Turbocharged molecular discovery of OLED emitters: From high-throughput quantum simulation to highly efficient TADF devices

Rafael Gómez-Bombarelli[a], Jorge Aguilera-Iparraguirre[a], Timothy D Hirzel[a], Dong-Gwang Ha[b],
Markus Einzinger[b], Tony Wu[b], Marc A Baldo[b], Alán Aspuru-Guzik*[a]
[a]Dept. of Chemistry and Chemical Biology, Harvard University, 12 Oxford St, Cambridge MA 02138; [b]Department of Electrical Engineering and Computer Science, 77 Massachusetts Avenue, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA

ABSTRACT

Discovering new OLED emitters requires many experiments to synthesize candidates and test performance in devices. Large scale computer simulation can greatly speed this search process but the problem remains challenging enough that brute force application of massive computing power is not enough to successfully identify novel structures. We report a successful High Throughput Virtual Screening study that leveraged a range of methods to optimize the search process. The generation of candidate structures was constrained to contain combinatorial explosion. Simulations were tuned to the specific problem and calibrated with experimental results. Experimentalists and theorists actively collaborated such that experimental feedback was regularly utilized to update and shape the computational search. Supervised machine learning methods prioritized candidate structures prior to quantum chemistry simulation to prevent wasting compute on likely poor performers. With this combination of techniques, each multiplying the strength of the search, this effort managed to navigate an area of molecular space and identify hundreds of promising OLED candidate structures. An experimentally validated selection of this set shows emitters with external quantum efficiencies as high as 22%.

Keywords: TADF, DFT, high-throughput, simulation, machine-learning, emitter

1. INTRODUCTION

Novel thermally assisted delayed fluorescence (TADF) materials afford highly efficient organic light-emitting diodes (OLEDs) by harvesting triplet excitons without the need for heavy metals to enhance intersystem crossing.[1-3] By engineering small energy gaps between the lowest excited singlet and triplet states of the emitter (ΔE_{ST}), random thermal fluctuations can be leveraged to produce reverse intersystem crossing of excitons.

High-Throughput Virtual Screening (HTVS) is a powerful tool that greatly accelerates molecular discovery by combining accurate simulation and large scale computing systems.[4] In this work we merge rational molecular design through HTVS with experimental realization to uncover novel highly efficient donor-acceptor TADF materials with a special focus on blue emitters.

The key components of the HTVS process are

- *Candidate libraries* that merge chemical intuition and automated combinatorial generation.
- *Quantum chemical calculation pipelines* to predict key molecular parameters such as emission color or singlet-triplet gap using density functional theory.
- *Supervised Machine Learning* (ML) methods trained on calculated data to automatically prioritize promising candidates.
- *Experiment-theory collaboration* that leverages experimental expertise to drive the computational discovery tool.

This computational platform provides lead candidate molecules that are then synthesized, chemically characterized and tested in optoelectronic devices.

---

* aspuru@chemistry.harvard.edu; phone (617) 384-8188 http://aspuru.chem.harvard.edu/

Organic Light Emitting Materials and Devices XX, edited by Franky So, Chihaya Adachi,
Jang-Joo Kim, Proc. of SPIE Vol. 9941, 99410A · © 2016 SPIE ·
CCC code: 0277-786X/16/$18 · doi: 10.1117/12.2236966

We have pre-screened nearly 2 million molecules with ML. TD-DFT quantum calculations have been performed for over 450,000 candidates. Among these, we have identified hundreds of compounds suitable for high-efficiency TADF emitters.

Various lead candidates have been synthesized and characterized in OLED devices. To assess the predictive power of our approach tested candidates cover a range from pure TADF character (small prompt component, large delayed component) to pure fluorescent (large prompt component, small delayed component). We report compounds with optimal balance between prompt and delayed components reaching external quantum efficiencies as high as 22%.

2. APPROACH

2.1 Candidate libraries

Candidate libraries were generated using customized software that relies on the open-source cheminformatics package RDKit.[5] We used a restricted combinatorial generator that created chemical structures for candidate OLEDs by linking donor bridge and acceptor units. We chose a pool of over 100 donors, over 100 acceptors, and 7 bridges.

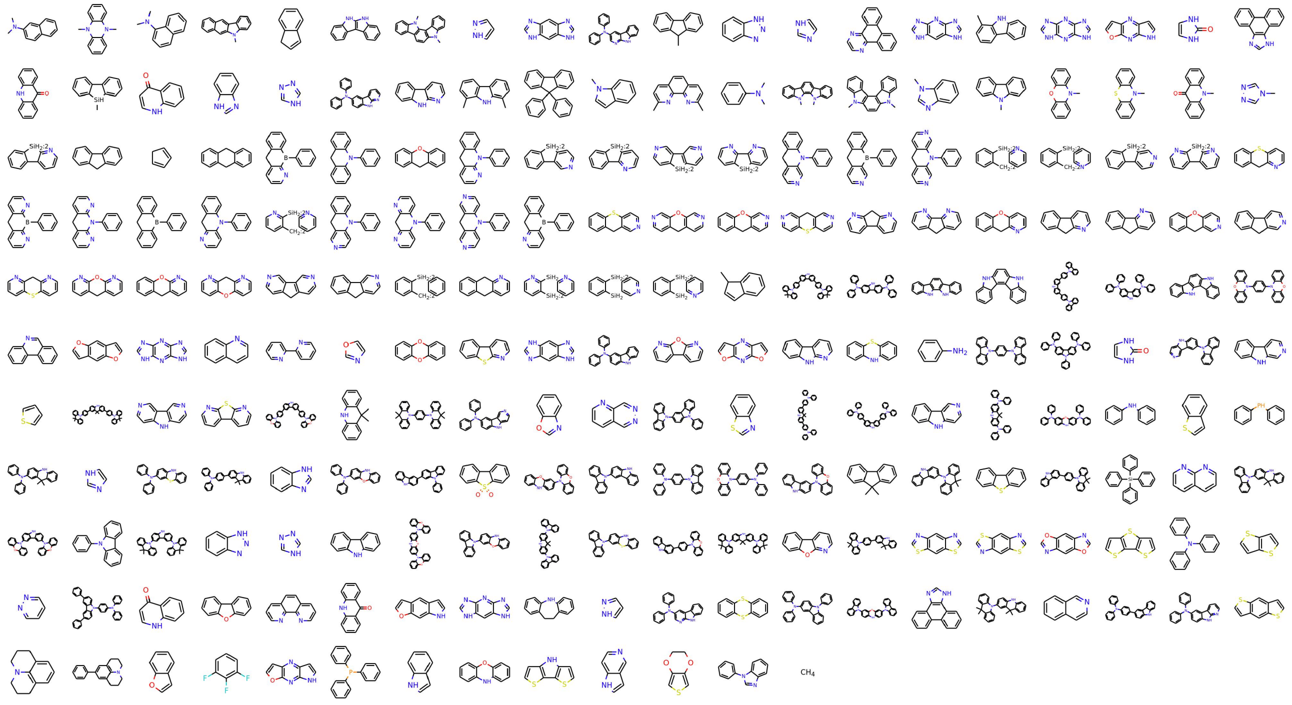


Figure 1. Donor fragments used in the combinatorial generation of OLED candidates.

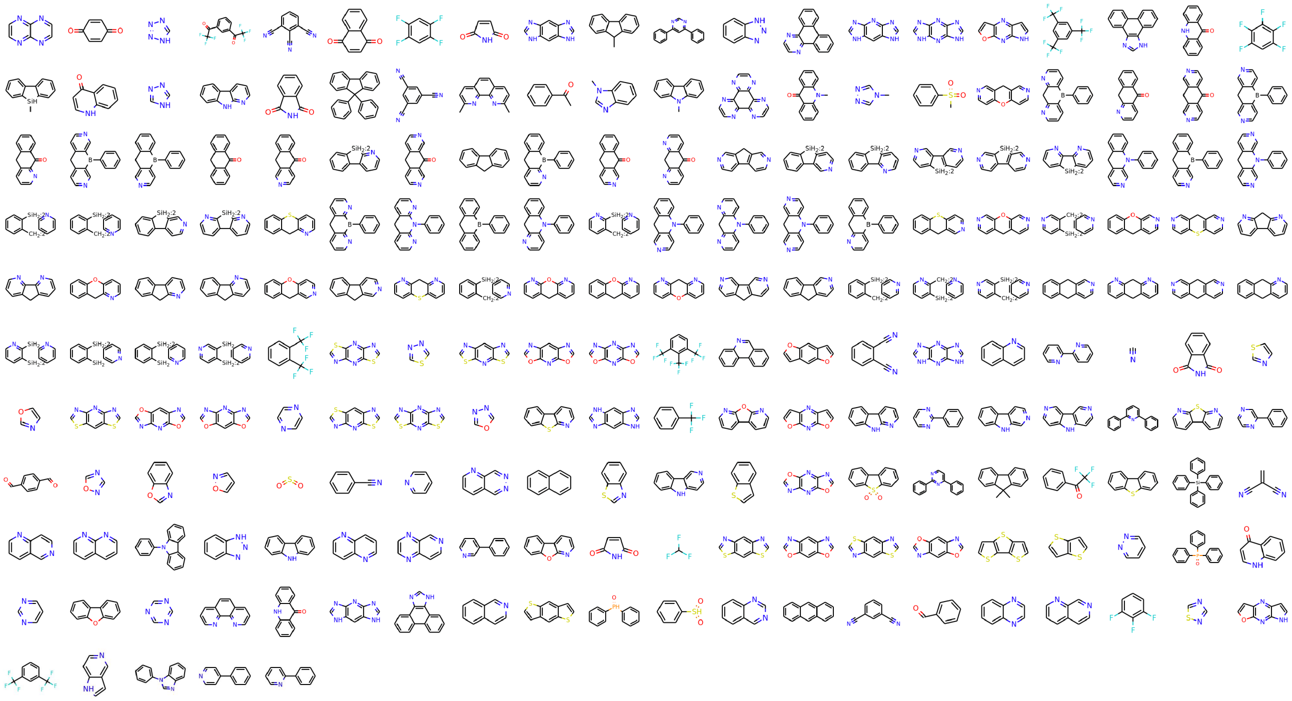


Figure 2. Acceptor fragments used in the combinatorial generation of OLED candidates.

2.2 Quantum chemical calculation pipelines

TADF emitters need to optimize radiative decay to obtain high fluorescent efficiencies and, at the same time, the RISC quantum yields. These two rates can be related to maximizing the oscillator strength (*f*) of the $S_1 \rightarrow S_0$ transition and minimizing the energy gap between the lowest-energy triplet and singlet (ΔE_{ST}). These two properties are readily available through time-dependent TD-DFT calculations. To predict the delayed decay rate that is responsible for the enhanced quantum yield of TADF molecules we used the following equation, deduced from a simple kinetic model of TADF.

$$k_{TADF} = \frac{2\pi e^2 n^3 f}{\varepsilon_0 m c \lambda^2} \frac{1}{1+3\exp(\Delta E_{ST}/T)}$$

After initial conformer generation at the MMFF94 level of theory, we performed consecutive PBE/6-31G(d) and B3LYP/6-31G(d) geometry refinements. We benchmarked the B3LYP, ω-B97xD, BH&HLYP, M06-2X, PBE0, LC-ωPBE0, CAM-B3LYP in the gas phase and in continuum solvent model (ϵ = 3). We also tested the effect of using excited state equilibrium geometries both on the singlet S_1 TD-DFT/B3LYP/6-31G(d) and on the triplet T_1 UDFT/B3LYP/6-31G(d) surface.

2.3 Supervised machine learning

For faster screening, we built a supervised model whose input is a molecule and whose output is a prediction of the TD-DFT-level k_{TADF}. A two-layer deep neural network with 250 hidden units in each layer was built using autograd. We used circular fingerprints as the input features. The system was initially trained on a set of 50,000 TD-DFT calculations for randomly selected molecules. From then on, we followed an epsilon-greedy strategy: 9 out of 10 times, the molecule selected for quantum chemical calculation would be the one for which the neural network predicted the highest k_{TADF} value. The remaining 10% were selected randomly to provide consistent coverage of the molecular search space. The neural network was retrained periodically.

2.4 Experiment-theory collaboration

A website was created to share the leading candidates as predicted by the computational engine and allow the experimental collaborators to discuss the molecules via a comments about molecules and to rate (positive, neutral, negative) molecules based on expected synthetic accessibility, performance, and stability. Between 2 and 10 synthetic organic chemists and device engineers rated each candidate molecule. The web tool was key to leveraging the combined chemical intuition of the team and allowed an international team to quickly collaborate for the selection and feedback process.

On a monthly basis, the best performers located by the computing engine in the previous month were filtered out and collected into a batch. This batch was defined in the website and shared with the whole team. This initiated the voting process. A subset of the collaborators would then meet in person to review the batch and votes. The voting process prior to the in-person discussion helped drive the discussion towards the most interesting candidates in the batch-- often the most popular candidates or controversial candidates that had opposing votes.

Feedback from collaborators was utilized in three key ways. First, the combinatorial generator could be updated. For example, theorists could add new fragments, or black-list weak bonds or unwanted sub-structures. Second, the simulations could be changed. Less commonly, the simulation process was updated to improve performance for certain properties or to accommodate new chemistries. Third, discussions could provoke fresh attempts at finding synthetic pathways to particularly desirable molecules.

2.5 Synthesis and device characterization

Lead candidates were synthesized using common synthetic approaches and characterized optically and electrochemically. Emitters were characterized in optoelectronic devices whose structure was as follows: indium tin oxide (ITO, 160nm), 4,4′-cyclohexylidenebis[*N*,*N*-bis(4-methylphenyl)benzenamine] (TAPC, 150nm), 10% emissive molecules doped in bis[2-(diphenylphosphino)phenyl] ether oxide (DPEPO, 30nm), DPEPO (5nm), 1,3,5-tris(3-pyridyl-3-phenyl)benzene (TmPyPb, 45 nm), lithium fluoride (LiF, 1nm) and aluminum (Al, 100nm). (See [6] for details)

3. RESULTS AND DISCUSSION

3.1 Quantum chemical predictions

In Figure 1 we report the predicted properties of over 400,000 molecules in our OLED database. In addition to thoroughly charting chemical space, the plots show the performance of various reported TADF molecules, the predicted properties of the four lead molecules characterized in this work, and the expected parameters for yet-untested promising compounds.

A significant tradeoff between f and ΔE_{ST} can be seen in Fig 1 a), as has been reported in the literature before. The lead untested candidates predictions from our search (orange squares) show a similar distribution to previously reported emitters (red circles) as regards their f and ΔE_{ST} values and are expected to afford high efficiencies. Fig 1 b) and c) show that our lead predictions tend to include more sky-blue and deep-blue OLED emitters than the literature sample. This is due to the fact that our chemical libraries have focused on donor and acceptor fragments with high T_1 energies, and thus are strongly biased towards blue emitters.

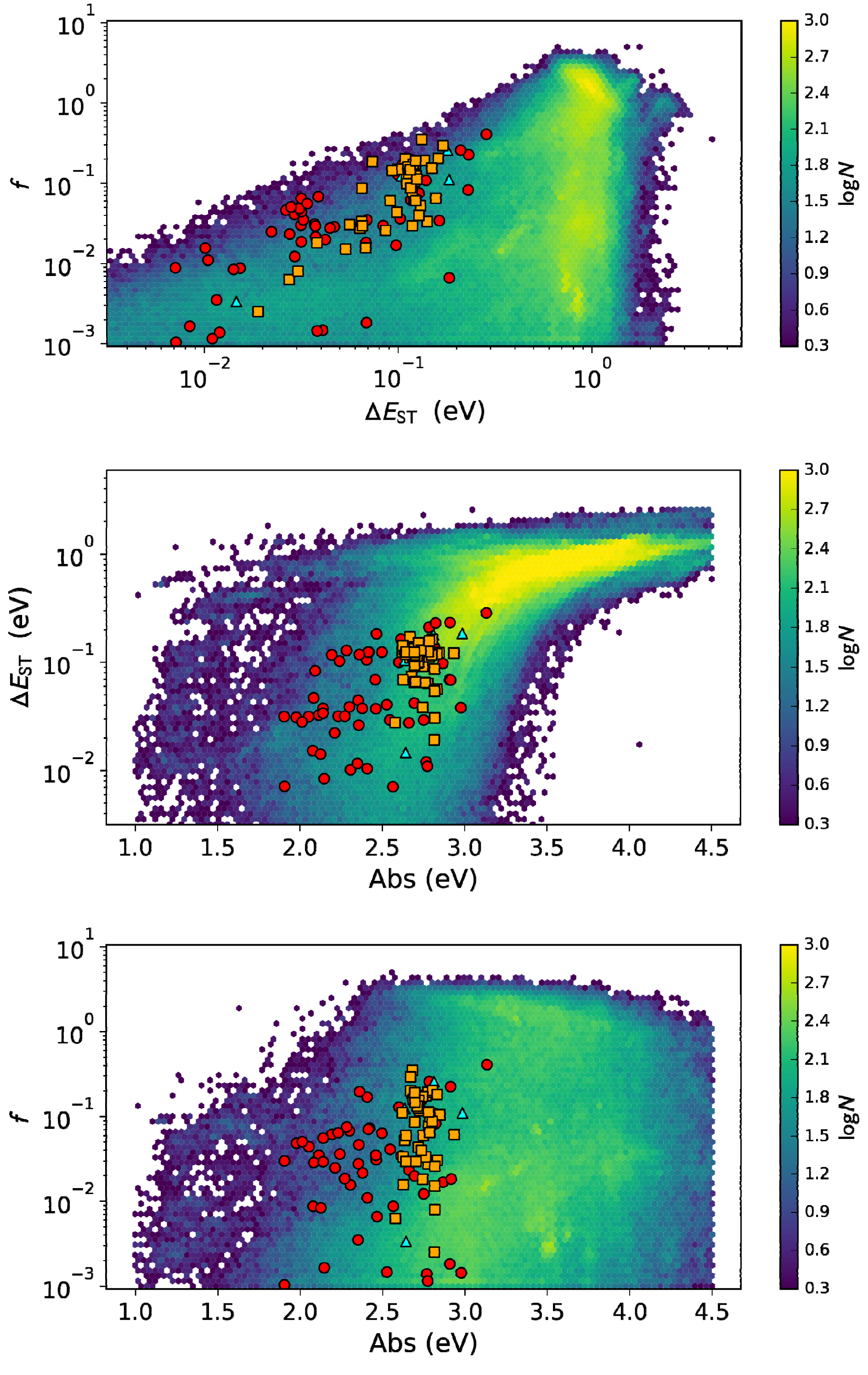
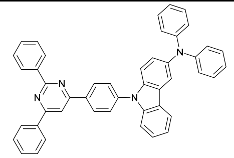
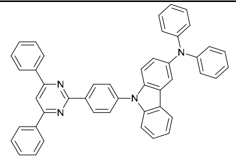
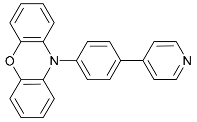
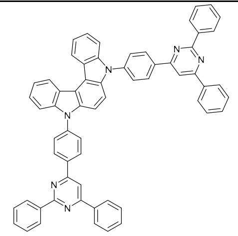


Figure 3. Density of candidate molecular materials as a function of their predicted properties. Known TADF emitters reported in the literature are shown as red dots, the four lead compounds tested in this work are reported as cyan triangles and the lead suggestions from our search are plotted as orange squares. a) strength (*f*) and singlet-triplet gap (ΔE_{ST}) b) Singlet-triplet gap (ΔE_{ST}) and gas-phase absorption (Abs) c) Density of candidate molecular materials as a function of their oscillator strength (*f*) and gas-phase absorption (Abs)

3.2 Experimental realization

The lead candidates tested in this work are reported in Table 1. They combine pyrimidine and pyridine acceptors with phenoxazine and carbazol-derivative donors. The molecules explore the balance between prompt emission and TADF character and achieve maximum external quantum efficiencies over 20%.

Table 1. Structure and metrics of performance of the four lead compounds tested in this work

Structure	Electroluminescence (nm)	EQE (max)	EQE (100 cd/cm^2)
	524	21	15.6
	520	13	6.8
	507	21	10.2
	501	12	7.6

3.3 Lead candidates for further exploration

In addition to the candidates that were tested empirically, we report a series on promising leads, with $f > 0.003$ and $\Delta E_{\mathrm{ST}} < 0.175$ eV. These molecules have not been realized to this date, and could be the subject of further experimental and theoretical study (Figure 4).

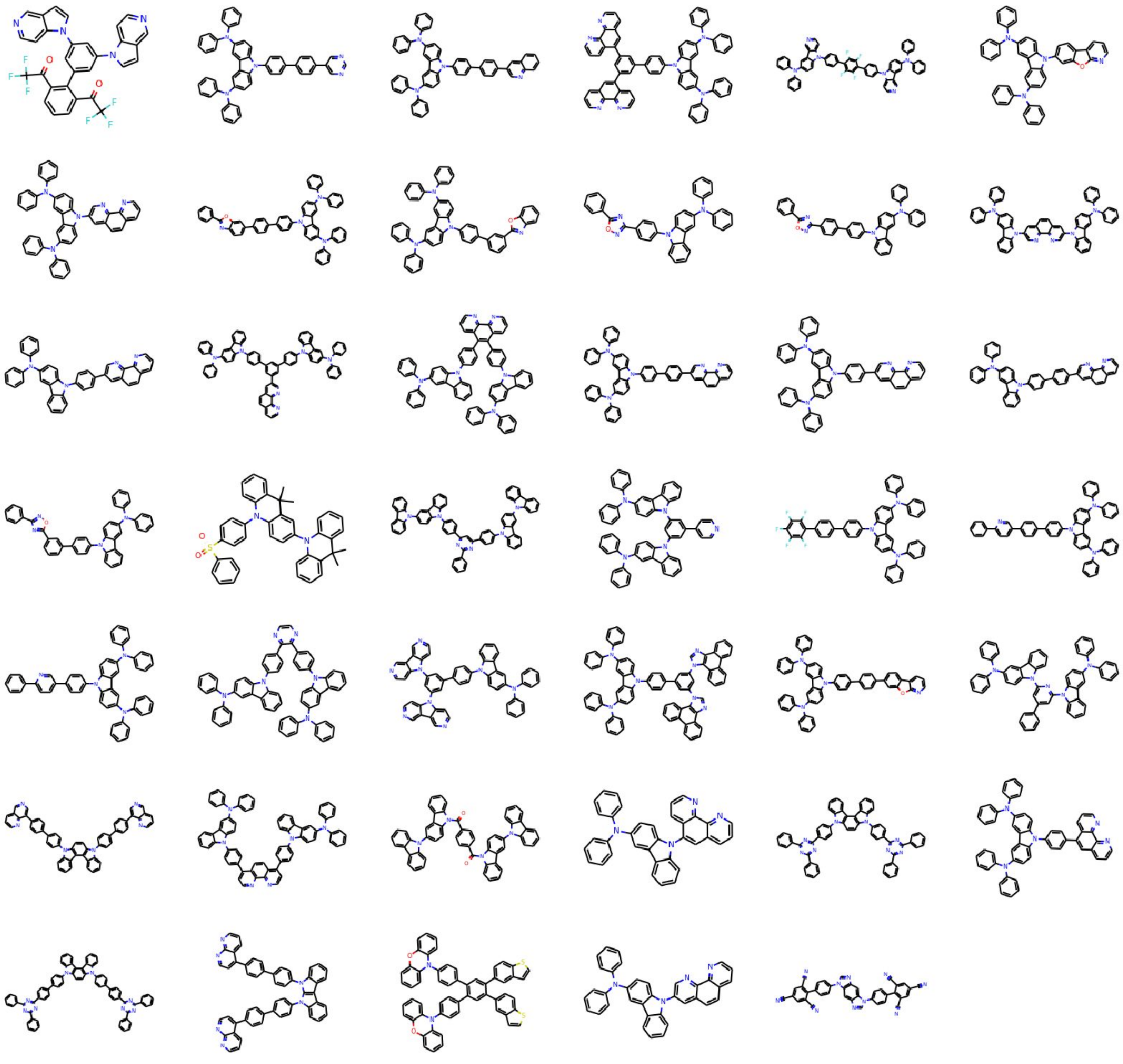


Figure 4. Promising OLED candidates with predicted absorption below 490 nm, predicted $\Delta E_{\mathrm{ST}} < 0.175\mathrm{eV}$ and $f > 0.003$.

4. CONCLUSION

We report a mixed theoretical and experimental study that combines multiple features of both simulation and laboratory work to produce a highly promising set of molecular candidates for TADF emitters. By successively sieving a set of over 2 million molecules, we have used machine learning, chemistry and chemical expertise to identify lead compounds. Four of these candidates have been realized experimentally and are reported to match the performance of known world-class TADF emitters, despite being entirely computationally derived. A small selection of untested chemical candidates is also put forward.

REFERENCES

[1] Uoyama, H., Goushi, K., Shizu, K., Nomura, H., and Adachi, C., "Highly efficient organic light-emitting diodes from delayed fluorescence," Nature*,* 492(7428), 234-+ (2012).

[2] Zhang, Q. S., Li, B., Huang, S. P., Nomura, H., Tanaka, H., and Adachi, C., "Efficient blue organic light-emitting diodes employing thermally activated delayed fluorescence," Nat Photonics*,* 8(4), 326-332 (2014).

[3] Hirata, S., Sakai, Y., Masui, K., Tanaka, H., Lee, S. Y., Nomura, H., Nakamura, N., Yasumatsu, M., Nakanotani, H., Zhang, Q. S., Shizu, K., Miyazaki, H., and Adachi, C., "Highly efficient blue electroluminescence based on thermally activated delayed fluorescence," Nature Materials*,* 14(3), 330-336 (2015).

[4] Pyzer-Knapp, E. O., Suh, C., Gómez-Bombarelli, R., Aguilera-Iparraguirre, J., and Aspuru-Guzik, A., "What Is high-throughput virtual screening? A perspective from organic materials discovery," Annual Review of Materials Research*,* 45, 195-216 (2015).

[5] Tosco, P., Stiefl, N., and Landrum, G., "Bringing the MMFF force field to the RDKit: implementation and validation," J Cheminformatics*,* 6, (2014).

[6] Gómez-Bombarelli, R., Aguilera-Iparraguirre, J., Hirzel, T., Duvenaud, D. K., Maclaurin, D., Blood-Forsythe, M. A., Sik Chae, H., Einzinger, M., Ha, D.-G., Wu, T., Markopoulos, G., Jeon, S., Kang, H., Miyazaki, H., Numata, M., Kim, S., Huang, W., Ik Hong, S., Baldo, M., Adams, R. P., and Aspuru-Guzik, A., "Design of efficient molecular organic light-emitting diodes by a high-throughput virtual screening and experimental approach," Nature Materials, (2016).