
Mesta, Murat; Chang, Jinhyun; Shil, Suranjan; Thygesen, Kristian Sommer; García Lastra, Juan Maria

Published in:

Link to article, DOI:
10.1021/acs.jpca.9b02391

Publication date:
2019

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Murat Mesta,*† Jin Hyun Chang,† Suranjan Shil,† Kristian S. Thygesen,‡ and Juan Maria Garcia Lastra*,†

†Department of Energy Conversion and Storage, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
‡Center for Atomic-scale Materials Design, Department of Physics, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

E-mail: murme@dtu.dk; jmgla@dtu.dk
Abstract

The ability of donor–acceptor (D-A) type polymers to control the positions of highest occupied (HOMO) and lowest unoccupied (LUMO) molecular orbitals makes them a popular choice for organic solar cell applications. The alternating D-A pattern in a monomer leads to a weak electronic coupling between the constituent monomers within the polymer chain. Exploiting the weak electronic coupling characteristics, we developed a method to efficiently calculate (1) the electronic properties and (2) the optical gap of such polymer chains. The electronic properties (HOMO and LUMO energies, ionization potential, electron affinity and quasiparticle gap of an oligomer of any length up to an infinitely long polymer) of the D-A polymers are predicted by combining density functional theory calculation results and a tight-binding model. The weak electronic coupling implies that the optical gap of the polymer is size independent and thus, it can be calculated using a monomer. We validated the methods using a set of 104 polymers by checking the consistency where the electronic gap of a polymer is larger than the optical gap. Furthermore, we establish relationships between the results obtained from more accurate, yet slower methods (i.e., B3LYP functional, singlet-∆SCF) with the ones obtained from the faster counterparts (i.e., BLYP functional, triplet-∆SCF). Leveraging on the found relationships, we propose a way in which the electronic and optical properties of the polymers can be calculated efficiently while retaining high accuracy. The use of the tight-binding model combined with the approach to estimate more accurate results based on less expensive simulations is crucial in the applications where a large volume of computations needs to be carried out efficiently with sufficiently high accuracy, such as high-throughput computational screening or training a machine-learning model.

Introduction

Organic solar cells (OSC)\textsuperscript{1–3} present numerous advantages over inorganic photovoltaic technologies. Their low weight, mechanical flexibility, low cost and ability to work well under a
wide range of lighting conditions and orientations are some of the key advantages that make OSCs an attractive choice. Despite numerous advantages, their relatively low power conversion efficiency prevents them from having a dominant position in the current photovoltaic market. It was predicted a decade ago that, in principle, the power conversion efficiency of OSCs could reach up to 19%.\(^4,5\) A steady progress has been made towards the 19% efficiency limit,\(^6-15\) and the 13% and 17% barriers are recently surpassed using single-junction\(^16\) and tandem\(^17\) OSCs, respectively. However, current OSCs produced in a large scale still have a power conversion efficiency of less than 10%.\(^18\)

A further improvement of the power conversion efficiency requires optimization of several aspects of the OSC components. Two parameters that are directly correlated to achieving the maximal performance of OSCs are the edge of the optical spectra of the light absorbers and the energy level alignment between the donor and acceptor units of the OSCs. The use of conjugated donor–acceptor (D-A) polymers as both light absorbers and electron donors in OSCs is particularly attractive for tuning the two parameters. The monomer unit of a D-A polymer consists of one donor and one acceptor moiety. The highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) of the monomer are located at the donor and acceptor moiety, respectively. Thus, the energetic positions of the HOMO and LUMO levels, which eventually determines the edges of the optical spectrum, can be controlled independently of each other by choosing the appropriate donor and acceptor moieties.\(^19\)

Given the vast molecular space, estimated to exceed \(10^{60}\) molecules,\(^20\) a search of the ideal donor and acceptor moieties cannot be done using Edisonian trial-and-error experimental strategies. It is, therefore, necessary to use computational screening techniques to guide future experimental studies in order to save time and resources substantially. Many efforts have been made in the computational screening domain in recent years. The Harvard Clean Energy Project (CEP)\(^21\) compiled a library with properties of 2.6 million molecules as potential components for OSCs. Using different levels of density functional theory (DFT),\(^22\)
the CEP computed the HOMO and LUMO Kohn-Sham (KS) energies of each molecule and computed the quasiparticle gap as the energy difference between the LUMO and HOMO (i.e., $\Delta E_{\text{KS}}^{\text{QP}} = E_{\text{LUMO}} - E_{\text{HOMO}}$).

Other efforts were made to expedite the screening process by using approaches with a lower computational cost. Hutchinson et al. used semi-empirical methods, such as AM1,23 PM6,24 and ZINDO/S25 to obtain the HOMO and LUMO energy levels and the lowest-energy optical excitation.26 More recently, Zwijnenburg et al. proposed a density-functional tight-binding (DFTB) approach27 calibrated with respect to time-dependent DFT (TDDFT)28 calculations to obtain the ionization potential (IP), electron affinity (EA) and the lowest optical excitation energy of donor and acceptor moieties in polymers.29

In contrast to the approaches taken by CEP and Hutchinson et al. where the quasiparticle gap is obtained using the KS HOMO and LUMO energies, Zwijnenburg at al. calculated the IP and EA using the $\Delta\text{SCF}$ method (i.e., calculating the energy difference between the cationic/anionic state and the neutral state of the molecule at the relaxed geometry of the latter). The quasiparticle gap is then calculated using $\Delta E_{\Delta\text{SCF}}^{\text{QP}} = \text{IP} - \text{EA}$. The preliminary results from Zwijnenburg et al.29 show high speed and reliability of their method, which makes it suitable for a large scale screening study of polymers.

The present work is devoted to developing a new method to further expedite the ongoing search for the optimal conjugated D-A polymers for OSC applications. Through a systematic investigation of a set of 104 conjugated polymers (combining 8 donors and 13 acceptor units), we demonstrate that a tight-binding (TB) model can predict the $E_{\text{HOMO}}$, $E_{\text{LUMO}}$, IP and EA of an oligomer of an arbitrary length — including an infinitely long polymer — only using the DFT calculation results for a monomer (i.e., a unit consisting of one acceptor moiety and one donor moiety) and a dimer. In other words, the quasiparticle gap of an infinitely long polymer can be determined by extrapolating the DFT calculation results of monomer and dimer using a TB model. Here, it is assumed that the structure of the polymer is rigid, i.e., the long polymer is represented as a periodic sequence of monomers.
Following the previous works,\textsuperscript{21,26,29} the optical excitation is assumed to occur locally, involving only a monomer. With the assumption, we also calculate the lowest-energy optical excitation (optical gap) through the $\Delta$SCF method. We evaluate the validity of the quasiparticle gap of polymers calculated using the TB model by comparing them with the optical gap; the quasiparticle gap should always be larger than the optical gap. We show that only the quasiparticle gap obtained using the IP and EA values from the $\Delta$SCF calculations fulfills the condition. By contrast, the $\Delta E_{\text{QP}}^{\text{KS}}$ of polymers give a reasonable approximation to the optical gap calculated through a $\Delta$SCF due to a cancellation of errors, which aligns very well with a previous report.\textsuperscript{30}

**Methods**

**Polymer structures**

All of the molecular structures considered in this work are constructed by combining 8 donor and 13 acceptor moieties illustrated in Figure 1. A monomer consists of one donor-acceptor pair, and 104 types of monomers can be generated using the considered moieties. The moieties shown in Figure 1 are simplified versions of the moieties synthesized by Bundgaard et al.,\textsuperscript{31} which included long alkane side chains to control the polymer-polymer interactions. Neglecting the side chains is well justified here since the focus is on how the sequential structure of a single polymer affects the molecular-scale photoexcitation and intramolecular charge transfer, and the side chains do not influence the frontier orbitals of the polymer. However, we note that it is likely that the side chains will affect the polymer–polymer distance and the molecular orientations of polymers, which in turn could affect the electronic and optical properties in an indirect way. Despite the limitations due to the simplifications, we have made a comparison of the electronic and optical properties of the oligomers, which is found in supporting information.

A polymer chain constructed using one of the 104 monomer types — A1D1, a pair
Figure 1: Acceptor and donor moieties used for constructing the D-A polymers. These are simplified versions of the ones studied experimentally in ref. 31.
consisting of A1 acceptor and D2 donor — is shown in Figure 2a. Repetition of any monomer type results in an alternating pattern of a donor and acceptor moiety. It is pointed out that the optoelectronic properties of conjugated polymers are primarily governed by the frontier molecular orbitals, which may delocalize over single or several monomer units depending on the orbital overlap between the repeating monomers. The orbital overlap is strongly influenced by the distance between two conjugated segments. The alternating pattern of the donor and acceptor units as shown in Figure 2a introduces a spatial separation between the neighboring donors and acceptors in the chain, which assists in localizing the HOMO of a donor unit and the LUMO of an acceptor unit. The anticipated weak monomer-monomer overlap makes it plausible that a TB model including only nearest neighbour hopping could provide an accurate description of the electronic structure of long D-A polymers.

Figure 2: (a) Structure of an example oligomer (A1D1) with a donor (shaded) and acceptor units. The off-diagonal terms of the model Hamiltonian matrix (eq. 1) arising from overlapping HOMO and LUMO units are denoted as $\beta_H$ and $\beta_L$, respectively. (b) Schematic plot of the HOMO and LUMO energies as a function of a number of the repeating donor-acceptor units. The HOMO (LUMO) level shifts by $\beta_H$ ($\beta_L$) when a monomer is extended to a dimer, and the extent of the energy level shift converges to $2\beta_H$ ($2\beta_L$) for an infinite chain.
Tight-binding model

Direct analysis of the electronic structure of long polymer chains using DFT simulations is unfeasible for a screening study due to the prohibitively large number of atoms in the simulation cell. Therefore, it is crucial to develop a model that can reliably predict the properties of long polymer chains by extrapolating the properties of a monomer and dimer. Here, we propose that the electronic structure of a conjugated D-A polymer can be predicted using a simple TB model, which takes the energies derived from the DFT calculations of monomer and dimer as inputs.

In a TB model, the secular determinant of the TB Hamiltonian of the model system with \( n \) repeating units is given by

\[
\begin{vmatrix}
\langle \phi^1_i | H | \phi^1_i \rangle - \lambda_i & \beta_i & 0 & \cdots & 0 & 0 \\
\beta_i & \langle \phi^2_i | H | \phi^2_i \rangle - \lambda_i & \beta_i & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \beta_i & \langle \phi^n_i | H | \phi^n_i \rangle - \lambda_i
\end{vmatrix} = 0, \quad (1)
\]

where a superscript corresponds to the site number within the chain and a subscript \( i \) is a binary parameter that can be either H or L, which respectively correspond to HOMO and LUMO. \( \beta \) is the intermolecular hopping parameter and \( \lambda \) is the eigenvalue (HOMO or LUMO energy). We assume that the on-site energies are the same for all monomers in the oligomer, which means \( \langle \phi^1_i | H | \phi^1_i \rangle = \langle \phi^2_i | H | \phi^2_i \rangle = \cdots = \langle \phi^n_i | H | \phi^n_i \rangle \).

The hopping parameter, \( \beta \), accounts for the orbital energy shift due to the hybridization with the HOMO/LUMO orbitals on neighboring moieties, and the sum of \( \beta_H \) and \( \beta_L \) is related to the reduction of the energy gap from that of a monomer. There is no direct way of determining \( \beta \) from DFT calculations and thus, we identify \( \beta \) as the orbital energy shift when a monomer is extended to a dimer. A solution of the determinant in eq. 1 for a dimer \( n = 2 \) yields \( E_H^{n=2} = E_H^{n=1} + \beta_H \) and \( E_L^{n=2} = E_L^{n=1} - \beta_L \). Therefore, the exact value of \( \beta \) can be determined through the DFT simulations of a monomer and a dimer by calculating
Once the value of $\beta$ is known, the TB corrections of larger oligomers can be obtained from the solution of their corresponding secular equations. In general, the HOMO and LUMO energies of oligomers of the length $n = N$ are

$$E_{H}^{n=N} = E_{H}^{n=1} + 2\beta_{H} \cos\left(\frac{\pi}{N+1}\right)$$

(2)

and

$$E_{L}^{n=N} = E_{L}^{n=1} + 2\beta_{L} \cos\left(\frac{N\pi}{N+1}\right),$$

(3)

respectively. In the limit where $n = \infty$, $E_{H}^{n=\infty} = E_{H}^{n=1} + 2\beta_{H}$ and $E_{L}^{n=\infty} = E_{L}^{n=1} - 2\beta_{L}$.

We stress that we use the ionization energies (IP or EA) in the diagonal terms of eq. 1 when extrapolating the ionization energy values obtained from $\Delta$SCF calculations. The value of $\beta$ is then assigned to the change in the ionization energies when a monomer is extended to a dimer.

**Density functional theory calculations**

The ground-state geometries of all 104 types of monomers and their dimers are obtained by running geometry optimization calculations using Gaussian 09 software package. The Becke three-parameter Lee-Yang-Parr (B3LYP) hybrid functional and a TZVP basis set are used. This choice allows one to extrapolate the energies at the hybrid level for oligomers of any size. The validity of the TB model is assessed using the DFT calculations of oligomers consisting up to four repeating oligomer units (i.e., monomer, dimer, trimer and tetramer). However, the computational cost of the hybrid functional DFT simulations is prohibitively high to be performed on all of the 104 trimers and tetramers. Another set of simulations are performed at generalized gradient approximation (GGA) level using the BLYP functional for all of the monomers, dimers, trimers and tetratomers in order to bypass the constraints imposed by the high computational cost of B3LYP-level calculations. The validity of the TB model is assessed using the BLYP functional DFT calculations, and the accuracy of the
$\beta$ values obtained from the calculations using BLYP functional is evaluated by comparing them with the ones obtained from the calculations using B3LYP functional.

The quasiparticle gap is calculated using two different methods. The most straightforward method to treat the quasiparticle gap as the difference between the KS LUMO and HOMO energies such that $\Delta E_{\text{QP}}^{\text{KS}} = E_{\text{LUMO}} - E_{\text{HOMO}}$, and we refer to it as a Kohn-Sham quasiparticle gap. The second method is to calculate the gap using the IP and EA obtained via $\Delta\text{SCF}$ calculations.\textsuperscript{38,39} Specifically, IP and EA are determined by calculating the energy difference between the ionized and neutral states of the molecule at the geometry of the neutral state. That is, $\text{IP} = E(M^+) - E(M)$ and $\text{EA} = -[E(M^-) - E(M)]$, where $E(M)$, $E(M^-)$ and $E(M^+)$ are the total energies of neutral, anionic and cationic states of the system with the geometry of the neutral state, respectively. The quasiparticle gap is then calculated by subtracting EA from IP (i.e., $\Delta E_{\Delta\text{SCF}}^{\text{QP}} = \text{IP} - \text{EA}$), and we refer to this gap as a quasiparticle $\Delta\text{SCF}$ gap.

The lowest-energy optical excitation (optical gap) is another property of interest in D-A polymers for OSC applications. The optical gaps, $\Delta E_{\text{opt}}$, of monomers are calculated using the $\Delta\text{SCF}$ method by accounting only for the vertical excitation (i.e., the excited state electronic configuration is calculated using the ground-state geometry). The optical gap is calculated by promoting an electron from the HOMO to the LUMO and ensuring the SCF convergence with the ground-state geometry. Such an approach explicitly accounts for the effect of the electronic-hole interaction in the excited state. We assume that the optical excitation occurs locally within a monomer unit due to the strong electron-hole interaction (i.e., the Coulomb interaction between the exciton and the nearest-neighbor ground-state charge does not influence the exciton binding).

The vertical excitation can be calculated using the $\Delta\text{SCF}$ method by forcing the electronic configuration to have either a singlet or a triplet multiplicity. Both the singlet- and triplet-$\Delta\text{SCF}$ calculations are carried out using the ground-state geometries of the molecules determined using the B3LYP hybrid functional provided in the Gaussian 09 software pack-
The singlet-∆SCF calculations are carried out using Amsterdam Density Functional (ADF) software package as it supports the assignment of specific orbital occupations. The triplet exciton energies are computed using both ADF and Gaussian 09 in order to ensure that they yield the same results and that the results obtained using one code are compatible with the results using another. All of the singlet- and triplet-∆SCF calculations used the B3LYP functional. The energies obtained using the two software packages differ only by 10 meV on average with the maximum difference of 28 meV. The results obtained from the two packages are virtually the same, which allows one to gain insight on singlet-triplet splitting. In the case where the energies calculated using singlet- and triplet-∆SCF calculations deviate mostly by a fixed value (i.e., constant offset) as seen in another system, one can predict the singlet exciton energy based on the triplet-∆SCF calculation results.

Results and Discussions

Impact of the exchange-correlation functional on the orbital energy shift

The orbital energy shift when a monomer is extended to a dimer, \( \beta \), is calculated using both BLYP and B3LYP functionals. The \( \beta \) values obtained from the two functionals show a strong linear correlation with a Pearson correlation coefficient, \( R \), of 0.97 as shown in Figure 3. The \( \beta \) values calculated using the B3LYP functional (\( \beta^{B3LYP} \)) are around 1.2 times larger than those calculated using the BLYP functional (\( \beta^{BLYP} \)). The smaller \( \beta^{BLYP} \) values are expected to be originating from the self-interaction error (SIE) in the BLYP calculations: the SIE in the BLYP calculations is larger in the monomer calculations than in the dimer calculations, and the reduction of SIE counteracts the orbital energy shift by shifting the energy in the opposite direction (see inset in Figure 3).

The strong linear correlation between the \( \beta^{BLYP} \) and \( \beta^{B3LYP} \) values presents an opportunity to reduce the computational cost of future studies while retaining the accuracy of the
Figure 3: Change in the HOMO (square) and the LUMO (circle) energies when a monomer is extended to a dimer. The values obtained using the B3LYP ($x$-axis) functional are plotted against those obtained using the BLYP ($y$-axis) functional. Dashed lines represent the linear fit to the values obtained using the two functionals. The inset shows that the self-interaction error (SIE) decreases as the length of the oligomer chain is increased, which leads to an underestimation of the $\beta$ parameter value.
B3LYP calculations. The $\beta_{\text{B3LYP}}$ values can be estimated with a reasonably high accuracy via the relationship where $\beta_{\text{B3LYP}} \approx 1.2 \times \beta_{\text{BLYP}}$. Therefore, one can estimate the HOMO and LUMO energies of dimers at the B3LYP-level accuracy using a B3LYP calculation of a monomer (to get an accurate HOMO and LUMO energies of a monomer) and BLYP calculations of a monomer and a dimer (to calculate $\beta_{\text{BLYP}}$). Such a procedure eliminates the need for running B3LYP-level calculations on dimers, which is the most computationally expensive part.

**Accuracy of the TB model**

The HOMO and LUMO energy levels, IP and EA calculated using the TB model are compared against those calculated from the DFT simulation results in order to evaluate the accuracy of the TB model. The values calculated from the TB model are plotted against the ones calculated from the DFT simulations for all trimers and tetramers as shown in Figure 4. The HOMO and LUMO energies from the TB model agree well with those calculated from the DFT simulations with a mean absolute error of less than 50 meV. The largest error of $\sim 0.26$ eV is found in the HOMO energies of the trimer and tetramer of A13D8. The TB model predicts the IP and EA with high accuracy as well, which is not surprising since a linear dependence is expected between the IP/EA and HOMO/LUMO energies.\(^{41}\) Such a linear dependency is confirmed in our analysis as shown in Figure 5.

**$\Delta$SCF versus Kohn-Sham Energies**

The IP, EA and quasiparticle gap are plotted against $E_{\text{HOMO}}$, $E_{\text{LUMO}}$ and KS HOMO-LUMO gap, respectively, in Figure 5. The physical quantities of monomers extracted from the B3LYP calculations are shown on the left column of Figure 5 while the physical quantities of polymers extracted from the TB model based on the B3LYP-level data are shown on the right column of Figure 5. It can see from Figure 5a and b that the HOMO energy and IP have a strong linear correlation with $R \geq 0.94$. On the other hand, the linear correlation
Figure 4: Comparison of the values calculated from the TB model and DFT simulation results (at BLYP level) for the oligomers of three (solid square) and four (circle) repeating monomer units: (a) KS-HOMO energy, (b) KS-LUMO energy, (c) ∆SCF ionization potential, and (d) ∆SCF electron affinity.
between the EA and LUMO energy is weaker with $R = 0.86$ for monomers and $R = 0.92$ for polymers (see Figure 5c and d). The weaker correlation between LUMO energy and EA is expected to be coming from the polarization in the LUMO in the anionic calculation (i.e., the added electron in the SCF calculation of the anionic system does not distribute exactly in the same region as the LUMO of the neutral molecule). The weaker correlation between EA and LUMO leads to a weaker correlation between the $\Delta E_{\text{QP}}^{\text{QSCF}}$ and $\Delta E_{\text{KS}}^{\text{QSCF}}$ as shown in Figure 5e and f. A linear correlation still exists between $\Delta E_{\text{QP}}^{\text{QSCF}}$ and $\Delta E_{\text{KS}}^{\text{QSCF}}$. However, one cannot reliably convert the $\Delta E_{\text{KS}}^{\text{QSCF}}$ to $\Delta E_{\text{QP}}^{\text{QSCF}}$ as it was the case for $\beta_{\text{BLYP}}$ and $\beta_{\text{B3LYP}}$. In other words, one cannot reliably predict the $\Delta E_{\text{QP}}^{\text{QSCF}}$ based on $\Delta E_{\text{KS}}^{\text{QSCF}}$ without performing the $\text{QSCF}$ calculations. The lack of reliability in predicting the $\Delta E_{\text{QP}}^{\text{QSCF}}$ based on $\Delta E_{\text{KS}}^{\text{QSCF}}$ does not pose a significant issue because the computational cost of $\Delta E_{\text{QP}}^{\text{QSCF}}$ is not very high for monomers and dimers.

**Optical gap and its relationship with the quasiparticle gap**

The discussion so far has been limited to the quasiparticle and KS HOMO-LUMO gaps. The quasiparticle gap is often interpreted as the lowest optical excitation energy (optical gap) even though it does not explicitly account for the electron-hole interaction. The exciton binding energy lies in a range between 0.3 and 1.0 eV for conjugated D-A polymers (typically with a dielectric constant of around 4), which is much higher than those found in the traditional inorganic semiconductors used in photovoltaic applications. Therefore, the strong electron-hole interaction should be taken into account when describing the photoexcitation energy of the conjugated D-A polymers.

The optical gaps are calculated using two different methods: singlet-ΔSCF and triplet-ΔSCF. It is noted that it is more probable that the photoexcitation of a polymer with a weak spin-orbit coupling results in the generation of a singlet exciton. However, the triplet-ΔSCF is much more straightforward to perform because it does not require one to force any specific orbital occupation: the lowest triplet excitation energy can be calculated by
Figure 5: Physical quantities of monomers (left column) and polymers (right column) calculated based on B3LYP calculations and TB model based on B3LYP calculations (i.e., B3LYP calculations on monomers and $\beta_{\text{B3LYP}}$ estimated from the $\beta_{\text{BLYP}}$ values), respectively. The green texts and lines respectively represent the equations of the best linear fit and their plots. (a, b) IP vs. $E_{\text{HOMO}}$, (c, d) EA vs. $E_{\text{LUMO}}$ (e, f) $\Delta E^\text{QP}_{\text{SCF}}$ vs. $\Delta E^\text{QP}_{\text{KS}}$. 
calculating the ground state in the triplet spin configuration. The singlet-ΔSCF method, on the other hand, requires one to impose constraints on the orbital occupation instead of the spin configuration. The singlet-ΔSCF method suffers from the difficulty in achieving the convergence as it relaxes the “aufbau principle.” Consequently, the triplet-ΔSCF method is a much more attractive choice for screening studies that require a large number of calculations.

The optical gaps are calculated using both the singlet- and triplet-ΔSCF methods. We were able to achieve the convergence for 101 out of 104 monomers with the occupation constraint of one hole in the HOMO and one electron in the LUMO with singlet multiplicity due to the aforementioned challenges. The optical gaps calculated using the singlet-ΔSCF are plotted against those calculated using the triplet-ΔSCF method in Figure 6. The optical gaps calculated using the two methods show a strong linear correlation with $R = 0.97$: the singlet excitation energy is blue-shifted by 0.3 eV from the triplet excitation energy and the mean absolute error of the fit of 0.06 eV. Our observation agrees well with the previously reported results on porphyrins where the singlet-ΔSCF excitation energies were about 0.3 eV larger than the triplet-ΔSCF excitation energies. The strong linear correlation suggests that one can carry out the triplet-ΔSCF calculations and estimate the singlet-ΔSCF excitation energies with a reasonably high accuracy simply by adding 0.3 eV to the triplet-ΔSCF excitation energies. For the remainder of this section, we refer to the singlet-ΔSCF excitation energies as an optical gap as the photoexcitation is more likely to generate a singlet exciton.

An important feature of the optical gap that can be used to evaluate the validity of the quasiparticle gap of polymers is that the optical gap should always be smaller than the quasiparticle gap. The presence of the electron-hole interaction (i.e., exciton binding energy) makes the optical gap lower than the quasiparticle gap, and the exciton binding energy is known to be large for the conjugated D-A polymers (between 0.3 to 1.0 eV). Therefore, a basic test of any model consists of a comparison of the optical and quasiparticle gaps to ensure that the optical gap is smaller in all cases. It can be seen from Figure 7 that a
majority of the $\Delta E_{KS}^{QP}$ is smaller than the $\Delta E_{opt}^{\text{opt}}$, which is obviously unphysical. Therefore, the $\Delta E_{KS}^{QP}$ values must be interpreted with care. The $\Delta E_{\text{SCF}}^{QP}$, on the other hand, is always higher than the $\Delta E_{opt}^{\text{opt}}$. Furthermore, the singlet exciton binding energy obtained by taking the difference between the $\Delta E_{\text{SCF}}^{QP}$ and $\Delta E_{opt}^{\text{opt}}$ ranges from 0.20 to 1.08 eV for the polymers considered in this work, which agrees well with previous work\textsuperscript{43} where the exciton binding energies of conjugated D-A polymers were found to fall in a range between 0.3 and 1.0 eV.

The $\Delta E_{\text{SCF}}^{QP}$, $\Delta E_{KS}^{QP}$ for monomers and polymers and $\Delta E_{opt}^{\text{opt}}$ (assumed to be length independent) are shown in Figure 8. It can be seen from Figure 8 that the $\Delta E_{KS}^{QP}$ consistently underestimates the quasiparticle gap with respect to the $\Delta E_{\text{SCF}}^{QP}$, and the average difference is 2.45 eV for monomers and 0.90 eV for polymers. Several factors can contribute to such an underestimation. The aforementioned SIE in the GGA calculations overestimates the HOMO energy, which in turn results in an underestimation of the $\Delta E_{KS}^{QP}$. The use of a hybrid functional SIE tends to alleviate the problem, but the underestimation of the HOMO
Figure 7: $\Delta E^{\text{opt}}$ (Singlet-$\Delta$SCF) with respect to $\Delta E^{\text{QP}}_{\Delta \text{SCF}}$ (green) and $\Delta E^{\text{QP}}_{\text{KS}}$ (black) of polymers. The points on a dashed line corresponds to the cases where the optical and quasiparticle gaps are equal, and the points above (below) the dash line corresponds to the cases where the optical gap is larger (smaller) than the quasiparticle gap.

energy (compared to the IP, which should be a negative of the HOMO energy based on the Koopman’s theorem\textsuperscript{44,45}) can be seen in Figure 5a and b. A more severe underestimation of the $\Delta E^{\text{QP}}_{\text{KS}}$ for monomers compared to polymers aligns with the fact that the SIE is larger for the systems with a more localized charge density. A discrepancy in the charge distribution between the added electron in the anionic system and the LUMO orbital of the neutral system leads to an underestimation of the LUMO energy, which also contributes to the underestimation of the $\Delta E^{\text{QP}}_{\text{KS}}$. Finally, the derivative discontinuity\textsuperscript{46} of the electronic energy with respect to the integer number of electrons is another factor that can lead to an underestimation of the energy gap.

It is worth pointing out that the magnitude of the underestimation of the $\Delta E^{\text{QP}}_{\text{KS}}$ incidentally is similar to that of the exciton binding energy that the quasiparticle gap calculations are missing due to its variational nature and fictitious unoccupied orbitals.\textsuperscript{45,47} Consequently, the $E^{\text{QP}}_{\text{KS}}$ exhibits a reasonable agreement with the $\Delta E^{\text{opt}}$ as shown in Figure 9. The $\Delta E^{\text{opt}}$
Figure 8: $\Delta E_{\text{ASCF}}^{\text{QP}}$, $\Delta E_{\text{KS}}^{\text{QP}}$ and $\Delta E_{\text{opt}}$ of monomers and polymers. The values for monomers are obtained from the B3LYP calculations and values for polymers are obtained using TB model based on B3LYP calculations. The optical gaps for monomers are obtained from the singlet-$\Delta$SCF calculations with the B3LYP functional.
is blue-shifted from the polymer $\Delta E_{K_S}^{QP}$ by 0.19 eV while the polymer $E_{\Delta SCF}^{QP}$ is blue-shifted from the polymer $\Delta E_{K_S}^{QP}$ by 0.90 eV. Although the $\Delta E_{K_S}^{QP}$ is close to the $\Delta E_{opt}$, the low correlation ($R = 0.64$) makes it unsuitable to bypass the singlet- or triplet-$\Delta SCF$ calculations and predict the $\Delta E_{opt}$ from the $\Delta E_{K_S}^{QP}$ value.

![Figure 9: $E_{K_S}^{QP}$ with respect to $\Delta E_{opt}$ and $E_{\Delta SCF}^{QP}$](image)

**Summary and Conclusions**

The $E_{HOMO}$, $E_{LUMO}$, IP, EA and quasiparticle and optical gaps of a set of 104 conjugated D-A polymers have been systematically investigated. The focus of the present work is to develop a new method that can accelerate the computational search/screening process to find the optimal conjugated D-A polymers for the OSC application. The key findings of the present work are:

- The physical quantities calculated using the BLYP and B3LYP functionals exhibit a strong linear correlation, which allows one to accurately estimate the B3LYP calculation results using the BLYP calculations.
• The TB model can be constructed efficiently by using the B3LYP calculation results for the monomers as the basis values and estimate the B3LYP level intermolecular hopping parameter (off-diagonal terms in the model Hamiltonian) based on the BLYP calculations. This scheme can save the computational cost by an order of magnitude (see supporting information for more information).

• The TB model can accurately predict the $E_{\text{HOMO}}$, $E_{\text{LUMO}}$, IP, EA and quasiparticle gap of the oligomer of any size up to an infinitely long polymer.

• There is a systematic deviation (about 0.3 eV) between the triplet-ΔSCF excitation energies and the singlet-ΔSCF energies, which allows one to estimate the singlet-ΔSCF energies based on the triple-ΔSCF calculations.

• The quasiparticle gaps calculated using the IP and EA values from the ΔSCF simulations fulfills the condition that the quasiparticle gaps should always be larger than the optical gaps. In contrast, this condition is not fulfilled for the quasiparticle gaps calculated using the $E_{\text{HOMO}}$ and $E_{\text{LUMO}}$.

• The systematic errors in the calculations of the quasiparticle gap from the $E_{\text{HOMO}}$ and $E_{\text{LUMO}}$ cancel with the ignored exciton-binding energy to bring the $\Delta E_{\text{KS}}^{\text{QP}}$ in a reasonable agreement with the optical gap obtained from the singlet-ΔSCF calculations.

On the basis of these findings, we propose that future large-scale screening studies of D-A polymers can be further accelerated by using the TB model to estimate the energy level alignment in combination with the 0.3 eV bias to the triplet excitation energies to predict the singlet excitation energies. Although crude, one can alternatively use the HOMO-LUMO gap to estimate the singlet excitation energies if the ΔSCF calculations are unaffordable or inconvenient.

One of the main themes in contemporary computational research in the domain of materials science is to accelerate the evaluation of the electronic structure and the properties
that can be derived from it. Machine learning techniques have become very popular tool to achieve the acceleration, but at the cost of neglecting the physics and physical properties of the systems under study. Computational protocols with a physical insight, like the one presented here, alleviate this problem by helping to produce large data sets to train machine learning algorithms and to rationalize their outputs based on the particular physics of the system under study.

**Acknowledgement**

We acknowledge the financial support of the Innovation Fund Denmark through the project of title “Inks for large-scale processing of polymer solar cells“ (INKA) (Contract N. 5160-00008A). We also thank Eva Bundgaard for important discussions.

**References**


