

COMPUTATIONAL ANALYSIS OF 1,2-DIALKYNILPYRROLE ANALOGUES

HONORS THESIS

Presented to the Honors College of
Texas State University
in Partial Fulfillment
of the Requirements

for Graduation in the Honors College

by

Joshua Bondoc

San Marcos, Texas
May 2018

COMPUTATIONAL ANALYSIS OF 1,2-DIALKYNILPYRROLE ANALOGUES

by

Joshua Bondoc

Thesis Supervisor:

Sean M. Kerwin, Ph.D.
Department of Chemistry and Biochemistry

Approved:

Heather C. Galloway, Ph.D.
Dean, Honors College

Abstract

Computational chemistry can be used to quickly predict the properties of molecules. We are interested in studying the potential of 1,2-dialkynylpyrroles to undergo a thermal Bergman cyclization-triggered rearrangement to reactive diradical intermediates that display cytotoxic activity. These diradical intermediates have garnered much interest in the scientific community as potential anti-tumor drugs. A (U)B3LYP hybrid functional with a 6-31G** basis set was used to predict the electronic energies of each intermediate and transition state in the reaction. The electronic energies were used to construct a reaction coordinate that quantified the energy gap between these intermediates and transition states. The goal of this experiment was to create an analogue of a 1,2-dialkynylpyrrole that maximized the energy gap between the diradical singlet and the retro Bergman transition state while simultaneously minimizing the energy gap between the diradical singlet and the diradical triplet. An analogue that produced a diradical with these properties would be long lived and reactive. We discovered that the addition of electron withdrawing groups to the C3 carbon of the starting 1,2-dialkynylpyrrole lowered the singlet-triplet gap and increased the energy barrier for the retro Bergman transition state. Electron withdrawing groups placed in close proximity to the retro Bergman cleavage site may affect the energies of the bonding and antibonding orbitals associated with that bond. This study provides insight into the further optimization of diradicals to be used for therapeutic applications.

Background

An Introduction to Diradicals

Research in targeted cancer treatment therapy is becoming more popular in the field of oncology.¹ Many current chemotherapy regimens have immunosuppressive side effects, leading to secondary diseases that further compromise the health of the patient.¹ These immunosuppressive side effects are a result of the drug not being specific enough to tumor cells, leading to the unintentional destruction of healthy cells, resulting in a weakened immune system.¹ Scientists are constantly trying to create tumor selective drugs to reduce the side effects of chemotherapy.¹

An effective antitumor drug should be both selective and cytotoxic. The drug must be toxic enough to destroy tumor cells, but selective enough to only target those cells. Scientists have compounds that fit these criteria, however these drugs are only effective against a specific type of cancer.² One example is the drug inotuzumab ozagamicin, an antibody-drug conjugate that causes double-stranded DNA breaks in cancerous lymphoblasts and is used to treat patients suffering from Acute Lymphoblastic Leukemia.² While inotuzumab ozagamicin is effective in treating patients, its use is only restricted to that single type of cancer.² Scientists have yet to discover effective treatments for other types of cancer, many of which have no treatment at all.²

Diradicals are a class of cytotoxic compounds that are garnering much interest in the scientific community.³ Their ability to cause double-stranded DNA breaks makes diradicals extremely cytotoxic, and a potential candidate for use in anti-tumor drugs if more selective compounds can be synthesized.³ One of the biggest problems with diradicals, however, is their instability.⁴ Diradicals are typically seen as intermediate

compounds, such as in Bergman Cyclization, and only exist in a small window of time before decomposing.⁴ The ephemeral nature of diradicals undermines their inherent reactivity. Computational chemistry can be used to predict the properties of diradicals in a short amount of time, allowing scientists to optimize diradicals for their intended use.

The Becke, three-parameter, Lee-Yang-Parr (B3LYP) method is a hybrid functional theory that combines elements of Density Functional Theory (DFT) and Hartree-Fock (HF) Theory in order to solve the Schrodinger equation.⁵ The Schrodinger equation is a wave equation that solves for the wave functions of a molecule.⁵ Many electronic properties of a molecule can be predicted when the Schrodinger equation is solved, based on the resulting wave functions.⁵ In the case of diradicals, relative longevity and relative reactivity can be estimated, which is helpful to scientists when determining the effectiveness of a drug. The B3LYP method uses HF theory to calculate the exchange correlation energy and uses DFT to calculate the dynamic electron correlation energy, both of which are critical energy values needed to solve the Schrodinger Equation.⁵

Computational predictions have been used to evaluate diradicals for potential use in anticancer drugs.³ The computationally obtained energy diagram for the decomposition of 1,2-diethynyl-1*H*-imidazole to 6,7-dihydro-5*H*-cyclopenta[b]pyrazine that was predicted using B3LYP calculations, was in agreement with experimentally obtained kinetic data for the same decomposition. The similarity between the computational data and the experimental data supports the idea that B3LYP calculations can accurately predict the reaction coordinate of similar molecules. My thesis extends this idea to analogues of 1,2-diethynyl-1*H*-pyrrole, based on the assumption that these similar

analogues will produce experimental data that are in agreement with computational data generated from B3LYP calculations. By using computational data to predict experimental data, high throughput methods of drug analysis can be performed, which saves time and resources.

Certain aspects of the diradical energy diagram can be used to predict the diradical's properties. The energy gap between the diradical singlet intermediate and the retro Bergman transition state determines the longevity of the diradical.⁶ A high transition state barrier for the diradical increases its longevity due to the increase in energy required to reach the transition state.⁶ The singlet-triplet energy gap determines the propensity of the diradical to undergo hydrogen abstraction.⁶ A smaller singlet-triplet energy gap allows for an easier transition between the singlet and triplet states, thus increasing the reactivity of the triplet diradical.⁶ The goal of my ongoing project is to identify a diradical that maximizes the energy gap between the diradical singlet and the retro Bergman transition state and minimizes the singlet-triplet energy gap. Such a diradical would be both stable and long lived. Once an optimized diradical is identified, future experiments would be conducted to test cytotoxicity, and to increase selectivity against cancer cells.

An Introduction to Computational Chemistry

Hartree-Fock (HF) is a computational method that is used to gather theoretical data on the electronic structure of a molecule by accurately estimating its Schrodinger Equation. By doing so, the wave function of a molecule can be retrieved and then used to derive properties of that molecule, such as its electronic energy. However, accurately estimating the Schrodinger Equation is time consuming, even for powerful computers.

The HF method addresses this problem and makes certain approximations in order to approximate a solution to the Schrodinger Equation.

The first approximation that the HF method makes is the Born Oppenheimer Approximation, which states that the electronic motion, and the nuclear motion can be separated. While the Schrodinger equation uses one equation to account for both the electronic motion, and the nuclear motion, the Born Oppenheimer Approximation separates these factors into two separate equations. While the Born Oppenheimer Approximation creates more equations that need to be estimated for the Schrodinger Equation, it also reduces the complexity of these equations, leading to many simple equations, rather than one complex equation. However, the HF method is not perfect and will always underestimate the electronic energy of a molecule because it does not account for individual electron-electron repulsions.

The second approximation that the HF method makes is the Mean Field Approximation, which treats the forces from all electrons and nuclei on a target electron as a single average force. This approximation also reduces the number of calculations needed to accurately estimate the Schrodinger Equation, but reduces the accuracy of the calculation. Rather than an electron being subjected to a dynamic environmental force caused by other moving electrons, the HF method models electrons moving in a constant plane of force equal to the average forces from all electrons and nuclei. Where HF fails in calculating the dynamic electron correlation, it makes up for in being able to calculate the electron exchange correlation. The electron exchange correlation is the part of the Schrodinger Equation that accounts for all of the quantum mechanical interactions

between electrons. Because HF is a quantum model, it intrinsically takes quantum forces into consideration.

Density Functional Theory (DFT) is another computational method that is also used to analyze electronic structures by accurately estimating the Schrodinger Equation, similar to the HF method. DFT also makes the Born Oppenheimer Approximation, and the Mean Field Approximation, but estimates the Schrodinger Equation by using information from the electron density, rather than from the sum of many single electron wave functions, as seen in the HF method. DFT relies on the Hohenberg and Kohn Theories, which state that the ground state energy is a unique functional of the electron density, and that the electron density that minimizes the overall functional is the true ground state electron density. A functional is defined as a function of a function, similar to how a derivative is a function based off of another function. These theories imply that the Schrodinger Equation can be solved using the electron density alone, bypassing the need for multiple electron wavefunctions. Using electron density values to solve the Schrodinger equation also yields a more accurate value for the dynamic electron correlation, a value that is less accurate in the HF model. However, DFT fails where HF excels, as the DFT model relies on approximations for the electron exchange correlation. The DFT model has a problem opposite to that of the HF model.

The Becke, three-parameter, Lee-Yang-Parr (B3LYP) method is a hybrid functional that combines elements of DFT and HF methods to create a more balanced computational protocol that is both quick and accurate. By playing on the strengths of both methods, the B3LYP hybrid method is able to provide a more accurate estimation of the Schrodinger Equation. The basic B3LYP calculation only requires the cartesian

coordinates of the atoms that comprise the molecule as input, in addition to the overall charge and multiplicity of the atom. Different parameters can also be passed along with the standard B3LYP to conduct more specific calculations. The ultimate goal of the calculations was to produce a wavefunction that could reveal properties of the molecule being tested.

Methods

The variable that was tested in this experiment was an R group placed on the 3-carbon. R groups are defined as chemical structures attached to a parent molecule that serve as slight modifications to that parent. In order to analyze the effects of varying R groups on the longevity and reactivity of a diradical, a reaction coordinate was constructed from the electronic energies of each stage in the reaction, including intermediate structures and transition states. To obtain the electronic energy of a particular species in the reaction coordinate, an educated guess of the geometry was first needed. From this guess, the computer optimized that geometry to its most stable state. From the optimized geometry, the computer estimated that molecule's Schrodinger equation to retrieve the wave function of its optimized geometry. Finally, from this wave function, the electronic energy for that structure was identified.

All calculations were performed using Gaussian09 at the Beck, three-parameter, Lee-Yang-Parr (B3LYP) level of theory with a 6-31G** basis set. The computer used to perform these calculations was a 32-node Sun server.

The order of the structures as they appear in the reaction coordinate is as follows:

1. Eneidyne
2. Bergman Transition State
3. Diradical
4. Retro Bergman Transition State
5. Cumulene
6. Cumulene-Carbene Transition State
7. Carbene

Note that each calculation is performed separately and does not necessarily need to be completed in the order shown above.

Eneidyne Calculation

The geometry guess for the substituted enediyne was obtained by taking the unsubstituted enediyne and replacing the hydrogen on the 3-carbon with the new R group as depicted in Figure 1. A geometry optimization, followed by a frequency calculation, was performed on the guess geometry. The electronic energy was subsequently obtained from the output file of the frequency calculation.

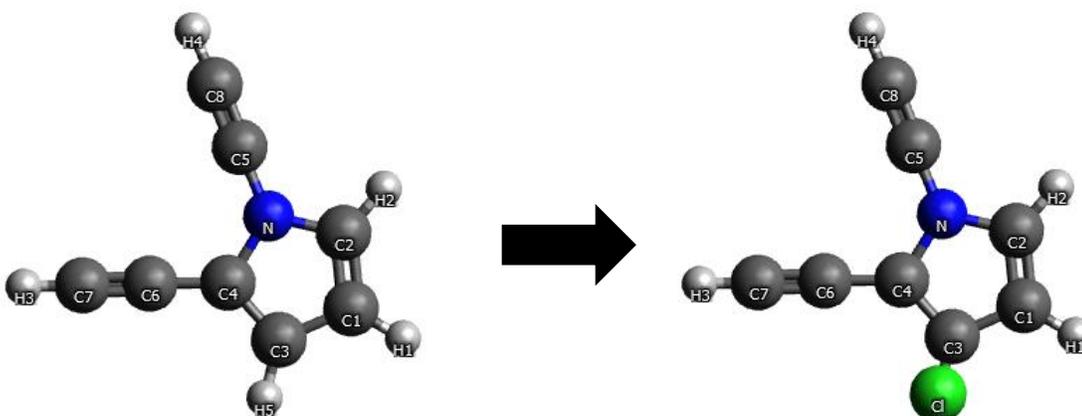


Figure 1. Substitution of the C3 hydrogen on the parent enediyne with an R group (Chlorine) using Avogadro editing and visualization software.

Bergman Transition State Calculation

To obtain the guess geometry for the Bergman Transition State, the optimized geometry of the enediyne was taken, and the distance between the two atoms involved in the cyclization step was incrementally shortened through a ModRedundant calculation. The ModRedundant calculation shortens or lengthens the distance between two specified atoms, performs a geometry optimization, and then retrieves the energy of that structure. The guess geometry of transition states for this system was defined as the geometry at which the distance between the two bond-forming/breaking atoms produced the highest electronic energy. A geometry optimization, followed by a frequency calculation, was performed on this guess geometry. The electronic energy was subsequently obtained from the output file of the frequency calculation.

Diradical Calculation

The form of the diradical in question must be considered when calculating its energy. Diradicals exist as a superposition of the singlet and triplet states, so when performing calculations on diradicals, it is important to separate the individual energy values of the singlet and triplet states, each of which have different geometries and, as a result, different wave functions. The difference between the singlet energy and the triplet energy defines the singlet-triplet gap and is used to determine the reactivity of the diradical as a whole.

It should also be noted that an unrestricted B3LYP (UB3LYP) level of theory was applied to all diradical calculations. Because B3LYP simplifies the Schrodinger equation by pairing electrons, wavefunctions that contain unpaired electrons are not possible.

Because diradicals contain unpaired electrons, UB3LYP must be applied to allow those wavefunctions to be possible.

The geometry guess for the substituted singlet diradical was obtained by taking the unsubstituted singlet diradical and replacing the hydrogen on the 3-carbon with the new R group as shown in Figure 2. A geometry optimization was performed. Next, the highest occupied molecular orbital (HOMO), and the lowest unoccupied molecular orbital (LUMO) were identified using Avogadro, a visualization software. In the following frequency calculation, a Guess=Alter parameter was passed to move one electron from the HOMO to the LUMO thereby forcibly inducing the spin-unpaired or “broken” spin state. The electronic energy was subsequently obtained from the output file of the frequency calculation.

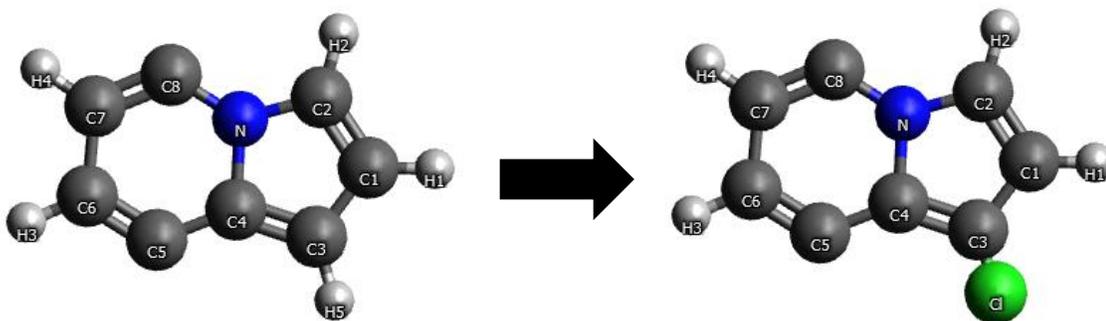


Figure 2. Substitution of the C3 hydrogen on the parent singlet diradical with an R group (Chlorine) using Avogadro editing and visualization software.

Unfortunately, singlet diradical wave functions are inherently contaminated with triplet characteristic; this is known as spin contamination. To correct this spin contamination, a triplet calculation was performed on the optimized geometry of the singlet diradical. The electronic energy from the resultant wave function was subtracted from that of the spin contaminated singlet. By subtracting the triplet contamination from

the contaminated singlet diradical, the electronic energy of the spin corrected singlet diradical was obtained.

The guess geometry for the substituted triplet diradical was obtained from the optimized geometry of the substituted singlet diradical. The geometry of the singlet, and the geometry of the triplet were assumed to be close, but not identical in structure. A geometry optimization, followed by a frequency calculation, was performed on the guess geometry. The electronic energy was subsequently obtained from the output file of the frequency calculation.

Retro Bergman Transition State Calculation

To obtain the guess geometry for the retro Bergman Transition State, the optimized geometry of the broken spin singlet diradical and its wavefunction were taken, and the distance between the two atoms involved in the bond cleavage step was incrementally lengthened through a ModRedundant calculation. Using this guess geometry, a single point calculation was performed, followed by the identification of the HOMO and the LUMO using Avogadro. In the following calculation, the Guess=Alter parameter was passed to move one electron from the HOMO to the LUMO, thereby forcibly inducing the broken spin state. A frequency calculation was performed on the guess geometry of the retro Bergman transition state using the broken spin wavefunction. The electronic energy was subsequently obtained from the output file of the frequency calculation.

Cumulene Calculation

The geometry guess for the substituted singlet diradical was obtained by taking the unsubstituted cumulene and replacing the hydrogen on the 3-carbon with the new R group as shown in Figure 3. A geometry optimization, followed by a frequency calculation, was performed on the guess geometry. The electronic energy was subsequently obtained from the output file of the frequency calculation.

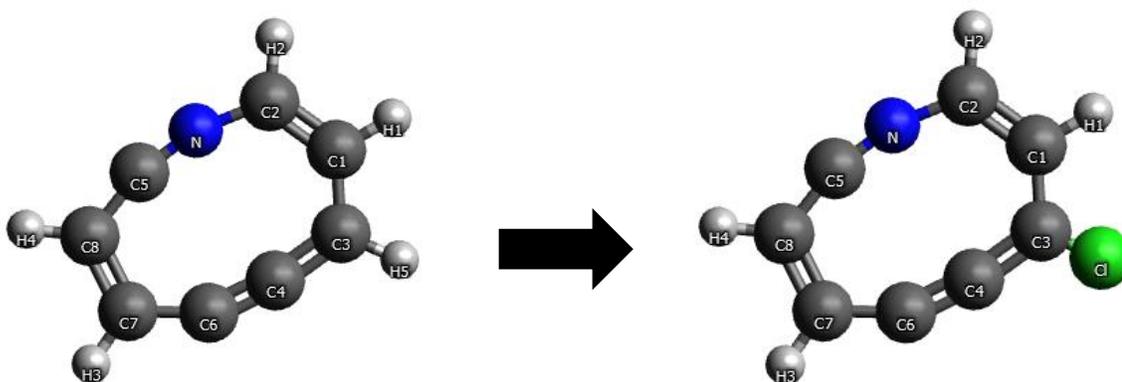


Figure 3. Substitution of the C3 hydrogen on the parent cumulene with an R group (Chlorine) using Avogadro editing and visualization software.

Cumulene-Carbene Transition State Calculation

The geometry guess for the cumulene-carbene transition state was obtained by taking the optimized geometry of the cumulene and compressing the atoms involved in the cyclization step through a ModRedundant calculation. A geometry optimization, followed by a frequency calculation, was performed on this guess geometry. The electronic energy was subsequently obtained from the output file of the frequency calculation.

Carbene

Carbenes, like diradicals, are a superposition of singlet and triplet states. As a result, two separate wavefunctions had to be obtained for the two separate states the carbene could exist in.

The geometry guess for the substituted singlet diradical was obtained by taking the unsubstituted carbene and replacing the hydrogen on the 3-carbon with the new R group as shown in Figure 4. A geometry optimization, followed by a frequency calculation was performed on this guess geometry. The electronic energy was subsequently obtained from the output file of the frequency calculation.

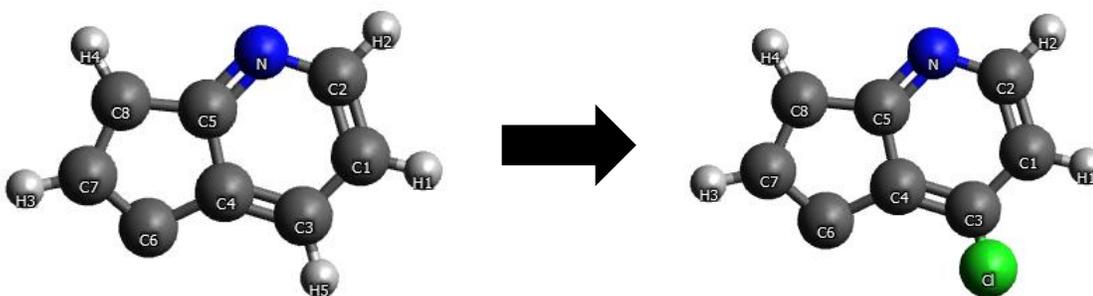


Figure 4. Substitution of the C3 hydrogen on the parent singlet carbene with an R group (Chlorine) using Avogadro editing and visualization software.

The guess geometry for the substituted triplet carbene was obtained from the optimized geometry of the substituted singlet carbene. The geometry of the singlet, and the geometry of the triplet were assumed to be close, but not identical in structure. A geometry optimization, followed by a frequency calculation, was performed on the guess geometry. The electronic energy was subsequently obtained from the output file of the frequency calculation.

Results

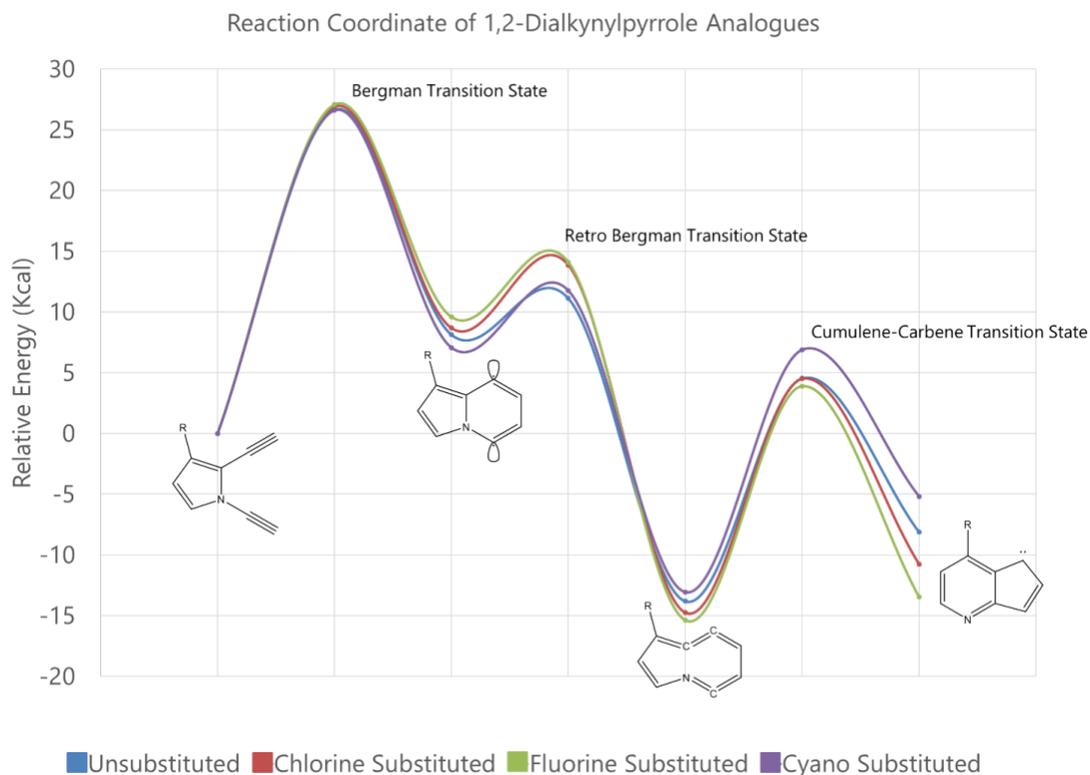


Figure 5. The reaction coordinate of each analogue was plotted. All predicted electronic energies were normalized to the enediyne.

The data show that substitution of the R group for electron withdrawing groups increases the singlet-transition state gap. The chlorine and fluorine substituted analogues both showed an increase in both the energy of the diradical as well as the retro Bergman transition state relative to the unsubstituted parent, but with a higher increase in the energy of the retro Bergman transition state relative to the diradical. The cyanide substituted analogue showed a decrease in the energy of the diradical and an increase in the energy of the retro Bergman transition state relative to the unsubstituted parent.

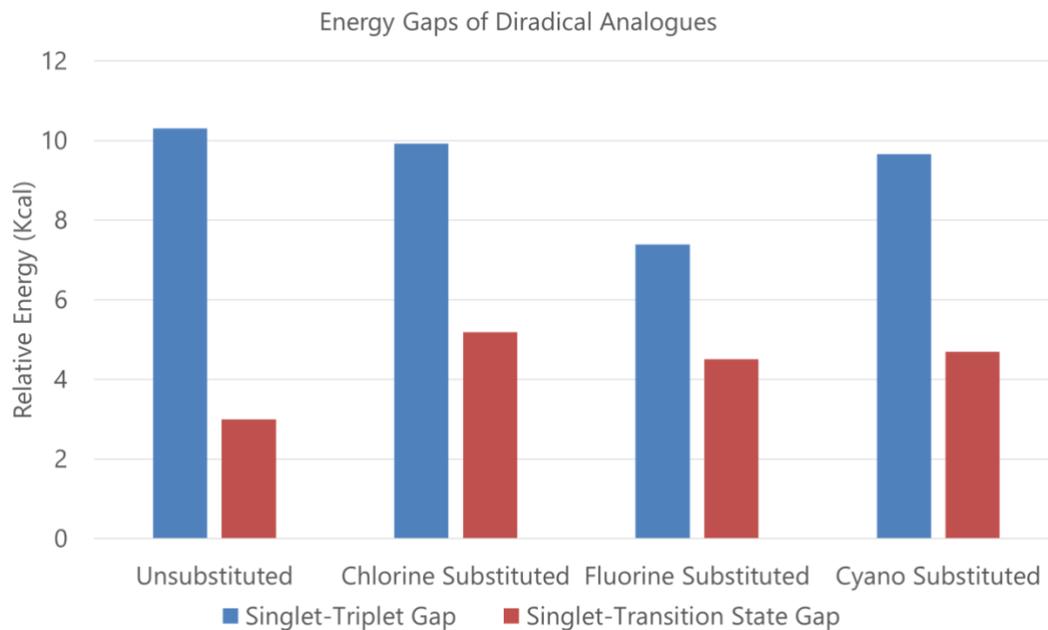


Figure 6. The predicted energy values for the singlet-triplet gap and the singlet-transition state gap were plotted for all analogues. All predicted electronic energies were normalized to the enediyne.

The data show that all three analogues both lowered the energy of the singlet-triplet gap and increased the energy of the singlet-transition state gap relative to the unsubstituted parent. The fluorine substituted analogue showed the greatest decrease in the singlet-triplet gap, while the chlorine substituted analogue showed the greatest increase in the singlet-transition state gap.

Discussion

Troubleshooting Calculations

The validity of the data depended on the retrieval of a stable wavefunction from the optimized geometry. Often times, the standard protocol was not sufficient enough to obtain a stable wave function, or even to obtain an optimized geometry. In these cases, certain modifications to the protocol were made.

In cases where an optimized geometry could not be obtained from the guess geometry, the `scf=qc` parameter was passed. This parameter modified the self-consistent field procedure to incorporate a quadratic convergence as opposed to the standard linear convergence. This modification allowed more difficult geometry optimizations to be conducted at the expense of making the calculation longer and more resource intensive.

Maintaining the broken spin state for the resultant wavefunction when performing calculations on the diradical and the retro Bergman transition state was necessary to obtain accurate data. In the standard protocol, the HOMO and LUMO were identified by hand through visualization software, and then switched using the `Guess=Alter` parameter. In cases where this protocol failed to produce a broken spin wavefunction, the `Guess=Mix` parameter was passed instead. This parameter used built in algorithms to computationally identify the HOMO and LUMO and switch them. This modified protocol would typically be able to identify the correct orbitals to switch when manual methods failed.

Following the retrieval of a wavefunction from the optimized geometry, it was important to check the stability of that wavefunction. Even though the optimized geometry produced a wavefunction, it was not guaranteed to be stable. To check the stability of the wavefunction, a single point calculation with the `Stable=Opt` parameter was performed on

that wavefunction. The Stable=Opt parameter not only checked the stability of the wavefunction, but also attempted to correct it if an instability was identified.

Using Electron Withdrawing Groups to Optimize Diradicals

The presence of an electron withdrawing group in close proximity to the C2-N bond that undergoes homolytic cleavage when going from the diradical to the cumulene seems to have an effect on the bonding and antibonding orbitals associated with that bond. Because the singlet-triplet gap was observed to be lower in energy when electron withdrawing groups were present, we can deduce that the energy of the singlet diradical's C2-N antibonding orbital was raised.

The presence of electron withdrawing groups increases the energy of the C2-N antibonding orbital. Increasing the energy of the antibonding orbital decreases the interaction between the electrons of the C2-N bond and the radical centers, which increases the energy of the singlet state. Because the triplet state is higher in energy than the singlet state, increasing the energy of the singlet state closes the energy gap between the singlet and the triplet, resulting in a lower singlet-triplet gap. Diradicals with a lower singlet-triplet gap are expected to be more reactive.

The presence of electron withdrawing groups also lowers the energy of the C2-N bonding orbital. Lowering the energy of the bonding orbital increases the energy required for homolytic cleavage of the C2-N bond, which increases the energy of the retro Bergman transition state. Increasing the energy of the retro Bergman transition state increases the singlet-transition state gap. Diradicals with a higher singlet-transition state gap are expected to be longer lived.

Future Studies

The extent of my research focused only on optimizing the diradical for predicted increases in stability and reactivity. Future studies would include synthesizing some of the analogues to confirm the viability of using the (U)B3LYP level of theory as an accurate predictor of the reaction coordinate for the Bergman and retro Bergman cyclization of 1,2-Dialkynylpyrrole analogues. Should the results of those future studies support the future use of the (U)B3LYP level of theory, the next step would be to screen synthesized analogues for cytotoxicity, before finally developing ways to increase selectivity towards cancer cells.

References:

1. Galluzzi, L., Buque, A., Kepp, O., Zitvogel, L., & Kroemer, G. (2015) Immunological effects of conventional chemotherapy and targeted anticancer agents. *Cancer cell*. 28, 690-714
2. Kantarjian, H., Ravandi, F., Short, N. J., Huang, X., Jain, N., Sasaki, K., ... & Alvarado, Y. (2018) Inotuzumab ozogamicin in combination with low-intensity chemotherapy for older patients with Philadelphia chromosome-negative acute lymphoblastic leukaemia: a single-arm, phase 2 study. *The Lancet Oncology*. 19, 240-248
3. Laroche, C., Li, J., Gonzales, C., David, W. M., & Kerwin, S. M. (2010) Cyclization kinetics and biological evaluation of an anticancer 1, 2-dialkynylimidazole. *Organic & biomolecular chemistry*. 8, 1535-1539
4. Logan, C. F., & Chen, P. (1996) Ab initio calculation of hydrogen abstraction reactions of phenyl radical and p-benzyne. *Journal of the American Chemical Society*. 118, 2113-2114
5. Houk, K. N., Liu, J., DeMello, N. C., & Condroski, K. R. (1997) Transition states of epoxidations: diradical character, spiro geometries, transition state flexibility, and the origins of stereoselectivity. *Journal of the American Chemical Society*. 119, 10147-10152
6. Chen, T., Zheng, L., Yuan, J., An, Z., Chen, R., Tao, Y., & Huang, W. (2015) Understanding the control of singlet-triplet splitting for organic exciton manipulating: a combined theoretical and experimental approach. *Scientific reports*. 5, 10923