Discovery of Crystallizable Organic Semiconductors with Machine Learning

Holly M. Johnson,[†,⊥] Filipp Gusev,[‡,¶,⊥] Jordan T. Dull,[†,⊥] Yejoon Seo,[§] Rodney D. Priestley,[§] Olexandr Isayev,[*,¶,‡] and Barry P. Rand[*,†,‖]

†*Department of Electrical and Computer Engineering, Princeton, New Jersey 08544, United States*
‡*Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States*
¶*Department of Chemistry, Mellon College of Science, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States*
§*Department of Chemical and Biological Engineering, Princeton, New Jersey 08544, United States*
‖*Andlinger Center for Energy and the Environment, Princeton, New Jersey 08544, United States*
⊥*These authors contributed equally to this work.*

Received April 16, 2024; E-mail: olexandr@olexandrisayev.com; brand@princeton.edu

Abstract: Crystalline organic semiconductors are known to have improved charge carrier mobility and exciton diffusion length in comparison to their amorphous counterparts. Certain organic molecular thin films can be transitioned from initially prepared amorphous layers to large-scale crystalline films via abrupt thermal annealing. Ideally, these films crystallize as platelets with long-range-ordered domains on the scale of tens to hundreds of microns. However, other organic molecular thin films may instead crystallize as spherulites or resist crystallization entirely. Organic molecules that have the capability of transforming into a platelet morphology feature both high melting point (T_m) and crystallization driving force (ΔG_c). In this work, we employed machine learning (ML) to identify candidate organic materials with the potential to crystallize into platelets by estimating the aforementioned thermal properties. Six organic molecules identified by the ML algorithm were experimentally evaluated; three crystallized as platelets, one crystallized as a spherulite, and two resisted thin film crystallization. These results demonstrate a successful application of ML in the scope of predicting thermal properties of organic molecules and reinforce the principles of T_m and ΔG_c as metrics that govern the crystallization of organic thin films.

Thin film devices composed of organic semiconductors (OSCs) have gained significant attention due to their compatibility with large area deposition, optoelectronic tunability, and mechanical flexibility.[1] Organic photovoltaic cells (OPVs) and organic light-emitting diodes (OLEDs) have been the most prominent; OPVs have recently reached efficiencies as high as 19.2%[2] while OLEDs have gained public acceptance in the display sector.[3] Typically, organic thin films are amorphous when incorporated into a device, despite crystalline organic films featuring improved exciton diffusion length[4,5] and charge carrier mobility,[6] often by several orders of magnitude. There are several methods for crystallizing organic thin films, including adding polymer or small-molecule additives to OSC solutions,[7] solvent vapor annealing,[8] organic epitaxial growth,[9] and abrupt thermal annealing.[10–15] In this work, we employed an abrupt thermal annealing technique to achieve organic thin-film crystallization.

Amorphous organic thin films fabricated via vacuum thermal deposition can transition into crystalline films upon annealing. The morphology of the crystals that are grown with this technique depend on the molecule itself and the experimental conditions used during fabrication. These conditions are found through experimental optimization and focus on factors such as the thickness of the organic layer, the presence of an organic underlayer, and annealing conditions, among other considerations.[10,11] There are different crystalline morphologies that result from this optimization, such as platelets or spherulites. From the point of view of electronic devices, it is more favorable to have a thin film crystallize as a platelet as spherulites have worse charge carrier mobility compared to platelet crystals[11] and contain many voids and pinholes that can lead to shunting and compromise device yield. Spherulite crystals can exhibit multiple morphologies, such as smooth gradients stemming from a single nucleation point or as sharper, needle-like crystals. Platelet crystals are large-area single crystals on the scale of tens to hundreds of microns across. The long-range order and few grain boundaries minimize deleterious effects such as exciton recombination[16] or carrier scattering.[12,13,17]

However, identifying OSCs that crystallize in this manner is not straightforward, making the pursuit of crystalline organic electronics challenging. Recently, we reported that the thermal properties of OSCs correlate with crystallization in either a platelet or spherulite morphology; the thermal properties guiding these trends are the material's crystallization driving force (ΔG_c) and melting point (T_m).[10] Materials with high ΔG_c and high T_m have a tendency to crystallize as platelets, and as such we can use these values as a guide for selecting organic molecules. Despite its usefulness as a predictive factor, ΔG_c is not a readily available value in databases containing information on OSC molecules and crystal structures, meaning that determination of the value of ΔG_c is still necessary.

To overcome these limitations, we employed machine learning (ML) to predict the thermal properties of OSCs, specifically T_m and ΔG_c. The application of ML in previous literature as applied to OSCs has focused on analyzing and predicting traits such as charge carrier mobility,[18,19] thermal conductivity,[20] static and dynamic disorder as it applies to charge transport,[21,22] and vibrational thermal characteristics such as entropy, specific heat, and dielectric function.[23] In reference to devices, ML has been used to predict the power conversion efficiency (PCE) of OPVs[24] and to aid in the design of more efficient OLEDs, such as identifying thermally activated delayed fluorescent (TADF) emitters.[25] The application of virtual screening, originally popularized in the

drug discovery field,[26–28] has found its use in other branches of chemical sciences.[29–31] Here, we applied descriptor-based ML models to screen commercially available virtual libraries for putative platelet-forming OSCs suitable for experimental validation. We experimentally assessed six organic molecules identified by the virtual screening and were able to crystallize three of these into films with large-area platelets, validating the ML predictions and reinforcing that thermal properties correlate to crystallization behavior.

A schematic of our virtual screening is shown in Figure 1, where several databases comprised of ≈462,000 commercially available organic molecules served as the starting point. In a number of subsequent steps, the database was screened for a combination of favorable chemical properties. Various factors were then taken into account to filter out materials incompatible with experimental constraints. Organic materials with a molecular weight (MW) of less than 300 g/mol were not considered, as materials with low MW have high vapor pressures that can linger in and contaminate vacuum chambers. Other experimental factors considered included the composition of the organic molecules, the number of rotatable bonds, the number of conjugated rings, and the aromatic proportions in each molecule that can give insight into its behavior as a potential crystallizing material and semiconductor. Having too many rotatable bonds inhibits crystallization, and having too few rotatable bonds is also not advantageous, as such molecules may crystallize upon deposition rather than in a post-deposition annealing step. Hence, we looked at materials that had at least three rotatable bonds (Figure 1), which has been shown previously to be the start of the platelet-forming region for organic molecules,[10] and limited the number of aromatic cycles by excluding fullerene-like molecules. We filtered the pool further by selecting molecules with three or more conjugated rings, which, due to the π-electron delocalization from p-orbital overlap caused by alternating single and double bonds, lends to semiconducting behavior.[32,33]

Application of these constraints reduced the number of candidate organic molecules to 7,742, forming a focused library. From here, ML modeling is employed to predict the T_m and ΔG_c of these molecules. We define ΔG_c as

$$\Delta G_c = -\frac{\Delta H_m}{T_m}(T_m - T_c) \quad (1)$$

where T_c is the material's cold crystallization temperature and ΔH_m is the enthalpy of melting.

The focused library (Figure 1) was further screened by ML models trained on thermal properties (See SI-1.2 Model development) for molecules with predicted $T_m > 500$ K and $\Delta G_c < -7.5$ kJ/mol. This resulted in 44 candidates belonging to the platelet-forming region. From these 44, the pool was narrowed down to 13 through consideration of commercial availability and price. Retrospectively, we also evaluated the pipeline's ability to recover those molecules that were previously shown to crystallize as platelets. Out of 6 reported platelet materials,[10] 4 were present in the virtual library. All 4 passed to the focused library stage.

Of these 13 identified molecules, 6 were chosen that fit the experimental parameters, had a molecular structure that aligned with device-building and crystal-forming qualities, and had high predicted values of T_m and ΔG_c. These six were rac-BINAP, TBT, spiro-TAD, TPB-Cz, 9DT, and CZBDF. Full chemical names are included in SI-1.4 Materials. See molecular structures in Figure 2.

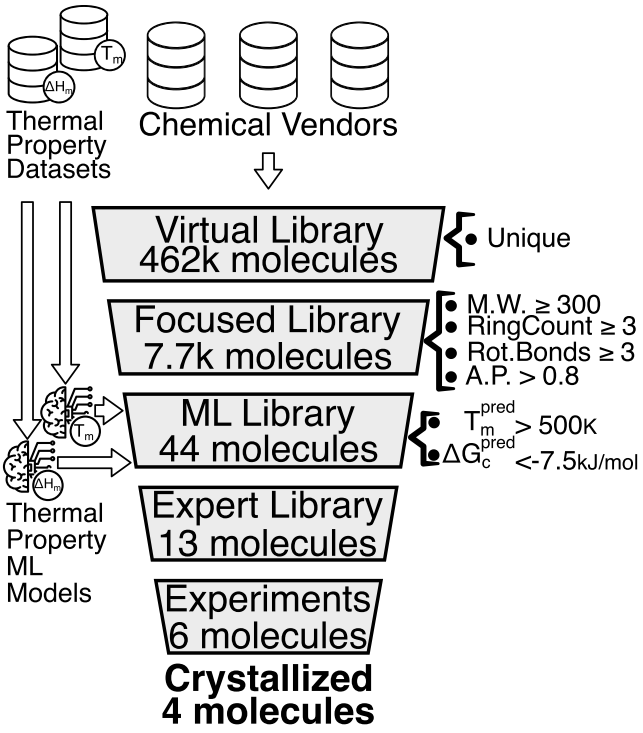

Figure 1. General scheme of the screening campaign: filters and ML-models applied on the corresponding stage of virtual screening shown are described in the text and SI.

Bulk differential scanning calorimetry (DSC) was performed on each material (Figure S2) to determine thermal properties for comparison to predicted values. The predicted values of T_m and ΔG_c and the experimentally derived thermal properties, T_m, T_c, and ΔH_m, are reported in Table 1. Taking the onset values of these properties, we calculated an experimental value of ΔG_c as expressed by Equation 1,[10] also reported in Table 1.

The DSC scan of TBT required further interpretation to calculate ΔG_c. The initial DSC scans of the as-received TBT showed multiple endothermic events. To gain a better understanding of the thermal behavior of this material, TBT was deposited onto a glass substrate in an effort to get rid of potential impurities and form an amorphous thin film. This film was scraped off to form a powder that was then used for DSC. In iterative rounds of heating and cooling, we saw two melting peaks during the first heating (Figure S1b), an indication that the material was not fully amorphous upon deposition and a possible sign of polymorphism,[34] and in following heating scans only the second peak appeared. As the first melting peak was only accessible from the as-deposited TBT, and that is the material used to fabricate and crystallize the thin film, we used the onset of the first melting peak to calculate the ΔG_c of TBT.

The initial DSC scans of CZBDF were featureless, with no thermal events up to 673.0 K (Figure S2f). To further analyze this material, a separate DSC system was used to rerun CZBDF at a higher temperature, recorded in Figure S3. This scan showed an onset melting peak at 714.8 K, which was quite far from the ML predicted T_m of 550.0 K. Crystallization was observed upon cooling, but not upon heating, so ΔG_c could not be calculated.

The process to probe thin film crystal motifs involves annealing films of various thickness at temperatures at and

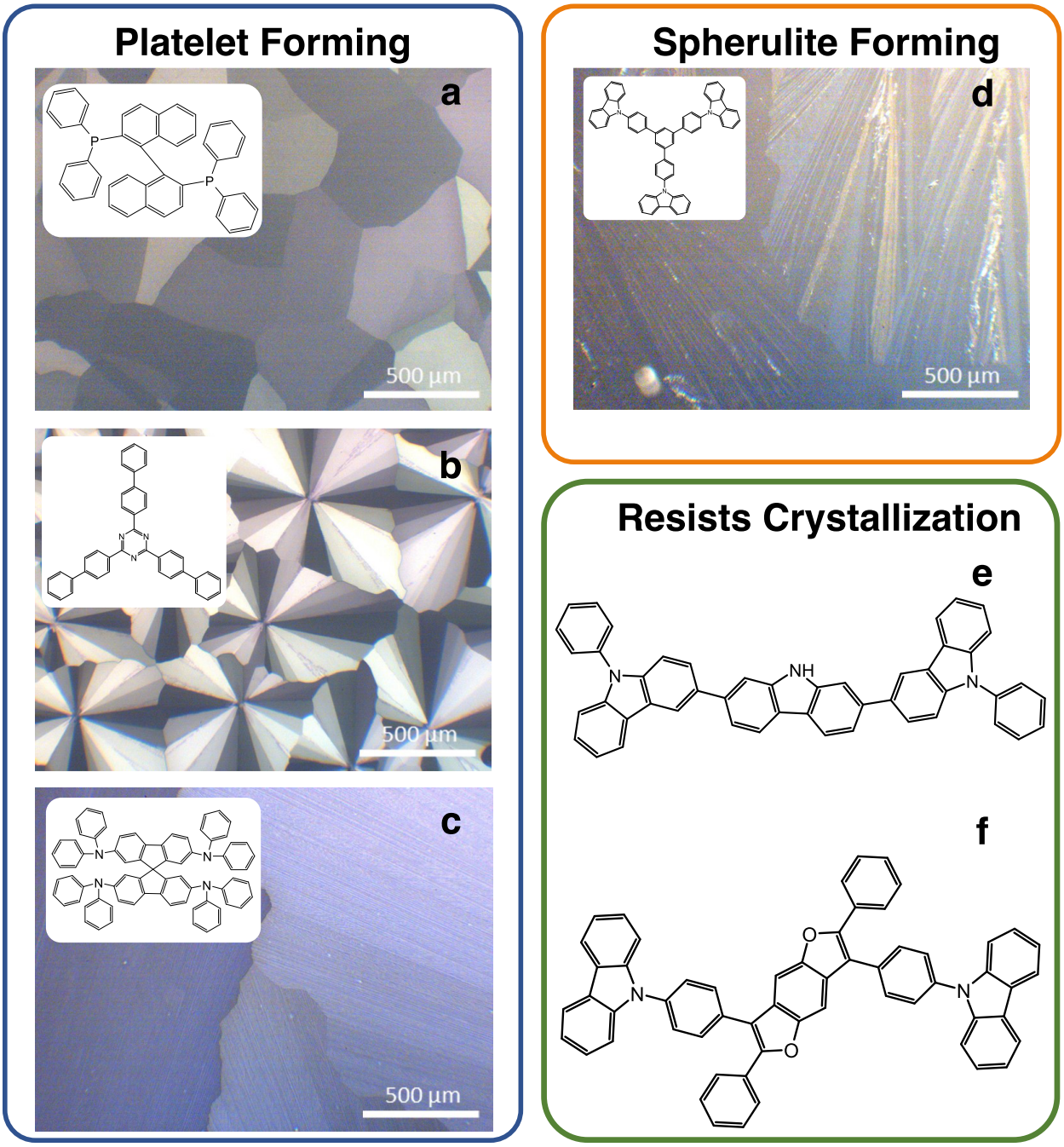

Figure 2. Polarized optical microscopy (POM) images inset with molecular structures of the materials used in this work, grouped by morphology. Here, (a) rac-BINAP, (b) TBT, (c) and spiro-TAD crystallized as platelets; (d) TPB-Cz crystallized as a spherulite; and (e) 9DT and (f) CZBDF resisted crystallization. The letters labeling each material correspond with the letters used in Table 1.

Table 1. Predicted and experimentally derived thermal properties of materials in this work. The experimental values were determined via DSC and based on the onset values of each thermal event. The DSC scan of CZBDF had no crystallization peak upon heating, explaining the absence of an experimental T_c and calculation of ΔG_c.

	Materials	T_m (K)		T_c (K)	ΔH_m (J/g)	ΔG_c (kJ/mol)	
		Predicted	Measured	Measured	Measured	Predicted	Measured
a	rac-BINAP	556.0	556.0	424.0	91.90	-9.09	-13.59
b	TBT	556.0	506.5	376.1	40.65	-9.05	-5.63
c	spiro-TAD	510.6	550.0	463.0	51.04	-8.20	-7.98
d	TPB-Cz	569.3	584.4	507.8	68.65	-8.36	-7.22
e	9DT	551.6	540.0	507.8	79.11	-8.35	-8.42
f	CZBDF	550.0	714.8	–	75.52	-9.60	–

above their glass transition temperature (T_g) as determined through DSC.[10,11] In the cases of TBT and spiro-TAD, 5 nm thick organic underlayers were added to aid crystallization.[11] These conditions are reported in Table S1. Polarized optical microscopy (POM) was performed on these films after annealing to determine morphology and coverage. The POM images of the optimized thin-film crystal growth for these six molecules are shown in Figure 2, grouped together by the type of crystallinity exhibited after optimization. Three of the six molecules investigated crystallized as platelets: rac-BINAP (Figure 2a), TBT (Figure 2b), and spiro-TAD (Figure 2c). The optimized rac-BINAP film displayed full-coverage platelet domains. The next molecule, TBT, crystallized in a unique manner with multiple single crystal domains growing outward from a single nucleation point reminiscent of spherulitic crystal growth. However, due to the limited number of crystal domains and the clear distinction between each domain, this motif is still considered a platelet. The last platelet-forming material, spiro-TAD, had the largest platelet crystal domains studied here, on the scale of millimeters. Only TPB-Cz crystallized as a spherulite (Figure 2d), with visibly rough, sharp, and needle-like crystals with many domains growing from the same nucleation point. Two molecules resisted crystallization: 9DT (Figure 2e) and CZBDF (Figure 2f). The first, 9DT, showed no evidence of crystallization through numerous optimizations, despite its DSC scan showing a clear glass transition and crystallization peak upon heating (Figure S2e). This phenomenon has been observed in a molecule we have previously investigated, di-[4-(N,N-di-p-tolylamino)-phenyl]cyclohexane (TAPC), which also resisted crystallization despite the presence of a crystallization peak in DSC.[10] The second material to resist crystallization, CZBDF, had an inconclusive POM (Figure S1a), so to confirm if the film was crystalline, X-ray diffraction (XRD) was performed on the annealed films. The XRD pattern showed one peak at $2\theta = 30.2°$, corresponding to the ITO substrate,[35] and no other peaks to indicate crystal growth (Figure S1b). Due to this, CZBDF was also defined as a molecule that resists crystallization. It is notable that a melting peak of CZBDF could not be observed until 714.8 K. This can give us further insight into the limits of the range of platelet forming materials as the melting point of CZBDF is significantly higher than the platelet-forming materials we have studied previously. As the T_m of CZBDF falls significantly out of the range of the other platelet forming molecules and the material was difficult to characterize, it should be considered an outlier and excluded from further analysis.

With these materials classified into their morphologies, we then combined these results with our previously reported thermal properties of 22 other organic molecules,[10] as shown in Figure 3 and color coded according to whether they form platelets, spherulites, or resist crystallization. The shaded regions represent the average and one standard deviation of the cumulative T_m and ΔG_c for each respective crystallization category. Each prominent data point shows the predicted and experimental T_m and ΔG_c for the six materials studied in this work and are shaped and color-coded to indicate the morphology of that particular molecule. The empty symbols indicate the ML predicted values, while the filled symbols represent experimental values. Lines guide the eye from the predicted to experimental results, with the exception of CZBDF which only uses the predictions. The 22 previously reported molecules[10] are included but greyed out, with the symbol corresponding to the crystal morphology.

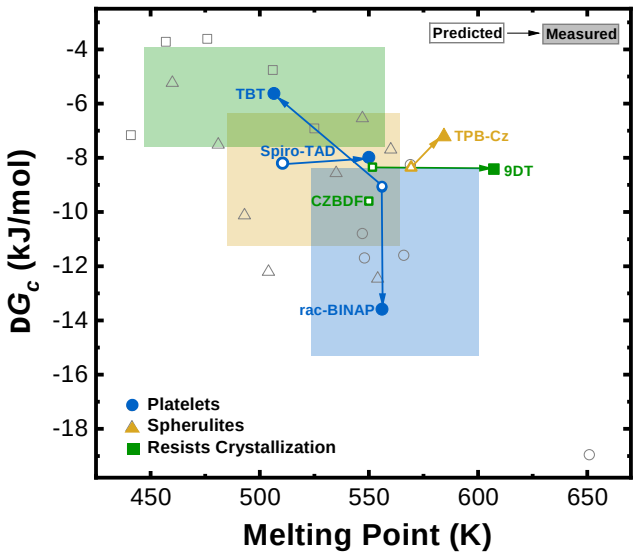

Figure 3. Predicted and measured crystallization driving force at T_c (ΔG_c) as a function of T_m. The arrows on each line point from the ML prediction (empty symbols) to the experimental findings (filled symbols). All other data on this graph come from previous experimental findings[10] and are included in the calculations of the shaded regions, representing the experimental average and one standard deviation of ΔG_c and T_m associated with each morphological trend. Experimental data for CZBDF is not included as its ΔG_c could not be determined.

The T_m and ΔG_c of platelet-forming organic molecules have been seen to be higher on average than the T_m and ΔG_c of spherulites, and markedly higher than organic molecules that resist crystallization,[10] though these regions did overlap and contain outliers. For example, TBT was predicted to fall into the average range of a platelet crystal, yet its experimentally measured thermal properties sited it in a region where we would expect it to resist crystallization. Even so, TBT successfully crystallized as a platelet. Similarly with 9DT, its predicted and experimentally derived T_m and ΔG_c fell within range of spherulitic or platelet growth, yet its film resisted crystallization. These regions, while valuable in producing trends of crystal morphology based on T_m and ΔG_c, are not strict boundaries, and further work is needed to fully understand if there are other reasons for exceptions.

We note that the prediction of T_m for organic molecules is a substantial challenge, and that there is a range of accuracy in the predicted T_m and ΔG_c when compared to the measured thermal properties. The overall RMS error in fully characterized materials is 29.5 K, which is in line with prior work, as the current state-of-the-art predictive models[36] have RMS errors 30-40 K. The model accuracy is limited by the extreme heterogeneity of the T_m data, decomposition upon melting, polymorphism in organic crystals, and amorphous forms. Such information is missing from the current databases, and represent areas where improvements can be made.

We have demonstrated successful application of ML in the scope of organic crystallization. Using ML, we screened nearly half a million commercially available organic molecules and selected several dozen that fit within the necessary property-composition space. Of these, six molecules were chosen based on qualities that aligned with what pre-

vious experimentation has shown to lead to organic crystalline thin films. Three of the six molecules crystallized as large-area platelets, one crystallized as a spherulite, and two resisted crystallization. This displayed a significant success rate of 50% in identifying organics that can crystallize as long-range-ordered platelet domains, the ideal form for future device applications, via the post-deposition annealing method. Machine learning is still a nascent field in its application to organic materials, hence, it is difficult to put current success in full context. However, the success rate substantially surpassed that of drug discovery virtual screening efforts, which rely on the guidance of experienced medicinal chemists for molecule selection to bind a specific protein.[37,38] In our study, we demonstrated that well-trained ML models are capable of accurately predicting thermal properties and replicating the expertise of experimentalists in choosing molecules for validation. This represents an opportunity for a significant shift in decision-making authority from human experts to algorithms. These capabilities mark a crucial advancement towards self-driving laboratories, illustrating the collaborative potential of machine and human intelligence.

Acknowledgement We acknowledge support for this work from the National Science Foundation (NSF) Designing Materials to Revolutionize and Engineer our Future (DMREF) program under awards #DMR-2323751 (Princeton) and DMR-2323749 (CMU). This work used Bridges2 at Pittsburgh Supercomputing Center (PSC) through allocation CHE200122 from the NSF Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program. The authors acknowledge the use of Princeton's Imaging and Analysis Center (IAC), which is partially supported by the Princeton Center for Complex Materials (PCCM), a National Science Foundation (NSF) Materials Research Science and Engineering Center (MRSEC; DMR-2011750).

References

(1) Forrest, S. R. The path to ubiquitous and low-cost organic electronic appliances on plastic. *Nature* **2004**, *428*, 911–918, DOI: 10.1038/nature02498, Number: 6986 Publisher: Nature Publishing Group.

(2) Zhu, L. et al. Single-junction organic solar cells with over 19% efficiency enabled by a refined double-fibril network morphology. *Nature Materials* **2022**, *21*, 656–663, DOI: 10.1038/s41563-022-01244-y, Number: 6 Publisher: Nature Publishing Group.

(3) Chen, H.-W.; Lee, J.-H.; Lin, B.-Y.; Chen, S.; Wu, S.-T. Liquid crystal display and organic light-emitting diode display: present status and future perspectives. *Light: Science & Applications* **2018**, *7*, 17168–17168, DOI: 10.1038/lsa.2017.168, Number: 3 Publisher: Nature Publishing Group.

(4) Lunt, R. R.; Benziger, J. B.; Forrest, S. R. Relationship between Crystalline Order and Exciton Diffusion Length in Molecular Organic Semiconductors. *Advanced Materials* **2010**, *22*, 1233–1236, DOI: 10.1002/adma.200902827, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.200902827.

(5) Najafov, H.; Lee, B.; Zhou, Q.; Feldman, L. C.; Podzorov, V. Observation of long-range exciton diffusion in highly ordered organic semiconductors. *Nature Materials* **2010**, *9*, 938–943, DOI: 10.1038/nmat2872, Number: 11 Publisher: Nature Publishing Group.

(6) Podzorov, V.; Menard, E.; Borissov, A.; Kiryukhin, V.; Rogers, J. A.; Gershenson, M. E. Intrinsic Charge Transport on the Surface of Organic Semiconductors. *Physical Review Letters* **2004**, *93*, 086602, DOI: 10.1103/PhysRevLett.93.086602, Publisher: American Physical Society.

(7) He, Z.; Asare-Yeboah, K.; Zhang, Z.; Bi, S. Manipulate organic crystal morphology and charge transport. *Organic Electronics* **2022**, *103*, 106448, DOI: 10.1016/j.orgel.2022.106448.

(8) Park, C.; Park, J. E.; Choi, H. C. Crystallization-Induced Properties from Morphology-Controlled Organic Crystals. *Accounts of Chemical Research* **2014**, *47*, 2353–2364, DOI: 10.1021/ar5000874, Publisher: American Chemical Society.

(9) Schreiber, F. Organic molecular beam deposition: Growth studies beyond the first monolayer. *physica status solidi (a)* **2004**, *201*, 1037–1054, DOI: 10.1002/pssa.200404334, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pssa.200404334.

(10) Dull, J. T.; Wang, Y.; Johnson, H.; Shayegan, K.; Shapiro, E.; Priestley, R. D.; Geerts, Y. H.; Rand, B. P. Thermal Properties, Molecular Structure, and Thin-Film Organic Semiconductor Crystallization. *The Journal of Physical Chemistry C* **2020**, *124*, 27213–27221, DOI: 10.1021/acs.jpcc.0c09408, Publisher: American Chemical Society.

(11) Fusella, M. A.; Yang, S.; Abbasi, K.; Choi, H. H.; Yao, Z.; Podzorov, V.; Avishai, A.; Rand, B. P. Use of an Underlayer for Large Area Crystallization of Rubrene Thin Films. *Chemistry of Materials* **2017**, *29*, 6666–6673, DOI: 10.1021/acs.chemmater.7b01143, Publisher: American Chemical Society.

(12) Euvrard, J.; Gunawan, O.; Kahn, A.; Rand, B. P. From Amorphous to Polycrystalline Rubrene: Charge Transport in Organic Semiconductors Paralleled with Silicon. *Advanced Functional Materials* **2022**, *32*, 2206438, DOI: 10.1002/adfm.202206438, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adfm.202206438.

(13) Fusella, M. A.; Brigeman, A. N.; Welborn, M.; Purdum, G. E.; Yan, Y.; Schaller, R. D.; Lin, Y. L.; Loo, Y.-L.; Voorhis, T. V.; Giebink, N. C.; Rand, B. P. Band-like Charge Photogeneration at a Crystalline Organic Donor/Acceptor Interface. *Advanced Energy Materials* **2018**, *8*, 1701494, DOI: 10.1002/aenm.201701494, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aenm.201701494.

(14) Fielitz, T. R.; Holmes, R. J. Crystal Morphology and Growth in Annealed Rubrene Thin Films. *Crystal Growth & Design* **2016**, *16*, 4720–4726, DOI: 10.1021/acs.cgd.6b00783, Publisher: American Chemical Society.

(15) Lee, H. M.; Moon, H.; Kim, H.-S.; Kim, Y. N.; Choi, S.-M.; Yoo, S.; Cho, S. O. Abrupt heating-induced high-quality crystalline rubrene thin films for organic thin-film transistors. *Organic Electronics* **2011**, *12*, 1446–1453, DOI: 10.1016/j.orgel.2011.05.015.

(16) Proctor, C. M.; Kuik, M.; Nguyen, T.-Q. Charge carrier recombination in organic solar cells. *Progress in Polymer Science* **2013**, *38*, 1941–1960, DOI: 10.1016/j.progpolymsci.2013.08.008.

(17) Choi, H. H.; Paterson, A. F.; Fusella, M. A.; Panidi, J.; Solomeshch, O.; Tessler, N.; Heeney, M.; Cho, K.; Anthopoulos, T. D.; Rand, B. P.; Podzorov, V. Hall Effect in Polycrystalline Organic Semiconductors: The Effect of Grain Boundaries. *Advanced Functional Materials* **2020**, *30*, 1903617, DOI: 10.1002/adfm.201903617, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adfm.201903617.

(18) Sokolov, A. N.; Atahan-Evrenk, S.; Mondal, R.; Akkerman, H. B.; Sánchez-Carrera, R. S.; Granados-Focil, S.; Schrier, J.; Mannsfeld, S. C.; Zoombelt, A. P.; Bao, Z.; Aspuru-Guzik, A. From computational discovery to experimental characterization of a high hole mobility organic crystal. *Nature Communications* **2011**, *2*, 437, DOI: 10.1038/ncomms1451.

(19) Tan, T.; Wang, D. Machine learning based charge mobility prediction for organic semiconductors. *The Journal of Chemical Physics* **2023**, *158*, 094102, DOI: 10.1063/5.0134379.

(20) Wu, S.; Kondo, Y.; Kakimoto, M.-a.; Yang, B.; Yamada, H.; Kuwajima, I.; Lambard, G.; Hongo, K.; Xu, Y.; Shiomi, J.; Schick, C.; Morikawa, J.; Yoshida, R. Machine-learning-assisted discovery of polymers with high thermal conductivity using a molecular design algorithm. *npj Computational Materials* **2019**, *5*, 66, DOI: 10.1038/s41524-019-0203-2.

(21) Pollice, R.; Friederich, P.; Lavigne, C.; Gomes, G. D. P.; Aspuru-Guzik, A. Organic molecules with inverted gaps between first excited singlet and triplet states and appreciable fluorescence rates. *Matter* **2021**, *4*, 1654–1682, DOI: 10.1016/j.matt.2021.02.017.

(22) Reiser, P.; Konrad, M.; Fediai, A.; Léon, S.; Wenzel, W.; Friederich, P. Analyzing Dynamical Disorder for Charge Transport in Organic Semiconductors via Machine Learning. *Journal of Chemical Theory and Computation* **2021**, *17*, 3750–3759, DOI: 10.1021/acs.jctc.1c00191, Publisher: American Chemical Society.

(23) Tawfik, S. A.; Isayev, O.; Spencer, M. J. S.; Winkler, D. A. Predicting Thermal Properties of Crystals Using Machine Learning. *Advanced Theory and Simulations* **2020**, *3*, 1900208, DOI: 10.1002/adts.201900208, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adts.201900208.

(24) Sahu, H.; Rao, W.; Troisi, A.; Ma, H. Toward Predicting Efficiency of Organic Solar Cells via Machine Learning and Improved Descriptors. *Advanced Energy Materials* **2018**, *8*, 1801032, DOI: 10.1002/aenm.201801032, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aenm.201801032.

(25) Gómez-Bombarelli, R. et al. Design of efficient molecular organic light-emitting diodes by a high-throughput virtual screening and experimental approach. *Nature Materials* **2016**, *15*, 1120–1127, DOI: 10.1038/nmat4717, Number: 10 Publisher: Nature Publishing Group.

(26) Shoichet, B. K. Virtual screening of chemical libraries. *Nature* **2004**, *432*, 862–865, DOI: 10.1038/nature03197.

(27) Sadybekov, A. A. et al. Synthon-based ligand discovery in virtual libraries of over 11 billion compounds. *Nature* **2022**, *601*, 452–459, DOI: 10.1038/s41586-021-04220-9.

(28) Lyu, J.; Wang, S.; Balius, T. E.; Singh, I.; Levit, A.; Moroz, Y. S.; O'Meara, M. J.; Che, T.; Algaa, E.; Tolmachova, K.; Tolmachev, A. A.; Shoichet, B. K.; Roth, B. L.; Irwin, J. J. Ultra-large library docking for discovering new chemotypes. *Nature* **2019**, *566*, 224–229, DOI: 10.1038/s41586-019-0917-9.
(29) Cheng, B.; Griffiths, R.-R.; Wengert, S.; Kunkel, C.; Stenczel, T.; Zhu, B.; Deringer, V. L.; Bernstein, N.; Margraf, J. T.; Reuter, K.; Csanyi, G. Mapping Materials and Molecules. *Accounts of Chemical Research* **2020**, *53*, 1981–1991, DOI: 10.1021/acs.accounts.0c00403.
(30) Back, S. et al. Accelerated chemical science with AI. *Digital Discovery* **2024**, *3*, 23–33, DOI: 10.1039/D3DD00213F.
(31) Butler, K. T.; Davies, D. W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine learning for molecular and materials science. *Nature* **2018**, *559*, 547–555, DOI: 10.1038/s41586-018-0337-2.
(32) Petty, M. C.; Nagase, T.; Suzuki, H.; Naito, H. In *Springer Handbook of Electronic and Photonic Materials*; Kasap, S., Capper, P., Eds.; Springer Handbooks; Springer International Publishing: Cham, 2017; pp 1–1, DOI: 10.1007/978-3-319-48933-9_51.
(33) Yang, Q.; Vriza, A.; Castro Rubio, C. A.; Chan, H.; Wu, Y.; Xu, J. Artificial Intelligence for Conjugated Polymers. *Chemistry of Materials* **2024**, DOI: 10.1021/acs.chemmater.3c02358, Publisher: American Chemical Society.
(34) Leitão, M.; Canotilho, J.; Cruz, M.; Pereira, J.; Sousa, A.; Redinha, J. Study of Polymorphism From DSC Melting Curves; Polymorphs of Terfenadine. *Journal of Thermal Analysis and Calorimetry* **2004**, *68*, 397–412, DOI: 10.1023/a:1016023315613, Publisher: Akadémiai Kiadó, co-published with Springer Science+Business Media B.V., Formerly Kluwer Academic Publishers B.V. Section: Journal of Thermal Analysis and Calorimetry.
(35) Kulkarni, A. K.; Schulz, K. H.; Lim, T. S.; Khan, M. Dependence of the sheet resistance of indium-tin-oxide thin films on grain size and grain orientation determined from X-ray diffraction techniques. *Thin Solid Films* **1999**, *345*, 273–277, DOI: 10.1016/S0040-6090(98)01430-8.
(36) Tetko, I. V.; Sushko, Y.; Novotarskyi, S.; Patiny, L.; Kondratov, I.; Petrenko, A. E.; Charochkina, L.; Asiri, A. M. How Accurately Can We Predict the Melting Points of Drug-like Compounds? *Journal of Chemical Information and Modeling* **2014**, *54*, 3320–3329, DOI: 10.1021/ci5005288.
(37) Korshunova, M.; Huang, N.; Capuzzi, S.; Radchenko, D. S.; Savych, O.; Moroz, Y. S.; Wells, C. I.; Willson, T. M.; Tropsha, A.; Isayev, O. Generative and reinforcement learning approaches for the automated de novo design of bioactive compounds. *Communications Chemistry* **2022**, *5*, 129, DOI: 10.1038/s42004-022-00733-0.
(38) Gutkin, E.; Gusev, F.; Gentle, F.; Ban, F.; Koby, S. B.; Narangoda, C.; Isayev, O.; Cherkasov, A.; Kurnikova, M. In silico screening of LRRK2 WDR domain inhibitors using deep docking and free energy simulations. *Chem. Sci.* **2024**, –, DOI: 10.1039/D3SC06880C, Publisher: The Royal Society of Chemistry.

TOC Graphic