X positions.

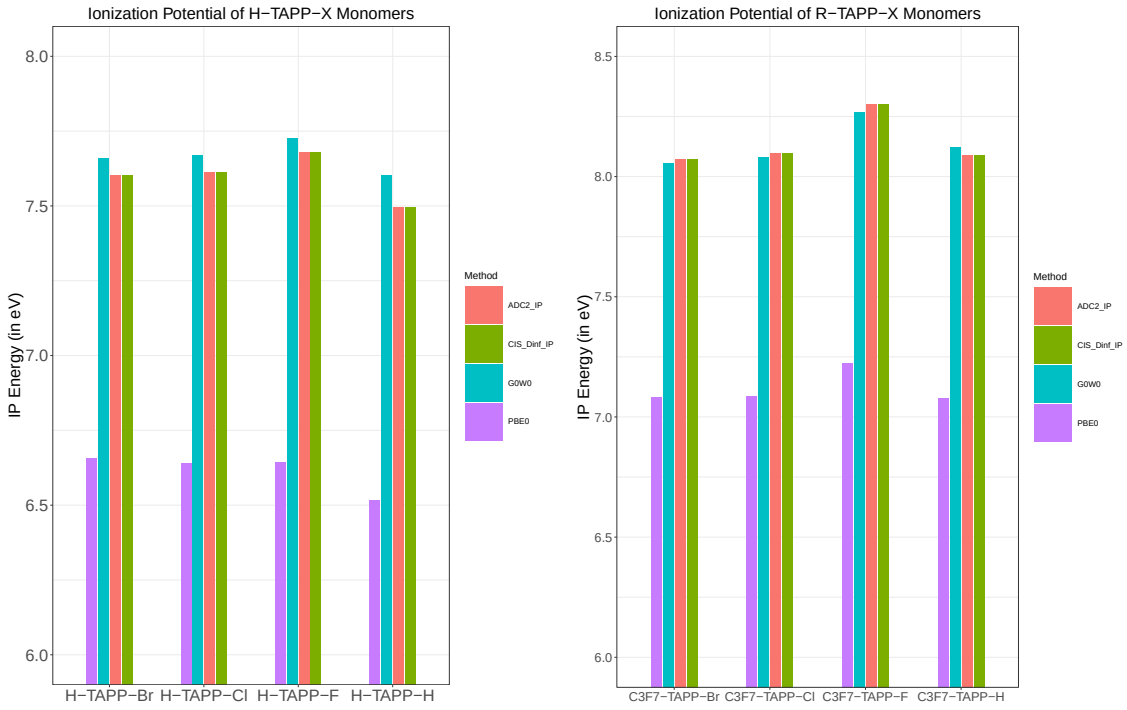


Fig. 8: Comparison of TAPP monomer vertical ionization potential calculated with different methods.

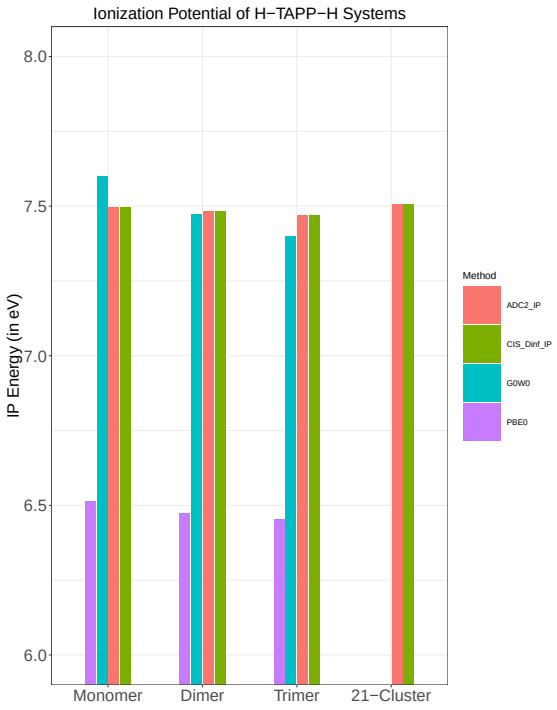


Fig. 9: Comparison of vertical ionization potential of H-TAPP-H cluster with different size, calculated using different methods.

1. Ionization potentials

Ionization potential results for different TAPP systems are collected in Table LVII and visualized in Fig 8. First of all, it can be seen that the result of the present work is in good agreement with G_0W_0 calculations, agreeing within about 0.05 eV for almost all cases. One exception is the H-TAPP-H for which the G_0W_0 method shows a value which is shifted by about 0.1 eV. The PBE0 results, i.e., the HOMO energies, show rather large errors with deviations of about 1 eV to G_0W_0 and the wavefunction methods, revealing that orbital energies can only be used as a qualitative measure if they have to be used.

The results also shows the influence of the different substituents upon the IP. While the substituents X has no significant effect on the IP, the substituents R yield an almost constant shift of $\sim$0.5 eV to higher IP values. In a single-particle picture, the substituents R can be understood to shift both the HOMO and the LUMO so that a significant shift for the IP (and EA) is obtained, while the lowest electronically excited states are not affected by different substituents R.

Table LVII also contains values for increased models of TAPP crystals, namely, the dimer,

trimer, and a cluster containing 21 H-TAPP-H molecules in total and it is visualized in Fig 9. For the dimer and trimer, the G_0W_0 method predicts a pronounced change in the IP, while the wavefunction methods, computed using FDE, yield a reduced dependence. This can, however, be related to the FDE approximation.

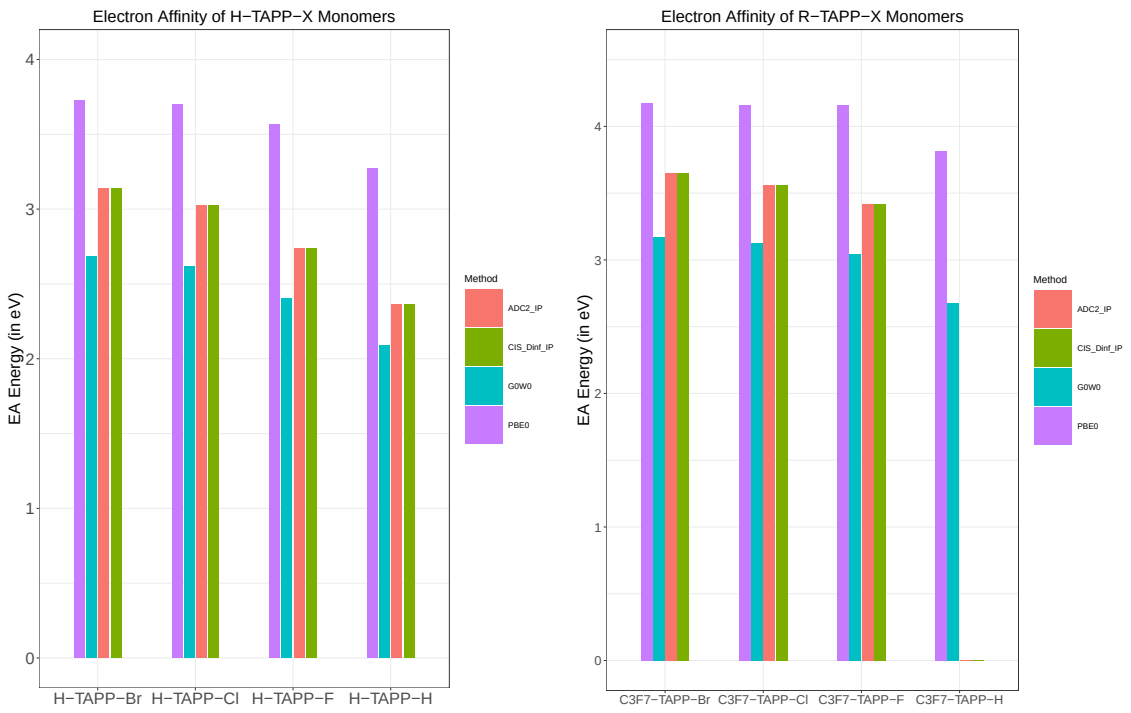


Fig. 10: Comparison of TAPP monomer vertical electron attachment calculated with different methods.

2. Electron attachment

Results for electron attachments for different TAPP systems are collected in Table LVIII and visualized in Fig 10. Similar to the IP case, it can be seen that the results of the present work are in good agreement with G_0W_0 calculations, agreeing with an almost constant shift of about 0.5 eV. The PBE0 results, i.e., the LUMO energies, show rather large errors with deviations of about 1 eV to G_0W_0 and the wavefunction methods, revealing that orbital energies can only be used as a qualitative measure if they can be used at all.

The results in Table LVIII also shows the influence of the different substituents upon the EA. In contrast to IPs, now the substituents X have a pronounced effect on the EA, while the substituents R yield again an almost constant shift to higher EA values.

Table LVIII also contains values for increased models of TAPP compounds, namely, the dimer, trimer, and a cluster containing 21 H-TAPP-H molecules in total. For the dimer and trimer, the G_0W_0 method predicts a pronounced change in the EA compared to the wavefunction methods, computed with FDE, yielding a reduced dependence. This can, however, be related to the FDE approximation. For example, calculating a dimer as an

Tab. VII: H-TAPP-H geometry-relaxed ionization and electron-attachment energies*

	IP		EA	
TAPP	IP-CIS(D_∞)	IP-ADC(2)	EA-CIS(D_∞)	EA-ADC(2)
Monomer	7.33 (-0.17)	7.33 (-0.17)	-2.74 (-0.38)	-2.74 (-0.38)
Dimer	7.31 (-0.17)	7.31 (-0.17)	-2.52 (-0.38)	-2.52 (-0.37)

* Results are given in eV

active subsystem in a cluster consisting of 28 H-TAPP-H molecules using ADC(2) FDE yields an EA value of -2.54 eV, which corresponds to a shift of about -0.4 eV compared to the monomer value of -2.14 eV. This value of -0.4 eV can be related to the strong shift of about -0.28 eV in case of supermolecular G_0W_0 , i.e., from -1.80 eV to -2.08 eV in the dimer without environment cluster contributions.

3. Geometry relaxation

Finally, we have studied the effect of geometry relaxation for the H-TAPP-H compound after electron loss and attachment. Results are collected in Table VII. This table shows that the geometry relaxation yields an energy decrease of about -0.2 eV in case of IPs and a slightly more pronounced energy decrease of about -0.4 eV in case of EAs. However, the charges do not lead to significant changes in the molecular geometry, which is desirable for efficient organic semiconductors.

V. BULK PHASE: GEOMETRIC ANALYSIS

A. Experimental and optimized structures: comparison and analysis

The present work is concerned with the structure determination of various TAPP compounds in the solid state. Table LIX gives information of all the TAPP compounds with their crystal structure and CCDC index number if experimental structure is available. This section will be structured as following. We start by assessing the accuracy of the computational setup by comparing geometric data of the computed results to the experimentally obtained crystal structures. After the validation of our approach, further optimization of all the TAPP structures listed in the Table LIX will be preformed using the same computational setting and the cell parameters of optimized structures will be compared with respect to the different substitution groups in order to study the influence of the functional group on the bare TAPP. After obtain all the optimized bulk structures, we also analysis the shifts between two layers in the bulk since this would potentially affect the properties related to the semiconducting ability such as charge mobility.

Tab. VIII: Geometric data of experimental and corresponding optimized structures for planar TAPPs, Angles are given in degrees, distances in Å.

Ortho	H-TAPP-H_8		C_3F_7-TAPP-H_8		C_2F_5-TAPP-H_4Cl_4		C_2F_5-TAPP-H_4Br_4	
	m*		t*		t*		t*	
	Exp. [129]	This work	Exp. [130]	This work	Exp. [130]	This work	Exp. [131]	This work
CCDC number	747374	-	844501	-	844502	-	911239	-
Inter-layer dist.	3.37	3.00	3.51	3.02	3.39	3.10	3.39	3.05
$\Delta_\perp$	1.54	3.71	0.20	1.06	1.03	1.17	1.17	1.21
$\Delta_\parallel$	0.19	0.09	3.42	3.47	3.45	3.32	3.49	3.38
θ_{inter}	1	0	1	7	2	12	2	12
θ_{intra}	2	1	0	5	2	6	2	7
α	90	90	111	114	82	83	91	91
β	97	89	91	85	82	83	95	95
γ	90	90	102	98	87	92	91	87
a	3.71	4.78	4.91	4.72	4.95	4.63	5.01	4.70
b	11.25	10.67	9.60	10.05	9.48	9.49	9.78	9.73
c	16.66	13.21	13.35	12.72	13.06	13.29	13.01	13.15

Ortho	C_3F_7-TAPP-H_4F_4		C_3F_7-TAPP-H_4Cl_4		C_3F_7-TAPP-H_4Br_4		C_3F_7-TAPP-H_4I_4	
	m*		t*		t*		t*	
	Exp. [38]	This work	Exp. [130]	This work	Exp. [131]	This work	Exp. [38]	This work
CCDC number	1406814	-	844503	-	911240	-	1406815	-
Inter-layer dist.	6.88	6.92	3.40	3.04	3.41	3.10	3.42	3.28
$\Delta_\perp$	5.55	5.67	1.04	1.14	1.02	1.08	0.94	1.03
$\Delta_\parallel$	5.04	4.93	3.45	3.36	3.49	3.44	3.59	3.57
θ_{inter}	2	1	2	13	3	16	0	18
θ_{intra}	1	0	2	7	1	10	2	12
α	90	90	103	104	104	105	106	107
β	95	95	95	95	95	95	94	94
γ	90	90	92	88	91	88	91	90
a	10.17	10.26	4.95	4.68	4.99	4.76	5.04	4.96
b	10.14	10.23	9.88	9.90	10.12	10.18	10.55	10.69
c	16.47	16.58	13.66	13.92	13.90	14.00	14.31	14.38

* Experimentally obtained crystal system

Tab. IX: Geometric properties of experimental and corresponding optimized structures for bay TAPP, Angles are given in degrees, distances in Å

Bay TAPP-H_4Cl_4	C_1F_3- Orthorhombic		C_2F_5- Monoclinic		C_3F_7- Triclinic	
	Exp.	This work	Exp.	This work	Exp.	This work
CCDC number	1908780 [132]	-	1908781 [132]	-	1908782 [132]	-
Inter-layer dist.	3.90	3.96	3.60	3.70	3.70	3.79
$\Delta_\perp$	0.30	0.47	0.75	0.73	0.44	0.48
$\Delta_\parallel$	5.10	5.01	4.96	4.95	5.47	5.54
θ_{inter}	10	7	8	7	13	11
θ_{intra}	9	7	7	7	11	7
θ_{bay}	53	52	55	56	55	55
α	90	90	90	90	68	69
β	90	90	98	97	83	84
γ	90	90	90	90	79	79
a	15.09	15.19	12.53	12.64	9.87	9.94
b	12.56	12.57	13.36	13.45	12.51	12.69
c	22.17	22.53	15.00	15.23	13.99	14.11

Tab. X: Geometric properties of experimental and corresponding optimized structures for all TAPP, Angles are given in degrees, distances in Å

All	C_3F_7-TAPP-Cl_8 Monoclinic		C_3F_7-TAPP-Cl_4Br_4 Monoclinic	
	Exp.	This work	Exp.	This work
CCDC number	1980430 [35]	-	1980431 [35]	-
Inter-layer dist.	4.67	4.86	4.43	4.46
$\Delta_\perp$	2.23	2.19	2.72	2.72
$\Delta_\parallel$	4.97	4.96	4.86	4.93
θ_{inter}	6	9	3	5
θ_{intra}	8	4	13	11
θ_{bay}	55	53	55	55
α	90	90	90	90
β	103	104	98	98
γ	90	90	90	90
a	14.85	15.17	13.37	13.51
b	15.61	15.78	15.88	15.99
c	13.41	13.60	16.04	16.12

Table VIII to Table X list the comparison between experimental structure and the optimized ones. In the case of the ortho TAPP, there is no constraint on the cell parameters when optimize the cell structures and the results show very little change compare to the experimental ones. The cell angles are kept constant in most of the cases, the largest change is 6° in case of the C_3F_7-TAPP-H_8. The changes of the cell length are also insignificant since the largest change happens on the H-TAPP-H_8 and only with 27% increase. As for the inter layer shift in different TAPP bulk, since all the optimization are unrestricted, all the TAPPs with monoclinic structures have reduced the inter-layer distance while all the TAPPs with triclinic ones have increased inter-layer distance. However for perpendicular and parallel shifts, except for the case of H-TAPP-H_8 and C_3F_7-TAPP-H_8, there is no noticeable changes, which indicate the stability of synthesised structure. Further analysis of thermostability will be discussed in the Chapter VI. The perpendicular shift change in case of C_3F_7-TAPP-H_8 in Table VIII can be explained by the large alkyl group force the molecules in the neighbouring layers move away in the perpendicular direction to avoid repulsive interaction. However the extremely pronounced perpendicular shift in case of the H-TAPP-H_8 in the same table is likely due to the rather large change on the cell parameters, which shows significant increase on the a direction. For the TAPP molecule itself, two most interesting geometric properties are inter- and intra-block angle. It is clear that for all the experimental structures, both angles are around 0° which are to be expected for the ortho-TAPPs. This agreements hold true for H-TAPP-H_8, C_3F_7-TAPP-H_8 and C_3F_7-TAPP-F_4H_4 while for the rest of the ortho-TAPPs, noticeable diverge around 10° can be observed. This effect, however, is not as pronounced in case of the TAPPs with halogenes in the bay position, most likely due to the steeper potential energy surface (PES).

In case of the bay TAPP, as shown in Table IX, all the cell parameters are almost kept constant. Even for the inter-layer distance and two different shifts, the change between experimental structures and optimized structures are negligible. In bay and all TAPP compounds, as shown in Fig 5, three different angles are present. Unlike the experimental ortho-TAPP compounds, bare TAPP in bay-TAPP molecules do not have flat surface and the experimental and optimized structures share the good agreements on all three angles.

Similar to the bay TAPP case, cell parameters in all TAPP compounds show great agreements between experimental and optimized results. Same agreements also occurs on inter-layer distance and inter-layer shifts, as well as all three intra-molecular angles.

In general, we observe a sufficient agreement of the measured and computed geometries. In particular when comparing different TAPP compounds among each other, the shifts from one compound to the other are reproduced.

B. Unit cell analysis

In this part the cell parameters of all the optimized TAPP compounds will be present in tables followed by the short analysis for each table. Only the optimized cell parameters will be discussed in this part, the comparison between the experimental and optimized cell parameters have been done in the previous section. All the tables are organized in such way that first TAPP-H_8 compounds will be discussed, followed by the ortho-, bay- and all-TAPP compounds. For each TAPP kind, based on the different alkyl substitution group, monoclinic and triclinic structures are listed.

a. TAPP-H_8 compounds

Tab. XI: Cell parameters for R-TAPP-H_8. Angles are given in degrees, distances in Å.

TAPP-H_8	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic*	monoclinic	monoclinic	monoclinic
Volume	672	1660	1686	1713
α	90	91	90	90
β	89	92	95	95
γ	90	89	90	90
a	4.78	10.26	10.18	10.21
b	10.67	9.81	10.15	10.24
c	13.21	16.50	16.39	16.45
Geom.	triclinic†	triclinic	triclinic	triclinic
Volume	-	433	502	547
α	-	77	80	114
β	-	83	83	85
γ	-	91	92	98
a	-	4.57	4.63	4.72
b	-	7.42	8.07	10.05
c	-	13.21	13.76	12.72

* Experimentally-obtained crystal system.

† Cell optimization not converged.

In case of the R-TAPP-H_8 compounds, optimized monoclinic structures shows the unit cells are cuboid for all the R from H to C_3F_7. In the triclinic cases, the unit cells are twisted and show no regularity. Changing the alkyl group however, shows little effect on monoclinic unit cells both angle and cell vector length wise except from H-TAPP-H_8 to the R-TAPP-H_8, which shows significant increase on the cell vector length. As for the triclinic cases the change of the alkyl group shows more noticeable effect on the size of the unit cells compare to the monoclinic cases. The size differences between the monoclinic and triclinic unit cells are largely due to the different amount atoms in the unit cells, which in monoclinic cases the atoms in the unit cells are almost as twice as in triclinic cases.

b. ortho compounds

Tab. XII: Cell parameters for ortho TAPP. Angles are given in degrees, distances in Å.

ortho TAPP-F_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic*
Volume	713	1713	1714	1733
α	90	90	90	90
β	84	94	95	95
γ	90	90	90	90
a	4.63	10.25	10.20	10.22
b	11.53	10.16	10.20	10.23
c	13.45	16.51	16.53	16.58
Geom.	triclinic	triclinic	triclinic	triclinic
Volume	598	524	534	594
α	106	82	82	103
β	98	84	83	95
γ	84	94	92	88
a	4.50	4.52	4.58	4.63
b	9.65	9.02	9.02	9.62
c	14.53	13.10	13.15	13.77

* Experimentally-obtained crystal system.

Table XII shows optimized cell information of the ortho TAPP-F_4H_4 compounds. Similar as in the TAPP-H_8 cases, for the monoclinic crystals even with the change of the alkyl groups the shape of the unit cells are kept since the unit cells angle changes very little. As for the unit cell vectors, except from the H-TAPP to R-TAPP, there is almost no change. However in the triclinic cases, unit cell size of R-TAPP shrink as cell vector b and c reduce in length compare to the H-TAPP-F_4H_4 and C_3F_7-TAPP-F_4H_4. The possible reason for this difference might be that C_1F_3- and C_2F_5-TAPP are artificially made from C_2F_5-TAPP-Cl_4H_4 while C_3F_7-TAPP-F_4H_4 is made from C_3F_7-TAPP-Cl_4H_4.

Tab. XIII: Cell parameters for ortho TAPP. Angles are given in degrees, distances in Å.

ortho TAPP-Cl_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
Volume	813	1805	1849	1889
α	90	90	90	90
β	96	95	96	98
γ	90	90	90	90
a	4.71	10.80	10.32	10.31
b	11.67	9.94	10.18	10.49
c	14.86	16.87	17.69	17.65
Geom.	triclinic	triclinic	triclinic*	triclinic*
Volume	398	589	575	622
α	110	104	83	104
β	101	95	83	95
γ	86	86	92	88
a	4.59	4.60	4.63	4.68
b	8.87	9.60	9.49	9.90
c	10.61	13.78	13.29	13.92

* Experimentally-obtained crystal system.

Ortho TAPP-Cl_4H_4 compounds shows the similar pattern as in ortho TAPP-F_4H_4, in which the unit cells of monoclinic crystals are increased with the increase of the alkyl group size while the angles of the unit cells are kept almost constant. However also similar to the ortho TAPP-F_4H_4 case, in the triclinic crystals, unit cell parameters have not shown any patterns rather change around C_2F_5-TAPP-Cl_4H_4.

Tab. XIV: Cell parameters for ortho TAPP. Angles are given in degrees, distances in Å.

ortho TAPP-Br_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
Volume	857	1863	1890	1918
α	90	90	90	90
β	99	95	97	99
γ	90	90	90	90
a	4.87	11.01	10.47	10.39
b	11.66	9.86	10.18	10.44
c	15.28	17.23	17.85	17.92
Geom.	triclinic†	triclinic	triclinic*	triclinic*
Volume	-	626	599	653
α	-	105	91	105
β	-	94	95	95
γ	-	86	87	88
a	-	4.67	4.70	4.76
b	-	9.99	9.73	10.18
c	-	13.95	13.15	14.00

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Again like the ortho TAPP-Cl_4H_4 cases, the monoclinic unit cells of ortho TAPP-Br_4H_4 show good patterns of increase size with alkyl group while maintain the cell shape. However in case of the triclinic crystals, C_2F_5-TAPP, similar as in TAPP-F_4H_4 and TAPP-Cl_4H_4 cases, shows decrease of unit cell vector b, which results the smaller unit cell size despite the relatively large alkyl groups size. The reason for the similarity between C_1F_3- and C_3F_7-TAPP unit cell might be the former is artificially built from the latter.

Tab. XV: Cell parameters for ortho TAPP. Angles are given in degrees, distances in Å.

ortho TAPP-I_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
Volume	967	2002	1971	2019
α	90	90	90	90
β	102	94	98	99
γ	90	90	90	90
a	5.23	11.46	10.77	10.54
b	11.87	9.93	10.17	10.39
c	15.93	17.67	18.17	18.67
Geom.	triclinic	triclinic	triclinic	triclinic*
Volume	487	682	701	728
α	103	107	107	107
β	106	95	94	94
γ	86	87	89	90
a	4.85	4.83	4.91	4.96
b	9.43	10.50	10.47	10.69
c	11.39	14.17	14.30	14.38

* Experimentally-obtained crystal system.

For the ortho TAPP-I_4H_4 crystals, unit cell size does not vary much for both monoclinic and increase weakly corresponding to the increased size of the alkyl groups in triclinic cases. Same as before mentioned other ortho TAPPs, monoclinic TAPP-I_4H_4 unit cell keep the cell shape intact after the optimization, which indicate the stability of the monoclinic structure.

As summary to the unit cell parameters of ortho TAPPs, the monoclinic structures show stability in their cell structures, despite the optimization process and different substitution groups the cell kept its shape. As for triclinic TAPPs, change of the alkyl group seems affect the very little on cell angles while slightly increase the cell vector length which also been observed in the monoclinic cases. The conclusion for the unit cell parameter analysis is that the increase size of the alkyl group will make the unit cell larger which is to be expected. Compare the cell size of the ortho monoclinic R-TAPP-X with same R group, it is clear to

see that with the halogen substitution increase in size the unit cell size would also increase while the cell angles do not change. Same trend can be observed in the triclinic R-TAPPs, the angle changes in the triclinic cases are mostly due to the use of different parent structures when triclinic crystals are artificially built.

Tab. XVI: Cell parameters for bay TAPP. Angles are given in degrees, distances in Å.

bay TAPP-F_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic†
Volume	1907	2358	2460	-
α	90	90	90	-
β	101	100	99	-
γ	90	90	90	-
a	11.62	12.63	12.39	-
b	12.71	13.05	13.20	-
c	13.17	14.51	15.22	-
Geom.	triclinic	triclinic	triclinic	triclinic
Volume	1376	1545	1556	1580
α	66	68	68	68
β	86	84	83	83
γ	77	78	78	78
a	9.59	9.78	9.84	9.83
b	11.38	12.32	12.26	12.49
c	14.14	14.20	14.29	14.15

* Experimentally-obtained crystal system.

† Cell optimization not converged.

c. bay compounds In the bay TAPP case, both monoclinic and triclinic crystal show that unit cell shape does not change with respect to the increased size of the alkyl substitutions, indicated by the cell angles which show no change despite the different alkyl groups in the TAPP. As the ortho TAPP case already shows, unit cell size only have noticeable change when change from H-TAPP to the C_1F_3-TAPPs in both the monoclinic and triclinic crystals. Among the R-TAPP with R from C_1F_3 to C_3F_7, unit cell size shows very little

change. Again the large unit cell difference between the monoclinic cell and triclinic cell is due to the different amount of atoms in each initial cells.

Tab. XVII: Cell parameters for bay TAPP. Angles are given in degrees, distances in Å.

bay TAPP-Cl_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic*	monoclinic
Volume	2098	2454	2569	3008
α	90	90	90	90
β	94	97	97	103
γ	90	90	90	90
a	12.35	12.79	12.64	14.93
b	11.95	12.90	13.45	15.26
c	14.26	14.99	15.23	13.57
Geom.	triclinic†	triclinic	triclinic	triclinic*
Volume	-	1598	1596	1624
α	-	69	68	69
β	-	83	84	84
γ	-	79	79	79
a	-	9.97	9.94	9.94
b	-	12.64	12.59	12.69
c	-	13.90	14.02	14.11

* Experimentally-obtained crystal system.

† Cell optimization not converged.

In case of the bay TAPP-Cl_4H_4, despite the unit cell size caused by the amount of atoms in the cell, both the monoclinic and triclinic crystals show no twist in shape after optimization with different alkyl groups. This further proof that change the alkyl group on the bare TAPP backbone structures would not affect the crystal structure of the compounds. On the side of the unit cell size, as to be expected with the increase of the alkyl groups the unit cell size also increase.

Tab. XVIII: Cell parameters for bay TAPP. Angles are given in degrees, distances in Å.

bay TAPP-Br_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic†
Volume	2163	2578	2684	-
α	90	90	90	-
β	92	97	97	-
γ	90	90	90	-
a	12.81	12.83	12.74	-
b	11.77	13.37	13.85	-
c	14.35	15.13	15.31	-
Geom.	triclinic	triclinic	triclinic	triclinic
Volume	1032	1634	1651	1675
α	50	69	69	68
β	83	83	85	85
γ	85	79	79	80
a	9.92	10.10	10.06	10.103
b	12.08	12.87	12.87	12.858
c	11.26	13.70	13.93	14.100

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Take a look at the optimized bay TAPP-Br_4H_4 crystal structures, it is clear that similar to the previous bay TAPPs the larger the alkyl substitution group, the larger the unit cell size. Similar to the ortho TAPPs, despite the change of the alkyl groups, both the monoclinic and triclinic TAPP compounds show no changes of unit cell shape, which is indicated by the cell angles.

Tab. XIX: Cell parameters for bay TAPP. Angles are given in degrees, distances in Å.

bay TAPP-I_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic†
Volume	2406	2814	2928	-
α	90	90	90	-
β	83	94	94	-
γ	90	90	90	-
a	12.70	13.17	13.40	-
b	12.09	13.61	14.32	-
c	15.79	15.73	15.30	-
Geom.	triclinic†	triclinic	triclinic	triclinic
Volume	-	1671	1799	1863
α	-	68	69	70
β	-	82	86	86
γ	-	81	81	81
a	-	10.33	10.35	10.48
b	-	13.44	13.290	13.32
c	-	13.21	14.190	14.42

* Experimentally-obtained crystal system.

† Cell optimization not converged.

In the bay TAPP-I_4H_4 cases, optimized structures show the different alkyl substitution have little to none affect on the shape of the unit cells, as suggested in the Table XIX. The increase of the unit cell size with respect to the increase of the alkyl groups size is as the same as we observed in the previous bay TAPP compounds.

d. orthorhombic bay compounds comparison

Tab. XX: Cell parameters for bay TAPP. Angles are given in degrees, distances in Å.

bay C_1F_3-TAPP	-F_4H_4	-Cl_4H_4	-Br_4H_4	-I_4H_4
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
Volume	2358	2454	2578	2814
α	90	90	90	90
β	100	97	97	94
γ	90	90	90	90
a	12.63	12.79	12.83	13.17
b	13.05	12.90	13.37	13.61
c	13.17	14.99	15.13	15.73
Geom.	triclinic	triclinic	triclinic	triclinic
Volume	1545	1598	1634	1671
α	68	69	69	68
β	84	83	83	82
γ	78	79	79	81
a	9.59	9.97	10.10	10.33
b	11.38	12.64	12.87	13.44
c	14.14	13.90	13.70	13.21
Geom.	orthorhomb.	orthorhomb.*	orthorhomb.	orthorhomb.
Volume	4176	4302	4504	4801
α	90	90	90	90
β	90	90	90	90
γ	90	90	90	90
a	14.82	15.19	15.63	16.22
b	12.56	12.57	12.45	12.25
c	22.43	22.53	23.16	24.17

* Experimentally-obtained crystal system.

In Table XX the bay TAPP compounds with orthorhombic crystal structures are listed in comparison with the monoclinic and triclinic structures. The obvious unit cell difference

between orthorhombic and other two structures are due to the very large amount of atoms in experimentally acquired cell information file. However if take a closer look it is clear that the experimental data of orthorhombic crystals are in very good agreement with the optimized results, all the cell angles are in 90°. Same as other bay TAPP compounds, with the increase of the alkyl group increase the unit cell size.

In summary of the bay TAPP compounds, they show the similar trends as ortho TAPPs, which the size of the alkyl group would affect the size of the unit cell. However in the case of the cell shape, changing the alkyl group always show no affect on the cell angle. By comparing the bay R-TAPP-X while keep R group unchanged, it is clear that both monoclinic and triclinic crystals show the increase of the unit cell size while the halogen substitution increases. As for the cell shape, change of the halogen atom would not affect this property for both the monoclinic and triclinic case.

e. all compounds

Tab. XXI: Cell parameters for all TAPP. Angles are given in degrees, distances in Å.

all TAPP-F_8	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
Volume	2576	2704	2838	2918
α	90	90	90	90
β	96	99	101	102
γ	90	90	90	90
a	13.44	14.06	14.47	14.77
b	15.20	15.10	15.21	15.14
c	12.69	12.88	13.14	13.33
Geom.	triclinic†	triclinic	triclinic†	triclinic
Volume	-	1619	-	1669
α	-	68	-	69
β	-	85	-	84
γ	-	79	-	80
a	-	9.96	-	10.07
b	-	12.52	-	12.63
c	-	14.30	-	14.35

* Experimentally-obtained crystal system.

† Cell optimization not converged.

In case of all TAPP, in which all the hydrogen atoms are replaced by the halogen atoms, similar pattern as in case of ortho and bay TAPP also shows which by increasing the size of the alkyl groups the unit cell size also increase. It is also clear that changing the alkyl group still not affect the cell shape because the cell angle does not change by replacing different alkyl groups.

Tab. XXII: Cell parameters for all TAPP. Angles are given in degrees, distances in Å.

all TAPP-Cl_8	H-†	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic*
Volume	-	3108	3162	3159
α	-	90	90	90
β	-	103	104	104
γ	-	90	90	90
a	-	15.00	15.09	15.17
b	-	15.75	15.90	15.78
c	-	13.52	13.59	13.60
Geom.	triclinic	triclinic	triclinic	triclinic
Volume	-	1570	1929	1855
α	-	67	71	70
β	-	82	87	86
γ	-	80	81	80
a	-	10.91	11.07	10.91
b	-	14.28	13.55	13.46
c	-	11.13	13.82	13.64

* Experimentally-obtained crystal system.

† Cell optimization not converged.

In all TAPP-Cl_8 crystal, only trcilinic show the same pattern as in all TAPP-F_8, with larger alkyl group increase the unit cell size. However in the monoclinic crystal, compare to the ortho and bay TAPP, cell angles β is no longer around 90°. All C_3F_7-TAPP-Cl_8 triclinic crystals shows slightly decrease on the cell size despite has the largest alkyl group. The cell angles however, show little change over the alkyl substitution in both the monoclinic and triclinic crystals.

Tab. XXIII: Cell parameters for all TAPP. Angles are given in degrees, distances in Å.

all TAPP-Br_8	H-†	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
Volume	-	3483	3462	3446
α	-	90	90	90
β	-	106	106	105
γ	-	90	90	90
a	-	15.65	15.82	15.75
b	-	16.59	16.46	16.50
c	-	13.96	13.83	13.73
Geom.	triclinic	triclinic	triclinic	triclinic
Volume	-	1705	2074	2011
α	-	70	71	71
β	-	85	88	87
γ	-	78	82	81
a	-	11.43	11.45	11.38
b	-	15.18	14.58	14.25
c	-	10.69	13.24	13.29

* Experimentally-obtained crystal system.

† Cell optimization not converged.

In case of the all TAPP-Br_8, in monoclinic case there is no sign the unit cell increase with the alkyl groups size, which might because the already large unit cell. However in triclinic case, from C_1F_3-TAPP to C_2F_5-TAPP there is a jump on unit cell size even the cell shape is kept unchanged.

Tab. XXIV: Cell parameters for all TAPP. Angles are given in degrees, distances in Å.

all TAPP-I_8	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
Volume	3072	4215	4174	4108
α	90	90	90	90
β	105	109	109	109
γ	90	90	90	90
a	17.45	17.42	17.59	17.42
b	11.65	17.67	17.70	17.72
c	15.65	14.46	14.22	14.08
Geom.	triclinic†	triclinic	triclinic†	triclinic
Volume	-	2025	-	2236
α	-	73	-	73
β	-	85	-	90
γ	-	79	-	81
a	-	12.01	-	12.07
b	-	15.99	-	15.91
c	-	11.26	-	12.37

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Similar as all TAPP-Cl_8 and all TAPP-Br_8, only the triclinic case shows the effect of the increased alkyl group over the unit cells. Alkyl groups show no affect on the unit cell parameters of the monoclinic all TAPP compounds both size-wise and shape-wise.

Tab. XXV: Cell information for all TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

all C_3F_7-TAPP	-Br_4Cl_4	-Cl_4Br_4
Geom.	monoclinic	monoclinic*
Volume	3493	3451
α	90	90
β	96	98
γ	90	90
a	13.61	13.51
b	15.88	15.99
c	16.27	16.12
Geom.	triclinic	triclinic
Volume	1957	1968
α	71	70
β	87	88
γ	80	81
a	11.11	11.21
b	14.10	13.90
c	13.46	13.58

* Experimentally-obtained crystal system.

There are also two extra all TAPP compounds which have been investigated, namely the C_3F_7-TAPP-Br_4Cl_4 and C_3F_7-TAPP-Cl_4Br_4. It is clearly shown that halogen substitution on the different position on TAPP backbone structure have no any meaningful influence on unit cell. In both cases, for both monoclinic and triclinic structure, the unit cells barely change either on size or on shape.

In summary, after eliminate the effect of parent structure on the artificially built ones, it is clear that despite the increased size of the alkyl group, unit cell only increase its size while kept its shape. The substitution of halogen atom also only affect the unit cell size by increase the length of the cell vectors while the cell angles are kept unchanged. It is also clear that interchange the substitution position of the halogen atoms will not affect the unit

cell in any meaningful way.

C. Inter-molecular shifts and intra-molecular angles

In this section the two geometric properties of the TAPP crystals will be discussed: inter-molecular shifts between two TAPP molecules in different layers, which are extracted from the crystal structures optimized from DFT calculations and the intra-molecular angles which indicates if and how the molecule twists. This section will be organized similar to the cell parameter discussion, TAPP compounds will be divided into four categories: TAPP-H_8, ortho TAPPs, bay TAPPs and all TAPPs. For each category, based on the halogen substitution, different TAPP-X compounds are listed. Within each compounds, all the inter-layer shift informations as well as intra-molecular angle informations are listed for all the crystal systems.

a. TAPP-H_8 compounds

Tab. XXVI: Cell information for R-TAPP-H. Angles are given in degrees, distances in Å, energies in kJ/mol.

TAPP-H_8	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic*	monoclinic	monoclinic	monoclinic
Volume	672	1660	1686	1713
Inter-layer dist.	3.02	7.00	6.98	6.71
$\Delta_\perp$	4.05	4.87	5.44	1.61
$\Delta_\parallel$	0.94	5.71	5.02	7.57
θ_{inter}	0	1	0	2
θ_{intra}	1	1	1	1
Geom.	triclinic†	triclinic	triclinic	triclinic*
Volume	-	433	502	595
Inter-layer dist.	-	2.91	2.97	3.02
$\Delta_\perp$	-	1.41	1.30	1.06
$\Delta_\parallel$	-	3.24	3.31	3.47
θ_{inter}	-	2	3	7
θ_{intra}	-	1	1	5

* Experimentally-obtained crystal system.

† Cell optimization not converged.

For the TAPP-H_8 compounds we notice that optimized structures, even without symmetry constraint, still show the flat molecular surface. All the intra-molecular angles are well within 10°, and similar as the ortho TAPP case, no bay position substitution means no need for θ_{bay}. It is also worth noticing that different TAPP compounds might be artificially built from different parent structures by tailoring the alkyl groups, and the optimization process would not bring the crystal geometry too far from their origin. So the most important information in the inter-molecular shift results are changes of shifts with respect to the different alkyl groups. There is no clear pattern for how the different alkyl groups would affect the packing arrangement of the monoclinic TAPP-H_8 molecules. However there is an indication that with increased size of the alkyl group, the molecules in different layers tend to avoid each other and without the large halogen atom, TAPP-H_8 molecules seem to

have more freedom on which direction they would take to distance themselves with each other. This freedom might be the reason why there seems no unified way on which properties, *i.e.* inter-layer distance, $\Delta_\perp$ or $\Delta_\parallel$, would change when smaller alkyl group replaced by the larger ones. Similarly to the monoclinic case, the triclinic TAPP-H_8 molecules show a bit less chaotic changes on the inter-layer shifts. With inter-layer distance and $\Delta_\parallel$ became larger when large alkyl groups substitute, $\Delta_\perp$ can vary a bit without cause the neighbour molecules repulse each other.

Tab. XXVII: Cell information for ortho TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

ortho TAPP-F_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic*
Volume	713	1713	1714	1733
Inter-layer dist.	2.89	6.92	6.78	6.92
$\Delta_\perp$	3.57	5.51	1.77	5.63
$\Delta_\parallel$	0.52	5.18	7.42	5.08
θ_{inter}	7	2	0	1
θ_{intra}	2	2	1	0
Geom.	triclinic	triclinic	triclinic	triclinic
Volume	598	524	534	594
Inter-layer dist.	2.83	2.88	2.95	2.97
$\Delta_\perp$	1.28	1.17	1.12	1.06
$\Delta_\parallel$	3.26	3.28	3.32	3.39
θ_{inter}	1	2	6	4
θ_{intra}	1	1	4	0

* Experimentally-obtained crystal system.

b. TAPP compounds with ortho substitution In the case of the ortho TAPP-F_4H_4, the intra-molecular angles show good agreement with our expectations and show the ortho TAPP compounds have flat molecular surface. Similar to the TAPP-H_8 case discussed above, monoclinic ortho TAPP-F_4H_4 show no clear pattern on how each molecules are moving away from each other despite the halogen substitution. For example, in the C_2F_5-TAPP-F_4H_4

case, molecules move away from each alone the same direction as the molecular axis, which indicate by the large $\Delta_{\parallel}$ value. This large shift on one direction leave the molecules shift only relatively small distance on the other two directions, namely the direction perpendicular to the molecular axis and the molecular surface, which indicated by the relatively smaller inter-layer distance and $\Delta_{\perp}$ value. On the other hand in the triclinic case, increase the size of the alkyl group affects the inter-layer distance as well as the $\Delta_{\parallel}$ to increase, which leaves the $\Delta_{\perp}$ relatively unaffected.

Tab. XXVIII: Cell information for ortho TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

ortho TAPP-Cl_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
Volume	813	1805	1849	1889
Inter-layer dist.	2.97	6.35	7.15	6.86
$\Delta_{\perp}$	3.64	6.51	5.59	6.13
$\Delta_{\parallel}$	0.40	5.83	4.92	4.66
θ_{inter}	5	1	3	1
θ_{intra}	6	0	1	0
Geom.	triclinic	triclinic	triclinic*	triclinic*
Volume	398	589	575	622
Inter-layer dist.	2.91	2.98	3.01	3.04
$\Delta_{\perp}$	1.54	1.06	1.17	1.14
$\Delta_{\parallel}$	3.20	3.34	3.32	3.36
θ_{inter}	2	12	12	13
θ_{intra}	1	8	6	7

* Experimentally-obtained crystal system.

In the ortho TAPP-Cl_4H_4 cases, the intra-molecular angles in the monoclinic crystals still show good agreement with the chemical intuition of the flat molecular surface. However in the triclinic cases, the inter-block angle θ_{inter} of R-TAPP-Cl_4H_4 are slightly larger than 10°, which results the molecular surface with minor twist. One of the possible reasons for this disagreement is the steep potential energy surface (PES), which the small change on

the molecule coordinates would results large change on the energy. This insensitivity of the energy with respect to the molecular coordinates would suggest that even with inter-block angle larger than 10°, TAPP molecule still lay on the bottom of the PES. On the aspect of inter-molecular shifts, consider the size of the Cl atom is much larger than F atom (about 100 times), ortho TAPP-Cl_4H_4 molecule have much less freedom on moving away from each other compare to TAPP-H_8 or even TAPP-F_4H_4. This restriction also shows on the inter molecular shift results, which compare to the TAPP-F_4H_4 case all three shifts are less diverse from C_1F_3-TAPP to C_3F_7-TAPP. The triclinic case however, all three shifts increase with increase size of the alkyl groups.

Tab. XXIX: Cell information for ortho TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

ortho TAPP-Br_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
Volume	857	1863	1890	1918
Inter-layer dist.	3.08	6.49	7.04	6.75
$\Delta_\perp$	3.73	6.70	5.99	6.45
$\Delta_\parallel$	0.60	5.85	4.91	4.54
θ_{inter}	9	1	2	1
θ_{intra}	0	0	1	1
Geom.	triclinic†	triclinic	triclinic*	triclinic*
Volume	-	626	629	653
Inter-layer dist.	-	3.03	3.05	3.10
$\Delta_\perp$	-	1.03	1.21	1.08
$\Delta_\parallel$	-	3.40	3.38	3.44
θ_{inter}	-	15	12	16
θ_{intra}	-	11	7	10

* Experimentally-obtained crystal system.

† Cell optimization not converged.

For ortho TAPP-Br_4H_4 molecule, same as TAPP-Cl_4H_4 case, molecules extracted from triclinic crystals show slightly twisted molecular surface while in monoclinic case the molecu-

lar surface is well flat. Compare to the ortho TAPP-F_4H_4 molecule, the twist of the molecular surface is more pronounced with heavier halogen atoms. By taking the closer look at the inter-layer shift, ortho TAPP-Br_4H_4 shows similarity with TAPP-Cl_4H_4 and TAPP-F_4H_4, which indicates molecules would move away from each other in general but not at the same time for all direction. Similarly to the previous ortho TAPP case, with increase size of the alkyl group, all three shifts from both C_1F_3-TAPP and C_2F_5-TAPP are increase, with only C_3F_7-TAPP as exception, which $\Delta_\perp$ shift decreases slightly.

Tab. XXX: Cell information for ortho TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

ortho TAPP-I_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
Volume	967	2002	1971	2019
Inter-layer dist.	3.37	6.61	7.01	6.79
$\Delta_\perp$	3.92	7.01	6.40	6.53
$\Delta_\parallel$	0.81	6.21	5.08	4.73
θ_{inter}	12	1	4	1
θ_{intra}	7	1	4	2
Geom.	triclinic	triclinic	triclinic	triclinic*
Volume	487	682	701	728
Inter-layer dist.	3.13	3.15	3.28	3.28
$\Delta_\perp$	1.93	1.03	1.03	1.03
$\Delta_\parallel$	3.16	3.51	3.57	3.57
θ_{inter}	0	15	17	18
θ_{intra}	2	11	12	12

* Experimentally-obtained crystal system.

In case of the ortho TAPP-I_4H_4 molecule, except the H-TAPP-I_4H_4, all the intra-molecular angles of monoclinic R-TAPP-I_4H_4 show great agreement with the assumption of flat ortho molecule surface. The triclinic molecules on the other hand agree with the previous TAPP-X_4H_4, such as TAPP-Br_4H_4 with minor twist on the molecular surface. When look into the inter-layer shifts, monoclinic molecules agrees with the previous analysis

which suggests that if TAPP molecule moves more on one direction, it might move less on other direction, thus the inconsistency of the different inter-layer shifts trends. Triclinic molecules however, show good agreement on all three directions.

In summary, for the ortho TAPP molecules, both monoclinic and triclinic crystals have flat molecule surface with the latter show little twist, which can be explained by the steep PES. It also suggests that by replace the H atom with heavier halogen atom, molecular surface would not be affected. As for the inter-layer shifts, both monoclinic and triclinic molecules agree that with larger alkyl group replace the H atom and smaller ones, molecules tend to move away from each other. The direction of these movement is unpredictable, yet the general principle is always applied, which suggest if molecules move more on one direction, distance change on other two directions would be less noticeable. Compare the effect of different halogen substitution with respect to the inter-layer shift, it can be concluded that there is no noticeable influence.

c. TAPP compounds with bay substitution

Tab. XXXI: Cell information for bay TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

bay TAPP-F_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic†
Volume	1907	2358	2460	-
Inter-layer dist.	3.76	3.59	3.62	-
$\Delta_\perp$	0.55	0.80	0.74	-
$\Delta_\parallel$	4.55	5.06	4.85	-
θ_{inter}	4	10	11	-
θ_{intra}	8	4	4	-
θ_{bay}	35	42	45	-
Geom.	triclinic	triclinic	triclinic	triclinic
Volume	1376	1545	1556	1580
Inter-layer dist.	3.62	4.08	3.68	4.15
$\Delta_\perp$	1.06	1.28	0.50	1.30
$\Delta_\parallel$	3.74	4.15	5.37	4.17
θ_{inter}	8	10	11	9
θ_{intra}	4	5	6	5
θ_{bay}	38	39	42	42

* Experimentally-obtained crystal system.

† Cell optimization not converged.

The difference between ortho and bay TAPPs, except the substitution position, is the latter have non-flat molecular surface, which shows clearly in the results here. Both the molecules from monoclinic and triclinic crystals show inter-block torsion angle θ_{inter} around 10°. The intra-block angle between halogen atom and its neighbouring H atom θ_{intra} gives the indication of twist within one building block of the TAPP molecule. Similar to the ortho TAPP molecules, θ_{intra} for TAPP-F_4H_4 compounds are relatively small due to the size of the F atom. With the halogen atom substitute on the bay site, it is also worth to check how the two halogen atom would affect each other, thus for the bay TAPP here and all TAPP later, θ_{bay} is introduced. As shown in the table, θ_{bay} even for the smallest halogen atom, is

still significantly larger compare to the θ_{intra}. On the side of the inter-layer shifts, even there is clear pattern on which direction TAPP molecules would move when larger alkyl groups take place, the parallel shift seems more pronounced for both the monoclinic and triclinic molecules.

Tab. XXXII: Cell information for bay TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

bay TAPP-Cl_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic*	monoclinic
Volume	2098	2454	2569	3008
Inter-layer dist.	4.20	4.17	3.70	4.28
$\Delta_\perp$	0.51	0.69	0.73	1.70
$\Delta_\parallel$	4.73	5.05	4.95	5.05
θ_{inter}	7	9	7	12
θ_{intra}	6	8	7	7
θ_{bay}	51	52	56	55
Geom.	triclinic†	triclinic	triclinic	triclinic*
Volume	-	1598	1596	1624
Inter-layer dist.	-	3.84	3.73	3.79
$\Delta_\perp$	-	0.41	0.45	0.48
$\Delta_\parallel$	-	5.47	5.47	5.54
θ_{inter}	-	9	10	11
θ_{intra}	-	8	7	7
θ_{bay}	-	55	55	55

* Experimentally-obtained crystal system.

† Cell optimization not converged.

For the bay TAPP-Cl_4H_4, similar as the bay TAPP-F_4H_4, that molecule surface only show minor twist around 10°. Even size of the Cl atom is noticeably larger compare to the F atom, the intra-block angle θ_{intra} changes little to reflex this effect for both the monoclinic and triclinic molecules. However for the angle between two halogen atoms in the bay position, compare to the TAPP-F_4H_4 θ_{bay} in TAPP-Cl_4H_4 is much larger due to the large size of

the Cl atom would results stronger repulsive interactions between two atoms. Same as in the previous cases, for TAPP-Cl_4H_4 from both the monoclinic and triclinic structures the inter-layer shifts are also more noticeable on the direction which parallel to the molecular axis.

Tab. XXXIII: Cell information for bay TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

bay TAPP-Br_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic†
Volume	2163	2578	2684	-
Inter-layer dist.	4.29	4.24	4.25	-
$\Delta_\perp$	0.56	0.69	0.82	-
$\Delta_\parallel$	4.96	5.02	4.95	-
θ_{inter}	7	7	8	-
θ_{intra}	10	9	9	-
θ_{bay}	56	55	56	-
Geom.	triclinic	triclinic	triclinic	triclinic
Volume	1032	1634	1651	1675
Inter-layer dist.	3.63	4.38	3.82	3.81
$\Delta_\perp$	0.30	1.32	0.43	0.42
$\Delta_\parallel$	5.48	4.27	5.59	5.59
θ_{inter}	15	11	11	10
θ_{intra}	7	7	8	8
θ_{bay}	56	58	58	58

* Experimentally-obtained crystal system.

† Cell optimization not converged.

The bay TAPP-Br_4H_4 molecules show similar pattern on the inter-block torsion angles as in previous bay TAPPs, which further support the assumption that size of the alkyl groups have little influences on the molecule surface. Compare the intra-block angle between halogen and H atoms shows further increase the size of the halogen atom would not affect this property. As for the θ_{bay}, due to the small difference between Br and Cl atoms, there is

little to none change of this property compare to the TAPP-Cl_4H_4 molecules. The inter-layer shifts also agree with the previous bay TAPP cases, in which the molecules from different layers mainly move alongside the direction parallel to the molecular axis to avoid each other.

Tab. XXXIV: Cell information for bay TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

bay TAPP-I_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic†
Volume	2406	2814	2928	-
Inter-layer dist.	4.50	4.49	3.96	-
$\Delta_\perp$	1.73	0.76	0.77	-
$\Delta_\parallel$	4.19	5.05	5.21	-
θ_{inter}	14	9	2	-
θ_{intra}	11	10	13	-
θ_{bay}	61	59	63	-
Geom.	triclinic†	triclinic	triclinic	triclinic
Volume	-	1671	1799	1863
Inter-layer dist.	-	3.90	4.54	4.14
$\Delta_\perp$	-	0.36	0.43	0.49
$\Delta_\parallel$	-	5.89	5.73	5.66
θ_{inter}	-	11	11	10
θ_{intra}	-	10	9	10
θ_{bay}	-	63	60	62

* Experimentally-obtained crystal system.

† Cell optimization not converged.

The bay TAPP-I_4H_4, similar to the other bay TAPP, shows molecular surface only with around 10° of twist. The intra-block angle θ_{intra} also shows agreement with bay TAPP-Cl_4H_4 and TAPP-Br_4H_4 molecules. The increased size of the I atom cause the angle between two halogen atoms θ_{bay} increase slightly (~7Å) compare to the previous cases. However despite the large size of the I atom, inter-layer shifts show the similar pattern as in TAPP-Cl_4H_4 and TAPP-Br_4H_4 without much of the change.

Tab. XXXV: Cell information for bay TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

bay C_1F_3-TAPP	-F_4H_4	-Cl_4H_4	-Br_4H_4	-I_4H_4
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
Volume	2358.468	2454.403	2578.271	2813.749
Inter-layer dist.	3.59	4.17	4.24	4.49
$\Delta_\perp$	0.80	0.69	0.69	0.76
$\Delta_\parallel$	5.06	5.05	5.02	5.05
θ_{inter}	10	9	7	9
θ_{intra}	4	8	9	10
θ_{bay}	42	52	55	59
Geom.	triclinic	triclinic	triclinic	triclinic
Volume	1545.307	1598.052	1634.228	1671.317
Inter-layer dist.	4.08	3.84	4.38	3.90
$\Delta_\perp$	1.28	0.41	1.32	0.36
$\Delta_\parallel$	4.15	5.47	4.27	5.89
θ_{inter}	10	9	11	11
θ_{intra}	5	8	7	10
θ_{bay}	39	55	58	63
Geom.	orthorhomb.	orthorhomb.*	orthorhomb.	orthorhomb.
Volume	4176	4302	4504	4801
Inter-layer dist.	3.50	3.96	4.03	3.78
$\Delta_\perp$	0.46	0.47	0.47	0.43
$\Delta_\parallel$	5.06	5.01	4.88	4.70
θ_{inter}	10	7	8	1
θ_{intra}	4	7	7	8
θ_{bay}	40	52	53	55

* Experimentally-obtained crystal system.

d. orthorhombic TAPP compounds with bay substitution

In this table the comparison between experimentally acquired orthorhombic bay TAPPs and the artificially built monoclinic and triclinic bay TAPPs are made. From this comparison it is clear that despite the size difference on unit cells, bay TAPP molecules extracted from orthorhombic crystals share the similar geometric properties as in other two kinds of TAPPs.

In summary of bay TAPPs, it is clear that both halogen substitution and alkyl group have little to none influence on the molecular surface, despite the different substitutional groups, molecular surface of the bay TAPPs always stay around 10°. Size effect of different halogen atoms only happens from changing F atom to other larger halogen atoms, Br, Cl and I atoms show very limited effect on both the θ_{intra} and θ_{bay} for all the crystal structures. As for the inter-layer shifts, in the monoclinic molecules it is clear that with the increase of the alkyl group size, inter-layer distance also increase. However for the parallel and perpendicular shift this tendency is much less pronounced with respect to the increase of the alkyl substitution group.

e. TAPP compounds with all substitution

Tab. XXXVI: Cell information for all TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

all TAPP-F_8	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
Volume	2576	2704	2838	2918
Inter-layer dist.	4.14	3.89	4.56	4.68
$\Delta_\perp$	1.64	2.16	2.20	2.14
$\Delta_\parallel$	5.05	5.05	4.90	4.96
θ_{inter}	17	14	18	20
θ_{intra}	2	2	1	1
θ_{bay}	39	38	40	40
Geom.	triclinic†	triclinic	triclinic†	triclinic
Volume	-	1619	-	1669
Inter-layer dist.	-	4.20	-	3.79
$\Delta_\perp$	-	1.38	-	0.53
$\Delta_\parallel$	-	4.24	-	5.44
θ_{inter}	-	16	-	15
θ_{intra}	-	2	-	0
θ_{bay}	-	42	-	41

* Experimentally-obtained crystal system.

† Cell optimization not converged.

In the case of the all TAPP, all the H atoms on the TAPP backbone are replaced by the halogen atoms, the inter-block angle which indicates the twist of the TAPP molecular surface reflex such complete substitution. As listed in Table XXXVI, even for the TAPP-F_8, both monoclinic and triclinic molecules have shown θ_{inter} larger than 10°, and in some cases even reach 20°. Due to the small size of the F atom, intra-block angle θ_{intra} for all TAPP-F_8 are well around 0°. Angle between the two halogen atoms compare to the corresponding bay TAPP molecules shows almost no change. For triclinic TAPPs the inter-layer shifts follows the same pattern as in the previous TAPPs, which means TAPP molecules move away from each other following the direction parallel to the molecular axis. However by look into the

monoclinic case, it shows the large shift also alongside the perpendicular direction about 2Å, which is not observed in both ortho and bay TAPPs.

Tab. XXXVII: Cell information for all TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

all TAPP-Cl_8	H-†	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic*
Volume	-	3108	3162	3159
Inter-layer dist.	-	4.26	4.75	4.86
$\Delta_\perp$	-	2.17	1.83	2.19
$\Delta_\parallel$	-	5.08	5.08	4.96
θ_{inter}	-	5	8	9
θ_{intra}	-	9	11	4
θ_{bay}	-	55	54	53
Geom.	triclinic	triclinic	triclinic	triclinic
Volume	-	1570	1929	1855
Inter-layer dist.	-	4.42	4.62	4.58
$\Delta_\perp$	-	0.19	1.94	1.71
$\Delta_\parallel$	-	6.62	4.26	4.31
θ_{inter}	-	22	17	19
θ_{intra}	-	5	10	4
θ_{bay}	-	56	57	56

* Experimentally-obtained crystal system.

† Cell optimization not converged.

In the all TAPP-Cl_8 case, the inter-block angle in the triclinic molecules are much larger than its monoclinic counterparts, with the former being around 20° while the latter around 10°. However for the intra-block and bay angles there are no such pattern. Compare to the TAPP-F_8, inter-block torsion angle has reduced in the monoclinic molecules while the other two have increased. In the triclinic molecules however, only the bay angle has been increased. The inter-layer shifts in the monoclinic case shows very limited changes from all TAPP-F_8 to all TAPP-Cl_8 with the exception of C_3F_7-TAPP. In case of the triclinic

molecules, both parallel and perpendicular shifts are varied from TAPP-F_8 molecules. This might be caused by the large size difference between F and Cl atoms.

Tab. XXXVIII: Cell information for all TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

all TAPP-Br_8	H-†	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
Volume	-	3483	3462	3446
Inter-layer dist.	-	5.02	4.54	4.49
$\Delta_\perp$	-	1.90	2.29	2.35
$\Delta_\parallel$	-	5.09	5.04	5.02
θ_{inter}	-	6	5	5
θ_{intra}	-	10	9	16
θ_{bay}	-	61	60	56
Geom.	triclinic	triclinic	triclinic	triclinic
Volume	-	1705	2074	2011
Inter-layer dist.	-	4.84	4.26	4.81
$\Delta_\perp$	-	2.16	0.60	1.97
$\Delta_\parallel$	-	4.93	6.52	4.61
θ_{inter}	-	18	19	21
θ_{intra}	-	7	2	6
θ_{bay}	-	62	61	62

* Experimentally-obtained crystal system.

† Cell optimization not converged.

For the all TAPP-Br_8 molecules, compare to the all TAPP-Cl_8, the intra-molecular angles have almost no change with inter-block angle around 10° to 20°, intra-block angle around 10° and bay angle around 60° for both monoclinic and triclinic structures. As for the inter-layer shifts, similar to the previous TAPP compounds including ortho and bay TAPPs, there seems no clear tendency on which direction molecule would move when large alkyl groups being added. Yet the same principle still applies that parallel and perpendicular shift are compensating each other, whenever molecule moves more on one direction, shift on the other

direction would be small.

Tab. XXXIX: Cell information for all TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

all TAPP-I_8	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
Volume	3072	4215	4174	4108
Inter-layer dist.	5.60	5.56	5.12	5.45
$\Delta_\perp$	2.80	2.72	2.12	2.10
$\Delta_\parallel$	5.70	4.83	4.68	4.69
θ_{inter}	8	6	3	5
θ_{intra}	12	6	13	25
θ_{bay}	69	63	67	61
Geom.	triclinic†	triclinic	triclinic†	triclinic
Volume	-	2025	-	2236
Inter-layer dist.	-	5.35	-	4.96
$\Delta_\perp$	-	2.24	-	2.38
$\Delta_\parallel$	-	5.18	-	5.37
θ_{inter}	-	15	-	18
θ_{intra}	-	10	-	9
θ_{bay}	-	68	-	67

* Experimentally-obtained crystal system.

† Cell optimization not converged.

In case of TAPP-I_8 molecule, intra-molecular angles of both structures show little differences compare to the TAPP-Br_8 or TAPP-Cl_8. It is also clear that different alkyl groups have no affect on how all TAPP molecules twist. As for the inter-layer shifts, they also show no clear pattern with respect to the change of alkyl groups. Compare to the previous all TAPP compounds such as TAPP-Br_8 or TAPP-Cl_8, in case of the inter-layer shifts TAPP-I_8 does not show any significant change on all three inter-layer shifts. The change of the alkyl groups also little to none effect on the inter-layer shifts.

Tab. XL: Cell information for all TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

all C_3F_7-TAPP	-Br_4Cl_4	-Cl_4Br_4
Geom.	monoclinic	monoclinic*
Volume	3493	3451
Inter-layer dist.	4.57	4.46
$\Delta_\perp$	2.69	2.72
$\Delta_\parallel$	4.88	4.93
θ_{inter}	3	5
θ_{intra}	15	11
θ_{bay}	58	55
Geom.	triclinic	triclinic
Volume	1957	1968
Inter-layer dist.	4.63	4.25
$\Delta_\perp$	1.83	0.65
$\Delta_\parallel$	4.60	6.12
θ_{inter}	18	22
θ_{intra}	6	1
θ_{bay}	61	58

* Experimentally-obtained crystal system.

After inspect every all TAPP-X_8 compounds, similar to the previous analysis TAPP-X_4Y_4 compounds are also being investigated here. In the triclinic cases, compare to the TAPP-Cl_8 molecule it is clear to see that mixing two halogen substitution have almost no effect on the intra-molecular angles while in the monoclinic case the intra-block torsion angle shows slightly more obvious increase compare to the all TAPP-Cl_8. When compare to the all TAPP-Br_8 molecule, there is however no noticeable changes among all the intra-molecular angles. Take a look at the inter-layer shifts, when compare to the all TAPP-Cl_8 molecule, monoclinic molecule shows tendency of move more on the direction perpendicular to the molecular axis and surface while no change on the direction parallel to the molecular axis. Compare to the all TAPP-Br_8 however, only shows the molecule move slightly more on the

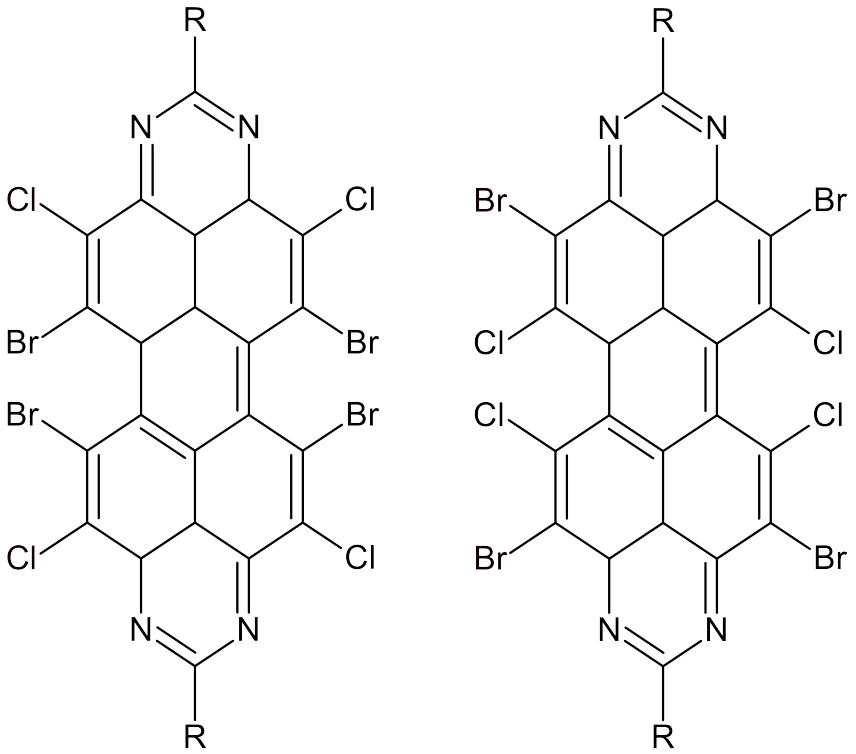


Fig. 11: R-TAPP-Br_4Cl_4 molecule with R as C_3F_7- (left) and R-TAPP-Cl_4Br_4 molecule with R as C_3F_7- (right).

direction perpendicular to the molecular axis while the rest two shifts show no change. As for the triclinic case, compare to the all TAPP-Cl_8 molecule, it seems that molecular shifts of TAPP-Cl_8 lies between TAPP-Br_4Cl_4 and TAPP-Cl_4Br_4. With the comparison to the TAPP-Br_8 case, it is clear that mixed TAPPs have smaller values on all three inter-layer shifts. It is safe to assume that the heavier halogen atom will have more influence on the packing arrangement of the all TAPP crystals.

In summary in the all TAPP compounds alkyl groups show very limited influence on the intra-molecular angles, while halogen substitutions on the other hand show great impact on these properties from TAPP-F_8 to TAPP-X_8 (X=Br, Cl, I and mix Cl with Br) for both monoclinic and triclinic structures. Inter-layer shifts on the monoclinic and triclinic structures however, show lack of the clear pattern with respect to the change of halogen and alkyl substitution. This might be explained as molecules on the neighbouring layers only sought to avoid each other, without constraint they can move freely on all three directions, thus results the lack of clear pattern. Take C_2F_5-TAPP-Br_8 as example, in the monoclinic case, with increased alkyl group size from C_1F_3 to C_2F_5, the inter-layer distance reduced,

the cause of this anti-intuitive results comes from the perpendicular shift, which increases from 1.90Åto 2.29Å. Lastly, all TAPP compounds with two mixed halogen substitution have been discussed and both the inter-layer shifts and intra-molecular angles show the geometric properties of such molecules are largely decided by the heavier atoms.

D. Inter-molecular interaction analysis: short contact

In this section the short contact information of the TAPP compounds will be discussed. The discussion will start with TAPP-H_8 compounds, followed by the ortho TAPPs, bay TAPPs and end with all TAPP compounds, which are shown in Fig 2. In each part all the TAPP derivatives will be discussed in both monoclinic and triclinic structures. For each TAPP compound, short contact is divided into three catalogues: same-stack means the short contact information is about the pair of molecules lay on top of each other; head-to-head configuration means the molecules in the dimer are on the same layer but shifts alongside the molecular axis; side-by-side configuration is the same as the head-to-head that molecules are located on the same layer but instead they shift perpendicular to the molecular axis. In each configuration, number of atom pairs in short contact are listed as well as the scaling factor and the size of the contact gap. As discussed in the short contact introduction part, scaling factor of 1.05 is used for all the TAPP compounds listed here.

a. TAPP-H_8 compounds

Tab. XLI: Short contact information for R-TAPP-H.

TAPP-H_8	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic*	monoclinic	monoclinic	monoclinic
same-stack	1.05/40/0.02	1.05/0/NE	1.05/0/NE	1.05/0/NE
head-to-head	1.05/3/0.05	1.05/4/0.15	1.05/4/0.13	1.05/6/0.04
side-by-side	1.05/4/0.22	1.05/0/NE	1.05/0/NE	1.05/0/NE
Geom.	triclinic†	triclinic	triclinic	triclinic*
same-stack	-	1.05/103/0.03	1.05/111/0.02	1.05/109/0.00
head-to-head	-	1.05/0/NE	1.05/0/NE	1.05/0/NE
side-by-side	-	1.05/0/NE	1.05/0/NE	1.05/10/0.27

* Experimentally-obtained crystal system.

† Cell optimization not converged.

In the case of the TAPP-H_8 compounds, it is clear that for the dimers of the same-stack configuration, both in the monoclinic and triclinic structures show the largest number of atom pairs in short contact. However for the other two configurations the number of short contact pairs are significantly reduced, especially for side-by-side configuration. When take a close look at the contact gap, it is also suggest that atoms from each molecules in the same-stack configuration are more close to each other than these from the other configuration, thus the smaller contact gap of same-stack dimer compare to other two dimers. This rather dramatic change on the number of atom pairs in short contact is to be expected, since in the head-to-head and side-by-side configurations only half of the each molecule in the dimer are close to each other, so the number of potentially able to be in short contact is already halved compare to the same-stack case. Another possible explanation for this large drop in number is that as discussed previous sections, TAPP molecules tend to move within the layer to avoid each other with increased size of the alkyl group. This tendency of moving away each other in either parallel or perpendicular direction would further reduce the number of the possible atoms pairs in short contact.

b. TAPP compounds with ortho substitution Table LX to LXIII show the short contact information of all the ortho TAPP compounds based on their halogen substitution. For the monoclinic case, it is shows two different patterns. With ortho H-TAPP compounds show

large amount of atom pairs in short contact, short contact pairs in R-TAPP compounds immediately reduced to 0 even for the smallest alkyl group that is C_1F_3. The inconsistency of C_3F_7-TAPP-F_4H_4 head-to-head configuration is due to the closeness between neighbouring stacks and the relatively large size of the C_3F_7 groups. The inconsistency C_2F_5-TAPP-F_4H_4 side-by-side configuration can be explained by the fact that the extremely small perpendicular shift in this compounds, which forced the atoms stay more closer than other structures. For the rest of the ortho TAPP compounds with monoclinic structures, only change from H-TAPP to R-TAPP will cause short contact in same-stack change while the head-to-head and side-by-side configurations mostly unaffected. The triclinic structures show similar pattern as in monoclinic TAPPs, with same-stack configuration has largest amount of short contact pairs and the rest two configurations show very little pairs in short contact. It is also clear in majority cases, changing the alkyl group as well as halogen substitutes would not affect the number of atom pairs in short contact. Even thought the atom size increases from fluorine to iodine, it is still not enough to effectively change the inter-molecular distance, thus having almost no influence on the number of short contact pairs. In case of the stabilities, the head-to-head short contacts are added to the side-by-side contacts and compared among the crystal systems. The crystal structure with the larger amount of short contacts is considered to be more favourable.

Tab. XLII: Short contact information for bay TAPP.

bay C_1F_3-TAPP	-F_4H_4	-Cl_4H_4	-Br_4H_4	-I_4H_4
Geom.	orthorhomb.	orthorhomb.*	orthorhomb.	orthorhomb.
same-stack	1.05/11/0.01	1.05/20/0.02	1.05/15/0.01	1.05/17/0.03
head-to-head	1.05/4/0.06	1.05/4/0.04	1.05/1/0.30	1.05/3/0.07
side-by-side	1.05/0/NE	1.05/0/NE	1.05/0/NE	1.05/5/0.12

* Experimentally-obtained crystal system.

c. TAPP compounds with bay substitution From Table LXIV to XLII the short contact information of bay TAPP compounds are listed. Compare the short contact number and gap values between monoclinic and triclinic structures, it is clear that the change of the structure type would no affect the short contact in any meaningful way. Similar to the ortho TAPP compounds, it is clear that for both the monoclinic and triclinic bay TAPP

compounds show little change on the number of atom pairs in short contact with respect to the change of the alkyl groups. The change in the short contact number however, is slightly more pronounced when change the H-TAPP to the R-TAPP, especially so if the halogen substitution is larger than F atom. Unlike in the ortho TAPP cases however, the number of atom pairs in short contact in bay TAPP compounds are very limited even in the same-stack configurations and there is no jump in number of short contact from H-TAPP to R-TAPP. The similarity between monoclinic and triclinic bay TAPP compounds is the result of their similarity in packing geometries. It is also shown that by changing halogen atom on the TAPP compounds, short contact number as well as gap change very little, which is in agreement with the previous geometric analysis that the packing geometries of bay TAPP rarely change with respect to the change of halogen substitutions. Further look into orthorhombic bay TAPP compounds shows even there is a little increase on the number of short contact atom pairs in the same-stack configuration, other two configurations also agree with the monoclinic and triclinic cases very well.

d. TAPP compounds with all substitution Table LXVIII to LXXII listed all the short contact information for the all TAPP compounds. Again similar to the ortho and bay TAPP compounds, all TAPPs show very little to no change in both the atom pairs in short contact as well as the short contact gap with respect to the change of the alkyl group. Even change from H-TAPP to R-TAPP there is no noticeable change on both the short contact number and the gap. Compare two different crystal structures within same all TAPP-X_8 also shows no change in both short contact properties. This can be explained similar as in ortho and bay TAPP case, which the similarity of short contact properties are tied the the similarity of the packing geometries between different TAPP compounds. Further comparison among different TAPP-X_8 compounds supports this argument, for they also show the non-change of the short contact properties with respect to the changing of halogen substitutes.

1. Short contact gap: case study

Tab. XLIII: Number of short contact pairs and gap Δ, in parentheses, for monoclinic and triclinic *ortho*-TAPP-F_4.

ortho-TAPP-H_4F_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	m	m	m	m*
head-to-head	2 (0.165)	5 (0.178)	1 (0.200)	6 (0.094)
side-by-side	3 (0.555)	2 (0.969)	2 (0.979)	4 (0.748)
Geom.	t	t	t	t
head-to-head	0 (–)	0 (–)	0 (–)	0 (–)
side-by-side	0 (–)	0 (–)	0 (–)	0 (–)
Pred.[a]	m	m	m	m

* Experimentally obtained crystal system.

[a] Theoretically resulting stable crystal system based.

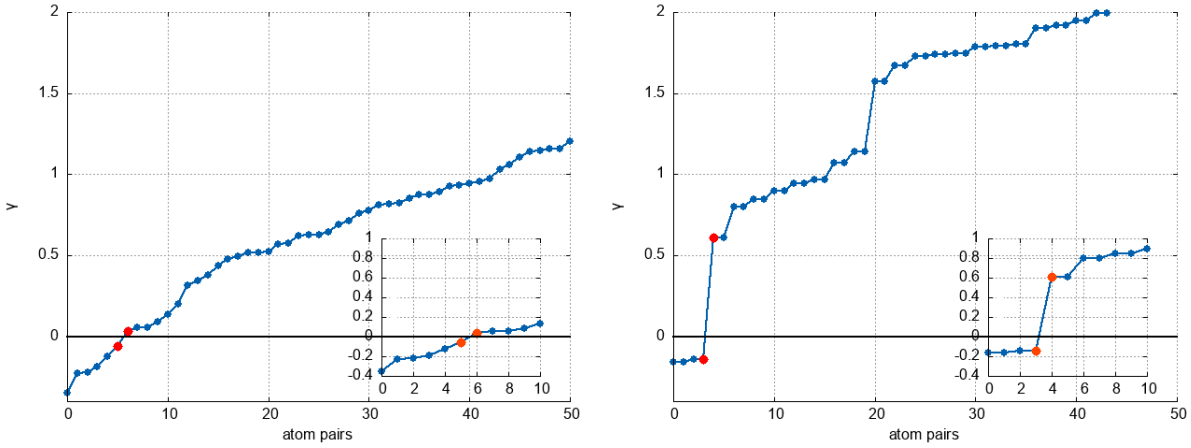


Fig. 12: Distribution of short-contact distance γ in Å for the closest 50 intermolecular atom pairs in case of monoclinic *ortho*-TAPP-H_4F_4. The red dots indicate the longest short distance, *i.e.* $\max \gamma_{ij}$, and the shortest long distance, *i.e.* $\min \gamma_{ab}$, defining the gap Δ around $\gamma = 0$. Left: head-to-head (gap 0.094 Å). Right: side-by-side (gap 0.748 Å).

Selected examples are given in Table XLIII, as obtained from the optimized geometries obtained in the present work. It can be seen that in case of triclinic crystals in case of *ortho*-TAPP-F_4H_4 no short contacts are obtained. In the monoclinic case, the H- and C_1F_3- yield a small number of short contacts while C_2F_5- and C_3F_7- lead to increased numbers of short contacts.

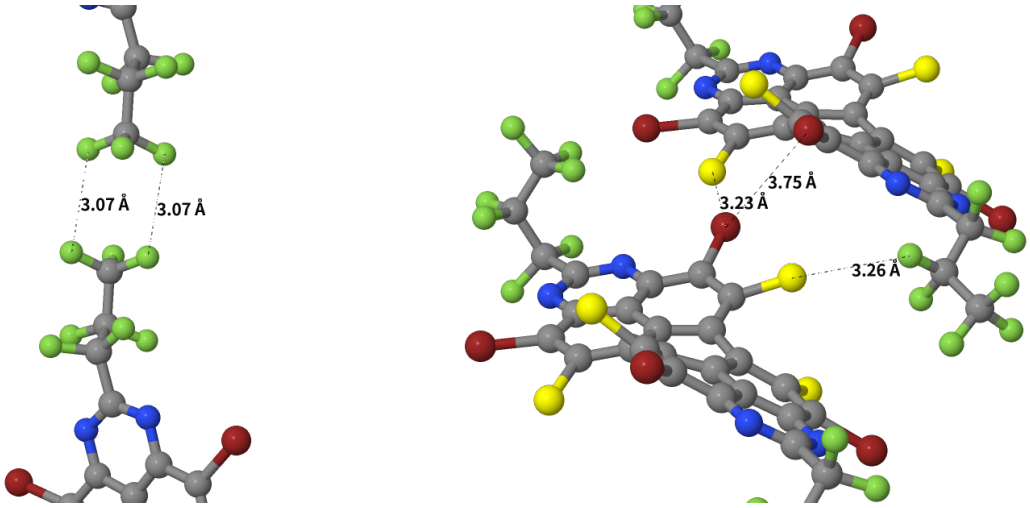


Fig. 13: Inter-atomic distance between neighbouring molecules using two triclinic *all*-C_3F_7-TAPP-Cl_4Br_4 dimer configurations as example: head-to-head(left) and side-by-side(right). Carbon atoms are displayed gray, nitrogen atoms blue, fluorine atoms green, and chlorine atoms yellow.

In order to analyze the geometries beyond only counting the amount of short contacts, the atom interactions and the distribution must be analyzed. One example of short contacts is displayed in Fig. 13 for *all*-C_3F_7-TAPP-Cl_4Br_4. On the left, the short contacts in the head-to-head are shown, on the right the side-by-side. The figure reveals that the short contacts are mostly interactions of atoms of the same kind, *i.e.* halogenes. Note however, that the short-contact criterion depends on the van-der-Waals radii so that atoms of similar distance might not be considered to be in short contact.

One important qualitative measure for the short contact analysis is in particular the gap Δ. This is illustrated in Fig. 12 for monoclinic *ortho*-TAPP-H_4F_4. The figure reveals that in case of the side-by-side configuration almost no gap Δ is obtained, in case of the side-by-side configuration a pronounced gap Δ of about 0.75 Å is obtained. In particular for an almost evenly distributed intermolecular distance for all atom pairs it is difficult to identify atoms in short contact unambiguously.

However, in several cases no short contacts are observed. For example, in case of monoclinic R-TAPP-H_8 no short contacts are obtained. Another example is given in Fig. 14 and 15. It can be seen that in these cases either the head-to-head or the side-by-side configuration contributes to short contacts. Therefore the head-to-head short contacts are added to

the side-by-side contacts and compared among the crystal systems. The crystal structure with the larger amount of short contacts is considered to be more favorable. Similarly to the thermodynamic stability, resulting data are collected in Table LXXIII.

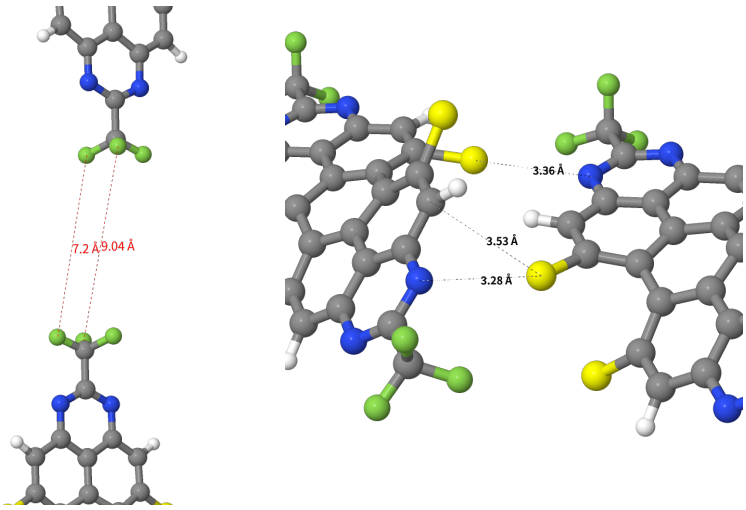


Fig. 14: Inter-atomic distance between neighbouring molecules using two triclinic *bay*-C_1F_3-TAPP-Cl_4 dimer configurations as example. Left: head-to-head. Right: side-by-side. Hydrogen atoms are displayed white, carbon atoms gray, nitrogen atoms blue, fluorine atoms green, and chlorine atoms yellow. Distances printed in black denote atoms being in short contact, distances printed in red denote atoms being not in short contact.

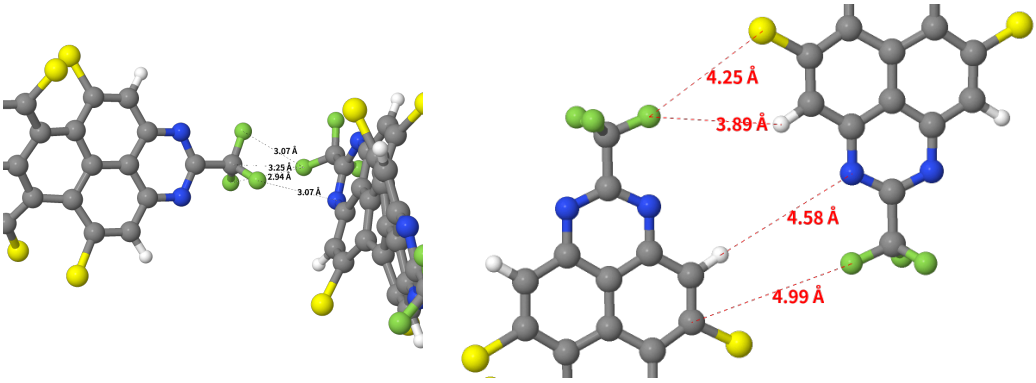


Fig. 15: Inter-atomic distance between neighbouring molecules using two orthorhombic *bay*-C_1F_3-TAPP-Cl_4 dimer configurations as example. Left: head-to-head. Right: side-by-side. Hydrogen atoms are displayed white, carbon atoms gray, nitrogen atoms blue, fluorine atoms green, and chlorine atoms yellow. Distances printed in black denote atoms being in short contact, distances printed in red denote atoms being not in short contact.

VI. BULK PHASE: ELECTRONIC GAP AND CRYSTAL STABILITY

In this section the electronic gap as well as the atomization energy are caluclated for every TAPP compound in order to get the better understanding of the semiconducting properties and also provide some insight in the molecular stability. In Table LXXIII the calculated thermodynamic and kinetic results for all the TAPP compounds with monoclinic, triclinic and orthorhombic crystals are been summarized. Letters are provided for those cases which provide the largest atomization energy per atom, in case of the kinetic stability those letters are given for which the sum of side and head short contacts give the largest number. Enthalpic results are considered equal if their difference is within 2 kJ/mol. Check and cross mark indicate if the calculated results agree with experimentally acquired ones, *i.e.* if the calculated results are in favour of monoclinic crystal while the synthesised crystal has triclinic structure, then the corresponding results will be marked with cross. We can see that for majority of the *ortho*- and *bay*-TAPP compounds, good agreement between calculated results and the experimental ones for both the thermodynamic and kinetic stability can be observed. While there are only two samples for the *all*-TAPP compounds, they all disagree with the calculated kinetic results.

A. Electronic gap

Following are the discussions of electronic gaps for different TAPP componds, each will be divided into vacuum molecule and bulk phase, each are calculated with both PBE and B3LYP functional. All the monoclinic results are followed by the energy difference between R-TAPP-X and H-TAPP-X compounds with same halogen atom X, shown in the (); while the triclinic results are followed by the energy difference between itself and corresponding monoclinic compounds, shown in {}.

1. TAPP-H compounds

Tab. XLIV: HOMO-LUMO gaps in eV for R-TAPP-H.

TAPP-H_8	Env.	Geom.	H-*	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	monoclinic	2.80	2.88 (+0.08)	2.87 (+0.08)	2.87 (+0.08)
PBE	vac	monoclinic	1.83	1.89 (+0.06)	1.89 (+0.06)	1.88 (+0.05)
B3LYP	bulk	monoclinic	1.75	1.90 (+0.15)	1.89 (+0.14)	1.88 (+0.13)
PBE	bulk	monoclinic	1.53	1.87 (+0.24)	1.86 (+0.23)	1.85 (+0.22)
TAPP-H_8	Env.	Geom.	H-†	C_1F_3-	C_2F_5-	C_3F_7-*
B3LYP	vac	triclinic	-	2.81 {-0.07}	2.83 {-0.04}	2.83 {-0.04}
PBE	vac	triclinic	-	1.84 {-0.05}	1.86 {-0.03}	1.85 {-0.03}
B3LYP	bulk	triclinic	-	2.93 {+1.03}	2.75 {+0.86}	2.90 {+1.02}
PBE	bulk	triclinic	-	2.30 {+0.43}	2.23 {+0.37}	2.29 {+0.44}

* Experimentally-obtained crystal system.

† Cell optimization not converged.

In case of the TAPP-H_8 compounds, both monoclinic bulk structure and vacuum molecules show similar pattern, that except change from H-TAPP-H_8 to R-TAPP-H_8, there is no real electronic gap change worth mentioning. However it is worth mention that both PBE and B3LYP calculate for bulk and vacuum molecule agree that replace H atom with alkyl group will cause electronic gap increase. This change of gap size is more pronounced in the bulk phase than vacuum. In case of the triclinic compounds, increase the size of alkyl groups alone shows the same pattern as in monoclinic case, which electronic gap hardly change with this increasing of substitution. Though for the molecule in vacuum case, both PBE and B3LYP calculation agree the electronic gap in triclinic molecules are lower than that of the monoclinic cases. This contradict the results of bulk phase calculation, in which results from both functional agree on increasing electronic gap. It is also worth mentioning that in both monoclinic and triclinic cases the PBE and B3LYP functional for bulk phase calculation give rather close results, while the vacuum molecule results show about 1 eV of energy difference.

2. TAPP compounds with ortho substitution

Tab. XLV: HOMO-LUMO gap in eV for ortho TAPP.

ortho TAPP-F_4H_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-*
B3LYP	vac	monoclinic	2.80	2.77 (-0.03)	2.69 (-0.11)	2.68 (-0.12)
PBE	vac	monoclinic	1.79	1.77 (-0.02)	1.70 (-0.09)	1.69 (-0.10)
B3LYP	bulk	monoclinic	1.42	1.70 (+0.28)	1.70 (+0.28)	1.69 (+0.27)
PBE	bulk	monoclinic	1.25	1.65 (+0.40)	1.65 (+0.40)	1.64 (+0.39)
ortho TAPP-Cl_4H_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	monoclinic	2.69	2.56 (-0.13)	2.57 (-0.12)	2.56 (-0.13)
PBE	vac	monoclinic	1.70	1.59 (-0.11)	1.60 (-0.10)	1.59 (-0.11)
B3LYP	bulk	monoclinic	1.72	2.29 (+0.57)	2.33 (+0.61)	2.31 (+0.59)
PBE	bulk	monoclinic	1.32	1.49 (+0.17)	1.52 (+0.20)	1.51 (+0.19)
ortho TAPP-Br_4H_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	monoclinic	2.46	2.53 (+0.07)	2.53 (+0.07)	2.54 (+0.08)
PBE	vac	monoclinic	1.51	1.56 (+0.05)	1.56 (+0.05)	1.57 (+0.06)
B3LYP	bulk	monoclinic	1.80	2.27 (+0.47)	2.28 (+0.48)	2.28 (+0.48)
PBE	bulk	monoclinic	1.36	1.46 (+0.10)	1.46 (+0.10)	1.46 (+0.10)
ortho TAPP-I_4H_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	monoclinic	2.43	2.45 (+0.02)	2.43 (+0.00)	2.44 (+0.01)
PBE	vac	monoclinic	1.46	1.46 (+0.00)	1.45 (-0.01)	1.45 (-0.01)
B3LYP	bulk	monoclinic	1.91	2.15 (+0.24)	2.14 (+0.23)	2.17 (+0.26)
PBE	bulk	monoclinic	1.40	1.34 (-0.06)	1.32 (-0.08)	1.33 (-0.07)

* Experimentally-obtained crystal system.

In Table XLV all the ortho TAPP compound of monoclinic structure are listed. For the PBE functional results the gap in the bulk phase is slightly smaller compare to the vacuum molecules, this is also supported by the results from results using B3LYP functional, even the gap energy drop in the latter case is much larger. Consider the different substitution group, it is clear that for vacuum molecules, with halogen atom change from F to I the electronic gap reduces. However in bulk phase, both PBE and B3LYP calculation show no change of the electronic gap from ortho H-TAPP-F_4H_4 to H-TAPP-I_4H_4, while only change

happens from TAPP-F_4H_4 to TAPP compounds with other halogen substitutions. For the different alkyl group, it is clear to see no electronic gap changes for molecules in vacuum with respect to the change of the alkyl group. As for the bulk phase cases, only energy jump happens during the transition from H-TAPP to R-TAPP compounds and between different alkyl groups there is almost no observable change on the electronic gaps. It is worth noticing that for ortho TAPP-F_4H_4 and TAPP-Cl_4H_4 molecule in vacuum, compounds without alkyl group have larger electronic gaps while in the rest ortho TAPPs this pattern is reversed.

Tab. XLVI: HOMO-LUMO gap in eV for ortho TAPP.

ortho TAPP-F_4H_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	triclinic	2.71 {-0.09}	2.68 {-0.09}	2.68 {-0.01}	2.68 {+0.00}
PBE	vac	triclinic	1.72 {-0.07}	1.70 {-0.07}	1.70 {+0.00}	1.69 {+0.00}
B3LYP	bulk	triclinic	2.61 {+1.19}	2.28 {+0.58}	2.30 {+0.60}	2.64 {+0.95}
PBE	bulk	triclinic	1.98 {+0.73}	1.93 {+0.28}	1.95 {+0.30}	1.99 {+0.35}
ortho TAPP-Cl_4H_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-*	C_3F_7-*
B3LYP	vac	triclinic	2.64 {-0.05}	2.65 {+0.09}	2.58 {+0.01}	2.58 {+0.02}
PBE	vac	triclinic	1.66 {-0.04}	1.67 {+0.08}	1.61 {+0.01}	1.61 {+0.02}
B3LYP	bulk	triclinic	2.22 {+0.50}	2.48 {+0.19}	2.38 {+0.05}	2.46 {+0.15}
PBE	bulk	triclinic	1.64 {+0.32}	1.87 {+0.38}	1.81 {+0.29}	1.85 {+0.34}
ortho TAPP-Br_4H_4	Env.	Geom.	H-†	C_1F_3-	C_2F_5-*	C_3F_7-*
B3LYP	vac	triclinic	-	2.55 {+0.02}	2.54 {+0.01}	2.54 {+0.00}
PBE	vac	triclinic	-	1.57 {+0.01}	1.57 {+0.01}	1.57 {+0.00}
B3LYP	bulk	triclinic	-	2.49 {+0.22}	2.33 {+0.05}	2.46 {+0.18}
PBE	bulk	triclinic	-	1.89 {+0.43}	1.76 {+0.30}	1.86 {+0.40}
ortho TAPP-I_4H_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-*
B3LYP	vac	triclinic	2.52 {+0.09}	2.45 {+0.00}	2.45 {+0.02}	2.45 {+0.01}
PBE	vac	triclinic	1.53 {+0.07}	1.46 {+0.00}	1.46 {+0.01}	1.46 {+0.01}
B3LYP	bulk	triclinic	1.99 {+0.08}	2.42 {+0.27}	2.41 {+0.27}	2.41 {+0.24}
PBE	bulk	triclinic	1.42 {+0.02}	1.82 {+0.48}	1.81 {+0.49}	1.80 {+0.47}

* Experimentally-obtained crystal system.

† Cell optimization not converged.

In Table XLVI all the ortho TAPP compound of triclinic structure are listed. In B3LYP calculations, electronic gaps of vacuum molecules are larger than gaps in bulk phase, however the trends are reversed for the PBE results. From orhto H-TAPP-F_4H_4 to H-TAPP-I_4H_4, both vacuum and bulk calculation agree that substitute with heavier halogen atom will reduce the electronic gap. But compare to the monoclinic results, triclinic bulk would results a slightly larger gap while for vacuum molecules the energy differences are not sound enough to make difference. For the R-TAPP-X_4H_4 however, vacuum molecules show no difference in electronic gaps with respect to different halogen substitutions while the bulk phase only show slight gap increase from TAPP-F_4H_4 to other halogen TAPPs. Similar to the H-TAPP-X_4H_4 cases, only triclinic in bulk phase show noticeable changes when compare to the corresponding monoclinic systems. On the other hand, changing alkyl group shows very little if not at all change on electronic gaps for all the triclinic ortho TAPP compounds.

3. TAPP compounds with bay substitution

In Table XLVII it is clear overall that changing of alkyl groups on the TAPP backbone introduces no change on the electronic gaps, with only exception is the bay TAPP-F_4H_4, which electronic gaps increase with the substitution of alkyl groups. It is also in agreement with the ortho TAPP compounds that molecules in vacuum have larger gap compare to the bulk phase. Unlike the ortho TAPP case, in monoclinic bay TAPP compounds, with substitute of heavier halogen atom the electronic gaps reduce in both H-TAPP and R-TAPP compounds.

Tab. XLVII: HOMO-LUMO gap in eV for bay TAPP.

bay TAPP-F_4H_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-†
B3LYP	vac	monoclinic	2.79	2.79 (+0.00)	2.76 (-0.03)	-
PBE	vac	monoclinic	1.79	1.79 (+0.00)	1.77 (-0.02)	-
B3LYP	bulk	monoclinic	2.25	2.48 (+0.23)	2.48 (+0.23)	-
PBE	bulk	monoclinic	1.44	1.58 (+0.14)	1.59 (+0.15)	-
bay TAPP-Cl_4H_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-*	C_3F_7-
B3LYP	vac	monoclinic	2.63	2.63 (+0.00)	2.62 (-0.01)	2.61 (-0.02)
PBE	vac	monoclinic	1.64	1.64 (+0.00)	1.63 (-0.01)	1.62 (-0.02)
B3LYP	bulk	monoclinic	2.39	2.41 (+0.02)	2.43 (+0.04)	2.45 (+0.06)
PBE	bulk	monoclinic	1.48	1.49 (+0.01)	1.51 (+0.03)	1.53 (+0.05)
bay TAPP-Br_4H_4	Env.	Geom.	H-†	C_1F_3-	C_2F_5-	C_3F_7-†
B3LYP	vac	monoclinic	2.60	2.59 (-0.01)	2.59 (-0.01)	-
PBE	vac	monoclinic	1.60	1.58 (-0.02)	1.58 (-0.02)	-
B3LYP	bulk	monoclinic	2.32	2.39 (+0.07)	2.37 (+0.05)	-
PBE	bulk	monoclinic	1.39	1.43 (+0.04)	1.41 (+0.02)	-
bay TAPP-I_4H_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-†
B3LYP	vac	monoclinic	2.49	2.46 (-0.03)	2.48 (-0.01)	-
PBE	vac	monoclinic	1.46	1.41 (-0.05)	1.43 (-0.03)	-
B3LYP	bulk	monoclinic	2.17	2.17 (+0.00)	‡	-
PBE	bulk	monoclinic	1.24	1.16 (-0.08)	1.13 (-0.11)	-

* Experimentally-obtained crystal system.

† Cell optimization not converged.

‡ Too many atoms in unit cell, calculating is not feasible even with ADMM approach

Tab. XLVIII: HOMO-LUMO gap in eV for bay TAPP.

bay TAPP-F_4H_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	triclinic	2.79 {+0.00}	2.77 {-0.02}	2.77 {+0.01}	2.76
PBE	vac	triclinic	1.79 {+0.00}	1.78 {-0.01}	1.78 {+0.01}	1.77
B3LYP	bulk	triclinic	2.54 {+0.29}	2.57 {+0.09}	2.57 {+0.09}	2.58
PBE	bulk	triclinic	1.72 {+0.28}	1.75 {+0.17}	1.75 {+0.16}	1.76
bay TAPP-Cl_4H_4	Env.	Geom.	H-†	C_1F_3-	C_2F_5-	C_3F_7-*
B3LYP	vac	triclinic	-	2.61 {-0.02}	2.61 {-0.01}	2.61 {+0.00}
PBE	vac	triclinic	-	1.63 {-0.01}	1.62 {-0.01}	1.62 {+0.00}
B3LYP	bulk	triclinic	-	2.46 {+0.05}	2.43 {+0.00}	2.43 {-0.02}
PBE	bulk	triclinic	-	1.60 {+0.11}	1.58 {+0.07}	1.58 {+0.05}
bay TAPP-Br_4H_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	triclinic	2.55 {-0.05}	2.56 {-0.03}	2.55 {-0.04}	2.55
PBE	vac	triclinic	1.56 {-0.04}	1.56 {-0.02}	1.56 {-0.02}	1.56
B3LYP	bulk	triclinic	2.34 {+0.02}	2.37 {-0.02}	2.35 {-0.02}	2.36
PBE	bulk	triclinic	1.52 {+0.13}	1.49 {+0.06}	1.47 {+0.06}	1.48
bay TAPP-I_4H_4	Env.	Geom.	H-†	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	triclinic	-	2.45 {-0.01}	2.44 {-0.04}	2.44
PBE	vac	triclinic	-	1.42 {+0.01}	1.40 {-0.03}	1.40
B3LYP	bulk	triclinic	-	2.20 {+0.03}	2.15	2.15
PBE	bulk	triclinic	-	1.25 {+0.09}	1.21 {+0.08}	1.21

* Experimentally-obtained crystal system.

† Cell optimization not converged.

In case of the triclinic bay TAPP compounds, substitution of alkyl groups function similarly as in monoclinic case, which they show almost no effect on the electronic gaps even between H-TAPP and R-TAPP compounds. Bay TAPP compounds with triclinic structures also in good agreement with monoclinic ones that molecules in vacuum also show larger electronic gap compare to the bulk phase. It is also clear that with the substitution of heavier halogen atoms, electronic gaps of bay TAPP compounds also reduce in both vacuum and bulk phase, similar as in the monoclinic case. However it is also worth noticing that compare to the monoclinic case, electronic gaps of the corresponding triclinic compounds change very little despite their differences in packing arrangements.

a. orthorhombic bay compounds Here all the electronic gaps of experimentally acquired orthorhombic bay TAPP compounds from vacuum and bulk phase are also listed. Due to the amount of atoms in the unit cell, bulk phase calculation with B3LYP functional using periodic DFT method is not possible. The PBE results for both the vacuum and the bulk phase however, also agree with the general trends of the previous bay TAPP compounds have show, with the halogen atom become heavier the electronic gaps become smaller. Also in agreement with the previous bay TAPP compounds, the electronic gaps from vacuum molecules are larger than that from the bulk phase. B3LYP results are only available for the vacuum molecules and despite the 1 eV energy shift compare to the PBE results, they also agrees with the trends which suggest the heavier halogen atom will results smaller electronic gap.

Tab. XLIX: HOMO-LUMO gap in eV for bay TAPP.

bay TAPP-F_4H_4	Env.	Geom.	C_1F_3-
B3LYP	vac	orthorhom.	2.76
PBE	vac	orthorhom.	1.77
B3LYP	bulk	orthorhom.	‡
PBE	bulk	orthorhom.	1.49
bay TAPP-Cl_4H_4	Env.	Geom.	C_1F_3-*
B3LYP	vac	orthorhom.	2.61
PBE	vac	orthorhom.	1.63
B3LYP	bulk	orthorhom.	‡
PBE	bulk	orthorhom.	1.43
bay TAPP-Br_4H_4	Env.	Geom.	C_1F_3-
B3LYP	vac	orthorhom.	2.55
PBE	vac	orthorhom.	1.54
B3LYP	bulk	orthorhom.	‡
PBE	bulk	orthorhom.	1.26
bay TAPP-I_4H_4	Env.	Geom.	C_1F_3-
B3LYP	vac	orthorhom.	2.32
PBE	vac	orthorhom.	1.25
B3LYP	bulk	orthorhom.	‡
PBE	bulk	orthorhom.	0.58

* Experimentally-obtained crystal system.

‡ Too many atoms in unit cell, calculating is not feasible even with ADMM approach

4. TAPP compounds with all substitution

In case of the all TAPP compounds, aside from the TAPP-X_8 the two different kind of TAPP-X_4Y_4 are also listed in order to better understand the affect of positioning different halogen atoms. First and foremost, it shows in the table that size of the alkyl groups does not affect the electronic gap in both the vacuum molecule and bulk phase. As also shows in the all TAPP-F_8 case that even change from H-TAPP to R-TAPP for molecules in vacuum,

electronic gap of these molecule changes almost shows no change at all while in the ortho and bay TAPP case there is a rather large energy jump for the gap. In agreement with other TAPP compounds, the monoclinic all TAPP compounds show similar drop in electronic gap when change from molecule in vacuum to the molecule in the bulk phase. With the halogen substitute become heavier, it is also clear in the Table L, electronic gaps drop in both vacuum and bulk phase. It is interesting to see that for the molecules in vacuum, both PBE and B3LYP calculation results in electronic gaps close to the unmixed all TAPP compounds. However in the bulk phase case, B3LYP calculation shows electronic gap of TAPP-Br_4Cl_4 is larger than TAPP-Br_8, which is larger than TAPP-Cl_4Br_4. PBE calculation further indicates that the electronic gap of TAPP-Cl_8 is the largest, followed by the TAPP-Br_4Cl_4, followed by the TAPP-Br_8, with the TAPP-Cl_4Br_4 is the smallest. Compare the two pairs of all TAPP compounds, TAPP-Cl_8 has larger gap than TAPP-Br_4Cl_4 and TAPP-Br_8 has gap larger than TAPP-Cl_4Br_4, it suggests that mix different halogen atom on bay position of given all TAPP compound might reduce the electronic gap of such compound.

Tab. L: HOMO-LUMO gap in eV for all TAPP.

all TAPP-F_8	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	monoclinic	2.61	2.57 (-0.04)	2.57 (-0.04)	2.57 (-0.04)
PBE	vac	monoclinic	1.63	1.60 (-0.03)	1.59 (-0.04)	1.59 (-0.04)
B3LYP	bulk	monoclinic	2.43	2.41 (-0.02)	2.43 (+0.00)	2.44 (+0.01)
PBE	bulk	monoclinic	1.55	1.53 (-0.02)	1.54 (-0.01)	1.54 (-0.01)
all TAPP-Cl_8	Env.	Geom.	H-†	C_1F_3-	C_2F_5-	C_3F_7-*
B3LYP	vac	monoclinic	-	2.39	2.38	2.38
PBE	vac	monoclinic	-	1.44	1.43	1.43
B3LYP	bulk	monoclinic	-	2.20	2.21	‡
PBE	bulk	monoclinic	-	1.33	1.34	1.34
all TAPP-Br_8	Env.	Geom.	H-†	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	monoclinic	-	2.36	2.35	2.35
PBE	vac	monoclinic	-	1.40	1.39	1.39
B3LYP	bulk	monoclinic	-	2.18	2.15	2.14
PBE	bulk	monoclinic	-	1.28	1.26	1.25
all TAPP-I_8	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	monoclinic	2.34	2.27 (-0.07)	2.27 (-0.07)	2.27 (-0.07)
PBE	vac	monoclinic	1.37	1.30 (-0.07)	1.28 (-0.09)	1.27 (-0.10)
B3LYP	bulk	monoclinic	2.09	×	1.96 (-0.13)	1.93 (-0.16)
PBE	bulk	monoclinic	1.20	1.08 (-0.12)	1.04 (-0.16)	1.02 (-0.18)
all TAPP-Br_4Cl_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	monoclinic	×	×	×	2.41
PBE	vac	monoclinic	×	×	×	1.44
B3LYP	bulk	monoclinic	×	×	×	2.21
PBE	bulk	monoclinic	×	×	×	1.32
all TAPP-Cl_4Br_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-*
B3LYP	vac	monoclinic	×	×	×	2.37
PBE	vac	monoclinic	×	×	×	1.41
B3LYP	bulk	monoclinic	×	×	×	2.06
PBE	bulk	monoclinic	×	×	×	1.19

* Experimentally-obtained crystal system.

Tab. LI: HOMO-LUMO gap in eV for all TAPP.

all TAPP-F_8	Env.	Geom.	H-†	C_1F_3-	C_2F_5-†	C_3F_7-
B3LYP	vac	triclinic	-	2.57 {-0.00}	-	2.57 {-0.00}
PBE	vac	triclinic	-	1.60 {-0.00}	-	1.59 {-0.00}
B3LYP	bulk	triclinic	-	2.34 {-0.07}	-	2.33 {-0.11}
PBE	bulk	triclinic	-	1.54 {+0.01}	-	1.55 {+0.01}
all TAPP-Cl_8	Env.	Geom.	H-†	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	triclinic	-	2.37 {-0.02}	2.38 {-0.00}	2.37 {-0.01}
PBE	vac	triclinic	-	1.42 {-0.02}	1.43 {-0.00}	1.42 {-0.01}
B3LYP	bulk	triclinic	-	2.21 {+0.01}	2.20 {-0.01}	2.18
PBE	bulk	triclinic	-	1.40 {+0.07}	1.39 {+0.05}	1.38 {+0.04}
all TAPP-Br_8	Env.	Geom.	H-†	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	triclinic	-	2.34 {-0.02}	2.34 {-0.01}	2.33 {-0.02}
PBE	vac	triclinic	-	1.39 {-0.01}	1.39 {-0.00}	1.38 {-0.01}
B3LYP	bulk	triclinic	-	2.12 {-0.06}	2.18 {+0.03}	2.16 {+0.02}
PBE	bulk	triclinic	-	1.32 {+0.04}	1.35 {+0.09}	1.34 {+0.09}
all TAPP-I_8	Env.	Geom.	H-†	C_1F_3-	C_2F_5-†	C_3F_7-
B3LYP	vac	triclinic	-	2.27 {+0.00}	-	2.25 {-0.02}
PBE	vac	triclinic	-	1.30 {+0.00}	-	1.28 {+0.01}
B3LYP	bulk	triclinic	-	2.03	-	2.06 {+0.13}
PBE	bulk	triclinic	-	1.09 {+0.01}	-	1.09 {+0.07}
all TAPP-Br_4Cl_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	triclinic	×	×	×	2.36 {-0.05}
PBE	vac	triclinic	×	×	×	1.41 {-0.03}
B3LYP	bulk	triclinic	×	×	×	2.18 {-0.03}
PBE	bulk	triclinic	×	×	×	1.36 {+0.04}
all TAPP-Cl_4Br_4	Env.	Geom.	H-	C_1F_3-	C_2F_5-	C_3F_7-
B3LYP	vac	triclinic	×	×	×	2.34 {-0.03}
PBE	vac	triclinic	×	×	×	1.39 {-0.02}
B3LYP	bulk	triclinic	×	×	×	×
PBE	bulk	triclinic	×	×	×	1.36 {+0.15}

* Experimentally-obtained crystal system.

For the all TAPP compounds listed in the Table LI, although without H-TAPP results for comparison, the results still clearly agree with the previous discussion, which concludes that change the alkyl group shows no effect on the electronic gaps of the TAPP compounds. Similar to other TAPP compounds, triclinic all TAPPs also show electronic gap decrease from molecules in vacuum to the bulk phase. In case of the halogenation effect, triclinic compounds also agree with the monoclinic ones that with the heavier halogen atom as substitution, electronic gaps decrease. However compare the electronic gaps between monoclinic and triclinic shows that there is no clear when the crystal structure changes. In case of the TAPP-X_4Y_4 compounds, it is shown that electronic gaps do not change in the vacuum molecules when halogen substitutes on the backbone TAPP. As for the bulk phase, electronic gap of all TAPP-Br_4Cl_4 calculated with B3LYP is almost the same as TAPP-Br_8 and TAPP-Cl_8. The PBE results also show no real difference between TAPP-Br_4Cl_4, TAPP-Cl_4Br_4, TAPP-Br_8 and TAPP-Cl_8 in bulk phase.

B. Atomization energy

In this section the thermodynamic stability of TAPP compounds based on their atomization energy will be discussed. This section is organized as following, first some small molecules are calculated with the same methods as TAPP to serve as the reference system, their calculated atomization will be compared with the experimentally acquire data in order to establish the reliability of the methods, then the atomization energy of each TAPP compounds will be listed and discussed. For each TAPP compound, monoclinic and triclinic structures are listed in the same table, and for each crystal type three different atomization energies are listed. Atomization energy for unit cell is calculated for all the atoms in one unit cell, this property would likely to give a indication of relative stability of the unit cell. Atomization energy per molecule is calculated for all the atoms in one TAPP molecule, this property gives indication on how thermodynamically stable a molecule is. Atomization energy per unit cell and per molecule are related to the size of cell and molecule, with larger system naturally the atomization energy also increase. In order to eliminate this size effect, atomization per atom for each TAPP compounds is also calculated and this would serve as the focus of our discussion.

1. Reference systems

Tab. LII: Reference calculation of atomization energy with small molecules as test cases, all the calculations are done with CP2K package using pseudopotential and GTH-TZVP basis set. Enrgies in kJ/mol, errors in percentage.

Molecule	Chemical formula	Calculated results	Error
Methane	CH_4	1,747	+5.1
Ethan	C_2H_6	2,983	+5.6
Propane	C_3H_8	4,207	+5.2
Benzene	C_6H_6	5,828	+5.5
Toluene	C_7H_8	7,075	+5.4
Ethylbenzene	C_8H_{10}	8,311	+5.4
Biphenyl	$C_{12}H_{10}$	11,168	+5.4
Naphthalene	$C_{10}H_8$	9,231	+5.4
Anthracene	$C_{14}H_{10}$	12,630	+5.4
Pyridine	C_5H_5N	5,377	+7.4
Pyrazine	$C_4H_4N_2$	4,921	+9.7
Pyrazole	$C_3H_4N_2$	4,194	+10.7

Here lists all the atomization energies of the reference molecules. The choice of the reference molecules are firstly based on the structural similarity between these molecules and the fracture of TAPP molecules. Secondly, the influence of the chain length and number of ring structure have on the atomization energy is also considered. The error is calculated as $((E - E_{\text{ref}})/E_{\text{ref}}) \times 100$, with E_{ref} taken from NIST Computational Chemistry Comparison and Benchmark Database `cccbdb.nist.gov`. As shown in the Table LII, in most of the cases the atomization energy calculated with our setting would results values only 5% larger than reference values, which consider the choice of methods and basis set is satisfied.

2. TAPP compounds

a. TAPP-H_8 compounds

Tab. LIII: Cell information for R-TAPP-H. Angles are given in degrees, distances in Å, energies in kJ/mol.

TAPP-H_8	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic*	monoclinic	monoclinic	monoclinic
AE / per unit cell	42,008	47,351	52,705	58,041
AE / per molecule	21,004	23,675	26,352	29,021
AE / per atom	583	564	549	537
Geom.	triclinic†	triclinic	triclinic	triclinic*
AE / per unit cell	-	23,981	26,627	29,257
AE / per molecule	-	23,981	26,627	29,257
AE / per atom	-	571	555	541

* Experimentally-obtained crystal system.

† Cell optimization not converged.

In Table LIII all the atomization energies of TAPP-H_8 for both the monoclinic and triclinic compounds are listed. It is clear to see that the atomization energy difference between unit cell and molecule is due to their size difference, i.e. different amount of atoms in each system. Thus it is more useful to average the atomization energy over all atoms, which would results the atomization energy per atom. It is therefore obverse that there is no clear preference toward any structure, be it monoclinic or triclinic, and the energy difference between atomization energy of given TAPP compound with different structures is small enough to be consider have no chemical importance since most of such difference are within 10 K/mol. It is also easy to notice that among each structure, changing alkyl groups would results gradually reduced atomization with the increase size of alkyl group.

b. TAPP compounds with ortho substitution Table LXXIV to Table LXXVII show the atomization energy information for all the *ortho*-TAPP compounds. It is in agreement with the TAPP-H_8 compounds that there is little to no change of atomization energy with respect to monoclinic and triclinic structures, which indicates there is no strong preference of one structure over another thermodynamically. It is also agree the conclusion of TAPP-H_8 which with larger alkyl substitution the atomization energy is reduce gradually. It is also noticed that with heavier halogen atom, the atomization enery also slowly decrease, however this

reduction is not very pronounced, with only about 10 kJ/mol between TAPP-X (X=F, Cl, Br and I, respectively) compounds.

c. TAPP compounds with bay substitution Table LXXVIII to Table LXXXI show the atomization energy of all the *bay*-TAPP compounds of both monoclinic and triclinic structures. Same as in *ortho*-TAPP cases, between monoclinic and triclinic crystals there is almost no change on atomization energy and this again support the argument that there is no thermodynamic preference between two structures. With the increase of the alkyl group on the TAPP, atomization energy gradually decrease as happens in the *ortho*-TAPPs. Even with the increase size of the halogen substitution, atomization energy also decrease in the *bay*-TAPP compounds, but the decrease still rather subtle as shown in the *ortho* cases.

Tab. LIV: Cell information for bay TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

bay C_1F_3-TAPP	-H_4F_4	-H_4Cl_4	-H_4Br_4	-H_4I_4
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
AE / per unit cell	95,327	93,256	92,269	91,344
AE / per molecule	23,832	23,314	23,067	22,836
AE / per atom	567	555	549	544
Geom.	triclinic	triclinic	triclinic	triclinic
AE / per unit cell	47,580	46,566	46,097	45,663
AE / per molecule	23,790	23,283	23,049	22,832
AE / per atom	566	554	549	544
Geom.	orthorhomb.	orthorhomb.*	orthorhomb.	orthorhomb.
AE / per unit cell	19,0638	18,6661	18,4737	18,2778
AE / per molecule	23,830	23,333	23,092	22,847
AE / per atom	567	556	550	544

* Experimentally-obtained crystal system.

d. orthorhombic bay-TAPP compounds Table LXXVIII to Table LIV show the atomization energies of the *bay*-TAPP compounds. Similar to the TAPP-H_8 and *ortho*-TAPP compounds, there is no worth mentioning changes between same TAPP compounds with monoclinic and triclinic structures. Orthorhombic *bay*-TAPP compounds also support the conclusion that change the structure type would not change atomization energy. As for the effect of changing the alkyl group on *bay*-TAPPs of different structure, atomization energy drops with increased size of alkyl group with almost same rate as in *ortho*-TAPP cases. In case of the halogenation effect, change halogen atom from F to I in *bay*-TAPP compounds, again the atomization energy shows about 10 kJ/mol decrease.

e. TAPP compounds with all substitution From Table LXXXII to Table LXXXVI all the atomization energies of the all TAPP compounds are listed, include the TAPP-X_4Y_4. Unsurprisingly, all TAPP compounds also following the same pattern as in previous cases, which shows change the structure would not affect the atomization energy in meaningful way. Similar energy drop around 10 kJ/mol can also be observed for all TAPP with different alkyl substitution groups. Same as bay or ortho TAPP compounds, heavier halogen atom as substitution group would also induce atomization energy drop around 10 kJ/mol. Mixed all TAPP compounds show atomization energy fall between all TAPP-Cl_8 and TAPP-Br_8.

VII. SUMMARY

In order to get a better understanding on the geometric and electronic properties of TAPP compounds with different substitution group and packing arrangement in bulk phase. New IP/EA method have been developed which can be used, with combination of FDE approach, for electronic property calculation of large chemical system such as TAPP clusters. Furthermore the periodic DFT method is very powerful when optimize the structure of the bulk phase system, which provide the great opportunity for analysis the geometric properties of the crystal structures of different TAPP compounds.

In the Section II several post-HF such as CC2, CIS(D_∞) and ADC(2) are given short theoretical discussion in order to establish the basic understanding of IP and EA methods used for the calculating the property of TAPP clusters later. Later in this section the new IP/EA method based on the combination of continuum orbital and ADC(2) scheme is derived and applied for testing reference systems and the TAPP variations with different substitution groups and system size. Similarly, conventional and periodic DFT methods, as well as the auxiliary density matrix method, are also discussed briefly for they are the main methods used in the studying of the bulk phase properties of the TAPP. It is also in this section the concept of modified short contact is introduced to serve as the tool to study the possible inter-molecular interactions between TAPP molecules in the bulk phase and the influences of the different substitution group hold for the packing arrangements of TAPP molecules.

In the Section IV, we report the derivation and implementation to compute vertical ionization potentials (IPs) and electron affinities (EAs) for embedded wavefunction methods as well as the corresponding analytical nuclear gradients for IP/EA-ADC(2) as well as IP/EA-CIS(D_∞). The methods were applied to selected case studies, with an emphasis on localizing the charges as introduced by the EOM-IP and EOM-EA ansatz in dimers. For the selected dimers, the error of the dimer shift in the supermolecular calculations is computed to be below 0.5 eV compared to supermolecular reference CCSD results for the distance at 3 Å, while for increased distances of 5 Å and 7 Å, the deviation to reference methods is found to be sufficiently small. The embedding approach localizes the charge onto one molecule, which has some conceptual advantages for studying charge transfer in molecular crystals. Additionally, it reduces significantly the computational costs, allowing us to study the TAPP

derivatives used as organic semiconductors in increased cluster models.

The FDE ansatz enforces a localization of the charge on one subsystem and can thus lead to significant differences with respect to supermolecular calculations, in particular for short distances. However, the FDE ansatz allows also to include, for example, two TAPP molecules in the active subsystem, allowing us to study long-range effects and also to investigate coupling in different directions.

The methods implemented in the present work are based on a previous implementation of analytical nuclear gradients for excited-state CC2 combined with FDE. While ADC(2) and CC2 are sufficiently accurate in most cases for purely valence-excited states, IP-ADC(2) and EA-ADC(2) can exhibit deviations in terms of absolute values of about 2 eV for IPs and EAs compared to IP-CCSD or EA-CCSD, respectively. Future work should address gradient contributions due to embedding potentials polarized by the electron change and corrections toward IP-CCSD[k]$_{CC2}$ or ADC(2)-x.

In the Section V the geometric properties of different TAPP crystal structures are discussed and comparison between some TAPP type with different substitution groups have been made. The geometric properties of different TAPP crystals in this section have been calculated with periodic DFT method without apply ADMM method, PBE functional and TZVP basis set are used for all the geometry optimization. The accuracy of the geometries is assessed by a comparison to experimentally obtained crystal data. It is shown that with TAPP type fixed, the changing of the substitution group holds very little influence over the overall geometric properties of the TAPP crystals. Cell parameters, i.e. cell angles and cell vectors, show tendency of increase unit cell size with larger substitution group but the unit cell shape mostly stay unchanged. As for the packing arrangement of TAPP crystals, namely the intra-molecular angles and the inter-molecular shifts, it is shown that larger substitution group would cause the TAPP molecules in neighbouring move away from each other, but the directions of such move is uncertain. For the intra-molecular angles though, the substitution group shows very limited changes with respect to the alkyl group but the influence of the halogen substitutes are more pronounced. Lastly the short contacts which provide us with some insight about the possible inter-atomic interaction between two neighbouring TAPP molecules are discussed.

As the TAPP compounds crystallize in different crystal systems, we used the geometries in a first step to estimate the stability of the crystal structures. For this, we used the

atomization energies as a measure for the thermodynamic stability and the concept of short contacts to estimate the kinetic stability, with the latter being discussed in Section V as mentioned before while the latter being discussed in Section VI.

As a last example, we discuss the electronic gap, *i.e.* the HOMO-LUMO gap, which can serve as a qualitative measure for semiconducting properties. We find that bulk properties might be significantly different from vacuum results. Additionally in some cases we observe a strong dependence on the crystal system while in other cases the crystal system has almost no influence. It is observed that the TAPP compounds in most cases show rather subtle changes on energy differences and a definite prediction remains challenging. In case of C_3F_7-TAPP-F_4, however, the energetic product is contradicting the experimental findings, which can only be explained taking into account the kinetic stability.

Appendices

Appendix A: IP/EA method

1. reference system

Tab. LV: Vertical IP reference results with supermolecular approach: benzene*

Distance	CC2-IP	ADC(2)-IP	CIS(D∞)-IP	G0W0	ΔCCSD(T)
3.0	7.589	7.640	7.640	7.802	8.104
3.5	8.137	8.188	8.189	8.309	8.609
4.0	8.442	8.494	8.494	8.585	8.888
4.5	8.616	8.669	8.670	8.740	9.047
5.0	8.717	8.773	8.773	8.831	9.141
7.5	8.877	8.938	8.938	8.974	9.284
Monomer	8.945	8.999	9.000	9.027	9.340

* all vaules are in eV

Tab. LVI: Vertical IP reference results with FDE approach: benzene*

Distance	CC2-IP	ADC(2)-IP	CIS(D∞)-IP	G0W0	ΔCCSD(T)
3.0	8.661	8.724	8.724	8.761	8.104
3.5	8.716	8.779	8.779	8.812	8.609
4.0	8.757	8.819	8.819	8.848	8.888
4.5	8.788	8.851	8.851	8.876	9.047
5.0	8.814	8.877	8.877	8.900	9.141
7.5	8.887	8.950	8.950	8.965	9.284
Monomer	8.945	8.999	9.000	9.027	9.340

* all vaules are in eV

2. TAPP compounds

Tab. LVII: R-TAPP-X vertical ionization potentials*

Molecule	PBE0	G_0W_0	IP-CIS(D_∞)	IP-ADC(2)
H-TAPP-H	6.52	7.60	7.50	7.50
H-TAPP-F	6.64	7.73	7.68	7.68
H-TAPP-Cl	6.64	7.67	7.61	7.61
H-TAPP-Br	6.66	7.66	7.60	7.60
C_3F_7-TAPP-H	7.08	8.12	8.09	8.09
C_3F_7-TAPP-F	7.23	8.27	8.30	8.30
C_3F_7-TAPP-Cl	7.09	8.08	8.10	8.10
C_3F_7-TAPP-Br	7.08	8.06	8.07	8.07
H-TAPP-H dimer	6.47 (-0.05)	7.47 (-0.13)	7.48 (-0.02)	7.48 (-0.02)
H-TAPP-H trimer	6.46 (-0.06)	7.40 (-0.20)	7.47 (-0.03)	7.47 (-0.03)
H-TAPP-H cluster	×	×	7.51 (+0.01)	7.51 (+0.01)

* Results are given in eV

Tab. LVIII: R-TAPP-X vertical electron attachment energy*

Molecule	PBE0	G_0W_0	EA-CIS(D_∞)	EA-ADC(2)
H-TAPP-H	-3.27 (-3.31)	-2.09 (-1.80)	-2.36 (-2.14)	-2.36 (-2.15)
H-TAPP-F	-3.57	-2.40	-2.74	-2.74
H-TAPP-Cl	-3.70	-2.62	-3.03	-3.03
H-TAPP-Br	-3.73	-2.69	-3.14	-3.14
C_3F_7-TAPP-H	-3.82	-2.68	-2.98	-2.98
C_3F_7-TAPP-F	-4.16	-3.05	-3.42	-3.42
C_3F_7-TAPP-Cl	-4.16	-3.12	-3.56	-3.56
C_3F_7-TAPP-Br	-4.17	-3.17	-3.65	-3.65
H-TAPP-H dimer	-3.41 (-3.46)	-2.34 (-2.08)	-2.35 (-2.13)	-2.35 (-2.13)
H-TAPP-H trimer	× (-3.21)	×	-2.33 (-2.11)	-2.33 (-2.11)
H-TAPP-H cluster	×	×	× (-2.17)	× (-2.18)

* Results are given in eV

Appendix B: Geometric analysis

1. Overview

Tab. LIX: Overview of monoclinic (m), triclinic (t), and orthorhombic (o) crystal structures computed in the present work and experimentally obtained crystal structures in the literature.

	Exp. CDCC[x]	This work		Exp. CDCC[x]	This work
H-TAPP-F_4	×	m,t	H-TAPP-H_8	m (747374)[a]	m,t[†]
H-TAPP-Cl_4	×	m,t	C_1F_3-TAPP-H_8	×	m,t
H-TAPP-Br_4	×	m,t[†]	C_2F_5-TAPP-H_8	×	m,t
H-TAPP-I_4	×	m,t	C_3F_7-TAPP-H_8	t (844501)[b]	m,t

	ortho-		*bay-*			*all-*	
	Exp.[x]	This work	Exp.[x]	This work		Exp.[x]	This work
C_1F_3-TAPP-F_4	×	m,t	×	o,m,t	C_1F_3-TAPP-F_8	×	m,t
C_1F_3-TAPP-Cl_4	×	m,t	o (1908780)[d]	o,m,t	C_1F_3-TAPP-Cl_8	×	m,t
C_1F_3-TAPP-Br_4	×	m,t	×	o,m,t	C_1F_3-TAPP-Br_8	×	m,t
C_1F_3-TAPP-I_4	×	m,t	×	o,m,t	C_1F_3-TAPP-I_8	×	m,t
C_2F_5-TAPP-F_4	×	m,t	×	m,t	C_2F_5-TAPP-F_8	×	m,t[†]
C_2F_5-TAPP-Cl_4	t (844502)[b]	m,t	m (1908781)[d]	m,t	C_2F_5-TAPP-Cl_8	×	m,t
C_2F_5-TAPP-Br_4	t (911239)[e]	m,t	×	m,t	C_2F_5-TAPP-Br_8	×	m,t
C_2F_5-TAPP-I_4	×	m,t	×	m,t	C_2F_5-TAPP-I_8	×	m,t[†]
C_3F_7-TAPP-F_4	m (1406814)[c]	m,t	×	m[†],t	C_3F_7-TAPP-F_8	×	m,t
C_3F_7-TAPP-Cl_4	t (844503)[b]	m,t	t (1908782)[d]	m,t	C_3F_7-TAPP-Cl_8	m (1980430)[f]	m,t
C_3F_7-TAPP-Br_4	t (911240)[e]	m,t	×	m[†],t	C_3F_7-TAPP-Br_8	×	m,t
C_3F_7-TAPP-I_4	t (1406815)[c]	m,t	×	m[†],t	C_3F_7-TAPP-I_8	×	m,t
					C_3F_7-TAPP-Br_4Cl_4	×	m,t
					C_3F_7-TAPP-Cl_4Br_4	m (1980431)[f]	m,t

[a] Ref. 129, [b] Ref. 130, [c] Ref. 38, [d] Ref. 132, [e] Ref. 131, [f] Ref. 35

[x] Entries of Cambridge Crystallographic Data Centre (CCDC) [128] in parentheses.

[†] Solid-state geometry optimization did not converge.

2. Short contact

a. TAPP compounds with ortho substitution

Tab. LX: Short contact information for ortho TAPP.

ortho TAPP-F_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic*
same-stack	1.05/59/0.10	1.05/0/NE	1.05/0/NE	1.05/0/NE
head-to-head	1.05/2/0.17	1.05/5/0.18	1.05/5/0.20	1.05/32/0.08
side-by-side	1.05/3/0.56	1.05/2/0.97	1.05/30/0.11	1.05/8/0.75
Geom.	triclinic	triclinic	triclinic	triclinic
same-stack	1.05/89/0.04	1.05/107/0.05	1.05/113/0.03	1.05/123/0.03
head-to-head	1.05/0/NE	1.05/0/NE	1.05/0/NE	1.05/0/NE
side-by-side	1.05/0/NE	1.05/0/NE	1.05/0/NE	1.05/0/NE

* Experimentally-obtained crystal system.

Tab. LXI: Short contact information for ortho TAPP.

ortho TAPP-Cl_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
same-stack	1.05/57/0.04	1.05/0/NE	1.05/0/NE	1.05/0/NE
head-to-head	1.05/8/0.08	1.05/3/0.07	1.05/4/0.09	1.05/4/0.06
side-by-side	1.05/0/NE	1.05/2/1.04	1.05/2/0.96	1.05/4/0.91
Geom.	triclinic	triclinic	triclinic*	triclinic*
same-stack	1.05/85/0.01	1.05/107/0.08	1.05/111/0.00	1.05/117/0.02
head-to-head	1.05/2/0.20	1.05/0/NE	1.05/0/NE	1.05/2/0.53
side-by-side	1.05/2/0.31	1.05/2/0.35	1.05/3/0.35	1.05/2/0.14

* Experimentally-obtained crystal system.

Tab. LXII: Short contact information for ortho TAPP.

ortho TAPP-Br_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
same-stack	1.05/55/0.09	1.05/0/NE	1.05/0/NE	1.05/0/NE
head-to-head	1.05/8/0.09	1.05/3/0.08	1.05/4/0.12	1.05/4/0.78
side-by-side	1.05/0/NE	1.05/2/0.95	1.05/2/0.92	1.05/3/0.23
Geom.	triclinic†	triclinic	triclinic*	triclinic*
same-stack	-	0.95/400/0.03	1.05/107/0.02	1.05/117/0.01
head-to-head	-	1.05/0/NE	1.05/1/0.74	1.05/12/0.14
side-by-side	-	1.05/15/0.20	1.05/5/0.42	1.05/18/0.20

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXIII: Short contact information for ortho TAPP.

ortho TAPP-I_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
same-stack	1.05/22/0.06	1.05/0/NE	1.05/0/NE	1.05/0/NE
head-to-head	1.05/0/NE	1.05/0/NE	1.05/4/0.09	1.05/5/0.55
side-by-side	1.05/0/NE	1.05/2/0.98	1.05/2/0.51	1.05/5/0.15
Geom.	triclinic	triclinic	triclinic	triclinic*
same-stack	1.05/68/0.03	1.05/101/0.03	1.05/97/0.05	1.05/84/0.00
head-to-head	1.05/0/NE	1.05/0/NE	1.05/0/NE	1.05/6/0.23
side-by-side	1.05/2/0.80	1.05/2/0.31	1.05/15/0.20	1.05/24/0.03

* Experimentally-obtained crystal system.

b. TAPP compounds with bay substitution

Tab. LXIV: Short contact information for bay TAPP.

bay TAPP-F_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic†
same-stack	1.05/8/0.07	1.05/8/0.11	1.05/8/0.08	-
head-to-head	1.05/0/NE	1.05/0/NE	1.05/1/1.11	-
side-by-side	1.05/4/0.19	1.05/0/NE	1.05/3/0.24	-
Geom.	triclinic	triclinic	triclinic	triclinic
same-stack	1.05/12/0.05	1.05/6/0.26	1.05/8/0.02	1.05/10/0.04
head-to-head	1.05/0/NE	1.05/0/NE	1.05/0/NE	1.05/2/0.62
side-by-side	1.05/2/0.12	1.05/0/NE	1.05/0/NE	1.05/2/0.39

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXV: Short contact information for bay TAPP.

bay TAPP-Cl_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic*	monoclinic
same-stack	1.05/6/0.16	1.05/12/0.08	1.05/12/0.02	1.05/11/0.03
head-to-head	1.05/0/NE	1.05/0/NE	1.05/1/1.11	1.05/3/0.03
side-by-side	1.05/3/0.09	1.05/2/0.26	1.05/2/0.07	1.05/3/0.22
Geom.	triclinic†	triclinic	triclinic	triclinic*
same-stack	-	1.05/12/0.11	1.05/16/0.01	1.05/10/0.02
head-to-head	-	1.05/0/NE	1.05/0/NE	1.05/2/0.61
side-by-side	-	1.05/3/0.10	1.05/4/0.10	1.05/0/NE

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXVI: Short contact information for bay TAPP.

bay TAPP-Br_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic†
same-stack	1.05/6/0.08	1.05/14/0.02	1.05/14/0.08	-
head-to-head	1.05/0/NE	1.05/0/NE	1.05/1/1.11	-
side-by-side	1.05/3/0.17	1.05/1/0.45	1.05/1/0.11	-
Geom.	triclinic	triclinic	triclinic	triclinic
same-stack	1.05/6/0.01	1.05/10/0.12	1.05/8/0.03	1.05/16/0.02
head-to-head	1.05/4/0.20	1.05/0/NE	1.05/0/NE	1.05/2/0.58
side-by-side	1.05/2/0.19	1.05/0/NE	1.05/7/0.15	1.05/2/0.22

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXVII: Short contact information for bay TAPP.

bay TAPP-I_4H_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic†
same-stack	1.05/4/0.09	1.05/8/0.04	1.05/8/0.09	-
head-to-head	1.05/0/NE	1.05/0/NE	1.05/1/1.11	-
side-by-side	1.05/2/0.19	1.05/1/0.65	1.05/1/1.11	-
Geom.	triclinic†	triclinic	triclinic	triclinic
same-stack	-	1.05/10/0.11	1.05/8/0.07	1.05/6/0.02
head-to-head	-	1.05/0/NE	1.05/0/NE	1.05/0/NE
side-by-side	-	1.05/6/0.03	1.05/2/0.13	1.05/6/0.10

* Experimentally-obtained crystal system.

† Cell optimization not converged.

c. TAPP compounds with all substitution

Tab. LXVIII: Short contact information for all TAPP.

all TAPP-F_8	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
same-stack	1.05/11/0.01	1.05/8/0.05	1.05/11/0.04	1.05/12/0.06
head-to-head	1.05/0/NE	1.05/0/NE	1.05/1/0.15	1.05/2/0.10
side-by-side	1.05/0/NE	1.05/4/0.10	1.05/0/NE	1.05/0/NE
Geom.	triclinic†	triclinic	triclinic†	triclinic
same-stack	-	1.05/16/0.08	-	1.05/10/0.10
head-to-head	-	1.05/0/NE	-	1.05/2/0.28
side-by-side	-	1.05/8/0.34	-	1.05/2/0.11

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXIX: Short contact information for all TAPP.

all TAPP-Cl_8	H-†	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic*
same-stack	-	1.05/10/0.03	1.05/11/0.04	1.05/10/0.04
head-to-head	-	1.05/0/NE	1.05/1/0.39	1.05/2/0.06
side-by-side	-	1.05/0/NE	1.05/3/0.56	1.05/6/0.25
Geom.	triclinic	triclinic	triclinic	triclinic
same-stack	-	1.05/12/0.03	1.05/10/0.11	1.05/12/0.13
head-to-head	-	1.05/0/NE	1.05/0/NE	1.05/2/0.52
side-by-side	-	1.05/3/0.18	1.05/3/0.14	1.05/4/0.06

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXX: Short contact information for all TAPP.

all TAPP-Br_8	H-†	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
same-stack	-	1.05/6/0.08	1.05/6/0.01	1.05/8/0.02
head-to-head	-	1.05/0/NE	1.05/0/NE	1.05/1/0.22
side-by-side	-	1.05/1/0.35	1.05/2/0.17	1.05/4/0.03
Geom.	triclinic	triclinic	triclinic	triclinic
same-stack	-	1.05/6/0.09	1.05/4/0.10	1.05/8/0.02
head-to-head	-	1.05/0/NE	1.05/0/NE	1.05/2/0.42
side-by-side	-	1.05/4/0.24	1.05/0/NE	1.05/0/NE

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXXI: Short contact information for all TAPP.

all TAPP-I_8	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
same-stack	1.05/5/0.08	1.05/7/0.05	1.05/9/0.05	1.05/8/0.06
head-to-head	1.05/6/0.12	1.05/0/NE	1.05/0/NE	1.05/1/0.18
side-by-side	1.05/2/0.59	1.05/1/0.89	1.05/1/0.68	1.05/2/0.28
Geom.	triclinic†	triclinic	triclinic†	triclinic
same-stack	-	1.05/4/0.12	-	1.05/8/0.09
head-to-head	-	1.05/0/NE	-	1.05/0/NE
side-by-side	-	1.05/2/0.24	-	1.05/2/0.66

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Appendix C: Bulk phase: electronic gap and crystal stability

Tab. LXXII: Short contact information for all TAPP.

all C_3F_7-TAPP	-Br_4Cl_4	-Cl_4Br_4
Geom.	monoclinic	monoclinic*
same-stack	1.05/14/0.02	1.05/18/0.03
head-to-head	1.05/0/NE	1.05/0/NE
side-by-side	1.05/2/0.04	1.05/3/0.38
Geom.	triclinic	triclinic
same-stack	1.05/6/0.10	1.05/8/0.09
head-to-head	1.05/2/0.48	1.05/2/0.27
side-by-side	1.05/2/0.48	1.05/3/0.16

* Experimentally-obtained crystal system.

1. Predication comparison with experimental structures

Tab. LXXIII: Overview of predicted stability for monoclinic (m), triclinic (t), and orthorhombic (o) crystal structures computed in the present work.

	Therm	Kin
H-TAPP-F_4	m,t	m
H-TAPP-Cl_4	t	m
H-TAPP-Br_4	m[a]	m[a]
H-TAPP-I_4	m,t	t

	Therm	Kin
H-TAPP-H_8	m[a]	m[a]
C_1F_3-TAPP-H_8	t	m
C_2F_5-TAPP-H_8	t	m,t
C_3F_7-TAPP-H_8	t ✓	m ✗

	ortho-		*bay-*	
	Therm	Kin	Therm	Kin
C_1F_3-TAPP-F_4	t	m	m,t,o	o
C_1F_3-TAPP-Cl_4	t	m	m,t,o ✓	o ✓
C_1F_3-TAPP-Br_4	t	t	m,t,o	m,o
C_1F_3-TAPP-I_4	t	m,t	m,t,o	o
C_2F_5-TAPP-F_4	t	m	m,t	m
C_2F_5-TAPP-Cl_4	t ✓	t ✓	m,t ✓	t ✗
C_2F_5-TAPP-Br_4	t ✓	t ✓	m,t	t
C_2F_5-TAPP-I_4	t	t	m,t	m,t
C_3F_7-TAPP-F_4	t ✗	m ✓	t[b]	t[b]
C_3F_7-TAPP-Cl_4	t ✓	t ✓	m,t ✓	m ✗
C_3F_7-TAPP-Br_4	t ✓	t ✓	t[b]	t[b]
C_3F_7-TAPP-I_4	t ✓	t ✓	t[b]	t[b]

	all-	
	Therm	Kin
C_1F_3-TAPP-F_8	m,t	t
C_1F_3-TAPP-Cl_8	m,t	t
C_1F_3-TAPP-Br_8	m,t	t
C_1F_3-TAPP-I_8	m,t	t
C_2F_5-TAPP-F_8	m[a]	m[a]
C_2F_5-TAPP-Cl_8	m,t	m
C_2F_5-TAPP-Br_8	m,t	m
C_2F_5-TAPP-I_8	m[a]	m[a]
C_3F_7-TAPP-F_8	m,t	t
C_3F_7-TAPP-Cl_8	m,t ✓	m ✗
C_3F_7-TAPP-Br_8	m,t	m
C_3F_7-TAPP-I_8	m,t	m
C_3F_7-TAPP-Br_4Cl_4	m,t	t
C_3F_7-TAPP-Cl_4Br_4	m,t ✓	t ✗

* Experimentally obtained crystal system.

[a] Geometry optimization not converged in case of triclinic polymorph.

[b] Geometry optimization not converged in case of monoclinic polymorph.

2. Atomization energy

a. TAPP compounds with ortho substitution

Tab. LXXIV: Cell information for ortho TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

ortho TAPP-H_4F_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic*
AE / per unit cell	42,329	47,623	52,977	58,312
AE / per molecule	21,164	23,812	26,489	29,156
AE / per atom	588	567	552	540
Geom.	triclinic	triclinic	triclinic	triclinic
AE / per unit cell	21,219	24,084	26,764	29,408
AE / per molecule	21,219	24,084	26,764	29,408
AE / per atom	589	573	558	545

* Experimentally-obtained crystal system.

Tab. LXXV: Cell information for ortho TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

ortho TAPP-H_4Cl_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
AE / per unit cell	41,457	46,704	52,060	57,393
AE / per molecule	20,729	23,352	26,030	28,697
AE / per atom	576	556	542	531
Geom.	triclinic	triclinic	triclinic*	triclinic*
AE / per unit cell	20,834	23,633	26,323	28,970
AE / per molecule	20,834	23,633	26,323	28,970
AE / per atom	579	563	548	536

* Experimentally-obtained crystal system.

Tab. LXXVI: Cell information for ortho TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

ortho TAPP-H_4Br_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
AE / per unit cell	40,959	46,263	51,598	56,940
AE / per molecule	20,480	23,131	25,799	28,470
AE / per atom	569	551	537	527
Geom.	triclinic†	triclinic	triclinic*	triclinic*
AE / per unit cell	-	23,374	26,068	28,716
AE / per molecule	-	23,374	26,068	28,716
AE / per atom	-	557	543	532

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXXVII: Cell information for ortho TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

ortho TAPP-H_4I_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
AE / per unit cell	40,457	45,855	51,188	56,533
AE / per molecule	20,228	22,927	25,594	28,266
AE / per atom	562	546	533	523
Geom.	triclinic	triclinic	triclinic	triclinic*
AE / per unit cell	20,304	23,112	25,793	28,460
AE / per molecule	20,304	23,112	25,793	28,460
AE / per atom	564	550	537	527

* Experimentally-obtained crystal system.

b. TAPP compounds with bay substitution

Tab. LXXVIII: Cell information for bay TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

bay TAPP-H_4F_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic†
AE / per unit cell	83,966	95,327	10,6022	-
AE / per molecule	20,991	23,832	26,506	-
AE / per atom	583	567	552	-
Geom.	triclinic	triclinic	triclinic	triclinic
AE / per unit cell	41,891	47,580	52,930	58,258
AE / per molecule	20,946	23,790	26,465	29,129
AE / per atom	582	566	551	539

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXXIX: Cell information for bay TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

bay TAPP-H_4Cl_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic*	monoclinic
AE / per unit cell	81,840	93,256	10,3986	11,4505
AE / per molecule	20,460	23,314	25,996	28,626
AE / per atom	568	555	542	530
Geom.	triclinic†	triclinic	triclinic	triclinic*
AE / per unit cell	-	46,566	51,911	57,255
AE / per molecule	-	23,283	25,956	28,628
AE / per atom	-	554	541	530

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXXX: Cell information for bay TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

bay TAPP-H_4Br_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic†
AE / per unit cell	80,882	92,269	10,3003	-
AE / per molecule	20,221	23,067	25,751	-
AE / per atom	562	549	536	-
Geom.	triclinic	triclinic	triclinic	triclinic
AE / per unit cell	40,478	46,097	51,441	56,777
AE / per molecule	20,239	23,049	25,720	28,389
AE / per atom	562	549	536	526

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXXXI: Cell information for bay TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

bay TAPP-H_4I_4	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic†
AE / per unit cell	80,009	91,344	10,2078	-
AE / per molecule	20,002	22,836	25,519	-
AE / per atom	556	544	532	-
Geom.	triclinic†	triclinic	triclinic	triclinic
AE / per unit cell	-	45,663	50,992	56,327
AE / per molecule	-	22,832	25,496	28,164
AE / per atom	-	544	531	522

* Experimentally-obtained crystal system.

† Cell optimization not converged.

c. TAPP compounds with all substitution

Tab. LXXXII: Cell information for all TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

all TAPP-F_8	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
AE / per unit cell	84,013	95,426	10,6104	11,6766
AE / per molecule	21,003	23,857	26,526	29,191
AE / per atom	583	568	553	541
Geom.	triclinic†	triclinic	triclinic†	triclinic
AE / per unit cell	-	47,693	-	58,369
AE / per molecule	-	23,846	-	29,184
AE / per atom	-	568	-	540

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXXXIII: Cell information for all TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

all TAPP-Cl_8	H-†	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic*
AE / per unit cell	-	91,716	10,2450	11,3137
AE / per molecule	-	22,929	25,612	28,284
AE / per atom	-	546	534	524
Geom.	triclinic	triclinic	triclinic	triclinic
AE / per unit cell	-	45,839	51,142	56,483
AE / per molecule	-	22,920	25,571	28,241
AE / per atom	-	546	533	523

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXXXIV: Cell information for all TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

all TAPP-Br_8	H-†	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
AE / per unit cell	-	89,786	10,0517	11,1220
AE / per molecule	-	22,446	25,129	27,805
AE / per atom	-	534	524	515
Geom.	triclinic	triclinic	triclinic	triclinic
AE / per unit cell	-	44,900	50,190	55,521
AE / per molecule	-	22,450	25,095	27,761
AE / per atom	-	535	523	514

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXXXV: Cell information for all TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

all TAPP-I_8	H-	C_1F_3-	C_2F_5-	C_3F_7-
Geom.	monoclinic	monoclinic	monoclinic	monoclinic
AE / per unit cell	76,872	88,076	98,759	10,9424
AE / per molecule	19,218	22,019	24,690	27,356
AE / per atom	534	524	514	507
Geom.	triclinic†	triclinic	triclinic†	triclinic
AE / per unit cell	-	44,051	-	54,683
AE / per molecule	-	22,026	-	27,342
AE / per atom	-	524	-	506

* Experimentally-obtained crystal system.

† Cell optimization not converged.

Tab. LXXXVI: Cell information for all TAPP. Angles are given in degrees, distances in Å, energies in kJ/mol.

all C_3F_7-TAPP	-Br_4Cl_4	-Cl_4Br_4
Geom.	monoclinic	monoclinic*
AE / per unit cell	11,2137	11,2214
AE / per molecule	28,034	28,054
AE / per atom	519	520
Geom.	triclinic	triclinic
AE / per unit cell	55,998	56,004
AE / per molecule	27,999	28,002
AE / per atom	519	519

* Experimentally-obtained crystal system.

[1] J. M. Shaw and P. F. Seidler, IBM J. Res. Dev. **45**, 3–9 (2001).

[2] J. L. Brédas, J. P. Calbert, D. A. da Silva Filho, and J. Cornil, PNAS **99**, 5804 (2002).

[3] Y. Zhang, L. Guo, X. Zhu, and X. Sun, Front. Chem. **8** (2020), 10.3389/fchem.2020.589207.

[4] S. Ahmad, J. Polym. Eng. **34**, 279 (2014).

[5] H. F. Haneef, A. M. Zeidell, and O. D. Jurchescu, J. Mater. Chem. C **8**, 759 (2020).

[6] J. D. Myers and J. Xue, Polym. Rev. **52**, 1 (2012).

[7] V. Coropceanu, J. Cornil, D. A. da Silva Filho, Y. Olivier, R. Silbey, and J.-L. Brédas, Chem. Rev. **107**, 926 (2007).

[8] C. Groves, Rep. Prog. Phys **80**, 026502 (2016).

[9] B. Lüssem, M. Riede, and K. Leo, Phys. Status Solidi A **210**, 9 (2013).

[10] H. Akamatu, H. Inokuchi, and Y. Matsunaga, Bull. Chem. Soc. Jpn **29**, 213 (1956).

[11] H. Akamatu and H. Inokuchi, J. Chem. Phys. **18**, 810 (1950).

[12] C. K. Chiang, C. R. Fincher, Y. W. Park, A. J. Heeger, H. Shirakawa, E. J. Louis, S. C. Gau, and A. G. MacDiarmid, Phys. Rev. Lett. **39**, 1098 (1977).

[13] S. E. Shaheen, D. S. Ginley, and G. E. Jabbour, MRS Bull. **30**, 10–19 (2005).

[14] G. Inzelt, M. Pineri, J. Schultze, and M. Vorotyntsev, Electrochim. Acta **45**, 2403 (2000).

[15] C. Wang, H. Dong, L. Jiang, and W. Hu, Chem. Soc. Rev. **47**, 422 (2018).

[16] J. T. E. Quinn, J. Zhu, X. Li, J. Wang, and Y. Li, J. Mater. Chem. C **5**, 8654 (2017).

[17] X. Zhang, H. Dong, and W. Hu, Adv. Mater. **30**, 1801048 (2018).

[18] S. Otep, T. Michinobu, and Q. Zhang, Sol. RRL **4**, 1900395 (2020).

[19] A. A. Virkar, S. Mannsfeld, Z. Bao, and N. Stingelin, Adv. Mater. **22**, 3857 (2010).

[20] C. R. Newman, C. D. Frisbie, D. A. da Silva Filho, J.-L. Brédas, P. C. Ewbank, and K. R. Mann, Chem. Mater. **16**, 4436 (2004).

[21] O. Ostroverkhova, Chem. Rev. **116**, 13279 (2016).

[22] L. Bai, P. Wang, P. Bose, P. Li, R. Zou, and Y. Zhao, ACS Appl. Mater. Interfaces **7**, 5056 (2015).

[23] S. A. Ayoub and J. B. Lagowski, Mater. Des. **156**, 558 (2018).

[24] R. S. Andre, R. C. Sanfelice, A. Pavinatto, L. H. Mattoso, and D. S. Correa, Mater. Des. **156**, 154 (2018).

[25] M. Hamilton, S. Martin, and J. Kanicki, IEEE Trans Electron Devices **51**, 877 (2004).

[26] A. G. MacDiarmid, Angew. Chem. Int. Ed. **40**, 2581 (2001).

[27] A. Mishra and P. Bäuerle, Angew. Chem. Int. Ed. **51**, 2020 (2012).

[28] K. M. Coakley and M. D. McGehee, Chem. Mater **16**, 4533 (2004).

[29] C. Yuan, Y. Zhou, Y. Zhu, J. Liang, S. Wang, S. Peng, Y. Li, S. Cheng, M. Yang, J. Hu, B. Zhang, R. Zeng, J. He, and Q. Li, Nat. Commun. **11**, 3919 (2020).

[30] G. Li, V. Shrotriya, J. Huang, Y. Yao, T. Moriarty, K. Emery, and Y. Yang, Nat. Mater. **4**, 864 (2005).

[31] T. Lei, J.-Y. Wang, and J. Pei, Chem. Mater. **26**, 594 (2014).

[32] J. E. Anthony, Chem. Rev. **106**, 5028 (2006).

[33] A. Facchetti, Chem. Mater. **23**, 733 (2011).

[34] W. Yuan, J. Cheng, X. Li, M. Wu, Y. Han, C. Yan, G. Zou, K. Müllen, and Y. Chen, Angew. Chem. Int. Ed. **59**, 9940 (2020).

[35] B. A. R. Günther, S. Höfener, R. Eichelmann, U. Zschieschang, H. Wadepohl, H. Klauk, and L. H. Gade, Org. Lett. **22**, 2298 (2020).

[36] L. Hahn, A. Hermannsdorfer, B. Günther, T. Wesp, B. Bühler, U. Zschieschang, H. Wadepohl, H. Klauk, and L. H. Gade, J. Org. Chem. **82**, 12492 (2017).

[37] T. Riehm, G. De Paoli, A. Konradsson, L. De Cola, H. Wadepohl, and L. Gade, Chem. Eur. J. **13**, 7317 (2007).

[38] L. Hahn, F. Maaß, T. Bleith, U. Zschieschang, H. Wadepohl, H. Klauk, P. Tegeder, and L. H. Gade, Chem. Eur. J. **21**, 17691 (2015).

[39] B. A. R. Günther, S. Höfener, R. Eichelmann, U. Zschieschang, H. Wadepohl, H. Klauk, and L. H. Gade, Org. Lett. **22**, 2298 (2020), pMID: 32118454.

[40] N. Wollscheid, B. Günther, V. J. Rao, F. J. Berger, J. L. P. Lustres, M. Motzkus, J. Zaumseil, L. H. Gade, S. Höfener, and T. Buckup, J. Phys. Chem. A **124**, 7857 (2020), pMID: 32962348.

[41] B. A. R. Günther, S. Höfener, U. Zschieschang, H. Wadepohl, H. Klauk, and L. H. Gade, Chem. Eur. J. **25**, 14669 (2019).

[42] S. Höfener, B. A. R. Günther, M. E. Harding, and L. H. Gade, J. Phys. Chem. A **123**, 3160 (2019).

[43] K. M. Pelzer, Á. Vázquez-Mayagoitia, L. E. Ratcliff, S. Tretiak, R. A. Bair, S. K. Gray, T. Van Voorhis, R. E. Larsen, and S. B. Darling, Chem. Sci. **8**, 2597 (2017).

[44] J. Gu, J. Phys. Conf. Ser. **2194**, 012024 (2022).

[45] S. Giannini and J. Blumberger, Acc. Chem. Res. **55**, 819 (2022).

[46] T. Stein, J. Autschbach, N. Govind, L. Kronik, and R. Baer, J. Phys. Chem **3**, 3740 (2012).

[47] S. Kümmel and L. Kronik, Rev. Mod. Phys. **80**, 3 (2008).

[48] J. P. Perdew and M. Levy, Phys. Rev. B **56**, 16021 (1997).

[49] M. Levy, J. P. Perdew, and V. Sahni, Phys. Rev. A **30**, 2745 (1984).

[50] D. P. Chong, O. V. Gritsenko, and E. J. Baerends, J. Chem. Phys. **116**, 1760 (2002).

[51] S. Goedecker and L. Colombo, Phys. Rev. Lett. **73**, 122 (1994).

[52] D. R. Bowler and T. Miyazaki, J. Condens. Matter Phys. **22**, 074207 (2010).

[53] S. Mohr, L. E. Ratcliff, L. Genovese, D. Caliste, P. Boulanger, S. Goedecker, and T. Deutsch, Phys. Chem. Chem. Phys. **17**, 31360 (2015).

[54] J. P. Perdew, Int. J. Quantum Chem. **28**, 497 (1985).

[55] J. P. Perdew, W. Yang, K. Burke, Z. Yang, E. K. U. Gross, M. Scheffler, G. E. Scuseria, T. M. Henderson, I. Y. Zhang, A. Ruzsinszky, H. Peng, J. Sun, E. Trushin, and A. Görling, PNAS **114**, 2801 (2017).

[56] Á. Morales-García, R. Valero, and F. Illas, J. Phys. Chem **121**, 18862 (2017).

[57] P. Borlido, J. Schmidt, A. W. Huran, F. Tran, M. A. L. Marques, and S. Botti, Npj Comput. Mater. **6**, 96 (2020).

[58] L. Reining, Wiley Interdiscip. Rev. Comput. Mol. Sci. **8**, e1344 (2018).

[59] H. Lambert and F. Giustino, Phys. Rev. B **88**, 075117 (2013).

[60] T. Kotani, M. van Schilfgaarde, and S. V. Faleev, Phys. Rev. B **76**, 165106 (2007).

[61] M. J. van Setten, F. Weigend, and F. Evers, J. Chem. Theory Comput. **9**, 232 (2013).

[62] F. Bruneval and X. Gonze, Phys. Rev. B **78**, 085125 (2008).

[63] H.-V. Nguyen, T. A. Pham, D. Rocca, and G. Galli, Phys. Rev. B **85**, 081101 (2012).

[64] T. Körzdörfer and N. Marom, Phys. Rev. B **86**, 041110 (2012).

[65] V. Atalla, M. Yoon, F. Caruso, P. Rinke, and M. Scheffler, Phys. Rev. B **88**, 165122 (2013).

[66] R. J. Harrison, J. Chem. Phys. **94**, 5021 (1991).

[67] A. Dreuw and M. Wormit, Wiley Interdiscip. Rev. Comput. Mol. Sci. **5**, 82 (2015).

[68] J. Schirmer, Phys. Rev. A **26**, 2395 (1982).

[69] J. F. Stanton and R. J. Bartlett, J. Chem. Phys. **98**, 7029 (1993).

[70] J. Geertsen, M. Rittby, and R. J. Bartlett, Chem. Phys. Lett. **164**, 57 (1989).

[71] O. Christiansen, P. Jørgensen, and C. Hättig, Int. J. Quantum Chem. **68**, 1 (1998).

[72] J. H. Starcke, M. Wormit, and A. Dreuw, J. Chem. Phys. **130**, 024104 (2009).

[73] A. Pershin and P. G. Szalay, J. Chem. Theory Comput. **11**, 5705 (2015).

[74] M. Head-Gordon, M. Oumi, and D. Maurice, Mol. Phys. **96**, 593 (1999).

[75] N. O. C. Winter, N. K. Graf, S. Leutwyler, and C. Hättig, Phys. Chem. Chem. Phys. **15**, 6623 (2013).

[76] C. Hättig, in *Response Theory and Molecular Properties (A Tribute to Jan Linderberg and Poul Jørgensen)*, Advances in Quantum Chemistry, Vol. 50, edited by H. Jensen (Academic Press, 2005) pp. 37–60.

[77] R. Send, M. Kühn, and F. Furche, J. Chem. Theory Comput. **7**, 2376 (2011).

[78] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).

[79] O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).

[80] A. Shee, T. Saue, L. Visscher, and A. Severo Pereira Gomes, J. Chem. Phys. **149**, 174113 (2018).

[81] S. Hirata, M. Nooijen, and R. J. Bartlett, Chem. Phys. Lett. **328**, 459 (2000).

[82] M. Nooijen and R. J. Bartlett, J. Chem. Phys. **102**, 3629 (1995).

[83] M. Musiał, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. **118**, 1128 (2003).

[84] J. F. Stanton and J. Gauss, J. Chem. Phys. **101**, 8938 (1994).

[85] T. A. Wesolowski and A. Warshel, J. Phys. Chem. **97**, 8050 (1993).

[86] A. Severo Pereira Gomes and C. R. Jacob, Annu. Rep. Prog. Chem., Sect. C: Phys. Chem. **108**, 222 (2012).

[87] C. R. Jacob and J. Neugebauer, Wiley Interdiscip. Rev. Comput. Mol. Sci. **4**, 325 (2014).

[88] T. A. Wesolowski, S. Shedge, and X. Zhou, Chem. Rev. **115**, 5891 (2015).

[89] T. A. Barnes, J. W. Kaminski, O. Borodin, and T. F. Miller, J. Phys. Chem **119**, 3865 (2015).

[90] S. J. Bennie, B. F. E. Curchod, F. R. Manby, and D. R. Glowacki, J. Phys. Chem **8**, 5559 (2017).

[91] X. Wen, D. S. Graham, D. V. Chulhai, and J. D. Goodpaster, J. Chem. Theory Comput. **16**, 385 (2020).

[92] L. Shi and A. P. Willard, Chem. Phys. **149**, 094110 (2018).

[93] M. Pinterić, S. Roh, S. Hammer, J. Pflaum, M. Dressel, and E. Uykur, J. Mater. Chem. C

10, 5582 (2022).

[94] H. J. Wörner, C. A. Arrell, N. Banerji, A. Cannizzo, M. Chergui, A. K. Das, P. Hamm, U. Keller, P. M. Kraus, E. Liberatore, P. Lopez-Tarifa, M. Lucchini, M. Meuwly, C. Milne, J.-E. Moser, U. Rothlisberger, G. Smolentsev, J. Teuscher, J. A. van Bokhoven, and O. Wenger, Struct. Dyn. **4**, 061508 (2017).

[95] J.-L. Brédas, D. Beljonne, V. Coropceanu, and J. Cornil, Chem. Rev. **104**, 4971 (2004).

[96] J. Schiedt, W. J. Knott, K. Le Barbu, E. W. Schlag, and R. Weinkauf, J. Chem. Phys. **113**, 9470 (2000).

[97] N. Ando, M. Mitsui, and A. Nakajima, J. Chem. Phys. **128**, 154318 (2008).

[98] J. F. Stanton and J. Gauss, J. Chem. Phys. **111**, 8785 (1999).

[99] A. Köhn and C. Hättig, J. Chem. Phys. **119**, 5021 (2003).

[100] C. Hättig and A. Köhn, J. Chem. Phys. **117**, 6939 (2002).

[101] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[102] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[103] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[104] S. Goedecker, M. Teter, and J. Hutter, Phys. Rev. B **54**, 1703 (1996).

[105] C. Hartwigsen, S. Goedecker, and J. Hutter, Phys. Rev. B **58**, 3641 (1998).

[106] M. Krack, Theor. Chem. Acc. **114**, 145 (2005).

[107] M. Guidon, J. Hutter, and J. VandeVondele, J. Chem. Theory Comput. **6**, 2348 (2010).

[108] S. Höfener, Int. J. Quantum Chem. **121**, e26351 (2021).

[109] S. Höfener, J. Comput. Chem. **35**, 1716 (2014).

[110] O. Vahtras, J. Almlöf, and M. Feyereisen, Chem. Phys. Lett. **213**, 514 (1993).

[111] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs, Chem. Phys. **242**, 652 (1995).

[112] C. Hättig, J. Chem. Phys. **118**, 7751 (2003).

[113] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[114] A. Lembarki and H. Chermette, Phys. Rev. A **50**, 5328 (1994).

[115] F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, Chem. Phys. Lett. **294**, 143 (1998).

[116] F. Weigend and R. Ahlrichs, Phys. Chem. Chem. Phys. **7**, 3297 (2005).

[117] F. Weigend, Phys. Chem. Chem. Phys. **8**, 1057 (2006).

[118] D. Rappoport and F. Furche, J. Chem. Phys. **133**, 134105 (2010).

[119] T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino,

R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A. S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack, and J. Hutter, J. Chem. Phys. **152**, 194103 (2020).

[120] J. VandeVondele and J. Hutter, J. Chem. Phys. **127**, 114105 (2007).

[121] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).

[122] G. J. O. Beran, Chem. Rev. **116**, 5567 (2016).

[123] "TURBOMOLE V7.2 2017, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from `http://www.turbomole.com`.".

[124] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[125] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[126] S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).

[127] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, J. Chem. Phys. **98**, 11623 (1994).

[128] C. R. Groom, I. J. Bruno, M. P. Lightfoot, and S. C. Ward, Acta Cryst. B **72**, 171 (2016).

[129] M. Matena, M. Stöhr, T. Riehm, J. Björk, S. Martens, M. Dyer, M. Persson, J. Lobo-Checa, K. Müller, M. Enache, H. Wadepohl, J. Zegenhagen, T. Jung, and L. Gade, Chem. Eur. J. **16**, 2079 (2010).

[130] S. C. Martens, U. Zschieschang, H. Wadepohl, H. Klauk, and L. H. Gade, Chem. Eur. J. **18**, 3498 (2012).

[131] S. Geib, U. Zschieschang, M. Gsänger, M. Stolte, F. Würthner, H. Wadepohl, H. Klauk, and L. H. Gade, Adv. Funct. Mater. **23**, 3866 (2013).

[132] B. A. R. Günther, S. Höfener, U. Zschieschang, H. Wadepohl, H. Klauk, and L. H. Gade, Chem. Eur. J. **25**, 14669 (2019).

ACKNOWLEDGEMENT

I would like to thank Dr. Sebastian Höfener first for all the help both in academic and in life general. The discussion concerning the frozen density embedding method in particular and theoretical chemistry in general are both inspiring, and the advises concerning how to properly conduct scientific work would beneficial me beyond this project.

I would also like to thank two of my former office-mate Nils Schieschke and Rodrigo Cortes. Thank you both for all the small and big talks, jokes and complains. Thank you Nils for the help on familiarize the working environments, Fortran coding and daily life in general. Thank you Rodrigo especially for the help before I submit my thesis, and all the moral lifting talk about human and the world in general, and computational chemistry in specific.

I would also like to thank everyone in our working group for creating the friendly working environment and all the life and academic help. Thank Prof. Dr. Klopper who gives me this opportunity to work on this project and always make sure everyone only need to focused on the academic work. Thank Mrs. Kühn for always takes care of all the administration matters. Thank you Ansgar Pausch, Angela Bihlmeier, Florian Rudolf Rehak, Max Karl Kehry and Michael Harding for all the discussion, technical help, talking, joking and laughing in these last few years.

Last but not least I would like to thank my family for their continues support during the time of this project and my better half Qi Yu for her understanding and moral support especially in the last two years.