Machine Learning Driven Optimization of Fe-Based TMCs for Photodynamic Therapy

Vladimir Manuilov[1] , **Antonio Francés-Monerris**[2] , **Abdelazim M. A. Abdelgawwad**[2] , **Daniel Escudero**[3] and **Ilya Makarov**[4,5,6]
[1]MIPT
[2]Universitat de València
[3]KU Leuven
[4]AIRI
[5]Research Center for AI, Innopolis
[6]ISP RAS

Abstract

Noble metal-based photoactive complexes have applications in photodynamic therapy (PDT), but their toxicity and high cost drive interest in sustainable and cheaper alternatives like iron-based compounds. In this paper, quantum chemistry and classical molecular dynamics were employed to characterize the photophysical properties and non-covalent interactions with DNA of two Fe(III) complexes. We explained the absorption of IR wavelength by bright ligand-to-metal transitions and showed that the complexes exhibit persistent, albeit modest, interaction with DNA. Building on these traditional simulation methods, we propose a conceptual ML-driven optimization module designed to refine the structure of iron complexes and enhance their photophysical features. While the framework is not yet implemented, we demonstrate that key properties relevant for PDT can be computationally evaluated, providing a foundation for future iterative optimization. The ML module integrates 3D molecular structures, simulation results, and quantum chemical insights to suggest modifications aimed at shifting the absorption spectrum more favorably into the visible range, improving their suitability for phototherapies.

1 Introduction

Solar energy is a promising alternative to fossil fuels due to its abundance and renewable nature [Kashyap and Ghasemi, 2020], and with the help of photovoltaics it can be converted to electricity [Nazeeruddin *et al.*, 2001] and chemical energy. Transition metal complexes (TMSc), especially those based on nobel metals, are widely used for solar cells, artificial photosynthesis, bioimaging, and photodynamic therapy due to their remarkable photophysical properties [Balzani *et al.*, 2007]. These precious metals, however, are very rare [Xia and Yang, 2017], which stimulates the search for alternative solutions, such as iron. Despite its abundance and lower toxicity, iron is inferior in photophysical properties, because iron-based complexes have lower ligand field and unstable excited states [Wenger, 2019]. Nonetheless, recent progress in prolonging excited state lifetimes in iron-based TMCs opens avenues for efficient photoactive complexes [Kjær *et al.*, 2019].

The photophysics of transition metal complexes plays a key role in their applications. Metal-organic complexes containing π -conjugated ligands have 4 main electronic excited states: ligand-centered (LC), metal-centered (MC), charge transfers from ligand to metal (LMCT), and from metal to ligand (MLCT) [Francés-Monerris *et al.*, 2019]. A strong ligand field in Ru complexes results in the destabilization of triplet and quintet MC states, which results in long-lived MLCT states. In the case of iron-containing complexes, the situation is different: due to the weak ligand field, the MC states are lower in energy, which leads to rapid (subpicosecond) deactivation of MLCT and the impossibility of effective use of such compounds in photoactive processes. Recent advances in ligand design, such as the use of strong σ -donor and π -acceptor properties, have extended MLCT lifetimes, making Fe-based complexes viable for photoactive applications. On the other hand, highly symmetric coordination spheres (close to the ideal octahedral disposition) achieved with proper chelating ligands could also destabilize excited-state deactivation through MC states. We rely on the latter strategy and consider two heteroleptic iron complexes (**C1** and **C2**) (Figure 1) synthetized and characterized by Carrillo and colleagues [Arcos, 2023].

One of the applications of photoactive molecules is PDT, a targeted cancer treatment, which is based on the combined action of a drug and molecular oxygen [Monro *et al.*, 2018]. Upon light irradiation, the excited state can be deactivated through electron or energy transfer to molecular oxygen, which results in reactive oxygen species (ROS). The highly active molecules are reactive towards DNA and cause cell apoptosis. TMCs are very suitable photosensitizers because of their stability, biocompatibility, high absorption within the relative rate of tissue transparency (700-1100 nm) and high triplet state quantum yield [Monro *et al.*, 2018].

Finding the optimal Fe-based photosensitizer for PDT [Bouché *et al.*, 2020] can be challenging due to the multiple objectives of the optimization problem and the

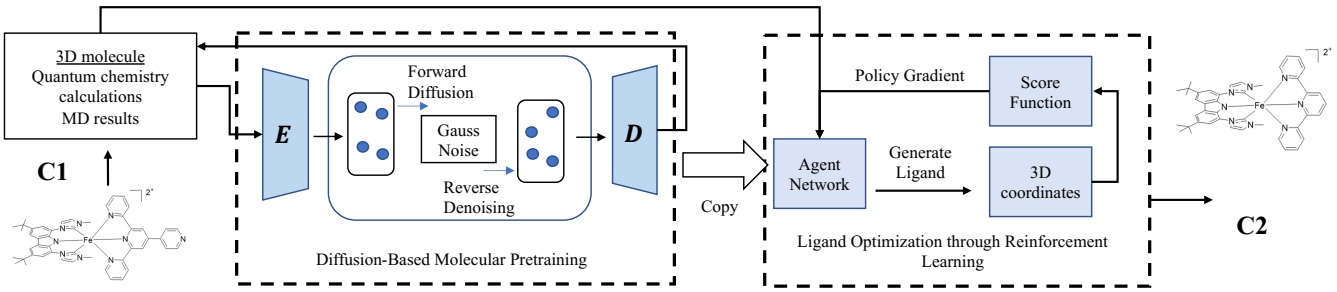

Figure 1: The proposed framework to optimize Fe-based TMCs

need for experimental validation. We propose a conceptual ML-driven optimization method to iteratively adjust the molecule. The goal is to achieve a shift in the absorption spectrum toward visible range, thereby enhancing performance in photodynamic therapy, while maintaining key attributes, such as DNA affinity.

2 Proposed Framework

In this section, we present our conceptual ML-driven framework for the iterative optimization of iron-based TMCs, aimed at shifting the absorption spectrum toward the visible range while preserving DNA affinity. While the diffusion-model-based agent has not yet been implemented, we outline a framework that leverages computationally accessible properties. The framework comprises two core modules: an agent module based on a conditioned diffusion model, and a scoring module, inspired by REINVENT-Transformer [Xu *et al.*, 2023]. The process begins with an initial Fe-based complex (e.g., **C1**), whose absorption wavelength and DNA binding properties are evaluated using established methods (TDDFT, MD). At the heart of the framework is an agent built on a pre-trained diffusion model, adapted from LigandDiff [Jin and Merz Jr, 2024]. The candidate structures are evaluated using TDDFT and MD, and a score is computed as the difference between the desired and the calculated properties. This score guides the agent in generating further modifications. To illustrate the framework, we present results from two iron-based complexes, **C1** and **C2**. **C2** is featuring a refined ligand and exhibits a shifted absorption wavelength consistent with our optimization targets, while retaining stable DNA affinity, serving as an example of a model output.

3 Results

3.1 Quantum Chemistry

A starting complex is iteratively optimized to shift the absorption wavelength while keeping key PDT features stable. We briefly describe results for complexes **C1** and **C2**.

One of the challenges computational chemists face when studying Fe-based TMCs is the correct determination of the ground state (GS) [Lawson Daku *et al.*, 2005]. For these 100-atom complexes, DFT is suitable, although functionals inconsistently predict the GS. To alleviate these deficiencies, Reiher and coworkers introduced a reparameterization of the B3LYP functional specifically for Fe complexes – B3LYP* functional with reduced (from 20 % to 15%) admixture of the Hartree-Fock exchange [Reiher, 2002; Reiher *et al.*, 2001]. B3LYP and B3LYP* functionals in combination with different basis sets reveal S_0 and D_1 as the GS for both **C1** and **C2** systems, respectively. The GAUSSIAN 16 [Frisch *et al.*, 2016] software package has been used in all quantum mechanical calculations.

Functional	VAE (eV)	λ (nm)	f	$\langle \hat{S}^2 \rangle$
BLYP	1.6848	735.89	0.1808	0.891
HCTH	1.6784	738.70	0.1194	1.509
M06	1.6899	733.66	0.1480	0.934
M06-2X	2.7039	458.53	0.0911	1.483
PBE0	1.8645	664.96	0.1341	1.145
B3LYP	1.7062	726.68	0.1515	1.031
B3LYP*	1.6634	745.37	0.1484	1.140
CAM-B3LYP	2.152	576.22	0.1294	0.919

Table 1: Benchmark of performance of TDDFT functionals for the D_5 state of **C2**. The 6-31G* basis set has been used throughout.

We benchmarked various exchange-correlation functionals and basis sets for **C1**'s key spectroscopic characteristics against B3LYP*. We have studied the following functionals in the benchmark: BLYP, HCTH, M06, M06-2X, PBE0, B3LYP, and CAM-B3LYP, and compiled vertical absorption energies (VAEs), oscillator strengths (f) and the expectation value of the total spin operator. In the case of the **C2** complex, the brightest state D_5 lies in the infra-red region, f = 0.1484 and λ = 745 nm (Table 1). The expectation value for the total spin operator is 1.14, which is comparable with an ideal value of 0.75 for a doublet. The results for the **C1** complex are similar: the brightest state is again D_5, f =0.1614 and λ = 757 nm. The total spin operator is 1.325, indicative of a slightly larger mixing with higher spin states as compared to the bright state of **C2**.

To understand the nature of the involved excited states, their density matrices obtained with the TD-B3LYP*/6-31G* method were analyzed utilizing the TheoDORE package [Plasser, 2020]. Results for **C2** are presented in Figure 2. The excitation energies and oscillator strengths are given as colored shading. The representation shows that the brightest state is D_5, which has a strongly pronounced LMCT character

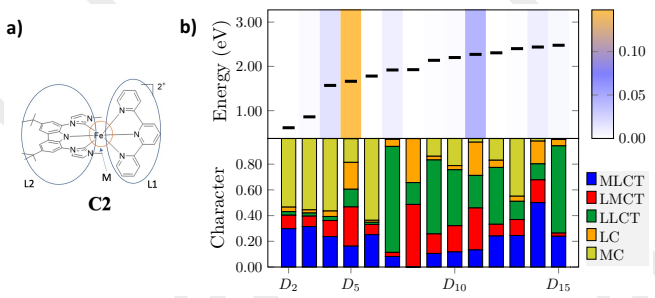

Figure 2: Fragment-based analysis of excited-states in **C2** using TheoDORE reveals key LMCT contributions in C2. Molecular structure (a) of the C2 molecule and its decomposition into three fragments; (b) excitation energies and oscillator strengths and excited-state character, automatically assigned using the CT numbers

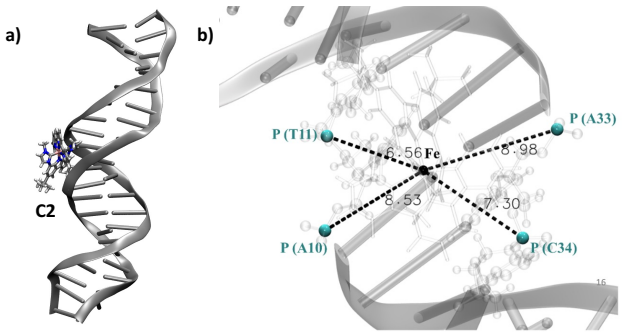

Figure 3: Distances between Fe atom of **C2** complex and P remain stable during the simulation.

and mixing with LC and LLCT, as shown by the CT numbers. The impact of the pyridine substitution from **C2** to **C1** on the absorption properties is small.

3.2 Molecular Dynamics

MD simulations rely on a force field to compute forces and equations of motion to describe particle dynamics. The force field must be parametrized, typically against experimental or high-level computational data. We assigned standard AMBER atom types to all ligands and unique atom types to each Fe-ligating atom, and generated the corresponding parameters with VFFDT [Zheng *et al.*, 2016]. All-atom classical MD simulations of complex **C2** with a DNA oligonucleotide were conducted to investigate non-covalent interactions by using Amber software [Case *et al.*, 2023]. We visually inspected equilibrium dynamics and used Umbrella Sampling to determine the free binding energy. A 0.5 microsecond equilibruim dynamics simulation showed that **C2** interacts with the DNA minor groove for 451.2 ns and remains unbound for 48.8 ns. Figure 3a shows the tert-butyl ligand in the minor groove, with a methyl group oriented toward the DNA. Figure 3b displays the Fe–P distances. Electrostatic forces dominate, as **C2**'s di-cation interacts with the negatively charged phosphate backbone. Small Fe–P fluctuations indicate a stable interaction.

Binding free energy (ΔG) was calculated by comparing

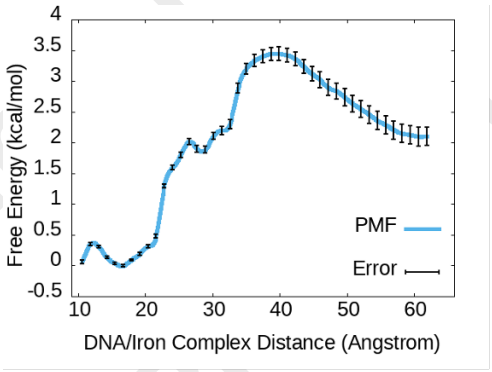

Figure 4: The reconstructed PMF for the DNA and **C2** demonstrates modest affinity

the free energies of the bound state (15–20 Å, where free energy is minimal) and the unbound state (50 Å, with no contact). The frequency diagram for 42 windows showed excellent sample overlap, validating the simulation (bias: $1\,\text{kcal}\cdot\text{mol}^{-1}\cdot\text{Å}^{-2}$). Figure 4's PMF indicates a favorable interaction with an estimated (ΔG) of 2 kcal/mol, corresponding to 96 % binding at 310.15 K

The unbound-to-bound transition has a $\sim$ 1.5 kcal/mol barrier (peaking at $\sim$ 40 Å), while the reverse path has a $\sim$ 3.5 kcal/mol barrier; both are accessible at 310.15 K.

Despite modest binding affinity, **C2** remains close to DNA, suggesting that efficiently generated short-lived ROS could reach DNA for PDT.

4 Conclusion

In summary, this study presents an conceptual computational framework that combines quantum chemical calculations, molecular dynamics simulations, and a machine learning-driven optimization module to design and refine Fe-based transition metal complexes for PDT. Our analysis reveals that the near-IR absorption is primarily governed by the LMCT states. Furthermore, the molecular dynamics simulations indicate that while the complexes exhibit modest binding affinities to the DNA, they remain in close proximity, a key feature for effective ROS mediated damage in PDT.

The proposed ML-driven framework, though not yet implemented, demonstrates a feasible approach to iteratively tailor ligand structures by leveraging rich simulation data. Our work establishes a foundation for future optimization, showing that key PDT-relevant properties can be computationally evaluated and used to guide molecular design.

Overall, our work not only advances the fundamental understanding of the photophysical and binding properties of Fe-based complexes but also lays the foundation for future studies aimed at optimizing these compounds for practical PDT applications.

Acknowledgements

The work of Ilya Makarov on Section 1 was supported by the Ministry of Economic Development of the RF (code 25-139-66879-1-0003).

References

[Arcos, 2023] Ulises Alonso Carrillo Arcos. *Iron-based photoactive complexes: control of geometry and electronic properties*. PhD thesis, Université de Lorraine, 2023.

[Balzani *et al.*, 2007] Vincenzo Balzani, Giacomo Bergamini, Sebastiano Campagna, and Fausto Puntoriero. *Photochemistry and photophysics of coordination compounds: overview and general concepts*. Springer, 2007.

[Bouché *et al.*, 2020] Mathilde Bouché, Cécilia Hognon, Stéphanie Grandemange, Antonio Monari, and Philippe C Gros. Recent advances in iron-complexes as drug candidates for cancer therapy: Reactivity, mechanism of action and metabolites. *Dalton Transactions*, 49(33):11451–11466, 2020.

[Case *et al.*, 2023] David A Case, Hasan Metin Aktulga, Kellon Belfon, David S Cerutti, G Andrés Cisneros, Vinícius Wilian D Cruzeiro, Negin Forouzesh, Timothy J Giese, Andreas W Götz, Holger Gohlke, et al. Ambertools. *Journal of chemical information and modeling*, 63(20):6183–6191, 2023.

[Francés-Monerris *et al.*, 2019] Antonio Francés-Monerris, Philippe C Gros, Xavier Assfeld, Antonio Monari, and Mariachiara Pastore. Toward luminescent iron complexes: unravelling the photophysics by computing potential energy surfaces. *ChemPhotoChem*, 3(9):666–683, 2019.

[Frisch *et al.*, 2016] MJ Frisch, GW Trucks, HB Schlegel, GE Scuseria, MA Robb, JR Cheeseman, G Scalmani, V Barone, GA Petersson, H Nakatsuji, et al. Gaussian16; gaussian, inc. *Wallingford, CT*, 2016.

[Jin and Merz Jr, 2024] Hongni Jin and Kenneth M Merz Jr. Liganddiff: de novo ligand design for 3d transition metal complexes with diffusion models. *Journal of Chemical Theory and Computation*, 2024.

[Kashyap and Ghasemi, 2020] Varun Kashyap and Hadi Ghasemi. Solar heat localization: concept and emerging applications. *Journal of materials chemistry A*, 8(15):7035–7065, 2020.

[Kjær *et al.*, 2019] Kasper Skov Kjær, Nidhi Kaul, Om Prakash, Pavel Chábera, Nils W Rosemann, Alireza Honarfar, Olga Gordivska, Lisa A Fredin, Karl-Erik Bergquist, Lennart Häggström, et al. Luminescence and reactivity of a charge-transfer excited iron complex with nanosecond lifetime. *Science*, 363(6424):249–253, 2019.

[Lawson Daku *et al.*, 2005] Latévi Max Lawson Daku, Alfredo Vargas, Andreas Hauser, Antony Fouqueau, and Mark Earl Casida. Assessment of density functionals for the high-spin/low-spin energy difference in the low-spin iron (ii) tris (2, 2'-bipyridine) complex. *ChemPhysChem*, 6(7):1393–1410, 2005.

[Monro *et al.*, 2018] Susan Monro, Katsuya L Colon, Huimin Yin, John Roque III, Prathyusha Konda, Shashi Gujar, Randolph P Thummel, Lothar Lilge, Colin G Cameron, and Sherri A McFarland. Transition metal complexes and photodynamic therapy from a tumor-centered approach: challenges, opportunities, and highlights from the development of tld1433. *Chemical reviews*, 119(2):797–828, 2018.

[Nazeeruddin *et al.*, 2001] Mohammad K Nazeeruddin, Peter Pechy, Thierry Renouard, Shaik M Zakeeruddin, Robin Humphry-Baker, Pascal Comte, Paul Liska, Le Cevey, Emiliana Costa, Valery Shklover, et al. Engineering of efficient panchromatic sensitizers for nanocrystalline tio2-based solar cells. *Journal of the American Chemical Society*, 123(8):1613–1624, 2001.

[Plasser, 2020] Felix Plasser. Theodore: A toolbox for a detailed and automated analysis of electronic excited state computations. *The Journal of chemical physics*, 152(8), 2020.

[Reiher *et al.*, 2001] Markus Reiher, Oliver Salomon, and Bernd Artur Hess. Reparameterization of hybrid functionals based on energy differences of states of different multiplicity. *Theoretical Chemistry Accounts*, 107:48–55, 2001.

[Reiher, 2002] Markus Reiher. Theoretical study of the fe (phen) 2 (ncs) 2 spin-crossover complex with reparametrized density functionals. *Inorganic chemistry*, 41(25):6928–6935, 2002.

[Wenger, 2019] Oliver S Wenger. Is iron the new ruthenium? *Chemistry–A European Journal*, 25(24):6043–6052, 2019.

[Xia and Yang, 2017] Younan Xia and Xuan Yang. Toward cost-effective and sustainable use of precious metals in heterogeneous catalysts. *Accounts of Chemical Research*, 50(3):450–454, 2017.

[Xu *et al.*, 2023] Pengcheng Xu, Tao Feng, Tianfan Fu, Siddhartha Laghuvarapu, and Jimeng Sun. Molecular de novo design through transformer-based reinforcement learning. *arXiv preprint arXiv:2310.05365*, 2023.

[Zheng *et al.*, 2016] Suqing Zheng, Qing Tang, Jian He, Shiyu Du, Shaofang Xu, Chaojie Wang, Yong Xu, and Fu Lin. Vffdt: a new software for preparing amber force field parameters for metal-containing molecular systems. *Journal of chemical information and modeling*, 56(4):811–818, 2016.