Theoretical studies on the Annexin A1-S100A11 Heterotetramer and Platinum-Mediated C-H Activation Pathways

By

Wesley Sanchez

July 2023

Director of Thesis: Dr. Yumin Li

Major Department: Chemistry

Computational chemistry is a powerful tool for characterizing chemical phenomena when they are difficult to observe in laboratory settings or simply elusive. By employing the fundamental equations of nature, computational chemistry simulates a system's behavior through efficient computer programs and algorithms. In this work, two computational approaches were used to investigate biologically important systems.

In Chapter 3, molecular dynamics simulations were performed on the Annexin A1-S100A11 heterotetramer. Annexins are important calcium binding proteins which have implications in calcium-regulation and numerous pathologies including many types of cancer. Annexin-A1 is known to interact with S100A11 to form a heterotetramer which is involved in regulating EGFR and thus tumor growth. In this work, the Annexin A1-S100A11 heterotetramer was modeled for the first time and subjected to multiple MD simulations to characterize the complex on the atomic level. Principle component analysis revealed three conformations of the Annexin N-terminus stemming from a kink formed through W12's interaction with M63. We identified a consistent, stable binding mode between the first 11 residues of A1 and S100A1 which was consistent to available crystal structures for A1-S100A11 and A2-S100A10. This work suggests that this stable binding mode is potentially a theme for other Annexin-S100

complexes and that the flexibility of the A1-ND affords multiple possible conformations of the A1t.

In Chapters 4 and 5, quantum mechanics calculations were performed in conjunction with the Nudged Elastic Band method to investigate the mechanisms of platinum mediated C-H activation. Direct conversion of C-H bonds into other bond types is inherently difficult owing to the inertness of standard C-H bonds. In recent years, transition metal-catalyzed C-H functionalization has grown in popularity due to its numerous applications in pharmaceuticals and chemical industry. Due to their versatile electronic properties, platinum complexes have found applications ranging from OLED technology to medical imaging and anticancer treatments. In this work, computational approaches were employed on two pathways: 1) Pt-catalyzed C-H acylation of 2-(2-methylphenoxy)-pyridine and 2) solvent-controlled sp^2/sp^3 C-H activation of N-methyl-N-phenyl-6-(1H-pyrazol-1-yl)pyridine-2-amine. In the former, ligand exchange for cis isomers was found to be kinetically favored which was congruent with experimental observations. In the latter, several key intermediates were identified for solvent-controlled cycloplatination reactions. With a catalytic water molecule, a significant reduction in energy barrier was observed for sp^2 reactions. Additionally, water was observed to facilitate a square pyramidal intermediate via oxidative addition to the platinum center in sp^3 pathways.

Theoretical studies on the Annexin A1-S100A11 Heterotetramer and Platinum-Mediated C-H Activation Pathways

A Master's Thesis

Presented to the Faculty in the Department of Chemistry

East Carolina University

In Partial Fulfilment of the Requirements for the Degree

Master's of Science Degree in Chemistry

By

Wesley Andres Sanchez

July 2023

Director of Thesis: Yumin Li, PhD

Thesis Committee Members:

Libero Bartolotti, PhD

Shouquan Huo, PhD

Xiaoping Pan, PhD

Acknowledgements

First and foremost, I would like to thank my thesis advisor Dr. Yumin Li, whom I have worked with since Fall 2020, for her endless support throughout my time as a graduate student. She has taught me countless lessons not only about chemistry but also about life. I joined her lab without knowing what computational chemistry was as a field let alone that Linux was an operating system. Despite all of my mistakes along the way, she has always been kind and patient and helped me grow as a researcher.

I would also like to thank my committee, Dr. Shouquan Huo, Dr. Xiaoping Pan, and Dr. Libero Bartolotti for their insights into my research. I would like to especially thank Dr. Bartolotti for his mentorship and enthusiasm for science. I would like to thank him for always being available to talk about science or help our lab with whatever technical problems we experienced. I especially enjoyed our long talks about tech support, quantum chemistry, and Southern cooking. Without him, I would have never found my passion for electronic structure theory and coding, and for that I will be forever grateful.

I am also thankful for Dr. Joi Walker, Dr. Eli Hvastkovs, Dr. Brian Love, and all the other faculty whom I have collaborated with over the years. I would also like to thank the current members of the Li lab for their help and support: Samuel Lindsay and Austin Gates. I thank both of them for being good friends and colleagues. Last but not least, I would like to express my gratitude for my mother, brother, and all of the friends who have supported me along the way. Thank you for all the love and support.

Table of Contents

Title Page i
Copyright Page ii
Acknowledgements iii
Table of Tables vi
Table of Figures vii
Table of Equations viii
CHAPTER 1: INTRODUCTION 1
1.1 Introduction 1
1.2 Annexin Proteins 2
1.2.1 Annexin A1 2
1.2.2 Annexin and S100 Proteins 3
1.3 C-H Functionalization 4
1.3.1 Methods of Functionalization 6
1.3.2 Transition Metal Mediated Functionalization 7
1.3.3 Platinum Complexes 10
CHAPTER 2: METHODOLOGY 12
2.1 Molecular Dynamic Theory 12
2.1.1 Potential Energy Function 12
2.1.2 Integration and Timestep 13
2.1.3 Langevin Dynamics and Ensembles 15
2.2 Electron Structure Theory 15
2.2.1 Hartree-Fock Theory 17
2.2.2 Density Functional Theory 21
2.3 Potential Energy Surfaces 25
2.4 Nudged Elastic Band 25
CHAPTER 3: MODELING THE ANNEXIN A1-S100A11 HETEROTETRAMER: A MOLECULAR DYNAMICS INVESTIGATION OF STRUCTURE AND CORRELATED MOTION 27
3.1 Abstract 27
3.2 Computational Methods 28

3.2.1 Modeling the Human A1-S100A11(A1t) Tetramer 28
3.2.2 Molecular Dynamics Simulation 29
3.2.3 Simulation Analysis and Figure Generation 30
3.3 Results and Discussion 30
3.3.1 A1t Structural Conformation 31
3.3.2 Secondary Structure Analysis 35
3.3.3 Strong Interactions 38
3.3.4 Calcium Binding 43
3.3.5 Correlated Motions 48
3.3.6 Conclusions 51
CHAPTER 4: THEORETICAL STUDIES ON SQUARE PLANAR PLATINUM COMPLEXES 53
4.1 Background 53
4.2 Methodology 53
4.3 Computational Results 55
4.3.1 Optimization of Structures 55
4.3.2 Mapping Reactions via NEB 57
4.4 Discussion 63
CHAPTER 5: THEORETICAL PROBE INTO THE SOLVENT-SELECTIVE NATURE OF C-H ACTIVATION IN CYCLOPLATINATION REACTIONS 65
5.1 Background 65
5.2 Methodology 66
5.3 Results 68
5.3.1 Proposed Mechanism 68
5.3.2 Geometry Optimization 68
5.3.3 Critical Field Analysis 71
5.3.4 NEB Pathways 72
5.4 Energy Diagrams 86
5.5 Conclusions 87
REFERENCES 90

Table of Tables

Table 1.1 5
Table 3.1 39
Table 3.2 41
Table 3.3 42
Table 3.4 44
Table 3.5 45
Table 3.6 46
Table 4.1 56
Table 4.2 59
Table 4.3 62
Table 5.1 70
Table 5.2 76
Table 5.3 77
Table 5.4 78
Table 5.5 79
Table 5.6 82
Table 5.7 83
Table 5.8 85
Table 5.9 86

Table of Figures

Figure 1.1 7
Figure 1.2 9
Figure 1.3 11
Figure 3.1 30
Figure 3.2 31
Figure 3.3 32
Figure 3.4 33
Figure 3.5 36
Figure 3.6 48
Figure 4.1 53
Figure 4.2 58
Figure 4.3 58
Figure 4.4 59
Figure 4.5 60
Figure 4.6 61
Figure 4.7 62
Figure 4.8 63
Figure 5.1 66
Figure 5.2 68
Figure 5.3 70
Figure 5.4 72
Figure 5.5 74
Figure 5.6 76
Figure 5.7 78
Figure 5.8 79
Figure 5.9 80
Figure 5.10 82
Figure 5.11 83
Figure 5.12 85
Figure 5.13 88
Figure 5.14 89

Table of Equations

2.1..... 12
2.2..... 13
2.3..... 13
2.4..... 14
2.5..... 14
2.6..... 15
2.7..... 16
2.8..... 16
2.9..... 16
2.10..... 17
2.11..... 17
2.12..... 17
2.13..... 17
2.14..... 18
2.15..... 18
2.16..... 18
2.17..... 18
2.18..... 18
2.19..... 19
2.20..... 19
2.21..... 19
2.22..... 19
2.23..... 19
2.24..... 19
2.25..... 19
2.26..... 20
2.27..... 20
2.28..... 20
2.29..... 20
2.30..... 20

2.31..... 21
2.32..... 21
2.33..... 21
2.34..... 23
2.35..... 23
2.36..... 23
2.37..... 23
2.38..... 23
2.39..... 24
2.40..... 24
2.41..... 24
2.42..... 24
2.43..... 24
2.44..... 24
2.45..... 26

CHAPTER 1: INTRODUCTION

1.1 Introduction

The overall goal in computational chemistry is to apply the fundamental equations governing molecular structure to accurately describe chemical phenomena. While traditional wet-lab experiments readily provide tangible evidence and quantitative measurements, they often face difficulty answering how or why a process occurs. For instance, the delicate process of crystallography requires a perfect concoction of solvents to procure a structure with the best x-ray matrix, not necessarily the most feasible biological structure. Furthermore, experimentation may characterize organic pathways in terms of starting, final, and exceptionally stable intermediate structures but face difficulty determining or explaining the transitions between these states. For this reason, computational chemistry is often applied in tandem with wet-lab experimentation to achieve a more robust understanding of chemical systems. Computational approaches utilize efficient computers thus are limited based on computational resources and system size. Two major branches of computational chemistry arise based on the size of the system being analyzed. For large macroscopic molecules such as proteins, micelles, vesicles, etc, molecular dynamics simulations are employed to visualize a system as it evolves subject to classical mechanics and Newton's Second Law. Smaller systems, i.e. organic molecules, allow for more exact calculations rooted in quantum mechanics and electron structure theory. In this work, both quantum and classical computational methodologies were employed to study biologically important molecules and pathways. In Chapter 3, molecular dynamics simulations were performed on the Annexin A1-S100A11 heterotetramer to investigate its structure and dynamics. Chapters 4 and 5 involve ab initio calculations in conjunction with the Nudged Elastic Band method to investigate mechanisms of platinum structures in the context of C-H activation.

1.2 Annexin Proteins

The Annexin family of proteins dates to 1977 when Carl et al. isolated Synexin, now known as Annexin A7. Since then, thousands of Annexins have been discovered across Eukaryota, however, there are only 12 members of the Annexin family present in vertebrates (Annexins A1-A11, A13) with A12 being unassigned.[1,2] The key feature of the annexin family is calcium-dependent binding with membranes and other proteins. Each annexin is composed of a highly conserved core domain (CD) and an N-terminal domain (ND) unique to each family member. Five α-helices make up each of the four homologous repeats in the core domain with the exception of Annexin A6, which has eight repeats.[1,3–5] The core-domain is curved, with its two faces being referred to as the concave and convex faces. The concave face contains the ND which serves as the site for interactions with S100 proteins. Additionally, phosphorylation and glycosylation sites found within the concave face have been shown to regulate interactions with other proteins[3,6,7] The convex face contains Ca^{2+} binding sites and the primary location of annexin-membrane interactions.[3,7]

1.2.1 Annexin A1

Annexin A1 (A1), a particularly important member of the Annexin family, contains 346 amino acids (~38 kDa) and a 44 amino acid ND.[1,3] A1 can bind up to eight calcium ions on its convex protein surface.[3] Prior to the binding of calcium, the ND exists in a buried state within Repeat III of the core domain.[1,3,8] Upon calcium binding at Repeat III, the ND is ejected from the core domain (CD).[1,3,9] Following calcium binding, the released ND can interact with cell membranes, bind with phospholipid bilayers, and perform additional biological functions. Compared to other Annexin proteins, the A1 ND is considerably longer. Additionally, the conformational shift of the ND in the presence of calcium is a characteristic phenomenon which

to our knowledge, is unique to A1. Consequently, the A1 ND is very important to understanding the function of the protein and a focus of many research efforts. [1,3]

A1 is most notably involved in inflammatory responses and apoptotic pathways.[3] For instance, A1 has been linked to Alzheimer's disease by its ability to reduce the inflammation marker secretion from Amyloid β plaques. In mice treated with human-recombinant A1, the leakage of the blood-brain barrier was reversed which improved amyloid plaque clearance. With a decrease in amyloid plaque, vascular structure was restored showing promise of A1 in the treatment or mediation of neurodegenerative diseases.[10] The misregulation of A1 has been found to be present in many forms of cancer including breast cancer, gastric cancer, glioblastoma and certain types of carcinoma.[8] This may be attributed to the involvement of A1 in pathways regulating EGFR. A1 is also related to the inherent resistance of recurrent ER positive breast cancer to tamoxifen therapy.[11]

1.2.2 Annexin and S100 Proteins

Annexin family proteins and S100 proteins form heterotetramers via interactions between the NDs of annexins and the hydrophobic pockets of S100 proteins upon Ca^{2+} binding.[4] The S100 family is comprised of 10-12 kDa proteins found exclusively in vertebrates. S100 proteins consist of two unique EF-hand Ca^{2+} binding motifs connected via a hydrophobic linker.[12,13] The C-terminal Ca^{2+} binding loop is made of 12 amino acids and has a higher Ca^{2+} affinity compared to the 12 amino acid N-terminal Ca^{2+} binding loop.[12] S100s are implicated in regulatory functions, Ca^{2+} homeostasis, and cellular motion. Additionally, S100s have been linked to various inflammatory and neurological conditions/diseases.[13] S100A11 is found cellularly as an antiparallel, non-covalent homodimer. Dimerization of S100 proteins occurs between helices I and IV of two S100 monomers. Upon calcium binding, the S100 dimers create hydrophobic cleft

sites for ligand interactions. Like A1, misregulation of S100A11 is a feature of various cancers. For instance, its upregulation is correlated to worse prognoses due to its involvement in tumor cell proliferation and apoptosis.[14] In the study conducted by Wang, H et al., cell proliferation and tumorigenesis of glioma were considerably inhibited when S100A11 was downregulated.[14]

Like the previously mentioned A1 functions, the A1-S100A11 heterotetramer (A1t) is proposed to regulate calcium-dependent biological processes. A1t has been implicated in the development of keratinocytes and their cornified envelopes.[15] A1t has also been shown to bind and inhibit phospholipase A2 which is pivotal for keratinocyte growth.[16] Furthermore, the cleavage of A1 prevents its association to S100A11 resulting in the unregulated growth of model squamous carcinomas.[17] In addition, A1t has been observed in early endosomes, and the binding of A1 to S100A11 helps regulate the degradation of EGFR via lysosomal degradation.[18] Malfunctions in the EGFR pathway can result in tumor formation, suggesting that A1t is involved in tumorigenesis.

1.3 C-H Functionalization

In industrial chemistry, the chemical feedstock used to synthesize organic materials is largely driven by economic factors. Currently, olefins serve as precursors to many high-demand organic compounds such as acrylic acid, acrylonitrile, and cumene. Compared to olefins, alkanes are even cheaper than alkenes since they are readily accessible from natural gas and crude oil. As a result, there is a great economic incentive to developing methods to directly convert these cheap chemicals into various popular chemicals used in the chemical industry.[19,20] Unfortunately, direct conversion of these C-H bonds into other chemical bonds (C-X bonds) is a challenging task and one of the major focuses in modern organic chemistry. This is in large part due to the inert nature of normal C-H bonds.[21,22] For instance, methane's homolytic bond dissociation

enthalpy is 440 kJ/mol at 25ºC with a pKa of 48.[22] Thus, product C-X bonds from homolytic pathways will be weaker than those of the reactants.

Another hindrance for hydrocarbon functionalization is the selectivity of C-H bonds in large, complicated molecules. Shown in **Table 1,** the bond dissociation energies of C-H bonds for primary, secondary, and tertiary C-H bonds are within ~5 kcal mol.[23] Since these BDE are very close, especially relative to the energy required for these reactions to take place, discriminating between these bonds can be very challenging. Furthermore, radical and electrophilic pathways involving alkanes favor reactions attacking tertiary C-H bonds over primary and secondary due to their relative stabilities (3º > 2º > 1º) which can lower the yield for a desired product.

Table 1.1: Select C-H, C-X Bond Dissociation Energies[23]

Bond	BDE (kcal/mol)	Bond	BDE (kcal/mol)
CH_3-H	105	C-C	90
Me-CH_2-H	101.1	C-O	92
(Me_2)CH-H	98.6	C-N	85
(Me_3)C-H	96.5	C-F	115
C_6H_5-H	113	C-Cl	84
HCC-H	133	C-Br	72

Traditionally, these issues have been circumvented by 1) careful selection of pre-functionalized starting materials and 2) sequential addition/modification of functional groups in a selective fashion. These traditional synthetic schemes often consist of numerous steps and chemical transformations to form a desired product. Conversely, direct C-H activation and functionalization allow for a direct transformation of a C-H bond into a C-C, C-O, or C-N bond type without altering additional functional groups present in the molecule. Compared to direct

C-H functionalization, the traditional approach adds a significant number of steps to a synthesis which results in lower yields, more waste, and in many cases, expensive and/or environmentally unfriendly catalysts/starting materials.[24] As a result, developing and characterizing methods to selectively functionalize C-H bonds has emerged as a leading focus of modern organic chemistry.[25]

1.3.1 Methods of Functionalization

There are 3 main ways to functionalize C-H bonds. The first approach takes advantage of radical species following hydrogen abstraction. In this form of activation, a proton and an electron are removed from a molecule resulting in a radical and a radical cation.[25] The radical species can also be oxidized to form additional cations. These radical species serve as important intermediates in C-H functionalization pathways. C-H functionalization through radical formation is especially favorable in intramolecular reactions due to the lower energy barrier afforded by structural proximity. Additionally, these intramolecular reactions allow ligands to be developed in a way that mitigates the regioselectivity issues previously described. Similarly, C-H functionalization can occur via intramolecular metal-catalyzed transfer of carbenes/nitrenes. These reactive species typically react with metal atoms to form a carbenoid prior to C-H activation. By forming a transient intermediate, desired product bonds are placed in geometric proximity. This helps improve selectivity when compared to the carbene/nitrene form.[26,27] Consequently, carbenoid insertion methods are especially powerful for stereospecific syntheses as seen in the chiral rhodium C-H functionalization used in the synthesis of Ritalin.[27] Lastly, transition-metal catalysis is arguably the most powerful method used to functionalize C-H bonds. This form of C-H functionalization involves the activation of sp^2 and sp^3 C-H bonds through the formation of ringed metallocycle intermediates. In this work, we will focus on platinum

mediated C-H functionalization.

1.3.2 Transition Metal Mediated Functionalization

The use of transition metal catalysts has demonstrated great promise in alleviating the challenges of selectivity in C-H functionalization of alkanes. In fact, these homogeneous catalysts can be modified to yield products following a reversed preference in terms of regioselectivity (1>2>3) and stability (stronger bond activated in presence of weaker bond). Broadly, the mechanism of C-H activation can be classified as either an outer sphere or inner sphere mechanism. The previously described metalloradical and metal carbene/nitrene insertion C-H functionalization fall under the outer sphere classification. Transition metal C-H activation is classified under inner sphere C-H activation. Transition metal C-H activation is believed to occur through either oxidative addition, sigma bond metathesis, or electrophilic substitution as shown in **Figure 1.1**.

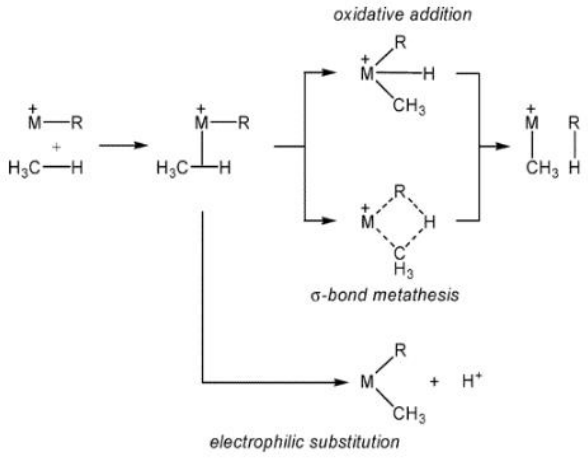


Figure 1.1: The general mechanism by which C-H activation occurs. After formation of the sigma-complex, the reaction undergoes either 1) electrophilic substitution or 2) oxidative addition/sigma-bond metathesis.

The first step in transition-metal-mediated C-H activation is the association of a hydrocarbon to the metal forming a so-called sigma complex. This intermediate serves as a crucial crossroad from which other mechanisms progress. The bonds between the metal and this intermediate are much weaker than the ligands which they replace.[21,23,28,29] Thus, the coordination of the hydrocarbon to the metal is thermodynamically unfavorable and an important place of study for mechanisms.

Another way to view the pathway of C-H activation is through the charge transfer between C-H bonds and the transition metal. This transfer can happen in the forward direction (filled C-H σ bond to an empty metal dσ orbital) or in the reverse direction (metal dπ to C-H σ*

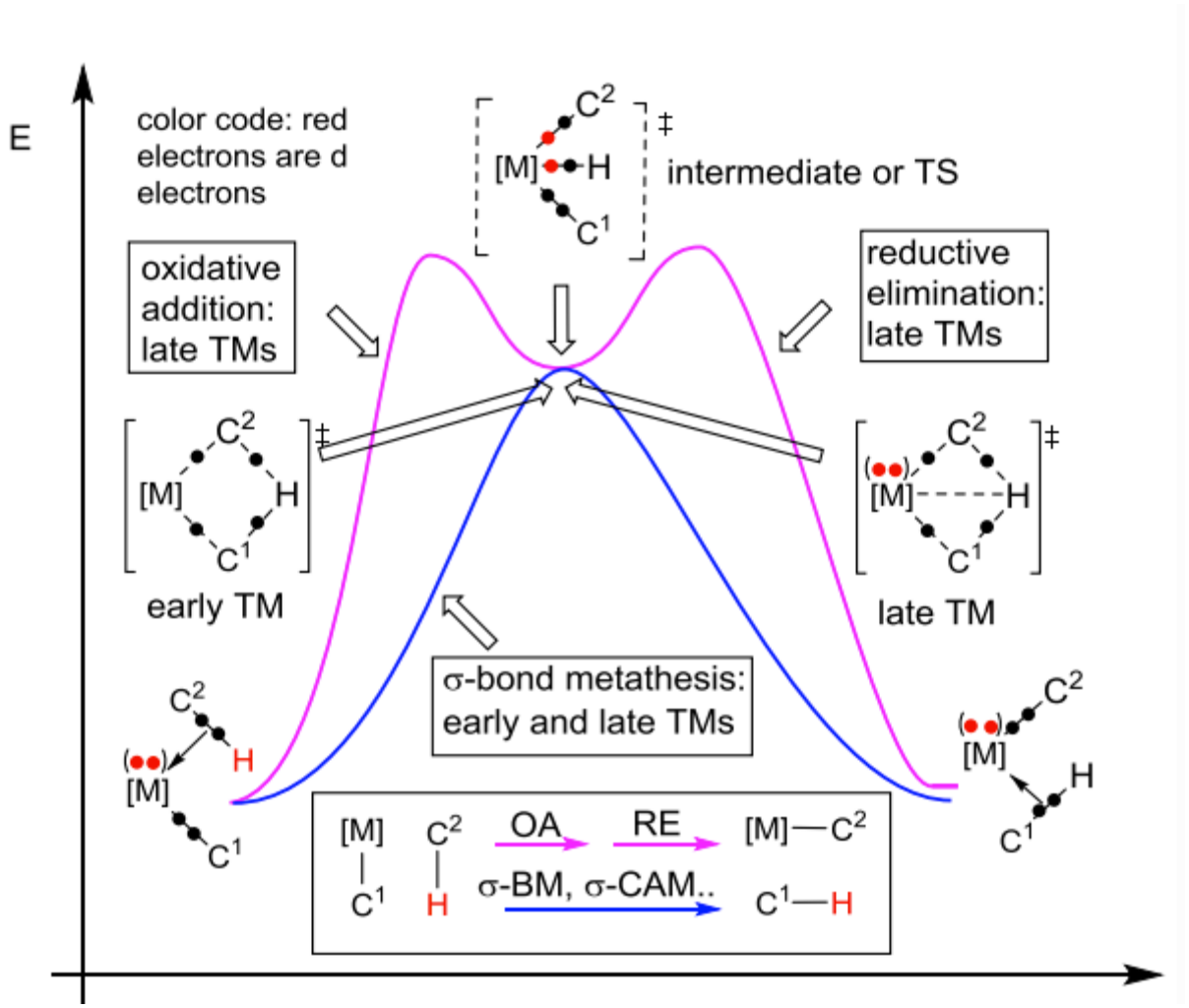


Figure 1.2: Reaction coordinate diagram showing the different mechanisms by which C-H functionalization may occur when mediated by transition metals.

orbital).[30] Thus, the electronic character of the transition metal has a significant impact on the mechanism of C-H activation as shown in **Figure 1.2.**

While early transition metals typically activate C-H bonds via σ-bond metathesis, late transition metals such as platinum may activate C-H bonds by both σ-bond metathesis and redox mechanisms. Both mechanisms progress through the σ-complex; however, whether the σ-bond species acts as a transition state or stable intermediate is dependent on the electronic properties of the metal-ligand complex. Thus, understanding the mechanism of transition metal C-H

activation is critical for ligand design. By carefully tailoring ligands/reaction conditions, it may be feasible to selectively harness specific intermediates and have greater control over product formation. As a result, great emphasis has been put on characterizing the mechanisms of transition metal catalysis.

1.3.3 Platinum Complexes

Pt-complexes have been a focus in C-H activation since the field's earliest mechanistic studies. Not only have platinum complexes served as important anticancer treatments, but they are also involved in various photophysical applications such as DNA probes and medical imaging technologies due to their versatile electronic properties. Due to its heavy atom effect, platinum increases the rate of intersystem crossing through spin-orbit coupling effects leading to populated triplet spin states. In organic light-emitting diodes, both singlet and triplet excited states are generated, thus a triplet emitter could theoretically increase quantum efficiencies up to 100%.[31] As a result, phosphorescent Pt-complexes show promise in future OLED technologies. The stellar intersystem crossing exhibited by these complexes allows both spin states emitted through OLEDs to be harvested.[32,33] Research has shown that cyclometalated platinum (II) complexes offer stronger photochemical reactivity and higher quantum yields than other metal complexes even at room temperature.[32] Luckily, these cyclometalated platinum complexes are efficiently synthesized via intramolecular Pt-mediated C-H activation.[34]

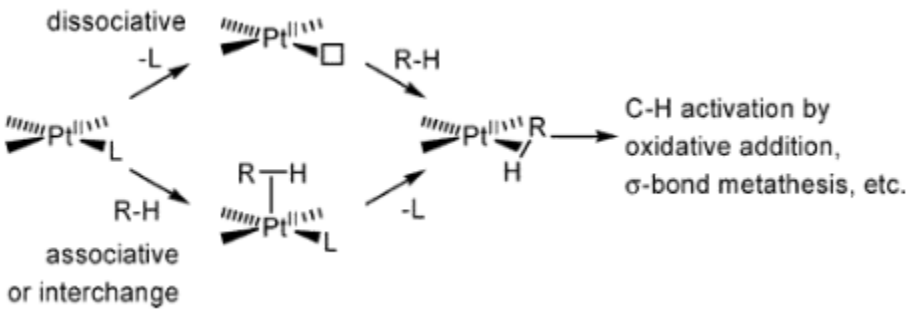


Figure 1.3: Formation of the sigma-complex in square planar platinum complexes.

Current literature suggests that sigma-complex formation in square planar Pt(II) complexes forms as a result of either a dissociative or concerted pathway as shown in **Figure 1.3.**[28,35,36] Fully characterizing the mechanisms by which Pt acts as a catalyst is important to better understanding its capabilities and uses in C-H functionalization. However, due to platinum's various electronic states as well as the convoluted mechanisms of C-H activation, this can be difficult and often requires computational approaches.

CHAPTER 2: METHODOLOGY

2.1 Molecular Dynamic Theory

In computational chemistry, large molecular systems i.e. proteins, membranes, micelles, and vesicles are frequently studied using molecular dynamics simulations. Molecular dynamics theory (MD) operates under classical approximations to model many bodied systems. Through classical mechanics, MD treats a system as a group of beads connected via springs which evolve over time through Newton's second law of motion

$$F_i = m\ddot{x}_i = -\frac{\partial U(r)}{\partial r_i} \qquad \textbf{0.1}$$

where *F* is the force acting on a particle of mass *m* with an acceleration given by the second derivative of its position *x* with respect to time; *F* can be equivalently expressed as the partial derivative of the potential energy function U*(r)* describing the system with respect to position.[37] The goal of MD is to simulate the evolution of a system through time. Solving the coupled differential equations and integrating them over time systematically updates the positions and velocities of every particle. The force acting on the updated system can be subsequently evaluated using the potential of the new coordinate system thus continuing the cycle. This iterative process results in movie-like snapshots of the system evolving over time which is known as a trajectory. [38,39]

2.1.1 Potential Energy Function

While there exists a variety of potential energy functions for various systems, the forcefield describing biomolecular systems within AMBER is given by

$$U(r) = \sum_{bonds} k(r - r_0)^2 + \sum_{angles} k(\theta - \theta_0)^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\varphi - \gamma)] + \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right] + \sum_{i<j} \left[\frac{q_i q_j}{\varepsilon R_{ij}}\right] \quad \textbf{0.2}$$

where the potential energy is a function of bond stretches and bends, dihedrals, nonbonded interactions, and electrostatic interactions.[40,41] The first two terms describe bond stretches and bends. These terms are calculated by a quadratic function of distances and angles which appropriately describe the (typically) unstrained structure of proteins. The dihedral term is described through a Fourier series where n is the number of bonds, γ is the phase, and V_n is the torsional potential associated with the type of dihedral. Nonbonded interactions are calculated via the 12- 6 function Lennard-Jones potential. Electrostatic interactions are calculated with the Classical Coulomb potential with point charges q, interatomic distances R, and dielectric permittivity ε. In other programs, the potential energy functions may include additional terms to describe energy contributions of hydrogen bonds. In AMBER, hydrogen bonds are described through the nonbonded and electrostatic terms. [38]

2.1.2 Integration and Timestep

Prior to MD simulation, a system's initial coordinates and velocities must be assigned. These coordinates are often provided by experimental data previously acquired via crystallography, NMR, etc. which can be obtained from various databases. Initial velocities in this work were randomly assigned using a Boltzmann distribution:

$$P(v_i) = \left(\frac{m_i}{2\pi k_b T}\right)^{1/2} e^{-\frac{m_i {v_i}^2}{k_b T}} \quad \textbf{2.3}$$

From these initial velocities/coordinates, Newton's equations can be solved via numerical integration to generate a system defined by a new set of coordinates and velocities. While several integration methods exist, one common algorithm is the velocity Verlet algorithm shown below

$$r_i(t + \Delta t) = r_i(t) + v_i(t)\Delta t + \frac{1}{2}a_i(t)(\Delta t)^2 \quad \textbf{2.4}$$

$$v_i(t + \Delta t) = v_i(t) + \frac{1}{2}[a_i(t) + a_i(t + \Delta t)]\Delta t \quad \textbf{2.5}$$

where $v_i(t + \Delta t)$ and $a_i(t + \Delta t)$ are the updated velocities and accelerations respectively with a timestep of Δt. The defining feature of this method is the calculation of both position and velocity with the same time variable. Additionally, the velocity Verlet explicitly calculates updated velocity values which are crucial in MD integration. Implementation of this algorithm is relatively simple. Following the calculation of the updated coordinates, new velocities are then determined which are used to then calculate the force.

The timestep of a MD simulation is correlated to both accuracy and speed of a simulation. Smaller timesteps improve the accuracy of simulations. However, this is more computationally expensive regarding time and computer storage. Larger timesteps improve the speed of computation but induce system instability. In general, it is recommended to use a 1fs timestep. Various methods have been developed to afford larger time steps such as hydrogen repartitioning.[38,41–43]

2.1.3 Langevin Dynamics and Ensembles

Per thermodynamics, macroscopic properties i.e. temperature volume and pressure define a system's macroscopic state. On the other hand, MD simulations operate on microstates which involve the velocities and positions of particles within the system. The macroscopic manifestations and particle dynamics are mapped via ensembles. Ensembles are defined to be weighted averages of velocities and positions which correspond to a macrostate. In this work, an NPT ensemble with fixed N, P, and T will be used to better mimic experimental conditions. Thus, it is important for the appropriate ensemble to be generated for a pre-defined PVT. In MD software, Langevin dynamics are implemented which essentially adds a friction term to Newton's equation (Eq. 1) to enhance system stability.[38,41,42]

2.2 Electron Structure Theory

In quantum mechanics, all a system's information is encoded in its wave function (ψ). Although the wave function alone does not necessarily carry any physical significance, it is a mathematical description of state which is proportional to the probability of the system existing at a specific location when squared ($\int_{-\infty}^{\infty}|\psi(x)|^2 dx = 1$). The crux of electronic structure theory involves solving the time independent Schrödinger equation (TISE) to describe the motion/position of the electrons and nuclei of a system. The full molecular TISE is shown below

$$\widehat{H}\psi = E\psi \qquad \textbf{2.6}$$

where $\widehat{H}$ is the Hamiltonian operator and E is the energy eigenvalue associated with the molecular wavefunction (ψ). The molecular Hamiltonian provides a quantum mechanical description of interacting electrons and nuclei in a system and takes the following form in atomic units:

$$\hat{H} = -\sum_{A}^{nuc} \frac{1}{2M_A} \nabla_A^2 - \frac{1}{2} \sum_{i}^{elec} \nabla_i^2 - \sum_{A}^{nuc} \sum_{i}^{elec} \frac{Z_A}{r_{Ai}} + \sum_{A>B}^{nuc} \frac{Z_A Z_B}{R_{AB}} + \sum_{i>j}^{elec} \frac{1}{r_{ij}} \quad \textbf{2.7}$$

or in simplified notation:

$$\hat{H} = \hat{T}_N(R) + \hat{T}_e(r) + V_{eN}(r, R) + V_{NN}(R) + V_{ee}(r) \quad \textbf{2.8}$$

where $\hat{T}_N$ and $\hat{T}_e$ are the kinetic energy operators for nuclei and electrons respectively; ∇_x^2 is the Laplacian operator acting with respect to a particle's position. Potential energy operators for nucleus-electron repulsion ($\hat{V}_{eN}$), nuclei ($\hat{V}_{NN}$), and electrons ($\hat{V}_{ee}$) are given by the remaining terms with nuclear charge given by *Z* and mass by M. r_{Ai} denotes the distance between electron *i* and nucleus A, R_{AB} denotes the distance between nuclei *A* and B, and r_{ij} denotes the distance between electrons *i* and *j*. Unfortunately, exact solutions for the TISE are only available for Hydrogenic atoms; thus, approximations must be employed for many-electron systems. The methods of finding these approximate solutions to the Schrödinger equation fall into two trains of thought: wavefunction-based electron structure theory and density functional theory.

$$\psi_{Total} = \psi_{Electrons}(R_A; r_i)\psi_{Nuclei}(R_A) \quad \textbf{2.9}$$

The theory discussed in subsequent sections will concern finding the solution to the electronic component of the wavefunction via the electronic Hamiltonian which can be written as

$$\hat{H} = \hat{T}_e(r) + V_{eN}(r;R) + V_{ee}(r) \quad \textbf{2.10}$$

$$T_e(r) = -\frac{1}{2}\sum_{i}^{elec} \nabla_i^2 \quad \textbf{2.11}$$

$$V_{Ne}(r) = -\sum_{A}^{nuc}\sum_{i}^{elec} \frac{Z_A}{r_{Ai}} \quad \textbf{2.12}$$

$$V_{ee}(r) = \sum_{i>j}^{elec} \frac{1}{r_{ij}} \quad \textbf{2.13}$$

It is important to note that the electronic wavefunction is explicitly dependent on the coordinates of electrons. Because each electron is viewed in a field of nuclei, there is a parametric dependence on nuclear coordinates. Thus, solving the electronic Schrödinger equation yields the potential felt by the nuclei. Subsequent evaluation for each nuclei in the mean field of electrons, one can calculate nuclear motion. In other words, the electronic energy of a system ultimately acts as the potential energy surface (PES) of a molecule which is of particular importance for determining reaction pathways and discussed later.[38,45,46]

2.2.1 Hartree-Fock Theory

Simplest of the wavefunction theory methods, the Hartree-Fock method is a rudimentary way of solving the electronic Schrödinger equation and is fundamental to many of the more sophisticated methods used in modern computational chemistry. Exact solutions to the electronic Schrödinger equation in which orbitals are the true eigenfunctions of the full Hamiltonian are only possible for Hydrogenic systems while multi-electron systems require approximations. In Hartree-Fock Theory, multi-electron wavefunctions are assumed to take the form of orbital products

$$\chi_i(\mathrm{x}_j) = \varphi_i(\mathrm{r}_j)\sigma_i(\omega_j) \quad \textbf{2.14}$$

where $\varphi_i(\mathrm{r}_j)$ represents the spatial component and $\sigma_i(\omega_j)$ represents the spin component of orbitals. These orbital products are expressed as Slater determinants (SD)

$$\psi = \frac{1}{\sqrt{N!}}\begin{vmatrix} \chi_1(\mathrm{x}_1) & \chi_2(\mathrm{x}_1) & \cdots & \chi_N(\mathrm{x}_1) \\ \chi_1(\mathrm{x}_2) & \chi_2(\mathrm{x}_2) & \cdots & \chi_N(\mathrm{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathrm{x}_N) & \chi_2(\mathrm{x}_N) & \ldots & \chi_N(\mathrm{x}_N) \end{vmatrix} \quad \textbf{2.15}$$

with $\frac{1}{\sqrt{N!}}$ acting as a normalization constant. Using SDs to represent wavefunctions has several important consequences. Since electrons are fermions, their wavefunctions are required by quantum mechanics to be antisymmetric with the interchange of any two electrons which is easily handled by SDs. Secondly, each electron in a Slater determinant is indiscernible from other electrons in a system meaning that every electron is connected to every orbital, thus abiding by quantum nature. Additionally, assuming HF wavefunctions to be expressed by SDs permits the calculation of their energies by the variational theorem which calculates the orbitals which minimize energy.

$$E[\tilde{\psi}] = \langle\tilde{\psi}|\hat{H}|\tilde{\psi}\rangle \quad \textbf{2.16}$$

$$\tilde{\psi} = \psi + \partial\psi \quad \textbf{2.17}$$

which expands the energy functional (Eq 15) to

$$E[\psi + \partial\psi] = E_0[\psi] + \partial E[\psi] \ldots \quad \textbf{2.18}$$

HFT is interested in cases where the first variation ($\partial E[\psi]$) disappears. At these stationary points, eigenvalues of the functional are the actual eigenstates of the wavefunction showing the basic idea behind solving the Schrodinger equation. The Hartree-Fock Energy functional (Eq 17) can be expanded in terms of its one and two electron operators as

$$E_{HF} = \langle \chi_i | h | \chi_j \rangle + \sum_{ij} \langle \chi_i \chi_j | \chi_i \chi_j \rangle - \langle \chi_i \chi_j | \chi_j \chi_i \rangle \quad \textbf{2.19}$$

$$h = -\frac{1}{2}\nabla_i^2 - \sum_A^{nuc} \frac{Z_A}{r_{iA}} \quad \textbf{2.20}$$

$$\mathscr{H} = \langle \chi_i | h | \chi_j \rangle = \int \chi_i^*(x1) h(x1) \chi_i(x1) \quad \textbf{2.21}$$

$$\mathcal{J} = \langle \chi_i \chi_j | \chi_i \chi_j \rangle = \iint \chi_i^*(x1) \chi_j^*(x2) \, \frac{1}{r_{12}} \chi_i \chi_j \mathrm{dx1dx2} \quad \textbf{2.22}$$

$$\mathcal{K} = \langle \chi_i \chi_j | \chi_j \chi_i \rangle = \iint \chi_i^*(x1) \chi_j^*(x2) \, \frac{1}{r_{12}} \chi_j(x1) \chi_i(\mathrm{x2}) \mathrm{dx1dx2} \quad \textbf{2.23}$$

As previously mentioned, the purpose of HFT is to minimize energy via atomic orbitals. It is convenient to assume the orbitals are orthonormal and will remain so after minimization. Using this as a constraint, The Hartree-Fock energy functional can be minimized via Lagrange's method of undetermined multipliers shown below:

$$\mathcal{L}[\chi] = E_{HF} - \sum_{ij} \epsilon_{ij} (\langle \chi_i | \chi_i \rangle - \delta_{ij}) \quad \textbf{2.24}$$

Where ϵ_{ij} are the Lagrange coeffiients and indices *i* and *j* represent the overlap between two orbitals. Setting $\delta\mathcal{L} = 0$, the Hartee fock equation becomes the following eigenvalue equation:

$$h(x_1)\chi_i(x_1) + \sum \left[\int dx_2 |\chi_j|^2 \frac{1}{r_{12}} \right] \chi_i(x_1) - \sum \left[\int dx_2 \chi_j^*(x_2) \chi_i(x_2) \frac{1}{r_{12}} \right] \chi_j(x_1) = \sum_j \epsilon_{ij} \chi_i(x_1) \quad \textbf{2.25}$$

The canonical form of the Hartree Fock equation (2.25) results from diagonalization of the epsilon matrix with a unitary transformation U shown below:

$$\chi'_i(x_1) = \sum_j U_{ji}\chi_j(x_1)) \tag{2.26}$$

Using equation 2.26, the canonical form of the Hartree Fock equations can be rewritten using the Coulomb and exchange operators.

$$\left[h(x_1) + \sum \mathcal{J}_j(x1) - \sum \mathcal{K}_j(x1)\right]\chi_i(x_1) = \epsilon_i\chi_i(x_1) \tag{2.27}$$

$$\mathcal{J}_j(x1)\chi_i(x_1) = \langle\chi_i\chi_j|\chi_i\chi_j\rangle = \left[\int \mathrm{dx2}\chi_j^*(x2)\frac{1}{r_{12}}\chi_j\right]\chi_i(x_1) \tag{2.28}$$

$$\mathcal{K}_j(x1)\chi_i(x_1) = \langle\chi_i\chi_j|\chi_j\chi_i\rangle = \left[\int \mathrm{dx2}\chi_j^*(x2)\,\frac{1}{r_{12}}\chi_i\right]\chi_j(x_1) \tag{2.29}$$

The part in brackets is known as the Fock operator *f(x_1)*. Although exact solutions to this integro-differential equation are possible analytically, a much more practical approach developed by Roothan represents spatial orbitals in terms of finite basis functions which simplifies the differential equation into matrix algebra. With these basis functions, the spatial orbitals take the form

$$\chi_i = \sum_b^N C_{bi}B_b \tag{2.30}$$

Which allows Hartree-Fock to be solved as the following matrix equation

$$\sum_{b}^{N} F_{\mu b} C_{bi} = \epsilon_i \sum_{b}^{N} S_{\mu b} C_{bi} \quad \textbf{2.31}$$

$$F_{\mu b} = \int dx_1 B_b^*(x_1) f(x1) B_b(x_1) \quad \textbf{2.32}$$

$$S_{\mu b} = \int dx_1 B_b^*(x_1) B_b(x_1) \quad \textbf{2.33}$$

where ϵ_i forms a diagonal matrix of energy values. The overlap integral $S_{\mu b}$ vanishes upon transformation to an orthogonal basis to yield a simple linear algebra eigenvalue equation. However, since $F_{\mu b}$ depends on itself, the Hartree-Fock method is a self-consisent-field method and requires iterative solutions.

While Hartree-Fock is computationally facile, it does so at the expense of electronic correlation. The approximations in Hartree-Fock neglect electron-electron interactions as a consequence of its electrons moving in a one-bodied field. Many methods known as post-Hartree Fock methods have been developed which directly include the effects of electron-electron interactions. Unfortunately, these methods do so at computational cost, so they are sometimes impractical for larger systems.[38,45–49]

2.2.2 Density Functional Theory

Density functional theory has proven to be one of the most powerful and successful methods to determine electronic structures in both chemistry and physics. Unlike HFT, DFT accounts for electron correlation at a similar computational cost. This is afforded by employing electron density as DFT's fundamental variable. This is starkly different than post-Hartree Fock

methods which use complex wavefunctions. What this implies is that instead of working with an energy functional of a wavefunction of 3N dimensions, DFT's functional depends only on the spatial/spin coordinates defining a system's electron density.

The use of total density in place of the wavefunction is validated through the two revolutionary theorems of Walter and Kohn.

Theorem 1: *The external potential v(r) is determined, within a trivial additive constant, by electron density p(r).*

As shown in HFT **Eq 9**, the v(r) is defined for a specific arrangement of nuclei per the parametric dependence on nuclear coordinates; the remaining electronic properties depend the number of electrons. The theorem postulates that the wavefunction is determined by density since $p(r) = N \int d^3 r_N \psi^*(r \dots r_N)\, \psi(r \dots r_N)$. Thus, there is a 1-1 map connecting v(r) and p(r) which in turn allows for the evaluation of the exact electronic energy of a system. By proving the wavefunction can be defined through electron density, the theory also proves that a system is completely described through its electron density. Being able to use electron density in quantum mechanics is conceptually advantageous compared to the preternatural nature of wavefunctions since density is measurable.

The second Hohenberg and Kohn theorem proves the ground state density may be determined through the variational principle. Through the variational principle, Hohenberg and Kohn derived an energy functional which was a function of electron density.

$$E_{el} = E[p(r)] \tag{2.34}$$

$$E[p(r)] = \langle \Psi_0[p(r)] | \hat{T}_e + \hat{V}_{ee} + \hat{V}_{ne} | \Psi_0[p(r)] \rangle \tag{2.35}$$

$$= \hat{T}_e[p(r)] + \hat{V}_{ee}[p(r)] + \int v_{ne} p(r) dr \tag{2.36}$$

$$= \hat{T}_e[p(r)] + J[p(r)] + E_{xc} + \int v_{ne} p(r) dr \tag{2.37}$$

$$J[p(r)] = \frac{1}{2} \int \int \frac{p(r)p(r')}{|r - r'|} dr dr' \tag{2.38}$$

The first three terms of **Eq. 2** are commonly referred to as the Hohenberg-Kohn density functional and consist of kinetic energy and electron-electron interactions. The second part involves the nuclear field and its interactions with the electrons. From the equation it is apparent only the second term involves nuclear contribution to the energy; therefore, the Hohenberg-Kohn energy functional is universal and may be applied to any given system. Despite providing a means to calculate the ground state energy via density, the theorem fails to provide functionals for the kinetic energy and exchange terms rendering its direct application unfeasible since approximations to the former are crude at best.

In 1965, Kohn-Sham circumvented the problem of an unknown kinetic energy functional by constructing a wavefunction of non-interacting electrons which has an identical density to the system of interacting electrons. In other words, a fictitious Hamiltonian was built using noninteracting electrons of the same ground state density as the real Hamiltonian. By doing

so, the fictitious Hamiltonian could be constructed in terms of one-electron operators acting on orbitals (SDs). In this formulation, the interacting system's kinetic energy is constructed using orbitals similar to HFT

$$T_s[p] = \sum_{\sigma}^{\alpha,\beta} \sum_{i}^{n} \left\langle \varphi \middle| -\frac{1}{2}\nabla^2 \middle| \varphi \right\rangle \quad \textbf{2.39}$$

With

$$p(r) = \sum_{\sigma}^{\alpha,\beta} \sum_{i}^{n} |\varphi(r)|^2 \quad \textbf{2.40}$$

Similar to HFT, the KS kinetic energy functional can be substituted into the Hohenberg-Kohn equations and minimized under the constraint that electrons are conserved

$$E_{KS} = \hat{T}_s[p(r)] + J[p(r)] + E_{xc} + \int v_{ne} p(r) dr \quad \textbf{2.41}$$

$$\delta\{E_{KS}[p] - \mu\left[\int p(r)\, dr - N\right]\} = 0 \quad \textbf{2.42}$$

$$\mu = \frac{\delta E_{KS}[p]}{\delta p(r)} = T_s[p(r)] + \frac{\delta F_{KS}[p]}{\delta p(r)} = \frac{\delta T_S[p]}{\delta p(r)} + V_{KS} \quad \textbf{2.43}$$

$$V_{KS}[r] = v_s + \int \frac{p(r')}{|r - r'|} dr' + \frac{\delta E_{xc}[p]}{\delta p(r)} \quad \textbf{2.44}$$

Unfortunately, in the formulation of the Coulomb potential, electron densities experiences repulsion with itself; this is a major discrepancy between HF and DFT. In DFT, multi-electron effects are accounted for in the exchange-correlation term which is the functional derivative of the exchange-correlation energy. Thus, the self-interaction error is accounted for in the exchange-correlation term. In principle, DFT exact in theory, meaning that an exact E_{xc} yields an exact ground state density. Unfortunately, the functional for E_{xc} is unknown and approximations must be made.[38,48–52]

2.3 Potential Energy Surfaces

As previously mentioned, a quintessential part of electron structure theory was the BOE which separated the TISE into its respective electronic and nuclear components. By focusing on solving the electronic problem and adding nuclear-nuclear energy classically, a potential energy surface (PES) on which nuclei move is obtained. Since the PES is calculated over all possible permutations of nuclear coordinates, it is a function of various geometries in Cartesian space and thus has a dimentionality of 3N-6 with 1 coordinate reserved for energy. Simply stated, the PES is a complete description of a molecule's available conformations as well as the paths connecting these accessible states.

Thus, PES are powerful tools for characterizing chemical systems since their critical points are chemically relevant. Per elementary Calculus, stationary points along the PES occur when the molecular force vanishes. Stationary points are subsequently evaluated with the Hessian matrix which is the collection of second derivatives which describes the curvature of the PES. Stable chemical structures (reactants, products, intermediates) whose Hessian is positive definite are minima. These stable wells are connected via saddle points. Saddle points whose Hessian contains 1 negative eigenvalue are referred to as transition states and are the lowest energy barrier path between two minima.[38,48,53]

2.4 Nudged Elastic Band

There are many methods for optimizing molecular structures such that they are stationary points on a PES; however, they do not all ensure that the points lie on the same PES. The Nudged Elastic Band (NEB) method is one reliable method for determining saddle points and the minimum energy pathway (MEP) connecting a given reactant and product. In this method, a continuous set of images are generated connecting optimized reactants and

products. These images are subjected to both a spring force and a molecular force. The molecular force is simply the gradient of an images energy with respect to its nuclear coordinates. These forces are projected out into their parallel and perpendicular components which constitute the total force shown below

$$F_i = F_i^s|_{||} - \nabla V(R_i)|_{\perp} \quad \textbf{2.45}$$

Where F_i is the total force acting on an image, $F_i^s|_{||}$ is the parallel force acting along the tangent, and $\nabla V(R_i)|_{\perp}$ is the force perpendicular to the tangent. This decoupling of the true force into parallel and perpendicular components is important as the parallel force, also known as the spring force, does not interfere with energy minimization. The parallel force acts as rubber bands connecting each image to maintain continuity and equal spacing of the energy pathway. The perpendicular force drives the energy pathway towards the MEP.

In theory, the MEP represents the PES connecting the given reactants and productus. Thus, its maxima and minima will correspond to transition states and intermediates respectively. The MEP generated via NEB contains a great deal of information which is important to characterizing a chemical system. For instance, the highest peak on the MEP corresponds to the activation energy of the transition from which kinetics can be calculated. While the NEB will reach the MEP connecting its endpoints when $\nabla V(R_i)$=0, it provides no estimate on how long this process will take. Thus, it is more practical to use a nearly converged NEB calculation to provide an excellent initial guess for a TS and subsequently refine it using other methods.

CHAPTER 3: MODELING THE ANNEXIN A1-S100A11 HETEROTETRAMER: A MOLECULAR DYNAMICS INVESTIGATION OF STRUCTURE AND CORRELATED MOTION

3.1 Abstract

Annexin A1 (A1) has been shown to form a tetrameric complex (A1t) with S100A11 which is implicated in calcium homeostasis and EGFR pathways. Previously, molecular dynamics (MD) simulations have been performed on monomeric A1. Lewis et al. investigated the conformational changes of A1 during the ejection of the ND from the CD.[54] Simpkins et al. performed MD simulations to study S27 mutations in the ND of A1 and their implications on structural organization and membrane aggregation.[55] In this work, we studied the structure and dynamics of the complete A1t with the intent of providing detailed structural and atomistic information to better understand the interactions within the A1t and guide future experimental studies.

For the first time, a full-length model of the human A1t was constructed and subjected to multiple Molecular Dynamic simulations for several hundred nanoseconds each. These simulations yielded three structures for the A1 N-terminus (ND) which were identified via principal component analysis. The orientations and interactions of the first 11 A1-ND residues for all three structures were conserved, and their binding modes were strikingly similar to those of the Annexin A2 N-terminus in the Annexin A2-p11 tetramer. In this study, we provided detailed atomistic information for the A1t. Strong interactions were identified within the A1t between the A1-ND and both S100A11 monomers. Residues M3, V4, S5, E6, L8, K9, W12, E15, and E18 of A1 were the strongest interactions between A1 and the S100A11 dimer. The different conformations of the A1t were attributed to the interaction between W12 of the A1-ND with M63 of S100A11 which caused a kink in the A1-ND. Cross-correlation analysis revealed

strong correlated motion throughout the A1t. Strong positive correlation was observed between the ND and S100A11 in all simulations regardless of conformation. This work suggests that the stable binding of the first 11 residues of A1-ND to S100A11 is potentially a theme for Annexin-S100 complexes and that the flexibility of the A1-ND allows for multiple conformations of the A1t.

3.2 Computational Methods

3.2.1 Modeling the Human A1-S100A11(A1t) Tetramer

To build the structure of the human A1t, the complete A1 structure previously generated by our group and the PDB structures 2LUC (PDB ID: 2LUC) and 1QLS (PDB ID: 1QLS) were used.[54,56,57] The complete A1 structure was previously constructed using the PDB structures 1MCX (PDB ID: 1MCX) and 1HM6 (PDB ID: 1HM6).[9,58] 1MCX is a calcium bound X-ray structure for the sus scrofa A1 monomer without the N-terminal. 1HM6 is a calcium-free X-ray structure for the full-length sus scrofa A1 monomer. As described by Lewis et al., the construction of the complete A1 structure used the core domain of 1MCX and the ND from 1HM6; then the ND was fused to the core domain.[54] 2LUC is an NMR structure of a S100A11 dimer in the presence of calcium. 1QLS is a crystal structure of S100A11 in complex with the first 11 residues of the A1-ND in the presence of calcium ions. Each structure was compared with its respective human wild-type sequence via EMBOSS Needle.[59] Differences were identified and modified to match the human wild-type sequence using Modeller.[60]

To build A1t, two copies of the 1QLS structures were aligned to the 2LUC S100A11 dimer. The NDs of two A1 monomers were then aligned to the A1-ND residues already present in the two 1QLS structures. All protein residues of 1QLS were then removed while leaving the calcium ions from 1QLS in the structure. Calcium ions were added to A1 by

first aligning two 1AIN (PDB ID: 1AIN) structures to each A1 monomer and then removing 1AIN protein residues. 1AIN is a calcium bound crystal structure of human annexin A1.[61] The resulting structure was a full length, calcium bound model of human A1t.

3.2.2 Molecular Dynamics Simulation

The topology and initial coordinate files for molecular dynamics (MD) simulation were generated using the tleap module of AmberTools 18.[62,63] The FF14SB force field was used along with the calcium ion parameters of Bartolotti et al.[64–66] Prior to simulation, proteins were solvated explicitly using TIP3P waters in an octahedral box with 16 Å between the protein and box boundary.[64] Sufficient chloride and sodium ions were added to neutralize any non-zero charges and bring the salt concentration to 0.15 M.

All MD simulations and calculations were performed on a 4-RTX 2080 Ti workstation housed at East Carolina University. Production run was performed using the pmemd.CUDA module within AMBER 18 while all other steps were carried out utilizing the pmemd.MPI module.[63] For all simulations, periodic boundary condition was utilized along with a 10Å non-bonded cutoff to limit the van der Wall non-bonded interactions. A 2 fs timestep was achieved through the constraint of bonds involving hydrogen atoms via the SHAKE algorithm.[67] In this work, three MD simulations were performed on A1t model. Prior to each MD simulation, the system was subjected to two steps of energy minimization. For the first step of minimization, the steepest descent method was used for 2500, 4000, and 5000 steps for each simulation respectively. For the second step of minimization, the conjugate gradient method was used for 7500 steps in all three simulations.

Following minimization, the systems were warmed in two steps. The first step scaled the temperature from 0-100 K over 5 ps, and in the second step, temperature was scaled from 100-

300 K over 50 ps. Finally, isothermal-isobaric (NPT) ensemble simulations were performed for 720, 600, and 560 nanoseconds.

3.2.3 Simulation Analysis and Figure Generation

Simulation trajectories were analyzed using the cpptraj and ptraj modules of AMBER18.[68,69] Interaction energies were calculated using the MMGBSA.py script.[70] All visualization and data manipulation were performed using Python, Pymol, and R.[63,71]

3.3 Results and Discussion

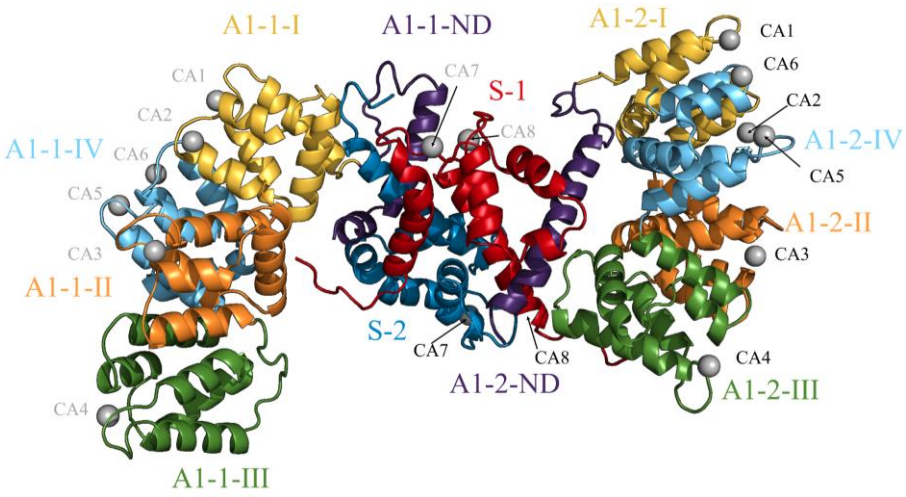


Figure 3.1 The cartoon representation of the simulated A1t. Two S100A11 monomers (S-1 and S-2) are in blue and red. The N-terminal domains (ND) of A1 are in purple, and the four structural repeats of A1 core domain (I-IV) are in yellow, orange, green, and cyan. The bound calcium ions for the A1t are shown in silver. A1-1 and S-1 calcium ions are labeled in silver while calcium ions bound to A1-2 and S-2 are labeled in black.

3.3.1 A1t Structural Conformation

Three simulations were performed on the A1t model for 720, 600, and 560 nanoseconds. These simulations differed only by the number of initial energy minimization steps as described in section **3.2.2 Molecular Dynamics Simulation**. The stabilities of these simulations were determined by the conservation of energy of the simulation system and the maintenance of constant average temperature. One of the simulated structures of A1t was shown in **Figure 3.1**. For all three simulations, the distance between two monomers of A1 within A1t and the root

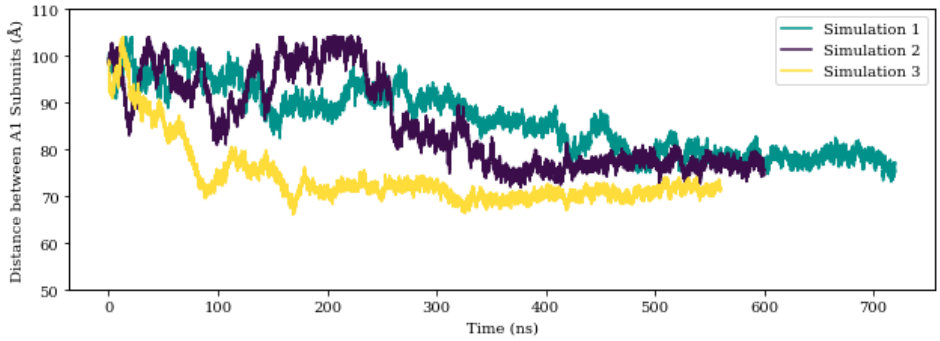


Figure 3.2 Distance between the centers of mass of two A1 subunits.

mean square coordinate deviation (RMSD) of α-carbons were calculated via the ptraj module of AMBER18 (see **Figures 3.2 and 3.3**). As shown in **Figure 3.2**, the three simulations converged. The distances between two A1 monomers in A1t tetramer stabilized at ~75 Å, which is similar to the experimental Cryo-EM results reported by Lambert et al. In their study, Lambert et al. investigated the organization of annexin and S100 proteins in the event of membrane aggregation/junctions.[72] Their Cryo-EM studies found that the A2-S100A10 tetramer induced junctions between two membranes. The Cryo-EM density analysis of the junction between two membranes indicated that the distance between two A2 monomers in A2-S100A10 tetramer is about 60Å.[72] Our A1t adopted a conformation with ~75Å between A1 monomers. Due to the

relative lengths of A1-ND (44 residues) and A2-ND (32 residues), it is reasonable for A1t to adopt a conformation which is slightly larger than that of the A2-S100A10 tetramer reported in the literature.

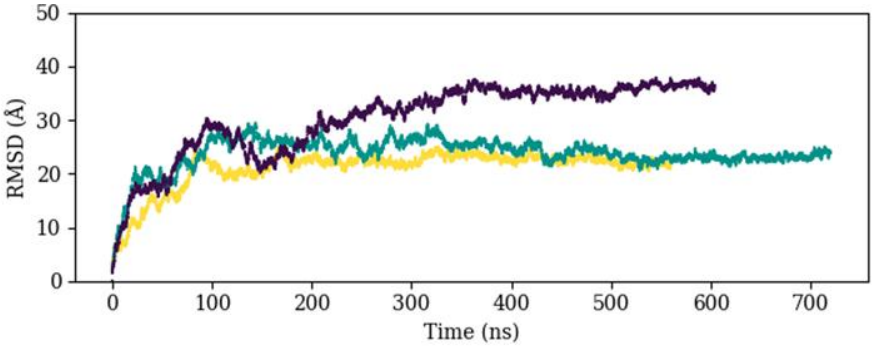


Figure 3.3 RMSD of α-carbons for the simulations.

3.3.1.1 PCA Analysis of A1 N-terminal domain (ND) within the A1t

Orientations of the A1-ND in A1t were explored using principal component analysis (PCA). The trajectories were aligned based on the S100A11 dimer, and PCA was performed for the first 23 residues of both A1-NDs of all three simulations. The data collected for the six NDs were combined into a single PCA plot where the trajectory frames were projected onto the first two principal components. A plot of the projections along with a kernel density estimate (KDE) overlay is shown in **Figure 3.4A**. From the KDE overlay, three regions with high densities of structures were identified. The most representative structures were extracted from the three regions (see **Figure 3.4 B-D**).

The three representative structures were colored based on their total interaction energies

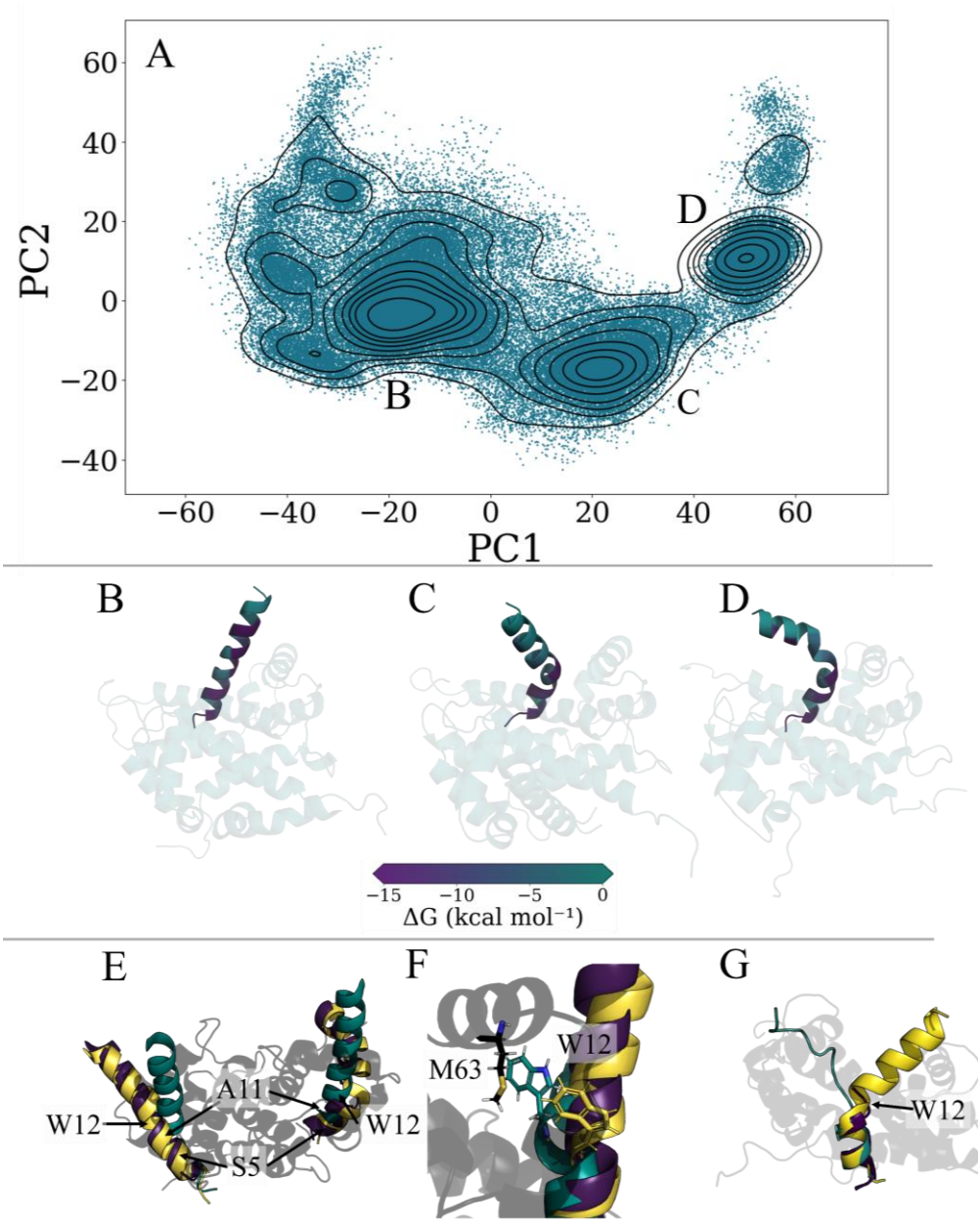


Figure 3.4. A) Scatter plot of projections of the trajectories for the N-terminal domain onto the first two principal components with a kernel density estimate overlay. B-D) Three representative structures. Color scale represents the strength of each individual residue's total interaction with the S100A11 dimer. E) Alignment for all six A1-ND in complex with the S100A11 dimer for representative structures B (yellow), C (teal), and D (purple). F) Close-up view of the difference in side-chain interaction between A1 W12 and S100A11 M63; linear conformation (yellow and purple) and kinked conformation (teal). G) Alignment of the A1-ND of Structure C (Yellow), 1QLS (purple), and 4HRE (teal).

with the S100A11 core (see **Figure 3.4 B-D**, **Table 3.1**). Structures C and D adopted a similar conformation which had a kink at W12 (see **Figure 3.4C, 3.4D**). Structures C and D differed only by the angle of the kink which resulted in slightly different interactions. In Structure B, the ND was a continuous helix with no kink (see **Figure 3.4B**).

Inspection of the side chain interactions between A1-ND and S100A11 for both kinked structures revealed a W12-M63 interaction between A1-ND and S100A11(see **Figure 3.4F**). The angled alpha helix placed W12 of A1 near the M63 sulfur of S100A11. This interaction was not possible for the linear conformation as shown in **Figure 3.4F**.

3.3.1.2 A1t Similarities to Crystal Structures

Figure 3.4E highlights the similarities of the first 11 residues of the simulated structures. For all 6 simulated NDs, the conformation of the first 11 residues was conserved as a helix. After residue 11, the structures diverged depending on the presence of the W12 kink. Comparison to the 1QLS crystal structure (**Figure 3.4G** in purple) revealed that the binding mode of the first 11 residues of our simulated A1-NDs was consistent with that of the A1-ND present in the 1QLS crystal structure. As previously described, 1QLS is the crystal structure of S100A11 in complex with the first 11 residues of the A1-ND in the presence of calcium. In the 1QLS structure, only the first 11 residues of the A1-ND were crystallized; this implicated that those remaining residues were more flexible and less stable in structure which was reflected in our simulations. Furthermore, the interaction between A2-ND and p11 within the A2-p11 tetramer (PDB: 4HRE) was also very similar to the interaction between A1 and S100A11 in our simulated A1t tetramer (**Figure 3.4G**). The first 10 residues of A2 from the x-ray structure (PDB Code: 4HRE, **Figure 3.4G** in teal) of the A2-p11 tetramer adopted a helical structure. Starting at residue 11, the A2-ND adopted a loop conformation until Repeat I of the A2 core domain. This was strikingly

consistent with our A1t structure.[73] Additional ongoing studies in our lab have observed the loss of A1-ND helical structure for simulations in the multi-μs timescale for A1 monomer upon calcium binding. This leads us to believe that eventually the A1-ND of our A1t may also lose some helical structure following its stable binding segment (first 11 residues) and become a disordered loop similar to the 4HRE structure.

Based on these observations, it appears that the first ~11 residues of Annexin NDs play a key binding role in forming Annexin-S100 tetramers. We suggest that the binding mode of the first 11 residues of Annexin NDs will be consistent for all the Annexin-S100 tetramer combinations and a common feature among these complexes.

3.3.2 Secondary Structure Analysis

Secondary structure analysis was performed to examine the overall change in A1 and S100A11 structures upon tetramer formation. A1 calcium binding sites are found at calcium binding loops connecting the helixes of its conserved core.

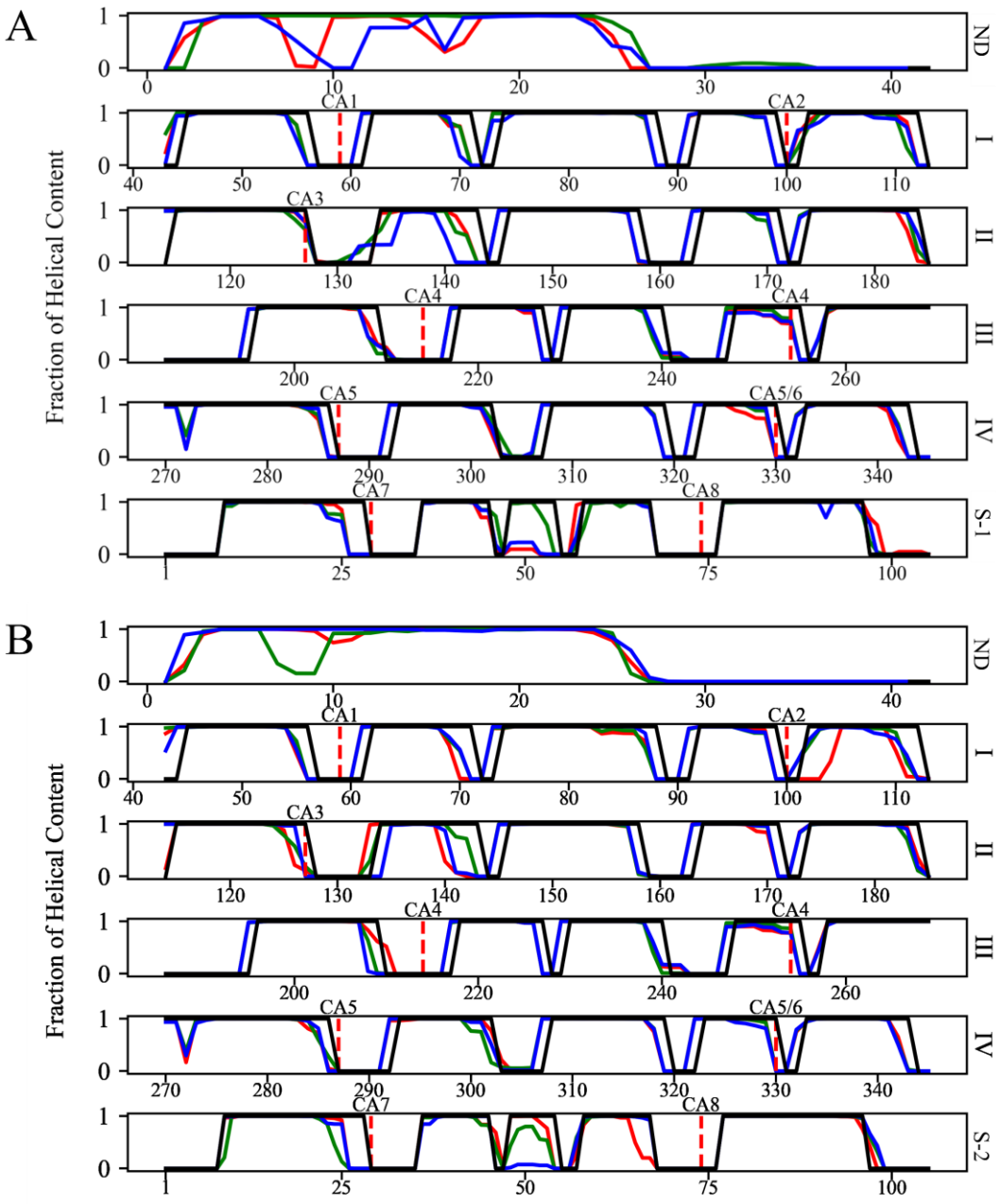


Figure 3.5. Fraction of Alpha helix content of the three simulations (Simulation 1:green, Simulation 2:blue, Simulation 3:red) for A1t divided into its components A) A1-1/S-1 and B) A1-2/S-2. Black represents the helix content from the 1MCX/2LUC PDB structures. Content is divided based on repeat, and calcium locations are shown with the dashed red lines.

Secondary structure analysis was performed using cpptraj. The fractions of alpha helix content for the three simulations were overlapped for A1-1/S-1 **Figure 3.5A** and A1-2/S-2

Figure 3.5B. Additionally, the fractions of alpha helix content for crystal structures 1MCX and 2LUC were overlapped with A1-1/A1-2 and S-1/S-2, respectively, to track the overall structural change upon tetramer formation. It should be noted that the N-terminus is not found in the 1MCX structure, thus the plot in black begins at residue 41. Furthermore, human A1 contains 6 calcium ions as opposed to the porcine annexin (1MCX), which contains 8 calcium ions. Consequently, calcium labels and numbers are based on the 6 binding sites of human annexin which are shared between the structures.

The simulated and crystal structures were mostly consistent in their secondary structures with no drastic conformational changes occurring within the A1 core domain upon A1 and S100A11 association. Analysis showed both A1-1 and A1-2 maintained all five alpha helices for each of the four Repeats (I-IV) in all three simulations. Additionally, the A1-1 and A1-2 calcium binding loops were preserved in the A1t.

The most significant differences between the simulated tetramer and the crystal structure were observed in the S100A11 dimer and the A1-NDs. A loss of helical content occurred at the Calcium 7 binding loop (EF1 calcium binding loop in S100A11). This expansion in the Calcium 7 binding loop was present for all three simulations. Additionally, a decrease in alpha helix content was observed at the site of S100A11 and A1 binding (S100A11 residues 48-55 and A1 residues 1-11). The loss of alpha helical structure within the A1-NDs was responsible for creating kinks in the A1-NDs. The different locations of the kinks contributed to the different ND conformations seen in **Figure 3.4**. In all three simulations, one A1-ND formed a kink while the other A1-ND retained its alpha helix. Additionally, variable levels of alpha helix structure were observed from residues 48-55 within the S100A11 regions.

3.3.3 Strong Interactions

3.3.3.1 Interactions between Representative N-terminal domain Structures and S100A11

MMGBSA calculations were performed to determine the strong interactions between 1) each representative structure from **Figure 3.4B-D** and the S100A11 dimer and 2) each subunit of A1t. All calculations were performed using the MMGBSA.py script of AMBER18 with a modified salt concentration of 0.15M.[40] For the energy calculations of the A1t subunits, 200 frames were sampled over the final 20 ns of each simulation. This ensured that each simulation had reached a stable conformation prior to energy calculations. The total energy contributions between each residue of the three representative structures with their respective S100A11 dimers were reported in **Table 3.1**. Consistent interaction energies were observed between the A1-ND and the S100A11 core for all three representative structures. Specifically, the interactions involving the first 12 residues were conserved. The following residues of the A1-ND were identified to have the strongest binding interactions with S100A11: W12, F7, V4, L8, S5, A2, and M3.

Table 3.1: Total interaction energies (in kcal/mol) of the first 12 residues of A1-ND with the S100A11 complex.

Residue	Structure B	Structure C	Structure D
A2	***-10.21***	***-13.64***	***-20.18***
M3	***-9.51***	***-12.80***	***-16.73***
V4	***-16.40***	***-20.08***	***-20.60***
S5	***-10.28***	***-13.25***	***-16.00***
E6	-1.44	-0.46	-1.04
F7	***-16.94***	***-14.63***	***-10.25***
L8	***-14.80***	***-16.42***	***-15.38***
K9	-9.11	-9.56	-11.76
Q10	-3.58	-4.33	-3.07
A11	-3.83	-5.08	-4.66
W12	***-32.56***	***-28.56***	***-22.15***

The residue-residue interaction pairs between S100A11 (S-1/S-2) and the A1-ND were reported in **Table 3.2**. For all three structures, the A1-ND had strong interactions with both S100A11 subunits. The strongest interactions involved either members of the first 12 A1-ND residues or acidic A1 residues (E-15 and E-18). Bolded in **Table 3.2**, A1-W12 interactions were similar between all three structures. However, Structures 2 and 3 had the additional strong interaction between W12 of A1 and M63 of S-2 which was not present for Structure B.

3.3.3.2 Interactions within the A1t

The 10 strongest residue-residue interactions between each tetramer component were summarized in **Table 3.3**. Strong interactions were reported for A1 monomers and both S100A11 subunits. Both S-1 and S-2 residues E11, I14, and E15 had multiple strong interactions with A1-1 and A1-2 ND residues respectively. The strongest interactions of the A1-S100A11 dimer pairs (A1-1:S-1 and A1-2:S-2) were highly consistent and involved Annexin ND residues, specifically, the first 11 residues (M3, V4, S5, E6, L8, K9). A1-S5 has been shown to be critical

for tetramer formation, as its phosphorylation inhibits A1-S100A11 association.[74] S5 of both A1-1 and A1-2 had strong interactions with E11 of S-1 and S-2 respectively. This interaction was present for all three structures demonstrating its importance within the A1t.

For A1-2:S-1 and A1-1:S-2, S100A11 residues K97 and M94 were important interaction sites. However, interacting residues within A1-1 and A1-2 differed for these regions. A1-2:S-1 involved residues 12-20 of the ND and Repeat II/IV residues (R213/R292) of A1-2. A1-1:S-2 relied on residues F7, W12, and Y39 of the ND and Repeat I residues (Q79, Q86). Additionally, A1-1:S-2 included the W12-M63 interaction which was reflected in the kinked conformation of A1-1. This structural feature facilitated interactions between Repeats I of A1-1 and S-2 which accounted for the asymmetric interactions and motion of the A1t (see section **3.3.5 Correlated Motions).**

Table 3.2: Pairwise MMGBSA data for the interaction of A1-ND and the S-100A11 core for the three representative structures. W12 interactions are in bold.

Structure B				Structure C				Structure D			
S-1/S-2	ND Residue	S100 Residue	TOTAL	S-1/S-2	ND Residue	S100 Residue	TOTAL	S-1/S-2	ND Residue	S100 Residue	TOTAL
S-2	E15	K66	-9.16	S-1	E18	K97	-9.14	S-2	E18	K97	-12.49
S-2	E18	R62	-9.15	S-2	A2	E15	-7.55	S-2	A2	E15	-10.59
S-2	E18	K66	-7.53	S-2	V4	E15	-6.83	S-1	A2	E15	-9.57
S-1	E15	K66	-5.73	S-2	S5	E11	-6.17	S-1	E18	K97	-8.19
S-1	E18	K66	-4.52	***S-1***	***W12***	***Q54***	***-5.41***	S-2	V4	E15	-6.30
S-1	S5	E11	-4.51	S-1	A2	E15	-5.33	S-2	S5	E11	-6.14
S-1	Q19	R62	-4.19	S-2	M3	E15	-4.56	S-2	Y21	K97	-5.18
S-2	***W12***	***Q54***	***-4.18***	***S-1***	***W12***	***N53***	***-4.02***	S-1	M3	E15	-4.94
S-1	***W12***	***F50***	***-3.87***	***S-2***	***W12***	***M63***	***-3.50***	S-2	M3	E15	-4.60
S-2	***W12***	***F50***	***-3.65***	S-1	F13	N53	-3.31	S-1	S5	E11	-4.56
S-2	A2	E15	-3.46	S-2	F7	S94	-3.27	S-1	E15	K97	-4.36
S-1	A2	E15	-3.46	S-1	Q10	N53	-3.26	S-1	V4	E15	-4.09
S-2	E15	R62	-3.38	S-1	S5	E11	-3.23	S-1	Y21	K97	-3.84
S-2	Q19	R62	-3.22	S-1	V4	E15	-3.08	***S-1***	***W12***	***N53***	***-3.62***
S-1	N16	Q54	-3.11	S-1	L8	F50	-2.92	***S-1***	***W12***	***K55***	***-3.51***
S-1	V4	E11	-2.81	S-1	F7	S94	-2.87	***S-2***	***W12***	***M63***	***-3.20***
S-2	L8	F50	-2.72	S-1	V4	E11	-2.46	S-1	L8	F50	-3.05
S-1	E15	R62	-2.46	S-2	N16	K66	-2.45	S-1	F13	N53	-2.95
S-1	F7	F95	-2.42	S-2	V4	E11	-2.43	S-2	K9	N53	-2.89
S-2	K9	N53	-2.38	S-2	F7	C91	-2.25	S-1	V4	E11	-2.57
S-2	F7	C91	-2.36	S-1	M3	F95	-2.11	S-1	Q10	N53	-2.48
S-2	F7	S94	-2.34	***S-1***	***W12***	***K55***	***-2.07***	***S-1***	***W12***	***Q54***	***-2.40***
S-1	V4	E15	-2.22	S-2	K9	F50	-1.98	S-1	T24	R104	-2.38
S-1	E18	R62	-2.18	S-1	K9	A49	-1.94	S-1	F7	C91	-2.32
S-2	V4	E15	-2.18	S-2	Q19	K66	-1.81	S-2	K9	F50	-2.13
S-1	***W12***	***L87***	***-2.11***	S-1	K9	N53	-1.73	S-2	V4	E11	-1.95
S-2	***W12***	***V59***	***-2.10***	S-1	F7	C91	-1.63	***S-2***	***W12***	***F50***	***-1.81***
S-1	F7	C91	-2.06	S-2	L8	E11	-1.59	S-2	L8	E11	-1.71
S-2	S5	E11	-2.02	***S-1***	***W12***	***V59***	***-1.55***	S-1	Y21	L96	-1.70
S-1	Q10	S94	-1.99	S-2	L8	L47	-1.50	S-2	F7	C91	-1.59

In **Figure 3.6A**, positive correlation was observed between A1-1 Repeat I and the S100A11 core while A1-2 displayed negative correlation between the same regions. This appeared to stem from the orientation of A1-1 with respect to the S100A11 dimer. As shown in

Table 3.3. The 10 strongest residue-residue interactions (in kcal/mol) between each A1 monomer and S100A11 monomer. **Δ**G AVG represents **Δ**G in kcal/mol averaged over the final 20 ns.

A1-1	A1-2	**Δ**G AVG	S-1	S-2	**Δ**G AVG
Q86	A2	-0.23	T8	E46	-12.21
L85	A2	-0.03	T10	E46	-11.01
K26	A2	-0.02	E46	T10	-6.38
M2	A2	-0.02	E9	E46	-6.11
K90	A2	-0.01	E9	S16	-4.83
K53	A2	-0.01	E9	R12	-4.13
A83	A2	-0.01	E46	T8	-3.91
K26	Q19	-0.01	K23	E9	-3.78
P32	A2	-0.01	H92	Q22	-3.49
Y21	A2	-0.01	H92	F77	-3.45

A1-1	S-1	**Δ**G AVG	A1-1	S-2	**Δ**G AVG
S5	E11	-4.41	Q86	K97	-5.71
V4	E11	-2.63	Q86	V99	-4.53
V4	E15	-2.28	W12	M63	-3.91
M3	E15	-1.86	F7	M94	-3.11
L8	E11	-1.56	Y39	K97	-3.10
V4	I14	-1.19	Q86	R104	-2.96
L8	I14	-0.81	S45	K97	-2.89
L8	T10	-0.75	L85	P100	-2.79
E190	M1	-0.58	Q79	M94	-2.59
V4	I18	-0.47	Y39	D93	-2.56

A1-2	S-1	**Δ**G AVG	A1-2	S-2	**Δ**G AVG
N16	K66	-10.43	E6	E11	-2.78
R213	T105	-7.01	V4	A2	-2.35
E20	R62	-6.94	M3	M1	-2.03
Q19	K66	-6.59	S5	E15	-1.96
F13	F50	-5.03	S5	E11	-1.76
R292	K97	-4.99	V4	I4	-1.73
E196	M94	-4.18	M3	A2	-1.44
E20	V59	-3.75	K9	E11	-1.25
A11	M94	-3.64	E6	I4	-1.21
A2	N31	-3.05	K9	I14	-1.14

Figure 3.6B, Repeat I of A1-1 was more closely associated to S-1 and S-2 than that of A1-2. Consequently, A1-1 Repeat I moved as a unit with the A1-ND-S100A11 complex hence the positive correlation. This demonstrates the importance of the A1-ND for the global motion of the A1t complex.

3.3.4 Calcium Binding

It is well-known that calcium binding is required for the formation of all Annexin-S100 heterotetramers except for the A2/S100A10 heterotetramer.[75] The calcium coordination was evaluated for the three simulations using cpptraj. Coordination sites were defined as oxygen atoms within 3.5 Å of calcium atoms. To screen for stable coordination sites, only oxygen atoms coordinated for at least 10% of the simulation were reported in **Table 3.4**. Simulated coordination sites for A1-1 and A1-2 were compared to the calcium coordination in the 1MCX crystal structure. 1QLS was used to compare calcium coordination sites for S100A11 residues. All 3 simulations were relatively consistent in their results. The coordination data for one simulation is reported below while the remaining results are in shown in **Table 3.4-3.6**.

Table 3.4: Calcium coordination comparison for simulated A1t and 1MCX and 1QLS. The fraction columns represent the fraction of the simulation during which the coordination was present.

Calcium	Type	Repeat	Crystal Comparison	A1-1/S-1 Trajectory	Fraction	A1-2/S-2 Trajectory	Fraction
1	II	I	-	G59@O	0.783	G59@O	0.998
			V60@O	V60@O	0.784	V60@O	0.998
			E62@E1	E62@OE1	0.964	E62@OE1	0.999
			E62@E2	E62@OE2	0.922	E62@OE2	1
			-	D334@OD1	0.195	-	-
			-	D334@OD2	0.414	-	-
2	III	I	K97@O	K97@O	0.166	K97@O	0.11
			L100@O	L100@O	0.139	-	-
			-	E105@OE1	0.838	E105@OE1	0.91
			E105@OE2	E105@OE2	0.85	E105@OE2	0.81
			-	E106@OE1	0.67	E106@OE1	0.754
			-	E106@OE2	0.692	E106@OE2	0.763
3	II	II	-	A126@O	0.63	-	-
			M127@O	-	-	-	-
			G129@O	G129@O	0.226	-	-
			G131@O	G131@O	0.4	-	-
			-	D133@OD1	0.877	-	-
			-	D133@OD2	0.502	-	-
			-	E134@OE1	0.376	-	-
			-	E134@OE2	0.326	E134@OE2	0.766
			-	-	-	E136@OE1	0.829
			D171@OD1	D171@OD1	0.879	D171@OD1	0.995
			D171@OD2	D171@OD2	0.935	D171@OD2	0.977
4	II	III	-	G210@O	0.114	G210@O	0.516
			-	E211@OE1	0.883	E211@OE1	0.632
			-	E211@OE2	0.891	E211@OE2	0.75
			R213@O	-	-	-	-
			G215@O	-	-	-	-
			E255@OE1	E255@OE2	0.997	E255@OE1	0.986
			E255@OE2	E255@OE1	0.985	E255@OE2	0.991
5	II	IV	M286@O	-	-	M286@O	0.975
			G288@O	-	-	G288@O	0.761
			G290@O	-	-	G290@O	0.762
			-	D329@OD1	0.627	-	-
			-	D329@OD2	0.578	-	-
			E330@OE1	E330@OE1	0.776	E330@OE1	0.996
			E330@OE2	E330@OE2	0.888	E330@OE2	0.992
6	III	IV	L328@O	L328@O	1	L328@O	0.716
			T311@O	T331@O	0.991	T331@O	0.711
			E336@OE1	E336@OE1	0.23	E336@OE1	0.925
			-	E336@OE2	0.804	E336@OE2	0.371
7	III	S100A11	D28@OD2	D28@OD2	0.999	D28@OD2	1
			N30@O	D28@OD1	0.999	D28@OD1	0.997
			K33@O	T33@O	0.998	T33@O	0.999
			E38@OE1	E38@OE1	0.886	E38@OE1	0.944
			E38@OE2	E38@OE2	0.883	E38@OE2	0.95
			-	Q74@OE1	0.109	-	-
			-	-	-	T105@O	0.51
8	III	S100A11	D68@OD1	D68@OD1	0.721	D68@OD1	0.683
			-	D68@OD2	0.283	D68@OD2	0.322
			D70@OD2	N70@OD1	0.979	N70@OD1	0.638
			-	D72@OD1	0.372	D72@OD1	0.526
			D72@OD2	D72@OD2	0.976	D72@OD2	0.996
			Q74@O	Q74@O	1	Q74@O	1
			-	D76@OD1	0.373	D76@OD1	0.505
			-	D76@OD2	0.628	D76@OD2	0.467
			E79@OE1	E79@OE1	1	E79@OE1	1
			E79@OE2	E79@OE2	1	E79@OE2	0.994

Table 3.5: Calcium Coordination for Simulation 2

Calcium	Type	Repeat	Crystal Comparison	A1-1/S-1 Trajectory	Fraction	A1-2/S-2 Trajectory	Fraction
1	II	I	-	G59@O	0.902	G59@O	1
			V60@O	V60@O	0.901	V60@O	1
			E62@E1	E62@OE1	1	E62@OE1	1
			E62@E2	E62@OE2	1	E62@OE2	0.999
			-	D334@OD1	0.522	D334@OD1	0.488
			-	D334@OD2	0.359	-	-
2	III	I	K97@O	-	-	K97@O	0.957
			L100@O	-	-	L100@O	0.746
			-	H-103@O	0.644	-	-
			-	E105@OE1	0.891	E105@OE1	0.629
			E105@OE2	E105@OE2	0.902	E105@OE2	0.758
			-	E106@OE1	0.907	-	-
			-	E106@OE2	0.912	-	-
			-	-	-	T132@O	0.638
3	II	II	M127@O	M127@O	0.781	-	-
			G129@O	G129@O	0.51	-	-
			-	-	-	L130@O	0.22
			G131@O	G131@O	0.479	-	-
			-	-	-	E134@OE1	0.611
			-	-	-	E134@OE2	0.803
			D171@OD1	D171@OD1	0.995	D171@OD1	0.993
			D171@OD2	D171@OD2	0.997	D171@OD2	0.972
4	II	III	-	G210@O	0.86	G210@O	0.548
			-	E211@OE1	0.587	E211@OE1	0.638
			-	E211@OE2	0.631	E211@OE2	0.789
			R213@O	-	-	-	-
			G215@O	-	-	-	-
			E255@OE1	E255@OE1	0.987	E255@OE1	0.991
			E255@OE2	E255@OE2	0.983	E255@OE2	0.991
5	II	IV	M286@O	M286@O	0.153	M286@O	0.615
			G288@O	G288@O	0.142	G288@O	0.572
			G290@O	G290@O	0.144	G290@O	0.57
			E255@OE1	-	-	-	-
			E255@OE2	-	-	-	-
			-	D329@OD1	0.65	D329@OD1	0.234
			-	D329@OD2	0.531	D329@OD2	0.173
			-	E330@OE1	0.982	E330@OE1	0.996
				E330@OE2	0.329		
6	III	IV	L328@O	L328@O	1	L328@O	0.404
			T311@O	T331@O	0.998	T331@O	0.4
			E336@OE1	E336@OE1	1	E336@OE1	0.696
			-	-	-	E336@OE2	0.928
7	III	S100	-	A25@O	0.2	-	-
			-	K27@O	0.227	-	-
			-	D28@OD1	0.776	D28@OD1	0.986
			D28@OD2	D28@OD2	0.806	D28@OD2	0.679
			N30@O	-	-	-	-
			K33@O	-	-	T33@O	0.673
			E38@OE1	E38@OE1	0.802	E38@OE1	0.893
			E38@OE2	-	-	E38@OE2	0.921
8	III	S100	D68@OD2	D68@OD1	0.911	D68@OD1	1
			D70@OD2	D68@OD2	0.758	-	-
			D72@OD2	-	-	D72@OD1	0.995
			-	D76@OD1	0.932	D72@OD2	0.992
			Q74@O	Q74@O	1	Q74@O	1
			-	-	-	D76@OD1	0.365
			-	-	-	D76@OD2	0.363
			E79@OE1	E79@OE1	1	E79@OE1	1
			E79@OE2	E79@OE2	1	E79@OE2	1

Table 3.6: Calcium Coordination for Simulation 3

Calcium	Type	Repeat	Crystal Comparison	A1-1/S-1 Trajectory	Fraction	A1-2/S-2 Trajectory	Fraction
1	II	I	-	E59@O	0.957	G59@O	1
			V60@O	V60@O	0.958	V60@O	1
			E62@E1	E62@OE1	0.999	E62@OE1	1
			E62@E2	E62@OE2	1	E62@OE2	1
			-	D334@OD1	0.943	D334@OD1	0.202
			-	-	-	D334@OD2	0.716
2	III	I	K97@O	-	-	K97@O	0.101
			L100@O	-	-	-	-
			-	E105@OE1	0.833	E105@OE1	0.952
			E105@OE2	E105@OE2	0.938	E105@OE2	0.949
			-	E106@OE1	0.733	E106@OE1	0.598
			-	E106@OE2	0.608	E106@OE2	0.64
3	II	II	-	-	-	A126@O	1
			M127@O	M127@O	0.818	-	-
			G129@O	G129@O	0.708	G129@O	0.215
			G131@O	G131@O	0.495	G131@O	0.209
			-	-	-	D133@OD1	0.792
			-	-	-	D133@OD2	0.906
			D171@OD1	D171@OD1	0.887	D171@OD1	0.998
			D171@OD2	D171@OD2	0.985	D171@OD2	0.997
4	II	III		G210@O	0.119	G210@O	0.116
			-	E211@OE2	0.929	E211@OE2	0.659
			-	E211@OE1	0.896	E211@OE1	0.657
			R213@O	-	-	-	-
			G215@O	-	-	-	-
			E255@OE1	E255@OE1	0.983	E255@OE1	0.709
			E255@OE2	E255@OE2	0.955	E255@OE2	0.706
5	II	IV	M286@O	M286@O	0.793	M286@O	0.948
			G288@O	G288@O	0.749	G288@O	0.843
			G290@O	G290@O	0.749	G290@O	0.863
			E255@OE1	E330@OE1	0.987	E330@OE1	0.997
			E255@OE2	E330@OE2	0.996	E330@OE2	0.999
			-	D329@OD1	0.105	-	-
6	III	IV	L328@O	L328@O	0.988	L328@O	1
			T331@O	T331@O	0.981	T331@O	0.991
			E336@OE1	E336@OE1	0.235	-	-
			-	E336@OE2	0.794	E336@OE2	0.972
7	III	S100	-	K24@O	0.77	-	-
			-	D28@OD1	0.991	D28@OD1	0.999
			D28@OD2	D28@OD2	0.989	D28@OD2	0.991
			N30@O	-	-	-	-
			-	-	-	T33@OG1	0.378
			K33@O	T33@O	0.91	T33@O	0.998
			E38@OE1	E38@OE1	0.806	E38@OE1	0.574
			E38@OE2	E38@OE2	0.806	E38@OE2	0.579
			-	Q74@OE1	0.108	Q74@OE1	0.416
8	III	S100	-	-	-	D68@OD1	0.577
			D68@OD2	D68@OD1	0.997	-	-
			-	D68@OD2	0.863	D68@OD2	0.441
			-	N70@O	0.476	-	-
			D70@OD2	N70@OD1	0.793	-	-
			-	D72@OD1	0.8	D72@OD1	0.989
			D72@OD2	D72@OD2	0.879	D72@OD2	0.996
			Q74@O	Q74@O	0.996	Q74@O	1
			-	-	-	D76@OD1	0.47
			-	-	-	D76@OD2	0.502
			E79@OE1	E79@OE1	0.845	E79@OE1	1
			E79@OE2	E79@OE2	0.794	E79@OE2	0.995

structures indicating that the simulations performed in this study were reasonable. By comparison, we observed that the simulated calcium coordinations for A1-1 and A1-2 were consistent. Similarly, the simulated calcium coordinations for S-1 and S-2 were also consistent. Particularly, we noticed that the X-ray coordinations for S100A11 calcium ions 7 and 8 were present for at least 60% of the simulation.

Overall, our results show good agreement with available experimental crystal structures. Experimental conditions may influence calcium coordination numbers for structures determined through X-ray crystallography. Our simulated results were not expected to be identical to X-ray data as MD simulations investigate the dynamics of each coordination. In this work, we observed the dynamic interactions between calcium and the A1t and provided detailed atomistic information pertaining to the calcium coordination within the A1t. The calcium binding information we reported may be valuable for future mutation studies investigating the impact of calcium binding on the biological functions of the A1t.

3.3.5 Correlated Motions

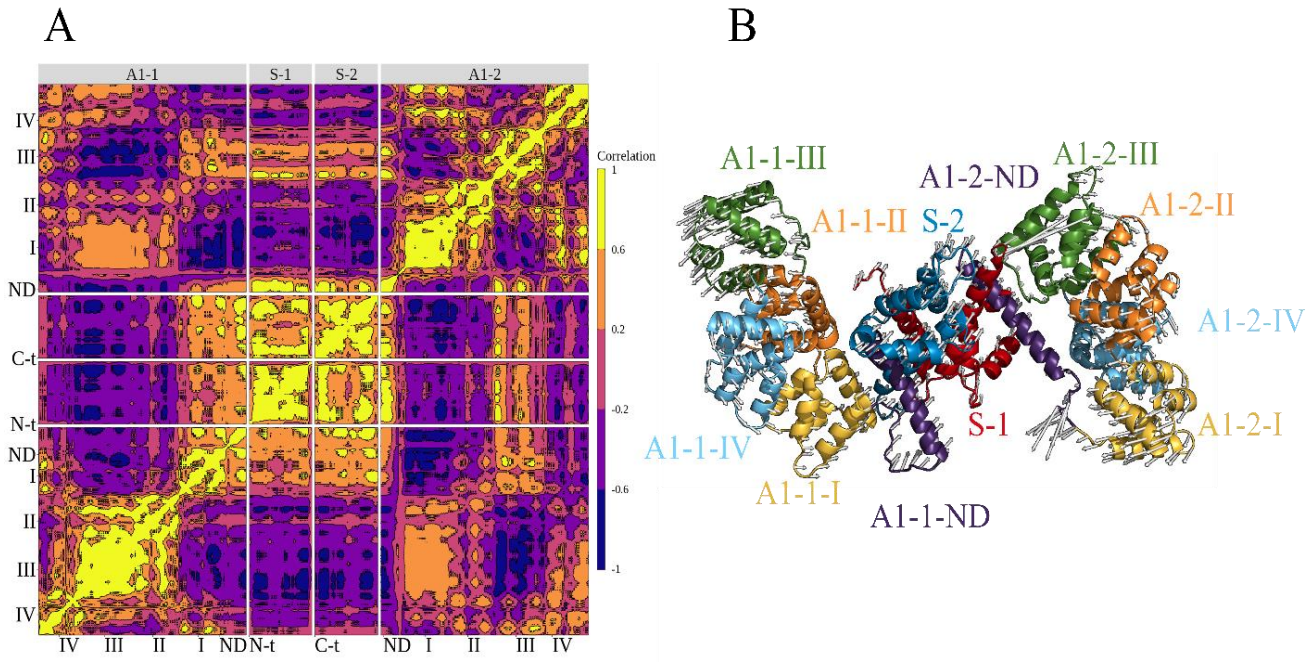


Figure 3.6 A) Cross-correlation heatmap of alpha carbons for A1t complex. Dark purple represents anticorrelated motion while yellow is representative of positively correlated motion. B) Graphical depiction of the first normal mode.

Cross-correlation analysis has been an important tool for assessing structural domain dynamic correlation during MD simulations. In this work, cross correlation analysis was employed to explore the correlated motions within A1t. Cross correlation analysis between the α-carbons of all residues was performed via the cpptraj module of AMBER.

The correlation heatmaps generated for all three simulations were consistent in their overall trends. One plot is shown in **Figure 3.6A** while the other plots are found in **Figure 3.7-3.8**. Overall, the heatmaps demonstrated correlated motion throughout A1t. A prominent feature shared by all three correlation maps was the strong positive correlation between the S100A11 dimer and the NDs of both A1-1 and A1-2. Shown in **Figure 3.6A**, the S100A11 dimer was negatively correlated to Repeat II in both A1-1 and A1-2. The asymmetric nature of A1t was

highlighted when comparing the correlations of A1 Repeats I and III with the S100A11 dimer. For A1-1, Repeat I was positively correlated and Repeat III was negatively correlated with the S100A11 dimer. Conversely, Repeats I and III of A1-2 exhibited negative and positive correlation respectively.

Within each A1 monomer, Repeats I and IV, II and III were positively correlated. However, Repeat III was negatively correlated with Repeat I. Between the two A1 monomers, strong negative correlations were observed between A1-1 Repeats II and III together and Repeat III of A1-2, and between A1-ND and Repeat I together and Repeat I of A1-2. This long-distance phenomenon was visualized through the depiction of first normal modes (**Figure 3.6B**). The normal modes depicted a rocking motion of the tetramer with the A1-1 Repeats II and III moving in the opposite directions of A1-2 Repeats II and III and A1-1 Repeat I moving in opposite directions of A1-2 Repeat I. The A1 region in close association with the S100A11 core, the ND of A1, was positively correlated to both S-1 and S-2, and they moved together.

The remaining heatmaps (**Figure 3.7-3.8**) also reflected the asymmetric correlation within A1t albeit through different A1 repeats. This was a consequence of the multiple orientations of A1 monomers allowed by the different A1-ND conformations. However, the defining trend seen in all our A1t conformations was the strong positive correlation between the A1-ND and the S100A11 dimer. Based on the similarity to the A2t (see **Section 3.3.1: A1t Structure and Conformation**) and the consistency among the three simulated structures of the A1t, we believe this stable, correlated complex between Annexin-NDs and the S100 proteins will be a theme among the other Annexin-S100 tetramers.

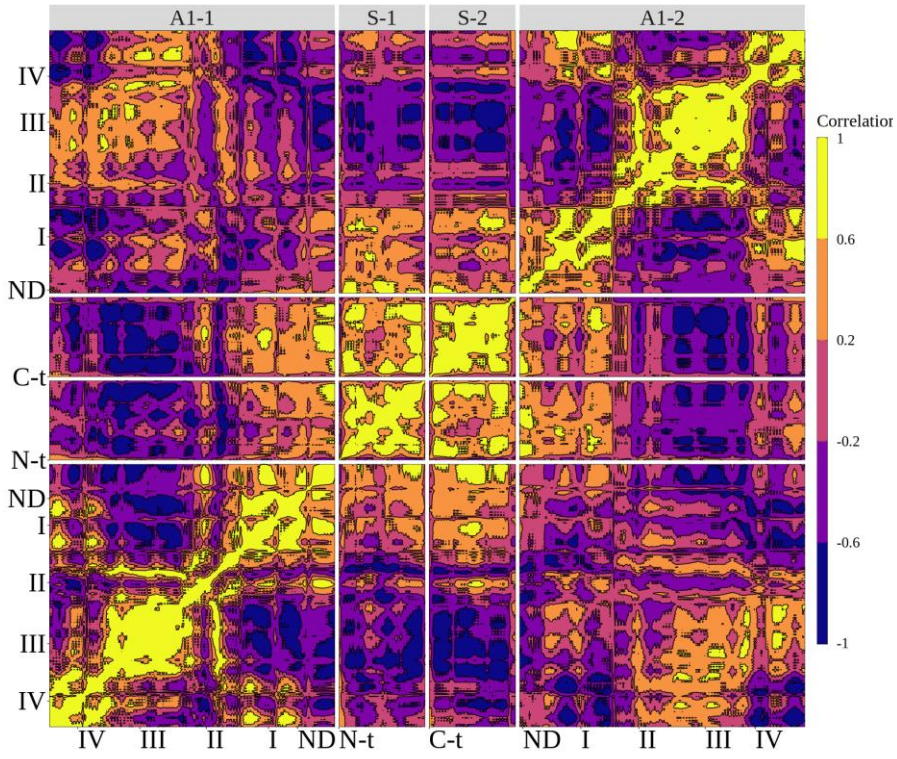


Figure 3.7 Cross-Correlation heatmap procured from alpha carbons for the A1t Simulation 3. Dark purple represents anti correlated motion while yellow is representative of positively correlated motion.

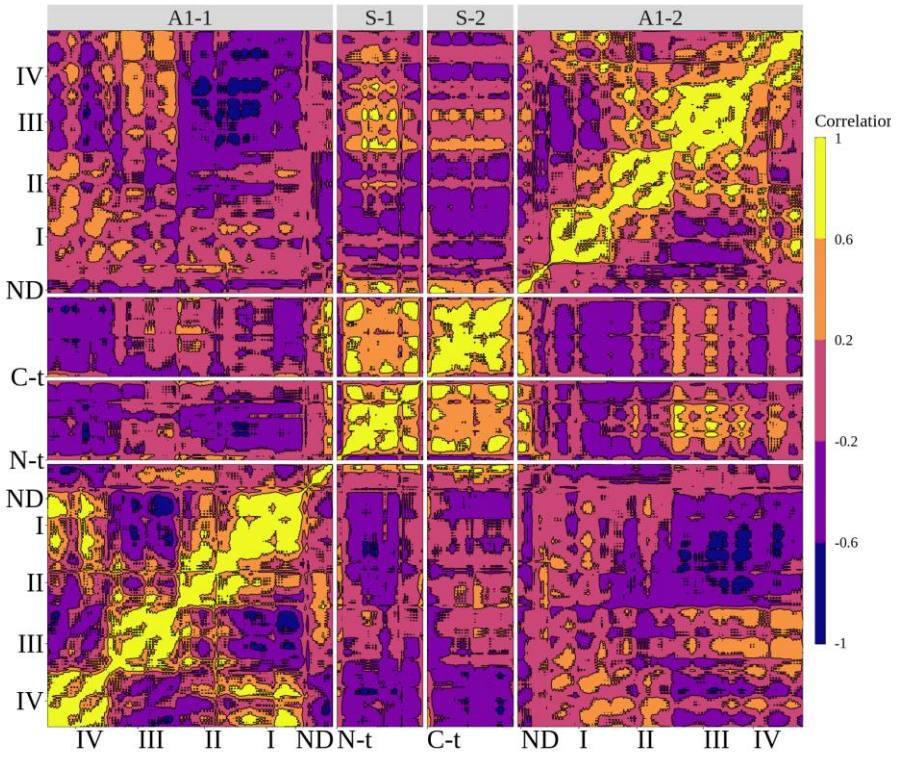


Figure 3.8 Cross-Correlation heatmap procured from alpha carbons for the A1t Simulation 2. Dark purple represents anti correlated motion while yellow is representative of positively correlated motion.

3.3.6 Conclusions

In this study, multiple molecular dynamics simulations were performed on a complete model of the A1-S100A11 heterotetramer (A1t) for the first time. Three simulations resulted in stable structures of A1t with ~75Å distances between two A1 monomers. A1t structures adopted a range of conformations afforded by the flexibility of the A1-ND. When comparing these structures, the orientations of the first 11 residues of the NDs, with reference to the S100A11-dimer, were conserved. Furthermore, the S100A11 interaction with the first 11 residues of the simulated A1-ND were consistent with both the crystal structure of A1-ND and S100A11 in 1QLS and the crystal structure of A2 and S100A10 in 4HRE. PCA was used to identify 3 representative structures of the A1-ND which adopted two different conformations. These conformations differed by the presence of a kink in alpha helix structure at W12 of A1 which resulted in a new interaction between W12 and M63 of S100A11. MMGBSA results show strong, consistent interaction energies for all three representative structures for the first 11 residues of the A1-ND (A2, M3, V4, S5, F7, L8) with the S100A11 core. Additionally, strong interactions were identified between the A1-NDs and both S100A11 subunits present in the A1t including the A1-S5 and S100A11-E11 interaction. A1-S5 has been demonstrated in the literature to be important for the association of A1 and S100A11. Calcium binding of the A1t was evaluated and compared to available crystal structures showing strong agreement and stable coordination especially within the S100A11 dimer. The detailed coordination data reported in this work may be important for future mutation studies on the A1t. Furthermore, strong, correlated motion was observed throughout the A1t. Despite having different orientations of the A1 monomers, all three simulations displayed strong positive correlation between A1-NDs and the S100A11 core. These results not only demonstrate the importance of the A1-ND for the

structure and dynamics of the A1t, but also provide insight into Annexin-S100 complexes. Based on the conserved interactions of the first 11 residues of A1 between our different structures and their consistency with available crystal structures for the A2t, we believe that the stable complex between Annexin ND and S100 dimer core will be a common feature of other Annexin-S100 complexes.

CHAPTER 4: THEORETICAL STUDIES ON SQUARE PLANAR PLATINUM COMPLEXES

4.1 Background

A previous study by Barnham et al. investigated the catalytic cycle of platinum-catalyzed C-H acylation of 2-aryloxypyridines. The preliminary ligand exchange step for this study is shown in **Figure 4.1.** In the study, ligand exchange as well as subsequent intramolecular C-H activation steps for cis isomers exhibited higher reactivities than those of trans isomers. Interestingly, the cis-trans isomerization was demonstrated to be competitive throughout the catalytic cycle suggesting the potential for multiple rapid reaction pathways all starting from the cis-$Pt(PhCN)_2Cl_2$.[76] As previously discussed, C-H activation in square planar platinum compounds can occur through a variety of mechanisms. Of particular interest, stable five-coordinated intermediates may result from oxidative addition pathways which may be difficult to isolate experimentally. In this chapter, a hybrid HF-DFT approach was employed to investigate the first step of ligand exchange shown in **Figure 4.1.**

4.2 Methodology

All calculations were performed with Gaussian 09 software on a computer cluster of 640 processors housed at East Carolina University, and initial structures were built using GaussView09.[77] Geometry optimizations, energy calculations, and harmonic frequencies were

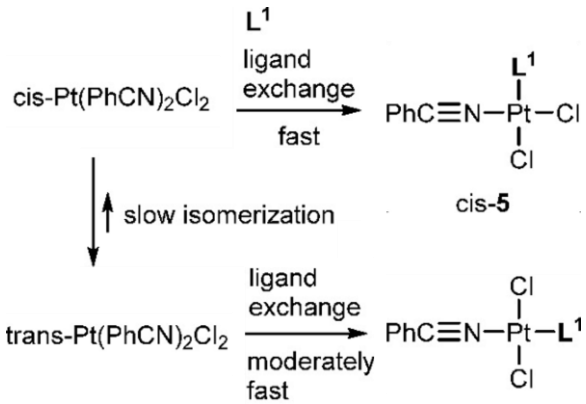


Figure 4.1. Previous mechanism from study from Barnham et al.

calculated using the M06-2x new hybrid meta exchange correlation functional.[78] Our lab has previously compared a variety of functionals on similar platinum systems, and M06-2x has been the optimal method for mechanistic studies of this type of platinum system. Calculations used the def2-svp basis set including f and g polarization functions for platinum atoms. All other atoms were optimized with the modest 6-31G** basis set.[79,80] Solvent effects were modeled using the polarizable continuum model within the self-consistent reaction field method. Dichloromethane was selected as the solvent to approximate experimental conditions.[81]

The Nudged Elastic Band (NEB) method was employed to determine the potential energy surface connecting two optimized reactants and products. The NEB method utilizes a series of images connecting the two end points. For each image, the total force (gradient) is calculated and then projected into its components which are either parallel to the plane of the images or perpendicular. Additional force constants (spring forces) are added to the parallel force to maintain continuity and equal spacing between the images while the perpendicular forces are minimized which drives the series of images to the minimum energy path. NEB calculations were performed on 6 reaction paths each using 20 images. Optimization during the NEB calculations was performed via the FIRE optimizer. NEB calculations usually take several hundred cycles to fully converge and yield a transition state. However, it is often more practical to run the NEB until forces around the transition state are sufficiently converged. Better estimates of the transition states were obtained by using the climbing image feature within the NEB program; this feature allows the highest energy image to be as close to the saddle point as possible. NEB calculations employed the same level of theory and PCM solvation as geometry calculations.[53]

Transition states were optimized from the NEB pathway using both the dimer method developed by Henkelman and the Berny Algorithm. The dimer method is a local gradient following algorithm which is easily implemented in NEB calculations. Instead of calculating the full Hessian, the dimer method only calculates the lowest eigenpair. In terms of an NEB, the eigenvector is defined by two adjacent images nearest the saddle point. This vector is subsequently rotated until finding the lowest curvature mode; the algorithm then moves in this direction until a saddle point is reached.[82]

Vibrational frequencies and force constants were calculated for all optimized structures and transition states at their same levels of theory. Vibrational frequencies were calculated from the second derivative of molecular energy with respect to the nuclear coordinates. Frequency calculations provided zero-point energy corrections as well as molecular normal modes. Stable structures were confirmed by the absence of negative frequencies while transition states were confirmed by the presence of one imaginary frequency.

4.3 Computational Results

4.3.1 Optimization of Structures

Two systems each differing in their model substrates (L1) were studied in this work: Model System and Complete System. In the Model System, the associating ligand was pyridine while the Full System used 2-(2-methylphenoxy)pyridine. For the Model and Complete Systems, both cis and trans pathways were investigated. The starting structures for these pathways were cis-Bis(benzonitrile)chloroplatinum and trans-Bis(benzonitrile)chloroplatinum respectively. The corresponding product structures involved replacing one benzonitrile ligand with their respective L1s while conserving stereochemistry.

Prior to NEB studies, geometry optimization was performed to locate stable structures for the reactants and products without the presence of L1. For both systems, the trans reactants and products were the more stable isomer as indicated in **Table 4.1.**

Structure	Reactants	Products (Model)	Products (Full)
Cis	-1688.178941	-1612.027853	-1957.363617
Trans	-1688.184047	-1612.033785	-1957.368443

Table 4.1 Reactant and Product zero-point corrected energies for the cis/trans Model and Full systems.

Product and reactant complexes (including L1s) were minimized and used for the start and endpoints for NEB calculations. The structures for these structures are shown in **Figures 4.4** and **4.7.** In the model cis/trans systems, the products and reactants formed near perfect square planar structures with respect to the platinum centers. The trans reactant adopted a conformation with the phenyl ligands in the same plane as each other; however, the phenyl ligands rotated to be perpendicular in the product structure. Reactant bond lengths for Pt-N1 and Pt-N2 were 1.990 and 1.991 Å respectively. Pt-N2 Product bond length increased to 2.019 Å and the leaving N1 group was replaced with a Pt-N3 bond of length 2.029 Å. For the cis structures, the phenyl groups for both products and reactants were rotated due to steric effects. For the trans reactants, the Pt-C1 and Pt-Cl2 bonds were both 2.361 Å and 2.361 and 2.362 Å respectively for the products.

The geometries of the complete cis/trans systems were very similar to the model system. However, both cis and trans isomers trans adopted starting structures with both benzonitrile groups in plane with each other. Coordination of 2-(2-methylphenoxy)pyridine yielded product

structures with pyridine rotated to be perpendicular to the phenyl group of the remaining benzonitrile moiety. All structures adopted near perfect square planar geometries.

4.3.2 Mapping Reactions via NEB

NEB was performed to map the potential energy surfaces connecting the optimized reactant and product complexes. These PES were utilized to help identify saddle point structures and determine the activation barrier for the corresponding reaction steps. Geometries and structures for the Model and Complete systems are shown in **Figures 4.4 and 4.7** and **Table 4.2** and **4.3**.

4.3.2.1 Model System

The NEB pathway for the model cis pathway is shown in **Figure 4.2.** The NEB calculations on this reaction step produced a modestly steep MEP with a single transition state. The increase in potential energy was identified to be due to the dissociation of the Pt-N1 bond and simultaneous association of the Pt-N3 bond. The geometry of its transition state suggested a concerted mechanism with **A** being the point at which the Pt-N3 bond is on the cusp of formation (**Rc**-N1: 1.990 Å, **A-**N1: 2.382 Å, **A-**N3: 2.528 Å). The TS was optimized as shown in **Figure 4.4** and the activation barrier was determined to be 16.78 kcal/mol.

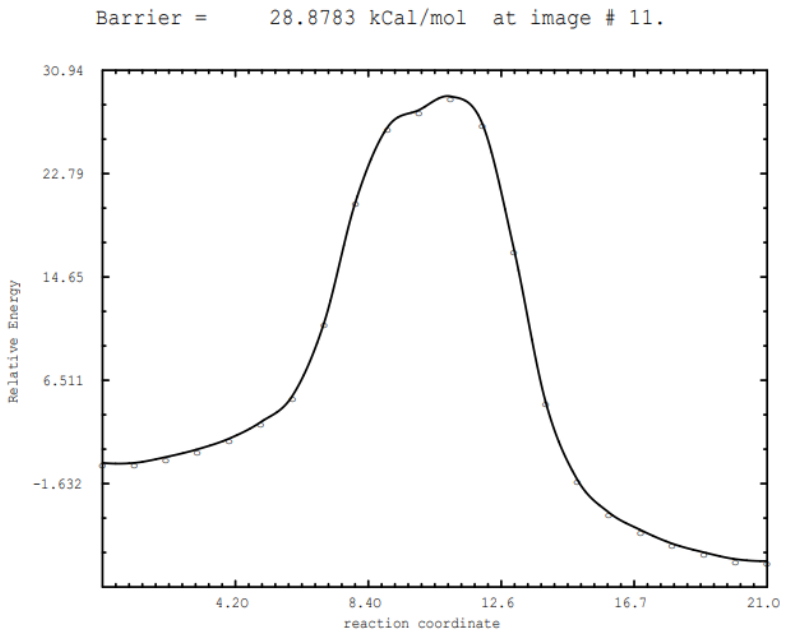


Figure 4.2 NEB potential energy curve for the cis pathway of the Model System

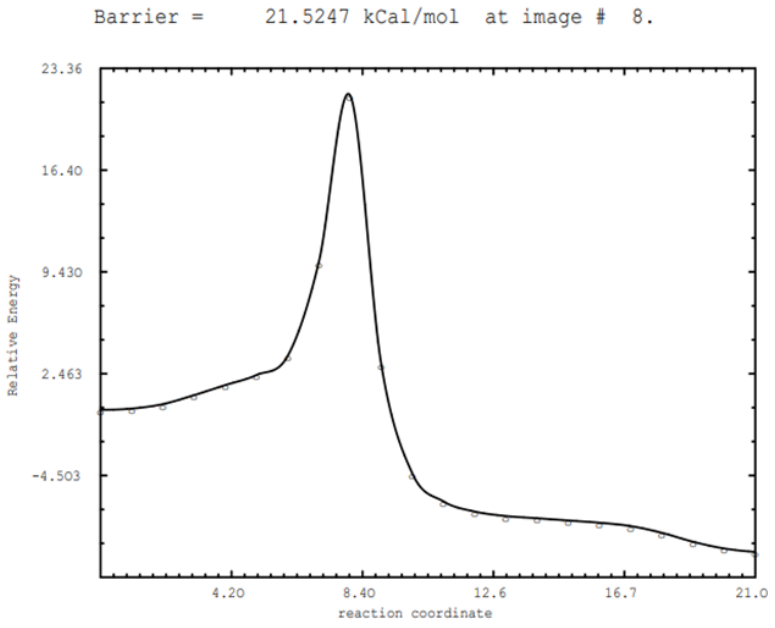


Figure 4.3 NEB potential energy curve for the trans pathway of the Model System.

The NEB pathway for the model trans system is shown in **Figure 4.3.** The MEP of this pathway displayed a rapid increase in potential energy. Like its corresponding cis pathway, the model trans system also underwent a concerted mechanism. The rapid change in energy was attributed to the simultaneous dissociation of Pt-N1 bond and the formation of the Pt-N3 bond (**Figure 4.4** and **Table 4.2**). The TS was optimized to yield a barrier of 16.05 kcal/mol.

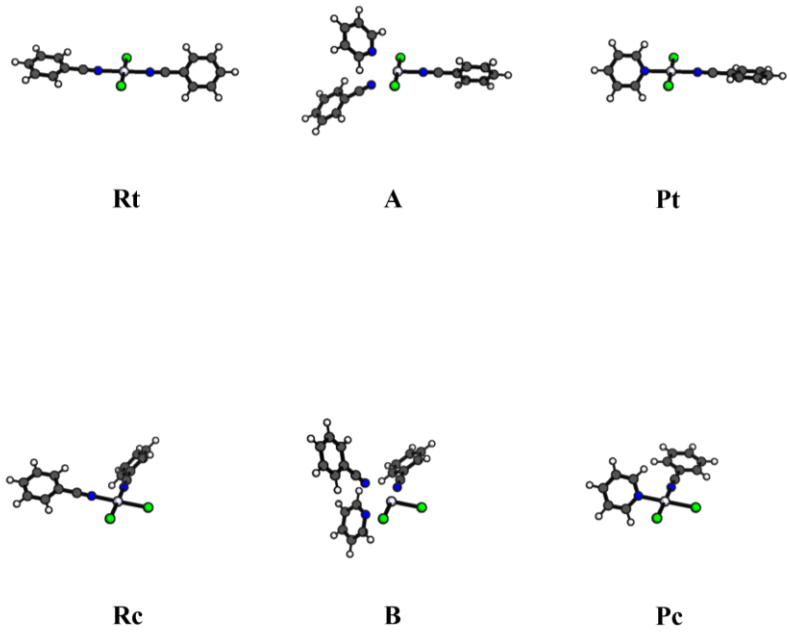


Figure 4.4. Optimized structures for Full System cis and trans pathways (**Figures 4.2** and **4.3**). Atoms bonded to platinum/in the coordination sphere are labeled.

Table 4.2. Bond Lengths (Å) and angles (degrees) for optimized structures.

	Rt	**A**	**Pt**	**Rc**	**B**
Pt-Cl1	2.361	2.354	2.361	2.328	2.329
Pt-Cl2	2.361	2.356	2.362	2.328	2.324
Pt-N1	1.990	2.382		2.025	2.410
Pt-N2	1.991	1.948	2.019	2.022	2.002
Pt-N3		2.528	2.029		2.605
Cl1-Pt-Cl2	179.547	179.591	179.873	91.498	91.868
Cl1-Pt-N1	90.285	89.403		89.218	89.258
Cl1-Pt-N2	90.025	90.248	89.402	178.955	178.123
Cl1-Pt-N3		90.008	90.124		91.395

4.3.2.2 Full System

The NEB for the full cis system (**Figure 4.5**) showed a gradual increase in potential energy and a broad peak similar to the PES of the modeled cis system. This reaction step was determined to progress via a concerted mechanism with the simultaneous dissociation of the benzonitrile group and the association of the ligand substrate. Transition state optimization led to the structure shown in **Figure 4.7** with an energy barrier of 18.65 kcal/mol.

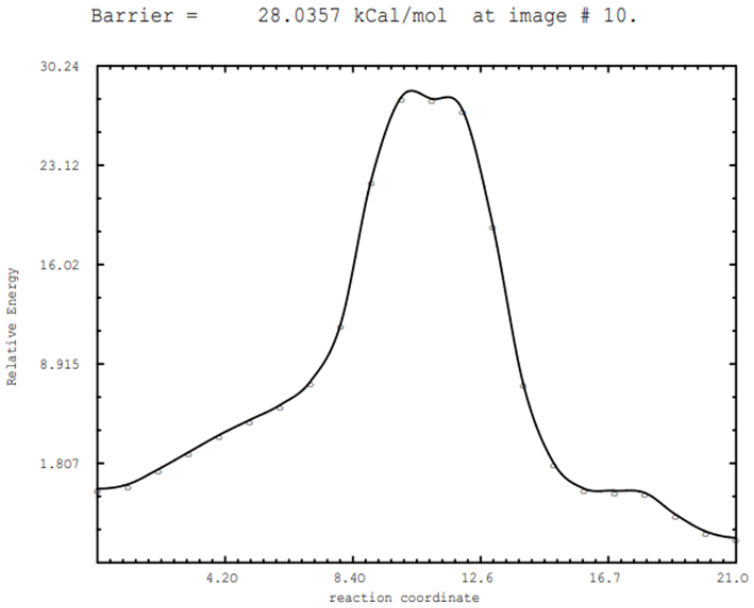


Figure 4.5 NEB potential energy curve for the cis pathway of the Full System

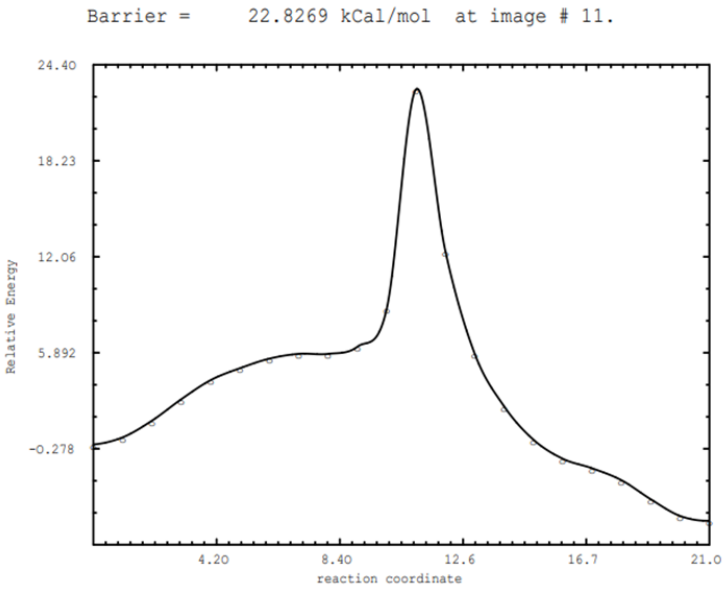


Figure 4.6 NEB potential energy curve for the trans pathway of the Full System.

The NEB for the full trans is shown in **Figure 4.6**. Like its model counterpart, the PES displayed a sharp increase in energy near the TS. However, it exhibited a much more gradual increase at the start of the reaction coordinate. The TS was optimized (**Figure 4.7**) to determine a barrier of 20.13 kcal/mol.

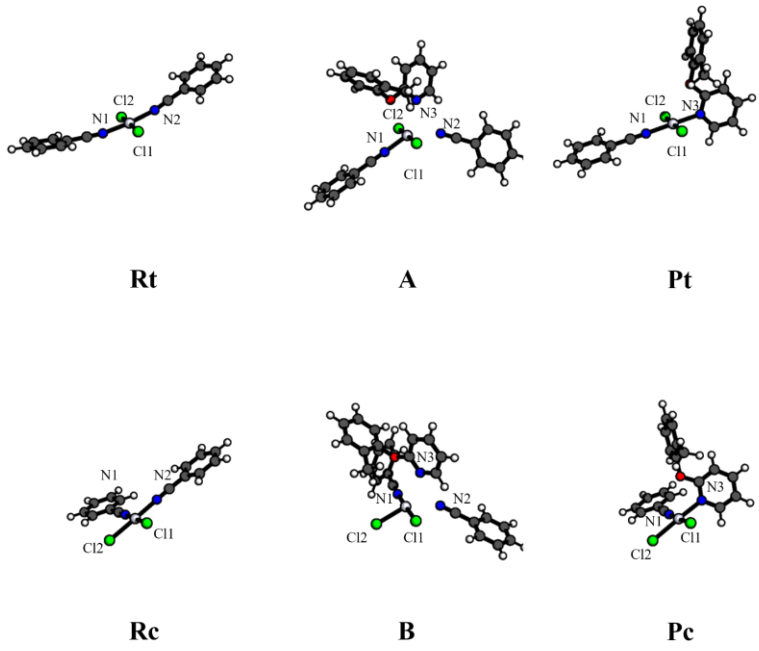


Figure 4.7 Optimized structures for Full System cis and trans pathways (**Figures 4.5** and **4.6**). Atoms bonded to platinum/in the coordination sphere are labeled.

Table 4.3 Bond Lengths (Å) and angles (degrees) for optimized structures.

	RT	**A**	**Pt**	**RC**	**B**	**PC**
Pt-Cl1	2.364	2.356	2.364	2.331	2.336	2.335
Pt-Cl2	2.358	2.355	2.363	2.331	2.319	2.353
Pt-N1	1.988	1.953	2.009	2.016	1.994	2.016
Pt-N2	1.996	2.408		2.021	2.469	
Pt-N3		2.535	2.039		2.620	2.069
Cl1-Pt15-Cl2	176.298	179.071	178.913	91.655	91.898	92.037
Cl1-Pt15-N1	89.700	89.432	90.491	179.431	178.110	178.204
Cl1-Pt15-N2	90.478	87.959		88.733	87.134	
Cl1-Pt15-N3		93.716	89.255		88.639	88.572

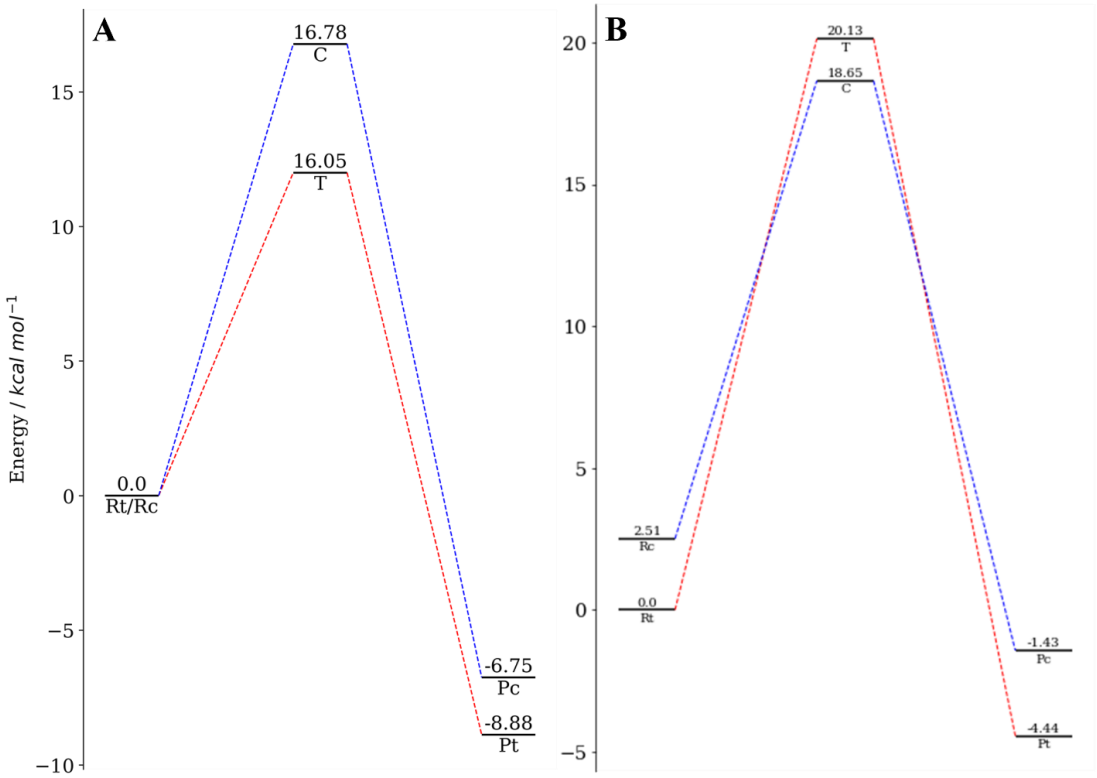


Figure 4.8 Energy diagrams for the Model System (**A**) and the Full System (**B**)**.** Cis and trans isomers are colored blue and red respectively.

4.4 Discussion

Based on the literature for square planar- Pt(II) complexes, the nucleophilic addition of the L1s was also considered. Attempts to optimize the hypothesized 5-coordinated intermediates structures for this pathway were not fruitful. These calculations either resulted in the reactants or products regardless of the basis set employed. Additionally, NEB calculations did not show these structures to exist on the potential energy surface connecting the precatalyst and product structures for any of the systems studied. From the calculated MEPs, the mechanism of ligand substitution was determined to progress through a concerted (single step) mechanism. For Full

and Model systems, the transition state for cis and trans structures involved trigonal bipyramidal geometries.

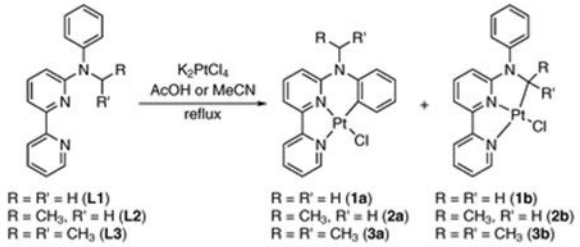
Comparison of the cis/trans pathways for the full system agreed with experimental results from Barnham et al's study. The energy barrier for the cis pathway was 16.14 kcal/mol compared to the 20.13 kcal/mol barrier of the trans pathway. Comparing product energies shows that the trans products are significantly more stable thus the thermodynamic products. Barnham et al found the cis-$Pt(PhCN)_2Cl_2$ was much more reactive than the trans isomer, thus the cis-ligand exchange occurred at a much faster rate. The lower energy barrier calculated in this study is congruent with their results. Additionally, we confirmed that the trans products are much more stable; thus, the formation of trans products may proceed under thermodynamic control.

Comparing the pathways using benzonitrile and pyridine as model substrates suggest that the relative reactivity of cis/trans starting materials is ligand dependent. With ligand substitution of pyridine, the trans pathway had lower energy structures for both the transition state and products when compared to the cis pathway. However, the ligand substitution process of the larger/more bulky 2-(2-methylphenoxy)pyridine resulted in the cis- pathway being kinetically favored.

CHAPTER 5: THEORETICAL PROBE INTO THE SOLVENT-SELECTIVE NATURE OF C-H ACTIVATION IN CYCLOPLATINATION REACTIONS

5.1 Background

Previous works of Dr. Huo led to an observed solvent-controlled selectivity of sp^2 and sp^3 C-H bond activation by K_2PtCl_4 complexes as shown in **Table 1**.[83] Ligand L1 has two possible locations for cyclometalation to occur: 1) the methyl group (1b) or 2) the N-phenyl group(1a). Generally, sp^2 C-H bonds are more reactive which would suggest cyclometalation to occur at one of the ortho locations of the phenyl ring. Instead, refluxing in acetic acid resulted in 96% sp^3 activation at the methyl group. This experiment was repeated using acetonitrile as a solvent and exclusively sp^2 C-H was activated. Kinetic isotope studies were performed to elucidate the selectivity in sp^2/sp^3 C-H activation. When refluxing 1a in AcOD, H/D exchange was observed between ortho positions on the N-phenyl ring as well as the methylene group suggesting there was some equilibrium between the two products. However, when AcOD reflux was performed on 1b, no H/D exchange occurred for the ortho-phenyl positions. These results suggest that the sp2 activation barrier is significantly smaller than sp^3 activation. Previous DFT calculations were performed on both products of the reaction and show that the sp^3 product is more stable suggesting that it is the thermodynamic product.



	In AcOH			In MeCN		
Product	*t*/h	Ratio [a]	Yield [b]	*t*/h	Ratio [a]	Yield [b]
1a∶**1b**	24	4∶96	73%	72	100∶0	73%
2a∶**2b**	48	3∶97	63%	48	93∶7	70%
3a∶**3b**	48	0∶100	38%	72	70∶30	36%

a Isomeric ratio determined by proton NMR spectra of the crude products. b Isolated yield of the pure major
.

Figure 5.1 Cyclometalation reactions of different ligands with K_2PtCl_4 with different solvents.

Another unpublished study in the Huo lab investigated the competitive sp^2/sp^3 C-H activation for the N-methyl-N-phenyl-6-(1H-pyrazol-1-yl)pyridine-2-amine ligand. Interestingly, the reaction did not occur in acetonitrile. Instead, reflux in acetic acid resulted in both sp^2/sp^3 products. Refluxing with 3N HCl resulted in sp^2 C-H product formation exclusively while using a solvent mixture of AcOH and HCl yielded the sp^3 product with ~80% selectivity. The proposed mechanism of this reaction carried out in acetic acid is shown in **Figure 4**. In this study, a theoretical approach using the Nudged Elastic Band method was employed to investigate the mechanism of the solvent-controlled switch in C-H activation observed in the experimental results of Dr. Huo's studies.

5.2 Methodology

All calculations were performed using Gaussian 16 software on a computer cluster of 640 processors located at East Carolina University.[84] Prior to computational study, a functional comparison was performed. M06, M06-2x, MN15, and M06-L were compared using water,

acetic acid, and gas phase as the solvent systems.[78] M06-L with water PCM was chosen for this computational study. Starting structures were assembled and visualized using GaussView09, Avogadro, and Pymol.[71,85,86] Geometry optimizations, reaction energetics, and frequency calculations were calculated using the def2-TZVP basis set including both *f* and g polarization functions for Pt atoms.[79] The modest 6-31G** basis was used for the remaining atoms. Implicit solvent effects using the polarizable continuum model within the self-consistent reaction field method for all calculations.[80] As previously mentioned, the PCM model used predominantly in this study was water.[81]

Geometry optimizations for all structures were confirmed through vibrational frequency calculations. Frequency calculations were performed at the same level of theory as its preceding optimization calculation. These calculations provided normal mode information which was used to determine the viability of transition states and provide zero-point energy corrections. Stable minima were confirmed via positive definite hessians and stable transition states were confirmed with the presence of a single negative frequency.

The NEB method is a chain-of-states method connecting optimized reactants and products by a potential energy surface. Successive minimization along the NEB path in theory yields the minimum energy path from which intermediate and transition state structures can be identified. In this work, NEB calculations used a string of 20 or 40 images depending on reaction step complexity. NEB calculations were performed using the FIRE optimizer to rapidly converge. In order to better approximate transition state structures, the climbing image feature in the NEB program was utilized. All NEB calculations used the M06-L functional along with the def2-TZVP (Platinum) and 6-31G** (all other atoms) as well as the water PCM model.

Several different methods were used to determine the transition states from NEB calculations. Henkelman's DIMER method was utilized to directly calculate saddle points from adjacent NEB images. DIMER calculations were performed at identical levels of theory as geometry optimizations. Additionally, transition state methods available within G16 were used (TS, QST2). All calculated TS were confirmed by frequency calculations as described previously.

5.3 Results

5.3.1 Proposed Mechanism

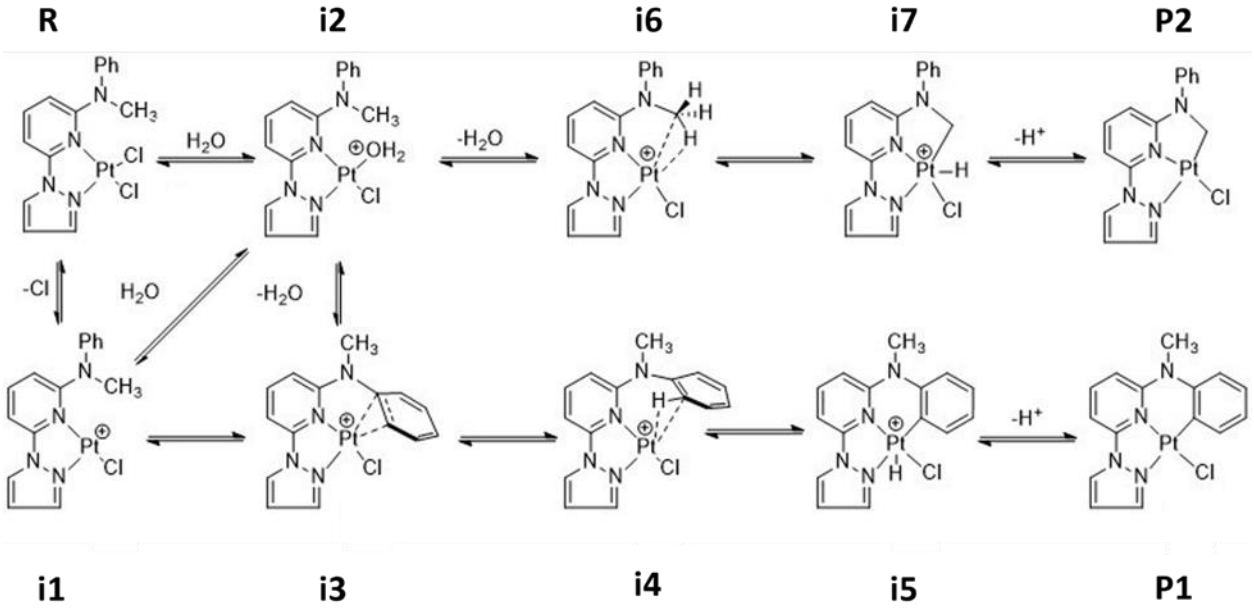


Figure 5.2 Proposed potential reaction pathways for the sp^2 and sp^3 C-H activation of the N-methyl-N-phenyl-6-(1H-pyrazol-1-yl)pyridine-2-amine ligand.

5.3.2 Geometry Optimization

Prior to NEB calculations, the structures from the proposed mechanism were optimized using maug-cc-pvtz/def2-tzvp basis sets to determine their validities. The structures for the

proposed structures are shown in **Figure 5.2**. All the intermediates from the proposed mechanism were successfully optimized, with the exceptions of **i1** and **i4,** and confirmed via positive definite frequency calculations. Selected bond distances and angles are shown in **Table 5.1**. For NEB analysis reactant and product molecules are required to have identical numbers of atoms. Thus, reactants and product complexes were also optimized including water, hydrochloric acid, and acetic acid molecules as well.

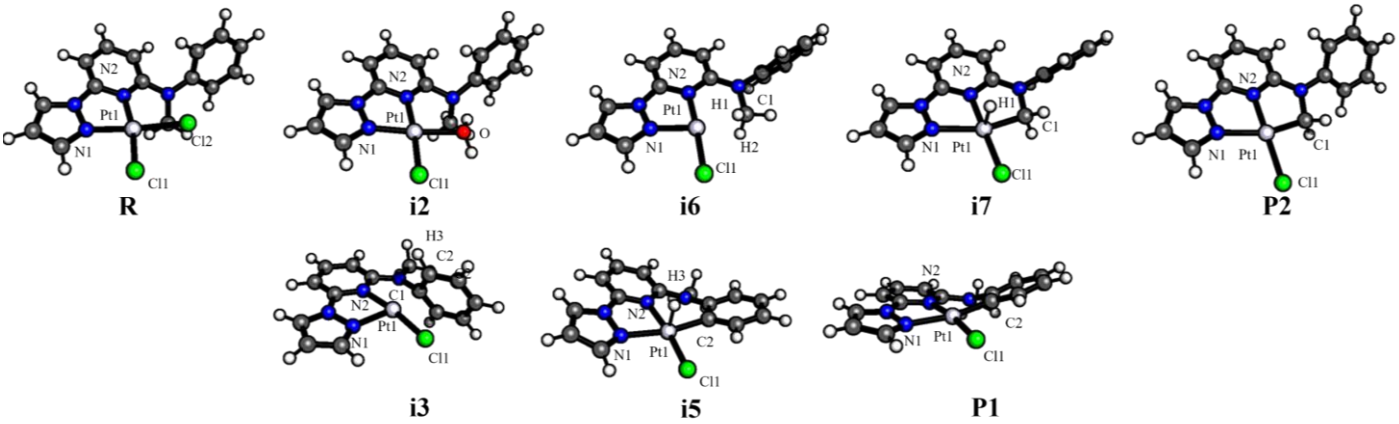


Figure 5.3 Optimized structures for the proposed pathway (Figure 5.1). Atoms bonded to platinum/ in the coordination sphere are labeled.

Table 5.1 Bond lengths (Å) and bond angles (degrees) of optimized structures

Structure	R	i2	i3	i5	i6	i7	P1	P2
Pt-Cl1	2.333	2.319	2.346	2.341	2.341	2.345	2.363	2.367
Pt-Cl2	2.350							
Pt-N1	1.999	1.961	2.039	2.160	1.977	2.232	2.127	2.187
Pt-N2	2.098	2.073	1.992	2.048	1.987	1.986	2.025	1.958
Pt-C1					2.399	2.034		2.016
Pt-C2			2.235					
Pt-C3			2.315	2.023			1.990	
Pt-H1					1.861	1.511		
Pt-H2					2.457			
Pt-O1		2.113						
Pt-H3			2.712	1.511				
Cl1-Pt-Cl2	88.690							
Cl1-Pt-N1	92.275	95.609	93.003	92.640	96.576	103.556	92.685	102.304
Cl1-P1-N2	170.397	174.692	171.498	170.965	176.749	179.172	170.996	179.327
Cl1-Pt-C1						97.381		97.268
Cl1-Pt-C2			109.944					
Cl1-Pt-C3			93.420	95.836			95.752	
Cl1-Pt-H1					91.567	89.408		
Cl1-Pt-H2					99.861			
Cl1-Pt-H3			105.503	88.936				
Cl1-Pt-O1		85.654						

5.3.3 Critical Field Analysis

Initial NEB studies incorporating the coordination of water to platinum were not fruitful. These initial studies either yielded potential energy curves exceeding 60 kcal/mol or resulted in unfeasible image paths. Critical Field Analysis (CFA) was performed to validate a solvent independent mechanism. As previously mentioned, one of the powers of DFT is its ability to paint an intuitive picture of a quantum system using electron density. One example used in this work is critical field analysis. In CFA, molecules are partitioned into subsystems which are surrounded by "zero-flux" regions. The implication of these partitions is that electron density is highest at nuclei and dissipates exponentially with respect to distance. Critical field analysis evaluates the points where the density gradients vanish. While the process of performing CFA is mathematically abstract, the critical points obtained are chemically relevant. There are 4 possible critical points in 3 dimensional space, but the critical points which are of concern in this work are the (3,-3), (3,-1),and (3,+1) critical points. (3,-3) critical points are points at which electron density is at a maximum in all 3 spatial dimensions and are known as nuclear attractors. (3,-1) critical points are indicative of bonds; they are maximums in two dimensions and minima in the third. Likewise, ring critical points, (3,+1), are points where electron density is at a minimum in 2 dimensions and a maximum in the third.[87,88]

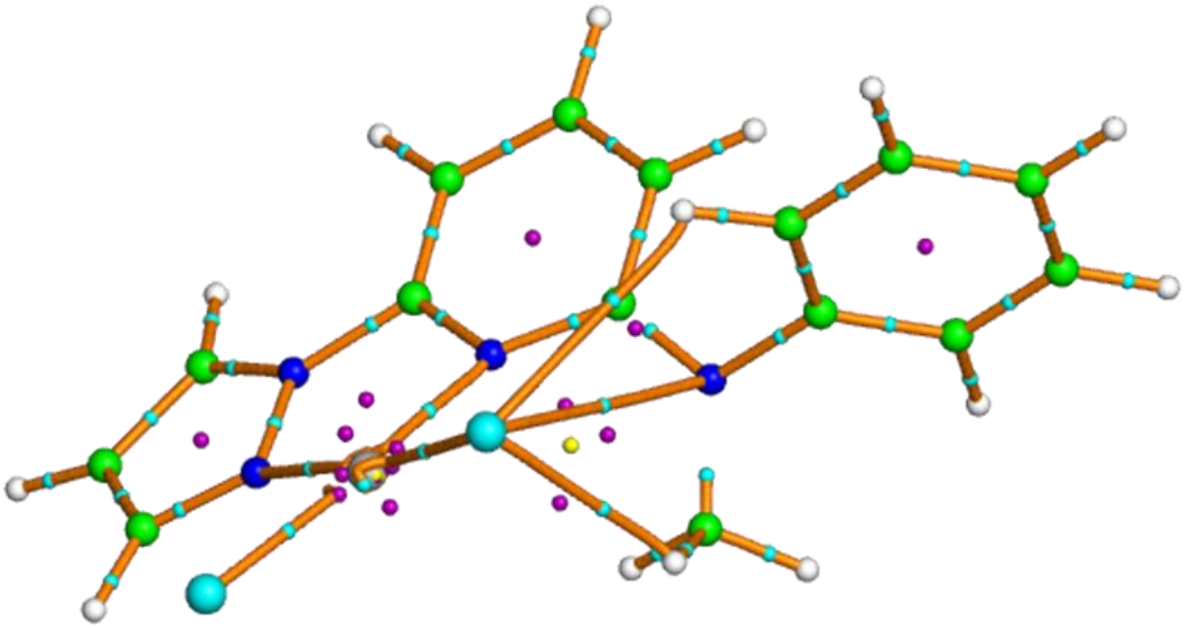


Figure 5.4 Critical Field analysis of structure R. Bonds are rendered in orange with (3,-1) critical points shown in cyan.

CFA was employed on the structure **R** to assess its reactivity. (3,-1) bond critical points were identified to connect the departing chlorine to both sp^2 and sp^3 hydrogens which are involved in C-H activation. This suggests that the intramolecular H-bonds between the hydrogens and the leaving chlorine may be responsible for a reaction which is independent of solvation. Thus, in addition to solvent-catalyzed pathways, we also studied the pathway using no explicit solvent molecules.

5.3.4 NEB Pathways

NEB calculations were performed to find the minimum energy pathways along with their corresponding saddle points connecting the optimized reactant and product molecules. This was done with the hope to elucidate the mechanism of the solvent-controlled C-H activation reaction observed by Barnham et al. Due to the varying reaction conditions/solvents, NEB calculations

were performed on sp^3/sp^2 pathways for several different systems. Our general method of mapping reaction pathways started with NEBs of 20 or 40 images connecting reactants and products. These reactants-to-product NEBs were then broken into smaller NEB depending on the complexity and intermediates present.

5.1.1.1. No Solvent

The first NEB calculations were performed to study the reaction mechanism involving no solvent catalysis in which one of the coordinated chlorines reacted with the sp^2/sp^3 hydrogen to form **P1/P2** and a hydrochloric acid molecule. The potential energy curve generated for the C-H sp^3 pathway with no explicit solvent is shown in **Figure 5.5**.

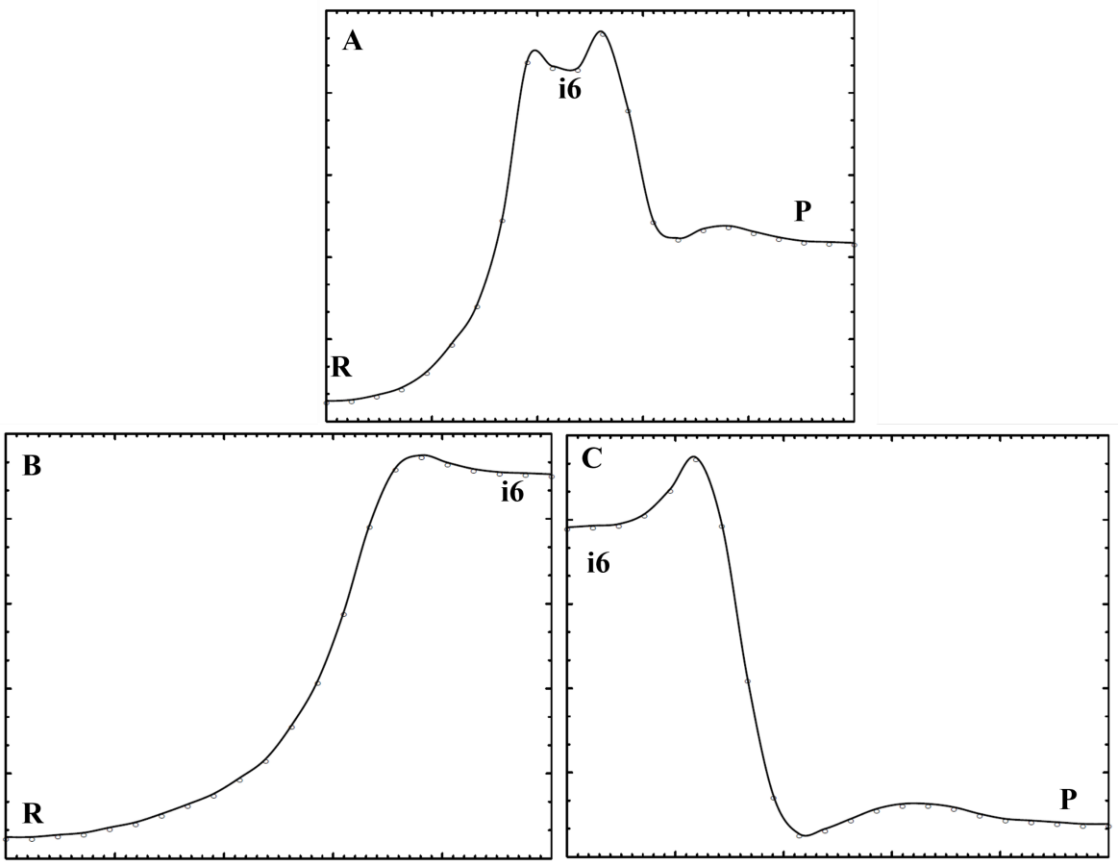


Figure 5.5 NEB simulated MEP for the sp^3 no solvent pathway from **R→P2 (A), R→i6 (B)**, and **i6→P2 (C).**

The NEB from **R->P2 (Figure 5.5 A)** produced a MEP showing the presence of a single intermediate which optimized to give the structure for **i6**. Structural information for all structures identified from **Figure 5.5** is shown in **Figure 5.6** and **Table 5.2** and their energetics in **Table 5.3**. Formation of **i6** occurred following the complete dissociation of the Pt-Cl2 bond. The intermediate **i6** is a highly reactive intermediate (25.63 kcal/mol higher than **R**) and exists in a very shallow energy well with the transition state (**TS 6|P2)** leading to product formation only having an energy barrier of 0.441 kcal/mol. This is reflected in the geometry of transition state **TS R|6** and **6|P2** which closely resemble the geometry of **i6**. Thus, it was determined that **TS R|6**

leading to the formation of the complex contributes to nearly all the energy barrier 26.784 kcal/mol. From **TS R|6** to **6|P2**, the bond distance between both the activated sp^3 hydrogen (H1) and its carbon (C1) gradually decreases. While in **TS 6|P2** the Pt-H1 bond distance reaches 1.674 (indicative of a η^2-complex), the oxidative addition of H to Pt to form a 5 coordinated square pyramidal complex does not occur to form **i7** as hypothesized. Instead, H1 is abstracted by the spectator Cl ion to directly form **P2**.

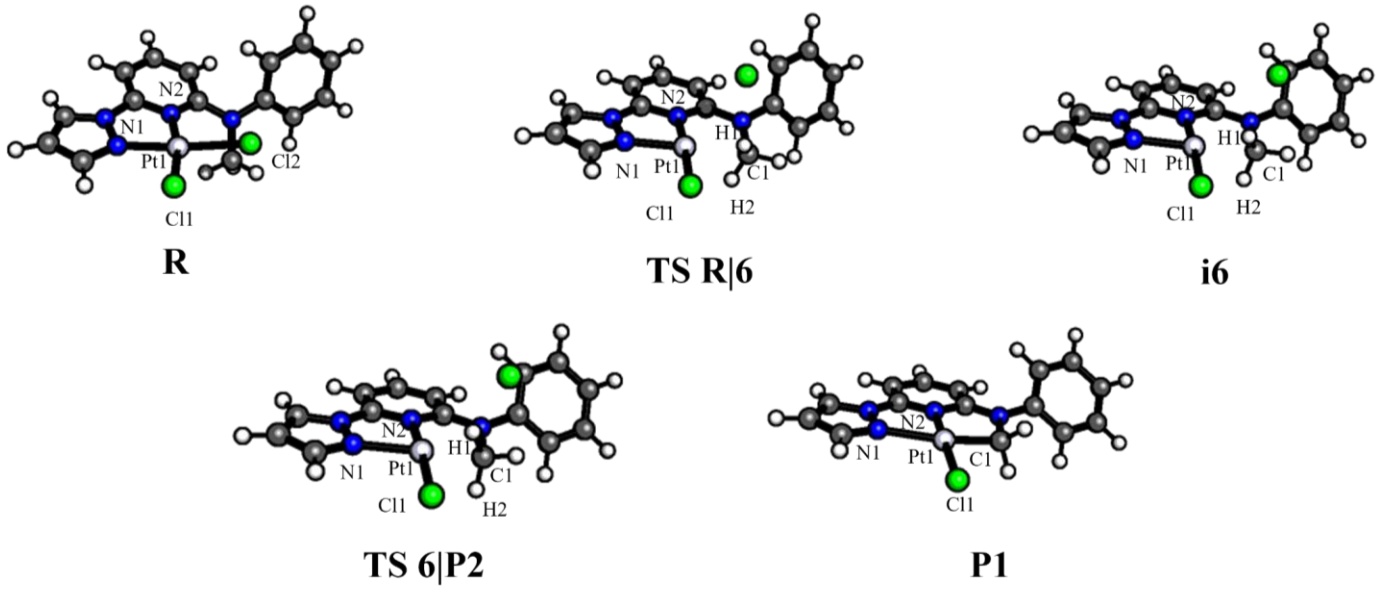


Figure 5.6 Optimized structures for the no solvent sp^3 NEB pathway (**Figure 5.4**). Atoms bonded to platinum/ in the coordination sphere are labeled.

	RT	**A**	**Pt**	**RC**	**B**	**PC**
Pt-Cl1	2.364	2.356	2.364	2.331	2.336	2.335
Pt-Cl2	2.358	2.355	2.363	2.331	2.319	2.353
Pt-N1	1.988	1.953	2.009	2.016	1.994	2.016
Pt-N2	1.996	2.408		2.021	2.469	
Pt-N3		2.535	2.039		2.620	2.069
Cl1-Pt15-Cl2	176.298	179.071	178.913	91.655	91.898	92.037
Cl1-Pt15-N1	89.700	89.432	90.491	179.431	178.110	178.204
Cl1-Pt15-N2	90.478	87.959		88.733	87.134	
Cl1-Pt15-N3		93.716	89.255		88.639	88.572

Table 5.2 Bond Lengths (Å) and angles (degrees) for optimized structures

Structure	**R**	**TS r|6**	**i6**	**TS 6|P2**	**P2**
Pt-Cl1	2.344	2.349	2.353	2.363	2.396
Pt-Cl2	2.361				
Pt-N1	2.004	1.979	1.998	2.087	2.197
Pt-N2	2.112	2.014	1.992	1.984	1.961
Pt-C1		2.591	2.361	2.147	2.017
Pt-H1		2.002	1.847	1.674	
Pt-H2		2.621	2.451	2.442	

Table 5.3 Energetics (HF energies, zero-point energy correction factors, zero-point corrected energies) for the structures of the sp^3 no solvent pathway. Energies units are in Hartree/Particle.

	HF	ZPC	HF-ZPC
R	-1838.869	0.271	-1838.599
R|6	-1838.825	0.270	-1838.556
i6	-1838.827	0.290	-1838.558
6|P2	-1838.822	0.285	-1838.557
P2	-1838.849	0.266	-1838.583

The MEP for the intramolecular sp2 mechanism (**Figure 5.7**) also showed the presence of a single intermediate (**i4**) following optimization of images. A very broad MEP was determined to connect structures **R** and **i3**. The steep initial increase in potential energy was determined to be due to the dissociation of Cl2, similar to the sp^3 pathway. The **TS R|3** was determined to have a reaction barrier of 25.33 kcal/mol with a frequency of -35.5 which is in the noise range. The structure of **TS R|3** was the proposed intermediate **i1**. The gradual decrease in potential energy following **TS R|3** was attributed to the rotation of the phenyl group to form **i3**. **i3** is an η^2-structure of Pt with the Pt-C2 Pt-C3 bond distances of 2.671Å and 2.245Å respectively. Compared to the sigma complex of the sp^3 pathway (**i6**), **i3** appears to be a relatively stable intermediate (14.07 kcal/mol above **R**). The pathway from **i3** to **P1** was connected via a sharp increase in potential energy. This was attributed to the simultaneous association of the phenyl ring to Pt (formation of Pt-C3) and abstraction of H3 by Cl2. Interestingly, optimization of **TS 3|P1** yielded the proposed η^2-structure for **i4**. **TS 3|P1** had a very strong negative frequency (-

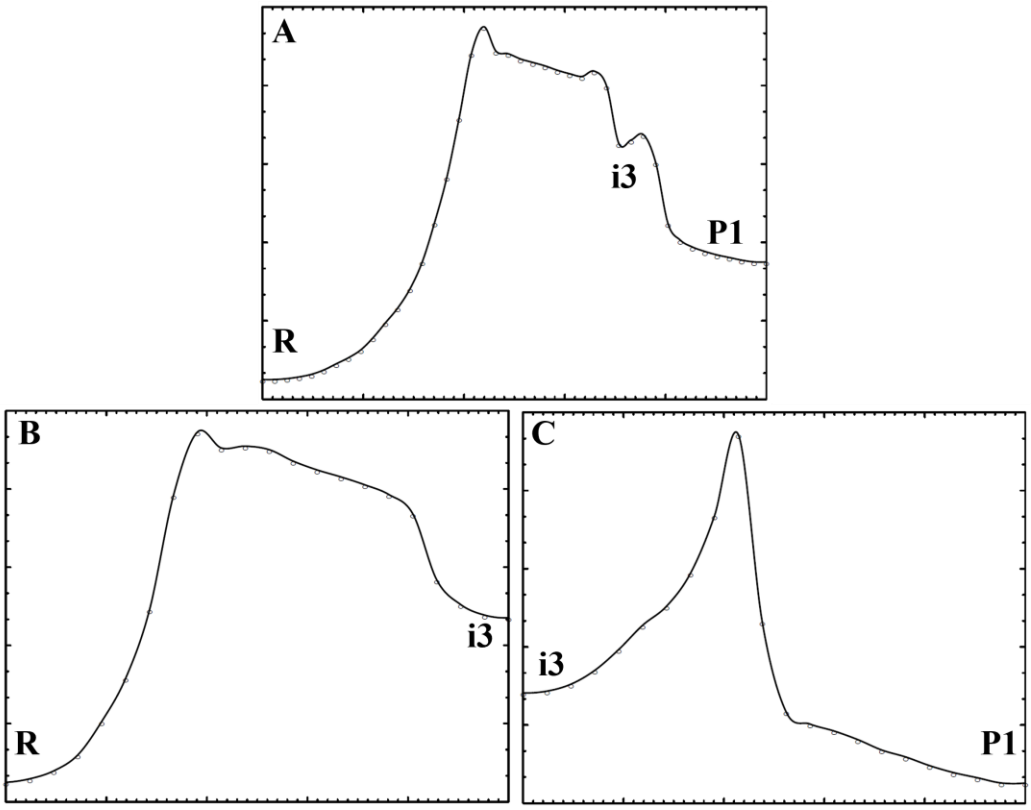


Figure 5.7 NEB simulated MEP for the sp^2 no solvent pathway from **R→P1 (A)**, **R→i3 (B)**, and **i3→P1 (C)**

Table 5.4 Energetics (HF energies, zero-point energy correction factors, zero-point corrected energies) for the structures of the sp^2 no solvent pathway. Energies units are in Hartree/Particle.

	HF	ZPC	HF-ZPC
R	-1838.869	0.271	-1838.598
R|3	-1838.828	0.270	-1838.558
i3	-1838.846	0.270	-1838.576
3|P1	-18.284	0.264	-1838.572
P1	-1838.853	0.266	-1838.586

1031.7079) and its barrier was 2.75 kcal/mol. Similar to the sp^3 no solvent pathway, the oxidative addition of H3 to form **i5** did not lie on the MEP generated through NEB.

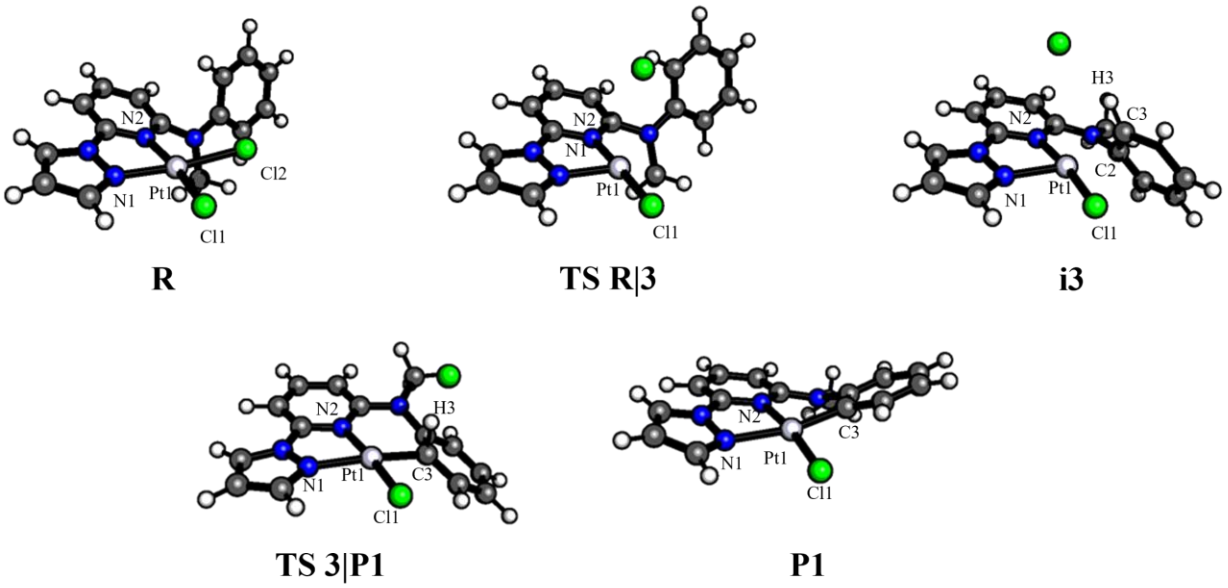


Figure 5.8 Optimized structures for the no solvent sp^2 NEB pathway (**Figure 5.6**). Atoms bonded to platinum/ in the coordination sphere are labeled.

	RT	**A**	**Pt**	**RC**	**B**	**PC**
Pt-Cl1	2.364	2.356	2.364	2.331	2.336	2.335
Pt-Cl2	2.358	2.355	2.363	2.331	2.319	2.353
Pt-N1	1.988	1.953	2.009	2.016	1.994	2.016
Pt-N2	1.996	2.408		2.021	2.469	
Pt-N3		2.535	2.039		2.620	2.069
Cl1-Pt15-Cl2	176.298	179.071	178.913	91.655	91.898	92.037
Cl1-Pt15-N1	89.700	89.432	90.491	179.431	178.110	178.204
Cl1-Pt15-N2	90.478	87.959		88.733	87.134	
Cl1-Pt15-N3		93.716	89.255		88.639	88.572

Table 5.5 Bond Lengths (Å) and angles (degrees) for optimized structures.

Structure	**R**	**TS R|3**	**i3**	**TS 3|P1**	**P1**
Pt-Cl1	2.343400435	2.333346547	2.356342614	2.364241364	
Pt-Cl2	2.358895143				
Pt-N1	2.004180055	1.937411807	2.022850694	2.06132895	
Pt-N2	2.102955734	2.051260641	2.029462097	2.042348502	
Pt-C2			2.671449717		
Pt-C3			2.244828881	2.062638042	
Pt-H3			2.545117906	2.203855317	
Cl1-Pt-Cl2	88.59649729				
Cl1-Pt-N1	92.48351606	98.15564601	92.87389239	92.62576697	93.39667925

5.1.1.2. Water Catalyzed

Experimental results used hydrochloric acid as a solvent for synthesis of the cycloplatinum complexes. Thus, NEB calculations were also performed using an explicit water molecule to catalyze the removal of the sp^2/sp^3 hydrogens. The potential energy curve for the water catalyzed C-H sp^3 pathway is shown in **Figure 5.9** and the corresponding optimized structures and structural features in **Figure 5.10** and **Table 5.6**.

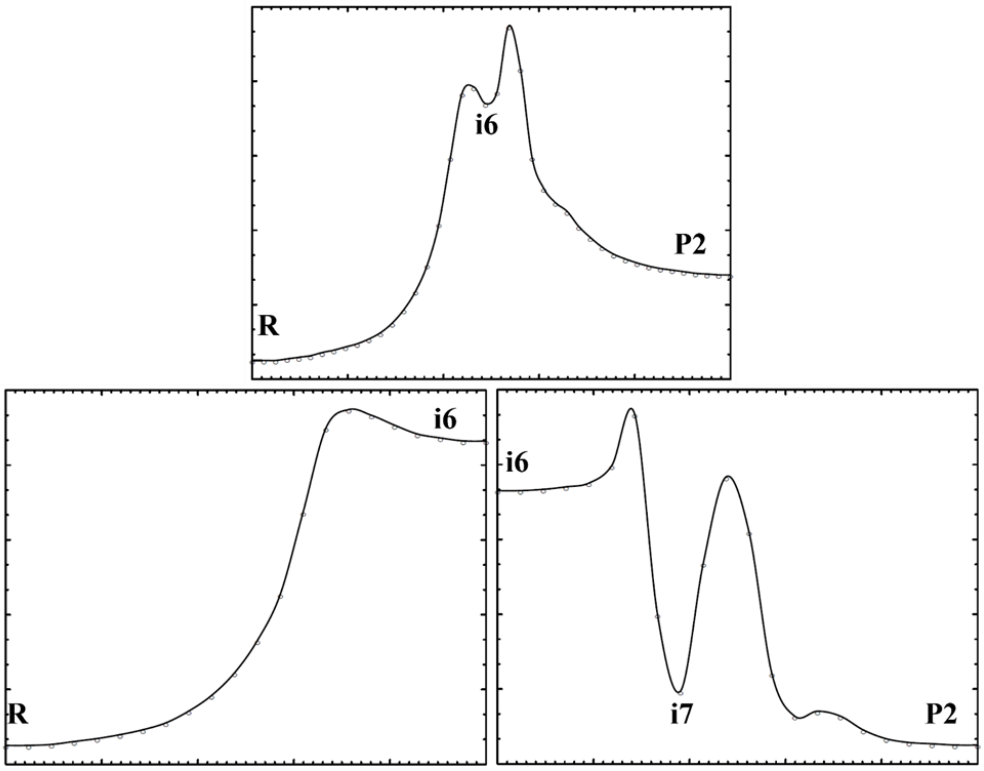


Figure 5.9 NEB simulated MEP for the sp^3 water catalyzed pathway from **R→P1 (A)**, **R→i3 (B)**, and **i3→P1 (C)**

The initial NEB from **R** to **P2** (**Figure 5.9A**) showed the presence of a single intermediate (**i6**) like the previous NEB curves. Formation of **i6** occurred through dissociation of Cl2 and had a reaction barrier of 25.313 kcal/mol. Interestingly, the NEB from **i6→ P2** revealed an additional intermediate; optimization yielded the structure of **i7**. **i7** formed from the oxidative addition of H1 to Pt to form a stable square pyramidal intermediate. The transition state **6|7** had a reaction barrier of 1.39 kcal/mol. The frequency of **TS 6|7** was very strong (-662.454). **TS 7|P2** closely resembled **TS 6|7** and had a barrier of 11.39 kcal/mol from **i7.**

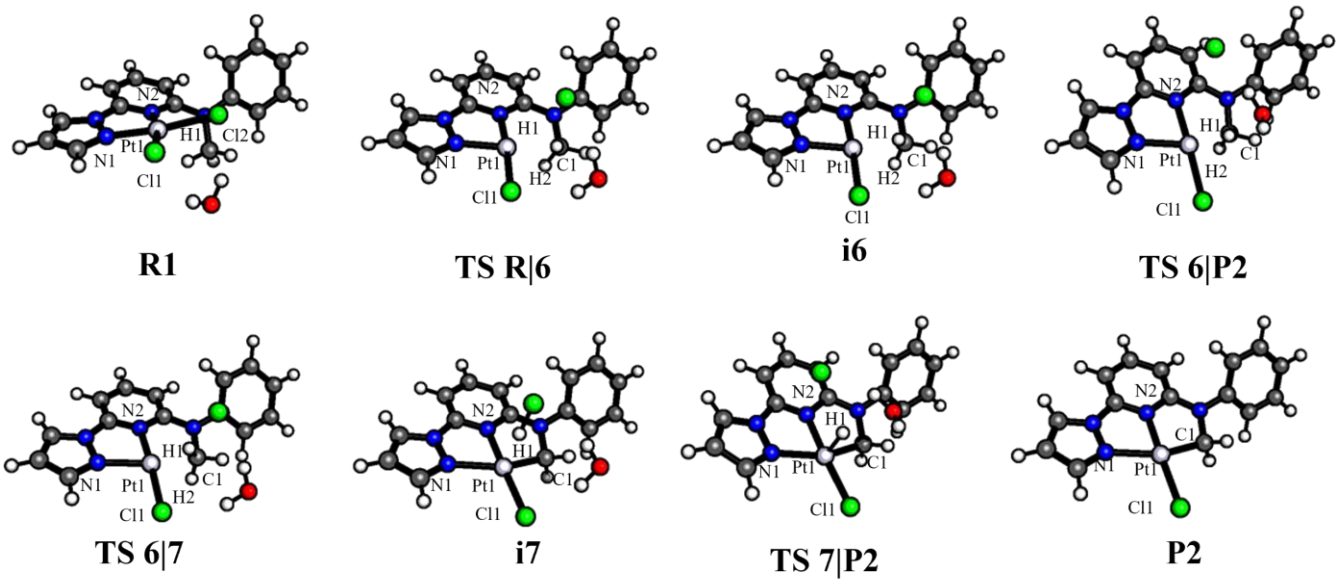


Figure 5.10 Optimized structures for the no solvent sp^3 NEB pathway (**Figure 5.8**). Atoms bonded to platinum/ in the coordination sphere are labeled.

	RT	**A**	**Pt**	**RC**	**B**	**PC**
Pt-Cl1	2.364	2.356	2.364	2.331	2.336	2.335
Pt-Cl2	2.358	2.355	2.363	2.331	2.319	2.353
Pt-N1	1.988	1.953	2.009	2.016	1.994	2.016
Pt-N2	1.996	2.408		2.021	2.469	
Pt-N3		2.535	2.039		2.620	2.069
Cl1-Pt15-Cl2	176.298	179.071	178.913	91.655	91.898	92.037
Cl1-Pt15-N1	89.700	89.432	90.491	179.431	178.110	178.204
Cl1-Pt15-N2	90.478	87.959		88.733	87.134	
Cl1-Pt15-N3		93.716	89.255		88.639	88.572

Table 5.6 Bond Lengths (Å) and angles (degrees) for optimized structures.

Structure	**R**	**TS R|6**	**i6**	**TS 6|7**	**TS 6|P2**	**i7**	**TS 7|P2**	**P2**
Pt-Cl1	2.350	2.355	2.359	2.368	2.360	2.376	2.357	2.390
Pt-Cl2	2.371							
Pt-N1	2.002	1.977	1.998	2.097	2.105	2.227	2.242	2.197
Pt-N2	2.101	2.019	1.991	1.986	1.985	1.979	1.986	1.962
Pt-C1		2.632	2.363	2.142	2.137	2.024	2.032	2.017
Pt-H1		2.015	1.835	1.651	1.643	1.841	1.534	
Pt-H2		2.690	2.456	2.429	2.449			
Cl1-Pt-Cl2	88.328							

Table 5.7 Energetics (HF energies, zero-point energy correction factors, zero-point corrected energies) for the structures of the sp^3 water catalyzed pathway. Energies units are in Hartree/Particle.

	HF	ZPC	HF-ZPC
R	-1915.298	0.295	-1915.003
R|6	-1915.256	0.294	-1914.962
i6	-1915.260	0.294	-1914.966
6|7	-1915.253	0.290	-1914.964
i7	-1915.280	0.289	-1914.991
7|P2	-1915.264	0.291	-1914.973
P2	-1915.284	0.291	-1914.993

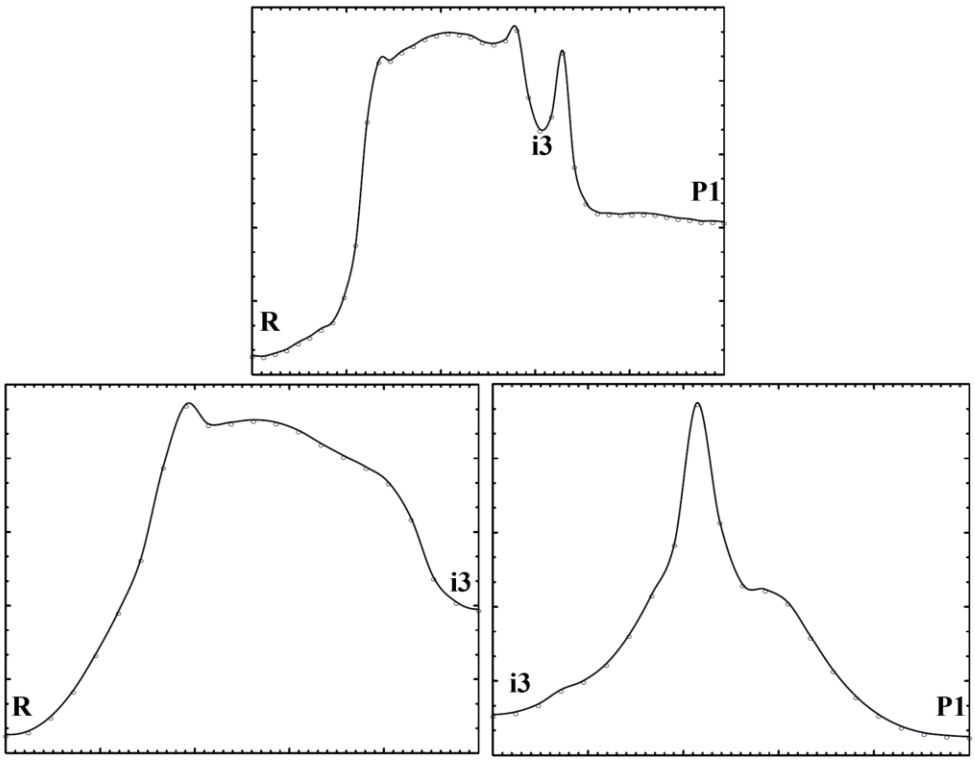


Figure 5.11 NEB simulated MEP for the sp^2 water catalyzed pathway from **R→P1 (A)**, **R→i3 (B)**, and **i3→P1 (C)**

.

The sp^2 water-catalyzed pathway was very similar when compared to the intramolecular pathways. A single intermediate (**i3**) was on the pathway connecting reactants and products. **TS R|3** and **3|P1** were optimized to resembe **i1** and **i4** from the proposed mechanism respectively. However, when comparing energetics, water was found to greatly reduce the energy of activation. The reaction step from **R→i3** only had a energy barrier of 22.12 kcal/mol, ~3 kcal/mol lower than the intramolecular reaction mechanism. The energy barrier for **i3→P1** was 6 kcal/mol which is higher than the corresponding intramolecular reaction step. This could be attributed to a much lower energy **i3** structure in the water catalyzed system. Although the barrier from **i3→P1** was larger, it was still lower than the RDS of **R→ i3**.

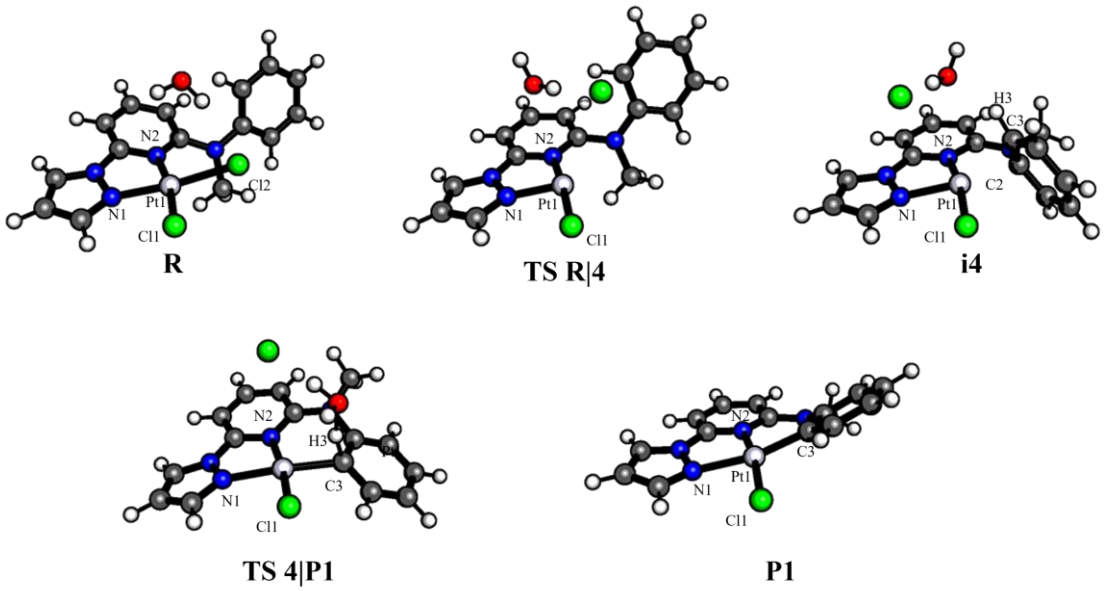


Figure 5.12 Optimized structures for the no solvent sp^2 NEB pathway (**Figure 5.8**). Atoms bonded to platinum/ in the coordination sphere are labeled.

	RT	A	Pt	RC	B	PC
Pt-Cl1	2.364	2.356	2.364	2.331	2.336	2.335
Pt-Cl2	2.358	2.355	2.363	2.331	2.319	2.353
Pt-N1	1.988	1.953	2.009	2.016	1.994	2.016
Pt-N2	1.996	2.408		2.021	2.469	
Pt-N3		2.535	2.039		2.620	2.069
Cl1-Pt15-Cl2	176.298	179.071	178.913	91.655	91.898	92.037
Cl1-Pt15-N1	89.700	89.432	90.491	179.431	178.110	178.204
Cl1-Pt15-N2	90.478	87.959		88.733	87.134	
Cl1-Pt15-N3		93.716	89.255		88.639	88.572

Table 5.8 Bond Lengths (Å) and angles (degrees) for optimized structures.

Structure	R	TS R|5	i4	TS 4|P1	P1
Pt-Cl1	2.342	2.345	2.357	2.365	2.366
Pt-Cl2	2.368				
Pt-N1	2.003	1.929	2.054	2.063	2.143
Pt-N2	2.109	2.059	1.992	2.040	2.035
Pt-C2			2.234		
Pt-C3			2.314	2.065	1.993

	HF	ZPC	HF-ZPC
R	-1915.299	0.295	-1915.003
R|3	-1915.263	0.295	-1914.968
i3	-1915.280	0.294	-1914.986
3|P2	-1915.266	0.290	-1914.976
P2	-1915.282	0.292	-1914.990

Table 5.9 Energetics (HF energies, zero-point energy correction factors, zero-point corrected energies) for the structures of the sp^2 water catalyzed pathway. Energies units are in Hartree/Particle.

5.4 Energy Diagrams

Energy diagrams (**Figures 5.13** and **5.14**) compared the MEP of sp^2/sp^3 C-H activation for the cycloplatination reaction using different solvent systems. It is important to note that the sp^3 products (**P2**) are the more stable product; the energetics of the **P1/P2** used in the NEB calculations are complexes including solvent molecules, leaving groups, and spectator ions which are used to generate the best pathways. Thus, they are local minima and are not necessarily the most stable arrangement of molecules. sp^2 pathways using no explicit solvent and water catalysis (blue in **Figures 5.13** and **5.14** respectively) both form a single intermediate (**i3**) through chlorine dissociation. The transition state connecting **i3** and **P1** was optimized to yield the η^2-structure for **i4** from the proposed mechanism. The mechanism of the sp^3 pathways (red in **Figures 5.13** and **5.14**) exhibited more mechanistic variability. With no explicit solvent, dissociation of the Cl ion resulted in a single reactive intermediate **i6**. However, a catalytic water allowed for the oxidative addition of H1 from **i6** to form the stable square pyramidal intermediate **i7** prior to **P2**.

As predicted in experimental studies, the sp^2 pathways exhibited smaller energy barriers with **P1** being the kinetic product. These smaller barriers were most ready seen in the formation of their sigma complex intermediates (**i3**) when compared to the respective sp^3 pathways. Using no explicit solvent, the energy difference between **R→i6** and **R→i3** was a modest 1.32 kcal/mol. However, with a catalytic water molecule, the barrier **R→i3** was significantly lower than that of **R→i6 (**3.65 kcal/mol**).**

5.5 Conclusions

In this work, a theoretical approach was applied to investigate the selective sp^2/sp^3 C-H activation of cyclometalated platinum complexes. NEB simulations were performed on potential reaction pathways using no-explicit solvent and a catalytic water molecule were performed. The results suggest that sp^2 C-H activation in our system proceeds through two elementary steps: 1) chlorine dissociation to form a stable trigonal bipyramidal η^2-benzylplatinum intermediate, 2) dehydrogenation of sp^2-H. Catalytic water significantly reduces the energy barrier for the formation of the benzylplatinum intermediate. The mechanism for the sp^3 C-H activation is not as straightforward. Using no solvent catalysis, the sp^3 pathway consisted on two elementary steps: 1) chlorine dissociation to form a highly reactive platinum intermediate 2) direct dehydrogenation of the sp^3 hydrogen. On the other hand, using a catalytic water molecule stabilized an additional intermediate. The water-catalyzed sp^3 pathway consisted of 1) chlorine dissociation to form the highly reactive platinum intermediate 2) oxidative addition of sp^3 H to platinum to complete the 5-membered platinacycle to form a square pyramidal intermediate and 3) dehydrogenation of the platinum coordinated hydrogen for form the product. The sp^2 C-H activation pathways exhibited lower energy barriers but less stable products. In each of these

pathways, the key rate limiting step is the chlorine dissociation/formation of hydrocarbon-platinum intermediates.

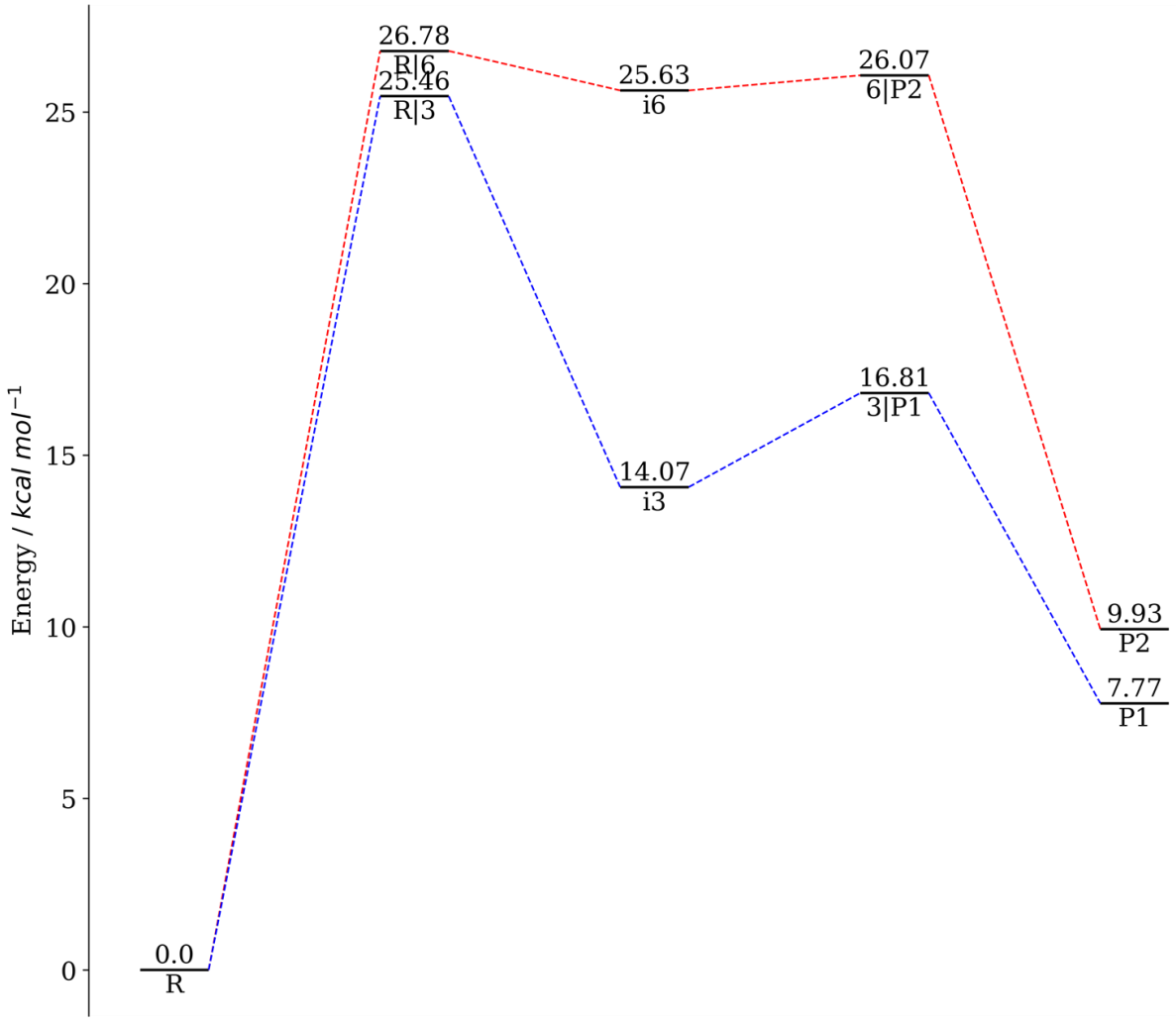


Figure 5.13 Energetics (HF energies, zero-point energy correction factors, zero-point corrected energies) for the structures of the sp^2 water catalyzed pathway. Energies units are in Hartree/Particle.

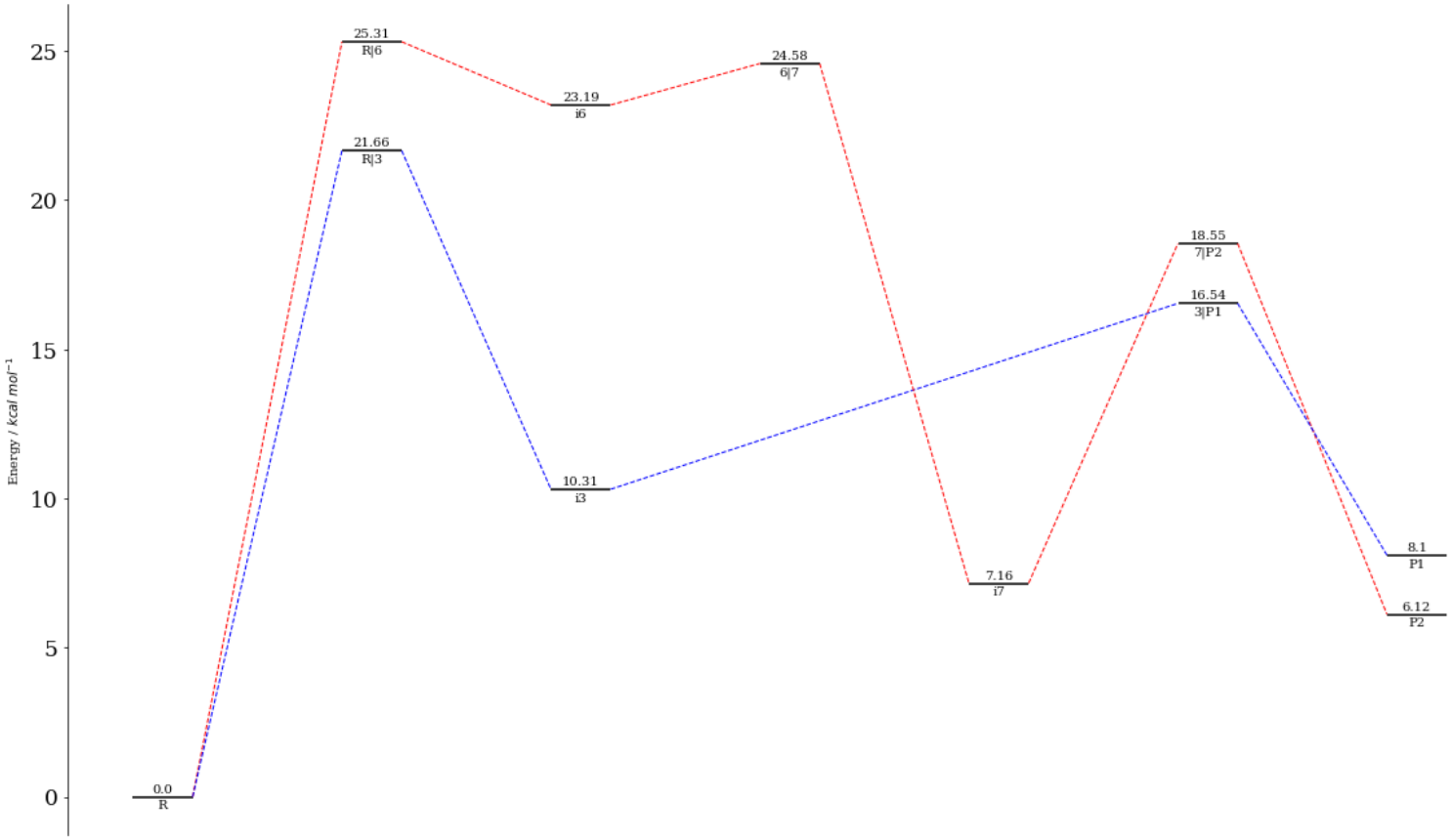


Figure 5.14 Energetics (HF energies, zero-point energy correction factors, zero-point corrected energies) for the structures of the sp^2 water catalyzed pathway. Energies units are in Hartree/Particle.

REFERENCES

(1) Han, P. F.; Che, X. Da; Li, H. Z.; Gao, Y. Y.; Wei, X. C.; Li, P. C. Annexin A1 Involved in the Regulation of Inflammation and Cell Signaling Pathways. *Chinese Journal of Traumatology - English Edition*. Elsevier B.V. April 1, 2020, pp 96–101. https://doi.org/10.1016/j.cjtee.2020.02.002.

(2) Mirsaeidi, M.; Gidfar, S.; Vu, A.; Schraufnagel, D. Annexins Family: Insights into Their Functions and Potential Role in Pathogenesis of Sarcoidosis. *J. Transl. Med.* **2016**, *14* (1), 1–9. https://doi.org/10.1186/s12967-016-0843-7.

(3) Fu, Z.; Zhang, S.; Wang, B.; Huang, W.; Zheng, L.; Cheng, A. Annexin A1: A Double-Edged Sword as Novel Cancer Biomarker. *Clinica Chimica Acta*. Elsevier B.V. May 1, 2020, pp 36–42. https://doi.org/10.1016/j.cca.2020.01.022.

(4) Rintala-Dempsey, A. C.; Rezvanpour, A.; Shaw, G. S. S100-Annexin Complexes-Structural Insights. https://doi.org/10.1111/j.1742-4658.2008.06654.x.

(5) Rosengarth, A.; Luecke, H. A Calcium-Driven Conformational Switch of the N-Terminal and Core Domains of Annexin A1. *J. Mol. Biol.* **2003**, *326* (5), 1317–1325. https://doi.org/10.1016/S0022-2836(03)00027-5.

(6) Seemann, J.; Weber, K.; Osborn, M.; Parton, R. G.; Gerke, V. The Association of Annexin I with Early Endosomes Is Regulated by Ca2+ and Requires an Intact N-Terminal Domain. *Mol. Biol. Cell* **1996**, *7* (9), 1359–1374. https://doi.org/10.1091/MBC.7.9.1359.

(7) D'acquisto, F.; Perretti, M.; Flower, R. J. Annexin-A1: A Pivotal Regulator of the Innate and Adaptive Immune Systems. https://doi.org/10.1038/bjp.2008.252.

(8) Araújo, T. G.; Mota, S. T. S.; Ferreira, H. S. V.; Ribeiro, M. A.; Goulart, L. R.; Vecchi, L.

Annexin A1 as a Regulator of Immune Response in Cancer. *Cells* **2021**, *10* (9). https://doi.org/10.3390/CELLS10092245.

(9) Rosengarth, A.; Luecke, H. A Calcium-Driven Conformational Switch of the N-Terminal and Core Domains of Annexin A1. *J. Mol. Biol.* **2003**, *326* (5), 1317–1325. https://doi.org/10.1016/S0022-2836(03)00027-5.

(10) Ries, M.; Watts, H.; Mota, B. C.; Lopez, M. Y.; Donat, C. K.; Baxan, N.; Pickering, J. A.; Chau, T. W.; Semmler, A.; Gurung, B.; Aleksynas, R.; Abelleira-Hervas, L.; Iqbal, S. J.; Romero-Molina, C.; Hernandez-Mir, G.; D'amati, A.; Reutelingsperger, C.; Goldfinger, M. H.; Gentleman, S. M.; Van Leuven, F.; Solito, E.; Sastre, M. Annexin A1 Restores Cerebrovascular Integrity Concomitant with Reduced Amyloid-β and Tau Pathology. *Brain* **2021**, *144* (5), 1526–1541. https://doi.org/10.1093/brain/awab050.

(11) De Marchi, T.; Timmermans, A. M.; Smid, M.; Look, M. P.; Stingl, C.; Opdam, M.; Linn, S. C.; Sweep, F. C. G. J.; Span, P. N.; Kliffen, M.; van Deurzen, C. H. M.; Luider, T. M.; Foekens, J. A.; Martens, J. W.; Umar, A. Annexin-A1 and Caldesmon Are Associated with Resistance to Tamoxifen in Estrogen Receptor Positive Recurrent Breast Cancer. *Oncotarget* **2016**, *7* (3), 3098. https://doi.org/10.18632/ONCOTARGET.6521.

(12) Schäfer, B. W.; Heizmann, C. W. The S100 Family of EF-Hand Calcium-Binding Proteins: Functions and Pathology. *Trends Biochem. Sci.* **1996**, *21* (4), 134–140. https://doi.org/10.1016/s0968-0004(96)80167-8.

(13) He Jingjing Li AE Shunyan Weng AE Mingfa Li AE Yan Yu, H. A. S100A11: Diverse Function and Pathology Corresponding to Different Target Proteins. **2009**. https://doi.org/10.1007/s12013-009-9061-8.

(14) Wang, H.; Yin, M.; Ye, L.; Gao, P.; Mao, X.; Tian, X.; Xu, Z.; Dai, X.; Cheng, H.

S100A11 Promotes Glioma Cell Proliferation and Predicts Grade-Correlated Unfavorable Prognosis. *Technol. Cancer Res. Treat.* **2021**, *20*, 1–9. https://doi.org/10.1177/15330338211011961.

(15) Miwa, N.; Uebi, T.; Kawamura, S. S100-Annexin Complexes-Biology of Conditional Association. https://doi.org/10.1111/j.1742-4658.2008.06653.x.

(16) Sakaguchi, M.; Huh, N.-H. S100A11, a Dual Growth Regulator of Epidermal Keratinocytes. https://doi.org/10.1007/s00726-010-0747-4.

(17) Sakaguchi, M.; Murata, H.; Sonegawa, H.; Sakaguchi, Y.; Futami, J. I.; Kitazoe, M.; Yamada, H.; Huh, N. H. Truncation of Annexin A1 Is a Regulatory Lever for Linking Epidermal Growth Factor Signaling with Cytosolic Phospholipase A2 in Normal and Malignant Squamous Epithelial Cells. *J. Biol. Chem.* **2007**, *282* (49), 35679–35686. https://doi.org/10.1074/jbc.M707538200.

(18) Poeter, M.; Radke, S.; Koese, M.; Hessner, F.; Hegemann, A.; Musiol, A.; Gerke, V.; Grewal, T.; Rescher, U. Disruption of the Annexin A1/S100A11 Complex Increases the Migration and Clonogenic Growth by Dysregulating Epithelial Growth Factor (EGF) Signaling. *Biochim. Biophys. Acta - Mol. Cell Res.* **2013**, *1833* (7), 1700–1711. https://doi.org/10.1016/j.bbamcr.2012.12.006.

(19) Tang, X.; Jia, X.; Huang, Z. Challenges and Opportunities for Alkane Functionalisation Using Molecular Catalysts. *Chem. Sci.* **2018**, *9* (2), 288–299. https://doi.org/10.1039/c7sc03610h.

(20) Hermans, I.; Spier, E. S.; Neuenschwander, U.; Turrà, N.; Baiker, A. Selective Oxidation Catalysis: Opportunities and Challenges. *Top. Catal.* **2009**, *52* (9), 1162–1174. https://doi.org/10.1007/s11244-009-9268-3.

(21) Lersch, M.; Tilset, M. Mechanistic Aspects of C−H Activation by Pt Complexes. **2005**. https://doi.org/10.1021/cr030710y.

(22) Labinger, J. A.; Bercaw Arnold, J. E.; Beckman, M. *Understanding and Exploiting C-H Bond Activation*; 2002.

(23) J. Blanksby, S.; Barney Ellison, G. Bond Dissociation Energies of Organic Molecules. *Acc. Chem. Res.* **2003**, *36* (4), 255–263. https://doi.org/10.1021/ar020230d.

(24) Dalton, T.; Faber, T.; Glorius, F. C−H Activation: Toward Sustainability and Applications. *Cite This ACS Cent. Sci* **2021**, *7*, 245–261. https://doi.org/10.1021/acscentsci.0c01413.

(25) Yi, H.; Zhang, G.; Wang, H.; Huang, Z.; Wang, J.; Singh, A. K.; Lei, A. Recent Advances in Radical C−H Activation/Radical Cross-Coupling. **2017**. https://doi.org/10.1021/acs.chemrev.6b00620.

(26) Godula, K.; Sames, D. C-H Bond Functionalization in Complex Organic Synthesis.

(27) Davies, H. M. L.; Manning, J. R.; Ml, M.; Ml, C. Catalytic C-H Functionalization by Metal Carbenoid and Nitrenoid Insertion. **2008**. https://doi.org/10.1038/nature06485.

(28) Ortuño, M. A.; Conejero, S.; Lledós, A. True and Masked Three-Coordinate T-Shaped Platinum(II) Intermediates. *Beilstein J. Org. Chem* **2013**, *9*, 1352–1382. https://doi.org/10.3762/bjoc.9.153.

(29) D. Jones, W. Isotope Effects in C−H Bond Activation Reactions by Transition Metals. *Acc. Chem. Res.* **2002**, *36* (2), 140–146. https://doi.org/10.1021/ar020148i.

(30) Roudesly, F.; Oble, J.; Poli, G. Metal-Catalyzed C[Sbnd]H Activation/Functionalization: The Fundamentals. *J. Mol. Catal. A Chem.* **2017**, *426*, 275–296. https://doi.org/10.1016/j.molcata.2016.06.020.

(31) Sheng, C.-X.; Singh, S.; Gambetta, A.; Drori, T.; Tong, M.; Tretiak, S.; Vardeny, Z. V. Ultrafast Intersystem-Crossing in Platinum Containing p-Conjugated Polymers with Tunable Spin-Orbit Coupling. **2013**. https://doi.org/10.1038/srep02653.

(32) Brooks, J.; Babayan, Y.; Lamansky, S.; Djurovich, P. I.; Tsyba, I.; Bau, R.; Thompson, M. E. Synthesis and Characterization of Phosphorescent Cyclometalated Platinum Complexes. **2002**. https://doi.org/10.1021/ic0255508.

(33) Baldo, M. A.; O'brien, D. F.; You, Y.; Shoustikov, A.; Sibley, S.; Thompson, M. E.; Forrest, S. R. Highly Efficient Phosphorescent Emission from Organic Electroluminescent Devices. **1998**.

(34) Huo, S. Platinum in Chemistry: An Adventure from Phosphorescent Materials to Catalytic CÀH Functionalization. **2018**. https://doi.org/10.1002/tcr.201800019.

(35) Aime, S.; Botta, M.; Fasano, M.; Paula, M.; Marques, M.; Geraldes, C. F. G. C.; Pubanz, D.; Merbach, A. E. *Conformational and Coordination Equilibria on DOTA Complexes of Lanthanide Metal Ions in Aqueous Solution Studied by 1 H-NMR Spectroscopy*; 1997.

(36) Jørgensen, C. K. Ligand Substitution Processes. *J. Franklin Inst.* **1966**, *282* (4), 262–263. https://doi.org/10.1016/0016-0032(66)90569-2.

(37) Padhi, A. K.; Janežič, M.; Zhang, K. Y. J. Molecular Dynamics Simulations: Principles, Methods, and Applications in Protein Conformational Dynamics. *Adv. Protein Mol. Struct. Biol. Methods* **2022**, 439–454. https://doi.org/10.1016/B978-0-323-90264-9.00026-X.

(38) Christopher J Cramer. *Essentials of Computational Chemistry*, Second.; John Wiley & Sons Ltd: Chichester, West Sussex, 2004.

(39) Filipe, H. A. L.; Loura, L. M. S. Molecular Dynamics Simulations: Advances and

Applications. *Molecules* **2022**, *27* (7), 37–47. https://doi.org/10.3390/molecules27072105.

(40) Onufriev, A.; Bashford, D.; Case, D. A. Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model. *Proteins Struct. Funct. Bioinforma.* **2004**, *55* (2), 383–394. https://doi.org/10.1002/PROT.20033.

(41) David A. Case; Walker, R. C.; II, T. E. C.; Simmerling, C. Amber 2022 Reference Manual. **2022**.

(42) Pal, S.; Ray, B. C. Basics of Molecular Dynamics Simulation. *Mol. Dyn. Simul. Nanostructured Mater.* **2020**, 171–198. https://doi.org/10.1201/9780429019845-5.

(43) Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2015**, *11* (4), 1864–1874. https://doi.org/10.1021/ct5010406.

(44) Terme, D.; Beitrag, S.; Grund, D.; Standpunkte, V. Der Physik 1. **1927**.

(45) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications, Inc.: Mineola, New York, 1996.

(46) Helgaker, T.; Jorgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons, Ltd, 2000.

(47) DeRocco, A. G. *Modern Quantum Chemistry*; 1967; Vol. 35. https://doi.org/10.1119/1.1973756.

(48) Sherril, D. *Notes on Quantum Chemistry*. http://vergil.chemistry.gatech.edu/notes/index.html.

(49) Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press; Clarendon Press, 1994.

(50) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136* (3B).

https://doi.org/10.1103/PhysRev.136.B864.

(51) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140* (4A). https://doi.org/10.1103/PhysRev.140.A1133.

(52) Geerlings, P.; Chamorro, E.; Kumar, P.; Frank, C.; José, D. P.; Liu, S.; Morell, C.; Toro, A.; Alberto, L.; Paul, V. Conceptual Density Functional Theory : Status , Prospects , Issues. *Theor. Chem. Acc.* **2020**, *139* (2), 1–18. https://doi.org/10.1007/s00214-020-2546-7.

(53) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. A Climbing Image Nudged Elastic Band Method for Finding Saddle Points and Minimum Energy Paths. *J. Chem. Phys.* **2000**, *113* (22), 9901. https://doi.org/10.1063/1.1329672.

(54) Lewis, K.; Lindsay, S.; Li, Y. Nudged Elastic Band Study on the N-Terminal Domain Conformational Pathways of Annexin A1 from a Buried State to an Exposed State. *J. Phys. Chem. B* **2019**, *123* (48), 10163–10170. https://doi.org/10.1021/acs.jpcb.9b08120.

(55) Simpkins, B.; Donohue, M. P.; Li, Y. Molecular Dynamic Studies on the Impact of Mutations on the Structure, Stability, and N-Terminal Orientation of Annexin A1: Implications for Membrane Aggregation. *Proteins Struct. Funct. Bioinforma.* **2014**, *82* (12), 3327–3334. https://doi.org/10.1002/PROT.24684.

(56) Hung, K. W.; Chang, Y. M.; Yu, C. NMR Structure Note: The Structure of Human Calcium-Bound S100A11. *J. Biomol. NMR* **2012**, *54* (2), 211–215. https://doi.org/10.1007/S10858-012-9661-2/FIGURES/2.

(57) Réty, S.; Osterloh, D.; Arié, J. P.; Tabaries, S.; Seeman, J.; Russo-Marie, F.; Gerke, V.; Lewit-Bentley, A. Structural Basis of the Ca2+-Dependent Association between S100C (S100A11) and Its Target, the N-Terminal Part of Annexin I. *Structure* **2000**, *8* (2), 175–

184. https://doi.org/10.1016/S0969-2126(00)00093-9.

(58) Rosengarth, A.; Gerke, V.; Luecke, H. X-Ray Structure of Full-Length Annexin 1 and Implications for Membrane Aggregation. *J. Mol. Biol.* **2001**, *306* (3), 489–498. https://doi.org/10.1006/JMBI.2000.4423.

(59) Madeira, F.; Pearce, M.; Tivey, A. R. N.; Basutkar, P.; Lee, J.; Edbali, O.; Madhusoodanan, N.; Kolesnikov, A.; Lopez, R. Search and Sequence Analysis Tools Services from EMBL-EBI in 2022. *Nucleic Acids Res.* **2022**, *50* (W1), gkac240–gkac240. https://doi.org/10.1093/NAR/GKAC240.

(60) Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinformatics* **2016**, *54*, 5.6.1. https://doi.org/10.1002/CPBI.3.

(61) Weng, X.; Luecke, H.; Song, I. S.; Kang, D. S.; Kim, S. S.-H.; Huber4, R. *Crystal Structure of Human Annexin I at 2.5 A Resolution*; Cambridge University Press, 1993; Vol. 2.

(62) Courbet, A.; Hansen, J.; Hsia, Y.; Bethel, N.; Park, Y.-J.; Xu, C.; Moyer, A.; Boyken, S. E.; Ueda, G.; Nattermann, U.; Nagarajan, D.; Silva, D.; Sheffler, W.; Quispe, J.; Nord, A.; King, N.; Bradley, P.; Veesler, D.; Kollman, J.; Baker, D. Computational Design of Mechanically Coupled Axle-Rotor Protein Assemblies. *Science (80-.).* **2022**, *376* (6591), 383–390. https://doi.org/10.1126/SCIENCE.ABM1183.

(63) D.A. Case, H.M. Aktulga, K. Belfon, I.Y. Ben-Shalom, J.T. Berryman, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, and P. A. K. Amber 2022. University of California,San Francisco 2022.

(64) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water

Models at 298 K. *J. Phys. Chem. A* **2001**, *105* (43), 9954–9960. https://doi.org/10.1021/JP003020W.

(65) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713. https://doi.org/10.1021/ACS.JCTC.5B00255/SUPPL_FILE/CT5B00255_SI_001.PDF.

(66) Bartolotti, L. J.; Pedersen, L. G.; Charifson, P. S. Long Range Nonbonded Attractive Constants for Some Charged Atoms. *J. Comput. Chem.* **1991**, *12* (9), 1125–1128. https://doi.org/10.1002/JCC.540120910.

(67) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327–341. https://doi.org/10.1016/0021-9991(77)90098-5.

(68) Roe, D. R.; Cheatham III, T. E. Parallelization of CPPTRAJ Enables Large Scale Analysis of Molecular Dynamics Trajectory Data. **2018**. https://doi.org/10.1002/jcc.25382.

(69) R. Roe, D.; E. Cheatham III, T. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084–3095. https://doi.org/10.1021/ct400341p.

(70) R. Miller III, B.; Dwight McGee Jr., T.; M. Swails, J.; Homeyer, N.; Gohlke, H.; E. Roitberg, A. MMPBSA.Py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8* (9), 3314–3321. https://doi.org/10.1021/ct300418h.

(71) WL, D. The PyMOL Molecular Graphics System. *CCP4 Newsletter On Protein Crystallography*. Schrödinger, LLC 2002, pp 82–92.

(72) Lambert, O.; Gerke, V.; Bader, M. F.; Porte, F.; Brisson, A. Structural Analysis of

Junctions Formed between Lipid Membranes and Several Annexins by Cryo-Electron Microscopy. *J. Mol. Biol.* **1997**, *272* (1), 42–55. https://doi.org/10.1006/JMBI.1997.1183.

(73) Oh, Y. S.; Gao, P.; Lee, K. W.; Ceglia, I.; Seo, J. S.; Zhang, X.; Ahn, J. H.; Chait, B. T.; Patel, D. J.; Kim, Y.; Greengard, P. SMARCA3, a Chromatin-Remodeling Factor, Is Required for P11-Dependent Antidepressant Action. *Cell* **2013**, *152* (4), 831–843. https://doi.org/10.1016/J.CELL.2013.01.014.

(74) Dorovkov, M. V; Kostyukova, A. S.; Ryazanov, A. G. Phosphorylation of Annexin A1 by TRPM7 Kinase: A Switch Regulating the Induction of an r-Helix. **2011**, *50*, 2187–2193. https://doi.org/10.1021/bi101963h.

(75) Weisz, J.; Uversky, V. N. Zooming into the Dark Side of Human Annexin-S100 Complexes: Dynamic Alliance of Flexible Partners. *Int. J. Mol. Sci.* **2020**, *21* (16), 1–30. https://doi.org/10.3390/ijms21165879.

(76) Barham, A.; Neu, J.; Canter, C. L.; Pike, R. D.; Li, Y.; Huo, S. Isomerization-Induced Multiple Reaction Pathways in Platinum-Catalyzed C-H Acylation Reaction of 2-Aryloxypyridines. *Organometallics* **2021**, *40* (18), 3158–3169. https://doi.org/10.1021/acs.organomet.1c00377.

(77) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenb, D. J. Gaussian 09, Revision D.01. Gaussian, Inc.: Wallingford,CT 2009.

(78) Zhao, Y.; Truhlar, D. G. 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.* **2008**, *120* (1–3), 215–241. https://doi.org/10.1007/s00214-007-0310-x.

(79) Weigend, F.; Ahlrichs, R. 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.* **2005**, *7* (18), 3297–3305. https://doi.org/10.1039/B508541A.

(80) Petersson, G. A.; Al-Laham, M. A. A Complete Basis Set Model Chemistry. II. Open-shell Systems and the Total Energies of the First-row Atoms. *J. Chem. Phys.* **1998**, *94* (9), 6081. https://doi.org/10.1063/1.460447.

(81) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105* (8), 2999–3093. https://doi.org/10.1021/cr9904009.

(82) Henkelman, G.; Jónsson, H. A Dimer Method for Finding Saddle Points on High Dimensional Potential Surfaces Using Only First Derivatives. *J. Chem. Phys.* **1999**, *111* (15), 7010. https://doi.org/10.1063/1.480097.

(83) Garner, A. W.; Harris, C. F.; Vezzu, D. A. K.; Pike, R. D.; Huo, S. Solvent-Controlled Switch of Selectivity between Sp2 and Sp3 C-H Bond Activation by Platinum(II). *Chem. Commun.* **2011**, *47* (6), 1902–1904. https://doi.org/10.1039/c0cc04581k.

(84) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A. Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasing, D. J. Gaussian 16. Gaussian, Inc.: Wallingfort CT 2016.

(85) Avogadro: An Open-Source Molecular Builder and Visualization Tool. Version 1.95.

(86) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminform.* **2012**, *4* (17).

(87) Bader, R. F. W. Bond Paths Are Not Chemical Bonds. *J. Phys. Chem. A* **2009**, *113* (38), 10391–10396. https://doi.org/10.1021/jp906341r.

(88) Bader, R. Richard F. W. Bader - Atoms in Molecules_ A Quantum Theory (International Series of Monographs on Chemistry)-Oxford University Press, USA (1994).Pdf.