THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCL Letter

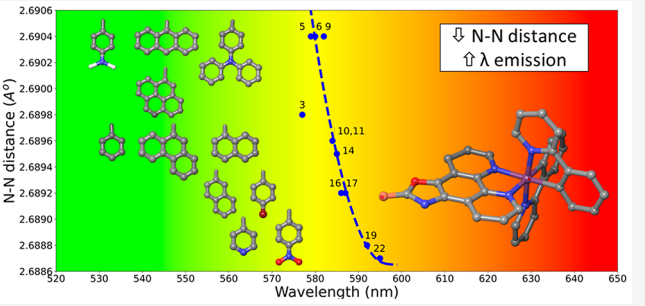
Combined Machine Learning, Computational, and Experimental Analysis of the Iridium(III) Complexes with Red to Near-Infrared Emission

Anas Karuth, Gerardo M. Casanola-Martin, Levi Lystrom, Wenfang Sun, Dmitri Kilin, Svetlana Kilina, and Bakhtiyor Rasulev*

Cite This: *J. Phys. Chem. Lett.* 2024, 15, 471−480 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Various coordination complexes have been the subject of experimental and theoretical studies in recent decades because of their fascinating photophysical properties. In this work, a combined experimental and computational approach was applied to investigate the optical properties of monocationic Ir(III) complexes. An interpretative machine learning-based quantitative structure−property relationship (ML/QSPR) model was successfully developed that could reliably predict the emission wavelength of the Ir(III) complexes and provide a foundation for the theoretical evaluation of the optical properties of Ir(III) complexes. A hypothesis was proposed to explain the differences in the emission wavelengths between structurally different individual Ir(III) complexes. The efficacy of the developed model was demonstrated by high R^2 values of 0.84 and 0.87 for the training and test sets, respectively. It is worth noting that a relationship between the N−N distance in the diimine ligands of the Ir(III) complexes and emission wavelengths is detected. This effect is most probably associated with a degree of distortion in the octahedral geometry of the complexes, resulting in a perturbed ligand field. This combined experimental and computational approach shows great potential for the rational design of new Ir(III) complexes with the desired optical properties. Moreover, the developed methodology could be extended to other transition-metal complexes.



In recent decades, various transition-metal coordination complexes have been the subject of extensive studies because of their fascinating photophysical properties. Among these transition-metal complexes, Ir(III) complexes are particularly interesting due to the strong spin−orbit coupling induced by the Ir(III) ion, which gives rise to an efficient intersystem crossing and high phosphorescence quantum yield at room temperature.[1−3] Because of these characteristics, Ir(III) complexes have been widely studied for use in organic light-emitting devices (OLEDs),[4,5] light-emitting electrochemical cells (LEECs),[6,7] luminescent biological labeling reagents,[8] chemosensors,[9] and upconversion applications.[10] In addition, Ir(III) complexes were found to exhibit efficient reverse saturable absorption (RSA),[11−14] which is an important property for optical switching, optical limiting, optical rectification, and laser pulse shaping and compression applications.[15,16]

Various structural modifications of the Ir(III) complexes were used in an attempt to tune the ground- and excited-state properties to shift the absorption and phosphorescence of these complexes to the near-infrared (NIR) spectral regions. For instance, the optical properties of Ir(III) complexes can be altered by using different types of diimine or cyclometalating ligands or varying the substituents on the ligands.[12,17,18] Extensive computational and electrochemical studies show that the lowest unoccupied molecular orbital (LUMO) of these Ir(III) complexes is typically located on the diimine (N^N) ligand, while the highest occupied molecular orbital (HOMO) is delocalized on the cyclometalating (C^N) ligands and the d orbital of the Ir(III) center.[11] Thus, an electron-donating or -withdrawing substituent on the N^N ligand can lead to a hypsochromic or bathochromic shift of the lowest-energy charge transfer optical transition by destabilization or stabilization of the excited electron state, while the hole state is negligibly affected.[19]

In addition to the electron-withdrawing or -donating groups, introducing π-conjugated substituents onto the N^N ligand could alter the nature of the lowest triplet excited state and

Received: September 9, 2023
Revised: December 19, 2023
Accepted: December 27, 2023

https://doi.org/10.1021/acs.jpclett.3c02533
J. Phys. Chem. Lett. 2024, 15, 471−480

thus impact the emission characteristics. Sun and co-workers demonstrated that incorporating π-conjugated fluorenyl and benzothiazolyl substituents onto the N^N ligands switched the lowest triplet excited state to mainly the N^N ligand-localized $^3\pi,\pi^*$ state, which gave rise to longer triplet lifetimes and strong phosphorescence in comparison to the values of those complexes without these π-conjugated substituents on the N^N ligands.[20]

Extending π-conjugation can also be achieved via benzannulation on the N^N or C^N ligands of the Ir(III) complexes.[11,21] Depending on the benzannulation site on the N^N ligand, the Ir(III) complexes can exhibit a blue- or red-shift of the lowest-energy absorption and emission bands.[22] This is attributed to the site-dependent destabilization or stabilization of the excited electron state upon benzannulation on the N^N ligand of the Ir(III) complexes,[23] while benzannulation on the C^N ligands either decreased the emission energy or had a negligible effect on the emission energy depending on the site of benzannulation.[11]

Despite these abundant experimental and computational studies of the optical properties of Ir(III) complexes,[13,20,21,33] establishing the relationship between the structures of the ligands and the emission properties of the Ir(III) complexes for a rational design of complexes with desired emission properties is still a complicated task due to scarce and insufficient experimental publicly available data. In this regard, machine learning (ML) methods can be very helpful in finding structure−property correlations in Ir(III) complexes, which can combine various data and extract the essential parameters. Thus, data-driven modeling techniques like quantitative structure−property relationship (QSPR) can be used to establish mathematical correlations between the optical properties of the Ir(III) complexes and their various structural features like the π-conjugation of the ligands (N^N and C^N ligands), benzannulation on the different sites of the ligands, electron-withdrawing or -donating substituents, etc. QSPR modeling has been successfully used to model structures and the ground-state electronic properties for polymers, organic and inorganic compounds, semiconductors, and organometallic complexes.[36−38] A QSPR has also been successfully applied to generate a toxicity model for organometallic complexes.[39] The relative binding affinity of diverse organometallic estradiol derivatives has been modeled using the QSPR approach, as well.[40] In addition, attempts to learn the transition density for phosphorescence emission and excited states in general were made.[41] Recently, two studies related to structure−property relationship analysis for a series of Ir(III) complexes were reported.[42,43] Bernhard and co-workers investigated the emission properties of a large set of Ir(III) complexes by applying a combinatorial synthesis and linear structure−emission wavelength relationship analysis,[42] while Kulik and co-workers discovered a low-cost ML prediction of the excited-state properties of Ir(III) phosphors using an artificial neural network (ANN) model.[43] However, both structure−property models in these two studies have some drawbacks due to the complexity of the ML algorithms used, limiting their use for structural analysis. No study has yet correlated the phosphorescence properties of the Ir(III) complexes to the structures of the complexes and/or ligands with a specific structure-based mechanistic explanation and a transparent, reproducible model. Deriving a transparent mathematical relationship between the emission or absorption wavelengths of the Ir(III) complexes and the structural features of the ligands in Ir(III) complexes from QSPR models can help in the development of Ir(III) complexes with predetermined optical properties.

In this work, we collected 47 different Ir(III) complexes with diverse structural features that were characterized experimentally by Sun and co-workers previously[13,17,21,23−35] to reveal the relationship between the emission energy and the structural feature and then to develop a QSPR model. The investigated Ir(III) complexes have diverse electron-withdrawing or -donating substituents, varying degrees of π-conjugation on C^N and/or N^N ligands, and benzannulation on the different sites of the ligands. In this study, we hope to predict the triplet emission wavelength of a set of Ir(III) complexes with structural complexities by using the unique combination of experiment, quantum chemistry, and machine learning-based QSPR methods, as illustrated in Figure 1. An

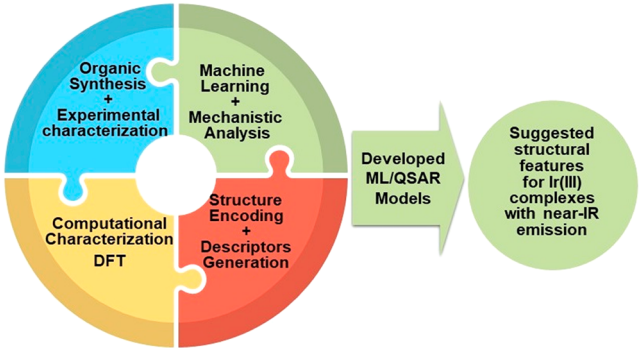


Figure 1. Scheme of the workflow applied to investigate structurally different Ir(III) complexes by combined experimental, computational, and machine learning methods.

interpretative ML-based QSPR model that can reliably predict the emission wavelength of the Ir(III) complexes on the basis of their structural modifications and provide the foundation for the theoretical evaluation of the optical properties of the Ir(III) complexes was successfully developed. It is worth noting that a relation between a decrease in the N−N distance of a pair of nitrogen atoms coordinated to the Ir(III) metal center and the red-shift of the emission wavelength is revealed. Interestingly, this relationship has never before been reported in the literature. We associate these correlations with a degree of distortion in the octahedral structures of the complexes, which results in a perturbed ligand field. Through our combined experimental and QSPR-based computational approach, we aim to pave a pathway for a cost-effective and efficient development of NIR-emitting complexes for photonic and biophotonic applications.

In this work, we have investigated in-house data of 47 Ir(III) complexes with diverse structural features that were synthesized and characterized experimentally and computationally to reveal a relationship between the emission values and the structural features and to develop a QSPR model. The data for these 47 complexes were selected to develop a machine learning and cheminformatics technique for further virtual screening purposes. Also, it was important to collect data that were measured under the same experimental conditions to improve the accuracy of the models and lower the probability of unnecessary deviations that usually come with data from different groups and experimental conditions.

https://doi.org/10.1021/acs.jpclett.3c02533
J. Phys. Chem. Lett. 2024, 15, 471−480

Table 1. Models (1−7) and Associated Features and Descriptors of Each Model

	descriptors	$R^2_{training}$	$RMSE_C$	Q^2_{LOO}	$CCC_{training}$	F test	R^2_{test}	$RMSE_{test}$
model 1	GATS4i	0.55	0.10	0.50	0.51	40.50	0.67	0.08
model 2	SpMin3_Bh(m) F08[N−N]	0.81	0.09	0.74	0.62	31.00	0.75	0.07
model 3	SpMin3_Bh(m) Eig04_EA(dm) B08[N−N]	0.84	0.08	0.78	0.71	31.09	0.87	0.06
model 4	SpMin3_Bh(m) Chi0_EA(dm) Mor29s F08[N−N]	0.89	0.08	0.84	0.79	34.54	0.84	0.06
model 5	GATS5i SpMin3_Bh(m) Mor28i Mor29s F08[N−N]	0.92	0.07	0.86	0.81	32.10	0.80	0.06
model 6	SM1_Dz(i) GATS5i SpMin3_Bh(m) Mor28i Mor29s F08[N−N]	0.93	0.07	0.87	0.85	33.20	0.77	0.06
model 7	ZM2Kup GATS5i SpMin3_Bh(m) Mor28m Mor21s Mor29s F08[N−N]	0.95	0.06	0.91	0.86	30.80	0.68	0.06

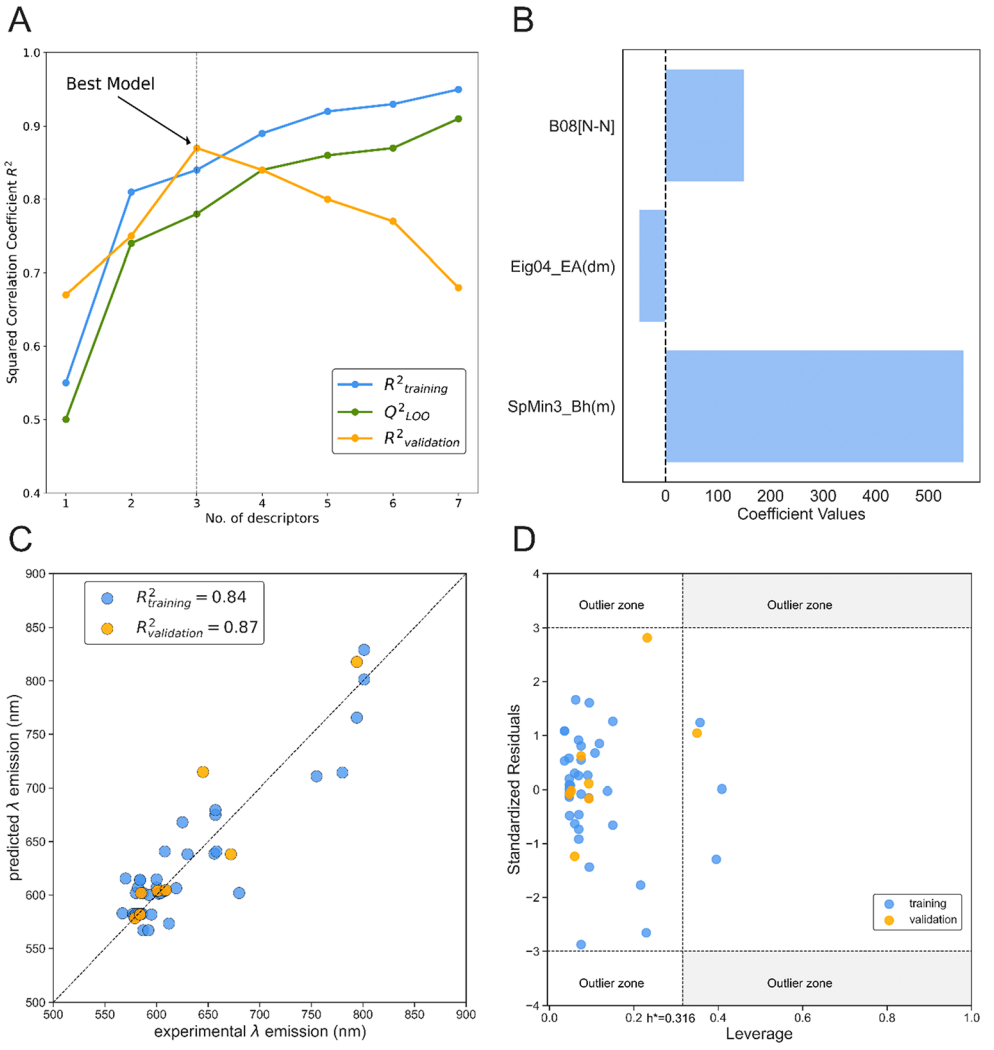


Figure 2. (A) Statistical data for correlation coefficients R^2 and Q^2 for models 1−7 for training and external (test) sets. (B) Magnitude of the influence of different descriptors of the three-variable model on the emission wavelength (λ, nanometers) according to eq 7. (C) Correlation plot of the observed and predicted values of emission wavelength (nanometers). (D) Williams plot of standardized residual vs leverage.

The development of the model was conducted systematically from the first model to the seventh model by accessing the regression coefficient of the training set in each model, followed by validation of the test set (Table 1).

Table 1 shows that the R^2_{train} values show a marching trend with an increase in the number of variables in the developed model. The same behavior is exhibited by R^2_{test} but up to three variables only, and then it showed plummeting behavior that manifests that the predictive capability of the model beyond the three variables is decreased due to overfitting issues. At the same time, the predictive coefficients (R^2 and Q^2) of the training set increase for the higher-magnitude models (4−7), while its predictive capability for the test set declines. This trend is illustrated in Figure 2A. Overall, the three-variable model shows a good combination of high R^2_{train} and $R^2_{test/validation}$ values and, therefore, is selected as the best and most accurate model among all seven.

The selected three-variable model predicts the emission wavelength (nanometers) according to eq 1:

$$\lambda_{emission} = 567.03 \times \text{SpMin3\_Bh(m)} - 50.23 \times \text{Eig04\_EA(dm)} + 149.5 \times \text{B08[N−N]} - 537.26 \quad (1)$$

https://doi.org/10.1021/acs.jpclett.3c02533
J. Phys. Chem. Lett. 2024, 15, 471−480

where n = 46, R^2_{train} = 0.84, R^2_{test} = 0.87, Q^2_{LOO} = 0.78, $RMSE_{train}$ = 26.72, $RMSE_{cv}$ = 31.2, $RMSE_{test}$ = 29.3, and F = 61.3. The descriptors are the specific structure-related indexes and listed in Table 2, while the obtained statistical parameters are explained in Materials and Methods and Table 1.

Table 2. Descriptors for the Three-Variable Model, Their Types, and Short Descriptions

descriptor	descriptor information	type
SpMin3_Bh(m)	smallest eigenvalue *n*, three of the burden matrix weighted by mass	two-dimensional matrix
Eig04_EA(dm)	eigenvalue *n*, four from the edge adjacency mat. weighted by the dipole moment	electrotopological
B08[N−N]	presence or absence of N−N at topological distance 8	two-dimensional atom pair

The descriptors of the three-variable model listed in Table 2 can be broadly classified as two-dimensional (2D) matrix, electrotopological, and atom pair descriptors. The model shows that the emission wavelength (nanometers) strongly correlates with molecular descriptors such as the weighted mass, weighted dipole moment, and N−N topological distance in the ligands of investigated Ir(III) complexes. Moreover, the mass or bulkiness of the ligand significantly influences the emission wavelength, thereby contributing to the optical properties of the complexes.

Figure 2B shows the magnitude of the influence of various descriptors on the emission wavelength (λ in nanometers) according to the three-variable model (eq 1). The positive or negative influence of each descriptor on λ is represented with a corresponding coefficient value. SpMin3_Bh(m) and B08[N−N] have a positive impact on the emission wavelength (λ in nanometers) of the Ir(III) complexes, indicating the red-shift of wavelength (λ in nanometers). The model shows that the larger mass of the ligand and the larger topological distance between the N atoms in the ligand can red-shift the emission wavelength; on the contrary, the larger dipole moment of the ligand tends to reduce the emission wavelength (λ in nanometers).

The predictive ability of the three-variable model (R^2 = 0.84 for the training set, and R^2 = 0.87 for the test set) shows a very good correlation between the predicted and observed emission wavelengths for a set of 46 different Ir(III) complexes (Figure 2C). Thus, the three-variable model is very robust. It is worth noting that complex **20** was excluded because it was not part of the applicability domain for the model (see Figure 2D, Williams plot).

In Figure 2C, the Williams plot for the three-variable model is presented. The plot allows a graphical detection of both the outliers for the response and the structurally different chemicals in the model. It consists of plotting the standardized residuals on the Y-axis and the leverage value on the X-axis. Observation with standardized residuals greater than the (-3σ; $+3\sigma$) range is considered an outlier response. The leverage value represents the degree of influence that the structure of each chemical has on the model. The Williams plot shown in Figure 2D confirms that all 46 compounds are located within the $\pm 3\sigma$ of error limit, indicating all complexes considered for prediction in the three-variable model fall within the applicability domain.

For further validation of the three-variable model, we show the y-scrambling plot in Figure S1. The y-scrambling plot is considered as a validation technique that confirms the

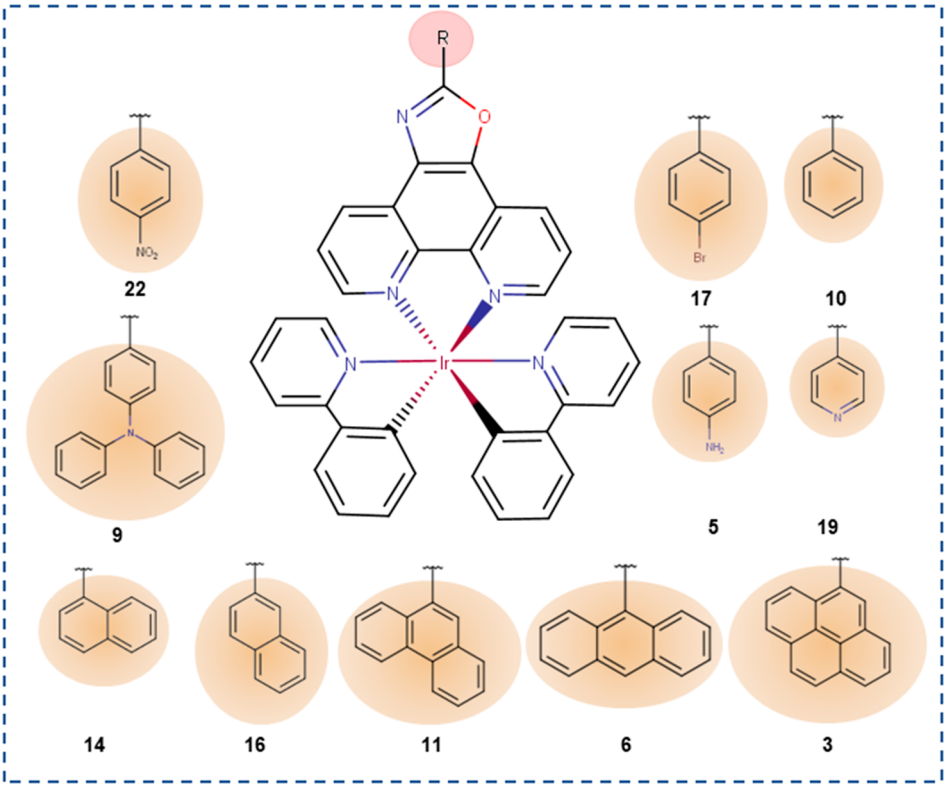


Figure 3. Subset of structures for Ir(III) complexes used to illustrate that a topological parameter, such as the N−N distance between the coordinated nitrogens in the oxazolo[4,5-*f*][1,10]phenanthroline ligand, directly corelates to the emission wavelength of these complexes.[13]

https://doi.org/10.1021/acs.jpclett.3c02533
J. Phys. Chem. Lett. 2024, 15, 471−480

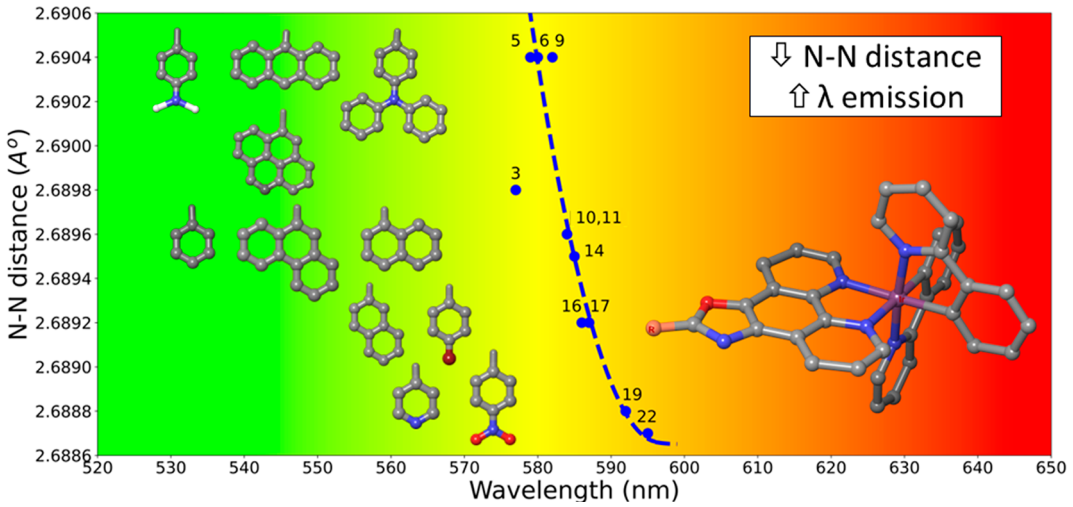


Figure 4. Schematic representation of the relationship between the N−N topological distance and the wavelength of emission (colored) for the selected complexes (the structures of the fragments of these complexes are given on the left side of the figure, in parallel).[13]

robustness and uniqueness of the best selected model. The *y*-scrambling plot is obtained by scrambling the *Y* values (experimental emission wavelength) randomly. In this randomization process, 2000 simulations per model are conducted. None of the random sets (artificially created sets) show any good correlation (chance correlation) in comparison with the selected three-variable model. The R^2 and Q^2 values of the selected model are significantly higher than all other R^2 and Q^2 values for simulated random sets (Figure S1). Thus, the *y*-scrambling plot confirms that the developed model is a case of robustness and high correlation rather than a mere coincidence. The essence of these findings is that three detected physicochemical and electrotopological features of the complexes (Table 2) play a pivotal role in determining the emission energy of Ir(III) complexes. We also verified that random splits and cross-validation considering the entire data set yield similar accuracy, and the results are listed in Table S2.

One of the topological parameters from our model affecting the optical performance of Ir(III) complexes is the distance between the nitrogens coordinated to the metal center (N−N topological distance in Table 2). To better understand this correlation, we closely consider 11 cationic Ir(III) complexes bearing similar N^N ligands coordinated to the metal center but with varied electron-donating or electron-withdrawing substituents (complexes **5**, **9**, **10**, **17**, and **22**) at position 4 of the 2-phenyl ring or different aromatic substituents with varied conjugation lengths, which are introduced by the number of joined aromatic rings in the ligand (complexes **3**, **6**, **11**, **14**, **16**, and **19**) on the oxazolo[4,5-*f*][1,10]phenanthroline ligand. Schematic structures of these complexes are shown in Figure 3.

The N−N topological distance for these Ir(III) complexes as a function of their corresponding emission wavelength is shown in Figure 4, demonstrating the inverse relationship between the N−N distance and the emission wavelength in all 11 complexes. The change in the N−N topological distance is associated with both electron-withdrawing and -donating groups and the conjugation length, which is introduced by the number of joined aromatic rings in the substituents. In particular, the N−N distance is reduced with an increase in the electron-withdrawing ability for complexes **22**, **17**, and **19** (NO_2, Br, and pyridyl groups, respectively) and with a decrease in the conjugated length of substituents in complexes **3**, **6**, **11**, **14**, **16**, and **19**. A decrease in the N−N distance is accompanied by an increase in the emission wavelength (increased red-shift). This correlation between the N−N distance of nitrogens coordinated with the metal center and the red-shift in the emission energy can be rationalized by geometrical distortions in the octahedral ligand field contributed by substituents (e.g., shortening of the coordinated bonds laying at the same surface). This shortening of coordinated bonds results in stronger stabilization of the hole state contributing to the lowest MLCT transitions, compared to less distorted cases of complexes with electron-donating ability (**5**, **9**, and **10**) and greater π-conjugation of substituents (**3**, **6**, and **11**). Thus, the complexes with shorter N−N distances show larger red-shifts in their triplet emission energies. To validate the MLR model's utility for simulating photoemitting properties, we collected a subset of Ir complexes from another data set,[33] not part of any original training or test set, and the emission wavelengths were predicted using the developed MLR model (eq 1); we found that the predictions were quite accurate, according to the original model's performance level (see Table S3).

This study combines experimental characterization, computational modeling, and machine learning approaches to study the emission energies of 47 Ir(III) complexes. A robust and interpretative machine learning-based QSPR three-variable model was developed with the following performance parameters: R^2 = 0.84 for the training set, and R^2 = 0.87 for the test set. The selected descriptors show that the emission wavelength strongly correlates with the following molecular descriptors: weighted mass, weighted dipole moment, and N−N topological distance in the ligands of the investigated Ir(III) complexes. In other words, the mass and bulkiness of the ligand, its dipole moment, and specific N−N topological distance in the ligand significantly influence the emission wavelength. These findings reveal the structural parameters of Ir(III) complexes that affect the emission energies and allow for a reliable prediction of the emission energies of new Ir(III) complexes and other octahedral transition-metal complexes.

On the basis of the developed ML/QSPR model, a hypothesis was formulated to explain the differences in emission energies among structurally different Ir(III) complexes. It was discovered that the emission energies are mainly

https://doi.org/10.1021/acs.jpclett.3c02533
J. Phys. Chem. Lett. 2024, 15, 471−480

related to the N−N distance between a pair of nitrogen atoms coordinated to the Ir(III) center, which is associated with the degree of distortion in the octahedral structure of the complexes and induces a perturbed ligand field. The combined experimental and computational approach demonstrated in this study sheds light on how to effectively design new organometallic complexes with the desired optical properties.

MATERIALS AND METHODS

The sequence of the tasks in this work to develop the predictive ML/QSPR model is based on the following steps: (1) data set selection and/or curation, (2) molecular structure construction, (3) three-dimensional geometry optimization using quantum-chemical methods, (4) generation of a set of numerical descriptors (features), (5) variable selection step or data reduction, (6) development and selection of models, and (7) validation of the developed models and predictability evaluation.[35,44]

Data Set Selection: Emission Data Set of Ir(III) Complexes. The experimental values of the emission wavelengths of 47 different Ir(III) complexes are collected from in-house data published in various papers (Table 3).[13,17,21,23−35] The collected data set of these complexes was split into training and test sets (80% and 20%, respectively); the external set was used to validate the developed model. The specific procedure was applied within the splitting of the data set into training and test sets, where the data set is arranged in ascending order, and every fifth molecule was taken from the data set to create an external test set.

Molecular Structure Generation: Density Functional Theory (DFT) Structure Optimization. All investigated complexes were optimized at the DFT level using Gaussian and 09 and Gaussian 16.[46] Thus, the geometries of investigated systems were optimized in their singlet and triplet spin configurations by DFT, applying the hybrid Perdew, Burke, and Ernzerh functional (PBE0)[47,48] and the LANL2DZ[49] basis sets assigned for Ir(III) and the 6-31G* basis set for all of the remaining atoms.[50,51] It worth noting that a more advanced DFT methodology is instrumental in improving the accuracy of calculations. Thus, it was shown that the functionals accounting for dispersion (such as wB97XD, M06, and M06L) and extended basis sets improve the structures of Ru(II) complexes, in particular that of the coordination center.[52,53] However, our previous investigations of Ir(III) complexes have shown that minor changes in the geometry of Ir(III) complexes caused by more advanced functionals and basis sets have an insignificant effect on their absorption spectra. In contrast, the choice of applied functional strongly impacts the optical spectra. For example, long-range corrected functionals, such as wB97XD, usually exhibit over-blue-shifted absorption bands compared to related experimental spectra.[54−56] Therefore, we have chosen the hybrid PBE0 functional for our calculations, because it shows a good compromise in the accuracy of the ground- and excited-state calculations, as evidenced by comparison with experimental data.[13,17,21,23−34]

In this study, the effective core potential (ECP) basis, such as LANL2DZ for transition metals, was applied. At the same time, all-electron basis sets were used for all other non-transition-metal atoms. This combination of basis sets applied here is a common practice in computations of systems containing transition metals.[57] Thus, ECPs are parametrized to implicitly take into account the scalar relativistic (SR) effects, which is important in application to transition metals. It is worth noting that the two-electron spin−orbit coupling integrals are implemented to an effective one-electron operator applied by Koseki et al. for the elements across the periodic table as an effective nuclear charge,[58] which are used in the LANL2DZ basis set mentioned above.

Here, dichloromethane was used as the solvent within the conductor-like polarizable continuum model (CPCM)[59] reaction field approach. This methodology is described in detail in our previous publications and has been shown to provide good agreement with experimental structural and optical data of reported Ir(III) complexes.[13,17,21,23−34]

Table 3. Data Set of 47 Investigated Ir(III) Complexes

complex	status	experimental λ_{em} (nm)	predicted λ_{em} (nm)	$\Delta\lambda_{em}$ (nm)	ref
1	training	567	582.91	15.91	32
2	training	570	615.48	45.48	30
3	training	577	582.34	5.34	13
4	training	578	580.07	2.07	32
5	prediction	579	578.37	−0.63	13
6	training	580	582.34	2.34	13
7	training	580	601.87	21.87	30
8	training	582	606.97	24.97	35
9	training	582	582.34	0.34	13
10	prediction	584	581.77	−2.23	13
11	training	584	582.34	−1.66	13
12	training	584	614.09	30.09	32
13	training	584	614.09	30.09	31
14	training	585	582.34	−2.66	13
15	prediction	585	601.87	16.87	30
16	training	586	582.34	−3.66	13
17	training	587	567.03	−19.97	13
18	training	587	601.87	14.87	30
19	training	592	567.03	−24.97	13
20	excluded	593			21
21	training	593	600.03	7.03	32
22	training	595	581.77	−13.23	13
23	training	600	606.97	6.97	35
24	training	600	614.66	14.66	32
25	prediction	601	604.00	3.00	29
26	training	602	601.37	−0.63	29
27	training	604	601.87	−2.13	30
28	training	608	604.00	−4.00	35
29	training	608	640.85	32.85	17
30	prediction	609	604.56	−4.44	35
31	training	612	573.38	−38.62	29
32	training	619	606.40	−12.60	31
33	training	625	668.21	43.21	21
34	training	630	638.16	8.16	31
35	prediction	645	714.70	69.70	21
36	training	656	638.72	−17.28	31
37	training	657	675.01	18.01	17
38	training	657	679.55	22.55	21
39	training	658	640.85	−17.15	17
40	prediction	672	638.16	−33.84	31
41	Training	680	601.87	−78.13	30
42	training	755	710.73	−44.27	21
43	training	780	714.14	−65.86	21
44	training	794	765.62	−28.38	25
45	prediction	794	817.79	23.79	25
46	training	801	801.25	0.25	25
47	training	801	829.13	28.13	25

https://doi.org/10.1021/acs.jpclett.3c02533
J. Phys. Chem. Lett. 2024, 15, 471−480

Generation of the Descriptors. A specific set of molecular features and/or descriptors was calculated for all complexes in this work; descriptors are mathematical representations of chemical information contained in a structure of an investigated large molecule. The structures in this study were constructed using ChemSketch,[60] followed by DFT optimization (see Molecular Structure Generation: Density Functional Theory (DFT) Structure Optimization), where the optimized singlet structure was then used for further steps by cheminformatics methods to generate descriptors, by applying Dragon 6.[61] Initially, Dragon 6 can generate 4885 various molecular descriptors. These descriptors are divided into 29 logical blocks, where each block is divided into sub-blocks. The 4885 descriptors correspond to zero-, one-, two-, and three-dimensional structure-based indexes, including functional groups and molecular property descriptors.[62] After the constant and near-constant descriptors had been filtered out (>0.95 similarity, i.e., >95% of values between a pair of descriptors are similar), ∼2863 descriptors per organometallic structure were used for further steps in structure−property modeling.

ML/QSPR Model Development and Validation. ML-based QSPR modeling is applied to determine the relationship between the physical property (NIR emission values) of a molecular system and its structure. Then ML-based QSPR model is used to predict the physical property of new compounds before expensive experimental synthesis and physicochemical testing.

In this work, the correlation between the NIR emission and the structural features of the Ir(III) complexes was developed by using the approach that combines a genetic algorithm (GA)[63] for variable selection and multiple linear regression (MLR) for final ML/QSAR model generation. In fact, a GA-based variable selection method has been applied in many QSAR studies.[36,45,65] The applied GA variable selection in this study started with a population of 500 random models, 5000 iterations, and a mutation probability of 85%. The model development was done by applying QSARINS version 2.24.[64] The systematic study performed here consists of several QSPR models (from one to seven variables in each selected best model), followed by comprehensive statistical analysis with accuracy estimation by a squared correlation coefficient R^2, "leave-one-out" squared coefficient Q^2, root-mean-square error (RMSE), and Fisher coefficient (F) to find the best ML/QSPR model.

To ensure the reliability of the data produced by the models, the validation of the ML/QSPR model is very important. Thus, internal and external validations are considered to be necessary for checking the robustness of the model. In the internal validation process, the final set of QSPR models was developed by using the GA-MLR approach and then tested by applying the cross-validation "leave-one-out" technique Q^2_{LOO}. Thus, we utilized the following equations to calculate correlation coefficient R^2 (eq 2) and RMSE as measures of the goodness of fit for each developed model (eq 3).

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i^{obs} - y_i^{pred})^2}{\sum_{i=1}^{n} (y_i^{obs} - \tilde{y}^{obs})^2} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i^{obs} - y_i^{pred})^2}{n}} \quad (3)$$

In the leave-one-out (LOO) method of cross validation (CV) step, the procedure for removing a molecule, in which the model is applied to the set and validated against the individual molecules, is performed for the entire training set. Once the procedure is completed, the mean Q^2 value is reported. This CV procedure is applied to estimate the predictive ability of the model and to avoid overfitting of the model. Cross-validated coefficient Q^2_{LOO} was calculated according to eq 4:

$$Q^2_{LOO} = 1 - \frac{\sum_{i=1}^{n} (y_i^{obs} - y_i^{predcv})^2}{\sum_{j=1}^{n} (y_j^{obs} - \tilde{y}^{obs})^2} \quad (4)$$

Next, external validation of the model is very important. The external validation is checked for its ability to predict new compounds that are not included in the training set. This is done by applying the model obtained on the basis of the training set to the external (test) set. The external predictivity of each model is represented by R^2_{test}, which is calculated in a manner similar to eq 2, but for the test set. In addition, the applicability domain (AD) for the developed models was also calculated, by applying the leverage approach, to verify the predictive reliability for a certain data set.[66] The Williams plot was used to visualize the applicability domain of the ML/QSPR models. In the Williams plot for the specific model, the data point based on standardized cross-validated residuals (RES) versus leverage (Hat diagonal) values (HAT) depicts both the response outliers, as Y outliers, and structurally influential compounds, as X outliers.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.3c02533.

Tables S1−S3 and Figures S1−S4 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Bakhtiyor Rasulev − *Coatings and Polymeric Materials, North Dakota State University, Fargo, North Dakota 58108, United States;* orcid.org/0000-0002-7845-4884; Email: bakhtiyor.rasulev@ndsu.edu

Authors

Anas Karuth − *Coatings and Polymeric Materials, North Dakota State University, Fargo, North Dakota 58108, United States*

Gerardo M. Casanola-Martin − *Coatings and Polymeric Materials, North Dakota State University, Fargo, North Dakota 58108, United States;* orcid.org/0000-0003-0383-2032

Levi Lystrom − *Department of Chemistry and Biochemistry, North Dakota State University, Fargo, North Dakota 58108, United States;* orcid.org/0000-0001-6369-8643

Wenfang Sun − *Department of Chemistry and Biochemistry, North Dakota State University, Fargo, North Dakota 58108, United States; Department of Chemistry and Biochemistry, The University of Alabama, Tuscaloosa, Alabama 35487, United States;* orcid.org/0000-0003-3608-611X

Dmitri Kilin − *Department of Chemistry and Biochemistry, North Dakota State University, Fargo, North Dakota 58108, United States;* orcid.org/0000-0001-7847-5549

https://doi.org/10.1021/acs.jpclett.3c02533
J. Phys. Chem. Lett. 2024, 15, 471−480

Svetlana Kilina − *Department of Chemistry and Biochemistry, North Dakota State University, Fargo, North Dakota 58108, United States;* orcid.org/0000-0003-1350-2790

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.3c02533

Author Contributions

A.K., G.M.C.-M., and B.R. designed and implemented the methodology. W.S. carried out the empirical testing of the emission properties of the organometallic complexes. L.L. designed and optimized the molecular complexes for the calculations. L.L., D.K., and S.K. performed quantum-mechanical calculations and analyzed the computational results. A.K., G.M.C.-M., and B.R. selected the optimal structural descriptors, developed and validated the ML/QSPR model, prepared the draft version of the manuscript, and discussed the ML/QSPR results. S.K., W.S., D.K., and B.R. discussed the obtained results and edited the manuscript.

Notes

The authors declare no competing financial interest.

Additional data, including the *xyz* coordinates of the optimized Ir(III) complexes, are available at https://github.com/bakhras/Ir-III-complexes-data.

■ ACKNOWLEDGMENTS

The authors acknowledge the main support from the National Science Foundation (NSF) under Grant CHE-1800476. This work is also supported in part by NSF MRI Grant OAC-2019077 and ND EPSCoR Grant IIA-1355466 and by the State of North Dakota. The authors also acknowledge the support of the U.S. Department of Energy (DOE) under Grant DE-SC0022239 for developing the improved feedback loop between ML-based predicted data and DFT-based calculations of structural and optical properties. The authors thank Prof. Paola Gramatica for generously providing a free license for the QSARINS software. Supercomputing support (Thunder cluster) from CCAST HPC System at North Dakota State University is acknowledged.

■ REFERENCES

(1) Radwan, Y. K.; Maity, A.; Teets, T. S. Manipulating the Excited States of Cyclometalated Iridium Complexes with β-Ketoiminate and β-Diketiminate Ligands. *Inorg. Chem.* **2015**, *54* (14), 7122−7131.

(2) Goswami, S.; Sengupta, D.; Paul, N. D.; Mondal, T. K.; Goswami, S. Redox Non-Innocence of Coordinated 2-(Arylazo) Pyridines in Iridium Complexes: Characterization of Redox Series and an Insight into Voltage-Induced Current Characteristics. *Chem. - Eur. J.* **2014**, *20* (20), 6103−6111.

(3) Gärtner, F.; Cozzula, D.; Losse, S.; Boddien, A.; Anilkumar, G.; Junge, H.; Schulz, T.; Marquet, N.; Spannenberg, A.; Gladiali, S.; et al. Synthesis, Characterisation and Application of Iridium(III) Photosensitisers for Catalytic Water Reduction. *Chem. - Eur. J.* **2011**, *17* (25), 6998−7006.

(4) Fan, C.; Zhu, L.; Jiang, B.; Li, Y.; Zhao, F.; Ma, D.; Qin, J.; Yang, C. High Power Efficiency Yellow Phosphorescent OLEDs by Using New Iridium Complexes with Halogen-Substituted 2-Phenylbenzo-[*d*]thiazole Ligands. *J. Phys. Chem. C* **2013**, *117* (37), 19134−19141.

(5) Tsuboyama, A.; Iwawaki, H.; Furugori, M.; Mukaide, T.; Kamatani, J.; Igawa, S.; Moriyama, T.; Miura, S.; Takiguchi, T.; Okada, S.; et al. Homoleptic Cyclometalated Iridium Complexes with Highly Efficient Red Phosphorescence and Application to Organic Light-Emitting Diode. *J. Am. Chem. Soc.* **2003**, *125* (42), 12971−12979.

(6) Bolink, H. J.; Cappelli, L.; Coronado, E.; Grätzel, M.; Ortí, E.; Costa, R. D.; Viruela, P. M.; Nazeeruddin, M. K. Stable Single-Layer Light-Emitting Electrochemical Cell Using 4,7-Diphenyl-1,10-phenanthroline-bis(2-phenylpyridine)iridium(III) Hexafluorophosphate. *J. Am. Chem. Soc.* **2006**, *128* (46), 14786−14787.

(7) Sun, L.; Galan, A.; Ladouceur, S.; Slinker, J. D.; Zysman-Colman, E. High stability light-emitting electrochemical cells from cationic iridium complexes with bulky 5,5′ substituents. *J. Mater. Chem.* **2011**, *21* (44), 18083.

(8) Lo, K. K.-W.; Chung, C.-K.; Lee, T. K.-M.; Lui, L.-H.; Tsang, K. H.-K.; Zhu, N. New Luminescent Cyclometalated Iridium(III) Diimine Complexes as Biological Labeling Reagents. *Inorg. Chem.* **2003**, *42* (21), 6886−6897.

(9) Ma, D.-L.; Lin, S.; Wang, W.; Yang, C.; Leung, C.-H. Luminescent chemosensors by using cyclometalated iridium(iii) complexes and their applications. *Chemical science* **2017**, *8* (2), 878−889.

(10) Sun, J.; Wu, W.; Guo, H.; Zhao, J. Visible-Light Harvesting with Cyclometalated Iridium(III) Complexes Having Long-Lived ^{3} IL Excited States and Their Application in Triplet−Triplet-Annihilation Based Upconversion. *Eur. J. Inorg. Chem.* **2011**, *2011* (21), 3165−3173.

(11) Liu, B.; Lystrom, L.; Kilina, S.; Sun, W. Effects of Varying the Benzannulation Site and π Conjugation of the Cyclometalating Ligand on the Photophysics and Reverse Saturable Absorption of Monocationic Iridium(III) Complexes. *Inorg. Chem.* **2019**, *58* (1), 476−488.

(12) Dragonetti, C.; Righetto, S.; Roberto, D.; Ugo, R.; Valore, A.; Fantacci, S.; Sgamellotti, A.; De Angelis, F. Cyclometallated iridium(iii) complexes with substituted 1,10-phenanthrolines: a new class of highly active organometallic second order NLO-phores with excellent transparency with respect to second harmonic emission. *Chem. Commun.* **2007**, No. 40, 4116.

(13) Zhu, X.; Cui, P.; Kilina, S.; Sun, W. Multifunctional Cationic Iridium(III) Complexes Bearing 2-Aryloxazolo[4,5-*f*][1,10]-phenanthroline (N̂N) Ligand: Synthesis, Crystal Structure, Photophysics, Mechanochromic/Vapochromic Effects, and Reverse Saturable Absorption. *Inorg. Chem.* **2017**, *56* (22), 13715−13731.

(14) Liu, B.; Lystrom, L.; Cameron, C. G.; Kilina, S.; McFarland, S. A.; Sun, W. Monocationic Iridium(III) Complexes with Far-Red Charge-Transfer Absorption and Near-IR Emission: Synthesis, Photophysics, and Reverse Saturable Absorption. *Eur. J. Inorg. Chem.* **2019**, *2019* (16), 2208−2215.

(15) Acharyya, J. N.; Desai, N. R.; Gangineni, R. B.; Vijaya Prakash, G. Effect of Photonic Cavity Interactions on Femtosecond Multi-photon Optical Nonlinear Absorptions from Bi2O3-Based One-Dimensional Photonic Crystal. *ACS Photonics* **2022**, *9* (6), 2092−2100.

(16) Yogeswari, C.; Hijas, K. M.; Girisun, T. C. S.; Nagalakshmi, R. Intensity dependent nonlinear absorption switching behavior of electrospun meta-nitroaniline nanofiber. *Opt. Mater.* **2020**, *100*, 109691.

(17) Zhu, X.; Lystrom, L.; Kilina, S.; Sun, W. Tuning the Photophysics and Reverse Saturable Absorption of Heteroleptic Cationic Iridium(III) Complexes via Substituents on the 6,6′-Bis(fluoren-2-yl)-2,2′-biquinoline Ligand. *Inorg. Chem.* **2016**, *55* (22), 11908−11919.

(18) Valore, A.; Cariati, E.; Dragonetti, C.; Righetto, S.; Roberto, D.; Ugo, R.; De Angelis, F.; Fantacci, S.; Sgamellotti, A.; Macchioni, A.; et al. Cyclometalated IrIII Complexes with Substituted 1,10-Phenanthrolines: A New Class of Efficient Cationic Organometallic Second-Order NLO Chromophores. *Chem. - Eur. J.* **2010**, *16* (16), 4814−4825.

(19) Langhals, H. Color Chemistry. Synthesis, Properties and Applications of Organic Dyes and Pigments. 3rd revised edition. By Heinrich Zollinger. *Angew. Chem., Int. Ed.* **2004**, *43* (40), 5291−5292.

(20) Sun, W.; Zhang, B.; Li, Y.; Pritchett, T. M.; Li, Z.; Haley, J. E. Broadband Nonlinear Absorbing Platinum 2,2′-Bipyridine Complex

Bearing 2-(Benzothiazol-2′-yl)-9,9-diethyl-7-ethynylfluorene Ligands. *Chem. Mater.* **2010**, *22* (23), 6384−6392.
(21) Liu, B.; Lystrom, L.; Kilina, S.; Sun, W. Tuning the Ground State and Excited State Properties of Monocationic Iridium(III) Complexes by Varying the Site of Benzannulation on Diimine Ligand. *Inorg. Chem.* **2017**, *56* (9), 5361−5370.
(22) Forde, A.; Lystrom, L.; Sun, W.; Kilin, D.; Kilina, S. Improving Near-Infrared Emission of meso-Aryldipyrrin Indium(III) Complexes via Annulation Bridging: Excited-State Dynamics. *J. Phys. Chem. Lett.* **2022**, *13* (39), 9210−9220.
(23) Wang, C.; Lystrom, L.; Yin, H.; Hetu, M.; Kilina, S.; McFarland, S. A.; Sun, W. Increasing the triplet lifetime and extending the ground-state absorption of biscyclometalated Ir(III) complexes for reverse saturable absorption and photodynamic therapy applications. *Dalton Trans.* **2016**, *45* (41), 16366−16378.
(24) Lu, T.; Wang, C.; Lystrom, L.; Pei, C.; Kilina, S.; Sun, W. Effects of extending the π-conjugation of the acetylide ligand on the photophysics and reverse saturable absorption of Pt(II) bipyridine bisacetylide complexes. *Phys. Chem. Chem. Phys.* **2016**, *18* (41), 28674−28687.
(25) Wang, L.; Yin, H.; Cui, P.; Hetu, M.; Wang, C.; Monro, S.; Schaller, R. D.; Cameron, C. G.; Liu, B.; Kilina, S.; et al. Near-infrared-emitting heteroleptic cationic iridium complexes derived from 2,3-diphenylbenzo[g]quinoxaline as in vitro theranostic photodynamic therapy agents. *Dalton Transactions* **2017**, *46* (25), 8091−8103.
(26) Pritchett, T. M.; Ferry, M. J.; Mott, A. G.; Shensky, W.; Haley, J. E.; Liu, R.; Sun, W. Long-lifetime reverse saturable absorption in a bipyridyl platinum(II) complex bearing naphthalimidylethynyl-substituted fluorenylacetylide ligands. *Opt. Mater.* **2015**, *39*, 195−198.
(27) Pei, C.; Cui, P.; McCleese, C.; Kilina, S.; Burda, C.; Sun, W. Heteroleptic cationic iridium(III) complexes bearing naphthalimidyl substituents: synthesis, photophysics and reverse saturable absorption. *Dalton Trans.* **2015**, *44* (5), 2176−2190.
(28) Wang, L.; Cui, P.; Kilina, S.; Sun, W. Toward Broadband Reverse Saturable Absorption: Investigating the Impact of Cyclometalating Ligand π-Conjugation on the Photophysics and Reverse Saturable Absorption of Cationic Heteroleptic Iridium Complexes. *J. Phys. Chem. C* **2017**, *121* (10), 5719−5730.
(29) Li, Y.; Dandu, N.; Liu, R.; Kilina, S.; Sun, W. Synthesis and photophysics of reverse saturable absorbing heteroleptic iridium(III) complexes bearing 2-(7-R-fluoren-2′-yl)pyridine ligands. *Dalton Trans.* **2014**, *43* (4), 1724−1735.
(30) Li, Z.; Cui, P.; Wang, C.; Kilina, S.; Sun, W. Nonlinear Absorbing Cationic Bipyridyl Iridium(III) Complexes Bearing Cyclometalating Ligands with Different Degrees of π-Conjugation: Synthesis, Photophysics, and Reverse Saturable Absorption. *J. Phys. Chem. C* **2014**, *118* (49), 28764−28775.
(31) Liu, R.; Dandu, N.; Chen, J.; Li, Y.; Li, Z.; Liu, S.; Wang, C.; Kilina, S.; Kohler, B.; Sun, W. Influence of Different Diimine (N^N) Ligands on the Photophysics and Reverse Saturable Absorption of Heteroleptic Cationic Iridium(III) Complexes Bearing Cyclometalating 2-{3-[7-(Benzothiazol-2-yl)fluoren-2-yl]phenyl}pyridine (C^N) Ligands. *J. Phys. Chem. C* **2014**, *118* (40), 23233−23246.
(32) Li, Y.; Dandu, N.; Liu, R.; Li, Z.; Kilina, S.; Sun, W. Effects of Extended π-Conjugation in Phenanthroline (N^N) and Phenylpyridine (C^N) Ligands on the Photophysics and Reverse Saturable Absorption of Cationic Heteroleptic Iridium(III) Complexes. *J. Phys. Chem. C* **2014**, *118* (12), 6372−6384.
(33) Liu, B.; Monro, S.; Lystrom, L.; Cameron, C. G.; Colón, K.; Yin, H.; Kilina, S.; McFarland, S. A.; Sun, W. Photophysical and Photobiological Properties of Dinuclear Iridium(III) Bis-tridentate Complexes. *Inorganic chemistry* **2018**, *57* (16), 9859−9872.
(34) Sun, W.; Pei, C.; Lu, T.; Cui, P.; Li, Z.; McCleese, C.; Fang, Y.; Kilina, S.; Song, Y.; Burda, C. Reverse saturable absorbing cationic iridium(III) complexes bearing the 2-(2-quinolinyl)quinoxaline ligand: effects of different cyclometalating ligands on linear and nonlinear absorption. *J. Mater. Chem. C* **2016**, *4* (22), 5059−5072.
(35) Li, Y.; Dandu, N.; Liu, R.; Hu, L.; Kilina, S.; Sun, W. Nonlinear Absorbing Cationic Iridium(III) Complexes Bearing Benzothiazolylfluorene Motif on the Bipyridine (N^N) Ligand: Synthesis, Photophysics and Reverse Saturable Absorption. *ACS Appl. Mater. Interfaces* **2013**, *5* (14), 6556−6570.
(36) Jabeen, F.; Chen, M.; Rasulev, B.; Ossowski, M.; Boudjouk, P. Refractive indices of diverse data set of polymers: A computational QSPR based study. *Comput. Mater. Sci.* **2017**, *137*, 215−224.
(37) Karuth, A.; Alesadi, A.; Vashisth, A.; Xia, W.; Rasulev, B. Reactive Molecular Dynamics Study of Hygrothermal Degradation of Crosslinked Epoxy Polymers. *ACS Applied Polymer Materials* **2022**, *4* (6), 4411−4423.
(38) Varghese, P. J. G.; David, D. A.; Karuth, A.; Manamkeri Jafferali, J. F.; P. M, S. B.; George, J. J.; Rasulev, B.; Raghavan, P. Experimental and Simulation Studies on Nonwoven Polypropylene-Nitrile Rubber Blend: Recycling of Medical Face Masks to an Engineering Product. *ACS Omega* **2022**, *7* (6), 4791−4803.
(39) Toropov, A. A.; Toropova, A. P.; Benfenati, E. QSAR-modeling of toxicity of organometallic compounds by means of the balance of correlations for InChI-based optimal descriptors. *Molecular Diversity* **2010**, *14* (1), 183−192.
(40) Dems, M. A. E.; Laib, S.; Latelli, N.; Ouddai, N. A DFT-based Quantitative structure activity relationship Study of organometallic estradiol derivatives. *J. Chem. Pharm. Sci.* **2017**, *10*, n/a.
(41) Sifain, A. E.; Lystrom, L.; Messerly, R. A.; Smith, J. S.; Nebgen, B.; Barros, K.; Tretiak, S.; Lubbers, N.; Gifford, B. J. Predicting phosphorescence energies and inferring wavefunction localization with machine learning. *Chemical Science* **2021**, *12* (30), 10207−10217.
(42) DiLuzio, S.; Mdluli, V.; Connell, T. U.; Lewis, J.; VanBenschoten, V.; Bernhard, S. High-Throughput Screening and Automated Data-Driven Analysis of the Triplet Photophysical Properties of Structurally Diverse, Heteroleptic Iridium(III) Complexes. *J. Am. Chem. Soc.* **2021**, *143* (2), 1179−1194.
(43) Terrones, G. G.; Duan, C.; Nandy, A.; Kulik, H. J. Low-cost machine learning prediction of excited state properties of iridium-centered phosphors. *Chemical Science* **2023**, *14* (6), 1419−1433.
(44) Chen, M.; Jabeen, F.; Rasulev, B.; Ossowski, M.; Boudjouk, P. A computational structure−property relationship study of glass transition temperatures for a diverse set of polymers. *J. Polym. Sci., Part B: Polym. Phys.* **2018**, *56* (11), 877−885.
(45) Ahmed, L.; Rasulev, B.; Turabekova, M.; Leszczynska, D.; Leszczynski, J. Receptor- and ligand-based study of fullerene analogues: comprehensive computational approach including quantum-chemical, QSAR and molecular docking simulations. *Organic & Biomolecular Chemistry* **2013**, *11* (35), 5798.
(46) Christenholz, C. L.; Obenchain, D. A.; Peebles, R. A.; Peebles, S. A. Rotational Spectroscopic Studies of C−H···F Interactions in the Vinyl Fluoride···Difluoromethane Complex. *J. Phys. Chem. A* **2014**, *118* (9), 1610−1616.
(47) Ernzerhof, M.; Scuseria, G. E. Assessment of the Perdew−Burke−Ernzerhof exchange-correlation functional. *J. Chem. Phys.* **1999**, *110* (11), 5029−5036.
(48) Perdew, J. P.; Yang, W.; Burke, K.; Yang, Z.; Gross, E. K. U.; Scheffler, M.; Scuseria, G. E.; Henderson, T. M.; Zhang, I. Y.; Ruzsinszky, A.; et al. Understanding band gaps of solids in generalized Kohn-Sham theory. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114* (11), 2801−2806.
(49) Wadt, W. R.; Hay, P. J. *Ab initio* effective core potentials for molecular calculations. Potentials for main group elements Na to Bi. *J. Chem. Phys.* **1985**, *82* (1), 284−298.
(50) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **1982**, *77* (7), 3654−3665.
(51) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72* (1), 650−654.

https://doi.org/10.1021/acs.jpclett.3c02533
J. Phys. Chem. Lett. 2024, 15, 471−480

(52) Minenkov, Y.; Singstad, Å.; Occhipinti, G.; Jensen, V. R. The accuracy of DFT-optimized geometries of functional transition metal compounds: a validation study of catalysts for olefin metathesis and other reactions in the homogeneous phase. *Dalton Transactions* **2012**, *41* (18), 5526−5541.
(53) Tekarli, S. M.; Drummond, M. L.; Williams, T. G.; Cundari, T. R.; Wilson, A. K. Performance of Density Functional Theory for 3d Transition Metal-Containing Complexes: Utilization of the Correlation Consistent Basis Sets. *J. Phys. Chem. A* **2009**, *113* (30), 8607−8614.
(54) Zhu, X.; Liu, B.; Cui, P.; Kilina, S.; Sun, W. Multinuclear 2-(Quinolin-2-yl)quinoxaline-Coordinated Iridium(III) Complexes Tethered by Carbazole Derivatives: Synthesis and Photophysics. *Inorg. Chem.* **2020**, *59* (23), 17096−17108.
(55) Peverati, R.; Truhlar, D. G. Quest for a universal density functional: the accuracy of density functionals across a broad spectrum of databases in chemistry and physics. *Philos. Trans A Math Phys. Eng. Sci.* **2014**, *372* (2011), 20120476.
(56) Qi, S.-C.; Hayashi, J.-i.; Zhang, L. Recent application of calculations of metal complexes based on density functional theory. *RSC Adv.* **2016**, *6* (81), 77375−77395.
(57) Yang, Y.; Weaver, M. N.; Merz, K. M., Jr. Assessment of the "6-31+G** + LANL2DZ" Mixed Basis Set Coupled with Density Functional Theory Methods and the Effective Core Potential: Prediction of Heats of Formation and Ionization Potentials for First-Row-Transition-Metal Complexes. *J. Phys. Chem. A* **2009**, *113* (36), 9843−9851.
(58) Koseki, S.; Gordon, M. S.; Schmidt, M. W.; Matsunaga, N. Main Group Effective Nuclear Charges for Spin-Orbit Calculations. *J. Phys. Chem.* **1995**, *99* (34), 12764−12772.
(59) Barone, V.; Cossi, M.; Tomasi, J. Geometry optimization of molecular structures in solution by the polarizable continuum model. *J. Comput. Chem.* **1998**, *19* (4), 404−417.
(60) *ChemSketch*; Advanced Chemistry Development, Inc.: Toronto, ON, 2015.
(61) Todeschini, R.; Consonni, V.; Mauri, A.; Pavan, M. *Dragon Software for the Calculation of Molecular Descriptors, Version 6 for Windows*; Talete SRL: Milan, 2014.
(62) Todeschini, R.; Consonni, V. *Molecular Descriptors for Chemoinformatics*; Wiley, 2009.
(63) Devillers, J. Genetic Algorithms in Computer-Aided Molecular Design. In *Genetic Algorithms in Molecular Modeling*; Elsevier, 1996; pp 1−34.
(64) Gramatica, P.; Cassani, S.; Chirico, N. QSARINS-chem: Insubria datasets and new QSAR/QSPR models for environmental pollutants in QSARINS. *J. Comput. Chem.* **2014**, *35* (13), 1036−1044.
(65) Karuth, A.; Alesadi, A.; Xia, W.; Rasulev, B. Predicting glass transition of amorphous polymers by application of cheminformatics and molecular dynamics simulations. *Polymer* **2021**, *218*, 123495.
(66) Gramatica, P.; Giani, E.; Papa, E. Statistical external validation and consensus modeling: A QSPR case study for Koc prediction. *Journal of Molecular Graphics and Modelling* **2007**, *25* (6), 755−766.

https://doi.org/10.1021/acs.jpclett.3c02533
J. Phys. Chem. Lett. 2024, 15, 471−480