PAPER

Check for updates

Computational screening of bioinspired mixed ionic-electronic conductors

Tristan Stephens-Jones [ID] and Micaela Matta [ID] *

In recent years, organic mixed conducting polymers and small molecules have shown great potential in bioelectronics, neuromorphic devices and transient electronics. Current mixed conducting materials are mostly derived from pre-existing semiconductors functionalised with polar ethylene glycol side chains; however, these materials still exhibit limited biocompatibility and degradability. Here, we develop a computational/*in silico* screening pipeline to investigate the potential of bioinspired building blocks as next-generation materials for organic mixed ionic-electronic conductors (OMIECs). Leveraging sustainable design principles and predictors for electronic charge transport and aggregation/conformational order, we compare two approaches to discover potential new mixed conductors: a computational funnel and a genetic algorithm. We apply and evaluate both approaches against a chemical design space created by matching conjugated fragments from the literature on organic semiconductors, hydrolysable linkers and bioinspired fragments, for a total of almost 25 000 unique combinations. Our study demonstrates that, despite the limited chemical diversity of our dataset, both approaches successfully discover many potential donor–linker–acceptor (D–L–A) systems with promising features, namely: low HOMO–LUMO gap, high inter-ring planarity, and low reorganisation energy. We then down-select a few promising D–L–A systems and symmetrically extend their conjugation to obtain small-molecule prototypes, which show competitive reorganisation energies (as low as 123 meV). We propose that this workflow could be applied to larger datasets and tailored to discover novel chemical motifs for OMIECs and other applications.

Introduction

Organic mixed ionic-electronic conductors (OMIECs) are π-conjugated small molecules, oligomers or polymers capable of conducting ionic and electronic charges.[1] While two-component OMIECs such as PEDOT : PSS are made of separate semiconducting and electrolyte species, single-component OMIECs feature both a conjugated backbone (for efficient electronic charge transport) and polar or charged groups (to promote ion doping).[2–4]

The rapid development of transient, edible or compostable electronic devices[5] is generating high demand for low-cost, sustainable, environmentally benign and (bio)degradable organic semiconductors.[6–11] Thus, semiconducting materials that satisfy at least some of the 4 Bs (biosourced, bioderived, biodegradable and/or bioresorbable), *e.g.* featuring bioderived conjugated groups or hydrolysable linkers have begun to garner attention in organic electronics.[10]

In the field of OMIECs, eumelanin remains one of the few biomaterials with demonstrated protonic/electronic conductivity.[12–17] Eumelanin is a biomaterial obtained by the oxidative polymerisation of 5,6-dihydroxyindole (DHI) and 5,6-dihydroxy-indole-2-carboxylic acid (DHICA), known for its broadband optical absorption and antioxidant activity.[12]

Computational material screening approaches have been used extensively to find new semiconductor materials with
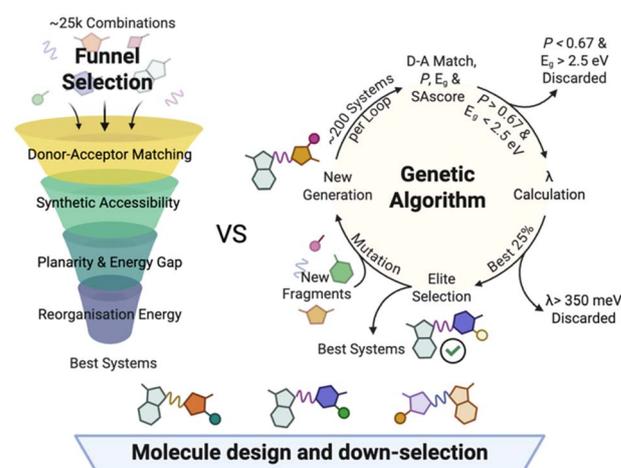

Fig. 1 Computational funnel (left) and genetic algorithm (right) approaches compared in this study, showing the key descriptors used in the screening process.

King's College, Department of Chemistry, Strand Campus, East Wing, WC2R 2LS, UK. E-mail: micaela.matta@kcl.ac.uk

10346 | *J. Mater. Chem. A*, 2026, **14**, 10346–10354

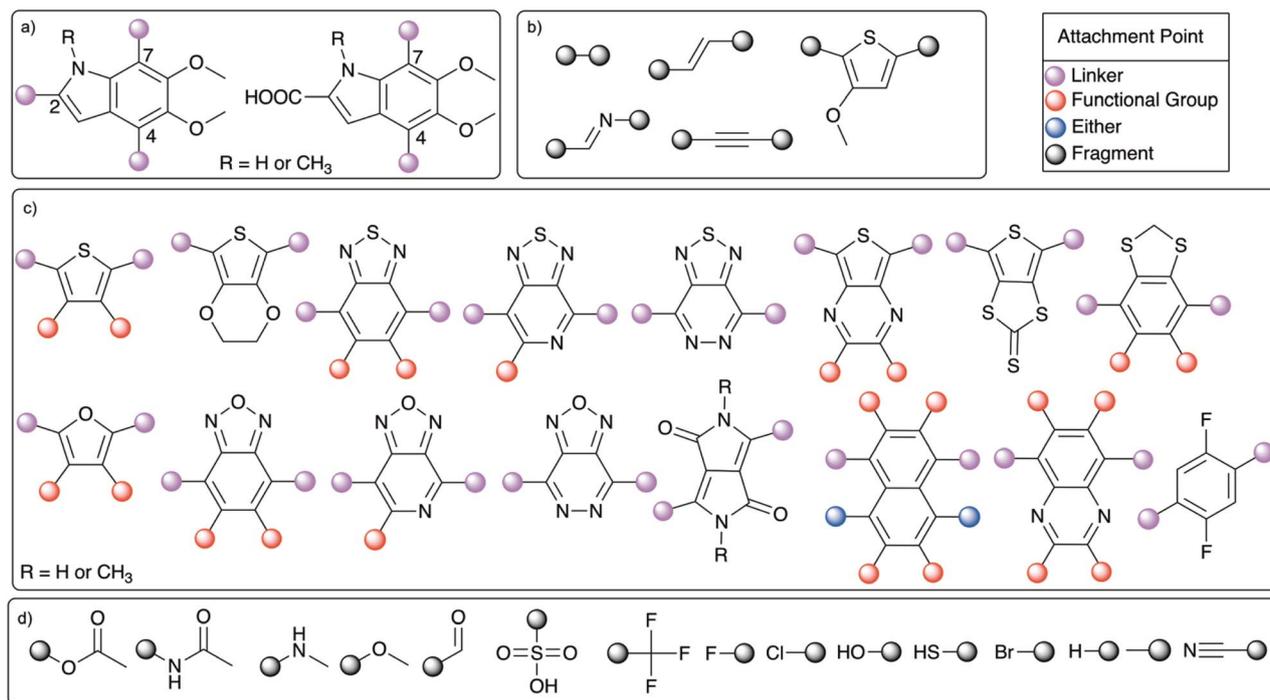This journal is © The Royal Society of Chemistry 2026

Fig. 2 List of heteroaromatic fragments, linkers and functional groups combined in this work, with their attachment points highlighted by spheres. (a) Eumelanin-inspired fragments derived by DHI and DHICA monomers; (b) linkers used to join fragments from (a) and (c); (c) fragments taken from literature on organic semiconductors; (d) functional groups used to functionalise the π-conjugated fragments from (c).

desirable optoelectronic properties,[18–21] however they have yet to be leveraged to discover novel mixed conductors incorporating sustainable design principles.

In this work, we propose a computational screening pipeline that repurposes and combines existing descriptors towards the design and screening of bioinspired mixed conductors (Fig. 1). As a proof-of-concept, we design new mixed conductors by pairing 5 eumelanin-inspired (3 DHI- and 2 DHICA-derived) building blocks with 17 π-conjugated, heteroaromatic fragments selected through a comprehensive literature search (top-down). The latter are functionalised with 15 different functional groups, yielding a total of 744 possible isolated donor–acceptor (D–A) pairs (Fig. 2). These pairs are then combined using 5 different linkers selected to maintain conjugation and promote system degradation e.g. via hydrolysis. This results in a chemical space of 24 990 possible donor–linker–acceptor (D–L–A) systems.

As more D–L–A systems are filtered out of the funnel, the computational cost of the calculated descriptors increases. Using a D–A design principle to pair the fragments based on their energy level matching; we then filter promising systems by revaluating (i) their synthetic accessibility, (ii) energy gap (E_g), (iii) inter-ring planarity (P) and (iv) internal reorganisation energy (λ_i).

Results and discussion

Donor–acceptor matching

The D–A design, widely used in organic semiconductors,[22–25] allows a low HOMO–LUMO gap, E_g by selecting donor and

acceptor fragments with energy level matching that falls within the range of organic semiconductors.[26,27] For a D–A system, E_g can be approximated as the difference between the HOMO and LUMO of isolated fragments:[26]

$$E_{g,est} \cong D_{HOMO} - A_{LUMO} \quad (1)$$

This eliminates the need of evaluating E_g for all possible D–A systems, reducing D–A pairing to performing a single optimisation on each fragment separately. The D–A matching prediction is performed at a low computational cost (B3LYP/3-21G*), which is sufficient to achieve a reliable result[28] without the risk of excluding promising candidates.

Due to DHI and DHICA having a shallow HOMO and higher LUMO (Fig. 3a), eumelanin-inspired fragments generally behave as donors when paired with other π-conjugated fragments (Fig. 3b). Hence, in this study the donor fragment is always eumelanin-inspired, paired with a range of different acceptor fragments.

The E_g range for OMIECs operating in aqueous electrolytes is 1–2.5 eV, as this is the electrochemical window of water.[29] Predicting $E_{g,est}$ from isolated fragments can potentially fail to account for increased conjugation or conformational/steric effects introduced when joining D and A fragments with different linkers. Therefore, at this stage a more permissive filter (0.5 < $E_{g,est}$ < 3.2 eV) is applied.

Out of 714 potential D–A pairs, 403 showed suitable $E_{g,est}$. These are then joined, using a series of 5 linkers, to form 14 363 unique D–L–A systems. Single, double or triple bonds, as well as imine and methoxythiophene linkers are selected with the aim
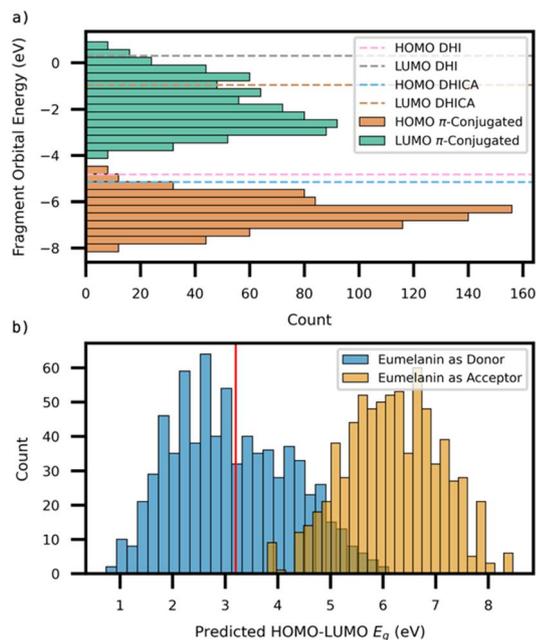
Fig. 3 (a) Distribution of HOMO and LUMO energies across all fragments in the dataset. Dashed lines indicate the frontier orbitals of DHI and DHICA. (b) $E_{g,est}$ distribution; the red line indicates the 3.2 eV threshold used to filter the dataset.

to maintain or increase conjugation length and planarity. Without this initial D–A matching, the total number of possible D–L–A systems would be 24 990: this initial filter thus reduces the total number of geometry optimisations needed at the next stage by 42.5% and only requires the geometry optimisation of 359 fragments to reduce the dataset by 10 627 systems.

Synthetic accessibility prediction

All the fragments in our dataset are – in isolation – synthetically available, however some of the functionalised fragments or D–L–A systems might be potentially challenging to synthesise. To focus our investigation on the most accessible systems, we calculate their Synthetic Accessibility score (SAscore).[30,31] The SAscore ranges from 1–10, where 1 denotes the highest synthesisability and 10 the lowest; we used a SAscore cutoff of 5. The SAscore distribution of our eumelanin-inspired dataset shows good predicted synthetical accessibility, with a range between 1.8–3.8 (Fig. S2), thus no systems are excluded on the basis of their SAscore alone. As the chemical space explored in this work includes only one type of bioinspired fragment, SAscore has limited predictive power; we anticipate that the synthesisability filter will gain importance in our workflow once our dataset is expanded to include a wider range of chemistries.

Planarity threshold and E_g with different linkers

In organic semiconductors, conformational order and inter-ring planarity play a key role in increasing conjugation, π stacking order and charge transport.[30] Electronic transport proceeds via inter-chain and intra-chain mechanisms, with intra-chain transport relying on orbital delocalisation between

neighbouring monomers and inter-chain relying on effective π-stacking.[2,31,32] Planar, π-conjugated systems can achieve high inter- and intra-chain charge mobility, due to larger carrier delocalisation lengths, lower band-gaps and stronger π–π interactions.[33]

Following from the work of Y. Che et al.,[34] we define an inter-ring planarity descriptor as:

$$P = \cos^2 \theta \qquad (2)$$

where $P = 1$ corresponds to a torsional angle θ of 0° or ±180°, which maximises the degree of π-orbital overlap along the conjugated backbone.[35] A P threshold of 0.67, corresponding to $\theta = 35°$, was chosen to filter the dataset. This threshold, which roughly corresponds to the optimum dihedral angle between two thiophene rings in gas phase,[36] was chosen because E_g of OMIECs has been shown to increase monotonically from 0° to 90°, with 38.9° corresponding to an inflection point.[34,37] In the case of a single bond linker, θ is defined as the D–A dihedral angle (see Fig. S1a). For all other linkers, θ is calculated as the average between the D–L and the L–A dihedrals (see Fig. S1b).

The geometry optimisation of the 14 363 D–L–A systems provides information on both P and E_g. Fig. 4 shows that introducing different conjugated linkers is an effective strategy to increase P and conjugation in D–L–A systems, resulting in a lower E_g. Linkers are also effective at minimising steric hindrance caused by side groups on neighbouring fragments. 3354 systems with $P > 0.67$ and $E_g \leq 2.5$ eV are retained through the funnel and selected for reorganisation energy calculation.

The imine linker appears in most of the systems with low E_g and high P, that is 36.8% of the systems selected for reorganisation energy calculations. The second most frequent linker was thiophene, found in 25.6% of the filtered systems. In general, all linkers that decrease steric clashes between the donor and acceptor fragments by increasing their separation show higher P and lower E_g. Both imine and thiophene linkers can also induce noncovalent interactions (e.g., N⋯H or S⋯O), which may also contribute to increase P.
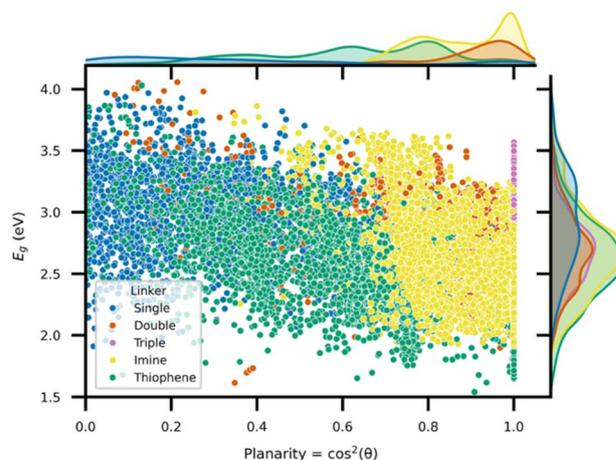

Fig. 4 Effect of five different linkers on the planarity score and E_g. Of the 14 363 D–L–A systems surveyed at this stage, 3354 of them fit within the filter regime of $P \geq 0.67$ and $E_g \leq 2.5$ eV.

10348 | J. Mater. Chem. A, 2026, 14, 10346–10354

This journal is © The Royal Society of Chemistry 2026

We then examine the most common bonding patterns across the best performing D–L–A systems. 48.9% of them feature DHI bonded at position 2 of the indole; this can be explained by the fact that this position is the furthest away from the methoxy side chains bonded in the 5 and 6 position of DHI and DHICA, therefore greatly reducing the possibility of steric clash. As DHICA bears a carboxylic group in position 2, its presence in the potential reorganisation systems is greatly reduced (17.1%) in comparison to DHI. Position 2 is reported to be the easiest to functionalise.[38]

Reorganisation energy

The internal reorganisation energy (λ_i) is related to the change in energy and geometry when a system becomes a charged species. λ_i is a key predictor of (and inversely proportional to) electronic charge mobility.[39–42] The cationic and anionic internal reorganisation energies, λ_i+ and λ_i^-, can inform on the tendency of each D–L–A system to behave either as a p-type or n-type semiconductor, respectively; we consider promising systems those with either λ_i+ or λ_i^- falling below the 250 meV threshold, E_λ. Internal reorganisation energies (we henceforth omit the subscript) were calculated using Nelson's 4-point method (see SI) at the B3LYP/6-31G* level of theory (see Computational Methods).[43,44]

Fig. 5 shows the reorganisation energy data for the remaining 3350 D–L–A systems. Out of the 133 systems with $\lambda^{+/-} < 250$ meV: 114 have $\lambda^- < E_\lambda$ and $\lambda^+ > E_\lambda$ and can thus be considered as potential n-type semiconductors; 11 have potential p-type behaviour with $\lambda^+ < E_\lambda$; 8 show potential ambipolar behaviour, having both $\lambda^+ < E_\lambda$ and $\lambda^- < E_\lambda$.

We then examine the makeup of our 133 surviving D–L–A systems in terms of linker and functional group occurrences (see Fig. S9). Both double bond and thiophene linkers appear equally, each making up 45 of the selected systems, so total 90 of the 133 surviving systems. Imine bond and triple bond linkers account for 27 and 16 of the total, respectively. No single bond linkers appear amongst the top performing D–L–A systems.
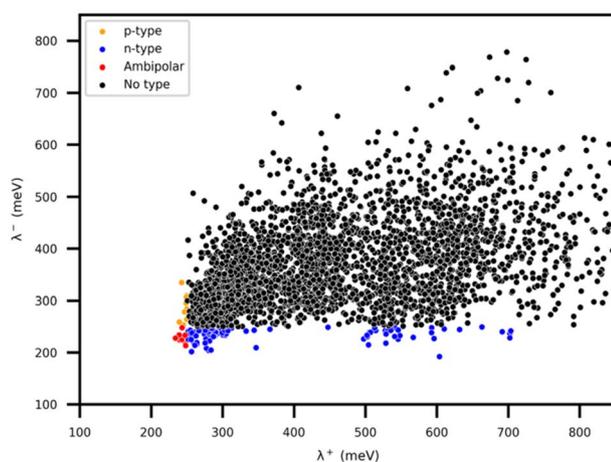

Fig. 5 Reorganisation energies of 3350 D–L–A systems. 133 fall within the E_λ threshold. Of these, 11 are labelled p-type ($\lambda^+ \leq E_\lambda$ orange), 114 n-type ($\lambda^- \leq E_\lambda$, blue) and 8 ambipolar ($\lambda^+ \leq E_\lambda$ and $\lambda^- \leq E_\lambda$, red). 3217 systems fall outside of the filter range (black).

Functional group variety in the dataset decreases sharply when looking at systems with low $\lambda^{-/+}$. The nitrile (15.4%), carbonyl (14.1%) and sulfonic acid (12.5%) functional groups are the most prevalent in the 3350 structures selected for the λ calculation (3 structures failed to converge). After the final filter, most of the D–L–A systems (66.9%) contain a nitrile group: of these, 79 are n-type and 3 are p-type (Fig. S11).

The significant decrease in the variety of functional groups, and the prevalence of n-type candidates amongst the 133 top-performing D–L–A systems can mainly be attributed to the lack of diversity in the original dataset. Only acceptor fragments with strong electron withdrawing substituents and deep LUMO can form a suitable match with melanin donors in terms of achieving the target E_g; when taking into account reorganisation energy, we observe an even stronger bias towards nitrile functional groups, which are well-known to decrease λ^-.[45]

In summary, the funnel takes 24 990 potential D–L–A systems and finds 133 potential D–L–A systems that have promising electronic properties ($E_g \leq 2.5$, $P \geq 0.67$ and $\lambda^{-/+} \leq 250$ meV) for use as small-molecule OMIECs. The funnel approach was useful to test how geometric and electronic descriptors can be used to filter a new chemical space; however, it still requires a large amount of costly DFT calculations – particularly when dealing with much larger datasets than the one being tested here. Therefore, we propose using a genetic algorithm (GA) to explore the chemical space more efficiently.[46,47] GAs have been used extensively in drug discovery, but have yet to be applied to OMIECs discovery.[46,48–56]

Genetic algorithm (GA) approach

The GA iterates over a set of D–L–A systems and tries to find the top performing structures by randomly mutating the best systems within each loop; the GA flowchart is shown in Fig. 6a.

Starting from the same combinatorial chemical space explored in the funnel (Fig. 2), the GA randomly assembles a pool of 201 D–L–A systems to be evaluated. Using the HOMO and LUMO energies of the isolated monomers, D–A matching is carried out, as in the funnel, using the same 3.2 eV threshold. Any D–A pairs that pass this cutoff are then optimised as a complete D–L–A system, to obtain P, E_g and SAscore.

For each D–L–A system, these 3 descriptors are assigned a relative ranking between 1 (best) and 201 (worst). The systems are then ranked according to the total cost function:

$$C_{elite} = C_P + C_{Eg} + 0.25 \cdot C_{SA} \qquad (3)$$

where C_P, C_{Eg} and C_{SA} are the rankings for P, E_g and SAscore respectively. Weights of 1 are applied to P and E_g and 0.25 to SAscore (see SI Fig. S14).

The top 25% ranked systems which also satisfied the cutoffs of $P \geq 0.67$ and $E_g \leq 2.5$ eV is then selected for reorganisation energy calculation. Any system with $\lambda^{-/+} \leq 350$ meV at this stage is defined as the 'elite' for the current loop, and stored with the elites of previous loops. A more permissive threshold (higher than the desirable range of ≤ 250 meV) was chosen, to account for the effect of random mutations which decrease the $\lambda^{-/+}$ in successive loops. As done with the other descriptors,
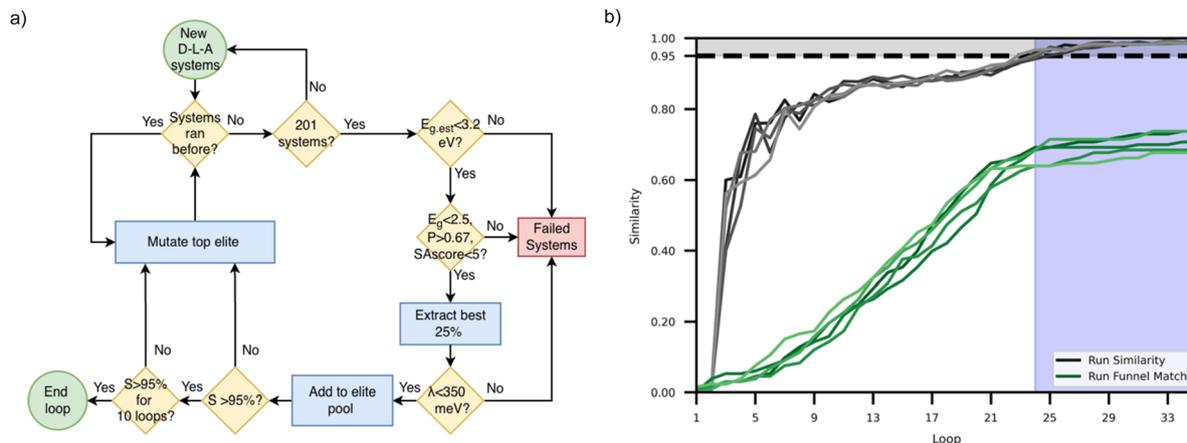
a)

b)

Fig. 6 (a) Genetic algorithm workflow. (b) Similarity as a function of the number of loops for 5 independent GA runs. The grey lines indicate the loop S for the 5 GA runs, the green lines indicate the S match of the systems discovered with a $P > 0.67$, $E_g < 2.5$ eV and $\lambda^{-/+} < 250$ meV compared to the total found in the funnel. The black dotted line at 0.95 shows the S threshold. The blue box highlights the average 10 consecutive loops where $S > 0.95$.

reorganisation energies are then assigned a relative ranking between 1 (lowest, best) and 50 (highest, worst).

The elite mutation stage gathers all the elite systems with $P > 0.67$, $E_g \leq 2.5$ eV, and $\lambda^{-/+} \leq 350$ meV and ranks them using the cost function:

$$C_{\text{mut}} = C_{\text{elite}} + 2 \cdot (C_\lambda^- + C_\lambda^+) \tag{4}$$

where C_λ^- and C_λ^+ denote the ranking coefficients assigned to λ^- and λ^+. Greater weight was applied to the $\lambda^{-/+}$ cost function to reflect its importance for the final system selection (see also S14b). The GA then randomly mutates either the donor, acceptor or linker for each elite; all have an equal chance of being selected for mutation. In the early stages of the GA, the number of such elite systems is below the chosen threshold of 20; in order to keep the GA run size constant at 201, a variable number of novel systems is added to the mutated elites for each loop. The novelty of each new system is determined by comparing its canonical SMILES to those of previously calculated systems. If the system has not previously appeared, it is then added to the population.

Loop similarity, overlap and efficiency

The GA was tested on 5 independent runs, each starting from a random set of 201 D–L–A systems. Overlap, O, is defined as:

$$O = \frac{R_X \cap R_Y}{R_X} \tag{5}$$

where R_X and R_Y denote 2 starting sets for 2 different GA runs, X and Y. Table S15 reports that the largest O is 0.025, showing that the starting point of the 5 GA runs is significantly different. This demonstrates that the GA can discover the same top performing structures from significantly different starting points.

The similarity metric, S, is used to calculate how similar the current loop's elite is to the previous one. S is defined as:

$$S = \frac{E_{t-1} \cap E_t}{E_t} \tag{6}$$

where E_t and E_{t-1} represent elite systems found in the current loop t and the previous loop, $t − 1$. The GA continues until S plateaus at 0.95 or higher for 10 consecutive runs.[46]

Overall, each GA run discovered on average 94 D–L–A systems satisfying the conditions: $P \geq 0.67$, $E_g \leq 2.5$ eV and $\lambda^{-/+} \leq E_\lambda$. These top performers were 97 for run 1, 94 for run 2, 91 for run 3, 98 for run 4 and 89 for run 5 (green lines in Fig. 6b). The overlap for the top-performing structures is shown in Fig. S18, including the overlap with the 133 structures identified in the funnel approach. The average O between each of the 5 GA runs and the funnel is 0.705. A total of 82 systems appear in all 5 GA top performers and the funnel, which represents 61.65% of the top-performing systems discovered *via* the funnel approach.

The computational efficiency, ϵ, of each stage of the GA is calculated as:

$$\varepsilon = \left(\frac{n_{\text{succ}}}{n_{\text{tot}}}\right) \times 100 \tag{7}$$

where n_{succ} is the number of systems calculated by the funnel or GA that meets the filter criteria, while n_{tot} is the total number of systems calculated at that stage. The funnel has a 1.5% higher ϵ for P, E_g and SAscore calculations, however the benefit of the GA approach is apparent when considering $\lambda^{-/+}$ calculations. The results in Table S19 show that the GA has an average ϵ of 8.2% for the most computationally demanding stage of the screening, compared to the funnel approach (4.0%).

Systems discovered

While the combined 5 GA repeats were able to find 113 out of 133 best performers discovered by the funnel, 20 systems do not appear in any of the GA repeats but appear in the funnel (see Fig. S20). For the rest of this study, we focus only on the 82 D–L–A systems that are discovered by all GA iterations and the funnel; 71 are classified as n-type, 6 as p-type, and 5 as ambipolar.

The 'missing' 20 systems can be attributed to the SAscore parameter, which was weighted 0.25 in the GAs. When

10350 | *J. Mater. Chem. A*, 2026, **14**, 10346–10354

This journal is © The Royal Society of Chemistry 2026

Table 1 Design strategy, HOMO−LUMO gap E_g and reorganisation energies λ^+ and λ^- for 12 small molecule candidates designed from 7 of the 82 top performers. The molecule design motifs highlight the alternance of donor (D), linker (L, L1, L2) and acceptors (A), while colors are used to highlight identical D−L−A parent systems

ID	Design	E_g (eV)	λ^+ (meV)	λ^- (meV)
1	D-L-A-L-D	1.50	367	198
2	D-L1-A-L2-A-L1-D	1.59	287	139
3	D-L-A-L-D	1.51	304	173
4	D-L1-A-L2-A-L1-D	1.51	206	137
5	D-L-A-L-D	1.61	437	199
6	D-L-A-L-L-A-L-D	1.44	252	156
7	D-L-A-L-D	1.66	261	192
8	D-L-A-L-L-A-L-D	1.51	213	147
9	D-L1-A-L2-A-L1-D	1.62	204	144
10	D-L-A-L-D	1.61	398	194
11	D-L-A-L-L-A-L-D	1.48	212	123
12	D-L-A-L-D	1.87	412	210

examining the funnel, the average SAscore for the top performing 133 systems is 3.18. For the 82 systems found across all 5 GA repeats, the average SAscore is 3.15. In contrast, for the 20 missing structures that none of the GA runs identified, the average SA score is 3.28. This result shows that the GA weighting biases top performing systems for better synthesisability, however at the cost of reducing the number of top performers discovered.

Additionally, the discrepancy in discovered systems can also be attributed to the 25% threshold on top performers in the GA. This constraint only takes the top 25% of systems from each loop, so it may cause promising systems to be 'missed'. However, each GA still discovers >60% of the potential top performers with considerably fewer calculations, demonstrating the strength of this approach in future studies.

Down-selected D−L−A systems

All of the 82 D–L–A systems discovered by both the screening funnel and GA are bonded at position 2 of the indole, and thus do not lend themselves to polymerisation. Out of these, 7 were selected for further study (highlighted in Fig. S21), extended and symmetrised to design potential small-molecule mixed conductors. Side chain engineering is known to affect the ionic/electronic conductivity,[57–60] solubility, processability and aggregation behaviour of OMIECs and, generally, organic semiconductors;[28,61–66] in this study, methoxy groups were used to mimic the functionalisation with polar oligoethylene glycol side chains as commonly seen in OMIEC design.[67]

The design motif, E_g, λ^+ and λ^- of each molecule are reported in Table 1. Molecules 1–11 maintain the 3-methoxythiophene linkers, while 2, 4, 6, 9 and 11 feature additional linkers (2 and 9: triple bond; 6, 8, 11: methoxybithiophene; 10: thiophene) to

force P onto the system, increase conjugation, and further reduce E_g and $\lambda^{-/+}$. 12 features an imine linker.

Across the parent D–L–A systems discovered, no structure has $\lambda^{-/+}$. <200 meV (see Table S22). The 12 molecules derived from these systems all show markedly lower E_g, λ^- and λ^+, with 11 out of 12 of them having at least one of λ^- or λ^+ < 200 meV. This confirms symmetric conjugation extension as an effective strategy to lower reorganisation energy. Notably, these results shows that our reorganisation energy threshold on smaller D–L–A systems can effectively be used to discover promising molecular designs. Interestingly, 4, 8, 9 and 11 have low enough λ^+ and λ^- to be promising as ambipolar small-molecule OMIECs.

The HOMO and LUMO distribution plots (Fig. S23–S34) provide information on orbital localisation and charge distribution. For all molecules, the HOMOs are delocalised across the whole conjugated backbone, while LUMOs are mostly centred on the acceptor fragments, as commonly seen in small-molecule semiconductors.

In future studies, molecular dynamic simulations will be used to study selected molecular aggregates, functionalised with polar side chains, in different electrolytes. Polar side chains are known to influence aggregation, π-stacking and backbone planarity.[68] Orbital localisation patterns will be used to evaluate π-stacking effectiveness towards generating charge percolation pathways in molecular aggregates. Finally, interactions between electrolyte and polar side chains in aqueous electrolytes will provide information on ionic transport and doping.

Conclusion

This proof-of-concept study validated the use of a low-cost computational screening technique to select novel organic mixed conductors that leverage sustainable molecular design,

This journal is © The Royal Society of Chemistry 2026

J. Mater. Chem. A, 2026, **14**, 10346–10354 | **10351**

and a pool of descriptors for aggregation/order and charge transport. While the size of this aggregation is modest, our proof-of-concept shows that low-cost energetic and structure-based descriptors can discover promising novel systems with low E_g, high P and low $\lambda^{-/+}$.

The screening funnel was able to discover 133 bioinspired D–L–A systems with promising energetic and structural parameters for mixed conduction, notably low band gap and reorganisation energy (0.532%) in a chemical space consisting of 24 990 potential D–L–A combinations. The GA approach showed that random mutations of top performing structures can greatly reduce the number of calculations required to discover promising systems.

While our chemical space was not so large as to make a funnel approach prohibitive, the GA still outperformed the funnel, discovering on average 70% of top performers at a much lower computational cost. Applying more stringent filters for E_g, P and $\lambda^{-/+}$ values would also be an effective way to decrease the number of calculations to run for larger datasets.

From this pool, we down-selected 7 D–L–A promising systems and designed 12 candidate systems showing competitive charge transport metrics. Future studies will investigate a much wider chemical space of bioinspired building blocks and π-conjugated heteroaromatic fragments by mining different databases.[69]

Computational methods

RDKit[70] was used to build the D–L–A dataset and combine fragments. The D–A matching and initial E_g estimation were performed at the B3LYP/3-21G* level of theory. B3LYP was chosen due to its demonstrated reliability in estimating E_g at a modest computational cost.[71]

The geometry optimisation of the D–L–A systems is performed in 2 steps. Firstly, starting from the D–L–A SMILES, 100 conformers are generated and optimised using the MMFF94 force field as implemented in the cheminformatics package RDKit. The lowest energy conformer is then optimised at the B3LYP/6-31G* level of theory, yielding an updated estimate of E_g. Further, the data shows the importance of carrying out a geometry optimisation but the value of using a forcefield (MMFF94), implemented within RDKit, to predict the lowest energy conformer (Fig. S35–S38).

Geometry optimisation, HOMO–LUMO gap and reorganisation energy calculations were carried out at the B3LYP/6-31G* level of theory. All DFT calculations were performed using GAUSSIAN[72] and Psi4.[73] cclib[74] was used to calculate the Mulliken population analysis.

Author contributions

TSJ: data acquisition, analysis, software, visualisation, writing – first draft and editing. MM: conceptualisation, project management, software, writing – first draft and editing.

Conflicts of interest

There are no conflicts to declare.

Data availability

Example notebooks, datasets and python code are available at https://github.com/matta-research-group/funnel_paper. The python code used for job submission and data extraction is available at https://github.com/matta-research-group/QCflow. The python code implementing the GA is available at https://github.com/matta-research-group/MithrilMolGA.

Supplementary information (SI): additional plots detailing dataset makeup for all funnel filters; genetic algorithm run statistics and efficiency comparison; chemical structures and orbital delocalisation plots for 12 molecules designed from 7 down-selected systems. See DOI: https://doi.org/10.1039/d5ta10351g.

Acknowledgements

References

1 J. D. Myers and J. Xue, *Polym. Rev.*, 2012, **52**, 1–37.
2 B. D. Paulsen, K. Tybrandt, E. Stavrinidou and J. Rivnay, *Nat. Mater.*, 2020, **19**, 13–26.
3 S. Inal, J. Rivnay, A.-O. Suiu, G. G. Malliaras and I. McCulloch, *Acc. Chem. Res.*, 2018, **51**, 1368–1376.
4 S. T. M. Tan, A. Gumyusenge, T. J. Quill, G. S. LeCroy, G. E. Bonacchini, I. Denti and A. Salleo, *Adv. Mater.*, 2022, **34**, 2110406.
5 S. Pradhan, A. K. Brooks and V. K. Yadavalli, *Mater. Today Bio*, 2020, **7**, 100065.
6 H. Liu, R. Jian, H. Chen, X. Tian, C. Sun, J. Zhu, Z. Yang, J. Sun and C. Wang, *Nanomaterials*, 2019, **9**, 950.
7 E. S. Hosseini, S. Dervin, P. Ganguly and R. Dahiya, *ACS Appl. Bio Mater.*, 2021, **4**, 163–194.
8 W. Li, Q. Liu, Y. Zhang, C. Li, Z. He, W. C. H. Choy, P. J. Low, P. Sonar and A. K. K. Kyaw, *Adv. Mater.*, 2020, **32**, 2001591.
9 M. Yu, Y. Peng, X. Wang and F. Ran, *Adv. Funct. Mater.*, 2023, 2301877.
10 J. Tropp and J. Rivnay, *J. Mater. Chem. C*, 2021, **9**, 13543–13556.
11 H. Tran, V. R. Feig, K. Liu, H.-C. Wu, R. Chen, J. Xu, K. Deisseroth and Z. Bao, *ACS Cent. Sci.*, 2019, **5**, 1884–1891.
12 J. V. Paulin and C. F. O. Graeff, *J. Mater. Chem. C*, 2021, **9**, 14514–14531.
13 M. Reali, P. Saini and C. Santato, *Mater. Adv.*, 2021, **2**, 15–31.
14 E. Vahidzadeh, A. P. Kalra and K. Shankar, *Biosens. Bioelectron.*, 2018, **122**, 127–139.
15 J. Wünsche, Y. Deng, P. Kumar, E. Di Mauro, E. Josberger, J. Sayago, A. Pezzella, F. Soavi, F. Cicoira, M. Rolandi and C. Santato, *Chem. Mater.*, 2015, **27**, 436–442.
16 A. B. Mostert, *Polymers*, 2021, **13**, 1670.
17 P. Meredith and T. Sarna, *Pigment Cell Res.*, 2006, **19**, 572–594.

10352 | *J. Mater. Chem. A*, 2026, **14**, 10346–10354

This journal is © The Royal Society of Chemistry 2026

18 J. T. Blaskovits, R. Laplaza, S. Vela and C. Corminboeuf, *Adv. Mater.*, 2023, **36**, 2305602.

19 S. Luo, T. Li, X. Wang, M. Faizan and L. Zhang, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2021, **11**, e1489.

20 C. Kunkel, J. T. Margraf, K. Chen, H. Oberhofer and K. Reuter, *Nat. Commun.*, 2021, **12**, 2422.

21 X. Xie and A. Troisi, *J. Phys. Chem. Lett.*, 2023, **14**, 4119–4126.

22 M. C. Scharber and N. S. Sariciftci, *Adv. Mater. Technol.*, 2021, **6**, 2000857.

23 N. S. Babu, *Des. Monomers Polym.*, 2021, **24**, 330–342.

24 S. Wang, M. Wang, X. Zhang, X. Yang, Q. Huang, X. Qiao, H. Zhang, Q. Wu, Y. Xiong, J. Gao and H. Li, *Chem. Commun.*, 2013, **50**, 985–987.

25 N. A. Kukhta, A. Marks and C. K. Luscombe, *Chem. Rev.*, 2022, **122**, 4325–4355.

26 D. Hashemi, X. Ma, R. Ansari, J. Kim and J. Kieffer, *Phys. Chem. Chem. Phys.*, 2019, **21**, 789–799.

27 B.-G. Kim, X. Ma, C. Chen, Y. Ie, E. W. Coir, H. Hashemi, Y. Aso, P. F. Green, J. Kieffer and J. Kim, *Adv. Funct. Mater.*, 2013, **23**, 439–445.

28 J. T. Blaskovits, M. Fumanal, S. Vela, R. Fabregat and C. Corminboeuf, *Chem. Mater.*, 2021, **33**, 2567–2575.

29 A. Shafiee, E. Ghadiri, J. Kassis, D. Williams and A. Atala, *Micromachines*, 2020, **11**, 105.

30 T. Dong, L. Lv, L. Feng, Y. Xia, W. Deng, P. Ye, B. Yang, S. Ding, A. Facchetti, H. Dong and H. Huang, *Adv. Mater.*, 2017, **29**, 1606025.

31 N. E. Jackson, K. L. Kohlstedt, B. M. Savoie, M. Olvera de la Cruz, G. C. Schatz, L. X. Chen and M. A. Ratner, *J. Am. Chem. Soc.*, 2015, **137**, 6254–6262.

32 S. Fratini, M. Nikolka, A. Salleo, G. Schweicher and H. Sirringhaus, *Nat. Mater.*, 2020, **19**, 491–502.

33 B. Liu, D. Rocca, H. Yan and D. Pan, *JACS Au*, 2021, **1**, 2182–2187.

34 Y. Che and D. F. Perepichka, *Angew. Chem., Int. Ed.*, 2021, **60**, 1364–1373.

35 L. Dou, Y. Liu, Z. Hong, G. Li and Y. Yang, *Chem. Rev.*, 2015, **115**, 12633–12665.

36 J. Torras, *J. Chem. Educ.*, 2023, **100**, 395–401.

37 R. Gutzler, *Phys. Chem. Chem. Phys.*, 2016, **18**, 29092–29100.

38 Z. Huang, O. Kwon, H. Huang, A. Fadli, X. Marat, M. Moreau and J.-P. Lumb, *Angew. Chem., Int. Ed.*, 2018, **57**, 11963–11967.

39 K.-H. Lin and C. Corminboeuf, *Phys. Chem. Chem. Phys.*, 2020, **22**, 11881–11890.

40 C.-P. Hsu, *Phys. Chem. Chem. Phys.*, 2020, **22**, 21630–21641.

41 Y. Shi, Y. Chang, K. Lu, Z. Chen, J. Zhang, Y. Yan, D. Qiu, Y. Liu, M. A. Adil, W. Ma, X. Hao, L. Zhu and Z. Wei, *Nat. Commun.*, 2022, **13**, 3256.

42 O. López-Estrada, H. G. Laguna, C. Barrueta-Flores and C. Amador-Bedolla, *ACS Omega*, 2018, **3**, 2130–2140.

43 S. Nelsen, S. Blackstock and Y. Kim, *J. Am. Chem. Soc.*, 1987, **109**, 677–682.

44 K. Chen, C. Kunkel, K. Reuter and J. T. Margraf, *Digit. Discov.*, 2022, **1**, 147–157.

45 A. Casey, S. D. Dimitrov, P. Shakya-Tuladhar, Z. Fei, M. Nguyen, Y. Han, T. D. Anthopoulos, J. R. Durrant and M. Heeney, *Chem. Mater.*, 2016, **28**, 5110–5120.

46 B. L. Greenstein, D. C. Elsey and G. R. Hutchison, *J. Chem. Phys.*, 2023, **159**, 091501.

47 Y. Lee, K. Choi and C. Kim, *arXiv*, 2021, preprint, arXiv:.2112.03518, DOI: 10.48550/arXiv.2112.03518.

48 Y. Tang, R. Moretti and J. Meiler, *J. Chem. Inf. Model.*, 2024, **64**, 1794–1805.

49 X. Wang, T. Chen and Y. Gao, in *Isbe 2011: 2011 International Conference on Biomedicine and Engineering*, ed. M. Zhou, Int Industrial Electronic Center, Sham Shui Po, 2011, vol 1, pp. 10–13.

50 R. V. Devi, S. S. Sathya and M. S. Coumar, *Curr. Comput.-Aided Drug Des.*, 2021, **17**, 445–457.

51 A. Tripp and J. M. Hernández-Lobato, *arXiv*, 2023, preprint, arXiv:2310.09267, DOI: 10.48550/arXiv.2310.09267.

52 J. Zhou, Y. Huang, A. Boromand, K. Noori, L. Purvis, C. Oh, L. Lu, Z. W. Ulissi, V. Gharakhanyan and X. Zhang, *RSC Adv.*, 2025, **15**, 43161–43173.

53 P. C. Jennings, S. Lysgaard, J. S. Hummelshøj, T. Vegge and T. Bligaard, *Npj Comput. Mater.*, 2019, **5**, 46.

54 M. Cieślak, J. Łęski, O. Krzysztyńska-Kuleta, J. Kalinowska-Tłuścik and T. Danel, *J. Chem. Inf. Model.*, 2025, **65**, 7811–7816.

55 J. Fang, C. Mao, Y. Zhu, X. Chen, Y. Huang, W. Ding, C.-Y. Hsieh and Z. Ma, *J. Chem. Inf. Model.*, 2025, **65**, 8168–8180.

56 R. Özçelik, H. Brinkmann, E. Criscuolo and F. Grisoni, *J. Chem. Inf. Model.*, 2025, **65**, 7352–7372.

57 Y. Sun, J. Luo, S. Cai, C. Deng, Q. Peng, Y. Shi, H. Li, J. Chen and J. Ding, *Chin. J. Chem.*, 2025, **43**(23), 3065–3074.

58 Z. Wang, X. Ge, X. Gong, Y. Wang and X. Liu, *ACS Appl. Polym. Mater.*, 2025, **7**, 8007–8021.

59 J. Mei and Z. Bao, *Chem. Mater.*, 2014, **26**, 604–615.

60 Y. He, N. A. Kukhta, A. Marks and C. K. Luscombe, *J. Mater. Chem. C*, 2022, **10**, 2314–2332.

61 C. R. Bridges, M. J. Ford, E. M. Thomas, C. Gomez, G. C. Bazan and R. A. Segalman, *Macromolecules*, 2018, **51**, 8597–8604.

62 S. P. O. Danielsen, C. R. Bridges and R. A. Segalman, *Macromolecules*, 2022, **55**, 437–449.

63 S. Himmelberger, D. T. Duong, J. E. Northrup, J. Rivnay, F. P. V. Koch, B. S. Beckingham, N. Stingelin, R. A. Segalman, S. C. B. Mannsfeld and A. Salleo, *Adv. Funct. Mater.*, 2015, **25**, 2616–2624.

64 V. Ho, B. W. Boudouris and R. A. Segalman, *Macromolecules*, 2010, **43**, 7895–7899.

65 M. Wykes, B. Milián-Medina and J. Gierschner, *Front. Chem.*, 2013, **1**, 35.

66 Y. Wan, F. Ramirez, X. Zhang, T.-Q. Nguyen, G. C. Bazan and G. Lu, *Npj Comput. Mater.*, 2021, **7**, 1–9.

67 S. Moro, N. Siemons, O. Drury, D. A. Warr, T. A. Moriarty, L. M. A. Perdigão, D. Pearce, M. Moser, R. K. Hallani, J. Parker, I. McCulloch, J. M. Frost, J. Nelson and G. Costantini, *ACS Nano*, 2022, **16**, 21303–21314.

68 S.-H. Kang, D. Lee, H. Kim, W. Choi, J. Oh, J. H. Oh and C. Yang, *ACS Appl. Mater. Interfaces*, 2021, **13**, 52840–52849.

69 M. Sorokina, P. Merseburger, K. Rajan, M. A. Yirik and C. Steinbeck, *J. Cheminf.*, 2021, **13**, 2.

This journal is © The Royal Society of Chemistry 2026

J. Mater. Chem. A, 2026, **14**, 10346–10354 | **10353**

70 G. Landrum, P. Tosco, B. Kelley, R. Rodriguez, D. Cosgrove, R. Vianello, P. G. sriniker, G. Jones, N. Schneider, E. Kawashima, D. Nealschneider, A. Dalke, M. Swain, B. Cole, S. Turk, A. Savelev, A. Vaucher, M. Wójcikowski, *et al.*, 2024, *rdkit/rdkit: 2024_09_2 (Q3 2024) Release*, (Release_2024_09_2), Zenodo, DOI: 10.5281/zenodo.13990314.

71 S. Tortorella, M. M. Talamo, A. Cardone, M. Pastore and F. De Angelis, *J. Phys.:Condens. Matter*, 2016, **28**, 074005.

72 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, *Gaussian 16*Gaussian, Inc., Wallingford CT, 2019.

73 D. G. A. Smith, L. A. Burns, A. C. Simmonett, R. M. Parrish, M. C. Schieber, R. Galvelis, P. Kraus, H. Kruse, R. Di Remigio, A. Alenaizan, A. M. James, S. Lehtola, J. P. Misiewicz, M. Scheurer, R. A. Shaw, J. B. Schriber, Y. Xie, Z. L. Glick, D. A. Sirianni, J. S. O'Brien, J. M. Waldrop, A. Kumar, E. G. Hohenstein, B. P. Pritchard, B. R. Brooks, H. F. Schaefer, A. Y. Sokolov, K. Patkowski, A. E. DePrince, U. Bozkaya, R. A. King, F. A. Evangelista, J. M. Turney, T. D. Crawford and C. D. Sherrill, *J. Chem. Phys.*, 2020, **152**, 184108.

74 N. M. O'boyle, A. L. Tenderholt and K. M. Langner, *J. Comput. Chem.*, 2008, **29**, 839–845.

10354 | *J. Mater. Chem. A*, 2026, **14**, 10346–10354

This journal is © The Royal Society of Chemistry 2026