

Supporting information

**Multi-level Δ -learning for Predicting Experimental Radiative Decay Rate Constant
of Phosphorescent Platinum(II) Complexes**

Shuai Wang¹, ChiYung Yam^{2,3,#}, LiHong Hu⁴, Faan-Fung Hung^{1,2}, Shuguang Chen^{1,2}, Chi-Ming

Che^{1,2,&} and GuanHua Chen^{1,2}*

¹Department of Chemistry, The University of Hong Kong, Pokfulam, Hong Kong SAR, China

²Hong Kong Quantum AI Lab Limited, Pak Shek Kok, Hong Kong SAR, China

³Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China,

Shenzhen, 518000, China

⁴School of Information Science and Technology, Northeast Normal University, Changchun, 130117,

China

*Email: ghc@everest.hku.hk (G.H.C.)

#Email: yamcy@uestc.edu.cn (C.Y.Y)

&Email: cmche@hku.hk (C.M.C.)

Contents

1. Generation details.....	3
2. ML hyperparameters tuning details	5
3. Performance and discussions of k_r prediction models.....	7
4. Discussion of the data	13
5. External evaluation of optimal experimental k_r prediction models.....	15
6. Partial code of rMolCLR.....	16
7. Different partitioning ratio and learning curve for k_r^f prediction.....	17
8. Computational details for k_r^f and the features employed for predicting k_r^e	19
9. Feature importance analysis.....	22
10. Discussions about the 51 top candidates.....	24
11. Benchmark against additional functionals	29
12. References	32

1. Generation details

With respect to the generation strategy, the basic route is to substitute the H atoms with R-groups on the core structures at the highlighted sites. For cores structures B, C, F, J, and K, the generations are in a symmetric manner, which means the substitution at the symmetric sites will be executed with the same R-groups. For core A, the substitutions at the bottom-left and bottom-right are kept symmetrically. For the cores D, E, and L, the two substitutions within one ring will be conducted using the same R-group. In the remaining G, H and I core structures, each H atom(s) at the substitution sites can be changed with any R-group in the dataset. Besides, the number of substitutions in one core structure will not exceed the cube of the number of the R-groups.

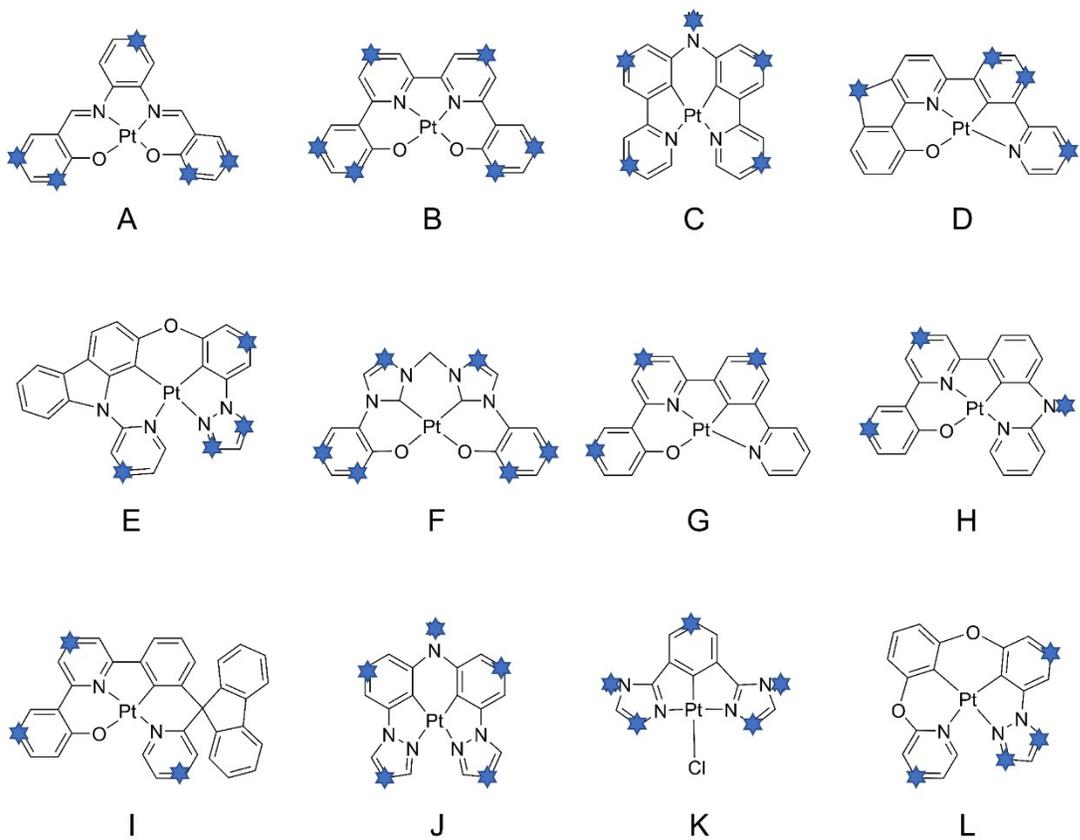


Fig. S1. Detailed substitution sites in the core structures

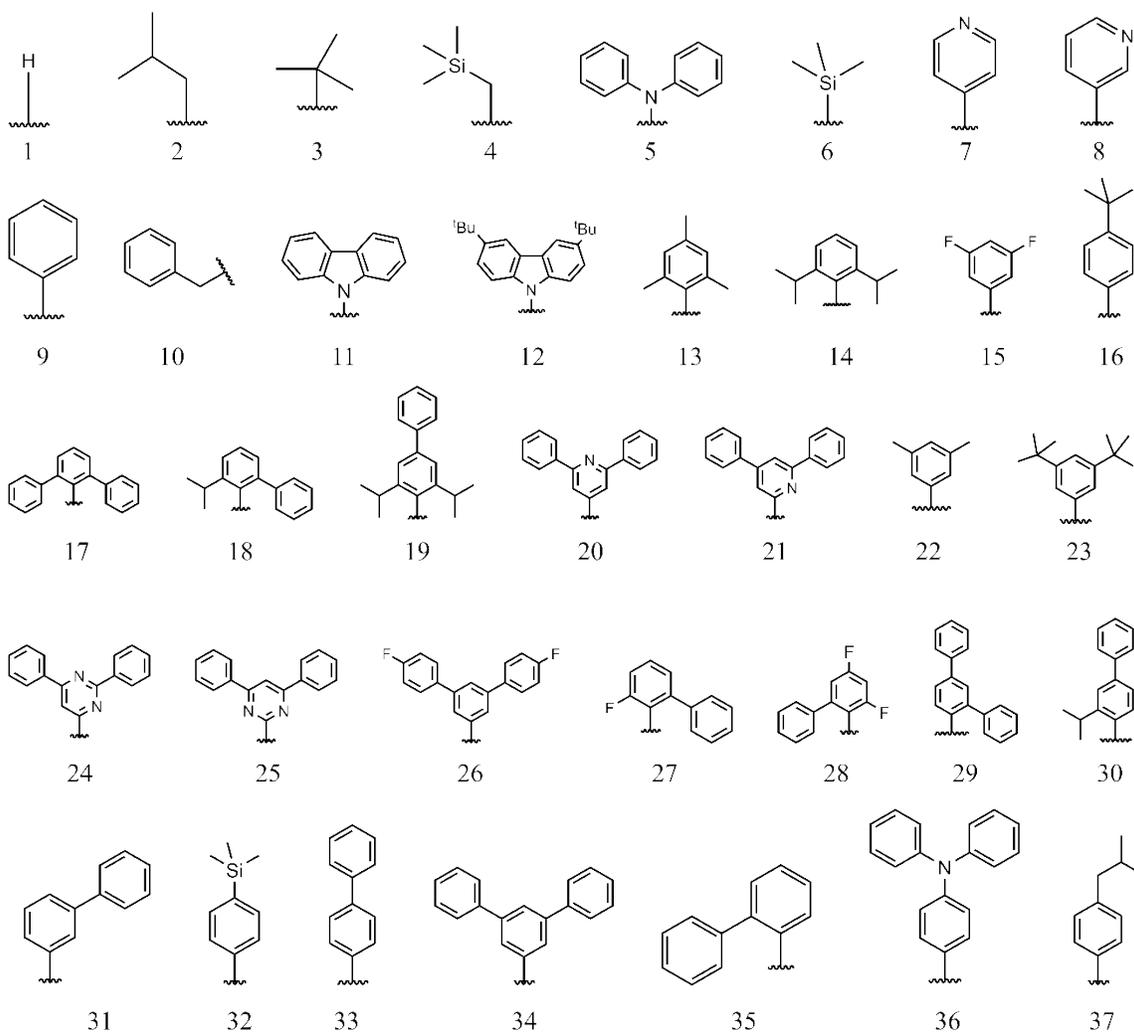


Fig. S2. All R-groups for generating molecules

2. ML hyperparameters tuning details

Hyperopt¹ was used to obtain the best hyperparameters for different ML algorithms.

XGBoost was trained under the following predefined ranges:

'gamma':(0, 0.5), 'max_depth': (3, 11), 'min_child_weight': (1, 20),

'colsample_bytree': (0.5, 1), 'subsample': (0.5, 1), 'learning_rate': (0.001, 0.2),

'n_estimators': (10, 500), 'max_delta_step': (0.5, 1), 'reg_alpha': (0, 0.5),

'reg_lambda': (0.5, 1), 'scale_pos_weight': (0, 0.2)

SVM was trained under the following predefined ranges:

'C': (1e-5, 1e2), 'gamma': (1e-5, 1e2), 'epsilon': (1e-5, 1)

KNN (weights='uniform') was trained under the following predefined ranges:

'weights': ('uniform'), 'n_neighbors': (5, 20), 'leaf_size': (1, 20)

LightGBM was trained under the following predefined ranges:

'num_leaves': (2, 13), 'learning_rate': (0.00001, 0.2), 'min_child_samples': (0, 50),

'max_depth': (0, 13), 'n_estimators': (10, 500), 'bagging_fraction': (0.5, 1)

KRR was trained under the following predefined ranges:

'alpha': (0, 3)

Adaboost was trained under the following predefined ranges:

'learning_rate': (0.00001, 0.2), 'n_estimators': (10, 500)

3. Performance and discussions of k_r prediction models

Table S1. The optimal four models' performance of predicting first-principle calculated k_r with 32 latent vectors based on GCN encoder*

Model type	Training			Testing		
	RMSE	MAE	R ²	RMSE	MAE	R ²
AdaBoost	0.42±0.02	0.35±0.01	0.79±0.01	0.41±0.02	0.35±0.01	0.75±0.09
XGB	0.21±0.00	0.16±0.00	0.94±0.00	0.31±0.03	0.19±0.01	0.62±0.08
LGBM	0.22±0.01	0.16±0.01	0.93±0.01	0.29±0.03	0.20±0.02	0.66±0.07
KNN	0.14±0.00	0.09±0.00	0.98±0.00	0.30±0.03	0.17±0.01	0.65±0.07

*Errors are measured in s⁻¹ in log scale.

Table S2. The optimal four models' performance of predicting first-principle calculated k_r with 64 latent vectors based on GCN encoder*

Model type	Training			Testing		
	RMSE	MAE	R ²	RMSE	MAE	R ²
KNN	0.14±0.00	0.09±0.00	0.98±0.00	0.26±0.03	0.16±0.01	0.81±0.04
LGBM	0.35±0.01	0.24±0.01	0.84±0.01	0.38±0.04	0.25±0.01	0.61±0.10
RF	0.36±0.01	0.22±0.01	0.84±0.01	0.37±0.03	0.23±0.01	0.63±0.08
XGB	0.20±0.00	0.15±0.00	0.95±0.00	0.35±0.04	0.24±0.01	0.68±0.08

*Errors are measured in s⁻¹ in log scale.

Table S3. The optimal four models' performance of predicting first-principle calculated k_r with 128 latent vectors based on GCN encoder*

Model type	Training			Testing		
	RMSE	MAE	R ²	RMSE	MAE	R ²
KNN	0.15±0.01	0.09±0.00	0.97±0.00	0.26±0.02	0.16±0.01	0.87±0.02
LGBM	0.07±0.01	0.04±0.00	0.99±0.00	0.24±0.03	0.16±0.01	0.87±0.04
RF	0.33±0.01	0.20±0.00	0.86±0.01	0.31±0.02	0.18±0.01	0.83±0.03
XGB	0.16±0.01	0.11±0.00	0.96±0.00	0.26±0.01	0.17±0.01	0.86±0.01

*Errors are measured in s⁻¹ in log scale.

Table S4. The optimal four models' performance of predicting first-principle calculated k_r with 256 latent vectors based on GCN encoder*

Model type	Training			Testing		
	RMSE	MAE	R ²	RMSE	MAE	R ²
KNN	0.12±0.00	0.08±0.00	0.98±0.00	0.21±0.01	0.14±0.00	0.75±0.02
LGBM	0.12±0.01	0.08±0.00	0.98±0.00	0.21±0.02	0.15±0.01	0.75±0.05
RF	0.26±0.01	0.17±0.00	0.93±0.00	0.19±0.01	0.14±0.00	0.79±0.02
XGB	0.15±0.01	0.10±0.00	0.97±0.00	0.20±0.01	0.15±0.01	0.78±0.03

*Errors are measured in s⁻¹ in log scale.

Table S5. The optimal four models' performance of predicting first-principle calculated k_r with 512 latent vectors based on GCN encoder*

Model type	Training			Testing		
	RMSE	MAE	R ²	RMSE	MAE	R ²
KNN	0.14±0.00	0.09±0.00	0.98±0.00	0.25±0.02	0.15±0.01	0.76±0.04
LGBM	0.25±0.01	0.18±0.01	0.92±0.01	0.20±0.01	0.15±0.01	0.86±0.02
RF	0.26±0.00	0.17±0.00	0.92±0.00	0.20±0.01	0.14±0.00	0.86±0.01
XGB	0.18±0.00	0.13±0.00	0.96±0.00	0.21±0.01	0.15±0.01	0.87±0.01

*Errors are measured in s⁻¹ in log scale.

Table S6. The optimal four models' performance of predicting first-principle calculated k_r with 32 latent vectors based on GIN encoder*

Model type	Training			Testing		
	RMSE	MAE	R ²	RMSE	MAE	R ²
KNN	0.27±0.01	0.17±0.00	0.90±0.01	0.42±0.03	0.24±0.01	0.76±0.03
LGBM	0.22±0.01	0.16±0.01	0.92±0.01	0.44±0.04	0.27±0.02	0.73±0.05
RF	0.30±0.01	0.19±0.00	0.89±0.01	0.46±0.02	0.25±0.01	0.75±0.04
XGB	0.29±0.01	0.20±0.00	0.87±0.01	0.44±0.02	0.28±0.01	0.75±0.03

*Errors are measured in s⁻¹ in log scale.

Table S7. The optimal four models' performance of predicting first-principle calculated k_r with 64 latent vectors based on GIN encoder*

Model type	Training			Testing		
	RMSE	MAE	R ²	RMSE	MAE	R ²
AdaBoost	0.40±0.01	0.33±0.01	0.77±0.01	0.55±0.04	0.44±0.03	0.71±0.10
LGBM	0.22±0.01	0.17±0.00	0.93±0.00	0.38±0.04	0.23±0.01	0.75±0.06
RF	0.34±0.01	0.21±0.00	0.85±0.01	0.41±0.02	0.22±0.01	0.68±0.04
XGB	0.18±0.00	0.13±0.00	0.95±0.00	0.35±0.03	0.19±0.01	0.79±0.05

*Errors are measured in s⁻¹ in log scale.

Table S8. The optimal four models' performance of predicting first-principle calculated k_r with 128 latent vectors based on GIN encoder*

Model type	Training			Testing		
	RMSE	MAE	R ²	RMSE	MAE	R ²
KNN	0.14±0.01	0.09±0.00	0.98±0.00	0.24±0.02	0.15±0.01	0.79±0.04
LGBM	0.13±0.00	0.10±0.00	0.98±0.00	0.24±0.02	0.17±0.01	0.81±0.03
RF	0.32±0.01	0.20±0.01	0.88±0.01	0.26±0.01	0.17±0.01	0.76±0.02
XGB	0.19±0.00	0.13±0.00	0.95±0.00	0.22±0.01	0.16±0.01	0.83±0.02

*Errors are measured in s⁻¹ in log scale.

Table S9. The optimal four models' performance of predicting first-principle calculated k_r with 256 latent vectors based on GIN encoder*

Model type	Training			Testing		
	RMSE	MAE	R ²	RMSE	MAE	R ²
KNN	0.16±0.00	0.10±0.00	0.97±0.00	0.27±0.02	0.16±0.01	0.69±0.04
LGBM	0.16±0.01	0.12±0.00	0.96±0.00	0.22±0.01	0.17±0.01	0.83±0.02
RF	0.31±0.01	0.19±0.01	0.90±0.01	0.22±0.01	0.16±0.01	0.84±0.02
XGB	0.14±0.00	0.10±0.00	0.97±0.00	0.23±0.02	0.18±0.01	0.78±0.03

*Errors are measured in s⁻¹ in log scale.

Table S10. The optimal four models' performance of predicting first-principle calculated k_r with 512 latent vectors based on GIN encoder*

Model type	Training			Testing		
	RMSE	MAE	R ²	RMSE	MAE	R ²
KNN	0.14±0.00	0.08±0.00	0.98±0.00	0.22±0.01	0.14±0.00	0.77±0.03
LGBM	0.16±0.01	0.10±0.00	0.96±0.00	0.22±0.01	0.15±0.01	0.77±0.02
RF	0.31±0.01	0.19±0.00	0.88±0.01	0.20±0.01	0.15±0.01	0.81±0.02
XGB	0.20±0.01	0.13±0.00	0.94±0.00	0.21±0.01	0.15±0.01	0.79±0.02

*Errors are measured in s⁻¹ in log scale.

Table S11. The optimal models' performance of predicting experimental k_r with unsupervised learned latent vectors based on different encoders and dimensions of latent vectors*

Model	Settings of rMolCLR	Independent testing set		
		RMSE	MAE	R ²
RF	GCN_32	0.34±0.01	0.26±0.01	0.17±0.03
RF	GCN_64	0.33±0.01	0.25±0.01	0.24±0.04
AdaBoos	GCN_128	0.31±0.01	0.25±0.01	0.36±0.05
RF	GCN_256	0.32±0.01	0.27±0.01	0.28±0.03
KNN	GCN_512	0.31±0.00	0.25±0.00	0.27±0.03
RF	GIN_32	0.34±0.01	0.27±0.01	0.17±0.05
KNN	GIN_64	0.33±0.01	0.27±0.01	0.20±0.04
XGB	GIN_128	0.32±0.01	0.26±0.01	0.30±0.07
KNN	GIN_256	0.31±0.01	0.25±0.01	0.30±0.04
KNN	GIN_512	0.29±0.01	0.22±0.01	0.36±0.05

*Errors are measured in s⁻¹ in log scale.

For the first-principles calculated radiative decay rate constants, our models were rigorously trained using 10-fold cross-validation, resulting in an R² value exceeding 0.75 with very low standard deviations. These metrics underscore the stability and reliability of the models employed.

Regarding the predictions of experimental radiative decay rate constants, our models have demonstrated excellent predictive performance on both independent and external testing sets, which were not included in the training process. This validation further underscores the reliability and generalization capabilities of the multi-level Δ -learning models.

Furthermore, to demonstrate the high radiative rate constants of the three promising candidates, we selected a Pt-NCCN complex² as a reference due to its similar molecular structure to the aforementioned candidates. This reference Pt-complex, denoted as Pt-ref, exhibits a high experimental rate constant (Exp. k_r) value of $1.30 \times 10^5 \text{ s}^{-1}$ in CH_2Cl_2 . The three promising candidates (Pt-1st to Pt-3rd) and the Pt-ref complex share the same core structure, featuring a tetradentate NCCN ligand with two oxygen bridges, as illustrated in Fig. S3. The radiative decay-related properties obtained from TDDFT calculations are collected and compared in Table S12.

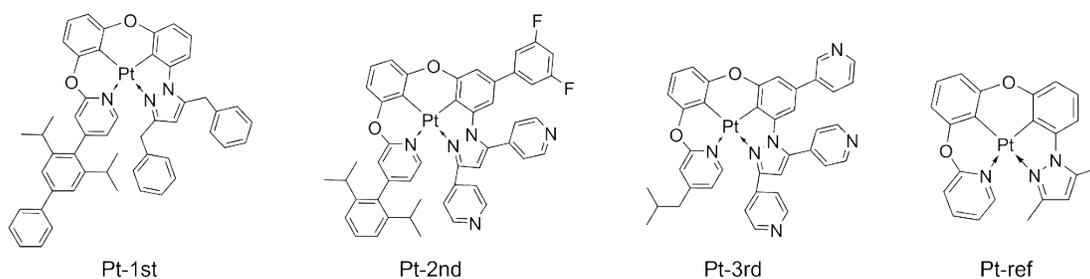


Fig. S3. Structures of the promising candidates and the reference complex

Table S12. Computational comparisons based on TDDFT calculations*

	ML-pred k_r	Exp. k_r	f	μ	$H_{\text{SOC_}S_1_T_1}$	S_1 - T_1 gap	ν	Medium
Pt-1st	5.36	-	0.00069	11.54	336.23	0.20	2.61	CH_2Cl_2
Pt-2nd	5.32	-	0.00079	11.32	276.11	0.26	2.21	CH_2Cl_2
Pt-3rd	5.30	-	0.00095	12.56	276.74	0.25	2.19	CH_2Cl_2
Pt-ref	5.17	5.11	0.00062	10.37	322.21	0.21	2.73	CH_2Cl_2

*ML-pred k_r and Exp. k_r represent ML-predicted k_r values by the proposed protocol and experimental k_r values, respectively, both measured in s^{-1} in log scale. f is oscillator strength from T_1 to S_0 state. μ denotes the transition dipole moment from S_1 to S_0 state, measured in Debye. $H_{\text{SOC_}S_1_T_1}$ stands for the spin-orbit coupling constant from S_1 to T_1 state. S_1 - T_1 gap represents the energy gap between S_1 and T_1 state while ν represents the emission energy from T_1 to S_0 state, both measured in eV.

In terms of evaluating Pt-ref, our protocol achieves outstanding performance with a tiny error of 0.06 when compared with the experimental observations, underscoring its precision in predicting Pt-NCCN complexes with oxygen bridges, including the promising complexes (Pt-1st to Pt-3rd). Moreover, the calculation of k_r values relies on the oscillator strength (f) and emission energy (ν), as described in Equation 1 of the Supplementary Information. Notably, the promising candidates (Pt-1st to Pt-3rd) all exhibit higher f values compared to Pt-ref. The oscillator strength (f) is primarily determined by three factors: the transition dipole moment from the S_1 to S_0 state (μ), the spin-orbit coupling constant from the S_1 to T_1 state ($H_{\text{SOC}}_{S_1-T_1}$), and the energy gap between the S_1 and T_1 states (S_1 - T_1 gap). Specifically, both μ and $H_{\text{SOC}}_{S_1-T_1}$ positively correlate with f , while the S_1 - T_1 gap negatively contributes to f . By comparing the simulated results of these three factors, it is evident that Pt-1 to Pt-3 exhibit superior performance compared to Pt-ref. Considering the emission energy from T_1 to S_0 state, Pt-1st closely matches Pt-ref, while Pt-2 and Pt-3 show lower values. In summary, due to the capacity and robustness of the multi-level Δ -learning protocol and substantial advantages on oscillator strength and transition dipole moment, as well as superior or comparable performance in spin-orbit coupling constant, S_1 - T_1 gap, and emission energy, the three promising complexes are expected to exhibit high k_r values from a computational standpoint. Next, we will collaborate with experimental groups to validate the high radiation performance of these promising candidates.

4. Discussion of the data

The experimental radiative decay rate constant (k_r) data used in our study has been carefully taken from peer-reviewed literature and has undergone rigorous preprocessing before being incorporated into the training procedures. To ensure data integrity, we carefully evaluated the consistency of reported k_r values. In cases where identical samples were cited across different publications but exhibited significant discrepancies, we excluded those entries from our dataset to maintain a homogeneous set of values. For instance, the k_r and PLQY of complex Pt[ppy-O-POPy] are reported to be $3.15 \times 10^5 \text{ s}^{-1}$ and 0.63, respectively in Ref.² while they are $4.15 \times 10^5 \text{ s}^{-1}$ and 0.83, respectively in Ref.³. This preprocessing confirmed that the selected experimental values are representative and exhibit a reasonable degree of consistency across the dataset. Furthermore, all experimental k_r values were obtained from the ratio of photoluminescence quantum yield to lifetime. The measurements of photoluminescence quantum yield and lifetime follow standard methods and well-established operating procedures using photoluminescence spectrometers and emission lifetime instruments. Thus, the experimental k_r values can be considered homogeneous with low observational errors.

In terms of first-principles calculated k_r values, the details of computational methods, including the choice of functional and basis set, is consistent with those employed in prior studies⁴. Currently, methodologies for evaluating radiative decay rate constants vary, and consistency in their application remains a topic of ongoing discussion in the field.

However, we believe that the general trends and correlations observed between our first-principles calculated results and experimental data are acceptable and meaningful.

In fact, for high-throughput virtual screening, our primary objective was to efficiently identify potential candidates from a comprehensive database of Pt structures. The subsequent step, which focuses on the accurate prediction of experimental k_r values, is crucial for validating and refining the results for the most promising candidates. This two-step approach enhances the reliability of our predictions and ensures that we can effectively prioritize compounds with desirable properties.

Table S13. First-principles calculated k_r (in s^{-1}) distribution

k_r range	Count
[1, 10)	1
[10, 100)	0
[100, 1000)	15
[1000, 10000)	18
[10000, 100000)	131
[100000, 1000000)	303
[1000000, 10000000)	1

Table S13 illustrates the distribution of first-principles calculated k_r (k_r^f) and it is evident that most calculated k_r values span in the range between 10^4 and $10^6 s^{-1}$. With respect to experimental k_r distributions, please refer to the previous paper⁴.

5. External evaluation of optimal experimental k_r prediction models

The same division of data as previous work⁴ into training, testing, and external testing (24 out-of-sample Pt-complexes) is adopted. Three top models on the independent testing set are selected for external generalization testing. Table S14 shows the evaluation results of the generalization and the best performance on the external set is plotted in Fig. S4. Among the three generalization evaluation results, the model with semi-supervised predicted k_r^f from GCN encoder and 256 latent vectors demonstrates the optimal generalization and performs the best on both the independent testing and external testing sets, thus is chosen as the optimal model for k_r^e predictions.

Table S14. Generalization performance evaluation*

Model type	Settings of rMolCLR	External testing set		
		RMSE	MAE	R ²
XGB	GCN_256	0.218	0.187	0.658
XGB	GIN_32	0.220	0.189	0.651
RF	GIN_512	0.270	0.213	0.536

*Errors are measured in s⁻¹ in log scale.

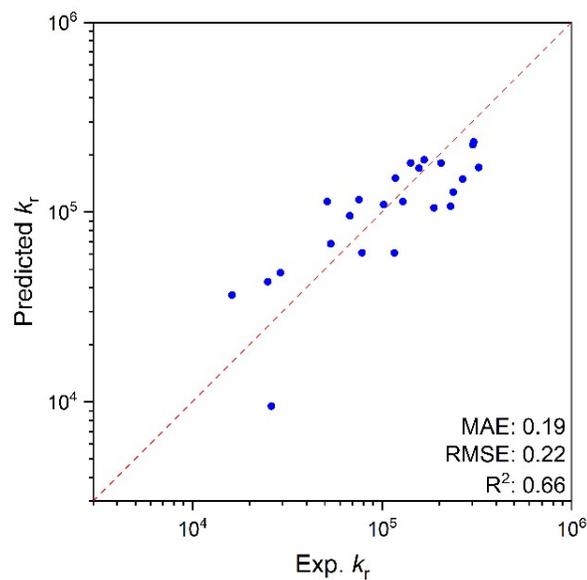


Fig. S4. XGB performances of experimental k_r (in s^{-1}) prediction model on the external testing set with errors measured in s^{-1} in log scale

6. Partial code of rMolCLR

The upgraded framework considers the coordination bonds for all the Pt-complexes, whether involving tridentate or tetradentate ligands, as well as the specific types of coordination atoms. In the code, `fromatoms = (6, 7, 8)` corresponds with the atoms C, N and O, which coordinate with the central metal atom (Pt). Here, we don't simply take all the bonds linking with Pt as coordination bonds, but the bonds involving the atoms with abnormal valence and special cases (such as $Pt - C \equiv N$). In this way, rMolCLR can better showcase the intrinsic nature of Pt(II) complexes. All the code can be found in the data availability section.

```

def set_dative_bonds(mol, fromAtoms=(6, 7, 8)):
    """ convert some bonds to dative

    Replaces some single bonds between metal (Pt) and other atoms with atomic numbers in fromAtoms
    with dative bonds. The replacement is done when the atom with unusual valence and special occasions.

    Returns the modified molecule.

    """
    pt = Chem.GetPeriodicTable()
    rwmol = Chem.RWMol(mol)
    rwmol.UpdatePropertyCache(strict=False)
    metals = [at for at in rwmol.GetAtoms() if is_transition_metal(at)]

    for metal in metals:
        for nbr in metal.GetNeighbors():
            if nbr.GetAtomicNum() in (7,):
                nbr.SetFormalCharge(0)

            if nbr.GetAtomicNum() in fromAtoms and \
                nbr.GetExplicitValence()!=pt.GetDefaultValence(nbr.GetAtomicNum()) and \
                rwmol.GetBondBetweenAtoms(nbr.GetIdx(),metal.GetIdx()).GetBondType() == Chem.BondType.SINGLE:
                rwmol.RemoveBond(nbr.GetIdx(),metal.GetIdx())
                rwmol.AddBond(nbr.GetIdx(),metal.GetIdx(),Chem.BondType.DATIVE)

    return rwmol

def set_formal_charge(mol):
    rwmol = Chem.RWMol(mol)
    rwmol.UpdatePropertyCache(strict=False)
    for atom in rwmol.GetAtoms():
        if atom.GetFormalCharge() != 0:
            atom.SetFormalCharge(0)
    return rwmol

```

Fig. S5. Partial code to recognize coordination bonds for Pt-complexes in rMolCLR

7. Different partitioning ratio and learning curve for k_r^f prediction

Table S15. Model performance based on different partitioning ratio

Model	Partition	test_R ²	test_MAE	test_RMSE
XGB	6 : 4	0.774	0.239	0.392
XGB	7 : 3	0.817	0.204	0.312
XGB	8 : 2	0.778	0.147	0.195
XGB	9 : 1	0.881	0.108	0.144

*Errors are measured in s⁻¹ in log scale.

When adjusting the portioning ratios of training set (from 6:4 to 9:1), no big change for correlation coefficients (all around 0.8 except 9:1), which demonstrates great robustness of the model across various data distributions. With respect to the MAE and RMSE, the errors go down when increasing the training size and decreasing the testing size, which is

common for ML models. Taking both the model performance and further generalization for consideration, the partitioning ratio of 8:2 is selected.

The learning curve is plotted in Figure S6, which indicates the XGB model learns well with increasing the training data when keeping the test set constant (20% of the whole dataset via SPXY partitioning method). In Figure S6, RMSE values are plotted based on the incremental increase in the number of training samples (from 20% to 100% of the training set, with a 10% interval). The RMSE values drop down clearly when enlarging the training set which demonstrates the XGB models learns well with the training data.

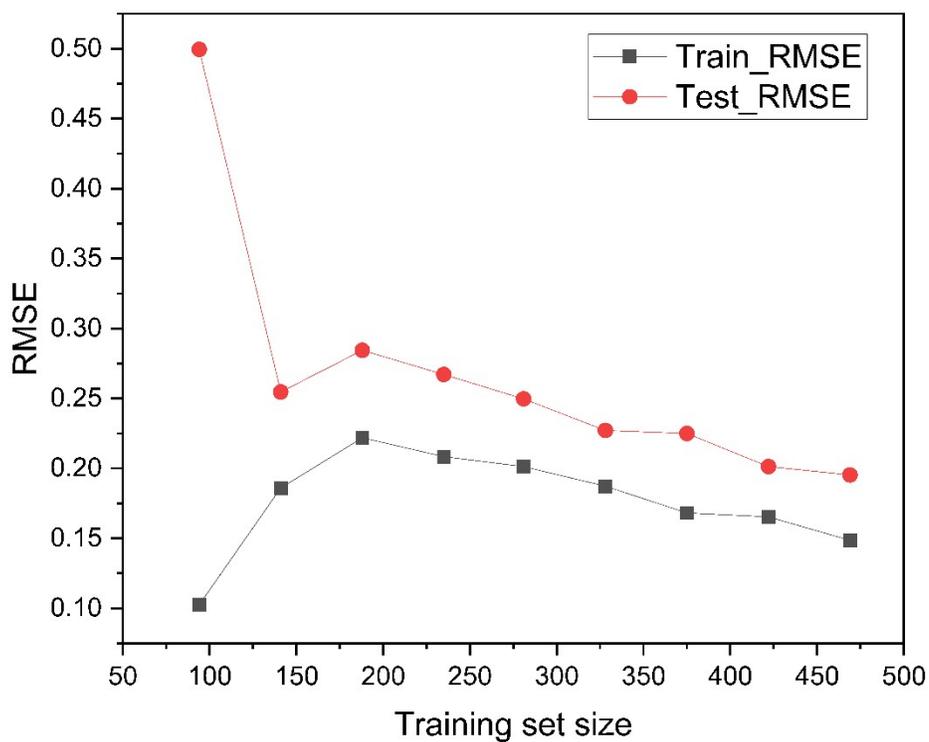


Fig. S6. Learning curve for k_r^f predictions (in s^{-1} in log scale)

8. Computational details for k_r^f and the features employed for predicting k_r^e

The Gaussian16 program package⁵ was used for all geometry optimizations, while the ADF2021 package⁶ was employed to calculate the phosphorescence processes of the Pt(II) complexes. The ground state (S_0) and excited states (T_1) structures for all the Pt-complexes in the datasets were optimized by DFT and TDDFT, respectively, with B3LYP functional⁷. Relativistic effects for the Pt atom were accounted for using the Stuttgart basis set⁸ and pseudopotential, while 6-31G* atomic basis set was applied for other atoms.^{9,10} Solvation effects were included through polarizable continuum model (PCM)¹¹ which corresponds to various testing environments.

Subsequently, the phosphorescence emission properties were computed using ADF2021 package⁶ based on the optimized T_1 structures. Spin-orbit coupling (SOC) was treated as a perturbation based on scalar relativistic orbitals, employing a triple-zeta polarized Slater-type basis set and PBE0 functional utilized in the TDDFT calculations.^{12,13} The COSMO continuum solvation model was employed to consider matrix effects in relevant experimental testing media.¹⁴ The variation in the choice of functional, basis set, and solvation model between the two procedures arises from the lack of identical options available in both Gaussian 16 and ADF 2021 with selections consistent with those used in previous studies⁴. With the emission energy and oscillator strength obtained from the above calculations, k_r^f can be calculated as follows:

$$k_r^f = \frac{2\pi\nu^2 e^2}{\epsilon_0 m c^3} f \quad (1)$$

where ν is the emission energy from the lowest triplet state (T_1) to the ground state (S_0); e denotes the elementary electric charge; ϵ_0 is the vacuum permittivity; m represents the mass of electrons; c is the speed of light, and f is the oscillator strength of transition from T_1 state to S_0 state.

Table S16. Features employed for predicting k_r^e

Features	Description
ν	Emission energy from the T_1 state to the S_0 state
coor_bond_length (N)	Coordinate bond lengths for complexes, the shortest one to longest corresponds with N from 1 to 4
coor_bond_type (N)	Types of coordination (Pt-C, Pt-N, Pt-O, and Pt-Cl); The order of this series of features is correlated to coor_bond_length(N)
ρ_{Pt}	Average electron density at Pt atom
ρ_{coor} (N)	Average electron density at the four coordination atoms
$H_{T_1_S_0}$	Spin-orbit coupling constant between T_1 state and S_0 state
$H_{T_1_S_1}$	Spin-orbit coupling constant between T_1 state and S_1 state
R_EH_excited state ^{a/b}	Charge-transfer descriptor interpreted in terms of the

electron-hole distance in a given excitation. “a” means calculation based on literatures¹⁵⁻¹⁸, “b” means calculation based on reference¹⁹. Small value indicates short-range excitations.

LAMBDA_excited state	Charge-transfer descriptor measures the spatial overlap in a given excitation. ²⁰ Small value signifies a long-range excitation
CT_excited state	Charge-transfer character ¹⁵⁻¹⁸ , 1 for completely charge-separated states; 0 for locally excited excitonic states
HOMO	Highest occupied molecular orbital energy
LUMO	Lowest unoccupied molecular orbital energy
μ	Molecular dipole moment
f	Oscillator strength of radiative transition from T ₁ state to S ₀ state
ML-predicted k_r	Radiative decay rate constant predicted by rMolCLR
refractive index	Refractive index to reflect the experiment testing condition

9. Feature importance analysis

All the features employed for predicting k_r^c are illustrated in section 8. As shown in Figure S7, through shapley additive explanations²¹ (SHAP) importance analysis, it is evident that emission energy has the most significant positive impact on the model. This implies that higher emission energy (indicated by red for high values and blue for low values in Figure S7) leads to a greater influence on the model output (k_r^c), which is consistent with the equation for k_r^f calculation. Likewise, the length of the shortest coordination bond (cood_bond_length (1)), oscillator strength (f), average electron density of the Pt atom (density_Pt), and ML-predicted k_r value (ML_pred_ k_r) all contribute positively to uplifting k_r^c . On the contrary, charge-transfer descriptors in excitations from the ground state (S_0) to the first singlet state (S_1) (LAMBDA_ S_1) and the third singlet state (S_3) (LAMBDA_ S_3) exhibit a clear negative impact on k_r^c . The descriptors in the LAMBDA series measure the spatial overlap in the specified excitations.²⁰ In detail, a low LAMBDA value represents a small overlap (i.e. long-range excitation), resulting in a relatively large transition dipole moment for S_0 to S_1 and S_3 . This contributes positively to the transition dipole moment from the lowest triplet state (T_1) to S_0 , thereby elevating the radiative decay process.²² The aforementioned observations may assist in designing and discovering new emitters with high k_r^c .

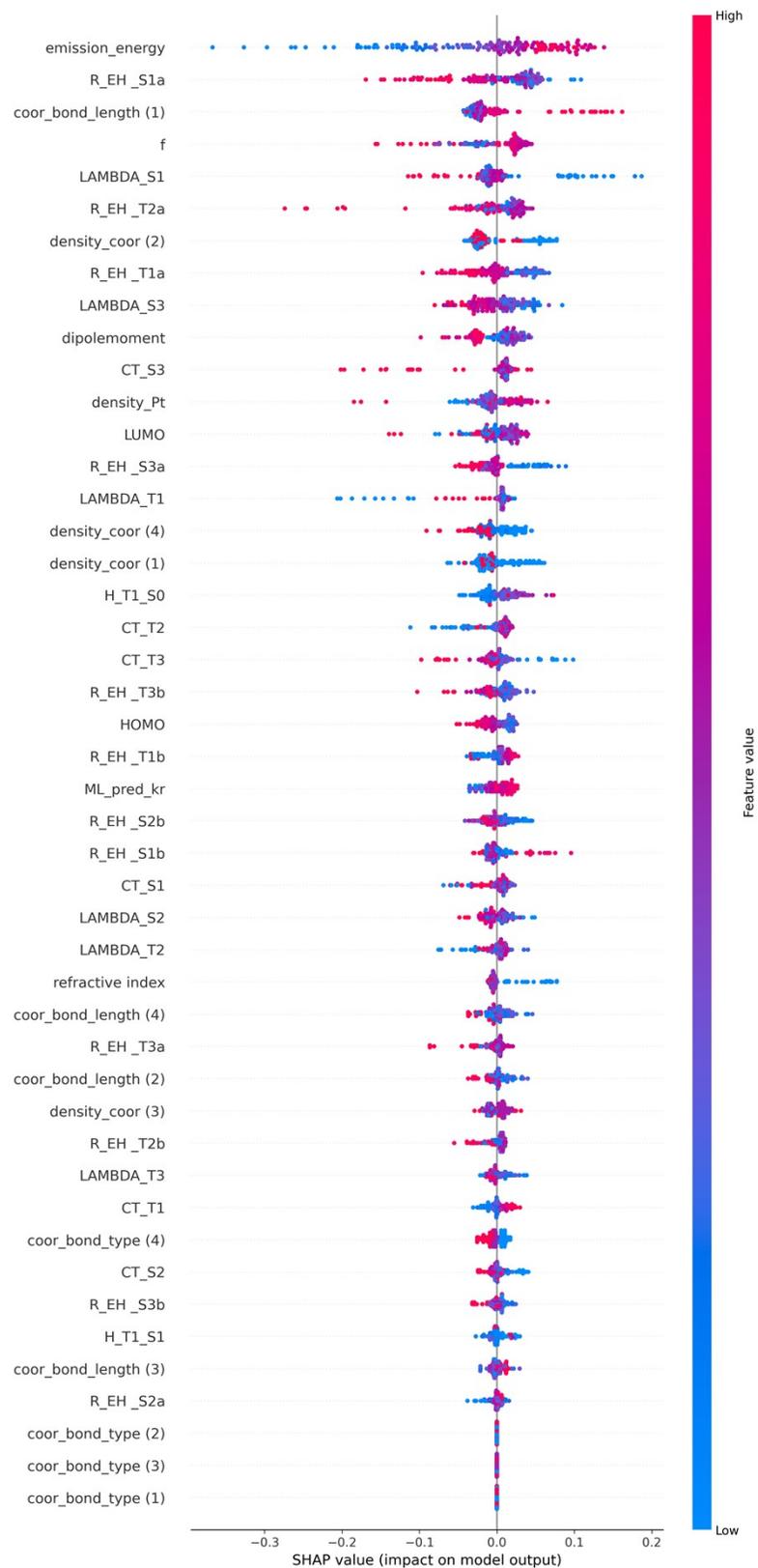
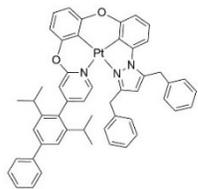
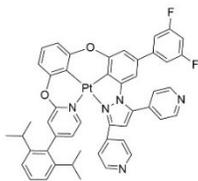


Fig. S7. SHAP analysis for experimental k_r predictions based on optimal XGB model

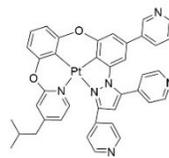
10. Discussions about the 51 top candidates



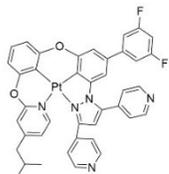
$2.31 \times 10^5 \text{ s}^{-1}$



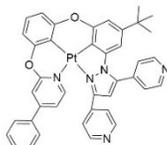
$2.10 \times 10^5 \text{ s}^{-1}$



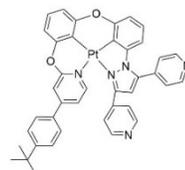
$2.01 \times 10^5 \text{ s}^{-1}$



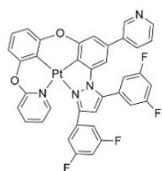
$1.97 \times 10^5 \text{ s}^{-1}$



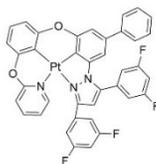
$1.95 \times 10^5 \text{ s}^{-1}$



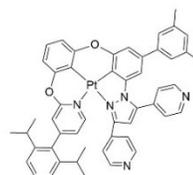
$1.89 \times 10^5 \text{ s}^{-1}$



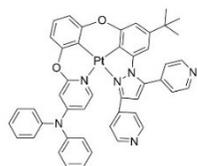
$1.82 \times 10^5 \text{ s}^{-1}$



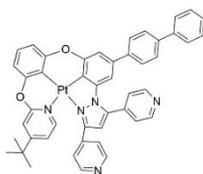
$1.73 \times 10^5 \text{ s}^{-1}$



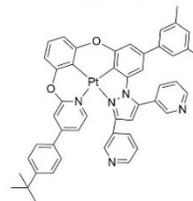
$1.64 \times 10^5 \text{ s}^{-1}$



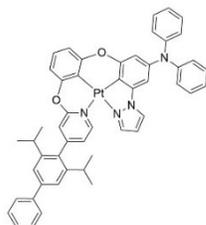
$1.52 \times 10^5 \text{ s}^{-1}$



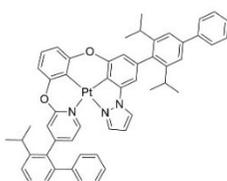
$1.46 \times 10^5 \text{ s}^{-1}$



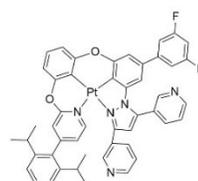
$1.33 \times 10^5 \text{ s}^{-1}$



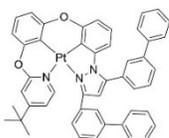
$1.30 \times 10^5 \text{ s}^{-1}$



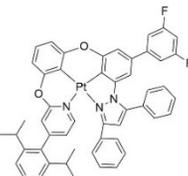
$1.29 \times 10^5 \text{ s}^{-1}$



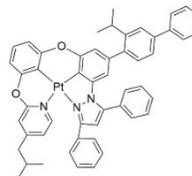
$1.10 \times 10^5 \text{ s}^{-1}$



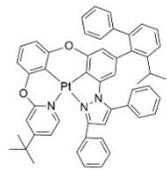
$1.02 \times 10^5 \text{ s}^{-1}$



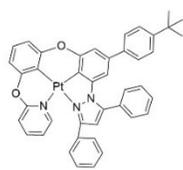
$9.71 \times 10^4 \text{ s}^{-1}$



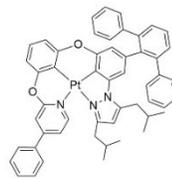
$9.53 \times 10^4 \text{ s}^{-1}$



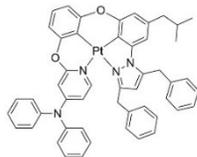
$9.18 \times 10^4 \text{ s}^{-1}$



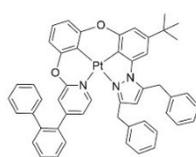
$8.98 \times 10^4 \text{ s}^{-1}$



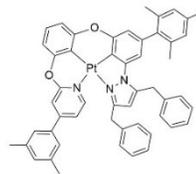
$8.88 \times 10^4 \text{ s}^{-1}$



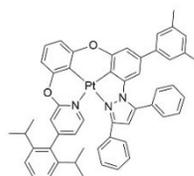
$8.59 \times 10^4 \text{ s}^{-1}$



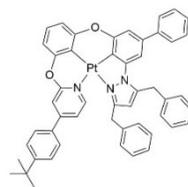
$8.57 \times 10^4 \text{ s}^{-1}$



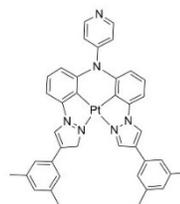
$8.27 \times 10^4 \text{ s}^{-1}$



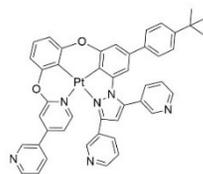
$7.70 \times 10^4 \text{ s}^{-1}$



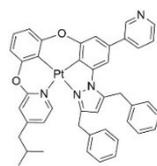
$7.69 \times 10^4 \text{ s}^{-1}$



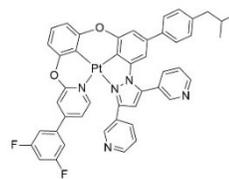
$7.18 \times 10^4 \text{ s}^{-1}$



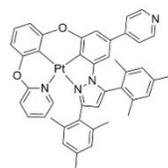
$6.75 \times 10^4 \text{ s}^{-1}$



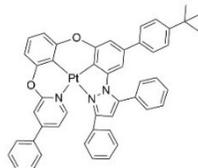
$6.63 \times 10^4 \text{ s}^{-1}$



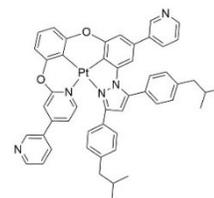
$6.39 \times 10^4 \text{ s}^{-1}$



$6.37 \times 10^4 \text{ s}^{-1}$



$6.36 \times 10^4 \text{ s}^{-1}$



$6.26 \times 10^4 \text{ s}^{-1}$

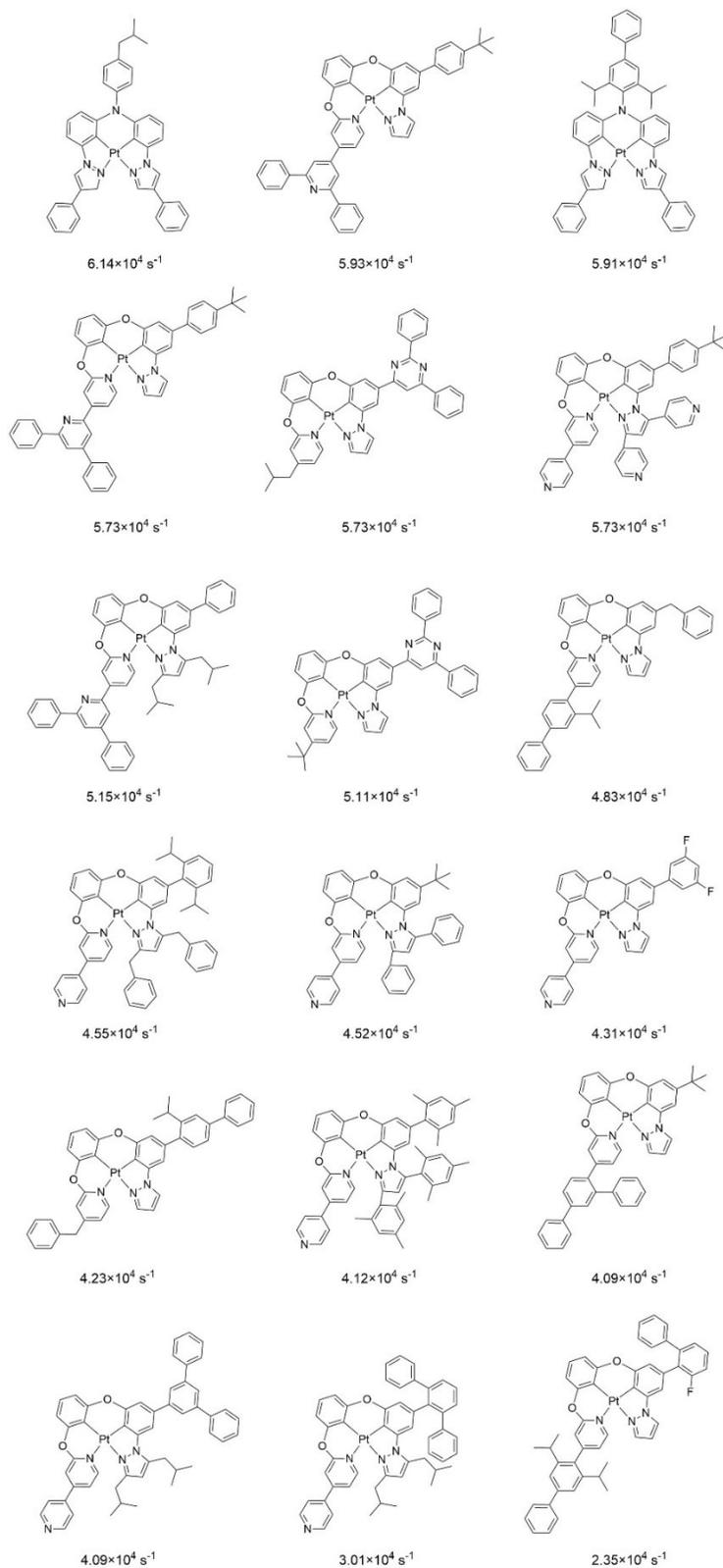


Fig. S8. The structures of 51 top candidates with their predicted k_t^e values

Among all 51 candidates with performance predicted in solvent (CH_2Cl_2), 41 candidates achieve higher k_r^e values than $5 \times 10^4 \text{ s}^{-1}$. More importantly, 16 candidates are with good performance on k_r^e (more than 10^5 s^{-1}). Last but absolutely not least, k_r^e values of 3 excellent candidates are over $2 \times 10^5 \text{ s}^{-1}$, which are rare in the Pt-emitters field, demonstrating the efficacy of the proposed protocol for finding novel high-performance Pt-emitters.

One commonality is that all these complexes are in Pt-NCCN coordination type with majority in core L and minority in core J. While the other commonality can be found among the R-groups occurred in these complexes (Figure S9). It can be observed that most R-groups are with phenyl and/or pyridinyl groups. The phenyl group is generally considered an inductively withdrawing group, because of the higher electronegativity of sp^2 carbon atoms, and a resonance donating group, due to the ability of its π system to donate electron density when conjugation is possible. Large phenyl groups may increase dihedral angles of the complex, increasing the steric hindrance and mitigating the stacking of Pt-complexes, and in turn facilitating the radiative decay process. The pyridinyl has a conjugated system of six π electrons that are delocalized over the ring. In contrast to phenyl groups, the electron density is not evenly distributed over the ring, reflecting the negative inductive effect of the nitrogen atom. For this reason, pyridinyl groups usually increase dipole moment which is a benefit for improving k_r .

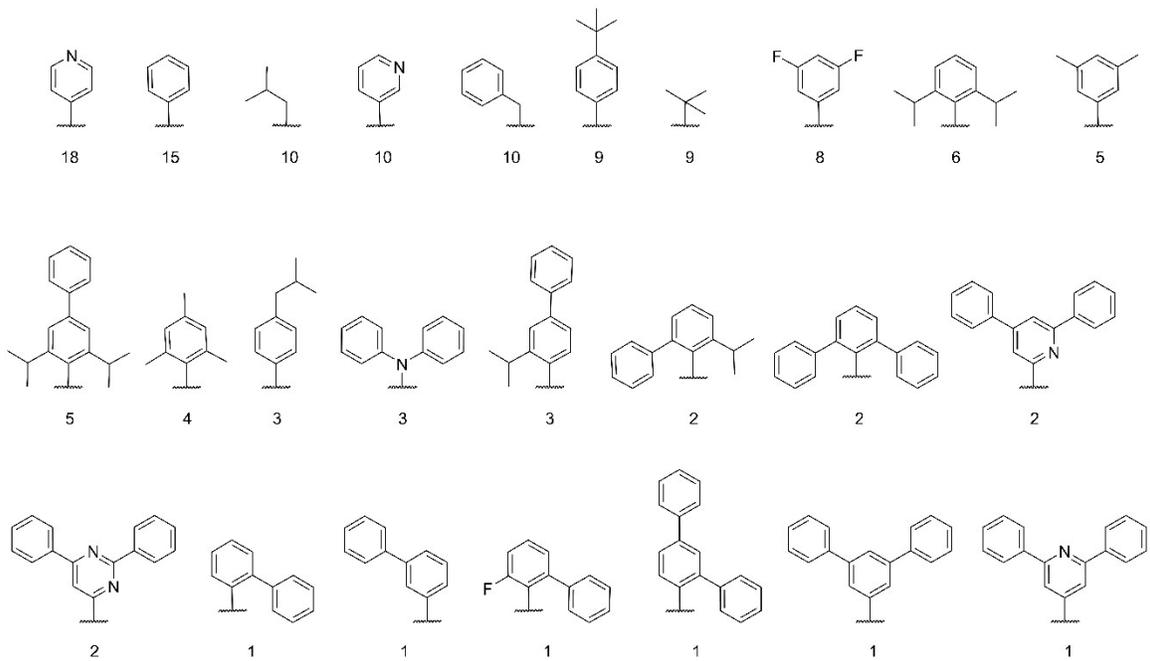


Fig. S9. R-groups collections of 51 top candidates in descending order of count

11. Benchmark against additional functionals

The phosphorescence emission properties were computed using ADF2021 package⁶ based on the optimized T_1 structures. To benchmark the simulations against additional functionals, M06-2X²³ and CAM-B3LYP²⁴ with TZP basis set¹² are employed to calculate the k_r values, allowing a comparison with the results obtained from PBE0²⁵. As a result, the simulation results of 139 samples, compared with experimental data, are presented in Table S17 and Fig. S10.

Table S17. Benchmarks for k_r simulations using various functionals

Functionals	RMSE*	MAE*
CAM-B3LYP	312553.3	155250.9
M06-2X	263660.2	113893.1
PBE0	257202.9	125091.3

*Errors are measured in s^{-1} .

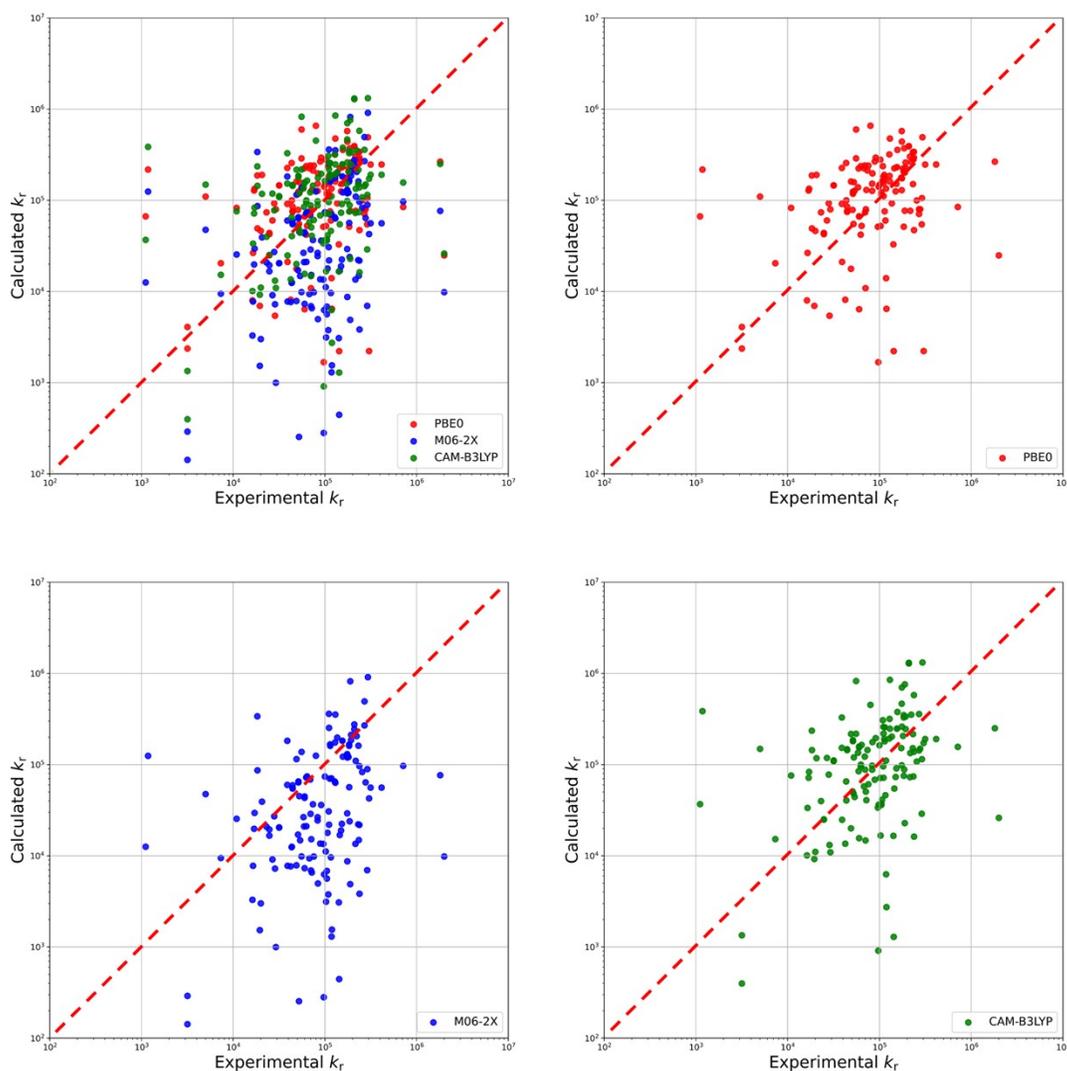


Fig. S10. Comparisons of k_r simulations (in s^{-1}) across various functionals

Overall, the limited accuracy of the three functionals in simulating the k_r values highlights the inherent complexities involved in calculating the emission properties of platinum emitters. In this context, the multi-level Δ -learning method may offer significant support. Notably, PBE0 and M06-2X exhibit comparable performance in simulating k_r values, surpassing that of CAM-B3LYP. In this study, we employed the PBE0 functional in

conjunction with the Δ -learning approach, achieving high accuracy in k_r predictions. Looking ahead, we intend to explore additional functionals and basis sets, develop more representative features, and implement more advanced algorithms to further improve predictions concerning the emission properties, particularly k_r , of metal complexes.

12. References

- 1 J. Bergstra, B. Komer, C. Eliasmith, D. Yamins and D. D. Cox, *Comput. Sci. Discov.*, 2015, **8**, 014008.
- 2 K. Li, G. S. M. Tong, Q. Wan, G. Cheng, W.-Y. Tong, W.-H. Ang, W.-L. Kwong and C.-M. Che, *Chem. Sci.*, 2016, **7**, 1653–1673.
- 3 T. Fleetham, G. Li and J. Li, *Adv. Mater.*, 2017, **29**, 1601861.
- 4 S. Wang, C. Yam, S. Chen, L. Hu, L. Li, F.-F. Hung, J. Fan, C.-M. Che and G. Chen, *J. Comput. Chem.*, 2024, **45**, 321–330.
- 5 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, Revision A.03*, Gaussian, Inc., Wallingford CT, 2016.
- 6 E.J. Baerends, T. Ziegler, J. Autschbach, D. Bashford, A. Bérces, F.M. Bickelhaupt, C. Bo, P.M. Boerrigter, L. Cavallo, D.P. Chong, L. Deng, R.M. Dickson, D.E. Ellis, M. van Faassen, L. Fan, T.H. Fischer, C. Fonseca Guerra, M. Franchini, A. Ghysels, A. Giammona, S.J.A. van Gisbergen, A.W. Götz, J.A. Groeneveld, O.V. Gritsenko, M. Grüning, S. Gusarov, F.E. Harris, P. van den Hoek, C.R. Jacob, H. Jacobsen, L. Jensen, J.W. Kaminski, G. van Kessel, F. Kootstra, A. Kovalenko, M.V. Krykunov, E. van Lenthe, D.A. McCormack, A. Michalak, M. Mitoraj, S.M. Morton, J. Neugebauer, V.P. Nicu, L. Noodleman, V.P. Osinga, S. Patchkovskii, M. Pavanello, P.H.T. Philipsen, D. Post, C.C. Pye, W. Ravenek, J.I. Rodríguez, P. Ros, P.R.T. Schipper, H. van Schoot, G. Schreckenbach, J.S. Seldenthuis, M. Seth, J.G. Snijders, M. Solà, M. Swart, D. Swerhone, G. te Velde, P. Vernooijs, L. Versluis, L. Visscher, O. Visser, F. Wang, T.A. Wesolowski, E.M. van Wezenbeek, G. Wiesenekker, S.K. Wolff, T.K. Woo and A.L. Yakovlev, *ADF2021, SCM*, Vrije Universiteit, Amsterdam, The Netherlands, 2021.
- 7 B. Michlich, A. Savin, H. Stoll and H. Preuss, *Chem. Phys. Lett.*, 1989, **157**, 200–206.
- 8 J. M. L. Martin and A. Sundermann, *J. Chem. Phys.*, 2001, **114**, 3408–3420.
- 9 R. Ditchfield, W. J. Hehre and J. A. Pople, *J. Chem. Phys.*, 1971, **54**, 724–728.
- 10 P. C. Hariharan and J. A. Pople, *Theoret. Chim. Acta*, 1973, **28**, 213–222.
- 11 B. Mennucci, *WIREs Comput. Mol. Sci.*, 2012, **2**, 386–404.
- 12 E. Van Lenthe and E. J. Baerends, *J. Comput. Chem.*, 2003, **24**, 1142–1156.
- 13 F. Wang and T. Ziegler, *J. Chem. Phys.*, 2005, **123**, 154102.
- 14 C. C. Pye and T. Ziegler, *Theor Chem Acc*, 1999, **101**, 396–408.
- 15 F. Plasser, S. A. Bäßler, M. Wormit and A. Dreuw, *J. Chem. Phys.*, 2014, **141**, 024107.
- 16 F. Plasser and H. Lischka, *J Chem Theory Comput*, 2012, **8**, 2777–2789.
- 17 S. A. Mewes, J.-M. Mewes, A. Dreuw and F. Plasser, *Phys. Chem. Chem. Phys.*, 2016, **18**, 2548–2563.

- 18 F. Plasser, M. Wormit and A. Dreuw, *J. Chem. Phys.*, 2014, **141**, 024106.
- 19 C. A. Guido, P. Cortona, B. Mennucci and C. Adamo, *J. Chem. Theory Comput.*, 2013, **9**, 3118–3126.
- 20 M. J. G. Peach, P. Benfield, T. Helgaker and D. J. Tozer, *J. Chem. Phys.*, 2008, **128**, 044118.
- 21 S. M. Lundberg and S.-I. Lee, *Neural Information Processing Systems*, 2017.
- 22 G. S. M. Tong, P. K. Chow, W.-P. To, W.-M. Kwok and C.-M. Che, *Chem. Eur. J.*, 2014, **20**, 6433–6443.
- 23 Y. Zhao and D. G. Truhlar, *Theor Chem Account*, 2008, **120**, 215–241.
- 24 T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51–57.
- 25 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158–6170.