Computational screening of chiral organic semiconductors: Exploring side-group functionalisation and assembly to optimise charge transport

Julia A. Schmidt,[†,‖] Joseph A. Weatherby,[‡] Isaac J. Sugden,[¶] Alejandro Santana Bonilla,[†] Francesco Salerno,[†,‖] Matthew J. Fuchter,[†,‖] Erin R. Johnson,[‡] Jenny Nelson,[§,‖] and Kim E. Jelfs[*,†,‖]

† *Department of Chemistry, Molecular Sciences Research Hub, Imperial College London, White City Campus, London W12 0BZ, UK*

‡ *Department of Chemistry, Dalhousie University, 6274 Coburg Rd, Halifax, Nova Scotia, B3H 4R2, Canada*

¶ *Department of Chemical Engineering, Imperial College London, London, SW7 2AZ, UK*

§*Department of Physics, Imperial College London, London, SW7 2AZ, UK*

‖ *Centre for Processable Electronics, Imperial College London, London, SW7 2AZ, United Kingdom.*

E-mail: k.jelfs@imperial.ac.uk

Abstract

Molecular materials are challenging to design as their packing arrangements, and hence their properties, are subject to subtle variations in the interplay of soft intermolecular interactions. Rational design of new molecular materials with tailored properties

is currently hampered by the difficulty to predict how a candidate molecule will pack in space and how to control the particular polymorph obtained experimentally. Here, we develop a rapid screening approach to aid the material design process, which is then applied to predict the charge-transfer properties of 1344 helicene compounds that have potential as organic electronic materials. Our approach bridges the gap between single-molecule design, molecular assembly, and the resulting charge-carrier mobilities. We find that fluorination significantly improves electron transport in the molecular material by over 200%, while side groups containing triple bonds largely lead to improved transfer integrals. We validate our screening approach through the use of full crystal structure prediction for the most promising compounds, to confirm the presence of favourable packing motifs that maximise charge mobility.

1 Introduction

Organic molecular materials hold promise as low-cost organic semiconductors (OSCs) due to their favourable magnetic, electrical, and optical properties. Development of OSCs offers technologies with new functionality, including as flexible, ultra-thin displays, biodegradable electronics, and energy-harvesting smart materials. OSC performance depends on the ease with which charge-carriers can move across the π-stacked systems of the component molecules, which is predominantly influenced by how they are arranged within the solid material,[1,2] as well as the energies of their frontier orbitals. The soft intermolecular interactions in molecular materials typically give rise to many isolable solid forms, or polymorphs, within a small energy window, the structures of which are difficult to predict from first principles. Changes in molecular packing between different polymorphs can result in large variations in physical properties.

Materials scientists are continually searching for organic semiconducting molecules with solid forms that give rise to high charge-carrier mobilities and good processability.[2,3] Previous methods to screen small molecules for potential suitability as OSCs include data mining,[4,5]

atomistic substitutions,[6] and machine-learning approaches.[7,8] For example, Kunkel *et al.* screened the Cambridge Structural Database (CSD) to obtain statistically significant structure-property relationships, with certain scaffolds (side groups) leading to consistently improved charge-transport properties.[4] Also, atomistic substitution leading to a modified polymorph landscape, and thus changed charge-transport behaviour, has been experimentally investigated by Sorli *et al.*[6] From a computational perspective, Pulido *et al.* combined crystal structure and property prediction to build energy-structure-function maps that describe the possible polymorphs of a candidate molecule and their respective transport properties.[9,10]

Recently, crystal structure prediction (CSP) has found increasing use in the materials community.[11–14] First-principles CSP refers to the computational prediction of the likely crystal structures of a molecule, starting from nothing more than the atom connectivity. The most common strategy is a multidimensional search of the crystal-energy landscape for local minima corresponding to different crystal structures.[15,16] All tentative crystal structures are then ranked by their lattice[17] or free energy to determine the thermodynamically most stable structures using classical force fields.[18] The lowest-energy crystal structures can be re-ranked more accurately, for example using periodic dispersion-corrected DFT calculations at 0 K.[19,20] Finally, thermal free-energy corrections can be applied to the lowest-energy candidates.[21,22] The relationship between molecular structure, crystal packing, low-frequency vibrations and charge transport still poses many open questions and there would be significant value in developing a framework to design these materials.[23,24]

We are particularly interested in chiral helicene molecules as potential OSCs. Helicenes are axially fused benzene rings, as shown for [6]helicene in Figure 1a. Due to the chirality present, a [6]helicene molecular crystal can either contain only one chiral form (enantiopure) or molecules of both handedness (racemic). Small-molecule chirality can be used in organic semiconductors to provide added opportunities from chiral composition-dependent changes to bulk properties,[25] chirality-driven crystallisation of racemic over enantiopure candidates,[26] and the control of spin for usage in chiral-induced spin selectivity (CISS) filters.[27] Helicenes

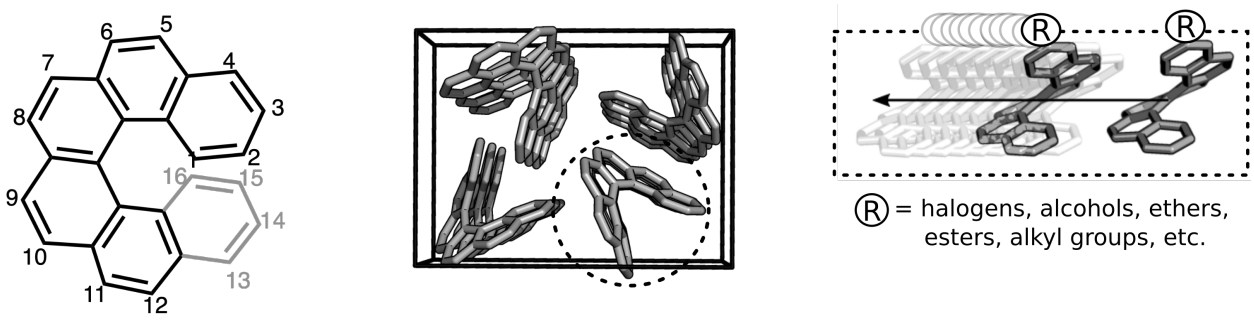

Figure 1: (a) Structure of the *M*-[6]helicene backbone, which is mono- and symmetrically di-substituted with different functional groups in all positions (1-16) during screening. (b) The experimental crystal structure of [6]helicene (front view), with the circled translational motif in the plane of the page. (c) The translational motif of the functionalised [6]helicene (side view) for the particular case of monosubstitution in the 3-position.

have already been successfully tested for applications as chiroptical switches,[28] in circularly polarised (CP) OLEDs,[29] in CP light photodetectors,[30] and as transistor materials.[31] This suggests future usage of helicenes in a large range of potential applications. For example, in OLED screen development, chiral materials have generated interest for use in the preparation of CP-OLEDs. CP-OLEDs provide a means to address energy losses by commonly used antiglare filters, which eliminate glare from external light sources (*e.g.* sunlight).[25]

Previously, we performed CSP studies to predict the carbo[6]helicene and aza[6]helicene crystal-energy landscapes and found that the chirality of the crystal (enantiopure *vs.* racemic) can drastically affect the corresponding charge-carrier mobilities.[31,32] For carbo[6]helicene, we structurally classified polymorphs into recurring packing motifs. A particular 'translational motif' (Figure 1c) was found with the highest probability across the entire crystal-energy landscape, being observed in 63% of the hypothetical low-energy polymorphs.[32] Most recently, Salerno *et al.* studied the effect of the nitrogen substitution position on charge transport for the most commonly observed [6]helicene crystal structure (CSD code: HEXHEL, Figure 1b),[33] and found that this same 'translational motif' gave rise to the highest electron-transfer integrals across the majority of aza[6]helicene isomers.

While one may design a material that has promising properties, unless the thermodynamic

viability and thus the synthetic route to the material has been considered, then it is unlikely that the material will actually be obtained.[34] Instead, one must screen thousands of potential materials, considering their viability and then their properties.[35] Thus, in the case of organic semiconductors, if we wish to 'design' a molecule with a target property, we must consider the solid-state packing. In practical terms, to include the effect of packing within a high-throughput computational study, as we attempt to do here, it is essential to make some approximations and consider a subset of possible structural degrees of freedom.

In this work, we leverage the ubiquity of the translational motif for helicene molecules to screen for high charge mobility across ∼1300 molecules. We focus on investigating the helicenes as n-type semiconductors, examining their electron mobilities, because [6]helicenes tend to be better electron-transporting than hole-transporting materials based on our computational predictions.[32] Since CSP is too computationally exhaustive to screen for large numbers of new molecular materials, the computed transport across a translated-dimer model serves as a proxy to predict charge transport in the solid material. Computational screening via the simplified dimer model, based on one of the most common [6]helicene dimer motifs, elucidates (i) a molecule's tendency to adopt the chosen translational motif, (ii) its suitability to maximize charge mobility, (iii) effects of substitution position for a given functional group, and (iv) whether asymmetric or symmetric disubstitution is preferable. For the most promising compounds, we perform a full CSP search and compute the electron mobilities for the lowest energy polymorphs, validating our approach. The overarching goal of this research is to uncover structure-property relationships to aid the design of new electronic materials with tailored properties.

2 Experimental Section

Compound Screening

Functional-group selection and dimer generation

The *M*-[6]helicene backbone was systematically monosubstituted in all 16 possible hydrogen positions (1-16, as shown in Figure 1a), and disubstituted in all 8 possible symmetric configurations (1,16-, 2,15-, 3,14-, 4,13-, 5,12-, 6,11-, 7,10-, and 8,9-disubstituted), to study the effect of terminal functionalisation on electron transport. We selected a range of common functional groups, such as halogens, alcohols, esters, ethers, carboxylic acids, phenyl groups, and alkyl groups, with the full set of substituents shown in Figures S3-S4. We focused on side groups of small to medium size compared to the size of the [6]helicene backbone, such that interactions involving the [6]helicene backbone are still the main factors determining molecular assembly. The total numbers of substitution patterns (24) and functional groups (56) resulted in a total of 1344 (24×56) compounds studied.

To screen 1344 synthetically unknown [6]helicenes for their suitability as OSC materials, we first assume that the electronic coupling observed across one single translational dimer is sufficient to qualitatively assess the relative transport properties of the candidate molecular materials. The translational motif (Figure 1c) in [6]helicenes was first identified by Rice *et al.* and is more common than other dimer motifs, such as the interlocked dimer, the herringbone dimer, or the back-to-back dimer (Figure S1).[32] The translational dimer is, as the name suggests, a purely homochiral (*M*- or *P*-handedness only) motif, where one molecule is translated with respect to the other along a vector perpendicular to the plane of the helix, and is characterised by good π-π overlap of the adjacent [6]helicene backbones (Figure 1c). This translational motif was chosen due to its common occurrence in previously studied [6]helicene polymorphs[31–33] and across [6]helicene molecules deposited in the CSD (see Section S1.2, Table S1 and Figure S2) - indeed, it is the most commonly occurring 1-dimensional motif in these systems. It is, therefore, an informed guess that a new hypothetical [6]helicene might

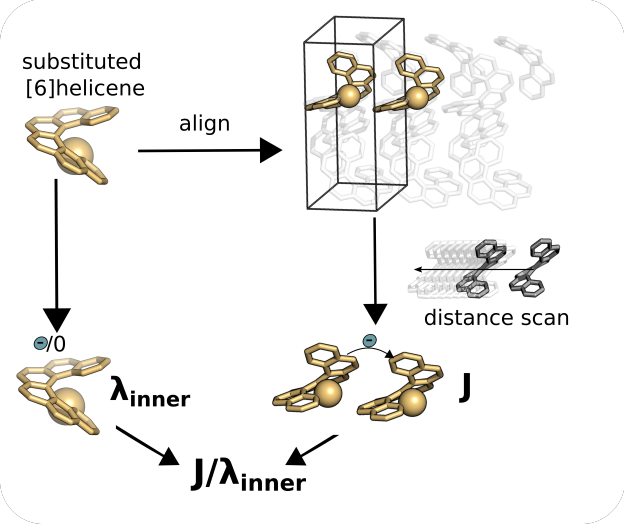

Figure 2: Overview of the screening procedure. A side group of choice is attached to the helicene backbone (top left) and the molecular geometry optimized. The substituted [6]helicene is then mapped back onto the unit cell, with the translational dimer shown along the *c*-axis, in the plane of the page. An intermolecular distance scan for this dimer is performed by modifying the *c*-axis length. The transfer integral (J) is computed at the minimum-energy distance for the dimer, while the inner reorganisation energy (λ_{inner}) is computed from only the optimised single-molecule geometries for the neutral and anionic states. The ratio of these quantities is calculated as one measure of OSC suitability.

contain the translational packing motif in low-energy, synthetically achievable, polymorphs (somewhere in its crystal-energy landscape). We test this assumption later in our work by performing full CSP searches to determine the packing motifs observed in the polymorphic landscapes of the most promising candidate molecules. The choice of translational dimer is further supported by previous work on TIPS-Pn thin films in which the strongest vibrational mode corresponds to the translational vibration of the pentacene backbone along the main stacking axis (a translational motif).[36]

The generalised screening procedure is illustrated in Figure 2. A single *M*-carbo[6]helicene molecule was extracted from the experimental carbo[6]helicene crystal structure (CSD code: HEXHEL, Figure 1b) and either monosubstituted or symmetrically disubstituted with a new side group. A short molecular dynamics (MD) simulation with the OPLS3e force field[37] was performed to obtain a realistic starting geometry for density-functional theory (DFT) optimisation, which is important in the case of flexible side groups that can adopt multiple

conformers. The full MD details and sampling information can be found in Section S2.1. The geometry of the substituted [6]helicene molecule was then optimised at the B3LYP[38]/6-31G(d) level of theory using Gaussian16.[39] The optimised molecular geometry was superimposed back onto the original [6]helicene position inside the experimentally observed HEXHEL crystal structure using a python open-source toolkit for cheminformatics, (RDKit).[40]

Translational dimers, based on those observed in the experimental carbo[6]helicene crystal structure, were extracted for each of the substituted [6]helicene molecules. As the intermolecular distance for the unsubstituted [6]helicene is not necessarily appropriate for the functionalised molecules, we performed a one-dimensional intermolecular distance scan along the c axis of the crystal (Figure 3) from 6 Å to 14 Å, in 0.1 Å increments, for each compound. The resulting minimum-energy distance was then used in all further property computations. The semi-empirical tight-binding method GFN-xTB[41] was chosen as a cheap, but powerful, compromise to perform the distance scans. High-level DLPNO$-$CCSD(T),[42] or even conventional DFT with large basis sets and a Grimme-D3 dispersion correction,[43] is too computationally costly to apply to 1344 molecules. Thus, we benchmarked the accuracy of the semiempirical GFN-xTB and the OPLS3e force field, compared to DLPNO$-$CCSD(T),[42] for the potential energy scan of carbo[6]helicene (Figure 3) and two terminally substituted [6]helicenes (Figures S7 and S8). We found that the minimum energy distances, d_{min}, at the DLPNO-CCSD(T) and GFN-xTB levels of theory consistently varied only by around 0.1 Å, validating the choice of GFN-xTB.

Property prediction

For each molecular dimer at its minimum-energy separation, we computed the transfer integrals by projecting the computed orbitals of the dimer onto the unperturbed localized orbitals of the individual molecules.[32,44] The transfer integral is a measure of the electron (or

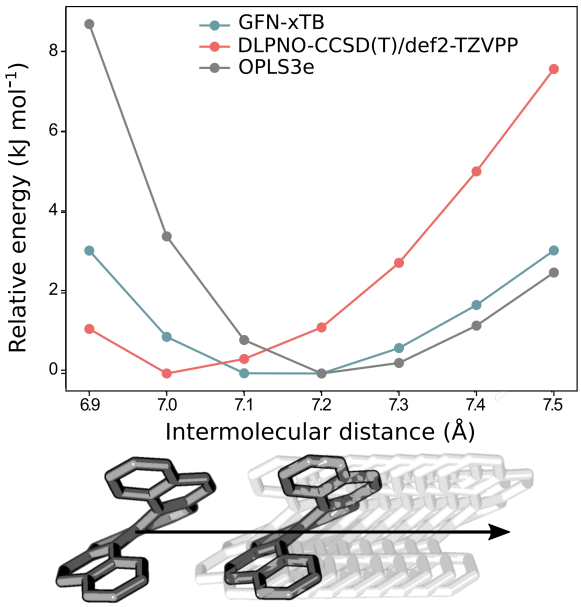

Figure 3: Intermolecular distance scans for the carbo[6]helicene translational dimer at three different levels of theory: the DLPNO-CCSD(T) reference, the GFN-xTB semi-empirical method, and the OPLS3e force field. A rigid scan along the *c*-axis of the HEXHEL crystal structure (as shown by the arrow) was performed, with the lowest-energy dimer highlighted in dark grey.

hole) coupling between two adjacent interacting molecules:

$$J_{\text{elec}} = \langle i|\hat{\mathbf{F}}|j\rangle, \tag{1}$$

where $\hat{\mathbf{F}}$ denotes the Fock matrix of the dimer system, while $|i\rangle$ and $|j\rangle$ represent the orbitals localized on the molecules i and j, respectively. All transfer integrals were computed with B3LYP/6-31G(d) using Gaussian16.[39] This level of theory has been proven successful in adequately describing the transfer integrals for [6]helicenes[32] and other organic semiconductors.[44] For full details of the transfer integrals and how they are employed using non-adiabatic, semi-classical Marcus theory, see Section S3.4.

A concern with using the translational dimer model as a proxy to screen for high bulk charge mobility is the distance dependence of the computed properties. It is commonly understood that transfer integrals exponentially decrease with the intermolecular distance, and

this is seen in this work as well for the three test cases of [6]helicene, 2,15-dicyano[6]helicene, and 2,15-diethynyl[6]helicene (Figures S6, S7b, and S8b). This raises the question of whether the transfer integrals should be normalised by the distance, to enable comparison across compounds and deconvolute the impact of substitution from the distance dependence. However, the optimized dimer separations are our best approximation to the realistic intermolecular distances found in the bulk materials, and these will necessarily vary with molecular size and charge distribution, depending on the side groups. Thus, we do not normalise the transfer integrals, and only compare results at each dimer's minimum-energy separation.

In addition to computing the transfer intergrals, the inner electronic reorganisation energy $\lambda_{\text{inner}}^{\text{elec}}$ was computed for each individual molecule using four-point-method:[45]

$$\lambda_{\text{inner}}^{\text{elec}} = E_n^- - E_c^- + E_n^0 - E_c^0. \tag{2}$$

Here, single-point energies were computed for the anionic or neutral states ($-$ or 0, as indicated by the the superscripts) at either the charged or neutral optimized geometries (c or n, as denoted by the subscripts). The reorganisation energy is a quantitative measure of the ability of a single molecule to accommodate a negative charge, and the energetic penalty for conversion between the charged to the neutral geometries.

For a total of 56 of the 1344 molecules (4%), at least one of the single-point energy calculations required to compute the inner reorganisation energy did not converge. These examples were excluded from the subsequent analysis. However, to make sure that we do not miss any promising molecules, we used machine-learning regression methods to predict the missing inner reorganisation energies based on our existing data. We compared both regression using Xgboost[46] and kernel ridge regression techniques[47] on Morgan radial fingerprints[48] and managed to describe up to 90% of the variance using the latter (see Section S2.4).

There are several ways that the OSC suitability of a candidate molecule can be assessed based on the properties we have computed. Some approaches focus more on variance of

J rather than the inner reorganisation energy, or *vice versa*.[4] To combine the intra- and intermolecular aspects of charge transport in a single parameter, we choose the ratio $J_{\text{elec}}/\lambda_{\text{inner}}^{\text{elec}}$ to quantify the relative trends and identify promising candidate molecules. The higher the ratio, the higher the suitability of the molecule as an electron-transporting OSC. The same methodology could also be applied to hole transport, but for the purpose of this study we focus on electron transport only.

As an alternative screeening metric, we also consider the electron mobilities, μ, obtained from non-adiabatic Marcus theory in the low-field limit. While a number of theoretical models for describing charge transport in different regimes exist, Marcus theory[49,50] is one of the most popular methods to describe charge transport in OSCs.[24] It expresses charge hopping between weakly coupled sites with localised charges (hopping regime, $|J_{ij}|<<\lambda$) and complements band transport theory, which instead assumes delocalised charges in the low-temperature limit (band regime).[24] Assuming a periodic one-dimensional array of helicenes, regularly spaced along the axis that defines the translational dimer, the charge-carrier mobility along the chain can be calculated as an approximation to the mobility in the bulk material. In the low-field limit, the mobilities can be obtained from the transfer integral (J), the reorganisation energy (λ), and the intermolecular centre-of-mass distance (a), as

$$\mu = \frac{|J|^2}{\hbar}\sqrt{\frac{\pi}{\lambda k_B T}}\left(\frac{ea^2}{k_B T}\right)\exp\left(-\frac{\lambda}{4k_B T}\right), \tag{3}$$

with more detail given in Section S2.5.

The absolute value of the mobility depends on the reorganisation energy, one component of which, the outer-sphere reorganisation energy, cannot in general be calculated explicitly for the structure of the individual molecule. We have used a conservatively low value of 0.3 eV for the outer-sphere reorganisation and computed the inner reorganisation energy in the gas phase for each molecule (Eq. 2). As reorganisation energies should not be strongly dependent on morphology,[51] we have used this simplified, constant outer reorganisation energy

across all crystal morphologies. The value of our calculated mobilities is uncertain to within approximately the same factor for all systems, as the outer-sphere reorganisation energy is in essence a scaling parameter that would affect all systems to a similar degree, and thus not change any relative trends. In the screening, we only consider the inner reorganisation energy ($\lambda = \lambda_{\text{inner}}^{\text{elec}}$) in our J/λ values, so those ratios will be overestimates of the true values. We note that, whilst we cannot expect the absolute values to be correct, we can compare relative trends between systems for which the same outer reorganisation value is used. While the assumption of weak coupling (hopping regime) may not be valid in every case, it is most likely to be valid for low J and, hence, low J/λ, so the screening methods should still be good at eliminating unpromising structures. We will discuss with care in the results section those systems that may not be adequately described by the hopping regime, but we always consider high J values to be a necessary condition for high mobility.

Crystal Structure Prediction

For the four most promising compounds with high $J_{elec}/\lambda_{inner}^{elec}$ ratios and electron mobilities, full CSP searches were performed to study the low-energy polymorphs observed on the crystal-energy landscape. This allowed validation of our assumption that the translational motif used for screening will occur in low-energy polymorphs of the candidate molecules. CSP also allows us to verify our dimer-model results by more sophisticated computational determination of the charge-carrier mobilities for the lowest-energy solid forms. Computational investigation of the polymorphic landscape and the associated charge mobilities rapidly provides information for new helicene compounds that could only be attained experimentally by a lengthy synthesis of the material, followed by (re)crystallisations and single-crystal mobility measurements.

The polymorphic energy landscapes of the top four candidate molecules identified in the screening process were generated using CrystalPredictorII,[16] assuming a single molecule in the asymmetric unit ($Z'=1$) and searching the most common chiral and achiral space groups (see Section S3.1 for details). The energies of the putative crystalline structures were subsequently

reranked using distributed multipole analysis (DMACRYS[52]). For all unique crystal structures within 15 kJ mol^{-1} of the energy minimum, single-point energies were computed using dispersion-corrected density-functional theory. Specifically, the SIESTA[53] implementation of the B86bPBE[54,55] functional and the exchange-hole dipole-moment (XDM) dispersion correction[56] was used with the double-zeta plus polarization (DZP) basis set.[57] Single-point energies from B86bPBE-XDM/DZP have been shown to rapidly provide improved crystal-energy landscapes compared to distributed multiple analysis,[20] while full projector augmented-wave calculations using XDM are highly accurate for the intermolecular interactions within helicene crystals.[31,32] Costly periodic relaxations of the crystals with DFT-XDM were omitted, due to time and computational cost constraints. Any future in-depth computational study of any of the screened polymorphs identified as interesting should begin with a full relaxation of the crystal structures before calculation of charge mobility properties. A more detailed description of the CSP methodology can be found in Section S3.

For the 10 lowest-energy DFT-ranked crystal structures of each compound, the charge-carrier mobilities were computed using non-adiabatic semi-classical Marcus theory, as described in Section S3.4. In contrast to the low-field limit mobility calculations based on a single dimer performed in the screening section, we computed the charge transport routes across the material in all three dimensions, for all dimers within a 15 Å cutoff, thereby allowing a penalty for the outer reorganisation energy, to yield a much more realistic representation of the charge-carrier mobility. Thus, in this second half of the discussion, we computed μ based on the CSP-generated crystals (Figure S32), and we included an outer reorganisation energy of $\lambda_{outer} = 0.3$ eV as explained above. Finally, when comparing crystal mobilities obtained for our new candidates with previous helicene results by Rice *et al.*[32], we recomputed the mobilities using a $\lambda_{outer} = 0.3$ eV, to ensure that our mobilities can be qualitatively compared.

3 Results and Discussion

OSC Suitability Screening

The transport properties of 1344 mono- and di-substituted helicene molecules were screened for potential OSC suitability using two different assessment criteria: (i) the straightforward $J_{\text{elec}}/\lambda_{\text{inner}}^{\text{elec}}$ ratio (hereafter referred to as J/λ) and (ii) the charge-carrier mobility, μ, along the translational chain in the low-field limit. Both approaches contain the transfer integral and inner reorganisation energy; the mobility also explicitly considers the intermolecular centre-of-mass distance of the dimer. In general, we aim to maximise the charge transfer integral, J_{elec}, while minimising the inner reorganisation energy, λ_{inner}^{elec}, to yield optimal charge transport.

We note that we are assuming a hopping regime, which requires low J/λ values to be valid. If anything, our approach, with λ only accounting for the internal reorganisation energy, is likely to overestimate J/λ, as an additional outer reorganisation energy to the denominator would decrease the J/λ values. Nevertheless, it appears likely that some of our systems with higher J/λ values would no longer be within a hopping regime. If not in the hopping regime, we would expect the systems to be in a 'transient localisation' regime rather than in a band transport regime. However, in either case, the mobility should increase with the power of J so that high J values remain a necessary condition for good charge transport and, thus, remain a useful screening parameter.[58] Figure S16 shows the distribution of J values across the systems.

This screening provided insight into the impacts of side groups and substitution position(s) on electron transport. The calculated J/λ ratios are shown in Figures 4 and 5. The raw data is also available at `https://data.hpc.imperial.ac.uk/resolve/?doi=7858`. Notably, there is no side group which consistently outperforms the carbo[6]helicene example. This is likely due to the steric hindrance of the translational dimer for some substitution positions. The J/λ ratio generally increases towards the limiting value for the bare helicene (from left

to right, Figure 4), because some substitution positions along the backbone are consistently disadvantageous for charge transport (independent of the side group attached), while in some positions, if the correct side group is attached, the terminally substituted [6]helicene has a larger intermolecular electronic coupling integral and thus a higher J/λ ratio. Hence, the unfavourable substitution positions tend to be poor for any side group. However, for the better substitution positions, if a more favourable side group is attached, the J/λ ratio increases in some, but not all, positions (thus there is a larger spread of J/λ values). Only a minority (0.4%) of the J/λ ratios computed for substituted [6]helicenes exceed the [6]helicene reference. This finding is of value in suggesting that, while it may seem intuitively promising, the functionalisation of [6]helicenes to improve electron transport is not often a fruitful approach in practice. This emphasises the complexity of the multivariable problem of optimising charge transport in organic semiconductors.

Overall, both screening metrics revealed three largely similar classes of terminal substituents, two of which give rise to electron transport: (i) small electronegative side groups (*e.g.* halogens, CN), that give the best transport across all substitution patterns; (ii) moderately sized acyclic side groups with unsaturated bonds (diacetylene, triacetylene) that give rise to good transport for some substituent positions, usually manifesting as large outliers; and (iii) bulky alkyl substituents, cyclic or acyclic, that result in very poor transport. The same trends are observed for μ as for the simpler J/λ ratio, but are less pronounced. Across all substitution patterns and positions, we observe that smaller side groups, which do not heavily perturb the geometry of the [6]helicene backbone, perform better than bulky substituents (Figure 4). Bulky substituents are more likely to alter the inner strain and interplanar angle (the dihedral angle between the two terminal rings) of the [6]helicene backbone, and thus affect electron delocalisation across the backbone.

Despite their larger size, side groups containing double or triple bonds result in transfer integrals that are higher than expected from the general trend of an exponential decrease with increased molecular separation (Figure S15). In particular, triple bonds lead to favourably low

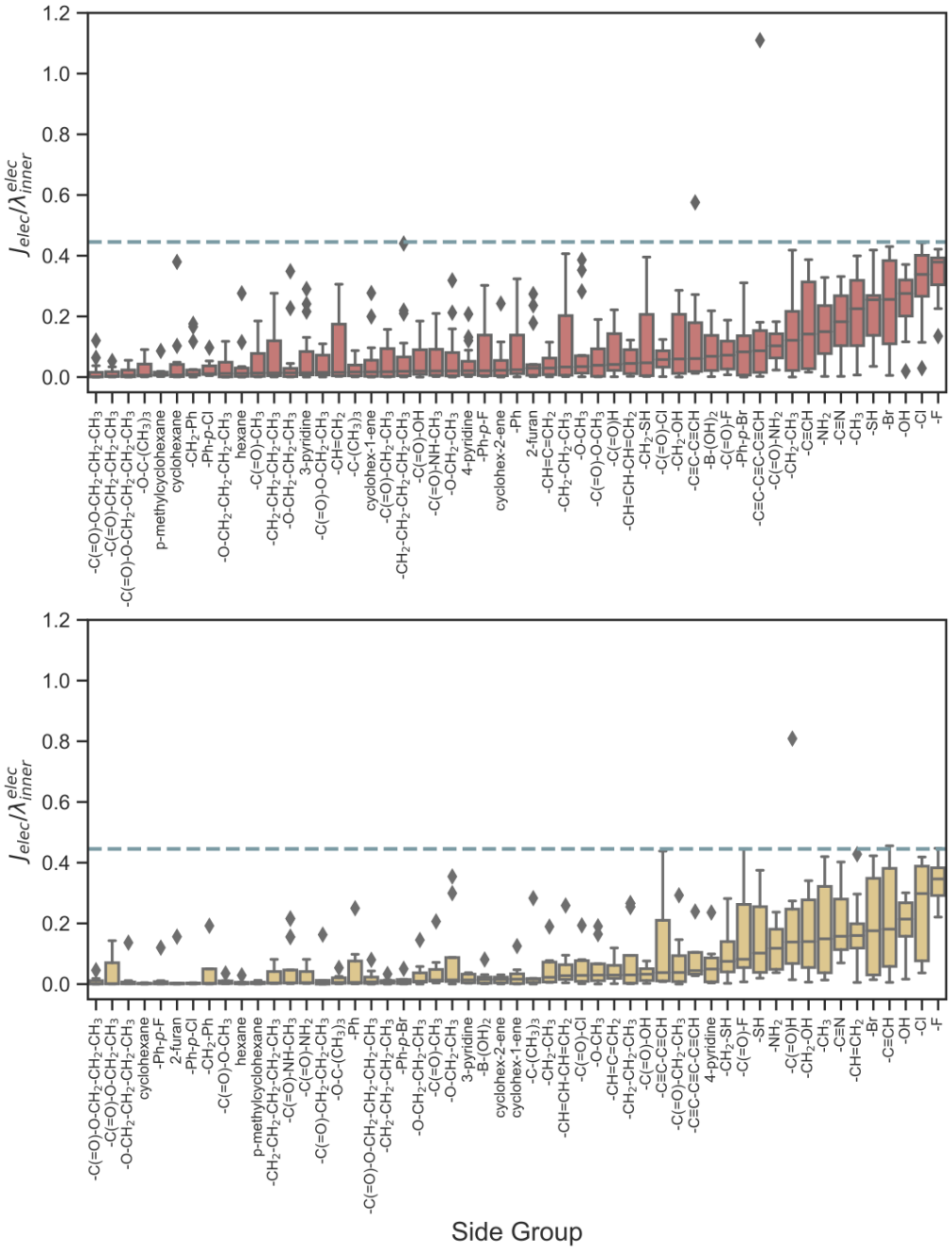

Figure 4: Boxplots illustrating the OSC suitability of functionalised [6]helicene molecules, using $J_{\mathrm{elec}}/\lambda_{\mathrm{inner}}^{\mathrm{elec}}$ as the target property to be maximized. Results are shown for monosubstitution (top, red) or symmetric disubstitution (bottom, yellow) of the [6]helicene backbone. For each side group, 16 (monosubstituted) or 8 (disubstituted) data points were obtained, corresponding to all possible backbone sites. The size of the box is the interquartile range, containing 50% of the ratios, and the horizontal line inside each box is the median value. The whiskers indicate the range of ratios obtained outside the middle 50% of data points, with the exception of some large outliers, shown as grey diamonds. The dashed line is the unsubstituted [6]helicene reference. Figures S9 and S16 show the different components (λ, J) of this data separately.

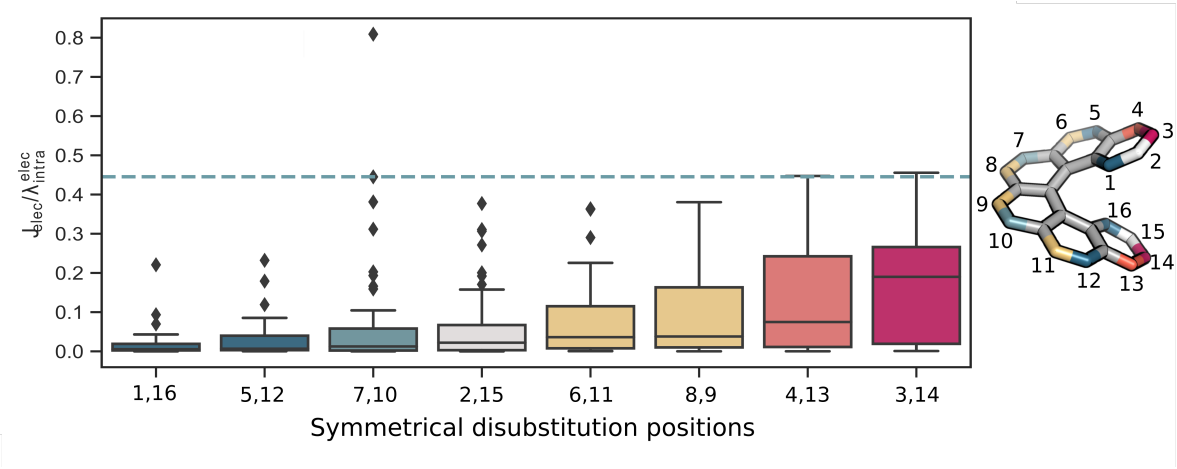

Figure 5: Boxplots illustrating ranges of the computed $J_{\text{elec}}/\lambda_{\text{inner}}^{elec}$ ratios spanning all 56 side groups. Results are shown for the 8 symmetrical disubstitution patterns of the functionalized [6]helicene molecules, sorted in order of ascending median, and the colouring ranges from good charge transport (red) to no charge transport (blue). Significant outliers are shown as grey diamonds and the dashed line is the unsubstituted [6]helicene reference. Analogous results for the monosubsituted [6]helicenes are shown in Figure S22.

inner reorganisation energies, which shows that the ability to accommodate a negative charge is enhanced for these substituted [6]helicenes. This behaviour is suggested by the LUMO being delocalised over the extended π-system introduced by the triacetylene unit (Figure S27).We find that this trend increases with the number of triple bonds per side group used for disubstitution (Figure S9 bottom). Functionalisation strategies involving the inclusion of the triple-bonded 6,13-bis(triisopropylsilylethynyl) (TIPS) functional group to fine tune electronic properties have been used previously for other well-studied OSC molecular cores, such as pentacene[59] or dinaphtho-fused *s*-indacenes.[60] TIPS side groups can increase a molecule's solubility, as well as directing the molecular packing, but without losing intermolecular electronic coupling at the same time. Molecular packing can be controlled by modifying the shape and size of the triple-bond-containing substituents. Depending on the ratio of the diameter of the side group to the pentacene backbone, different packing arrangements with altered intermolecular interaction strengths are observed.[61,62] Triple-bond containing side groups (*e.g.* TIPS-Pn, TMTES-Pn, diF-TESADT) can influence the distribution of intermolecular and intramolecular vibrational modes to minimise non-local electron-phonon

coupling processes along the π-π stacking direction (translational motif) that reduce the overall mobility.[36] A longer discussion of the effect of π-system size on λ_{inner} can be found in Section S2.6.1.

Fluorine substituents outperform any other side group for both mono- and disubstituted [6]helicenes when considering J/λ ratios. This finding is consistent with previous research in leveraging electron-withdrawing halogens to lower the LUMO energy and the electron-injection barrier, such that electron and hole transport occurred simultaneously.[63–65] Fluorination has been previously applied to enhance transport in other organic semiconductors, *e.g.,* arylene diimides,[66,67] naphthalene,[68] or functionalised acene derivatives,[61] such as pentacene[69–71] or TIPS-pentacene.[71] The impact of fluorination on reducing intermolecular spacing has been investigated for many OSC molecules, including carbon diimides,[72] naphthalene,[68] and benzene.[73]

Overall, we found that asymmetrical monosubstitution results in higher mean, median, and maximum transfer integrals than symmetrical disubstitution. Monosubstitution of [6]helicenes has been less experimentally investigated, because the quickest and most common synthetic route to [6]helicenes is photocyclisation,[74] which gives symmetrically disubstituted products. The best side-group placements ranked by median were found to be the 3,14- and 4,13-positions for disubstitution and the 3-, 9-, and 13- positions for monosubstitution across all the side groups assessed (Figures 5 and S22).

The top-ranked helicenes for OSC performance, obtained from both metrics, are given in Table 1. As stated above, the fluorine substituent consistently gives the highest mean screening metrics, with 15-fluoro[6]helicene and 4,13-difluoro[6]helicene being the best performers for mono- and di-substitution, respectively (13-fluoro[6]helicene is the second best mono-fluorination position). The other top-ranked cases consist of large outliers from the general trends shown in Table 1, and Figures 4 and 5. Across both screening metrics, the top three performing disubstituted helicenes were 7,10-dialdehyde, 3,14-diethynyl, and 4,13-difluoro[6]helicene. Similarly, both measures gave 6-triacetylene and 7-diacetylene and 2-

chloro[6]helicene as the best monosubstituted helicenes using the J/λ ratio, and the maximum low-field mobility. Giving consideration to the fact that the 6-triacetylene system with a high J/λ value may no longer be in the hopping regime, we note that that system still has a high J value (see Table 1 and Figure S16) that would be promising for charge transport in a 'transient localisation' regime. Thus, the selection of systems for further investigation would remain the same. In an ideal world, for systems where the hopping regime may no longer be valid, one would want to fully consider the factors that would be important in the 'transient localisation' regime. Namely, one would want information on the disorder in J resulting from thermal fluctuations, the degree of anisotropy in these fluctuations, and the consequent effect on reducing mobility.[75] It could, for example, be true that there are large thermal fluctuations in the 6-triacetylene system, causing anisotropic transport and thus reduced mobility. To have a better insight into this, we would ideally want to calculate the intermolecular vibrational modes in each system, but this requires computationally demanding calculations where the full crystal structure is known, and is thus clearly beyond this screening study of over 1300 systems where the crystal structures are not *a priori* known.

Please refer to the supporting information, section S2.6, for an extended discussion of the screening results, covering how molecular structure affects electron transport (S2.6.1), the separation dependence of transfer integrals (S2.6.2), whether terminal substitution improves transfer integrals (S2.6.3), and whether triple bonds are good for transport (S2.6.4).

Analysis of the CSP results

From the initial screening of OSC suitability, fluorine was found to be the best side group overall, independent of the substitution pattern and the exact functional group position. Thus, the highest performing disubstituted fluoro[6]helicenes, 4,13-difluoro[6]helicene, and the second highest and structurally related monofluorinated [6]helicene, 13-fluoro[6]helicene, were selected for CSP analysis. Additionally, the two extreme outlier cases in Table 1, 7,10-dialdehyde[6]helicene and 6-triacetylene[6]helicene, were also selected for CSP. The structures

Table 1: Top three candidates for the mono- and disubstituted [6]helicenes according to both OSC measures: $J_{elec}/\lambda_{inner}^{elec}$ ratio and low-field limit charge-carrier mobility. The electron mobility μ_{elec} is given in $\times 10^{-3} cm^2\ V^{-1}\ s^{-1}$. The relevant J_{elec} values are stated in meV.

Rank by $J_{elec}/\lambda_{inner}^{elec}$	Monosubstition	$J_{elec}/\lambda_{inner}^{elec}$	Disubstitution	$J_{elec}/\lambda_{inner}^{elec}$
1	6-triacetylene	1.11	7,10-dialdehyde	0.81
2	7-diacetylene	0.58	3,14-diethynyl	0.46
3	2-chloro	0.44	4,13-difluoro	0.45
6	15-fluoro	0.42		
12	13-fluoro	0.42		
Rank by μ_{elec}	Monosubstitution	μ_{elec}	Disubstitution	μ_{elec}
1	6-triacetylene	1.46	7,10-dialdehyde	0.75
2	7-diacetylene	0.38	3,14-diethynyl	0.28
3	2-chloro	0.26	4,13-difluoro	0.27
8	15-fluoro	0.24		
14	13-fluoro	0.24		
Rank by J_{elec}	Monosubstitution	J_{elec}	Disubstitution	J_{elec}
1	6-triacetylene	251	7,10-dialdehyde	191
2	7-diacetylene	135	7,10-diacylfluoride	118
3	15-fluoro	74	5,11-diacylfluoride	79
4	14-propylether	73	4,13-difluoro	78
29	13-fluoro	71		

of these four compounds are shown in Figure 6. For each of these candidate molecules, no solid-state structure has ever been synthesised, nor have mobilities been measured. However, our previous work has shown that CSP can provide predictions of experimentally realised helicene crystal structures and we can qualitatively compare crystal mobilities using hopping theory, including explaining experimentally obtained properties.[31,32] We note the caution with the application of the hopping regime to the 6-triacetylene case mentioned already above. To investigate further the capability of our approach to identify promising OSC molecular materials, we generated the low-energy crystal packings for the four selected candidates using CSP. The crystal-energy landscapes and resulting low-energy polymorphs are shown in Figure 7. In general terms, the lower in energy a polymorph is, the more likely it should be synthetically achievable.

As seen in Figure 7, the translational motif was repeatedly found in the crystal-energy landscape of the 13-fluoro[6]helicene and 4,13-difluoro[6]helicene molecules, with 68% and

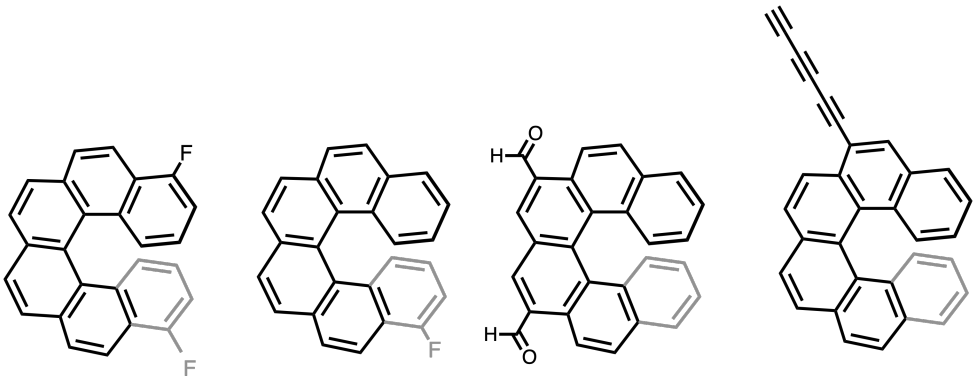

Figure 6: Four promising OSC molecular material candidates identified by computational screening. From left to right: 4,13-difluoro[6]helicene, 13-fluoro[6]helicene, 7,10-dialdehyde[6]helicene, and 6-triacetylene[6]helicene.

46% frequency, respectively. For the cases with bulkier side groups, 7,10-dialdehyde and 6-triacetylene[6]helicenes, the translational motif was observed less frequently (18% and 22%, respectively). Amongst the 10 lowest-energy crystals in each CSP landscape, the translational motif was observed 100% (4,13-difluoro), 70% (13-fluoro), 30% (6-triacetylene) and 30% (7,10-dialdehyde) of the time, compared to the [6]helicene reference (90%). Fluorination does not affect the overall shape of the molecule and, thus, similar packing motifs (translational in one plane and back-to-back in the other plane) are observed as for [6]helicene. However, for the flexible dialdehyde and the large asymmetric triacetylene side group, the single-molecule geometry is significantly different from the [6]helicene backbone. Thus, it is not surprising that other motifs are accessed, with the translational motif seen less frequently. Additionally, due to the altered single-molecule geometry, the exact overlap angle of the translational motif and other motifs is changed slightly. It is, unfortunately, not the case that the translational dimer motif is guaranteed to be found, but it is plausible that a structure containing the motif can be found via a targeted experimental polymorph screening, and it does remain a frequently occurring 1-dimensional packing motif in the helicene systems. From our previous CSP studies with helicenes, motifs frequently occurring in the calculated CSP landscapes are observed in experimentally obtained crystal packings[32] and explain the experimentally observed properties of these systems.[31]

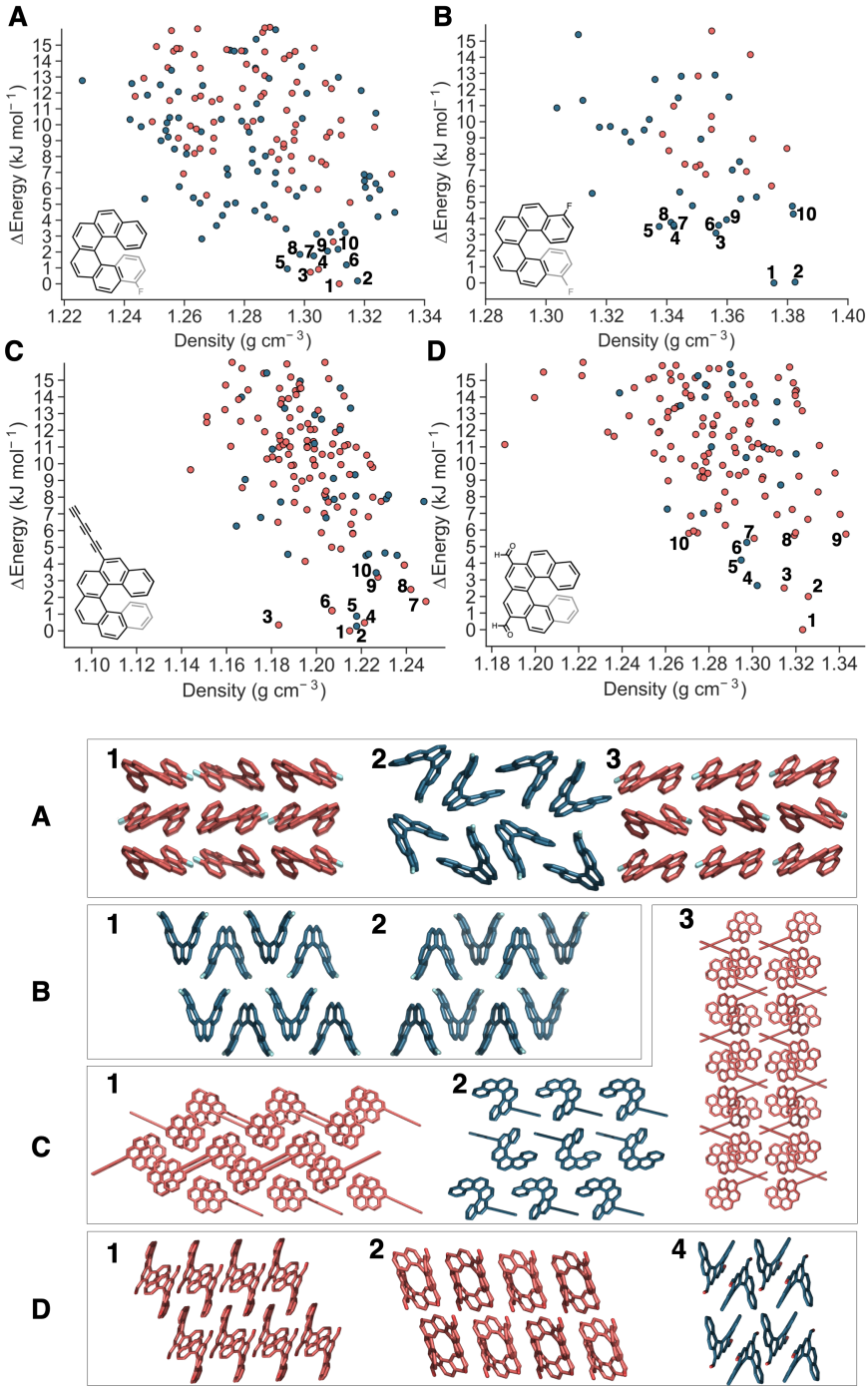

Figure 7: Top: Crystal-energy landscapes for 13-fluoro[6]helicene (**A**), 4,13-difluoro[6]helicene (**B**), 6-triacetylene[6]helicene (**C**) and 7,10-dialdehyde[6]helicene (**D**). Data points are shown in blue (red) if they do (do not) contain the translational dimer motif. Low-energy crystals are numbered in order of ascending relative energy from the global minimum. The electron mobilities for the 10 lowest-energy polymorphs are reported in Figures 8, 9 and S32. Bottom: low-energy crystal packings, corresponding to the labeled structures above. Labeled crystal packings not shown here can be found in Figure S30. Oxygen atoms are shown in red, and fluorine in light blue.

Figure 8 shows the electron mobilities across the 10 lowest-energy polymorphs of the global minimum for the CSP landscapes for the 4 molecules (*i.e.* those the most likely to be experimentally obtainable). Notably, in the case of fluorination, the translational-dimer motif gives rise to either the highest or second-highest transfer integral in the crystal polymorphs. Additionally, for both fluorinated [6]helicene CSP landscapes, the majority of the lowest-energy crystals contain the translational motif, while this motif is less prevalent in the higher-energy crystals. Across all 4 molecules, the translational motif coincides with comparably high mobilities. The distributions of the electron mobilities across the 10 lowest energy polymorphs (in all planes of each polymorph) for each of the 4 CSP systems and for carbo[6]helicene are given in Figure 9. Both 6-triacetylene and 7,10-dialdehyde[6]helicene have a mean and maximum electron mobility significantly lower than the bare backbone. This suggests that neither of these modifications to the backbone are likely to enhance charge transport. In contrast, both fluorination examples have a higher mean and maximum electron mobility than for the bare backbone. Thus, it is the fluorinated systems where we are most hopeful of the synthesis of a polymorph with an increased charge-carrier mobility compared to the bare [6]helicene backbone.

4,13-difluoro[6]helicene

For 4,13-difluoro[6]helicene, CSP found two polymorphs that are significantly lower in energy than the remaining candidate structures on the crystal-energy landscape (crystals **1** and **2** in Figure 7B), and both contain the targeted translational motif. In fact all 10 lowest-energy crystals contain the translational motif. The dimer model predicted a minimum-energy separation of 7.2 Å for 4,13-difluoro[6]helicene, with a transfer integral of J_{elec}=78 meV. These properties are in excellent agreement with the CSP results, where the translational dimer gave rise to J_{elec}=77 meV at an intermolecular distance of 7.3 Å for both low-energy polymorphs. This transfer intergral is of the same order of magnitude as the [6]helicene reference dimer (78 meV). For the lowest-energy crystal (**1**), the translational dimer gave rise to the highest

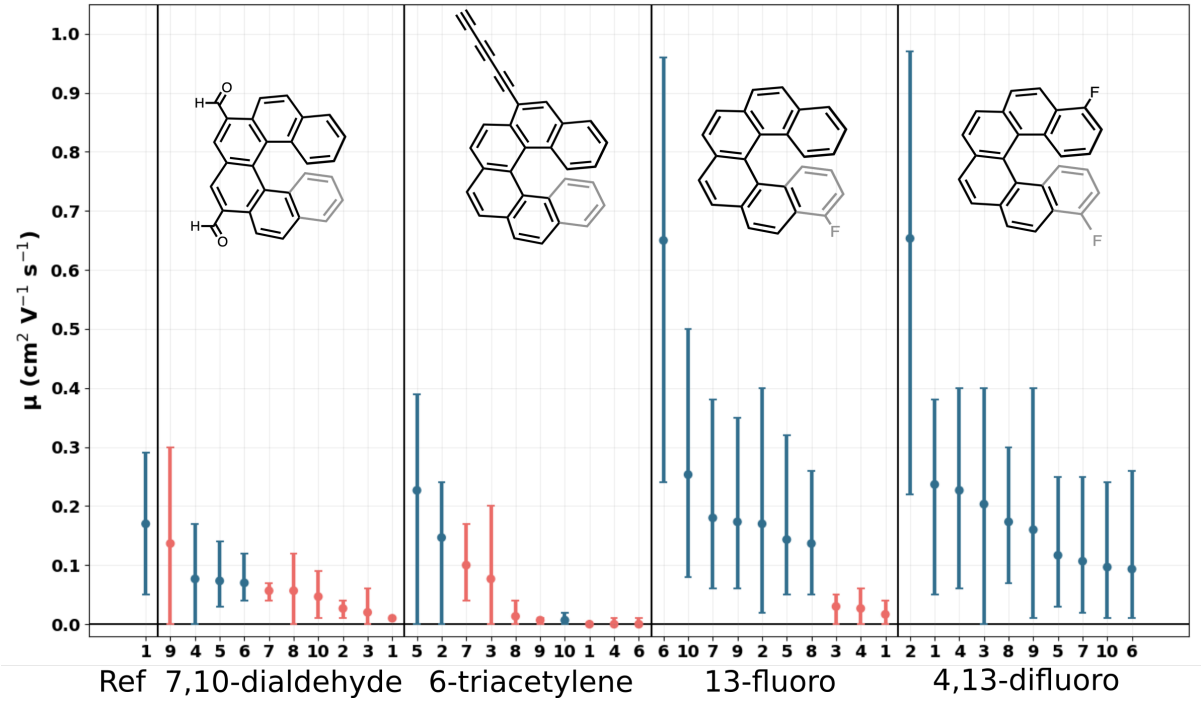

Figure 8: Minimum, maximum, and average electron mobilities, μ, computed for each of the 10 lowest-energy polymorphs identified from the CSP searches. Data for structures with (without) the translational motif are shown in blue (red). Results for the reference [6]helicene crystal (HEXHEL) are shown on the far left. The equivalent plot showing the 10 lowest energy polymorphs for each system is in Figure S32.

transfer integral observed. Conversely, in crystal **2**, the back-to-back dimer (Figure S1 VI,VII) produced a higher transfer integral, of 133 meV, giving rise to a much higher overall electron mobility for this material (Figure S32). This compound is a particularly promising OSC candidate, as both low energy polymorphs give mobilities higher than the unsubstituted [6]helicene reference. While form **2** gives a significantly higher mobility than form **1**, these polymorphs are predicted to be nearly degenerate and form **2** could potentially become the more stable were a higher level of theory used for energetic reranking of the CSP polymorphs.

13-fluoro[6]helicene

Monofluorination gave rise to 176 low-energy polymorphs within 15 kJ mol^{-1} of the global minimum for 13-fluoro[6]helicene (Figure 7A), which is over three times as many as found for 4,13-difluoro[6]helicene. The asymmetry of 13-fluoro[6]helicene, relative to the symmetrically

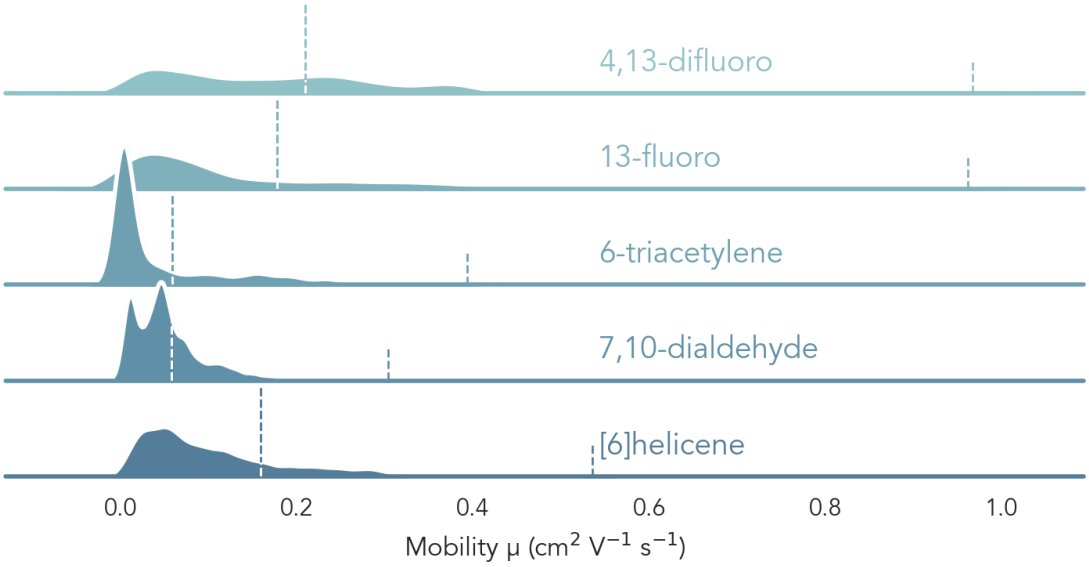

Figure 9: Distribution of the calculated electron mobilities observed across the ten lowest energy crystals for each molecule where CSP was carried out ([6]helicene results as previously reported in Ref. 32). Maximum and average electron mobilities are indicated with dashed lines. Fluorination has the potential to improve charge transport in [6]helicenes, whereas 6-triacetylene and 7,10-dialdehyde[6]helicene perform worse than [6]helicene. The mobility plots for individual crystals can be found in Figures 8 and S32.

disubstituted molecule, stochastically increases the number of permutations of homo and heterochiral crystalline packings possible. The majority of 13-fluoro[6]helicene chiral crystals contained the translational motif (Figure S31 I). Seven of the ten lowest-energy polymorphs contained the translational motif at an average separation of $\bar{d}$=7.35 Å. For the two lowest-energy crystals containing the translational motif, **2** and **5**, the intermolecular distances (7.3 Å) and maximum transfer intergrals (72 meV) for this motif are again in excellent agreement with the dimer screening prediction (d_{min}=7.3 Å and J_{elec}=71 meV).

The mobilities of the ten lowest-energy crystal structures for 13-fluoro[6]helicene reveal the influence of the translational motif on electron transport. The averaged maximum transfer integral of the polymorphs containing the translational motif (crystals **2**, **5-10**, $\bar{J}_{\mathrm{elec}}$=80 meV) is significantly higher than the average obtained for the crystals without the

translational motif (crystals **1**, **3** and **4**, $\bar{J}_{\text{elec}}$=24 meV). Similarly, the averaged maximum electron mobilities are significantly improved if the translational motif is present, with $\bar{\mu}_{\text{max}}$ computed to be 0.452 and 0.05 cm^2 V^{-1} s^{-1} for crystals with and without the translational motif, respectively. The crystal with the highest mobility is polymorph **6**, which has J_{max}=128 meV and $\mu_{\text{max}} = 0.96$ cm^2 V^{-1} s^{-1}, leading to a 231% increase in the charge-carrier mobility compared to the [6]helicene reference structure[33] (0.29 cm^2 V^{-1} s^{-1}). While form **6** is not the thermodynamically most stable of the polymorphs considered in the mobility calculations, this polymorph is <2 kJ mol^{-1} above the energetic minimum and may be isolable experimentally. Higher levels of theory, or inclusion of thermal free-energy corrections, could also substantially reorder the low-energy structures.

6-triacetylene[6]helicene

In this case, we can observe how our dimer screening model performs for helicenes with an altered overall molecular shape due to a large asymmetric perturbation to the backbone. In our initial screening, the 6-triacetylene[6]helicene translational dimer was found to have the shortest intermolecular distance (5.7 Å) of all compounds considered. CSP confirms this abnormally short intermolecular distance, predicting values of 5.5 Å for forms **2** and **5** (Table S3). However, due to the large side group, the translational dimer motif was present at a much lower frequency, occurring in 22% of the crystal structures. Where present, this motif did lead to promising charge-mobility, with transfer integrals of J_{elec}= 125 meV for form **2** and J_{elec}= 126 meV for form **5**, 76-77% higher than the [6]helicene reference (65 meV). This increase is potentially caused by the parallel π-π stacking of the triacetylene side groups and helicene backbone at short intermolecular distances.

Crystal structures **5** and **2**, containing the translational dimer, produce the highest μ_{max} and $\bar{\mu}$ across the 6-triacetylene polymorphs (Figure 9). The translational motif in crystal **6** gives rise to a lower mobility; however, as noted above, this crystal contains a heterochiral stacked dimer, with the side groups attached to opposite ends of the backbone, pointing in

opposite directions. Despite the short intermolecular distances, only form **5** gives a maximum mobility (0.39 cm^2 V^{-1} s^{-1}) higher than the [6]helicene reference, and this is significantly lower than the results obtained for both the mono- and difluoro[6]helicenes. Consequently, crystal mobilities indicate poor electron transport for this screening candidate.

As discussed earlier, the high J/λ values for this system are suggestive of the fact that it may not operate in the hopping regime, but more likely in the 'transient localisation' regime. Ideally, to understand the charge transport in this system, we would need to calculate the intermolecular vibrational modes to get an insight into thermal fluctuations that may reduce charge mobility. However, given there are already indications that this system is not promising at this point, the fact that these calculations are computationally demanding, and the CSP suggests multiple plausible low-energy polymorphs that would need to be considered, we do not believe this is a worthwhile effort, and certainly beyond the scope of this study.

7,10-dialdehyde[6]helicene

The CSP search for the flexible 7,10-dialdehyde[6]helicene molecule resulted in new packing motifs, most noticeably back-to-back dimers with the backbones overlapping to generate more of a network-type architecture (shown for form **1** in Figure 7D). Also, the 7,10-dialdehyde[6]helicene is altered by the orientation of the carbonyl oxygens with respect to the initial gas phase dimers. Thus, the translational motif is slightly altered, but still found in 18% of the crystal structures and 30% of the 10 lowest-energy structures. None of the ten lowest-energy crystals lead to improved results compared to [6]helicene. The 7,10-dialdehyde[6]helicene crystals perform even worse than the 6-triacetylene[6]helicene crystals, with lower average and maximum mobilities observed (Figures 9 and S32). Whilst the vast majority of crystals do not show any measurable electron transport, the translational dimer is only observed across polymorphs that exhibit some weak transport, with *e.g.* crystals **5** and **4** having maximum mobilities of 0.14 and 0.17 cm^2 V^{-1} s^{-1}, respectively, compared to crystals **1**, **2**, and **3**, which have very low mobilities in the range of 0.01-0.06 cm^2 V^{-1} s^{-1}.

4 Conclusions

In this work, 1344 monosubstituted and symmetrically disubstituted [6]helicenes were screened to assess their potential as organic semi-conducting (OSC) materials. The initial screening was based on properties of a specific homochiral translated molecular dimer, which is the most common motif found in low-energy polymorphs of carbo[6]helicene, and is also prevalent in other helicene structures present in the Cambridge Structural Database. This dimer-based approach is orders of magnitude cheaper computationally than performing full crystal structure prediction (CSP) workflows for all candidates. However, due to the simplifications required in a screening approach for a complex multivariable problem such as this, it could misclassify systems that do not present with the assumed packing motif, or it may lead to incorrect ranking of systems that lie outside the assumed hopping regime of charge transport. Nonetheless, the fluorine side group, predicted by the screening to be the most promising substituent to maximise OSC performance metrics overall, was found to have promising electron mobilities when the predicted crystal polymorphs were analysed.

In the 10 lowest-energy polymorphs of the CSP landscapes for the four most promising systems, the translational motif was found in 100% (4,13-difluoro), 70% (13-fluoro), 30% (6-triacetylene) and 30% (7,10-dialdehyde) of the structures. This is encouraging for the use of our dimer-based model, in that we could hope an experimental polymorph screening study could lead to the targeted motif. The intermolecular distances for the translational motif found in the lowest-energy crystal polymorphs were found to closely match the minimum-energy distances obtained from semi-empirical calculations on the isolated molecular dimers. Even the unusually shorter intermolecular distance (5.7 Å) in the 6-triacetylene[6]helicene dimer was confirmed through CSP. Although the translational-dimer motif did not necessarily yield the maximum transfer integral in the molecular crystals, its occurrence in low-energy polymorphs was strongly correlated to good electron transport properties. We accept that using J and J/λ as metrics for charge carrier mobility is a simplification and will be inadequate for strongly coupled systems. However, it serves as a first indicator of structures with promising

transport properties. To fully characterise structures identified as promising from the screening approach, a phonon calculation of the relevant crystal structure would be an obvious next step. The dimer assumption is not perfect, and the frequency with which it was found for different systems does underline that there is still work to be done in finding a more general screening approach that is computationally viable for screening more than 1300 molecular systems, while also trying to factor in the complex interplay between molecular structure, solid-state structure, and charge-transport properties in the area of organic semiconductors. We do propose, however, that our approach is an improvement upon completely ignoring solid-state packing effects and performing computational screening only at the single-molecule level.

Comparing the present results to those for other [6]helicene OSCs, where we have used the same methodology to calculate electron transport indicators,[31,32] one can clearly see that terminal substitution best improves the suitability of [6]helicenes for electron transport, while nitrogen substitution in aza[6]helicenes favours hole transport over electron transport. The maximum electron mobility calculated in aza[6]helicene crystals using our approach did not exceed 0.30 $cm^2\ V^{-1}\ s^{-1}$ (4-aza[6]helicene: 0.26 $cm^2\ V^{-1}\ s^{-1}$).[33] Here, computational screening resulted in the identification of new candidate molecules with maximum electron mobilities more than three times greater than this limit using the same approach. Notably, some, but not all, 13-fluoro and 4,13-difluoro[6]helicene polymorphs outperform the unsubstituted carbo[6]helicene reference structure, with maximum mobilities of 0.96, and 0.97 $cm^2\ V^{-1}\ s^{-1}$ respectively. Generally, of the four compounds carried forward from screening to CSP, the lowest-energy 4,13-difluoro[6]helicene polymorphs most consistently exceeded the reference. The symmetric 4,13-difluoro[6]helicene is our most promising identified OSC compound, with both low-energy polymorphs identified from the CSP posessing high electron mobilities; its synthesis should be feasible via a Wittig procedure starting from 2-fluorobenzaldehyde. Future work could include varying the degree of fluorination (tetra-, hexa- and octafluorination *etc.*) to see if this would further increase charge-carrier mobilities, similar to pentacenes.[70]

Overall, our screening approach to guide the design of new organic semiconducting molecular materials shows promise, even though it is short of a 'universal solution' in organic semiconductor screening. Structural classification of the intermolecular packing into motifs was shown to be a powerful tool in the design of new functionalised materials, capable of providing observable crystal structure-mobility relationships. We hope that this approach can be developed further to aid in the extremely complex problem of accelerating the discovery of molecular materials with targeted properties.

Conflicts of interest

There are no conflicts to declare.

Acknowledgement

We acknowledge funding from the European Research Council under FP7 (CoMMaD, ERC Grant No. 758370), and the Engineering Research Council and Physical Sciences Research Council (EPSRC) (EP/P005543/1), including the UKs HEC Materials Chemistry Consortium (EP/L000202/1) for time on the UK supercomputer, ARCHER. K.E.J. thanks the Royal Society for a University Research Fellowship and for an Enhancement Award 2017 (J.A.S.). J.N. acknowledges funding from the European Research Council (CAPaCITy, Action No. 742708). M.J.F. acknowledges funding from the Centre for Processable Electronics (EP/L016702/1) and his EPSRC fellowship (EP/R00188X/1). The authors acknowledge funding from the Natural Sciences and Engineering Research Council (NSERC) of Canada to E.R.J. and from the Government of Nova Scotia for a fellowship to J.A.W. We also thank Compute Canada for computational resources. I.J.S. the EPSRC (EP/J014958/1, EP/J003840/1, EP/P022561/1 and EP/P020194) and the support of Eli Lilly. The authors thank Dr. Javier Segarra-Martí, Dr. Louise Price and Prof. S. Claire Adjiman for useful discussions.

References

(1) Bässler, H.; Köhler, A. Charge transport in organic semiconductors. *Top. Curr. Chem.* **2012**, *312*, 1–65.

(2) Petty, A. J.; Ai, Q.; Sorli, J. C.; Haneef, H. F.; Purdum, G. E.; Boehm, A.; Granger, D. B.; Gu, K.; Rubinger, C. P. L.; Parkin, S. R.; Graham, K. R.; Jurchescu, O. D.; Loo, Y. L.; Risko, C.; Anthony, J. E. Computationally aided design of a high-performance organic semiconductor: The development of a universal crystal engineering core. *Chem. Sci.* **2019**, *10*, 10543–10549.

(3) Ruiz, C.; García-Frutos, E. M.; Hennrich, G.; Gómez-Lor, B. Organic semiconductors toward electronic devices: High mobility and easy processability. *J. Phys. Chem. Lett.* **2012**, *3*, 1428–1436.

(4) Kunkel, C.; Schober, C.; Margraf, J. T.; Reuter, K.; Oberhofer, H. Finding the Right Bricks for Molecular Legos: A Data Mining Approach to Organic Semiconductor Design. *Chem. Mater.* **2019**, *31*, 969–978.

(5) Nematiaram, T.; Padula, D.; Landi, A.; Troisi, A. On the Largest Possible Mobility of Molecular Semiconductors and How to Achieve It. *Adv. Funct. Mater.* **2020**, *30*, 2001906.

(6) Sorli, J. C.; Ai, Q.; Granger, D. B.; Gu, K.; Parkin, S.; Jarolimek, K.; Telesz, N.; Anthony, J. E.; Risko, C.; Loo, Y. L. Impact of Atomistic Substitution on Thin-Film Structure and Charge Transport in a Germanyl-ethynyl Functionalized Pentacene. *Chem. Mater.* **2019**, *31*, 6615–6623.

(7) Musil, F.; De, S.; Yang, J.; Campbell, J. E.; Day, G. M.; Ceriotti, M. Machine learning for the structure-energy-property landscapes of molecular crystals. *Chem. Sci.* **2018**, *9*, 1289–1300.

(8) Atahan-Evrenk, S.; Atalay, F. B. Prediction of Intramolecular Reorganization Energy Using Machine Learning. *J. Phys. Chem. A* **2019**, *123*, 7855–7863.

(9) Pulido, A.; Chen, L.; Kaczorowski, T.; Holden, D.; Little, M. A.; Chong, S. Y.; Slater, B. J.; McMahon, D. P.; Bonillo, B.; Stackhouse, C. J.; Stephenson, A.; Kane, C. M.; Clowes, R.; Hasell, T.; Cooper, A. I.; Day, G. M. Functional materials discovery using energy-structure-function maps. *Nature* **2017**, *543*, 657–664.

(10) Campbell, J. E.; Yang, J.; Day, G. M. Predicted energy-structure-function maps for the evaluation of small molecule organic semiconductors. *J. Mater. Chem. C* **2017**, *5*, 7574–7584.

(11) Motherwell, W. D.; Ammon, H. L.; Dunitz, J. D.; Dzyabchenko, A.; Erk, P.; Gavezzotti, A.; Hofmann, D. W.; Leusen, F. J.; Lommerse, J. P.; Mooij, W. T.; Price, S. L.; Scheraga, H.; Schweizer, B.; Schmidt, M. U.; Van Eijck, B. P.; Verwer, P.; Williams, D. E. Crystal structure prediction of small organic molecules: A second blind test. *Acta Crystallogr. Sect. B Struct. Sci.* **2002**, *58*, 647–661.

(12) Day, G. M.; Motherwell, W. D.; Ammon, H. L.; Boerrigter, S. X.; Della Valle, R. G.; Venuti, E.; Dzyabchenko, A.; Dunitz, J. D.; Schweizer, B.; Van Eijck, B. P.; Erk, P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W.; Leusen, F. J.; Liang, C.; Pantelides, C. C.; Karamertzanis, P. G.; Price, S. L.; Lewis, T. C.; Nowell, H.; Torrisi, A.; Scheraga, H. A.; Arnautova, Y. A.; Schmidt, M. U.; Verwer, P. A third blind test of crystal structure prediction. *Acta Crystallogr. Sect. B Struct. Sci.* **2005**, *61*, 511–527.

(13) Day, G. M.; Cooper, T. G.; Cruz-Cabeza, A. J.; Hejczyk, K. E.; Ammon, H. L.; Boerrigter, S. X.; Tan, J. S.; Della Valle, R. G.; Venuti, E.; Jose, J.; Gadre, S. R.; Desiraju, G. R.; Thakur, T. S.; Van Eijck, B. P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W.; Neumann, M. A.; Leusen, F. J.; Kendrick, J.; Price, S. L.; Misquitta, A. J.; Karamertzanis, P. G.; Welch, G. W.; Scheraga, H. A.; Arnautova, Y. A.;

Schmidt, M. U.; Van De Streek, J.; Wolf, A. K.; Schweizer, B. Significant progress in predicting the crystal structures of small organic molecules - A report on the fourth blind test. *Acta Crystallogr. Sect. B Struct. Sci.* **2009**, *65*, 107–125.

(14) Bardwell, D. A.; Adjiman, C. S.; Arnautova, Y. A.; Bartashevich, E.; Boerrigter, S. X.; Braun, D. E.; Cruz-Cabeza, A. J.; Day, G. M.; Della Valle, R. G.; Desiraju, G. R.; Van Eijck, B. P.; Facelli, J. C.; Ferraro, M. B.; Grillo, D.; Habgood, M.; Hofmann, D. W.; Hofmann, F.; Jose, K. V.; Karamertzanis, P. G.; Kazantsev, A. V.; Kendrick, J.; Kuleshova, L. N.; Leusen, F. J.; Maleev, A. V.; Misquitta, A. J.; Mohamed, S.; Needs, R. J.; Neumann, M. A.; Nikylov, D.; Orendt, A. M.; Pal, R.; Pantelides, C. C.; Pickard, C. J.; Price, L. S.; Price, S. L.; Scheraga, H. A.; Van De Streek, J.; Thakur, T. S.; Tiwari, S.; Venuti, E.; Zhitkov, I. K. Towards crystal structure prediction of complex organic compounds - A report on the fifth blind test. *Acta Crystallogr. Sect. B Struct. Sci.* **2011**, *67*, 535–551.

(15) Day, G. M.; Motherwell, W. D.; Jones, W. A strategy for predicting the crystal structures of flexible molecules: The polymorphism of phenobarbital. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1693–1704.

(16) Karamertzanis, P. G.; Pantelides, C. C. Ab initio crystal structure prediction. II. Flexible molecules. *Mol. Phys.* **2007**, *105*, 273–291.

(17) Beyer, T.; Lewis, T.; Price, S. L. Which organic crystal structures are predictable by lattice energy minimisation? *CrystEngComm* **2001**, *3*, 178–212.

(18) Hoja, J.; Ko, H. Y.; Neumann, M. A.; Car, R.; DiStasio, R. A.; Tkatchenko, A. Reliable and practical computational description of molecular crystal polymorphs. *Acta Hortic. Sin.* **2019**, *45*, eaau3338.

(19) Brandenburg, J. G.; Grimme, S. *Top. Curr. Chem.*; Springer, 2014; Vol. 345; pp 1–24.

(20) Leblanc, L. M.; Johnson, E. R. Crystal-energy landscapes of active pharmaceutical ingredients using composite approaches. *CrystEngComm* **2019**, *21*, 5995–6009.

(21) Nyman, J.; Day, G. M. Static and Lattice Vibrational Energy Differences Between Polymorphs. *CrystEngComm* **2015**, *17*, 5154–5165.

(22) Whittleton, S. R.; Otero-de-la Roza, A.; Johnson, E. R. Exchange-Hole Dipole Dispersion Model for Accurate Energy Ranking in Molecular Crystal Structure Prediction II: Non-Planar Molecules. *J. Chem. Theory Comput.* **2017**, *13*, 5332–5342.

(23) Sosorev, A. Y.; Maslennikov, D. R.; Kharlanov, O. G.; Chernyshov, I. Y.; Bruevich, V. V.; Paraschuk, D. Y. Impact of Low-Frequency Vibrations on Charge Transport in High-Mobility Organic Semiconductors. *Phys. status solidi (RRL)–Rapid Res. Lett.* **2019**, *13*, 1800485.

(24) Gryn'Ova, G.; Lin, K. H.; Corminboeuf, C. Read between the Molecules: Computational Insights into Organic Semiconductors. *J. Am. Chem. Soc.* **2018**, *140*, 16370–16386.

(25) Brandt, J. R.; Salerno, F.; Fuchter, M. J. The added value of small-molecule chirality in technological applications. *Nat. Rev. Chem.* **2017**, *1*, 45.

(26) Carpenter, J. E.; Grünwald, M. Heterogeneous interactions promote crystallization and spontaneous resolution of chiral molecules. *J. Am. Chem. Soc.* **2020**, *142*, 10755–10768.

(27) Kiran, V.; Mathew, S. P.; Cohen, S. R.; Hernández Delgado, I.; Lacour, J.; Naaman, R. Helicenes—A new class of organic spin filter. *Adv. Mater.* **2016**, *28*, 1957–1962.

(28) Isla, H.; Crassous, J. Helicene-based chiroptical switches. *Comptes Rendus Chim.* **2016**, *19*, 39–49.

(29) Jhulki, S.; Mishra, A. K.; Chow, T. J.; Moorthy, J. N. Helicenes as All-in-One Organic Materials for Application in OLEDs: Synthesis and Diverse Applications of Carbo-and Aza [5] helical Diamines. *Chem. Eur. J.* **2016**, *22*, 9375–9386.

(30) Yang, Y.; Da Costa, R. C.; Fuchter, M. J.; Campbell, A. J. Circularly polarized light detection by a chiral organic semiconductor transistor. *Nat. Photonics* **2013**, *7*, 634–638.

(31) Yang, Y.; Rice, B.; Shi, X.; Brandt, J. R.; Correa da Costa, R.; Hedley, G. J.; Smilgies, D.-M.; Frost, J. M.; Samuel, I. D. W.; Otero-de-la Roza, A.; Johnson, E. R.; Jelfs, K. E.; Nelson, J.; Campbell, A. J.; Fuchter, M. J. Emergent Properties of an Organic Semiconductor Driven by its Molecular Chirality. *ACS Nano* **2017**, *11*, 8329–8338.

(32) Rice, B.; LeBlanc, L. M.; Otero-De-La-Roza, A.; Fuchter, M. J.; Johnson, E. R.; Nelsona, J.; Jelfs, K. E. Correction: A computational exploration of the crystal energy and charge-carrier mobility landscapes of the chiral [6]helicene molecule. *Nanoscale* **2018**, *10*, 1865–1876.

(33) Salerno, F.; Rice, B.; Schmidt, J. A.; Fuchter, M. J.; Nelson, J.; Jelfs, K. E. The influence of nitrogen position on charge carrier mobility in enantiopure aza[6]helicene crystals. *Phys. Chem. Chem. Phys.* **2019**, *21*, 5059–5067.

(34) Szczypiński, F. T.; Bennett, S.; Jelfs, K. E. Can we predict materials that can be synthesised? *Chem. Sci.* **2020**,

(35) Jansen, M.; Schön, J. C. Rational development of new materials–putting the cart before the horse? *Nat. Mater.* **2004**, *3*, 838.

(36) Illig, S.; Eggeman, A. S.; Troisi, A.; Jiang, L.; Warwick, C.; Nikolka, M.; Schweicher, G.; Yeates, S. G.; Geerts, Y. H.; Anthony, J. E.; Sirringhaus, H. Reducing dynamic disorder in small-molecule organic semiconductors by suppressing large-amplitude thermal motions. *Nat. Commun.* **2016**, *7*, 1–10.

(37) Roos, K.; Wu, C.; Damm, W.; Reboul, M.; Stevenson, J. M.; Lu, C.; Dahlgren, M. K.; Mondal, S.; Chen, W.; Wang, L.; Abel, R.; Friesner, R. A.; Harder, E. D. OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *J. Chem. Theory Comput.* **2019**, *15*, 1863–1874.

(38) Becke, A. D. A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

(39) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian16 Revision C.01. 2016.

(40) Landrum, G. RDKit Documentation. *Read. Writ.* **2011**, *1*, 1–79.

(41) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB - An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.

(42) Neese, F. Software update: the ORCA program system, version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1327.

(43) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(44) Xie, X.; Santana-Bonilla, A.; Troisi, A. Nonlocal electron–phonon coupling in prototypical molecular semiconductors from first principles. *J. Chem. Theory Comput.* **2018**, *14*, 3752–3762.

(45) Sakanoue, K.; Motoda, M.; Sugimoto, M.; Sakaki, S. A molecular orbital study on the hole transport property of organic amine compounds. *J. Phys. Chem. A* **1999**, *103*, 5551–5556.

(46) Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. 2016; pp 785–794.

(47) Robert, C. Machine Learning, a Probabilistic Perspective . 2014.

(48) Rogers, D.; Hahn, M. Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742–754.

(49) Marcus, R. A. Chemical and electrochemical electron-transfer theory. *Annu. Rev. Phys. Chem.* **1964**, *15*, 155–196.

(50) Marcus, R. A. Electrostatic free energy and other properties of states having nonequilibrium polarization. I. *J. Chem. Phys.* **1956**, *24*, 979–989.

(51) Andrienko, D.; Kirkpatrick, J.; Marcon, V.; Nelson, J.; Kremer, K. Structure–charge mobility relation for hexabenzocoronene derivatives. *Phys. status solidi* **2008**, *245*, 830–834.

(52) Price, S. L.; Leslie, M.; Welch, G. W.; Habgood, M.; Price, L. S.; Karamertzanis, P. G.; Day, G. M. Modelling organic crystal structures using distributed multipole and polarizability-based model intermolecular potentials. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8478–8490.

(53) Soler, J. M.; Artacho, E.; Gale, J. D.; García, A.; Junquera, J.; Ordejón, P.; Sánchez-Portal, D. The SIESTA method for ab initio order-N materials simulation. *J. Phys. Condens. Matter* **2002**, *14*, 2745.

(54) Becke, A. D. On the large-gradient behavior of the density functional exchange energy. *J. Chem. Phys.* **1986**, *85*, 7184–7187.

(55) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(56) Johnson, E. R. In *Non-covalent Interact. Quantum Chem. Phys.*; Otero-de-la Roza, A., DiLabio, G. A., Eds.; Elsevier, 2017; Chapter 5, pp 169–194.

(57) Leblanc, L. M.; Weatherby, J. A.; Otero-De-La-Roza, A.; Johnson, E. R. Non-Covalent Interactions in Molecular Crystals: Exploring the Accuracy of the Exchange-Hole Dipole Moment Model with Local Orbitals. *J. Chem. Theory Comput.* **2018**, *14*, 5715–5724.

(58) Fratini, S.; Mayou, D.; Ciuchi, S. The transient localization scenario for charge transport in crystalline organic materials. *Adv. Funct. Mater.* **2016**, *26*, 2292–2315.

(59) Griffith, O. L.; Anthony, J. E.; Jones, A. G.; Lichtenberger, D. L. Electronic properties of pentacene versus triisopropylsilylethynyl-substituted pentacene: Environment-dependent effects of the silyl substituent. *J. Am. Chem. Soc.* **2010**, *132*, 580–586.

(60) Zeidell, A. M.; Jennings, L.; Frederickson, C. K.; Ai, Q.; Dressler, J. J.; Zakharov, L. N.; Risko, C.; Haley, M. M.; Jurchescu, O. D. Organic semiconductors derived from dinaphtho-fused s-indacenes: How molecular structure and film morphology influence thin-film transistor performance. *Chem. Mater.* **2019**, *31*, 6962–6970.

(61) Dong, H.; Wang, C.; Hu, W. High performance organic semiconductors for field-effect transistors. *Chem. Commun.* **2010**, *46*, 5211–5222.

(62) Wang, C.; Liu, Y.; Ji, Z.; Wang, E.; Li, R.; Jiang, H.; Tang, Q.; Li, H.; Hu, W. Cruciforms: assembling single crystal micro-and nanostructures from one to three dimensions and their applications in organic field-effect transistors. *Chem. Mater.* **2009**, *21*, 2840–2845.

(63) Zhang, B.; Kan, Y.-H.; Geng, Y.; Duan, Y.-A.; Li, H.-B.; Hua, J.; Su, Z.-M. An efficient strategy for improving carrier transport performance – Introducing fluorine into aryl substituted tetracene. *Org. Electron.* **2013**, *14*, 1359–1369.

(64) Tang, M. L.; Oh, J. H.; Reichardt, A. D.; Bao, Z. Chlorination: a general route toward electron transport in organic semiconductors. *J. Am. Chem. Soc.* **2009**, *131*, 3733–3740.

(65) Li, M.-J.; Long, M.-Q.; Chen, K.-Q.; Xu, H. Fluorination effects on the electronic transport properties of dithiophene-tetrathiafulvalene (DT-TTF) molecular junctions. *Solid State Commun.* **2013**, *157*, 62–67.

(66) Jones, B. A.; Facchetti, A.; Wasielewski, M. R.; Marks, T. J. Tuning orbital energetics in arylene diimide semiconductors. Materials design for ambient stability of n-type charge transport. *J. Am. Chem. Soc.* **2007**, *129*, 15259–15278.

(67) Di Donato, E.; Fornari, R. P.; Di Motta, S.; Li, Y.; Wang, Z.; Negri, F. n-Type charge transport and mobility of fluorinated perylene bisimide semiconductors. *J. Phys. Chem. B* **2010**, *114*, 5327–5334.

(68) Maiti, B.; Schubert, A.; Sarkar, S.; Bhandari, S.; Wang, K.; Li, Z.; Geva, E.; Twieg, R. J.; Dunietz, B. D. Enhancing charge mobilities in organic semiconductors by selective fluorination: a design approach based on a quantum mechanical perspective. *Chem. Sci.* **2017**, *8*, 6947–6953.

(69) Koch, N.; Vollmer, A.; Duhm, S.; Sakamoto, Y.; Suzuki, T. The effect of fluorination on pentacene/gold interface energetics and charge reorganization energy. *Adv. Mater.* **2007**, *19*, 112–116.

(70) Kleemann, H.; Schuenemann, C.; Zakhidov, A. A.; Riede, M.; Lüssem, B.; Leo, K. Structural phase transition in pentacene caused by molecular doping and its effect on charge carrier mobility. *Org. Electron.* **2012**, *13*, 58–65.

(71) Tang, M. L.; Reichardt, A. D.; Wei, P.; Bao, Z. Correlating carrier type with frontier molecular orbital energy levels in organic thin film transistors of functionalized acene derivatives. *J. Am. Chem. Soc.* **2009**, *131*, 5264–5273.

(72) Purdum, G. E.; Telesz, N. G.; Jarolimek, K.; Ryno, S. M.; Gessner, T.; Davy, N. C.; Petty, A. J.; Zhen, Y.; Shu, Y.; Facchetti, A.; Collis, G. E.; Hu, W.; Wu, C.; Anthony, J. E.; Weitz, R. T.; Risko, C.; Loo, Y. L. Presence of Short Intermolecular Contacts Screens for Kinetic Stability in Packing Polymorphs. *J. Am. Chem. Soc.* **2018**, *140*, 7519–7525.

(73) Mottishaw, J. D.; Sun, H. Effects of Aromatic Trifluoromethylation, Fluorination, and Methylation on Intermolecular π–π Interactions. *J. Phys. Chem. A* **2013**, *117*, 7970–7979.

(74) Shen, Y.; Chen, C.-F. Helicenes: synthesis and applications. *Chem. Rev.* **2012**, *112*, 1463–1535.

(75) Nematiaram, T.; Troisi, A. Strategies to reduce the dynamic disorder in molecular semiconductors. *Mater. Horiz.* **2020**, *7*, 2922–2928.