Supplementary Information

Customized Carbon Dots with Predictable Optical Properties Synthesized at Room Temperature Guided by Machine Learning

Qin Hong,‡ Xiao-Yuan Wang,‡ Ya-Ting Gao, Jian Lv, Bin-Bin Chen, Da-Wei Li, and Ruo-Can Qian*

Key Laboratory for Advanced Materials & School of Chemistry and Molecular Engineering, East China University of Science and Technology, Shanghai 200237, P. R. China

*E-mail: ruocanqian@ecust.edu.cn

Supplementary Note 1. Reaction parameters for the preparation of 25 kinds of CDs.

Supplementary Table S1

Reaction parameters for the preparation of 25 kinds of CDs.

Supplementary Note 2. Pairwise Correlation between Original Selected Features of CDs

Supplementary Figure S1

The heatmap of the Pearson's correlation coefficient matrix among the original selected features of CDs (from A to F: dielectric constant, viscosity, density, melting point, boiling point, and refractive index).

Supplementary Note 3. Python Code of Principal Component Analysis (PCA)

Supplementary Figure S2

PCA code for the extraction of PC1 and PC2.

Supplementary Note 4. The Performance of Different Models for the Prediction of FL Intensity

Supplementary Figure S3

(a) Coefficient of determination (R^2) and (b) Pearson's correlation coefficient (r) of different models for the prediction of FL intensity (from A to F: DT, KNN, CNN, RF, SVM, and XGBoost).

Supplementary Note 5. The Correlation between CD Emission Centers and Different Reaction Parameters

Supplementary Figure S4

Box plots showing emission centers of CDs under different reaction parameters (from a to c: M_{PBQ}, V_{EDA}, and reaction time).

Supplementary Note 6. The Performance of Different Models for the Prediction of Emission Centers

Supplementary Figure S5

(a) Coefficient of determination (R^2) and (b) Pearson's correlation coefficient (r) of different models for the prediction of emission centers (from A to F: DT, KNN, CNN, RF, SVM, and XGBoost).

Supplementary Note 7. Decision Trees of the Visualized XGBoost Model

Supplementary Figure S6 and S7

The structure of decision tree 2 and 3 randomly selected by the XGBoost model. The nodes were represented by oval, and the leaves were represented by rectangle.

Supplementary Note 8. The Schematic Illustration of the CD Formation

Supplementary Figure S8

The schematic illustration of the CD formation.

Supplementary Note 9. The Comparison of Predicted Optical Properties and Experimental Optical Properties of CDs Synthesized in EtOH Guided by the XGBoost Model

Supplementary Figure S9

(a) Bar graph, (b) Point plot, and (c) Box plot of predicted FL intensity and experimental FL intensity of CDs synthesized in EtOH (FL unit: x 10^3 a.u.)**.** (e) Bar graph, (f) Point plot, and (g) Box plot of predicted emission wavelength and experimental emission wavelength of CDs synthesized in EtOH**.**

Supplementary Note 10. The Comparison of Predicted Optical Properties and Experimental Optical Properties of CDs Synthesized in Methanol and Acetonitrile Guided by the XGBoost Model

Supplementary Figure S10 and S11

(a) Bar graph, (b) Point plot, and (c) Box plot of predicted FL intensity and experimental FL intensity of CDs synthesized in different solvents (methanol and acetonitrile, FL unit: x 10^3 a.u.). (e) Bar graph, (f) Point plot, and (g) Box plot of predicted emission wavelength and experimental emission wavelength of CDs synthesized in different solvents (methanol and acetonitrile)**.**

Supplementary Note 11. The Synthesis Conditions and Corresponding FL Intensity of CDs Synthesized in Different Solvents

Supplementary Table S2, S3, and S4

The synthesis conditions and corresponding FL intensity of synthesized in different solvents (H_2O, DMF, and DMSO) with high FL intensity predicted by the XGBoost model.

Supplementary Note 12. The Synthesis Conditions of Four CDs Screened by the XGBoost Model

Supplementary Table S5

The synthesis conditions of four CDs with good optical properties in different solvents screened by the XGBoost model.

Supplementary Note 13. C1s, N1s and O1s Spectra of XPS for CDs Synthesized in Different Solvents

Supplementary Figure S12, S13, S14, and S15

(a) C1s, (b) N1s and (c) O1s spectra of XPS for CDs synthesized in different solvents (H_2O, EtOH, DMF, and DMSO).

Supplementary Note 14. The FL Spectra of Different CDs Responding to Fe^{3+} Ions in the Aqueous Solution

Supplementary Figure S16

The FL spectra of different CDs prepared in (a) H_2O, (b) EtOH, (c) DMF, (d) DMSO at the absence and presence of Fe^{3+} in the aqueous solution (1 mM).

Supplementary Note 15. The Response Time of CDs-1 toward Fe^{3+} Ions

Supplementary Figure S17

(a) The FL spectra of CDs-1 before and after adding the Fe^{3+} ions at different times. (b) FL intensity of CDs-1 before and after adding the Fe^{3+} ions at different times (1mM).

Supplementary Note 16. The FL Spectra of CDs-1 Solutions in Presence of Different Metal Ions

Supplementary Figure S18

The FL spectra of CDs-1 mixed with different metal ions (10 mM).

Supplementary Note 17. The FL Spectra of CDs-1 Solutions Mixed with Different Concentrations of Fe^{3+}

Supplementary Figure S19

The FL spectra of CDs-1 solutions with different final Fe^{3+} concentrations of (0-200 μM).

Supplementary Note 18. The UV Spectra of CDs-1 Solutions Mixed with Different Concentrations of Fe^{3+}

Supplementary Figure S20

The UV spectra of CDs-1 after the addition of different concentrations of Fe^{3+} ions. The embedded figure shows the absorbance in the range from 300 nm to 540 nm (1 mM).

Supplementary Note 19. The FL Lifetime of CDs-1 Solutions Mixed with Different Concentrations of Fe^{3+}

Supplementary Figure S21

The FL lifetime of CDs-1 before and after mixing with different concentrations of Fe^{3+} ions.

Supplementary Note 20. MTT assay evaluating the toxicity of CDs-3.

Supplementary Figure S22

Column plot of cell viability after incubating with different concentrations of CDs-3 solutions (10^{-4} - 0.2 mg/mL).

Supplementary Note 1. Reaction parameters for the preparation of 25 kinds of CDs.

Table S1. Reaction parameters for the preparation of 25 kinds of CDs.

C_{PBQ}(mg/mL) / V_{EDA}(μL)	**1**	**2**	**3**	**4**	**5**
10	1	2	3	4	5
20	6	7	8	9	10
30	11	12	13	14	15
40	16	17	18	19	20
50	21	22	23	24	25

Supplementary Note 2. Pairwise Correlation between Original Selected Features of CDs

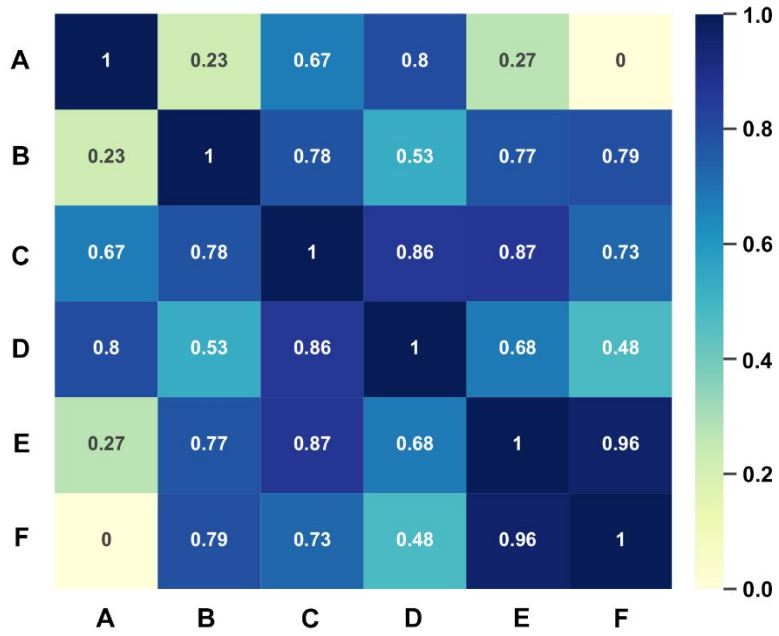

Figure S1. The heatmap of the Pearson's correlation coefficient matrix among the original selected features of CDs (from A to F: dielectric constant, viscosity, density, melting point, boiling point, and refractive index).

Supplementary Note 3. Python Code of Principal Component Analysis (PCA)

PCA can covert the correlative features into new independent variables by orthogonal transformation, which enables the extraction of the input information to the maximum extent. The ratio of principal component 1 (PC1) occupying the previous six features is 0.799 and the ratio of principal component 2 (PC2) is 0.188 Input: dielectric constant, viscosity, density, melting point, boiling point, and refractive index of four solvents (H_2O, EtOH, DMF, and DMSO). Output: PC1 and PC2.

PCA.py > ...

```python
import numpy as np
import pandas as pd
from sklearn.decomposition import PCA
import matplotlib.pyplot as plt

filename = "CDs.xlsx"
filepath = "ML"
a = pd.read_excel(filename,skiprows=1,header = None)
A = a.values
mean = A.mean(axis=0)
C = A - mean
model = PCA(n_components=2)
model.fit(C)
C_new = model.transform(C)
print(C_new)
print("ratio:",model.explained_variance_ratio_)
df=pd.DataFrame(C_new)
df.to_excel(filepath+filename,header=False,index=False)

```

Figure S2. PCA code for the extraction of PC1 and PC2.

Supplementary Note 4. The Performance of Different Models for the Prediction of FL Intensity

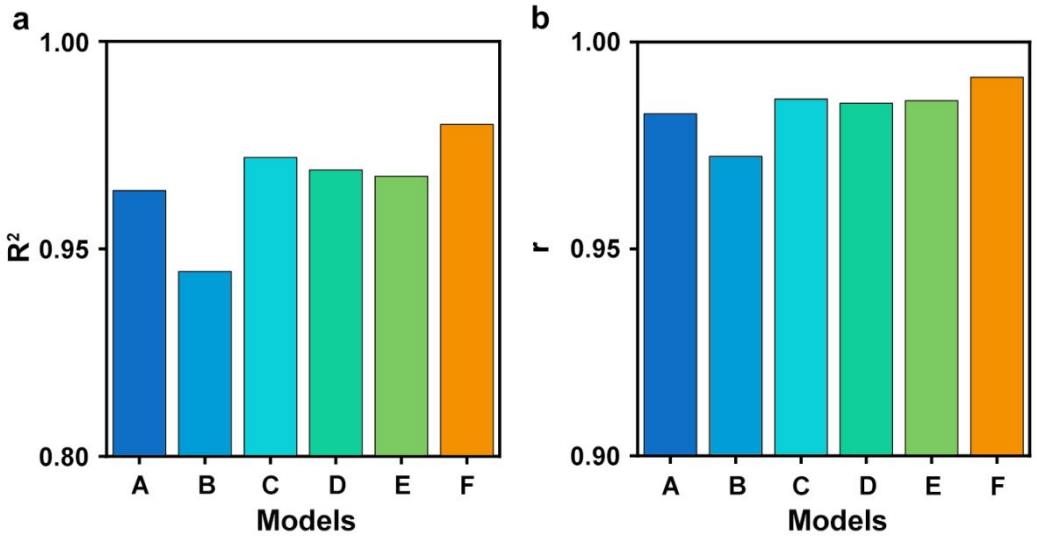

Figure S3. (a) Coefficient of determination (R^2) and (b) Pearson's correlation coefficient (r) of different models for the prediction of FL intensity (from A to F: DT, KNN, CNN, RF, SVM, and XGBoost).

Supplementary Note 5. The Correlation between CD Emission Centers and Different Reaction Parameters

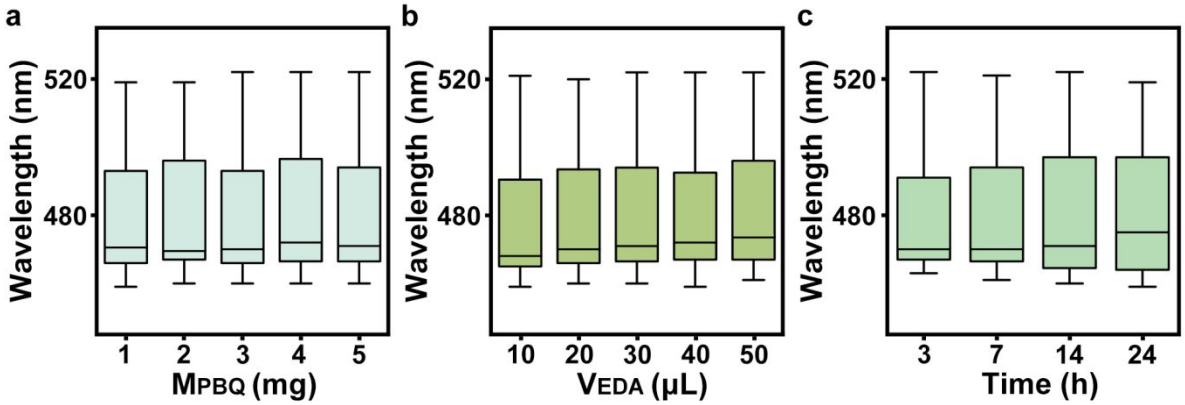

Figure S4. Box plots showing emission centers of CDs under different reaction parameters (from a to c: M_{PBQ}, V_{EDA}, and reaction time).

Supplementary Note 6. The Performance of Different Models for the Prediction of Emission Centers

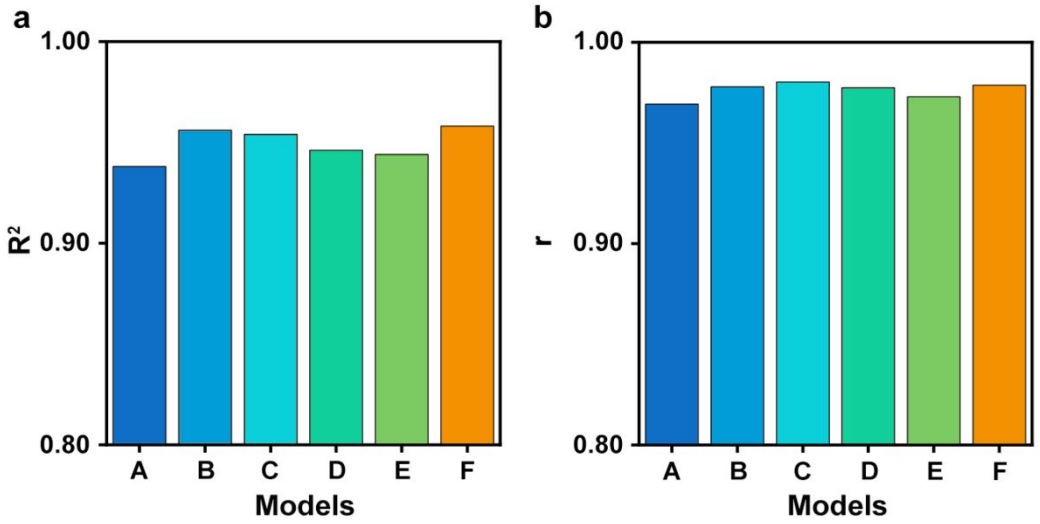

Figure S5. (a) Coefficient of determination (R^2) and (b) Pearson's correlation coefficient (r) of different models for the prediction of emission centers (from A to F: DT, KNN, CNN, RF, SVM, and XGBoost).

Supplementary Note 7. Decision Trees of the Visualized XGBoost Model

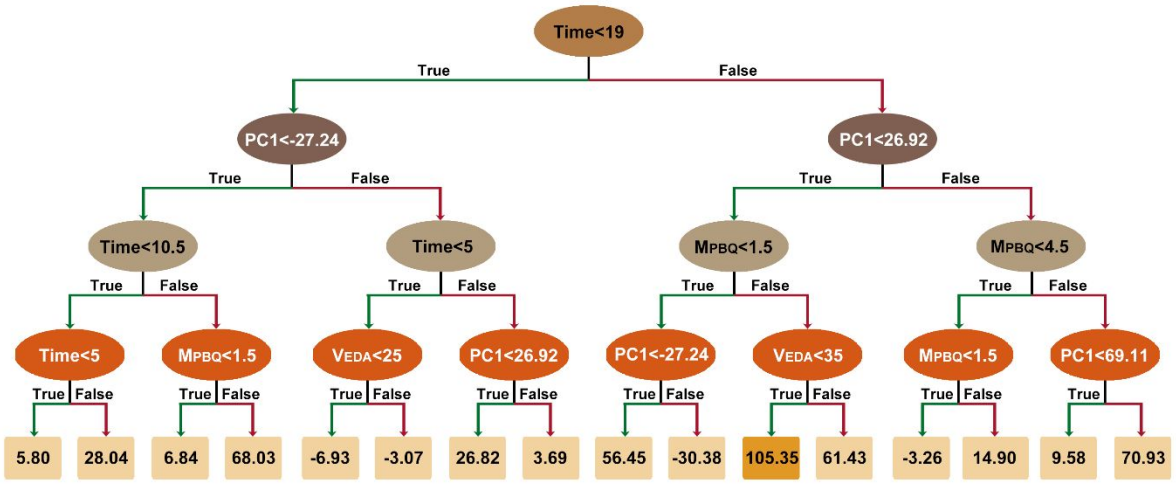

Figure S6. The structure of decision tree 2 randomly selected by the XGBoost model. The nodes were represented by oval, and the leaves were represented by rectangle.

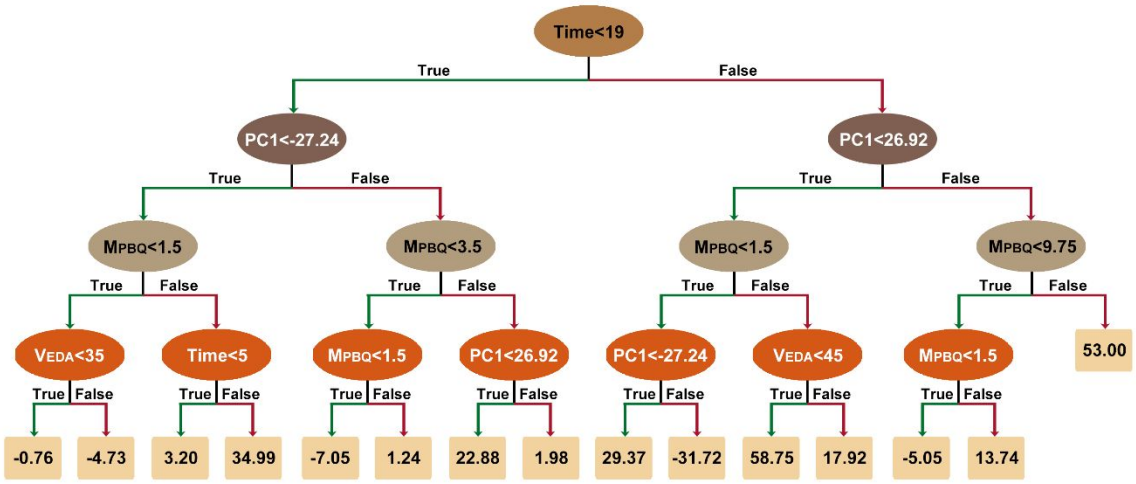

Figure S7. The structure of decision tree 3 randomly selected by the XGBoost model. The nodes were represented by oval, and the leaves were represented by rectangle.

Supplementary Note 8. The Schematic Illustration of the CD Formation

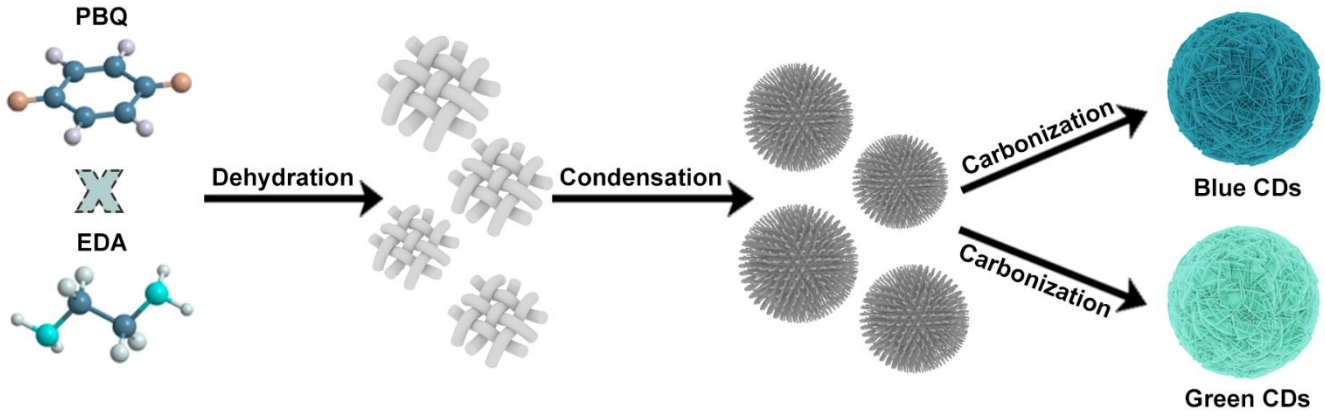

Figure S8. The schematic illustration of the CD formation.

Supplementary Note 9. The Comparison of Predicted Optical Properties and Experimental Optical Properties of CDs Synthesized in EtOH Guided by the XGBoost Model

The root mean square errors (rmse) between the predicted and experimental values of FL intensity and emission center are 139.31, and 1.85, respectively.

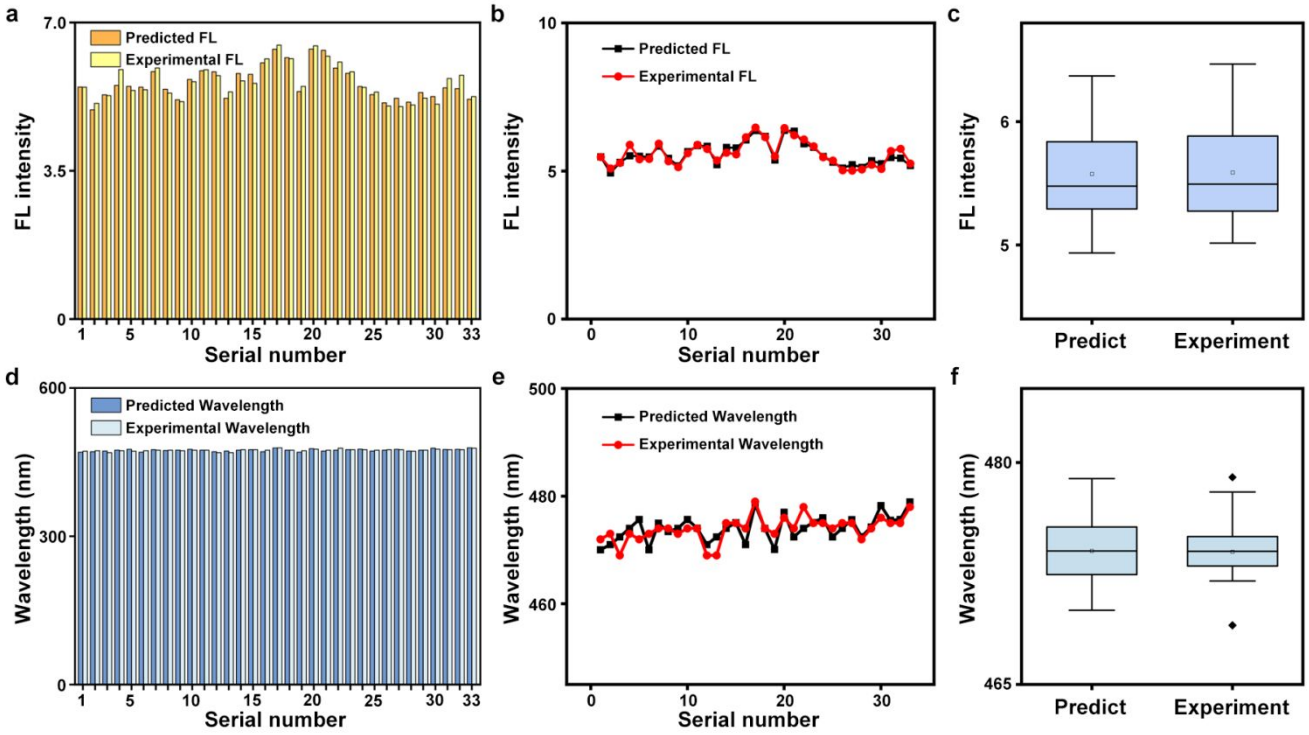

Figure S9. (a) Bar graph, (b) Point plot, and (c) Box plot of predicted FL intensity and experimental FL intensity of CDs synthesized in EtOH (FL unit: x 10^3 a.u.). (e) Bar graph, (f) Point plot, and (g) Box plot of predicted emission wavelength and experimental emission wavelength of CDs synthesized in EtOH**.**

Supplementary Note 10. The Comparison of Predicted Optical Properties and Experimental Optical Properties of CDs Synthesized in Methanol and Acetonitrile Guided by the XGBoost Model

The root mean square errors (rmse) between the predicted and experimental values of FL intensity and emission center are 146.87, and 1.13, respectively.

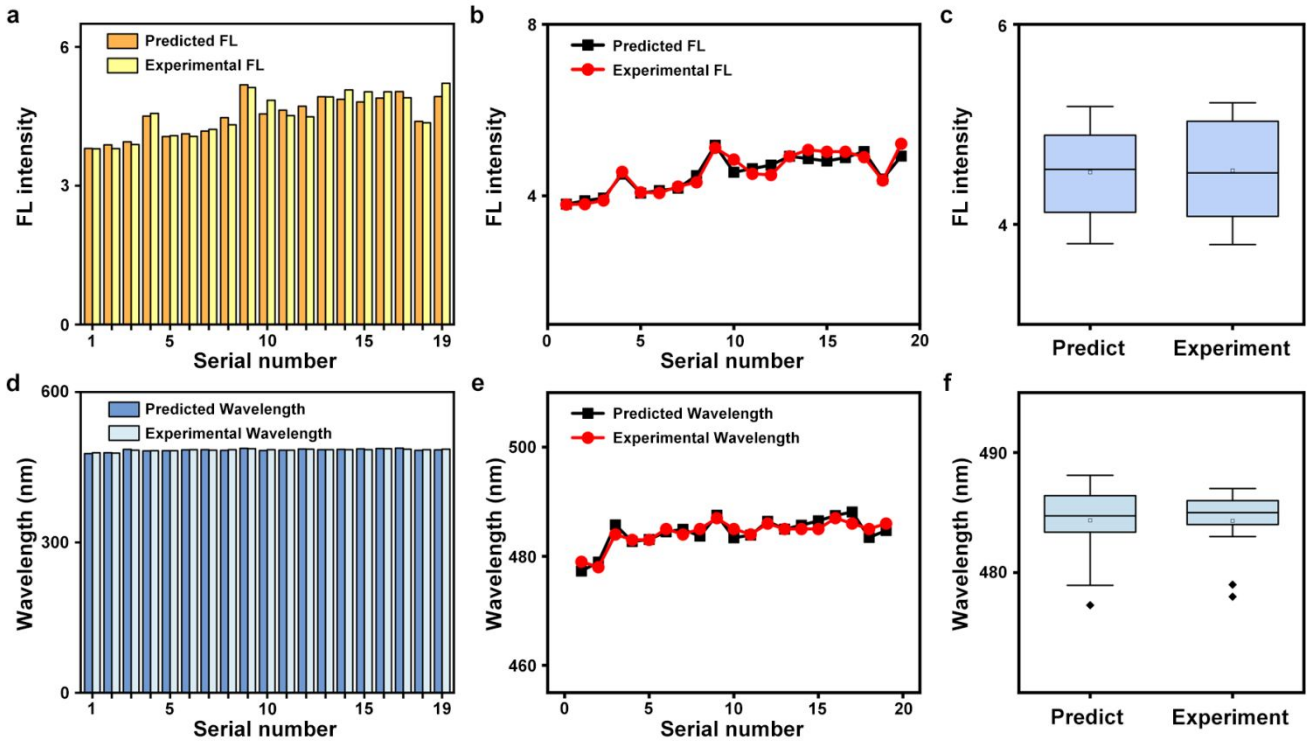

Figure S10. (a) Bar graph, (b) Point plot, and (c) Box plot of predicted FL intensity and experimental FL intensity of CDs synthesized in methanol (FL unit: x 10^3 a.u.). (e) Bar graph, (f) Point plot, and (g) Box plot of predicted emission wavelength and experimental emission wavelength of CDs synthesized in methanol**.**

The root mean square errors (rmse) between the predicted and experimental values of FL intensity and emission center are 224.78, and 1.85, respectively.

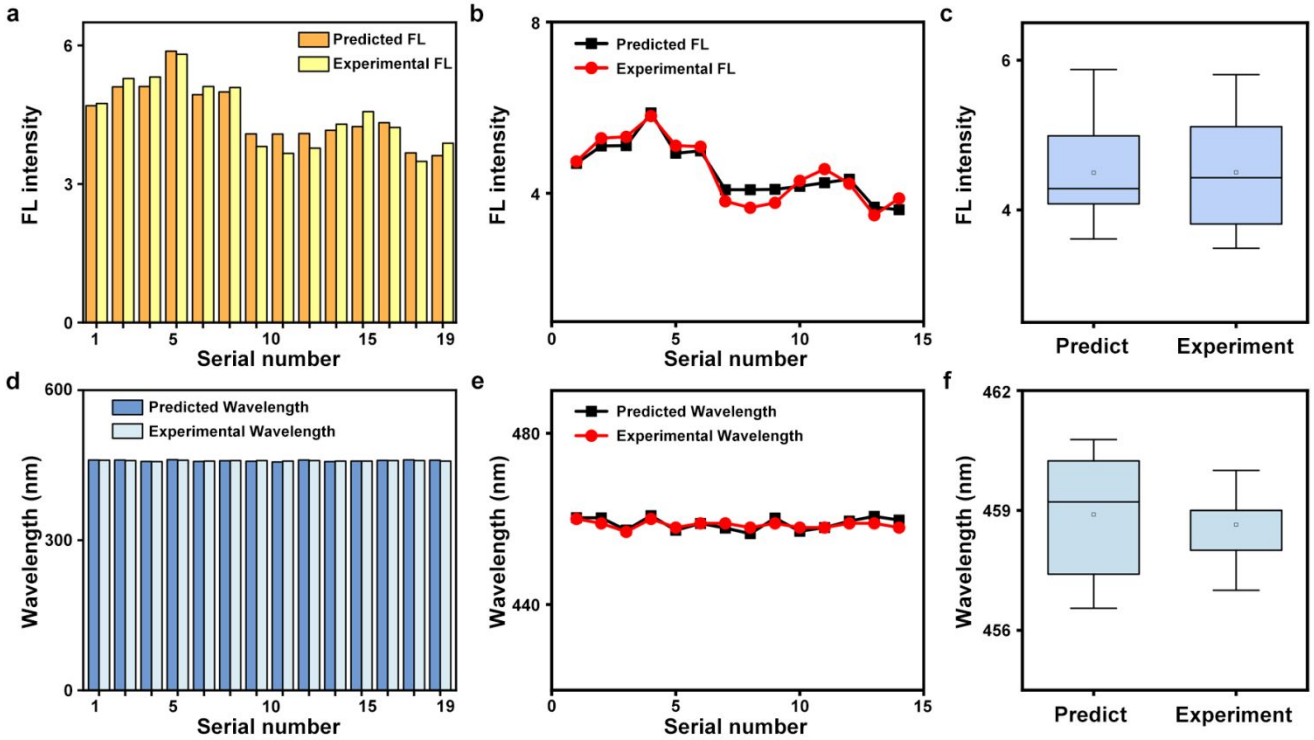

Figure S11. (a) Bar graph, (b) Point plot, and (c) Box plot of predicted FL intensity and experimental FL intensity of CDs synthesized in acetonitrile (FL unit: x 10^3 a.u.). (e) Bar graph, (f) Point plot, and (g) Box plot of predicted emission wavelength and experimental emission wavelength of CDs synthesized in acetonitrile**.**

Supplementary Note 11. The Synthesis Conditions and Corresponding FL Intensity of CDs Synthesized in Different Solvents

Table S2. The synthesis conditions and corresponding FL intensity of 12 kinds of CDs synthesized in H_2O with high FL intensity predicted by the XGBoost model.

	M_{PBQ} (mg)	V_{EDA} (μL)	Time (h)	Solvent	Predicted FL intensity	Tested FL intensity
1	6.5	5	24	H_2O	637.5863	385.8
2	7	5	24	H_2O	651.74365	887.8
3	7.5	5	24	H_2O	653.18756	455.4
4	8	5	24	H_2O	619.47003	693.6
5	8.5	5	24	H_2O	585.63086	461.6
6	9	5	24	H_2O	553.24646	308.6
7	9.5	5	24	H_2O	523.71265	319.8
8	10	5	24	H_2O	491.7528	418.9
9	8.5	10	24	H_2O	370.90143	159.2
10	9	10	24	H_2O	372.8996	135.2
11	9.5	10	24	H_2O	378.29974	165.7
12	10	10	24	H_2O	385.5193	339

Table S3. The synthesis conditions and corresponding FL intensity of 15 kinds of CDs synthesized in DMF with high FL intensity predicted by the XGBoost model.

	M_{PBQ} (mg)	V_{EDA} (μL)	Time (h)	Solvent	Predicted FL intensity	Tested FL intensity
1	6	5	24	DMF	4239.504	3622
2	6.5	5	24	DMF	4388.9595	3832
3	7	5	24	DMF	4538.415	4305
4	7.5	5	24	DMF	4645.428	4643
5	8	5	24	DMF	4651.628	4587
6	8.5	5	24	DMF	4657.6313	4599
7	9	5	24	DMF	4663.635	4619
8	9.5	5	24	DMF	4669.6377	4913
9	10	5	24	DMF	4668.725	4948
10	8	10	24	DMF	3967.3328	3868
11	8.5	10	24	DMF	4088.4668	3516
12	9	10	24	DMF	4209.5996	3800
13	9.5	10	24	DMF	4330.733	4010
14	10	10	24	DMF	4451.866	4071
15	10	15	24	DMF	3304.1377	3253

Table S4. The synthesis conditions and corresponding FL intensity of 15 kinds of CDs synthesized in DMSO with high FL intensity predicted by the XGBoost model.

	M_{PBQ} (mg)	V_{EDA} (μL)	Time (h)	Solvent	Predicted FL intensity	Tested FL intensity
1	6	5	24	DMSO	2700.2434	2566
2	6.5	5	24	DMSO	2849.6987	3149
3	7	5	24	DMSO	2999.1543	3131
4	7.5	5	24	DMSO	3106.1672	2668
5	8	5	24	DMSO	3112.3672	3360
6	8.5	5	24	DMSO	3118.37	2829
7	9	5	24	DMSO	3124.3735	2715
8	9.5	5	24	DMSO	3130.377	2224
9	10	5	24	DMSO	3129.4639	3126
10	8	10	24	DMSO	2463.0078	3004
11	8.5	10	24	DMSO	2584.1416	2618
12	9	10	24	DMSO	2705.2744	3221
13	9.5	10	24	DMSO	2826.408	2958
14	10	10	24	DMSO	2947.541	3502
15	10	15	24	DMSO	1252.5854	1460

Supplementary Note 12. The Synthesis Conditions of Four CDs Screened by the XGBoost Model

Table S5. The synthesis conditions of four CDs with good optical properties in different solvents screened by the XGBoost model.

	M_{PBQ} (mg)	V_{EDA} (μL)	Time (h)	Solvent
CDs-1	7	5	24	H_2O
CDs-2	6.5	25	24	EtOH
CDs-3	9.5	5	24	DMF
CDs-4	10	5	24	DMSO

Supplementary Note 13. C1s, N1s and O1s Spectra of XPS for CDs Synthesized in Different Solvents

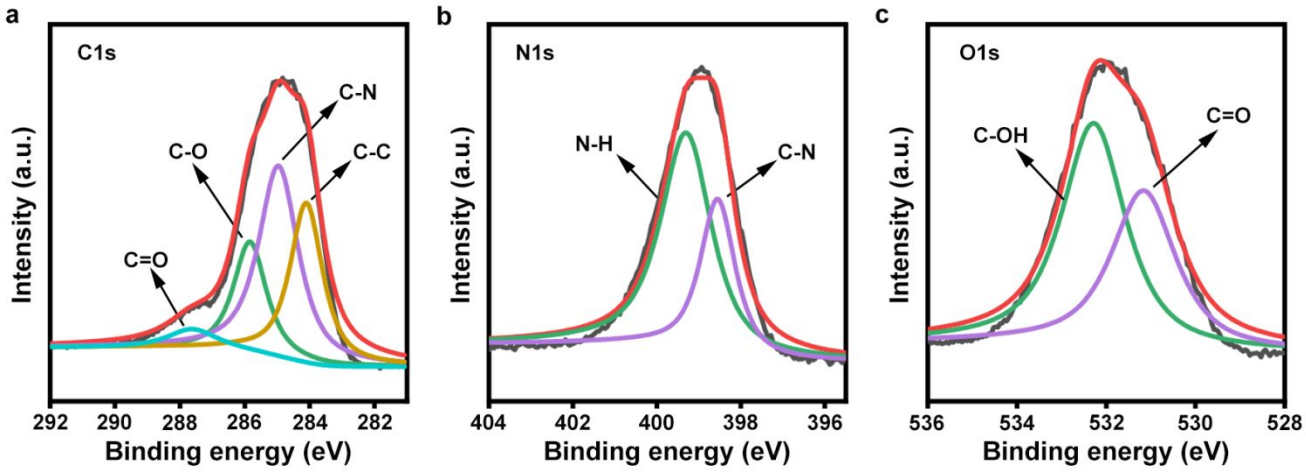

Figure S12. (a) C1s, (b) N1s and (c) O1s spectra of XPS for CDs synthesized in H_2O.

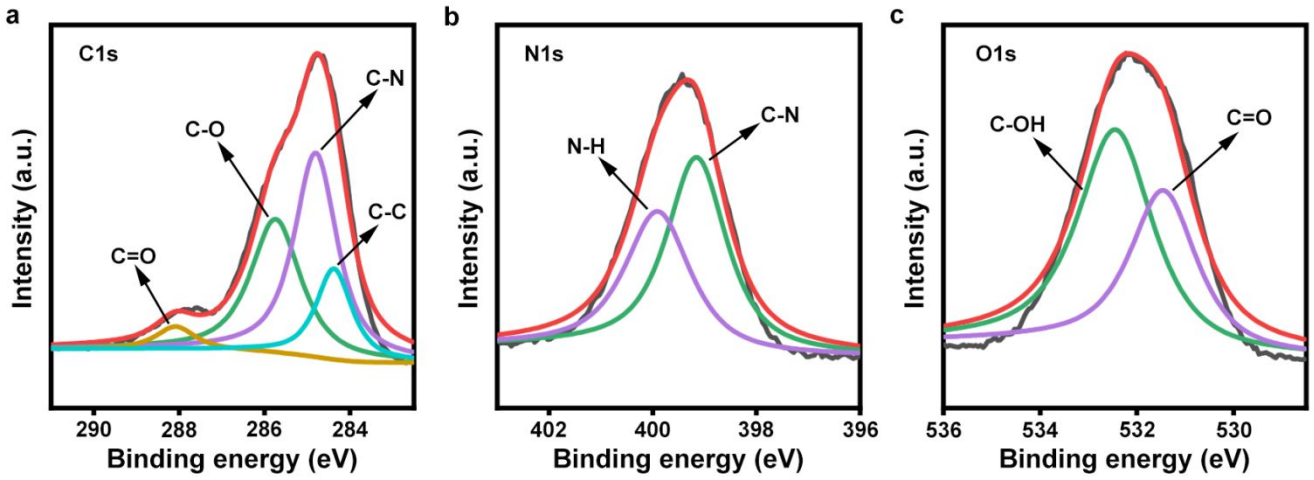

Figure S13. (a) C1s, (b) N1s and (c) O1s spectra of XPS for CDs synthesized in EtOH.

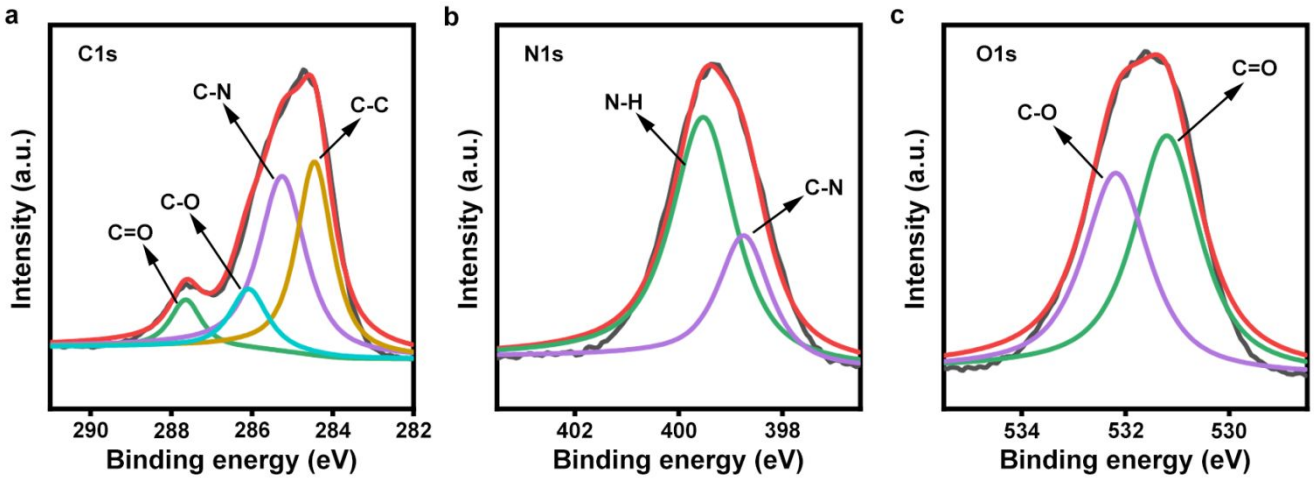

Figure S14. (a) C1s, (b) N1s and (c) O1s spectra of XPS for CDs synthesized in DMF.

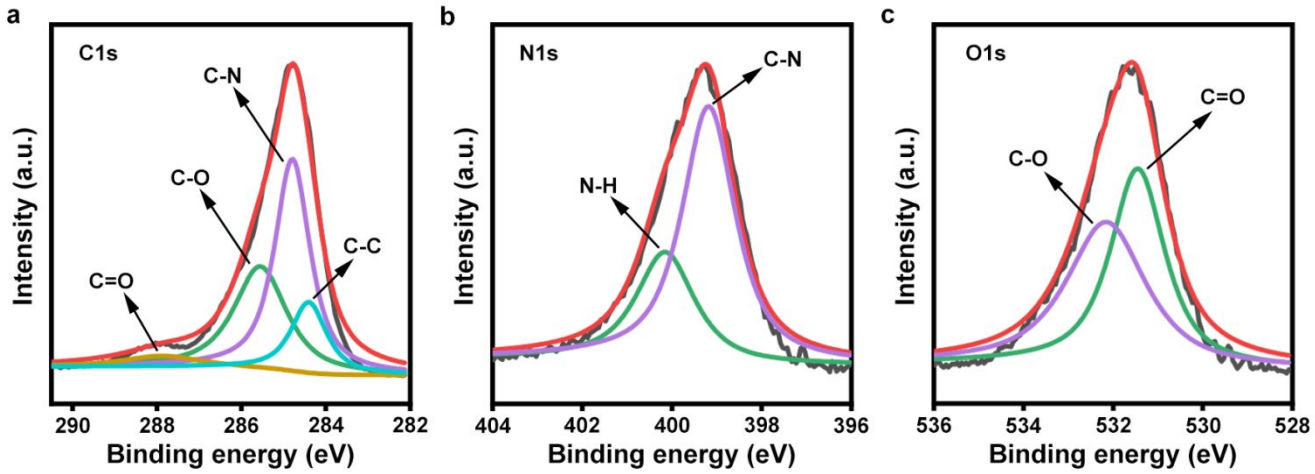

Figure S15. (a) C1s, (b) N1s and (c) O1s spectra of XPS for CDs synthesized in DMSO.

Supplementary Note 14. The FL Spectra of Different CDs Responding to Fe^{3+} Ions in the Aqueous Solution

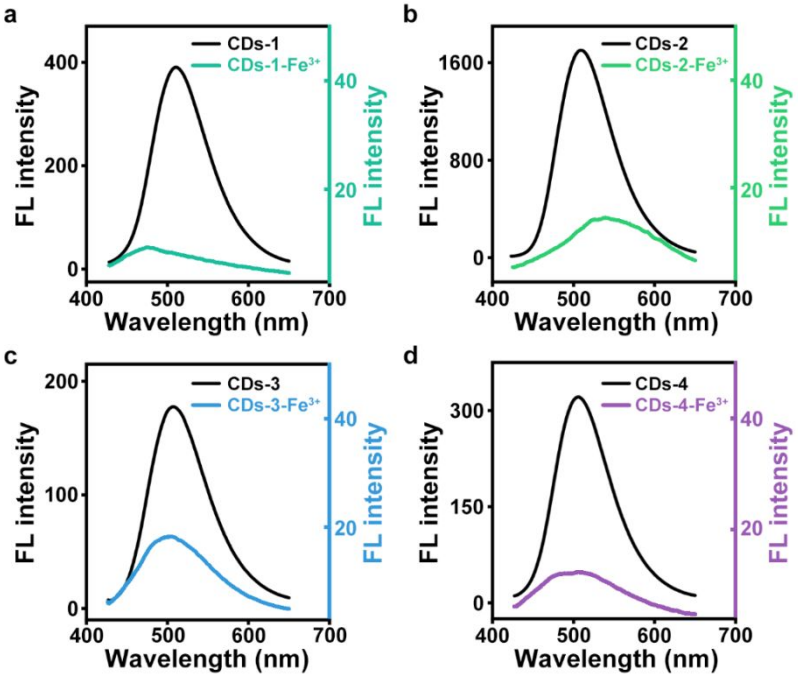

Figure S16. The FL spectra of different CDs prepared in (a) H_2O, (b) EtOH, (c) DMF, (d) DMSO at the absence and presence of Fe^{3+} in the aqueous solution (1 mM).

Supplementary Note 15. The Respond Time of CDs-1 toward Fe^{3+} Ions

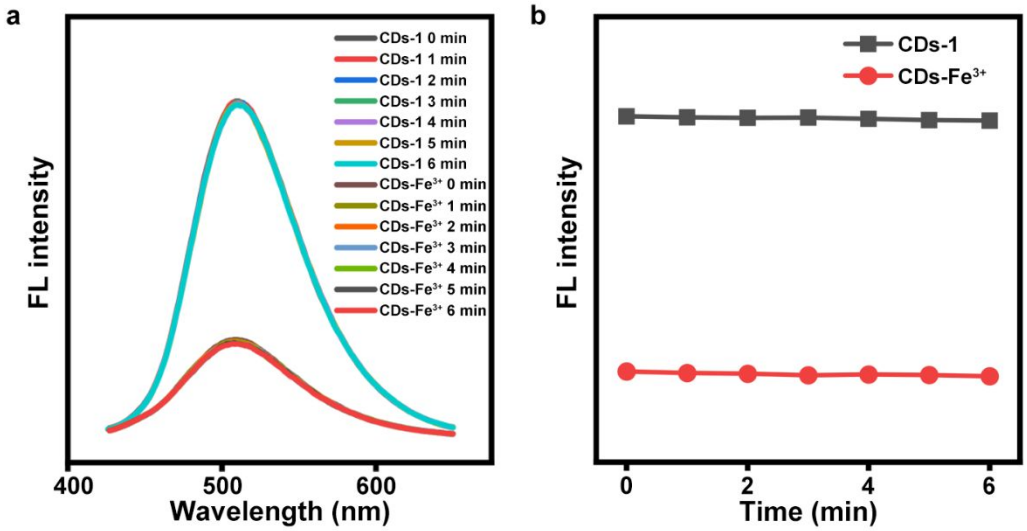

Figure S17. (a) The FL spectra of CDs-1 before and after adding the Fe^{3+} ions at different times. (b) FL intensity of CDs-1 before and after adding the Fe^{3+} ions at different times (1mM).

Supplementary Note 16. The FL Spectra of CDs-1 Solutions in Presence of Different Metal Ions

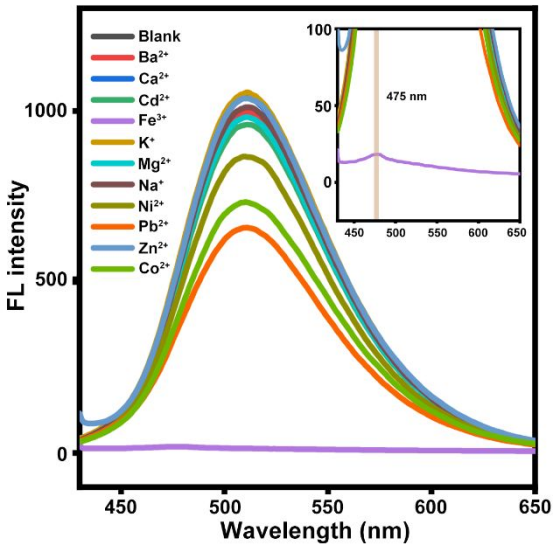

Figure S18. The FL spectra of CDs-1 mixed with different metal ions (10 mM).

Supplementary Note 17. The FL Spectra of CDs-1 Solutions Mixed with Different Concentrations of Fe^{3+}

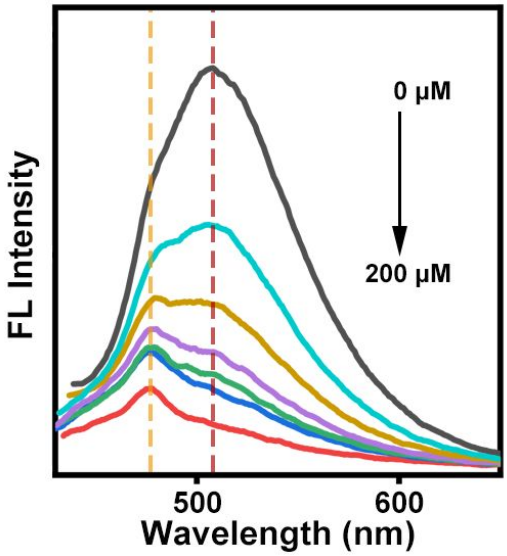

Figure S19. The FL spectra of CDs-1 solutions with different final Fe^{3+} concentrations of (0-200 μM).

Supplementary Note 18. The UV Spectra of CDs-1 Solutions Mixed with Different Concentrations of Fe^{3+}

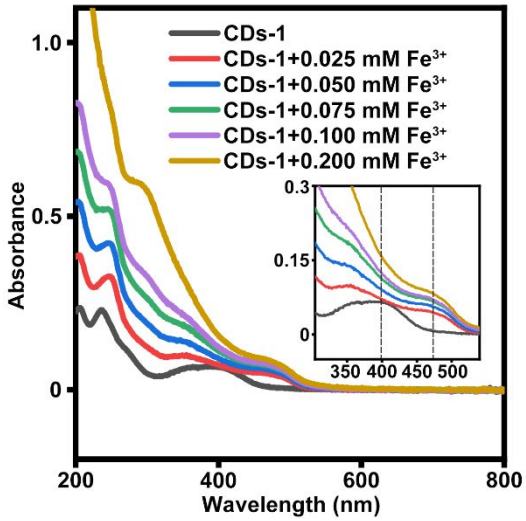

Figure S20. The UV spectra of CDs-1 after the addition of different concentrations of Fe^{3+}. The embedded figure shows the absorbance in the range from 300 nm to 540 nm (1mM).

Supplementary Note 19. The FL Lifetime of CDs-1 Solutions Mixed with Different Concentrations of Fe^{3+}

To investigate the mechanism of Fe^{3+} ions quenching the fluorescence of CDs-1, the fluorescence lifetime of CDs-1 solution before and after the addition of Fe^{3+} ions was tested. As shown in Figure S21, the fluorescence life of CDs-1 mixed with different concentrations of Fe^{3+} ions hardly change compared to CDs without Fe^{3+} ions. According to the Stern-Volmer equation:

$$F_0/F = 1 + K_{SV}[Q]$$

Where F_0 and F represent the FL intensity before and after adding Fe^{3+} ions, respectively, K_{SV} is the quenching constant, and $[Q]$ is the Fe^{3+} concentration. The linear relationship between the FL ratio (F_0/F) in the absence and presence of Fe^{3+} ions and Fe^{3+} concentration was constructed ($F_0/F = 1.69 + 0.034\ [Fe^{3+}]/\mu M$, $R^2 = 0.999$; Figure 5a iii). Therefore, the K_{SV} is calculated to be 0.034.

K_{SV} is applied to distinguish the quenching type containing the static quenching and the dynamic quenching effects (SQE and DQE). In DQE, K_{SV} is estimated by the following equation:

$$K_{SV} = k_q \times \tau_0$$

Where τ_0 is the FL lifetime of CDs-1 without Fe^{3+} ions, and k_q is the quenching rate constant. The results show k_q equals to 9.27×10^{12} $M^{-1} \cdot s^{-1}$ (K_{SV} = 0.034 μM^{-1}, τ_0 = 3.72 ns), which exceed the diffusion-controlled process value 1.0×10^{10} $M^{-1} \cdot s^{-1}$ for a molecule in solution and further indicate the FL quenching induced by Fe^{3+} ions is due to SQE rather than DQE. [1]

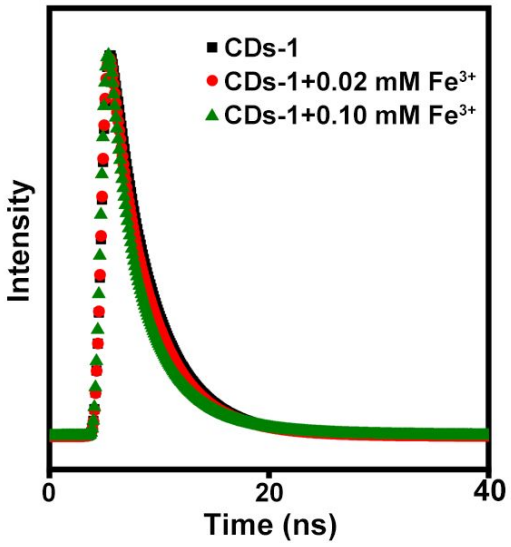

Figure S21. The FL lifetime of CDs-1 before and after mixing with different concentrations of Fe^{3+} ions (blank-3.72 ns, 0.02 mM Fe^{3+}-3.75 ns, and 0.10 mM Fe^{3+}-3.67 ns).

Supplementary Note 20. MTT assay evaluating the toxicity of CDs-3.

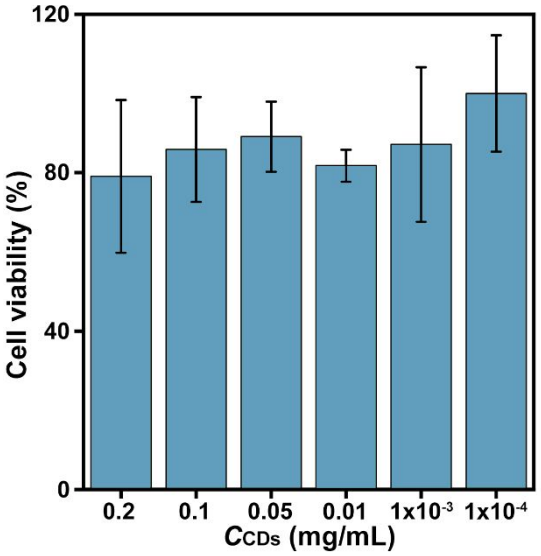

Figure S22. Column plot of cell viability after incubating with different concentrations of CDs-3 solutions (10^{-4} - 0.2 mg/mL).

References

[1] Gao, Y. T.; Chen, B. B.; Jiang, L.; Lv, J.; Chang, S.; Wang, Y.; Qian, R. C.; Li, D. W.; Hafez, M. E. Dual-emitting carbonized polymer dots synthesized at room temperature for ratiometric fluorescence sensing of Vitamin B12. *ACS Appl. Mater. Interfaces* **2021**. DOI: 10.1021/acsami.1c12993.