Supporting Information:
Harnessing Generative AI for Efficient Organic Materials Discovery in Low-Data Regimes

Jun Hyeong Kim[a‡], Kyunghoon Lee[a‡], Hyeonsu Kim[a,b], MinSoo Kang[c], Suk-Ku Chang[d], Yinglan Jin[d], Dongwook Kim[*e], and Woo Youn Kim[*a]

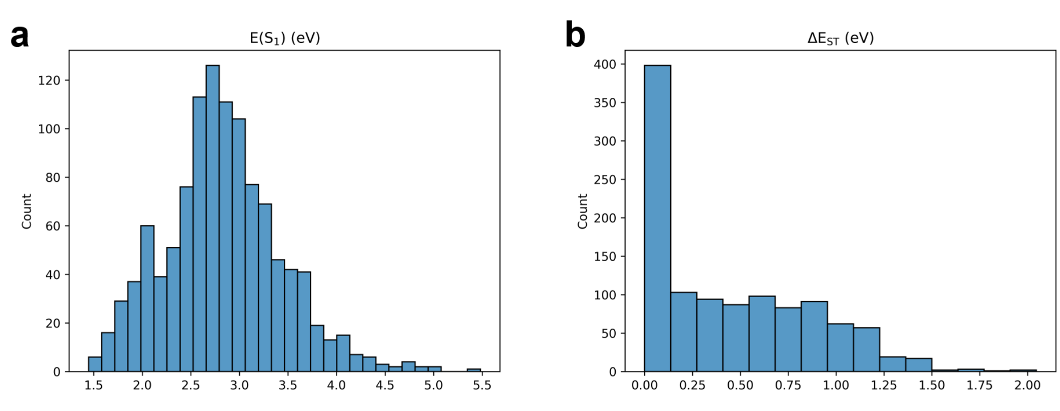
[a]Department of Chemistry, KAIST, 291, Daehak-ro, Yuseong-gu, Daejeon, 34141, Republic of Korea,
[b] Current address : Simulation Group, Samsung SDI, Samsung-ro, Yeongtong-gu, Suwon-si,
16678, Gyeonggi-do, Republic of Korea,
[c] Material Development Department, SK Materials JNC, 5609 Dongtangiheung-ro,
Hwa-sung City, 18469, Gyeonggi-do, Republic of Korea,
[d] UNIPlus, 40 Omokcheon-ro 152beon-gil, Suwon-si, 16642, Gyeonggi-do, Republic of Korea,
[e] Department of Chemistry, Kyonggi University, 154-42 Gwanggyosan-ro, Yeongtong-gu,
Suwon-si, 16227, Gyeonggi-do, Republic of Korea,
[‡] These authors contributed equally to this work,
[*]Corresponding authors; E-mail: dongwook-kim@kgu.ac.kr; wooyoun@kaist.ac.kr.

January 27, 2026

Contents

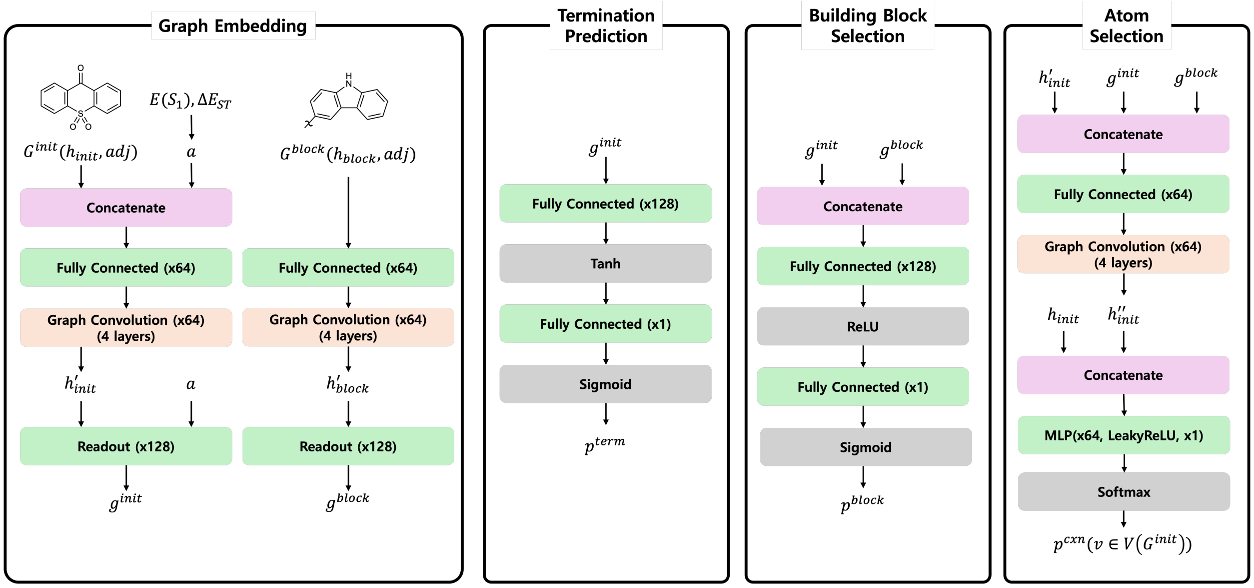
List of Tables

1 Property distribution of training dataset



Supplementary Figure 1: Property distribution of the training dataset. a) Distribution of vertical singlet excitation energies ($E(S_1)$, S_1 energy) in the training set. b) Distribution of singlet-triplet energy gaps (ΔE_{ST}) for the same set of OLED molecules.

2 Model Architecture and Hyperparameters



Supplementary Figure 2: A schematic of the model architecture

For the GNN model, each molecule was represented as a graph where atoms are nodes and bonds are edges. The features used to describe each atom (node) are detailed in Table 1. Edge features were represented by the molecule's adjacency matrix, which indicates connectivity between atoms. Furthermore, a self-loop was added for each atom.

Supplementary Table 1: Atom features used in the GNN model.

Feature name	**Description**	**Values**
Period	The period of the atom	`[0, 1, 2, 3, 4, 5]`
Group	The number of outer-shell electrons	`[0, 1, ..., 8]`
Degree	The number of bonds connected to the atom	`[0, 1, ..., 5, other]`
Valence	The total valence of the atom	`[0, 1, ..., 8, other]`
Formal Charge	The formal charge of the atom	`[-1, -2, 1, 2, 0, other]`
Number of Hydrogens	The number of bonded hydrogen atoms	`[0, 1, ..., 4, other]`
Hybridization	The hybridization state of the atom	`[SP, SP2, SP3, SP3D, SP3D2, other]`
BRICS Index	Index of the BRICS fragment (building block only)	`[0, 1, ..., 16]`

During training, we maximize p^{term}, p^{block}, and p^{cxn} using cross-entropy-based losses. For p^{term}, we apply binary cross-entropy loss, as it is a binary classification task. In contrast, p^{block} and p^{cxn} are multi-class classification tasks, and thus require softmax cross-entropy loss. However, p^{block} involves up to 589 building block classes, which can hamper learning due to the large output space. To address this, the original BBAR work adopts a negative sampling strategy, in which a subset of other fragments (excluding the correct one) is sampled as negative examples during training. We follow the same approach here and refer to the original BBAR paper for further details on the training procedure.

The total loss is expressed as a linear combination (often with equal weights) of the three loss terms. We trained the model for 20 epochs. During training, the batch size was set to 64 and the maximum number of atoms per molecule was set to 93. Finally, we randomly split the dataset into training, validation, and test sets in a 75:15:10 ratio.

3 Quantitative Evaluation of Chemical Novelty and Structural Diversity

To assess the chemical diversity of the molecules (G) generated by the BBAR model relative to the training set (T), we performed a two-part quantitative analysis. Each part is described in detail below:

3.1 Novelty and Similarity to Training Data

First, we measured the structural similarity between the training set and the generated molecules using Tanimoto similarity based on Morgan fingerprints (radius 2, 2048 bits). Here, $fp(x)$ denotes the fingerprint of molecule x, and $Sim(A, B)$ represents the Tanimoto similarity between the fingerprints of molecules A and B.

- **Novelty (↑)**: The fraction of generated molecules not identical to any instance in T based on canonical SMILES strings. A value of 1.0 indicates that all generated molecules are novel relative to the training data.

$$\text{Novelty} = 1 - \frac{|G \cap T|}{|G|} \tag{1}$$

- **Average Maximum Similarity (↓)**: Measures the proximity of each generated candidate to its nearest neighbor in T at both whole-molecule and Murcko scaffold levels. Lower values indicate higher structural novelty. In terms of molecule, the metric is measured as

$$\text{Avg. Max. Sim.(Mol)} = \frac{1}{|G|} \sum_{g \in G} \max_{t \in T} Sim(fp(g), fp(t)) \tag{2}$$

  and in terms of the Murcko scaffold, it is computed as

$$\text{Avg. Max. Sim.(Murcko)} = \frac{1}{|G|} \sum_{s_g \in S_G} \max_{s_t \in S_T} Sim(fp(s_g), fp(s_t)) \tag{3}$$

  where S_G and S_T refers to the sets of Mucrko scaffolds extracted from G and T.

- **Internal Diversity (↑)**: Measures the degree of structural dispersion within the generated set by calculating the average pairwise distance. Higher values indicate that the model avoids mode collapse (producing same molecules) and maintains a diverse output.

$$\text{Internal Diversity} = 1 - \frac{2}{|G|(|G|-1)} \sum_{i<j}^{|G|} Sim(fp(g_i), fp(g_j)) \tag{4}$$

Supplementary Table 2 summarizes the values of four metrics for the BBAR-generated molecules across the blue, green, and red sets. The results indicate that the BBAR model effectively explores novel chemical space, as evidenced by novelty scores approaching 1.0. The moderate average maximum similarities at the molecular (0.399–0.444) and scaffold (0.460–0.487) levels confirm that the generated candidates are structurally distinct from the training instances rather than simple duplicates. Furthermore, the internal diversity of the generated sets (0.804–0.809) is comparable to that of the training set (0.8146), indicating that the distribution of pairwise molecular similarities is well preserved and that the model avoids mode collapse.

Supplementary Table 2: Novelty and similarity metrics comparing the Training set and BBAR-generated sets. The arrows indicate the direction of better performance with respect to chemical diversity (↑: higher is better, ↓: lower is better).

Group	Novelty (↑)	Avg. Max Mol. Sim. (↓)[a]	Avg. Max Scaf. Sim. (↓)[b]	Internal Diversity (↑)
Train	-	-	-	0.815
Blue	0.996	0.444	0.487	0.808
Green	1.000	0.420	0.473	0.809
Red	1.000	0.399	0.460	0.804

[a] Avg. Max Mol. Sim.: Average Maximum Tanimoto Similarity at the whole-molecule level to the training set.
[b] Avg. Max Scaf. Sim.: Average Maximum Tanimoto Similarity at the Murcko scaffold level to the training set.

3.2 Core Scaffold Variety and Building Block Statistics

Second, we analyzed the diversity of core structural motifs and the frequency with which the model utilized the building block (BB) library. This analysis provides insight into whether the model actively incorporates a wide range of building blocks to generate structurally diverse molecules. To evaluate the model, we utilized two metrics defined as below:

- **Scaffold Uniqueness (↑)**: The ratio of unique Murcko scaffolds to the total number of Murcko scaffolds. A value near 1.0 signifies that nearly every generated molecule possesses a unique core framework.
$$\text{Scaffold Uniqueness} = \frac{N_{\text{unique scaffolds}}}{N_{\text{valid scaffolds}}} \tag{5}$$
- **Top N Dominance (↓)**: The cumulative frequency of the N most frequent structural units. Lower values represent a more decentralized and utilization of diverse building blocks, avoiding over-reliance on a few dominant fragments.
$$\text{Top N Dominance} = \sum_{k=1}^{N} \text{Frequency}(\text{Item}_k) \tag{6}$$

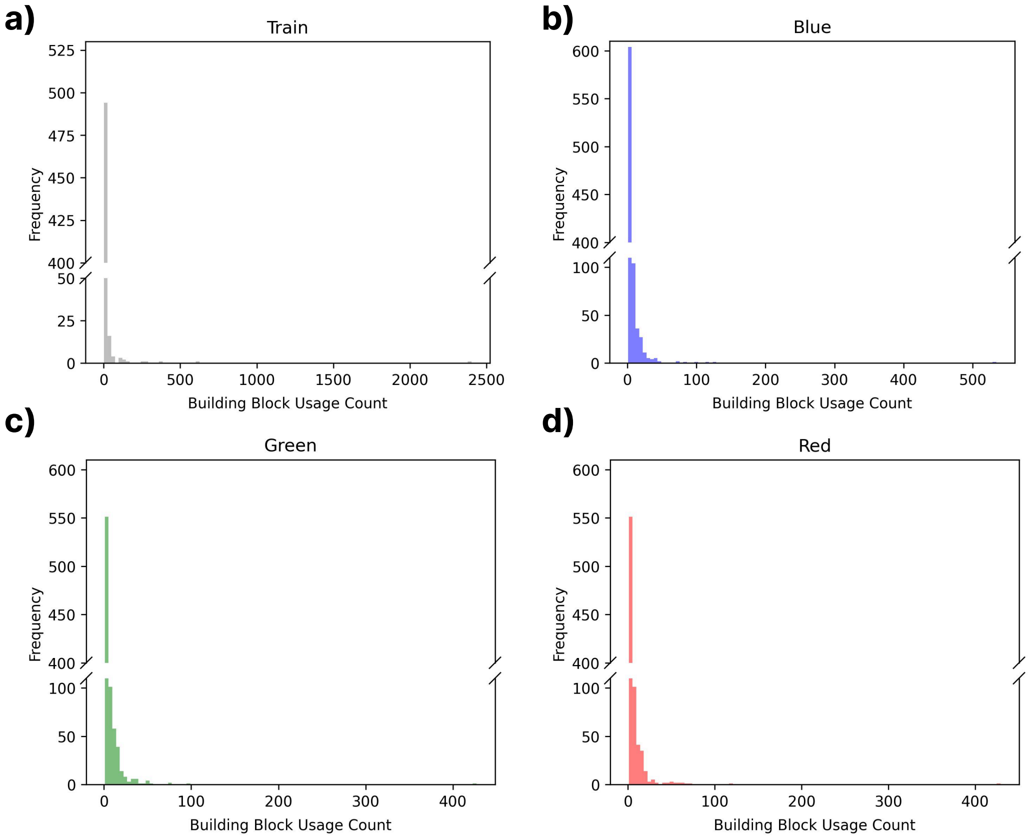
Supplementary Table 3: Scaffold uniqueness and building block usage metrics. The arrows indicate the direction of better performance with respect to chemical diversity (↑: higher is better, ↓: lower is better).

Group	Scaffold Uniqueness (↑)	Top 10 B.B. Dom. (↓)[a]	Top 25 B.B. Dom. (↓)[a]	Top 50 B.B. Dom. (↓)[a]
Train	0.858	63.075	73.213	80.983
Blue	0.997	23.345	32.675	42.551
Green	0.997	18.959	29.118	39.802
Red	1.000	21.541	33.112	43.145

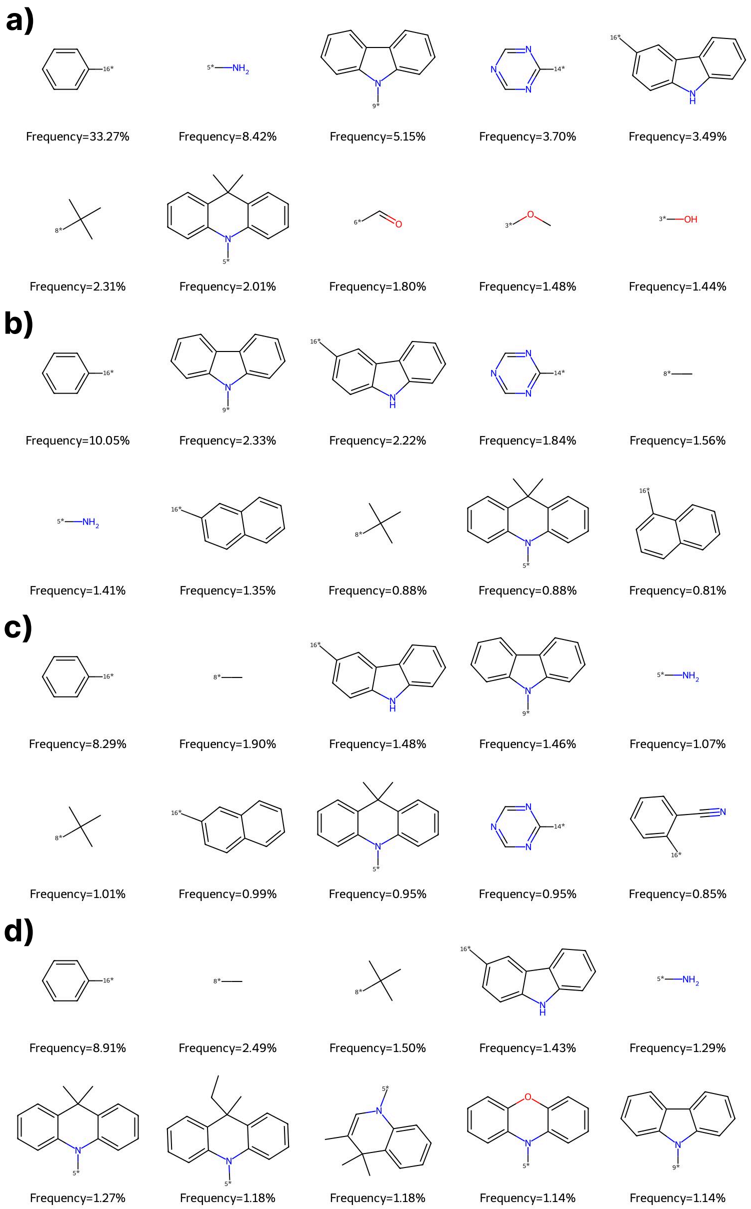
[a] Top k B.B. Dom.: Cumulative frequency of the k most frequent BRICS building blocks, expressed as a percentage of the total fragment occurrences in each set.

Supplementary Figure 3 shows the histogram plot for the frequency statistics of building block utilization, and Supplementary Table 3 summarizes the values of two metrics (scaffold uniqueness and top-K building block dominance) for the BBAR-generated molecules across the blue, green, and red sets. As summarized in Supplementary Table 3, BBAR achieves near-perfect scaffold uniqueness (≥ 0.997), reflecting its ability to generate diverse molecular skeletons. Moreover, with respect to the top-K building block dominance, while the training set is heavily biased toward a few privileged building blocks (top 10 dominance: 63.075%), the BBAR model significantly decentralizes fragment usage (top 10 dominance: 20–23%). The structures of those top 10 building blocks along with their frequency rates within the training set and each generated candidate set are provided in Supplementary Figure 4

Supplementary Figure 3 shows a histogram plot of building block usage frequencies, and Supplementary Table 3 summarizes two key metrics (scaffold uniqueness and top$K$ building block dominance) for the BBAR-generated molecules across the blue, green, and red sets. As shown in Supplementary Table 3, BBAR achieves near-perfect scaffold uniqueness (≥ 0.997), reflecting its ability to generate diverse molecular skeletons. Furthermore, in terms of top$K$ building block dominance, the training set is heavily biased toward a few privileged building blocks (top 10 dominance: 63.075%), whereas BBAR significantly decentralizes fragment usage (top 10 dominance: 20–23%). The structures of the top 10 building blocks, along with their usage rates in the training and generated sets, are provided in Supplementary Figure 4.

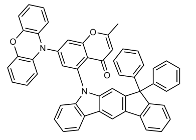


Supplementary Figure 3: Histograms of building block usage counts for (a) the training set, (b) blue, (c) green, and (d) red TADF candidate sets. The x-axis represents the number of times a specific building block was used within the set, and the y-axis indicates the frequency of building blocks with that usage count.

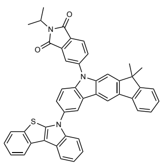


Supplementary Figure 4: Top 10 most frequent building blocks and their frequencies for (a) the training sets, (b) blue, (c) green, and (d) red TADF candidate sets.

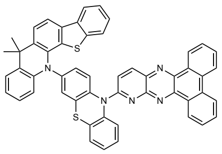
4 Selected TADF candidates and their synthesized derivatives



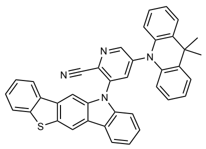
$E(S_1)$: 2.413 eV
ΔE_{ST} : 0.095 eV
$f(S_1)$: 0.0366



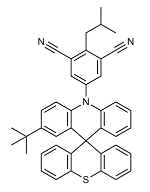
$E(S_1)$: 2.359 eV
ΔE_{ST} : 0.063 eV
$f(S_1)$: 0.0304



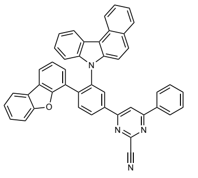
$E(S_1)$: 2.244 eV
ΔE_{ST} : 0.031 eV
$f(S_1)$: 0.0051



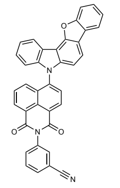
$E(S_1)$: 2.438 eV
ΔE_{ST} : 0.020 eV
$f(S_1)$: 0.0002



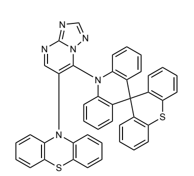
$E(S_1)$: 2.404 eV
ΔE_{ST} : 0.036 eV
$f(S_1)$: 0.0011



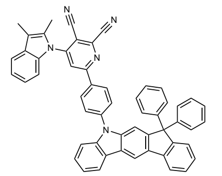
$E(S_1)$: 2.476 eV
ΔE_{ST} : 0.043 eV
$f(S_1)$: 0.0047



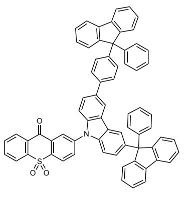
$E(S_1)$: 2.270 eV
ΔE_{ST} : 0.194 eV
$f(S_1)$: 0.0715



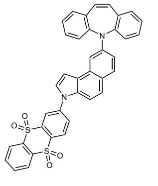
$E(S_1)$: 2.386 eV
ΔE_{ST} : 0.020 eV
$f(S_1)$: 0.0003



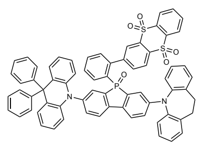
$E(S_1)$: 2.405 eV
ΔE_{ST} : 0.195 eV
$f(S_1)$: 0.2222



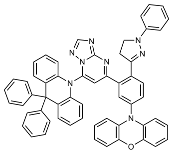
$E(S_1)$: 2.206 eV
ΔE_{ST} : 0.107 eV
$f(S_1)$: 0.0255



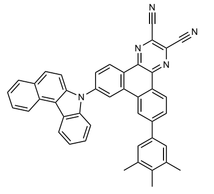
$E(S_1)$: 2.370 eV
ΔE_{ST} : 0.091eV
$f(S_1)$: 0.0189



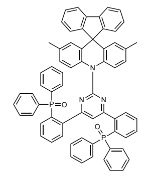
$E(S_1)$: 2.394 eV
ΔE_{ST} : 0.006 eV
$f(S_1)$: 0.0012



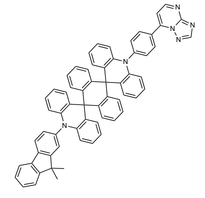
$E(S_1)$: 2.497 eV
ΔE_{ST} : 0.058 eV
$f(S_1)$: 0.0006



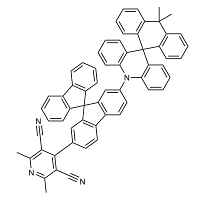
$E(S_1)$: 2.295 eV
ΔE_{ST} : 0.106 eV
$f(S_1)$: 0.0404



$E(S_1)$: 2.498 eV
ΔE_{ST} : 0.053 eV
$f(S_1)$: 0.0024

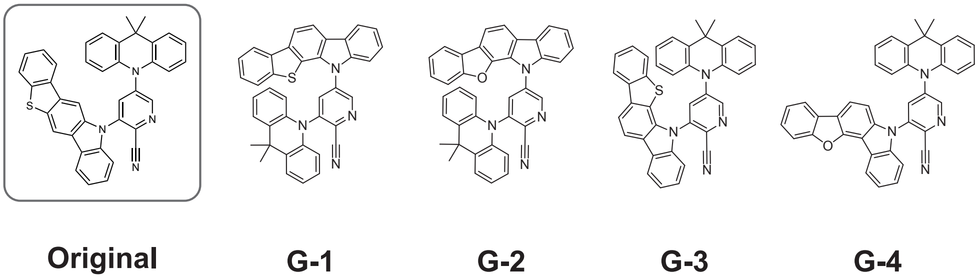


$E(S_1)$: 2.397 eV
ΔE_{ST} : 0.007 eV
$f(S_1)$: 0.0001



$E(S_1)$: 2.458 eV
ΔE_{ST} : 0.003 eV
$f(S_1)$: 0.0001

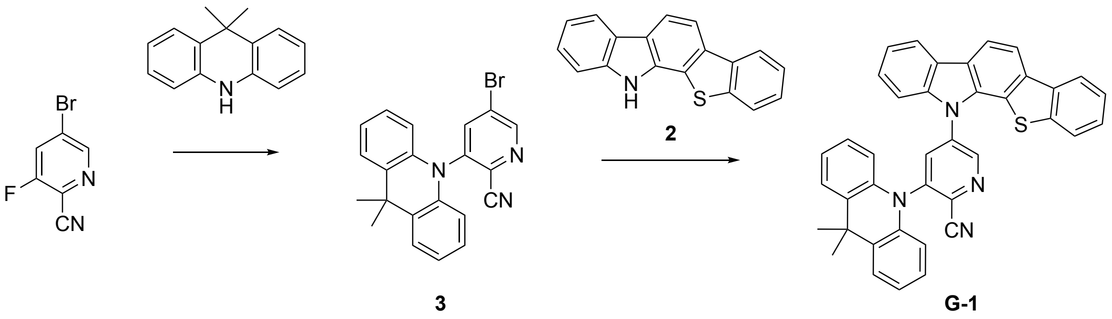
Supplementary Figure 5: 17 TADF candidates after hybrid screening and their quantum chemical properties



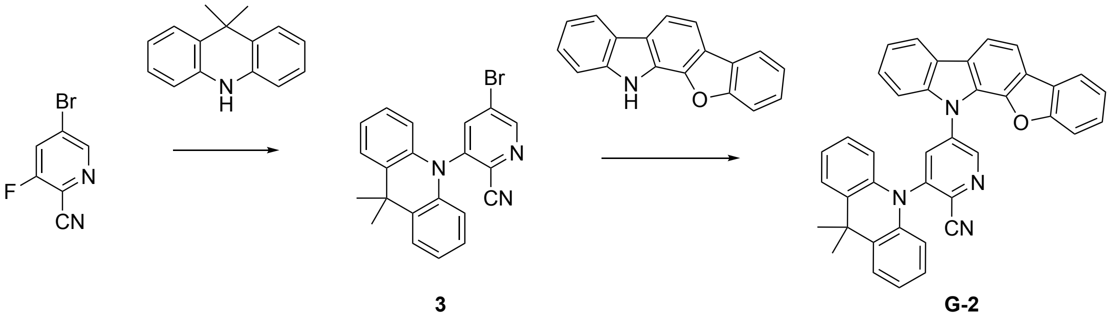
Supplementary Figure 6: Chemical structures of the original TADF candidate selected from BBAR (left) and four synthesized derivatives (G-1 to G-4). Considering building block availability and synthetic costs, we explored structural derivatives with minimal modifications.

5 Synthesis

5.1 Synthesis of G-1 and G-2

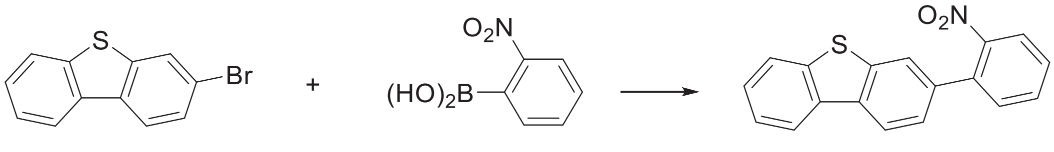


Supplementary Figure 7: Synthetic scheme of G-1



Supplementary Figure 8: Synthetic scheme of G-2

5.1.1 3-(2-nitrophenyl)dibenzo[b,d]thiophene (1)

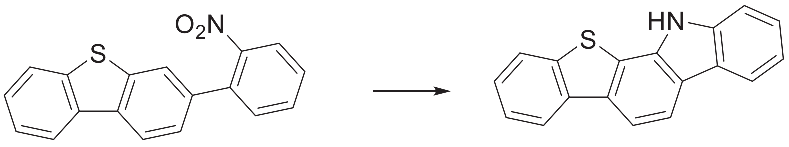


1

Supplementary Figure 9: Synthetic scheme of 3-(2-nitrophenyl)dibenzo[b,d]thiophene (1)

3-Bromodibenzo[b,d]thiophene (5 g, 19 mmol) and (2-nitrophenyl)boronic acid (6.34 g, 38 mmol), [1,1'-Bis(diphenylphosphino)ferrocene]dichloropalladium(II) (0.776 g, 0.95 mmol), and potassium phosphate (12.1 g, 57 mmol) were added to a mixture of 1,4-dioxane (80 ml) and H_2O (1.8 ml). The reaction mixture was stirred under a nitrogen atmosphere and refluxed for about 24 hours. After cooling to room temperature, the organic layer was separated and dried over anhydrous Na_2SO_4. The organic phase was concentrated with silica gel in a rotary evaporator to prepare a solid column. Column chromatography was performed using a developing solvent (dichloromethane 10% / hexane 90%) to remove impurities, yielding 3-(2-nitrophenyl)dibenzo[b,d]thiophene as a product (5 g, 86% yield).

5.1.2 12H-benzo[4,5]thieno[2,3-a]carbazole (2)

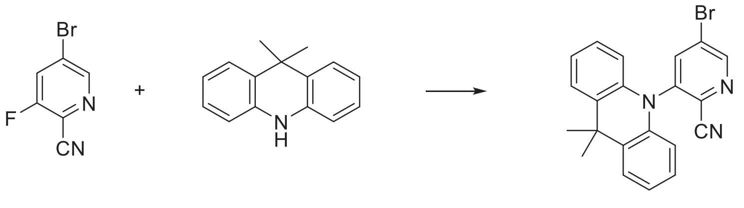


1

2

Supplementary Figure 10: Synthetic scheme of 12H-benzo[4,5]thieno[2,3-a]carbazole (2)

3-(2-nitrophenyl)dibenzo[b,d]thiophene (5 g, 16.37 mmol) and PPh_3 (12.88 g, 49.12 mmol) were added to 1,2-dichlorobenzene (81 ml) under a nitrogen atmosphere, and the mixture was refluxed with stirring for approximately 18 hours. After cooling to room temperature, the organic layer was separated and dried over anhydrous Na_2SO_4. The organic phase was concentrated with silica gel using a rotary evaporator to prepare a solid column. Column chromatography was performed using a developing solvent (dichloromethane 20% / hexane 80%) to remove impurities, yielding 12H-benzo[4,5]thieno[2,3-a]carbazole as a product (3.12 g, 70% yield).



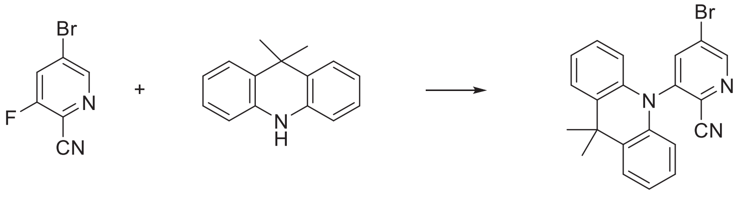
3

Supplementary Figure 11: Synthetic scheme of 5-bromo-3-(9,9-dimethylacridin-10(9H)-yl)picolinonitrile (3)

5.1.3 5-bromo-3-(9,9-dimethylacridin-10(9H)-yl)picolinonitrile (3)

NaH (1.08 g, 22.5 mmol) was suspended in DMF (30 ml) in an ice bath. Under a nitrogen atmosphere, 9,9-dimethyl-9,10-dihydroacridine (3.4 g, 16.4 mmol), diluted in DMF (30 ml), was slowly added dropwise. After stirring for 30 minutes, 5-bromo-3-fluoropicolinonitrile (3 g, 14.9 mmol), diluted in DMF (30 ml), was slowly added dropwise. The mixture was stirred at room temperature for approximately 18 hours. The reaction mixture was cooled in an ice bath, then water was added to quench the reaction. The organic layer was separated and dried over anhydrous Na_2SO_4. The organic phase was concentrated with silica gel using a rotary evaporator to prepare a solid column. Column chromatography was performed using a developing solvent (dichloromethane 10% / hexane 90%) to remove impurities, yielding 5-bromo-3-(9,9-dimethylacridin-10(9H)-yl)picolinonitrile as a product (3.6 g, 61% yield).

5.1.4 5-(12H-benzo[4,5]thieno[2,3-a]carbazol-12-yl)-3-(9,9-dimethylacridin-10(9H)-yl)picolinonitrile (G-1)



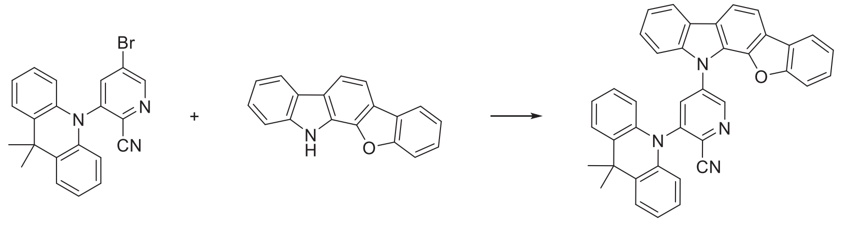
3

Supplementary Figure 12: Synthetic scheme of 5-(12H-benzo[4,5]thieno[2,3-a]carbazol-12-yl)-3-(9,9-dimethylacridin-10(9H)-yl)picolinonitrile (G-1)

5-Bromo-3-(9,9-dimethylacridin-10(9H)-yl)picolinonitrile (300 mg, 0.77 mmol), 12H-benzo[4,5]thieno[2,3-a]carbazole (252 mg, 0.92 mmol), Tris-(dibenzylideneacetone)dipalladium(0) (70 mg, 0.077 mmol), tri-tert-butylphosphonium tetrafluoroborate (45 mg, 0.154 mmol), and sodium tert-butoxide (148 mg, 1.54 mmol) were added to toluene (24 ml) under a nitrogen atmosphere. The mixture was refluxed with stirring for approximately 18 hours. After cooling to room temperature, the organic layer was separated and dried over anhydrous Na_2SO_4. The organic phase was concentrated with silica gel using a rotary evaporator to prepare a solid column. Column chromatography was performed using a developing solvent (dichloromethane 50% / hexane 50%) to remove impurities, yielding G-1 as a product (290 mg, 65% yield).

^{1}H NMR (400 MHz, $CDCl_3$): 9.21 (d, 1H), 8.26∼8.22 (m, 3H), 8.19∼8.17 (m, 2H), 8.01∼7.98 (m, 2H), 7.75 (d, 1H), 7.43(t, 2H), 7.14 (dt, 2H), 7.08 (dt, 2H), 6.43 (dd, 2H), 1.74 (s, 6H).

5.1.5 3-(2-nitrophenyl)dibenzo[b,d]thiophene (G-2)



3

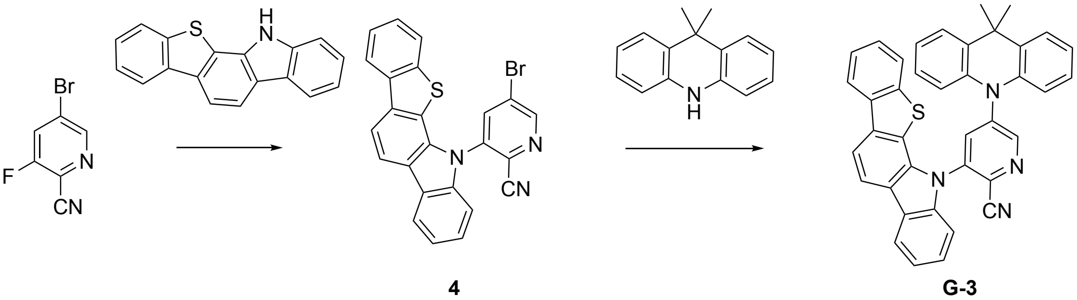
G-2

Supplementary Figure 13: Synthetic scheme of 3-(2-nitrophenyl)dibenzo[b,d]thiophene (G-2)

Intermediate 5-Bromo-3-(9,9-dimethylacridin-10(9H)-yl)picolinonitrile (300 mg, 0.77 mmol), 12H-benzofuro[2,3-a]carbazole (218 mg, 0.847 mmol), Tris(dibenzylideneacetone)dipalladium(0) (70 mg, 0.077 mmol), tri-tert-butylphosphonium tetrafluoroborate (45 mg, 0.154 mmol), and sodium tert-butoxide (148 mg, 1.54 mmol) were added to toluene (24 ml) under a nitrogen atmosphere. The mixture was refluxed with stirring for approximately 18 hours. After cooling to room temperature, the organic layer was separated and dried over anhydrous Na_2SO_4. The organic phase was concentrated with silica gel using a rotary evaporator to prepare a solid column. Column chromatography was performed using a developing solvent (dichloromethane 50% / hexane 50%) to remove impurities, yielding G-2 as a product (314 mg, 72% yield).

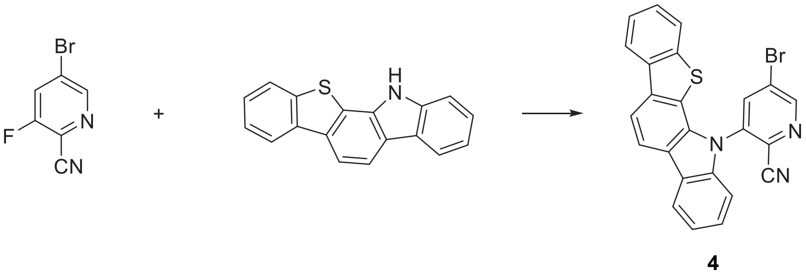
^{1}H NMR (400 MHz, $CDCl_3$): 9.37 (d, 1H), 8.36 (d, 1H), 8.24 (d, 1H), 8.14(d, 1H), 8.01∼7.99 (m, 1H), 7.95 (d, 1H), 7.69 (d, 1H), 7.56∼7.52 (m, 3H), 7.48 (t, 1H), 7.37∼7.35 (m, 2H), 7.27 (m, 1H), 7.15∼7.05 (m, 4H), 6.42 (d, 2H), 1.75 (s, 6H).

5.2 Synthesis of G-3



Supplementary Figure 14: Synthetic scheme of G-3

5.2.1 3-(12H-benzo[4,5]thieno[2,3-a]carbazol-12-yl)-5-bromopicolinonitrile (4)



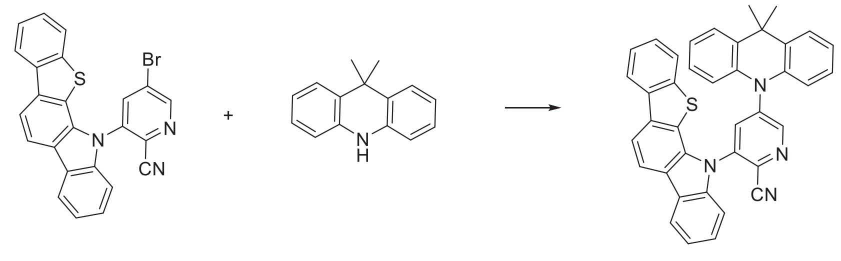
4

Supplementary Figure 15: Synthetic scheme of 3-(12H-benzo[4,5]thieno[2,3-a]carbazol-12-yl)-5-bromopicolinonitrile (4)

NaH (108 mg, 2.25 mmol) was suspended in DMF (3 ml) in an ice bath. Under a nitrogen atmosphere, 12H-benzo[4,5]thieno[2,3-a]carbazole (492 mg, 1.8 mmol), diluted in DMF (3 ml), was slowly added dropwise. After stirring for 30 minutes, 5-bromo-3-fluoropicolinonitrile (300 mg, 1.5 mmol), diluted in DMF (3 ml), was slowly added dropwise. The mixture was stirred at room temperature for approximately 18 hours. The reaction mixture was cooled in an ice bath, and water was added to quench the reaction. The organic layer was separated and dried over anhydrous Na_2SO_4. The organic phase was concentrated with silica gel using a rotary evaporator to prepare a solid column. Column chromatography was performed using a developing solvent (dichloromethane 10% / hexane 90%) to remove impurities, yielding 3-(12H-benzo[4,5]thieno[2,3-a]carbazol-12-yl)-5-bromopicolinonitrile as a product (512 mg, 75% yield).

5.2.2 3-(12H-benzo[4,5]thieno[2,3-a]carbazol-12-yl)-5-(9,9-dimethylacridin-10(9H)-yl)picolinonitrile (G-3)

Intermediate 3-(12H-benzo[4,5]thieno[2,3-a]carbazol-12-yl)-5-bromopicolinonitrile (500 mg, 1.1 mmol), 9,9-dimethyl-9,10-dihydroacridine (253 mg, 1.21 mmol), Tris(dibenzylideneacetone)dipalladium(0) (100 mg, 0.11 mmol), tri-tert-butylphosphonium tetrafluoroborate (64 mg, 0.22 mmol), and sodium tert-butoxide (211 mg, 2.2 mmol) were added to toluene (40 ml) under a nitrogen atmosphere. The mixture was refluxed with



4

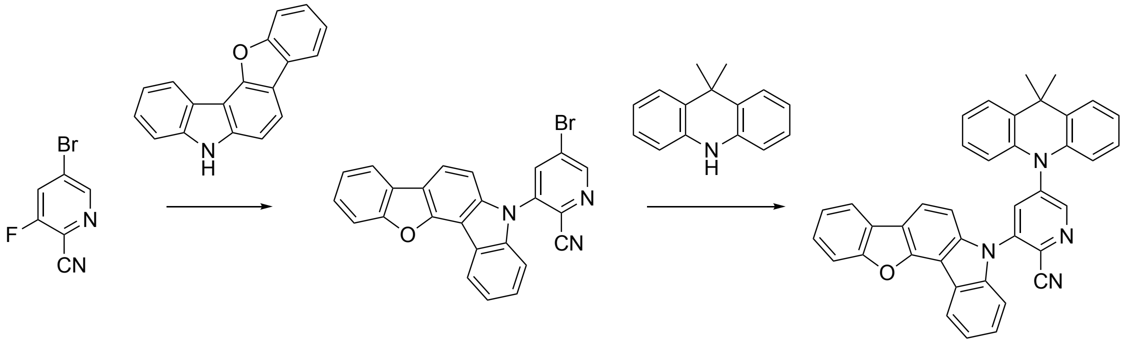
G-3

Supplementary Figure 16: Synthetic scheme of 3-(12H-benzo[4,5]thieno[2,3-a]carbazol-12-yl)-5-(9,9-dimethylacridin-10(9H)-yl)picolinonitrile (G-3)

stirring for approximately 18 hours. After cooling to room temperature, the organic layer was separated and dried over anhydrous Na_2SO_4. The organic phase was concentrated with silica gel using a rotary evaporator to prepare a solid column. Column chromatography was performed using a developing solvent (dichloromethane 60% / hexane 40%) to remove impurities, yielding G-3 as a product (460 mg, 74% yield).

^{1}H NMR (400 MHz, $CDCl_3$): 9.09 (d, 1H), 8.26∼8.20 (m, 3H), 8.16 (d, 1H), 7.90 (d, 1H), 7.76 (d, 1H), 7.52∼7.50 (m, 3H), 7.48∼7.46 (m, 1H), 7.44∼7.38 (m, 2H), 7.27∼7.23 (m, 3H), 7.22∼7.14 (m, 4H), 1.61 (s, 6H).

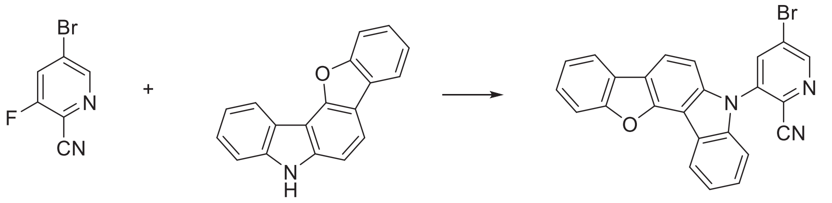
5.3 Synthesis of G-4



5 **G-4**

Supplementary Figure 17: Synthetic scheme of G-4

5.3.1 3-(5H-benzofuro[3,2-c]carbazol-5-yl)-5-bromopicolinonitrile (5)



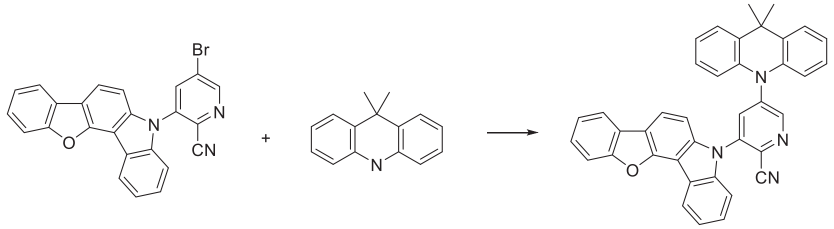
5

Supplementary Figure 18: Synthetic scheme of 3-(5H-benzofuro[3,2-c]carbazol-5-yl)-5-bromopicolinonitrile (5)

NaH (179 mg, 3.73 mmol) was suspended in DMF (3 ml) in an ice bath. Under a nitrogen atmosphere, 5H-benzofuro[3,2-c]carbazole (705 mg, 2.74 mmol), diluted in DMF (3 ml), was slowly added dropwise. After stirring for 30 minutes, 5-bromo-3-fluoropicolinonitrile (500 mg, 2.49 mmol), diluted in DMF (3 ml), was slowly added dropwise. The mixture was stirred at room temperature for approximately 18 hours. The reaction mixture was cooled in an ice bath, and water was added to quench the reaction. The organic layer was separated and dried over anhydrous Na_2SO_4. The organic phase was concentrated with silica gel using a rotary evaporator to prepare a solid column. Column chromatography was performed using a developing solvent (dichloromethane 10% / hexane 90%) to remove impurities, yielding 3-(5H-benzofuro[3,2-c]carbazol-5-yl)-5-bromopicolinonitrile as a product (773 mg, 71% yield).

5.3.2 3-(5H-benzofuro[3,2-c]carbazol-5-yl)-5-(9,9-dimethylacridin-10(9H)-yl) picolinonitrile (G-4)

Intermediate 3-(5H-benzofuro[3,2-c]carbazol-5-yl)-5-bromopicolinonitrile (400 mg, 0.91 mmol), 9,9-dimethyl-9,10-dihydroacridine (209 mg, 1 mmol), Tris(dibenzylideneacetone)dipalladium(0) (83 mg, 0.091 mmol), tri-tert-butylphosphonium tetrafluoroborate (53 mg, 0.182 mmol), and sodium tert-butoxide (175 mg, 1.82 mmol) were added to toluene (24 ml) under a nitrogen atmosphere. The mixture was refluxed with stirring



5

G-4

Supplementary Figure 19: Synthetic scheme of 3-(5H-benzofuro[3,2-c]carbazol-5-yl)-5-(9,9-dimethylacridin-10(9H)-yl)picolinonitrile (G-4)

for approximately 18 hours. After cooling to room temperature, the organic layer was separated and dried over anhydrous Na_2SO_4. The organic phase was concentrated with silica gel using a rotary evaporator to prepare a solid column. Column chromatography was performed using a developing solvent (dichloromethane 50% / hexane 50%) to remove impurities, yielding G-4 as a product (430 mg, 83% yield).

^{1}H NMR (400 MHz, $CDCl_3$): 8.96 (d, 1H), 8.57 (d, 1H), 8.02∼7.99 (m, 2H), 7.95 (d, 1H), 7.76 (d, 1H), 7.54∼7.47 (m, 5H), 7.42 (t, 1H), 7.33 (d, 1H), 7.26∼7.23 (m, 3H), 7.22$\tilde{7}$.16 (m, 4H), 1.61 (s, 6H).

6 Photophysical and Device Characterization

6.1 Photophysical characterization

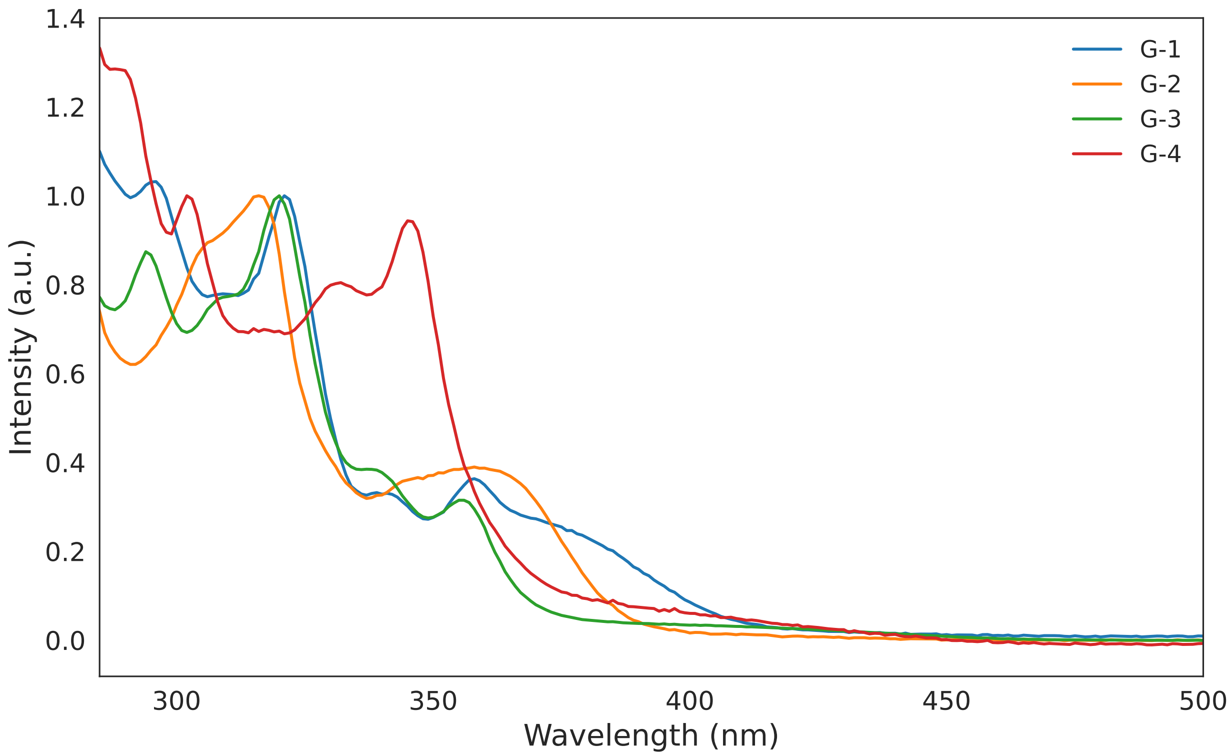
The ultraviolet-visible (UV-Vis) absorption and solution photoluminescence (PL) emission spectra of the materials were measured in dilute toluene solutions at a concentration of 1×10^{-5} M. The UV-Vis absorption spectra were obtained using an EVOLUTION 201 UV-Visible Spectrophotometer, while the PL emission spectra were recorded on a PerkinElmer FL 8500 spectrometer. Photoluminescence quantum yields were determined in dilute toluene solutions at a concentration of 1×10^{-6} M using a Hamamatsu Absolute C11347. Transient PL decay measurements were carried out with a Hamamatsu Compact Fluorescence Lifetime Spectrometer C11367. Additionally, the highest occupied molecular orbital and lowest unoccupied molecular orbital energy levels were estimated by differential pulse voltammetry using a BioLogic SP-50 system.

6.2 OLED device fabrication and characterization

Pre-patterned 150 nm-thick ITO-glass substrates were treated with dry cleaning (nitrogen plasma treatment) process. The OLED devices ITO (120 nm) / hole injection layer, (10 nm), OST01 and p-doped (10 wt%, NDP series, purchased from Novaled AG) / hole Transporting layer (90 nm), OST01 / Electron blocking layer (5 nm), SiCzCz, 9-(3-(triphenylsilyl)phenyl)-9H-3,9'-bicarbazole / EML (30 nm 1 wt%), SiCzCz:SiTrzCz2 7:3, GD 1 wt% (30 nm), SiTrzCz2 is 9,9'-(6-(3-(triphenylsilyl)phenyl)-1,3,5-triazine-2,4-diyl)bis(9H-carbazole) / HBL (5 nm), mSiTrz, / ETL (30 nm), OSE02: Liq 2:1, Liq is 8-hydroxyquinolinolato-lithium / EIL (1 nm), LiF / Al (150 nm) were fabricated by the process where the organic layers were consecutively deposited on the ITO-glass substrates at a rate of $0.1 \sim 1$ Å/s and Al electrodes were thermally evaporated at a rate of 5 Å/s under high vacuum conditions ($< 1.0 \times 10^{-6}$ torr). The device area (2×2 mm^2) is defined by the overlap between the anode and cathode electrodes. All OLED devices were encapsulated in a UV curable resin with a glass lid in an N_2 filled glovebox before device measurement. The current density-voltage-luminance (I-V-L) characteristics and electroluminescence spectra were measured with a source meter (2635B, Keithley) and spectroradiometer (CS2000 Konica Minolta).

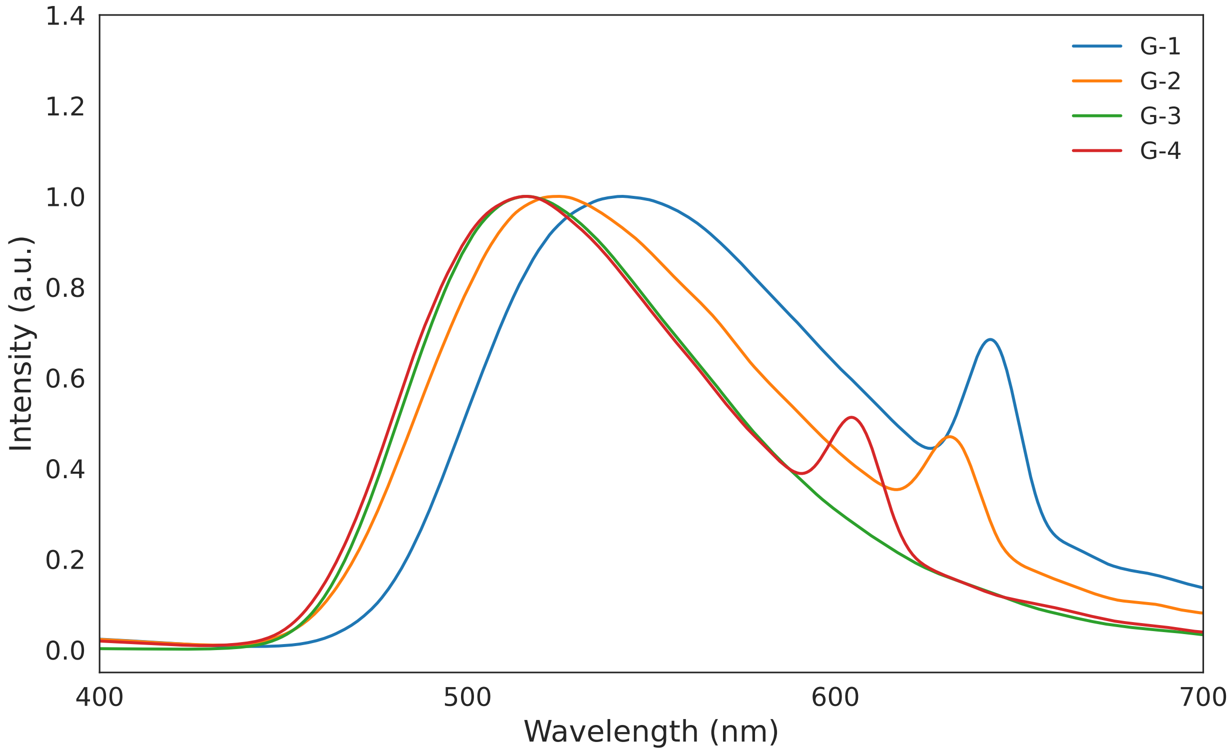
Due to confidentiality constraints, the exact molecular structures of OST01 and OSE02 cannot be disclosed.

6.3 Absorption spectra



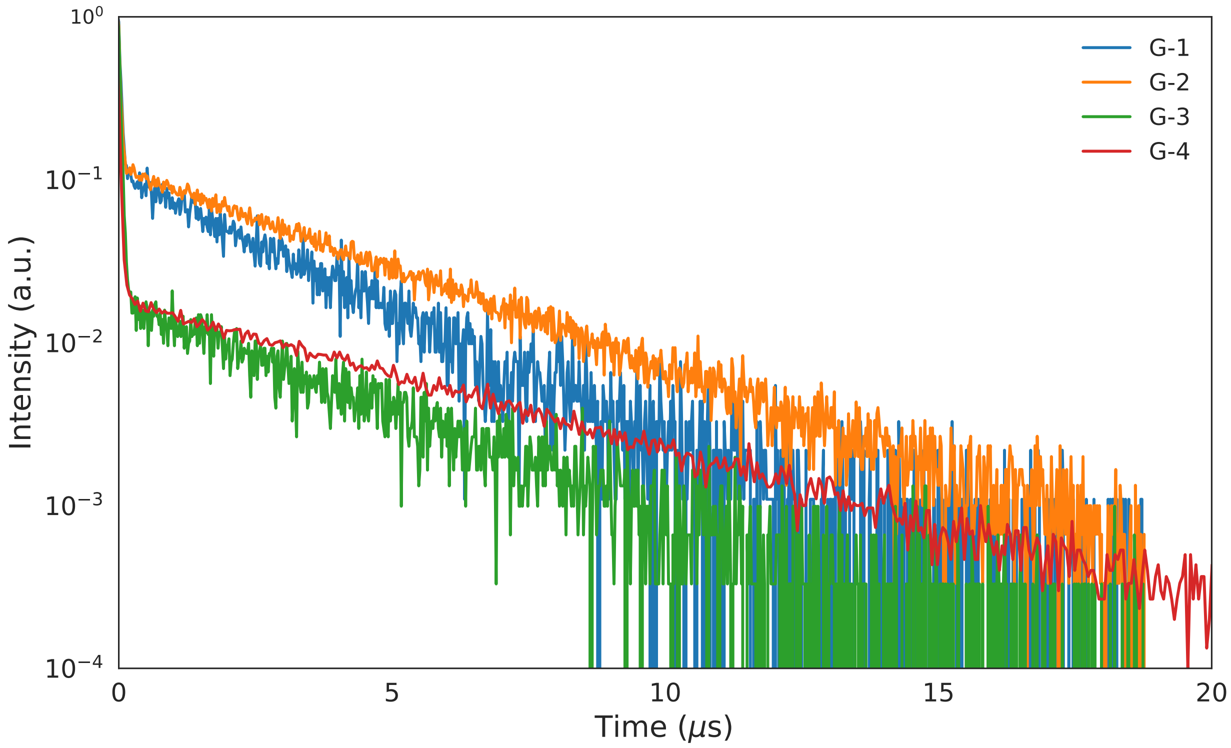
Supplementary Figure 20: Absorption spectra of synthesized TADF materials in dilute toluene solution (1×10^{-5} M)

6.4 Photoluminescence spectra



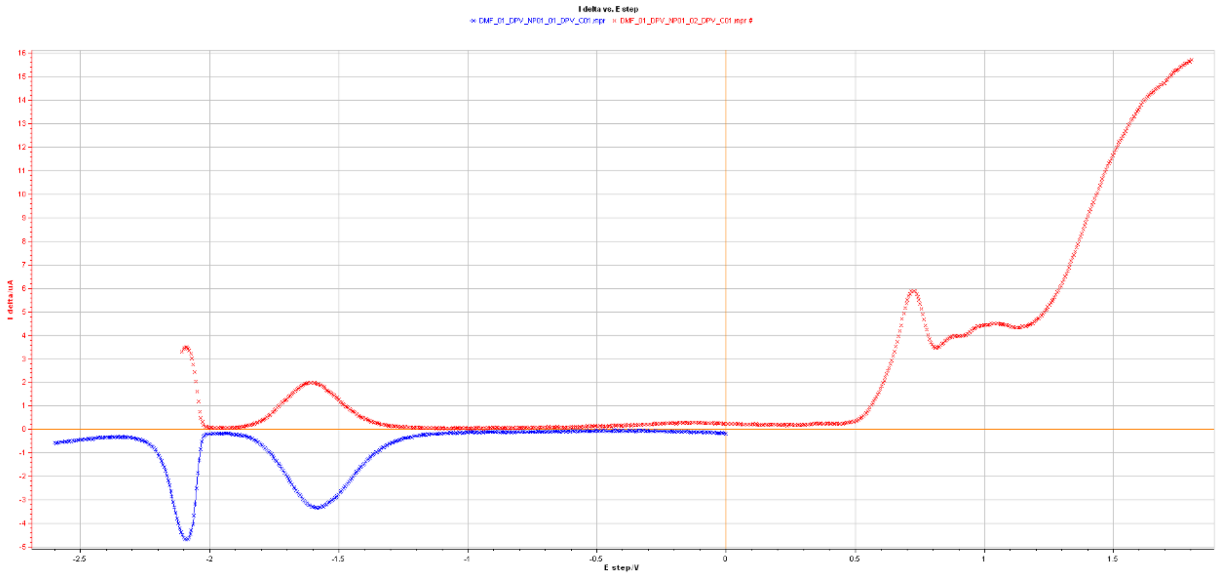
Supplementary Figure 21: Photoluminescence spetra of synthesized TADF materials in dilute toluene solution ($1{\times}10^{-5}$ M)

6.5 Transient photoluminescence spectra

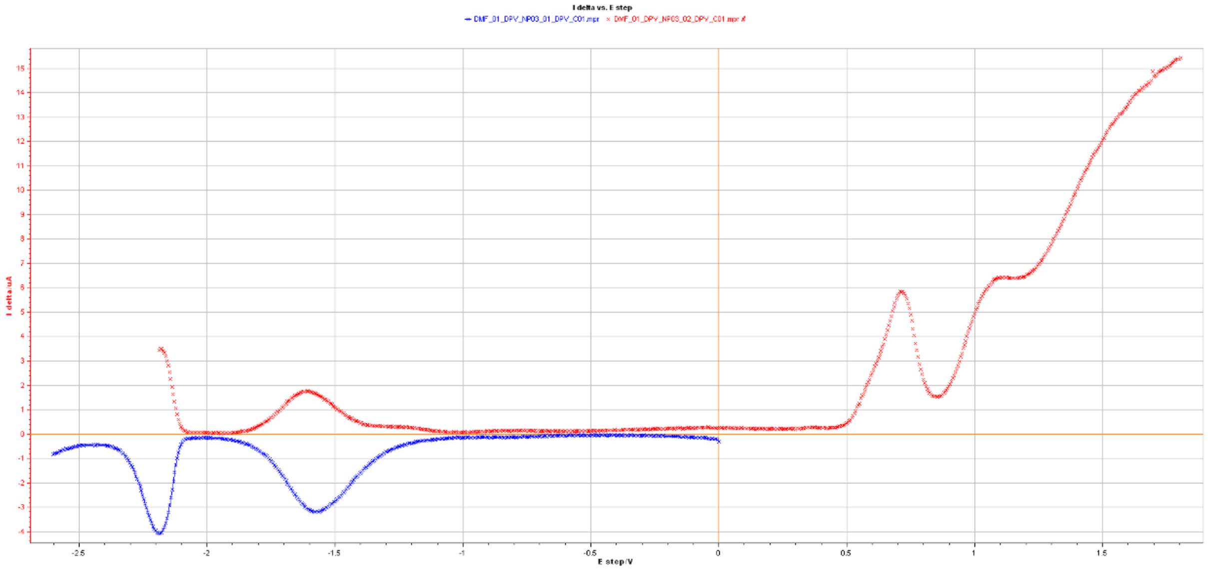


Supplementary Figure 22: Transient Photoluminescence decay of synthesized TADF materials in dilute toluene solution (1×10^{-6} M)

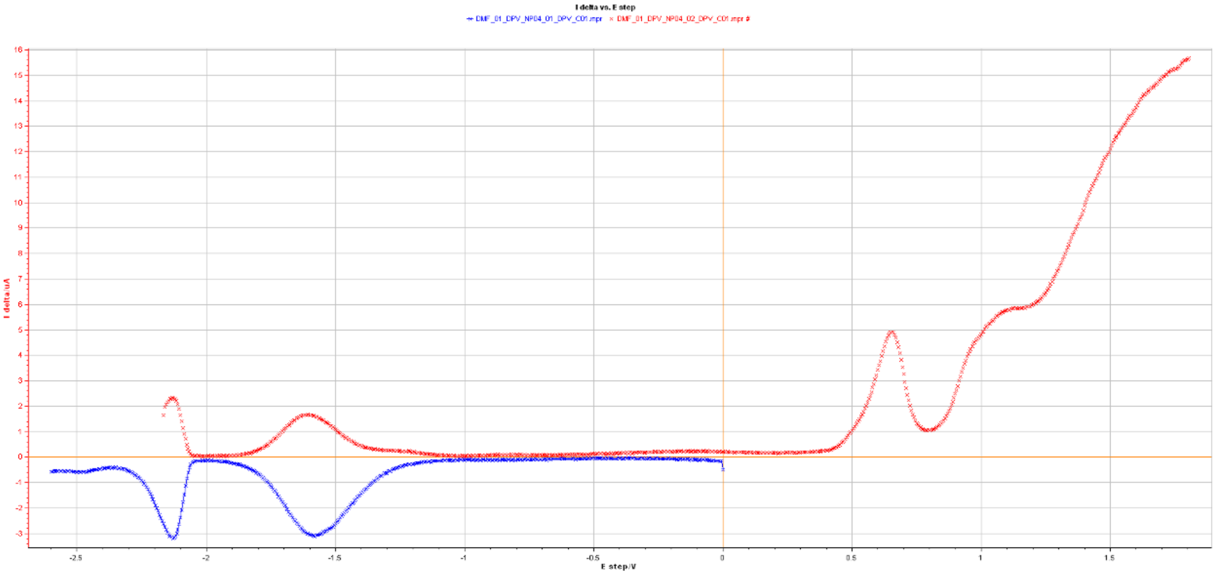
6.6 Differential pulse voltammetry Curve



Supplementary Figure 23: Differential pulse voltammetry Curve of G-1



Supplementary Figure 24: Differential pulse voltammetry Curve of G-2



Supplementary Figure 25: Differential pulse voltammetry Curve of G-3

Supplementary Table 4: Photophysical and device-level properties of green TADF materials.

	Theory[a]		Experiments						
	λ_{PL} [nm]	$\Delta\mathrm{E}_{\mathrm{ST}}$ [eV]	λ_{abs} [nm][b]	λ_{PL} [nm] [b]	τ_{PF} [μs][c]	τ_{DF} [μs][d]	Φ_{PL} [%][e]	λ_{EL} [nm][f]	η_{ext} (%)[f,g]
G-1	516.9	0.033	282.5, 321.1	541.6	0.045	2.570	25	510	9.16
G-2	499.0	0.034	283.3, 316.0	526.5	0.008	3.580	36	493	7.43
G-3	508.2	0.051	294.3, 319.8	515.6	0.046	3.193	33	497	10.85
G-4	508.8	0.020	284.9, 302.3	515.3	0.019	4.654	37	493	11.22

[a] Calculated at the level of B3LYP/6-31G(d); [b] Measured in diluted toluene solution (1×10^{-5} M);
[c] Prompt fluorescence lifetime in toluene (1×10^{-6} M); [d] Delayed fluorescence lifetime in toluene (1×10^{-6} M);
[e] PLQY in toluene (1×10^{-6} M); [f] Measured in device; [g] EQE at 1,000 cd·m^{-2}.