UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

COMPUTATIONAL ANALYSIS OF SUBSTITUENT EFFECT ON EMISSION WAVELENGTHS AND CHEMICAL REACTION RATES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

VARDHAN SATALKAR
Norman, Oklahoma
2021

COMPUTATIONAL ANALYSIS OF SUBSTITUENT EFFECT ON EMISSION WAVELENGTHS AND CHEMICAL REACTION RATES

A DISSERTATION APPROVED FOR THE
SCHOOL OF ARTS AND SCIENCES

BY THE COMMITTEE CONSISTING OF

Dr. Yihan Shao, Chair

Dr. Wai-Tak Yip, Co-Chair

Dr. Daniel Glatzhofer

Dr. Bin Wang

Dedication

To the fireflies,

twinkling among leaves,

make the stars wonder!

-Rabindranath Tagore

Acknowledgments

I am deeply indebted to Dr. Yihan Shao for his guidance, encouragement, patience, and many hours of counsel throughout my doctoral research. The research work with Dr. Shao has been both stimulating and rewarding. Moreover, I would like to to deliver other special acknowledgment to my current as well as former graduate committee members for their helpful suggestions and advice during research investigations.

I want to mention my sincere thanks to our Computational Chemistry: Applications, Theory and Softwares (CC-ATS) group members, departmental staff, and fellow graduate students of the Chemistry Department, whose friendliness and helpfulness will never be forgotten. I would also like to acknowledge financial and academic resources received from the Chemistry Department and the University of Oklahoma. Finally, I would like to deliver my gratitude to my family and friends whose support and patience made my graduate work possible.

Table of Contents

List of Tables

List of Figures

Abstract

Substituent groups are widely used to take advantage of their electronic and steric properties. For example, substituents with an electron-donating or electron-withdrawing nature have been utilized to design new materials with effective photophysical properties. Additionally, bulky substituents have been used to perform regioselective transformations. The effect of various substituents can be indirectly measured from observables, such as absorption/emission wavelength and chemical reaction rates. To obtain a deeper understanding of the photophysical and chemical properties of a substituted molecule at the molecular level, a systematic theoretical investigation on both molecular and its substituent fragment would be highly beneficial.

In this work, the effect of an electron-donating and electron-withdrawing substituent group on the electron distribution of several chromophores or reagents are studied using a relatively simple quantum mechanics-based method based on the absolutely localized molecular orbitals (ALMO) for analyzing substituent effect. Our work showed that the substitution effects might be readily decomposed into electrostatics, polarization, and charge-transfer interactions between the substituent group and the substrate molecule.

Specifically, in the case of oxyluciferin analogs, out-of-phase mixing of occupied-occupied orbitals can raise the HOMO energy of the chromophore, whereas in-phase virtual-virtual orbital mixing can lower the LUMO energy. Separately, the chemical reactivity trends observed in the CH activation reactions of β-TMS fluorobenzene were explained by computing the reaction barrier heights and performing the ALMO analysis.

Keywords

Substituent Effect, Luciferin-luciferase Bioluminescence, Bio-imaging, CH Activation reactions, Computational Chemistry

Chapter 1

Introduction

1.1 Substituent Effects

A substituent is a discrete atom or molecular fragment attached to a substrate molecule to modulate its physical and chemical properties. In synthetic chemistry, the incorporation of substituent groups on a substrate molecule is essential to design new library compounds with distinct properties and to modulate chemical reactions in predictable ways (1; 2). Moreover, with the help of chemical substitutions, we can also further understand chemical and biological mechanisms. Therefore, till this date, substituent effects have been widely utilized in chemistry (3). It is common to use substituent groups based on their electronic and steric properties (4). For example, the substituents with electron-donating or electron-withdrawing nature have been utilized to design new materials with unique photophysical properties (3). Additionally, it is typical to use bulky substituents to perform regioselective transformations. Experimentally, the effect of substituents can be indirectly measured from observables, such as absorption/emission wavelengths and reaction rates.

1.1.1 Hammett Equations

Derick reported the first account of the substituent effect on chemical reactivity (5), where he described the polar nature of the substituents using dissociation constant. However, Brönsted and coworkers proposed a more quantitative approach (6), where they correlated the logarithmic function of the rate of reaction with the electrophilicity or nucleophilicity of the substituent. The concept was termed Linear Free energy relationships (LFERs). The above approach is also known as the Brönsted catalysis equation that shows the relation of the catalyst efficiency with a strength of acid or base (7). Furthermore, Hammett and coworkers studied the ionization of substituted benzoic acid (2; 8). Based on the reaction

conditions and the position of the substituents, they classified the observed substituent effect relative to H atom in terms of empirical descriptors, such as σ_p, σ_m, σ_p^-, and σ_m^-. For example, the σ_p constants were defined as:

$$\sigma_p = \log K_p - \log K_H \tag{1.1}$$

where K_H and K_p are the ionization constant of the benzoic acid and K_p are the rate constant of *para*-substituted benzoic acid, respectively, at room temperature in an aqueous solution (2; 8).

1.1.2 Aromaticity Indices

In addition to Hammett's postulates, several other theories and models were developed to quantify the substituent effects. For example, the inductive effect scale designed by Exner and Böhm (9) was employed to account for the inductive and resonance effects of polar substituent groups.

Krygowski and coworkers focused on the interconnection between substituent effect and aromaticity (7). They developed a para delocalization index (PDI), which showed a strong correlation with substituent constants. Similar to PDI, several other aromaticity indices such as nucleus independent chemical shift (NICS), harmonic oscillator model of aromaticity (HOMA), substituent effect stabilization energy (SESE), and aromatic stabilization energy (ASE) were developed for monosubstituted benzene (7).

1.1.3 Molecular Electrostatic Potential (MESP)

In 1972, Tomasi and coworkers (10) derived an electronic distribution method based on MESP. However, their model was limited to unconjugated and small cyclic molecules. Later, the MESP analysis was extended to larger and conjugated systems (11). Politzer and coworkers derived a quantity called the average local ionization energy $I(r)$ to characterize the locations of the least tightly bound electrons in a molecule (12). The smallest $I(r)$ value

on the surface of a molecule, I_{min}, has been shown to be a useful index to measure or predict the electron-donating and -accepting strength of substituents on the benzene ring.

1.1.4 Absolutely-Localized Molecular Orbital Analysis (ALMO)

Although the Hammett parameters have been widely utilized in chemistry research (4; 13; 14), they do not always capture the effect of substitutents in conjugated chemical systems accurately (14). Such evident shortcomings of these empirical descriptors can be attributed to the lack of a rigorous theoretical foundation.

Therefore, to obtain a deeper understanding of photophysical and chemical properties of substituted molecules at the molecular level, a systematic theoretical investigation on both molecule and its substituent fragment is highly desirable. The fundamental chemical concepts, such as covalent bonding, induction effect, and hyperconjugation, can be very well described using quantum mechanics-based qualitative molecular orbital (MO) theory (15). The key perturbational MO concepts, such as in-phase mixing/out-of-phase orbital mixing, have been used to qualitatively describe small molecular systems, such as H_2 or HF molecule (15).

In order to measure electronic and photophysical properties of a chemical system, excited-state quantum mechanical methods, especially the time-dependent density functional theory (TDDFT), are used routinely (16; 17). These calculations are usually based on canonical MOs that are delocalized over entire substituted systems instead of fragment orbitals from either the substrate or the substituent. As such, it is not easy to pinpoint the effect of substituent from standard molecular orbital theory calculations. In this dissertation, we sought an orbital interaction picture between fragment orbitals (from the substrate and substituent, in the vein of perturbational MO theory). Specifically, a newly developed absolutely localized molecular orbital-based energy decomposition analysis (ALMO-EDA) from our group (18) was used to analyze intramolecular interactions. Within this analysis,

we can turn on/off certain interactions between substituent and substrate fragments of the molecule to assess corresponding energy components.

In Chapters 2 and 3 of this dissertation, we utilized our ALMO-based orbital interaction analysis tool to understand interactions between electron-donating (or electron-withdrawing) substituents with well-known chromophore substrates, such as oxyluciferin, curcumin, and their analogs (19; 20). In Chapter 4, we applied the same analysis tool on the CH activation reaction to explain the hyperconjugation effect due to β-silicon substituent. Overall, our orbital interaction analysis led to important chemical insights that could benefit the development of future new chromophores and help explain the role of β-silicon substituent in hyperconjugation.

1.2 Computational Chemistry Methods

Computational chemistry is the branch of theoretical chemistry, in which quantum mechanics-based numerical methods are implemented on computers to solve fundamental and applied problems in chemistry. For example, quantum chemistry-based software Q-CHEM (21) can be used to calculate the solvation free energy of a drug molecule during the ligand optimization, which can be considered as a routine application of computational chemistry. In this chapter, we will briefly introduce quantum chemistry and its implementation on computers.

In quantum mechanics, a state of a N-electron M-nuclei system is denoted by a wave function $\Psi(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N;\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_M)$, which is a complex valued square-integrable in the Hilbert space. When it is multiplied with the complex conjugate:

$$\begin{aligned}\Psi^*(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N;\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_M)\Psi(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N;\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_M)\\ = |\Psi(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N;\mathbf{R}_1,,\mathbf{R}_2,\cdots,\mathbf{R}_M)|^2 \qquad (1.2)\end{aligned}$$

the probability of finding the system in a volume of element $d\mathbf{r}_1 d\mathbf{r}_1 \cdots d\mathbf{r}_N$ at a particular nuclear geometry, $\{\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_M\}$, is given by:

$$P(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N) d\mathbf{r}_1 d\mathbf{r}_1 \cdots d\mathbf{r}_N = |\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N; \mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_M)|^2 d\mathbf{r}_1 d\mathbf{r}_1 \cdots d\mathbf{r}_N \tag{1.3}$$

The most important equation in quantum chemistry is the non-relativistic time–independent Schrödinger equation:

$$\hat{H}\Psi = E\Psi \tag{1.4}$$

In the above equation, the Hermitian operator $\hat{H}$ is the energy operator. When it acts on the state wave function Ψ, it provides the eigenvalues and orthogonal eigenfunctions of this equation.

Based on their masses, nuclei move slowly relative to electrons. This gives rise to the hypothesis that considers the electrons to be a flux in a non-variable field of constant nuclei coordinates. This was termed as the Born-Oppenheimer (BO) approximation. Based on the above approximation, the Schrödinger equation is to be divided into nuclear and electronic fragments, with the latter subjected to a non-variable nuclear coordinates $\{\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_M\}$,

$$\hat{H}_N \Psi_N(\mathbf{R}) = E_N \Psi_N(\mathbf{R}) \tag{1.5}$$

$$\hat{H}_{el}(\mathbf{R}) \Psi_{el}(\mathbf{r}; \mathbf{R}) = E_{el}(\mathbf{R}) \Psi_{el}(\mathbf{r}; \mathbf{R}) \tag{1.6}$$

where the electronic part of the Hamiltonian (H_{el}) (in a.u.) can be written as follows:

$$\hat{H}_{el}(\mathbf{R}) = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} + \sum_{A>B} \frac{Z_A Z_B}{R_{AB}} \tag{1.7}$$

Thanks to the Born-Oppenheimer approximation, one could ignore the nuclear kinetic energy from the H_{el} and consider the inter nuclei repulsion as a non-variable in Eq. 1.7.

1.2.1 The orbital approximation

The orbital approximation presumes the orbital as a wave function for every electron in a complex molecule. These orbitals are usually considered as linear combinations of atomic orbitals i.e. basis functions, namely:

$$\psi_p(\mathbf{r}) = \sum_{\mu} C_{\mu p} \phi_{\mu}(\mathbf{r}) \tag{1.8}$$

With the help of such a basis set approximation, the construction of $\psi_p(\mathbf{r})$ is possible. The problem of identifying a set of suitable orbitals is tantamount to finding coefficients $c_{\mu i}$ for every occupied MO $\psi_i(\mathbf{r})$. To avoid the high computational cost, the Gaussian-like orbitals $\phi_{\mu}(\mathbf{r})$ are used. After finding solutions to the Dirac equation, the spin of the electron can be obtained as an additional quantum number. This spin part is later multiplied by the spatial orbital part of the Schrödinger equation.

$$\chi(x) = \psi(r)\alpha(\omega) \tag{1.9}$$

According to the Pauli exclusion principle, the antisymmetric wave function of any two electrons can be written in the Slater determinant form.

There are two ways to solve the Schrödinger equation: (1) variational method; and (2) the perturbation method. The $\hat{H}_0$ does not change in the case of the variational methods. The energy of the known wave function should be the upper limit for the guess wave function in the variational principle. In the case of the perturbation method, one has to add the $\hat{H}_0$ and perturbation $\hat{H}^{'}$.

1.2.2 The Hartree-Fock method

The most important approximation method to the Schrödinger equation is the Hartree-Fock (HF) equation. This equation was derived from the variational method. In general,

HF equation is implemented to determine the lowest energy of Slater determinant *via* suitable MOs. In this method, the multiple wave functions are represented as a single Slater determinant. The Fock-operator can be written as:

$$\hat{f}(1) = -\frac{1}{2}\nabla_1^2 - \sum_A^M \frac{Z_A}{r_{iA}} + \sum_j^N \hat{J}_j(1) - \hat{K}_j(1)] \tag{1.10}$$

In the above equation, the kinetic energy term, potential energy term, repulsion term, and exchange-correlation terms are shown. MOs are eigenfunctions to the Fock operator. In simple terms, the HF equation reduces down to n single-electron problems compared to the n-electron problem. With each spin orbital defined through the basis set (*i.e.* atomic orbitals) expansion, these Roothan-Hall equations can be written down as

$$\hat{f}(1)\psi_j = \varepsilon_j \psi_j \tag{1.11}$$

$$\mathbf{FC} = \varepsilon \mathbf{SC} \tag{1.12}$$

where **S** is the overlap matrix between each pair of the spin orbitals, $S_{\mu\nu}$, and **C** contains all basis set coefficients $C_{\mu p}$. These equation will be solved iteratively.

The lack of the electron correlation term is the major drawback of the HF method. Multiple Slater determinants obtained from various HF-based MOs can be utilized to get solutions that recover the electron correlation. However, this task is computationally expensive for complex chemical systems.

1.2.3 Density Functional Theory

Kohn–Sham density functional theory (KS–DFT) (22) has been one of the most popular chemical reaction analysis and calculation methods in the past two decades. The aspect of considering dynamic correlation effects has been one of the major reasons for its dominance over Hartree-Fock, despite other similarities. KS–DFT has been the backbone for the majority of calculations in this dissertation. According to Kohn and Hohenberg, DFT

is built upon two theorems which state that: (1) electron density of a system of interacting particles can be uniquely determined; and (2) the electron density that minimizes the energy of the overall function is the true electron density corresponding to the full solution of the Schrödinger equation (23).

Despite the use of MOs, KS-DFT doesn't fall in the category of traditional wave function theory. In rare cases, MOs (ψ_i) are used to determine the kinetic energy density (24) as follows:

$$\tau_\sigma = \frac{1}{2} \sum_{i}^{\text{Occupied}} |\nabla \psi_{i\sigma}(\mathbf{r})|^2 \tag{1.13}$$

1.2.4 Excited Electronic States Methods

1.2.4.1 Determinantal Representation Methods

An excited electronic state at a given molecular geometry can be generated by means of various processes such as absorption of light and activation due to chemical reaction. For a molecule with an even number of electrons, its ground state wavefuncion as a single Slater determinant can be shown as:

$$|^1\Phi_0\rangle = |\psi_0^2 \psi_2^2 ... \psi_{N/2}^2\rangle \tag{1.14}$$

In the determinantal representation method, an excited electronic state is obtained where at least one electron is not occupying a low–energy orbital at a given molecular geometry. For instance, an excited state is generated by the removal of the one-electron from the highest occupied molecular orbital (HOMO) of the ground state and its placement into a higher energy orbital, ψ_a.

$$\Psi_{N/2}^a = |\psi_0^2 \psi_2^2 ... \psi_{N/2}^1 \overline{\psi}_a^1\rangle \tag{1.15}$$

If an orbital is not doubly occupied, it contains an α spin function unless there is a bar over it, in which case it is multiplied by a β spin function. The Slater determinant in Eq. 1.15 is an equal combination of singlet and triplet states.

A singlet wave function can be written as the combination of determinants having electrons of opposite spin occupying orbitals N/2 and *a* respectively:

$$^1\Psi^a_{N/2} = \frac{1}{\sqrt{2}}|\psi^2_0\psi^2_2...\psi^1_{N/2}\overline{\psi}^1_a\rangle + \frac{1}{\sqrt{2}}|\psi^2_0\psi^2_2...\overline{\psi}^1_{N/2}\psi^1_a\rangle \tag{1.16}$$

whereas that of the triplet wave function is:

$$^3\Psi^a_{N/2} = \frac{1}{\sqrt{2}}|\psi^2_0\psi^2_2...\psi^1_{N/2}\overline{\psi}^1_a\rangle - \frac{1}{\sqrt{2}}|\psi^2_0\psi^2_2...\overline{\psi}^1_{N/2}\psi^1_a\rangle \tag{1.17}$$

When such a singlet or triplet excited state is generated, the occupation number of the newly occupied orbital (a) is increased by 1, and the occupation number of HOMO is decreased by 1.

1.2.4.2 Time-Dependent Density Functional Theory

Time-dependent density functional theory (TDDFT) is widely utilized in quantum chemistry as a useful tool for obtaining excited state energies and wavefunction. In this section, we will briefly describe the TDDFT method. Furthermore, we will account for the advantages and challenges involved in making accurate predictions for excited-state properties.

Two Hohenberg-Kohn (HK) theorems lead to the formulation of TDDFT. The first theorem concerns the relation between the exact electron density and exact ground state wave function $\Psi[\rho](\mathbf{r})$.

$$\Psi(\mathbf{r},t) = \Psi[\rho(t)](\mathbf{r})e^{-i\alpha(t)} \tag{1.18}$$

The second theorem establishes that the variational principle, i.e., the electronic energy computed with trial density, is always higher than the total energy obtained with the exact density.

The general time–dependent Kohn-Sham equation is

$$i\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \hat{H}(\mathbf{r},t)\Psi(\mathbf{r},t) \tag{1.19}$$

with the electron density given as

$$\rho(\mathbf{r}_1,t) = \int |\Psi(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3...\mathbf{r}_N,t)|^2 d\mathbf{r}_2...d\mathbf{r}_N \tag{1.20}$$

A time-dependent single-particle wave function with M basis functions can be written as

$$\psi_p(\mathbf{r},t) = \sum_{\mu=1}^{M} C_{\mu p}(t)\phi_\mu(\mathbf{r}) \tag{1.21}$$

Then the time-dependent equation becomes

$$i\frac{\partial}{\partial t}\mathbf{C} = \mathbf{F}^{\mathrm{KS}}\mathbf{C} \tag{1.22}$$

In the above equation, the p^{th} column of the matrix $\mathbf{C}$ contains the time-dependent expansion coefficients of $\psi_p(\mathbf{r},t)$. Moreover, $\mathbf{F}^{\mathrm{KS}}$ is the matrix form of the time-dependent KS operator in the given basis. By multiplying equation 1.22 with $\mathbf{C}^{\dagger}$ then subtracting from the resultant equation of its Hermitian transpose, one obtains

$$\sum_q \left(F_{pq}P_{qr} - P_{pq}F_{qr}\right) = i\frac{\partial}{\partial t}P_{pr} \tag{1.23}$$

where density matrix $\mathrm{P}_{pr} = \left(\mathbf{C}_{\mathrm{o}}\mathbf{C}_{\mathrm{o}}^{\dagger}\right)_{pr} = \sum_i^{N/2} C_{pi}(t)C_{ri}^{*}(t)$ is related to electron density *via* following equation.

$$\rho(r,t) = \sum_{p,q}^{M}\sum_{i}^{N/2} C_{pi}(t)C_{qi}^{*}(t)\psi_p(\mathbf{r})\psi_q^{*}(\mathbf{r}) = \sum_{p,q}^{M} P_{pq}\psi_p(\mathbf{r})\psi_q^{*}(\mathbf{r}) \tag{1.24}$$

From equation 1.23, one can obtain excitation energies and oscillator strengths using two ways. First, we could propagate the time-dependent KS wave function. Second, we could use the linear response of the time-dependent KS equation. In this dissertation, we employed the latter, namely the LR-TDDFT formalism.

The LR-TDDFT working equation is

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^{*} & \mathbf{A}^{*} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \tag{1.25}$$

where elements of the matrices A and B are given as

$$A_{ia,jb} = \delta_{ij}\delta_{ab}\left(\varepsilon_a - \varepsilon_i\right) + \langle ia|jb\rangle - C_{\mathrm{HF}}\langle ij|ab\rangle + \langle ia|f_{xc}|jb\rangle \tag{1.26}$$

$$B_{ia,jb} = \langle ia|bj\rangle - C_{\mathrm{HF}}\langle ib|aj\rangle + \langle ia|f_{xc}|bj\rangle \tag{1.27}$$

In hybrid functionals, HF exchange through -$C_{\mathrm{HF}}\langle ij|ab\rangle$ and -$C_{\mathrm{HF}}\langle ib|aj\rangle$. For example, C_{HF} is equal to 0.2 for the widely used B3LYP functional.

1.2.5 Absolutely-localized Frontier Orbital Analysis

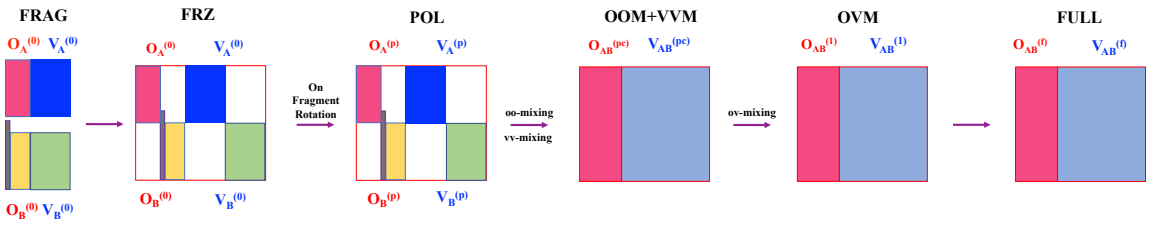


Figure 1.1: ALMO-based scheme for analyzing substituent effects on fluorophore frontier orbitals. For oxyluciferin analogs, for instance, "A" and "B" denote the OLU and substituent fragments, respectively. The vertical bar in light gray corresponds to the link C–X orbital between two fragments (*i.e.* the single bond connecting the OLU core and the substituent.

The ALMO-based energy decomposition analysis procedure (25; 26; 27; 28) was previously extended to handle covalently bonded molecular fragments (18). Most recently, this analysis tool was employed to elucidate the underlying physics of frontier molecular orbital engineering involved in the design of thermally activated delayed fluorescence (TADF) emitters (29). Below we will use oxyluciferin analogs as examples to briefly summarize the analysis procedure (see Figure 1.1) as implemented within a development version of Q-Chem 5.0 (21).

Fragment (FRAG) State: Firstly, we divided a substituted oxyluciferin analog (OLU–Sub) into two fragments — the oxyluciferin core (OLU) and a substituent group (Sub) — to construct the "isolated" fragment orbitals. In order to have both fragments retain a closed-shell electronic structure, the main fragment, in this case, oxyluciferin, was saturated with a hydrogen link atom (Fragment A: OLU-H). Meanwhile, the substituent group was terminated with a phenyl ring (Fragment B: Sub-Ar). Molecular orbitals of both fragments were only allowed to be spanned by atomic orbital basis functions of the corresponding fragment. After performing self-consistent field (SCF) calculations on each fragment, orbital localization was carried out to locate the localized C–Sub bond orbital (connecting the

aryl group and the substituent group) as well as the localized OLU–H bond orbital. These two localized bond orbitals were then tailored and kept frozen during the subsequent SCF calculations on fragments A and B to obtain all other absolutely localized fragment orbitals on OLU and Sub. For further details, we refer the reader to the Supplementary Information of Ref. 18.

Frozen (FRZ) State: In order to account for the effects of permanent electrostatics and exchange-repulsion on frontier orbitals, two sets of fragment orbitals (with the OLU–H bond orbital removed) were combined together to obtain a one-particle density matrix. There is a possibility to further separate the permanent electrostatics and exchange-repulsion terms (26), but this is not pursued in this work because our focus lies with the frontier orbitals.

Polarized (POL) State: In this state, two sets of frozen fragment orbitals are allowed to polarize each other while restricted to stay absolutely localized on their respective fragments. Further orbital relaxations on each fragment thus include only occupied-virtual orbital rotations on each fragment while leaving the C–Sub bond frozen.

OOM +VVM State: Diagonalizing the non-vanishing occupied-occupied block of the Fock matrix within the basis of supersystem polarized orbitals, pseudocanonical occupied orbitals and their energies can be obtained (see Eqs. 9 and 10 in Ref. 18). The occupied-occupied mixing can lead to an increase in the HOMO energy due to an out-of-phase combination of fragment occupied orbitals. In order to rigorously capture the virtual-virtual orbital mixing effect, we diagonalize the corresponding block of the Fock matrix within the basis of projected virtual orbitals to obtain pseudocanonical virtual orbitals and their corresponding energies (see Eqs. 11–13 in Ref. 18). The in-phase combination of fragment frontier virtual orbitals can result in a noticeable lowering of the LUMO energy.

OVM State: After oo- and vv-mixing, a single diagonalization of the Fock matrix is performed to obtain mixing of occupied-virtual orbitals. This step causes a rearrangement of the electron density across OLU and Sub fragments and potentially produces a net charge

transfer. Unlike oo- and vv-mixing, the ov-mixing step can potentially lower the HOMO energy and elevate the LUMO energy.

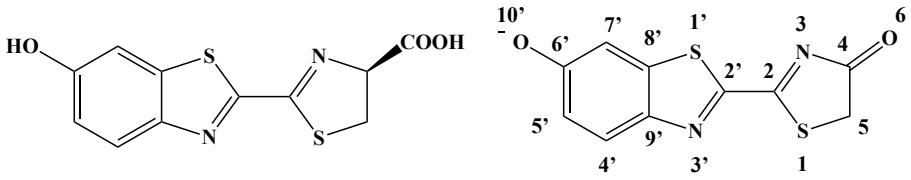
Full State: In the final step after the OVM stage, further orbital rotations will be performed to converge the energy to the final minimum. Fully converged orbitals can then be correlated to polarized occupied and virtual orbitals (see Eqs. 14–15 in Ref. 18).

Chapter 2

Substitution Effect in Oxyluciferin Analogs

2.1 Background

Firefly bioluminescence has fascinated the humankind for thousands of years. But it was not until the late 1950s (30) when the light-emitting process was discovered to arise from the luciferase enzyme and its natural substrate, which is now commonly called firefly luciferin. The substrate was first purified by Bitler and McElory (30) and then identified by White, McCapra, and Field (31; 32) to be D-2-(6-hydroxy-2-benzothiazolyl)-Δ^2-thiazoline-4-carboxylic acid (LUC, see Figure 2.1a). This allowed them to propose a now widely-accepted bioluminescence mechanism: In the belly of firefly beetles, the luciferase enzyme catalyzes a reaction of firefly oxyluciferin with ATP and O_2 molecules to produce an electronically excited oxyluciferin anion (Figure 2.1b), whose relaxation yields the characteristic green-yellow light emission. While there were earlier suggestions that the keto–enol tautomerism of oxyluciferin might play a role in the light emission, the keto form of oxyluciferin (OLU), Figure 2.1b) has been established as the sole light emitter *via* both experimental studies by Branchini and coworkers (33; 34) and theoretical modeling by Liu, Fang, Lindh, and others (35; 36; 37; 38; 39).



a) Firefly D-Luciferin **b) Oxyluciferin**

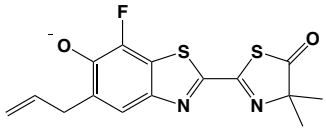
Figure 2.1: The chemical structures of firefly luciferin (LUC) and oxyluciferin (OLU).

With firefly bioluminescence (fBL) assays routinely employed in biochemistry research, there is a constant push to shift the emission color from the common green-yellow to red

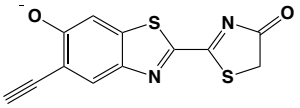
and near-infrared regions. Luminescence with a longer wavelength is less harmful and can penetrate much deeper into human/animal tissues due to the reduced scattering effect. There are two general approaches to module the fBL emission lengths: chemical modifications to the oxyluciferin fluorophore and environmental changes such as amino acid mutations of the luciferase enzyme.

The oxyluciferin emitter can be modified in at least four ways to achieve a substantial redshift in the bioluminescence wavelength. Firstly, the C4', C5', and C7' sites of the benzothiazole ring can be functionalized with one or more substituents, leading to analogs such as 7'-F-5'-allyl-oxyluciferin (40), 5'-alkynyl-oxyluciferin (41), 5'- and 7'-Br-oxyluciferin (40; 42), 7'-Cl-oxyluciferin (40), and 7'-allyl-oxyluciferin (43). Secondly, a conjugated bridge (such as one or multiple double bonds or a conjugate ring) can be inserted between the benzothiazole and thiazole moieties to increase the conjugation length, leading to infra-oxyluciferin (44), PhOH-Luc (45), and other compounds. Thirdly, the benzothiazole ring can be modified, with examples such as D-Quinolyl-oxyluciferin (46). Finally, following the design of 6'-Amino-oxyluciferin (47), many other groups replaced the 6'-hydroxyl group with other linear or cyclic amino groups, leading to AkaLumine (48; 49; 50), CycLuc10 (51), Cyb-oxyluciferin (51), NH_2-NpLH2 (52), and DmaV-oxyluciferin (53).

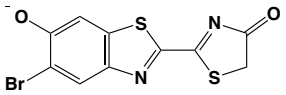
A second approach to tune the bioluminescence emission wavelength is to genetically modify the luciferase enzyme (54; 55; 56; 57). This was inspired by the natural variation of bioluminescence emission colors (58; 59; 60; 61): north American firefly *Photinus pyralis* (λ_{em} = 560 nm) (62), click beetle *Pyrophorus plagiophthalamus* (λ_{em} = 590 nm) (63), and railroad worm (λ_{em} = 630 nm) (64), all with the same oxyluciferin substrate. In the early 2000s, Branchini and coworkers developed red-emitting (614 nm) mutant of *Photinus pyralis* luciferase (56). Later, Kato and coworkers reported luciferase mutant of *Luciola cruciata*, which emits in both yellow-green (560 nm) and red region (613 nm) using oxyluciferin as substrate (54). Prescher and coworkers genetically engineered luciferases to facilitate the bioluminescence emission of the oxyluciferin analogs with C4' and C7'



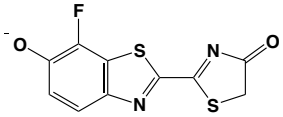
7’-F-5’-allyl-oxyluciferin (617 nm)



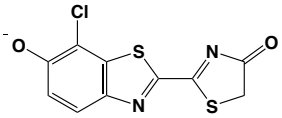
5’-alkynyl-oxyluciferin (610 nm)



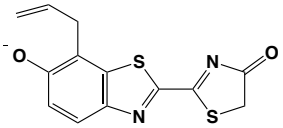
5’-Br-oxyluciferin (620 nm)



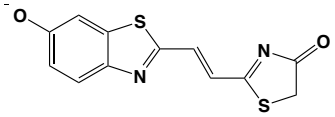
7’-F-oxyluciferin (562 nm)



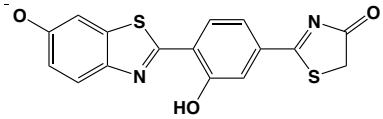
7’-Cl-oxyluciferin (570 nm)



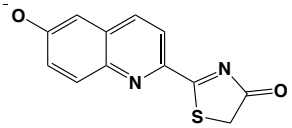
7’-allyl-oxyluciferin (614 nm)



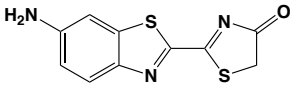
Infra-oxyluciferin (706 nm)



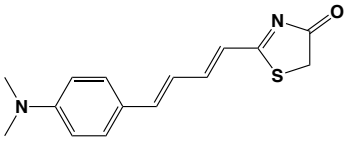
PhOH-oxyluciferin (608 nm)



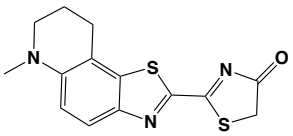
D-Quinolyl-oxyluciferin (608 nm)



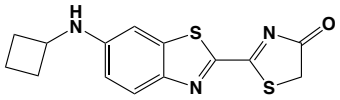
6’-amino-oxyluciferin (605 nm)



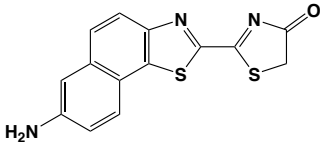
AkaLumine-oxyluciferin (668 nm)



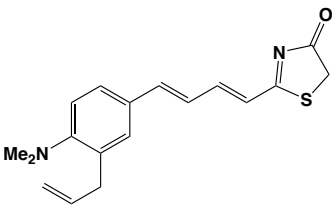
Cyc10-oxyluciferin (642 nm)



Cyb-oxyluciferin (603 nm)



NH_2-Np-oxyluciferin (728 nm)



DmaV-oxyluciferin (685 nm)

Figure 2.2: Selected oxyluciferin analogs and their bioluminescence wavelengths (in nm). For PhOH-oxyluciferin, its bioluminescence wavelength was measured with the G2 mutant of luciferase. Bioluminescence wavelength of NH_2-NpLH_2-oxyluciferin was reported with the CBR2opt luciferase mutant.

substituents (42; 65; 66). More recently, Miyawaki and coworkers performed laboratory-directed evolution to develop mutant Akaluc luciferases that are most compatible with the AkaLumine luciferin substrate shown in Figure 2.2 (50).

In this work, we will focus primarily on the first approach to modulate the emission wavelength, namely *via* chemical modifications to the oxyluciferin fluorophore. Specifically, we will perform computational modeling of the effect of various substituent groups on the fluorophore. We note that several computational studies have been performed to elucidate bioluminescence mechanism (35; 38) as well as to design novel oxyluciferin analogs (67; 68; 69; 70). In 2014, Liu and coworkers reported TDDFT absorption and fluorescence wavelengths of oxyluciferin analogs, where the 6'-hydroxy group of the benzothiazole was replaced with electron-donating substituents (67). In particular, OMe and $N(CH_3)_2$ caused a red-shift in the emission wavelength due to the hyperconjugation effect of the methyl groups, whereas N(CH3)CH=CH2 and $N(C_6H_5)_2$ groups red-shifted the emission thanks to the electron-donating conjugation effect. Later, Liu and coworkers carried out an even more comprehensive computational study of oxyluciferin analogs (68). They reviewed eight classes of analogs and performed an elaborate screening based on their fluorescence wavelengths obtained from gas-phase TDDFT calculations. The most red-shifted oxyluciferin analog, which was called a star, was then further analyzed in conjunction with wild-type luciferase and its mutants using QM/MM calculations. Most recently, Navizet and coworkers reported a good agreement between experimental bioluminescence wavelengths of oxyluciferin analogs and the corresponding computed emission wavelengths based on QM/MM calculations (69). In their study, the authors investigated benzothiphene-Oxy (nitrogen atom replaced by a methine), dihydropyrrolone-Oxy(sulfur atom replaced by methylene), and allylbenzothiazole-Oxy (introduction of an allyl group on C7' site) analogs and probed the effects of these substitutions within the protein environment. Some oxyluciferin analogs were also suggested as effective light-emitting materials in OLED applications. For instance, Min and coworkers replaced the benzothiazolyl ring

of oxyluciferin with phenyl, anthryl, or pyrenyl rings and computed the excited state properties of those analogs (70). Through that, they predicted that oxyluciferin analogs emit with a variety of colors, such as blue, green, orange, and red.

Inspired by all these experimental and computational research on oxyluciferin analogs, we aimed in this chapter to gain a further understanding of the substituent effects on the oxyluciferin molecule. Specifically, we substituted the OLU structure at three different sites, C7', C5' and C4' in Figure 2.1, with 14 functional groups, and performed fragment-based frontier-orbital analysis. These included several electron-donating groups with a varying Hammett σ_p constant (1): $CHCH_2$(-0.04) $>$ OH (-0.37) $>$ NH_2(-0.66) $>$ $NHCH_3$ (-0.70) $>$ $N(CH_3)_2$ (-0.83). Also included were multiple electron-withdrawing groups: NO_2(0.78) $>$ CN(0.66) $>$ CF_3(0.54) $>$ $COCH_3$(0.50) $>$ Br(0.23) $\sim$ Cl(0.23) $\sim$ CCH(0.23) $>$ I(0.18) $>$ F(0.06).

Out of the compounds to be analyzed in this chapter, several were already studied in the past. Urano and coworkers synthesized 7'-fluoro-oxyluciferin and 7'-chloro-oxyluciferin analogs (40). The oxyluciferin analog with an alkynyl substituent at C5' were developed and tested by Prescher and coworkers (41). The same research group later also reported brominated oxyluciferin analogs using C4', C5', and C7' sites (42). In the aforementioned computational study by Liu and coworkers (68), the effects of fluorine and hydroxyl groups at C4' and C7' positions of oxyluciferin structure were analyzed (See Table B.2).

Herein, a systematic computational modeling was performed to explore the effect of electron-donating and electron-withdrawing substituents on the excited state properties of oxyluciferin analogs. In particular, the absolutely localized molecular orbitals (ALMO)-based analysis was carried out on selected mono-substituted oxyluciferin analogs to obtain a physically meaningful picture of frontier orbital evolution (18). With the aid of such a molecular orbital-based picture, it would be possible to engineer frontier orbitals of oxyluciferin analogs to design an effective bio-imaging toolkit.

2.2 Computational Methods and Details

2.2.1 Ground State and Excited State Calculations

All quantum mechanical electronic ground-state calculations were performed using the density functional theory (DFT) method (71; 72; 73), whereas excited-state calculations were carried out using the time-dependent density functional theory (TDDFT) method (74; 75; 76).

Functionals. In the main text of this article, calculation results (absorption, emission, and ALMO-based frontier orbital analysis) would be presented using the ωB97X-D functional (77). Four other density functionals, namely B3LYP (78; 79; 80), PBE0 (81), M06-2X (24), and CAM-B3LYP functionals (82), were also used for TDDFT absorption wavelength calculations, while CAM-B3LYP was also used for emission wavelength calculations.

Basis sets and integration grids. Most of the DFT and TDDFT absorption/emission results were obtained using 6-311++G** basis set (83), and SG-1 numerical integration grid (84), which is a pruned Lebedev grid with 50 radial shells and 194 angular points (85). Additionally, several other basis sets, such as 6-311+G* and aug-cc-pVTZ (86; 87) were used to compute TDDFT absorption wavelengths of oxyluciferin. As shown in Table B.3, relative to 6-311++G** values, the absorption wavelength of oxyluciferin computed using the aug-cc-pVTZ basis set changed by no more than 3 nm. The effect of using a larger integration grid is also marginal, with the computed absorption wavelengths changing by no more than 0.1 nm between SG-1 and a finer (99, 590) grid. In ALMO-based frontier orbital analysis calculations, the 6-31+G* basis set (88; 89) was employed.

Implicit solvent models. In order to account for solvent effect on oxyluciferin anion, we used the polarizable continuum model (PCM) for water molecules, specifically the C-PCM model (90; 91), as implemented in Q-CHEM using a smoothly-discretized solvent cavity with Bondi radii and a scaling factor of 1.2 (92). For absorption energy calculations,

ground-state geometry optimization was performed using PCM, and then vertical excitation energies were computed using TDDFT and the non-equilibrium linear-response PCM model (93; 94; 95). According to the results tabulated in Tables B.4 and B.5, the TDDFT absorption wavelengths obtained by two solvent models differed by 2.3 nm. To optimize excited state energies and obtain emission wavelengths, TDDFT calculations were carried out with equilibrium PCM.

Geometry optimization. The convergence criteria for ground state and excited state geometry optimizations were set to 3×10^{-3} a.u. for maximum force, 1×10^{-6} a.u. for RMS force, 6×10^{-6} a.u. for maximum displacement and 1.2×10^{-3} a.u. for RMS displacement. The vibrational analyses were performed at the ground state, which indicated that the geometries are located at minima. The geometry of OH and $NHCH_3$ substituted OLU rotamers were considered and listed in Figure B.2.

2.3 Results and Discussions

2.3.1 Computed Emission Wavelengths

For oxyluciferin analogs with a single substituent group attached at C7', C5', or C4' site, our computed emission wavelengths and oscillator strengths are summarized in Table 2.1.

First, let us compare these wavelengths against experimental values when available. Earlier, Urano and coworkers reported the fluorescence wavelength of 7'-F-oxyluciferin to be 540.0 nm, which is 10 nm longer than the emission wavelength of 7'-Cl-oxyluciferin (530 nm) in the basic medium (pH 10.0) (40). Our TD-ωB97X-D calculations predicted slightly longer emission wavelengths of both analogs: 7'-F-oxyluciferin (584.2 nm) and 7'-Cl-oxyluciferin (563.1 nm), but with a correct trend in the spectral shift between the two analogs. In a separate work, Prescher and coworkers found the fluorescence wavelength of C7', C5', and C4'-bromine substituted oxyluciferin to be comparable to D-luciferin (42). From our TDDFT calculations, we also observed marginally blue-shifted

Table 2.1: Computed emission wavelengths (λ_{em}, in nm) and oscillator strengths (f) of oxyluciferin and its analogs in aqueous solution at TD-ωB97X-D/6-311++G**/C-PCM level of theory.

Group	site C7'			site C5'			site C4'		
	λ_{em}	$\Delta\lambda_{em}$	f	λ_{em}	$\Delta\lambda_{em}$	f	λ_{em}	$\Delta\lambda_{em}$	f
NO_2	490.5	-88.6	1.232	583.7	4.6	1.009	654.0	74.9	0.635
CN	517.7	-61.4	1.279	564.6	-14.5	1.204	587.4	8.3	1.164
CH_3CO	519.5	-59.6	1.288	589.7	10.6	1.165	598.2	19.1	1.158
CF_3	528.3	-50.8	1.304	560.7	-18.4	1.186	575.0	-4.1	1.218
I	559.0	-20.1	1.269	571.7	-7.4	1.342	581.9	2.8	1.186
CCH	561.0	-18.1	1.236	580.6	1.5	1.337	589.5	10.4	1.181
Br	561.1	-18.0	1.267	569.6	-9.5	1.325	579.0	-0.1	1.191
Cl	563.7	-15.4	1.266	568.0	-11.1	1.303	577.9	-1.2	1.192
H	579.1	0.0	1.270	579.1	0.0	1.270	579.1	0.0	1.270
F	584.2	5.1	1.230	563.2	-15.9	1.297	571.2	-7.9	1.200
$CHCH_2$	592.1	13.0	1.232	597.5	18.4	1.349	599.8	20.7	1.107
OH-2	662.7	83.6	1.089	586.0	6.9	1.388	584.2	5.1	1.164
OH	665.8	86.7	1.007	574.1	-5.0	1.383	579.7	0.6	1.174
NH_2	825.8	246.7	0.793	640.9	61.8	1.402	605.7	26.6	1.031
$N(CH_3)_2$	882.0	302.9	0.719	645.3	66.2	1.428	608.4	29.3	1.048
NCH_3H	888.5	309.4	0.741	651.6	72.5	1.384	596.4	17.3	0.993
$NHCH_3$	925.8	346.7	0.658	689.5	110.4	1.367	614.6	35.5	1.031

bromo-oxyluciferin analogs: compared with the emission wavelength of unsubstituted oxyluciferin, the blue shift is 18.0 nm with bromination at C7', 9.5 nm at C5', and 0.1 nm at the C4' site.

As a separate validation of TDDFT calculations on oxyluciferin analogs, we also computed the fluorescence wavelength of seven C6' oxyluciferin analogs in the PBS buffer and compared it to the experimental results reported by Miller and coworkers (96) (see Table B.1). Figure B.1 showed a good correlation (R^2 = 0.973; y= 0.992x-0.258) between our computed TDDFT emission wavelengths and the experimental values. Together, this confirmed that TDDFT could be used to predict spectral shifts of substituted oxyluciferin analogs reliably.

2.3.1.1 Site C7'

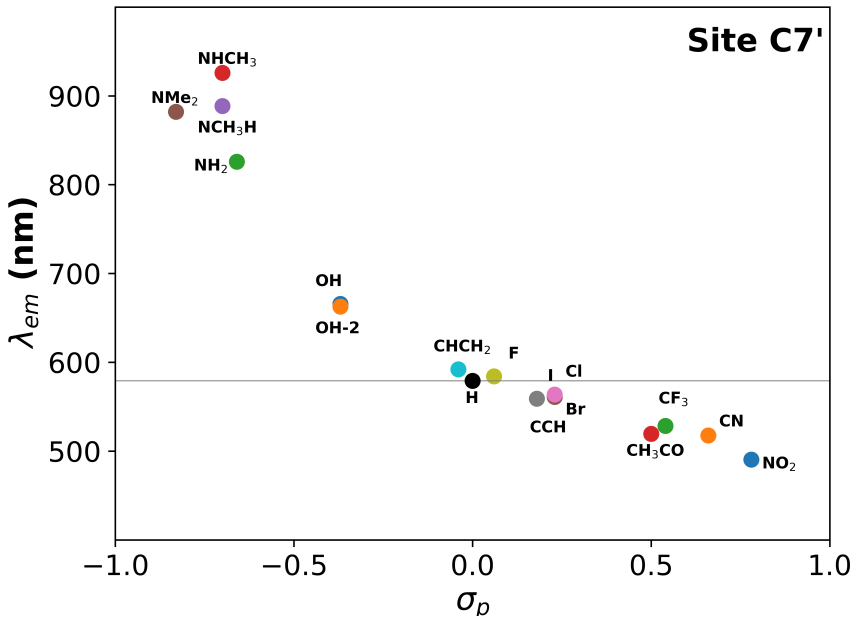


Figure 2.3: Computed emission wavelengths (λ_{em}) of C7'-functionalized oxyluciferin analogs versus the Hammett σ_p constants of the corresponding substituent groups. Emission wavelengths were obtained from TD-ωB97X-D/6-311++G**/C-PCM calculations.

Upon their attachment to C7' site, most substituent groups were predicted in Table 2.1 and Figure 2.3 to cause a substantial shift in the emission wavelength of the chromophore. Moreover, the spectral shift was strongly correlated with the electron-donating or withdrawing capability of each substituent group, as measured by the Hammett σ_p constants.

Among the electron-donating groups, the 7'-NH_2 group was predicted to yield an emission wavelength of 825.8 nm, which amounted to a substantial redshift of 246.7 nm. Even longer wavelengths were obtained for 7'-NCH_3H-oxyluciferin (885.2 nm), its 7'-$NHCH_3$-oxyluciferin conformer (925 nm, with the methyl group pointing away from the 6'-oxygen atom), and 7'-$N(CH_3)_2$-oxyluciferin (882.0 nm). In these molecules, the methyl group(s) affected the electronic structure mainly through hyperconjugation (see Figure B.9), namely the methyl group(s) of $NHCH_3$ and $N(CH_3)_2$ directly contributed *p*-like orbitals to the HOMO of the modified chromophores. Specifically, due to the out-of-phase combinations between these *p*-like orbitals and the HOMO of the OLU core, the HOMO energy of the full system got raised from around -6.1715 to around -6.0137 eV, thus narrowing the HOMO-LUMO gap and causing a redshift. Unlike these nitrogen-based electron-donating groups, a smaller redshift was obtained for the 7'-hydroxyl analog (662.7 nm and 665.8 nm), which has two conformers (OH and OH–2) with the H atom pointing away from the 6'–oxygen atom in the latter conformation, as well as the 7'-vinyl analog (592.1 nm). This trend in the emission wavelength can be traced to the relatively lower HOMO energies of these two oxyluciferin analogs, -6.5797 eV and -6.6096 eV, respectively.

In contrast, the attachment of electron-withdrawing groups (except for fluorine substituent) resulted in blue-shifted emission peaks. The 7'-NO_2-oxyluciferin analog was found to have the shortest emission wavelength of 490.5 nm. The CN and CH_3CO groups at the C7' site of oxyluciferin yielded emission wavelengths of 517.7 nm and 519.5 nm, respectively. A slightly longer emission wavelength (528.5 nm) was predicted for the 7'-CF_3-oxyluciferin analog. The effect of all four halogens was marginal. Amongst, the I-analog (559.0 nm), Br-analog (561.1 nm), and Cl-analog (563.7 nm) displayed a blue-shift of 20.1

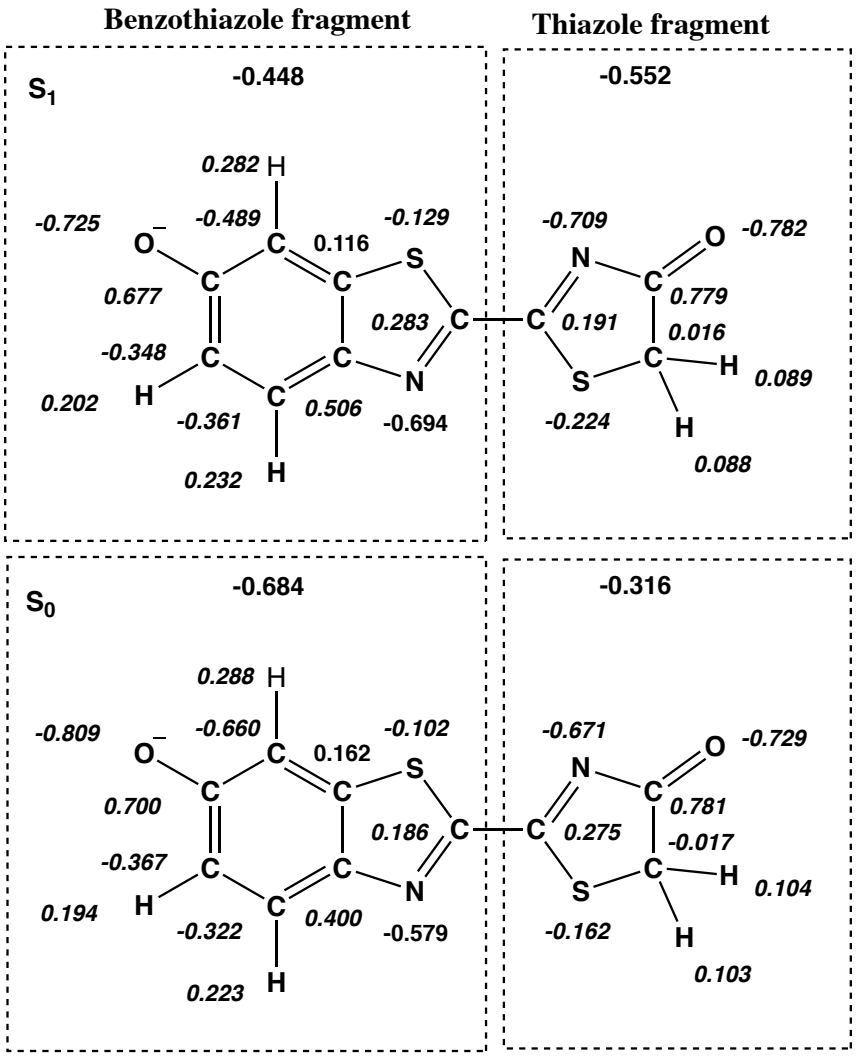


Figure 2.4: ESP charges for (the relaxed density of) S_1 (upper panel) and S_0 (lower panel) states obtained from TD-ωB97X-D/6-311++G**/C-PCM calculations.

nm, 18 nm, and 15.4 nm in the oxyluciferin emission spectra, respectively. On the other hand, the oxyluciferin analog with C7'-fluorine substitution (584.2 nm) was predicted to have an emission peak slightly redshifted by 5.1 nm when compared to oxyluciferin.

Towards an understanding of the emission wavelength shifts caused by these groups, one might consider the charge-transfer (CT) character of the S_1 state of unsubstituted oxyluciferin anion (37; 97). As shown in Figure 2.4, upon the $S_0 \rightarrow S_1$ excitation of the oxyluciferin anion, the net electrostatic potential fitted (ESP) charge transfer (from the benzothiazolyl ring to the thiazolyl ring) was predicted to be 0.236 e^- with TD-ωB97X-D calculations. Intuitively, the addition of an electron-donating/withdrawing group on the benzothiazole ring would likely facilitate/hinder the charge transfer to the thiazole ring during the $S_0 \rightarrow S_1$ transition, thus reducing/increasing the excitation energy.

2.3.1.2 Site C5'

In contrast to the C7' site, most of the electron-donating and withdrawing groups had a much smaller impact on the emission wavelengths. Furthermore, as shown in Figure 2.5, there is no obvious correlation between the computed emission wavelengths and the Hammett σ_p constants of the substituents. Out of all substituent groups, only the amino-based electron-donating groups have a notable effect on the emission wavelength of C5'–analogs. For instance, The 5'-$NHCH_3$-oxyluciferin analog was found to have the longest emission wavelength of 689.5 nm. Its 5'-NCH_3H conformer yielded a shorter emission wavelength of 651.6 nm. Similarly, functionalization with the $N(CH_3)_2$ and NH_2 groups resulted in redshifted oxyluciferin analogs emitting at 645.3 nm and 640.9 nm, respectively. The substitution effect of all other electron-donating substituents at the C5' site of oxyluciferin was marginal. For instance, the 5'-(OH)-oxyluciferin analog (574.1 nm) caused a slight

blue-shift of 5.0 nm compared to oxyluciferin, whereas it is another conformer 5'-(OH-2)-oxyluciferin resulted in a redshift of 6.0 nm. The emission wavelength of the 5'-vinyl-oxyluciferin analog was predicted to be 597.5 nm, which is 19.5 nm longer than that of the OLU anion.

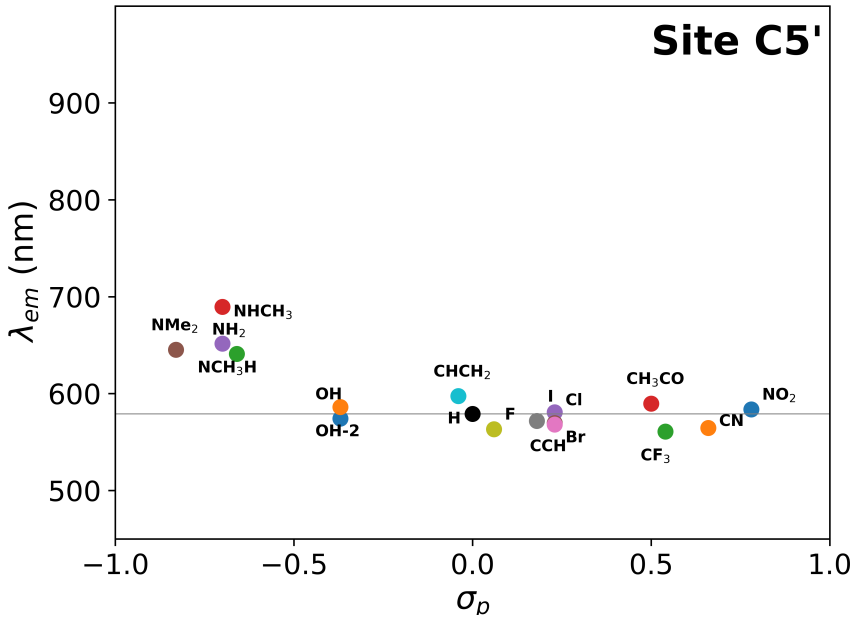


Figure 2.5: Computed emission wavelengths (λ_{em}) of C5'-functionalized oxyluciferin analogs versus the Hammett σ_p constants of the corresponding substituent groups. Emission wavelengths were obtained from TD-ωB97X-D/6-311++G**/C-PCM calculations.

Unlike the effects of having electron-withdrawing groups at C7', substitution by these groups at the C5' site did not cause a significant blue-shift in the emission spectra. The substitution of cyano (CN), trifluoro (CF_3) groups at the C5' site resulted in marginally blue-shifted oxyluciferin analogs emitting at 564.6 nm and 560.7 nm, respectively. The substitution of all halogens (F, Cl, Br, and I) at the C5' site also resulted in slightly blue-shifted oxyluciferin analogs. On the other hand, the oxyluciferin analogs such as 5'-CH_3CO-oxyluciferin (589.7 nm), 5'-NO_2-oxyluciferin (583.7 nm), and 5'-ethynyl-oxyluciferin (580.6

nm) led to a small redshift (10.6 nm, 4.6 nm, and 1.5 nm, respectively) in the emission wavelengths.

Clearly, these observed variations in emission wavelengths, which are somewhat irregular, cannot be explained by how the charge transfer associated with the $S_0 \rightarrow S_1$ transition of the oxyluciferin anion is facilitated or hindered by these substituents.

2.3.1.3 Site C4'

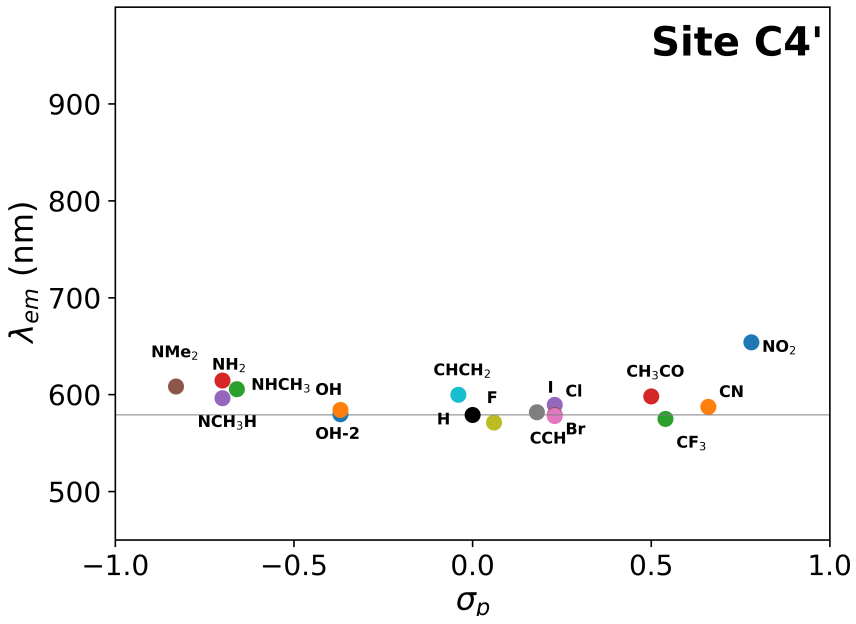


Figure 2.6: Computed emission wavelengths (λ_{em}) of C4'-functionalized oxyluciferin analogs versus the Hammett σ_p constants of the corresponding substituent groups. Emission wavelengths were obtained from TD-ωB97X-D/6-311++G**/C-PCM calculations.

Similar to the case of site C5', there also lacks a clear correlation between the emission wavelengths and Hammett constants for C4'-analogs (see Figure 2.6). The addition of electron-donating groups, such as $NHCH_3$, $N(CH_3)_2$ and NH_2, at the C4' site of oxyluciferin resulted in computed emission wavelengths in the red to the orange region at

614.5/596.4 nm, 608.4 nm, and 605.7 nm, respectively. The 4'-OH-oxyluciferin (584.2 nm) and 4'-(OH-2)-oxyluciferin (579.7 nm) analogs displayed a smaller redshift (5.1 nm, 0.6 nm) relative to the oxyluciferin anion. The oxyluciferin analog with a C4'-vinyl group was predicted to emit at 599.8 nm, which is 20.7 nm longer than the emission wavelength of OLU.

Very surprisingly, *the 4-nitro-oxyluciferin analog was also shown to produce a red-shifted emission wavelength of 654.0 nm.* This is counter-intuitive because, as we mentioned earlier, an electron-withdrawing group on the benzothiazole ring would likely hinder the net partial charge transfer ($0.0682e^-$ for 4-nitro-OLU versus $0.236e^-$ for OLU) to the thiazole ring during the $S_0 \rightarrow S_1$transition and thus should cause a blue shift.

Oxyluciferin analogs with a few other electron-withdrawing groups, such as 4'-CH_3CO-oxyluciferin (598.2 nm), 4'-CN-oxyluciferin (587.4 nm), 4'-CCH-oxyluciferin (589.5 nm), and 4'-I-oxyluciferin (581.9 nm) also emit at slightly longer wavelengths than oxyluciferin. Meanwhile, other oxyluciferin analogs with electron-withdrawing groups, such as 4'-CF_3-oxyluciferin (575.0 nm), 4'-F-oxyluciferin (571.2 nm), and 4'-Cl-oxyluciferin (577.9 nm), were found to be marginally blue-shifted.

Putting together the observations at the three sites, it is clear that the Hammett constants only manage to explain the substitution effects on the spectroscopic properties of a few selected oxyluciferin analogs (mostly C7'-analogs). Most intriguingly, the redshift caused by the electron-withdrawing NO_2 group in C4'-nitro-oxyluciferin cannot be rationalized based on the empirical Hammett constants. Therefore, in order to gain an understanding of this, below, we will pursue an analysis of the frontier orbitals of oxyluciferin analogs.

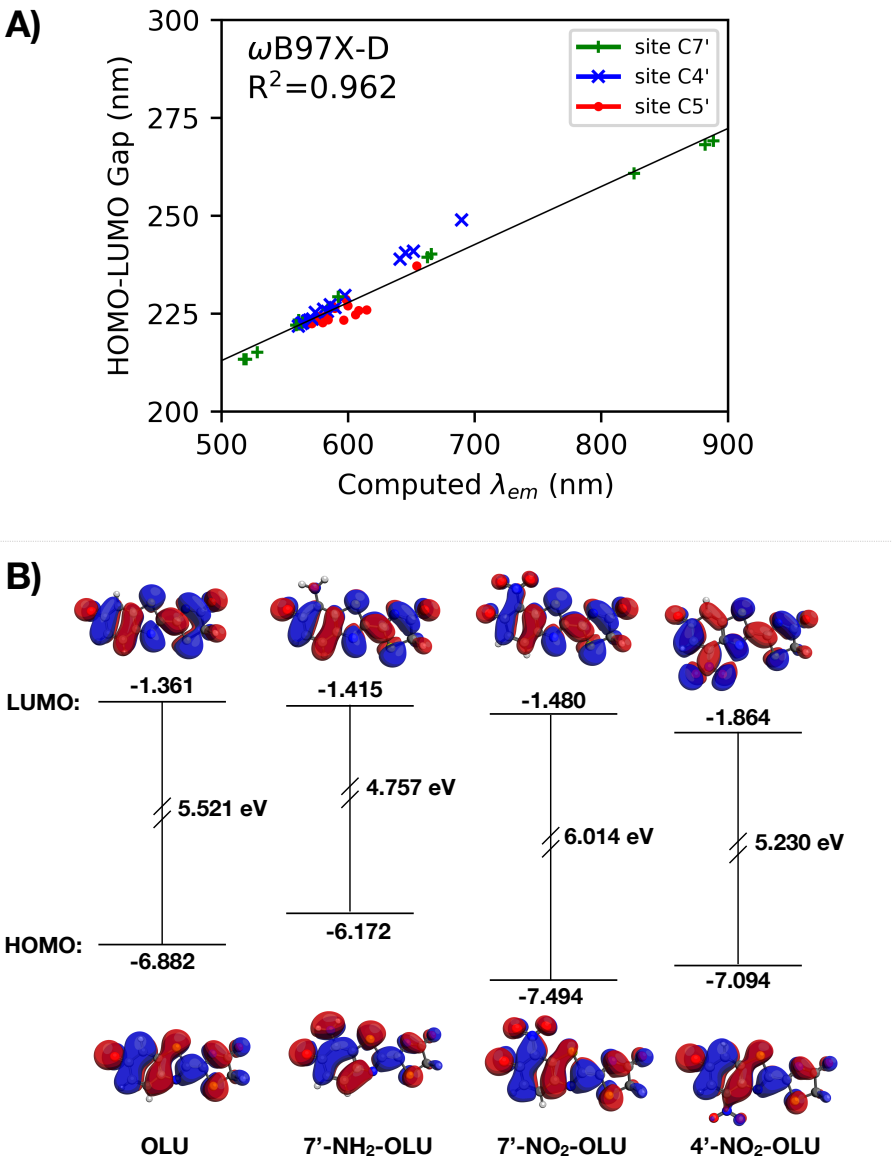


Figure 2.7: (A) HOMO-LUMO gaps (converted to nm) vs TDDFT emission wavelengths (in nm) of oxyluciferin analogs. (B) HOMO and LUMO orbital contours for oxyluciferin and selected analogs. All results were collected from ωB97X-D/6-311++G**/C-PCM(water) calculations.

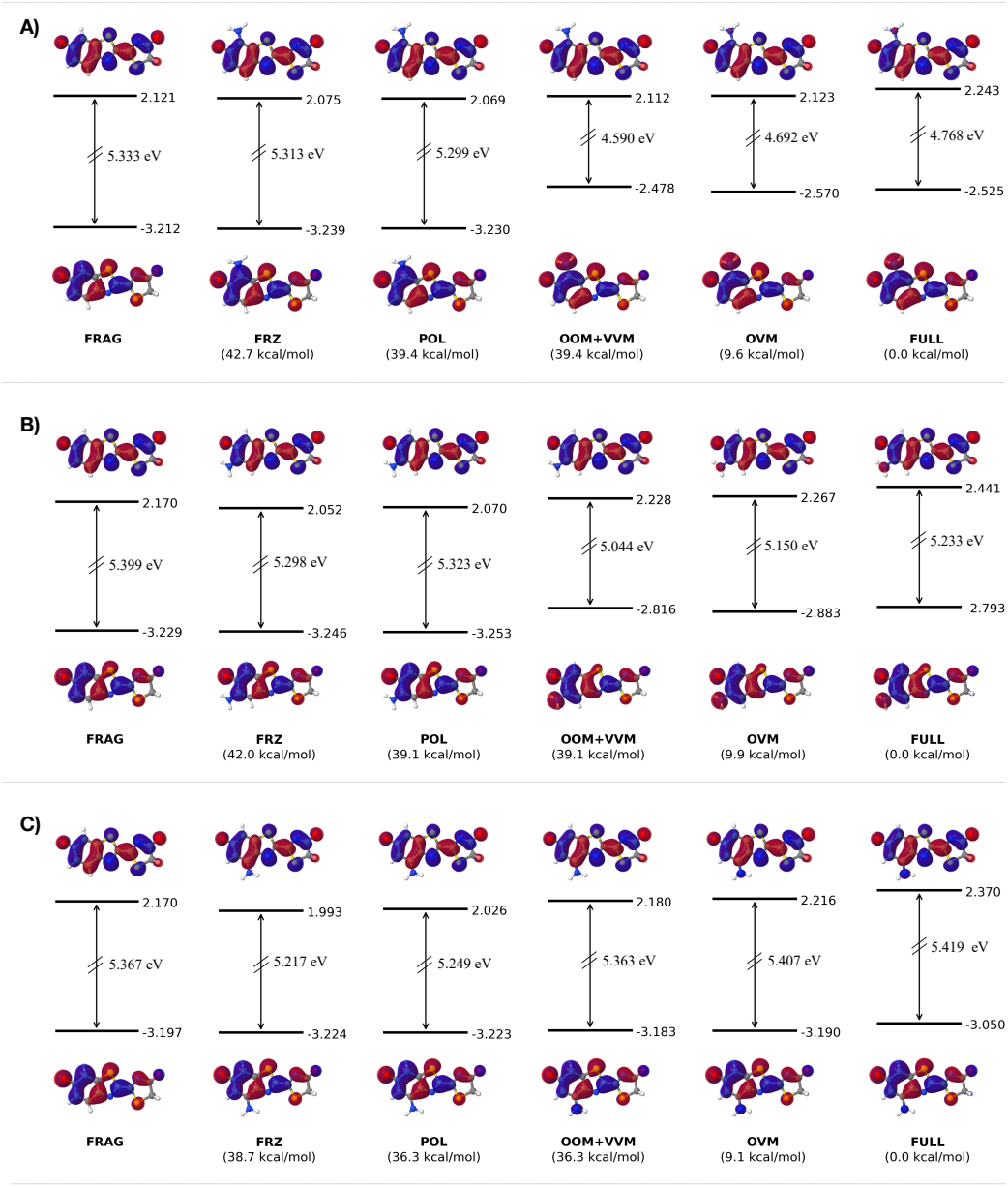


Figure 2.8: Effect of the amino group on the HOMO and LUMO of oxyluciferin when attached to the C7' (panel A), C5' (panel B), and C4' (panel C) sites. All results were collected from ALMO-based analysis using the ωB97X-D functional and 6-31G(d) basis set. The orbital energy values were reported in eV. The energies of the intermediate states relative to the full SCF energy of each analog were shown in the parentheses.

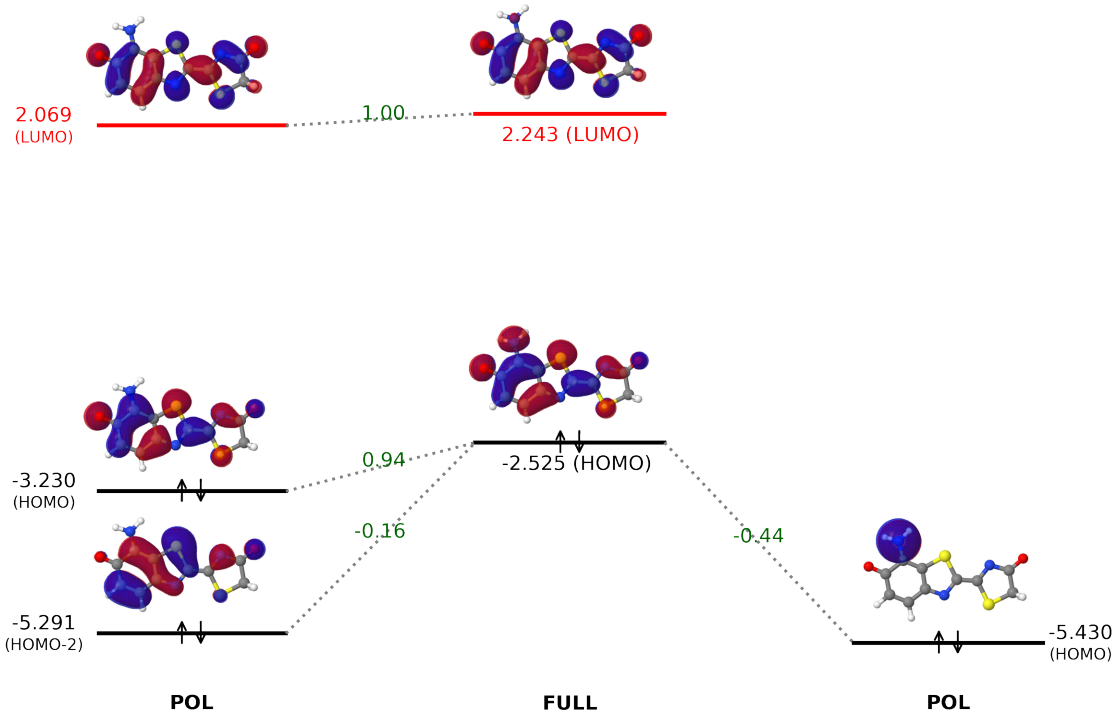


Figure 2.9: Interactions between polarized oxyluciferin and amino orbitals to yield the fully converged MOs of 7'-amino-oxyluciferin. The calculations for orbital interaction analysis were performed in the gas–phase using the ωB97X-D/6-31G(d) level of theory. All orbital energies are in eV.

2.3.2 Frontier Orbital Analysis

2.3.2.1 Correlation Between HOMO–LUMO Gap and Emission Wavelength

For oxyluciferin and its analogs, the HOMO–LUMO transition dominates the TDDFT $S_1 \rightarrow S_0$ emission. Specifically, the amplitude for HOMO→LUMO excitation was computed to be 0.9848 for the unsubstituted oxyluciferin anion. Among the analogs, the smallest HOMO→LUMO amplitude of 0.9807 was obtained for 7'-vinyl-oxyluciferin, whereas the largest amplitude of 0.9891 was found for the C4'-nitro-oxyluciferin analog (refer to Table B.12).

Given such dominance by the HOMO→LUMO excitation, we plotted the computed HOMO-LUMO gaps (converted to nm) of oxyluciferin analogs and their emission wavelengths in Figure 2.7A. A significant correlation (R^2= 0.962) between the HOMO–LUMO gaps and emission wavelengths was observed when the ωB97X-D functional was used. Similarly, the computed HOMO-LUMO gaps using the CAM-B3LYP functional also exhibited a remarkable correlation (R^2 = 0.967) in Figure B.8 with the emission wavelengths. This suggests that, in order to interpret the shifts in the emission wavelengths among various oxyluciferin analogs, it is essential to understand the variation in their HOMO-LUMO gaps.

Below we shall focus on the effects of NH_2 and NO_2 groups, both as representative substituents, at the C7'/C5'/C4' sites of oxyluciferin. As shown in Figure 2.7B, the incorporation of the NH_2 group at the C7' site of oxyluciferin reduced the HOMO-LUMO gap from 5.521 to 4.757 eV. Moreover, such a narrowing of 0.764 eV of the gap was caused mostly by the elevation of the HOMO energy by 0.710 eV (from -6.882 to -6.172 eV), while the LUMO energy was lowered only by 0.054 eV (from -1.361 to -1.415 eV). In contrast, the attachment of the NO_2 group at the C7'-site increased the HOMO-LUMO gap from 5.521 eV to 6.014 eV. The nitro group caused such an increase of 0.493 eV to the HOMO-LUMO gap by lowering the HOMO energy by 0.612 eV and the LUMO energy by only

0.119 eV. Surprisingly, when NO_2 group was placed at the C4' site, the HOMO-LUMO gap got reduced by 0.291 eV. For that analog, both HOMO and LUMO were lowered in their energy, but it was the latter (LUMO) that was subjected to a larger shift of -0.503 eV.

This brings up some intriguing questions: (a) *how do the amino group orbitals interact with the fluorophore to raise the HOMO energy?* (b) *why do the nitro group have the opposite effect, namely lowering LUMO energy?* and (c) *what is causing the aforementioned dramatically different effects upon the attachment of the nitro group at C4' and C7' sites?* This calls for a qualitative analysis of the underlying interactions between fluorophore and substituent orbitals, which will be provided in the next two subsections within the framework of ALMO analysis as described in Section 1.2.5.

2.3.2.2 Effect of amino group on three sites of OLU

Within the ALMO analysis, FRZ, POL, and OVM intermediate states provide variational upper bound to the FULL SCF energy. The total polarization energy (i.e., the difference between FRZ and POL state) varied from -2.4 kcal/mol for C4'-, -2.9 kcal/mol for C5'-, to -3.3 kcal/mol for C7'-amino substitution. This suggests that the amino group does not significantly stabilize the analogs through the polarization effect at the three sites. As shown in Figure 2.8, the difference between the HOMO energies in the FRAG state and FRZ state is marginal for all three amino-substituted analogs, indicating that the effect of permanent electrostatic/Pauli repulsion is not significant between the amino group and oxyluciferin. Shape-wise, HOMOs, and LUMOs of these three analogs remained essentially the same in the FRAG, FRZ, and POL states. As shown in Figure 2.8, the extension of both orbitals to the amino group occurred, if at all, through OOM. Specifically, in the case of C7'- and C5'-amino substitutions, the HOMO energy was increased when moving from the POL state to the OOM+VVM state by 0.752 and 0.437 eV, respectively. After such a large increase in the HOMO energy due to occupied-occupied orbital mixing, the HOMO energy converges further to the FULL state through the occupied-virtual orbital mixing (OVM).

In contrast to the significant changes in HOMO energies associated with the C7'- and C5'-amino analogs, the HOMO energy of the C4'-amino-oxyluciferin only changed minimally (< 0.05 eV) from the POL state to the OOM+VVM state.

On the other hand, the LUMOs and their energies did not change dramatically across the entire ALMO analysis procedure for all three amino-substituted oxyluciferin fluorophores. In the case of C5'- and C4'-amino substitutions, the LUMO energy increased by roughly 0.158 and 0.154 eV, respectively, due to virtual-virtual mixing. Meanwhile, the LUMO energy only increased by 0.043 eV for C7'-amino substitution.

To further understand the effect of amino group on the oxyluciferin substrate, we projected the frontier orbitals of the fully-converged frontier orbitals of C7'-amino-oxyluciferin onto its absolutely-localized polarized fragment orbitals. In Figure 2.9, the interactions between polarized amino and oxyluciferin fragment orbitals that give rise to the full-system HOMO and LUMO were shown along with the corresponding projection coefficients. Clearly, the amino group HOMO undergoes an out-of-phase mixing with the oxyluciferin occupied orbitals (HOMO and HOMO-2), which led to an overall increase in the full-system HOMO energy. As such, *a destructive (i.e. anti-bonding-like) interaction of the HOMO of a fluorophore with the highest occupied orbital of an electron-donating group is the key to the elevation of the HOMO energy*.

The impact of occupied-occupied orbital mixing on the HOMO-LUMO gap of amino-substituted OLU analogs is more clearly shown in the left column of Figure 2.12. Clearly, the strength of the net impact follows the order C7' $>$ C5' $>$ C4', which coincides with the weight of their respective carbon atomic orbitals within the OLU HOMO. Namely, the OLU HOMO, as shown in Figs. 2.8 and 2.9, receives the largest contribution from C7'-carbon, followed by C5'-carbon, while C4'-carbon has the smallest contribution. Accordingly, the HOMO energy is raised the most with the C7'-amino group and the least with the C4'-amino group. To design new fluorophores featuring smaller HOMO-LUMO gaps,

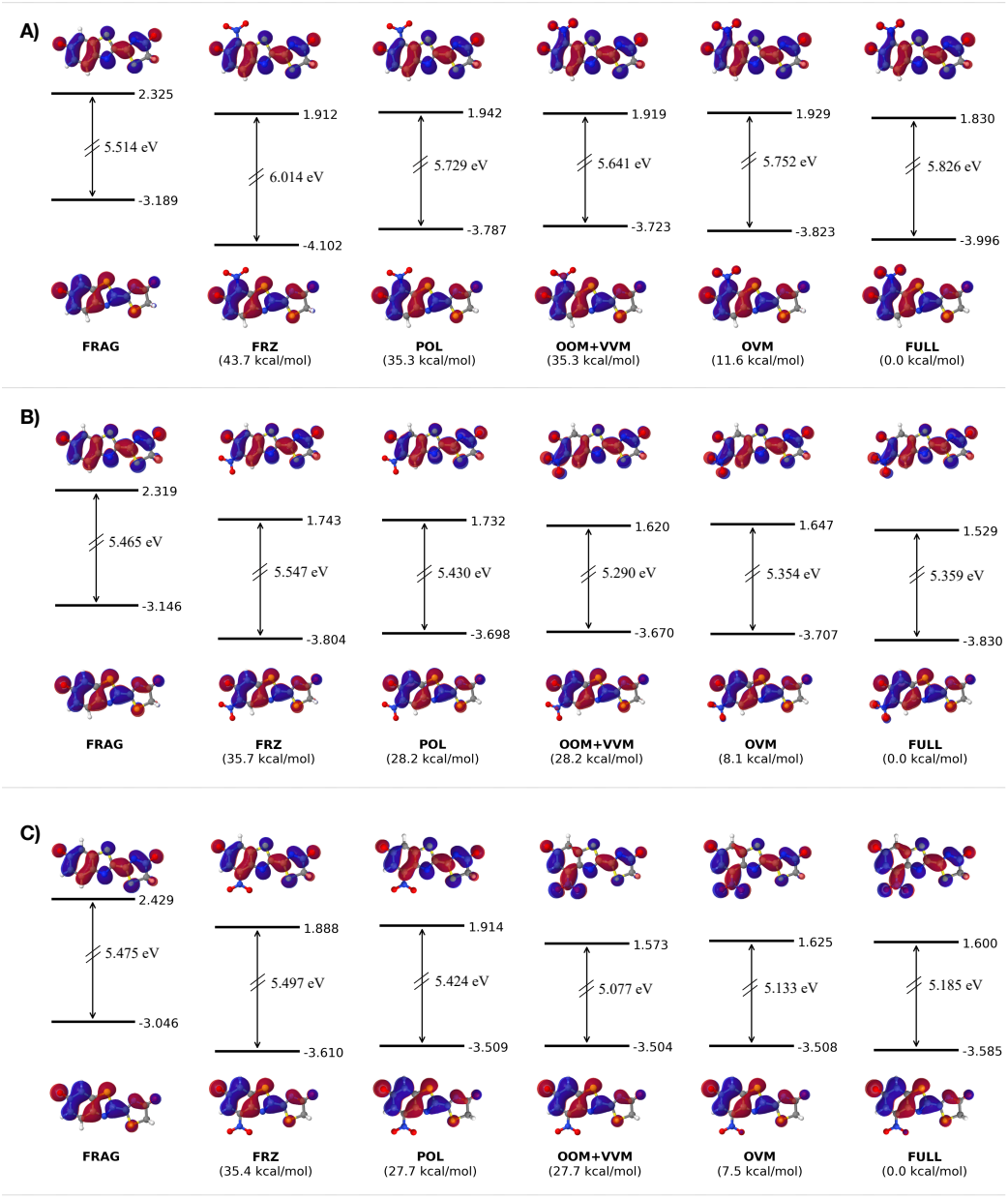


Figure 2.10: Effect of the nitro group on the HOMO and LUMO of oxyluciferin when attached to the C7' (panel A), C5' (panel B), and C4' (Panel C) sites. All results were collected from gas–phase ALMO-based calculations using the ωB97X-D functional and 6-31G(d) basis set. The orbital energy values were reported in eV. The energies of the intermediate states relative to the full SCF energy of each analog were shown in the parentheses.

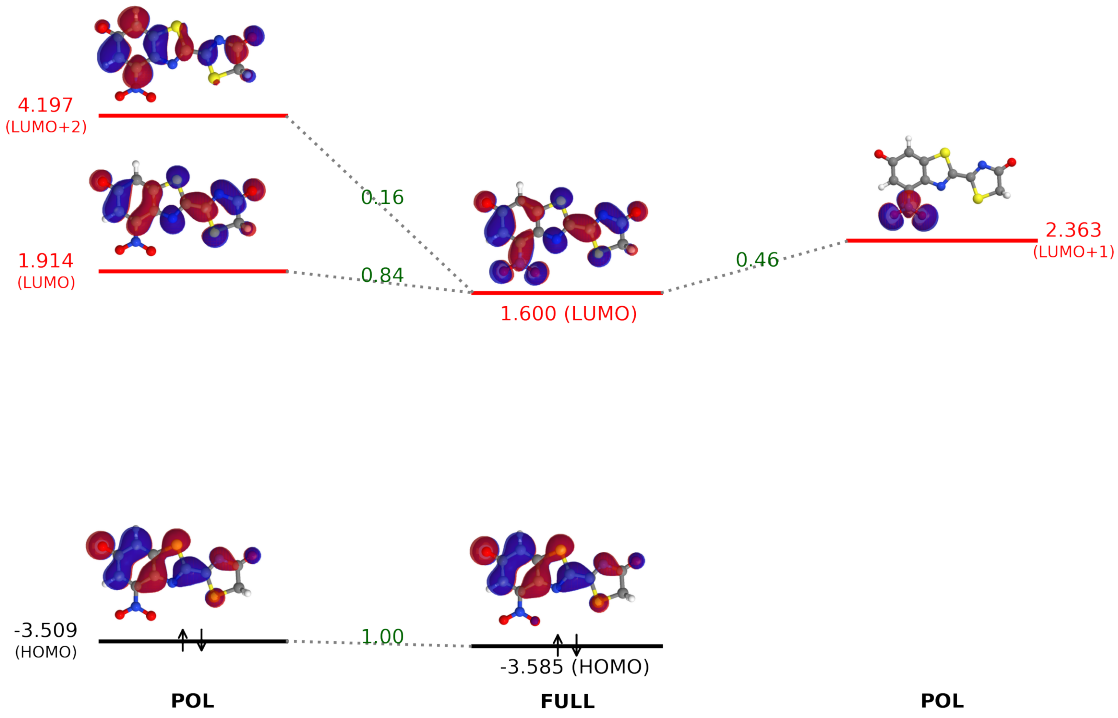


Figure 2.11: Interaction between polarized oxyluciferin and amino orbitals to produce fully converged MOs of 4'-nitro-oxyluciferin. The calculations for orbital interaction analysis were performed using the ωB97X-D/6-31G(d) level of theory. All orbital energies are in eV.

therefore, *one should attach strong electron-donating groups to atoms where HOMO has a large weight*.

2.3.2.3 Effect of nitro group on three sites of OLU

Unlike the amino group, the nitro group shifted the HOMO and LUMO energies substantially from the FRAG to the FRZ state. Specifically, the HOMO energy was lowered by 0.913, 0.658, and 0.564 eV, respectively, for C7'-, C5'-, and C4'-nitro-oxyluciferin, as shown in Figure 2.10. Meanwhile, the LUMO energy was reduced by 0.413, 0.576, and 0.541 eV, respectively. Such a significant contribution from permanent electrostatics/Pauli repulsion was expected for the nitro substituent, which has a larger dipole moment due to the presence of two electronegative oxygen atoms.

For C7'-nitro substitution, the HOMO energy was raised from -4.102 eV (FRZ) to -3.787 eV (POL) due to the polarization interaction. From this point onward, the HOMO energy gradually shifted back to the full-SCF value of -3.996 eV. Even smaller polarization effect on the HOMO energy was observed for C4' and C5'-nitro-substituted analogs (-3.804 $\rightarrow$ -3.698 eV and -3.610 $\rightarrow$-3.509 eV, respectively).

As far as LUMO is concerned, the largest virtual-virtual mixing occurred on the C4'-nitro-oxyluciferin analog, where the LUMO energy was reduced by 0.341 eV (1.914 $\rightarrow$ 1.573 eV). This is the primary cause for the narrowing of the HOMO-LUMO gap of this analog. In contrast, nitro groups at the C7' and C5' sites reduced the LUMO energy by only 0.023 and 0.112 eV, respectively, through the virtual-virtual orbital mixing. The LUMO orbitals then remained more or less constant after the VVM stage.

The orbital interaction analysis shown in Figure 2.11 revealed that the LUMO+1 orbital of the nitro group was combined with OLU's LUMO and LUMO+2. The resultant virtual orbital has lower energy due to the constructive (*i.e.* bonding-like) combination. Among the three sites, the C4'-carbon has the most substantial contribution to the isolated (and polarized) fluorophore LUMO and thus lowered the LUMO energy the most. It thus

caused the most significant reduction to the HOMO-LUMO gap, as shown in Figs. 2.12 and 2.10. Therefore, *in frontier molecular orbital engineering, one can attach strong electron-withdrawing groups with unoccupied π^* orbitals* to sites where the LUMO has a significant weight in order to lower the LUMO energy.

In Figure 2.12, the substituent effect of fluorine, another electron-withdrawing group, is shown.

2.4 Conclusions

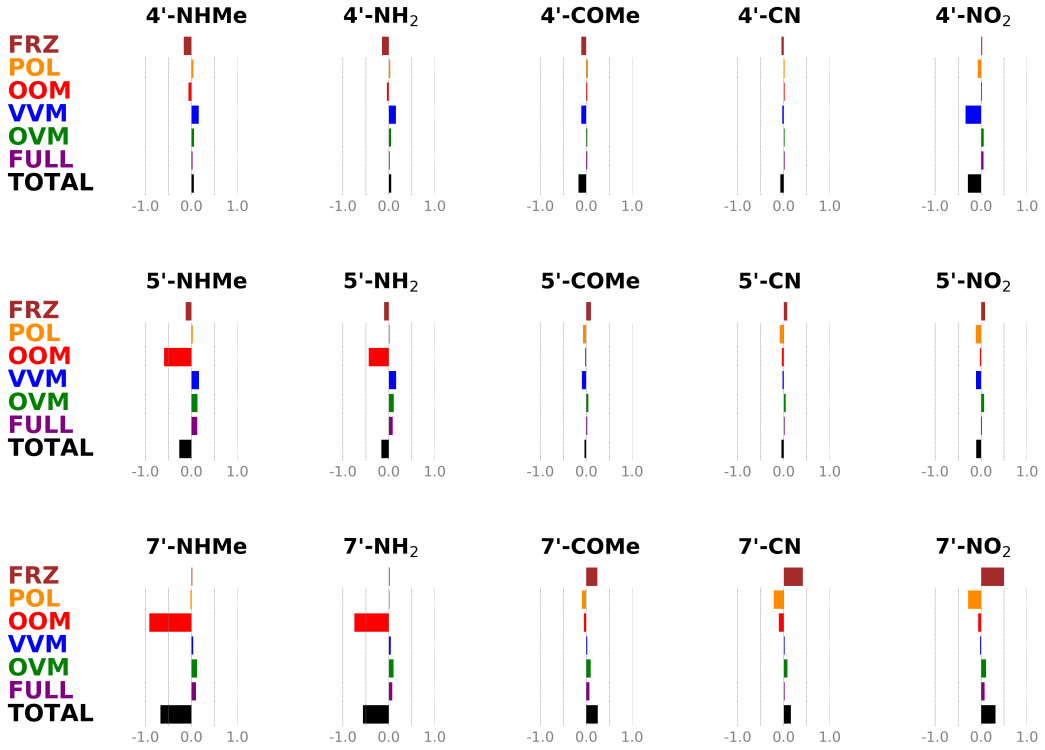


Figure 2.12: Decomposition of the effects of the NHMe, amino, acetyl, cyano, and nitro groups on the HOMO and LUMO of the oxyluciferin system obtained from the ALMO-based analysis. All data were computed at the ωB97X-D/6-31G(d) level of theory. Energy values are in eV.

Through a systematic computational study of oxyluciferin analogs with substituent groups at C7', C5', and C4'-sites, we found that

- Substitution of EDGs (such as NH_2, NHMe, OH groups) at the C7' position of OLU caused a redshift in the computed emission wavelength, while EWGs (such as NO_2 and CN groups) blue-shift the emission wavelength. There was a strong correlation between the empirical Hammett parameters σ_p and computed λ_{em} for C7'-OLU analogs.

- There is no correlation between σ_p and computed λ_{em} for C4'- and C5'-oxyluciferin analogs. Most surprisingly, the OLU analog with the NO_2 group at C4' position was predicted to emit at a significantly longer wavelength (by 74.9 nm) compared to that of OLU.

- The oo-mixing between the HOMO of the OLU core and the highest-lying occupied orbital of EDGs (such as the amino group) raised the HOMO level and significantly reduced the HOMO-LUMO gap. On the other hand, the vv-mixing between OLU's LUMO and the lowest-lying virtual orbital of EWGs (such as the nitro group) lowered the LUMO level and also caused a red-shifted emission.

To summarize, the overall design strategies are: strong electron-donating groups (such as amino and dimethylamino groups) should be placed at sites with a substantial contribution to the fluorophore HOMO, whereas strong electron-withdrawing groups (such nitro and cyano groups) should be attached to carbon atoms with a considerable LUMO contribution.

Chapter 3

Substituent Effect in Aβ Peptide Fluorescent Probes

Note: The following chapter is reprinted by permission from Springer Nature: Springer, Theoretical Chemistry Accounts, (Vardhan Satalkar et al, *Theor. Chem. Acc.* **138**(2),pp 29 (98), [COPYRIGHT](2019)

3.1 Background

Alzheimer's Disease (AD) is a progressive and fatal neurodegenerative disease. AD research is crucial because it affects a large percentage of the senior population and currently there are neither preventative measures nor effective treatments for the disease (99; 100). It has been known for many years that Aβ plaques are produced during different stages of AD. However, the exact role of Aβ plaques in the progression of AD had been a subject of dispute (101; 102), largely because their abundance found in postmortem histological studies correlate poorly with the severity of the disease. One plausible explanation is that Aβ plaques can be categorized into "silent" (mature) plaques and "active" newly formed ones. The latter is regarded to be more toxic and leads to cell deaths through inhibiting mitochondrial function, calcium hemostasis, and other neuronal cell functions (103).

It has been suggested that Aβ aggregates should be studied in conjunction with oxidative stresses (104; 105; 106; 107), because a higher reactive oxygen species (ROS) level is expected in AD brains than healthy ones. ROS can be produced during the formation and growth of new "active" Aβ plaques, specifically through the Fenton reaction of Aβ with metal ions during Aβ cross-linking. When ROS is accumulated, it can prompt a buildup of inflammatory cytokines, which then attracts microglia to surround the plaques (108). With the production of microglia, more ROS is released (109). Consequently, relatively high concentrations of ROS can be viewed as a signature feature of "active" Aβ plaques, and its

measurement provides a convenient technique to differentiate these plaques from "silent" ones (110).

To address the many challenges facing the AD research, many efforts have been devoted to the detection and imaging of Aβ aggregates and, more recently, ROS in the brain (see Ref. (111) for a review). Thioflavin (112) and other early probes emit visible light, which poorly penetrates the tissue and blood brain barrier. More recently, several near infrared fluorescence (NIRF) probes were designed to target Aβ fibrils, such as NIAD-**4** (113), BODIPY derivatives (114), MCAAD-**3** (115), DANIRs molecules (116), and a series of curcumin analogues (CRANAD-X) (20; 117; 118; 119; 110; 120; 121; 122). We note that computational modeling were carried out previously on Thioflavin (123), NIAD-**4** (124), DANIRs (125), and potential new NIRF probes (126). These calculations provided valuable insight into the photophysical properties of these fluorophores and their binding to Aβ fibrils.

In this work, we will focus on a series of boron-containing curcumin analogues (see Figure 3.1) mentioned above (20; 117; 118; 119; 110; 127), which are called CRANAD molecules, for detecting Aβ aggregates and ROS. Among these molecules, CRANAD-**54** is the smallest. CRANAD-**2** combines a pull force from an electron accepting difluoroboryl (-BF_2) group and push forces from electron-donating dimethylamino groups at both ends of the curcumin molecule, which led to a red shifted emission relative to the curcumin molecule (20). CRANAD-**2** molecule binds only to insoluble Aβ aggregates but not to soluble Aβs. To improve the binding, one (or both) of the hydrophobic phenyl moieties in CRANAD-2 can be replaced with a moderately hydrophilic pyridyl moiety. This leads to CRANAD-**3** (118) and CRANAD-**58** (117), which displayed fluorescence responses to both soluble Aβ oligomers and insoluble aggregates. Another molecule, CRANAD-**28** (117), was designed as a two-photon imaging probe.

Of particular interest are a pair of molecules, CRANAD-**5** and CRANAD-**61**, (110) which were developed recently as on-off NIRF probe molecules for sensing ROSs in AD

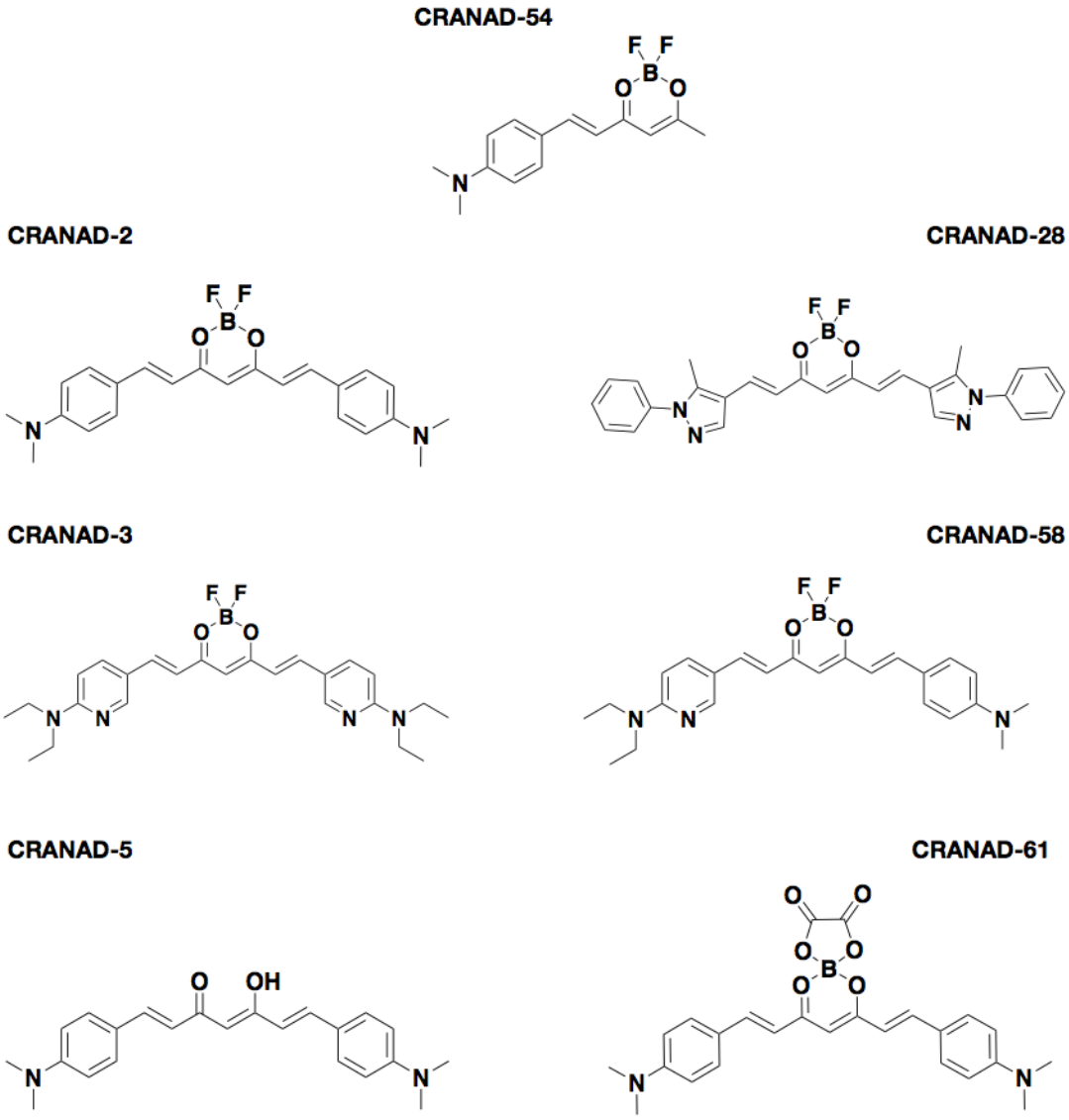


Figure 3.1: CRANAD molecules (reprinted by permission from Springer Nature: Springer, Theoretical Chemistry Accounts).

brains. A standalone CRANAD-**61** fluorophore emits in the near-infrared (NIR) region. But it can readily react with hydrogen peroxide and other ROS in living cells and animals, leading to the removal of the oxalate moiety. The molecule is converted to CRANAD-**5**, whose emission blue-shifts to the visible region. So an injection of CRANAD-**61** into the brain and a subsequent time-course measurement of fluorescent emission intensities at two different wavelengths will yield information about the ROS concentration in the brain (110).

In this article, a systematic computational modeling on CRANAD molecules will be reported. Specifically, we will investigate: (1) if the widely-used time-dependent density functional theory (TDDFT) calculations can be used to reliably predict the spectral shifts among CRANAD molecule; (2) if the absorption/emission wavelengths of these molecules correlate well with the energy gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO); and (3) what causes the difference in the emission wavelengths between CRANAD-**5** and CRANAD-**61** molecules?

We note that, very recently, Fages and coworkers performed a combined experimental and computational study of curcumoids with the borondifluoride moiety (127), where they focused on the effect of introducing different functional groups at the terminal and meso positions of the curcuminoid backbone. One of their analogs (**13** in Ref. (127)), which contains a dimethylamino group at both terminals, was referred to as CRANAD-**2** earlier (20) and in this paper. In their work, Fages *et al.* compared observed emission energies, electrochemical gaps (based on oxidation/reduction potentials from cyclic voltametry measurements), and computed HOMO-LUMO gaps, and found a good correlation among these data. Compared to their work, we will study the electronic structure of four other borondifluoride-containing CRANAD molecules (**3**, **28**, **54**, and **58**). In addition, through studying a boron-oxalate-containing CRANAD molecule (**61**) and its unsubstituted counterpart (**5**), it will help us understand the effect of different boron-containing groups, such

as $-BF_2$ and $-B(C_2O_4)$, on the ground and excited state properties of curcuminoid fluorophores.

3.2 Computational Details and Methods

All QM electronic ground-state calculations were performed using the density functional theory (DFT) method (73; 128) whereas excited-state calculations were carried out using the time-dependent density functional theory (TDDFT) method (17; 129). Version 4.4 of the Q-CHEM software package (21) was used.

Functionals. Four density functionals, including B3LYP (85; 130; 131), PBE0 (132), M06-2X (133), and ωB97X-D (134) were used for absorption energy calculations; while only B3LYP, PBE0, and ωB97X-D functionals were used in emission energy calculations (because TDDFT analytical nuclear gradient for meta functionals has yet to be implemented in Q-CHEM). The performance of density functionals in TDDFT calculations have been well documented (135; 136; 137; 138). In general, hybrid functionals such as B3LYP and PBE0 are expected to perform well for non-charge-transfer valence excited states. Meanwhile, M06-2X and ωB97X-D functionals, with higher ratios of long-range Hartree-Fock exchange (54% and 100%, respectively), usually provide better descriptions for charge-transfer states.

Basis set and integration grid. Most of the results were obtained using a triple-zeta basis, 6-311++G** (83), and SG-1 numerical integration grid (84), which is a pruned Lebedev grid with 50 radial points and 194 angular points (78). As shown in Table A.1 in the appendix, relative to 6-311++G** values, TDDFT absorption wavelengths (using the ωB97X-D functional) for three representative molecules (CRANAD-**5**, **54**, and **61**) changed by no more than 1.3 nm with the 6-311+G(2d, p) basis, which was used in Ref. (127), and by no more than 8.2 nm with cc-pVTZ or aug-cc-pVTZ basis set. The effect of using larger integration grid is also small — the computed absorption wavelengths

of three molecules in Table A.2 changed by no more than 1.3 nm between SG-1 and the much more demanding (99, 590) grids.

Implicit solvent model. To account for the solvent effect, we employed the polarizable continuum model (PCM) (90) for water solvent molecules, specifically the C-PCM model (139; 81) (as implemented in Q-CHEM (21; 92) using Bondi radii and a 1.2 scaling factor). For a comparison, some results using the integral equation formalism of PCM (IEF-PCM) (140) are shown in Tables A.3 and A.4 of the appendix, where the two solvent models led to differences of no more than 2.4 nm in the computed absorption wavelengths, and no more than 0.6 nm in the emission wavelengths. For absorption energy calculations, ground-state geometry optimization was performed using equilibrium PCM, then vertical excitation energies were computed using TDDFT and the non-equilibrium linear-response PCM model (93; 94). For a comparison, TDDFT absorption energies using a non-equilibrium perturbative state-specific model (141) are shown in Table A.5, which differ by no more than 3.0 nm from values in Table 3.1 obtained using the linear-response model. For emission energy calculations, excited-state geometry optimization was performed using TDDFT energies using the equilibrium linear-response PCM model.

Validation with curcumin conformers. As an initial validation of our computational approach, we computed the ground-state energy (using the ωB97X-D functional) and absorption/emission energies of six conformers of the curcumin molecule. As shown in Figure A.1, the ground-state energy varied by up to 3.3 kcal/mol among these conformers, while the absorption energy changed by no more than 0.010 eV (ranging from 3.2805 to 3.2904 eV) and the emission energy by no more than 0.017 eV (ranging from 2.5716 to 2.5886 eV). Such small variation among planar curcumin conformers is consistent with the observation by Benassi *et al* based on B3LYP/6-31+G* and B3LYP/6-311+G** calculations (142).

3.3 Results and Discussions

3.3.1 From curcumin to CRANAD molecules

The natural curcumin molecule contains hydroxyl and methoxy groups at both terminals. In most CRANAD molecules shown in Figure 3.1, the hydroxyl group is replaced with dimethylamino or diethylamino groups, while the methoxy group was removed from all molecules. The effect of such systematic structural substitutions on the electronic transition of curcumin molecule can be seen in Figure A.2 of the appendix.

Functional groups with different electron-donating strengths, as measured by their Hammett σ_p constants (1) [–H (0.0); –OH (-0.37); –NH_2 (–0.66); -NMe_2 (-0.83); –NEt_2 (-0.72)], were adopted in our calculations. For analogs without the methoxy groups (left column in Figure A.2), their computed absorption and emission energies clearly red shifted with stronger electron-donating groups at the *para* position. The same trend was observed with analogs with methoxy groups (right column).

The effect of the methoxy group is more subtle. With small groups (–H, –OH, and –NH_2) at the *para* position, the addition of a methoxy group at the neighboring (*meta*) position led to a decrease of about 0.04 or 0.05 eV in the computed excitation energies. With more bulky groups, –NMe_2 and –NEt_2, the addition of a methoxy group led to a less planar structures and, therefore, to an increase of about 0.12–0.15 eV in the computed absorption energies. For the ground state of curcumin molecule containing adjacent methoxy and dimethylamino groups, both methyl groups were found to be out of the aryl plane with computed C–N–C–C dihedral angles of 14° and 64°. These angles are reduced to around 7°, in the absence of the methoxy group.

3.3.2 Keto-enol Tautomerization of CRANAD-5

In addition to the enol form, which was observed in X-ray structures (143), the curcumin molecule can undergo keto-enol tautomerization (121; 122; 144; 145; 146; 147). Its diketo

form comprises of a methylene group (–CH_2–), which is reactive in basic conditions due to the presence of two adjacent carbonyl groups. According to previous computational studies, the enol form was predicted to be about 7.00 kcal/mol lower then the diketo form (146), with the latter adopting a non-planar structure. In agreement with these studies, our calculation indicated that the enol form is 2.40 – 5.63 kcal/mol lower in energy depending on the functional choice (Table A.6). This indicated that the enol form of curcumin is predominant in the condensed phase. In addition, as shown in Table A.6, the non-planar diketo form is predicted to have a vertical excitation energy that is 0.5–0.7 eV higher than the enol form.

Among the seven CRANAD molecules in this study (Figure 3.1), CRANAD-**5** is the only molecule that can also undergo keto-enol tautomerization. Our DFT calculations indicated, like the curcumin molecule, the enol form of CRANAD-**5** is also the energetically favorable conformer in the condensed phase. The keto-enol energy difference was predicted in Table A.7 to be 0.92 kcal/mol at the ωB97X-D/6-311++G**//C-PCM level of theory, and 3.41 kcal/mol with the B3LYP functional. As shown in Figure A.3, the enol form of CRANAD-5 is nearly planar, with a O1–C1–C2–C3 dihedral angle of 0.3° and a C1–C2–C3–O2 angle of -0.3° within the ωB97X-D/6-311++G**/C-PCM level of theory. The diketo form of CRANAD-5 is non-planar (Figure A.4), where the O1-C1–C2–C3 angle is found to be 103.6° and the C1–C2–C3–O2 angle is -7.3°.

3.3.3 Absorption Wavelengths

Experimental absorption wavelengths for seven CRANAD molecules are listed in Table 3.1. The shortest absorption wavelengths was found for the shortest molecule: CRANAD-**54** (490 nm). The wavelength was slightly longer with CRANAD-**28** (498 nm) and CRANAD-**5** (530 nm). Much longer wavelengths were measured for other four molecules: CRANAD-**3** (610 nm); CRANAD-**58** (630 nm); CRANAD-**2** (645 nm); and CRANAD-**61** (675 nm). All absorption energies were measured in the PBS buffer, which was approximated as water solvent in our computational study.

Table 3.1: Absorption wavelengths (in nm) of CRANAD-X molecules from TDDFT/C-PCM calculations using different functionals and 6-311++G** basis set.

Molecule	2	3	5	28	54	58	61
Expt.	645.0	610.0	530.0	498.0	490.0	630.0	657.0
B3LYP	593.1	573.7	534.2	470.4	473.6	585.0	610.0
PBE0	566.9	547.3	508.6	453.9	458.7	559.2	584.8
ωB97X-D	477.1	457.8	415.0	397.2	414.3	469.0	498.6
M06-2X	507.3	489.6	436.7	412.1	432.1	501.5	532.8

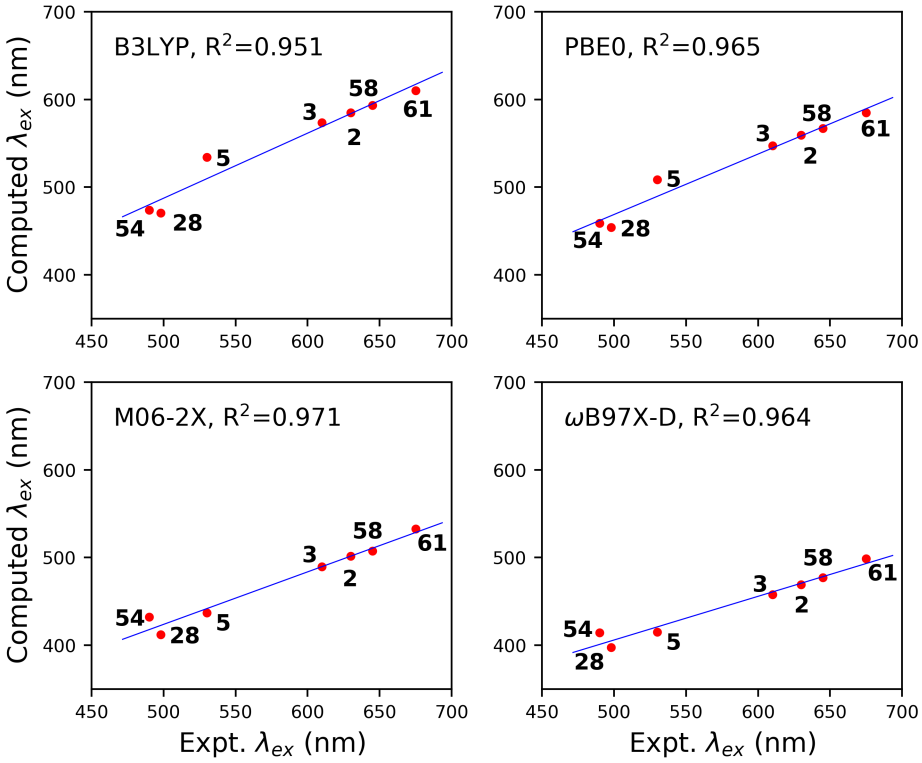


Figure 3.2: Computed TDDFT absorption wavelengths (in nm) *vs* experimental values for CRANAD molecules (reprinted by permission from Springer Nature: Springer, Theoretical Chemistry Accounts).

Computed TDDFT absorption wavelengths using four different functionals (B3LYP, PBE0, M06-2X and ωB97X-D) are also listed in Table 3.1. They are plotted against experimental values in Figure 3.2. It is clear that TDDFT calculations reproduced the absorption spectral shifts rather well (Figures A.6 and A.7). Standard hybrid functionals, B3LYP and PBE0, yielded wavelengths closer to experimental values, and the corresponding R^2 values are 0.951 and 0.965, respectively. On the other hand, M06-2X and ωB97X-D functionals led to systematically shorter wavelengths, but with R^2 values of 0.971 and 0.964, a strong correlation was also observed between computed and experimental absorption wavelengths.

It is interesting to note that in Table 3.1 that, in almost all cases, TDDFT calculations consistently yielded shorter wavelengths than experimental values. The only exception was observed for CRANAD-**5**, where computed B3LYP absorption wavelength is longer by 4.2 nm than the experimental peak at 530 nm. This exception might arise from structural flexibilities of CRANAD-**5**. As the only CRANAD molecule in this study containing a β-keto enol group, CRANAD-**5** can undergo keto-enol tautomerization, intramolecular proton transfer, or hydroxyl deprotonation.

Keto-enol tautomerization of CRANAD-**5** can lead to absorption at a longer wavelength. As mentioned in Section IIIB above, the (non-planar) diketo form was found to be only 3.41 kcal/mol (or 0.92 kcal/mol) higher than the enol form based on B3LYP (or ωB97X-D) calculations. TDDFT calculations results in Table A.7 showed that the absorption wavelengths of the diketo form were predicted to be 50–100 nm shorter than those of the enol form.

Intramolecular proton transfer can lead to a small red shift in the absorption wavelength. To study this effect, potential energy scan was performed with the proton moving between the two oxygen atoms (O1 and O2) in the enol form of CRANAD-**5**. As shown in Figure A.5, when the proton is moved from one oxygen atom towards the equidistant point, the ground-state energy was increased by about 2.5 kcal/mol. This is accompanied by a 0.05

eV reduction of the excitation energy, which amounts to a 7 nm increase in the absorption wavelength.

On the other hand, hydroxyl deprotonation can cause a blue-shift in the absorption wavelength. Upon the deprotonation of the hydroxyl group, CRANAD-**5** becomes an anion. As shown in the next subsection, the lowest-energy excitation involves the transfer of a small amount of electronic charge from the terminal groups towards the middle. This excitation will be hindered by more negative charges on the oxygen atom after hydroxyl deprotonation.

3.3.4 Wavefunction Analysis

In TDDFT calculations, the lowest-energy excitation of CRANAD molecules were found to be dominated by an excitation from HOMO to LUMO. So it would be worthwhile to examine the computed HOMO-LUMO gaps of these molecules (Figure A.8). As is well known, the values of computed HOMO-LUMO gaps strongly depend on the functional in use, and, in general, their values increases with the ratio of the long-range Hartree-Fock exchange for the functional. This can be easily be seen in Figure 3.3, where HOMO-LUMO gaps (y-axis values) follow the trend: B3LYP $<$ PBE0 $<$ M06-2X $<$ ωB97X-D. With each of these functionals, there is a strong correlation between HOMO-LUMO gaps and vertical excitation energies (x-axis values), with the correlation coefficients ranging from 0.938 (B3LYP) to 0.978 (M06-2X). This allows us to explain the variation of excitation energies of CRANAD molecules through studying the changes in their HOMOs and LUMOs.

Of particular interest to us is the red shift observed in the absorption wavelength from CRANAD-**5** to CRANAD-**61**. As shown in Figure 3.4, the HOMO of CRANAD-**5** appeared very similar to that of CRANAD-**61**. These two Kohn-Sham orbitals have nearly the same energies, with values of -0.2548 a.u. (CRANAD-**5**) and -0.2564 a.u. (CRANAD-**61**), at the ωB97X-D/6-311++G** level of theory.

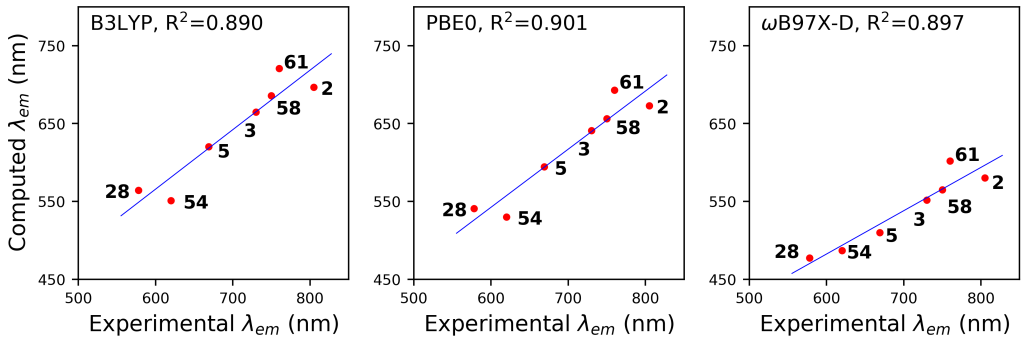


Figure 3.3: HOMO-LUMO gaps *vs* TDDFT absorption excitation energies (in nm), both computed at ground-state equilibrium geometries (reprinted by permission from Springer Nature: Springer, Theoretical Chemistry Accounts).

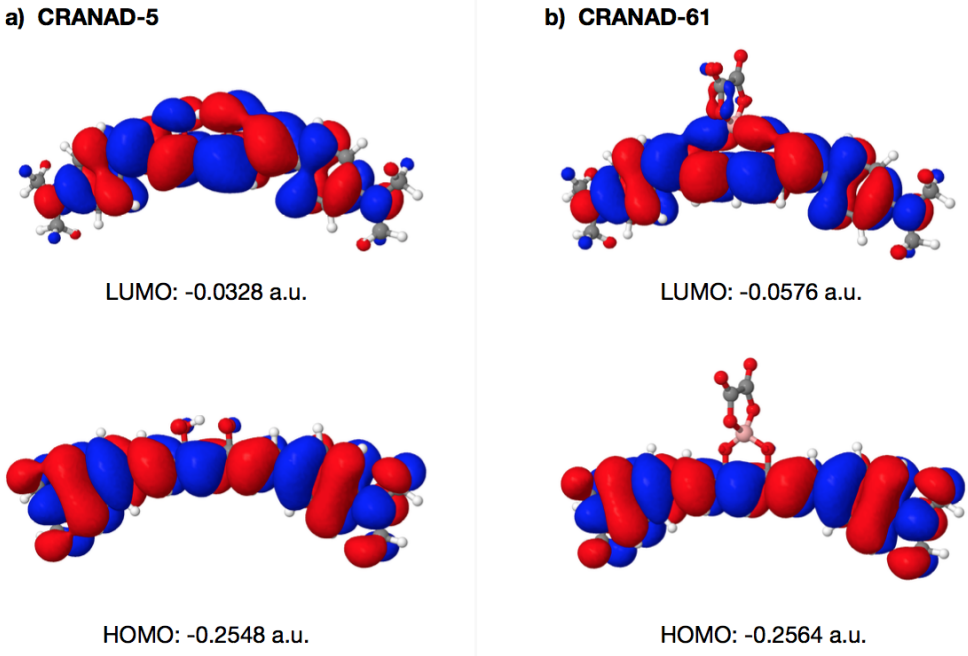


Figure 3.4: HOMO and LUMO orbitals of CRANAD-**5** and CRANAD-**61** at the ωB97X-D/6-311++G** level of theory (reprinted by permission from Springer Nature: Springer, Theoretical Chemistry Accounts).

On the other hand, Figure 3.4 shows a small but noticeable difference in the LUMOs of these two molecules. There is a small contribution from the oxalate moiety to the LUMO of CRANAD-**61**, which making the orbital slightly more delocalized than the LUMO of CRANAD-**5**. As a result, the LUMO energy of CRANAD-**61**, -0.0576 a.u., was substantially lower than the -0.0328 a.u. LUMO energy value for CRANAD-**5**. This led to a smaller HOMO-LUMO gap for CRANAD-**61** (0.1988 a.u.) than the gap of CRANAD-**5** (0.2220 a.u.), which in turn led to a relatively lower excitation energy (and longer absorption wavelength) for CRANAD-**61**.

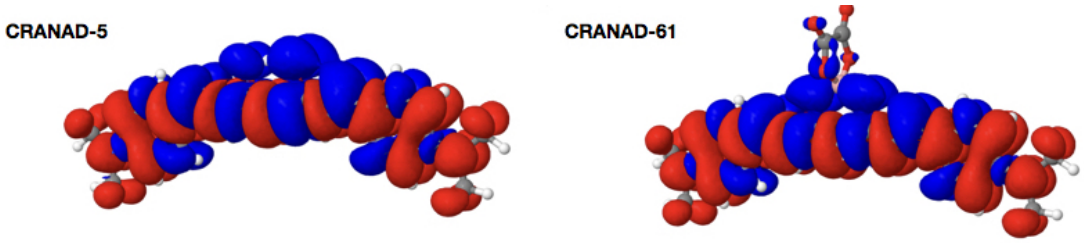


Figure 3.5: Unrelaxed difference density of CRANAD-**5** and CRANAD-**61** from TDDFT/ωB97X-D/6-311++G** calculations (reprinted by permission from Springer Nature: Springer, Theoretical Chemistry Accounts).

The more delocalization nature of the electron excitation can also be seen in Figure 3.5, which shows the unrelaxed difference densities (i.e. difference in the ground-state and excited-state electronic densities) of two molecules. For CRANAD-**61**, a small amount of electron density was transferred to the oxalate moiety upon the vertical excitation, and the actual amount was found to be -0.0177 e (Merz-Kollman ESP charge) and -0.0145 e (Mulliken population) from TDDFT/ωB97X-D/6-311++G** calculations.

3.3.5 Emission Wavelengths

Experimental and computed emission wavelengths for CRANAD molecules are tabulated in Table 3.2, and their values are also plotted in Figure 3.6. Similar to the case of absorption calculations, TDDFT calculations using hybrid functionals B3LYP and PBE0 produced

Table 3.2: Emission wavelengths (in nm) of CRANAD-X molecules from TDDFT/C-PCM calculations using different functionals and 6-311++G** basis set (reprinted by permission from Springer Nature: Springer, Theoretical Chemistry Accounts).

Molecule	2	3	5	28	54	58	61
Expt.	805.0	730.0	669.0	578.0	620.0	750.0	760.0
B3LYP	696.7	664.7	620.3	564.4	551.0	685.8	720.8
PBE0	672.9	640.9	594.3	541.0	530.0	656.4	693.0
ωB97X-D	604.9	574.0	529.4	494.0	511.3	588.7	628.8

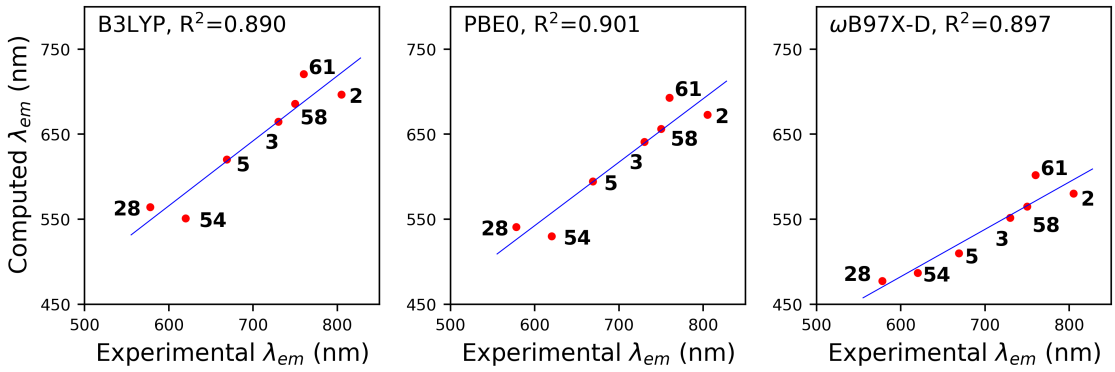


Figure 3.6: Computed TDDFT emission wavelengths (in nm) *vs* experimental values for CRANAD molecules (reprinted by permission from Springer Nature: Springer, Theoretical Chemistry Accounts).

emission wavelengths closer to experimental values, while TDDFT calculations using the range-separated ωB97X-D functional tended to overestimate the excitation energies and thus underestimate the wavelengths. On the other hand, with all three functionals (B3LYP, PBE0, and ωB97X-D), TDDFT calculations well reproduced the experimentally observed spectral shift in the emission wavelengths among CRANAD molecules. The R^2 values were found to be 0.89-0.90 for all three (B3LYP, PBE0 and ωB97X-D) functionals. This reaffirms that TDDFT calculations with these functionals can be employed to predict emission spectral shifts of fluorescent probes such as CRANAD molecules.

3.3.6 Torsional energy profiles of CRANAD-54

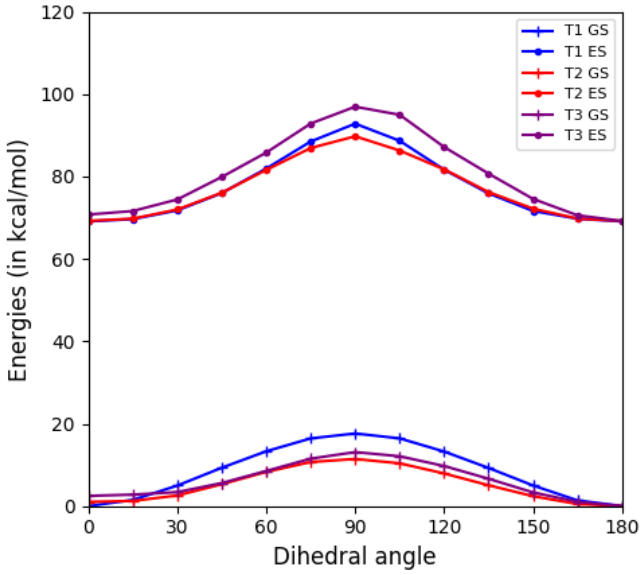


Figure 3.7: Torsional energy profiles for CRANAD-**54** from three potential energy scans, with respect to T1, T2 and T3 angles. Ground state energies were optimized with ωB97X-D/6-311++G** with T1 or T2 or T3 constrained to a given value. Excited-state energy for each constrained geometry were computed using TDDFT and ωB97X-D/6-311++G** (reprinted by permission from Springer Nature: Springer, Theoretical Chemistry Accounts).

Recently Segado, Benassi and Barone (142) studied the torsional energy profile of o-(p-dimethyl-amino-stryryl)methylpridinium cation (DASPMI). For this cationic molecule, torsional motions around three single bonds led to multiple minima on the excited-state potential energy surface, which causes nonemissive twisted intramolecular charge transfer (TICT) and multifluorescence.

To study whether TICT might affect CRANAD molecules, we followed their work and explored torsional energy profiles of CRANAD-**54**, which is the smallest molecule in our set of molecules. Figure 3.7 clearly shows that, with each of the three torsional angles (T1, T2, and T3, see Figure A.11), both ground-state and excited state energies have a minimum around 180°. Therefore, TICT is unlikely to affect CRANAD-**54**. This also explains why geometry optimizations (for both absorption and emission calculations) led to nearly planar structures for the CRANAD molecules (see Tables A.8 and A.9).

Our torsional scan also showed the effect of boron-containing groups, such as $–BF_2$ and $–B(C_2O_4)$, on the rigidity of the half-curcumin scaffold in CRANAD-**54**. As shown in Figure A.11 and Table A.11, in comparison to the keto-enol half-curcumin, both BF_2 and $B(C_2O_4)$ groups increased the ground-state barrier for all three rotations, which thus made these half-curcumin derivatives more rigid.

3.4 Conclusions

Through performing TDDFT calculations of the lowest-energy electronic transition of CRANAD molecules, the following conclusions were reached:

- Computed TDDFT absorption and emission wavelengths well reproduced the spectral shifts among CRANAD molecules, with R^2 values of above 0.95 for absorption wavelengths and above 0.72 for emission wavelengths.

- Computed vertical excitation energies were found to correlate well with the HOMO-LUMO gaps. This supports the use of HOMO and LUMO energies in the analysis

of the absorption/emission spectra shifts and in the design of new curcumin-based fluorescent molecules.

- Between the recently proposed pair of curcumin derivatives, CRANAD-**5** and CRANAD-**61**, the LUMO energy was lowered in CRANAD-**61** due to a contribution from the oxalate moiety, which extended the delocalization of the orbital.

At the same time, with a red-shift in the electronic transition energies and a related reduction in the HOMO-LUMO gaps,

Chapter 4

Substitution Effect of β-silicon group in C–H Activation Reaction

4.1 Background

On account of the complex nature of organic substrates, several types of C–H bonds exist in chemical structures. Even though these C–H bonds are generally non-reactive, several organometallic or organic catalysts were developed during the last couple of decades to activate rather stable C–H bonds to functionalize complex organic molecules (148; 149; 150; 151). This has led to several important organic scaffolds developed by utilizing C–H insertion strategies such as oxidative addition, single electron transfer, σ-bond metathesis (152), and coordination assisted control.

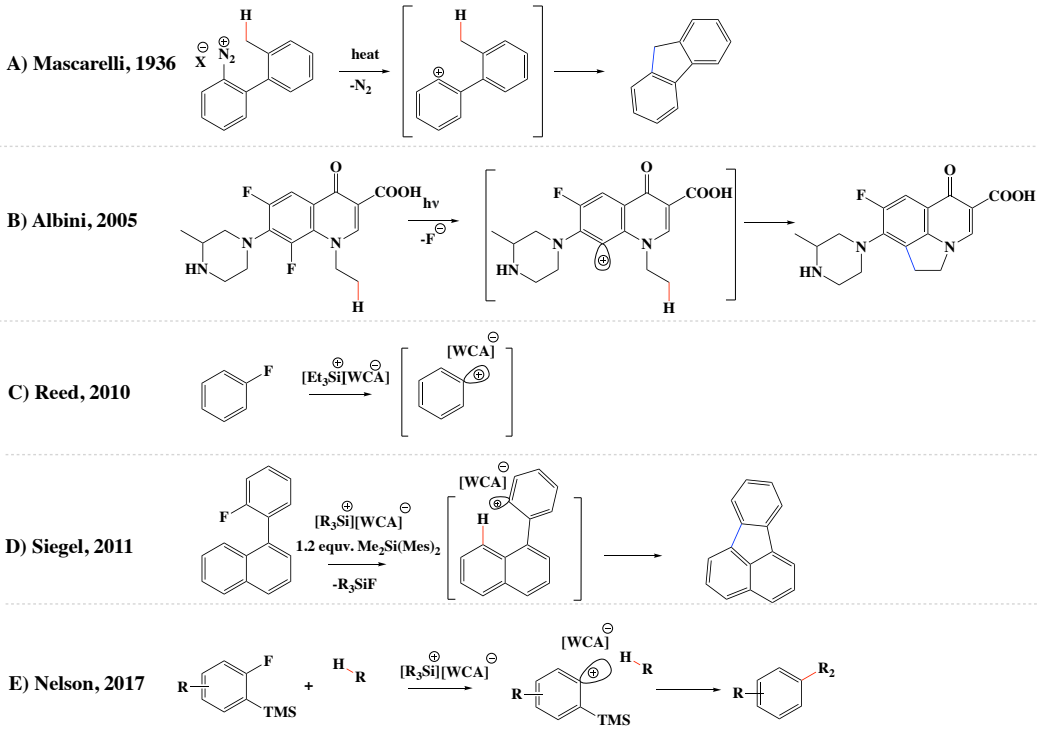


Figure 4.1: Reactions involving aryl cations as a key intermediate

Amongst, one of the widely-used strategies to functionalize a rather non-reactive C–H chemical bond is the coordination assisted control. In general, coordination assisted control can be achieved using a variety of directing groups present in the chemical species. The weak coordination of such directing groups leads to a metal chelating complex, which can be pivotal in the control of selectivity and the rate of product formation. Early efforts from Yu and coworkers highlighted the significance and applicability of the weakly coordinating directing groups (150). In order to gain detailed mechanistic insights about the fluoride abstraction and C–H insertion steps, several experimental and computational studies have been reported recently (153; 154). In particular, a regioselective C–H arylation using aryl halides can be achieved through coordination-assisted control. For instance, the weak chemical coordination of a directing group to cationic arene π-complex can be pivotal in the determination of regioselectivity and the rate of the C–H activation reaction (153).

The first evidence of such aryl cation utility in C–H insertion was reported by Reed and coworkers (155). Specifically, they analyzed C-–F activation in fluorobenzene, in which silylium carboranes are formed as an intermediate along with weakly coordinating anion (WCA, carborane anion). Subsequently, Albini and coworkers performed light-catalyzed selective heterolytic defluorination of 6,8-difluoroquinolone leading to C–H insertion reaction (156). The intermediate of this reaction was believed to be cationic in nature (156). Later, Siegel and coworkers showed that the participation of a phenyl cation intermediate in intramolecular Friedel–Crafts reactions of aryl fluorides (157). However, in these applications, the high reactivity of aryl cations resulted in poor selectivity for the desired products.

In a substantial advance that overcame the above limitation, Nelson and coworkers reported a selective arylation of strong sp^3/sp^2 C–H and C–F bonds with the help of β-trimethylsillane (TMS) directing group *via* carborane catalysis (153). The role of β-TMS group was found to be rather substantial because in the absence of the β-silicon directing

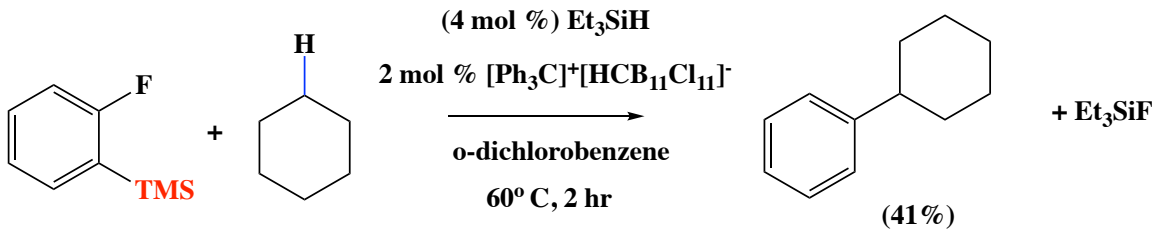


Figure 4.2: Arylation of cyclohexane using at fluoroarene in ODCB solvent.

group, the C–H activation reaction did not proceed in quantitative yields. In their experimental studies, the authors reported that the *ortho*-TMS-aryl fluoride underwent the CH activation cycle with a 55% yield. In contrast, the *para*-TMS and *meta*-TMS substituted aryl fluorides did not undergo any C–H insertion reaction. Nelson and coworkers also suggested that C–H activation *via* the aryne intermediate pathway was highly unlikely based on their experimental results (153).

To better understand these experimental findings, Li and coworkers performed DFT calculations to study both C-F activation and C–H insertion steps as well as the effect of WCA on the reaction mechanism (158). They also carried out the non-covalent interaction (NCI) analysis to account for the weak coordination between aryl cation and WCA. Although they mentioned the key importance of the formation of a β-silicon stabilized Wheland intermediate, a detailed analysis of the hyperconjugation effect of β-silicon during the fluorine abstraction step was not provided. In a related study, Houk and coworkers elucidated the C–H insertion reaction mechanism of vinyl cations formed during silylium-WCA catalysis with the help of DFT calculations and MD simulations (154).

Towards a better understanding of the enhancement of the reactivity of aryl cation by the β-silicon group, here we report a detailed computational analysis. Specifically, we identified the transition state structures and thus obtained reaction barriers for *ortho*, *meta*, and *para*-TMS-AR molecules. Furthermore, ALMO-EDA analysis was performed to quantify

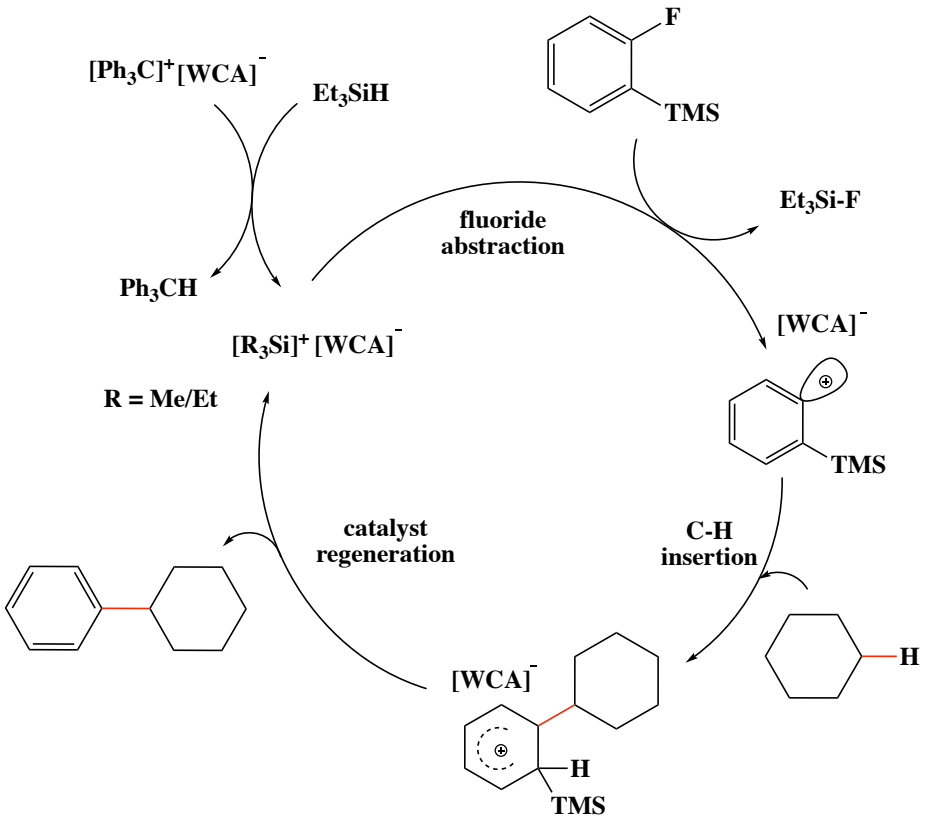


Figure 4.3: Proposed catalytic cycle[a] for the fluoride abstraction of β-TMS-fluorobenzene and the subsequent C–H insertion reaction with cyclohexane.

[a]Catalytic cycle was reported by Nelson and co-workers in Ref. 153.

the effect of the TMS directing group during the fluoride abstraction step, where the corresponding transition state was found to be stabilized the most by a charge transfer from the β-TMS group to arene cations.

4.2 System setup and Computational Details

In order to set up for the fluoride abstraction reaction, three components: 1) cationic triethyl silane $[Et_3Si]^+$; 2) anionic carborane $[WCA]^-$; and 3) neutral trimethyl silane (TMS) substituted fluorobenzene substrates were constructed using the IQmol software (TMS-C_6H_4F first, then Et_3Si, and then WCA). The intermediate structure (INT_1) from the fluoride abstraction step was used as the starting point of the subsequent C–H insertion step where the by-product of fluoride abstraction step, namely Et_3SiF, was replaced with cyclohexane in the INT_1 structure.

In this work, density functional theory (DFT) (16) was adopted due to its accuracy and affordability in terms of computational cost. All calculations were performed using the Q-CHEM version 5.2 (21). All structures of reactants, products, transition states, and intermediates were optimized using ωB97X-D and M06-2X functionals with a 6-31G* basis set. Following the experimental protocol, *o*-dichlorobenzene (ODCB) was employed as a non-polar solvent and described with the C-PCM solvent model (90). Frequency calculations were also performed using the same level of theory to find the number of imaginary harmonic frequencies and to obtain the vibrational entropy contribution to the reaction free energy. All transition states were verified to have only one imaginary frequency. Intrinsic reaction coordinate (IRC) pathway was generated using the GRRM software (159). To examine the basis set effect, single-point energies of the reagents in the fluoride abstraction reaction were also computed using the def2-TZVP basis set (160).

To decipher the difference in the electronic structural properties of *ortho*-TMS -fluorobenzene, *meta*-TMS-fluorobenzene, and *para*-TMS-fluorobenzene, we divide these complexes into two fragments, TMS and benzene, which are linked together by a C–Si

bond. Then, we perform the ALMO analysis (as described in Section 1.2.5) on these complexes.

4.3 Results and Discussions

Computational results will be presented below in three subsections: 1) mechanism of fluoride abstraction; 2) mechanism of C–H insertion; and 3) effects of β-TMS as a substituent and WCA a catalyst.

4.3.1 Mechanism of Fluoride Abstraction

Key bond lengths for *ortho*-TMS, *meta*-TMS, and *para*-TMS functionalized fluorobenzene substrates in complex with $[Et_3Si]^+[WCA]^-$ in the fluoride abstraction were listed in Figure 4.4. In all three reactant geometries, the fluorobenzene substrate was found to be well separated from the active catalyst $[Et_3Si]^+[WCA]^-$. Meanwhile, the carbon-fluorine (C–F) bond for all three fluorobenzene substrates has a normal length: 1.35 Å for *meta*-TMS-fluorobenzene) and 1.36 Å for *ortho*- and *meta*-TMS-fluorobenzene.

As the reaction progresses, the cationic fragment of the active catalyst $[Et_3Si]^+$ approaches the fluorobenzene substrate and interacts with it to weaken the C–F bond. By the transition state (TS) structures, the C–F bond lengths were stretched to 2.50, 2.27, and 2.47 Å, respectively, for the *ortho*-, *meta*-, and *para* substrates (Figure 4.4). A further weakening of the aryl C–F bond increases its length eventually to 3.17 Å (*ortho*), 3.12 Å (*meta*), and 3.05 Å (*para*). Due to the dipolar nature of C–F bonds, this led to the generation of an aryl cation moiety also known as the Wheland intermediate (INT_1), which was stabilized by the anionic $[WCA]^-$ part of the active catalyst. Moreover, in this work, we also considered the possibility of C-F abstraction mechanism *via* benzyne intermediate formation route. The TMS-aryl bond was partially weakened from reactants to TS (from 1.89 Åto 1.93 Å). However, during the INT_1, the bond length reduces to 1.92 Å. The reaction energy

of the formation of benzyne INT was found to be +51.1 kcal/mol higher than much more stable aryl-cation INT_1 (Table C.1).

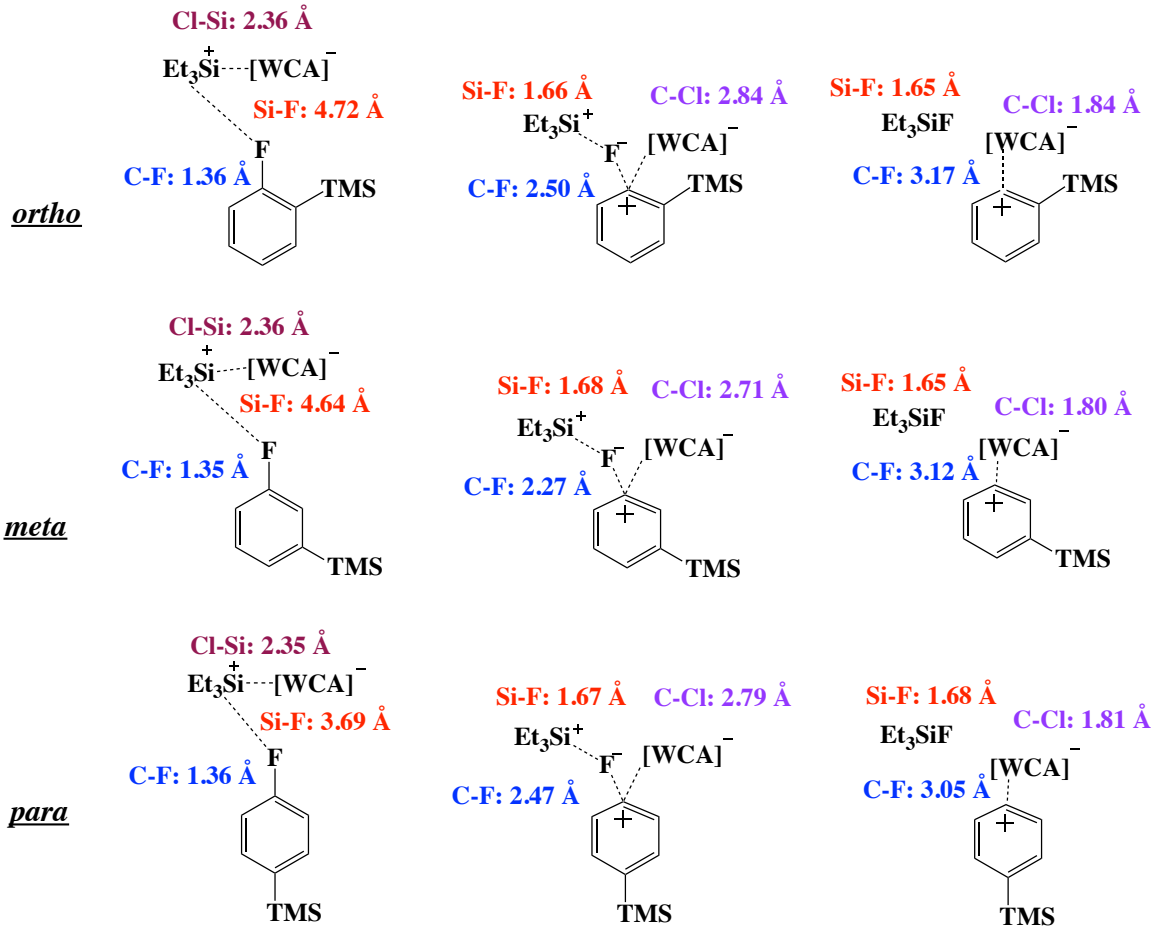


Figure 4.4: Optimized stationary structures for the fluoride abstraction reaction step for A) *ortho*-TMS-; B) *meta*-TMS-; c) *para*-TMS-fluorobenzene using ωB97X-D/6-31G* calculations with the IEF-PCM solvent model for *o*-dichlorobenzene.

As shown in Figure 4.5, the energetics for *ortho*, *meta* and *para*-TMS aryl fluoride reactions are significantly different from each other. For the *ortho*-TMS functionalized fluorobenzene substrate, the free energy barrier ($\Delta G^{\ddagger}$) was predicted to be 13.8 kcal/mol. For *meta*-TMS, and *para*-TMS substituted fluorobenzene substrates, the activation barriers were predicted to be 17.6 and 23.3 kcal/mol, which were 3.8 and 9.5 kcal/mol higher than that of *ortho*-TMS-fluorobenzene. As shown in Table 4.1, the reaction barrier energies

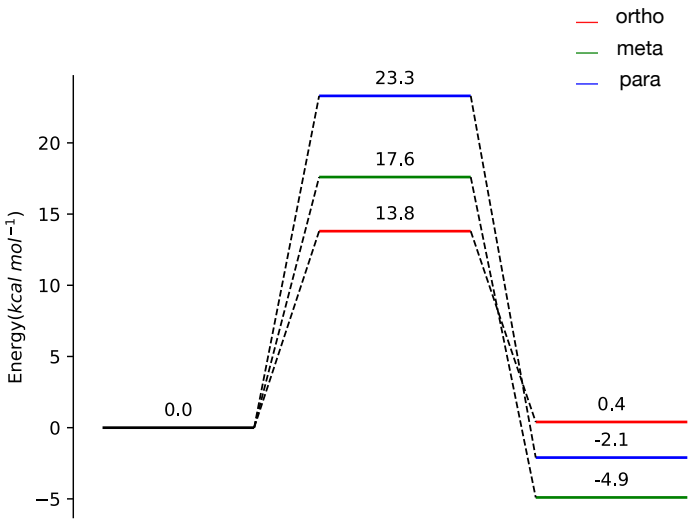


Figure 4.5: Reaction energetics for the fluoride abstraction reaction of *ortho*-, *meta*-, and *para*-TMS substrates as predicted from ωB97X-D/6-31G* calculations with the IEF-PCM solvent model for *o*-dichlorobenzene.

($\Delta E^{\ddagger}$) for *ortho*, *meta* and *para* substrates also followed the same trend as $\Delta G^{\ddagger}$. *In an excellent agreement with the experimental findings, our results suggested the ortho-TMS substrate is most likely to undergo a fluoride abstraction at the experimental condition (xx K), while the generation of aryl cations from meta-TMS and ortho-TMS substrates are much less likely to occur.*

Very interestingly, the computed reaction free energies displayed a very different trend. The reaction free energy of the *ortho*-TMS-fluorobenzene substrate was predicted to be 0.4 kcal/mol, which was higher than that of *para*-TMS fluorobenzene (-2.1 kcal/mol), and *meta*-TMS fluorobenzene (-4.9 kcal/mol). *Nearly zero reaction energy for the ortho-TMS substrate avoids an over-stabilization of the activated intermediate (INT_1) and thus facilitates the subsequent C–H insertion.*

The role of the TMS group was further investigated by analyzing charge populations of the optimized reactant, transition state, and INT_1 structures. As shown in Table 4.2, a larger positive charge on the TMS group for the TS structure would lead to a less charged and thus

Table 4.1: Computed activation barriers ($\Delta E^{\ddagger}$ and $\Delta G^{\ddagger}$) and reaction energies (ΔE and ΔG) for *ortho*, *meta*, and *para*-TMS substituted substrates during the fluoride abstraction at ωB97X-D/6-31G*/IEF-PCM level of theory. Also listed are results with the def2-TZVP basis set. All values are in kcal/mol.

Site	ωB97X-D/6-31G*				ωB97X-D/def2-TZVP	
	$\Delta E^{\ddagger}$	ΔE	$\Delta G^{\ddagger}$	ΔG	$\Delta E^{\ddagger}$	ΔE
ortho	15.6	-1.9	13.8	0.4	17.1	-3.0
meta	22.0	-4.3	17.6	-4.9	22.5	-6.4
para	28.4	1.2	23.3	-2.1	30.6	-0.6

Table 4.2: Computed ESP charges on the TMS and aryl cation fragments of reactant, TS and INT_1 structures during the fluoride abstraction reaction using ωB97X-D/6-31G* level of theory in *o*-dichlorobenzene using C-PCM solvent model.

Site	**TMS fragment**			**Aryl cation fragment**		
	Reactant	**TS**	**INT_1**	**Reactant**	**TS**	**INT_1**
ortho	0.321	0.453	0.389	-0.031	0.369	0.253
meta	0.236	0.383	0.316	0.013	0.476	0.042
para	0.215	0.360	0.266	-0.011	0.418	0.235

Table 4.3: Computed ESP charges on the WCA and Et_3SiF fragments of reactant, TS and INT_1 structures during the fluoride abstraction reaction using ωB97X-D/6-31G* level of theory in *o*-dichlorobenzene using C-PCM solvent model.

Site	**WCA fragment**			**Et_3SiF fragment**		
	Reactant	**TS**	**INT_1**	**Reactant**	**TS**	**INT_1**
ortho	-0.635	-0.882	-0.558	0.345	0.059	-0.084
meta	-0.691	-0.908	-0.381	0.442	0.049	0.022
para	-0.679	-0.872	-0.501	0.475	0.095	0.001

more stabilized aryl cation. As shown in Table 4.2, the TMS fragment stabilizes the TS structure the most as compared to reactant and INT_1 structures in all three (*ortho*, *meta* and *para*) case during the fluoride abstraction. In the transition state structure, the *ortho*-TMS group has the highest positive charge (+0.453) compared to *meta* (+0.383), and *para* TMS (+0.360) groups. Moreover, in the transition state structure, the ESP charge of the aryl cation fragment of *ortho*-substituted TMS fluorobenzene substrate was the least (+0.369) positive compared to aryl cation of the *meta*-TMS (+0.476) and *para*-TMS(+0.418) substrates. This way, these ESP charges indicate the degree of charge-transfer stabilization of aryl cation in the TS structure, which underlies the trend in experimental reactivities and computed activation barriers. This will be analyzed in more detail later through ALMO analysis.

In order to understand the role of the WCA anion, a similar charge population analysis was carried out on the fragment. In Table 4.3, the ESP charges on the WCA fragment were listed for the fluoride abstraction reaction with *ortho*, *meta*, and *para*-TMS substituted fluorobenzene substrates. From reactant to INT_1 state, the WCA anion loses more electrons (-0.310 = -0.381-(-0.691)) to the *meta*-TMS-Ar fragment than *ortho*-TMS-Ar (-0.077 = -0.558-(-0.635)) and *para*-TMS-Ar (-0.178 = -0.501-(-0.679)) fragments. Based on the above ESP charges, it is clear that the charge transfer occurred from anionic WCA fragment

to the positively charged TMS-substituted aryl fragment. This charge transfer effect was further explored with help of EDA analysis in the last part of the results section.

Moreover, these results are in agreement with the trend in the computed reaction free energy with all three TMS-substituted fluorobenzenes. As the computed reaction free energy is slightly positive for the *ortho*-TMS functionalized fluorobenzene substrate, the INT_1 can be considered activated complex for the subsequent C–H insertion step on the potential energy surface.

4.3.2 Mechanism of C–H Insertion

As the cross-coupling reaction between INT_1 and benzene molecules has been previously reported by Li and co-workers (158), in this study, we focused on the cross-coupling reaction involving INT_1 and cyclohexane. This C–H insertion resulted in an overall 41% yield of cyclohexyl benzene as determined by GC-FID using nonane as an internal standard (153).

Table 4.4: Computed reaction barrier energies ($\Delta G^{\ddagger}$) and reaction free energies (ΔG) for C–H insertion along the cyclohexyl transfer (M_1) and hydride transfer (M_2) pathways from ωB97X-D/6-31G* calculations in *o*-dichlorobenzene using the IEF-PCM solvent model.

Functional		**INT_1**	**TS_2**	**INT_2**	**TS_3**	**INT_3**	**TS_4**	**P**
ωB97X-D	M_1	0.0	15.2	-25.3	-14.6	-45.6	-29.1	-38.7
	M_2	0.0	15.2	-25.2	-17.5	-45.1	-29.7	-38.7
M06-2X	M_1	0.0	14.2	-22.0	-13.6	-46.4	-30.1	-40.1
	M_2	0.0	15.7	-22.0	-16.7	-45.6	-30.0	-40.8

The reaction between INT_1 and cyclohexane was initiated as the cyclohexane (especially one of its C–H bonds) drew closer, thus disrupting the weak electrostatic interaction between [WCA] and *ortho*-TMS functionalized aryl cation of the Int_1. This led to the insertion of the aryl cation into cyclohexyl-H (C–H) bond to obtain INT_2 *via* TS_2. From

INT_1 to INT_2M_1, the cyclohexyl C–H bond length increased from 1.099 to 2.127 Å . The computed activation barrier for C–H insertion step is 15.2 kcal/mol. Not surprisingly, the C–H insertion reaction was found to be highly exothermic, with an energy difference of -25.2 kcal/mol between INT_1 and INT_2, due to the formation of a carbon-carbon bond between cyclohexyl and aryl moieties (C–C bond length = 1.595 Å). With a bond length is slightly longer than typical C–C single bonds, INT_2 remains reactive and would be further stabilized in subsequent reactions.

Following INT_2, the reaction could have two probable pathways: 1) cyclohexyl transfer (M_1); and hydride transfer (M_2). While the activation barrier of the hydride transfer reaction (INT_2 to TS_3M_1) was shown in Figure 4.7 to be 7.7 kcal/mol, that of the cyclohexyl transfer was found to be slightly higher with a value of 10.6 kcal/mol (see Figure 4.6). Therefore, the hydride transfer pathway would be the minimum energy reaction pathway in this reaction mechanism, but the cyclohexyl transfer can also occur due to its lower reaction barrier. The last step of the catalytic cycle involves the elimination of the sacrificial TMS group to regain aromaticity of the cationic phenyl ring with an activation barrier of 16.5 and 15.4 kcal/mol, respectively. Overall, the free energy of the cross-coupling C–H insertion reaction was -38.7 kcal/mol.

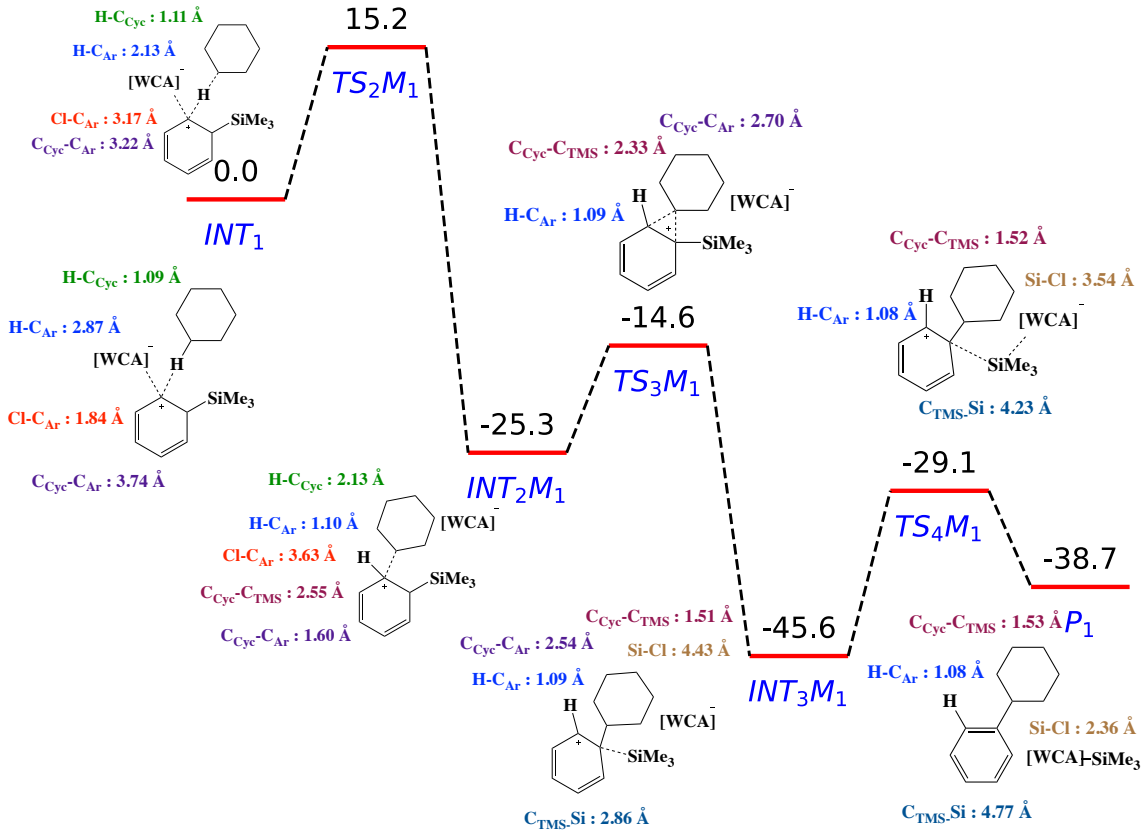


Figure 4.6: Proposed reaction mechanism for C–H insertion (cyclohexyl transfer pathway) from ωB97X-D/6-31G* calculations in *o*-dichlorobenzene using the IEF-PCM solvent model.

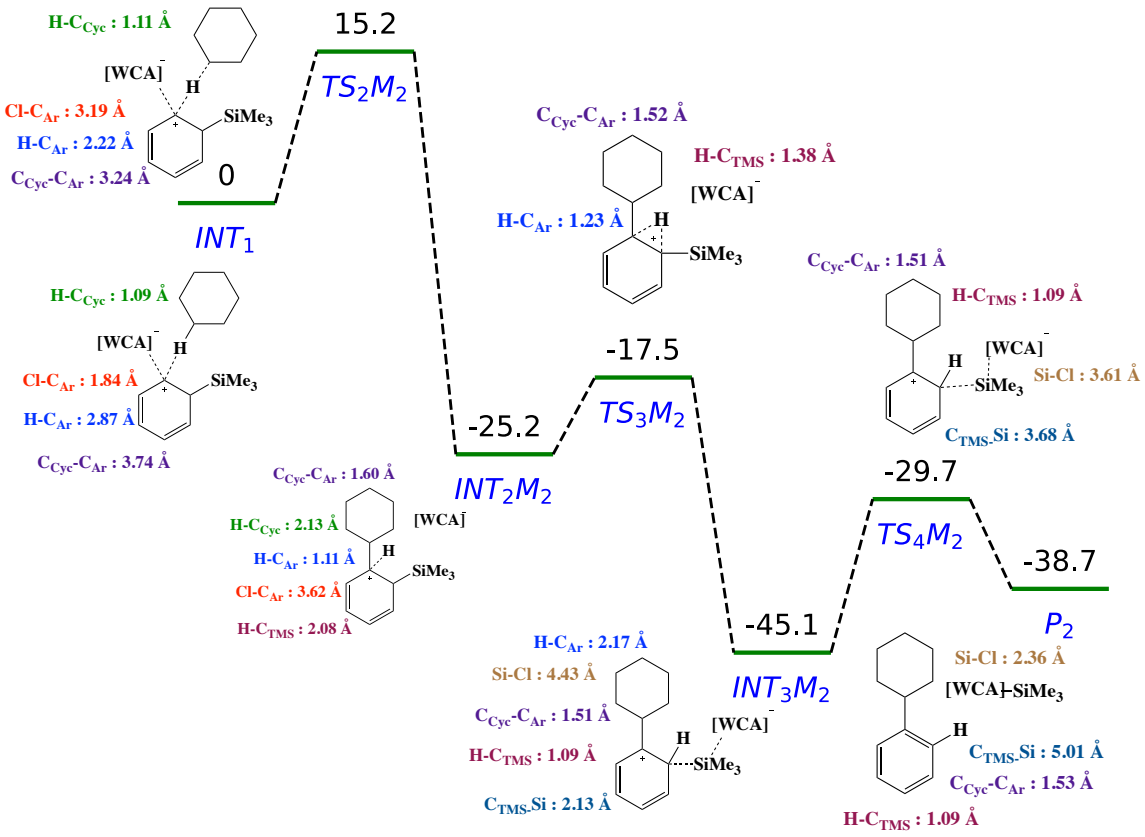


Figure 4.7: Proposed reaction mechanism for C–H insertion (hydride transfer pathway) from ωB97X-D/6-31G* calculations in *o*-dichlorobenzene using the IEF-PCM solvent model.

4.3.3 Energy Decomposition Analysis

4.3.3.1 ALMO-EDA for the Covalent Linkage

In order to gain a clearer picture in the β-silicon functionalized aryl cation system, the overall energy of the system is decomposed into POL and CT terms.

Table 4.5: Polarization (POL) and charge transfer (CT) interaction energies between the TMS group and aryl cation fragment at *ortho*, *meta* and *para*-TMS functionalized fluorobenzne during transition state as calculated using the ωB97X-D/6-31G* level of theory in gas phase. All values are in kcal/mol.

Site	Polarization(POL)	Charge Transfer(CT)
ortho	2.8	14.2
meta	1.9	11.1
para	1.4	9.2

The effect of *ortho*, *meta*, and *para*-TMS substituents on the aryl cation substrate obtained from ALMO-EDA analysis are listed in Table 4.5. In the case of *ortho*-TMS and *meta*-TMS substituent, the net polarization effect was computed to be 2.8 and 1.9 kcal/mol respectively. For *para*-TMS substituent, the polarization term was determined to be about 1.4 kcal/mol. *Ortho*-TMS substituent stabilized the overall energy by approximately 14.2 kcal/mol *via* charge transfer, whereas *meta*-TMS and *para*-TMS substituent only reduces the energy by 11.1 kcal/mol, and 9.2 kcal/mol, respectively (see Figure 4.8).

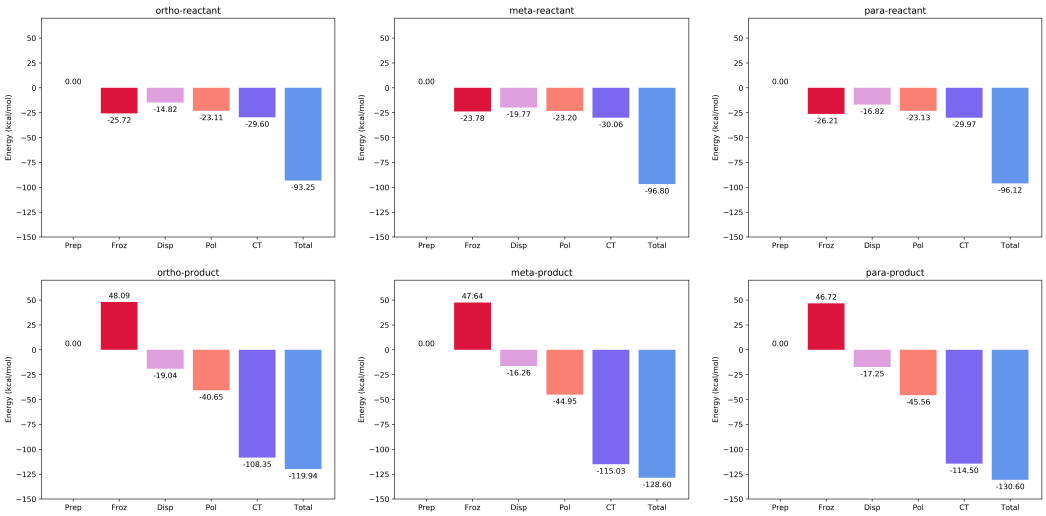
4.3.3.2 EDA for the Non-covalent Linkage

To evaluate the contribution of WCA anion to the stabilization of aryl-TMS cation at the INT_1 structure for the corresponding *ortho*, *meta*, and *para*-TMS substituted phenyl cation (see Figure 4.8). The total energy of the two fragment system was further decomposed into

Table 4.6: Computed energies of fragment AR-TMS (A) , WCA (B), complex (AB) and their interaction energy (in Hartree and in kcal/mol respectively). All single point energies are obtained at ωB97X-D/6-31G* level in *o*-dichlorobenzene using the IEF-PCM solvent model.

Site		**A**	**B**	**AB**	**Interaction Energy**	
		(a.u.)	(a.u)	(a.u)	(a.u.)	(kcal/mol)
	Reactants	-1266.9065	-5374.8659	-6641.8220	-0.0496	-31.1
ortho	TS	-1266.9058	-5374.8683	-6641.7971	-0.0230	-14.4
	Products	-1266.8675	-5374.8633	-6641.8251	-0.0943	-59.2
	Reactants	-1266.8976	-5374.8654	-6641.8197	-0.0566	-35.5
meta	TS	-1266.8887	-5374.8684	-6641.7846	-0.0275	-17.2
	Products	-1266.8556	-5374.8638	-6641.8266	-0.1070	-67.3
	Reactants	-1266.9037	-5374.8658	-6641.8227	-0.0532	-33.4
para	TS	-1266.8878	-5374.8683	-6641.7774	-0.0213	-13.4
	Products	-1266.8469	-5374.8639	-6641.8207	-0.1100	-69.0

Figure 4.8: EDA calculations on aryl-TMS cation and WCA anion fragments at INT_1 geometry during fluoride abstraction using ωB97X-D/6-31G* in gas phase.



five energy terms: preparation (Prep), frozen (Frz), dispersion (Disp), polarization (POL), and charge transfer (CT). The negative values of the overall interaction energies correspond to the stabilization interaction between the two fragments, which are in agreement with the computed reaction free energies (as listed in Table 4.6). The WCA anion stabilized the *meta* and *para*-TMS substituted aryl cation by -115.0 kcal/mol and -114.5 kcal/mol. In contrast, the WCA anion stabilized the *ortho*-TMS aryl cation only by -108.5 kcal/mol *via* charge transfer.

4.4 Conclusions

DFT calculations were performed to understand the mechanistic details of arylation for both fluoride abstraction of TMS-substituted fluorobenzene and subsequent C–H insertion with cyclohexane.

Key conclusions were as follows:

- Computed activation barriers for fluoride abstraction of TMS substituted fluorobenzenes followed the trend (*ortho* $<$ *meta* $<$ *para*). The results were in agreement with the experimental findings as reported by Nelson and co-workers.

- The hyperconjugation effect of the β-silicon group was crucial in stabilizing the aryl cation at the transition state structure during the fluoride abstraction step *via* charge transfer. The ALMO-EDA analysis showed that the stabilization effect of β-TMS substituent arose from a charge transfer interaction energy of 14.2 kcal/mol.

- ESP charge population and EDA analyses were performed at the INT_1 structure. The results suggested that the *meta*-substituted TMS aryl cation is subjected to the largest charge transfer interaction energy (-115.0 kcal/mol) from anionic WCA catalyst to the aryl cation, whereas the *ortho*-substituted TMS aryl cation exhibited a smaller charge transfer(-108.5 kcal/mol) from the WCA anion to aryl cation.

- The cross-coupling between aryl cation and cyclohexane was found to be thermodynamically favorable (ΔG = -38.5 kcal/mol). The C-C bond formation and re-aromatization were two major driving forces for the forward reaction.

- Post C–H insertion step, the reaction can proceed in two pathways: A) cyclohexyl transfer; B) hydride transfer. The latter has a slightly lower ($\sim$ 3 kcal/mol) activation barrier.

Reference List

[1] C. Hansch, A. Leo, and R.W. Taft, "A survey of Hammett substituent constants and resonance and field parameters", *Chem. Rev.* **91**(2), pp. 165 (1991).

[2] L.P. Hammett, "The Effect of Structure upon the Reactions of Organic Compounds. Benzene Derivatives", *J. Am. Chem. Soc.* **59**(1), pp. 96 (1937).

[3] R.G.W. Jinadasa, Y. Fang, S. Kumar, A.J. Osinski, X. Jiang, C.J. Ziegler, K.M. Kadish, and H. Wang, "β-Functionalized Push–Pull *opp* -Dibenzoporphyrins", *J. Org. Chem.* **80**(24), pp. 12076 (2015).

[4] T. Hirano, H. Nagai, T. Matsuhashi, Y. Hasumi, S. Iwano, K. Ito, S. Maki, H. Niwa, and V.R. Viviani, "Spectroscopic studies of the color modulation mechanism of firefly (beetle) bioluminescence with amino-analogs of luciferin and oxyluciferin", *Photochem. Photobiol. Sci.* **11**(8), pp. 1281 (2012).

[5] C.G. Derick, "Polarity of elements and radicals measured in terms of a logarithmic function of the ionization constant.", *J. Am. Chem. Soc.* **33**(7), pp. 1152 (1911).

[6] B. Brönsted and K. Pedersen, "Die katalytische zersetzung des nitramids und ihre physikalisch-chemische bedeutung", *Phys. Chem.* , p. 185 (1924).

[7] T.M. Krygowski and B.T. Stepień, "Sigma- and pi-electron delocalization: Focus on substituent effects", *Chem. Rev.* **105**(10), pp. 3482 (2005).

[8] L.P. Hammett, "Some relations between reaction rates and equilibrium constants", *Chem. Rev.* **17**(1), pp. 125 (1935).

[9] O. Exner and S. Böhm, "Bond angles and bond lengths in monosubstituted benzene and ethene derivatives: a comparison of computational and crystallographic results", *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **58**(5), pp. 877 (2002).

[10] R. Bonaccorsi, A. Pullman, E. Scrocco, and J. Tomasi, "The molecular electrostatic potentials for the nucleic acid bases: Adenine, thymine, and cytosine", *Theor. Chim. Acta.* **24**(1), pp. 51 (1972).

[11] E. Scrocco and J. Tomasi, "Electronic molecular structure, reactivity and intermolecular forces: an heuristic interpretation by means of electrostatic molecular potentials", *Adv. Quantum Chem.* **11**, pp. 115 (1978).

[12] P. Politzer, K.C. Daiker, and R.A. Donnelly, "Molecular electrostatic potentials: A new approach to the study of the metabolic and carcinogenic activities of hydrocarbons", *Cancer Lett.* **2**(1), pp. 17 (1976).

[13] G.S. Remya and C.H. Suresh, "Quantification and classification of substituent effects in organic chemistry: a theoretical molecular electrostatic potential study", *Phys. Chem. Chem. Phys.* **18**(30), pp. 20615 (2016).

[14] S. Böhm and O. Exner, "Basicity of carboxylic acids: resonance in the cation and substituent effects", *New J. Chem.* **29**(2), pp. 336 (2005).

[15] T.A. Albright, J.K. Burdett, and M.H. Whangbo, *Orbital interactions in chemistry*, Wiley Hoboken, New Jersey second edition edition (2013).

[16] E. Runge and E.K.U. Gross, "Density-functional theory for time-dependent systems", *Phys. Rev. Lett.* **52**(12), pp. 997 (1984).

[17] K. Burke, J. Werschnik, and E.K.U. Gross, "Time-dependent density functional theory: past, present, and future", *J. Chem. Phys.* **123**(6), pp. 062206 (2005).

[18] Y. Mao, M. Head-Gordon, and Y. Shao, "Unraveling substituent effects on frontier orbitals of conjugated molecules using an absolutely localized molecular orbital based analysis", *Chem. Sci.* **9**(45), pp. 8598 (2018).

[19] P. Naumov, Y. Ozawa, K. Ohkubo, and S. Fukuzumi, "Structure and spectroscopy of oxyluciferin, the light emitter of the firefly bioluminescence", *J. Am. Chem. Soc.* **131**(32), pp. 11590 (2009).

[20] C. Ran, X. Xu, S.B. Raymond, B.J. Ferrara, K. Neal, J. Brian, Z. Medarova, and A. Moore, "Design, synthesis, and testing of difluoroboron derivatized curcumins as near infrared probes for in vivo detection of amyloid-β deposits", *J. Am. Chem. Soc.* **131**, pp. 15257 (2009).

[21] Y. Shao, Z. Gan, E. Epifanovsky, A.T. Gilbert, M. Wormit, J. Kussmann, A.W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P.R. Horn, L.D. Jacobson, I. Kaliman, R.Z. Khaliullin, T. Kuś, A. Landau, J. Liu, E.I. Proynov, Y.M. Rhee, R.M. Richard, M.A. Rohrdanz, R.P. Steele, E.J. Sundstrom, H.L. Woodcock, P.M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G.J.O. Beran, Y.A. Bernard, E. Berquist, K. Brandhorst, K.B. Bravaya, S.T. Brown, D. Casanova, C.M. Chang, Y. Chen, S.H. Chien, K.D. Closser, D.L. Crittenden, M. Diedenhofen, R.A. DiStasio, H. Do, A.D. Dutoi, R.G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M.W. Hanson-Heine, P.H. Harbach, A.W. Hauser, E.G. Hohenstein, Z.C. Holden, T.C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R.A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C.M. Krauter, K.U. Lao, A.D. Laurent, K.V. Lawler, S.V. Levchenko, C.Y. Lin, F. Liu, E. Livshits, R.C. Lochan, A. Luenser, P. Manohar, S.F. Manzer, S.P. Mao, N. Mardirossian, A.V. Marenich, S.A. Maurer, N.J. Mayhall, E. Neuscamman, C.M. Oana, R. Olivares-Amaya, D.P. O'Neill, J.A. Parkhill, T.M. Perrine, R. Peverati, A. Prociuk, D.R. Rehn, E. Rosta, N.J. Russ, S.M. Sharada, S. Sharma, D.W. Small, A. Sodt, T. Stein, D. Stuck, Y.C. Su, A.J. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M.A. Watson, J. Wenzel, A. White, C.F. Williams, J. Yang, S. Yeganeh, S.R. Yost, Z.Q. You, I.Y. Zhang, X. Zhang, Y. Zhao, B.R. Brooks, G.K. Chan, D.M. Chipman, C.J. Cramer, W.A. Goddard, M.S. Gordon, W.J. Hehre, A. Klamt, H.F.

Schaefer, M.W. Schmidt, C.D. Sherrill, D.G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A.T. Bell, N.A. Besley, J.D. Chai, A. Dreuw, B.D. Dunietz, T.R. Furlani, S.R. Gwaltney, C.P. Hsu, Y. Jung, J. Kong, D.S. Lambrecht, W. Liang, C. Ochsenfeld, V.A. Rassolov, L.V. Slipchenko, J.E. Subotnik, T. Van Voorhis, J.M. Herbert, A.I. Krylov, P.M. Gill, and M. Head-Gordon, "Advances in Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package", *Mol. Phys.* **113**(2), pp. 184 (2015).

[22] W. Kohn and L.J. Sham, "Self-consistent equations including exchange and correlation effects", *Phys. Rev.* **140**(4A), pp. A1133 (1965).

[23] D.S. Sholl and J.A. Steckel, *Density functional theory: a practical introduction*, Wiley Hoboken, N.J. (2009).

[24] Y. Zhao and D.G. Truhlar, "The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals", *Theor. Chem. Acc.* **120**(1-3), pp. 215–241 (2008).

[25] R.Z. Khaliullin, E.A. Cobar, R.C. Lochan, A.T. Bell, and M. Head-Gordon, "Unravelling the origin of intermolecular interactions using absolutely localized molecular orbitals", *J. Phys. Chem. A* **111**(36), pp. 8753 (2007).

[26] P.R. Horn, Y. Mao, and M. Head-Gordon, "Probing non-covalent interactions with a second generation energy decomposition analysis using absolutely localized molecular orbitals", *Phys. Chem. Chem. Phys.* **18**(33), pp. 23067 (2016).

[27] Y. Mao, P.R. Horn, and M. Head-Gordon, "Energy decomposition analysis in an adiabatic picture", *Phys. Chem. Chem. Phys.* **19**(8), pp. 5944 (2017).

[28] Y. Mao, M. Loipersberger, P.R. Horn, A. Das, O. Demerdash, D.S. Levine, S.P. Veccham, T. Head-Gordon, and M. Head-Gordon, "From intermolecular interaction energies and observable shifts to component contributions and back again: A tale of variational energy decomposition analysis", *Annu. Rev. Phys. Chem.* **72**(1), pp. 641 (2021).

[29] Z. Pei, Q. Ou, Y. Mao, J. Yang, A.d.l. Lande, F. Plasser, W. Liang, Z. Shuai, and Y. Shao, "Elucidating the electronic structure of a delayed fluorescence emitter via orbital interactions, excitation energy components, charge-transfer numbers, and vibrational reorganization energies", *J. Phys. Chem. Lett.* **12**(11), pp. 2712 (2021).

[30] B. Bitler and W. McElroy, "The preparation and properties of crystalline firefly luciferin", *Arch. Biochem. Biophys.* **72**(2), pp. 358 (1957).

[31] E.H. White, F. McCapra, G.F. Field, and W.D. McElroy, "The structure and synthesis of firefly luciferin", *J. Am. Chem. Soc.* **83**(10), pp. 2402 (1961).

[32] E.H. White, F. McCapra, and G.F. Field, "The structure and synthesis of firefly luciferin", *J. Am. Chem. Soc.* **85**(3), pp. 337 (1963).

[33] B.R. Branchini, M.H. Murtiashaw, R.A. Magyar, N.C. Portier, M.C. Ruggiero, and J.G. Stroh, "Yellow-green and red firefly bioluminescence from 5,5-dimethyloxyluciferin", *J. Am. Chem. Soc.* **124**(10), pp. 2112 (2002).

[34] B.R. Branchini, T.L. Southworth, M.H. Murtiashaw, R.A. Magyar, S.A. Gonzalez, M.C. Ruggiero, and J.G. Stroh, "An alternative mechanism of bioluminescence color determination in firefly luciferase", *Biochemistry* **43**(23), pp. 7255 (2004).

[35] F. Liu, Y. Liu, L. De Vico, and R. Lindh, "Theoretical study of the chemiluminescent decomposition of dioxetanone", *J. Am. Chem. Soc.* **131**(17), pp. 6181 (2009).

[36] I. Navizet, Y.J. Liu, N. Ferré, H.Y. Xiao, W.H. Fang, and R. Lindh, "Color-tuning mechanism of firefly investigated by multi-configurational perturbation method", *J. Am. Chem. Soc.* **132**(2), pp. 706 (2010).

[37] S.F. Chen, Y.J. Liu, I. Navizet, N. Ferré, W.H. Fang, and R. Lindh, "Systematic theoretical investigation on the light emitter of firefly", *J. Chem. Theory Comput.* **7**(3), pp. 798 (2011).

[38] L. Yue, Y.J. Liu, and W.H. Fang, "Mechanistic insight into the chemiluminescent decomposition of firefly dioxetanone", *J. Am. Chem. Soc.* **134**(28), pp. 11632 (2012).

[39] Y.Y. Cheng and Y.J. Liu, "What exactly is the light emitter of a firefly?", *J. Chem. Theory Comput.* **11**(11), pp. 5360 (2015).

[40] H. Takakura, R. Kojima, T. Ozawa, T. Nagano, and Y. Urano, "Development of 5*prime*- and 7*ı*-substituted luciferin analogues as acid-tolerant substrates of firefly luciferase", *ChemBioChem* **13**(10), pp. 1424 (2012).

[41] R.C. Steinhardt, J.M. O'Neill, C.M. Rathbun, D.C. McCutcheon, M.A. Paley, and J.A. Prescher, "Design and synthesis of an alkynyl luciferin analogue for bioluminescence imaging", *Chem. Eur. J.* **22**(11), pp. 3671 (2016).

[42] R.C. Steinhardt, C.M. Rathbun, B.T. Krull, J.M. Yu, Y. Yang, B.D. Nguyen, J. Kwon, D.C. McCutcheon, K.A. Jones, F. Furche, and J.A. Prescher, "Brominated luciferins are versatile bioluminescent probes", *ChemBioChem* **18**(1), pp. 96 (2017).

[43] Y. Ikeda, T. Saitoh, K. Niwa, T. Nakajima, N. Kitada, S.A. Maki, M. Sato, D. Citterio, S. Nishiyama, and K. Suzuki, "An allylated firefly luciferin analogue with luciferase specific response in living cells", *Chem. Commun.* **54**(14), pp. 1774 (2018).

[44] A.P. Jathoul, H. Grounds, J.C. Anderson, and M.A. Pule, "A dual-color far-red to near-infrared firefly luciferin analogue designed for multiparametric bioluminescence imaging", *Angew. Chem. Int. Ed.* **53**(48), pp. 13059 (2014).

[45] Z. Yao, B.S. Zhang, R.C. Steinhardt, J.H. Mills, and J.A. Prescher, "Multicomponent bioluminescence imaging with a π-extended luciferin", *J. Am. Chem. Soc.* **142**(33), pp. 14080 (2020).

[46] B.R. Branchini, "Chemical synthesis of firefly luciferin analogs and inhibitors", *Meth Enzymol.* **305**, pp. 188 (2000).

[47] E.H. White, H. Wörther, H.H. Seliger, and W.D. McElroy, "Amino analogs of firefly luciferin and biological activity thereof", *J. Am. Chem. Soc.* **88**(9), pp. 2015 (1966).

[48] S. Iwano, R. Obata, C. Miura, M. Kiyama, K. Hama, M. Nakamura, Y. Amano, S. Kojima, T. Hirano, S. Maki, and H. Niwa, "Development of simple firefly luciferin analogs emitting blue, green, red, and near-infrared biological window light", *Tetrahedron* **69**(19), pp. 3847 (2013).

[49] M. Kiyama, S. Iwano, S. Otsuka, S.W. Lu, R. Obata, A. Miyawaki, T. Hirano, and S.A. Maki, "Quantum yield improvement of red-light-emitting firefly luciferin analogues for *in vivo* bioluminescence imaging", *Tetrahedron* **74**(6), pp. 652 (2018).

[50] S. Iwano, M. Sugiyama, H. Hama, A. Watakabe, N. Hasegawa, T. Kuchimaru, K.Z. Tanaka, M. Takahashi, Y. Ishida, J. Hata, S. Shimozono, K. Namiki, T. Fukano, M. Kiyama, H. Okano, S. Kizaka-Kondoh, T.J. McHugh, T. Yamamori, H. Hioki, S. Maki, and A. Miyawaki, "Single-cell bioluminescence imaging of deep tissue in freely moving animals", *Science* **359**(6378), pp. 935 (2018).

[51] W. Wu, J. Su, C. Tang, H. Bai, Z. Ma, T. Zhang, Z. Yuan, Z. Li, W. Zhou, H. Zhang, Z. Liu, Y. Wang, Y. Zhou, L. Du, L. Gu, and M. Li, "CybLuc: an effective aminoluciferin derivative for deep bioluminescence imaging", *Anal. Chem.* **89**(9), pp. 4808 (2017).

[52] M.P. Hall, C.C. Woodroofe, M.G. Wood, I. Que, M. van't Root, Y. Ridwan, C. Shi, T.A. Kirkland, L.P. Encell, K.V. Wood, C. Löwik, and L. Mezzanotte, "Click beetle luciferase mutant and near infrared naphthyl-luciferins for improved bioluminescence imaging", *Nat. Commun.* **9**(1), pp. 132 (2018).

[53] N. Kitada, T. Saitoh, Y. Ikeda, S. Iwano, R. Obata, H. Niwa, T. Hirano, A. Miyawaki, K. Suzuki, S. Nishiyama, and S.A. Maki, "Toward bioluminescence in the near-infrared region: tuning the emission wavelength of firefly luciferin analogues by allyl substitution", *Tetrahedron Lett.* **59**(12), pp. 1087 (2018).

[54] T. Nakatsu, S. Ichiyama, J. Hiratake, A. Saldanha, N. Kobashi, K. Sakata, and H. Kato, "Structural basis for the spectral difference in luciferase bioluminescence", *Nature* **440**(7082), pp. 372 (2006).

[55] T. Pozzo, F. Akter, Y. Nomura, A.Y. Louie, and Y. Yokobayashi, "Firefly luciferase mutant with enhanced activity and thermostability", *ACS Omega* **3**(3), pp. 2628 (2018).

[56] B.R. Branchini, T.L. Southworth, N.F. Khattak, E. Michelini, and A. Roda, "Red- and green-emitting firefly luciferase mutants for bioluminescent reporter applications", *Anal. Biochem.* **345**(1), pp. 140 (2005).

[57] B.R. Branchini, D.M. Ablamsky, M.H. Murtiashaw, L. Uzasci, H. Fraga, and T.L. Southworth, "Thermostable red and green light-producing firefly luciferase mutants for bioluminescent reporter applications", *Anal. Biochem.* **361**(2), pp. 253 (2007).

[58] K.R. Zinn, T.R. Chaudhuri, A.A. Szafran, D. O'Quinn, C. Weaver, K. Dugger, D. Lamar, R.A. Kesterson, X. Wang, and S.J. Frank, "Noninvasive bioluminescence imaging in small animals", *ILAR J.* **49**(1), pp. 103 (2008).

[59] O.A. Gandelman, V.L. Church, C.A. Moore, G. Kiddle, C.A. Carne, S. Parmar, H. Jalal, L.C. Tisi, and J.A.H. Murray, "Novel bioluminescent quantitative detection of nucleic acid amplification in real-time", *PLoS ONE* **5**(11), pp. e14155 (2010).

[60] A. Jathoul, E. Law, O. Gandelman, M. Pule, L. Tisi, and J. Murray, "Development of a pH-tolerant thermostable photinus pyralis luciferase for brighter *in vivo* imaging", In D. Lapota, editor, *Bioluminescence:Recent Advances in Oceanic Measurements and Laboratory Applications*. InTech (2012).

[61] B.R. Branchini, T.L. Southworth, D.M. Fontaine, A.L. Davis, C.E. Behney, and M.H. Murtiashaw, "A photinus pyralis and luciola italica chimeric firefly luciferase produces enhanced bioluminescence", *Biochemistry* **53**(40), pp. 6287 (2014).

[62] E. Conti, N.P. Franks, and P. Brick, "Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes", *Structure* **4**(3), pp. 287 (1996).

[63] K.V. Wood, Y.A. Lam, and W.D. McElroy, "Introduction to beetle luciferases and their applications", *J. Biolumin. Chemilumin.* **4**(1), pp. 289 (1989).

[64] V.R. Viviani, E.J.H. Bechara, and Y. Ohmiya, "Cloning, sequence analysis, and expression of active *Phrixothrix* railroad-worms luciferases: relationship between bioluminescence spectra and primary structures", *Biochemistry* **38**(26), pp. 8271 (1999).

[65] K.A. Jones, W.B. Porterfield, C.M. Rathbun, D.C. McCutcheon, M.A. Paley, and J.A. Prescher, "Orthogonal luciferase–luciferin pairs for bioluminescence imaging", *J. Am. Chem. Soc.* **139**(6), pp. 2351 (2017).

[66] C.M. Rathbun, W.B. Porterfield, K.A. Jones, M.J. Sagoe, M.R. Reyes, C.T. Hua, and J.A. Prescher, "Parallel screening for rapid identification of orthogonal bioluminescent tools", *ACS Cent. Sci.* **3**(12), pp. 1254 (2017).

[67] Y.Y. Cheng, J. Zhu, and Y.J. Liu, "Theoretical tuning of the firefly bioluminescence spectra by the modification of oxyluciferin", *Chem. Phys. Lett.* **591**, pp. 156 (2014).

[68] Y.Y. Cheng and Y.J. Liu, "Theoretical development of near-infrared bioluminescent systems", *Chem. Eur. J.* **24**(37), pp. 9340 (2018).

[69] M. Zemmouche, C. García-Iriepa, and I. Navizet, "Light emission colour modulation study of oxyluciferin synthetic analogues *via* QM and QM/MM approaches", *Phys. Chem. Chem. Phys.* **22**(1), pp. 82 (2020).

[70] C.G. Min, Y. Leng, X.K. Yang, S.J. Huang, and A.M. Ren, "Systematic color tuning of a family of firefly oxyluciferin analogues suitable for OLED applications", *Dyes Pigm.* **126**, pp. 202 (2016).

[71] R.G. Parr and W. Yang, *Density-functional theory of atoms and molecules*, International series of monographs on chemistry. Oxford Univ. Press New York, NY (1994).

[72] K. Burke, "Perspective on density functional theory", *J. Chem. Phys.* **136**(15), pp. 150901 (2012).

[73] A.D. Becke, "Perspective: fifty years of density-functional theory in chemical physics", *J. Chem. Phys.* **140**(18), pp. 18A301 (2014).

[74] M.E. Casida, "Time-Dependent Density-Functional Response Theory for molecules", In *Recent Advances in Density Functional Methods Part I.* , p. 155. World Scientific (1995).

[75] A. Dreuw and M. Head-Gordon, "Single-reference ab initio methods for the calculation of excited states of large molecules", *Chem. Rev.* **105**(11), pp. 4009 (2005).

[76] M. Casida and M. Huix-Rotllant, "Progress in Time-Dependent Density-Functional Theory", *Ann. Rev. Phys. Chem.* **63**(1), pp. 287 (2012).

[77] J.D. Chai and M. Head-Gordon, "Long-Range corrected hybrid density functionals with damped atom-atom dispersion corrections", *Phys. Chem. Chem. Phys.* **10**(44), pp. 6615 (2008).

[78] A.D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior", *Phys. Rev. A* **38**(6), pp. 3098 (1988).

[79] A.D. Becke, "A new mixing of Hartree-Fock and local density-functional theories", *J. Chem. Phys.* **98**(2), pp. 1372 (1993).

[80] C. Lee, W. Yang, and R.G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density", *Phys. Rev. B* **37**(2), pp. 785 (1988).

[81] C. Adamo and V. Barone, "Toward reliable density functional methods without adjustable parameters: the PBE0 model", *J. Chem. Phys.* **110**(13), pp. 6158 (1999).

[82] T. Yanai, D.P. Tew, and N.C. Handy, "A new hybrid exchange–correlation functional using the Coulomb-Attenuating Method (CAM-B3LYP)", *Chem. Phys. Lett.* **393**(1-3), pp. 51 (2004).

[83] R. Krishnan, J.S. Binkley, R. Seeger, and J.A. Pople, "Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions", *J. Chem. Phys.* **72**(1), pp. 650 (1980).

[84] P.M. Gill, B.G. Johnson, and J.A. Pople, "A standard grid for density functional calculations", *Chem. Phys. Lett.* **209**(5-6), pp. 506 (1993).

[85] A.D. Becke, "A multicenter numerical integration scheme for polyatomic molecules", *J. Chem. Phys.* **88**(4), pp. 2547 (1988).

[86] T.H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen", *J. Chem. Phys.* **90**(2), pp. 1007 (1989).

[87] R.A. Kendall, T.H. Dunning, and R.J. Harrison, "Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions", *J. Chem. Phys.* **96**(9), pp. 6796 (1992).

[88] W.J. Hehre, R. Ditchfield, and J.A. Pople, "Self-consistent molecular orbital methods. XII. Further extensions of gaussian-type basis sets for use in molecular orbital studies of organic molecules", *J. Chem. Phys.* **56**(5), pp. 2257 (1972).

[89] P.C. Hariharan and J.A. Pople, "The influence of polarization functions on molecular orbital hydrogenation energies", *Theor. Chim. Acta* **28**(3), pp. 213 (1973).

[90] J. Tomasi, B. Mennucci, and R. Cammi, "Quantum mechanical continuum solvation models", *Chem. Rev.* **105**(8), pp. 2999 (2005).

[91] B. Mennucci, "Polarizable continuum model", *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2**(3), pp. 386 (2012).

[92] A.W. Lange and J.M. Herbert, "A smooth, nonsingular, and faithful discretization scheme for polarizable continuum models: the switching/gaussian approach", *J. Chem. Phys.* **133**(24), pp. 244111 (2010).

[93] R. Cammi and B. Mennucci, "Linear response theory for the polarizable continuum model", *J. Chem. Phys.* **110**(20), pp. 9877 (1999).

[94] M. Cossi and V. Barone, "Time-dependent density functional theory for molecules in liquid solutions", *J. Chem. Phys.* **115**(10), pp. 4708 (2001).

[95] R. Improta, V. Barone, G. Scalmani, and M.J. Frisch, "A state-specific polarizable continuum model time dependent density functional theory method for excited state calculations in solution", *J. Chem. Phys.* **125**(5), pp. 054103 (2006).

[96] D.K. Sharma, S.T. Adams, K.L. Liebmann, and S.C. Miller, "Rapid access to a broad range of 6'-substituted firefly luciferin analogues reveals surprising emitters and inhibitors", *Org. Lett.* **19**(21), pp. 5836 (2017).

[97] C.I. Song and Y.M. Rhee, "Development of force field parameters for oxyluciferin on its electronic ground and excited states", *Int. J. Quantum Chem.* **111**(15), pp. 4091 (2011).

[98] V. Satalkar, T.A. Rusmore, E. Phillips, X. Pan, E. Benassi, Q. Wu, C. Ran, and Y. Shao, "Computational modeling of curcumin-based fluorescent probe molecules", *Theor. Chem. Acc.* **138**(2), pp. 29 (2019).

[99] J. Hardy and D.J. Selkoe, "The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics", *Science* **297**, pp. 353 (2002).

[100] D.J. Selkoe, "Resolving controversies on the path to Alzheimer's therapeutics", *Nat. Med.* **17**(9), pp. 1060 (2011).

[101] L.f. Lue, Y.m. Kuo, A.E. Roher, L. Brachova, Y. Shen, L. Sue, T. Beach, J.H. Kurth, R.E. Rydel, and J. Rogers, "Soluble Amyloid beta Peptide Concentration as a Predictor of Synaptic Change in Alzheimer's Disease", *Am. J. Pathol.* **155**, pp. 853 (1999).

[102] C.A. Mclean, R.A. Cherny, F.W. Fraser, S.J. Fuller, M.J. Smith, K. Beyreuther, A.I. Bush, and C.L. Masters, "Soluble pool of Aβ amyloid as a determinant of severity of neurodegeneration in Alzheimer's disease", *Ann. Neurol.* **46**, pp. 860 (1999).

[103] E. Giacobini and G. Gold, "Alzheimer disease therapy – moving from amyloid Aβ to tau", *Nat. Rev. Neurol.* **9**, pp. 677 (2013).

[104] D. Pratico, "Oxidative stress hypothesis in Alzheimer's disease: a reappraisal", *Trends Pharmacol. Sci.* **29**, pp. 609 (2008).

[105] H. Xie, S. Hou, J. Jiang, M. Sekutowicz, J. Kelly, and B.J. Bacskai, "Rapid cell death is preceded by amyloid plaque-mediated oxidative stress", *Proc. Nat. Acad. Sci. USA* **110**, pp. 7904 (2013).

[106] X. Wang, W. Wang, L. Li, G. Perry, H.g. Lee, and X. Zhu, "Oxidative stress and mitochondrial dysfunction in Alzheimer's disease", *Biochim. Biophys. Acta* **1842**, pp. 1240 (2014).

[107] G.H. Kim, J.E. Kim, S.J. Rhie, and S. Yoon, "The role of oxidative stress in neurodegenerative diseases", *Exp. Neurobiol.* **24**, pp. 325 (2015).

[108] Y. Huang, N. Erdmann, H. Peng, Y. Zhao, and J. Zheng, "The role of TNF related apoptosis-inducing ligand in neurodegenerative diseases.", *Cell. Mol. Immunol.* **2**, pp. 113 (2005).

[109] W.R. Markesbery, "Oxidative stress hypothesis in Alzheimer's disease", *Free Radic. Biol. Med.* **23**, pp. 134 (1997).

[110] J.J. Yang, X. Zhang, P. Yuan, J.J. Yang, Y. Xu, J. Grutzendler, Y. Shao, A. Moore, and C. Ran, "Oxalate-curcumin based probe for micro- and macro-imaging of reactive oxygen species in Alzheimer's Disease.", *Proc. Nat. Acad. Sci. USA* **114**, pp. 12384 (2017).

[111] V. Patel, X. Zhang, N.A. Tautiva, A.N. Nyabera, O.O. Owa, M. Baidya, H.C. Sung, P.S. Taunk, S. Abdollahi, S. Charles, R.A. Gonnella, N. Gadi, K.T. Duong, J.N. Fawver, C. Ran, T.O. Jalonen, and I.V.J. Murray, "Small molecules and Alzheimer's disease: misfolding, metabolism and imaging", *Curr. Alzheimer Res.* **12**, pp. 445 (2015).

[112] N. Amdursky, Y. Erez, and D. Huppert, "Molecular rotors: what lies behind the high Sensitivity of the thioflavin-T fluorescent", *Acc. Chem. Res.* **45**(9), pp. 1548 (2012).

[113] E.E. Nesterov, J. Skoch, B.T. Hyman, W.E. Klunk, B.J. Bacskai, and T.M. Swager, "*In vivo* optical imaging of amyloid aggregates in brain: design of fluorescent markers", *Angew. Chem. Int. Ed.* **44**, pp. 5452 (2005).

[114] M. Ono, H. Watanabe, H. Kimura, and H. Saji, "BODIPY-based molecular probe for imaging of cerebral β-amyloid plaques", *ACS Chem. Neurosci.* **3**, pp. 319 (2012).

[115] H. Fu, M. Cui, P. Tu, Z. Pan, and B. Liu, "Evaluation of molecules based on the electron donor-acceptor architecture as near-infrared β-amyloidal-targeting probes", *Chem. Commun.* **50**, pp. 11875 (2014).

[116] H. Fu, M. Cui, L. Zhao, P. Tu, K. Zhou, J. Dai, and B. Liu, "Highly sensitive near-infrared fluorophores for *in Vivo* detection of amyloidβ plaques in Alzheimer's disease", *J. Med. Chem.* **58**, pp. 6972 (2015).

[117] X. Zhang, Y. Tian, P. Yuan, Y. Li, M.A. Yaseen, J. Grutzendler, A. Moore, and C. Ran, "A bifunctional curcumin analogue for two-photon imaging and inhibiting crosslinking of amyloid beta in Alzherimer's disease", *Chem. Commun.* **50**, pp. 11550 (2014).

[118] X. Zhang, Y. Tian, C. Zhang, X. Tian, A.W. Ross, R.D. Moir, H. Sun, R.E. Tanzi, A. Moore, and C. Ran, "Near-infrared fluorescence molecular imaging of Aβ species and monitoring therapy in animal models of Alzheimer's disease", *Proc. Nat. Acad. Sci. USA* **112**, pp. 9734 (2015).

[119] J. Yang, J. Yang, S.H. Liang, Y. Xu, A. Moore, and C. Ran, "Imaging hydrogen peroxide in Alzheimer's disease *via* cascade signal amplification", *Sci. Rep.* **6**, pp. 35613 (2016).

[120] M. Martínez-Cifuentes, B. Weiss-López, and R. Araya-Maturana, "A computational study of structure and reactivity of N-substitued-4-piperidones curcumin analogues and their radical anions", *Molecules* **21**, pp. 1658 (2016).

[121] E. Ferrari, R. Benassi, M. Saladini, G. Orteca, Z. Gazova, and K. Siposova, "*In vitro* study on potential pharmacological activity of curcumin analogues and their copper complexes", *Chem. Biol. Drug Des.* **89**, pp. 411 (2017).

[122] G. Xu, D. Wei, J. Wang, B. Jiang, M. Wang, X. Xue, S. Zhou, B. Wu, and M. Jiang, "Dyes and pigments crystal structure, optical properties and biological imaging of two curcumin derivatives", *Dyes. Pigm.* **101**, pp. 312 (2014).

[123] R. Sabate, L. Rodriguez-santiago, M. Sodupe, S.J. Saupe, and S. Ventura, "Thioflavin-T excimer formation upon interaction with amyloid fibers", *Chem. Commun.* **49**, pp. 5745 (2013).

[124] F. Peccati, J. Hernando, L. Blancafort, X. Solans-Monfort, and M. Sodupe, "Disaggregation-induced fluorescence enhancement of NIAD-4 for the optical imaging of amyloid-beta fibrils", *Phys. Chem. Chem. Phys.* **17**(30), pp. 19718 (2015).

[125] F. Peccati, M. Widniewska, X. Solans-Monfort, and M. Sodupe, "Computational study on donor-acceptor optical markers for Alzheimer's disease: a game of charge transfer and electron delocalization", *Phys. Chem. Chem. Phys.* **18**, pp. 11634 (2016).

[126] N.A. Murugan, R. Zalesny, J. Kongsted, A. Nordberg, and H. Ågren, "Promising two-photon probes for *in vivo* detection of beta amyloid deposits", *Chem. Commun.* **50**, pp. 11694 (2014).

[127] G. Canard, M. Ponce-Vargas, D. Jacquemin, B. Le Guennic, A. Felouat, M. Rivoal, E. Zaborova, A. D'Aléo, and F. Fages, "Influence of the electron donor groups on the optical and electrochemical properties of borondifluoride complexes of curcuminoid derivatives: a joint theoretical and experimental study", *RSC Adv.* **7**, pp. 10132 (2017).

[128] R.G. Parr and W. Yang, *Density-functional theory of atoms and molecules*, Oxford University Press New York (1989).

[129] M.E. Casida and M. Huix-Rotllant, "Progress in time-dependent density-functional theory", *Ann. Rev. Phys. Chem.* **63**, pp. 287–323 (2012).

[130] A.D. Becke, "A new mixing of Hartree-Fock and local density functional theories", *J. Chem. Phys.* **98**, pp. 1372 (1993).

[131] C. Lee, W. Yang, and R.G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density", *Phys. Rev. B* **37**, pp. 785–789 (1988).

[132] C. Adamo and V. Barone, "Toward reliable density functional methods without addjustable parameters: the PBE0 model", *J. Chem. Phys.* **110**, pp. 6158 (1999).

[133] Y. Zhao and D.G. Truhlar, "The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functions", *Theor. Chem. Acc.* **120**, pp. 215 (2007).

[134] J.D. Chai and M. Head-Gordon, "Systematic optimization of long-range corrected hybrid density functionals", *J. Chem. Phys.* **128**, pp. 084106 (2008).

[135] M. Caricato, G.W. Trucks, M.J. Frisch, and K.B. Wiberg, "Electronic transition energies: A study of the performance of a large range of single reference density functional and wave function methods on valence and Rydberg states compared to experiment", *J. Chem. Theory Comput.* **6**, pp. 370 (2010).

[136] S.S. Leang, F. Zahariev, and M.S. Gordon, "Benchmarking the performance of time-dependent density functional methods", *J. Chem. Phys.* **136**, pp. 104101 (2012).

[137] M. Isegawa, R. Peverati, and D.G. Truhlar, "Performance of recent and high-performance approximate density functionals for time-dependent density functional theory calculations of valence and Rydberg electronic transition energies.", *J. Chem. Phys.* **137**, pp. 244104 (2012).

[138] A. Charaf-Eddin, A. Planchat, B. Mennucci, C. Adamo, and D. Jacquemin, "Choosing a functional for computing absorption and fluorescence band shapes with TD-DFT", *J. Chem. Theory Comput.* **9**, pp. 2749 (2013).

[139] T.N. Truong and E.V. Stefanovich, "A new method for incorporating solvent effect into the classical, *ab initio* molecular orbital and density functional theory frameworks for arbitrary shape cavity", *Chem. Phys. Lett.* **240**, pp. 253–260 (1995).

[140] J. Tomasi, B. Mennucci, and E. Cancès, "The IEF Version of the PCM Solvation Method: An Overview of a New Method Addressed to Study Molecular Solutes at the QM ab initio Level", *J. Mol. Struct. THEOCHEM* **464**, pp. 211 (1999).

[141] Z.Q. You, J.M. Mewes, A. Dreuw, and J.M. Herbert, "Comparison of the Marcus and Pekar partitions in the context of non-equilibrium, polarizable-continuum solvation models", *J. Chem. Phys.* **143**, pp. 1 (2015).

[142] M. Segado, E. Benassi, and V. Barone, "A "twist" on the interpretation of the multifluorescence patterns of DASPMI", *J. Chem. Theory Comput.* **11**, pp. 4803–4813 (2015).

[143] S.P. Parimita, Y.V. Ramshankar, S. Suresh, and T.N. Guru Row, "Redetermination of curcumin: (1E,4Z, 6E)-5-hydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)hepta-1,4,6-trien-3-one", *Acta Cryst.* **E63**, pp. o860 (2007).

[144] C.F. Chignell, P. Bilski, K.J. Reszka, A.G. Motten, R.H. Sik, and T.A. Dahl, "Spectral And photochemical properties of curcumin", *Photochem. Photobiol.* **59**, pp. 295–302 (1994).

[145] S.i. Kawano, Y. Inohana, Y. Hashi, and J.M. Lin, "Analysis of keto-enol tautomers of curcumin by liquid chromatography/mass spectrometry", *Chin. Chem. Lett.* **24**, pp. 685 (2013).

[146] L. Shen and H.F. Ji, "Theoretical study on physicochemical properties of curcumin", *Spectrochim. Acta A* **67**, pp. 619 (2007).

[147] S. Anjomshoa, M. Namazian, and M.R. Noorbala, "Is curcumin a good scavenger of reactive oxygen species? A computational investigation", *Theor. Chem. Acc.* **136**, pp. 1 (2017).

[148] A.E. Shilov and G.B. Shul'pin, "Activation of C-H bonds by metal complexes", *Chem. Rev.* **97**(8), pp. 2879 (1997).

[149] K.M. Engle, D.H. Wang, and J.Q. Yu, "Ligand-accelerated C-H activation reactions: evidence for a switch of mechanism", *J. Am. Chem. Soc.* **132**(40), pp. 14137 (2010).

[150] Y.F. Yang, X. Hong, J.Q. Yu, and K.N. Houk, "Experimental–computational synergy for selective Pd(II)-catalyzed C–H activation of aryl and alkyl groups", *Acc. Chem. Res.* **50**(11), pp. 2853 (2017).

[151] K. Liao, T.C. Pickel, V. Boyarskikh, J. Bacsa, D.G. Musaev, and H.M.L. Davies, "Site-selective and stereoselective functionalization of non-activated tertiary C–H bonds", *Nature* **551**(7682), pp. 609 (2017).

[152] D. Gallego and E.A. Baquero, "Recent advances on mechanistic studies on C–H activation catalyzed by base metals", *Open Chem.* **16**(1), pp. 1001 (2018).

[153] B. Shao, A.L. Bagdasarian, S. Popov, and H.M. Nelson, "Arylation of hydrocarbons enabled by organosilicon reagents and weakly coordinating anions", *Science* **355**(6332), pp. 1403 (2017).

[154] S. Popov, B. Shao, A.L. Bagdasarian, T.R. Benton, L. Zou, Z. Yang, K.N. Houk, and H.M. Nelson, "Teaching an old carbocation new tricks: intermolecular C–H insertion reactions of vinyl cations", *Science* **361**(6400), pp. 381 (2018).

[155] S. Duttwyler, C. Douvris, N.L.P. Fackler, F.S. Tham, C.A. Reed, K.K. Baldridge, and J.S. Siegel, "C–F activation of fluorobenzene by silylium carboranes: evidence for incipient phenyl cation reactivity", *Angew. Chem. Int. Ed.* **49**(41), pp. 7519 (2010).

[156] E. Fasani, M. Mella, D. Caccia, S. Tassi, M. Fagnoni, and A. Albini, "The photochemistry of lomefloxacin. An aromatic carbene as the key intermediate in photodecomposition", *Chem. Commun.* **14**, pp. 1329 (1997).

[157] O. Allemann, K.K. Baldridge, and J.S. Siegel, "Intramolecular C–H insertion vs. Friedel–Crafts coupling induced by silyl cation-promoted C–F activation", *Org. Chem. Front.* **2**(9), pp. 1018 (2015).

[158] X. Li, L. Sun, Q. Zhang, S. Li, Y. Wang, D. Wei, W. Zhang, and Y. Lan, "Mechanism and substituent effects of benzene arylation *via* a phenyl cation strategy: a density functional theory study", *ChemCatChem* **11**(20), pp. 5068 (2019).

[159] S. Maeda, K. Ohno, and K. Morokuma, "Systematic exploration of the mechanism of chemical reactions: the global reaction route mapping (GRRM) strategy using the ADDF and AFIR methods", *Phys. Chem. Chem. Phys.* **15**(11), pp. 3683 (2013).

[160] F. Weigend and R. Ahlrichs, "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy", *Phys. Chem. Chem. Phys.* **7**(18), pp. 3297 (2005).

Appendix A

Supplemental Calculation Results of CRANAD Molecules

A.1 Basis set, grid size, and solvent model effects

Table A.1: Basis-set-dependence of computed absorption wavelengths (in nm) of CRANAD molecules from TDDFT/TDA and TDDFT calculations using ωB97X-D functional, different basis sets, and C-PCM solvent model.

Basis Set	**5**		**54**		**61**	
	TDA	TDDFT	TDA	TDDFT	TDA	TDDFT
6-31G*	386.9	397.7	382.0	398.6	470.2	481.8
6-311+G*	401.6	410.0	396.2	414.1	482.9	495.6
6-311+G(2d,p)	403.6	410.0	395.6	413.5	483.1	495.3
6-311++G**	403.2	410.9	396.9	414.8	483.5	495.6
def2-TZVP	399.5	405.6	390.8	408.4	476.9	489.2
cc-pVDZ	390.8	401.6	386.9	404.0	471.1	482.8
cc-pVTZ	396.6	404.9	389.1	406.6	475.3	488.1
aug-cc-pVDZ	408.0	415.3	400.4	418.7	487.4	500.4
aug-cc-pVTZ	-	-	393.1	408.2	-	-

Table A.2: Integration-grid-dependence of computed bbsorption wavelengths (in nm) of CRANAD-X molecules from TDDFT/TDA and TDDFT calculations using ωB97X-D functional, 6-311++G** basis set, and C-PCM solvent model.

Grid	**5**		**54**		**61**	
	TDA	TDDFT	TDA	TDDFT	TDA	TDDFT
SG-1	403.2	410.9	396.9	414.8	483.5	495.6
SG-2	404.4	412.1	397.3	415.4	484.3	496.8
(50,194)	403.5	409.9	396.4	414.3	485.1	497.7
(75,302)	404.2	411.9	397.2	415.3	484.3	496.8
(99,590)	404.1	410.9	397.2	415.2	484.4	496.9

Table A.3: Solvent-model-dependence of computed absorption wavelengths (in nm) of CRANAD molecules from TDDFT/TDA and TDDFT calculations using ωB97X-D functional, 6-311++G** basis set, and different solvent models.

Method	**5**		**54**		**61**	
	TDA	TDDFT	TDA	TDDFT	TDA	TDDFT
C-PCM	404.3	410.7	397.3	415.3	483.5	495.6
IEF-PCM	403.3	412.2	394.9	413.4	481.6	496.6
Difference	1.0	-1.5	2.4	1.9	1.9	-1.0

Table A.4: Solvent-model-dependence of computed emission wavelengths (in nm) of CRANAD molecules from TDDFT/TDA and TDDFT calculations using ωB97X-D functional, 6-311++G** basis set, and different solvent models.

Method	**5**		**54**		**61**	
	TDA	TDDFT	TDA	TDDFT	TDA	TDDFT
C-PCM	428.2	440.3	486.7	512.7	602.0	630.9
IEF-PCM	428.1	440.2	486.3	512.3	601.6	630.5
Difference	0.1	0.1	0.4	0.4	-0.6	0.4

Table A.5: Computed absorption wavelengths (in nm) of CRANAD molecules from TDDFT/C-PCM-ptSS calculations using different functionals and 6-311++G** basis set.

molecule	2	3	5	28	54	58	61
Expt.	645.0	610.0	500.0	498.0	490.0	630.0	657.0
B3LYP	592.3	574.0	536.1	471.1	473.0	584.5	608.6
PBE0	566.8	548.2	510.8	454.5	458.9	559.5	584.3
ωB97X-D	479.0	459.8	417.3	397.6	417.3	471.0	500.3
M06-2X	509.5	491.8	439.8	412.7	434.8	503.6	534.0

A.2 Curcumin Conformers and Analogues

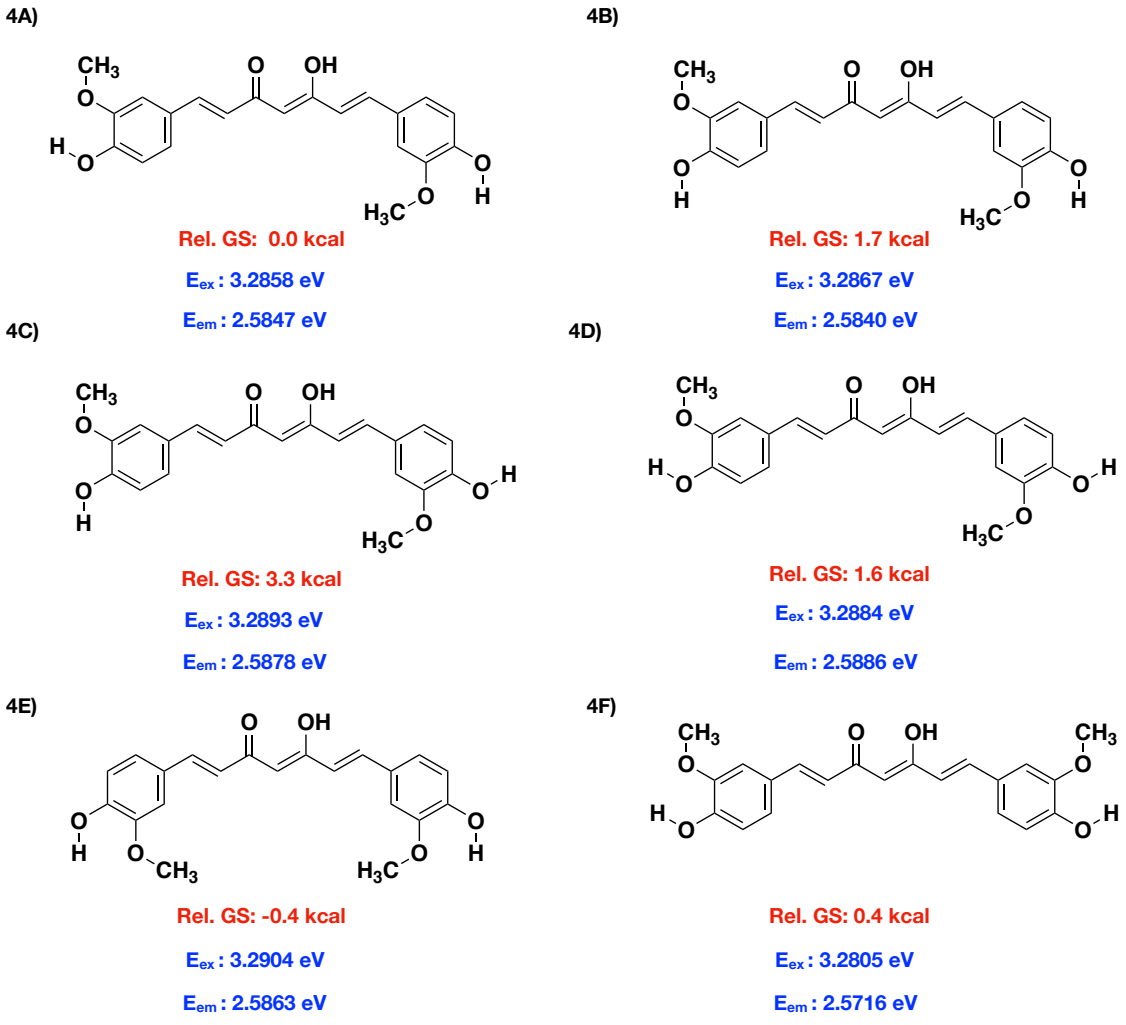


Figure A.1: Ground state energy (at ωB97X-D/6-311++G**/C-PCM level of theory) of several curcumin conformers relative to the reported crystal structure conformer (4A). Also shown are absorption and emission energies from TDDFT calculations.

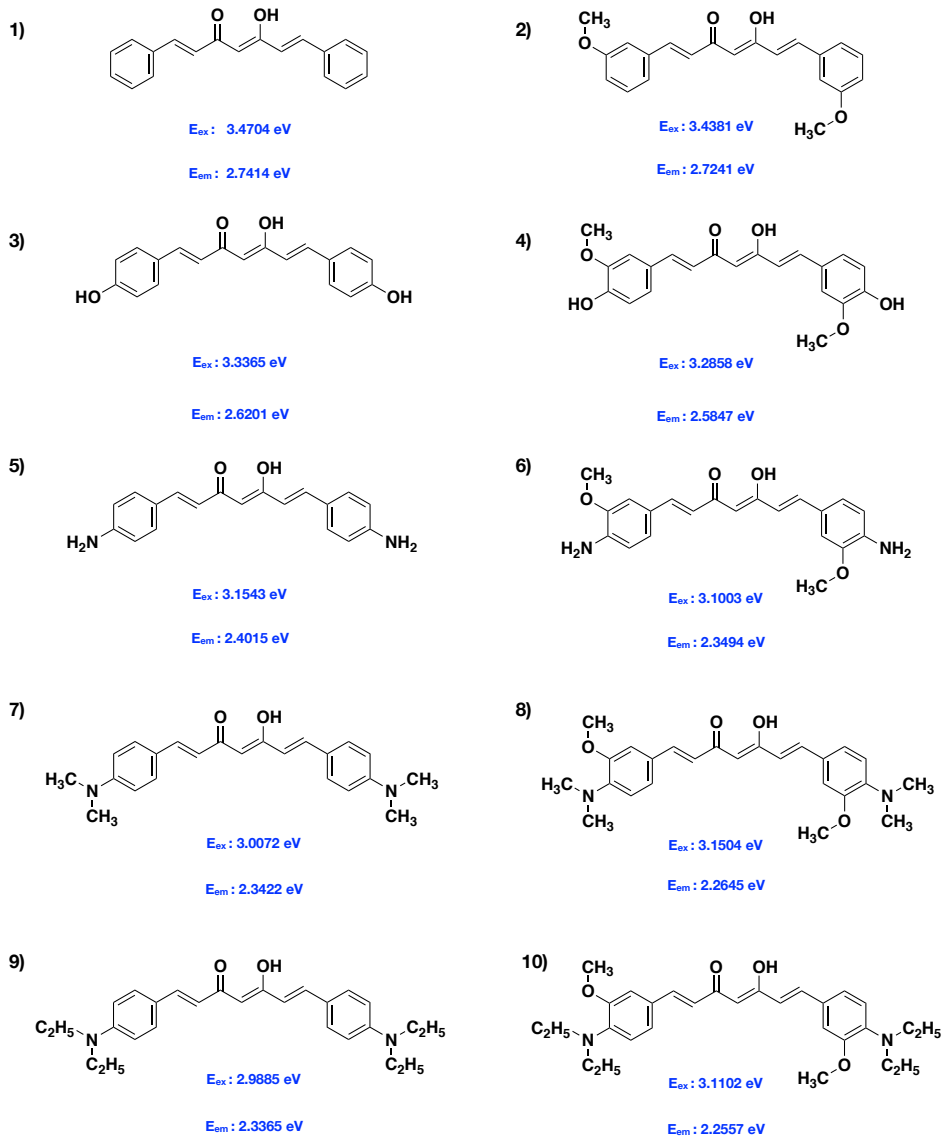


Figure A.2: Absorption and emission energies for curcumin and analogs from TDDFT calculations using ωB97X-D functional, 6-311++G** basis set, and C-PCM solvent model. Note that molecule 4 is the standard curcumin molecule, while molecule 7 is the same molecule as CRANAD-**5**.

Table A.6: Relative ground-state energies between the diketo and enol forms of curcumin from DFT calculations using four functionals, 6-311++G** basis set, and C-PCM solvent model. Absorption energies were computed using TDDFT calculations.

Functional	Tautomeric form	Rel. E_{gs} (in kcal/mol)	E_{ex} (in eV)	λ_{ex} (in nm)
B3LYP	enol	0.00	2.6520	467.6
	diketo	4.36	3.2818	377.8
PBE0	enol	0.00	2.7723	447.3
	diketo	5.63	3.4526	359.1
ωB97X-D	enol	0.00	3.2976	376.0
	diketo	2.40	3.8078	248.3
M06-2X	enol	0.00	3.2198	385.1
	diketo	3.84	3.8115	325.3

A.3 Keto-enol Tautomerization and Proton Transfer of CRANAD-5

Table A.7: Relative ground-state energies between the diketo and enol forms of CRANAD-**5** from DFT calculations using four functionals, 6-311++G** basis set, and C-PCM solvent model. Absorption energies were computed using TDDFT calculations.

Functional	Tautomeric form	Rel. E_{gs} (in kcal/mol)	E_{ex} (in eV)	λ_{ex} (in nm)
B3LYP	enol	0.00	2.3216	534.1
	diketo	3.41	2.8550	434.3
PBE0	enol	0.00	2.4382	508.6
	diketo	5.12	3.0742	403.4
ωB97X-D	enol	0.00	2.9957	413.9
	diketo	0.92	3.4460	359.8
M06-2X	enol	0.00	2.8373	437.0
	diketo	3.16	3.3623	368.8

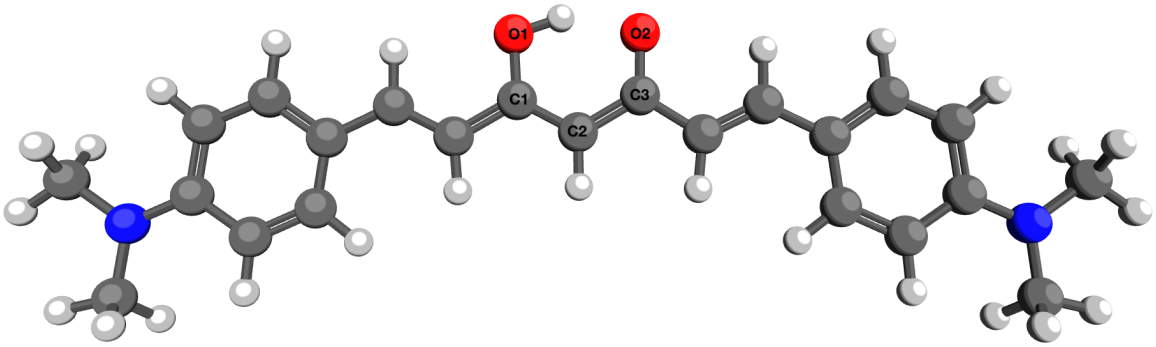


Figure A.3: Optimized ground-state geometry of the enol form of CRANAD-**5** at the ωB97X-D/6-311++G**/C-PCM level of theory. The molecule is nearly planar with a O1–C1–C2–C3 dihedral angle of 0.3° and a C1–C2–C3–O2 angle of -0.3°.

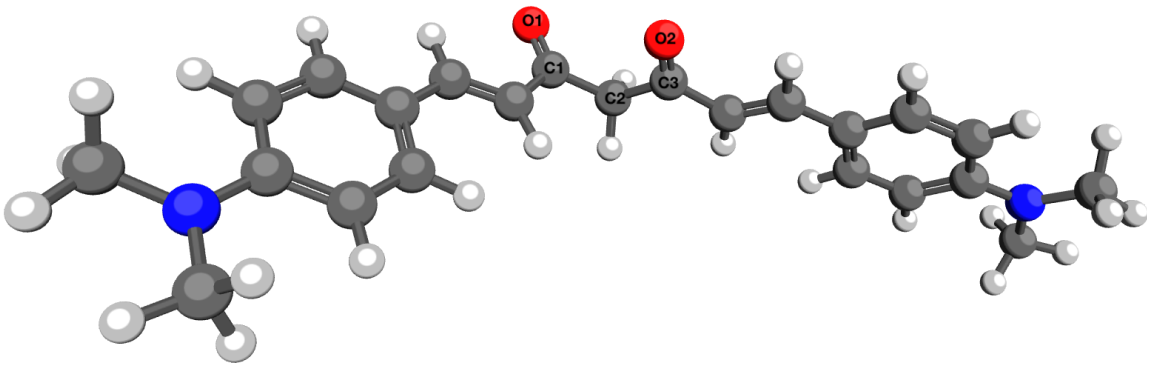


Figure A.4: Optimized ground-state geometry of the diketo form of CRANAD-**5** at the ωB97X-D/6-311++G**/C-PCM level of theory. The molecule is non-planar with a O1–C1–C2–C3 dihedral angle of 103.6°and a C1–C2–C3–O2 angle of -7.3°.

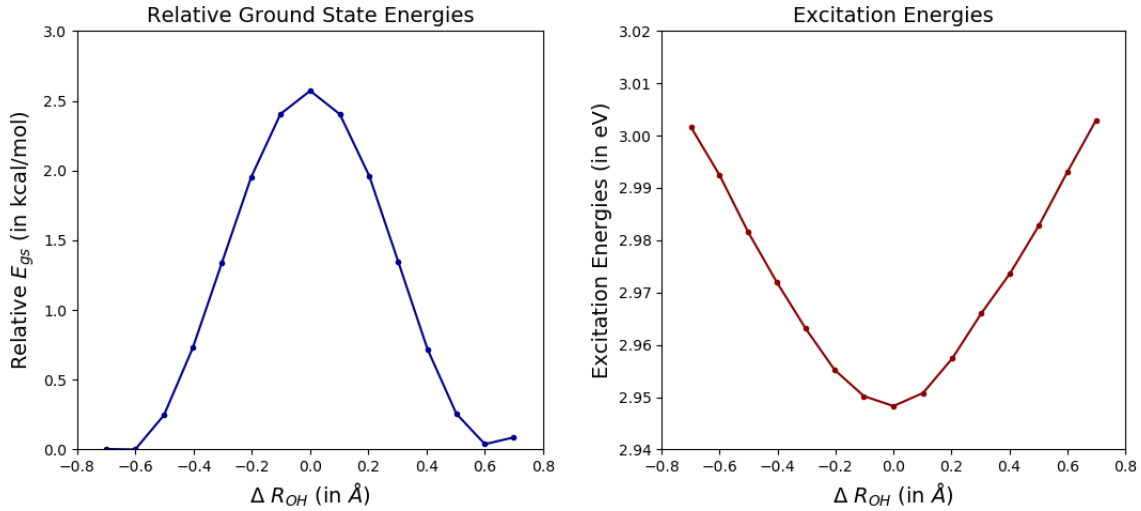


Figure A.5: Proton transfer barrier for CRANAD-**5** molecule. X-axis: Difference in the O-H distances, ΔR_{OH} = R_{O1-H} - R_{O2-H} (see Figure A.4). Left: Relative gas-phase ground-state energy at ωB97X-D/6-311++G** level of theory from restrained geometry optimization. Right: vertical S_0 to S_1 excitation energies.

A.4 Geometrical and Spectroscopical properties of CRANAD molecules

Table A.8: Dihedral angles of CRANAD molecules in their ground state (S_0) geometry from DFT calculations at ωB97X-D/6-311++G**/C-PCM level of theory.

Molecule	O1–C1–C2–C3	C1–C2–C3–O2	B1–O1–C1–C2
2	0.1	-0.1	0.2
3	4.7	-5.1	10.4
5	0.3	-0.3	-
28	5.0	-5.0	9.6
54	3.6	-3.7	7.5
58	5.2	-5.2	10.8
61	4.5	-4.5	8.1

Table A.9: Dihedral angles of CRANAD molecules in their first excited state (S_1) geometry from TDDFT calculations at ωB97X-D/6-311++G**/C-PCM level of theory.

Molecule	O1–C1–C2–C3	C1–C2–C3–O2	B1–O1–C1–C2
2	1.5	-1.5	18.0
3	2.4	-2.4	17.7
5	0.4	-0.4	-
28	2.3	-2.3	17.5
54	2.8	-4.1	15.8
58	2.8	-2.5	17.3
61	2.4	-2.4	15.0

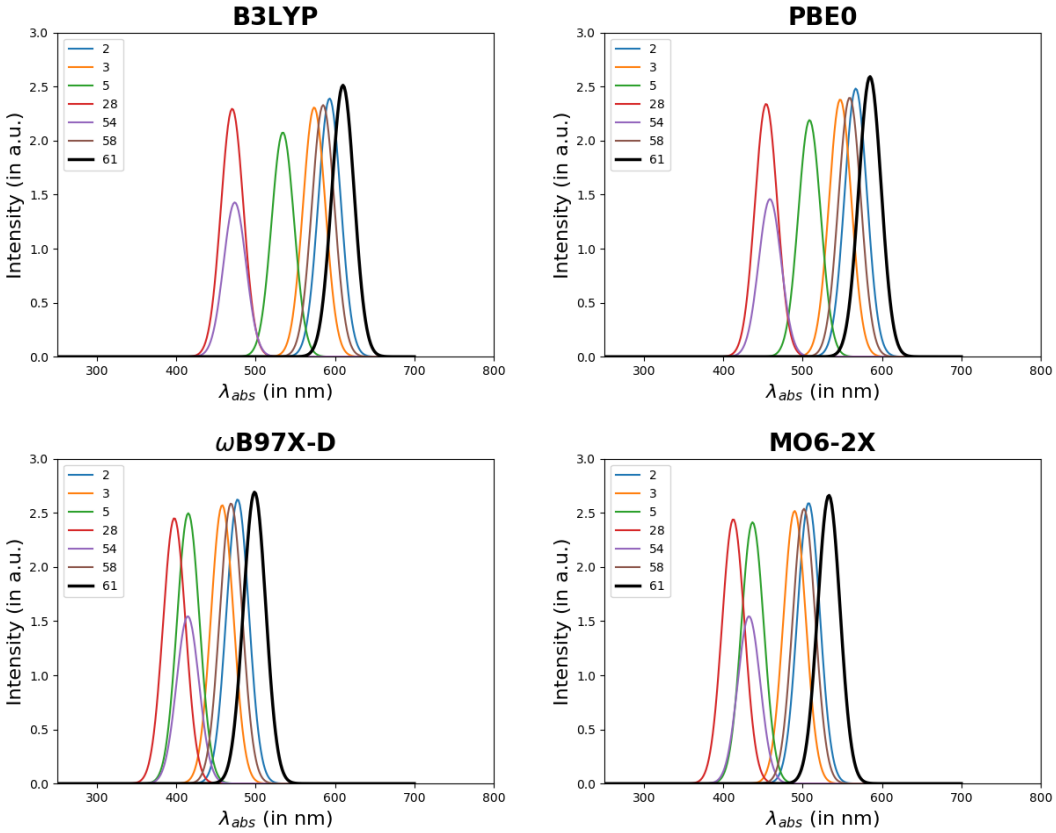


Figure A.6: Computed $S_0 \rightarrow S_1$ absorption wavelengths (in nm) of CRANAD molecules and intensities (in a.u.) from TDDFT calculations using different functionals, 6-311++G** basis set, and C-PCM solvent model.

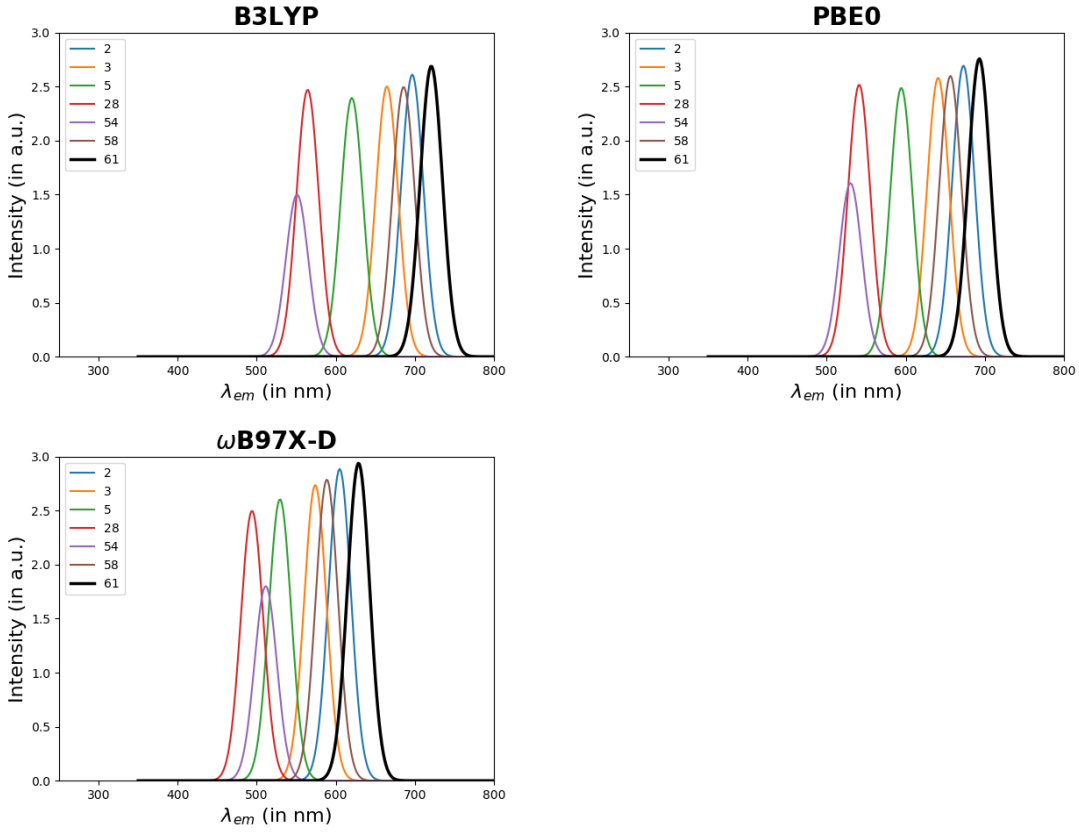


Figure A.7: Computed $S_1 \rightarrow S_0$ emission wavelengths (in nm) of CRANAD molecules and intensities (in a.u.) from TDDFT calculations using different functionals, 6-311++G** basis set, and C-PCM solvent model.

Table A.10: Computed absorption and emission wavelengths (in nm) of the enol, BF_2, BC_2O_6 analogs of CRANAD molecules from TDDFT calculations using ωB97X-D functional, 6-311++G** basis set, and C-PCM solvent model.

Analogs	2	3	28	54	58
			Absorption		
Enol	413.6	400.8	358.9	365.8	408.6
BF_2	477.2	457.3	398.2	415.4	468.9
BC_2O_6	496.8	474.8	410.8	431.6	488.0
			Emission		
Enol	529.3	509.5	450.5	465.8	530.3
BF_2	604.9	574.0	494.0	511.3	588.7
BC_2O_6	628.8	597.4	509.3	520.5	612.3

A.5 Wavefunction Analysis

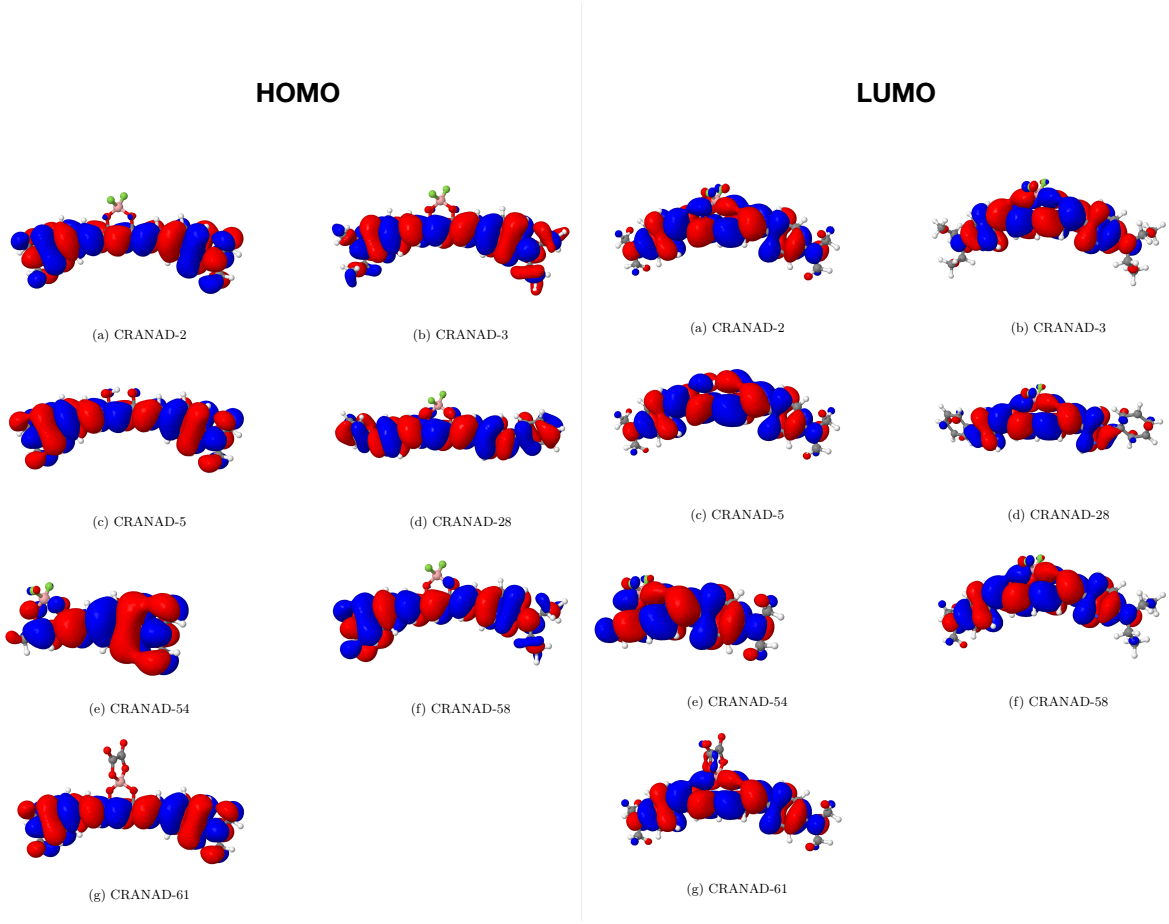


Figure A.8: HOMO and LUMO of CRANAD molecules from DFT calculations using ωB97X-D functional, 6-311++G** basis set, and C-PCM solvent model.
Isosurface value is 0.01.

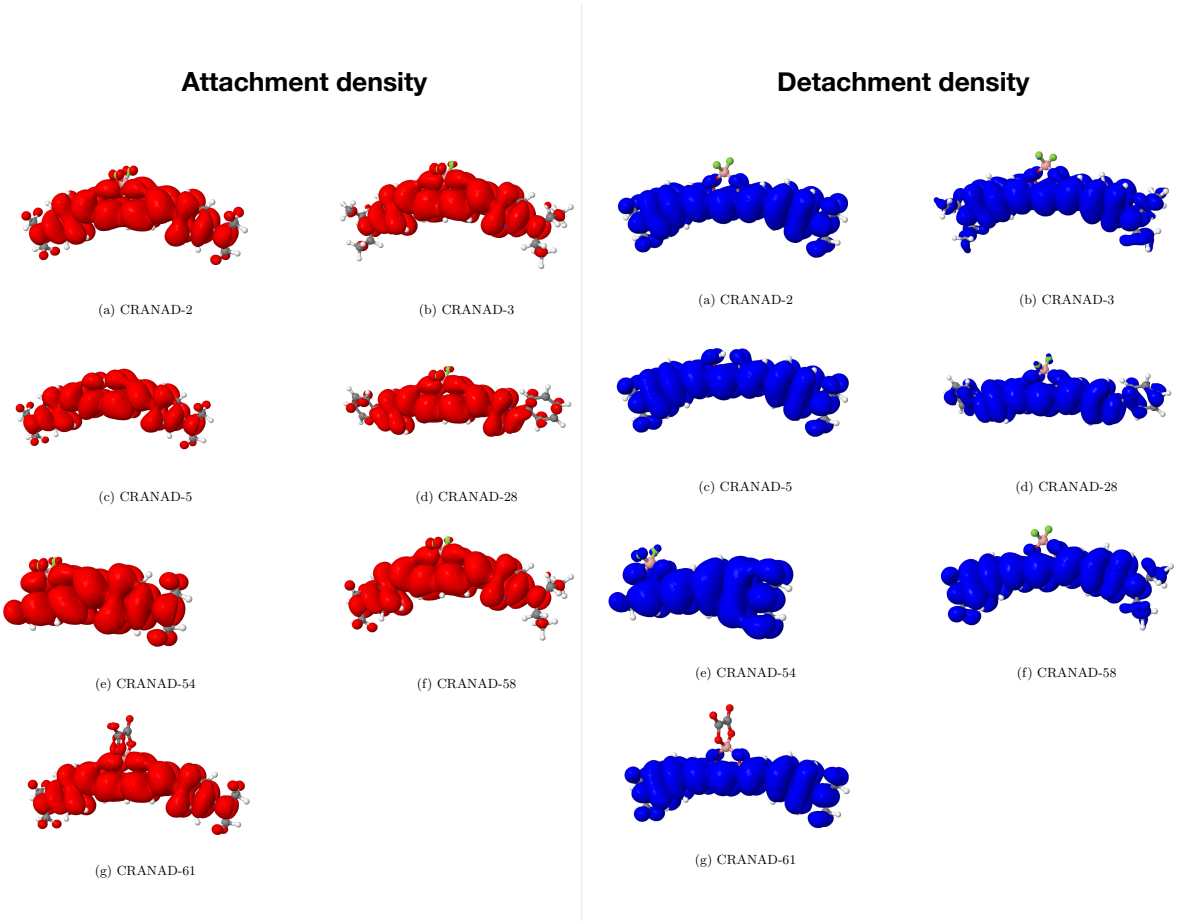


Figure A.9: Attachment and detachment densities of CRANAD molecules from TDDFT/TDA calculations using ωB97X-D functional, 6-311++G** basis set, and C-PCM solvent model. Isosurface value is 0.00004.

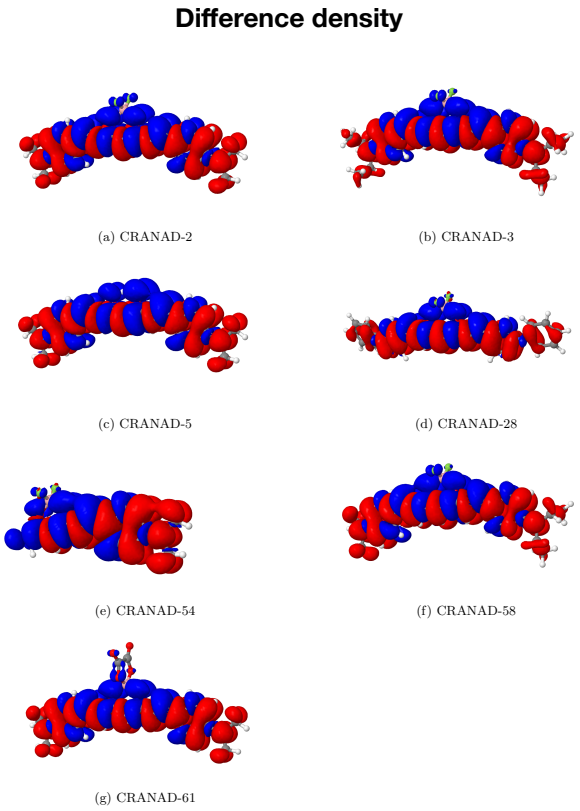


Figure A.10: Difference densities of CRANAD molecules from TDDFT/TDA calculations using ωB97X-D functional, 6-311++G** basis set, C-PCM solvent model. Isosurface value of 0.00004.

A.6 Torsional Energy Profiles

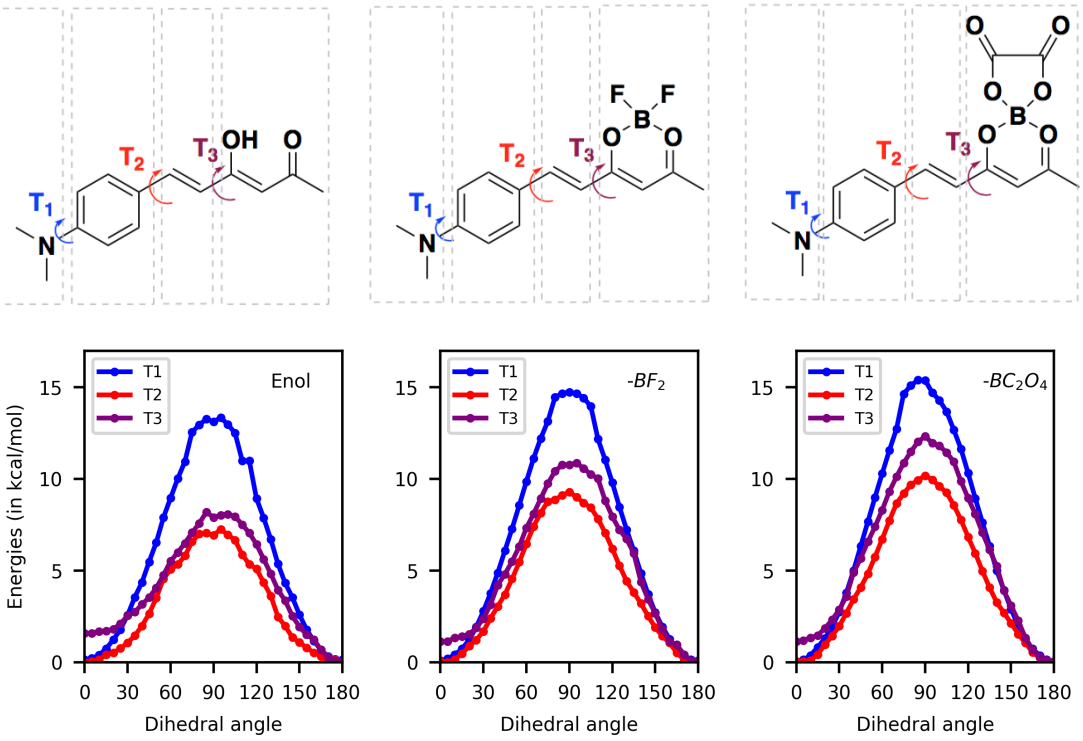


Figure A.11: Rotational barriers on the ground-state potential energy surface of CRANAD-54 (middle) and its enol and boron-oxalate analogs with dihedral angles T1, T2, and T3. This indicates a planar ground-state geometry (T1=T2=T3= 180^o) for CRANAD-54 and analogs, and slightly higher barriers with BF_2 and BC_2O_6 groups.

Table A.11: Barrier heights (in kcal/mol) with T1, T2, and T3 torsions on the ground-state PES of the CRANAD-54 molecule and its enol and boro-oxalate analogs.

Angle	54-BC_2O_6	54-BF_2	54-enol
T1	15.35	14.65	12.98
T2	10.17	9.23	6.92
T3	11.21	9.66	6.32

Appendix B

Supplemental Computational Results of Oxyluciferin Analogs

B.1 Comparison with experimental results

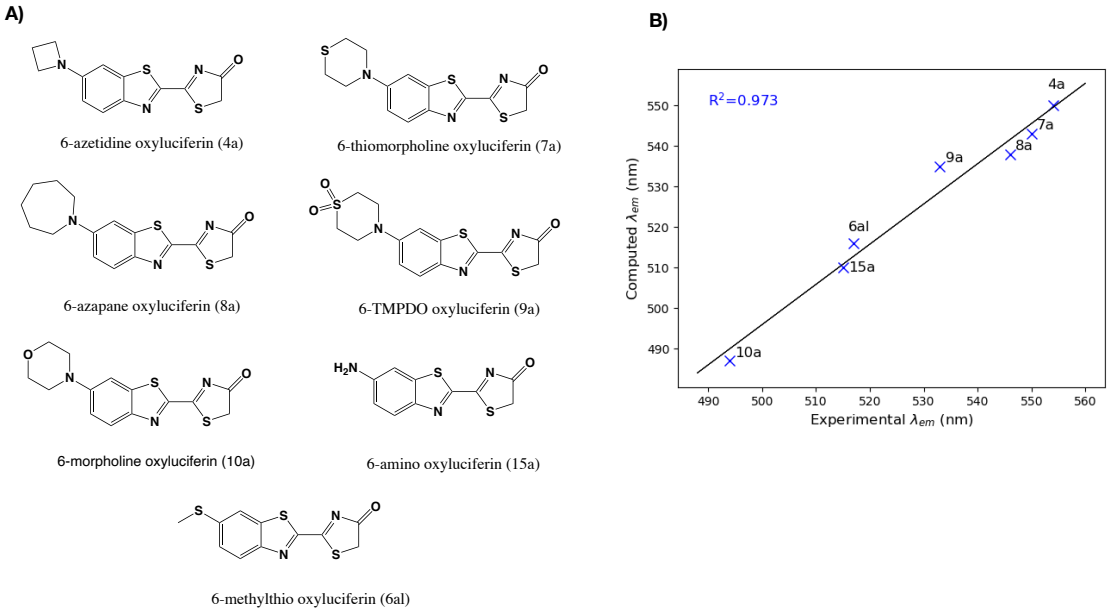


Figure B.1: (A) C6'-functionalized oxyluciferin analogs from Miller and coworkers;(B) Correlation ($R^2 = 0.973$ with slope=0.992, and y-intercept=-0.258) between experimentally measured emission wavelengths and computed TDDFT emission wavelengths obtained using the ωB97X-D/6-311++G** level of theory and the C-PCM solvent model.

Table B.1: Computed emission wavelengths (λ, in nm) of oxyluciferin analogs using the TDDFT/ωB97X-D/6-311++G** level of theory in C-PCM solvent model.

Compound	λ_{abs} (nm)		λ_{em} (nm)	
	Exp[a]	Comp	Exp[a]	Comp
6'-azetidine-oxyluciferin (4a)	369	391	554	550
6'-thiomorpholine-oxyluciferin (7a)	414	405	550	543
6'-azapane oxyluciferin	346	362	546	538
6'-TMPDO-oxyluciferin (9a)	354	347	533	535
6'-morpholine-oxyluciferin (10a)	344	344	494	487
6'-amino-oxyluciferin (15a)	359	343	515	510
6'-methylthio-oxyluciferin (6al)	351	365	517	516

[a] Experimental emission wavelengths as reported by Miller and coworkers (96).

Table B.2: Computed emission wavelengths (λ_{em}, in nm) of oxyluciferin analogs in gas phase at TDDFT/CAM-B3LYP/6-31+G* level of theory.

Substituent	Site					
	C4'		C5'		C7'	
	λ_{em}[a]	λ_{em}	λ_{em}[a]	λ_{em}	λ_{em}[a]	λ_{em}
H	490.9	489.1	490.9	489.1	490.9	489.1
F	491.6	489.9	478.8	477.1	505.0	503.0
OH	-	494.2[c]	469.2	467.5	-	557.9
HO	507.1	505.3[c]	-	480.5	609.0	605.3
C(CH)	501.9	496.9	503.2	501.3	498.3	499.5

[a] Computed emission wavelengths were reported by Liu and coworkers (68).

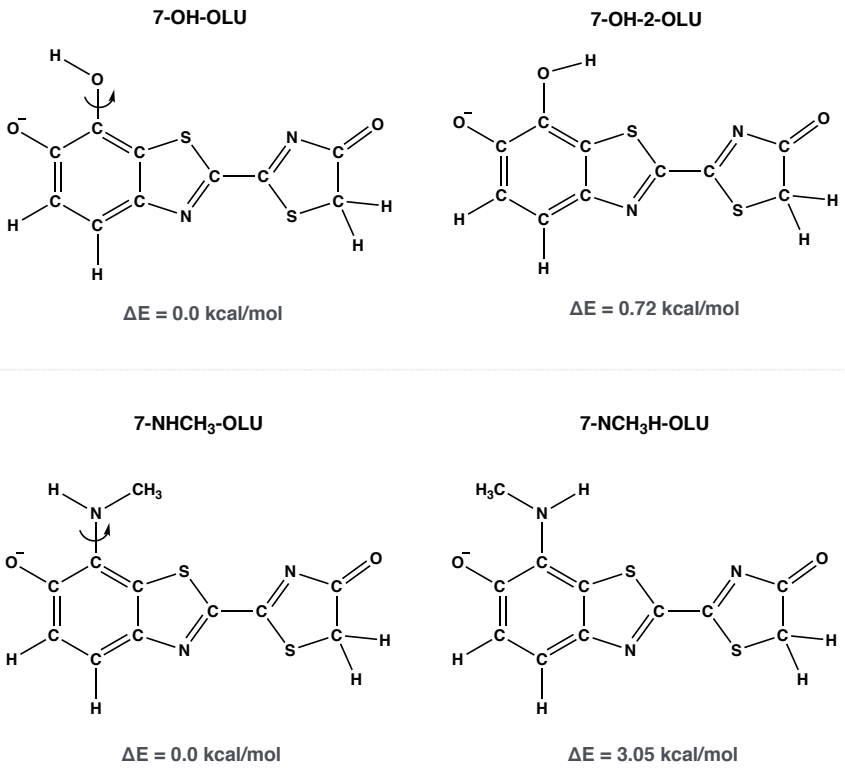


Figure B.2: Chemical structures and relative ground state energies of rotamers hydroxy(OH/OH-2) and $NHCH_3$/NCH_3H OLU analogs optimized using DFT/ωB97X-D/6-311++G**//C-PCM level of theory

B.2 Basis set and solvent model effects on absorption and emission wavelengths of oxyluciferin and its analogs

Table B.3: Basis-set-dependence of calculated absorption wavelengths (λ, in nm) of oxyluciferin molecule using both TDA/TDDFT and TDDFT calculations at ωB97X-D/6-311++G** level of theory using C-PCM solvent model.

Basis Set	Gas Phase		PCM	
	TDA	TDDFT	TDA	TDDFT
6-31G*	402.7	445.5	425.7	457.6
6-311+G*	415.0	456.4	424.4	451.4
6-311++G**	415.0	456.2	424.1	450.9
6-311++G(3df, 3pd)	417.0	457.1	424.4	451.2
def2-TZVP	411.2	452.1	422.5	450.3
cc-pVTZ	408.2	449.1	422.7	452.0
aug-cc-pVDZ	419.0	460.4	429.7	458.0
aug-cc-pVTZ	416.0	456.5	425.1	452.3

Table B.4: Solvent-model-dependence of computed absorption wavelengths (λ, in nm) of oxyluciferin molecule from TDDFT/TDA and TDDFT calculations using ωB97X-D functional, 6-311++G** basis set, and different solvent models.

Method	OLU	
	TDA	TDDFT
C-PCM	424.5	451.3
IEF-PCM	421.6	449.0
Difference	2.9	2.3

Table B.5: TDDFT absorption and emission wavelengths (λ_{abs}) and strengths (f) of phenolate keto OLU are computed using ωB97X-D/6-311++G** in various solvents using C-PCM solvent model.

Solvent (ε_{static})($\varepsilon_{optical}$)	Absorption		Emission	
	λ_{abs} (nm)	f	λ_{em} (nm)	f
gas phase (1)(1)	456	0.73	486	0.77
benzene (2.27)(2.25)	491	1.03	523	0.99
xylene (2.940)(1.50)	486	1.01	533	1.05
chloroform (4.710)(2.08)	476	0.98	549	1.12
propanol (19.019)(1.92)	460	0.93	573	1.23
methanol (32.63)(1.77)	453	0.90	577	1.24
acetonitrile (37.5)(1.81)	453	0.90	578	1.25
dimethylsulfoxide (46.83)(2.18)	456	0.92	579	1.25
water (78.3)(1.78)	451	0.73	580	1.26

B.3 Functional dependence

Table B.6: Computed absorption wavelengths (λ, in nm) and oscillator strengths (f) of oxyluciferin and its analogs in aqueous solution at TDDFT/ωB97X-D/6-311++G**//C-PCM level of theory.

Group	C7'			C5'			C4'		
	λ	$\Delta\lambda$	f	λ	$\Delta\lambda$	f	λ	$\Delta\lambda$	f
NO_2	381.9	-67.2	0.690	441.3	-7.8	0.570	458.7	9.6	0.550
CN	402.8	-46.3	0.890	422.3	-26.8	0.790	442.9	-6.2	0.790
CH3CO	407.0	-42.1	0.890	443.9	-5.2	0.810	453.0	3.9	0.820
CF3	412.1	-37.0	0.910	427.8	-21.3	0.810	439.0	-10.1	0.850
I	432.5	-16.6	0.880	435.7	-13.4	0.950	445.0	-4.1	0.820
Br	433.3	-15.8	0.890	435.2	-13.9	0.940	444.0	-5.1	0.830
Cl	434.1	-15.0	0.890	435.0	-14.1	0.920	443.3	-5.8	0.840
CCH	435.2	-13.9	0.860	442.9	-6.2	0.940	452.3	3.2	0.830
F	442.8	-6.3	0.880	436.0	-13.1	0.930	441.5	-7.6	0.860
H	449.0	0.0	0.920	449.1	0.0	0.925	449.1	0.0	0.925
$CHCH_2$	454.4	5.3	0.840	458.9	9.8	0.980	466.3	17.2	0.770
NMe_2	458.5	9.4	0.890	476.2	27.1	1.090	487.4	38.3	0.770
OH	478.5	29.4	0.690	444.7	-4.4	1.010	456.3	7.2	0.870
OH-2	486.5	37.4	0.790	459.7	10.6	1.040	459.4	10.3	0.870
NCH_3H	491.0	41.9	0.760	487.7	38.6	1.070	479.6	30.5	0.740
$NHCH_3$	506.8	57.7	0.670	472.0	22.9	1.070	489.8	40.7	0.810
NH_2	508.7	59.6	0.680	489.9	40.8	1.060	483.4	34.3	0.800

Table B.7: Computed absorption wavelengths (λ, in nm) and oscillator strengths (f) of oxyluciferin and its analogs in aqueous solution at TDDFT/B3LYP/6-311++G**//C-PCM level of theory.

Group	C7'			C5'			C4'		
	λ	$\Delta\lambda$	f	λ	$\Delta\lambda$	f	λ	$\Delta\lambda$	f
NO_2	456.6	-53.6	0.63	568.4	58.2	0.41	641.8	131.6	0.24
$COCH_3$	476.4	-33.8	0.76	544.7	34.5	0.62	528.4	18.2	0.68
CF_3	479.4	-30.8	0.81	514.1	3.9	0.65	515.0	4.8	0.70
CN	480.8	-29.4	0.75	521.7	11.5	0.63	525.2	15.0	0.66
Cl	507.2	-3.0	0.73	508.3	-1.9	0.79	516.4	6.2	0.69
Br	507.4	-2.8	0.72	510.1	-0.1	0.80	517.3	7.1	0.68
I	509.0	-1.2	0.70	512.2	2.0	0.82	518.3	8.1	0.67
H	510.2	0.0	0.79	510.2	0.0	0.79	510.2	0.0	0.79
F	513.2	3.0	0.72	500.4	-9.8	0.80	511.8	1.6	0.71
CCH	523.4	13.2	0.65	524.6	14.4	0.80	525.7	15.5	0.69
NMe_2	525.5	15.3	0.31	536.9	26.7	0.88	551.3	41.1	0.30
$CHCH_2$	550.4	40.2	0.60	527.7	17.5	0.88	534.9	24.7	0.62
OH	565.6	55.4	0.58	503.6	-6.6	0.95	516.7	6.5	0.69
OH	577.8	67.6	0.50	496.8	-13.4	0.96	515.6	5.4	0.70
NCH_3H	615.1	104.9	0.43	588.5	78.3	0.63	552.4	42.2	0.24
NH_2	627.2	117.0	0.42	523.7	13.5	0.92	541.2	31.0	0.43
$NHCH_3$	644.6	134.4	0.38	525.6	15.4	0.97	548.2	38.0	0.43

Table B.8: Computed absorption wavelengths (λ, in nm) and oscillator strengths (f) of oxyluciferin and its analogs in aqueous solution at TDDFT/M06-2X/6-311++G**//C-PCM level of theory.

Group	C7'			C5'			C4'		
	λ	$\Delta\lambda$	f	λ	$\Delta\lambda$	f	λ	$\Delta\lambda$	f
NO_2	389.2	-69.3	0.77	447.8	-10.7	0.62	463.5	5.0	0.63
CN	414.0	-44.5	0.88	433.4	-25.1	0.81	452.1	-6.4	0.81
$COCH_3$	417.6	-40.9	0.89	455.7	-2.8	0.82	468.4	9.9	0.80
CF_3	421.1	-37.4	0.91	437.3	-21.2	0.81	449.1	-9.4	0.85
I	444.8	-13.7	0.88	447.3	-11.2	0.95	455.4	-3.1	0.84
Br	445.2	-13.3	0.88	446.0	-12.5	0.95	454.9	-3.6	0.84
Cl	446.3	-12.2	0.89	445.8	-12.7	0.93	454.0	-4.5	0.85
CCH	448.5	-10.0	0.85	454.6	-3.9	0.95	461.6	3.1	0.85
F	453.7	-4.8	0.88	445.6	-12.9	0.94	451.4	-7.1	0.87
H	458.5	0.0	0.93	458.5	0.0	0.93	458.5	0.0	0.93
NMe_2	468.8	10.3	0.89	482.0	23.5	1.10	484.4	25.9	0.80
$CHCH_2$	469.9	11.4	0.83	469.9	11.4	0.99	476.0	17.5	0.79
OH	490.3	31.8	0.73	453.2	-5.3	1.02	466.1	7.6	0.89
HO	495.7	37.2	0.82	466.7	8.2	1.05	466.1	7.6	0.89
NCH_3H	504.7	46.2	0.77	482.0	23.5	1.10	478.0	19.5	0.80
NH_2	523.8	65.3	0.69	494.7	36.2	1.09	479.7	21.2	0.85
$NHCH_3$	528.4	69.9	0.66	503.5	45.0	1.07	497.7	39.2	0.79

Table B.9: Computed absorption wavelengths (λ, in nm) and oscillator strengths (f) of oxyluciferin and its analogs in aqueous solution at TDDFT/CAM-B3LYP/6-311++G**//C-PCM level of theory.

Group	C7'			C5'			C4'		
	λ	$\Delta\lambda$	f	λ	$\Delta\lambda$	f	λ	$\Delta\lambda$	f
NO_2	390.2	-69.2	0.708	453.0	-6.4	0.586	501.4	42.0	0.454
CN	413.9	-45.5	0.893	434.6	-24.8	0.801	453.7	-5.7	0.810
CH_3CO	417.0	-42.4	0.896	457.5	-1.9	0.809	462.7	3.3	0.848
Br	443.9	-15.5	0.903	447.2	-12.2	0.947	455.0	-4.4	0.849
I	444.0	-15.4	0.895	448.2	-11.2	0.965	456.2	-3.2	0.845
Cl	444.8	-14.6	0.907	446.6	-12.8	0.935	454.3	-5.1	0.858
CCH	448.3	-11.1	0.866	455.8	-3.6	0.960	462.7	3.3	0.857
F	452.1	-7.3	0.894	445.7	-13.7	0.938	452.0	-7.4	0.872
H	459.4	0.0	0.945	459.4	0.0	0.945	459.4	0.0	0.945
NMe_2	468.8	9.4	0.904	485.8	26.4	1.120	487.1	27.7	0.801
$CHCH_2$	469.6	10.2	0.847	473.2	13.8	0.999	476.3	16.9	0.798
OH	489.8	30.4	0.717	453.0	-6.4	1.017	464.6	5.2	0.889
HO	496.1	36.7	0.819	468.2	8.8	1.051	467.7	8.3	0.882
NCH3H	511.0	51.6	0.756	482.0	22.6	1.105	481.3	21.9	0.777
NHCH3	521.3	61.9	0.680	498.1	38.7	1.092	495.6	36.2	0.822
NH_2	567.9	108.5	0.588	484.1	24.7	1.100	488.8	29.4	0.812

Table B.10: Computed emission wavelengths (λ, in nm) and oscillator strengths (f) of oxyluciferin and its analogs in aqueous solution at TDDFT/CAM-B3LYP/6-311++G**//C-PCM level of theory.

Group	Site C7'			Site C5'			Site C4'		
	λ	$\Delta\lambda$	f	λ	$\Delta\lambda$	f	λ	$\Delta\lambda$	f
NO_2	673.2	97.5	0.587	590.3	14.6	0.983	499.2	-76.5	1.205
CN	584.9	9.2	1.172	564.2	-11.5	1.197	518.2	-57.5	1.279
CH3CO	598.0	22.3	1.164	588.3	12.6	1.165	518.7	-57.0	1.292
CF_3	573.3	-2.4	1.224	559.3	-16.4	1.177	526.3	-49.4	1.311
I	578.5	2.8	1.196	570.6	-5.1	1.345	557.6	-18.1	1.268
Br	576.9	1.2	1.198	568.1	-7.6	1.325	559.0	-16.7	1.268
Cl	575.9	0.2	1.198	566.8	-8.9	1.308	561.9	-13.8	1.261
CCH	587.5	11.8	1.184	580.0	4.3	1.338	562.4	-13.3	1.230
H	575.7	0.0	1.278	575.7	0.0	1.278	575.7	0.0	1.278
F	570.0	-5.7	1.201	560.5	-15.2	1.299	578.0	2.3	1.236
$CHCH_2$	598.0	22.3	1.112	596.4	20.7	1.355	594.4	18.7	1.212
OH-2	582.8	7.1	1.167	582.8	7.1	1.392	653.6	77.9	1.086
OH	578.3	2.6	1.176	569.7	-6.0	1.396	657.4	81.7	1.005
NH_2	604.6	28.9	1.036	637.9	62.2	1.392	814.6	238.9	0.774
NMe_2	607.0	31.3	1.049	653.3	77.6	1.407	872.6	296.9	0.696
NCH_3H	595.1	19.4	0.993	649.4	73.7	1.372	876.0	300.3	0.724
$NHCH_3$	613.2	37.5	1.031	684.7	109.0	1.352	916.5	340.8	0.629

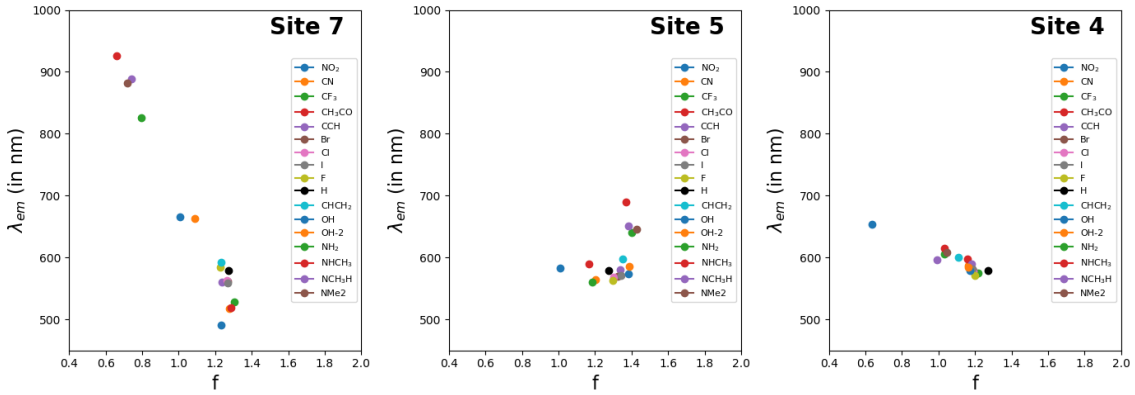


Figure B.3: Computed oscillator strengths and emission wavelengths of oxyluciferin analogs computed using TDDFT/ωB97X-D/6-311++G**//C-PCM

B.4 Effect of electron-donating and withdrawing groups on computed emission spectra of oxyluciferin analogs

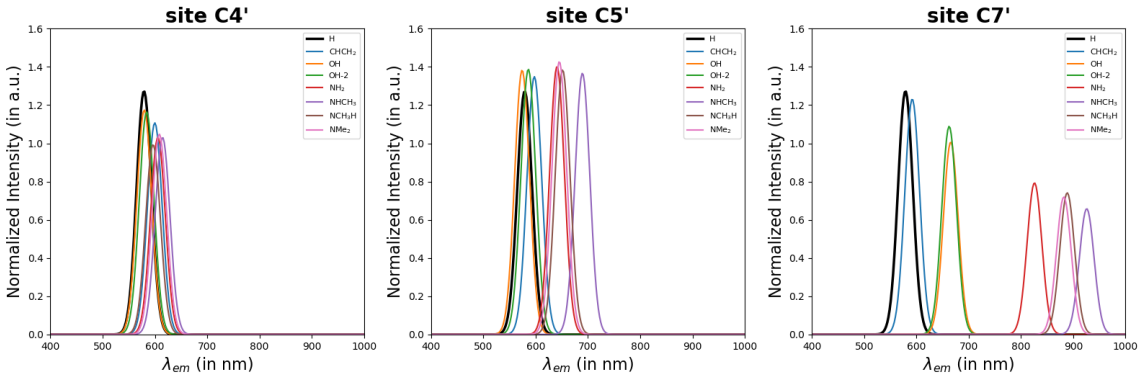


Figure B.4: Effect of electron-donating groups on the computed emission wavelengths (λ_{em}) and oscillator strengths for the first excited state of oxyluciferin analogs computed using TDDFT/ωB97X-D/6-311++G**//C-PCM

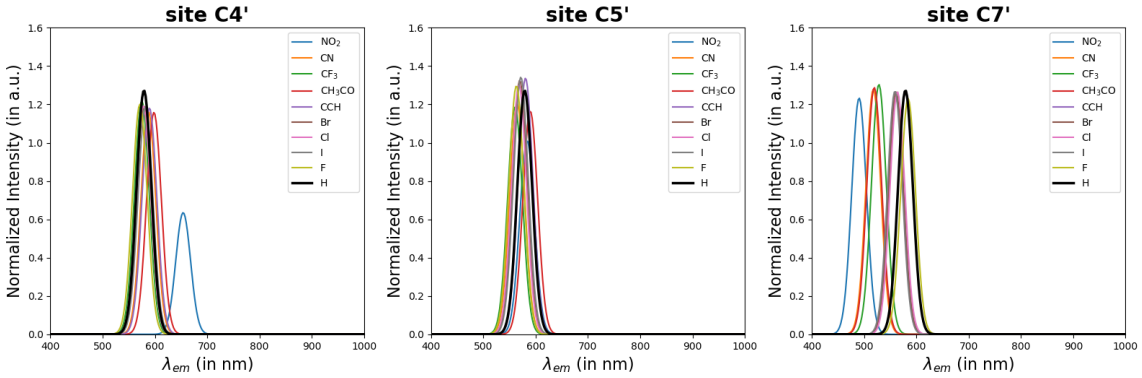


Figure B.5: Effect of electron-withdrawing groups on computed emission wavelengths (λ_{em}) and oscillator strengths for the first excited state of oxyluciferin analogs computed using TDDFT/ωB97X-D/6-311++G**//C-PCM

B.5 Wavefunction analysis

B.5.1 Charge population analysis

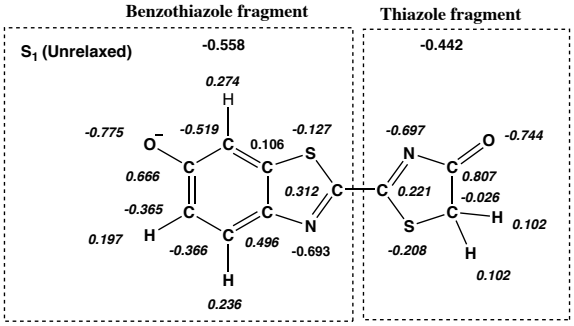


Figure B.6: ESP charges of the S1 state using the unrelaxed electron density obtained from TDDFT calculations with ωB97X-D functional, 6-311++G** basis, and C-PCM solvation model.

Table B.11: Computed Mulliken and ESP charges (in au) of thiazolone ring of oxyluciferin and its analogs using the TD-ωB97X-D/6-311++G**//C-PCM(water) level of theory.

Group	Site C4'		Site C5'		Site C7'	
	Mulliken	ESP	Mulliken	ESP	Mulliken	ESP
NO_2	-0.0769	-0.0682	-0.2350	-0.2239	-0.2501	-0.2544
CN	-0.2267	-0.2284	-0.2661	-0.2624	-0.2459	-0.2506
CH_3CO	-0.2045	-0.2026	-0.2496	-0.2485	-0.2426	-0.2549
CF_3	-0.2324	-0.2326	-0.2535	-0.2530	-0.2322	-0.2411
I	-0.2361	-0.2402	-0.2441	-0.2452	-0.2422	-0.2375
Br	-0.2326	-0.2382	-0.2420	-0.2425	-0.2409	-0.2356
Cl	-0.2325	-0.2386	-0.2405	-0.2404	-0.2403	-0.2357
CCH	-0.2210	-0.2264	-0.2477	-0.2461	-0.2488	-0.2534
F	-0.2325	-0.2402	-0.2272	-0.2307	-0.2356	-0.2351
H	-0.2270	-0.2362	-0.2270	-0.2362	-0.2270	-0.2362
$N(CH_3)_2$	-0.2040	-0.2136	-0.2054	-0.2079	-0.2607	-0.2685
$CHCH_2$	-0.2064	-0.2114	-0.2278	-0.2287	-0.2480	-0.2548
OH-2	-0.2145	-0.2260	-0.2142	-0.2172	-0.2641	-0.2597
OH	-0.2128	-0.2220	-0.2050	-0.2105	-0.2352	-0.2371
$NHCH_3$	-0.1886	-0.2015	-0.1993	-0.2117	-0.2803	-0.2789
NH_2	-0.1932	-0.2103	-0.2044	-0.2092	-0.2705	-0.2613
NCH_3H	-0.2081	-0.2247	-0.2036	-0.2114	-0.2545	-0.2455

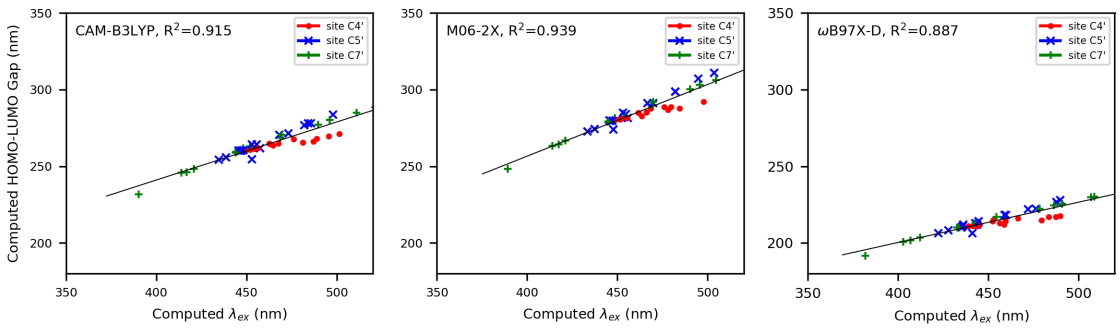


Figure B.7: HOMO-LUMO gap (in nm) vs TDDFT excitation wavelengths (in nm) of oxyluciferin analogs computed using three different functionals (CAM-B3LYP, M06-2X, and ωB97X-D) using 6-311++G** basis set and C-PCM solvent model.

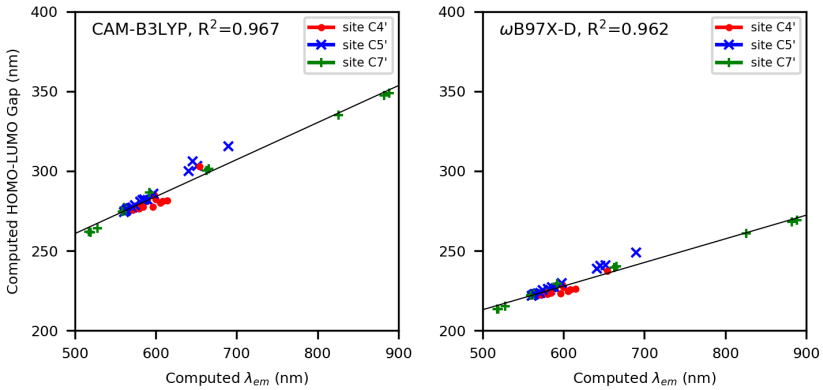


Figure B.8: HOMO-LUMO gap (in nm) vs TDDFT emission wavelengths (in nm) of oxyluciferin analogs computed using CAM-B3LYP and ωB97X-D functional using 6-311++G** basis set and C-PCM solvent model.

Table B.12: Computed HOMO (a.u.), LUMO (a.u.), HOMO-LUMO Gap (a.u), amplitude, and transition dipole moment (μ_{ge}) (debye) for C4'-functionalized oxyluciferin analogs using aqueous solution at TDDFT/ωB97X-D/6-311++G**//C-PCM(water) level of theory

Group	HOMO (a.u.)	LUMO (a.u.)	HOMO-LUMO Gap (a.u.)	Amplitude	μ_{ge} (Debye)
NO_2	-0.261	-0.069	0.192	0.989	3.696
CN	-0.261	-0.060	0.201	0.986	4.744
CH_3CO	-0.256	-0.056	0.199	0.986	4.775
CF_3	-0.259	-0.055	0.203	0.986	4.801
I	-0.257	-0.054	0.203	0.985	4.766
Br	-0.257	-0.054	0.204	0.985	4.765
Cl	-0.257	-0.054	0.204	0.985	4.762
CCH	-0.256	-0.054	0.201	0.985	4.787
F	-0.257	-0.052	0.205	0.985	4.750
NMe_2	-0.248	-0.046	0.202	0.984	4.582
$CHCH_2$	-0.253	-0.052	0.201	0.984	4.675
OH-2	-0.253	-0.049	0.204	0.984	4.731
OH	-0.253	-0.048	0.205	0.984	4.734
$NHCH_3$	-0.247	-0.045	0.202	0.985	4.568
NH_2	-0.248	-0.045	0.203	0.983	4.534
NCH_3H	-0.248	-0.044	0.204	0.984	4.414

Table B.13: Computed HOMO (a.u.), LUMO (a.u.), HOMO-LUMO Gap (a.u), amplitude, and transition dipole moment (μ_{ge}) (debye) for C5'-functionalized oxyluciferin analogs using aqueous solution at TDDFT/ωB97X-D/6-311++G**//C-PCM level of theory

Group	HOMO (a.u.)	LUMO (a.u.)	HOMO-LUMO Gap (a.u.)	Amplitude	μ_{ge} (Debye)
CCH	-0.256	-0.054	0.202	0.983	5.054
NH2	-0.236	-0.046	0.191	0.986	5.438
Br	-0.257	-0.053	0.204	0.984	4.984
CF3	-0.260	-0.055	0.205	0.985	4.679
Cl	-0.257	-0.053	0.204	0.984	4.935
CN	-0.263	-0.059	0.205	0.985	4.730
CH3CO	-0.258	-0.056	0.201	0.985	4.756
F	-0.256	-0.051	0.205	0.984	4.904
I	-0.257	-0.053	0.204	0.983	5.025
NHCH3	-0.229	-0.046	0.183	0.987	5.570
NCH3H	-0.234	-0.045	0.189	0.986	5.449
NMe2	-0.236	-0.046	0.190	0.983	5.508
NO2	-0.266	-0.064	0.202	0.985	4.403
OH	-0.251	-0.049	0.202	0.984	5.113
OH-2	-0.248	-0.048	0.201	0.984	5.174
CHCH2	-0.250	-0.052	0.198	0.982	5.151

Table B.14: Computed HOMO (a.u.), LUMO (a.u.), HOMO-LUMO Gap (a.u), amplitude, and transition dipole moment (μ_{ge}) (debye) for C7'-functionalized oxyluciferin analogs using aqueous solution at TDDFT/ωB97X-D/6-311++G**//C-PCM level of theory

Group	HOMO (a.u.)	LUMO (a.u.)	HOMO-LUMO Gap (a.u.)	Amplitude	μ_{ge} (Debye)
CCH	-0.256	-0.052	0.204	0.981	4.779
NH2	-0.227	-0.052	0.175	0.989	4.643
Br	-0.257	-0.052	0.205	0.983	4.838
CF3	-0.264	-0.052	0.212	0.983	4.762
Cl	-0.257	-0.052	0.205	0.984	4.846
CN	-0.268	-0.055	0.214	0.981	4.670
CH3CO	-0.263	-0.049	0.214	0.981	4.694
F	-0.254	-0.053	0.201	0.985	4.864
I	-0.258	-0.052	0.205	0.982	4.833
NHCH3	-0.221	-0.052	0.169	0.989	4.480
NCH3H	-0.221	-0.052	0.169	0.989	4.656
NMe2	-0.221	-0.051	0.170	0.989	4.570
NO2	-0.275	-0.054	0.221	0.977	3.696
OH	-0.243	-0.053	0.190	0.987	4.698
OH-2	-0.242	-0.051	0.190	0.987	4.874
CHCH2	-0.249	-0.051	0.199	0.981	4.901

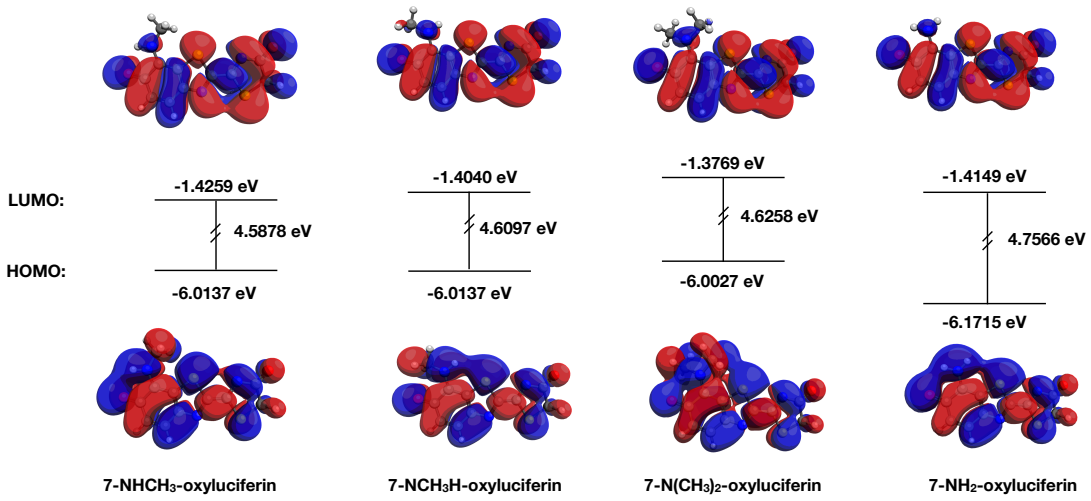


Figure B.9: HOMO and LUMO of oxyluciferin analogs substituted with $NHCH_3$, NCH_3H, $N(CH_3)_2$, and NH_2 groups at C7' site. All orbitals were computed from DFT calculations using the ωB97X-D functional, 6-311++G** basis set, and C-PCM solvent model.

B.5.2 ALMO-EDA analysis

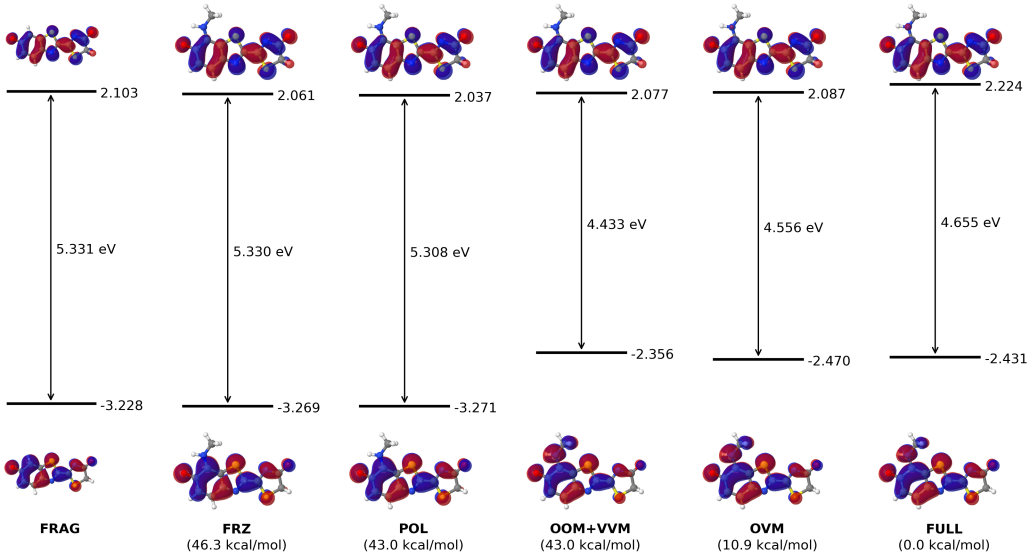


Figure B.10: Effect of $NH(CH_3)$ group on the HOMO and LUMO of oxyluciferin on C7' site. All results were collected from gas–phase ALMO-based calculations using the ωB97X-D functional and 6-31G(d) basis set. The orbital energy values were reported in eV. The relative energies of the intermediate states were shown in the parentheses.

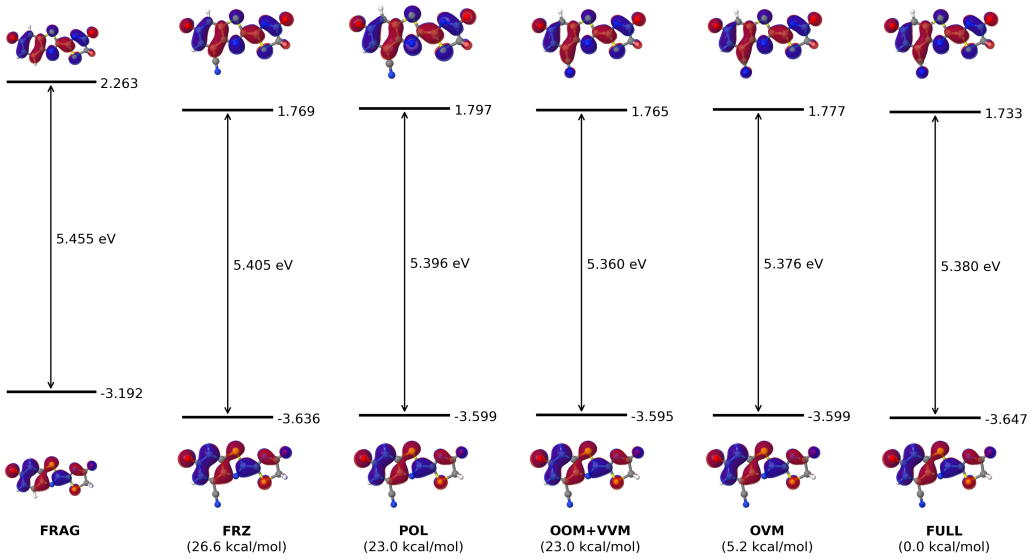


Figure B.11: Effect of CN group on the HOMO and LUMO of oxyluciferin on C7' site. All results were collected from gas–phase ALMO-based calculations using the ωB97X-D functional and 6-31G(d) basis set. The orbital energy values were reported in eV. The relative energies of the intermediate states were shown in the parentheses.

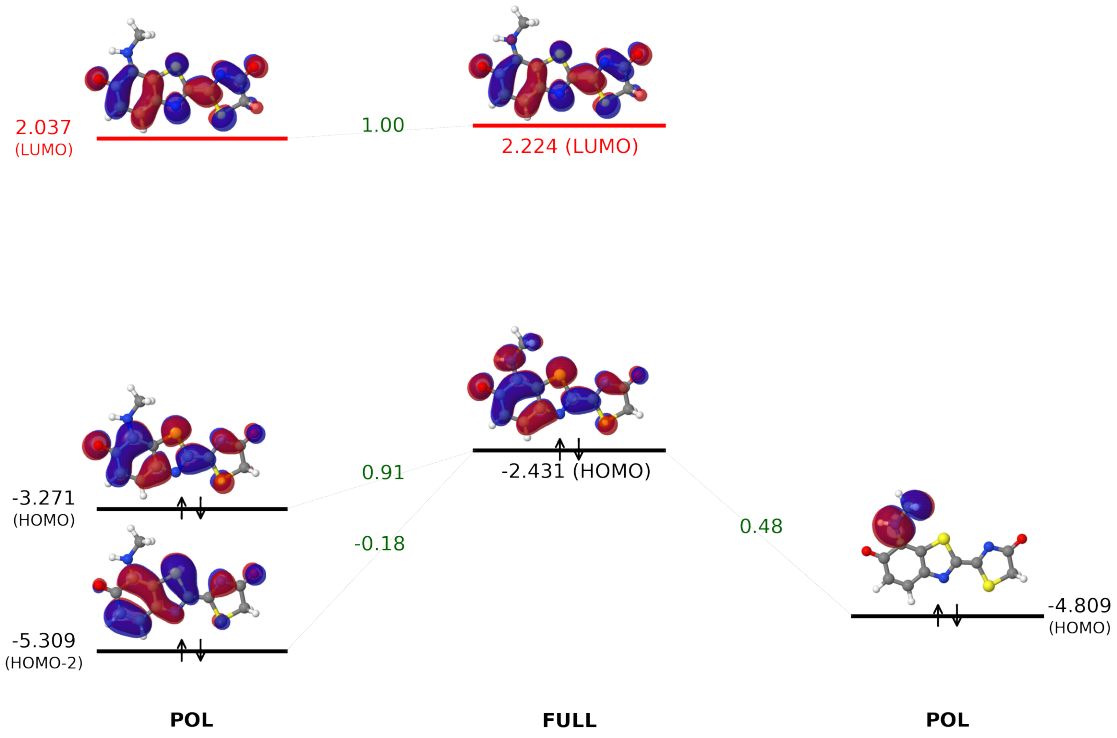


Figure B.12: Interaction between polarized oxyluciferin and amino orbitals to produce fully converged MOs of 7'-NH(CH_3)-oxyluciferin. Computed in the gas–phase using the ωB97X-D/6-31G(d) level of theory. All orbital energies are in eV.

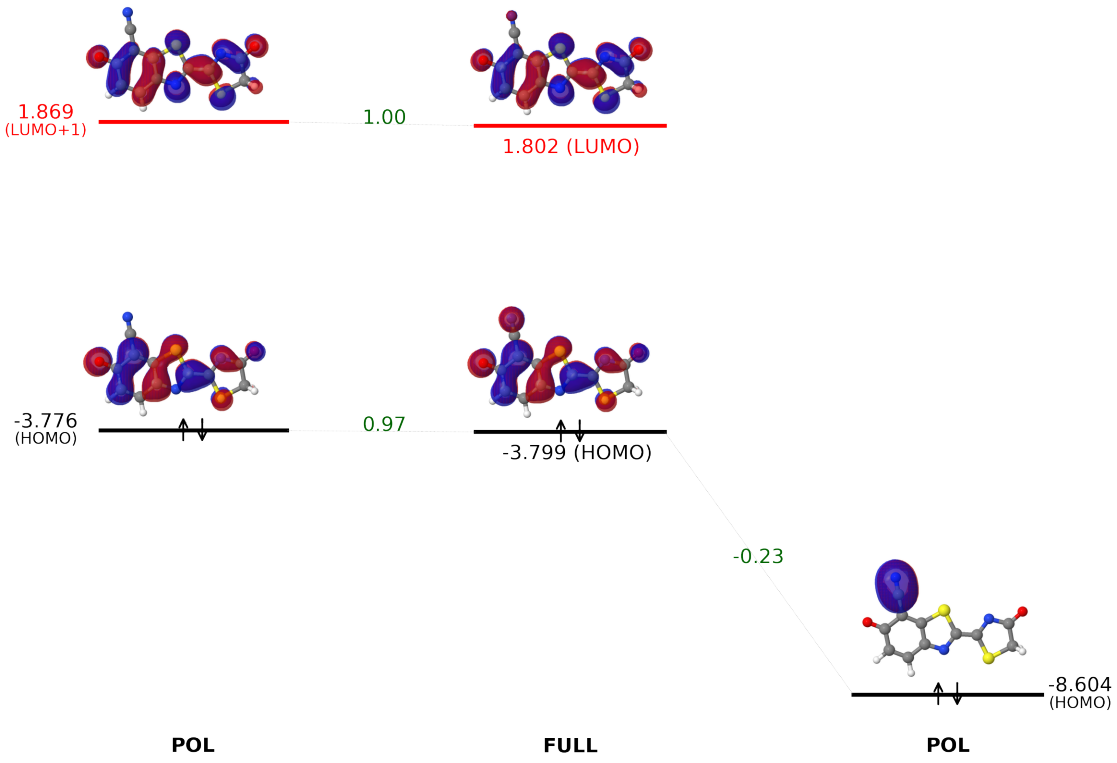


Figure B.13: Interaction between polarized oxyluciferin and amino orbitals to produce fully converged MOs of 7'-cyano-oxyluciferin. Computed in the gas–phase using the ωB97X-D/6-31G(d) level of theory. All orbital energies are in eV.

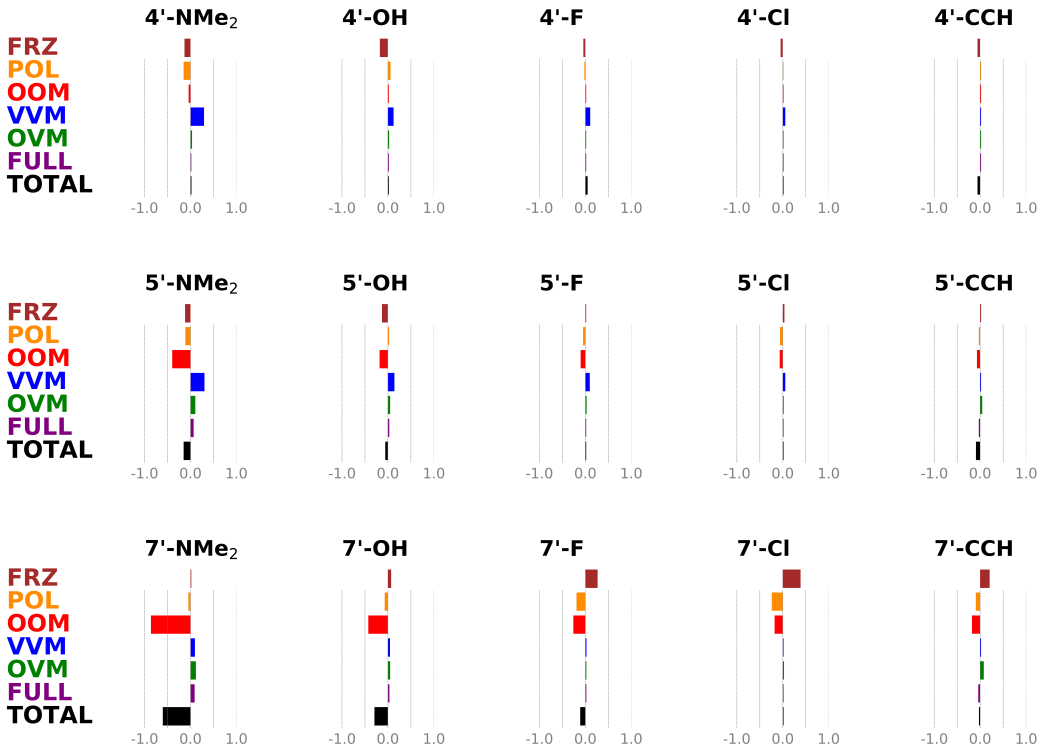


Figure B.14: Decomposition of the effects of the NMe_2, hydroxy, fluoro, Chloro, and ethenyl group on the HOMO and LUMO of the oxyluciferin system obtained from the ALMO-based analysis. All data were computed at the ωB97X-D/6-31G(d) level of theory. Energy values are in eV.

Appendix C

Supplemental Computational Results of C–H Activation Reaction

Table C.1: Relative energies of benzyne-INT$_1$ and aryl cation-INT$_1$ structures at ωB97X-D/6-31+G* level of theory using C-PCM (ODCB) solvation model.

Benzyne-INT (1) (a.u.)		Aryl cation-INT (2) (a.u.)	ΔE (2) - (1) (a.u.)	ΔE (2) - (1) (kcal/mol)
Benzyne	TMS-WCA	AR-TMS-WCA		
-230.8208	-5783.7556	-6014.6663	-0.0899	-56.4
-6014.5849		-6014.6663	-0.0814	-51.1