# Exploring the Claisen Rearrangement and Aldol Reaction "on Water" and in Catalytic Antibodies using QM/MM Simulations and the Examination of Ligand Induced Conformational Changes in Alkanesulfonate Monooxygenase using Molecular Dynamics

by

Kira A. Armacost

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
May 5, 2013



Approved by

Orlando Acevedo, Chair, Associate Professor of Chemistry
J. Vincent Ortiz, Ruth W. Molette Professor and Chairman
Konrad Patkowski, Assistant Professor of Chemistry
Michael Squillacote, Associate Professor of Chemistry

## Abstract

The “on water” environment, defined by the absence of water solubility of the reactants has been reported to provide increased rate accelerations, yields, and specificity for several types of organic reaction classes compared to organic solvents. The aromatic Claisen rearrangements of allyl *p*-R-phenyl ethers (R = $CH_3$, Br, and $OCH_3$) and allyl naphthyl ether, and another class of reaction, the aldol reaction have been investigated to determine their on water enhancements using QM/MM Monte Carlo calculations and free-energy perturbation theory. The aldol reaction was further studied in catalytic antibody 33F12 to determine the mechanistic properties and its comparison with the aldol reaction in condensed phase simulations. In another study, molecular dynamics calculations were performed on monooygenase systems to identify the protein fluctuations with various bound ligands (reduced flavin mononucleotide ($FMNH_2$), a C4a-peroxyflavin intermediate, alkanesulfonate, $FMNH_2$-alkanesulfonate, and C4a-peroxyflavin intemediate-alkanesulfonate) and an apo (free of ligand) structure. Salt bridge formations have been shown to participate in dynamic conformational changes for the ligand bound structures between Glu20-Arg297 and Asp111-Arg297. Also, Arg226 was confirmed to stabilize the peroxy group of the c4a-peroxyflavin ligand, and possibly help contributing to the sulfite release.

## Acknowledgements

My sincere and utmost appreciation is given to my advisor, Dr. Orlando Acevedo, for all the guidance he has given me through my graduate school career. I am extremely thankful for all the mentoring and support I have received from him. Next, I'd like to thank Dr. Vincent Ortiz, Dr. Konrad Patkowski, Dr. Michael Squillacote and Dr. Paul Cobine for serving as my committee members and for the support I've gotten in my dissertation process. To my past and present research lab mates, Samba, Billy, Caley, Symon, Nicole and Bin, I'd like to thank each one of you for helping with everything I've needed and being a great support system. Last, but definitely not least, I'd like to especially thank my dad, mom and brother who have dealt with all of my calls home and extreme levels of stress. Without my family, I wouldn't have gotten to where I am today.

# Table of Contents

# List of Tables

# List of Figures

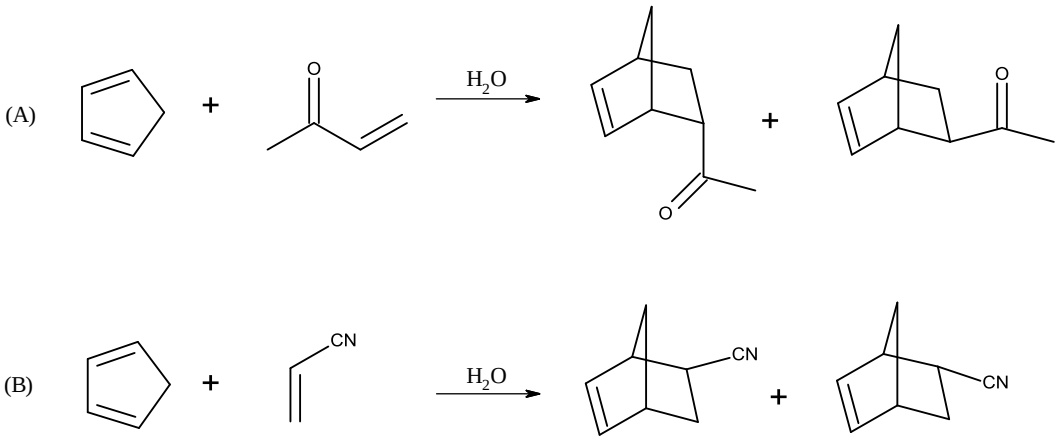
# List of Schemes

# Chapter 1

## Introduction

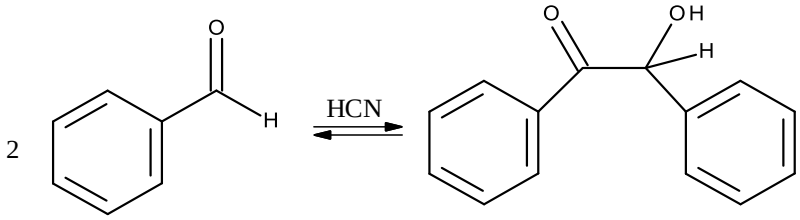
### 1.1 Chemical Reactions in Aqueous Solutions

Considerable attention has been directed towards the development of organic reactions that benefit from water in terms of enhanced stereoselectivities, yields, and rates; the list is extensive and reviews are available.[1-9] The interest in water as a viable alternative to organic solvents began with reports from Breslow and coworkers of significant rate accelerations for Diels-Alder reactions in aqueous solution. He proposed the hydrophobic effect as the main reason for rate accelerations, where nonpolar solutes move to a region that diminishes the accessible surface area. Several Diels-Alder reactions, e.g. cyclopentadiene with butenone (Scheme 1.1A) and cyclopentadiene with acrylonitrile (Scheme 1.1B) were found to have enhanced rates in aqueous solvent, compared to a hydrocarbon solvent or




**Scheme 1.1:** Diels-Alder reaction of (A) cyclopentadiene and butenone and (B) cyclopentadiene and acrylonitrile.

methanol.[10] These reactions were also performed by "salting out" the reaction in water by the use of lithium chloride, and greater rate enhancements were observed.[10-11] The Diels-Alder work was expanded to further include the benzoin condensation (Scheme 1.2), which also showed increased rates in water and with the "salting out"




**Scheme 1.2:** The Benzoin condensation.

agent.[11-12] That work and the potential advantages derived from an aqueous reaction medium in organic synthesis, e.g., safety, low cost, and ease of product isolation, has led to a widespread interest in the field. These include other pericyclic reactions including the Claisen rearrangement and cycloaddition reactions, along with oxidation and carbon-carbon bond forming reactions, including the aldol condensation.[12]

**1.2 "On Water" Reactions**

The solubility of organic compounds has been known to be necessary for reactions to go to completion. The phrase, "corpora non agunt nisi solute" has been used to describe this phenomenon, and is translated as, "substances do not interact unless dissolved."[13] Organic cosolvents are frequently utilized to increase the solubility of organic solutes in aqueous solutions[9] and consequently much of water's unique properties, i.e. hydrophobic effect, high polarity, and high cohesive energy density, are diminished or lost.[7] However, the assumption that solubility is required for efficient

catalysis has been challenged by Sharpless who recently reported large rate and yield increases comparable to organic solvents for a variety of reactions solely "on water".[13-14] "On water" is defined as a reaction environment that proceeds in an aqueous organic emulsion prepared by vigorously stirring insoluble reactants with water, whereas "in water" the reactants are dissolved homogeneously in water.[13, 15-16] The term "in the presence of water" has also been used interchangeably to describe "on water" conditions.[15]

In 2005, Narayan, Sharpless and coworkers described cycloaddition reactions, rearrangement reactions, and epoxide ring opening reactions that underwent "on water" rate accelerations. The general finding was that solubility of the substrate was not crucial for the reactions to occur.[13] A review paper was subsequently published by Chanda and Fokin describing many types of organic reactions occurring in an "on water" environment.[1] These reactions included nucleophilic substitution reactions, carbon-carbon bond forming reactions, and oxidation/reduction reactions, among others, which all undergo rate accelerations and/or enhanced stereoselectivities. In many cases, a significant rate increase for "on water" reactions versus solvent-free (or "neat") reactions indicates that the rate acceleration is not simply a consequence of increased concentration of the reacting species.[1]

While the molecular details behind "on water" reactivity are scarce, Jung and Marcus proposed that free OH groups from water molecules protruding into the organic phase play a key role in catalyzing reactions via the formation of hydrogen bonds.[17-18] Their computational work used transition state theory (TST) and UB3LYP/6-31+G(d)

energy minimizations on the cycloaddition of quadricyclane and dimethyl azodicarboxylate complexed to three water molecules (Figure 1). Their interpretations have greatly enhanced the general understanding of "on water" catalysis. However, a major drawback associated with using a "supermolecule complex" approach is its inability to account for hydrophobic effects.

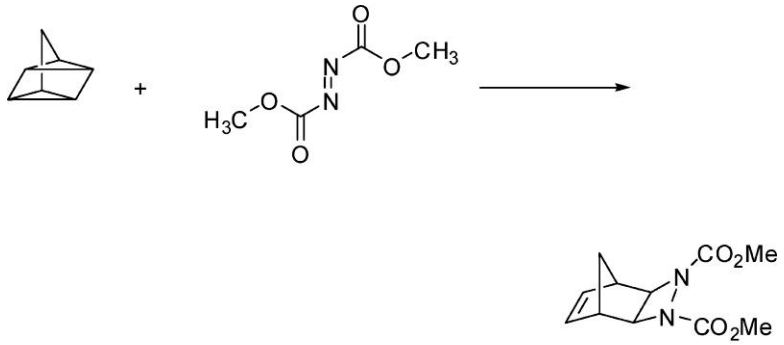



**Figure 1.1:** The reaction of quadricyclane and dimethyl azodicarboxylate (Adapted from Jung et al.[17])

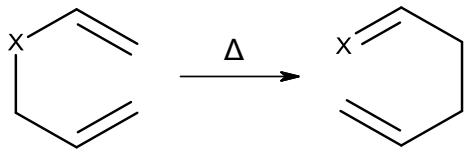
Mixed quantum and molecular mechanical (QM/MM) simulations were performed on the Diels-Alder reaction of cyclopentadiene and 1,4-napthoquinone, methyl vinyl ketone (MVK), and acrylonitrile in 2010 by Thomas et al., in order to elucidate the "on water" rate enhancements.[19] They concluded that the hydrogen bond acceptors oriented themselves in a way that they were either parallel to the surface of water, or were protruding into the water and OH bonds dangling at the water interface was unlikely to be the reason for catalysis.

QM/MM methodology has been utilized in this thesis to produce a more experimentally realistic environment featuring hundreds of water molecules with the reacting organic layer lying on top.[5] The reaction of interest in this thesis for clarifying

the “on water” enhancements are the Claisen rearrangements of allyl *p*-tolyl ether and allyl naphthyl ether and the aldol reaction between acetone and benzaldehyde.

### 1.3 The Claisen Rearrangement

The Claisen rearrangement was discovered in 1912 by L. Claisen and describes a [3,3] sigmatropic rearrangement of allyl vinyl ethers to γ,δ-unsaturated carbonyl compounds.[12, 20] Claisen described an isomerization of an allyl vinyl ether (where X could represent oxygen, nitrogen or sulfur) compound (Scheme 1.3).[20-21] The first account of the Claisen rearrangement was in 1912 where L. Claisen described the rearrangement of




X = O, N, S

**Scheme 1.3:** [3,3] sigmatropic rearrangement of an allyl vinyl ether.

an *O*-allylated acetoacetate with $NH_4Cl$ present.[20] Over the next few years, a series of research groups discovered that aliphatic substrates, not just aromatic substrates could also undergo a [3,3] sigmatropic rearrangement (see reference 23 for a review). The first Claisen rearrangement in aqueous media was the transformation of chorismate to prephenate (Figure 1.2). The reaction of chorismate to prephenate is an important reaction for generation of aromatic amino acids in bacterial, fungal and plant systems in the shikimate biosynthetic pathway.

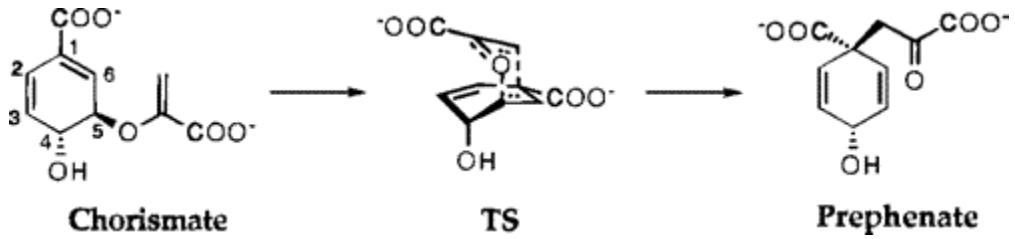
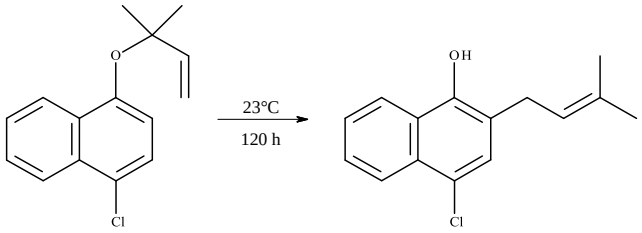



**Figure 1.2:** Conversion of chorsimate to prephenate. (Adapted from Carlson et al.[22])

Since the 1970's, the Claisen rearrangement has been extensively studied both experimentally and theoretically, and in the gas phase and solvent.[23-30]

Solvent dependent rate enhancements have been reported for the Claisen rearrangement of allyl *p*-tolyl ether by White et al. in 1970.[31-34] His group studied solvent effects and substituent effects in 17 solvents with 4 different substituents. Electron-withdrawing and electron-donating substituents were utilized to understand the effect that substituents have on the reaction rate.[32] For example, the rate of reaction increased by a factor of 100 upon going from the least polar solvent (tetradecane) to the most polar solvent (*p*-chlorophenol), and the allyl *p*-R-tolyl ether where R = $OCH_3$ was faster than when R= $NO_2$ in tetradecane, carbitol and a 28.5% ethanol-water mixture.[31-32] However, the rates in phenol and water are about the same and are 300 times greater than in vacuum for the ether rearrangement.[31] The solvent dependence of rates for Claisen reactions is complex and does not show simple increases with increasing solvent polarity. Since the solvent polarity can't fully explain the rate enhancements, researchers have explored other factors, including one where the solute does not dissolve in the solvent.

In 2005, Sharpless and coworkers discovered that the Claisen rearrangement underwent "on water" reactivity. The aromatic Claisen rearrangement of 1-chloro-4-[(1,1-dimethyl-2-propen-1-yl)oxy]-naphthalene (allyl naphthyl ether), (Scheme 1.4) demonstrated a significant increase in yield when performed in an "on water" environment (Table 1.1), but the exact reasoning behind the enhanced yield was unknown.

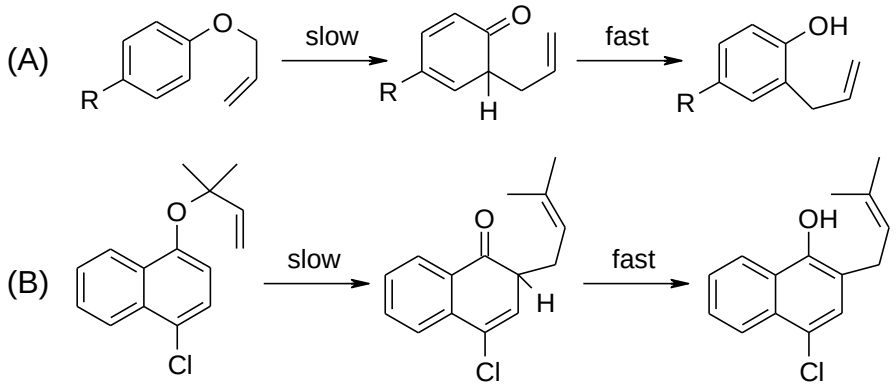



**Scheme 1.4:** The Claisen rearrangement of allyl naphthyl ether.

**Table 1.1:** Percent yields of different solvents for the Claisen rearrangement of allyl naphthyl ether.

| Solvent | Yield (%) |
|---|---|
| Toluene | 16 |
| DMF | 21 |
| $CH_3CN$ | 27 |
| MeOH | 56 |
| Neat | 73 |
| On $H_2O$ | 100 |

The reactions studied here for the elucidation of "on water" reactivity were the Claisen rearrangements of allyl *p*-R-phenyl ethers and allyl naphthyl ether (Scheme 1.5). Experimental rates and thermodynamic parameters have been published for the

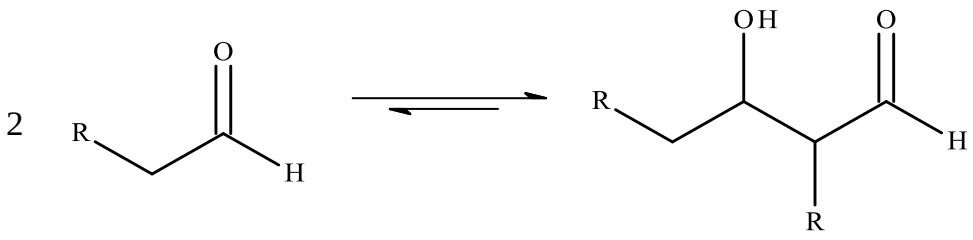



**Scheme 1.5:** Claisen rearrangement of (A) allyl *p*-R-phenyl ethers, R = $CH_3$, Br, and $OCH_3$, and (B) allyl naphthyl ether.

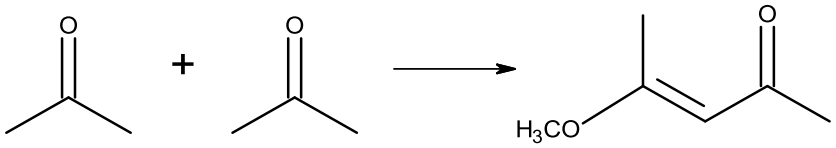
rearrangement of allyl *p*-tolyl ether (R = $CH_3$ in Scheme 1.5A) in the gas phase and in 17 solvents of different polarities.[31]

### 1.4 The Aldol Reaction

Another reaction class thought to benefit from an "on water" environment is the aldol reaction.[35] In the aldol reaction, a carbonyl that can undergo enolization reacts with the carbonyl of either an aldehyde or a ketone. The enol must have an acidic proton at its alpha position in order to undergo a nucleophilic attack at the carbonyl of the other aldehyde or ketone, as seen in Scheme 1.6.[36-37] In 1848, Kane described the first aldol reaction, where acetone reacted to generate mesityl oxide (Scheme 1.7). It was further described by Schmidt, Claisen, Claparede, with the first base-catalyzed aldol




**Scheme 1.6:** General aldol reaction.




**Scheme 1.7:** Reaction of acetone to mesityl oxide.

condensations of a ketone and aldehyde (Claisen-Schmidt condensation)[38-41] and Wurtz who reacted 3-hydroxybutanal with acetaldehyde, coining the name aldol condensation.[36, 42] One thing each of these had in common was a catalyst, which is necessary for the aldol reaction to occur.

The initial reaction conditions used an acid or base as the catalyst and these catalysts have been studied extensively.[43-44] Recently, the use of organocatalysts for the aldol reaction has been studied, mainly proline and prolinamide derivatives. Proline has been demonstrated as an effective catalyst for the intramolecular aldol reaction since the 1970's[45-48], and expanded more recently to include the intermolecular aldol reaction. In 2000, Barbas et al. developed a series of amino acid catalysts for the reaction of acetone and 4-nitrobenzaldehyde. This reaction proceeds through an enamine mechanism, just

like the Class I aldolase antibodies (Figure 1.3).[49] Step (a) involves a nucleophilic attack of the amino group on the carbonyl, next, (b) involves the dehydration of the intermediate, (c) tautomerization to the enamine, (d) the carbon-carbon bond formation, and the final steps, (e & f) where hydrolysis occurs of the iminium-aldol intermediate to give the aldol product.[50] The use of proline derivatives as catalysts is extremely beneficial, due to proline's nontoxicity, availability and ease of extraction.[50] Proline derivatives have also shown stereospecificity in the aldol reaction.[50-53]

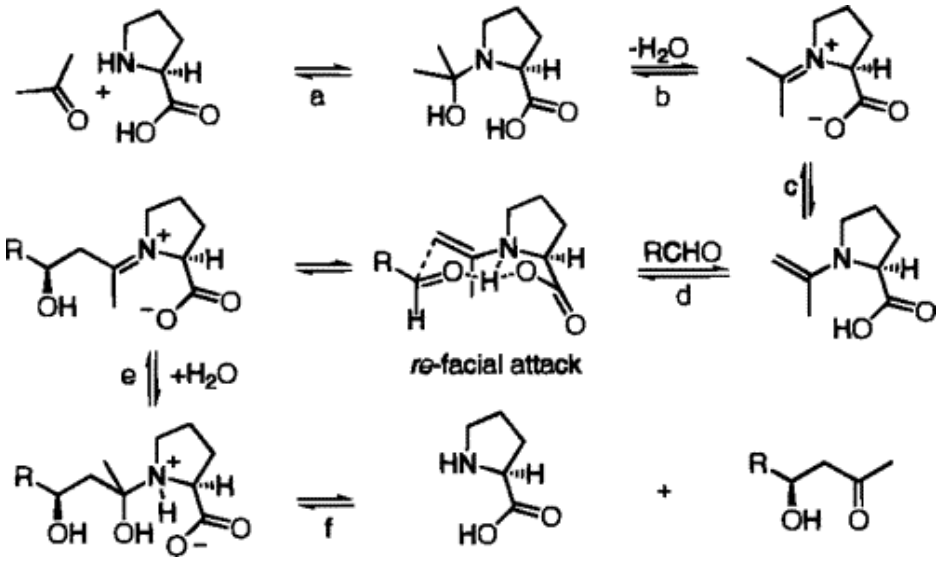
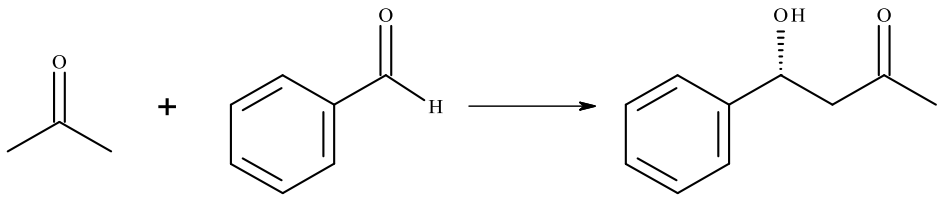



**Figure 1.3:** Enamine mechanism of acetone and 4-nitrobenzaldehyde. (Adapted from List et al.[50])

Singh et al. in 2009 developed a series of small organocatalysts for the enantioselective aldol reaction of benzaldehyde and acetone (Scheme 1.8). These catalysts were designed based on the idea that if the $pK_a$ of the hydroxyl group involved

in hydrogen bonding was increased, higher enantioselectivity would be noticed due to a compact transition structure.[35] The organocatalyst giving the best enantioselectivity




**Scheme 1.8:** Reaction of acetone and benzaldehyde.

was chosen and the conditions were further optimized (using different aldehydes and cyclic ketones), seen in Figure 1.4. A gem-diphenyl group was necessary for a compact transition state.[35]

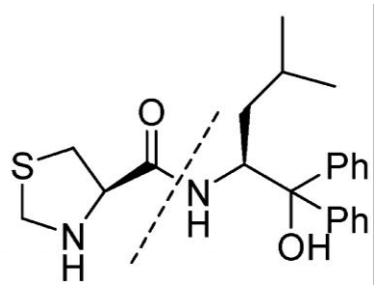



**Figure 1.4:** Hydrophobic organocatalyst with gem-diphenyl group. (Adapted from Vishnumaya et al.[35])

The transition structure of the enamine mechanism was initially thought to occur via the Zimmerman-Traxler six-membered ring model (Figure 1.5A), but DFT calculations later predicted a different model (Figure 1.5B). The Zimmerman-Traxler model has the nitrogen from the proline hydrogen bonding with the proton of the proton

transfer, whereas the DFT model does not agree.[35, 54-56] Instead, it has a nine-membered transition structure where the enamine is formed anti to the carboxylic group, and is stabilized through hydrogen bonding.[35, 54, 56]

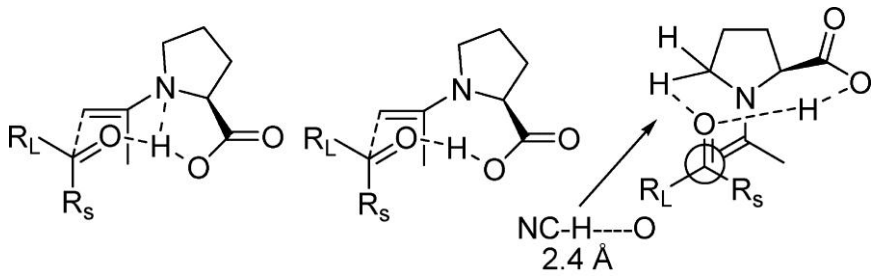



**Figure 1.5:** Transition state models of the enamine mechanism (A: Zimmerman-Traxler model and B: DFT models). (Adapted from Vishnumaya et al.[35])

To mimic the aldolase antibodies that catalyze aldol reactions in water, the organocatalysts developed must also perform well in water. For Singh et al.'s catalysts, water played a crucial role in increasing the enantioselectivity of the aldol reaction (Figure 1.6). As the equivalents of water increase to 50%, the enantioselectivity reached a maximum of 97%.

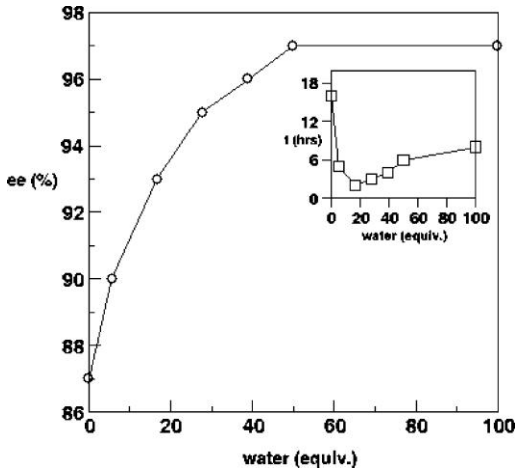



**Figure 1.6:** Role of water in the reaction of acetone and benzaldehyde with organocatalyst. (Adapted from Visnumaya et al.[35])

These enhanced enantioselectivities were due to hydrophobicity and were explained by transition state models (Figure 1.7).[35] C- C bond formation occurs from the *re* face (Figure 1.7B) due to favorable

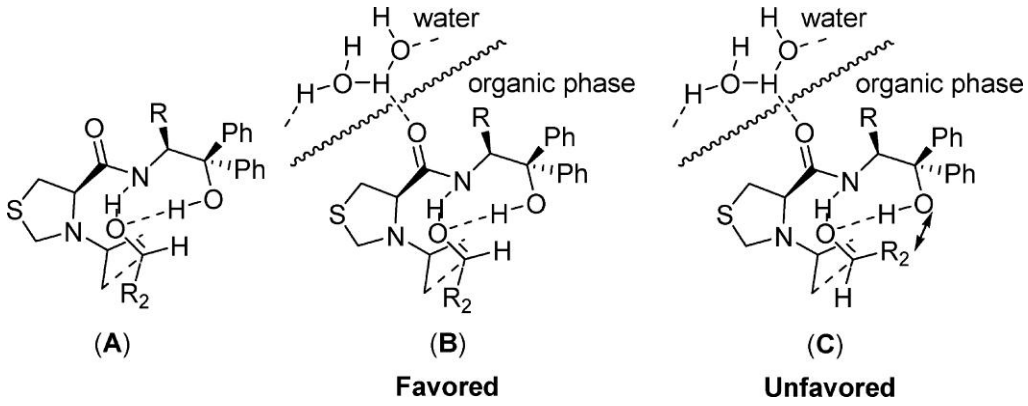
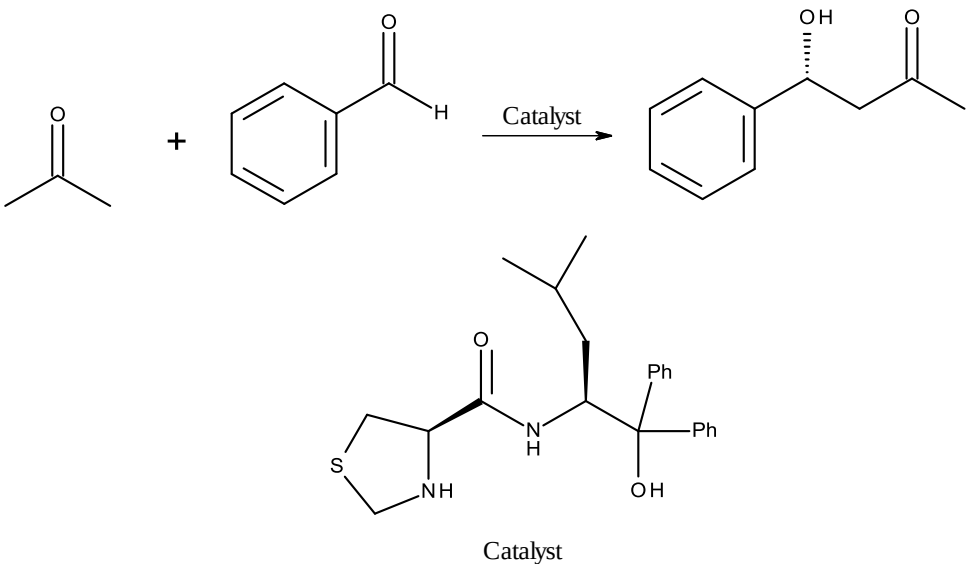



**Figure 1.7:** Transition state models of benzaldehyde and acetone with catalyst. (Adapted from Vishnumaya et al.[35])

hydrogen bonding between the NH and OH groups of the catalysts to the aldehyde oxygen. The *si* face (Figure 1.7C) is unfavorable due to the large steric interactions of the $R_2$ (benzene) and the OH of the catalyst. Figure 1.7 demonstrates an "on water" like interaction where the hydrogens at the surface of the water stabilize the amide oxygen via hydrogen bonding.[35]

To further characterize the "on water" reactivities, QM/MM/MC/FEP simulations (further explained in chapter 2.1) of the aldol reaction seen in scheme 1.9 were performed. The reactions were done in an "on water" environment, in water and in *n*-octanol. *n*-octanol was used to compare the reactivity of the organocatalyst to that of Class I aldolase antibodies compared to the intermolecular aldol reaction.




Catalyst

**Scheme 1.9:** Aldol reaction of acetone and benzaldehyde with chiral organocatalyst.

### 1.5 Catalytic Aldolase Antibody

Catalytic aldolase antibodies catalyze the aldol reaction via an enamine mechanism. Class I aldolases activate the substrate as an aldol donor by using the ε-

amino group of a lysine to form a Schiff base between the substrate and the amino group (Figure 1.8). Class II aldolases coordinate a metal to the carbonyl oxygen from the substrate to allow for enolate formation.[57]

The design of catalytic antibodies can be beneficial because they can be programmed to lower the activation energy of a reaction, resulting in an increased rate. Barbas et al. have developed a series of catalytic aldolase antibodies which produce a Schiff base (iminium ion) that lowers the activation energy for the abstraction of the proton from the α-carbon, forming an enamine (class I aldolase enzyme).[49] The carbon-carbon bond is formed between the enamine and aldehyde, similar to the enamine mechanism in Figure 1.3, step d .[57]

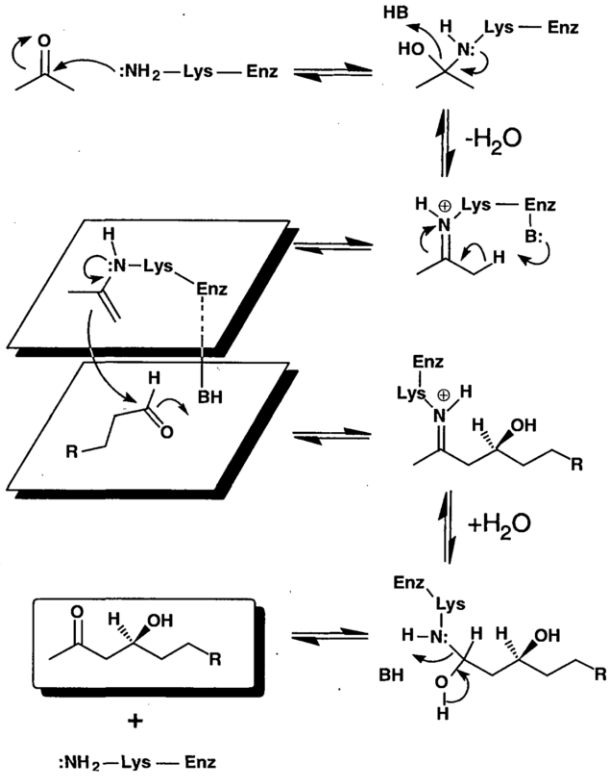



**Figure 1.8:** Mechanism of class-I aldolase antibodies (Enz=enzyme, B=base). (Adapted from Wagner et al.[57])

A *β*-diketone is used as a hapten (small molecule used to trigger an antibody to bind), which is able to trap a lysine in the active site, form the enamine, and overcome the activation barrier for the biological reaction. The two antibodies that were selected, 33F12 and 38C2 for the catalysis of the aldolase reaction gave the proper mechanism for enamine formation (Figure 1.9).[57-58] The ε-amino group from LysH93 reacts with a carbonyl from the *β*-diketone forming a *β*-keto hemiaminal intermediate. Tautomerization occurs after dehydration to a *β*-keto imine to an enaminone, causing the hapten to bind in the active site.[58] These antibodies catalyze a variety of aldol reactions

(ketone-ketone, ketone-aldehyde, aldehyde-ketone, aldehyde-aldehyde) intermolecular and intramolecular reactions and are very enantioselective.

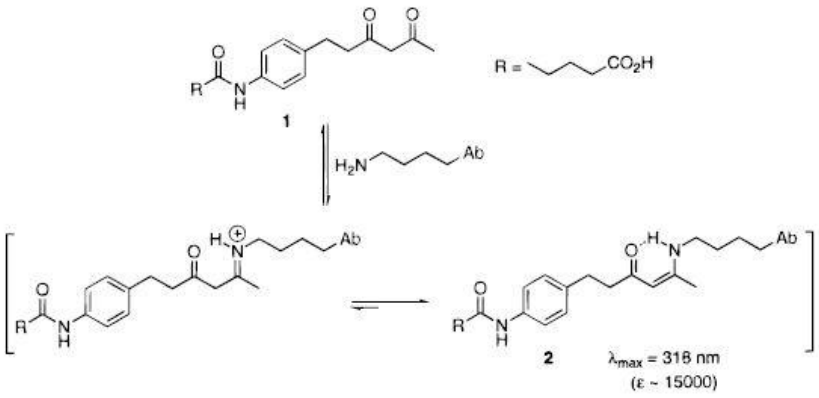
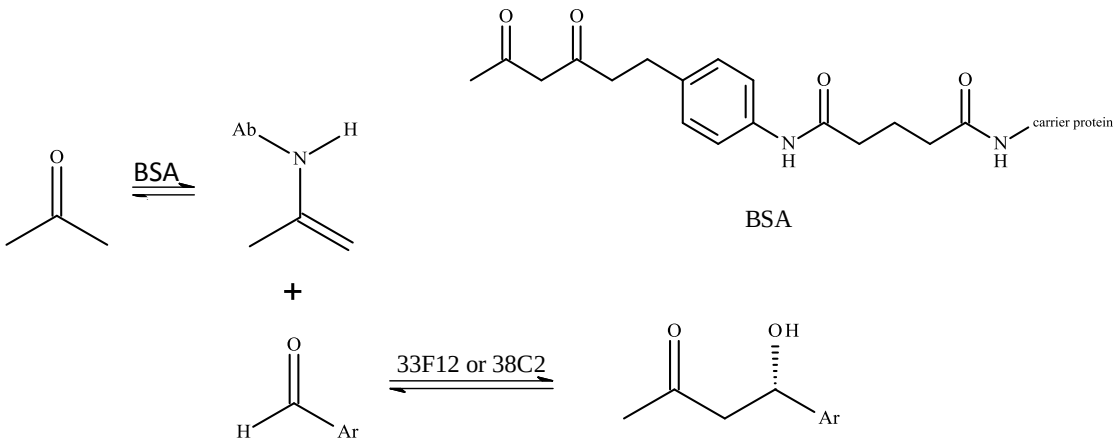



**Figure 1.9:** Mechanism for trapping Lys residue in active site by antibodies (Ab) 33F12 and 38C2. (Adapted from Hoffmann et al.[58])

These antibodies generated using the hapten BSA (bovine serum albumin) catalyze the reaction of a ketone and aromatic aldehyde to yield an (*S*)-aldol product (Scheme 1.10).[49, 57-58]




**Scheme 1.10:** Aldol reaction between a ketone and aromatic aldehyde using the antibodies 33F12 or 38C2 with BSA as the hapten.

The generation of an (*S*)-aldol product gave rise to the production of another class of antibodies able to generate an (*R*)-aldol product (93F3). 93F3 differs in its two lysine residues in the active site and the position of these lysine residues (LysL89 and LysL46).[49] Density functional theory (DFT) calculations were performed to locate the transition structure on a simplified structure of methylamine instead of the ε-amino group of the antibody (Figure 1.10). Docking was also performed for the four transition states

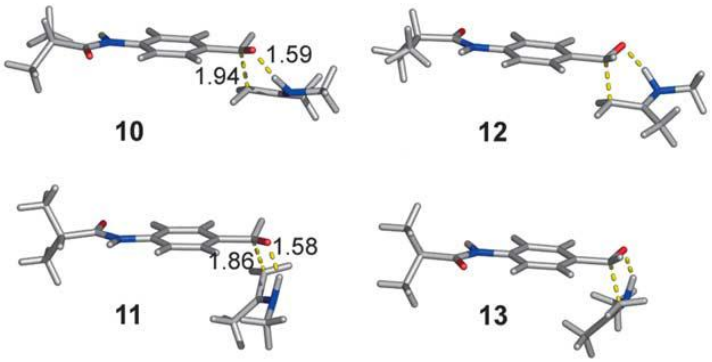
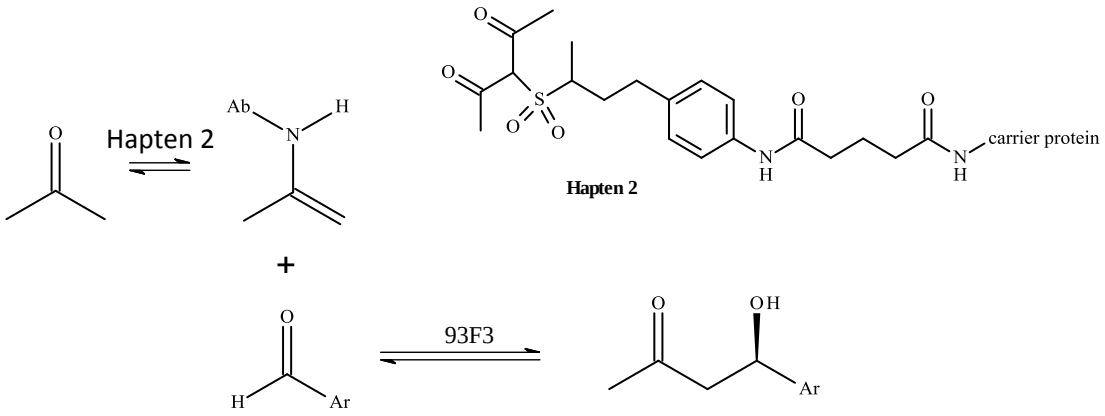



**Figure 1.10:** DFT calculations yielding the transition structures. (Adapted from Zhu et al.[49])

and antibody 33F12 and 93F3 showing the enantiopreference of 33F12 for the (*S*)-aldol product (transition structures 10 and 11) and 93F3 for the (*R*)-aldol product (transition structures 12 and 13) (Scheme 1.11).[49]




**Scheme 1.11:** Reaction of a ketone and aromatic aldehyde with antibody 93F3 generated by hapten 2.

Recent studies of catalytic aldolase antibodies have included chemically programmed antibodies (cpAbs). Antibody 38C2 was programmed to bind specifically to certain cells by reacting LysH93 with a β-diketone or acetone aldol adduct. A β-lactam is used in order to program 38C2.[59] Many other cpAbs have been developed by a similar method to the one above. These cpAbs can utilize integrins (which are receptors that bind to a cell that are important in processes occurring within the cell)[60], specifically αvβ3, αvβ6, and α5β1 for treatment of many human diseases including tumor cells for cancer.[61-62]

QM/MM/MC/FEP simulations were performed in this thesis for the aldol reaction between acetone and benzaldehyde in the catalytic aldolase antibody (33F12) to further

understand the reactivity and solvent effects associated with the aldol reaction. A detailed description of QM/MM coupled to MC and FEP is described in chapter 2.1. Results of the simulations for the Claisen rearrangement appear in chapter 3 and for the aldol reaction and catalytic antibody in chapter 4.

### 1.6 Alkanesulfonate Monooxygenase

Alkanesulfonate monooxygenase is a two-component, flavin-dependent enzyme located in *Escherichia coli*. One component, SsuE catalyzes the reduction of flavin mononucleotide (FMN) to its reduced flavin ($FMNH_2$), which is then transferred to the second component SsuD, using molecular oxygen to catalyze the cleavage of the alkanesulfonate to its corresponding aldehyde and sulfite (Figure 1.11).[63-69]

Two-component flavin-dependent enzymes catalyze a variety of processes including the biosynthesis of antibiotics, oxidation of long-chain alkanes, bioluminescence and desulfonation of sulfur containing compounds.[70] Each process includes a component available to reduce flavin, whether it be FMN or FAD (flavin adenosine dinucleotide), seen in Figure 1.12. Many of these enzymes use flavin as a substrate instead of using it as a prosthetic group.[68] Specifically, Figure 1.13 shows flavin mononucleotide monooxygenase systems which catalyze a wide variety of reactions.[70]

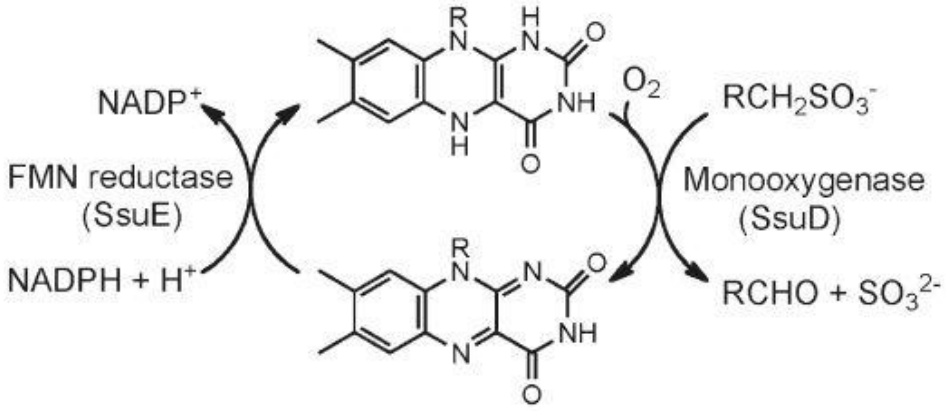



**Figure 1.11:** Desulfonation mechanism of alkanesulfonate monooxygenase. (Adapted from Carpenter et al.[67])

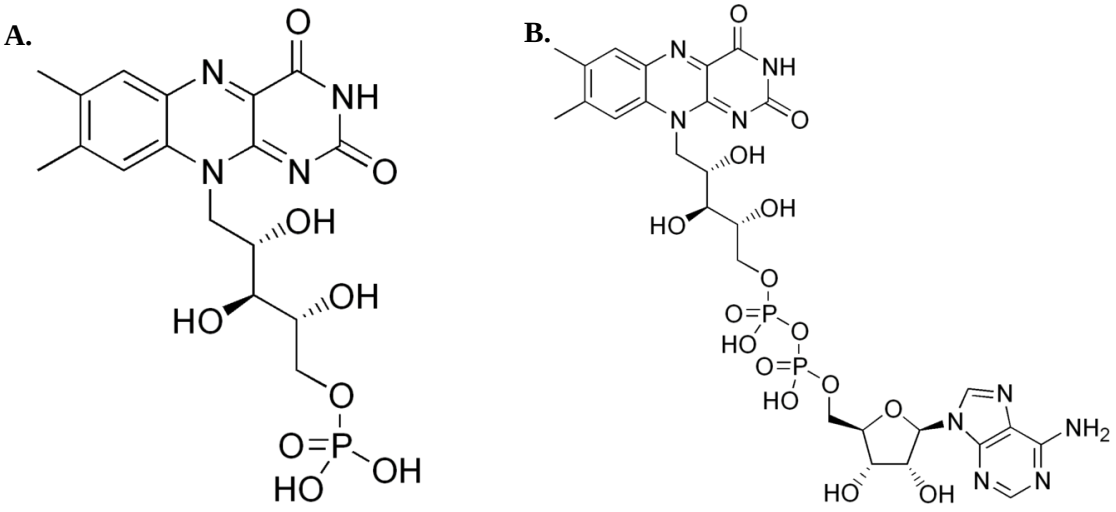



**Figure 1.12:** A: Flavin mononucleotide (FMN) and B: Flavin Adenoside Dinucleotide (FAD).

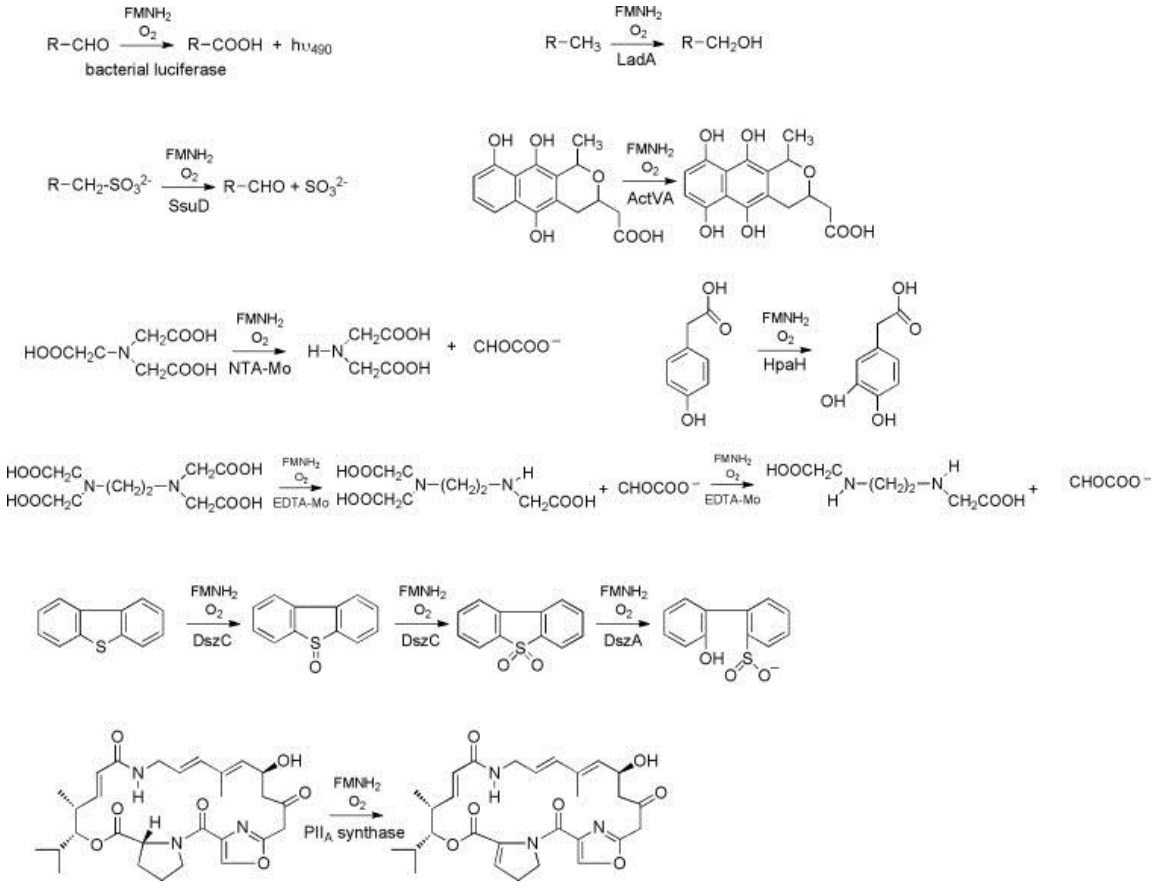



**Figure 1.13:** Flavin-dependent monooxygenase systems. (Adapted from Ellis[70])

SsuD is classified within the bacterial luciferase family, which includes bacterial luciferase and LadA, as seen in Figure 1.13. A triose phosphate isomerase (TIM)-barrel fold is found within this family, where the active site is located at the C-terminal end of the β-barrel (Figure 1.14).[63, 67] SsuD exists as a homotetramer, LadA is a homodimer with two identical subunits, and bacterial luciferase has α and β subunits in its dimer. The α subunit has an insertion region with a flexible loop, similar to the unresolved loop in SsuD. Amongst the TIM-barrel proteins, these loop regions either contribute to the active site of the enzyme or are involved in the binding of substrates.[67]

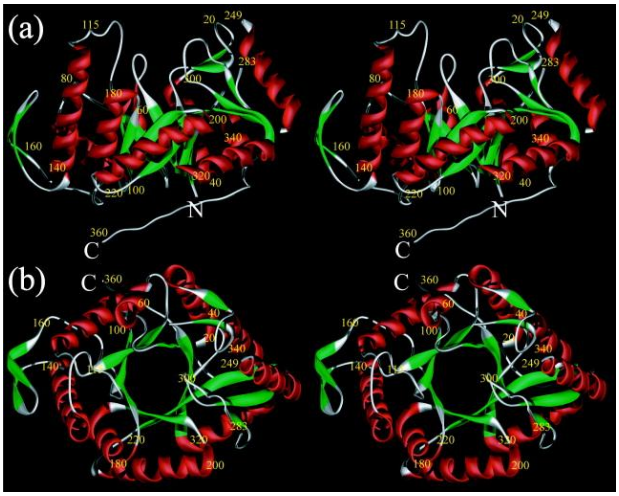



**Figure 1.14:** SsuD enzyme. (a) View perpendicular to β-barrel. (b) View along β-barrel axis. (Adapted from Eichhorn et al.[71])

Experimental studies have been performed extensively on the SsuE/SsuD two-component monooxygenase system to elucidate the properties of the enzyme.[63-69] The two component system acts as a dimer in solution. When SsuE was purified, there was no flavin bound, meaning the flavin is a substrate, and steady-state kinetic studies shows a preference for FMN over FAD. SsuE also shows preference for oxidized flavin over reduced flavin.[64, 68] The kinetic mechanism for the reduction of flavin by SsuE is seen in

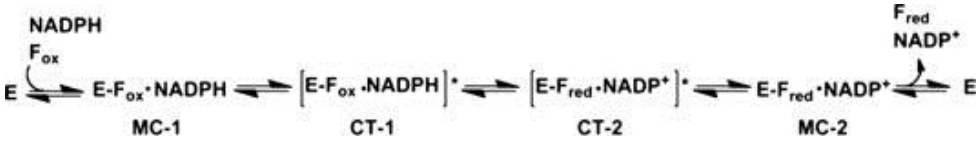



**Figure 1.15:** The kinetic mechanism for the reduction of flavin by SsuE. (Adapted from Ellis[68])

Figure 1.15, where oxidized flavin and NADPH first form a ternary complex (MC-1). Next, two charge-transfer complexes form, one with oxidized flavin (CT-1) and one with reduced flavin (CT-2). A hydride transfer occurs between CT-1 to CT-2, where a hydrogen from NADPH transfers to the flavin and reduces it. Then, CT-2 decays and the reduced ternary complex forms, where the reduced flavin and $NADP^+$ release, giving the enzyme.[68]

Several flavin mononucleotides can be utilized for biological reactions. The flavin that desulfonates alkanesulfonate to its aldehyde and sulfite is C4a-(hydro)peroxyflavin. C4a-(hydro)peroxyflavin and C4a-peroxyflavin are formed by the reaction of molecular oxygen with reduced flavin (Figure 1.16).[72] These

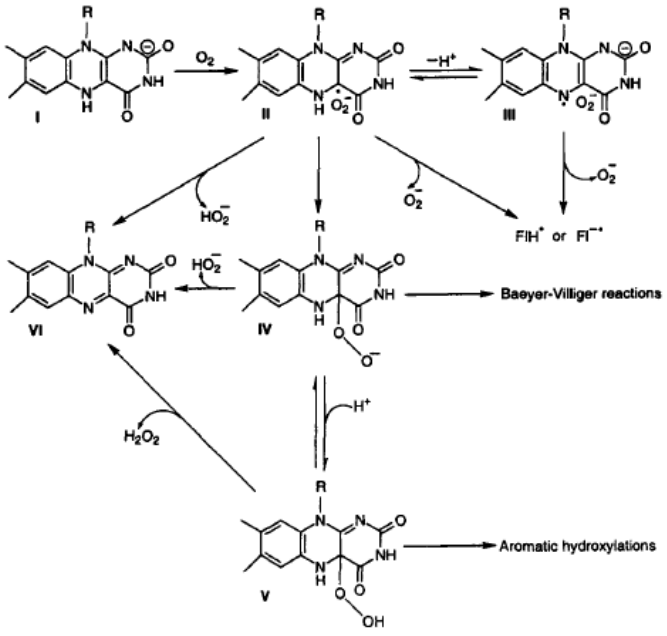
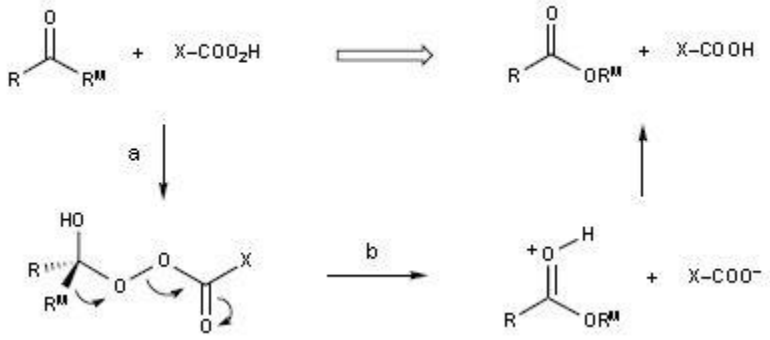
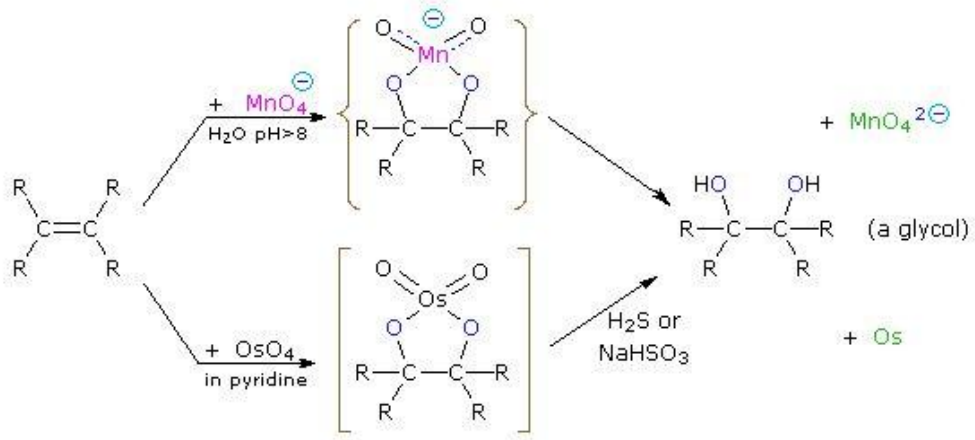



**Figure 1.16:** Mechanism for formation of C4a-(hydro)peroxyflavin and C4a-peroxy flavin intermediates. (Adapted from Massey[72])

intermediates can act as either an electrophile or a nucleophile, where electrophilic reactions occur via Baeyer-Villiger reactions (Scheme 1.12), and nucleophilic reactions occur as hydroxylation reactions (Scheme 1.13).[68, 72]




**Scheme 1.12:** General mechanism of a Baeyer-Villiger reaction. (Adapted from Strukul[73])




**Scheme 1.13:** Examples of a hydroxylation reaction. (Adapted from Reusch[74])

A Baeyer-Villiger type rearrangement occurs in one of the two mechanisms proposed by Ellis et al. describing the reaction of SsuD and C4a-(hydro)peroxyflavin and

C4a-peroxyflavin.[68] Reaction B in Figure 1.17 produces a C4a-peroxyflavin intermediate which attacks the sulfonate group of the alkanesulfonate to produce a peroxyflavin-organosulfonate intermediate. In step III, a Baeyer-Villiger like rearrangement occurs to release the sulfite group. Next, a proton is abstracted from the C1 carbon of the alkane group, which leads to cleavage of the oxygen-oxygen bond, leaving the aldehyde as a product.[68] Figure 1.17A shows the reaction of C4a-(hydro)peroxyflavin with SsuD. After the production of the C4a-(hydro)peroxyflavin intermediate, the active site base abstracts a proton from the C1 carbon of the alkanesulfonate to form a carbanion intermediate. A nucleophilic attack occurs between the carbanion intermediate and hydroperoxyflavin generating 1-hydroxyalkanesulfonate, which immediately decays to its corresponding aldehyde and sulfite.[68] The exact mechanism for sulfite production, however, is unknown.

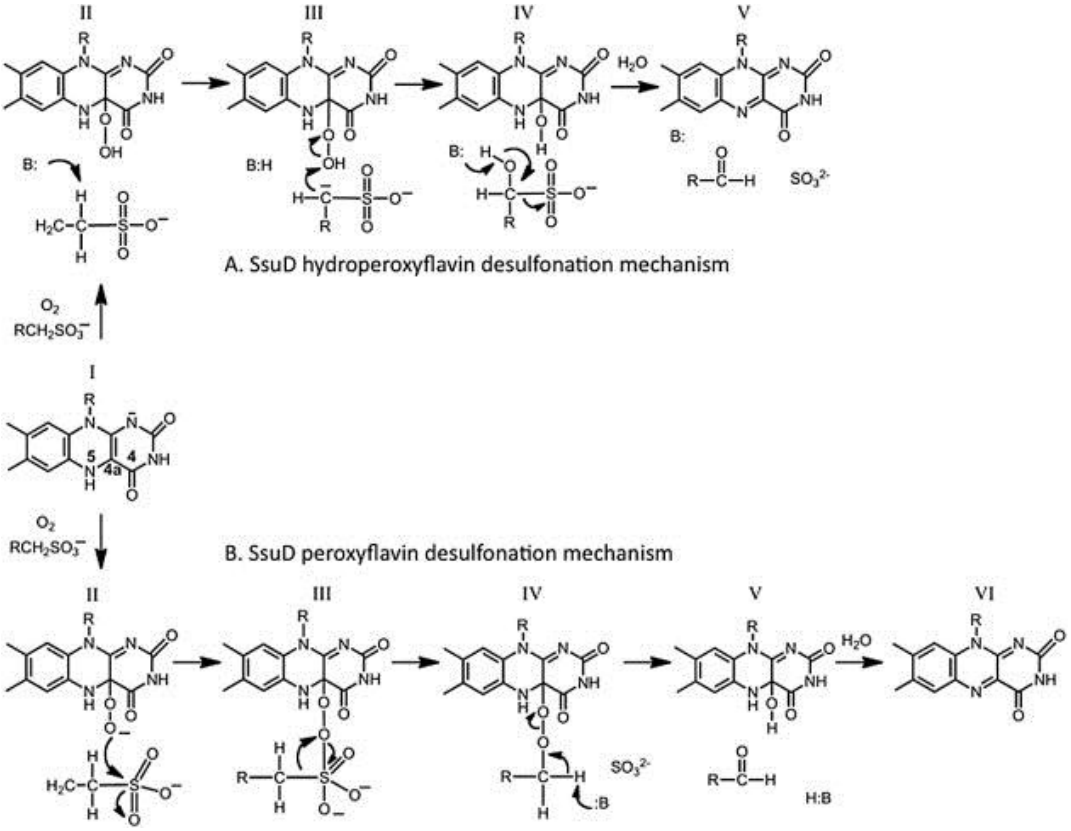



**Figure 1.17:** Catalytic mechanisms of C4a-(hydro)peroxyflavin (A) and C4a-peroxyflavin (B).

Molecular dynamics (MD) simulations have been carried out to elucidate the fluctuations of the enzyme as a response to different bound ligands and cofactors. The methodology is described in chapter 2, sections 2 and 3. Computational studies have been performed for SsuD with reduced flavin bound, SsuD with C4a-peroxyflavin bound, SsuD with octanesulfonate, SsuD with reduced flavin and octanesulfonate and SsuD with C4a-peroxyflavin and octanesulfonate. Also, a cofactor free (apo) enzyme was studied. Chapter 5 describes the results of the MD simulations.

# Chapter 2

## Methods

### 2.1 QM/MM

A combined quantum mechanical/molecular mechanical (QM/MM) approach has been applied to the Claisen rearrangement and the aldol reaction in conventional organic and aqueous solutions. The quantum mechanical (QM) aspect is the reacting species (or the solute) and the molecular mechanical (MM) portion is the solvent.[75]

The first QM/MM calculation was performed in 1976 by Warshel and Levitt.[76] They applied a QM/MM method to the entire enzyme-substrate-solvent system of the lysosome enzyme mechanism. The quantum mechanical term was restricted to only the atoms involved directly in the reaction, where the classical portion (MM region) included the energy and charge distribution of the system.[76-77] However, QM/MM methodology did not take off until the early 1990's, when the AM1 and MNDO semiempirical QM methods, coupled to the CHARMM force field were used to compare the results of condensed phase reactions to experimental and ab initio data.[77-78] Since then, many researchers have applied the QM/MM methodology to various substrate-solvent systems including small organic systems in solvent, proteins, and also, spectroscopic properties. Many different QM methods can be used, including semiempirical methods, density functional theory (DFT) methods, and post Hartree-Fock ab initio methods (Møller-Plesset perturbation (MP2) or coupled cluster theory). Force fields are used for the MM

portion and include the OPLS-AA force field, the AMBER force fields, or the CHARMM force fields.[77]

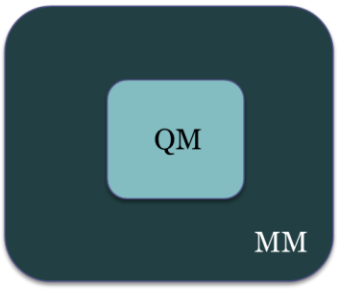



**Figure 2.1:** The QM region represents the bond making/bond breaking portion, where the MM represents the remainder of the system.

A schematic representation of the QM/MM method is seen in Figure 2.1. The QM region is where the bond making and bond breaking occurs, and the rest of the system is treated in the MM region.[79-80] The total energy ($E_{TOT}$) of the system can be represented by

$$E_{TOT} = E_{QM} + E_{MM} + E_{QM/MM} \tag{2.1}$$

where $E_{QM/MM}$ is the interaction energy between the QM and MM regions. When a solute is treated by QM and solvent by MM, the Hamiltonian for a QM/MM calculation is shown below,

$$H_{QM/MM} = -\sum_i \sum_m \frac{q_m}{r_{im}} + \sum_k \sum_m \left[ \frac{z_k q_m}{r_{km}} + 4\varepsilon_{km} \left( \frac{\sigma_{km}^{12}}{r_{km}^{12}} - \frac{\sigma_{km}^{6}}{r_{km}^{6}} \right) \right] \tag{2.2}$$

where *i* represents a QM electron, k is a QM nucleus and m is a MM nucleus with a partial charge of $q_m$. The first term in the Hamiltonian represents the electrostatic interaction between the MM nuclei and the QM electron. The Coulombic interactions of the QM nucleus and MM charges are represented by the second term, whereas the final term represents the 12-6 Lennard-Jones potential which is given in the MM force field.[79-81] $z_\alpha$ represents the nuclear charge k, $r_{im/km}$ is the interatomic distance between atoms i and m and k and m, ε is the well depth of the Lennard-Jones potential and σ is the collision diameter between atoms k and m.[79-80]

The total energy of the system is not as simple as the summation of the energies from the QM region and the energies from the MM region. The interaction energy of the QM and MM region must also be accounted for.[77] There are many procedures to account for all aspects but the procedure represented throughout the thesis is shown below,[82]

$$E_{QM/MM} = \sum_{a} E_{as} \tag{2.3}$$

$$E_{as} = \sum_{i}^{on\ s} \sum_{j}^{on\ a} \left[ \frac{q_i^* q_j e^2}{r_{ij}} + 4\epsilon_{ij} \left( \frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}} \right) \right]. \tag{2.4}$$

The interaction of the QM solutes and the classical solvent molecules represented by MM is used to calculate the QM/MM energy, where the summation is taken over atoms "i" of the solute (s) and "j" of the solvent (a). The partial charges, $q_i^*$, are calculated from the CM1A charge model of the AM1 wavefunction

$$q_i^* = \alpha q_i^{CM1A}, \tag{2.5}$$

where α represents the scaling factor of the charge model.[82] The AM1-CM1A charge model was developed by Truhlar et al. to yield more accurate partial charges. The dipole moments are modeled to reduce the average error to 0.3D for 195 neutral molecules in the gas phase.[82-83] The Jorgensen group then applied the CM1A charge model in QM/MM calculations to calculate free energies of hydration. The best free energies of hydration were found for 38 organic molecules when the CM1A charges were scaled by 1.14.[82, 84] The Lennard Jones potential, $4\varepsilon_{ij}\left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}}\right)$ is also included in equation 2.4, where σ represents the collision diameter of the atoms, ε, the well depth of the Lennard Jones potential, and $r_{ij}$, the distance between atoms i and j. Figure 2.2 represents the Lennard Jones potential, where the leftmost portion of the graph represents the repulsion of atoms *i* and *j*, the lowest energy represents the equilibrated bond length of atoms *i* and *j*, and the rightmost portion the separation of atoms *i* and *j*.

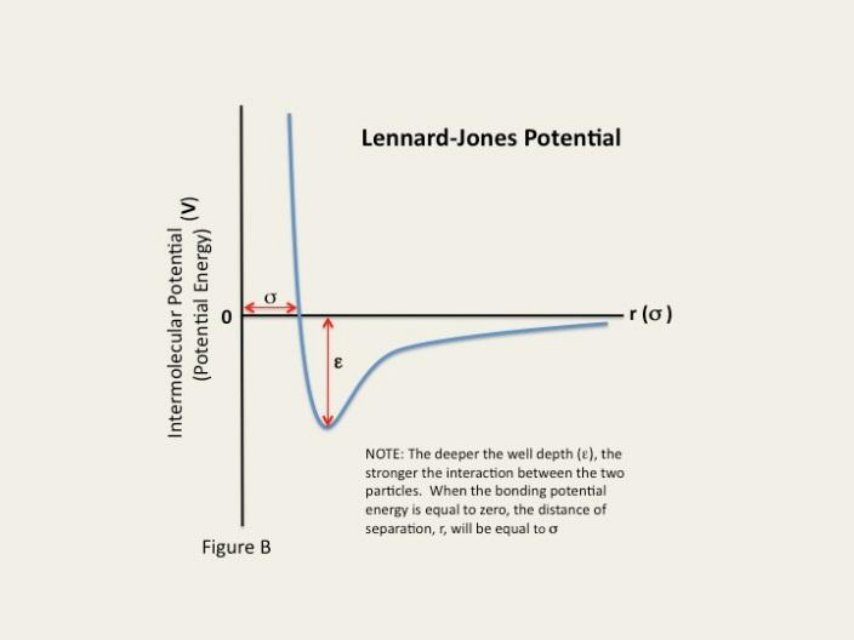



**Figure 2.2:** Lennard-Jones Potential graphical representation. (Adapted from Naeem[85])

Another method used for the interaction of the QM and MM region for large systems (e.g. proteins) is link atoms (Figure 2.3), which is used for the aldolase antibody simulations in chapter 4. Link atoms are used when the MM region is covalently bound to the QM region. Additional terms are added to compute the interaction energies between the QM and MM region where the bonds are cut by the boundary between the QM and MM region. This can be accomplished by a few techniques:

1. The Hamiltonian of the QM region ($H_{QM}$) is calculated at every bond with hydrogen involved in the cut bonds.
2. The MM bond stretching term is used to calculate the energies of the cut bonds, where a large force constant is imposed so the bond angle for MM atom, QM atom, and capped H remains zero.

3. Angle bending is calculated with the MM angle bending formula between two MM atoms and one QM atom.
4. Torsions are computed using standard MM formulas when dealing with one or two QM atoms and two or three MM atoms.
5. Charges on nearby atoms of the cut bonds can be zeroed, scaled or kept.
6. All other terms in the Hamiltonian of the MM and QM/MM region are calculated the same.

The boundaries are chosen to not include polarizable bonds, so the choice is typically a C-H bond due to their similar electronegativities.[79]

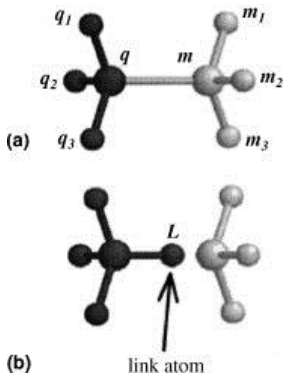



**Figure 2.3:** Description of the link atom for ethane, where two methyl fragments are generated and link atom (L) replaces the q-m bond (Adapted from Wang et al.[86]).

### 2.1.1. PDDG/PM3

The QM method primarily used in our work is the PDDG/PM3 semiempirical method developed by Jorgensen et al..[87-89] PDDG/PM3 is a modified version of the PM3

and MNDO semiempirical method with additional pairwise core repulsion functions.[87] PDDG/PM3 is an NDDO (neglect of diatomic differential overlap) method, which the NDDO methods are found to have large mean absolute errors (MAE) in their heats of formation for many closed-shell organic molecules. To correct for the MAE, the bond and group equivalents (BGE) method was used in PDDG/PM3,

$$\Delta H_f = E_{mol} + T/R + \sum_j a_j n_j. \quad (2.6)$$

where the molecular energy of the system is converted to the heat of formation by the addition of T/R, translation/rotational correction term (at 298K for nonlinear molecules T/R=4RT) along with the number of equivalents in a system ($a_j$) and the number of times they occur ($n_j$). BGE's consist of the energy occupying the atoms bound to each other and the alterations in the atom's chemical environment.[87, 90]

The addition of the BGE scheme to semiempirical methods vastly improved the MAE's, however it is unable to be added to mobile molecules on the reaction surface. In order to resolve this problem, Repasky et al. chose to modify the core repulsion function (CRF) of the semiempirical method,

$$CRF_{AM1,PM3}(A,B) = CRF_{MNDO}(A,B) + \left(\frac{Z_A Z_B}{R_{AB}}\right)\left[\sum_k a_{kA} e^{-b_{kA}(R_{AB}-c_{kA})^2} + \sum_k a_{kB} e^{-b_{kB}(R_{AB}-c_{kB})^2}\right]. \quad (2.7)$$

The summations were taken over k, which represents the number of Gaussians per element. These Gaussians consist of three parameters ($a_k$, $b_k$, $c_k$) that were added to the CRF of the MNDO semiempirical method. This modification of the CRF's is unable to differentiate between various functional groups, due to the Gaussians being held too closely to parent atoms.[87]

In 2000, Bernal-Uruchurtu et al. proposed a modified core repulsion function to improve the hydrogen bonding and energetics of molecular systems.[91-92] The CRF was modified for intermolecular atom pairs by the addition of α, β, χ, δ, and ε which are dependent on the atom types of atoms A and B.[90-91]

$$CRF = CRF_{MNDO} + \sum_{A,B}^{inter} \alpha_{AB} e^{-\beta_{AB} R_{AB}} + \frac{\chi_{AB}}{R_{AB}^6} + \frac{\delta_{AB}}{R_{AB}^8} + \frac{\varepsilon_{AB}}{R_{AB}^{10}}. \quad (2.8)$$

The Bond and Group Equivalents scheme also highlighted the importance of the atomic eisol parameter, which represents a gaseous atom's electronic energy of formation to give correct heats of formation. The eisol parameter is derived so it's dependent on "all one-center parameters within the semiempirical formalism."[87, 90] By varying the

functional groups, the eisol values became flawed, so within the BGE scheme, these values were reevaluated for organic atoms.[87] By reevaluating the eisol parameters and modifying the core repulsion function to include functional group information, the PDDG/PM3 semiempirical method was developed,

$$PDDG(A,B) = \sum_{A}\sum_{B>A}\frac{1}{n_A + n_B} x \left[\sum_{i=1}^{2}\sum_{j=1}^{2}\left(n_A P_{A_i} + n_B P_{B_i}\right)\exp\left(-10\text{Å}^2\left(R_{AB} - D_{A_i} - D_{B_j}\right)^2\right)\right]. \quad (2.9)$$

This formalism represents the sum of all atoms, looped over atoms A and B. Distance terms, $D_{A_1}$ and $D_{A_2}$ and preexponential factors, $P_{A_1}$ and $P_{A_2}$ compose the optimizable parameters for the above PDDG function. The PDDG function is weighted by the number of valence electrons on atoms A and B, $n_A$ and $n_B$, respectively. The NDDO methods can be improved by adding the above equation to the core repulsion functions of these methods.[87]

Another important aspect in developing a semiempirical method is including a set of parameters to reproduce experimental data. The two methods used to develop these parameters are reoptimization of the PM3 parameters to reduce overall error and using chemical knowledge to produce an accurate method by only optimizing a small set of parameters, which is similar to the AM1/MNDO method. These reoptimized parameters were chosen from the Dewar method parameters, including the Gaussian preexponential

and distance terms, along with the new parameters for the PDDG methodology. S represents the error function, which is optimized for every parameter to give a new

$$S = \sum_i w_i^2 \left(q_i^{calc} - q_i^{exp}\right)^2 + \sum_j w_j \frac{\delta E}{\delta x_j} + (\chi^2)^4 \quad (2.10)$$

parameter set, where $q_i$ represents the heats of formation, geometries including bond length, angle and dihedrals, ionization potentials and dipole moments. The second term represents the energy gradients with respect to their geometries$\left(\frac{\delta E}{\delta x_j}\right)$, weighting factors (w) , and the multiple regression fit ($\chi^2$) which generates eisol parameters. The step penalty given to data sets that generate poor energetic results is consequence of taking $\chi$ to the fourth power.[87] Repasky et al. calculated heats of formation and mean absolute errors for 622 organic molecules giving reduced MAE's for many different functional groups.[87]

### 2.1.2 Polarized Continuum Model (PCM)

The PCM method was developed by Miertus et al. to calculate the electrostatic interaction between a solute and a continuous polarizable medium.[93] The solute's molecular Hamiltonian is used, including the *ab initio* description of the solute, with the addition of point charges on the cavity wall of the solute. The wall is divided into surfaces, and each surface is given its own point charge. These point charges represent the polarization of the solvent, which are proportional to the magnitude of the electric field gradient at each surface.

$$\phi(r) = \phi_p(r) + \phi_\sigma(r) \tag{2.11}$$

represents the total electrostatic potential of each surface, where $\phi_p(r)$ is the potential from the solute, and $\phi_\sigma(r)$ is the potential from the charges on the surface.[80]

The surface cavity can be determined from the van der Waals radii of the atoms, where each fraction of the radii contributes to the cavity wall. The cavity surface is then divided to calculate the surface area of the wall. The point charge on each surface is determined by,

$$q_i = -\left[\frac{\varepsilon - 1}{4\pi\varepsilon}\right] E_i \Delta S \tag{2.12}$$

where ε represents the dielectric constant of the solvent, $E_i$ is the electric field gradient, and ΔS is the surface area of the divisions of the surface cavity.[80, 93] Execution of equation 2.12 produces an electrostatic potential, $\phi_\sigma(r)$, which is incorporated into the solute's Hamiltonian to determine the polarization of the solvent.[93] This method is continued until self-consistency is achieved. Variations of PCM have been adapted, including IEF-PCM, which eases the computation of gradients, allows for infinite periodic boundaries, and the ability to calculate liquid/liquid and liquid/vapor properties at the interface.[79, 94-96] Alternative ways to determine the surface cavity have also been developed, including CPCM. CPCM uses an infinite dielectric constant, where the surface cavity is in a conductor.[80, 97-99]

### 2.1.3 OPLS

The OPLS-AA force field describes the MM portion of the research performed in our work. There are two types of OPLS force fields, the united atom and the all atom. The united atom (UA) force field has sites for non-bonded interactions on all atoms, not including hydrogens except for those attached to phenyl carbons or heteroatoms. The all atom (AA) OPLS force field allows for more flexibility than the UA force field by addition of sites to all atoms. The all atom OPLS force field is more computationally expensive, but generally improves results.[100]

The OPLS-AA force field is represented by

$$E(\phi) = E_{bond}(\phi) + E_{angle}(\phi) + E_{n.b.}(\phi) + E_{torsion}(\phi), \tag{2.13}$$

where $E_{bond}$ is represented by,

$$E_{bond} = \sum_{bonds} K_r\left(r - r_{eq}\right)^2 \tag{2.14}$$

and $E_{angle}$ by,

$$E_{angle} = \sum_{angles} K_\theta\left(\theta - \theta_{eq}\right)^2. \tag{2.15}$$

$K_r$ and $K_\theta$ represent the force constants for the bond length and angles and r and θ represent the respective bond lengths and angles, respectively. These parameters were mainly taken from the AMBER all atom force field, except for some alkanes that were developed using the CHARMM/22 force field. The non-bonded terms are a combination

$$E_{ab} = \sum_{i}^{on\ a} \sum_{j}^{on\ b} \left[ q_i q_j e^2 / r_{ij} + 4\epsilon_{ij} \left( \sigma_{ij}^{12} / r_{ij}^{12} - \sigma_{ij}^{6} / r_{ij}^{6} \right) \right] f_{ij} \tag{2.16}$$

of Coulomb and Lennard-Jones terms. $E_{ab}$ represents the interaction energy between two atoms, a and b. $\sigma_{ij}$ and $\epsilon_{ij}$ follow the standard combining rules, where $\sigma_{ij} = \left(\sigma_{ii}\sigma_{jj}\right)^{1/2}$ and $\epsilon_{ij} = \left(\epsilon_{ii}\epsilon_{jj}\right)^{1/2}$. The scaling factor, $f_{ij}$ is typically 1.0, except for 1,4-intramolecular interactions, where $f_{ij}$ is 0.5. These parameters were adapted from the OPLS-UA force field. The torsional term is represented by equation 2.17, where $\varphi_i$ is the

$$E_{torsion} = \sum_{i} \frac{V_1^i}{2} [1 + \cos(\phi_i + f_i 1)] + \frac{V_2^i}{2} [1 - \cos(2\phi_i + f_i 2)] + \frac{V_3^i}{2} [1 + \cos(3\phi_i + f_i 3)]. \tag{2.17}$$

dihedral angle, $V_1$, $V_2$, and $V_3$ are the Fourier series coefficients, and f1, f2, and f3 are the phase angles for the system. The OPLS force field charges were determined empirically to replicate the charges of organic liquids.[100]

### 2.1.4 Free Energy Perturbation

Free energy perturbation (FEP) is used to calculate important thermodynamic parameters for chemical reactions.[101-102] FEP was introduced in 1954 by R. W. Zwanzig introducing the FEP equations, 2.18 and 2.19.

$$\Delta F = F_1 - F_0 = -k_b T \ln \langle exp \left[ \frac{-(E_1 - E_0)}{kT} \right] \rangle_0, \tag{2.18}$$

$$\Delta F = F_1 - F_0 = \langle V \rangle - \frac{1}{2k_b T}(\langle V^2 \rangle - \langle V \rangle^2) \tag{2.19}$$

The free energy (F) is related between the initial and final states of a system and an average of the energy changes between the two states.[101-103] In the first equation, the subscript, 0, represents averaging over the initial, unperturbed configurations of the system.[80] Zwanzig represented free energies as a power series, and generation of equation 2.19 developed, where $V^2$ is a perturbing potential generated by making the equation 2.18 second order.[103]

The type of free energy that is the function of a coordinate is the potential of mean force (PMF), which is defined as the surface of the free energy along a chosen coordinate.[80] It was first introduced in 1935 by J.G. Kirkwood and is defined as

$$\mathcal{W}(\xi) = \mathcal{W}(\xi^*) - k_B T \tag{2.20}$$

where $\xi^*$ and $\mathcal{W}$ ($\xi^*$) are constants, and

$$\langle \rho(\xi) \rangle = \frac{\int dR \delta\left(\xi'^{[R]} - \xi\right) e^{-U(R)/k_B T}}{\int dR e^{-U(R)/k_B T}}. \tag{2.21}$$

$\rho(\xi)$ is the average distribution function used to generate equation 2.20, which is developed from a Boltzmann weighted average. U(R) represents the total energy that is dependent on the coordinates, R, $\xi$ is the coordinate the PMF travels across, and the function $\xi$'(R) depends on the degrees of freedom in the system.[80, 104-105] $\delta\left(\xi'^{[R]} - \xi\right)$ represents the Dirac delta function for the coordinate, $\xi$.

The change in free energy equation from above can represent the chemical equilibria of a system where the two states of the system can be two separate molecules in two different media. Figure 2.4 is a representation of a thermodynamic cycle for

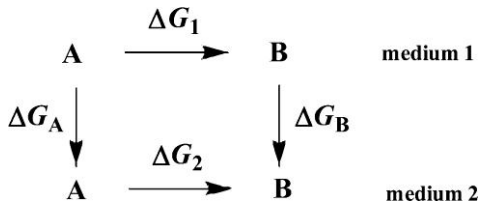



**Figure 2.4:** Thermodynamic cycle for molecules A and B in two different solvents. (Adapted from Jorgensen et al.[101])

perturbing molecule A to B in two solvents. Free energies are associated with each perturbation, and the overall free energy is calculated by,

$$\Delta\Delta G = \Delta G_2 - \Delta G_1 = \Delta G_B - \Delta G_A \tag{2.22}$$

where ΔΔG is the overall free energy calculated by the difference in the ΔG's of the two different peturbations.[101]

Free energy perturbation necessitates that the two systems are similar, otherwise the averaging process of the change in free energy equation will be extremely slow. In order to ensure an easy transition between states, a coupling parameter (χ) is introduced to scale parameters for the force field and geometries between states A and B.[101-102]

$$\chi_i = \lambda_i \chi_B + (1 - \lambda_i)\chi_A \tag{2.23}$$

When the two states are dramatically different, multiple simulations can be run for various values of $\lambda_i$ that are given values between the two states, 0 and 1. Double wide sampling is the method of choice, which is a type of sampling run in two directions,

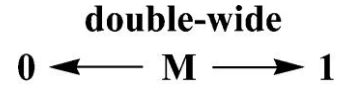



**Figure 2.5:** Representation of double wide sampling. (Adapted from Jorgensen et al.[101])

$\lambda_i \rightarrow \lambda_{i+1}$ and $\lambda_{i+1} \rightarrow \lambda_i$. The Zwanzig equation for free energy changes, equation 2.18, can be transformed into a new free energy equation for double wide sampling,

$$\Delta G^{DW} = kT \ln[\langle \exp(-(E_0 - E_M)/kT)\rangle_M / \langle exp(-(E_1 - E_M)/kT)\rangle_M]. \quad (2.24)$$

Double wide sampling requires a defined *M*,

$$E_M = \lambda_M E_1 + (1 - \lambda_M)E_0, \quad (2.25)$$

where $\lambda_M$ represents a parameter chosen to reduce the random error.[101]

### 2.1.5 Monte Carlo

Monte Carlo (MC) simulations are performed with each QM/MM calculation, where configurations of a system are created by making changes to the positions and orientations of the system.[80] The most widely-used description of Monte Carlo was developed in 1953 by Metropolis et al., with their application to a rigid-sphere system of 224 particles.[106-107] The MC method chooses its movements from a 3N dimensional space of positions, and in turn calculates potential energies of each of these positions.

Due to the large number of configurations that give an almost zero Boltzmann factor, Metropolis et al. developed their method to only generate configurations that have a large impact to the potential energy. The basic procedure for the Metropolis Monte Carlo (MC) method is the following: the result of the run is based only on the previous run and that each run is within a limited set of results. MC samples the whole system, solute and solvent included.

To implement MC, a new configuration is created by changing the Cartesian coordinates and/or internal coordinates within the structure and calculating the energy of the new system configuration. If the energy of the new configuration is lower than the previous configuration, then it is saved, but if it's higher, the Boltzmann factor, $\exp(-\Delta\mathcal{V}/k_B t)$ of the structure is compared to a random number between 0 and 1. If the Boltzmann factor is greater than the random number, then the structure is kept, but if it's lower, the structure is rejected and the preliminary configuration is saved for the next MC step,

$$rand(0.1) \leq \exp(-\Delta\mathcal{V}(r^N)/k_B t). \tag{2.26}$$

The size of each configurational change can be adjusted by manipulating the maximum displacement, $\delta r_{max}$. $\delta r_{max}$ was chosen to accept 40% of the total moves of the system.[80]

### 2.1.6 Periodic Boundary Conditions

Another method introduced in the 1953 Metropolis paper was the use of periodic boundary conditions. A periodic boundary condition is applied in order to minimize the

complexity of a simulation by using a small number of particles and treating them as if they were in a bulk fluid.[79-80, 106] Each box has a cutoff distance for evaluating non-bonded interactions. Figure 2.6 represents a 3x3x1 periodic box where the central box has eight identical neighboring boxes. If a solute or solvent molecule leaves the central

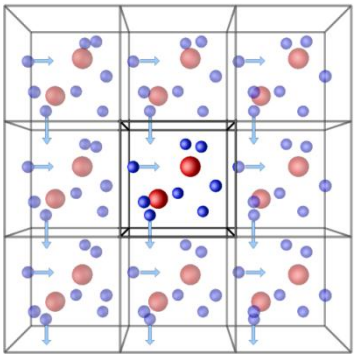


**Figure 2.6:** Periodic boundary conditions in two dimensions. (Adapted from LeRoux[108])

box, it is replaced by one from a surrounding box, so the central box stays constant through the simulation.[80] To find the coordinates of each particle, integral multiples of the box sides can be added or subtracted.[80] Different types of periodic boundary conditions are available, including cubic, hexagonal prism, truncated octahedron, rhombic dodecahedron and an elongated dodecahedron, and is chosen based on the geometry of the system of interest.[80]

## 2.2 Molecular Dynamics

Molecular dynamics (MD) simulations have been performed on the alkanesulfonate monooxygenase system described in chapter 1, section 1.6 to understand the dynamics of the enzyme. MD was run using components of the AMBER biomolecular package. MD simulations solve Newton's equations of motion for

molecules in a system.[79-80, 109-111] For a simple harmonic oscillator, Newton's equations of motion can be represented by a position vector (q), and a momentum vector (p),[79]

$$q(t_2) = q(t_1) + \int_{t_1}^{t_2} \frac{p(t)}{m} dt \tag{2.27}$$

$$p(t_2) = p(t_1) + m \int_{t_1}^{t_2} a(t) dt. \tag{2.28}$$

The relationship of two positions is represented by equation 2.27, where $p(t)/m$ denotes the velocity, t is the time, and m is the particle mass. Acceleration, as seen in equation 2.28, is used to demonstrate the relationship of two momentum vectors, where $a(t) = F/m$ and F is the force acting upon the vectors.[79,80] The force and position can be related by,

$$F_q = -\frac{\partial U}{\partial q} \tag{2.29}$$

where U represents the potential energy.[79] Analytical solutions can be found for the position vector and momentum vector by solving for a simple form of the potential energy of vibrations,

$$U(r_{AB}) = \frac{1}{2} k_{AB} \left(r_{AB} - r_{AB,eq}\right)^2. \tag{2.30}$$

The symbol k represents the second derivative of U, or the force constant of a spring, $r_{AB}$ is the bond length and $r_{AB,eq}$ represents the minimum bond length. The above equations are used when periodic boundary conditions are applied to the system being studied.[79]

Harmonic oscillators are simple systems, so when systems become more complex, the Euler approximation is used to transform equations 2.27 and 2.28 into ones for more complicated systems:

$$q(t + \delta t) = q(t) + \frac{p(t)}{m}\delta t \tag{2.31}$$

and

$$p(t + \delta t) = p(t) + ma(t)\delta t. \tag{2.32}$$

These equations are used for simulation of a trajectory, which trace the atoms positions over time, as seen in Figure 2.7.[112]

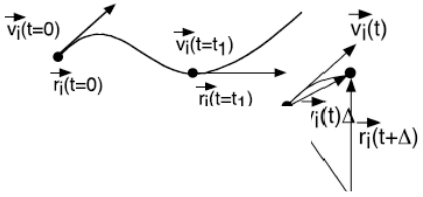



**Figure 2.7:** Atomic positions in a 3-dimensional space. (Adapted from Nakano[112])

Typically, the force will change when the position of the particle changes or when other particles interacting with the particle of interest changes.[80] In order to calculate the position, a Taylor expansion is used,

$$f(x_0 + h) = \sum_{n=0}^{\infty} \frac{h^n}{n!} \frac{d^n f}{dx^n}\Big|_{x=x_0} \tag{2.33}$$

$$= f(x_0) + h\frac{df}{dx}\Big|_{x=x_0} + \frac{h^2}{2}\frac{d^2 f}{dx^2}\Big|_{x=x_0} + \frac{h^3}{3!}\frac{d^3 f}{dx^3}\Big|_{x=x_0} + \cdots$$

which gives a function based off its derivative, $x_0$, at its distance from its atomic coordinate, $x_0$+h.[112] Euler's approximation is not a stable method to simulate a trajectory, due to a very small time step, so Newton's laws of motion must be integrated numerically.[80]

### 2.2.1 Integration Algorithms

The Verlet algorithm is one of the most accepted algorithms for integrating Newton's equations of motion in molecular dynamics simulations.[80,113] The new positions are determined from the positions of the previous step, $q(t - \delta t)$ and the positions and acceleration at time t,

$$q(t + \delta t) = 2q(t) - q(t - \delta t) + \delta t^2 a(t). \tag{2.34}$$

To calculate the velocities, two different methods can be used.[80] The first is to divide the difference in positions by $2\delta t$,

$$v(t) = [q(t + \delta t) - q(t - \delta t)]/2\delta t, \tag{2.35}$$

or by estimating them,

$$v\left(t+\tfrac{1}{2}\delta t\right)=[q(t+\delta t)-q(t)]/\delta t. \tag{2.36}$$

In equation 2.36, the velocity is estimated by using the half-step, $t+\frac{1}{2}\delta t$.

The Verlet algorithm has a few major disadvantages. Precision may be lost due to the addition of a small acceleration term $\delta t^2 a(t)$ to much larger terms, $2q(t)$ and $q(t-\delta t)$ in equation 2.32. This equation does not include the velocity, making it difficult to calculate. Also, the velocity for the previous step is not known until the new position of the next step is calculated. In order to calculate the position at the initial step, the Euler approximation for positions and momentum, represented in equations 2.31 and 2.32, is used.[79,80,112]

The leap frog algorithm was introduced in 1970 by R. W. Hockney to include the velocity explicitly into the Verlet algorithm.[114] The position vector, extended by the Taylor expansion, is shortened at the second order giving,

$$q(t+\delta t)=q(t)+v\left(t+\tfrac{1}{2}\delta t\right)\delta t \tag{2.37}$$

for positions, and

$$v\left(t+\tfrac{1}{2}\delta t\right)=v\left(t-\tfrac{1}{2}\delta t\right)+a(t)\delta t \tag{2.38}$$

for velocities.[79-80] Initially, the velocities are calculated from the velocity and acceleration at time, t. The positions are acquired from the previously calculated velocities (v), which, at time (t), can be calculated via,

$$v(t) = \frac{1}{2}\left[v\left(t + \frac{1}{2}\delta t\right) + v\left(t - \frac{1}{2}\delta t\right)\right]. \quad (2.39)$$

While the leap frog algorithm includes the velocities explicitly unlike the Verlet algorithm, the velocity and position calculations aren't coordinated. Also, the kinetic energy value in the total energy cannot be calculated at the same time the positions are computed.[80] By only including the second-order Taylor expansion terms, the integration of the numerical term of the trajectory becomes less stable, meaning the Verlet algorithm is unreliable.[79]

Since the leap-frog method doesn't calculate the vectors easily, Swope, Andersen, Berens and Wilson developed the velocity-Verlet algorithm in 1982.[115-116] The positions, equation 2.40, and velocities, equation 2.41 are calculated by,

$$q(t + \delta t) = q(t) + \delta t v(t) + \frac{1}{2}\delta t^2 a(t) \quad (2.40)$$

$$v(t + \delta t) = v(t) + \frac{1}{2}\delta t[a(t) + a(t + \delta t)]. \quad (2.41)$$

The positions at time, $t + \delta t$ are calculated with equation 2.40 while the velocities at an intermediate step are found using,

$$v\left(t + \frac{1}{2}\delta t\right) = v(t) + \frac{1}{2}\delta t a(t). \quad (2.42)$$

At time, $t + \delta t$, the forces and accelerations are calculated then the velocity of the new move is calculated via,

$$v(t+\delta t) = v\left(t+\tfrac{1}{2}\delta t\right) + \tfrac{1}{2}\delta t a(t+\delta t). \tag{2.43}$$

The kinetic and potential energies are both available at the end of the velocity-Verlet algorithm.[115-116] Figure 2.8 demonstrates the forms of the Verlet algorithm, including the original method, the leap frog method and the velocity form, where the kept variables are in the gray boxes.[116]

Stability of numerical integration is shown in the Runge-Kutta method and the Gear predictor-corrector method.[79,80,116,117-118] The Runge-Kutta method evaluates the gradient of the function at intermediate points before moving onto the next point in the trajectory.[79,118] The higher-order terms from the Taylor expansion are included in the Gear predictor-corrector method in order to predict the next step of the trajectory. Positions are corrected from the change in the actual and predicted steps.[79,117] These methods are extremely stable, but often not used for MD simulations due to their high computational cost.[79]

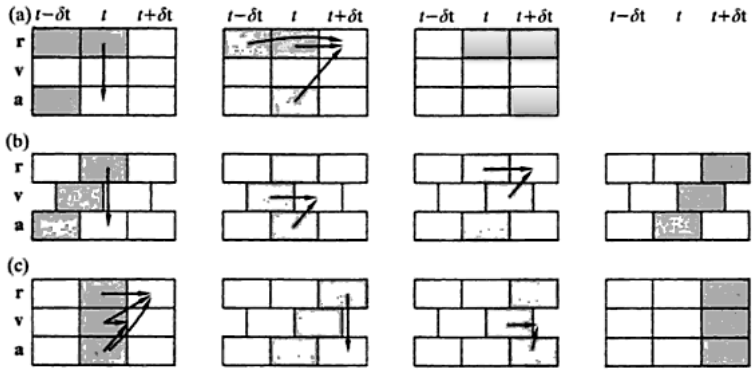



**Figure 2.8:** The forms of the Verlet algorithm. (a) Original. (b) Leap frog method. (c) Velocity Verlet algorithm. (Adapted from Allen et al.[116])

A time step is used for molecular dynamics simulations, which is represented by δt. The time step needs to be carefully chosen or errors may arise in not selecting enough of the phase space for the trajectory or with the integration algorithms. Typically, a time step is chosen to be roughly one-tenth the length of the shortest motion of the system for flexible systems, but for longer time steps this causes a problem. To remedy this problem, some degrees of freedom are constrained that have a minimal effect on the fluctuations of the system.[80]

### 2.2.2 SHAKE Algorithm

The SHAKE algorithm introduced in 1977 by Ryckaert, Ciccotti and Berendsen reduces the highest frequency systems by limiting the degrees of freedom.[79-80,119] The degrees of freedom (motions of molecules) with the highest frequencies are constrained by "fixing" the internal coordinates during the simulation. The type of constraint used by the SHAKE algorithm is the holonomic constraint (k), defined by,

$$f(q_1, q_2, q_3, \dots, t) = 0 \tag{2.44}$$

where $q_1, q_2, q_3$ represent the particle's coordinates which are independent of one another. In a system of N particles, the motions represent 3N degrees of freedom. The system is reduced to 3N-k degrees of freedom, when the holonomic constraints are added.[80]

The forces in Newton's equations of motion are manipulated when a constrained system is present. These forces include the ones representing the intra and intermolecular interactions and the forces corresponding to the constraints.

$$F_{ckx} = \lambda_k \frac{\delta \sigma_k}{\delta x} \tag{2.45}$$

represents the constraint force for a system in the x direction, where atoms *i* and *j* are a fixed bond. Here, $\lambda_k$ is the Lagrange multiplier, $\sigma_k$ is the constraint and *x* is a Cartesian coordinate of atom *i* or *j*.[80] The constraint ($\sigma_{ij}$) is represented as[80],

$$\sigma_{ij} = \left(r_i - r_j\right)^2 - d_{ij}^2 = 0, \tag{2.46}$$

where $d_{ij}$ is the distance constraint between atoms *i* and *j* and r represents the position constraints.[80] The force represented by equation 2.45 can be introduced into Verlet's algorithm as,

$$r_i(t + \delta t) = 2r_i(t) - r_i(t - \delta t) + \frac{\delta t^2}{m_i} F_i(t) + \sum_k \frac{\lambda_k \delta t^2}{m_i} r_{ij}(t) \tag{2.47}$$

which is used to solve for the Lagrange multiplier ($\lambda_k$) that will satisfy all constraints.[79-80]

### 2.2.3 AMBER Biomolecular Package

Molecular dynamics simulations carried out in our work use the AMBER biomolecular package originally developed in 1978 in Peter Kollman's laboratory.[120] It's currently maintained by over 40 researchers and collaborators in the science field. The overall figure that represents the functions performed for simulations can be seen in

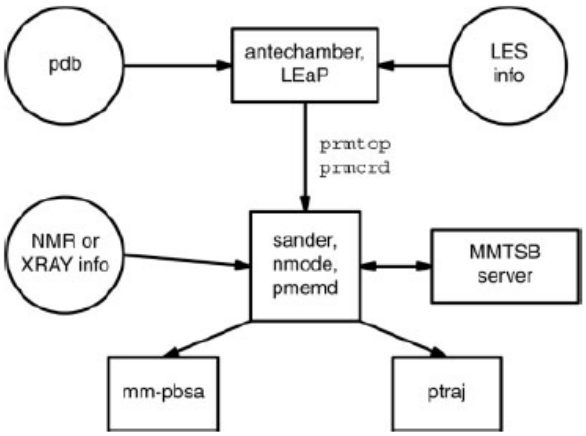



**Figure 2.9:** Flow chart for the AMBER program. (Adapted from Case et al.[121])

Figure 2.9.[121] In order to understand the concepts behind the AMBER biomolecular package, an understanding of molecular mechanics, similar to the QM/MM section, has to be known. The potential energy function in Amber is seen below,

$$V = \sum_{bonds} K_r\left(r - r_{eq}\right)^2 + \sum_{angles} K_\theta\left(\theta - \theta_{eq}\right)^2 + \sum_{dihedrals} \frac{1}{2} V_n[1 + \cos(n\phi - \gamma)] + \sum_{1<j} \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}}\right) + \sum_{H-bonds} \left(\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{10}}\right), \quad (2.48)$$

where $r_{eq}$ and $\theta_{eq}$ are equilibrium structural parameters, $K_r$, $K_\theta$, and $V_n$ are force constants, the periodicity of the system is represented by n, $\gamma$ is the phase angle, and nonbonded parameters are A, B and q. The nonbonded parameters are based off a Lennard-Jones 12-6 potential. The strength of Amber depends on the accuracy of equation 2.48 with respect

to conformational and non-bonded energies.[120] For hydrogen bonds, an added 10-12 Lennard-Jones potential can be used where C and D are parameters based on well depth of the potential energy function. A 10-12 Lennard Jones potential is used for some force fields due to its less steep curve for the repulsive portion of Figure 2.2.[80]

Input preparation starts with a protein databank file, pdb, which gives the starting coordinates of a biomolecule.[122] Antechamber is designed to develop force fields for residues not present in the libraries within AMBER, using the GAFF force field. The general AMBER force field (GAFF), developed by J. Wang et al., is a force field for organic molecules (ligands) within a protein.[121] The non-bonded parameters are taken from the traditional AMBER force field, and the RESP (restrained electrostatic potential fit) develops the partial charges of the atoms.[123] The bond lengths, angles and dihedrals are parameterized and the force constants from equation 2.48 are fit to reproduce the vibrational frequencies of the experimental and *ab* initio calculations. Next, LEaP (Link, Edit and Parm) is used to generate topology and starting coordinate files for use with the other components of the AMBER package.[120-122, 124] Within LEaP, a user can manipulate the molecule in many ways, including solvation, adding residues or atoms, adding ions, among others.[122] Solvation can occur via two periodic boundary conditions, the rectangular parallelpiped or the truncated octahedron.[120, 125] Once the preparation is complete, energy minimization and MD simulations can be performed.

Energy minimization and molecular dynamics simulations are carried out with PMEMD (particle mesh Ewald molecular dynamics). PMEMD, developed by Bob Duke at the National Institute of Environmental Health Sciences is based off the sander

function of the AMBER biomolecular package. Sander is a parallel program that communicates to a number of processors.[121] Each processor is assigned certain atoms, but is aware of the coordinates of all the atoms. The main difference between the sander method and the PMEMD method is improved parallelization.[124] The input and output generated with PMEMD is comparable to that of sander.[121]

The force field implemented in our work is given in equation 2.48. Long range electrostatic interactions are handled by the particle-mesh Ewald technique (PME), by using a slab fast Fourier transform (FFT).[124, 126] For PMEMD, the PME can be manipulated by allowing a block FFT instead of a slab FFT. The block FFT has the option for better parallelization across more processors, but is more computationally demanding. Depending on the amount of processors requested, PMEMD chooses between the block and slab FFT.[124] Energy minimization is performed by a conjugant gradient method with many steps to deplete the bad energy values of the structure. The integration algorithm for running our molecular dynamics simulations is the Verlet leap-frog method with SHAKE constraints for sustaining bond lengths.

Analysis of the trajectories generated by execution of PMEMD is carried out by the ptraj function of the AMBER biomolecular package. Ptraj was developed in order to alleviate the difficulties faced by molecular dynamics trajectories. Trajectories can be very large and require a great deal of time to analyze and they can be analyzed by a variety of different methods.[121] To implement, ptraj takes in a series of 3D atomic coordinates from the trajectory and analyzes them with the help of the topology file

generated by LEaP. The types of analysis that can be performed within ptraj include hydrogen bonding, clustering, entropy and distance calculations.[125]

Trajectories can be further analyzed by principal component analysis to reduce the dimensionality of the dataset and to calculate the root mean square fluctuations (RMSF) of the protein over the simulation length.

## 2.3 Principal Component Analysis

Principal component analysis (PCA) is a type of multivariate analysis that considers the relationship of two or more components as one variable to yield a statistical result.[127] The product of two small matrices, T and P', is used to provide an estimate of a data matrix, M.

$$M \cong T \times P' \tag{2.49}$$

T and P' are matrices used to portray the important data patterns of the data matrix. The matrix, T gives the object patterns of M and P' shows the variable patterns that complement T.[128] The results given for two components are separated via regression in order to establish the relationship between both components. Regression develops an equation for each component, which has the ability to give results in both directions (from one data set to the other and vice versa). However, a single equation is necessary to compare both sets of results.[127] The goals of principal component analysis are as follows,

1. Conversion of correlated variables to uncorrelated ones.
2. Finding linear combinations of a data set while maximizing the variability.
3. Reducing the dimensionality of a data set.[129-130]

Principal component analysis relies mainly on data variances. A linear function ($\alpha_1^T X$) is determined that exhibits the largest variance, noting that $\alpha_1$ is a vector of $b$ constants, " T " represents the transpose vector of α and X is a vector of b variables[129],

$$\alpha_1^T X = \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1b}x_b = \sum_{j=1}^{b} \alpha_{1j}x_j. \tag{2.50}$$

The next step is to search for an uncorrelated linear function to $\alpha_1^T X$, $\alpha_2^T X$ that exhibits the next largest variance, and so on until the last stage has a linear function with largest variance that is uncorrelated with $\alpha_1^T X, \alpha_2^T X, \ldots, \alpha_{m-1}^T X$. Each of these linear functions, represent principal components (PC) until *m* PC's have been found. The largest variations from the original variable are only seen in the first few principal components.[129] A linear dependency is a variable in a data set which can be described as a linear combination of another variable in the same data set,

$$x_1 = \beta_1 x_2 + \beta_2 x_3. \tag{2.51}$$

If none of the variables, *m*, are exact linear dependencies, then the number of principal components matches the number of variables. The dimensionality of the data set is reduced when there is a specific number of linear dependencies.[131]

Solving for principal components is a daunting task. Suppose a vector of random variables, X, has an unknown covariance matrix, S. For PC $m$, $z_m = \alpha_m^T X$, where $\alpha_m^T$ signifies the eigenvector of the covariance matrix which corresponds to the eigenvalue $\lambda_m$. Assuming the normalization of $\alpha_m$ ($\alpha_m^T \alpha_m = 1$), the variance of $z_m$ represents the eigenvalue ($\lambda_m$) of S.[129] The variance can be written as,

$$\frac{1}{n-1}\sum_{i=1}^{n}(\tilde{z}_{in} - \bar{z}_n)^2 \tag{2.52}$$

where $\tilde{z}_{in} = \alpha_n^T x_i$, representing the scores, $\alpha_n^T$ is a transpose vector of coefficients used to maximize the variation, and $\bar{z}_n$ is the average of all z values. The covariance matrix (S) can then be calculated from equation 2.52 and is defined as

$$S_{jk} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\tilde{x}_{ij} - \bar{x}_j\right)\left(\tilde{x}_{ik} - \bar{x}_k\right) \tag{2.53}$$

where

$$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n}\tilde{x}_{ij}\,, \qquad j = 1,2,\dots,m. \tag{2.54}$$

S can be reduced to

$$S = \frac{1}{n-1}X^T X \tag{2.55}$$

where X represents an $(n\ x\ m)$ matrix, or $\left(\tilde{x}_{ij} - \bar{x}_j\right)$.[129]

By reducing the covariance matrix to a diagonal matrix, the principal components can be found using

$$U^T S U = \lambda \tag{2.56}$$

where $\lambda$ represents the eigenvalues of S, and U represents the eigenvectors of S. The eigenvalues can be determined by

$$|S - \lambda I| = 0 \tag{2.57}$$

where I is the identity matrix,

$$\begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}. \tag{2.58}$$

Solving for λ, gives the eigenvalues of a given $(m\ x\ m)$ covariance matrix. The eigenvectors corresponding to the eigenvalues, can be found by

$$[S - \lambda I] t_i = 0 \tag{2.59}$$

and normalized by

$$u_i = \frac{t_i}{\sqrt{t_i' t_i}} \tag{2.60}$$

for *i*=1, 2, …, m.[127] The eigenvectors represent the orthonormal matrix

$$U = [u_1 | u_2] \tag{2.61}$$

The number of principal components to fully account for the variation of the data in the data matrix, X, varies depending on the data set. There are a few ways to correctly determine the number of PC's of interest. The first method is by examining the sum of the total percent variation by

$$t_m = 100 \sum_{k=1}^{n} l_k \Big/ \sum_{k=1}^{m} l_k \tag{2.62}$$

where *n* represents the number of PC's of interest, *m* represents the total number of PC's, *k* is the PC of interest, and $l_k$ is the variance of *k*. The correct number of PC's to choose represents 70-90% of the total variance of the data matrix, X.[129] Another method is the scree plot which is a plot of the eigenvalues versus its principal component ($l_k vs. n$). R. B. Cattell designed the scree plot in order to include 95% of the total variance.[132] He proposed that the last principal component of interest is where the line goes from steep to shallow, at the "elbow." An "elbow" typically forms in the plot of $l_k vs. n$ which designates the cutoff of principal components (Figure 2.10).[129, 132]

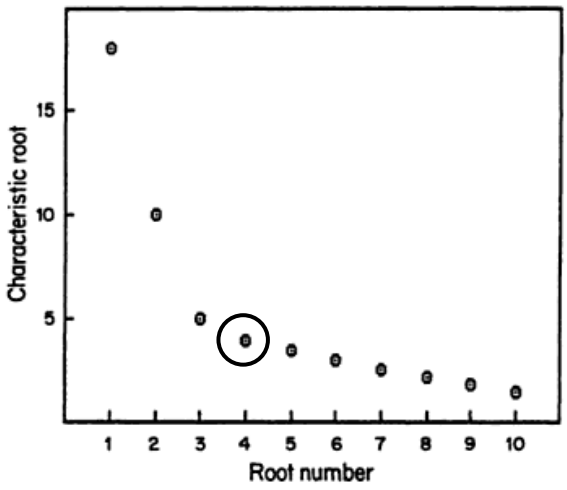



**Figure 2.10:** Scree plot (circle represents "elbow" region of plot). (Adapted from Kaiser[132])

H. F. Kaiser, in 1960, concluded that the best answer for choosing the number of factors (principal components) is any variance ($l_k$) larger than 1 should be retained, also known as Kaiser's rule.[129, 133] In some cases, an $l_k$ value of 1 may not keep enough principal components, so a lower cutoff can be applied where $l_k$ is less than 1. I. T. Jolliffe concluded that a variance of 0.7 is a better option than a variance of 1.[129, 134]

The correlations of variables in principal component analysis can be seen using score plots, which are scatter plots of two principal components. The three dimensional data generated from each principal component generates a plot of points looking similar to an egg shape. Within this plot, lines represent one-dimensional subspaces, and planes representing two-dimensional subspaces can be found. If this line or plane lies closest to the majority of the points, it represents the majority of the variance of the data set.[131]

For larger dimensions (greater than three-dimensional), PCA generates a lower dimension subspace (2D or 3D) and closely fits all dimensions to this lower subspace. The scores, or principal points (p), of an n-variate random vector (X) are represented by vectors,

$$\xi_1, \dots, \xi_n \epsilon \mathbb{R}^p \tag{2.63}$$

which minimize

$$E[D^2(X|\xi_1, \dots, \xi_n)]. \tag{2.64}$$

The squared Elucidean distance ($E[D^2]$) between the random vector and the $j^{th}$ random vector ($\xi_j$) is,

$$D^2(X|\xi_1, \dots, \xi_N) = \min_{1 \le j \le n} \left(X - \xi_j\right)^T \left(X - \xi_j\right). \tag{2.65}$$

Using this methodology, the principal points lie in the first few principal components subspace.[130, 135-136]

Principal component analysis can be performed with many different statistical programs. The program of choice for this research was Bio3D. Bio3D is a program that is able to analyze protein structures.[137] It is capable of reading and writing pdb structures, atom selection, reorientation, superposition, analysis of distance matrices, principal component analysis, among many other functions. It has allowed for computations of root mean square fluctuation (RMSF) of each principal component of

interest, and has given scree and score plots for the alkanesulfonate monooxygenase systems.

# Chapter 3

## Claisen Rearrangement: Insight into Solvent Effects and "on Water" Reactivity from QM/MM Simulations

An "on water" environment, defined by the absence of water solubility of the reactants, has been reported to provide increased rate accelerations, yields, and specificity for several types of organic reaction classes compared to organic solvents. The aromatic Claisen rearrangements of allyl *p*-R-phenyl ethers (R = $CH_3$, Br, and $OCH_3$) and allyl naphthyl ether have been investigated to determine the origin of the "on water" effect using QM/MM Monte Carlo calculations and free-energy perturbation theory. The simulations indicate that "on water" rate enhancements for the rearrangements are derived from the ability of the interfacial waters to stabilize the polar transition state via enhanced hydrogen bonding at the oil/water interface. The position and orientation of the aromatic ethers at the interface are crucial factors affecting solvent accessibility during the reaction pathway; computed solute-solvent energy pair distributions and radial distribution functions showed that hydrophobic substituents on the solute provided a more polar solvent environment than hydrophilic substituents by tilting the reacting oxygen toward the water surface. Calculations in 16 different solvents accurately reproduced the experimental trend of increased rates correlated to increasing solvent polarity. Hydrophobic effects did not provide a substantial contribution to the lowering of

**Reprinted with permission from Acevedo, O., Armacost, K. *J. Am. Chem. Soc.*, 2010, 132, 1966-1975. Copyright (2010), American Chemical Society.**

the free energy activation barrier (< 0.5 kcal/mol), and solvent polarizability via a polarizable force field was also found to be negligible in the observed rate accelerations. New insight into solvent effects for the Claisen rearrangement is presented herein, and a QM/MM approach for computing reactions on a liquid surface is highlighted.

### 3.1 Introduction

QM/MM calculations featuring the PDDG/PM3 semiempirical QM method coupled to free energy perturbation (FEP) theory and Monte Carlo (MC) simulations have been carried out for the four Claisen rearrangements shown for allyl R-*p*-tolyl ether (R = $CH_3$, $OCH_3$, Br) and allyl naphthyl ether (Scheme 1.5), "on water", "in water", and in 14 additional solvents, viz. *p*-chlorophenol, phenol, *p*-cresol, ethylene glycol, 2-aminoethanol, diethylene glycol monoethyl ether (carbitol), sulfolane, adiponitrile, propylene carbonate, n-decylamine, methanol, acetonitrile, DMF, and toluene. Activation barriers, solute-solvent interaction energies, and radial distribution functions have been computed with complete sampling of the solute to elucidate the role of solvation on the reactants and activated complexes. Substituent effects were investigated using R = $CH_3$, Br, and $OCH_3$ groups in allyl *p*-R-phenyl ethers (Scheme 1.5A).[32] The aromatic Claisen rearrangement of allyl naphthyl ether (Scheme 1.5B), reported by Sharpless and coworkers to provide better yields and rates "on water" when compared to typical organic solvents,[13] was also computed in this investigation. In addition, a detailed comparison of alternative methods to our QM/MM approach is given. Notably, high-level *ab initio* and density functional theory (DFT) calculations at the CBS-QB3 and B3LYP theory levels,

respectively, coupled to a continuum solvent model failed to predict the sequence of reaction rates in solution.[4, 13, 138]

New insight into "on water" Claisen rearrangement reactions is presented within chapter 3 section 3. In addition, solvent effects on aromatic Claisen rearrangements from different classes of solvent, e.g., protic, polar aprotic, and nonpolar, are discussed in detail and changes along the reaction pathway are fully characterized. To our knowledge, this work is the first QM/MM study of a reaction on a liquid surface featuring hundreds of solvent molecules. Reported benefits of an "on water" aqueous medium for multiple reaction types open up many additional systems for study and interpretation using the present QM/MM methodology, e.g., cycloadditions, Diels-Alder, and epoxide opening reactions.

**3.2 Methods**

Reactants and transition states were located in hundreds of explicit solvent molecules represented using the TIP4P water model[139] and the OPLS-AA force field for non-aqueous solutions.[100, 140] The solutes were treated using the PDDG/PM3 semiempirical QM method,[87-89] which has given excellent results for a wide variety of organic and enzymatic reactions in the solution-phase.[141-152] Computation of the QM energy and atomic charges is performed for each attempted move of the solute, which occurs every 100 configurations. For electrostatic contributions to the solute-solvent energy, CM3 charges[153] were obtained for the solute with a scaling factor of 1.14. Lennard-Jones interactions between solutes and solvent atoms were taken into account using OPLS parameters. This combination is appropriate for a PM3-based method as it

minimizes errors in the computed free energies of hydration.[154] Changes in free energy were calculated using free energy perturbation (FEP) theory in conjunction with NPT Metropolis Monte Carlo (MC) simulations at 25 °C and 1 atm. Each FEP calculation entailed ca. 40 million (M) configurations of equilibration and 20 M configurations of averaging using increments of 0.01 Å for the making/breaking bond distances (Figure 1).

To explore the effect of various solvents upon the Claisen rearrangement, new OPLS-AA solvent boxes were constructed in a fashion similar to previous work.[143] Briefly, the liquid-phase simulations were carried out by placing 400 solvent molecules at random positions in the simulation box. The boxes were then equilibrated at 25 °C for 100-400 million MC steps in the NPT ensemble. In the case of long-chain molecules, e.g., carbitol and n-decylamine, an initial temperature value of 1000 °C was applied prior to equilibration for 10 million configurations in the NVT ensemble to encourage a thorough mixing. The initial heating has been shown to provide improved convergence of MC simulations for long-chain alkanes and ionic liquids.[143, 155] The computed liquid densities and heats of vaporization ($\Delta H_{vap}$) for the solvent boxes were found to be in good agreement with experiment. A complete table comparing the computed and experimental densities and $\Delta H_{vap}$ is given in the Supporting Information as Table S1; OPLS atom types for all solvents simulated are given in Figure S1.

*BOSS* was used to carry out all QM/MM reactions from the stored and custom-made solvent boxes.[156] All solvents were fully flexible, i.e., all bond stretching, angle bending, and torsional motions were sampled, with the exception of water, acetonitrile,[157] and methanol[19] which already exist in *BOSS* as fixed united-atom solvents. Periodic

boundary conditions have been applied to tetragonal boxes containing between 400-750 solvent molecules. Solute-solvent and solvent-solvent cutoffs of 10-12 Å were employed with quadratic feathering of the intermolecular interactions within 0.5 Å of the cutoff. Adjustments to the allowed ranges for rotations, translations, and dihedral angle movements led to overall acceptance rates of 30 – 50% for new configurations. The ranges for bond stretching and angle bending were set automatically by the BOSS program on the basis of force constants and temperature. All QM/MM/MC calculations were run on a Linux cluster at Auburn University.

The complete basis set method CBS-QB3[158] was also used to characterize the transition structures and ground states in vacuum using Gaussian 09.[159] In a recent study, the CBS-QB3 method gave energetic results in the closest agreement to experiment for a set of 11 different pericyclic reactions compared to other *ab initio* and density functional theory (DFT) methods.[160] The CBS-QB3 calculations were used for geometry optimizations and computations of vibrational frequencies, which confirmed all stationary points as either minima or transition structures and provided thermodynamic corrections. The effect of solvent was approximated by subsequent optimization and vibrational frequency calculations using the polarizable continuum model (PCM)[161] and the B3LYP/6-311+G(2d,p) theory level.[162-163] All *ab initio* calculations were performed on computers located at the Alabama Supercomputer Center.

### 3.3 Structures

Configurationally-averaged free-energy maps, $\Delta G$ (kcal/mol), were computed for the aromatic Claisen rearrangements (Scheme 1.5) in 16 different solvents by perturbing the reacting bond distances $R_{OC}$ and $R_{CC}$ via potentials of mean force (PMF) calculations in increments of 0.05 Å (Figure 3.1). To locate the critical points more precisely, the regions surrounding the free-energy minima and maxima from the initial maps were explored using increments of 0.01 Å. This provided refined results for the geometries and activation barriers, as summarized in Tables 3.1-3.4 and Tables S2-S4 in Appendix I. All internal degrees of freedom, minus the $R_{OC}$ and $R_{CC}$ reaction coordinates, were fully sampled during the simulations. It was necessary to use multiple gas-phase optimized structures to provide good starting points for the reaction profile. Each solution-phase free energy perturbation (FEP) calculation required approximately 120 million single point QM calculations per free-energy map, illustrating the need for highly efficient QM methods.

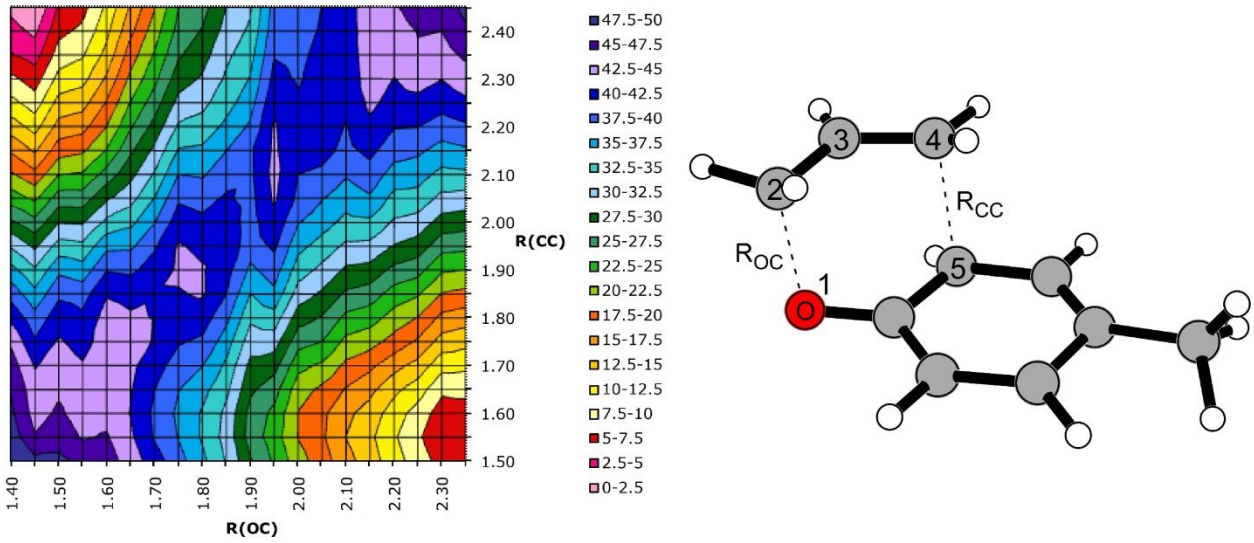



**Figure 3.1.** Free energy map (kcal/mol) computed for the "on water" Claisen rearrangement of allyl *p*-tolyl ether using the reaction coordinates $R_{OC}$ and $R_{CC}$. Illustrated structure is the transition structure from the QM/MM/MC simulations.

The "on water" reactions were carried out in a similar fashion to the homogeneous solution-phase calculations except that periodicity was removed from the z-axis of an explicit 500 water molecule box and the ether was placed on top of the z-axis; an equilibrated NVT water slab was used to prevent a very slow drift of the water molecules that reduces the exposed surface, i.e. stretch in the z-direction. An orientation analysis of multiple snapshots of allyl *p*-tolyl ether relative to the surface of the slab (perpendicular to the z-axis) found that the aromatic ring tended to lie essentially flat on the surface with a slight tilt towards the oxygen while the allyl side chain occupied the gas-phase (Figures 3.2 and Appendix I - S2). Hydrogen bonding was found to be primarily responsible for the orientation, i.e., favorable interactions between water and the π-system of the aromatic ring and the oxygen. A similar orientation was found for allyl naphthyl ether, where the fused aromatic rings lay essentially flat on the surface of

the water slab with a tilt towards the oxygen atom (Appendix I Figure S3).The addition of hydrophilic substituents on the phenyl ring, e.g., R = $OCH_3$ in the allyl *p*-R-phenyl ethers, strongly tilted the ring on the surface towards the substituent and as a result tilted the reacting oxygen away from the interfacial waters reducing the number of hydrogen bonds available (see "Radial Distributions Functions" section).

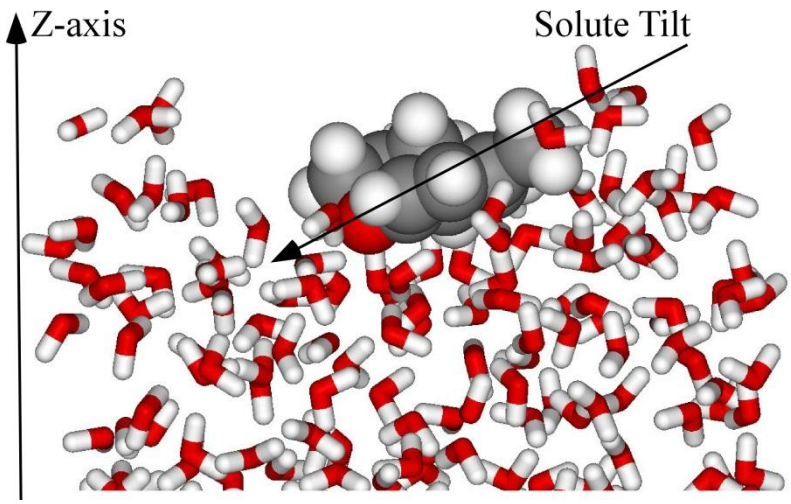



**Figure 3.2:** Illustration of the "on water" allyl *p*-tolyl ether transition structure from the QM/MM/MC Claisen rearrangement calculations.

The calculations showed that, although solvent effects can produce a rate acceleration of up to 100-fold,[31] the transition structure geometries do not show large variations among different solvents, e.g., ca. $R_{OC}$ of 1.86 ± 0.02 Å and $R_{CC}$ of 1.98 ± 0.02 Å for the allyl *p*-tolyl ether reactions (Table 3.2). Secondary deuterium kinetic isotope effect (KIE) experiments on related allyl vinyl (and phenyl) ethers also found the transition state to be solvent-insensitive with no significant bond breaking increase in aqueous media compared to nonpolar solvents.[164-165] CBS-QB3 calculations predicted an earlier activated complex for allyl *p*-tolyl ether with gas-phase $R_{OC}$ and $R_{CC}$ distances of

2.16 and 2.25 Å, respectively; B3LYP/6-311+G(2d,p) gave similar results, i.e. 2.17 and 2.25 Å. The trend of earlier transition structures for the *ab initio* and DFT calculations compared to PDDG/PM3 was also found in the other Claisen rearrangements studied (see Appendix I Tables S2-S4 for gas-and solution-phase geometries of the allyl *p*-R-phenyl ethers and allyl naphthyl ether). All theory levels used were consistent with KIE measurements which indicated that the activated complex should resemble the reactant with more C–O bond cleavage than C–C bond formation occurring.[164] In the case of allyl naphthyl ether, the $R_{OC}$ and $R_{CC}$ distances at the transition state were found to slightly increase with increasing solvent polarity, e.g., 1.83, 1.84, 1.86, 1.87, 1.89, and 1.90 Å for toluene, DMF, acetonitrile, methanol, "in water," and "on water," respectively, for the $R_{OC}$ distance (Appendix I, Table S4)

**Table 3.1:** Computed Bond Lengths (Å) for the Allyl *p*-Tolyl Ether Claisen Rearrangement Transition Structures at 25 °C and 1 atm.

| | $R_{OC}$ | $R_{CC}$ |
|---|---|---|
| *p*-chlorophenol | 1.88 | 2.00 |
| "in water" | 1.87 | 2.01 |
| phenol | 1.87 | 2.00 |
| *p*-cresol | 1.84 | 1.97 |
| ethylene glycol | 1.86 | 1.97 |
| "on water" | 1.86 | 1.97 |
| 2-aminoethanol | 1.84 | 1.98 |
| carbitol | 1.86 | 1.98 |
| sulfolane | 1.86 | 1.97 |
| adiponitrile | 1.87 | 1.97 |
| propylene carbonate | 1.86 | 1.99 |
| n-decylamine | 1.83 | 1.97 |

[a]From the 2D free-energy maps computed in the QM/MM/MC/FEP simulations.

### 3.4 Energetics

The absolute $\Delta G^{\ddagger}$ values at 25 °C computed from the QM/MM/MC simulations were overestimated by an average of ca. 14 kcal/mol compared to experiment at 170 °C; however, excellent computed $\Delta\Delta G^{\ddagger}$ (kcal/mol) values between different solvents were found (Table 3.2). Error ranges in the computed free-energy values have been estimated from fluctuations in the $\Delta G$ values for each FEP window using the batch means procedure with batch sizes of 0.5 million configurations; calculated errors in the free energies imply overall uncertainties in the $\Delta G^{\ddagger}$ of ca. 0.4 kcal/mol. The $\Delta G^{\ddagger}$ overestimation is a systematic error common in pericyclic reactions when employing a

semiempirical method.[166] Similar errors have been reported for the Claisen rearrangement of chorismate to prephenate using the AM1 method.[24] Cramer and Truhlar also reported an underestimation of the absolute rate accelerations for aqueous allyl vinyl ether and derivative Claisen rearrangements when employing AM1.[167] The authors concluded that a systematic underestimation of the charges at the transition states was responsible, but the analogous trends in charge reported from the AM1-SM1 model provided excellent relative solvation free energies for the reacting species.[167-168] It is important to note that the overestimation of the absolute $\Delta G^{\ddagger}$ value using semiempirical QM/MM methods is not limited to the Claisen rearrangement as similar findings have been reported for multiple Diels-Alder reactions, [145-146, 169] ene reactions,[142] and methyl transfer reactions.[170]

Potentially, a simple reparameterization of the PDDG/PM3 Hamiltonian by scaling the energies from points along the reaction coordinate could provide accurate $\Delta G^{\ddagger}$ values; however, the physical reasons for determining structures would be absolutely the same as the original Hamiltonian. Hence, there is no difference in leaving the Hamiltonian in its original form or in scaling the energies when one considers the relative energies computed. The close $\Delta\Delta G^{\ddagger}$ agreement indicates that solvent effects on the Claisen rearrangements are being appropriately modeled (Table 3.2). Notably, recent B3LYP/6-31G(d) calculations using the polarizable continuum model (PCM) failed to predict the accurate sequence of reaction rates for the allyl *p*-tolyl ether rearrangement in solution.[171] Higher theory levels coupled to improved PCM models gave similar results in this work (see Table 3.4 and the "Continuum Solvent Model" section). An additional

approach for the improvement of the $\Delta G^{\ddagger}$ values is the assignment of specific Lennard-Jones terms for every atom type in the Claisen rearrangements as opposed to the general OPLS empirical parameters chosen for each atom in the QM region, e.g., C, O, H. Optimization of the Lennard-Jones parameters embedded in the hydrid QM/MM potential, as pioneered by Gao, has been successful for multiple organic reaction studies.[172] However, preliminary testing found the results to be nearly identical to the present methodology.

**Table 3.2:** Absolute Free Energy of Activation, $\Delta G^{\ddagger}$ (kcal/mol), at 25 °C for the Claisen Rearrangement of Allyl *p*-Tolyl Ether and the $\Delta\Delta G^{\ddagger}$ Relative to "In Water" using QM/MM.

| | $\Delta G^{\ddagger}$ (calc)[a] | $\Delta G^{\ddagger}$ (exptl)[b] | $\Delta\Delta G^{\ddagger}$ (calc)[a] | $\Delta\Delta G^{\ddagger}$ (exptl)[b] |
|---|---|---|---|---|
| *p*-chlorophenol | 47.9 | 33.4 | -0.8 | -1.0 |
| "in water" | 48.8 | 34.4[c] | 0.0 | 0.0[c] |
| phenol | 49.1 | 34.4 | 0.4 | 0.0 |
| *p*-cresol | 48.7 | 34.7 | 0.0 | 0.3 |
| ethylene glycol | 48.7 | 34.7 | 0.0 | 0.3 |
| "on water" | 49.6 | - | 0.9 | - |
| 2-aminoethanol | 49.7 | 35.5 | 1.0 | 1.1 |
| carbitol | 50.0 | 35.9 | 1.3 | 1.5 |
| sulfolane | 50.5 | 36.3 | 1.8 | 1.9 |
| adiponitrile | 50.7 | 36.4 | 2.0 | 2.0 |
| propylene carbonate | 50.8 | 36.5 | 2.1 | 2.1 |
| *n*-decylamine | 51.4 | 37.1 | 2.7 | 2.7 |

[a] PDDG/PM3 and MC/FEP. [b] Ref. [31], 170°C; Exptl. error for $\Delta G^{\ddagger}$ is ca. ± 0.3 kcal/mol. [c]28.5% EtOH-$H_2O$.

Of particular interest is that the $\Delta\Delta G^{\ddagger}$ for the "on water" reaction was predicted to be only ca. 0.9 and 0.6 kcal/mol higher in energy than the same reaction homogeneously

dissolved in pure aqueous solution for the allyl *p*-R-phenyl ethers and allyl naphthyl ether, respectively (Tables 3.2-3.4). This prediction places the rate of reaction for the "on water" environment in-line with that of polar aprotic solvents and potentially faster than half of the solvents tested. The results are consistent with experimental observations by Sharpless on the allyl naphthyl ether Claisen rearrangement and with other recent work on cycloaddition and Wittig reactions.[13-14, 173-174]

### 3.5 Solvent Effects on the Rate of Reaction

In the Claisen rearrangement study of allyl *p*-tolyl ether by White and Wolfarth the rate of reaction was found to increase with increasing solvent polarity as defined using a *Z* factor[175] correlation at 25 °C.[31] However, no specific characteristic of the solvent polarity could be attributed to the rate accelerations. For example, while hydrogen bonding may play a large role and the fastest rates generally occurred in hydroxylic solvents, the reaction was also equally as fast in some nonhydroxylic solvents, e.g., propylene carbonate compared to 2-octanol. In addition, while increased solvent acidity also appeared to increase rates, no clear correlation exists, e.g., the reaction was a factor of 2.8 slower in octanoic acid as compared to *p*-chlorophenol.[31] The major difficulty associated with determining the origin of the solvent effects on the rate of the Claisen rearrangement is that solvent polarity can be a composite of many different properties, including solvent polarizability, hydrogen-bonding ability, and dipole moment. Resolving their individual contributions on the reaction is extremely difficult from an experimental perspective. However, detailed insight on the changes in solvation along the reaction

pathway is available from the present QM/MM/MC calculations. The relationship between the structures, solvent properties, rates, and energies is discussed below.

### 3.6 Solvent Polarization

Fixed charges on the solvent molecules, derived from the OPLS-AA force field,[100] were used to model the electrostatic interactions occurring between the solvent and solute in the QM/MM simulations; this approach utilizing non-polarizable potential functions for the solvent molecules has been successful in accurately reproducing condensed-phase effects on a large variety of organic reactions.[141] However, solvent polarizability is often cited as a necessary criterion for accurately reproducing solvent effects, particularly for low dielectric systems.[75, 176-178] To test the effect of solvent polarization upon the Claisen rearrangement, calculation of the allyl *p*-tolyl ether reaction in polarizable phenol was carried out. An approach identical to that reported in a recent paper by Jorgensen et al. was employed, in which inducible dipoles were added to non-hydrogen atoms and nonbonded interactions were handled using a OPLS/CM1A approach; improved interaction energies were reported for anion-phenol complexes when using this polarizable model.[179] The electric field that determines the inducible dipoles is computed from the permanent charges using Eq. 3.1, and the polarization energy is given by Eq. 3.2. The same first-order polarization model has been used in earlier studies with good success.[180-182] Values of 1.0 and 1.5 $Å^3$ for the polarizabilities, $\alpha_i$, on carbon and heteroatoms, respectively, were used.[179] The $\Delta G^{\ddagger}$ values computed for the allyl *p*-tolyl ether reaction using the polarized OPLS/CM1A versus unpolarized OPLS-AA solvent models were 49.2 and 49.1 kcal/mol, respectively. The transition structure geometries

were also essentially identical with $R_{OC}$ and $R_{CC}$ distances of 1.87 and 2.00 Å compared to 1.87 and 1.99 Å in the polarized and unpolarized phenol. Solvent polarization was not found to provide an essential contribution in the observed rate accelerations for the Claisen rearrangement, at least when considering phenol which should be a good candidate for polarization effects in view of its low dielectric constant (ε = 9.9).

$$\vec{\mu}_i = \alpha_i \vec{E}_i^0 \tag{3.1}$$

$$E_{pol} = -(1/2) \sum_i \vec{\mu}_i \cdot \vec{E}_i^0 \tag{3.2}$$

### 3.7 Substituent Effects on Energetics

Substituent effects on the Claisen rearrangement were studied with the inclusion of R = $OCH_3$, $CH_3$, and Br groups in the allyl *p*-R-phenyl ethers. The computed energies "in water" and "on water" were in good agreement with experimental free-energies of activation relative to $CH_3$ (Table 3.3), e.g., computed $\Delta\Delta G^{\ddagger}$ values of 0.4, 0.0, and -1.1 kcal/mol compared to experimental values of 0.5, 0.0, and -0.9 kcal/mol for the Br, $CH_3$, and $OCH_3$ reactions "in water", respectively. Correlation of experimental rate constants for *p*-substituted allyl phenyl ethers with different substituents has shown that electron-donating groups weaken the reacting carbon-oxygen bond and increase the rate for both thermal- and photo-Claisen rearrangements.[25, 183-184] However, R = Br in allyl *p*-R-phenyl ether has a reduced reaction rate compared to $OCH_3$ and $CH_3$. Bond dissociation energy (BDE) in phenol has been shown to be strongly dependent on substituents; experimental

measurements and theoretical work indicate that halogen substituents have only a minor effect on the BDE for cleavage reactions of phenolic ethers as compared to methoxy substituents.[185-186] This may originate from the strong electron-withdrawing inductive effect of Br outweighing its weak electron-donating resonance effect. However, the current aqueous phase simulations also suggest that site-specific hydrogen bonding between the reacting oxygen of the ether and the water molecules, and accessibility of the interfacial waters to the solute play a large role on the rate accelerations (see "Radial distribution functions" section).

**Table 3.3:** Absolute Free Energy of Activation, $\Delta G^{\ddagger}$ (kcal/mol), at 25 °C for the Claisen Rearrangement of Allyl *p*-R-Phenyl Ethers and the $\Delta\Delta G^{\ddagger}$ Relative to R = $CH_3$ using QM/MM simulations.

| R | $\Delta G^{\ddagger}$ (calc)[a] | $\Delta G^{\ddagger}$ (exptl)[b] | $\Delta\Delta G^{\ddagger}$ (calc)[a] | $\Delta\Delta G^{\ddagger}$ (exptl)[b] |
|---|---|---|---|---|
| "in water" | | | | |
| $OCH_3$ | 47.6 | 33.5 | -1.1 | -0.9 |
| $CH_3$ | 48.8 | 34.4 | 0.0 | 0.0 |
| Br | 49.1 | 34.9 | 0.4 | 0.5 |
| "on water" | | | | |
| $OCH_3$ | 48.5 | - | -1.1 | - |
| $CH_3$ | 49.6 | - | 0.0 | - |
| Br | 50.1 | - | 0.5 | - |

[a] PDDG/PM3/MC/FEP. [b] Exptl. values for 28.5% EtOH-$H_2O$ at 170 °C; Ref. [32].

Allyl naphthyl ether was also computed "in water," "on water," and in 4 additional solvents, i.e. methanol, acetonitrile, DMF, and toluene (Table 3.4). Activation barriers have not been reported experimentally; however, a qualitative representation of the rates and yields were given by Sharpless and coworkers.[13] The "on water"

environment for the allyl naphthyl ether rearrangement was observed to produce faster rates at 23 °C (5 days for complete reaction to occur) compared to neat conditions (6 days) and organic solvents (>6 days). A 100% yield was reported for the "on water" reaction and provided the best set of conditions despite the modest rate acceleration.[13] The current calculations are consistent with the experimental results predicting "on water" to have a faster rate than all other solvents computed with the exception of acetonitrile and the "in water" reactions. Acetonitrile used a 3-site united-atom model[157] which could potentially have affected the results, as most reactions computed used an all-atom OPLS model.[100] Accordingly, an all-atom acetonitrile solvent box was constructed in an identical fashion to the other solvents and used to recompute the Claisen rearrangement of allyl naphthyl ether. The $\Delta G^{\ddagger}$ was not dramatically changed as the predicted activation barrier fell within the computed error range of ± 0.4 kcal/mol ($\Delta G^{\ddagger}$ = 40.8 and 41.1 kcal/mol in the all-atom and united-atom $CH_3CN$ models, respectively).

**Table 3.4:** Computed Free Energy of Activation, $\Delta G^{\ddagger}$ (kcal/mol), at 25 °C for the Claisen Rearrangement of Allyl Naphthyl Ether and the $\Delta\Delta G^{\ddagger}$ Relative to "In Water."

| | $\Delta G^{\ddagger}$ (QM/MM)[a] | $\Delta\Delta G^{\ddagger}$ (QM/MM)[a] | $\Delta G^{\ddagger}$ (PCM)[b] | $\Delta\Delta G^{\ddagger}$ (PCM)[b] |
|---|---|---|---|---|
| "in water" | 40.7 | 0.0 | 20.0 | 0.0 |
| "on water" | 41.3 | 0.6 | - | - |
| $CH_3CN$ | 41.1[c] | 0.4 | 20.1 | 0.1 |
| MeOH | 43.5 | 2.8 | 20.1 | 0.1 |
| DMF | 43.9 | 3.2 | 20.1 | 0.1 |
| toluene | 42.5 | 1.8 | 21.8 | 1.8 |

[a] PDDG/PM3/MC/FEP. [b] Optimized using B3LYP/6-311+G(2d,p). [c] Using a united-atom OPLS solvent model. The all-atom version gave a $\Delta G^{\ddagger}$ of 40.8 kcal/mol.

### 3.8 Solute-Solvent Energy Pair Distributions

To elucidate the fact that reaction rates increase with increasing solvent polarity (as defined by Z values),[31] solute-solvent energy pair distributions were computed for all reactions studied. Solute-solvent energy pair distributions record the average number of solvent molecules that interact with the solute and their associated energy. The interaction energies are quantified by analyzing the QM/MM/MC results near the reactant and at the transition structure FEP windows. The results for the Claisen rearrangement of allyl *p*-tolyl ether in *p*-chlorophenol and n-decylamine are shown in Figure 3.3, results for the allyl *p*-R-phenyl ethers "in water" and "on water" are given in Figure 3.4, and the values for all other solvents and allyl naphthyl ether are found in Figures S4-S6 of Appendix I. Hydrogen bonding "in water" and "on water" and the most favorable electrostatic interactions in the non-aqueous solvents are reflected in the left-most region, with solute-solvent interaction energies more attractive than -4 kcal/mol. The large bands near 0 kcal/mol result from the many distant solvent molecules in the outer shells.

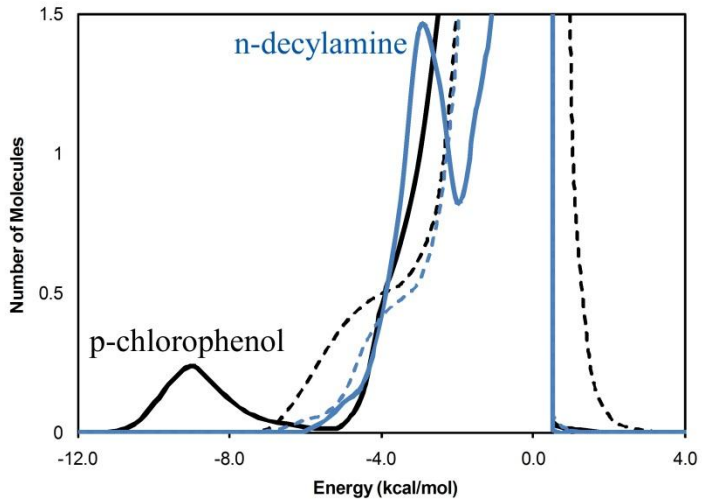



**Figure 3.3:** Solute-solvent energy pair distributions for the Claisen rearrangement of allyl *p*-tolyl ether in *p*-chlorophenol, transition structure (solid black) and reactant (dashed black), and in n-decylamine, transition structure (solid blue) and reactant (dashed blue), at 25 °C. The ordinate records the number of solvent molecules that interact with the solutes and their interaction energy on the abscissa.

It is immediately clear in viewing Figure 3.3 that there is a much greater gain in very favorable solute-solvent interactions when proceeding from the reactant to the transition structure in the most polar non-aqueous solvent studied, *p*-chlorophenol, compared to the least polar, n-decylamine. The low energy band in *p*-chlorophenol arises from favorable electrostatic interactions with the more dipolar transition structure, i.e. hydrogen bonding between the hydroxyl group of the solvent and the oxygen of the ether, and favorable π-π stacking between the aromatic rings; these interactions diminish dramatically in n-decylamine. Integration of the transition structure band from -12 to -5.5 kcal/mol yields 1.1 *p*-chlorophenol molecules and 0.5 n-decylamine molecules (expanding the integration to -4 kcal/mol increases the number of n-decylamine

molecules to 0.7). The average strength of the peak for *p*-chlorophenol at the transition structure is at -9 kcal/mol which weakens to -3 kcal/mol for n-decylamine. The results are consistent with experimental findings that transition state effects are primarily responsible for rate accelerations (determined by alkyl and aryl substitutions at the α and γ-positions in the allyl side chain).[187] Inspection of the solute-solvent energy pair distributions for the other 8 non-aqueous solvents finds a general trend of weaker and less specific dipole-dipole interactions between the solvent and the allyl *p*-tolyl ether transition structure as the solvent polarity diminishes (Figures S4 and S5 of Appendix I).

Of more relevance to the "on water" reactivity is the exact nature of the favorable solute-solvent interactions in the aqueous environments relative to the conventional organic solvents. Comparison of the "in water" and "on water" systems also found stronger bands at the activated complex, shown as solid red and black lines in Figure 3.4, relative to the reactants (dashed lines). Integration of these bands from -12 to -4 kcal/mol results in 2.2 and 2.1 water molecules interacting with the allyl *p*-tolyl ether reactant and transition state, respectively, "in water" and 1.9 and 2.0 "on water" (Table 3.5; integration from -12 to -3.5 also given). While the number of solute-solvent interactions are similar at the ground and transition states, the associated energies for those interactions are stronger at the transition structure (Figure 3.4). Destabilization of the reactant "on water," in addition to increased solute-solvent interactions at the transition structure, was found for the R = $OCH_3$ and Br substituted ethers with the number of waters interacting with the reactant decreasing to 1.0 and 1.4, respectively, compared to 1.9 for $CH_3$. The increase in solute-solvent interactions from the reactant to activated

complex is also larger for the "on water" environment compared to the "in water," particularly for allyl naphthyl ether which increases by ca. 1 water molecule (Table 3.5). The results suggest that while enhanced hydrogen bonding interactions at the transition state is the primary contributor towards a reduction in the activation barrier, destabilization of the reactant plays large role for the "on water" systems. The hydrogen bonds become stronger for the transition state owing to the enhanced charges at the reacting oxygen of the Claisen rearrangements (Table 3.6). The overall results supports the fact that hydrogen bonding is very sensitive to charge distributions,[47, 188] so it is more affected by the progression from the reactant to the activated complex in water than the weaker dipole-dipole interactions in the less polar solvents.

**Table 3.5:** Solute-Solvent Energy Pair Distributions for the Claisen Rearrangement of Allyl p-R-Phenyl Ethers and Allyl Naphthyl Ether for the Reactant (GS) and Transition Structure (TS) "in water" and "on water" integrated to -4.0 kcal/mol (and -3.5 kcal/mol in parenthesis)[a].

| R | "in water" GS | "in water" TS | "on water" GS | "on water" TS |
|---|---|---|---|---|
| $OCH_3$ | 1.7 (2.5) | 1.7 (2.3) | 1.0 (1.4) | 1.8 (2.2) |
| $CH_3$ | 2.2 (2.7) | 2.1 (2.5) | 1.9 (2.4) | 2.0 (2.3) |
| Br | 1.6 (2.1) | 2.1 (2.5) | 1.4 (1.8) | 1.8 (2.2) |
| naphthyl | 2.0 (2.6) | 2.6 (3.1) | 1.1 (1.5) | 2.0 (2.4) |

[a]From Figures 4 and S6.

**Table 3.6:** Charges for the Claisen Rearrangement Transition Structure (and Reactant in Parenthesis) of Allyl *p*-R-Phenyl Ether[a].

| R | O1 | C2 | C4 | C5 |
|---|---|---|---|---|
| "in water" | | | | |
| $OCH_3$ | -0.368 (-0.256) | -0.089 (-0.045) | -0.224 (-0.338) | -0.135 (-0.173) |
| $CH_3$ | -0.385 (-0.267) | -0.087 (-0.027) | -0.178 (-0.346) | -0.222 (-0.193) |
| Br | -0.402 (-0.248) | -0.099 (-0.061) | -0.149 (-0.332) | -0.243 (-0.214) |
| "on water" | | | | |
| $OCH_3$ | -0.380 (-0.256) | -0.100 (-0.082) | -0.186 (-0.316) | -0.169 (-0.257) |
| $CH_3$ | -0.383 (-0.255) | -0.064 (-0.058) | -0.165 (-0.341) | -0.248 (-0.199) |
| Br | -0.389 (-0.249) | -0.124 (-0.042) | -0.118 (-0.332) | -0.275 (-0.220) |

[a] CM3 charges scaled by 1.14 from QM/MM/MC simulations. Atom numbering scheme given in Figure 3.1.

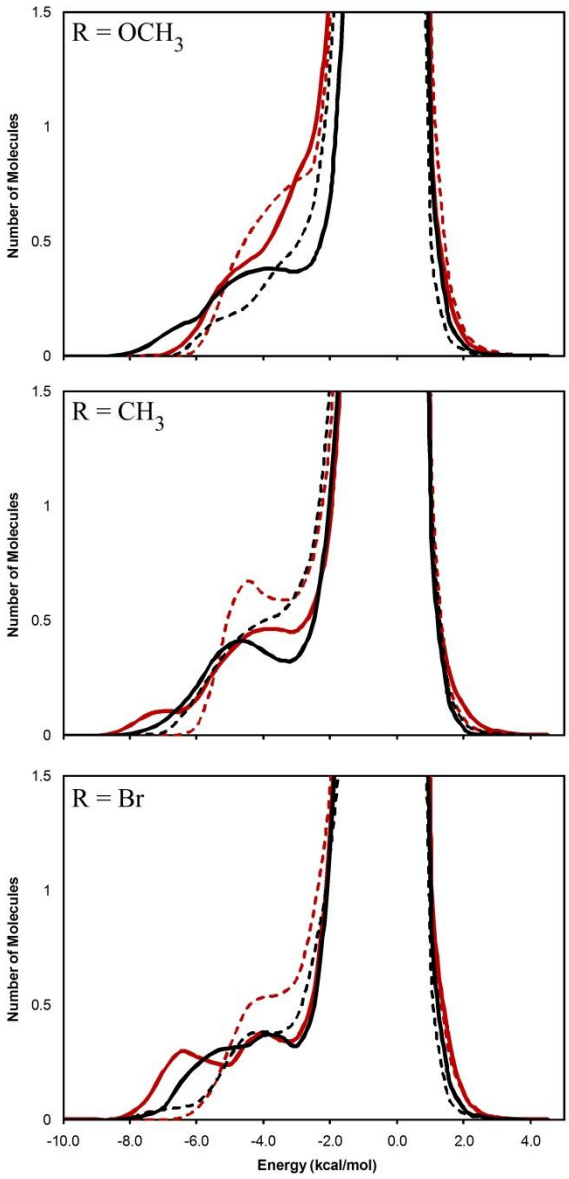



**Figure 3.4:** Solute-solvent energy pair distributions for the Claisen rearrangement of allyl *p*-R-phenyl ethers: "in water" transition structure (solid red), "in water" reactant (dashed red), "on water" transition structure (solid black), and "on water" reactant (dashed black) at 25 °C.

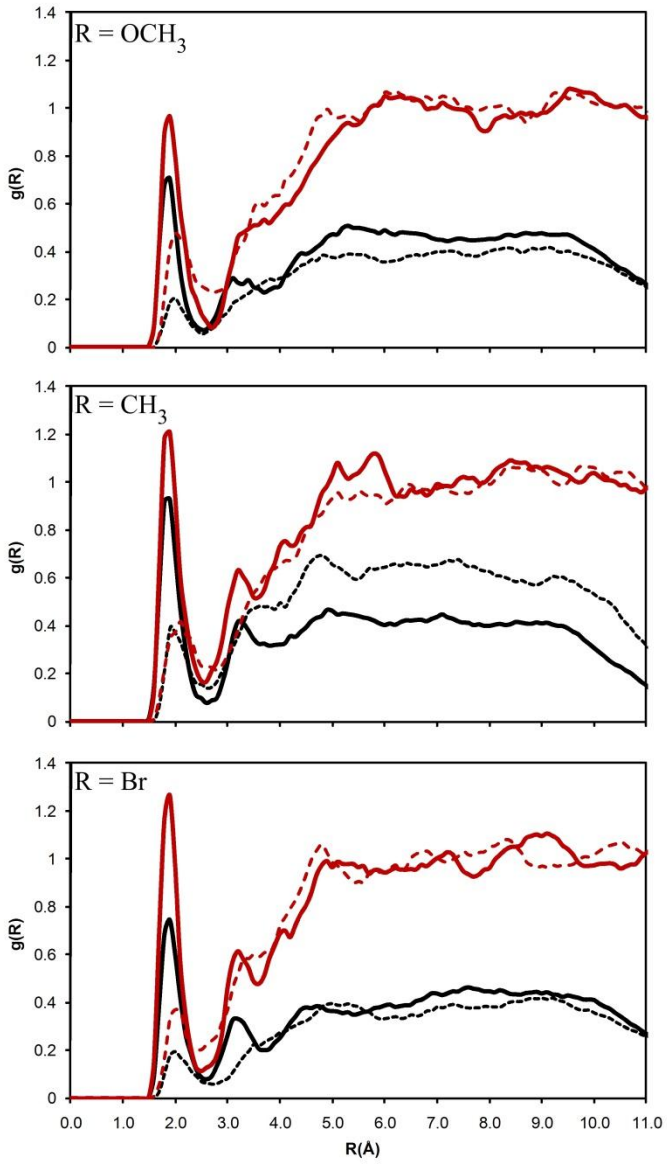



**Figure 3.5:** Computed O(ether)–H(water) radial distribution function for the Claisen rearrangement reactions of allyl *p*-R-phenyl ethers: "in water" transition structure (solid red), "in water" reactant (dashed red), "on water" transition structure (solid black), and "on water" reactant (dashed black) at 25 °C.

### 3.9 Radial Distribution Functions

The solute-solvent structure for the Claisen rearrangements "in water" and "on water" can be further characterized by radial distribution functions, $g(R)$. Hydrogen bonding between the reacting oxygen of the allyl *p*-R-phenyl ethers and the hydrogen of water, O(ether)–H(water), should yield contacts shorter than 2.7 Å. The corresponding $g_{OH}(R)$ gives the probability of finding a hydrogen of a water at a distance $R$ from the reacting oxygen of the ether. Accordingly, both the "in water" and "on water" rearrangement reactions show a well-defined first peak centered around 1.9 Å with a minimum around 2.5-2.7 Å that reflects the hydrogen bonds (Figure 3.5). In both cases, the hydrogen bonding is clearly greatest for the activated complex in all three substituents, however the "in water" transition structure shows a larger number of O(ether)–H(water) interactions with the homogeneous water molecules (red solid line in Figure 3.5) compared to the interfacial water interactions for the "on water" transition structure (solid black line). Integration of the peaks from 0 to 2.5-2.7 Å confirm a reduced number of site-specific O(ether)–H(water) hydrogen bonds for the allyl *p*-R-phenyl ether transition structures "on water" when compared to the "in water" environment, i.e. 1.3 from 1.7 for R = $OCH_3$, 1.9 from 2.0 for R = $CH_3$, and 1.6 from 1.8 for R = Br (Table 3.7).

The resultant destabilization of the transition structure "on water" correlates well with the predicted reduction in rates compared to the homogeneous aqueous conditions. Inspection of the orientations of the solutes at the interface found the aromatic ring to lie

flat on the water surface with a tilt towards the reacting oxygen throughout the reaction pathway when hydrophobic substituents, i.e. $CH_3$ and Br, were included on the ring. Site-specific interactions between the hydrophobic and hydrophilic substituents and the interfacial waters are predicted to act as an "anchor" for the molecules at the oil/water interface dictating their overall accessibility to the surface water. The more hydrophilic the substituent, the larger the tilt towards the substituent group on the water surface and, as a result, the larger the tilt away from the reacting oxygen (similar to a "see-saw" motion). The computed geometry orientations, radial distribution functions, and solute-solvent energy pair distributions all predict a greater loss of hydrogen bonds for the hydrophilic $OCH_3$ substituted rearrangement reaction "on water" compared to the hydrophobic $CH_3$ and Br groups. These results are consistent with recent experimental tilt-angle measurements of coumarins at the air/water interface using electronic sum frequency generation (ESFG) spectroscopy; the most hydrophobic coumarins were reported to experience an intermediate polar environment, whereas moderately water-soluble ones felt the most nonpolar environment.[189]

The interactions between O(ether)–H(water) in the reactant were also reduced for the "on water" reactions compared to "in water", however, the loss was more dramatic for the $OCH_3$–phenyl system compared to the more hydrophobic substituents. For example, integrating the first peak to 2.6–2.7 Å for the R = $OCH_3$ reactant yields 1.4 and 0.7 hydrogen bonds "in water" compared to "on water", respectively, while the number for R = $CH_3$ remains relatively constant at 1.2 and 1.1 (Table 3.7). Despite a loss in favorable site-specific hydrogen bonds for the more hydrophilic ether, the strong

electron-donating ability of $OCH_3$ relative to $CH_3$ provides an intramolecular stabilization at the transition state that is able to offset the loss of hydrogen bonding and provide an overall faster rate of reaction.

**Table 3.7:** Number of Water Molecules Interacting with the Reacting Oxygen of Allyl *p*-R-Phenyl Ethers and Allyl Naphthyl Ether for the Claisen Rearrangement Reactant (GS) and Transition Structure (TS) "in water" and "on water" from Radial Distribution Functions[a].

| R | "in water" GS | "in water" TS | "on water" GS | "on water" TS |
|---|---|---|---|---|
| $OCH_3$ | 1.4 | 1.7 | 0.5 | 1.3 |
| $CH_3$ | 1.2 | 2.0 | 1.1 | 1.9 |
| Br | 0.8 | 1.8 | 0.6 | 1.6 |
| naphthyl | 1.0 | 1.3 | 0.9 | 1.5 |

[a]Integration of the first peak of Figures 5 and S7 to 2.5-2.7 Å.

### 3.10 Hydrophobic Effects

To focus on the role hydrophobic effects play in the rate acceleration of aromatic Claisen rearrangements, the variation of the solute's solvent accessible surface area (SASA) at the ground and transition state were examined. The same pattern is observed for the allyl *p*-R-phenyl ethers and allyl naphthyl ether: the reactant has a SASA which has a larger exposed surface than the activated complex by ca. 20-25 $Å^2$. It is known that the change in SASA and free energy of hydration for hydrocarbon systems show a linear correlation with a proportionality constant of about 0.01 kcal/mol/$Å^2$.[169, 190] This suggests that the hydrophobic effect should not contribute more than 0.5 kcal/mol to the reduced free energies of activation for these rearrangements "in water." The results are consistent with early AM1-SM2 calculations by Cramer and Truhlar on allyl vinyl ether (and

derivatives) that predicted the accelerative hydrophobic effect to be small and not sensitive to structure.[167] Solutes utilizing the "on water" environment benefit less from the hydrophobic effect as the allyl hydrocarbon portion was found to minimize the interaction with water by occupying almost exclusively the gas-phase portion of the air/water interface. The ethers with hydrophilic substituents on the phenyl ring were particularly affected as the radial distribution calculations found a reduced number of water interactions with the allyl chain.

### 3.11 Continuum Solvent Models

The polarizable continuum model (PCM) presents an alternative method for exploring solvent effects.[161] However, *ab initio* and DFT calculations coupled to continuum solvent models have been consistently unsuccessful in predicting rate differences between protic and aprotic solvents for a large variety of organic reaction types.[141] For example, previous calculations on allyl *p*-tolyl ether using B3LYP/6-31G(d) and the integral equation formalism variant of PCM (IEFPCM) failed to predict the accurate sequence of reaction rates in five different solvents.[171] However, multiple improvements to the PCM framework[191-192] have been subsequently implemented into the newly released Gaussian 09 software to enhance accuracy.[159] For example, IEFPCM now uses Karplus and York's conductor screening model for building the solute's cavity and computing the reaction field.[193] In light of the implication of solvent on rates, the new IEFPCM methodology has been tested in this work in conjunction with B3LYP/6-311+G(2d,p) to optimize transition structures and ground states for the allyl *p*-R-phenyl ether and allyl naphthyl ether rearrangements (Tables 3.4, 3.8 and 3.9). Geometries for

the stationary points on the free-energy profiles were initially located in vacuum at the CBS-QB3 and B3LYP/6-311+G(2d,p) theory levels. The gas-phase free-energy activation barriers were found to be in reasonable agreement with experimentally measured values (Table 3.8). For example, $\Delta G^{\ddagger}$ values of 35.1 and 34.5 kcal/mol for the allyl *p*-tolyl ether rearrangement using CBS-QB3 and B3LYP/6-311+G(2d,p), respectively, were in good agreement with the experimentally reported value of 38.5 kcal/mol.[31] The PDDG/PM3 method systematically overestimated the gas-phase results in a similar magnitude compared to the solution-phase calculations (Tables 3.2 and 3.8).

**Table 3.8:** Gas-Phase Activation Barriers, $\Delta G^{\ddagger}$ (kcal/mol), at 25 °C for the Claisen Rearrangement of Allyl *p*-R-Phenyl Ethers (R = $OCH_3$, $CH_3$, and Br).

| $\Delta G^{\ddagger}$ | $OCH_3$ | $CH_3$[a] | Br |
|---|---|---|---|
| CBS-QB3 | 35.1 | 35.1 | 35.1 |
| B3LYP/6-311+G(2d,p) | 34.0 | 34.5 | 34.4 |
| PDDG/PM3 | 51.3 | 50.7 | 51.2 |

[a] Gas-phase $\Delta G^{\ddagger}$(exptl) = 38.5 kcal/mol at 170 °C; Ref. [31].

Geometry optimizations at the B3LYP/6-311+G(2d,p)/PCM level were then executed in five different solvents for the allyl naphthyl ether (Table 3.4) and in water and n-pentadecane for the allyl *p*-R-phenyl ether rearrangements (Table 3.9). The DFT/PCM model was unsuccessful in reproducing the observed changes in the rate of reaction. For example, the allyl naphthyl ether reaction was predicted to have essentially identical activation barriers, $\Delta G^{\ddagger}$ of ca. 20 kcal/mol, in water, methanol, acetonitrile, and DMF (Table 3.4), contrary to the experimentally reported differences in rates.[13] Specific

changes in hydrogen bonding are expected to be important along the reaction path, but they are not explicitly treated in the continuum model. The QM/MM/MC calculations clearly overcome this limitation (Tables 3.2-3.4). Ideally the use of B3LYP/6-311+G(2d,p) or higher theory levels coupled to the current QM/MM approach could provide reasonable quantitative accuracy for absolute activation barriers while correctly reproducing the solvent effects. Unfortunately, the use of DFT methods in the QM/MM framework is impractical in view of the need for thorough configurational sampling in the fluid simulations, i.e. 120 million QM single point energy evaluations per Claisen rearrangement free-energy map. An *ab initio* QM/MM approach that includes the solvent reaction coordinate, such as Gao's modern valence bond theory (MOVB) method[194] or Warshel's empirical valence bond (EVB) treatment,[195] may provide an alternative approach for computing "on water" reactions. Either method would be an interesting avenue to pursue in future studies in order to assess any potential solvent fluctuation differences between the "in water" and "on water" environments.

**Table 3.9:** Free Energy of Activation, $\Delta G^{\ddagger}$ (kcal/mol), at 25 °C for the Claisen Rearrangement of Allyl *p*-R-Phenyl Ethers in Solution using B3LYP/6-311+G(2d,p)/PCM.

| R | water (calc) | water (exptl)[a,b] | n-pentadecane (calc) | n-tetradecane (exptl)[a] |
|---|---|---|---|---|
| $OCH_3$ | 31.9 | 33.5 | 33.1 | 36.4 |
| $CH_3$ | 32.3 | 34.4 | 33.7 | 37.4 |
| Br | 32.1 | 34.9 | 33.6 | 37.7 |

[a]170 °C; Ref. [32]. [b] 28.5% EtOH-$H_2O$.

### 3.12 Conclusions

In summary, QM/MM/MC simulations have been carried out for the Claisen rearrangements of allyl *p*-R-phenyl ether, where R = $OCH_3$, $CH_3$, and Br, and allyl naphthyl ether "on water," "in water," and in 14 additional solvents yielding good accord between the computed and observed variations in the free energies of activation. Rate accelerations for the Claisen rearrangements are shown to be correlated to increasing solvent polarity and in the case of the "on water" environment are derived from the ability of the interfacial waters to stabilize the transition state. Destabilization of the reactant also played a large role for the "on water" systems. An orientation analysis of multiple snapshots of the reacting solute relative to the surface of the slab (perpendicular to the z-axis) found that the aromatic ring tended to lie essentially flat on the surface with a slight tilt towards the oxygen when hydrophobic substituents, i.e. $CH_3$ and Br, were included on the ring. The inclusion of hydrophilic substituents, i.e. $OCH_3$, on the ring tilted the reacting oxygen away from the interfacial waters and towards the substituent. Radial distribution functions and solute-solvent energy pair distributions confirm a reduction in the site-specific hydrogen bonding interactions with the surface waters as the aromatic ethers became more hydrophilic. These results are consistent with recent experimental tilt-angle measurements of coumarins at the air/water interface; the most hydrophobic coumarins were reported to experience an intermediate polar environment, whereas moderately water-soluble ones felt the most nonpolar environment.[189] Hydrophobic effects, where the solvent accessible surface area decreases in going from ground to transition state was determined to contribute no more than 0.5 kcal/mol

towards lowering the activation barrier in a homogeneous aqueous environment and were substantially diminished “on water” as much of the reacting allyl chain favored occupying the organic phase. Solvent polarizability was not found to provide a substantial contribution on the rates or geometry of the reaction. High-level *ab initio* and density functional theory (DFT) calculations at the CBS-QB3 and B3LYP/6-311+G(2d,p) theory levels, respectively, coupled to the polarizable continuum model (PCM) were carried out, but did not predict the correct sequence of reaction rates in solution.

Overall, the “on water” reactions benefited most from site-specific hydrogen bonding between the interfacial waters and the transition structure. The addition of hydrophobic substituents increased the polarity felt by the solute at the air/water interface which provided a superior environment for enhanced rate accelerations. Such information could possibly be exploited in the preparation of organic molecules via a safer and more environmentally-friendly “on water” environment.

# Chapter 4

## Understanding the Aldol Reaction from Catalytic Antibodies to "On Water" Organocatalysts using QM/MM Calculations

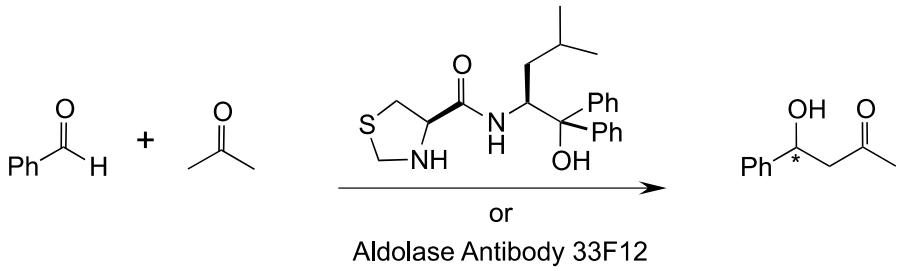
Of particular interest is the aldol reaction, whose rate is thought to benefit from hydrophobic effects in combination with specific hydrogen bonding. The 'on water' conditions and aldolase antibodies 33F12 provide excellent environments to study the effect of hydrophobicity on the rates, stereoselectivity, and the enamine mechanism of the aldol reaction. QM/MM and Monte Carlo simulations coupled to our highly efficient polynomial quadrature free-energy perturbation method were performed to clarify the enamine mechanism and the intermolecular interactions responsible for the amine catalysis and enantioselectivity using on water conditions and in aldolase antibody 33F12. The aldolase antibody proceeded through a stepwise mechanism with a transition structure bond distance for the carbon-carbon bond of 2.02Å. The C-C bond formation was the rate limiting step with a $\Delta\Delta G^{\ddagger}$ of 10.2 ± 1 kcal/mol, where the proton transfer occurs rapidly. The aldol reaction between acetone and benzaldehyde with the organocatalyst is presently being carried out 'on water', in water and in acetone for the *re* and *si* face. Transition structure geometries and free energies will be evaluated for these systems. Due to the enhanced enantioselectivities that have been shown for the aldol reaction of benzaldehyde and acetone with chiral organocatalysts, extensive analysis including solute-solvent energy pair distributions and radial distribution functions will be performed to examine the role water plays on the system.

## 4.1 Introduction

The enantioselective aldol reaction is an important C–C bond formation reaction that has received considerable attention following the development of small proline-derived enamine-based catalysts by List, Barbas, and Lerner.[183-184, 194, 196-198] While the use of (*S*)-proline to catalyze asymmetric intramolecular aldol reactions began in the 1970's,[47, 172, 188] more recent derivatives of (*S*)-proline and (*S*)-prolinamide for the direct intermolecular aldol reaction between two carbonyl molecules have been developed that deliver high yields, 99:1 diastereoselectivies, and enantioselectivities > 99% ee.[16, 18, 35, 158, 167-168, 174-175, 178, 195, 199-203] Polar organic solvents have been widely used as a suitable solvent for the aldol reaction; however, many contemporary investigations have favored water as the reaction medium due to its low cost, ease of product isolation, and unique influence on the stereoselectivity of enamine-based organocatalytic aldol reactions.[180-182, 204-205] Of particular interest are the recently reported aqueous-phase enantioselective catalysts[35, 201] by Singh and coworkers composed of a thiazolidine ring and a chiral diphenyl amino alcohol -- similar in structure to those first developed by Barbas[206] and Hayashi.[186, 207] For example, the reaction between acetone and benzaldehyde using a Vishnumaya and Singh organocatalyst (Scheme 4.1) was reported to complete in 6 hours in pure water compared to >16 hours in acetone and the subsequent product was obtained at 93% ee compared to 85% ee, respectively.[35]

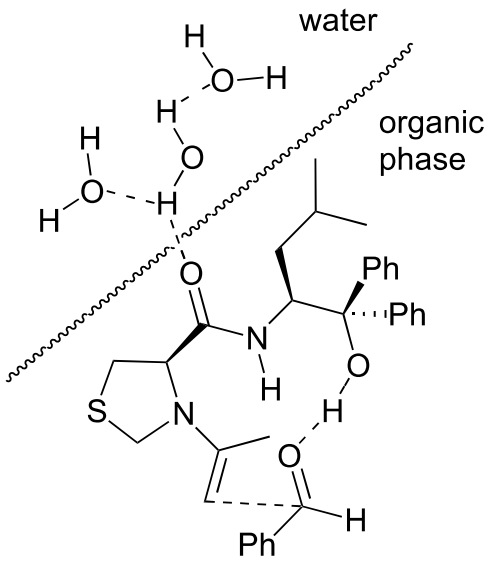
Singh and coworkers have proposed that the enamine-based aldol reaction proceeds via an "on water" environment that segregates the transition state away from the water molecules; this allows the reactants to benefit from the hydrophobic portion while

stabilizing the amide oxygen via specific hydrogen bonding with the surface waters (Scheme 4.2).[35, 201] "On water" conditions were first reported by Sharpless to give large rate increases and enhanced yields for a variety of reactions. They are defined as a




**Scheme 4.1:** Aldol reaction between acetone and benzaldehyde using an aqueous-phase organocatalyst or catalytic antibody 33F12.

reaction that proceeds in an aqueous organic emulsion prepared by vigorously stirring insoluble reactants with water.[13-14] In contrast, reactants "in water" are dissolved homogeneously in water primarily through the use of co-solvents.[15, 208] Superior yields and enantioselectivities for the direct aldol reaction of acyclic and cyclic ketones with different aldehydes in brine[35, 201] relative to in water suggests an important 'salting out effect' indicating that water itself provides more than simply a medium for reaction.

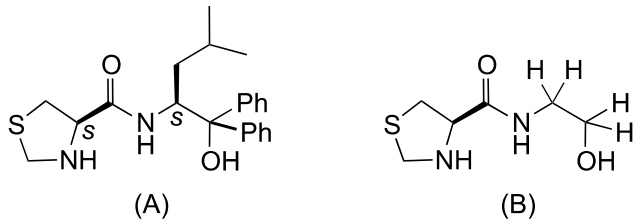



**Scheme 4.2:** Proposed transition state for an "on water" enamine-based aldol reaction.

Interestingly, catalysis of the aldol reaction via aldolase catalytic antibodies[57-58, 209-210] laid the foundation for the asymmetric proline-based organocatalysts; however, mechanistic study of aldolase antibodies has proven to be more limited and difficult as compared to the organocatalysts.[157, 167, 178] A recently resolved crystal structure for catalytic antibody 33F12 with a 1,3-diketone has allowed direct observation of the enamine for the first time.[211] As enamines are generally unstable in water, both the antibody and on water reactions appear to rely on hydrophobicity to catalyze the reaction. For example, the p$K_a$ of the ε-amino group of LysH93 (which reacts with acetone to form the enamine during reaction) has been estimated at 5.5 for antibody 33F12,[209] whereas the p$K_a$ in free solution is 10.5.[58] In water, typical amines would be protonated and ineffective for reaction. A detailed atomic-level study upon the aldol reaction in both

enzymatic and aqueous conditions would help elucidate the role of hydrophobic effects and specific hydrogen bonding upon the mechanism and selectivity.

Mixed quantum and molecular mechanical (QM/MM) calculations coupled to free energy perturbation theory and Monte Carlo sampling (MC/FEP) are carried out here to clarify the enamine mechanism and the intermolecular interactions responsible for the amine catalysis and enantioselectivity using on water conditions and in aldolase antibody 33F12. The enamine mechanism with multiple organocatalysts has been investigated using DFT methods by Houk and coworkers and provides a basis for comparison;[165, 190, 197, 212-216] however, the origin behind the on water and antibody enhancements has not been examined. The reaction between acetone and benzaldehyde (Scheme 4.1) was simulated in this study with fully explicit environments using the aqueous organocatalysts shown in Scheme 4.3 and with antibody 33F12. The direct aldol reaction between acetone and benzaldehyde provides an immediate comparison between the diverse reaction conditions due to the availability of experimental data in both the aqueous and protein environments.[49, 209, 211] Insight into the experimentally observed increase in enantioselectivity is given by computing the approach of the aldehyde from both the *re* and *si* faces. Additional calculations were performed in a homogenous water solution and acetone to clarify the role of solvent. This work formulates a comprehensive theory on how both the aqueous surface boundary and the aldolase active site use favorable hydrophobic effects and specific hydrogen bonding to enhance rates and selectivity for the aldol reaction.




**Scheme 4.3:** (A) Full organocatalyst and (B) reduced organocatalyst.

## 4.2 Methods

**4.2.1 Setup and Processing for Antibody 33F12**. Initial coordinates for the protein-based aldol system was obtained from a reported 1.9 Å crystal structure (PDB ID: 3FO9) of aldolase antibody 33F12 complexed with a 1,3-diketone hapten.[211] The raw PDB file was prepared for input to *MCPRO* through the *chop* program.[156] *Chop* outputs directives for the *pepz* program, which adds missing hydrogens, performs the residue truncation and capping, and converts the PDB file to a Z-matrix with OPLS atom typing that is suitable for input to *MCPRO*. Residues more than 15 Å from the binding site were removed for computational efficiency, which left one active site and 135 residues, including the light and heavy chains of the antibody nearest to the ligand. The protein fragments furthest from the active site were made neutral, so no counterions were added. Benzaldehyde and the enamine formed between acetone and LysH93 were built and optimally positioned in the binding site by overlaying with the existing hapten ligand. The initial structure was relaxed via a conjugate gradient optimization using a distance-dependent dielectric constant of 4*r*.

**4.2.2 QM/MM Method.** All solutes and reacting antibody side-chains were treated using the PDDG/PM3 semiempirical QM method, which has given excellent results for a wide variety of organic and enzymatic reactions in the solution-phase.[5, 75, 142, 191, 217] QM/MM calculations were run on a Linux cluster at Auburn University and on computers located at the Alabama Supercomputer Center. Protein simulations were carried out using the program *MCPRO* while the condensed-phase reactions utilized *BOSS*.[156]

*Condensed Phase*. For the "on water" reactions, reactants and transition states were located in an explicit 1220 water molecule box represented using the TIP4P water model.[139] Periodicity was removed from the z-axis of the water box and the solutes were placed on top of the z-axis; an equilibrated NVT water slab was used to prevent a very slow drift of the water molecules that reduces the exposed surface, i.e. stretch in the z-direction. Periodic boundary conditions have been applied to tetragonal boxes and solute-solvent and solvent-solvent cutoffs of 12 Å were employed with quadratic feathering of the intermolecular interactions within 0.5 Å of the cutoff. Computation of the QM energy and atomic charges is performed for each attempted move of the solute, which occurs every 100 configurations. For electrostatic contributions to the solute-solvent energy, CM3 charges[153] were obtained for the solute with a scaling factor of 1.14. Lennard-Jones interactions between solutes and solvent atoms were taken into account using OPLS parameters. This combination is appropriate for a PM3-based method as it minimizes errors in the computed free energies of hydration.[218] Changes in free energy were calculated using free energy perturbation (FEP) theory in conjunction with NPT Metropolis Monte Carlo (MC) simulations at 25 °C and 1 atm. Each FEP window

entailed ca. 40 million (M) MC configurations of equilibration and 20 M configurations of averaging.

A new OPLS-AA fully flexible solvent box for acetone was constructed in a fashion similar to previous work.[5, 143] Briefly, the liquid-phase simulations were carried out by placing 400 solvent molecules at random positions in the simulation box. The box was then equilibrated at 25 °C for 225 million MC steps in the NPT ensemble.

*Aldolase antibody 33F12*. In our QM/MM implementation of the aldol reaction, benzaldehyde, the enamine formed between acetone and LysH93, and TyrL36 were treated with the PDDG/PM3 semiempirical QM method. The reminder of the protein utilized the OPLS-AA force field.[219] The interactions of overlapped atoms in the QM and MM regions are described through "link atoms" using hydrogens in the QM calculation.[150] The connection of the QM and MM regions requires the inclusion of the classical bond stretching, angle bending and torsion terms if any MM atom is involved in the interaction. CM3 charges were obtained for the QM active site with a scaling factor of 1.12. At the beginning of the MC simulations, a water cap with 22 Å radius (ca. 900 TIP4P waters) was added. A half-harmonic potential with a force constant of 1.5 kcal $mol^{-1}$ $Å^{-2}$ was applied to water molecules at a distance greater than 22 Å. To ensure that the final orientations of the water molecules within the aldolase binding pocket were not an artifact of the water cap, multiple reaction pathway simulations with different caps were carried out.

**4.2.3 Monte Carlo Simulation Protocol.** All simulations were run at 25 °C using MC statistical mechanics. Adjustments to the allowed ranges for rotations, translations, and dihedral angle movements for the solution-based calculations led to overall MC acceptance rates of 30 – 50% for new configurations. The ranges for bond stretching and angle bending were set automatically by the *BOSS* program on the basis of force constants and temperature. For the antibody, only the bond angles and dihedrals of side chains of residues with any atom within 10 Å of the center of the system are varied. All degrees of freedom in the QM region are varied, except the carbon atom adjacent to the link atom and those involved in the reaction coordinates. Each simulation for a FEP window consists of 5 x $10^6$ (5M) configurations of solvent relaxation, where the water molecules were moved randomly while keeping the protein and substrate fixed, followed by 10M configurations of full equilibration where all degrees of freedom are varied, and 25M configurations of averaging, where all degrees of freedom were sampled and the free-energy changes were obtained. Stability in the computed energies, free energies, and volume were used to monitor convergence of the MC simulations. During the simulations, 10% of the attempted MC moves involve the active site and 1% the QM region. Each move of the QM region requires one self-consistent field QM calculation if the move is rejected and three if accepted (reference and two perturbed structures). Therefore, each protein FEP window requires ca. (0.5-1)M QM calculations; the total for this work exceeds 500M highlighting the need for very fast QM energy and charge evaluation.

**4.2.4 Polynomial Quadrature Method.** In our recent work elucidating the mechanism for fatty acid amide hydrolase (FAAH) and antibody 4B2, significant technical advances were reported for the treatment of proton transfer reactions.[150, 191] For a typical proton transfer, O–H···O' → O···H–O', it was found that the O···O' distance remains relatively constant and that $r$(O–H) – $r$(H–O') can be used to compute a 1-D PMF. This normally requires approximately 30 to 50 double-wide FEP windows using 0.02 Å $\Delta r$ increments, while ca. 900 windows would be needed for a 2-D PMF using two distances as reaction coordinates. However, free-energy changes for individual windows can be fitted almost perfectly by a 5$^{th}$ order polynomial. Analytical integration yields a sextic polynomial for the overall proton-transfer PMF. This allows for the accurate construction of the full PMF using only 7 FEP windows instead of the usual 50.[143, 191, 217] The largest deviation found between the approximate and the detailed calculation was 0.5 kcal/mol.

## 4.3 Results and Discussion

**4.3.1 Aldol Reaction in Antibody 33F12.** Potentials of mean force (PMF) calculations were used to build a free energy map, $\Delta G$ (kcal/mol), for the aldol reaction in antibody 33F12 using the reaction coordinates between the LysH93-enamine and benzaldehyde C-C bond ($R_{CC}$) (Figure 4.1) and the proton transfer, $R_1 - R_2$, between the TyrL36 and benzaldehyde oxygen atoms (see Figure 4.1). A fixed distance of $R_1 + R_2$ = 2.6 Å was established to be appropriate from test simulations and prior work.[150] Kinetic experiments have determined the C-C bond formation step highlighted in Scheme 4.4 to be both rate- and enantioselectivity-determining.[220] In addition, TyrL36 is believed to function as a

general acid and base in antibody 33F12 as shown in the projected mechanism (Scheme 4.5).[211] The role of TyrL36 has been supported by the TyrL36Phe mutant of aldolase antibody 38C2, which does not bind to the diketones, but yields an (*S*)-aldol product between a ketone and an aromatic aldehyde similar to 33F12.[211]

The enzymatic reaction was predicted to follow a stepwise mechanism where the enamine benzaldehyde C-C distance began at a separation of 2.22 Å as larger distances became energetically unfavorable due to poor steric interactions within the active site pocket. The $R_{CC}$ bond distance at the transition state was calculated to be 2.02 Å and proceeded to a geometry of 1.54 Å for the intermediate where the C-C bond is fully developed, but the proton transfer had not occurred (Figure 4.2). A calculated $\Delta G^{\ddagger}$ of 10.2 ± 1 kcal/mol was found to be rate-limiting and the subsequent proton transfer between the TyrL36 and benzaldehyde oxygens occurred rapidly with a $\Delta\Delta G^{\ddagger}$ of 2.1 kcal/mol. The energetics are consistent with a previously computed MP2/6-31G(d,p) value of 10.1 kcal/mol for the "theozyme" aldol reaction between a primary enamine and acetaldehyde using an implicit PCM solvent model (a permittivity value of $\varepsilon = 10.4$ was used for similarity to *n*-octanol medium).[221] The MP2 optimized transition state $R_{CC}$ distance of 1.944 Å was also reasonable compared to the current geometry value of 2.02 Å. DFT-based calculations resulted in similar energies and geometries.[212] The QM/MM $\Delta G$ of reaction computed was −49.2 kcal/mol between the enamine reactants and iminium cation intermediate.

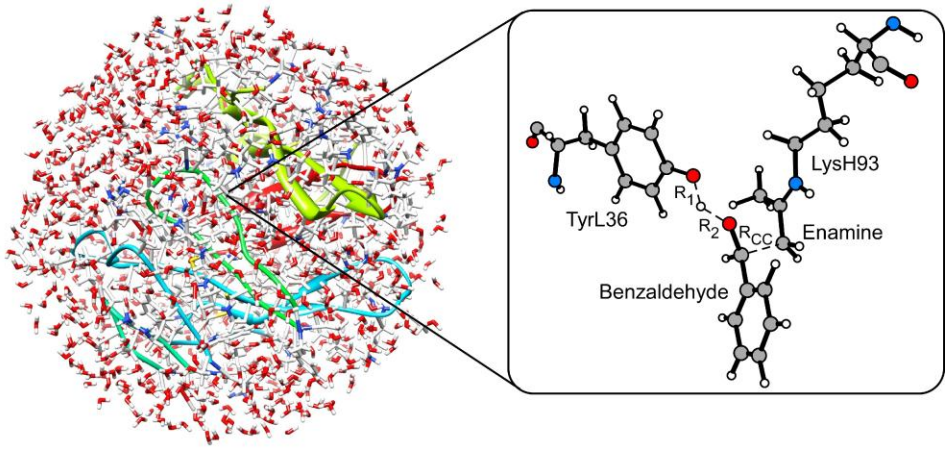
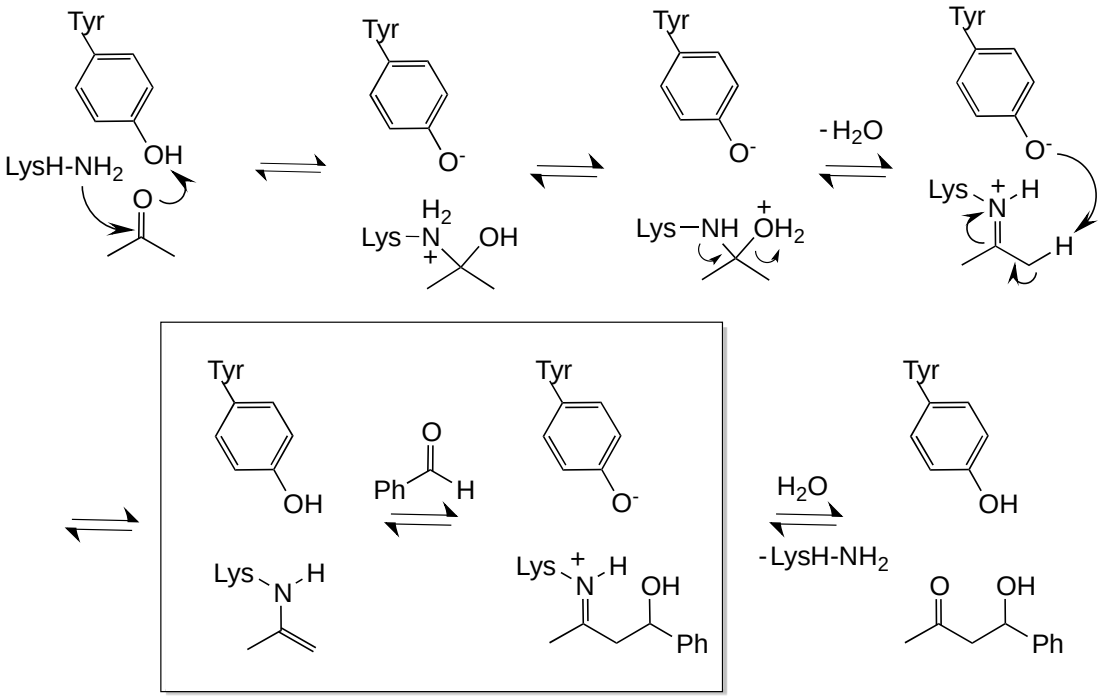



**Figure 4.1**: Illustration of reduced antibody 33F12 consisting of 135 residues, benzaldehyde, TyrL36, the enamine formed between acetone and LysH93, and a 22 Å water cap.




**Scheme 4.4:** Proposed enamine mechanism for aldol reaction in antibody 33F12.[211] The rate- and enantioselectivity-determining step is emphasized within the box.

A strong hydrogen bond network featuring TrpH95 and hydrophilic residues SerH100 and AsnL34 helped anchor the TyrL36 residue in an appropriate position to participate in a proton transfer and adopt the correct enantioselective *si*-face orientation for benzaldehyde with the LysH93-enamine (Figure 4.3). Inspection of the active site finds the *re*-face of benzaldehyde would incur poor steric interactions with TyrL36 and its supporting residues. Docking simulations of reduced DFT-based transition structures of methylamine in place of the ε-amino group into the 33F12 active site also supported a geometric origin for the enantiopreference of (*S*)-aldol products.[49] From the current simulations, TrpH95 points its hydrogen atom at the oxygen of TyrL36 at an average distance of 2.20 and 2.49 Å at the ground and transition states of the reaction pathway, respectively. The SerH100 forms tighter H-bonds with the TyrL36 oxygen at distances of 2.05 and 1.99 Å, for the reactants and transition states, respectively, while AsnL34 provides longer H-bond distances of 3.67 and 2.96 Å.

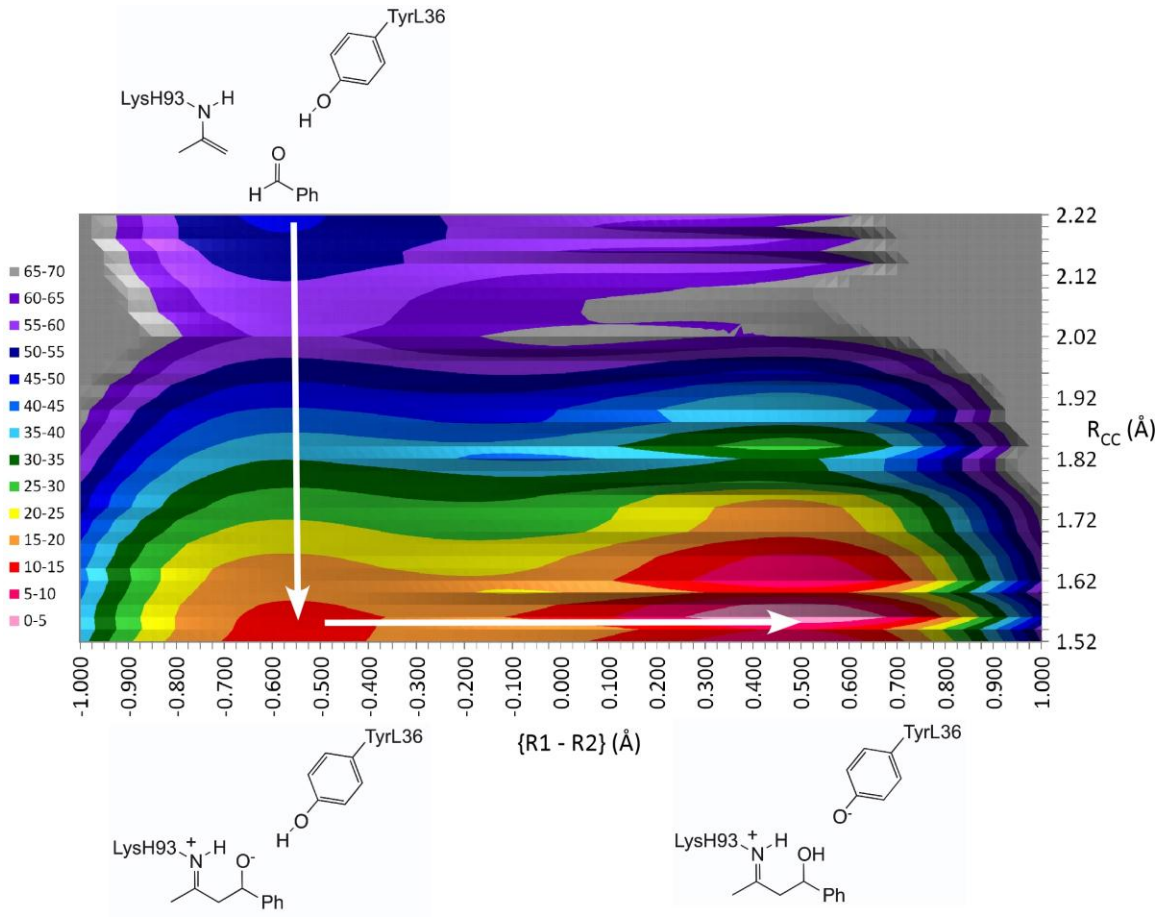



**Figure 4.2:** Free-energy profile (kcal/mol) for the aldol reaction between benzaldehyde, the enamine formed between acetone and LysH93, and TyrL36 in antibody 33F12. Maximum free-energy values are truncated to 70 kcal/mol for clarity.

The reactive LysH93 residue in catalytic antibody 33F12 is surrounded primarily by hydrophobic residues resulting in a microenvironment suggested to approximate that of *n*-octanol.[209] However, following the formation of the C-C bond and proton transfer, interaction with water is crucial for the hydrolysis of the iminium cation intermediate towards the final aldol adduct (Scheme 4.5). In general, the implications of water-mediated active sites are broad, as a study of theoretically designed and experimentally

tested retro-aldol enzymes found that the explicit inclusion of water molecules into the active site gave significant rate enhancements of up to four orders of magnitude greater than those that simply relied on charged side-chain networks.[222] In addition, our computational study of catalytic antibody 4B2 found the role of water to be essential in the catalysis of the mechanistically different Kemp elimination and allylic rearrangement reactions.[191]

A snapshot taken at the end of the Monte Carlo simulations near the transition state region found up to 6 water molecules within 5.0 Å of the benzaldehyde. One water molecule was located near the enamine throughout the reaction forming an average interaction distance of 2.20 Å with the nitrogen at the reactants and a longer value of 2.61 Å at the transition state. The distance between the same water molecule and the terminal reacting $sp^2$ carbon on the LysH93-enamine was 2.83 and 2.56 Å at the ground and transition states, respectively. In addition, a water molecule was found to hydrogen bond with the π-system of the aromatic benzaldehyde ring (Figure 4.3). There were no waters interacting directly with TyrL36 throughout the C-C bond forming process. Favorable site-specific electrostatic stabilization of the TyrL36 phenoxide was provided exclusively by the adjacent hydrophilic residues discussed previously. However, water may play a larger role during the other stages of the mechanism, e.g., the hydrolysis of the iminium.

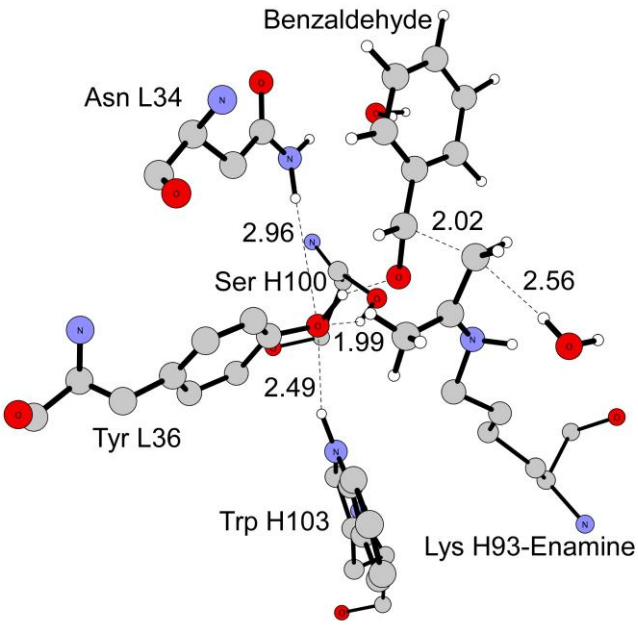



**Figure 4.3:** A close-up of the aldol reaction transition state between the LysH93-enamine and benzaldehyde in the active site of antibody 33F12 with 2 nearby water molecules retained. Average distances over the final 25 million configurations given in angstroms.

### 4.4 Future Work

**4.4.1 Aldol reaction with an Organocatalyst in the Gas Phase.** Free energies were computed for the aldol reaction between acetone and benzaldehyde with an organocatalyst for both the *re*- and *si*- face. A reduced organocatalyst was utilized (Scheme 4.3b) for comparison with the aldol reaction in solution. Reactant and transition structure geometries were found using the PDDG-PM3 semi-empirical method with Gaussian 09.[159] Similar to the aldolase antibody, two transition structures were found, where the $\Delta\Delta G_1^{\ddagger}$ (C-C bond formation) between the *re*- (Figure 4.4) and the *si*- (Figure 4.5) face was computed to be 4.9 kcal/mol and the $\Delta\Delta G_2^{\ddagger}$ (proton transfer) was computed

to be 1.1 kcal/mol. The enantiomeric excess was calculated to be 71.7%. The *re*- and *si*-faces are both present in the gas phase, whereas in the presence of water the reaction shows a high %ee for the *re*- face.[35]

$$er = e^{-\Delta\Delta G/RT} \tag{4.1}$$

$$\%ee = \frac{1-er}{1+er} * 100\% \tag{4.2}$$

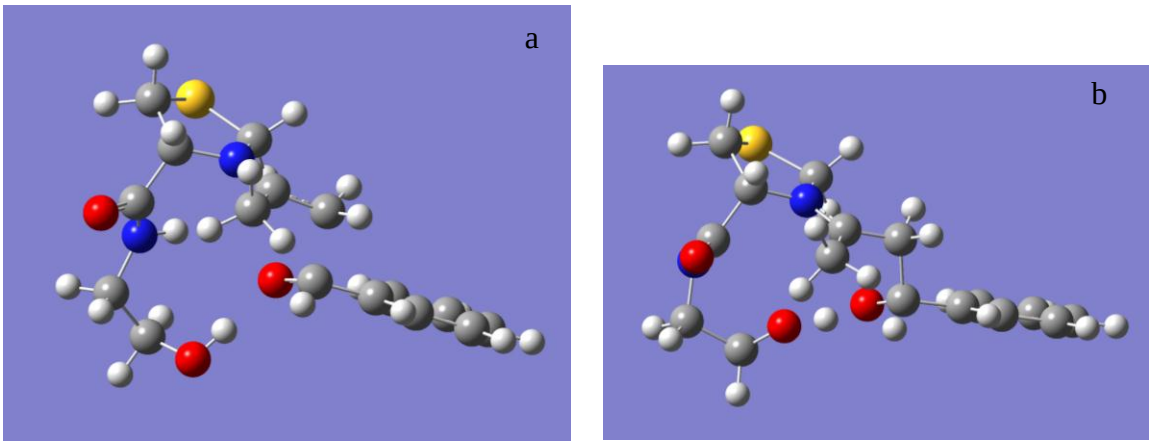



**Figure 4.4:** *re*- face transition state geometry (a) is C-C bond formation transition structure, and (b) is the proton transfer transition structure.

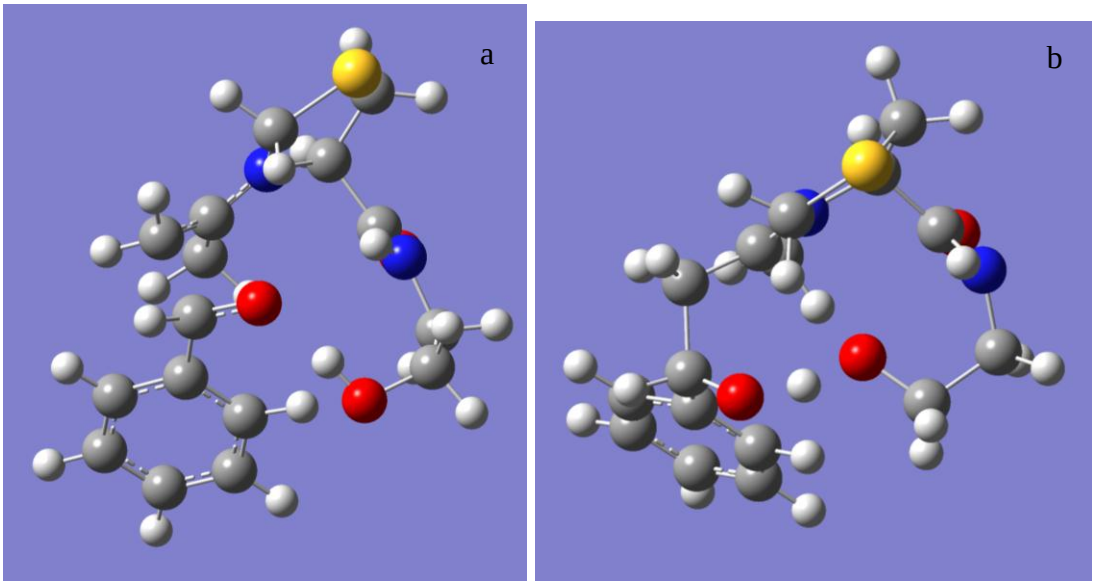



**Figure 4.5:** *si*- face transition state geometry (a) is C-C bond formation transition structure, and (b) is the proton transfer transition structure.

**4.4.2 Aldol Reaction with an Organocatalyst in Solution**. Free energy maps were computed for the aldol reaction between acetone and benzaldehyde using the organocatalyst[35] from Scheme 1 on water and in a neat reaction (acetone as solvent). Given the extensive computational resources required to properly sample the enamine-based aldol reaction in multiple solvents from the *re*- and *si*-faces of benzaldehyde, a reduced catalyst was utilized where the phenyl rings and isobutyl substituents were replaced with hydrogens (Scheme 4.4). It has been experimentally determined that stereospecific phenyl groups at the β–carbon atom are not essential for enhanced catalytic performance and the substituents' primary function is to form a concentrated organic phase with hydrophobic reactants.[35] Our on water QM/MM methodology will maintain both enamine-based systems at the surface of water throughout the simulations (Figure

4.6); however, statistical orientation calculations were performed for the full and reduced catalytic systems to compare similarities between the catalysts at the surface of the water. A plane was constructed between the O on benzaldehyde, N on the thiazolidine ring, and the C bonded to the hydroxyl group and the angle tilt was calculated with respect to the z-axis of the box of water. The average angles were 96 degrees for the full model and 80 degrees for the reduced model with approximately ± 17 degree deviations during the test simulations. In addition, the solute-solvent interactions between the enamine-based reactions transition structure and the water surface will be discussed once the future work is completed.

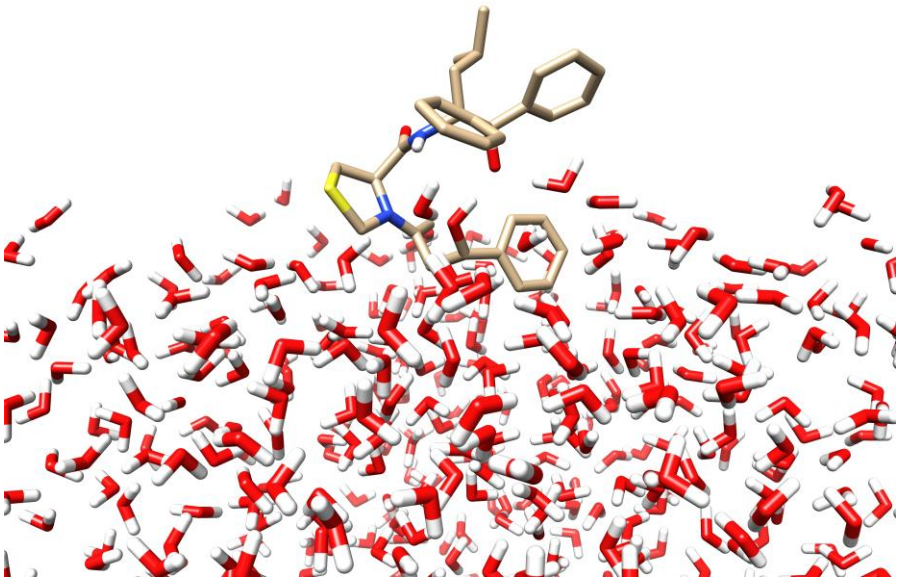


**Figure 4.6:** Illustration of the organocatalyst-enamine aldol reaction with benzaldehyde on water from the QM/MM MC/FEP simulations.

1D-PMF calculations are being carried out where the enamine-benzaldehyde C-C bond ($R_{CC}$) was perturbed in increments of 0.01 Å to a final 5 Å separation for in water,

on water, and in neat (acetone) solutions. Future comparisons of the solvents with the catalytic aldolase antibody in hopes to find that key water molecules provide favorable hydrogen bonding at the transition state, multiple effects may play a crucial role in the aldol reaction catalysis, e.g., orientation of the solute on the aqueous boundary, ground state destabilization, hydrophobic effects, and polarity of water at the surface.

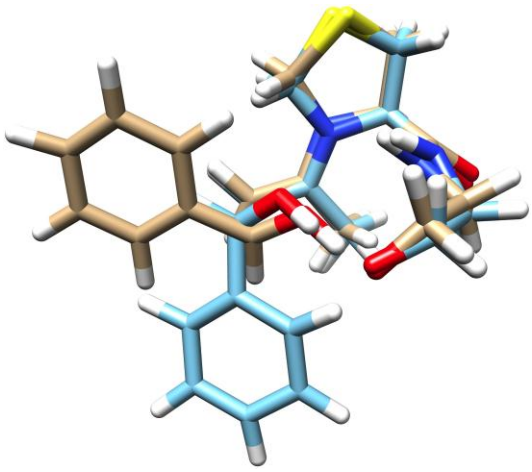


**Figure 4.7.** Overlay of the organocatalyst-enamine transition states with the favorable *re*-face (beige) and unfavorable *si*-face benzaldehyde (blue) conformations.

## 4.5 Conclusion

In summary, QM/MM/MC/FEP calculations have been carried out to compute the free energy values for the aldol reaction of acetone and benzaldehyde with the reduced organocatalyst "on water", in water and in acetone and for the catalytic aldolase antibody 33F12. The aldolase antibody yielded the enantiopreference for the (*S*)-aldol product due the geometric orientation of hydrogen bonding interactions from TrpH95, SerH100 and

AsnL34 with the proton transfer amino acid TyrL36. In addition, the *re*-face orientation of benzaldehyde addition during the aldol reaction would yield potentially unfavorable steric interactions with TyrL36. Hydrogen bonding played an important role in the catalytic antibody, therefore the reaction in the condensed phase should benefit from site-specific hydrogen bonding interactions as well. Gas phase simulations gave a low enantiomeric excess of 71.7% and future work of the aldol reaction in a neat solution, in water, and "on water" should elucidate the role of solvent on enantioselectivity preferences.

# Chapter 5

## Flavin-Induced Active Site Conformational Changes in Alkanesulfonate Monooxygenase Predicted from Molecular Dynamic Simulations

Alkanesulfonate monooxygenase (SsuD/E) is a two-component enzyme triggered in times of sulfur starvation. An apo SsuD crystal structure (PDB ID: 1M41) has been reported, however structures containing bound substrates is not known due to the large conformational changes occurring for the enzyme. Homology modeling was carried out to insert missing residues 250-283 into 1M41, and docking was used to insert octanesulfonate and the flavin cofactors into SsuD. Extensive molecular dynamics simulations and analysis were performed on the apo WT SsuD and WT SsuD with bound substrates: reduced flavin ($FMNH_2$), C4a-peroxyflavin ($FMNO^-$), octanesulfonate (OCS), $FMNH_2$-octanesulfonate ($FMNH_2$-OCS), and $FMNO^-$-octanesulfonate ($FMNO^-$-OCS). Residues were analyzed (Glu20, Asp111, Arg226, and Arg297) that have been shown to be part of the active site and/or mobile loop. Salt bridge formations were found to occur for the simulations either between Glu20 and Arg297 or Asp111 and Arg297 that controlled the open and closed conformations of the enzyme. Arg226 played an important role in stabilizing the OCS and peroxy group of $FMNO^-$ and potentially plays the role of a general acid in a Baeyer-Villiger-like rearrangement. Distance calculations showed the OCS coming within 5Å of the peroxy group, consistent with the proposed mechanism and experimental crystal structural distances in a similar Baeyer-Villiger Monooxygnease (MtmOIV) enzyme.

## 5.1 Introduction

Alkanesulfonate monooxygenase (SsuD), in *Escherichia coli*, is a two-component monooxygenase system (SsuD/E) capable of capturing sulfur from alkanesulfonate in times of sulfur starvation.[64-65, 68, 70, 223-226] SsuD/E belongs to a family of proteins having a triose phosphate isomerase (TIM)-barrel fold. Bacterial luciferase and LadA (long-chain alkane monooygenase) are the most similar enzymes of the TIM-barrel family.[223] Of the two components, SsuE shows preference for oxidized flavin mononucletide (FMN) and utilizes NADPH to reduce FMN to its reductase ($FMNH_2$). SsuD then generates the corresponding sulfite and aldehyde from $FMNH_2$, molecular oxygen and alkanesulfonate (Figure 1.11).[65, 68, 224-225] The reduced flavin in the presence of molecular oxygen generates C4a-(hydro)peroxyflavin and C4a-peroxyflavin intermediates, which are the appropriate flavins for sulfite production.[68]

The catalytic mechanism for sulfite production has been proposed to first bind $FMNH_2$, and that allows the mobile loop to close over the active site. A second conformational change is then believed to occur that allows octanesulfonate to bind.[65] These conformational changes essentially "crack" the crystal structure of SsuD. No known crystal structures for SsuD with bound substrates has been published to date. However, the crystal structure for closely related bacterial luciferase with a bound FMN[227] and LadA with a bound FMN[228] have been determined and their positions in the active site defined. The active site for bacterial luciferase contains the isoalloxazine ring of $FMNH_2$ which is coordinated to an Ala-Ala peptide bond, and the C4a position of the hydroperoxy group is stabilized by Cys106. Phe6 and Glu175 move depending on the

binding conformation of FMN.[227] LadA contains an active site His138 which has been proposed to stabilize the hydroperoxy group of the flavin and is located in the same spatial arrangement as bacterial luciferase. The active site for SsuD is thought to be at the C-terminal end of the $\beta$-barrel, similar to other TIM-barrel proteins. It has been proposed that residues His11, His228, His333, Tyr331, Arg297, Arg226 and Cys54 occupy the active site of SsuD and are in similar arrangement to bacterial luciferase and LadA. [65, 68, 223, 226, 229] Various mutagenesis studies have been performed to confirm the catalytic activity of the residues in the active site.[223, 225, 229]

A molecular dynamics study was performed by Ferrario et al. on SsuD without active site residues 250-283.[230] Their 20 ns MD simulations located a salt bridge between Arg297 and Asp111 for the closed SsuD bound with $FMNH_2$ and another one between Arg297 and Glu20 for open apo SsuD. Salt bridge formation between Arg297 and Asp111 for the closed conformation is proposed to prevent the exiting of the cofactor. The study did not examine the C4a-peroxyflavin intermediate or include an alkanesulfonate along with the flavin cofactor.

In our work, molecular dynamics simulations have been carried out for six SsuD systems: 1. apo SsuD, 2. SsuD with $FMNH_2$, 3. SsuD with the C4a-peroxyflavin intermediate, 4. SsuD with $FMNH_2$ and octanesulfonate, 5. SsuD with the C4a-peroxyflavin intermediate and octanesulfonate, and 6. SsuD with octanesulfonate. Extensive analysis, including principal component analysis (PCA), has been performed for relevant residues within the active site and the mobile loop to elucidate any conformational changes that occur in the enzyme with various bound substrates.

## 5.2 Methods

**5.2.1 Enzyme Setup.** Initial Cartesian coordinates for the SsuD monoxygenase enzyme-FMN systems were generated from a 2.3 Å resolution crystal structure (PDB entry 1M41). The reported structure contains residues 1 – 249 and 283 – 361 per monomer, as no clear density was available for internal residues 250 – 282 or C-terminal residues 362 – 380.[226] Residues 250-282 were constructed using a comparative modeling program MODELLER9.10.[231] The program generates a refined 3-D model of a given protein sequence (target) based primarily on its alignment to one or more proteins of known structure (templates). The templates used were the SsuD structure (PDB entry 1NQK) and the structure of the Luciferase-like Monooxygenase from Bacillus cereus (PDB entry 3RAO).

As there is no three-dimensional structure available for SsuD with substrates bound, flavin ligands were docked into the active site region of SsuD based on a superposition with coordinates from a long-chain monooxygenase (LadA) enzyme with a bound FMN (PDB entry 3B9O).[232] Recent studies have suggested that conserved residues among the flavin-dependent monooxygenases are structurally related and possess similar catalytic activities and should therefore provide a reasonable starting point for the simulations.[68, 223, 225] Docking studies were performed using AutoDock Vina.[233] A grid box was fit to cover the proposed active site (grid dimensions listed in Table GB of the Supporting Information), where 10-20 binding modes were analyzed to determine the most probable structure based on previous proposed catalytic sites.[68, 223,

[225-226] Alkanesulfonate was subsequently docked in the active site of SsuD with the reduced and peroxy flavins.

**5.2.2 MD Simulation Protocol.** MD simulations were performed with the apo WT SsuD and with five different combinations of ligands: reduced ($FMNH_2$), peroxy ($FMNO^-$), octanesulfonate (OCS), $FMNH_2$-octanesulfonate ($FMNH_2$-OCS), and $FMNO^-$-octanesulfonate ($FMNO^-$-OCS). In each case, SsuD was simulated as a monomer. The resulting structures were fed into the leap module of the AMBER12 molecular dynamics package[234] where the appropriate hydrogen atoms were added. The monomer was solvated in an orthorhombic box of TIP3P explicit water molecules extending at least 10 Å beyond the enzyme and sodium ions were added to maintain charge neutrality. The ff99SB[235] force field was used to construct the topology files for the protein, while the parameters for the ligands were obtained from the Generalized Amber Force Field.[236] Each ligand was optimized by carrying out a Monte Carlo conformational search using the BOSS program[237] and the OPLS force field.[238] The five lowest energy structures were re-optimized at the MP2/6-31G(d) theory level using Gaussian 09[159] and the most favorable MP2 structure was used to determine the restrained electrostatic potential (RESP)[239] partial atomic charges using HF/6-31G(d) and the AmberTools 1.5 antechamber module.

For each system, the initial structure was conjugate gradient (CG) minimized for 200 steps for the water molecules only, followed by 10,000 steps of CG optimization of the entire system to remove any bad contacts. Next, the full system was gradually heated, at constant NVT, from 0 K to 300 K during 50 ps of MD using the weak-coupling

algorithm with a temperature coupling value of 2.8 ps. The system was then switched to constant NPT conditions, using a coupling value of 2.0 ps for both temperature and pressure, at 300 K for 500 ps and then for an additional 500 ps of equilibration at constant NVT. Following equilibration, 300 ns of production data was collected at constant NVT for each protein complex using the GPU-accelerated version of AMBER12.[240] In order to verify at that the GPU-based simulations yielded reasonable results, a shorter simulation of 10 ns was carried out using the CPU code to check the Ewald error estimate and to ensure a stable system. All MD simulations carried out in this work utilized the particle mesh Ewald method[126] to compute the long-range Coulomb force, periodic boundary conditions with a nonbonded cutoff distance of 12.0 Å, and a time step of 1.0 fs.

**5.2.3 Principal Component Analysis**. Principal component analysis was used to examine the conformational fluctuations of SsuD. PCA reduces the dimensionality of a data set by generating a mass-weighted covariance matrix (S) (5.1) and diagonalizing it (5.2), yielding eigenvectors (or principal components).[127, 129]

$$S_{jk} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\tilde{x}_{ij} - \bar{x}_j\right)\left(\tilde{x}_{ik} - \bar{x}_k\right) \tag{5.1}$$

$$U^T S U = \lambda \tag{5.2}$$

$\tilde{x}_{ij}$ and $\tilde{x}_{ik}$ represent atomic coordinates, $\bar{x}_j$ and $\bar{x}_k$ are the averages, U are the eigenvectors of S, and λ are the associated eigenvalues. PCA was performed using the Bio3D package.[241] Root mean square fluctuations (RMSF) of the protein were modeled

from the $C_\alpha$ chain of the initial structure. The trajectories required for Bio3D were made using VMD.[242]

**5.2.4 Analysis.** Clustering, hydrogen bond analysis, distance calculations, RMSd, and entropy calculations were carried out using the ptraj and cpptraj analysis programs within AmberTools 1.5.[234] For the clustering calculations, the average-linkage algorithm was utilized.[243] All enzyme figures were made in Chimera.[244]

Calculations were carried out on computers located at the Alabama Supercomputer Center and Auburn University. GPU-accelerated calculations utilized NVIDIA Telsa M2070 and Kepler GeForce GTX 680 GPUs and the AMBER12 pmemd.cuda, which was compiled using the CUDA 4.0 and 5.0 SDK, respectively.

## 5.3 Results and Discussion

**5.3.1 Cluster Analysis.** Cluster analysis was performed to identify the ten major configurations for each 300ns (60000 frame) MD simulation. Table 1 identifies the main cluster and the length it occurred over the trajectory. A list of all the clusters for each simulation, along with the top three structures from each simulation can be found in the Supporting Information. To examine the similarity of the cluster structure to the initial structure of the trajectory, RMSd calculations were performed (Figure 5.1). The dominant apo (substrate unbound) structure occurred from 104-152ns and differed from the first frame with an average RMSd value of 3.78Å. However, the apo structure more closely resembles the starting structure than the other 5 structures with RMSd values of 4.96Å for $FMNH_2$, 4.29Å for $FMNO^-$, 5.91Å for OCS, 4.34Å for $FMNH_2$-OCS, and

3.98Å for $FMNO^-$-OCS (Table 5.1). Larger RMSd values for the bound systems result from the large conformational changes of SsuD when the substrates bind.

**Table 5.1**: Dominant clusters for SsuD unbound and with various bound ligands. The RMSds are the average values over the range the cluster occurs.

| Enzyme | Cluster | % Occupied | No. of Frames | Time Occurs (ns) | Average RMSd (Å) |
|---|---|---|---|---|---|
| Apo | 6 | 25.4% | 15212 | 104-152 | 3.78 |
| $FMNH_2$ | 9 | 36.2% | 21717 | 195-300 | 4.96 |
| $FMNO^-$ | 3 | 34.2% | 20499 | 23-127 | 4.29 |
| OCS | 8 | 43.4% | 26059 | 97-231 | 5.91 |
| $FMNH_2$-OCS | 9 | 65.6% | 39340 | 102-300 | 4.34 |
| $FMNO^-$-OCS | 3 | 37.5% | 22476 | 44-163 | 3.98 |

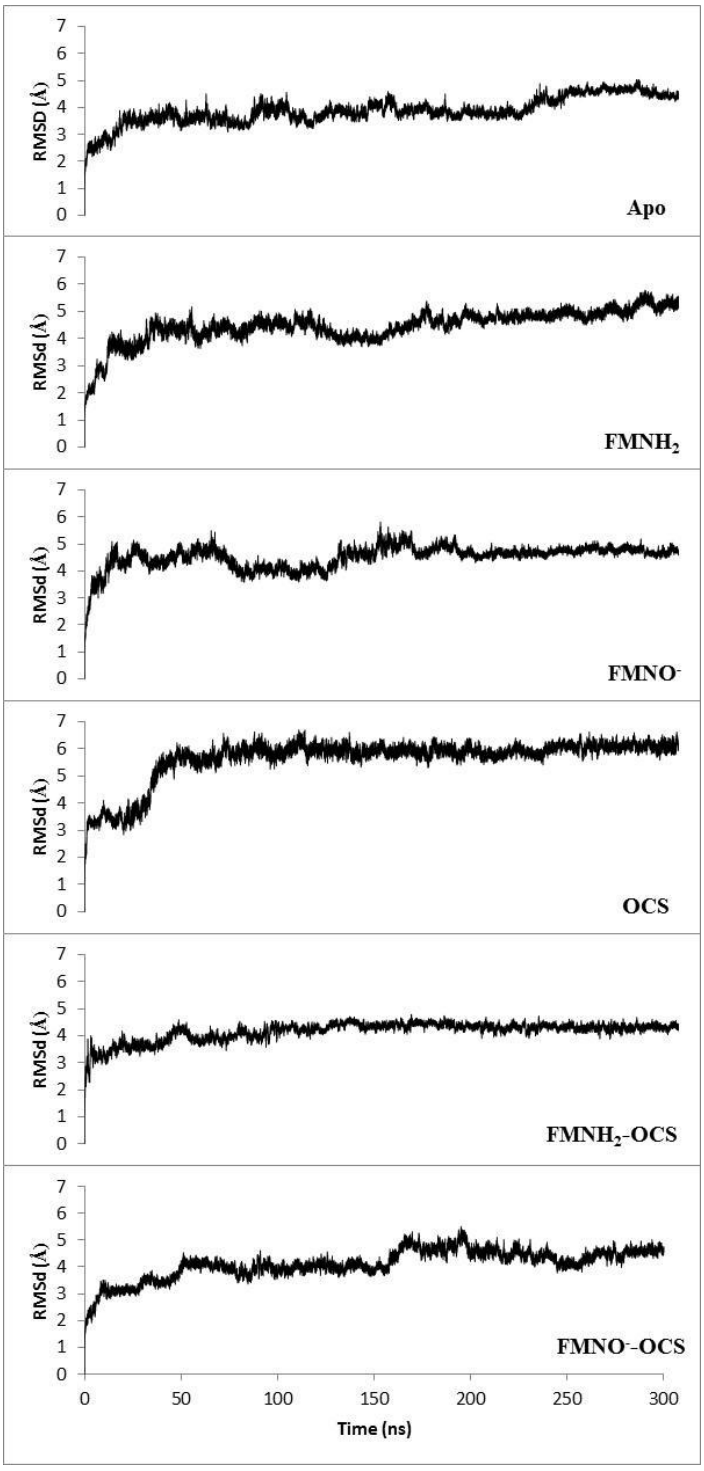



**Figure 5.1:** RMSd for WT SsuD unbound and bound with various substrates, with respect to the initial structure.

**5.3.2 Role of Arg297.** Arg297 is thought to play an important role in the active site of SsuD. It is a conserved residue amongst various organisms that all utilize SsuD for sulfite formation. Arg297 is an amino acid located in the mobile loop postulated to close over the active site after the ligands bind. Studies of bacterial luciferase proposed a major conformational change occurring with the addition of phosphate that is necessary for the isomerization reaction.[225, 245] SsuD may interact similarly to bacterial luciferase, where Arg297 interacts with the phosphate ($PO_4$) group of the flavin.[225] Distance calculations have been performed for all WT SsuD enzymes containing a flavin group (Figure 5.2). Arg297 interacts with the $PO_4$ group of $FMNH_2$ for approximately 36% of the simulation length (Figures 5.2 and 5.3, Table 5.1). However, the $FMNO^-$ and $FMNH_2$-OCS show little to no interaction with the residue (Figure 5.3). Arg297 is instead occupied by electrostatic interactions with Glu20 for the $FMNO^-$ and $FMNO^-$-OCS systems (Figure 5.4), while in the $FMNH_2$-OCS bound SsuD system, Arg297 has significant interactions occurring with Asp111 as confirmed by distance calculations (Figures 5.7 and 5.8, Figure 2-1 in Appendix II).

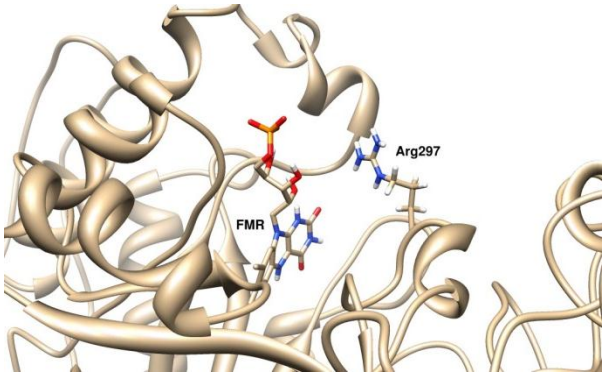



**Figure 5.2**: Cluster structure 9 from the $FMNH_2$ bound WT SsuD system highlighting the interacting distance between the phosphorous of $PO_4$ and $NH_2$ of Arg297.

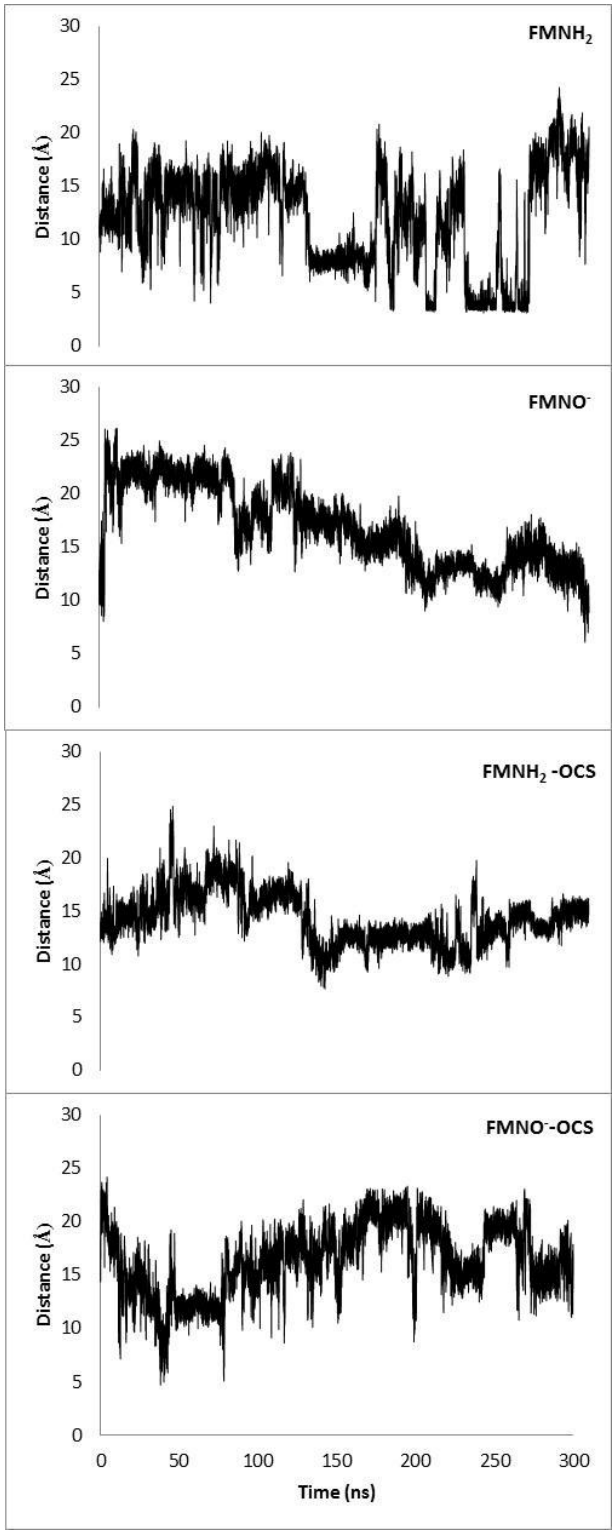



**Figure 5.3:** Distance calculations between the phosphorus of the flavin phosphate group with $NH_2$ of Arg297 for all WT SsuD structures containing a flavin group.

**5.3.3 Salt bridge formations between Glu20-Arg297 and Asp111-Arg297.** Recent molecular dynamics calculations proposed salt bridge formations occurring in SsuD. The "closed" conformation of SsuD included $FMNH_2$ and occurred with a salt bridge between Asp111 and Arg297. The salt bridge between Glu20 and Arg297 for the "open" conformer occurred for the apo structure. In order to elucidate these potential salt bridges, distance calculations were performed. Figure 5.4 shows the salt bridge formation of Glu20-Arg297 at an average distance between 2.65-3.0Å for a negligible 3 frames (15ps) for the apo structure as indicated by no hydrogen bonding interactions were apparent more than 0.1% of the simulation in Figure 5.5. The average distance then increases to 9-10Å for the next 240ns. The previous results discussed by Ferrario et al. showed a salt bridge conformation occurring after the first 200ps and continuing for the remainder of the 20ns simulation.[230] Their simulations had missing residues (250-283), which could possibly alter the salt bridge formations for residues near the unresolved mobile loop. The WT SsuD with bound $FMNH_2$ shows a salt bridge of ~3Å between Asp111 and Arg297 instead and it occurs in the simulation between 10-20 ns, earlier than in the apo structure (Figure 5.6). Hydrogen bonding analysis (Figure 5.7) indicates that the largest occupancy occurs with the $FMNH_2$ substrate over the other substrate combinations, but only occupies approximately 5-10% of the simulation. $FMNO^-$ shows less than 0.1% of the occupancy, which is consistent with a dominant salt bridge for the Glu20-Arg297 $FMNO^-$ structure (Figure 5.8).

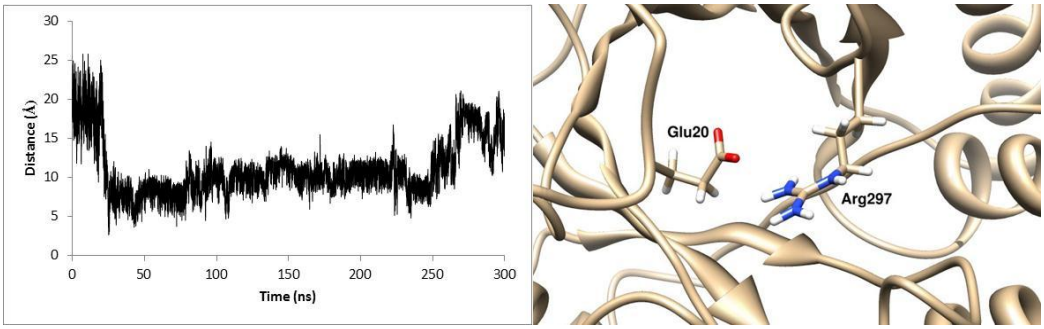



**Figure 5.4**: Salt bridge between Glu20 and Arg297 for WT SsuD with no bound substrate occurring at 30ns.

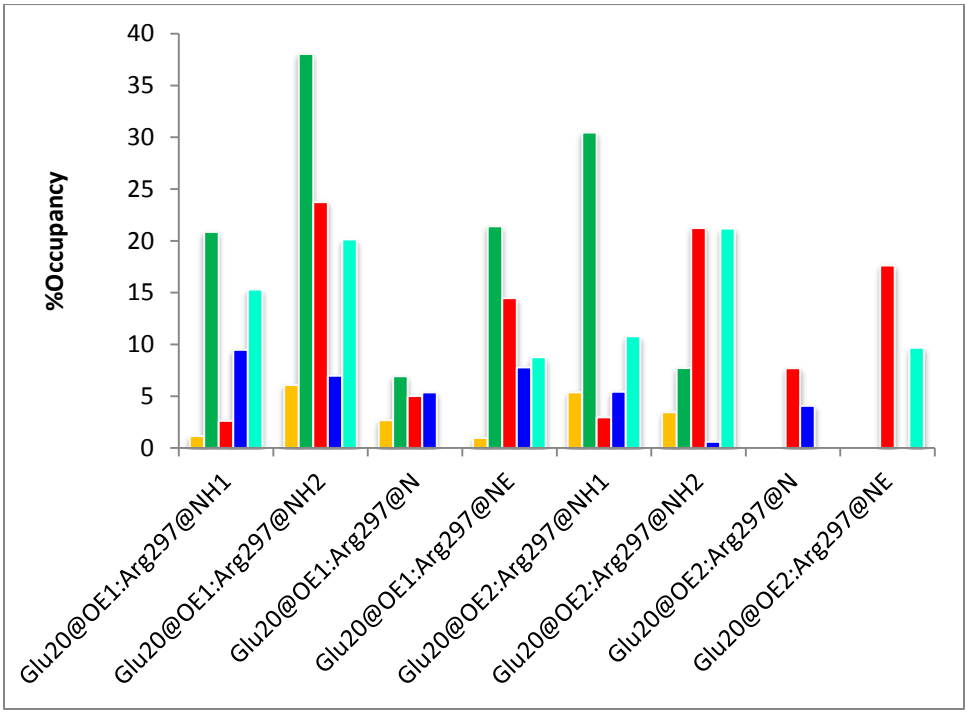



**Figure 5.5**: Hydrogen bonds formed between Glu20 and Arg297 for all simulations (No occupancy above 0.1% for the apo structure). ■ represents $FMNH_2$, ■ represents $FMNO^-$, ■ represents OCS, ■ represents $FMNH_2$-OCS and ■ represents $FMNO^-$-OCS.

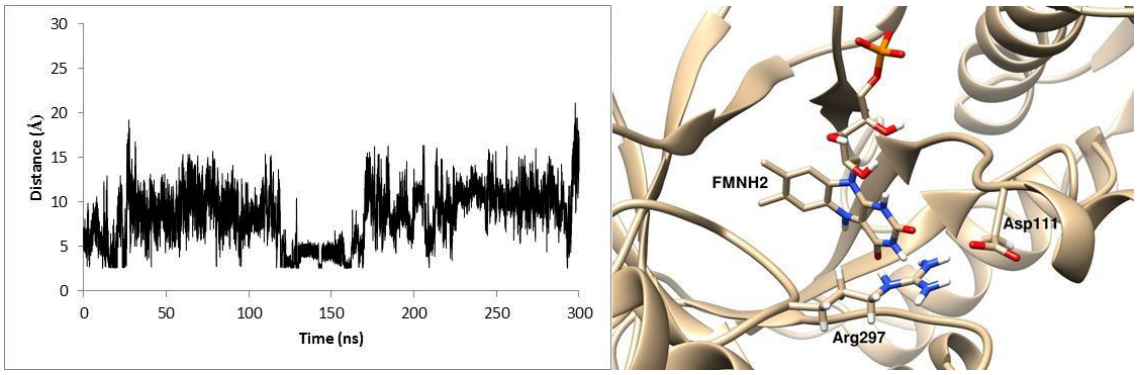



**Figure 5.6:** Salt bridge between Asp111 and Arg297 in WT SsuD with docked $FMNH_2$. Illustrated structure given at 11.8ns.

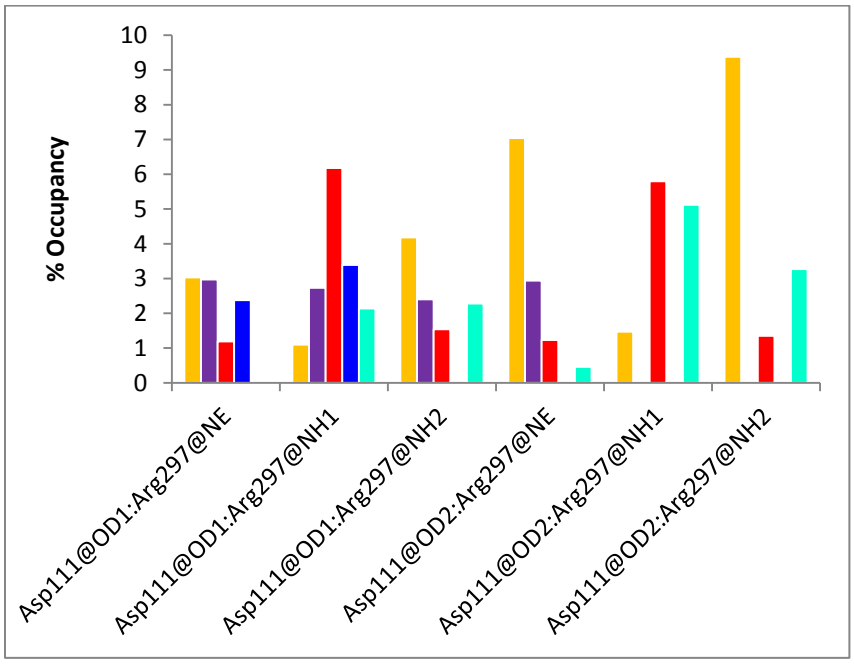



**Figure 5.7**: Hydrogen bonds formed between Asp111 and Arg297 for all simulations (No occupancy above 0.1% for the $FMNO^-$ structure). represents $FMNH_2$, represents apo, represents OCS, represents $FMNH_2$-OCS and represents $FMNO^-$-OCS.

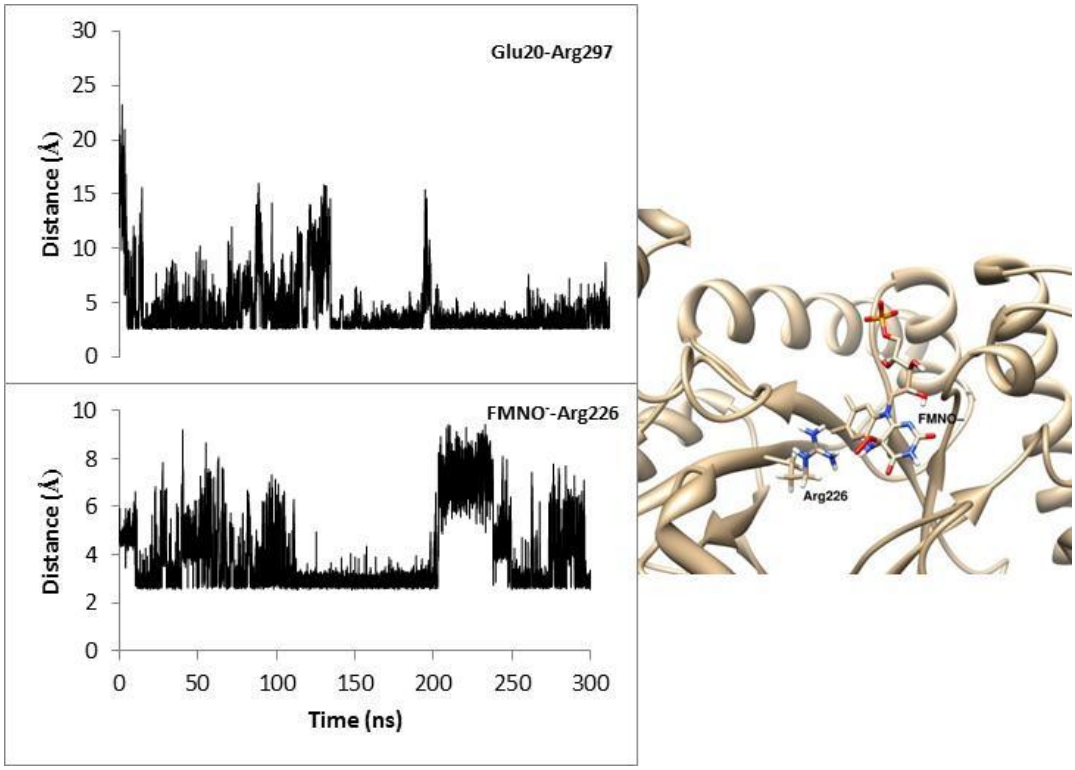



**Figure 5.8:** Distance calculations showing a salt bridge between Glu20 and Arg297 occurring over a large portion of the $FMNO^-$ simulation and electrostatic interactions between the peroxy group of $FMNO^-$ and $NH_2$ of Arg226, represented by the illustrated cluster structure for the enzyme.

To further examine the $FMNO^-$ and $FMNO^-$-OCS SsuD systems in the large occupancy of hydrogen bonds for Glu20-Arg297 (Figure 5.5), additional distance calculations were carried out. The main cluster structure for the $FMNO^-$ and $FMNO^-$-OCS systems determined that Arg226 interacts substantially with the peroxy group of the flavin (Figures 5.8 and 5.9). The distance calculations found a close interaction distance of 2-3Å occurring over much of the simulation. Recent studies have shown that Arg226 plays two roles in SsuD, stabilization of the peroxy group and protonation of the peroxy group causing a conformational change releasing water and oxidized flavin (FMN).[229]

The results from the distance and clustering calculations are consistent with these experimental findings.

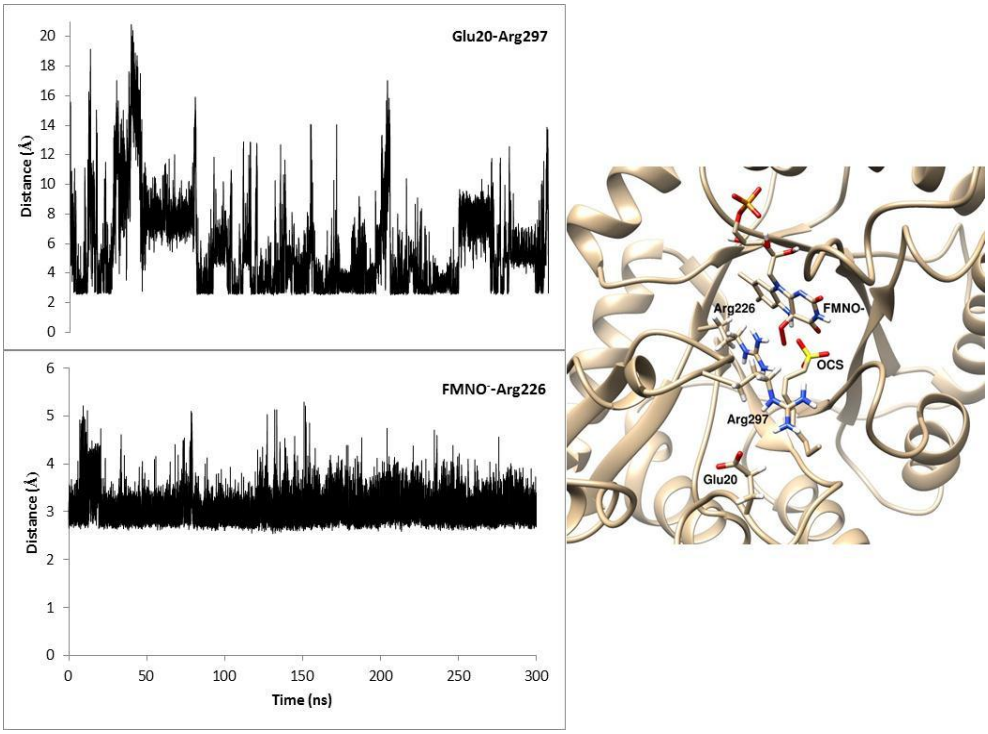



**Figure 5.9:** Distance calculations showing a salt bridge between Glu20 and Arg297 occurring over a large portion of the $FMNO^-$-OCS simulation and electrostatic interactions between the peroxy group of $FMNO^-$-OCS and $NH_2$ of Arg226, as shown in both the distance calculation and the illustrated cluster structure for the enzyme.

**5.3.4 Octanesulfonate and its role in SsuD.** The clustering analysis performed on the SsuD-octanesulfonate system is given in Table 5.1 and Figure 5.11. Distance calculations were performed between the sulfur of octanesulfonate and $NH_2$ group of Arg226 for each of the simulations containing OCS. The distance calculations found a close interaction distance of 4-5Å over most of the simulation for OCS bound SsuD system, but an even closer interaction of 2-3Å with $FMNH_2$-OCS. The flavin group in the $FMNH_2$-OCS system is held in place by hydrogen bonding interactions between Tyr128

and Ser110 and the oxygens of the flavin for much of the simulation (Figure 5.10). The $FMNH_2$-OCS cluster structure in Figure 5.11 shows the position of the OCS, which occurs 65.6% of the simulation, with minimal movement over the remaining simulation (Figure 2-8 in Appendix II).

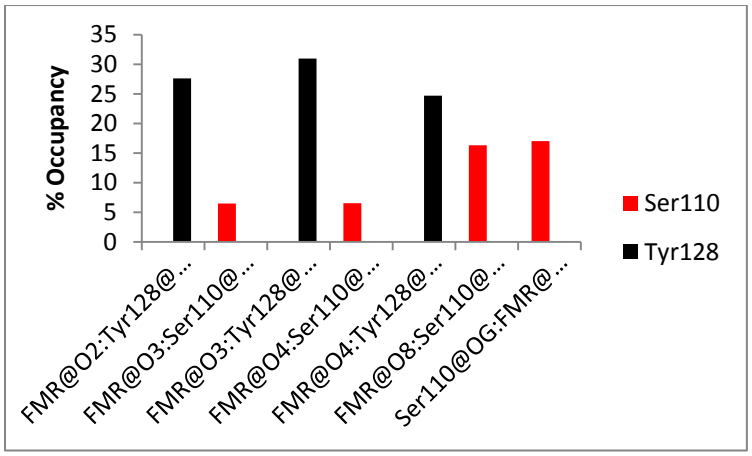



**Figure 5.10:** All hydrogen bond interactions occupying above 3% of the simulation time for the $FMNH_2$ in the $FMNH_2$ bound and $FMNH_2$-OCS bound SsuD systems.

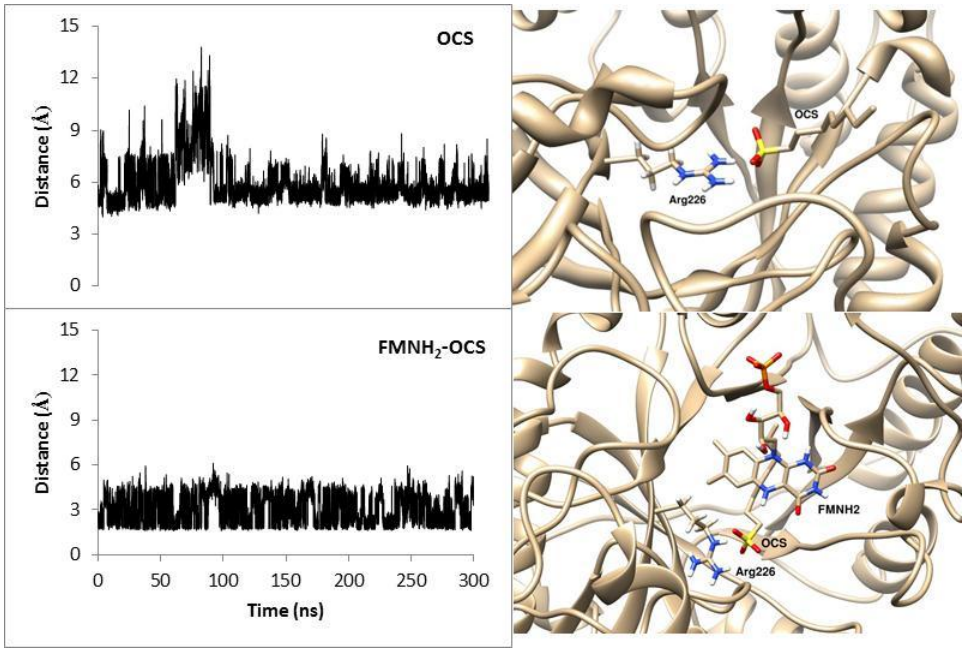



**Figure 5.11:** Distance calculations for WT SsuD containing octanesulfonate as a substrate for the sulfure of OCS and $NH_2$ of Arg226 with the dominant cluster structures for the OCS and $FMNH_2$-OCS simulation illustrated.

The $FMNO^-$-OCS simulation did not show an interaction between Arg226 and octanesulfonate (Figure 2-3 in Appendix II). However, it did find the peroxy group of the flavin to be in close proximity to Arg226 and octanesulfonate. The proposed mechanism to release sulfite in SsuD has the peroxy group coming into contact with octanesulfonate and an active site base (Figure 1.17B). Although the active site base is not known, the reaction is proposed to proceed through a Baeyer-Villiger-like rearrangement.[68] A distance calculation was performed for the peroxy group of the flavin and the sulfur of the octanesulfonate to examine their distance over the course of the simulation (Figure 5.12). As a point of comparison, the Baeyer-Villiger Monooxygenase enzyme (MtmOIV) has a reported distance of ~5.3Å between the carbon of the substrate and the C4a position of the flavin ring from crystal structure data.[246] The distance for the

second most dominant cluster in $FMNO^-$-OCS system shows the C-1 position of OCS to be 5.28Å from the C4a position of the flavin ring. This similar distance in the SsuD simulations suggests the reaction could occur as proposed in Figure 1.17B.

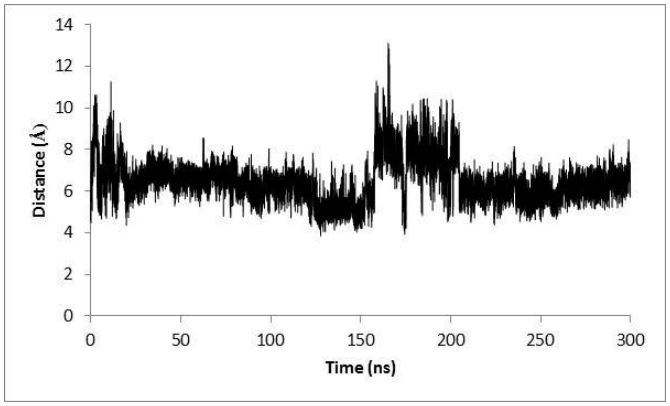



**Figure 5.12**: Distance calculation of the peroxy group of the flavin group with octanesulfonate for the $FMNO^-$-OCS SsuD simulation.

**5.3.5 Conformational fluctuations occurring in SsuD.** The major conformational changes that occur throughout the simulation can be further examined by principal component analysis. The first three principal components (PC) were taken for each WT SsuD structure and compared with the apo structure. These PC's are shown to give the most variation from the initial structure, however only account for 24-26% of the total variance (Figure 5.13). The height of each bar in Figures 5.14-5.18 represents the relative displacement of each residue described by a given PC and is depicted as atomic displacements from the starting structure. The movement from the crystal structure should be different when bound substrates are present. The root mean square fluctuations (RMSF) are depicted for the first three principal components of $FMNH_2$ (Figure 5.14),

FMNO$^-$ (Figure 5.15), OCS (Figure 5.16), $FMNH_2$-OCS (Figure 5.17) and FMNO$^-$-OCS (Figure 5.18) relative to the apo structure.

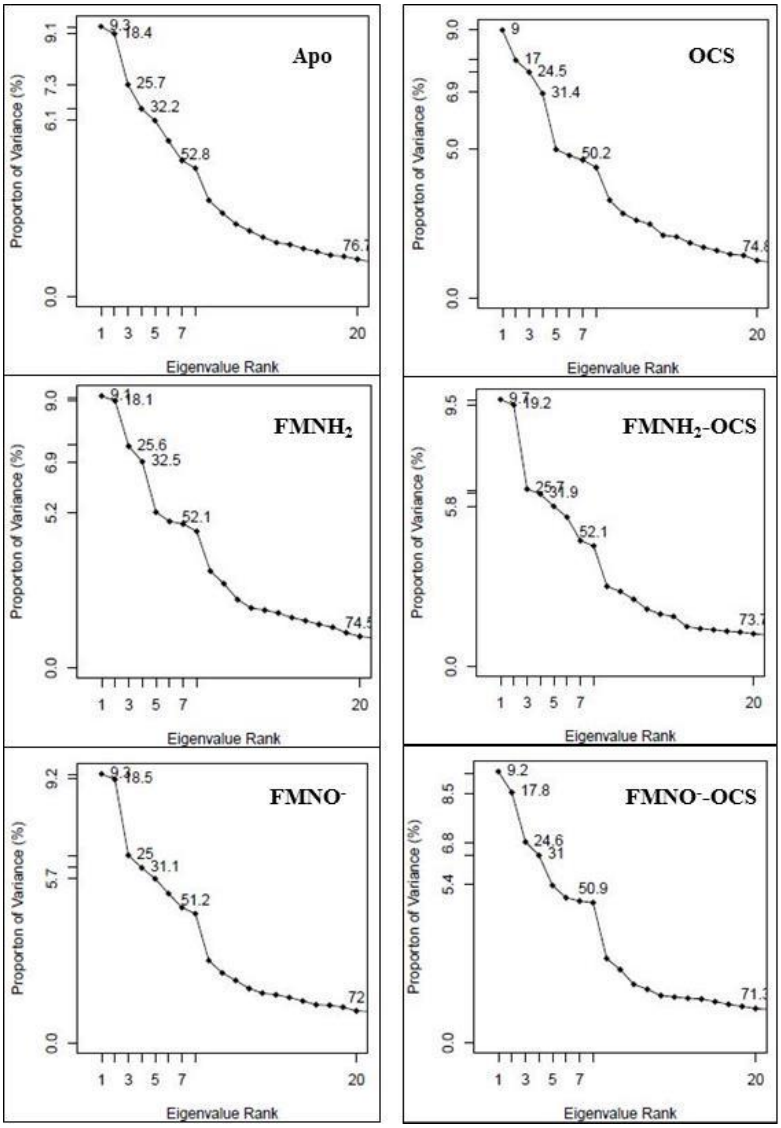



**Figure 5.13:** Eigenvalue scree plot indicating the total percent variance for each principal component. Numbers next to points represent the sum of the total variance at the given PC.

The fluctuations of interest are Glu20-Arg297, Asp111-Arg297, Arg226-OCS, Arg226-flavin, and Arg297-flavin which have been proposed to be in close interaction with each other. The mobile loop, which contains Arg297, identified by the dynamic salt bridge formations observed in the distance calculations has increased fluctuations for each of the bound simulations compared to the apo structure. The negative Δd values in Table 5.3 represent the difference in the fluctuations between the apo and the bound systems for the salt bridges between Glu20/Asp111 and Arg297. For all systems except the salt bridge for Asp111-Arg297 in the OCS structure, more fluctuations occurred for the bound structures than the apo for the salt bridges. The change was more drastic for the $FMNO^-$-OCS SsuD system than any of the other simulations, possibly accounting for the two conformational changes proposed.

Arg226 plays an important role in stabilization of the octanesulfonate and the flavin cofactors, noticed in the distance calculations. Fluctuations were monitored for Arg226 with both the octanesulfonate and flavin. The fluctuational changes were measured for the $FMNO^-$, $FMNH_2$-OCS, and $FMNO^-$-OCS SsuD bound systems, relative to $FMNH_2$ (Table 5.3). Arg226 has been shown to be within close interaction distance of 2-5Å with the flavin cofactor for the $FMNH_2$, $FMNO^-$, $FMNH_2$-OCS, and $FMNO^-$-OCS simulations as shown by distance calculations (Figures 5.9, 5.11, 5.12). As expected, minor fluctuational differences were noticed for the $FMNO^-$, $FMNO^-$-OCS and $FMNH_2$-OCS systems with Arg226 relative to $FMNH_2$. Interactions of Arg226 and the octanesulfonate group show more fluctuations occurring for the $FMNO^-$-OCS bound SsuD system than the $FMNH_2$-OCS system. This is a result of the peroxy group of the

FMNO$^-$-OCS system interacting with the octanesulfonate substrate (Figure 5.12) and Arg226 (Figure 5.9) for the majority of the simulation. $FMNH_2$-OCS does not show as many fluctuations as the FMNO$^-$-OCS, demonstrating the necessity of the peroxy group of the flavin for the proposed conformational changes in SsuD.

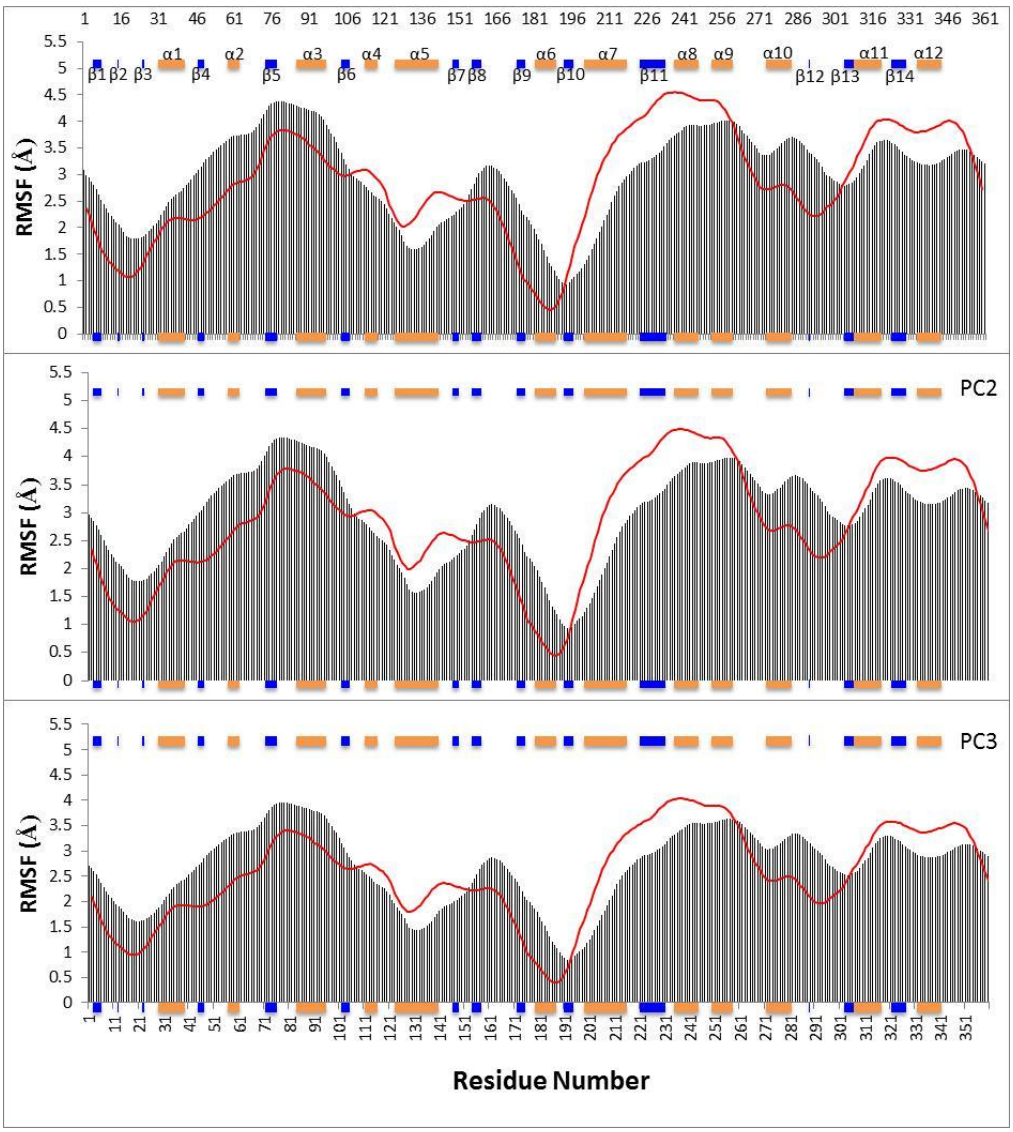



**Figure 5.14**: Root mean square fluctuations of $FMNH_2$ (solid black) with the apo structure (solid red). The secondary structure is located with blue and orange bars (alpha helices are in orange and beta sheets are in blue). PC represents the first three dominant principal components.

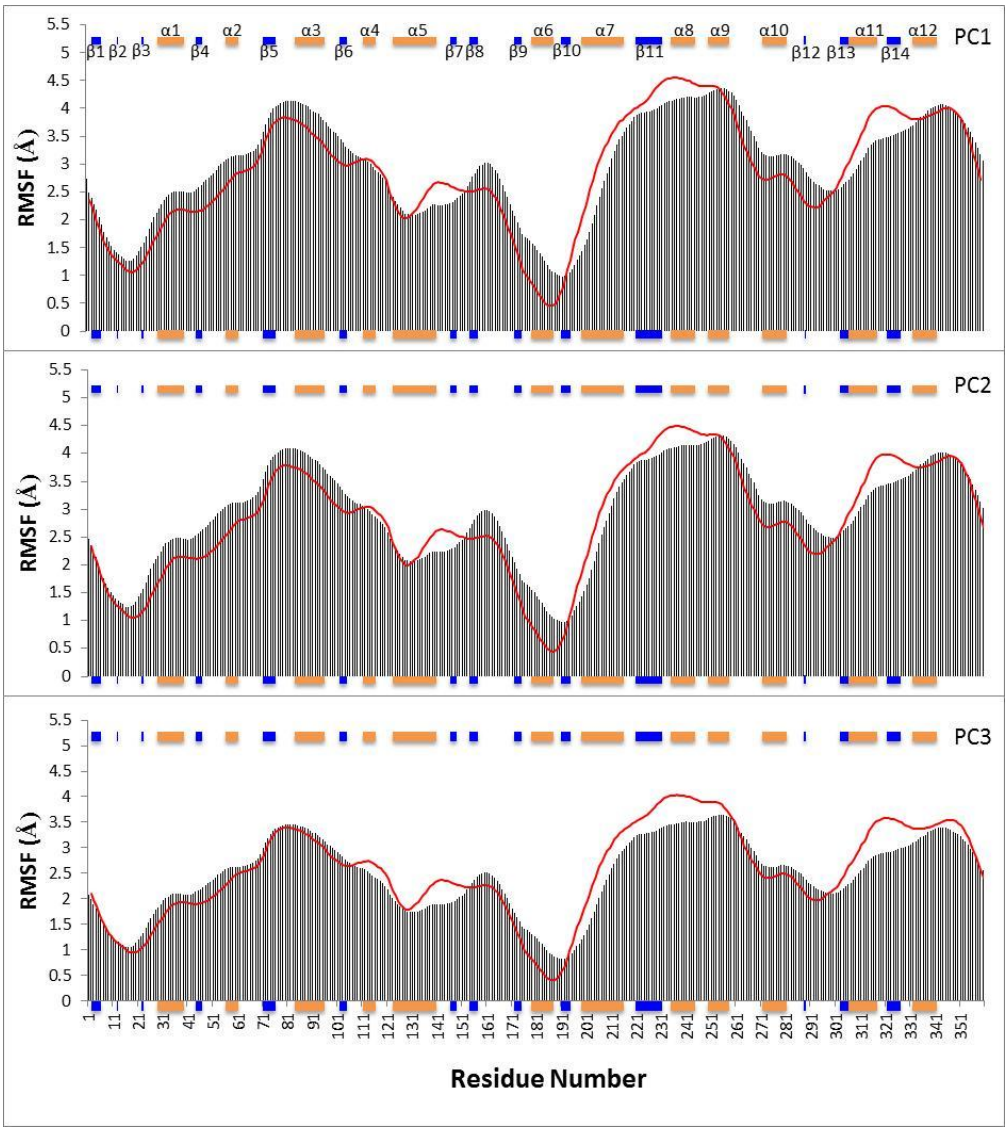



**Figure 5.15:** Root mean square fluctuations of $FMNO^{\cdot}$ (solid black) compared to the apo structure (solid red). The secondary structure is located with blue and orange bars (alpha helices are in orange and beta sheets are in blue). PC represents the first three dominant principal components.

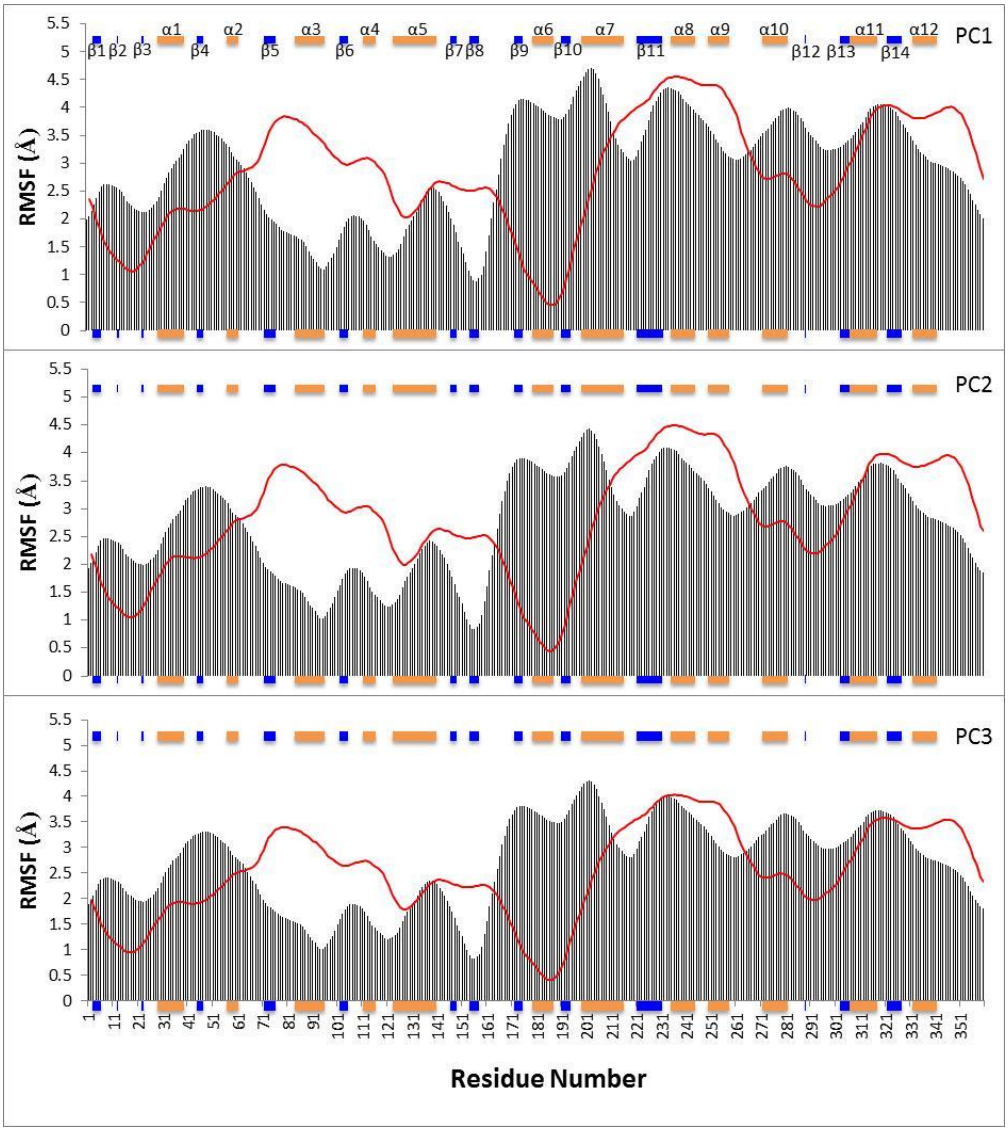



**Figure 5.16:** Root mean square fluctuations of OCS (solid black) compared to the apo structure (solid red). The secondary structure is located with blue and orange bars (alpha helices are in orange and beta sheets are in blue). PC represents the first three dominant principal components.

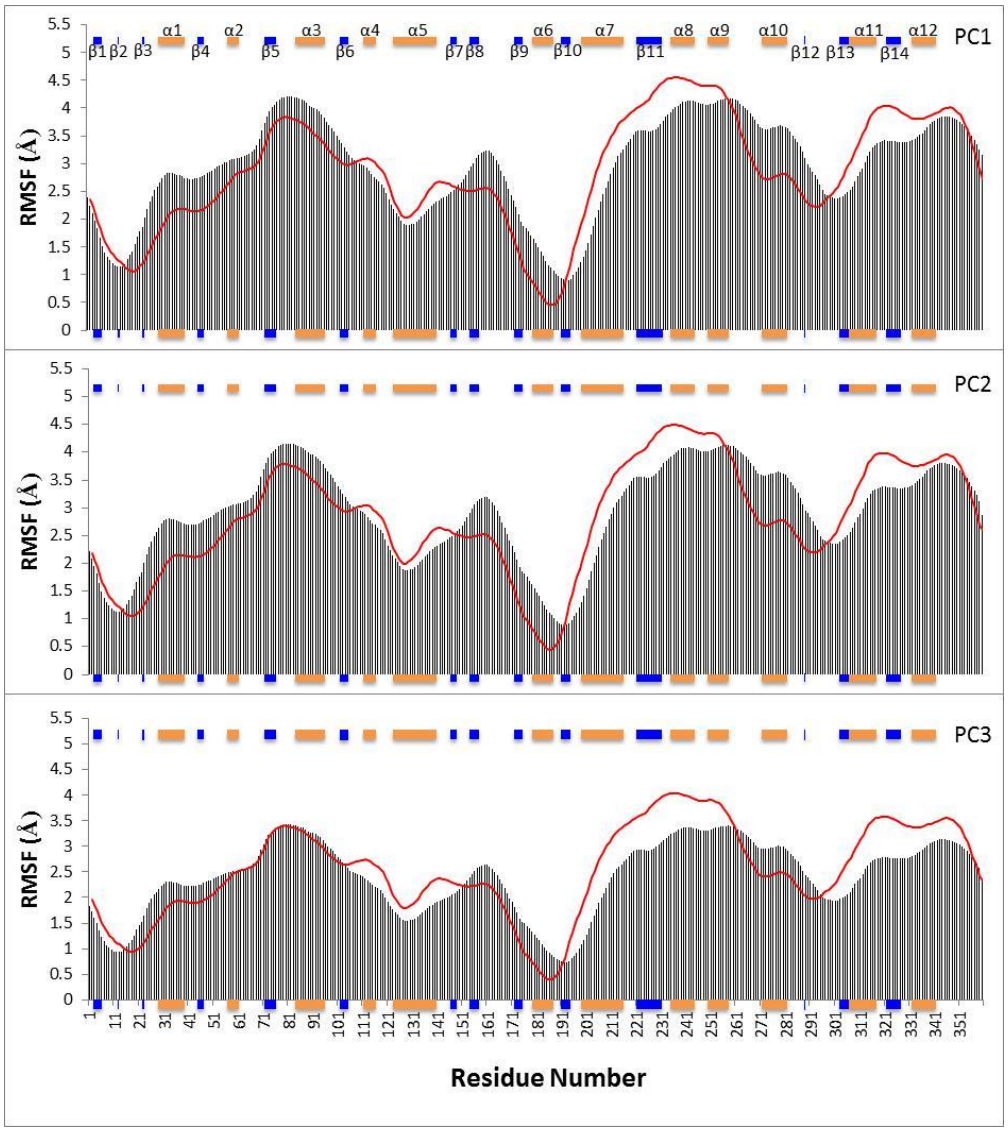



**Figure 5.17:** Root mean square fluctuations of $FMNH_2$-OCS (solid black) compared to the apo structure (solid red). The secondary structure is located with blue and orange bars (alpha helices are in orange and beta sheets are in blue). PC represents the first three dominant principal components.

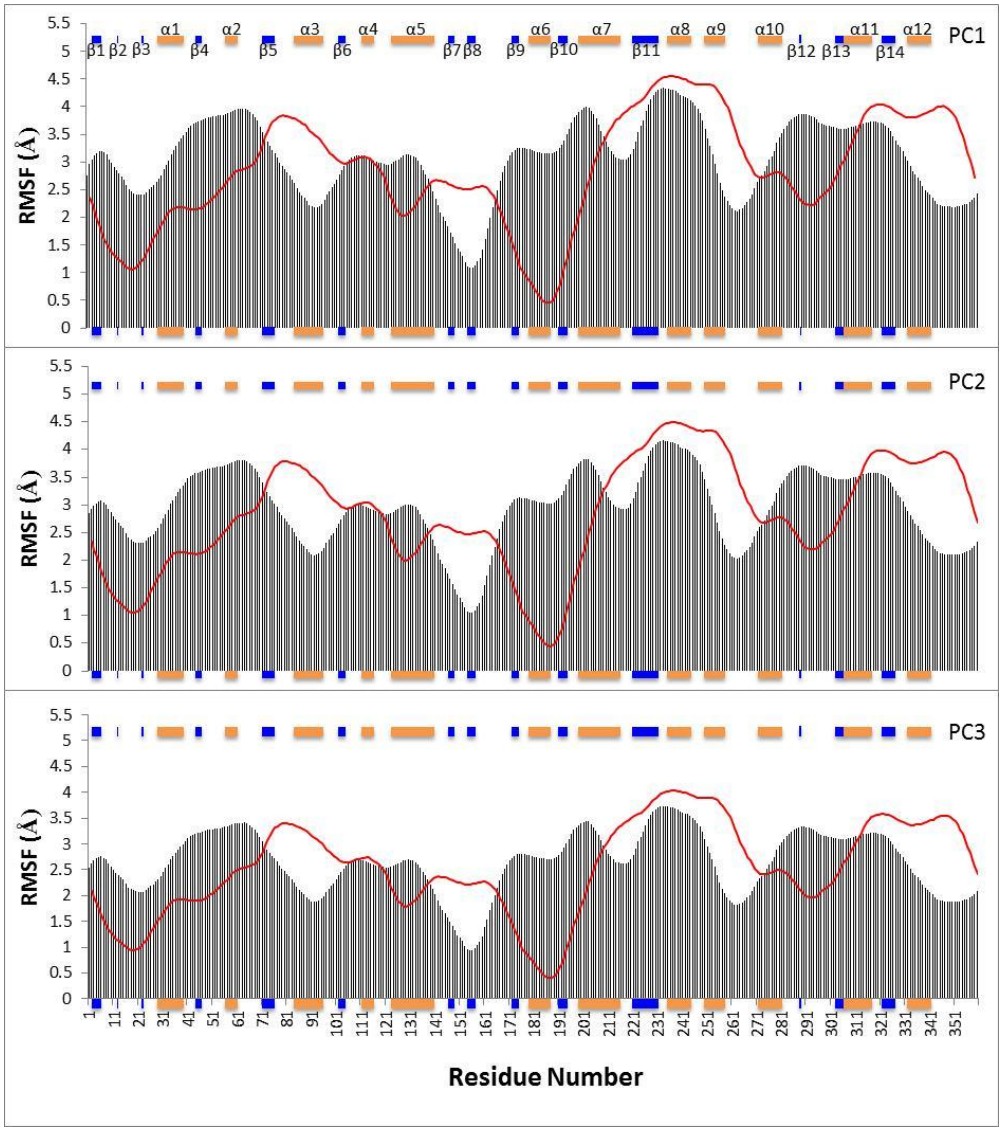



**Figure 5.18:** Root mean square fluctuations of $FMNO^{\cdot}$-OCS (solid black) compared to the apo structure (solid red). The secondary structure is located with blue and orange bars (alpha helices are in orange and beta sheets are in blue). PC represents the first three dominant principal components.

**Table 5.2**: RMSF for interactions listed (average RMSF of residues).

| Interaction | RMSF Values (Å) | | | | | |
|---|---|---|---|---|---|---|
| | Apo | $FMNH_2$ | $FMNO^-$ | OCS | $FMNH_2$-OCS | $FMNO^-$-OCS |
| Glu20-Arg297 | 1.7 | 2.4 | 1.9 | 2.7 | 2.0 | 3.1 |
| Asp111-Arg297 | 2.7 | 2.9 | 2.8 | 2.6 | 2.8 | 3.4 |
| Arg297-Flavin | -- | 3.1 | 2.6 | -- | 2.7 | 3.2 |
| Arg226-Flavin | -- | 3.3 | 3.3 | -- | 3.2 | 3.3 |
| Arg226-OCS | -- | -- | -- | 2.9 | 3.1 | 2.6 |

**Table 5.3**: Average difference in fluctuations of interactions with comparison of each bound structure (b) to the unbound (apo) structure (a). The Arg297-Flavin and Arg226-Flavin (b) differences are compared to $FMNH_2$ (a) and Arg226-OCS (b) is compared to OCS (a).

| Interaction | $\Delta d\ (\overline{RMSF_a} - \overline{RMSF_b})$ | | | | |
|---|---|---|---|---|---|
| | $FMNH_2$ | $FMNO^-$ | OCS | $FMNH_2$-OCS | $FMNO^-$-OCS |
| Glu20-Arg297 | -0.7 | -0.2 | -1.0 | -0.3 | -1.4 |
| Asp111-Arg297 | -0.2 | -0.1 | 0.1 | -0.1 | -0.7 |
| Arg297-Flavin | -- | 0.5 | -- | 0.4 | -0.1 |
| Arg226-Flavin | -- | 0.0 | -- | 0.1 | 0.0 |
| Arg226-OCS | -- | -- | -- | -0.2 | 0.3 |

## 5.4 Conclusion

Molecular dynamics simulations were carried out for six bound and unbound SsuD structures containing different combinations of flavin cofactors and octanesulfonate. Salt bridge formations played a large role in the simulations. The apo and $FMNO^-$ bound SsuD systems had salt bridge formations for Glu20-Arg297 occurring at interaction distances of 2.9-3Å. The salt bridge for $FMNO^-$ occurred for a large portion of the simulation. The $FMNH_2$, OCS, and $FMNH_2$-OCS bound SsuD systems had a different Asp111-Arg297 salt bridge formation ranging in distances of 2-3Å. Interestingly, the $FMNO^-$-OCS bound SsuD structure fluctuated back and forth between the two salt bridge formations, with Asp111-Arg297 dominating in the first 150ns and Glu20-Arg297 fluctuating over the course of the entire simulation. These findings were further analyzed by PCA and RMSF calculations to view the conformational changes occurring during the simulation. Major conformational changes occurred for each bound SsuD system for the proposed salt bridges, with a larger fluctuation for all systems compared to apo except the Asp111-Arg297 octanesulfonate bound simulation. These results are consistent with the proposed mobile loop closing over the active site induced by the substrates. The fluctuations were more apparent for the doubly bound systems, $FMNH_2$-OCS and $FMNO^-$-OCS bound SsuD systems, which are proposed to have two conformational changes.[65]

The Arg226 residue played an important role in stabilizing the peroxy group of $FMNO^-$ and octanesulfonate. The $FMNO^-$ and $FMNO^-$-OCS bound SsuD systems had interaction distances of 2-4Å between the peroxy group of the flavin and the $NH_2$ group

of Arg226. In the SsuD bound systems containing octanesulfonate, Arg226 was within interaction distances of 2-5Å to the sulfur of octanesulfonate, with the exception of $FMNO^{-}$-OCS, where the distance between Arg226 and octanesulfonate was 10-20Å. Arg226 is concluded to be an important active site residue in stabilization of the peroxyflavin intermediate and could potentially act as an active site acid where a proton is donated to the peroxyflavin intermediate for the proposed Baeyer-Villiger-like mechanism.[68] To further confirm the Baeyer-Villiger reaction triggering sulfite release in SsuD, distance calculations were run to examine the interaction distance of octanesulfonate and the peroxy group of $FMNO^{-}$. The distance between the peroxy group and octanesulfonate fluctuated over the simulation, but the two came in contact as close as 3.8Å. A crystal structure of a similar Baeyer-Villiger Monooxygenase enzyme (MtmOIV) reported a distance of ~5.3Å from the C4a position of the flavin ring and the C1 position of the substrate, similar to our computed distance of 5.28Å.

Fluctuations of Arg226 and the proposed interactions of the flavin cofactor and octanesulfonate via PCA were also examined. The Arg226-flavin fluctuations were similar for the systems including a flavin cofactor, consistent with the strong interactions seen in the distance calculations. However, when octanesulfonate was present, the fluctuations were more dramatic for the $FMNO^{--}$OCS bound SsuD system compared to the $FMNH_2$-OCS SsuD bound system. These results demonstrate the need for the c4a-peroxyflavin instead of the reduced flavin for the proposed conformational changes occurring in SsuD. Future QM/MM simulations will investigate the reaction pathway present in the SsuD with bound c4a-peroxyflavin and octanesulfonate.

**References**


1. Chanda, A.; Fokin, V. V., Organic Synthesis "On Water". *Chem. Rev.* **2009,** *109*, 725-748.

2. Gruttadauria, M.; Giacalone, F.; Noto, R., Water in Stereoselective Organocatalytic Reactions. *Adv. Synth. Catal.* **2009,** *351*, 33-57.

3. Li, C.-J.; Chen, L., Organic Chemistry in Water. *Chem. Soc. Rev.* **2006,** *35*, 68-82.

4. Pirrung, M. C., Acceleration of Organic Reactions through Aqueous Solvent Effects. *Chem. Eur. J.* **2006,** *12*, 1312-1317.

5. Acevedo, O.; Armacost, K., Claisen Rearrangements: Insight into Solvent Effects and "on Water" Reactivity from QM/MM Simulations. *J. Am. Chem. Soc.* **2010,** *132*, 1966-1975.

6. Buldu, M., Young children's perceptions of scientists: A preliminary study. *Educational Research* **2006,** *48* (1), 121-132.

7. Lindström, U. M., Stereoselective Organic Reactions in Water. *Chem. Rev.* **2002,** *102*, 2751-2772.

8. Heylin, M., Salaries And Jobs. *C & E News* **2006,** *84* (45), 43-46.

9. Weingart, P.; Muhl, C.; Pansegrau, P., Of power maniacs and unethical geniuses: science and scientists in fiction film. *Public Understanding of Science* **2003,** *12* (3), 279-287.

10. Rideout, D. C.; Breslow, R., Hydrophobic Acceleration of Diels-Alder Reactions. *J. Am. Chem. Soc.* **1980,** *102* (26), 7816-7817.

11. Breslow, R., Hydrophobic Effects on Simple Organic Reactions in Water. *Acc. Chem. Res.* **1991,** *24* (6), 159-164.

12. *Organic Reactions in Water*. Blackwell Publishing: Oxford, 2007.

13. Narayan, S.; Muldoon, J.; Finn, M. G.; Fokin, V. V.; Kolb, H. C.; Sharpless, K. B., "On Water": Unique Reactivity of Organic Compounds in Aqueous Suspension. *Angew. Chem. Int. Ed.* **2005,** *44*, 3275-3279.

14. Narayan, S.; Fokin, V. V.; Sharpless, K. B., Chemistry 'on water' - organic synthesis in aqueous suspension. In *Organic Reactions in Water*, Blackwell Publishing Ltd.: Oxford, UK, 2007; pp 350-365.

15. Hayashi, Y., In Water or in the Presence of Water? . *Angew. Chem. Int. Ed.* **2006,** *45* (48), 8103-8104.

16. Zhao, J. H., L.; Jiang, J.; Tang, Z.; Cun, L.; Gong, L., Organo-Catalyzed Highly Diastereo- and Enantio-selective Direct Aldol Reactions in Water. *Tetrahedron Lett.* **2008,** *49*, 3372-3375.

17. Jung, Y.; Marcus, R. A., On the Theory of Organic Catalysis “on Water”. *J. Am. Chem. Soc.* **2007,** *129*, 5492-5502.

18. Tang, Z.; Jiang, F.; Cui, X.; Gong, L.; Mi, A.; Jiang, Y.; Wu, Y., Enantioselective Direct Aldol Reactions Catalyzed by L-Prolinamide Derivatives. *Proc. Nat. Acad. Sci. USA* **2004,** *101*, 5755-5760.

19. Thomas, L. L.; Tirado-Rives, J.; Jorgensen, W. L., Quantum Mechanical/Molecular Mechanical Modeling Finds Diels-Alder Reactions Are

Accelerated Less on the Surface of Water Than in Water. *J. Am. Chem. Soc.* **2010,** *132*, 3097-3104.

20. Claisen, L., *Berichte* **1912,** *45*, 3157-3166.

21. Castro, A. M. M., Claisen Rearrangement over the Past Nine Decades. *Chem. Rev.* **2004,** *104*, 2939-3002.

22. Carlson, H. A.; Jorgensen, W. L., Monte Carlo Investigations of Solvent Effects on the Chorismate to Prephenate Rearrangement. *J. Am. Chem. Soc.* **1996,** *118*, 8475-8484.

23. Severance, D. L.; Jorgensen, W. L., Effects of hydration on the Claisen rearrangement of allyl vinyl ether from computer simulations. *J. Am. Chem. Soc.* **1992,** *114* (27), 10966-10968.

24. Repasky, M. P.; Guimarães, C. R. W.; Chandrasekhar, J.; Tirado-Rives, J.; Jorgensen, W. L., Investigation of Solvent Effects for the Claisen Rearrangement of Chorismate to Prephenate: Mechanistic Interpretation via Near Attack Conformations. *J. Am. Chem. Soc.* **2002,** *125* (22), 6663-6672.

25. Pincock, A. L.; Pincock, J. A.; Stefanova, R., Substituent Effects on the Rate Constants for the Photo-Claisen Rearrangement of Allyl Aryl Ethers. *J. Am. Chem. Soc.* **2002,** *124* (33), 9768-9778.

26. Majumdar, K. C.; Alam, S.; Chattopadhyay, B., Catalysis of the Claisen rearrangement. *Tetrahedron* **2008,** *64*, 597-643.

27. Gao, J., Combined QM/MM Simulation Study of the Claisen Rearrangement of Allyl Vinyl Ether in Aqueous Solution. *J. Am. Chem. Soc.* **1994,** *116* (4), 1563-1564.

28. Gajewski, J. J.; Brichford, N. L., Secondary Deuterium Kinetic Isotope Effects in the Aqueous Claisen Rearrangement: evidence against an Ionic Transition State. *J. Am. Chem. Soc.* **1994,** *116* (7), 3165-3166.

29. Gajewski, J. J., The Claisen Rearrangement. Response to Solvents and Substituents: The Case for Both Hydrophobic and Hydrogen Bond Acceleration in Water and for a Variable Transition State. *Acc. Chem. Res.* **1997,** *30* (5), 219-225.

30. Cramer, C. J.; Truhlar, D. G., What Causes Aqueous Acceleration of the Claisen Rearrangement? *J. Am. Chem. Soc.* **1992,** *114* (23), 8794-8799.

31. White, W. N.; Wolfarth, E. F., The o-Claisen rearrangement. VIII. Solvent effects. *J. Org. Chem.* **1970,** *35* (7), 2196-2199.

32. White, W. N.; Wolfarth, E. F., ortho-Claisen rearrangement. IX. Effect of solvent on the substituent effect. *J. Org. Chem.* **1970,** *35* (10), 3585.

33. White, W. N.; Slater, C. D., The ortho-Claisen Rearrangement. V. The Products of Rearrangement of Allyl m-X-Phenyl Ethers. *J. Org. Chem.* **1961,** *26*, 3631-3638.

34. White, W. N.; Gwynn, D.; Schlitt, R.; Girard, C.; Fife, W., The ortho-Claisen Rearrangement. I. The Effect of Substituents on the Rearrangement of Allyl p-X-Phenyl Ethers. *J. Am. Chem. Soc.* **1958,** *80* (13), 3271-3277.

35. Vishnumaya, M. R.; Singh, V. K., Highly Efficient Small Organic Molecules for Enantioselective Direct Aldol Reaction in Organic and Aqueous Media. *J. Org. Chem.* **2009,** *74* (11), 4289-4297.

36. Nielsen, A. T.; Houlihan, W. J., The Aldol Condensation. *Organic Reactions* **2011**, 1-438.

37. *Modern Aldol Reactions*. Wiley-VCH: Weinheim, 2004; Vol. 1 and 2.

38. Claisen, L.; Claparede, A., Condensationen von Ketonen mit Aldehyden. *Ber. Dtsch. Chem. Ges.* **1881,** *14*, 349.

39. Schmidt, J. G., Ueber die Einwirkung von Aldehyd auf Furfurol. *Ber. Dtsch. Chem. Ges.* **1880,** *13*, 2342.

40. Haas, G. W.; Gross, M. L., Gas-Phase Base-Catalyzed Claisen-Schmidt Reactions of the Acetone Enolate Anion with Various Para-Substituted Benzaldehydes. *J. Am. Soc. Mass Spectrom.* **1996,** *7*, 82-92.

41. Dhakshinamoorthy, A.; Alvaro, M.; Garcia, H., Claisen-Schmidt Condensation Catalyzed by Metal-Organic Frameworks. *Adv. Synth. Catal.* **2010,** *352*, 711-717.

42. Wurtz, C. A., *Ber. Dtsch. Chem. Ges.* **1872,** *5*, 326.

43. Heathcock, C. H., *Comprehensive Organic Synthesis*. Pergamon: Oxford, 1991; Vol. 2.

44. House, H. O., *Modern Synthetic Reactions*. 2nd ed. ed.; W. A. Benjamin: Menlo Park, 1972.

45. List, B., Proline-Catalyzed Asymmetric Reactions. *Tetrahedron* **2002,** *58*, 5573-5590.

46. Hajos, Z. G.; Parrish, D. R., Stereocontrolled Synthesis of trans-Hydrindan Steroidal Intermediates. *J. Org. Chem.* **1973,** *38*, 3239-3243.

47. Hajos, Z. G.; Parrish, D. R., Asymmetric Synthesis of Bicyclic Intermediates of Natural Product Chemistry. *J. Org. Chem.* **1974,** *39*, 1615-1621.

48. Eder, U.; Sauer, G.; Wiechert, R., New Type of Asymmetric Cyclization to Optically Active Steroid CD Partial Structures. *Angew. Chem. Int. Ed. Engl.* **1971,** *10*, 496-497.

49. Zhu, X.; Tanaka, F.; Hu, Y.; Heine, A.; Fuller, R.; Zhong, G.; Olson, A. J.; Lerner, R. A.; Barbas III, C. F.; Wilson, I. A., The Origin of Enantioselectivity in Aldolase Antibodies: Crystal Structure, Site-directed Mutagenesis, and Computational Analysis. *J. Mol. Biol.* **2004,** *343*, 1269-1280.

50. List, B.; Lerner, R. A.; Barbas III, C. F., Proline-Catalyzed Direct Asymmetric Aldol Reactions. *J. Am. Chem. Soc.* **2000,** *122*, 2395-2396.

51. Sakthivel, K.; Notz, W.; Bui, T.; Barbas III, C. F., Amino Acid Catalyzed Direct Asymmetric Aldol Reactions: A Bioorganic Approach to Catalytic Asymmetric Carbon-Carbon Bond-Forming Reactions. *J. Am. Chem. Soc.* **2001,** *123*, 5260-5267.

52. Notz, W.; Tanaka, F.; Barbas III, C. F., Enamine-Based Organocatalysis with Proline and Diamines: The Development of Direct Catalytic Asymmetric Aldol, Mannich, Michael and Diels-Alder Reactions. *Acc. Chem. Res.* **2004,** *37* (580-591).

53. List, B., Enamine Catalysis Is a Powerful Strategy for the Catalytic Generation and Use of Carbanion Equivalents. *Acc. Chem. Res.* **2004,** *37*, 548-557.

54. List, B.; Hoang, L.; Martin, H. J., New Mechanistic Studies on the Proline-Catalyzed Aldol Reaction. *Proc. Nat. Acad. Sci. USA* **2004,** *101*, 5839.

55. Zimmerman, H. E.; Traxler, M. D., The Stereochemistry of the Ivanov and Reformatsky Reactions. I. *J. Am. Chem. Soc.* **1957,** *79*, 1920-1923.

56. Bahmanyar, S.; Houk, K. N.; Martin, H. J.; List, B., Quantum Mechanical Predictioins of the Stereoselectivites of Proline-Catalyzed Asymmetric Intermolecular Aldol Reactions. *J. Am. Chem. Soc.* **2003,** *125*, 2475-2479.

57. Wagner, J.; Lerner, R. A.; Barbas III, C. F., Efficient Aldolase Catalytic Antibodies That Use the Enamine Mechanism of Natural Enzymes. *Science* **1995,** *270*, 1797-1800.

58. Hoffmann, T.; Zhong, G.; List, B.; Shabat, D.; Anderson, J.; Gramatikova, S.; Lerner, R. A.; Barbas III, C. F., Aldolase Antibodies of Remarkable Scope. *J. Am. Chem. Soc.* **1998,** *120* (12), 2768-2779.

59. Gavrilyuk, J. I.; Wuellner, U.; III, C. F. B., b-Lactam-based Approach for the Chemical Progamming of Antibody 3C2. *Bioorg. Med. Chem. Lett.* **2009,** *19*, 1421-1424.

60. Barczyk, M.; Carracedo, S.; Gullberg, D., Integrins. *Cell Tissue Res.* **2010,** *339*, 269-280.

61. Goswami, R. K.; Bajjuri, K. M.; Forsyth, J. S.; Das, S.; Hassebpflug, W.; Huang, Z.-Z.; Lerner, R. A.; Flding-Habermann, B.; Sinha, S. C., Chemically Programmed Antibodies Targeting Multiple Alpha(v) Integrins and Their Effects on Tumor-Related Functions in Vitro. *Bioconjugate Chemistry* **2011,** *22*, 1535-1544.

62. Goswami, R. K.; Liu, Y.; Liu, C.; Lerner, R. A.; Sinha, S. C., Synthesis and Evaluation of the Aldolase Antibody-Derived Chemical-Antibodies Targeting a5b1 Integrin. *Molecular Pharmaceutics* **2012,** *ASAP*.

63. Eichorn, E., C. A. Davey, D. F. Sargent, T. Leisinger, T. J. Richmond, Crystal Structure of *Escherichia coli* Alkanesulfonate Monooxygenase SsuD. *J. Mol. Biol.* **2002,** *324*, 457-468.

64. Eichhorn, E.; van der Ploeg, J. R.; Leisinger, T., Characterization of a Two-Component Alkanesulfonate Monooxygenase from *Escherichia coli*. *Journal of Biological Chemistry* **1999,** *274*, 26639-26646.

65. Zhan, X.; Carpenter, R. A.; Ellis, H. R., Catalytic Importance of the Substrate Binding Order for the $FMNH_2$-Dependent Alkanesulfonate Monooxygenase Enzyme. *Biochemistry* **2008,** *47*, 2221-2230.

66. Carpenter, R. A.; Zhan, X.; Ellis, H. R., Catalytic Role of a Conserved Cysteine Residue in the Desulfonation Reaction by the Alkanesulfonate Monooxygeanse Enzyme. *Biochimica et Biophysica Acta* **2010,** *1804*, 97-105.

67. Carpenter, R. A., J. Xiong, J. M. Robbins, H. R. Ellis, Functional Role of a Conserved Arginine Residue Located on a Mobile Loop of Alkanesulfonate Monooxygenase. *Biochemistry* **2011,** *50*, 6469-6477.

68. Ellis, H. R., Mechanism for Sulfur Acquisition by the Alkanesulfonate Monooxygenase System. *Bioorganic Chemistry* **2011,** *39*, 178-184.

69. Gao, B.; Ellis, H. R., Mechanism of Flavin Reduction in the Alkanesulfonate Monooxygenase System. *Biochimica et Biophysica Acta* **2007,** *1774*, 359-367.

70. Ellis, H. R., The FMN-Dependent Two-Component Monooxygenase Systems. *Archives of Biochemistry and Biophysics* **2010,** *497*, 1-12.

71. Eichhorn, E.; Davey, C. A.; Sargent, D. F.; Leisinger, T.; Richmond, T. J., Crystal Structure of *Escheria coli* Alkanesulfonate Monooxygenase SsuD. *J. Mol. Biol.* **2002,** *324*, 457-468.

72. Massey, V., The Chemical and Biological Versatility of Riboflavin. *Biochem. Soc. Trans.* **2000,** *28*, 283-296.

73. Strukul, G., Transition Metal Catalysis in the Baeyer-Villiger Oxidation of Ketones. *Angew. Chem. Int. Ed.* **1998,** *37*, 1198-1209.

74. Reusch, W. Oxidations. http://www2.chemistry.msu.edu/faculty/reusch/VirtTxtJml/addene2.htm#add4b.

75. Acevedo, O.; Jorgensen, W. L., Advances in Quantum and Molecular Mechanical (QM/MM) Simulations for Organic and Enzymatic Reactions. *Acc. Chem. Res.* **2010,** *43*, 142-151.

76. Warshel, A.; Levitt, M., Theoretical Studies of Enzymic Reactions : Dielectric, Electrostatic and Steric Stabilization of Reaction of Lysozyme. *J. Mol. Biol.* **1976,** *103*, 227-249.

77. Senn, H. M.; Thiel, W., QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed.* **2009,** *48*, 1198-1229.

78. Field, M. J.; Bash, P. A.; Karplus, M., A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynammics Simulations. *J. Comput. Chem.* **1990,** *11*, 700-733.

79. Cramer, C., *Essentials of Computational Chemistry*. John Wiley & Sons: West Sussex, 2004.

80. Leach, A. R., *Molecular Modelling, Principles and Applications*. 2nd ed.; Pearson Prentice Hall: Essex, 2001.

81. Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M., Frontier Bonds in QM/MM Methods: A Comparison of Different Approaches. *J. Phys. Chem. A* **2000,** *104* (8), 1720-1735.

82. Kaminski, G. W.; Jorgensen, W. L., A Quantum Mechanical and Molecular Mechanical Method Based on CM1A Charges: Applications to Solvent Effects on Organic Equilibria and Reactions. *J. Phys. Chem.* **1998,** *102*, 1787-1796.

83. Storer, J. W.; Giesen, D. J.; Cramer, C. C.; Truhlar, D. G., Class IV Charge Models: A New Semiempirical Approach in Quantum Chemistry. *J. Comput.-Aided Mol. Des.* **1995,** *9*, 87-110.

84. Udier-Blagovic, M.; Morales de Tirado, P.; Pearlman, S. A.; Jorgensen, W. L., Accuracy of Free Energies of Hydration Using CM1 and CM3 Atomic Charges. *J. Comput. Chem.* **2004,** *25*, 1322-1332.

85. Naeem, R. Lennard-Jones Potential. chemwiki.ucdavis.edu/Physical_Chemistry/Quantum_Mechanics/Intermolecular_Forces/Lennard-Jones_Potential.

86. Wang, C. Y.; Zhang, X., Multiscale Modeling and Related Hybrid Approaches. *Curr. Opin. Solid State Mater. Sci.* **2006,** *10*, 2-14.

87. Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L., PDDG/PM3 and PDDG/MNDO: Improved semiempirical methods. *J. Comput. Chem.* **2002,** *23*, 1601-1622.

88. Tubert-Brohman, I.; Guimarães, C. R. W.; Repasky, M. P.; Jorgensen, W. L., Extension of the PDDG/PM3 and PDDG/MNDO semiempirical molecular orbitial methods to the halogens. *J. Comput. Chem.* **2003,** *25*, 138-150.

89. Tubert-Brohman, I.; Guimarães, C. R. W.; Jorgensen, W. L., Extension of the PDDG/PM3 Semiempirical Molecular Orbital Method to Sulfur, Silicon, and Phosphorus. *J. Chem. Theory Comput.* **2005,** *1*, 817-823.

90. Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L., Improved Semiempirical Heats of Formation Through the Use of Bond and Group Equivalents. *J. Comput. Chem.* **2002,** *23*, 498-510.

91. Bernal-Uruchurtu, M. I.; Ruiz-Lopez, M. F., Basic Ideas for the Correction of Semiempirical Methods Describing H-Bonded Systems. *Chem. Phys. Lett.* **2000,** *330*, 118-124.

92. Bernal-Uruchurtu, M. I.; Martins-Costa, M. T. C.; Millot, C.; Ruiz-Lopez, M. F., Improving Description of Hydrogen Bonds at the Semiempirical Level: Water–Water Interactions as Test Case. *J. Comput. Chem.* **2000,** *21*, 572-581.

93. Miertus, S.; Scrocco, E.; Tomasi, J., Electrostatic Interaction of a Solute with a Continuum. A Direct Utilization of *ab initio* Molecular Potentials for the Prevision of Solvent Effects. *Chem. Phys.* **1981,** *55*, 117-129.

94. Cossi, M.; Scalmani, G.; Rega, N.; Barone, V., New Developments in the Polarizable Continuum Model for Quantum Mechanical and Classical Calculations on Molcules in Solution. *J. Chem. Phys.* **2002,** *117*, 43-54.

95. Cossi, M., Continuum Solvation Model for Infinite Periodic Systems. *Chem. Phys. Lett.* **2004,** *384*, 179-184.

96. Frediani, L.; Cammi, R.; Corni, S.; Tomasi, J., A Polarizable Continuum Model for Molecules at Diffuse Interfaces. *J. Chem. Phys.* **2004,** *120*, 3893-3907.

97. Klamt, A.; Schuurmann, G., COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and its Gradient. *Journal of the Chemistry Society, Perkin Transations* **1993,** *2*, 799-805.

98. Klamt, A.; Jonas, V.; Burger, T.; Lohrenz, J. C. W., Refinements and Parametrisation of COSMO-RS. *J. Phys. Chem.* **1998,** *102*, 5704-5085.

99. Klamt, A., Conductor-like Screening Model for Real Solvent: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995,** *99*, 2224-2235.

100. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J., Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996,** *118* (45), 11225-11236.

101. Jorgensen, W. L.; Thomas, L. L., Perspective on Free-Energy Perturbation Calculations for Chemical Equilibria. *J. Chem. Theory Comput.* **2008,** *4*, 869-876.

102. Jorgensen, W. L.; Ravimohan, C., Monte Carlo simulation of differences in free energies of hydration. *J. Chem. Phys.* **1985,** *83* (6), 3050-3054.

103. Zwanzig, R. W., High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954,** *22*, 1420-1426.

104. Kirkwood, J. G., Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935,** *3,* 300-313.

105. Roux, B., The Calculation of the Potential of Mean Force Using Computer Simulations. *Computer Physics Communications* **1995,** *91*, 275-281.

106. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E., Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* **1953,** *21,* 1087-1092.

107. Ceperley, D. M. In *Metropolis Methods for Quantum Monte Carlo Simulations,* AIP Conference Proceedings, Gubernatis, J. E., Ed. 2003; pp 85-98.

108. Le Roux, S., Petkov, V. Model Box Periodic Boundary Conditions - P. B. C. isaacs.sourceforge.net/phys/pbc.html.

109. van Gunsteren, W. F.; Berendsen, H. J. C., Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry. *Angew. Chem. Int. Ed. Engl.* **1990,** *29,* 992-1023.

110. Haile, J. M., *Molecular Dynamic Simulation: Elementary Methods*. John Wiley & Sons: New York, 1990.

111. Hoover, W. G., *Lecture Notes in Physics: Molecular Dynamics*. Springer-Verlag: Berlin, 1986.

112. Nakano, A., Molecular Dynamics. University of Southern California.

113. Verlet, L., Computer Experiments on Classical Fluids. *Physical Review* **1967,** *159*, 98-103.

114. Hockney, R. W., The Potential Calculation and Some Applications. *Methods in Computational Physics* **1970,** *9,* 136-211.

115. Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R., A Computer-Simulation Method for the Calculation of Equilibrium-Constants for the Formation of Physical Clusters of Molecules—Application to Small Water Clusters. *J. Chem. Phys.* **1982,** *76,* 637-649.

116. Allen, M. P.; Tildesley, D. J., *Computer Simulation of Liquids*. Oxford University Press: Oxford, 1987.

117. Gear, C. W., *Numerical Initial Value Problems in Ordinary Differential Equations*. Prentice Hall: Englewood Cliffs, 1971.

118. Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T., *Numerical Recipes*. Cambrdige University Press: New York, 1986.

119. 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. *Journal of Computational Physics* **1977,** *23*, 327-341.

120. Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham III, T. E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P., AMBER, A Package of Computer Programs For Applying Molecular Mechanics, Normal Mode Analysis, Molecular Dynamics and Free Energy Calculations to Simulate the Structural and Energetic Properties of Molecules. *Computer Physics Communications* **1995,** *91*, 1-41.

121. Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr., K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005,** *26*, 1668-1688.

122. Case, D. A.; Darden, T. A.; Cheatham, I., T. E. ; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M. J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *Amber 12 Reference Manual*, University of California, San Francisco, 2012.

123. Wang, J. W., R. M.; Caldwell, J. W.; Kollamn, P. A.; Case, D. A., Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004,** *25*, 1157–1174.

124. Salomon-Ferrer, R.; Case, D. A.; Walker, R. C., An Overview of the Amber Biomolecular Simulation Package. *WIREs Comput Mol Sci* **2012**.

125. D.A. Case, T. A. D., T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R.; Luo, R. C. W., W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra,; J. Swails, A. W. G., I. Kolossváry, K.F.Wong, F. Paesani, J. Vanicek, R.M.Wolf, J. Liu,; X. Wu, S. R. B., T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G.; Cui, D. R. R., D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T.; Luchko, S. G., A. Kovalenko, and P.A. Kollman University of California, San Francisco, 2012.

126. Darden, T.; York, D.; Pedersen, L., Particle Mesh Ewald: An N-log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993,** *98*, 10089-10092.

127. Jackson, J.; Edward, A., *A User's Guide to Principal Components*. John Wiley & Sons: Hoboken, 2003.

128. Wold, S.; Esbensen, K.; Geladi, P., Principal Component Analysis. *Chemometrics and Intelligent Laboratory Systems* **1987,** *2*, 37-52.

129. Jolliffe, I. T., *Principal Component Analysis*. Springer Science and Business Media: New York, 2004.

130. Flury, B., *Common Principal Components and Related Multivariate Models*. John Wiley & Sons: New York, 1988.

131. Dunteman, H. G., *Principal Component Analysis*. Sage Publications: Newbury Park, 1989.

132. Cattell, R. B., The Scree Test for the Number of Factors. *Multivariate Behavorial Research* **1966,** *1*, 245-276.

133. Kaiser, H. F., The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement* **1960,** *XX*, 141-151.

134. Jolliffe, I. T., Discarding Variablies in a Principal Component Analysis. I: Artifical Data. *Appl. Statist.* **1972,** *21*, 160-173.

135. Flury, B., Principal Points. *Biometrika* **1990,** *77*, 33-41.

136. Flury, B., Estimation of Principal Points. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **1993,** *42*, 139-151.

137. Grant, B.; Rodrigues, A. P. C.; ElSawy, K. M.; McCammon, J. A.; Caves, L. S. D., Bio3d: An R Package for the Comparative Analysis of Protein Structures. *Bioinformatics* **2006,** *22*, 2695-2696.

138. Chanda, A. V. F., Organic Synthesis "On Water". *Chem. Rev.* **2009,** *109*, 725-748.

139. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983,** *79*, 926-935.

140. Jorgensen, W. L.; Tirado-Rives, J., Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Nat. Acad. Sci. USA* **2005,** *102*, 6665-6670.

141. Halgren, T. A., The representation of van der Waals (vdW) interactions in molecular mechanics force fields: potential form, combination rules, and vdW parameters. *J. Am. Chem. Soc.* **1992,** *114* (20), 7827-7843.

142. Sheppard, A. N.; Acevedo, O., Multidimensional Exploration of Valley-Ridge Inflection Points on Potential Energy Surfaces. *J. Am. Chem. Soc.* **2009,** *131*, 2530-2540.

143. Sambasivarao, S. V.; Acevedo, O., Development of OPLS-AA Force Field Parameters for 68 Unique Ionic Liquids. *J. Chem. Theory Comput.* **2009,** *5*, 1038-1050.

144. Acevedo, O.; Squillacote, M. E., A New Solvent-Dependent Mechanism for a Triazolinedione Ene Reaction. *J. Org. Chem.* **2008,** *73*, 912-922.

145. Acevedo, O.; Jorgensen, W. L., Understanding Rate Accelerations for Diels-Alder Reactions in Solution Using Enhanced QM/MM Methodology. *J. Chem. Theory Comput.* **2007,** *3*, 1412-1419.

146. Acevedo, O.; Jorgensen, W. L.; Evanseck, J. D., Elucidation of Rate Variations for a Diels-Alder Reaction in Ionic Liquids from QM/MM Simulations. *J. Chem. Theory. Comput.* **2007,** *3* (1), 132-138.

147. Acevedo, O.; Jorgensen, W. L., Cope Elimination: Elucidation of Solvent Effects from QM/MM Simulations. *J. Am. Chem. Soc.* **2006,** *128*, 6141-6146.

148. Acevedo, O.; Jorgensen, W. L., Medium Effects on the Decarboxylation of a Biotin Model in Pure and Mixed Solvents from QM/MM Simulations. *J. Org. Chem.* **2006,** *71*, 4896-4902.

149. Acevedo, O.; Jorgensen, W. L., Solvent Effects on Organic Reactions from QM/MM Simulations. In *Annual Reports in Computational Chemistry*, Spellmeyer, D., Ed. Elsevier: 2006; Vol. 2, pp 263-278.

150. Tubert-Brohman, I.; Acevedo, O.; Jorgensen, W. L., Elucidation of Hydrolysis Mechanisms for Fatty Acid Amide Hydrolase and Its Lys142Ala Variant via QM/MM Simulations. *J. Am. Chem. Soc.* **2006,** *128*, 16904-16913.

151. Acevedo, O.; Jorgensen, W. L., Influence of Inter- and Intramolecular Hydrogen Bonding on Kemp Decarboxylations from QM/MM Simulations. *J. Am. Chem. Soc.* **2005,** *127* (24), 8829-8834.

152. Acevedo, O.; Jorgensen, W. L., Solvent Effects and Mechanism for a Nucleophilic Aromatic Substitution from QM/MM Simulations. *Org. Lett.* **2004,** *6* (17), 2881-2884.

153. Thompson, J. D.; Cramer, C. J.; Truhlar, D. G., Parameterization of charge model 3 for AM1, PM3, BLYP, and B3LYP. *J. Comput. Chem.* **2003,** *24* (11), 1291-1304.

154. Blagović, M. U.; Morales de Tirado, P.; Pearlman, S. A.; Jorgensen, W. L., Accuracy of Free Energies of Hydration from CM1 and CM3 Atomic Charges. *J. Comput. Chem.* **2004,** *25*, 1322-1332.

155. Zimmerli, U.; Parrinello, M.; Koumoutsakos, P., Dispersion corrections to density functionals for water aromatic interactions. *J. Chem. Phys.* **2004,** *120*, 2693-2699.

156. Jorgensen, W. L.; Tirado-Rives, J., Molecular Modeling of Organic and Biomolecular Systems Using BOSS and MCPRO. *J. Comput. Chem.* **2005,** *26*, 1689-1700.

157. Bertelsen, S.; Jørgensen, K. A., Organocatalysis—after the gold rush. *Chem. Soc. Rev.* **2009,** *38*, 2178-2189.

158. Tang, Z.; Jiang, F.; Yu, L.-T.; Cui, X.; Gong, L.-Z.; Mi, A.-Q.; Jiang, Y.-Z.; Wu, Y.-D., Novel Small Organic Molecules for a Highly Enantioselective Direct Aldol Reaction. *J. Am. Chem. Soc.* **2003,** *125* (18), 5262-5263.

159. Frisch, M. J. T., 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.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida,

M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Gaussian, Inc.: Wallingford, CT, 2009.

160. Guner, V.; Khuong, K. S.; Leach, A. G.; Lee, P. S.; Bartberger, M. D.; Houk, K. N., A Standard Set of Pericyclic Reactions of Hydrocarbons for the Benchmarking of Computational Methods: The Performance of ab Initio, Density Functional, CASSCF, CASPT2, and CBS-QB3 Methods for the Prediction of Activation Barriers, Reaction Energetics, and Transition State Geometries. *J. Phys. Chem. A* **2003,** *107*, 11445-11459.

161. Szabo, A.; Ostlund, N. S., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, Inc.: 1996.

162. Yau, H. M.; Howe, A. G.; Hook, J. M.; Croft, A. K.; Harper, J. B., Solvent reorganisation as the driving force for rate changes of Menschutkin reactions in an ionic liquid. *Org. Biomol. Chem.* **2009,** *7*, 3572-3575.

163. D'Anna, F.; Frenna, V.; Noto, R.; Pace, V.; Spinelli, D., Can the Absence of Solvation of Neutral Reagents by Ionic Liquids Be Responsible for the High Reactivity in

Base-Assisted Intramolecular Nucleophilic Substitutions in These Solvents? *J. Org. Chem.* **2005,** *70* (7), 2828-2831.

164. Ho, M.-C.; Ménétret, J.-F.; Tsuruta, H.; Allen, K. N., The origin of the electrostatic perturbation in acetoacetate decarboxylase. *Nature* **2009,** *459*, 393-399.

165. Allemann, C.; Uma, J. M.; Houk, K. N., Computational investigations of the stereoselectivities of proline-related catalysts for aldol reactions. *J. Mol. Catal. A: Chem.* **2010,** *324*, 31-38.

166. Sattelmeyer, K. W.; Tubert-Brohman, I.; Jorgensen, W. L., NO-MNDO: Reintroduction of the Overlap Matrix into MNDO. *J. Chem. Theory Comput.* **2006,** *2*, 413-419.

167. Barbas III, C. F., Organocatalysis Lost: Modern Chemistry, Ancient Chemistry, and an Unseen Biosynthetic Apparatus. *Angew. Chem. Int. Ed.* **2008,** *47*, 42-47.

168. Mukherjee, S.; Yang, J. W.; Hoffmann, S.; List, B., Asymmetric Enamine Catalysis. *Chem. Rev.* **2007,** *107* (12), 5471-5569.

169. Chandrasekhar, J.; Shariffskul, S.; Jorgensen, W. L., QM/MM Simulations of Cycloaddition Reactions in Water: Contribution of Enhanced Hydrogen Bonding at the Transition State to the Solvent Effects. *J. Phys. Chem. B* **2002,** *106*, 8078-8085.

170. Gunaydin, H.; Acevedo, O.; Jorgensen, W. L.; Houk, K. N., Computation of Accurate Barriers for Methyl Transfer Reactions of Sulfonium and Ammonium Salts in Aqueous Solution. *J. Chem. Theory Comput.* **2007,** *3*, 1028-1035.

171. Irani, M.; Haqgu, M.; Talebi, A.; Gholami, M. R., A joint experimental and theoretical study of kinetic and mechanism of rearrangement of allyl p-tolyl ether *J. Mol. Struct. (THEOCHEM)* **2009,** *893* (1-3), 73-76.

172. Eder, U.; Sauer, G.; Wiechert, R., New Type of Asymmetric Cyclization to Optically Active Steroid CD Partial Structures. *Angew. Chem. Int. Ed.* **1971,** *10* (7), 496-497.

173. Krenske, E. H.; Houk, K. N., Aromatic Interactions as Control Elements in Stereoselective Organic Reactions. *Acc. Chem. Res.* **2012**.

174. Machajewski, T. D.; Wong, C., The Catalytic Asymmetric Aldol Reaction. *Angew. Chem. Int. Ed.* **2000,** *39*, 1352-1374.

175. Dziedzic, P. Z., W.; Hafren, J.; Cordova, A., The Small Peptide-Catalyzed Direct Asymmetric Aldol Reaction in Water. *Org. Biomol. Chem.* **2006,** *4*, 38-40.

176. Hispanics in the United States.

177. Wick, C. D.; Kuo, I.-F. W.; Mundy, C. J.; Dang, L. X., The Effect of Polarizability for Understanding the Molecular Structure of Aqueous Interfaces. *J. Chem. Theory Comput.* **2007,** *3*, 2002-2010.

178. Melchiorre, P.; Marigo, M.; Carlone, A.; Bartoli, G., Asymmetric Aminocatalysis - Gold Rush in Organic Chemistry. *Angew. Chem. Int. Ed.* **2008,** *47* (33), 6138-6171.

179. Jorgensen, W. L.; Jensen, K. P.; Alexandrova, A. N., Polarization Effects for Hydrogen-Bonded Complexes of Substituted Phenols with Water and Chloride Ion. *J. Chem. Theory Comput.* **2007,** *3*, 1987-1992.

180. Guizzetti, S.; Benaglia, M.; Raimondi, L.; Celentano, G., Enantioselective Direct Aldol Reaction “on Water” Promoted by Chiral Organic Catalysts. *Org. Lett.* **2007,** *9* (7), 1247-1250.

181. Ribas-Arino, J.; Carvajal, M. A.; Chaumont, A.; Masia, M., Unraveling the Role of Water in the Stereoselective Step of Aqueous Proline-Catalyzed Aldol Reactions. *Chem. Eur. J.* **2012,** *18*, 15868–15874.

182. Aratake, S.; Itoh, T.; Okano, T.; Usui, T.; Shoji, M.; Hayashi, Y., Small organic molecule in enantioselective, direct aldol reaction “in water”. *Chem. Commun.* **2007**, 2524-2526.

183. Notz, W.; Tanaka, F.; Barbas III, C. F., Enamine-Based Organocatalysis with Proline and Diamines: The Development of Direct Catalytic Asymmetric Aldol, Mannich, Michael, and Diels−Alder Reactions. *Acc. Chem. Res.* **2004,** *37* (8), 580-591.

184. List, B., Enamine Catalysis Is a Powerful Strategy for the Catalytic Generation and Use of Carbanion Equivalents. *Acc. Chem. Res.* **2004,** *37* (8), 548-557.

185. Chang, T.-M.; Dang, L. X., Recent Advances in Molecular Simulations of Ion Solvation at Liquid Interfaces. *J. Am. Chem. Soc.* **2006,** *106* (4), 1305-1322.

186. Hayashi, Y.; Aratake, S.; Okano, T.; Takahashi, J.; Sumiya, T.; Shoji, M., Combined Proline-Surfactant Organocatalyst for the Highly Diastereo- and Enantioselective Aqueous Direct Cross-Aldol Reaction of Aldehydes. *Angew. Chem. Int. Ed.* **2006,** *45* (33), 5527-5529.

187. White, W.; Slater, C.; Fife, W., Communications. The Electronic Nature of the Transition State of the Claisen Rearrangement. *J. Org. Chem.* **1961,** *26*, 627-628.

188. Cohen, N., Asymmetric induction in 19-norsteroid total synthesis. *Acc. Chem. Res.* **1976,** *9* (11), 412-417.

189. Angelini, G.; Maria, P. D.; Chiappe, C.; Fontana, A.; Gasbarri, C.; Siani, G., The Base-Catalyzed Keto−Enol Interconversion of 2-Nitrocyclohexanone in Ionic Liquids. *J. Org. Chem.* **2009,** *74* (17), 6572-6576.

190. Zhao, Q.; Lam, Y.-h.; Kheirabadi, M.; Xu, C.; Houk, K. N.; Schafmeister, C. E., Hydrophobic Substituent Effects on Proline Catalysis of Aldol Reactions in Water. *J. Org. Chem.* **2012,** *77*, 4784-4792.

191. Acevedo, O., Role of Water in the Multifaceted Catalytic Antibody 4B2 for Allylic Isomerization and Kemp Elimination Reactions. *J. Phys. Chem. B* **2009,** *113*, 15372-15381.

192. Anastas, P.; Eghbali, N., Green Chemistry: Principles and Practice. *Chem. Soc. Rev.* **2010,** *39*, 301-312.

193. Jurecka, P.; Sponer, J.; Cerný, J.; Hobza, P., Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* **2006,** *8*, 1985-1993.

194. List, B., Proline-catalyzed asymmetric reactions. *Tetrahedron* **2002,** *58* (28), 5573-5590.

195. MacMillan, D. W. C., The advent and development of organocatalysis. *Nature* **2008,** *455*, 304-308.

196. List, B.; Lerner, R. A.; Barbas III, C. F., Proline-Catalyzed Direct Asymmetric Aldol Reactions. *J. Am. Chem. Soc.* **2000,** *122* (10), 2395-2396.

197. Allemann, C.; Gordillo, R.; Clemente, F. R.; Cheong, P. H.-Y.; Houk, K. N., Theory of Asymmetric Organocatalysis of Aldol and Related Reactions:Rationalizations and Predictions. *Acc. Chem. Res.* **2004,** *37* (8), 558-569.

198. Bock, D. A.; Lehmann, C. W.; List, B., Organocatalysis Special Feature: Crystal structures of proline-derived enamines. *Proc. Nat. Acad. Sci. USA* **2010,** *107* (48), 20636-20641.

199. Bisai, V.; Bisai, A.; Singh, V. K., Enantioselective organocatalytic aldol reaction using small organic molecules. *Tetrahedron* **2012,** *68*, 4541-4580.

200. Dondoni, A.; Massi, A., Asymmetric Organocatalysis: From Infancy to Adolescence. *Angew. Chem. Int. Ed.* **2008,** *47*, 4638–4660.

201. Maya, V.; Raj, M.; Singh, V. K., Highly Enantioselective Organocatalytic Direct Aldol Reaction in an Aqueous Medium. *Org. Lett.* **2007,** *9* (13), 2593-2595.

202. Raj, M.; Vishnumaya; Ginotra, S. K.; Singh, V. K., Highly Enantioselective Direct Aldol Reaction Catalyzed by Organic Molecules. *Org. Lett.* **2006,** *8* (18), 4097-4099.

203. Tang, Z.; Yang, Z.-H.; Chen, X.-H.; Cun, L.-F.; Mi, A.-Q.; Jiang, Y.-Z.; Gong, L.-Z., A Highly Efficient Organocatalyst for Direct Aldol Reactions of Ketones with Aldedydes. *J. Am. Chem. Soc.* **2005,** *127* (25), 9285-9289.

204. Tang, G.; Hu, X.; Altenbach, H. J., L-Prolinamide derivatives as efficient organocatalysts for aldol reactions on water. *Tetra. Lett.* **2011,** *52* (52), 7034-7037.

205. Bhowmick, S.; Bhowmick, K. C., Catalytic asymmetric carbon–carbon bond-forming reactions in aqueous media. *Tetrahedron Asymmetry* **2011,** *22* (23), 1945-1979.

206. Mase, N.; Nakai, Y.; Ohara, N.; Yoda, H.; Takabe, K.; Tanaka, F.; Barbas III, C. F., Organocatalytic Direct Asymmetric Aldol Reactions in Water. *J. Am. Chem. Soc.* **2006,** *128* (3), 734-735.

207. Hayashi, Y.; Sumiya, T.; Takahashi, J.; Gotoh, H.; Urushima, T.; Shoji, M., Highly Diastereo- and Enantioselective Direct Aldol Reactions in Water. *Angew. Chem. Int. Ed.* **2006,** *45* (6), 958-961.

208. Klijn, J. E.; Engberts, J. B. F. N., Organic chemistry: Fast reactions 'on water'. *Nature* **2005,** *435* (9), 746-747.

209. Barbas III, C. F.; Heine, A.; Zhong, G.; Hoffmann, T.; Gramatikova, S.; Björnestedt, R.; List, B.; Anderson, J.; Stura, E. A.; Wilson, I. A.; Lerner, R. A., Immune Versus Natural Selection: Antibody Aldolases with Enzymic Rates But Broader Scope. *Science* **1997,** *278*, 2085-2092.

210. Zhong, G.; Hoffmann, T.; Lerner, R. A.; Danishefsky, S.; Barbas III, C. F., Antibody-Catalyzed Enantioselective Robinson Annulation. *J. Am. Chem. Soc.* **1997,** *119* (34), 8131-8132.

211. Zhu, X.; Tanaka, F.; Lerner, R. A.; Barbas III, C. F.; Wilson, I. A., Direct Observation of an Enamine Intermediate in Amine Catalysis. *J. Am. Chem. Soc.* **2009,** *131* (51), 18206-18207.

212. Bahmanyar, S.; Houk, K. N., Transition States of Amine-Catalyzed Aldol Reactions Involving Enamine Intermediates: Theoretical Studies of Mechanism, Reactivity, and Stereoselectivity. *J. Am. Chem. Soc.* **2001,** *123* (45), 11273-11283.

213. Bahmanyar, S.; Houk, K. N.; Martin, H. J.; List, B., Quantum Mechanical Predictions of the Stereoselectivities of Proline-Catalyzed Asymmetric Intermolecular Aldol Reactions. *J. Am. Chem. Soc.* **2003,** *125* (9), 2475-2479.

214. Clemente, F. R.; Houk, K. N., Computational Evidence for the Enamine Mechanism of Intramolecular Aldol Reactions Catalyzed by Proline. *Angew. Chem. Int. Ed.* **2004,** *43* (43), 5766-5768.

215. Clemente, F. R.; Houk, K. N., Theoretical Studies of Stereoselectivities of Intramolecular Aldol Cyclizations Catalyzed by Amino Acids. *J. Am. Chem. Soc.* **2005,** *127* (32), 11294-11302.

216. Zhang, X.; Houk, K. N., Acid/Base Catalysis by Pure Water: The Aldol Reaction. *J. Org. Chem.* **2005,** *70* (24), 9712-9716.

217. Allen, C.; Sambasivarao, S. V.; Acevedo, O., An Ionic Liquid Dependent Mechanism for Base Catalyzed β-Elimination Reactions from QM/MM Simulations. *J. Am. Chem. Soc.* **2013,** *135*, 1065-1072.

218. Vilseck, J. Z.; Sambasivarao, S. V.; Acevedo, O., Optimal Scaling Factors for CM1 and CM3 Atomic Charges in Aqueous RM1-Based Simulations. *J. Comput. Chem.* **2011,** *32*, 2836-2842.

219. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L., Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with

Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001,** *105*, 6474-6487.

220. Reymond, J.-L., Stereoselectivity of aldolase catalytic antibodies. *J. Mol. Catal. B: Enzym.* **1998,** *5*, 331-337.

221. Arnó, M.; Domingo, L. R., Theozyme for antibody aldolases. Characterization of the transition-state analogue. *Org. Biomol. Chem.* **2003,** *1*, 637-643.

222. Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas III, C. F.; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D., De Novo Computational Design of Retro-Aldol Enzymes. *Science* **2008,** *319*, 1387-1391.

223. Carpenter, R. A.; Zhan, X.; Ellis, H. R., Catalytic Role of a Conserved Cysteine Residue in the Desulfonation Reaction by the Alkanesulfonate Monooxygenase Enzyme. *Biochimica et Biophysica Acta* **2010,** *1804*, 97-105.

224. Ellis, H. R.; Gao, B., Mechanism of Flavin Reduction in the Alkanesulfonate Monooxygenase System. *Biochimica et Biophysica Acta* **2007,** *1774*, 359-367.

225. Carpenter, R. A.; Xiong, J.; Robbins, J. M.; Ellis, H. R., Functional Role of a Conserved Arginine Residue Located on a Mobile Loop of Alkanesulfonate Monooxygenase. *Biochemistry* **2011,** *50*, 6469-6477.

226. Eichhorn, E.; Davey, C. A.; Sargent, D. F.; Leisinger, T.; Richmond, T. J., Crystal Structure of *Escherichia coli* Alkanesulfonate Monooxygenase SsuD. *Journal of Molecular Biology* **2002,** *324*, 457-468.

227. Campbell, Z. T.; Weichsel, A.; Montfort, W. R.; Baldwin, T. O., Crystal Structure of the Bacterial Luciferase/Flavin Complex Provides Insight into the Function of the b Subunit. *Biochemistry* **2009,** *48*, 6085-6094.

228. Li, L.; Liu, X.; Yang, W.; Xu, F.; Wang, W.; Feng, L.; Bartlam, M.; Wang, L.; Rao, Z., Crystal Structure of Long-Chain Alkane Monooxygnease (LadA) in Complex with Coenzyme FMN: Unveiling the Long-Chain Alkane Hydroxylase. *Journal of Molecular Biology* **2008,** *376*, 453-465.

229. Robbins, J. M.; Ellis, H. R., Identification of Critical Steps Governing the Two-Component Alkanesulfonate Monooxygenase Catalytic Mechanism. *Biochemistry* **2012,** *51*, 6378-6387.

230. Ferrario, V.; Braiuca, P.; Tessaro, P.; Knapic, L.; Gruber, C.; Pleiss, J.; Ebert, C.; Eichhorn, E.; Gardossi, L., Elucidating the Structural and Conformational Factors Responsible for the Activity and Substrate Specificity of Alkanesulfonate Monooxygenase. *Journal of Biomolecular Structure and Dynamics* **2012,** *30*, 74-88.

231. Eswar, N.; Marti-Renom, A.; Webb, B.; Madhusudhan, M. S.; Eramian, D.; Shen, M.; Pieper, U.; Sali, A., Comparative Protein Structure Modeling with MODELLER. In *Current Protocols in Bioinformatics*, John Wiley & Sons, Inc.: 2006; pp 5.6.1-5.6.30.

232. Li, L.; Liu, X.; Yang, W.; Xu, F.; Wang, W.; Feng, L.; Bartlam, M.; Wang, L.; Rao, Z., Crystal Structure of Long-Chain Alkane Monooxygenase (LadA) in Complex with Coenzyme FMN: Unveiling the Long-Chain Alkane Hydroxylase. *J. Mol. Biol.* **2008,** *376*, 453-465.

233. Trott, O.; Olson, A. J., AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization and Multithreading. *Journal of Computational Chemistry* **2010,** *31*, 455-461.

234. Case, D. A.; Darden, T. A.; Cheatham, I., T. E. ; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M. J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER12*, University of California, San Francisco, 2012.

235. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Structure, Function and Bioinformatics* **2006,** *65*, 712-725.

236. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and Testing of a General Amber Force Field. *Journal of Computational Chemistry* **2004,** *25*, 1157-1174.

237. Jorgensen, W. L., BOSS - Biochemical and Organic Simulation System. In *Encyclopedia of Computational Chemistry*, Schleyer, R., Ed. John Wiley & Sons Ltd.: Athens, USA, 1998; Vol. 5, pp 3281-3285.

238. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J., Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996,** *118*, 11225-11236.

239. Woods, R. J.; Chappelle, R., Restrained Electrostatic Potential Atomic Partial Charges for Condensed-Phase Simulations of Carbohydrates. *J. Mol. Struct. (THEOCHEM)* **2000,** *527*, 149-156.

240. Goetz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C., Routine Microsecond Molecular Dynamics Simulations wither AMBER - Part I: Generalized Born. *J. Chem. Theory Comput.* **2012,** *8*, 1542-1555.

241. Grant, B. J.; Rodrigues, A. P. C.; ElSawy, K. M.; McCammon, J. A.; Caves, L. S. D., Bio3d: an R Package for the Comparative Analysis of Protein Structures. *Structural Bioinformatics* **2006,** *22*, 2695-2696.

242. Humphrey, W.; Dalke, A.; Schulten, K., VMD - Visual Molecular Dynamics. *J. Molec. Graphics* **1996,** *14*, 33-38.

243. Shao, J.; Tanner, S. W.; Thompson, N.; Cheatham III, T. E., Clustering Molecular Dynamics Trajectories: 1. Characterizing the Performance of Different Clustering Algorithms. *Journal of Chemical Theory and Computation* **2007,** *3*, 2312-2334.

244. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera - A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004,** *25*, 1605-1612.

245. Holzman, T. F.; Baldwin, T. O., The Effects of Phosphate on the Structure and Stability of the Luciferase from *Beneckea harvei, Photobacterium fischeri,* and *Photobacterium phosphoreum*. *Biochem. Biophys. Res. Commun.* **1980,** *94*, 1199-1206.

246. Beam, M. P.; Bosserman, M. A.; Noinaj, N.; Wenhenkel, M.; Rohr, J., Crystal Structure of Baeyer-Villiger Monooxygenase MtmOIV, the Key Enzyme of the Mithramycin Biosynthetic Pathway. *Biochemistry* **2009,** *48*, 4476-4487.

# Appendix I

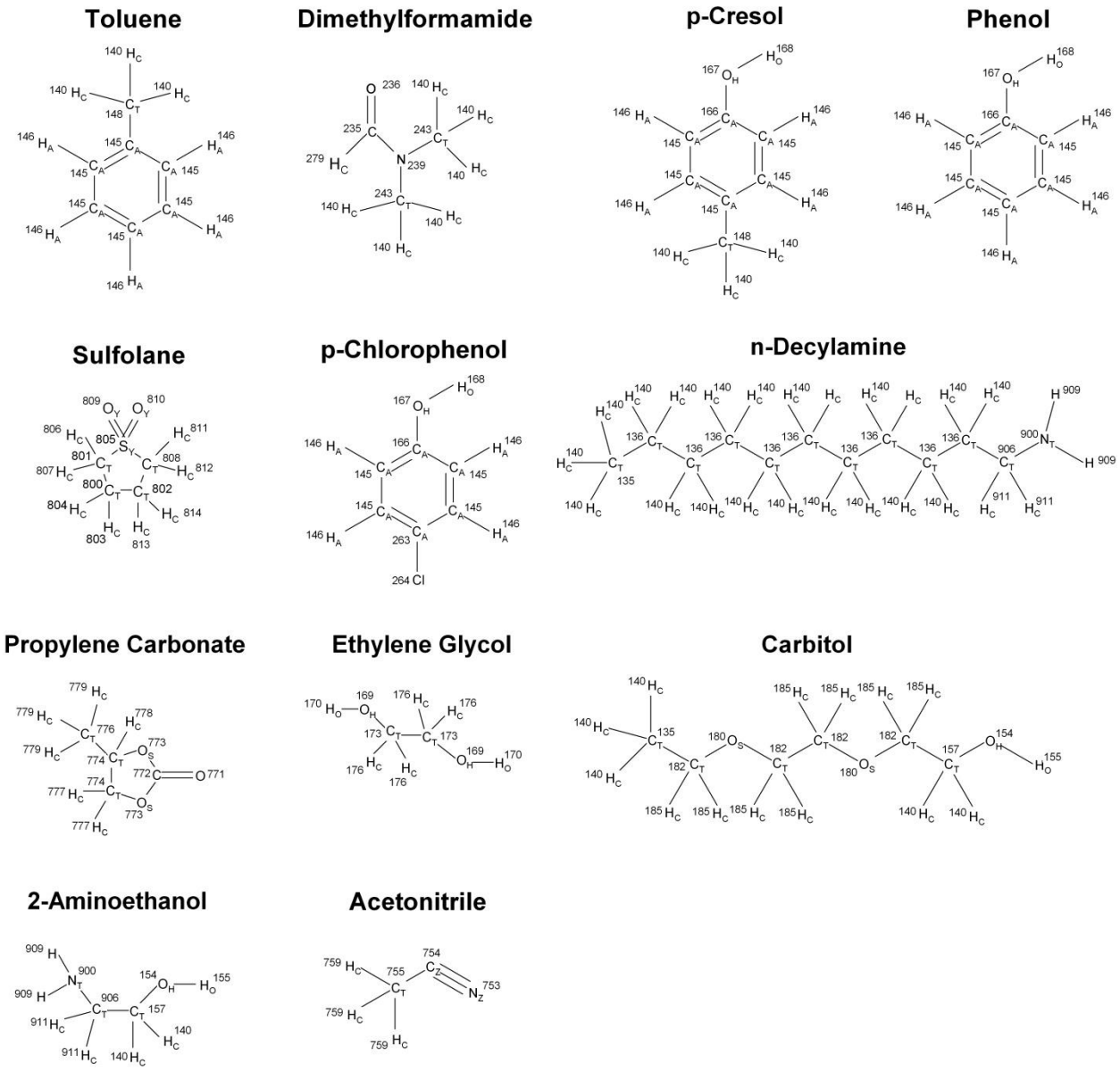



**Figure S1. OPLS-AA atom types for solvents used in simulations.**

**Table S1.** Density (g/cm$^3$) and Heat of Vaporization (kcal/mol) for All-Atom OPLS Solvent Models at 25 °C and 1 atm.

| | ρ(calc) | ρ(exptl) | ref. | $\Delta H_v$(calc) | $\Delta H_v$(exptl) | ref. |
|---|---|---|---|---|---|---|
| *p*-chlorophenol | 1.25 | 1.2727[a] | 1 | 15.5 | 15.4 | 2 |
| phenol | 1.06 | 1.058[b] | 3 | 14.8 | 13.82 | 3 |
| *p*-cresol | 1.02 | 1.0341[c] | 4 | 16.0 | 14.8[d] | 5 |
| ethylene glycol | 1.09 | 1.10998 | 6 | 16.7 | 15.7 | 7 |
| 2-aminoethanol | 0.95 | 1.0118 | 8 | 12.3 | 12.1[e] | 5 |
| carbitol | 0.94 | 0.983082 | 9 | 14.5 | 12.5[f] | 5 |
| sulfolane | 1.22 | 1.2614[g] | 10 | 17.7 | 17.5[c] | 11 |
| adiponitrile | 0.94 | 0.9585 | 12 | 17.2 | 14.0[i] | 5 |
| propylene carbonate | 1.24 | 1.1978 | 13 | 16.7 | 17.05 | 14 |
| *n*-decylamine | 0.75 | 0.7933[c] | 15 | 14.0 | 12.5[j] | 5 |
| DMF | 0.91 | 0.9445 | 16 | 10.8 | 11.2 | 5 |
| toluene | 0.86 | 0.86675[c] | 17 | 9.2 | 9.1 | 5 |

[a]28 °C. [b]Based on values at 20 and 40 °C. [c]20 °C. [d]50 °C. [e]14 °C. [f]60 °C. [g]30 °C. [h]0 °C. [i]90 °C. [j]152 °C.


References:

(1) Badachhape, R. B.; Gharpurey, M. K.; Biswas, A. B. *J. Chem. Eng. Data* **1965**, *10*, 143-145.

(2) Verevkin, S. P.; Emel'yanenko, V. N.; Klamt, A. *J. Chem. Eng. Data* **2007**, *52*, 499-510.

(3) Riddick, J. A.; Bunger, W. B.; Sakano, T. K. *Techniques of Chemistry, Vol. II: Organic Solvents, Physical Properties and Methods of Purification*; 4th ed.; Ed.; Wiley: New York, 1986.

(4) Schoorl, N. *Recueil des Travaux Chimiques des Pays-Bas et de la Belgique* **1929**, *48*, 935.

(5) Chickos, J. S.; William E. Acree, J. *J. Phys. Chem. Ref. Data* **2003**, *32*, 519-878.

(6) Kinart, C. M.; Klimczak, M.; Kinart, W. J. *J. Mol. Liq.* **2009**, *145*, 8-13.

(7) Verevkin, S. P. *Fluid Phase Equilib.* **2004**, *224*, 23-29.

(8) Gubskaya, A. V.; Kusalik, P. G. *J. Phys. Chem. A* **2004**, *108*, 7151-7164.

(9) El-Hefnawy, M.; Sameshima, K.; Matsushita, T.; Tanaka, R. *Bull. Chem. Soc. Jpn.* **2006**, *79*, 845-856.

(10) Awwad, A. M.; Al-Dujaili, A. H.; Salman, H. E. *J. Chem. Eng. Data* **2002**, *47*, 421-424.

(11) Walker, E. E. *J. Appl. Chem.* **1952**, *2*, 470-481.

(12) Sears, P. G.; Caruso, J. A.; Popov, A. I. *J. Phys. Chem.* **1967**, *71*, 905-909.

(13) Moumouzias, G.; Ritzoulis, G. *J. Chem. Eng. Data* **1997**, *42*, 710-713.

(14) Nasirzadeh, K.; Neueder, R.; Kunz, W. *J. Chem. Eng. Data* **2005**, *50*, 26-28.

(15) Kagan, Y. B.; Bashkirov, A. N.; Kliger, G. A.; Chou, C.-T.; Mak, N. E. *Neftekhimiya* **1961**, *1*, 555-563.
(16) Nikam, P. S.; Kharat, S. J. *J. Chem. Eng. Data* **2005**, *50*, 455-459.
(17) Silva, A. A.; Reis, R. A.; Paredes, M. L. L. *J. Chem. Eng. Data* **2009**, *54*, 2067-2072.

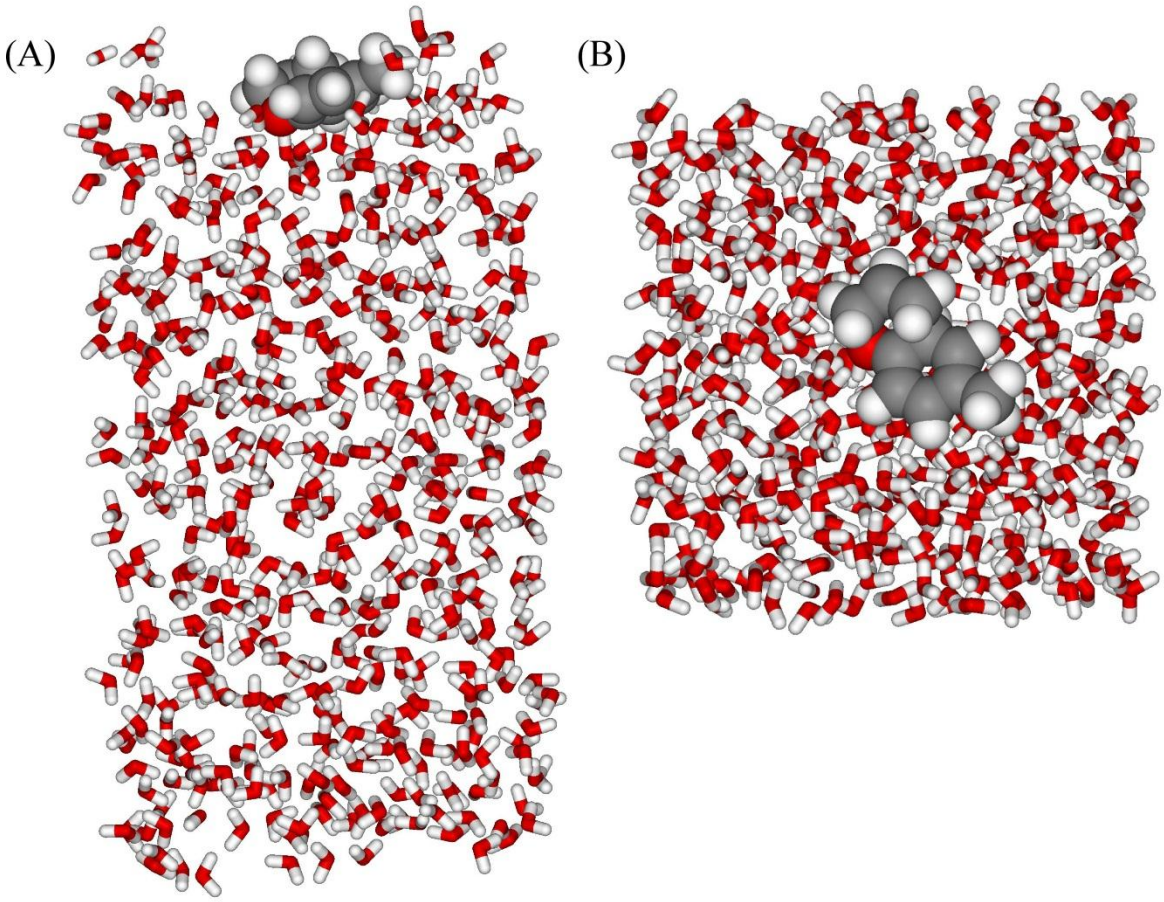



**Figure S2. (A) Side and (B) overhead illustration of the "on water" allyl *p*-tolyl ether transition structure from the QM/MM/MC Claisen rearrangement calculations.**

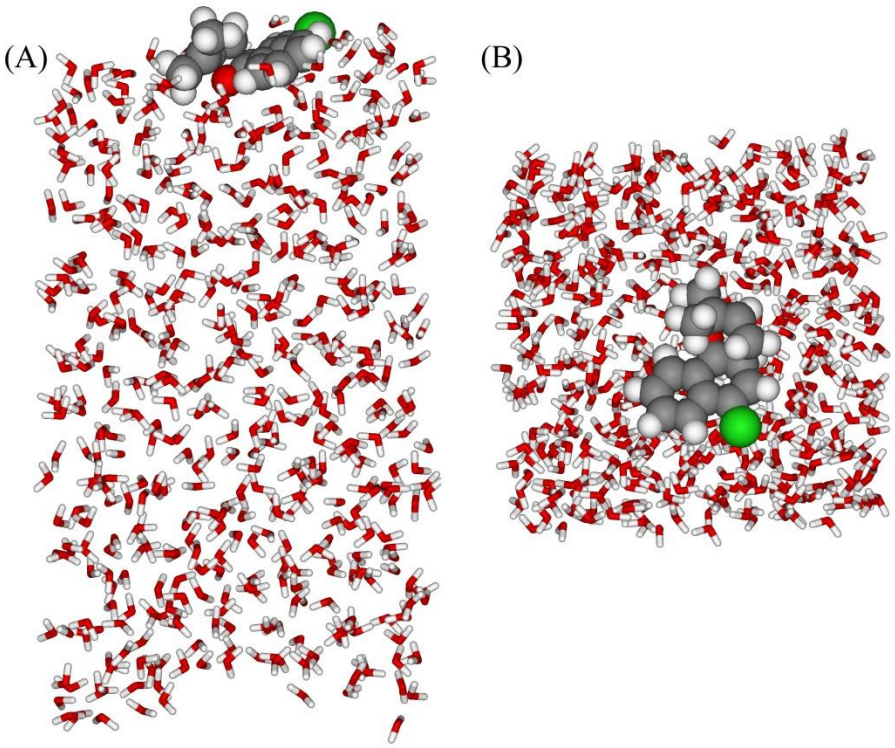



**Figure S3. (A) Side and (B) overhead illustration of the "on water" allyl naphthyl ether transition structure from the QM/MM/MC Claisen rearrangement calculations.**

**Table S2.** Computed Bond Lengths (Å) for Gas-Phase Allyl *p*-R-Phenyl Ether Transition Structures (R = $OCH_3$, $CH_3$, and Br)[a].

| | $OCH_3$ | $CH_3$ | Br |
|---|---|---|---|
| CBS-QB3 | 2.22, 2.25 | 2.16, 2.25 | 2.19, 2.25 |
| B3LYP[b] | 2.23, 2.25 | 2.17, 2.25 | 2.20, 2.25 |
| PDDG/PM3 | 1.86, 1.98 | 1.85, 1.98 | 1.86, 1.99 |

[a]$R_{OC}$ and $R_{CC}$ values, respectively. [b]6-311+G(2d,p).

**Table S3.** Computed Bond Lengths (Å) for Solution-Phase Allyl *p*-R-Phenyl Ether Transition Structures (R = $OCH_3$, $CH_3$, and Br)[a].

| | $OCH_3$ | $CH_3$ | Br |
|---|---|---|---|
| Water | 2.28, 2.31 | 2.21, 2.32 | 2.25, 2.31 |
| n-Pentadecane | 2.24, 2.27 | 2.19, 2.27 | 2.21, 2.27 |

[a]$R_{OC}$ and $R_{CC}$ values, respectively, from B3LYP/6-311+G(2d,p).

**Table S4.** Computed Bond Lengths (Å) for Solution-Phase Allyl Naphthyl Ether Transition Structures[a.]

| | DFT/PCM[b] | QM/MM[c] |
|---|---|---|
| "In Water" | 2.36, 2.52 | 1.89, 2.01 |
| "On Water" | - | 1.90, 2.00 |
| MeOH | 2.36, 2.51 | 1.87, 1.99 |
| $CH_3CN$ | 2.36, 2.51 | 1.86, 2.00 |
| DMF | 2.36, 2.51 | 1.84, 1.95 |
| Toluene | 2.31, 2.47 | 1.83, 1.97 |

[a]$R_{OC}$ and $R_{CC}$ values, respectively. [b]Optimized using B3LYP/6-311+G(2d,p). [c]PDDG/PM3 and MC/FEP.

**Table S5.** Charges for the Claisen Rearrangement of Allyl p-Tolyl Ether[a].

| | O1 | C2 | C4 | C5 |
|---|---|---|---|---|
| *p*-chlorophenol | -0.377 (-0.278) | -0.088 (-0.063) | -0.174 (-0.350) | -0.175 (-0.194) |
| "in water" | -0.385 (-0.267) | -0.087 (-0.027) | -0.178 (-0.346) | -0.222 (-0.193) |
| phenol | -0.385 (-0.280) | -0.093 (-0.068) | -0.211 (-0.342) | -0.201 (-0.214) |
| *p*-cresol | -0.386 (-0.269) | -0.108 (-0.048) | -0.194 (-0.343) | -0.136 (-0.273) |
| ethylene glycol | -0.384 (-0.272) | -0.108 (-0.049) | -0.169 (-0.336) | -0.200 (- |
| "on water" | -0.383 (-0.255) | -0.064 (-0.058) | -0.165 (-0.341) | -0.248 (-0.199) |
| 2-aminoethanol | -0.383 (-0.283) | -0.114 (-0.063) | -0.186 (-0.351) | -0.198 (-0.219) |
| carbitol | -0.401 (-0.261) | -0.072 (-0.042) | -0.173 (-0.335) | -0.190 (-0.195) |
| sulfolane | -0.379 (-0.280) | -0.100 (-0.035) | -0.205 (-0.340) | -0.163 (-0.218) |
| adiponitrile | -0.412 (-0.278) | -0.104 (-0.045) | -0.191 (-0.327) | -0.194 (-0.185) |
| propylene | -0.384 (-0.263) | -0.113 (-0.038) | -0.193 (-0.350) | -0.212 (-0.199) |
| *n*-decylamine | -0.375 (-0.233) | -0.053 (-0.061) | -0.147 (-0.331) | -0.208 (-0.211) |

[a]CM3 charges scaled by 1.14 from QM/MM/MC simulations. Atom numbering scheme given in Figure 1.

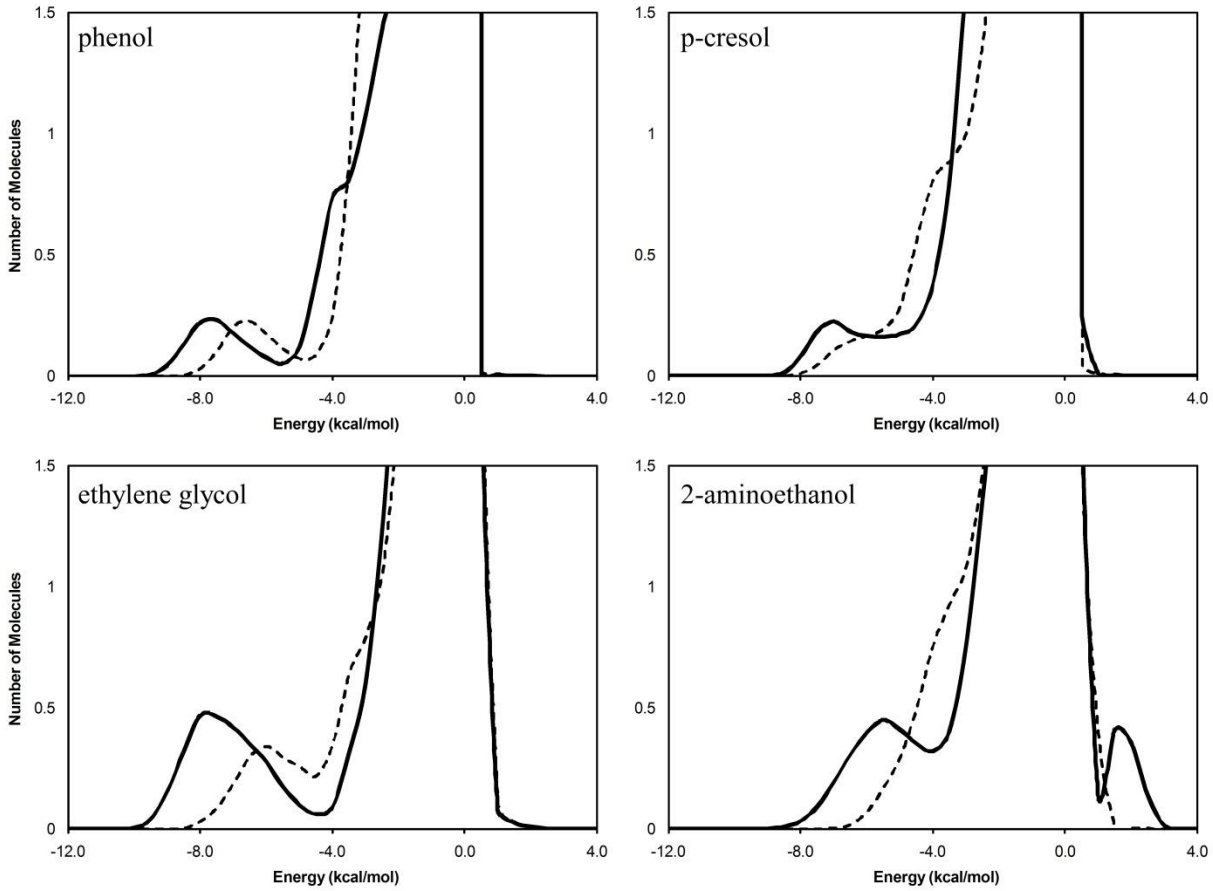



**Figure S4. Solute-solvent energy pair distributions for the Claisen rearrangement of allyl *p*-tolyl ether in phenol, p-cresol, ethylene glycol, and 2-aminoethanol, transition state (solid black) and reactant (dashed black) at 25 °C. The ordinate records the number of solvent molecules that interact with the solutes and their interaction energy on the abscissa. Units for ordinate are number of molecules per kcal/mol.**

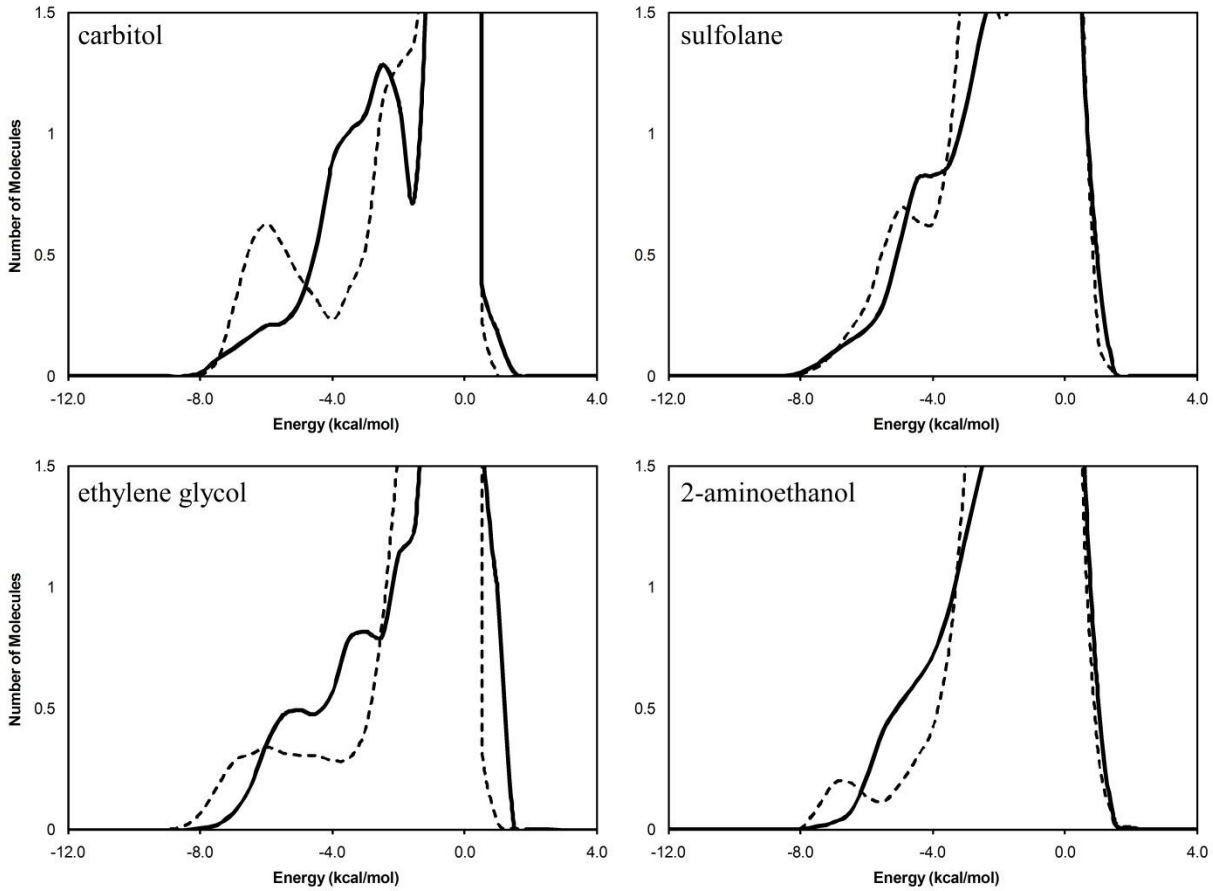



**Figure S5. Solute-solvent energy pair distributions for the Claisen rearrangement of allyl *p*-tolyl ether in carbitol, sulfolane, ethylene glycol, and 2-aminoethanol, transition state (solid black) and reactant (dashed black) at 25 °C.**

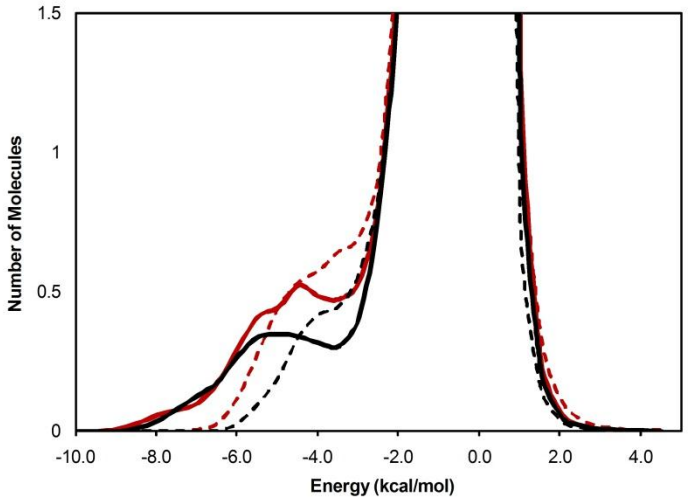



**Figure S6. Solute-solvent energy pair distributions for the Claisen rearrangement of allyl naphthyl ether: "in water" transition state (solid red), "in water" reactant (dashed red), "on water" transition state (solid black), and "on water" reactant (dashed black) at 25 °C.**

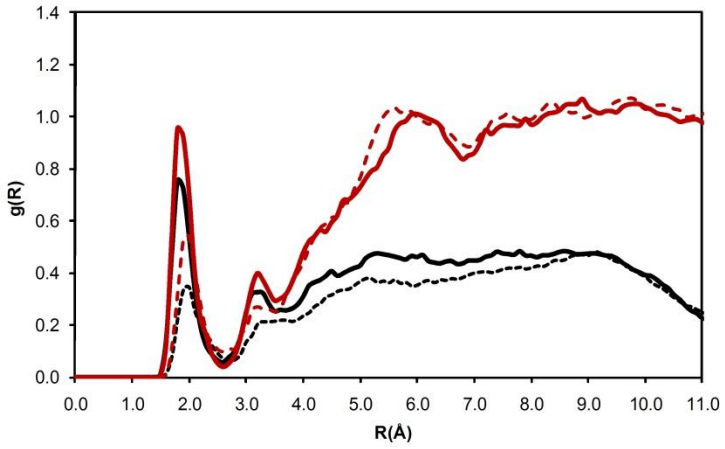



**Figure S7. Computed O(ether)–H(water) radial distribution function for the Claisen rearrangement reactions of allyl naphthyl ether: "in water" transition state (solid red), "in water" reactant (dashed red), "on water" transition state (solid black), and "on water" reactant (dashed black) at 25 °C.**

**<u>Condensed-Phase</u> B3LYP/6-311+G(2d,p) allyl *p*-tolyl ether optimizations in PCM:**

**Water transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -321.7564 | 85.6775 | 112.0578 |
| Red. masses -- | 7.5323 | 2.9970 | 1.2348 |
| Frc consts -- | 0.4594 | 0.0130 | 0.0091 |
| IR Inten -- | 92.6723 | 1.3783 | 0.4509 |

Sum of electronic and zero-point Energies= -463.384638
Sum of electronic and thermal Energies= -463.374321
Sum of electronic and thermal Enthalpies= -463.373376
Sum of electronic and thermal Free Energies= -463.420110


1\1\ASN_X86_64-DMC14\FTS\RB3LYP\6-311+G(2d,p)\C10H12O1\AUBOXA\16-Aug-2009\1\\# B3LYP/6-311+G(2d,p) opt=(TS,noeigen,z-matrix) scrf=(pcm,solvent=water)\\Claisen rearrangement allyl p-tolyl ether GS\\0,1\C\O,1,r21\C,2,r32,1,a321\C,3,r43,2,a432,1,d4321,0\C,4,r54,3,a543,2,d5432,0\C,5,r65,4,a654,3,d6543,0\C,6,r76,1,a761,2,d7612,0\C,1,r81,2,a812,3,d8123,0\C,7,r97,6,a976,1,d9761,0\C,8,r108,1,a1081,2,d10812,0\C,9,r119,7,a1197,6,d11976,0\H,3,r123,2,a1232,1,d12321,0\H,3,r133,2,a1332,1,d13321,0\H,4,r144,3,a1443,2,d14432,0\H,5,r155,4,a1554,3,d15543,0\H,5,r165,4,a1654,3,d16543,0\H,6,r176,1,a1761,2,d17612,0\H,7,r187,6,a1876,1,d18761,0\H,10,r1910,8,a19108,1,d191081,0\H,8,r208,1,a2081,2,d20812,0\H,11,r2111,9,a21119,7,d211197,0\H,11,r2211,9,a22119,7,d221197,0\H,11,r2311,9,a23119,7,d231197,0\\r21=1.28089795\r32=2.21516118\r43=1.37497845\r54=1.3878699\r65=2.31680892\r76=1.4177673\r81=1.43456495\r97=1.37447343\r108=1.36850621\r119=1.50598617\r123=1.08064661\r133=1.08154631\r144=1.0851403\r155=1.08136754\r165=1.08215151\r176=1.08143224\r187=1.08484741\r1910=1.08565788\r208=1.08419801\r2111=1.09390465\r2211=1.0940131\r2311=1.0906175\a321=104.10650681\a432=96.89241933\a543=121.12439754\a654=100.00544624\a761=120.48627805\a812=121.89669344\a976=121.46636696\a1081=121.09780506\a1197=122.20515721\a1232=86.05117539\a1332=98.28931111\a1443=118.75420108\a1554=120.57136027\a1654=120.55216764\a1761=117.9087508\a1876=118.68937849\a19108=119.32289081\a2081=117.92329961\a21119=111.2151083\a22119=111.12784084\a23119=111.34246848\d4321=53.1502216\d5432=-69.62713742\d6543=77.10532108\d7612=183.99632794\d8123=96.83908025\d9761=-6.61336399\d10812=-181.67155937\d11976=-178.81175184\d12321=-67.65550896\d13321=176.09243941\d14432=96.13108968\d15543=-182.84181414\d16543=-22.37637567\d17612=-13.27528603\d18761=174.56492534\d191081=-177.71041069\d20812=0.86416043\d211197=-120.26605712\d221197=-239.20768346\d231197=0.32100882\\Version=AM64L-G09RevA.01\State=1-A\HF=-463.574778\RMSD=4.317e-09\RMSF=2.819e-05\Dipole=1.1605909,0.31539,-1.2842923\Quadrupole=-3.4647315,4.0064832,-0.5417517,1.2365072,4.9599968,1.01899\PG=C01 [X(C10H12O1)]\\@


**Water ground state**


1\1\ ASN_ALTIX-ALTIX8\Freq\RB3LYP\6-311+G(2d,p)\C10H12O1\AUBOXA\27-Aug

-2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcall,noeigen,z-matrix) scrf=(pc
m,solvent=water)\\Claisen rearrangement allyl p-tolyl ether GS\\0,1\C\
O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B5,2,A4,3
,D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0\C,9,B9,
7,A8,6,D7,0\C,9,B10,7,A9,6,D8,0\H,3,B11,2,A10,1,D9,0\H,3,B12,2,A11,1,D
10,0\H,4,B13,3,A12,2,D11,0\H,5,B14,4,A13,3,D12,0\H,5,B15,4,A14,3,D13,0
\H,6,B16,1,A15,2,D14,0\H,7,B17,6,A16,1,D15,0\H,10,B18,9,A17,7,D16,0\H,
8,B19,1,A18,2,D17,0\H,11,B20,9,A19,7,D18,0\H,11,B21,9,A20,7,D19,0\H,11
,B22,9,A21,7,D20,0\\B1=1.371\B2=1.4368\B3=1.5006\B4=1.3274\B5=1.3988\B
6=1.3855\B7=1.3925\B8=1.4009\B9=1.3918\B10=1.5088\B11=1.093\B12=1.092\
B13=1.0871\B14=1.0835\B15=1.0855\B16=1.0835\B17=1.085\B18=1.0848\B19=1
.08\B20=1.0935\B21=1.0938\B22=1.091\A1=119.7894\A2=112.847\A3=124.5156
\A4=115.522\A5=120.3113\A6=125.2643\A7=121.501\A8=117.3274\A9=120.9707
\A10=110.8462\A11=103.4632\A12=115.2385\A13=121.4453\A14=121.7236\A15=
118.8718\A16=118.9448\A17=119.4414\A18=121.2719\A19=111.3953\A20=111.3
607\A21=111.2053\D1=76.8241\D2=-136.85\D3=-175.8308\D4=-178.4252\D5=5.
4341\D6=-0.1162\D7=-0.1162\D8=-180.149\D9=-49.3725\D10=195.3896\D11=45
.7108\D12=183.283\D13=2.7803\D14=1.3193\D15=179.8914\D16=180.2447\D17=
-2.1192\D18=57.4804\D19=-61.9405\D20=-182.173\\Version=IA64L-G09RevA.0
2\State=1-A\HF=-463.6256898\RMSD=4.761e-09

**Pentadecane transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -349.0637 | 87.5641 | 116.1994 |
| Red. masses -- | 8.0612 | 3.0727 | 1.3361 |
| Frc consts -- | 0.5787 | 0.0139 | 0.0106 |
| IR Inten -- | 61.3166 | 0.8144 | 0.3590 |

Sum of electronic and zero-point Energies= -463.378933
Sum of electronic and thermal Energies= -463.368684
Sum of electronic and thermal Enthalpies= -463.367739
Sum of electronic and thermal Free Energies= -463.414272

1\1\ASN_X86_64-DMC14\FTS\RB3LYP\6-311+G(2d,p)\C10H12O1\AUBOXA\16-Aug-2
009\1\\# B3LYP/6-311+G(2d,p) opt=(TS,noeigen,z-matrix) scrf=(pcm,solve
nt=n-Pentadecane)\\Claisen rearrangement allyl p-tolyl ether\\0,1\C\O,
1,r21\C,2,r32,1,a321\C,3,r43,2,a432,1,d4321,0\C,4,r54,3,a543,2,d5432,0
\C,5,r65,4,a654,3,d6543,0\C,6,r76,1,a761,2,d7612,0\C,1,r81,2,a812,3,d8
123,0\C,7,r97,6,a976,1,d9761,0\C,8,r108,1,a1081,2,d10812,0\C,9,r119,7,
a1197,6,d11976,0\H,3,r123,2,a1232,1,d12321,0\H,3,r133,2,a1332,1,d13321
,0\H,4,r144,3,a1443,2,d14432,0\H,5,r155,4,a1554,3,d15543,0\H,5,r165,4,
a1654,3,d16543,0\H,6,r176,1,a1761,2,d17612,0\H,7,r187,6,a1876,1,d18761
,0\H,10,r1910,8,a19108,1,d191081,0\H,8,r208,1,a2081,2,d20812,0\H,11,r2
111,9,a21119,7,d211197,0\H,11,r2211,9,a22119,7,d221197,0\H,11,r2311,9,
a23119,7,d231197,0\\r21=1.27510012\r32=2.1849625\r43=1.37448374\r54=1.
3900044\r65=2.27405653\r76=1.42025411\r81=1.43651867\r97=1.37247144\r1
08=1.3668305\r119=1.50603932\r123=1.08061407\r133=1.08120664\r144=1.08
531856\r155=1.08166047\r165=1.08230189\r176=1.08132649\r187=1.08519461
\r1910=1.08589534\r208=1.08393058\r2111=1.09419974\r2211=1.09077295\r2
311=1.09400825\a321=104.46282028\a432=97.21598571\a543=121.17875967\a6
54=100.51542072\a761=120.34255127\a812=121.99203064\a976=121.48165522\

a1081=121.1023816\a1197=122.22667802\a1232=86.3798784\a1332=97.9161453
1\a1443=118.52704575\a1554=120.40683695\a1654=120.45625686\a1761=117.4
3774837\a1876=118.75099381\a19108=119.35715882\a2081=117.66906674\a211
19=111.23541824\a22119=111.37343092\a23119=111.23355119\d4321=53.21366
678\d5432=-68.7205623\d6543=76.14253067\d7612=184.11684731\d8123=97.11
336859\d9761=-7.31472908\d10812=-181.4242875\d11976=-178.88235784\d123
21=-67.71050271\d13321=176.18876133\d14432=95.75075019\d15543=-182.428
4543\d16543=-24.25332639\d17612=-15.3566327\d18761=173.67494721\d19108
1=-177.45389138\d20812=1.5504231\d211197=-238.79911775\d221197=-359.25
117101\d231197=-119.70789259\\Version=AM64L-G09RevA.01\State=1-A\HF=-4
63.5692848\RMSD=9.508e-09\RMSF=2.418e-05\Dipole=0.8280602,0.2267685,-1
.0240892\Quadrupole=-3.4534875,3.681177,-0.2276894,1.0325072,4.0493163
,0.8169742\PG=C01 [X(C10H12O1)]\\@

**Pentadecane ground state**

1\1\ ASN_ALTIX-ALTIX7\FOpt\RB3LYP\6-311+G(2d,p)\C10H12O1\AUBOXA\16-Aug
-2009\1\\# B3LYP/6-311+G(2d,p) opt=(noeigen,z-matrix) scrf=(pcm,solven
t=n-Pentadecane)\\Claisen rearrangement allyl p-tolyl ether GS\\0,1\C\
O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B5,2,A4,3
,D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0\C,9,B9,
7,A8,6,D7,0\C,9,B10,7,A9,6,D8,0\H,3,B11,2,A10,1,D9,0\H,3,B12,2,A11,1,D
10,0\H,4,B13,3,A12,2,D11,0\H,5,B14,4,A13,3,D12,0\H,5,B15,4,A14,3,D13,0
\H,6,B16,1,A15,2,D14,0\H,7,B17,6,A16,1,D15,0\H,10,B18,9,A17,7,D16,0\H,
8,B19,1,A18,2,D17,0\H,11,B20,9,A19,7,D18,0\H,11,B21,9,A20,7,D19,0\H,11
,B22,9,A21,7,D20,0\\B1=1.37047354\B2=1.43145604\B3=1.50138929\B4=1.326
67608\B5=1.39812148\B6=1.38441642\B7=1.39182048\B8=1.40048574\B9=1.390
72709\B10=1.50857657\B11=1.09411898\B12=1.09258861\B13=1.08686875\B14=
1.08322973\B15=1.08563905\B16=1.08326445\B17=1.08513893\B18=1.08488607
\B19=1.07992219\B20=1.09369011\B21=1.09383002\B22=1.09111196\A1=119.79
062036\A2=113.14356967\A3=124.5586415\A4=115.47206184\A5=120.26314739\
A6=125.28403526\A7=121.52652554\A8=117.34313027\A9=120.94583482\A10=11
0.8583157\A11=103.54830245\A12=115.13402758\A13=121.54873547\A14=121.7
1388787\A15=118.71865389\A16=118.94992073\A17=119.420329\A18=121.18831
468\A19=111.43686281\A20=111.41759378\A21=111.2262615\D1=78.34008315\D
2=-139.04781726\D3=-180.52384166\D4=-178.80601971\D5=0.33582095\D6=-0.
04479166\D7=-0.18772344\D8=-180.10574659\D9=-47.85555738\D10=196.90181
76\D11=43.45985804\D12=183.2887134\D13=2.90627096\D14=0.89355114\D15=1
79.85537141\D16=180.36436396\D17=-2.25357685\D18=58.83478265\D19=-60.7
5786957\D20=-180.94390129\\Version=IA64L-G09RevA.01\State=1-A\HF=-463.
6223832\RMSD=4.745e-09\RMSF=3.591e-05\Dipole=0.5180847,0.0487195,0.043
9895\Quadrupole=1.8907515,-3.4456261,1.5548747,-0.1535922,2.4692664,0.
5227612\PG=C01 [X(C10H12O1)]\\@

**<u>Gas-Phase</u> B3LYP/6-311+G(2d,p) allyl *p*-tolyl ether optimizations:**

**Transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -362.6151 | 88.1706 | 118.2188 |
| Red. masses -- | 8.3002 | 3.1345 | 1.3135 |
| Frc consts -- | 0.6430 | 0.0144 | 0.0108 |

IR Inten -- 45.9409 0.6198
0.2867

Sum of electronic and zero-point Energies= -463.375767
Sum of electronic and thermal Energies= -463.365546
Sum of electronic and thermal Enthalpies= -463.364602
Sum of electronic and thermal Free Energies= -463.411052


1\1\ASN_X86_64-DMC14\FTS\RB3LYP\6-311+G(2d,p)\C10H12O1\AUBOXA\27-Aug-2
009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,TS,noeigen,z-matrix)\\Claisen
rearrangement allyl p-tolyl ether GS\\0,1\C\O,1,r21\C,2,r32,1,a321\C,
3,r43,2,a432,1,d4321,0\C,4,r54,3,a543,2,d5432,0\C,5,r65,4,a654,3,d6543
,0\C,6,r76,1,a761,2,d7612,0\C,1,r81,2,a812,3,d8123,0\C,7,r97,6,a976,1,
d9761,0\C,8,r108,1,a1081,2,d10812,0\C,9,r119,7,a1197,6,d11976,0\H,3,r1
23,2,a1232,1,d12321,0\H,3,r133,2,a1332,1,d13321,0\H,4,r144,3,a1443,2,d
14432,0\H,5,r155,4,a1554,3,d15543,0\H,5,r165,4,a1654,3,d16543,0\H,6,r1
76,1,a1761,2,d17612,0\H,7,r187,6,a1876,1,d18761,0\H,10,r1910,8,a19108,
1,d191081,0\H,8,r208,1,a2081,2,d20812,0\H,11,r2111,9,a21119,7,d211197,
0\H,11,r2211,9,a22119,7,d221197,0\H,11,r2311,9,a23119,7,d231197,0\\r21
=1.27206324\r32=2.16967867\r43=1.37416685\r54=1.39108217\r65=2.2535564
4\r76=1.42148316\r81=1.43752927\r97=1.37148591\r108=1.36596237\r119=1.
50604385\r123=1.08057078\r133=1.08103946\r144=1.08541397\r155=1.081822
27\r165=1.08238253\r176=1.08132556\r187=1.08535643\r1910=1.08601454\r2
08=1.08385325\r2111=1.09429413\r2211=1.09085653\r2311=1.09409039\a321=
104.66943503\a432=97.40676846\a543=121.20277539\a654=100.76636499\a761
=120.26633928\a812=122.04526395\a976=121.48233148\a1081=121.1062802\a1
197=122.23022873\a1232=86.54437984\a1332=97.70516781\a1443=118.4242837
8\a1554=120.30701764\a1654=120.39498445\a1761=117.18265605\a1876=118.7
8227534\a19108=119.37942847\a2081=117.53439606\a21119=111.27710617\a22
119=111.38818908\a23119=111.26362198\d4321=53.22110035\d5432=-68.23466
103\d6543=75.62892032\d7612=184.18617446\d8123=97.28569116\d9761=-7.68
00297\d10812=-181.2990651\d11976=-178.81274895\d12321=-67.74479763\d13
321=176.21690106\d14432=95.60628272\d15543=-182.23258445\d16543=-25.20
467891\d17612=-16.38192044\d18761=173.30344928\d191081=-177.39863555\d
20812=1.81636774\d211197=-238.8390999\d221197=-359.2586607\d231197=-11
9.68166293\\Version=AM64L-G09RevA.02\State=1-A\HF=-463.5662132\RMSD=4.
756e-09\RMSF=1.761e-05\Dipole=0.6936024,0.1836034,-0.8858663\Quadrupol
e=-3.4031018,3.5149056,-0.1118039,0.9322978,3.564965,0.7100033\PG=C01
[X(C10H12O1)]\\@


**Ground state**


1\1\ASN_X86_64-DMC35\FOpt\RB3LYP\6-311+G(2d,p)\C10H12O1\AUBOXA\27-Aug-
2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,noeigen,z-matrix)\\Claisen r
earrangement allyl p-tolyl ether GS\\0,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2
,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B5,2,A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C,
1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0\C,9,B9,7,A8,6,D7,0\C,9,B10,7,A9,6,
D8,0\H,3,B11,2,A10,1,D9,0\H,3,B12,2,A11,1,D10,0\H,4,B13,3,A12,2,D11,0\
H,5,B14,4,A13,3,D12,0\H,5,B15,4,A14,3,D13,0\H,6,B16,1,A15,2,D14,0\H,7,
B17,6,A16,1,D15,0\H,10,B18,9,A17,7,D16,0\H,8,B19,1,A18,2,D17,0\H,11,B2
0,9,A19,7,D18,0\H,11,B21,9,A20,7,D19,0\H,11,B22,9,A21,7,D20,0\\B1=1.37
051125\B2=1.42843381\B3=1.50181118\B4=1.32634768\B5=1.39767864\B6=1.38
38879\B7=1.3914508\B8=1.4002385\B9=1.39020414\B10=1.50842615\B11=1.094
71253\B12=1.09291242\B13=1.08680739\B14=1.08312054\B15=1.08567947\B16=
1.08308789\B17=1.0851777\B18=1.08489503\B19=1.07993442\B20=1.09374272\
B21=1.09390513\B22=1.09118573\A1=119.78644584\A2=113.2639436\A3=124.61

163847\A4=115.43777168\A5=120.23512682\A6=125.29160034\A7=121.53597003
\A8=117.35696725\A9=120.9340276\A10=110.90985681\A11=103.59249741\A12=
115.05944918\A13=121.58978958\A14=121.72216036\A15=118.65005248\A16=11
8.95411646\A17=119.41934156\A18=121.17379866\A19=111.46461397\A20=111.
44322283\A21=111.23948807\D1=78.7808566\D2=-140.23558898\D3=-181.85789
942\D4=-178.9154162\D5=-1.12990552\D6=-0.0260852\D7=-0.21823306\D8=-18
0.14079847\D9=-47.41847038\D10=197.32681672\D11=42.24069467\D12=183.30
826441\D13=2.92648354\D14=0.75863744\D15=179.84596341\D16=180.41066008
\D17=-2.32820254\D18=58.731158\D19=-60.92617622\D20=-181.07714707\\Ver
sion=AM64L-G09RevA.02\State=1-A\HF=-463.6205405\RMSD=9.325e-09\RMSF=3.
223e-05\Dipole=0.4542391,0.0487756,0.0530927\Quadrupole=1.8227869,-3.1
713808,1.3485939,-0.0285141,2.2203557,0.4919149\PG=C01 [X(C10H12O1)]\\

**<u>Gas-Phase</u> PDDG-PM3 allyl *p*-tolyl ether optimizations:**

**Transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -953.9422 | 50.2400 | 96.4726 |
| Red. masses -- | 10.4383 | 1.0436 | 3.1546 |
| Frc consts -- | 5.5966 | 0.0016 | 0.0173 |
| IR Inten -- | 107.0059 | 0.2187 | 0.3238 |

| | |
|---|---|
| Sum of electronic and zero-point Energies= | 0.253188 |
| Sum of electronic and thermal Energies= | 0.263208 |
| Sum of electronic and thermal Enthalpies= | 0.264152 |
| Sum of electronic and thermal Free Energies= | 0.217604 |

1\1\ ASN_ALTIX-ALTIX9\FTS\RPDDG\ZDO\C10H12O1\AUBOXA\10-Aug-2009\1\\# P
DDG opt=(TS,noeigen,z-matrix)\\Claisen rearrangement allyl p-tolyl eth
er GS\\0,1\C\O,1,r21\C,2,r32,1,a321\C,3,r43,2,a432,1,d4321,0\C,4,r54,3
,a543,2,d5432,0\C,5,r65,4,a654,3,d6543,0\C,6,r76,1,a761,2,d7612,0\C,1,
r81,2,a812,3,d8123,0\C,7,r97,6,a976,1,d9761,0\C,8,r108,1,a1081,2,d1081
2,0\C,9,r119,7,a1197,6,d11976,0\H,3,r123,2,a1232,1,d12321,0\H,3,r133,2
,a1332,1,d13321,0\H,4,r144,3,a1443,2,d14432,0\H,5,r155,4,a1554,3,d1554
3,0\H,5,r165,4,a1654,3,d16543,0\H,6,r176,1,a1761,2,d17612,0\H,7,r187,6
,a1876,1,d18761,0\H,10,r1910,8,a19108,1,d191081,0\H,8,r208,1,a2081,2,d
20812,0\H,11,r2111,9,a21119,7,d211197,0\H,11,r2211,9,a22119,7,d221197,
0\H,11,r2311,9,a23119,7,d231197,0\\r21=1.2803281\r32=1.84925069\r43=1.
39953088\r54=1.40398369\r65=1.98162851\r76=1.44203004\r81=1.44896765\r
97=1.35977214\r108=1.35715968\r119=1.48536115\r123=1.10451951\r133=1.1
0029349\r144=1.09959613\r155=1.09519907\r165=1.09883723\r176=1.1107718
2\r187=1.10176063\r1910=1.10186014\r208=1.10154318\r2111=1.0980201\r22
11=1.09771325\r2311=1.09741139\a321=107.52122593\a432=99.56186481\a543
=118.89926928\a654=99.42349817\a761=117.68015398\a812=120.58649573\a97
6=121.01409757\a1081=120.25188952\a1197=121.41105116\a1232=100.4546289
5\a1332=102.05455286\a1443=119.92164447\a1554=118.30997525\a1654=119.6
2681816\a1761=118.31974551\a1876=118.04102606\a19108=120.64872943\a208
1=118.98020641\a21119=112.95679126\a22119=111.5250245\a23119=110.82697
484\d4321=56.19630062\d5432=-69.7555399\d6543=73.186764\d7612=178.9524
2208\d8123=100.50079079\d9761=-6.78374392\d10812=-176.57239368\d11976=
-177.67903547\d12321=-66.24889498\d13321=177.82084071\d14432=95.913986

25\d15543=-178.60296046\d16543=-34.35254671\d17612=-34.95245762\d18761
=174.53435897\d191081=-177.67828583\d20812=5.72435881\d211197=-371.234
68977\d221197=-492.17616969\d231197=-251.0395933\\Version=IA64L-G09Rev
A.01\State=1-A\HF=0.0658361\RMSD=7.638e-09\RMSF=2.697e-06\Dipole=0.589
1268,0.0006633,-0.8216877\PG=C01 [X(C10H12O1)]\\@

**Ground state**

1\1\ ASN_ALTIX-ALTIX9\FOpt\RPDDG\ZDO\C10H12O1\AUBOXA\10-Aug-2009\1\\#
PDDG opt=(noeigen,z-matrix)\\Claisen rearrangement allyl p-tolyl ether
GS\\0,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,
1,B5,2,A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D
6,0\C,9,B9,7,A8,6,D7,0\C,9,B10,7,A9,6,D8,0\H,3,B11,2,A10,1,D9,0\H,3,B1
2,2,A11,1,D10,0\H,4,B13,3,A12,2,D11,0\H,5,B14,4,A13,3,D12,0\H,5,B15,4,
A14,3,D13,0\H,6,B16,1,A15,2,D14,0\H,7,B17,6,A16,1,D15,0\H,10,B18,9,A17
,7,D16,0\H,8,B19,1,A18,2,D17,0\H,11,B20,9,A19,7,D18,0\H,11,B21,9,A20,7
,D19,0\H,11,B22,9,A21,7,D20,0\\B1=1.37529093\B2=1.41094283\B3=1.501738
41\B4=1.32994848\B5=1.40561168\B6=1.3877791\B7=1.39829829\B8=1.3959668
6\B9=1.39201535\B10=1.48536405\B11=1.12177705\B12=1.10651376\B13=1.100
45224\B14=1.08707247\B15=1.09059126\B16=1.10192414\B17=1.10147593\B18=
1.10136607\B19=1.10404309\B20=1.09779419\B21=1.09733948\B22=1.09780199
\A1=118.18776753\A2=111.90504016\A3=121.13736501\A4=113.72298187\A5=11
9.04770962\A6=125.35142265\A7=120.35462429\A8=120.196644\A9=119.798275
5\A10=114.59096774\A11=99.21950742\A12=116.92476953\A13=122.15130673\A
14=122.60667766\A15=121.42500042\A16=119.71828507\A17=120.04685574\A18
=121.82467378\A19=112.22574368\A20=110.8043627\A21=112.32473336\D1=81.
46057122\D2=-130.86288891\D3=-187.52775467\D4=-179.9461586\D5=-7.90383
441\D6=0.24405512\D7=-0.71288971\D8=-181.19951384\D9=-47.69261285\D10=
198.04672807\D11=49.07176597\D12=181.55435631\D13=2.05549967\D14=0.210
63278\D15=180.41201202\D16=180.96597558\D17=-1.31519485\D18=30.720897\
D19=-88.70284139\D20=-208.22361896\\Version=IA64L-G09RevA.01\State=1-A
\HF=-0.0152394\RMSD=5.746e-09\RMSF=5.866e-06\Dipole=0.447283,-0.033663
7,-0.0769343\PG=C01 [X(C10H12O1)]\\@

**<u>Gas-Phase</u> CBS-QB3 allyl p-tolyl ether optimizations:**

**Transition structure**

Temperature= 298.150000 Pressure=
1.000000
E(ZPE)= 0.188535 E(Thermal)=
0.198839
E(SCF)= -460.554006 DE(MP2)= -
1.867767
DE(CBS)= -0.176643 DE(MP34)= -
0.060049
DE(CCSD)= -0.076017 DE(Int)=
0.060657
DE(Empirical)= -0.084593
CBS-QB3 (0 K)= -462.569884 CBS-QB3 Energy= -
462.559579
CBS-QB3 Enthalpy= -462.558635 CBS-QB3 Free Energy= -
462.605241
1\1\ ASN_ALTIX-ALTIX7\Mixed\CBS-QB3\CBS-QB3\C10H12O1\AUBOXA\11-Aug-200
9\0\\# CBS-QB3 opt=(calcfc,TS,noeigen,z-matrix)\\Claisen rearrangement
allyl p-tolyl ether GS\\0,1\C,0,0.,0.,0.\O,0,0.,0.,1.2708\C,0,2.08959

52756,0.,1.8295133291\C,0,2.4754914457,1.0939894057,1.0876236296\C,0,2
.4869589939,1.0530769466,-0.3053284966\C,0,0.2849461209,1.1973847996,-
0.7535670389\C,0,0.1951897135,1.191059217,-2.1746212196\C,0,-0.1511487
376,-1.2104801042,-0.7664910087\C,0,-0.0033870543,0.025585482,-2.87450
65318\C,0,-0.1698689531,-1.1858936013,-2.1339418844\C,0,-0.0730927841,
-0.0054543258,-4.3804746761\H,0,2.1387027642,-0.9990712167,1.418310935
4\H,0,1.9570489588,0.0685025124,2.9009750035\H,0,2.5099039614,2.063374
5832,1.575992503\H,0,2.7590546302,1.9323198493,-0.8755347513\H,0,2.602
2463016,0.1092808311,-0.8244591343\H,0,0.2202707239,2.1381421208,-0.22
27816157\H,0,0.3033992142,2.1295416244,-2.7102866701\H,0,-0.3296350389
,-2.1098401088,-2.6832273648\H,0,-0.3046362833,-2.1317761685,-0.215314
644\H,0,0.697179285,-0.6612474089,-4.8003537797\H,0,0.0661645729,0.990
0904749,-4.8062760714\H,0,-1.0399186574,-0.3865046806,-4.7255447528\\V
ersion=IA64L-G09RevA.01\State=1-A\HF/CbsB3=-460.5540064\E2(CBS)/CbsB3=
-2.0444104\CBS-Int/CbsB3=0.0606571\OIii/CbsB3=14.6102304\MP2/CbsB4=-46
2.029577\MP4(SDQ)/CbsB4=-462.0896256\MP4(SDQ)/6-31+G(d')=-461.993425\Q
CISD(T)/6-31+G(d')=-462.0694422\CBSQB3=-462.5698836\

**Ground state**

Temperature= 298.150000 Pressure=
1.000000
 E(ZPE)= 0.190770 E(Thermal)=
0.201786
 E(SCF)= -460.643268 DE(MP2)= -
1.828726
 DE(CBS)= -0.174473 DE(MP34)= -
0.075187
 DE(CCSD)= -0.066335 DE(Int)=
0.059779
 DE(Empirical)= -0.085634
 CBS-QB3 (0 K)= -462.623075 CBS-QB3 Energy= -
462.612058
 CBS-QB3 Enthalpy= -462.611114 CBS-QB3 Free Energy= -
462.661130
 1\1\ ASN_ALTIX-ALTIX7\Mixed\CBS-QB3\CBS-QB3\C10H12O1\AUBOXA\11-Aug-200
 9\0\\# CBS-QB3 opt=(calcfc,noeigen,z-matrix)\\Claisen rearrangement al
 lyl p-tolyl ether GS\\0,1\C\O,1,1.3701\C,2,1.429,1,119.8468\C,3,1.5043
 ,2,113.243,1,79.2668,0\C,4,1.3285,3,124.6467,2,-140.7361,0\C,1,1.4003,
 2,115.3757,3,-183.8689,0\C,6,1.3857,1,120.2675,2,-179.1248,0\C,1,1.394
 3,2,125.3671,3,-3.3122,0\C,7,1.4026,6,121.4835,1,0.0136,0\C,9,1.3927,7
 ,117.4081,6,-0.2206,0\C,9,1.5102,7,120.899,6,-180.1321,0\H,3,1.0962,2,
 110.9759,1,-46.8911,0\H,3,1.0937,2,103.5789,1,197.7577,0\H,4,1.0874,3,
 114.8701,2,41.8548,0\H,5,1.0838,4,121.6628,3,183.3711,0\H,5,1.0865,4,1
 21.7366,3,2.9844,0\H,6,1.0835,1,118.3715,2,0.5459,0\H,7,1.0858,6,119.0
 067,1,179.851,0\H,10,1.0855,9,119.4005,7,180.4254,0\H,8,1.0806,1,121.1
 145,2,-2.2207,0\H,11,1.0947,9,111.4613,7,59.0461,0\H,11,1.0949,9,111.4
 433,7,-60.622,0\H,11,1.092,9,111.2522,7,-180.7743,0\\Version=IA64L-G09
 RevA.01\State=1-A\HF/CbsB3=-460.6432676\E2(CBS)/CbsB3=-2.0031992\CBS-I
 nt/CbsB3=0.0597785\OIii/CbsB3=14.7899135\MP2/CbsB4=-462.0776589\MP4(SD
 Q)/CbsB4=-462.1528463\MP4(SDQ)/6-31+G(d')=-462.0556437\QCISD(T)/6-31+G
 (d')=-462.1219792\CBSQB3=-462.6230751\

**<u>Condensed-Phase</u> B3LYP/6-311+G(2d,p) allyl *p*-$OCH_3$-phenyl ether optimizations in PCM:**

**Water transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -294.8042 | 55.5767 | 92.7257 |
| Red. masses -- | 7.1909 | 3.2332 | 3.4270 |
| Frc consts -- | 0.3682 | 0.0059 | 0.0174 |
| IR Inten -- | 72.3838 | 4.3050 | 3.0028 |

Sum of electronic and zero-point Energies= -538.610369
Sum of electronic and thermal Energies= -538.599033
Sum of electronic and thermal Enthalpies= -538.598088
Sum of electronic and thermal Free Energies= -538.647494

1\1\ ASN_ALTIX-ALTIX8\FTS\RB3LYP\6-311+G(2d,p)\C10H12O2\AUBOXA\27-Aug-2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,TS,noeigen,z-matrix) scrf=(pcm,solvent=water)\\Claisen rearrangement allyl p-OCH3 ether TS\\0,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B5,2,A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0\C,8,B9,1,A8,2,D7,0\H,3,B10,2,A9,1,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B12,3,A11,2,D10,0\H,5,B13,4,A12,3,D11,0\H,5,B14,4,A13,3,D12,0\H,6,B15,1,A14,2,D13,0\H,7,B16,6,A15,1,D14,0\H,10,B17,8,A16,1,D15,0\H,8,B18,1,A17,2,D16,0\O,9,B19,7,A18,6,D17,0\C,20,B20,9,A19,7,D18,0\H,21,B21,20,A20,9,D19,0\H,21,B22,20,A21,9,D20,0\H,21,B23,20,A22,9,D21,0\\B1=1.27630878\B2=2.28167534\B3=1.37152505\B4=1.39079848\B5=1.44034406\B6=1.4109891\B7=1.43463697\B8=1.37656547\B9=1.37007894\B10=1.08070573\B11=1.08173415\B12=1.08535317\B13=1.08109226\B14=1.0820242\B15=1.08130062\B16=1.08340597\B17=1.08153862\B18=1.08382947\B19=1.36574379\B20=1.42851647\B21=1.08731434\B22=1.0925833\B23=1.09265772\A1=103.52417448\A2=95.28875324\A3=121.49484047\A4=121.82812264\A5=120.84871929\A6=122.06067632\A7=120.42954519\A8=121.76205755\A9=84.51775039\A10=99.56615713\A11=118.73325324\A12=120.48688055\A13=120.43375753\A14=117.69594775\A15=120.23558217\A16=119.18488246\A17=117.9436226\A18=117.00059778\A19=118.97120756\A20=105.66882618\A21=111.31087242\A22=111.29645954\D1=52.4953598\D2=-69.17777663\D3=-78.78857701\D4=-174.13085822\D5=96.97284085\D6=-7.5191815\D7=176.85407417\D8=-68.42635557\D9=175.40710915\D10=97.58064173\D11=178.18087743\D12=-22.10517233\D13=-11.96045412\D14=174.00163285\D15=-177.62178152\D16=-0.25520886\D17=-178.35017691\D18=182.11074665\D19=179.0542234\D20=-62.35536961\D21=60.42070074\\Version=IA64L-G09RevA.02\State=1-A\HF=-538.8050862\RMSD=3.785e-09\RMSF=4.258e-05\Dipole=0.9517656,-0.4036861,-1.4446835\Quadrupole=-3.2145581,5.0426801,-1.828122,2.801659,6.4356757,5.5246552\PG=C01 [X(C10H12O2)]\\@

**Water ground state**

1\1\ASN_X86_64-DMC14\FOpt\RB3LYP\6-311+G(2d,p)\C10H12O2\AUBOXA\27-Aug-2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,noeigen,z-matrix) scrf=(pcm,solvent=water)\\Claisen rearrangement allyl p-OCH3 ether GS\\0,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B5,2,A4,3,D3

,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0\C,9,B9,7,A
8,6,D7,0\H,3,B10,2,A9,1,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B12,3,A11,2,D10,
0\H,5,B13,4,A12,3,D11,0\H,5,B14,4,A13,3,D12,0\H,6,B15,1,A14,2,D13,0\H,
7,B16,6,A15,1,D14,0\H,10,B17,9,A16,7,D15,0\H,8,B18,1,A17,2,D16,0\O,9,B
19,7,A18,6,D17,0\C,20,B20,9,A19,7,D18,0\H,21,B21,20,A20,9,D19,0\H,21,B
22,20,A21,9,D20,0\H,21,B23,20,A22,9,D21,0\\B1=1.37424583\B2=1.43570576
\B3=1.50074634\B4=1.32742752\B5=1.39302703\B6=1.39129043\B7=1.39647387
\B8=1.39579755\B9=1.39386582\B10=1.09333399\B11=1.092042\B12=1.0871211
\B13=1.08347951\B14=1.0854224\B15=1.08332024\B16=1.08078709\B17=1.0833
5901\B18=1.07987455\B19=1.37071542\B20=1.42572441\B21=1.09334544\B22=1
.08754571\B23=1.09335642\A1=119.6992485\A2=112.85911797\A3=124.5622715
1\A4=115.82170886\A5=121.04864933\A6=125.12779314\A7=119.88221328\A8=1
19.08787885\A9=110.92459745\A10=103.48964621\A11=115.21612881\A12=121.
44871633\A13=121.72928308\A14=118.94997761\A15=118.91597417\A16=118.99
494094\A17=121.26480558\A18=124.62735809\A19=118.0267342\A20=111.26076
501\A21=105.92752188\A22=111.27411969\D1=76.14895675\D2=-136.68191506\
D3=185.23303857\D4=-178.25466225\D5=6.67980875\D6=-0.1596771\D7=-0.077
60896\D8=-50.06669886\D9=-165.27738132\D10=45.79623649\D11=-176.784500
51\D12=2.70095962\D13=1.42151253\D14=179.829785\D15=-179.73496297\D16=
-2.19234171\D17=180.04165078\D18=0.29914331\D19=61.00382784\D20=-180.2
0082552\D21=-61.40285098\\Version=AM64L-G09RevA.02\State=1-A\HF=-538.8
565843\RMSD=5.884e-09\RMSF=3.733e-05\Dipole=0.0143257,0.091792,0.03800
35\Quadrupole=4.1717807,-5.3088395,1.1370588,-0.6351973,8.2017152,0.04
04963\PG=C01 [X(C10H12O2)]\\@

**Pentadecane transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -324.7767 | 55.9490 | 93.6223 |
| Red. masses -- | 7.7016 | 3.1862 | 3.6189 |
| Frc consts -- | 0.4786 | 0.0059 | 0.0187 |
| IR Inten -- | 46.4568 | 3.0840 | 1.7685 |

Sum of electronic and zero-point Energies= -538.603422
Sum of electronic and thermal Energies= -538.592147
Sum of electronic and thermal Enthalpies= -538.591203
Sum of electronic and thermal Free Energies= -538.640445

1\1\ ASN_ALTIX-ALTIX2\FTS\RB3LYP\6-311+G(2d,p)\C10H12O2\AUBOXA\27-Aug-
2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,TS,noeigen,z-matrix) scrf=(p
cm,solvent=n-Pentadecane)\\Claisen rearrangement allyl p-OCH3 ether TS
\\0,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B
5,2,A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0
\C,8,B9,1,A8,2,D7,0\H,3,B10,2,A9,1,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B12,3
,A11,2,D10,0\H,5,B13,4,A12,3,D11,0\H,5,B14,4,A13,3,D12,0\H,6,B15,1,A14
,2,D13,0\H,7,B16,6,A15,1,D14,0\H,10,B17,8,A16,1,D15,0\H,8,B18,1,A17,2,
D16,0\O,9,B19,7,A18,6,D17,0\C,20,B20,9,A19,7,D18,0\H,21,B21,20,A20,9,D
19,0\H,21,B22,20,A21,9,D20,0\H,21,B23,20,A22,9,D21,0\\B1=1.27064406\B2
=2.24634246\B3=1.37107078\B4=1.39274245\B5=1.44327911\B6=1.41379454\B7
=1.43659683\B8=1.37438159\B9=1.36880617\B10=1.08058742\B11=1.08132608\
B12=1.08552436\B13=1.08139355\B14=1.0822016\B15=1.08117698\B16=1.08341
126\B17=1.08183985\B18=1.08366976\B19=1.36669099\B20=1.42350417\B21=1.
08770471\B22=1.09369216\B23=1.09354758\A1=104.01185348\A2=95.73452828\

A3=121.55737664\A4=121.81479693\A5=120.72971429\A6=122.0974582\A7=120.
43894536\A8=121.73591272\A9=84.90707487\A10=98.96648163\A11=118.500040
64\A12=120.3527359\A13=120.31315473\A14=117.19347054\A15=120.45332664\
A16=119.19826777\A17=117.66485724\A18=116.99917675\A19=119.01840018\A2
0=105.72134676\A21=111.50821628\A22=111.43370477\D1=52.47698308\D2=-68
.39163549\D3=-77.65143313\D4=-174.19551829\D5=97.51141018\D6=-8.136593
24\D7=177.20867076\D8=-68.53248314\D9=175.43064493\D10=97.08232767\D11
=178.61308168\D12=-23.896446\D13=-14.24129489\D14=173.1203346\D15=-177
.3489257\D16=0.52077246\D17=-178.03207321\D18=184.04769581\D19=178.010
90763\D20=-63.45755333\D21=59.39490495\\Version=IA64L-G09RevA.02\State
=1-A\HF=-538.7983321\RMSD=8.924e-09\RMSF=2.085e-05\Dipole=0.63828,-0.3
523424,-1.1734914\Quadrupole=-3.3074873,4.4114408,-1.1039535,2.357042,
5.4416804,4.7111344\PG=C01 [X(C10H12O2)]\\@

**Pentadecane ground state**

1\1\ASN_X86_64-DMC24\FOpt\RB3LYP\6-311+G(2d,p)\C10H12O2\AUBOXA\27-Aug-
2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,noeigen,z-matrix) scrf=(pcm,
solvent=n-Pentadecane)\\Claisen rearrangement allyl p-OCH3 ether GS\\0
,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B5,2
,A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0\C,
9,B9,7,A8,6,D7,0\H,3,B10,2,A9,1,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B12,3,A1
1,2,D10,0\H,5,B13,4,A12,3,D11,0\H,5,B14,4,A13,3,D12,0\H,6,B15,1,A14,2,
D13,0\H,7,B16,6,A15,1,D14,0\H,10,B17,9,A16,7,D15,0\H,8,B18,1,A17,2,D16
,0\O,9,B19,7,A18,6,D17,0\C,20,B20,9,A19,7,D18,0\H,21,B21,20,A20,9,D19,
0\H,21,B22,20,A21,9,D20,0\H,21,B23,20,A22,9,D21,0\\B1=1.37387893\B2=1.
43016433\B3=1.50171994\B4=1.32673949\B5=1.3921464\B6=1.3906664\B7=1.39
605592\B8=1.39547059\B9=1.3928978\B10=1.09449549\B11=1.09262918\B12=1.
08689522\B13=1.08323014\B14=1.08566049\B15=1.08311847\B16=1.0809575\B1
7=1.08313028\B18=1.07984696\B19=1.37057536\B20=1.42057643\B21=1.094363
97\B22=1.08793661\B23=1.09435789\A1=119.70023665\A2=113.22487694\A3=12
4.57331964\A4=115.7677265\A5=120.97700528\A6=125.16911323\A7=119.94487
763\A8=119.07780426\A9=110.91996977\A10=103.56496565\A11=115.13292435\
A12=121.55157176\A13=121.71735443\A14=118.7975457\A15=118.89940633\A16
=118.81196232\A17=121.14833545\A18=124.69813679\A19=118.04636563\A20=1
11.40006499\A21=105.97549421\A22=111.40085064\D1=77.4036086\D2=-139.23
916072\D3=180.48128618\D4=-178.67778232\D5=1.49687745\D6=-0.05928042\D
7=-0.17321627\D8=-48.82608089\D9=-164.01958729\D10=43.23423709\D11=-17
6.7382605\D12=2.86464789\D13=0.98644782\D14=179.84619927\D15=-179.5898
5313\D16=-2.30292059\D17=179.99808695\D18=-0.01410358\D19=61.22024806\
D20=-180.00659469\D21=-61.23427664\\Version=AM64L-G09RevA.02\State=1-A
\HF=-538.85211\RMSD=3.576e-09\RMSF=4.498e-05\Dipole=0.0041314,0.057627
9,0.034101\Quadrupole=3.5760283,-4.6239746,1.0479463,0.0621205,7.01133
68,0.4787642\PG=C01 [X(C10H12O2)]\\@

**<u>Gas-Phase</u> B3LYP/6-311+G(2d,p) allyl *p*-$OCH_3$-phenyl ether optimizations:**

**Transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -339.7861 | 52.4592 | 94.2021 |

```
 Red. masses --     7.9783                3.1243
3.8926
 Frc consts  --     0.5427                0.0051
0.0204
 IR Inten    --    34.4252                2.7633
1.2223

 Sum of electronic and zero-point Energies=          -538.599462
 Sum of electronic and thermal Energies=             -538.588205
 Sum of electronic and thermal Enthalpies=           -538.587261
 Sum of electronic and thermal Free Energies=        -538.636493


 1\1\ASN_X86_64-DMC40\FTS\RB3LYP\6-311+G(2d,p)\C10H12O2\AUBOXA\27-Aug-2
 009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,ts,noeigen,z-matrix)\\Claisen
  rearrangement allyl p-OCH3 ether TS\\0,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,
 2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B5,2,A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C
 ,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0\C,8,B9,1,A8,2,D7,0\H,3,B10,2,A9,1
 ,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B12,3,A11,2,D10,0\H,5,B13,4,A12,3,D11,0
 \H,5,B14,4,A13,3,D12,0\H,6,B15,1,A14,2,D13,0\H,7,B16,6,A15,1,D14,0\H,1
 0,B17,8,A16,1,D15,0\H,8,B18,1,A17,2,D16,0\O,9,B19,7,A18,6,D17,0\C,20,B
 20,9,A19,7,D18,0\H,21,B21,20,A20,9,D19,0\H,21,B22,20,A21,9,D20,0\H,21,
 B23,20,A22,9,D21,0\\B1=1.2676997\B2=2.22870814\B3=1.37074166\B4=1.3937
 8995\B5=1.44480119\B6=1.41524522\B7=1.43748589\B8=1.37319147\B9=1.3681
 7638\B10=1.08050092\B11=1.08114032\B12=1.08560871\B13=1.08155881\B14=1
 .08230346\B15=1.08121577\B16=1.08339927\B17=1.08201767\B18=1.08364429\
 B19=1.36739579\B20=1.42025205\B21=1.08800062\B22=1.09433677\B23=1.0941
 6958\A1=104.23855835\A2=95.9851775\A3=121.56988606\A4=121.78425527\A5=
 120.6510938\A6=122.12749697\A7=120.43984256\A8=121.7228073\A9=85.09188
 656\A10=98.67886911\A11=118.40835084\A12=120.27110302\A13=120.23400181
 \A14=116.93947823\A15=120.57494657\A16=119.19834435\A17=117.51472589\A
 18=116.96224822\A19=119.05248168\A20=105.74897447\A21=111.62777602\A22
 =111.54756582\D1=52.44018994\D2=-67.94678471\D3=-77.09103073\D4=-174.1
 8560633\D5=97.80142604\D6=-8.47581865\D7=177.37022088\D8=-68.6007071\D
 9=175.41547324\D10=96.90019835\D11=178.83735581\D12=-24.8173958\D13=-1
 5.33567284\D14=172.7474206\D15=-177.28099844\D16=0.83672889\D17=-177.8
 9717502\D18=184.45596651\D19=177.86323544\D20=-63.62570621\D21=59.2482
 3683\\Version=AM64L-G09RevA.02\State=1-A\HF=-538.7944475\RMSD=4.110e-0
 9\RMSF=3.847e-05\Dipole=0.5217175,-0.3245573,-1.0237742\Quadrupole=-3.
 3006227,4.1136438,-0.8130211,2.113529,4.8610852,4.2493831\PG=C01 [X(C1
 0H12O2)]\\@
```

**Ground state**

```
 1\1\ASN_X86_64-DMC14\FOpt\RB3LYP\6-311+G(2d,p)\C10H12O2\AUBOXA\27-Aug-
 2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,noeigen,z-matrix)\\Claisen r
 earrangement allyl p-OCH3 ether GS\\0,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2,
 A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B5,2,A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C,1
 ,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0\C,9,B9,7,A8,6,D7,0\H,3,B10,2,A9,1,D
 8,0\H,3,B11,2,A10,1,D9,0\H,4,B12,3,A11,2,D10,0\H,5,B13,4,A12,3,D11,0\H
 ,5,B14,4,A13,3,D12,0\H,6,B15,1,A14,2,D13,0\H,7,B16,6,A15,1,D14,0\H,10,
 B17,9,A16,7,D15,0\H,8,B18,1,A17,2,D16,0\O,9,B19,7,A18,6,D17,0\C,20,B20
 ,9,A19,7,D18,0\H,21,B21,20,A20,9,D19,0\H,21,B22,20,A21,9,D20,0\H,21,B2
 3,20,A22,9,D21,0\\B1=1.37385965\B2=1.42715055\B3=1.50216636\B4=1.32639
 518\B5=1.39164248\B6=1.39034701\B7=1.39583967\B8=1.39526614\B9=1.39242
 544\B10=1.09506675\B11=1.09299716\B12=1.08683245\B13=1.0831237\B14=1.0
 8569611\B15=1.08297084\B16=1.08105244\B17=1.08297884\B18=1.07988444\B1
```

9=1.37044965\B20=1.41753622\B21=1.09496127\B22=1.08819962\B23=1.094949
24\A1=119.69168406\A2=113.32089048\A3=124.63569517\A4=115.7374817\A5=1
20.93581351\A6=125.17667772\A7=119.96556045\A8=119.08882739\A9=110.974
43774\A10=103.63273423\A11=115.04632671\A12=121.60163631\A13=121.71618
439\A14=118.71260531\A15=118.89453654\A16=118.71190184\A17=121.1044504
8\A18=124.74241227\A19=118.0708885\A20=111.48462772\A21=105.99591758\A
22=111.4936814\D1=78.4505065\D2=-140.21656246\D3=178.24959325\D4=-178.
90956216\D5=-0.97202359\D6=-0.00907062\D7=-0.21313713\D8=-47.76376637\
D9=-162.99795034\D10=42.21008718\D11=-176.7395382\D12=2.87959062\D13=0
.74584536\D14=179.84885846\D15=-179.54333341\D16=-2.2791415\D17=179.95
872807\D18=0.08353174\D19=61.2157193\D20=-180.01352494\D21=-61.2417598
6\\Version=AM64L-G09RevA.02\State=1-A\HF=-538.8495367\RMSD=3.213e-09\R
MSF=3.573e-05\Dipole=0.000282,0.0385623,0.0311899\Quadrupole=3.2659555
,-4.2412685,0.975313,0.3081183,6.3357235,0.6120333\PG=C01 [X(C10H12O2)
]\\@

**<u>Gas-Phase</u> PDDG/PM3 allyl *p*-$OCH_3$-phenyl ether optimizations:**

**Transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -952.5158 | 48.3906 | 93.0768 |
| Red. masses -- | 10.4545 | 3.0826 | 4.8518 |
| Frc consts -- | 5.5885 | 0.0043 | 0.0248 |
| IR Inten -- | 99.9580 | 4.0025 | 2.7263 |

| | |
|---|---|
| Sum of electronic and zero-point Energies= | 0.208917 |
| Sum of electronic and thermal Energies= | 0.219646 |
| Sum of electronic and thermal Enthalpies= | 0.220590 |
| Sum of electronic and thermal Free Energies= | 0.172387 |

1\1\ ASN_ALTIX-ALTIX9\FTS\RPDDG\ZDO\C10H12O2\AUBOXA\10-Aug-2009\1\\# P
DDG opt=(ts,noeigen,z-matrix)\\Claisen rearrangement allyl p-OCH3 ethe
r TS\\0,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C
,1,B5,2,A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,
D6,0\C,8,B9,1,A8,2,D7,0\H,3,B10,2,A9,1,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B
12,3,A11,2,D10,0\H,5,B13,4,A12,3,D11,0\H,5,B14,4,A13,3,D12,0\H,6,B15,1
,A14,2,D13,0\H,7,B16,6,A15,1,D14,0\H,10,B17,8,A16,1,D15,0\H,8,B18,1,A1
7,2,D16,0\O,9,B19,7,A18,6,D17,0\C,20,B20,9,A19,7,D18,0\H,21,B21,20,A20
,9,D19,0\H,21,B22,20,A21,9,D20,0\H,21,B23,20,A22,9,D21,0\\B1=1.2789157
8\B2=1.85968532\B3=1.3982623\B4=1.40430347\B5=1.44233045\B6=1.43989513
\B7=1.44884956\B8=1.3696652\B9=1.357239\B10=1.10396575\B11=1.09980431\
B12=1.09977538\B13=1.09529495\B14=1.09866757\B15=1.11096353\B16=1.1023
1152\B17=1.10235397\B18=1.10204867\B19=1.37643927\B20=1.40813406\B21=1
.09134324\B22=1.10313237\B23=1.10331026\A1=107.50181815\A2=99.11541418
\A3=118.98724494\A4=119.85523464\A5=118.08633683\A6=120.52021338\A7=12
0.31330722\A8=120.75070877\A9=100.29671249\A10=102.01689134\A11=119.91
311012\A12=118.34954063\A13=119.64469112\A14=118.28630295\A15=117.4973
1224\A16=119.31921665\A17=118.89045887\A18=114.90258206\A19=117.461193
92\A20=100.6405063\A21=114.47107015\A22=114.86098148\D1=56.28654101\D2
=-69.95691387\D3=-70.68514574\D4=-180.87096592\D5=99.99002463\D6=-7.32

271832\D7=183.55215366\D8=-66.11817186\D9=177.91000242\D10=95.84881197
\D11=181.68342055\D12=-34.1489574\D13=-34.7866375\D14=174.21082042\D15
=-178.0037649\D16=5.79250113\D17=-177.29012406\D18=184.46996372\D19=17
7.54078037\D20=-65.01570422\D21=59.83176429\\Version=IA64L-G09RevA.01\
State=1-A\HF=0.0173684\RMSD=8.707e-09\RMSF=1.210e-05\Dipole=0.427798,-
0.4978881,-0.7004693\PG=C01 [X(C10H12O2)]\\@

**Ground state**

1\1\ ASN_ALTIX-ALTIX9\FOpt\RPDDG\ZDO\C10H12O2\AUBOXA\10-Aug-2009\1\\#
PDDG opt=(noeigen,z-matrix)\\Claisen rearrangement allyl p-OCH3 ether
GS\\0,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1
,B5,2,A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6
,0\C,9,B9,7,A8,6,D7,0\H,3,B10,2,A9,1,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B12
,3,A11,2,D10,0\H,5,B13,4,A12,3,D11,0\H,5,B14,4,A13,3,D12,0\H,6,B15,1,A
14,2,D13,0\H,7,B16,6,A15,1,D14,0\H,10,B17,9,A16,7,D15,0\H,8,B18,1,A17,
2,D16,0\O,9,B19,7,A18,6,D17,0\C,20,B20,9,A19,7,D18,0\H,21,B21,20,A20,9
,D19,0\H,21,B22,20,A21,9,D20,0\H,21,B23,20,A22,9,D21,0\\B1=1.37728858\
B2=1.41071057\B3=1.50200858\B4=1.33002887\B5=1.40247767\B6=1.38971511\
B7=1.39957736\B8=1.39885225\B9=1.40277589\B10=1.12182235\B11=1.1065834
7\B12=1.10051027\B13=1.08709474\B14=1.09057466\B15=1.10225117\B16=1.10
185371\B17=1.10215979\B18=1.10460154\B19=1.37601698\B20=1.40851119\B21
=1.10338571\B22=1.09123892\B23=1.1033775\A1=118.05355369\A2=111.869297
34\A3=121.12345882\A4=113.81170381\A5=119.52008828\A6=125.2760586\A7=1
19.54085497\A8=120.95238052\A9=114.64995872\A10=99.29008348\A11=116.93
185093\A12=122.1530228\A13=122.60849443\A14=121.40316669\A15=118.57168
536\A16=121.34569965\A17=121.64872443\A18=124.94597262\A19=116.9436702
4\A20=114.55329417\A21=100.89814548\A22=114.54378566\D1=82.24642083\D2
=-131.06139257\D3=170.91099549\D4=-180.2972671\D5=-9.58170545\D6=0.178
42303\D7=-0.31931853\D8=-46.87508043\D9=-161.20929532\D10=48.86732175\
D11=-178.39812242\D12=2.0651384\D13=-0.16934633\D14=180.10478333\D15=-
179.65771555\D16=-0.77023302\D17=179.75297036\D18=0.20689413\D19=62.06
830877\D20=-180.12490052\D21=-62.32562621\\Version=IA64L-G09RevA.01\St
ate=1-A\HF=-0.0643699\RMSD=3.515e-09\RMSF=4.887e-05\Dipole=0.0020609,-
0.0905499,0.0236985\PG=C01 [X(C10H12O2)]\\@

**<u>Gas-Phase</u> CBS-QB3 allyl *p*-$OCH_3$-phenyl ether optimizations:**

**Transition structure**

Temperature= 298.150000 Pressure=
1.000000
E(ZPE)= 0.193073 E(Thermal)=
0.204390
E(SCF)= -535.420220 DE(MP2)= -
2.099190
DE(CBS)= -0.200226 DE(MP34)= -
0.060954
DE(CCSD)= -0.081767 DE(Int)=
0.067262
DE(Empirical)= -0.095077
CBS-QB3 (0 K)= -537.697098 CBS-QB3 Energy= -
537.685781
CBS-QB3 Enthalpy= -537.684837 CBS-QB3 Free Energy= -
537.734113
1\1\ ASN_ALTIX-ALTIX7\Mixed\CBS-QB3\CBS-QB3\C10H12O2\AUBOXA\12-Aug-200

9\0\\# CBS-QB3 opt=(calcfc,ts,noeigen,z-matrix)\\Claisen rearrangement
allyl p-OCH3 ether TS\\0,1\C,0,0.,0.,0.\O,0,0.,0.,1.2664\C,0,2.148590
735,0.,1.8253292114\C,0,2.5007822183,1.0830926904,1.0574435798\C,0,2.4
800871166,1.0273361856,-0.3374891658\C,0,0.2814521745,1.1981994843,-0.
7640905439\C,0,0.1594811723,1.2009404975,-2.1764309142\C,0,-0.16155885
68,-1.2084461081,-0.7677638828\C,0,-0.0521697523,0.0256285667,-2.86057
08424\C,0,-0.2082494267,-1.1940439536,-2.1368922802\H,0,2.1648646347,-
1.0033699059,1.4223280301\H,0,2.0410571081,0.0835022833,2.8986285628\H
,0,2.550804687,2.0586545211,1.5326535935\H,0,2.7439132027,1.8990703582
,-0.9223524734\H,0,2.5897877859,0.0773512147,-0.8464422169\H,0,0.22926
66679,2.1384699218,-0.2314783142\H,0,0.2515036293,2.1221917244,-2.7398
583578\H,0,-0.3782375957,-2.1209725182,-2.6701495338\H,0,-0.3079619067
,-2.1307481161,-0.2167360767\O,0,-0.1158490408,0.1062747363,-4.2232018
98\C,0,-0.417773954,-1.0618890636,-4.972526109\H,0,-0.4445856946,-0.74
49189443,-6.0139262646\H,0,0.3529719941,-1.8320332557,-4.8540161643\H,
0,-1.3936076677,-1.4785721307,-4.6992455301\\Version=IA64L-G09RevA.01\
State=1-A\HF/CbsB3=-535.4202198\E2(CBS)/CbsB3=-2.2994164\CBS-Int/CbsB3
=0.0672623\OIii/CbsB3=16.4208572\MP2/CbsB4=-537.0548313\MP4(SDQ)/CbsB4
=-537.1157851\MP4(SDQ)/6-31+G(d')=-537.0194345\QCISD(T)/6-31+G(d')=-53
7.1012011\CBSQB3=-537.6970978\

**Ground State**

Temperature= 298.150000 Pressure= 1.000000
E(ZPE)= 0.195778 E(Thermal)= 0.207553
E(SCF)= -535.511966 DE(MP2)= -2.058733
DE(CBS)= -0.198270 DE(MP34)= -0.076659
DE(CCSD)= -0.071645 DE(Int)= 0.066392
DE(Empirical)= -0.096181
CBS-QB3 (0 K)= -537.751283 CBS-QB3 Energy= -537.739509
CBS-QB3 Enthalpy= -537.738564 CBS-QB3 Free Energy= -537.790079

1\1\ ASN_ALTIX-ALTIX7\Mixed\CBS-QB3\CBS-QB3\C10H12O2\AUBOXA\12-Aug-200
9\0\\# CBS-QB3 opt=(calcfc,noeigen,z-matrix)\\Claisen rearrangement al
lyl p-OCH3 ether GS\\0,1\C\O,1,1.3735\C,2,1.4276,1,119.7425\C,3,1.5047
,2,113.3025,1,79.117,0\C,4,1.3286,3,124.6722,2,-140.7537,0\C,1,1.3944,
2,115.6748,3,175.923,0\C,6,1.3921,1,120.9727,2,-179.1828,0\C,1,1.3987,
2,125.262,3,-3.5196,0\C,7,1.3981,6,119.9715,1,0.0611,0\C,9,1.3955,7,11
9.0282,6,-0.2375,0\H,3,1.0966,2,111.0462,1,-47.0573,0\H,3,1.0938,2,103
.6281,1,-162.3966,0\H,4,1.0875,3,114.8543,2,41.7979,0\H,5,1.0838,4,121
.6758,3,-176.6693,0\H,5,1.0865,4,121.7303,3,2.9508,0\H,6,1.0833,1,118.
4366,2,0.4994,0\H,7,1.0818,6,118.9068,1,179.8824,0\H,10,1.0833,9,118.4
235,7,-179.5288,0\H,8,1.0805,1,121.0634,2,-2.1338,0\O,9,1.3695,7,124.8
702,6,179.9276,0\C,20,1.4177,9,118.1708,7,0.0742,0\H,21,1.0965,20,111.
6642,9,61.2749,0\H,21,1.0891,20,105.945,9,-180.0048,0\H,21,1.0965,20,1
11.673,9,-61.2821,0\\Version=IA64L-G09RevA.01\State=1-A\HF/CbsB3=-535.
511966\E2(CBS)/CbsB3=-2.2570023\CBS-Int/CbsB3=0.0663924\OIii/CbsB3=16.
6115892\MP2/CbsB4=-537.1039607\MP4(SDQ)/CbsB4=-537.1806201\MP4(SDQ)/6-
31+G(d')=-537.0833017\QCISD(T)/6-31+G(d')=-537.1549468\CBSQB3=-537.751
2834\

**<u>Condensed-Phase</u> B3LYP/6-311+G(2d,p) allyl *p*-Br-phenyl ether optimizations in PCM:**

**Water transition structure**

```
                     1                      2                      3
                     A                      A                      A
 Frequencies --  -313.1168               65.7087
107.0725
 Red. masses --     7.3586                6.1530
3.0623
 Frc consts  --     0.4251                0.0157
0.0207
 IR Inten    --   100.8622                1.1446
0.8601

 Sum of electronic and zero-point Energies=          -2997.632240
 Sum of electronic and thermal Energies=             -2997.622157
 Sum of electronic and thermal Enthalpies=           -2997.621213
 Sum of electronic and thermal Free Energies=        -2997.668886
```

```
 1\1\ ASN_ALTIX-ALTIX7\FTS\RB3LYP\6-311+G(2d,p)\C9H9Br1O1\AUBOXA\27-Aug
 -2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,TS,noeigen,z-matrix) scrf=(
 pcm,solvent=water)\\Claisen rearrangement allyl p-Br ether TS\\0,1\C\O
 ,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B5,2,A4,3,
 D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0\C,8,B9,1
 ,A8,2,D7,0\H,3,B10,2,A9,1,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B12,3,A11,2,D1
 0,0\H,5,B13,4,A12,3,D11,0\H,5,B14,4,A13,3,D12,0\H,6,B15,1,A14,2,D13,0\
 H,7,B16,6,A15,1,D14,0\H,10,B17,8,A16,1,D15,0\H,8,B18,1,A17,2,D16,0\Br,
 9,B19,7,A18,6,D17,0\\B1=1.27558323\B2=2.25020601\B3=1.37282726\B4=1.38
 927028\B5=1.43837403\B6=1.41756801\B7=1.43623384\B8=1.36637919\B9=1.36
 872964\B10=1.08069265\B11=1.08151775\B12=1.08486247\B13=1.08130807\B14
 =1.08211323\B15=1.0809374\B16=1.08202267\B17=1.08241274\B18=1.08355759
 \B19=1.92318856\A1=103.76532353\A2=96.34541454\A3=121.07916769\A4=121.
 42812967\A5=120.69986111\A6=121.8435208\A7=119.27868133\A8=121.4451430
 1\A9=85.47548366\A10=98.23555402\A11=118.91612723\A12=120.57065793\A13
 =120.42087794\A14=118.04622128\A15=119.59912484\A16=120.34200954\A17=1
 18.07630795\A18=120.07532297\D1=52.96000607\D2=-70.16485642\D3=-77.036
 70929\D4=-175.61036368\D5=98.08027025\D6=-6.87204794\D7=178.11419894\D
 8=-67.87339499\D9=175.83426304\D10=96.45322469\D11=178.16876231\D12=-2
 1.67245114\D13=-13.8482187\D14=174.10137557\D15=-177.8023254\D16=0.650
 50189\D17=-178.66847203\\Version=IA64L-G09RevA.02\State=1-A\HF=-2997.7
 849479\RMSD=6.655e-09\RMSF=1.701e-05\Dipole=1.4043232,0.4028825,0.0191
 884\Quadrupole=-0.3516576,6.8562218,-6.5045643,1.8616293,6.2147531,1.1
 239735\PG=C01 [X(C9H9Br1O1)]\\@
```

**Water ground state**

```
 1\1\ ASN_ALTIX-ALTIX7\FOpt\RB3LYP\6-311+G(2d,p)\C9H9Br1O1\AUBOXA\27-Au
```

g-2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,noeigen,z-matrix) scrf=(pc
m,solvent=water)\\Claisen rearrangement allyl p-Br ether GS\\0,1\C\O,1
,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B5,2,A4,3,D3
,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0\C,9,B9,7,A
8,6,D7,0\H,3,B10,2,A9,1,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B12,3,A11,2,D10,
0\H,5,B13,4,A12,3,D11,0\H,5,B14,4,A13,3,D12,0\H,6,B15,1,A14,2,D13,0\H,
7,B16,6,A15,1,D14,0\H,10,B17,9,A16,7,D15,0\H,8,B18,1,A17,2,D16,0\Br,9,
B19,7,A18,6,D17,0\\B1=1.36391448\B2=1.44012091\B3=1.49973748\B4=1.3273
4074\B5=1.39897389\B6=1.38602031\B7=1.39539753\B8=1.39004143\B9=1.3840
9756\B10=1.09270801\B11=1.091545\B12=1.08691357\B13=1.0833923\B14=1.08
539817\B15=1.08285591\B16=1.08167939\B17=1.08180954\B18=1.07946146\B19
=1.92512017\A1=120.00937517\A2=112.67259329\A3=124.37775377\A4=115.397
03053\A5=120.62930554\A6=125.1448324\A7=119.29386798\A8=120.86895155\A
9=110.69729205\A10=103.32774295\A11=115.30432046\A12=121.42641284\A13=
121.72816891\A14=119.03130856\A15=119.92240141\A16=120.70261961\A17=12
1.42145418\A18=119.47654335\D1=77.76680907\D2=-135.93711847\D3=182.851
46207\D4=-178.79118851\D5=3.81534469\D6=-0.09046324\D7=-0.08483192\D8=
-48.37179484\D9=-163.58925248\D10=46.54672202\D11=-176.77236038\D12=2.
63459499\D13=0.98421482\D14=179.88345273\D15=-179.80645937\D16=-1.6730
8887\D17=179.93024407\\Version=IA64L-G09RevA.02\State=1-A\HF=-2997.837
0096\RMSD=6.705e-09\RMSF=3.472e-05\Dipole=0.8173359,0.0529192,1.278304
9\Quadrupole=4.6685749,-3.2155824,-1.4529925,-0.5131979,2.9274428,0.16
85785\PG=C01 [X(C9H9Br1O1)]\\@

**Pentadecane transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -341.4596 | 66.3913 | 108.5563 |
| Red. masses -- | 7.8941 | 6.1434 | 3.0412 |
| Frc consts -- | 0.5423 | 0.0160 | 0.0211 |
| IR Inten -- | 66.6504 | 0.6619 | 0.4990 |

Sum of electronic and zero-point Energies= -2997.626441
Sum of electronic and thermal Energies= -2997.616416
Sum of electronic and thermal Enthalpies= -2997.615471
Sum of electronic and thermal Free Energies= -2997.662990

1\1\ ASN_ALTIX-ALTIX8\FTS\RB3LYP\6-311+G(2d,p)\C9H9Br1O1\AUBOXA\27-Aug
-2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,TS,noeigen,z-matrix) scrf=(
pcm,solvent=n-Pentadecane)\\Claisen rearrangement allyl p-Br ether TS\
\0,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B5
,2,A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0\
C,8,B9,1,A8,2,D7,0\H,3,B10,2,A9,1,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B12,3,
A11,2,D10,0\H,5,B13,4,A12,3,D11,0\H,5,B14,4,A13,3,D12,0\H,6,B15,1,A14,
2,D13,0\H,7,B16,6,A15,1,D14,0\H,10,B17,8,A16,1,D15,0\H,8,B18,1,A17,2,D
16,0\Br,9,B19,7,A18,6,D17,0\\B1=1.2707572\B2=2.21383352\B3=1.37271496\
B4=1.39099523\B5=1.44115224\B6=1.42008314\B7=1.4380477\B8=1.36504634\B
9=1.36727401\B10=1.08064356\B11=1.08118001\B12=1.0851146\B13=1.0815809
8\B14=1.08226583\B15=1.08097118\B16=1.08225676\B17=1.08250396\B18=1.08
349668\B19=1.92020491\A1=104.33149152\A2=96.75345354\A3=121.17289248\A
4=121.3950646\A5=120.5591593\A6=121.90968256\A7=119.38144295\A8=121.45
580488\A9=85.86216762\A10=97.88480967\A11=118.65777892\A12=120.4257179

7\A13=120.33924145\A14=117.54649771\A15=119.74054674\A16=120.46833976\
A17=117.78857273\A18=120.18994536\D1=52.84263803\D2=-69.12181324\D3=-7
6.28634433\D4=-175.55649825\D5=98.32855954\D6=-7.57172754\D7=178.41369
902\D8=-68.07623971\D9=175.7847169\D10=96.1597455\D11=178.52087493\D12
=-23.5315062\D13=-15.84436501\D14=173.30307166\D15=-177.55501388\D16=1
.40180986\D17=-178.43319589\\Version=IA64L-G09RevA.02\State=1-A\HF=-29
97.7792831\RMSD=5.048e-09\RMSF=2.569e-05\Dipole=1.0300035,0.2952472,0.
0845213\Quadrupole=-1.0020915,6.1499802,-5.1478887,1.5832205,4.876553,
0.8423514\PG=C01 [X(C9H9Br1O1)]\\@

**Pentadecane ground state**

1\1\ASN_X86_64-DMC24\FOpt\RB3LYP\6-311+G(2d,p)\C9H9Br1O1\AUBOXA\27-Aug
-2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,noeigen,z-matrix) scrf=(pcm
,solvent=n-Pentadecane)\\Claisen rearrangement allyl p-Br ether GS\\0,
1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B5,2,
A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0\C,9
,B9,7,A8,6,D7,0\H,3,B10,2,A9,1,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B12,3,A11
,2,D10,0\H,5,B13,4,A12,3,D11,0\H,5,B14,4,A13,3,D12,0\H,6,B15,1,A14,2,D
13,0\H,7,B16,6,A15,1,D14,0\H,10,B17,9,A16,7,D15,0\H,8,B18,1,A17,2,D16,
0\Br,9,B19,7,A18,6,D17,0\\B1=1.36474334\B2=1.43466983\B3=1.50053309\B4
=1.32668281\B5=1.39821839\B6=1.38535657\B7=1.39461651\B8=1.39011319\B9
=1.38367652\B10=1.09394619\B11=1.09210815\B12=1.08675244\B13=1.0831396
2\B14=1.08561209\B15=1.08274616\B16=1.08168286\B17=1.08182286\B18=1.07
954733\B19=1.92181734\A1=120.00319776\A2=113.00501632\A3=124.43708507\
A4=115.33946782\A5=120.60077018\A6=125.20647035\A7=119.4131649\A8=120.
70244069\A9=110.7356156\A10=103.38861166\A11=115.21174182\A12=121.5254
6799\A13=121.73257024\A14=118.8524086\A15=120.01326242\A16=120.4974605
6\A17=121.31803072\A18=119.55639601\D1=77.83883208\D2=-138.13755389\D3
=180.44818612\D4=-178.94630834\D5=1.23138915\D6=-0.04506143\D7=-0.1540
5507\D8=-48.33768542\D9=-163.49710206\D10=44.32813407\D11=-176.7275472
9\D12=2.7502487\D13=0.79524209\D14=179.86853269\D15=-179.65272724\D16=
-2.0150264\D17=179.89345986\\Version=AM64L-G09RevA.02\State=1-A\HF=-29
97.8335078\RMSD=7.073e-09\RMSF=3.766e-05\Dipole=0.6803677,0.0853904,1.
1433377\Quadrupole=4.1758738,-2.8843544,-1.2915193,-0.1181244,2.496035
7,0.2835535\PG=C01 [X(C9H9Br1O1)]\\@

**<u>Gas-Phase</u> B3LYP/6-311+G(2d,p) allyl *p*-Br-phenyl ether optimizations:**

**Transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -355.0751 | 67.2298 | 109.1287 |
| Red. masses -- | 8.1302 | 6.1084 | 3.0787 |
| Frc consts -- | 0.6039 | 0.0163 | 0.0216 |
| IR Inten -- | 49.2709 | 0.4766 | 0.3917 |

Sum of electronic and zero-point Energies= -2997.623061

Sum of electronic and thermal Energies= -2997.613072
Sum of electronic and thermal Enthalpies= -2997.612127
Sum of electronic and thermal Free Energies= -2997.659551

1\1\ASN_X86_64-DMC35\FTS\RB3LYP\6-311+G(2d,p)\C9H9Br1O1\AUBOXA\27-Aug-2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,ts,noeigen,z-matrix)\\Claisen rearrangement allyl p-Br ether TS\\0,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B5,2,A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0\C,8,B9,1,A8,2,D7,0\H,3,B10,2,A9,1,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B12,3,A11,2,D10,0\H,5,B13,4,A12,3,D11,0\H,5,B14,4,A13,3,D12,0\H,6,B15,1,A14,2,D13,0\H,7,B16,6,A15,1,D14,0\H,10,B17,8,A16,1,D15,0\H,8,B18,1,A17,2,D16,0\Br,9,B19,7,A18,6,D17,0\\B1=1.26834089\B2=2.19580098\B3=1.37254073\B4=1.39194198\B5=1.44252276\B6=1.42131359\B7=1.43875713\B8=1.36448128\B9=1.36654107\B10=1.08059272\B11=1.08103673\B12=1.08526033\B13=1.08173828\B14=1.08236152\B15=1.08108807\B16=1.08237746\B17=1.08254664\B18=1.08350725\B19=1.91798654\A1=104.59729895\A2=97.00936554\A3=121.19728553\A4=121.36942419\A5=120.48737789\A6=121.95563665\A7=119.44715677\A8=121.45425629\A9=86.00832448\A10=97.73049582\A11=118.548946\A12=120.34043131\A13=120.28481574\A14=117.27826978\A15=119.81708714\A16=120.53839567\A17=117.63019748\A18=120.27272252\D1=52.72100661\D2=-68.51842408\D3=-75.93256424\D4=-175.46432555\D5=98.48769824\D6=-7.92153019\D7=178.50302872\D8=-68.22425662\D9=175.71914358\D10=96.09956673\D11=178.68126488\D12=-24.4706445\D13=-16.7349595\D14=172.96752182\D15=-177.48101635\D16=1.65645046\D17=-178.32098333\\Version=AM64L-G09RevA.02\State=1-A\HF=-2997.7760122\RMSD=6.691e-09\RMSF=3.159e-05\Dipole=0.8655578,0.2407995,0.0857266\Quadrupole=-1.290986,5.7454157,-4.4544296,1.4347849,4.2197481,0.7029945\PG=C01 [X(C9H9Br1O1)]\\@

**Ground state**

1\1\ASN_X86_64-DMC14\FOpt\RB3LYP\6-311+G(2d,p)\C9H9Br1O1\AUBOXA\27-Aug-2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,noeigen,z-matrix)\\Claisen rearrangement allyl p-Br ether GS\\0,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,B5,2,A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,0\C,9,B9,7,A8,6,D7,0\H,3,B10,2,A9,1,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B12,3,A11,2,D10,0\H,5,B13,4,A12,3,D11,0\H,5,B14,4,A13,3,D12,0\H,6,B15,1,A14,2,D13,0\H,7,B16,6,A15,1,D14,0\H,10,B17,9,A16,7,D15,0\H,8,B18,1,A17,2,D16,0\Br,9,B19,7,A18,6,D17,0\\B1=1.36561261\B2=1.43124701\B3=1.50109273\B4=1.3263419\B5=1.39769506\B6=1.3850398\B7=1.39412169\B8=1.39025079\B9=1.38355996\B10=1.09459023\B11=1.09250545\B12=1.08671313\B13=1.08302121\B14=1.08565642\B15=1.08270337\B16=1.08168793\B17=1.08183943\B18=1.07965043\B19=1.91932311\A1=119.99293701\A2=113.15849362\A3=124.49379527\A4=115.3113758\A5=120.57258884\A6=125.22301557\A7=119.49979246\A8=120.58105786\A9=110.79902126\A10=103.44276308\A11=115.13587152\A12=121.57248956\A13=121.73796404\A14=118.76014662\A15=120.0628763\A16=120.37571061\A17=121.27879184\A18=119.6130181\D1=78.29204969\D2=-139.75105215\D3=179.01399582\D4=-179.07991549\D5=-0.34843367\D6=-0.01941932\D7=-0.18376555\D8=-47.89057551\D9=-163.0592155\D10=42.74804636\D11=-176.67416383\D12=2.79409845\D13=0.6458598\D14=179.86285505\D15=-179.60542264\D16=-2.09262345\D17=179.87763633\\Version=AM64L-G09RevA.02\State=1-A\HF=-2997.8313686\RMSD=3.146e-09\RMSF=2.790e-05\Dipole=0.6039934,0.0855082,1.0319767\Quadrupole=3.8874641,-2.727072,-1.1603921,0.0548529,2.2929613,0.3048677\PG=C01 [X(C9H9Br1O1)]\\@

**<u>Gas-Phase</u> PDDG/PM3 allyl *p*-Br-phenyl ether optimizations:**

**Transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -950.1906 | 70.5923 | 115.0685 |
| Red. masses -- | 10.4435 | 5.8485 | 3.6053 |
| Frc consts -- | 5.5554 | 0.0172 | 0.0281 |
| IR Inten -- | 133.8776 | 0.2422 | 1.3392 |

| | |
|---|---|
| Sum of electronic and zero-point Energies= | 0.239501 |
| Sum of electronic and thermal Energies= | 0.249056 |
| Sum of electronic and thermal Enthalpies= | 0.250001 |
| Sum of electronic and thermal Free Energies= | 0.203559 |


1\1\ ASN_ALTIX-ALTIX9\FTS\RPDDG\ZDO\C9H9Br1O1\AUBOXA\10-Aug-2009\1\\#
PDDG opt=(ts,noeigen,z-matrix)\\Claisen rearrangement allyl p-Br ether
TS\\0,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,
1,B5,2,A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D
6,0\C,8,B9,1,A8,2,D7,0\H,3,B10,2,A9,1,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B1
2,3,A11,2,D10,0\H,5,B13,4,A12,3,D11,0\H,5,B14,4,A13,3,D12,0\H,6,B15,1,
A14,2,D13,0\H,7,B16,6,A15,1,D14,0\H,10,B17,8,A16,1,D15,0\H,8,B18,1,A17
,2,D16,0\Br,9,B19,7,A18,6,D17,0\\B1=1.27903124\B2=1.85763302\B3=1.3989
3352\B4=1.40338335\B5=1.44434623\B6=1.44273511\B7=1.45083612\B8=1.3523
8165\B9=1.35890029\B10=1.10440995\B11=1.10025411\B12=1.09992149\B13=1.
09525744\B14=1.09888166\B15=1.11083934\B16=1.1007642\B17=1.10125309\B1
8=1.10178818\B19=1.8212154\A1=107.53291818\A2=99.39560032\A3=118.96575
642\A4=119.50446896\A5=117.57941973\A6=120.39700351\A7=119.55923627\A8
=120.17825383\A9=100.23761629\A10=101.79681962\A11=119.92890004\A12=11
8.39709344\A13=119.72588108\A14=118.42657343\A15=119.33510671\A16=121.
97401779\A17=119.06482377\A18=119.75063404\D1=56.34806592\D2=-70.17935
692\D3=-70.44954831\D4=-181.17999284\D5=100.89078784\D6=-6.46059222\D7
=183.49887377\D8=-66.1347095\D9=177.93854048\D10=95.88895238\D11=181.4
7340321\D12=-33.67661072\D13=-34.9442034\D14=174.82226297\D15=-178.109
24653\D16=5.71496938\D17=-178.23653737\\Version=IA64L-G09RevA.01\State
=1-A\HF=0.0882281\RMSD=7.788e-09\RMSF=2.901e-05\Dipole=0.6531671,0.108
4938,-0.085613\PG=C01 [X(C9H9Br1O1)]\\@


**Ground state**


1\1\ ASN_ALTIX-ALTIX9\FOpt\RPDDG\ZDO\C9H9Br1O1\AUBOXA\10-Aug-2009\1\\#
PDDG opt=(noeigen,z-matrix)\\Claisen rearrangement allyl p-Br ether G
S\\0,1\C\O,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,1,
B5,2,A4,3,D3,0\C,6,B6,1,A5,2,D4,0\C,1,B7,2,A6,3,D5,0\C,7,B8,6,A7,1,D6,
0\C,9,B9,7,A8,6,D7,0\H,3,B10,2,A9,1,D8,0\H,3,B11,2,A10,1,D9,0\H,4,B12,
3,A11,2,D10,0\H,5,B13,4,A12,3,D11,0\H,5,B14,4,A13,3,D12,0\H,6,B15,1,A1
4,2,D13,0\H,7,B16,6,A15,1,D14,0\H,10,B17,9,A16,7,D15,0\H,8,B18,1,A17,2
,D16,0\Br,9,B19,7,A18,6,D17,0\\B1=1.37269636\B2=1.41157326\B3=1.501428
73\B4=1.32997675\B5=1.40816277\B6=1.38897231\B7=1.40049683\B8=1.387141
93\B9=1.38323989\B10=1.12184461\B11=1.10669205\B12=1.10053817\B13=1.08
722416\B14=1.09071895\B15=1.10232923\B16=1.10068311\B17=1.10052123\B18
=1.10436352\B19=1.81946775\A1=118.41563308\A2=111.95032082\A3=121.1048

3407\A4=113.49601299\A5=118.97517514\A6=125.19536914\A7=118.97740122\A
8=122.71728937\A9=114.51077731\A10=99.09816052\A11=116.93595105\A12=12
2.11485216\A13=122.64653999\A14=121.51433836\A15=120.95842095\A16=120.
19283217\A17=121.93052218\A18=118.56653955\D1=80.69378688\D2=-130.6589
686\D3=173.40296238\D4=-179.92654553\D5=-6.7344627\D6=0.12658186\D7=-0
.31033367\D8=-48.56498302\D9=-162.71207628\D10=49.44383275\D11=-178.32
981507\D12=2.05171791\D13=0.17828916\D14=180.03392167\D15=-179.6623434
7\D16=-1.30074565\D17=179.8322469\\Version=IA64L-G09RevA.01\State=1-A\
HF=0.0066063\RMSD=2.825e-09\RMSF=3.095e-05\Dipole=0.5488985,0.0053084,
0.6450878\PG=C01 [X(C9H9Br1O1)]\\@

**<u>Gas-Phase</u> CBS-QB3 allyl *p*-Br-phenyl ether optimizations:**

**Transition structure**

Temperature= 298.150000 Pressure=
1.000000
E(ZPE)= 0.151344 E(Thermal)=
0.161415
E(SCF)= -2993.294873 DE(MP2)= -
2.058238
DE(CBS)= -0.249698 DE(MP34)= -
0.053838
DE(CCSD)= -0.075530 DE(Int)=
0.067516
DE(Empirical)= -0.110685
CBS-QB3 (0 K)= -2995.624002 CBS-QB3 Energy= -
2995.613931
CBS-QB3 Enthalpy= -2995.612987 CBS-QB3 Free Energy= -
2995.660553
1\1\ ASN_ALTIX-ALTIX7\Mixed\CBS-QB3\CBS-QB3\C9H9Br1O1\AUBOXA\12-Aug-20
09\0\\# CBS-QB3 opt=(calcfc,ts,noeigen,z-matrix)\\Claisen rearrangemen
t allyl p-Br ether TS\\0,1\C,0,0.,0.,0.\O,0,0.,0.,1.2672\C,0,2.1134244
93,0.,1.8316405747\C,0,2.4925855449,1.0867392093,1.0784904983\C,0,2.50
40547699,1.0308291271,-0.3147409539\C,0,0.3058320541,1.1962150042,-0.7
53206551\C,0,0.210819422,1.2013545598,-2.1739237624\C,0,-0.1763817341,
-1.2098658726,-0.7647002447\C,0,-0.0074262038,0.0239904117,-2.83523527
73\C,0,-0.2003857189,-1.19603927,-2.1328198279\H,0,2.1509227314,-1.001
9266763,1.4262036901\H,0,1.9811901967,0.0777644668,2.9024074348\H,0,2.
5344059275,2.0603987942,1.5574653185\H,0,2.7842410486,1.9010325008,-0.
8946241297\H,0,2.6130999685,0.0798960545,-0.8220616868\H,0,0.250387006
6,2.1381981144,-0.2241650529\H,0,0.3269477823,2.1289889211,-2.72044101
34\H,0,-0.3756974925,-2.106501393,-2.6926681239\H,0,-0.3467000999,-2.1
27589077,-0.2130976575\Br,0,-0.0897036219,-0.0056738126,-4.7511399905\
\Version=IA64L-G09RevA.01\State=1-A\HF/CbsB3=-2993.294873\E2(CBS)/CbsB
3=-2.3079355\CBS-Int/CbsB3=0.0675157\OIii/CbsB3=19.1166565\MP2/CbsB4=-
2992.4211232\MP4(SDQ)/CbsB4=-2992.4749615\MP4(SDQ)/6-31+G(d')=-2992.36
00703\QCISD(T)/6-31+G(d')=-2992.4356002\CBSQB3=-2995.624002\

**Ground State**

Temperature= 298.150000 Pressure=
1.000000
E(ZPE)= 0.153808 E(Thermal)=
0.164407

```
 E(SCF)=               -2993.384619 DE(MP2)=                        -
2.020527
 DE(CBS)=                 -0.247497 DE(MP34)=                       -
0.068311
 DE(CCSD)=                -0.065804 DE(Int)=
0.066573
 DE(Empirical)=           -0.111710
 CBS-QB3 (0 K)=        -2995.678088 CBS-QB3 Energy=             -
2995.667488
 CBS-QB3 Enthalpy=     -2995.666544 CBS-QB3 Free Energy=        -
2995.716451
 1\1\ ASN_ALTIX-ALTIX7\Mixed\CBS-QB3\CBS-QB3\C9H9Br1O1\AUBOXA\12-Aug-20
 09\0\\# CBS-QB3 opt=(calcfc,noeigen,z-matrix)\\Claisen rearrangement a
 llyl p-Br ether GS\\0,1\C\O,1,1.3651048335\C,2,1.431943158,1,120.06512
 443\C,3,1.5035526785,2,113.13918604,1,78.77274033,0\C,4,1.3285681261,3
 ,124.51143641,2,-140.11155983,0\C,1,1.4004849131,2,115.23826984,3,177.
 05963781,0\C,6,1.3869087797,1,120.63496178,2,-179.26136323,0\C,1,1.396
 9966298,2,125.29664251,3,-2.45859785,0\C,7,1.3936973749,6,119.4010917,
 1,0.02266411,0\C,9,1.3871145128,7,120.65629969,6,-0.20131157,0\H,3,1.0
 960410243,2,110.85516515,1,-47.36476542,0\H,3,1.0932166942,2,103.42125
 823,1,-162.61710381,0\H,4,1.087317926,3,114.96638065,2,42.4855088,0\H,
 5,1.0837021397,4,121.6424912,3,-176.62515742,0\H,5,1.0864286108,4,121.
 75121562,3,2.83416311,0\H,6,1.0830837415,1,118.46794429,2,0.46310355,0
 \H,7,1.0821962991,6,120.20732335,1,179.86414235,0\H,10,1.0823459889,9,
 120.30562322,7,-179.58185585,0\H,8,1.0802779082,1,121.2102098,2,-2.026
 3367,0\Br,9,1.9190826964,7,119.57162019,6,179.85473008,0\\Version=IA64
 L-G09RevA.01\State=1-A\HF/CbsB3=-2993.3846189\E2(CBS)/CbsB3=-2.2680243
 \CBS-Int/CbsB3=0.0665729\OIii/CbsB3=19.2936315\MP2/CbsB4=-2992.4712895
 \MP4(SDQ)/CbsB4=-2992.5396001\MP4(SDQ)/6-31+G(d')=-2992.4236993\QCISD(
 T)/6-31+G(d')=-2992.4895037\CBSQB3=-2995.6780877\
```

**<u>Condensed-Phase</u> B3LYP/6-311+G(2d,p) allyl napthyl ether optimizations in PCM:**

**Water transition structure**

```
                     1                      2                      3
                     A                      A                      A
 Frequencies --  -163.2676               39.6816
64.2228
 Red. masses --     6.1512                3.6995
4.6621
 Frc consts  --     0.0966                0.0034
0.0113
 IR Inten    --    43.3667                0.0331
2.4457

 Sum of electronic and zero-point Energies=          -1115.966357
 Sum of electronic and thermal Energies=             -1115.950575
 Sum of electronic and thermal Enthalpies=           -1115.949631
 Sum of electronic and thermal Free Energies=        -1116.009390
```

1\1\ ASN_ALTIX-ALTIX7\FTS\RB3LYP\6-311+G(2d,p)\C15H15Cl1O1\AUBOXA\16-A
ug-2009\1\\# B3LYP/6-311+G(2d,p) opt=(TS,noeigen,z-matrix) scrf=(pcm,s
olvent=water)\\Allyl Naphthyl Ether TS\\0,1\O\C,1,B1\C,2,B2,1,A1\C,3,B
3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,5,B5,4,A4,3,D3,0\C,6,B6,5,A5,4,D4,0
\C,7,B7,6,A6,5,D5,0\C,8,B8,7,A7,6,D6,0\C,9,B9,8,A8,7,D7,0\C,10,B10,9,A
9,8,D8,0\Cl,5,B11,4,A10,3,D9,0\C,1,B12,2,A11,3,D10,0\C,3,B13,2,A12,1,D
11,0\C,13,B14,1,A13,2,D12,0\C,13,B15,1,A14,2,D13,0\C,13,B16,1,A15,2,D1
4,0\H,3,B17,2,A16,1,D15,0\H,4,B18,3,A17,2,D16,0\H,8,B19,7,A18,6,D17,0\
H,11,B20,10,A19,9,D18,0\H,9,B21,8,A20,7,D19,0\H,10,B22,9,A21,8,D20,0\H
,14,B23,3,A22,2,D21,0\H,14,B24,3,A23,2,D22,0\H,15,B25,13,A24,1,D23,0\H
,16,B26,13,A25,1,D24,0\H,16,B27,13,A26,1,D25,0\H,16,B28,13,A27,1,D26,0
\H,17,B29,13,A28,1,D27,0\H,17,B30,13,A29,1,D28,0\H,17,B31,13,A30,1,D29
,0\\B1=1.2720034\B2=1.42104018\B3=1.41440918\B4=1.36233612\B5=1.432115
22\B6=1.42595342\B7=1.40611811\B8=1.37828642\B9=1.40447772\B10=1.37750
302\B11=1.76788403\B12=2.36474616\B13=2.51632827\B14=1.39712291\B15=1.
49197248\B16=1.49362807\B17=1.0803117\B18=1.08256391\B19=1.0819912\B20
=1.08148992\B21=1.08338053\B22=1.08352363\B23=1.08131466\B24=1.0804785
2\B25=1.0853814\B26=1.09175002\B27=1.09686129\B28=1.08917543\B29=1.097
24182\B30=1.0892542\B31=1.08769255\A1=121.71438978\A2=121.47223705\A3=
120.61152005\A4=122.10852866\A5=117.91509114\A6=118.97694391\A7=121.30
180961\A8=119.75033734\A9=120.383283\A10=118.40006289\A11=109.08044405
\A12=90.29039621\A13=89.91213086\A14=96.86211964\A15=96.22662021\A16=1
18.22844662\A17=119.40507228\A18=117.86714754\A19=119.85198037\A20=120
.24148997\A21=119.93281524\A22=90.10440622\A23=92.58575038\A24=116.689
38499\A25=110.28721906\A26=109.51244396\A27=112.29294759\A28=108.23754
053\A29=110.9805966\A30=113.54062392\D1=182.11426444\D2=-3.95533042\D3
=1.74546961\D4=-0.15257301\D5=-181.76207927\D6=0.94854199\D7=-0.248316
31\D8=-0.36133455\D9=181.0676969\D10=-77.74111126\D11=72.67603122\D12=
52.56367431\D13=172.24977632\D14=-70.55597266\D15=-10.50657124\D16=176
.49249419\D17=180.9509489\D18=-179.4241663\D19=-180.21230269\D20=-180.
02312643\D21=-179.04748726\D22=64.83795623\D23=93.2240024\D24=43.22540
486\D25=160.40462532\D26=-78.91826627\D27=-179.45601098\D28=-61.971427
08\D29=61.21088961\\Version=IA64L-G09RevA.01\State=1-A\HF=-1116.221251
6\RMSD=9.978e-09\RMSF=4.534e-05\Dipole=0.3968784,1.8311218,-1.1561514\
Quadrupole=7.5209148,-4.8508094,-2.6701054,1.7235133,-1.8306177,-5.918
2745\PG=C01 [X(C15H15Cl1O1)]\\@

**Water ground state**

1\1\ ASN_ALTIX-ALTIX7\FOpt\RB3LYP\6-311+G(2d,p)\C15H15Cl1O1\AUBOXA\16-
Aug-2009\1\\# B3LYP/6-311+G(2d,p) opt=(noeigen,z-matrix) scrf=(pcm,sol
vent=water)\\Allyl Naphthyl Ether GS\\0,1\O\C,1,B1\C,2,B2,1,A1\C,3,B3,
2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,5,B5,4,A4,3,D3,0\C,6,B6,5,A5,4,D4,0\C
,7,B7,6,A6,5,D5,0\C,8,B8,7,A7,6,D6,0\C,9,B9,8,A8,7,D7,0\C,10,B10,9,A9,
8,D8,0\Cl,5,B11,4,A10,3,D9,0\C,1,B12,2,A11,3,D10,0\C,3,B13,2,A12,1,D11
,0\C,13,B14,1,A13,2,D12,0\C,13,B15,1,A14,2,D13,0\C,13,B16,1,A15,2,D14,
0\H,3,B17,2,A16,1,D15,0\H,4,B18,3,A17,2,D16,0\H,8,B19,7,A18,6,D17,0\H,
11,B20,10,A19,9,D18,0\H,9,B21,8,A20,7,D19,0\H,10,B22,9,A21,8,D20,0\H,1
4,B23,3,A22,2,D21,0\H,14,B24,3,A23,2,D22,0\H,15,B25,13,A24,1,D23,0\H,1
6,B26,13,A25,1,D24,0\H,16,B27,13,A26,1,D25,0\H,16,B28,13,A27,1,D26,0\H
,17,B29,13,A28,1,D27,0\H,17,B30,13,A29,1,D28,0\H,17,B31,13,A30,1,D29,0
\\B1=1.35826255\B2=1.37661954\B3=1.40993335\B4=1.36539524\B5=1.4205963
5\B6=1.43138618\B7=1.41569057\B8=1.37337873\B9=1.40889286\B10=1.372867
54\B11=1.76792058\B12=1.4686919\B13=3.75526458\B14=1.51348064\B15=1.53
465215\B16=1.52798368\B17=1.07759515\B18=1.0820264\B19=1.08080775\B20=
1.08140342\B21=1.08326121\B22=1.0833647\B23=1.08342935\B24=1.08360252\

B25=1.0873145\B26=1.09159755\B27=1.09021138\B28=1.09104387\B29=1.09072
726\B30=1.09091551\B31=1.08863717\A1=126.08222886\A2=120.46038249\A3=1
20.81164664\A4=121.43728776\A5=117.80716954\A6=118.80098637\A7=121.086
42948\A8=120.11461053\A9=120.34634205\A10=118.55547827\A11=124.5597800
4\A12=94.33907\A13=109.76511181\A14=102.08875712\A15=110.44106874\A16=
121.08487888\A17=119.05453834\A18=118.57052708\A19=120.0047636\A20=120
.01080448\A21=119.84998301\A22=118.41696713\A23=99.74517341\A24=113.72
100916\A25=110.01082251\A26=109.88571413\A27=110.59287532\A28=109.8663
7306\A29=109.30806646\A30=112.90091557\D1=180.89268863\D2=0.13476096\D
3=0.02091225\D4=-0.1797032\D5=-179.83812226\D6=-0.08131355\D7=0.005755
03\D8=0.06215856\D9=179.96729777\D10=-9.55258188\D11=-13.29865854\D12=
64.59037853\D13=179.35004699\D14=-63.74540731\D15=-0.11159944\D16=179.
96304722\D17=179.88191986\D18=-180.0447889\D19=-180.04954932\D20=-179.
97565312\D21=-122.47315518\D22=109.9010678\D23=39.81481271\D24=59.0860
6511\D25=179.05253459\D26=-61.00971732\D27=-170.60946728\D28=-52.34072
067\D29=68.62372124\\Version=IA64L-G09RevA.01\State=1-A\HF=-1116.25573
06\RMSD=5.862e-09\RMSF=2.282e-05\Dipole=0.7049765,0.2028748,-1.6083601
\Quadrupole=8.5983765,-4.9059154,-3.6924611,0.6926926,-3.2344287,-1.23
7381\PG=C01 [X(C15H15Cl1O1)]\\@

**Acetonitrile transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -164.7798 | 39.5560 | 64.3441 |
| Red. masses -- | 6.1727 | 3.6958 | 4.6549 |
| Frc consts -- | 0.0987 | 0.0034 | 0.0114 |
| IR Inten -- | 42.9371 | 0.0341 | 2.3489 |

| | |
|---|---|
| Sum of electronic and zero-point Energies= | -1115.966045 |
| Sum of electronic and thermal Energies= | -1115.950272 |
| Sum of electronic and thermal Enthalpies= | -1115.949328 |
| Sum of electronic and thermal Free Energies= | -1116.009056 |

1\1\ASN_X86_64-DMC14\FTS\RB3LYP\6-311+G(2d,p)\C15H15Cl1O1\AUBOXA\15-Au
g-2009\1\\# B3LYP/6-311+G(2d,p) opt=(TS,noeigen,z-matrix) scrf=(pcm,so
lvent=Acetonitrile)\\Allyl Naphthyl Ether TS\\0,1\O\C,1,B1\C,2,B2,1,A1
\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,5,B5,4,A4,3,D3,0\C,6,B6,5,A5,
4,D4,0\C,7,B7,6,A6,5,D5,0\C,8,B8,7,A7,6,D6,0\C,9,B9,8,A8,7,D7,0\C,10,B
10,9,A9,8,D8,0\Cl,5,B11,4,A10,3,D9,0\C,1,B12,2,A11,3,D10,0\C,3,B13,2,A
12,1,D11,0\C,13,B14,1,A13,2,D12,0\C,13,B15,1,A14,2,D13,0\C,13,B16,1,A1
5,2,D14,0\H,3,B17,2,A16,1,D15,0\H,4,B18,3,A17,2,D16,0\H,8,B19,7,A18,6,
D17,0\H,11,B20,10,A19,9,D18,0\H,9,B21,8,A20,7,D19,0\H,10,B22,9,A21,8,D
20,0\H,14,B23,3,A22,2,D21,0\H,14,B24,3,A23,2,D22,0\H,15,B25,13,A24,1,D
23,0\H,16,B26,13,A25,1,D24,0\H,16,B27,13,A26,1,D25,0\H,16,B28,13,A27,1
,D26,0\H,17,B29,13,A28,1,D27,0\H,17,B30,13,A29,1,D28,0\H,17,B31,13,A30
,1,D29,0\\B1=1.27180757\B2=1.42117392\B3=1.41446711\B4=1.3622862\B5=1.
43219374\B6=1.42588068\B7=1.40607178\B8=1.37827459\B9=1.40443471\B10=1
.37749142\B11=1.76774061\B12=2.36135218\B13=2.51415351\B14=1.39711743\
B15=1.49204677\B16=1.49374262\B17=1.08030943\B18=1.0825641\B19=1.08199
339\B20=1.08147957\B21=1.08338698\B22=1.08352842\B23=1.0813248\B24=1.0
8046809\B25=1.08539156\B26=1.09175102\B27=1.09685616\B28=1.08917666\B2
9=1.09724928\B30=1.08925023\B31=1.08770058\A1=121.70908676\A2=121.4680

2073\A3=120.61334773\A4=122.10869097\A5=117.91726635\A6=118.98090923\A
7=121.29638055\A8=119.75238277\A9=120.38394326\A10=118.3988852\A11=109
.12155023\A12=90.30010762\A13=89.97760201\A14=96.87779932\A15=96.20403
67\A16=118.20939847\A17=119.41205572\A18=117.86086726\A19=119.8582284\
A20=120.24028651\A21=119.93298977\A22=90.1665482\A23=92.53518335\A24=1
16.68300171\A25=110.2724907\A26=109.53086903\A27=112.28783946\A28=108.
25661788\A29=110.98049973\A30=113.53110551\D1=182.06202585\D2=-3.96881
367\D3=1.79224266\D4=-0.17871556\D5=-181.81366212\D6=0.9722631\D7=-0.2
4931199\D8=-0.3756094\D9=181.09346617\D10=-77.72834614\D11=72.62338127
\D12=52.58430118\D13=172.28632217\D14=-70.52646868\D15=-10.64327178\D1
6=176.47390417\D17=181.00193595\D18=-179.39956607\D19=-180.19612486\D2
0=-180.01920574\D21=-179.0136373\D22=64.88165022\D23=93.24339247\D24=4
3.04917008\D25=160.25320111\D26=-79.04887927\D27=-179.61350588\D28=-62
.10173495\D29=61.04947499\\Version=AM64L-G09RevA.01\State=1-A\HF=-1116
.2209591\RMSD=2.559e-09\RMSF=4.447e-05\Dipole=0.3882306,1.7962906,-1.1
473368\Quadrupole=7.4876912,-4.8441528,-2.6435384,1.6854238,-1.8114944
,-5.8277695\PG=C01 [X(C15H15Cl1O1)]\\@

**Acetonitrile ground state**

1\1\ ASN_ALTIX-ALTIX8\FOpt\RB3LYP\6-311+G(2d,p)\C15H15Cl1O1\AUBOXA\23-
Aug-2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,noeigen,z-matrix) scrf=(
pcm,solvent=Acetonitrile)\\Allyl Naphthyl Ether TS\\0,1\O\C,1,B1\C,2,B
2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,5,B5,4,A4,3,D3,0\C,6,B6
,5,A5,4,D4,0\C,7,B7,6,A6,5,D5,0\C,8,B8,7,A7,6,D6,0\C,9,B9,8,A8,7,D7,0\
C,10,B10,9,A9,8,D8,0\Cl,5,B11,4,A10,3,D9,0\C,1,B12,2,A11,3,D10,0\C,3,B
13,2,A12,1,D11,0\C,13,B14,1,A13,2,D12,0\C,13,B15,1,A14,2,D13,0\C,13,B1
6,1,A15,2,D14,0\H,3,B17,2,A16,1,D15,0\H,4,B18,3,A17,2,D16,0\H,8,B19,7,
A18,6,D17,0\H,11,B20,10,A19,9,D18,0\H,9,B21,8,A20,7,D19,0\H,10,B22,9,A
21,8,D20,0\H,14,B23,3,A22,2,D21,0\H,14,B24,3,A23,2,D22,0\H,15,B25,13,A
24,1,D23,0\H,16,B26,13,A25,1,D24,0\H,16,B27,13,A26,1,D25,0\H,16,B28,13
,A27,1,D26,0\H,17,B29,13,A28,1,D27,0\H,17,B30,13,A29,1,D28,0\H,17,B31,
13,A30,1,D29,0\\B1=1.35836647\B2=1.37657709\B3=1.40993477\B4=1.3653722
4\B5=1.4205883\B6=1.43134841\B7=1.41566144\B8=1.37334314\B9=1.40886811
\B10=1.37283553\B11=1.76780574\B12=1.46833953\B13=3.75543315\B14=1.513
51335\B15=1.53468275\B16=1.52803956\B17=1.07759052\B18=1.08202205\B19=
1.08079841\B20=1.08139452\B21=1.08326572\B22=1.08336803\B23=1.08342419
\B24=1.08360742\B25=1.08730706\B26=1.09159256\B27=1.090209\B28=1.09103
298\B29=1.09073237\B30=1.09090698\B31=1.0886355\A1=126.08008809\A2=120
.45656034\A3=120.81474716\A4=121.4326657\A5=117.81011043\A6=118.802843
39\A7=121.08445061\A8=120.11535832\A9=120.34692298\A10=118.55586773\A1
1=124.55423043\A12=94.33883105\A13=109.76852939\A14=102.09682082\A15=1
10.44622885\A16=121.0852835\A17=119.06035905\A18=118.56590801\A19=120.
01059158\A20=120.00988349\A21=119.85044828\A22=118.38057835\A23=99.826
95908\A24=113.71451137\A25=110.00547891\A26=109.89167304\A27=110.59110
044\A28=109.87321433\A29=109.30828804\A30=112.89580058\D1=180.90130502
\D2=0.12912152\D3=0.02343338\D4=-0.17827476\D5=-179.84444746\D6=-0.079
26551\D7=0.00639439\D8=0.06012606\D9=179.96928289\D10=-9.50103374\D11=
-13.33203654\D12=64.53579362\D13=179.29771438\D14=-63.79514129\D15=-0.
11800411\D16=179.96160144\D17=179.8866803\D18=-180.0428047\D19=-180.04
814071\D20=-179.97599337\D21=-122.47746712\D22=109.86074573\D23=39.730
546\D24=59.05026671\D25=179.02564525\D26=-61.03446794\D27=-170.6866741
1\D28=-52.41120517\D29=68.53824514\\Version=IA64L-G09RevA.01\State=1-A
\HF=-1116.2555757\RMSD=6.703e-09\RMSF=2.227e-05\Dipole=0.6992562,0.199
4157,-1.6003534\Quadrupole=8.5399315,-4.8687473,-3.6711842,0.6831564,-
3.2136536,-1.2275153\PG=C01 [X(C15H15Cl1O1)]\\@

**n,n-Dimethylformamide transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -164.7498 | 39.6887 | 64.3271 |
| Red. masses -- | 6.1725 | 3.7033 | 4.6541 |
| Frc consts -- | 0.0987 | 0.0034 | 0.0113 |
| IR Inten -- | 42.9461 | 0.0337 | 2.3514 |

Sum of electronic and zero-point Energies= -1115.966064
Sum of electronic and thermal Energies= -1115.950294
Sum of electronic and thermal Enthalpies= -1115.949349
Sum of electronic and thermal Free Energies= -1116.009070


1\1\ ASN_ALTIX-ALTIX7\FTS\RB3LYP\6-311+G(2d,p)\C15H15Cl1O1\AUBOXA\15-Aug-2009\1\\# B3LYP/6-311+G(2d,p) opt=(TS,noeigen,z-matrix) scrf=(pcm,solvent=n,n-DiMethylFormamide)\\Allyl Naphthyl Ether TS\\0,1\O\C,1,B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,5,B5,4,A4,3,D3,0\C,6,B6,5,A5,4,D4,0\C,7,B7,6,A6,5,D5,0\C,8,B8,7,A7,6,D6,0\C,9,B9,8,A8,7,D7,0\C,10,B10,9,A9,8,D8,0\Cl,5,B11,4,A10,3,D9,0\C,1,B12,2,A11,3,D10,0\C,3,B13,2,A12,1,D11,0\C,13,B14,1,A13,2,D12,0\C,13,B15,1,A14,2,D13,0\C,13,B16,1,A15,2,D14,0\H,3,B17,2,A16,1,D15,0\H,4,B18,3,A17,2,D16,0\H,8,B19,7,A18,6,D17,0\H,11,B20,10,A19,9,D18,0\H,9,B21,8,A20,7,D19,0\H,10,B22,9,A21,8,D20,0\H,14,B23,3,A22,2,D21,0\H,14,B24,3,A23,2,D22,0\H,15,B25,13,A24,1,D23,0\H,16,B26,13,A25,1,D24,0\H,16,B27,13,A26,1,D25,0\H,16,B28,13,A27,1,D26,0\H,17,B29,13,A28,1,D27,0\H,17,B30,13,A29,1,D28,0\H,17,B31,13,A30,1,D29,0\\B1=1.27184687\B2=1.42114282\B3=1.41446618\B4=1.36228872\B5=1.43218343\B6=1.4258893\B7=1.40607576\B8=1.37827128\B9=1.4044406\B10=1.37748962\B11=1.76775372\B12=2.36152863\B13=2.51428845\B14=1.39712689\B15=1.49204837\B16=1.49371002\B17=1.08030872\B18=1.08256405\B19=1.08199246\B20=1.08148079\B21=1.08338664\B22=1.08352795\B23=1.08132259\B24=1.0804698\B25=1.08539103\B26=1.09175158\B27=1.09685462\B28=1.0891773\B29=1.09725352\B30=1.08925841\B31=1.08769414\A1=121.70917088\A2=121.46873503\A3=120.61239221\A4=122.10850573\A5=117.91729366\A6=118.98046616\A7=121.29685104\A8=119.75224677\A9=120.38397216\A10=118.39895646\A11=109.11970961\A12=90.28830659\A13=89.97011913\A14=96.8848156\A15=96.21310445\A16=118.21131576\A17=119.41210792\A18=117.86090571\A19=119.85804079\A20=120.240424\A21=119.93278638\A22=90.16028154\A23=92.53596004\A24=116.68282052\A25=110.27479532\A26=109.52862241\A27=112.28781369\A28=108.25158111\A29=110.98018465\A30=113.53588205\D1=182.06914267\D2=-3.97148556\D3=1.7876375\D4=-0.17363168\D5=-181.80814361\D6=0.97093505\D7=-0.25055724\D8=-0.37361598\D9=181.0911206\D10=-77.73476851\D11=72.64185946\D12=52.56790217\D13=172.26612035\D14=-70.54611909\D15=-10.62618105\D16=176.47014746\D17=180.99761991\D18=-179.40184959\D19=-180.19863089\D20=-180.01898573\D21=-179.03288676\D22=64.86168755\D23=93.2517781\D24=43.0630605\D25=160.26710783\D26=-79.03823639\D27=-179.53590818\D28=-62.03339081\D29=61.12487201\\Version=IA64L-G09RevA.01\State=1-A\HF=-1116.2209809\RMSD=5.246e-09\RMSF=4.471e-05\Dipole=0.3886034,1.7990005,-1.1479791\Quadrupole=7.489572,-4.8455539,-2.6440181,1.6875189,-1.8123223,-5.8352525\PG=C01 [X(C15H15Cl1O1)]\\@

**Ground state**


1\1\ASN_X86_64-DMC32\FOpt\RB3LYP\6-311+G(2d,p)\C15H15Cl1O1\AUBOXA\24-A
ug-2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,noeigen,z-matrix) scrf=(p
cm,solvent=n,n-DiMethylFormamide)\\Allyl Naphthyl Ether TS\\0,1\O\C,1,
B1\C,2,B2,1,A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,5,B5,4,A4,3,D3,
0\C,6,B6,5,A5,4,D4,0\C,7,B7,6,A6,5,D5,0\C,8,B8,7,A7,6,D6,0\C,9,B9,8,A8
,7,D7,0\C,10,B10,9,A9,8,D8,0\Cl,5,B11,4,A10,3,D9,0\C,1,B12,2,A11,3,D10
,0\C,3,B13,2,A12,1,D11,0\C,13,B14,1,A13,2,D12,0\C,13,B15,1,A14,2,D13,0
\C,13,B16,1,A15,2,D14,0\H,3,B17,2,A16,1,D15,0\H,4,B18,3,A17,2,D16,0\H,
8,B19,7,A18,6,D17,0\H,11,B20,10,A19,9,D18,0\H,9,B21,8,A20,7,D19,0\H,10
,B22,9,A21,8,D20,0\H,14,B23,3,A22,2,D21,0\H,14,B24,3,A23,2,D22,0\H,15,
B25,13,A24,1,D23,0\H,16,B26,13,A25,1,D24,0\H,16,B27,13,A26,1,D25,0\H,1
6,B28,13,A27,1,D26,0\H,17,B29,13,A28,1,D27,0\H,17,B30,13,A29,1,D28,0\H
,17,B31,13,A30,1,D29,0\\B1=1.35835663\B2=1.37657998\B3=1.40993454\B4=1
.36537325\B5=1.42058861\B6=1.4313509\B7=1.41566362\B8=1.37334558\B9=1.
40886965\B10=1.37283767\B11=1.76781418\B12=1.46835672\B13=3.75524493\B
14=1.51350528\B15=1.53467947\B16=1.52803563\B17=1.07758964\B18=1.08202
213\B19=1.08079874\B20=1.08139499\B21=1.08326509\B22=1.08336746\B23=1.
08342516\B24=1.08360671\B25=1.08730781\B26=1.09159287\B27=1.09020866\B
28=1.09103333\B29=1.09073205\B30=1.09090731\B31=1.08863614\A1=126.0807
6788\A2=120.45683585\A3=120.81460038\A4=121.4329725\A5=117.80988564\A6
=118.80263683\A7=121.08464117\A8=120.11531859\A9=120.3468409\A10=118.5
5586117\A11=124.55674049\A12=94.34004045\A13=109.7676575\A14=102.09652
858\A15=110.44616063\A16=121.08528502\A17=119.05987732\A18=118.5662530
2\A19=120.01013262\A20=120.00994536\A21=119.85042246\A22=118.3833638\A
23=99.8147472\A24=113.71480797\A25=110.00600096\A26=109.89135645\A27=1
10.59099884\A28=109.87226463\A29=109.30810677\A30=112.89666991\D1=180.
90137406\D2=0.12976179\D3=0.02269221\D4=-0.17793412\D5=-179.84368684\D
6=-0.07980557\D7=0.00625504\D8=0.06070588\D9=179.96872586\D10=-9.50358
745\D11=-13.33067218\D12=64.53804679\D13=179.29997273\D14=-63.79263048
\D15=-0.11698402\D16=179.96182859\D17=179.88556472\D18=-180.04326272\D
19=-180.04833694\D20=-179.97567416\D21=-122.47544557\D22=109.87065635\
D23=39.74818494\D24=59.05945135\D25=179.03414306\D26=-61.02603938\D27=
-170.6721782\D28=-52.39741026\D29=68.55332946\\Version=AM64L-G09RevA.0
1\State=1-A\HF=-1116.2555872\RMSD=5.294e-09\RMSF=2.251e-05\Dipole=0.69
96979,0.1996664,-1.6009882\Quadrupole=8.5442687,-4.871527,-3.6727418,0
.6839074,-3.2146734,-1.2280711\PG=C01 [X(C15H15Cl1O1)]\\@


**Methanol transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -164.9371 | 40.0156 | 64.3469 |
| Red. masses -- | 6.1735 | 3.7120 | 4.6498 |
| Frc consts -- | 0.0990 | 0.0035 | 0.0113 |
| IR Inten -- | 42.9391 | 0.0342 | 2.3475 |

| | |
|---|---|
| Sum of electronic and zero-point Energies= | -1115.965989 |
| Sum of electronic and thermal Energies= | -1115.950222 |
| Sum of electronic and thermal Enthalpies= | -1115.949278 |
| Sum of electronic and thermal Free Energies= | -1116.008979 |

1\1\ ASN_ALTIX-ALTIX7\FTS\RB3LYP\6-311+G(2d,p)\C15H15Cl1O1\AUBOXA\13-A
ug-2009\1\\# B3LYP/6-311+G(2d,p) opt=(TS,noeigen,z-matrix) scrf=(pcm,s
olvent=Methanol)\\Allyl Naphthyl Ether TS\\0,1\O\C,1,B1\C,2,B2,1,A1\C,
3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,5,B5,4,A4,3,D3,0\C,6,B6,5,A5,4,D
4,0\C,7,B7,6,A6,5,D5,0\C,8,B8,7,A7,6,D6,0\C,9,B9,8,A8,7,D7,0\C,10,B10,
9,A9,8,D8,0\Cl,5,B11,4,A10,3,D9,0\C,1,B12,2,A11,3,D10,0\C,3,B13,2,A12,
1,D11,0\C,13,B14,1,A13,2,D12,0\C,13,B15,1,A14,2,D13,0\C,13,B16,1,A15,2
,D14,0\H,3,B17,2,A16,1,D15,0\H,4,B18,3,A17,2,D16,0\H,8,B19,7,A18,6,D17
,0\H,11,B20,10,A19,9,D18,0\H,9,B21,8,A20,7,D19,0\H,10,B22,9,A21,8,D20,
0\H,14,B23,3,A22,2,D21,0\H,14,B24,3,A23,2,D22,0\H,15,B25,13,A24,1,D23,
0\H,16,B26,13,A25,1,D24,0\H,16,B27,13,A26,1,D25,0\H,16,B28,13,A27,1,D2
6,0\H,17,B29,13,A28,1,D27,0\H,17,B30,13,A29,1,D28,0\H,17,B31,13,A30,1,
D29,0\\B1=1.27176287\B2=1.42119431\B3=1.41445498\B4=1.36229502\B5=1.43
219501\B6=1.42587447\B7=1.40606227\B8=1.37827374\B9=1.40443219\B10=1.3
7748764\B11=1.76771357\B12=2.36109245\B13=2.5139205\B14=1.39708013\B15
=1.49206007\B16=1.49377448\B17=1.08030543\B18=1.08256529\B19=1.0819935
4\B20=1.0814795\B21=1.08338888\B22=1.0835304\B23=1.08131998\B24=1.0804
6698\B25=1.08539829\B26=1.091757\B27=1.09685568\B28=1.08917459\B29=1.0
9725127\B30=1.0892507\B31=1.08769575\A1=121.70986744\A2=121.46840749\A
3=120.61417183\A4=122.10773109\A5=117.91778816\A6=118.98165097\A7=121.
29584148\A8=119.75242495\A9=120.38418584\A10=118.39883835\A11=109.1354
5934\A12=90.29115755\A13=90.00757374\A14=96.87087876\A15=96.1666923\A1
6=118.20813507\A17=119.41368832\A18=117.85870883\A19=119.85947431\A20=
120.24037085\A21=119.93288194\A22=90.16111928\A23=92.5435408\A24=116.6
8164168\A25=110.2642373\A26=109.53975664\A27=112.28882951\A28=108.2590
1884\A29=110.97419484\A30=113.53468735\D1=182.07043083\D2=-3.97500656\
D3=1.79046572\D4=-0.17438582\D5=-181.81269199\D6=0.97455794\D7=-0.2521
3516\D8=-0.3745001\D9=181.09319395\D10=-77.68499568\D11=72.65357133\D1
2=52.49538663\D13=172.20431417\D14=-70.62540635\D15=-10.62283662\D16=1
76.46461302\D17=181.00778985\D18=-179.39671539\D19=-180.19592261\D20=-
180.01631267\D21=-179.08891505\D22=64.80654889\D23=93.28385056\D24=42.
91927894\D25=160.12420405\D26=-79.16725851\D27=-179.50208721\D28=-61.9
9153217\D29=61.15411945\\Version=IA64L-G09RevA.01\State=1-A\HF=-1116.2
209096\RMSD=6.351e-09\RMSF=4.651e-05\Dipole=0.3878575,1.7907456,-1.145
7633\Quadrupole=7.4835714,-4.8458994,-2.6376721,1.681275,-1.8115134,-5
.8095498\PG=C01 [X(C15H15Cl1O1)]\\@

**Methanol ground state**

1\1\ ASN_ALTIX-ALTIX8\FOpt\RB3LYP\6-311+G(2d,p)\C15H15Cl1O1\AUBOXA\24-
Aug-2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,noeigen,z-matrix) scrf=(
pcm,solvent=Methanol)\\Allyl Naphthyl Ether TS\\0,1\O\C,1,B1\C,2,B2,1,
A1\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,5,B5,4,A4,3,D3,0\C,6,B6,5,A
5,4,D4,0\C,7,B7,6,A6,5,D5,0\C,8,B8,7,A7,6,D6,0\C,9,B9,8,A8,7,D7,0\C,10
,B10,9,A9,8,D8,0\Cl,5,B11,4,A10,3,D9,0\C,1,B12,2,A11,3,D10,0\C,3,B13,2
,A12,1,D11,0\C,13,B14,1,A13,2,D12,0\C,13,B15,1,A14,2,D13,0\C,13,B16,1,
A15,2,D14,0\H,3,B17,2,A16,1,D15,0\H,4,B18,3,A17,2,D16,0\H,8,B19,7,A18,
6,D17,0\H,11,B20,10,A19,9,D18,0\H,9,B21,8,A20,7,D19,0\H,10,B22,9,A21,8
,D20,0\H,14,B23,3,A22,2,D21,0\H,14,B24,3,A23,2,D22,0\H,15,B25,13,A24,1
,D23,0\H,16,B26,13,A25,1,D24,0\H,16,B27,13,A26,1,D25,0\H,16,B28,13,A27
,1,D26,0\H,17,B29,13,A28,1,D27,0\H,17,B30,13,A29,1,D28,0\H,17,B31,13,A
30,1,D29,0\\B1=1.35836737\B2=1.37657289\B3=1.40993898\B4=1.36536479\B5
=1.4205893\B6=1.43134012\B7=1.41565548\B8=1.37333953\B9=1.40886207\B10
=1.37283177\B11=1.76778836\B12=1.46828448\B13=3.75480946\B14=1.5135041
6\B15=1.53469465\B16=1.52805735\B17=1.07757972\B18=1.08202191\B19=1.08
079706\B20=1.08139368\B21=1.0832671\B22=1.08336926\B23=1.08342297\B24=
1.0836091\B25=1.0873058\B26=1.09159352\B27=1.09020935\B28=1.09103012\B

29=1.09073512\B30=1.09090526\B31=1.08863632\A1=126.08377807\A2=120.456
00876\A3=120.81594297\A4=121.43142864\A5=117.81073392\A6=118.80300463\
A7=121.08430312\A8=120.11540229\A9=120.34686235\A10=118.55624618\A11=1
24.55818558\A12=94.33418028\A13=109.77386024\A14=102.09675635\A15=110.
44331452\A16=121.08732466\A17=119.06069813\A18=118.56536746\A19=120.01
159919\A20=120.00964361\A21=119.85068273\A22=118.37198917\A23=99.83282
335\A24=113.71417886\A25=110.0043058\A26=109.89320853\A27=110.59094426
\A28=109.87550314\A29=109.30649654\A30=112.89700859\D1=180.8810416\D2=
0.12346597\D3=0.02630587\D4=-0.17679772\D5=-179.84748521\D6=-0.0781304
\D7=0.00669274\D8=0.05878608\D9=179.97230143\D10=-9.32258324\D11=-13.4
5869807\D12=64.42726042\D13=179.1897892\D14=-63.90741945\D15=-0.141804
96\D16=179.95891291\D17=179.88822119\D18=-180.04144149\D19=-180.047215
99\D20=-179.97620902\D21=-122.43092368\D22=109.90789063\D23=39.7392781
4\D24=59.03350319\D25=179.00967237\D26=-61.04898426\D27=-170.67280226\
D28=-52.39684071\D29=68.54837921\\Version=IA64L-G09RevA.01\State=1-A\H
F=-1116.2555492\RMSD=4.030e-09\RMSF=2.384e-05\Dipole=0.6989894,0.19612
63,-1.5995093\Quadrupole=8.532045,-4.8662291,-3.6658159,0.6747227,-3.2
124232,-1.2201454\PG=C01 [X(C15H15Cl1O1)]\\@

**Toluene transition structure**

| | 1 | 2 | 3 |
|---|---|---|---|
| | A | A | A |
| Frequencies -- | -192.4006 | 39.3517 | 65.4194 |
| Red. masses -- | 6.5899 | 3.6805 | 4.5933 |
| Frc consts -- | 0.1437 | 0.0034 | 0.0116 |
| IR Inten -- | 35.9419 | 0.0080 | 1.2915 |

| | |
|---|---|
| Sum of electronic and zero-point Energies= | -1115.960735 |
| Sum of electronic and thermal Energies= | -1115.945030 |
| Sum of electronic and thermal Enthalpies= | -1115.944086 |
| Sum of electronic and thermal Free Energies= | -1116.003623 |

1\1\ ASN_ALTIX-ALTIX7\FTS\RB3LYP\6-311+G(2d,p)\C15H15Cl1O1\AUBOXA\16-A
ug-2009\1\\# B3LYP/6-311+G(2d,p) opt=(TS,noeigen,z-matrix) scrf=(pcm,s
olvent=Toluene)\\Allyl Naphthyl Ether TS\\0,1\O\C,1,B1\C,2,B2,1,A1\C,3
,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,5,B5,4,A4,3,D3,0\C,6,B6,5,A5,4,D4
,0\C,7,B7,6,A6,5,D5,0\C,8,B8,7,A7,6,D6,0\C,9,B9,8,A8,7,D7,0\C,10,B10,9
,A9,8,D8,0\Cl,5,B11,4,A10,3,D9,0\C,1,B12,2,A11,3,D10,0\C,3,B13,2,A12,1
,D11,0\C,13,B14,1,A13,2,D12,0\C,13,B15,1,A14,2,D13,0\C,13,B16,1,A15,2,
D14,0\H,3,B17,2,A16,1,D15,0\H,4,B18,3,A17,2,D16,0\H,8,B19,7,A18,6,D17,
0\H,11,B20,10,A19,9,D18,0\H,9,B21,8,A20,7,D19,0\H,10,B22,9,A21,8,D20,0
\H,14,B23,3,A22,2,D21,0\H,14,B24,3,A23,2,D22,0\H,15,B25,13,A24,1,D23,0
\H,16,B26,13,A25,1,D24,0\H,16,B27,13,A26,1,D25,0\H,16,B28,13,A27,1,D26
,0\H,17,B29,13,A28,1,D27,0\H,17,B30,13,A29,1,D28,0\H,17,B31,13,A30,1,D
29,0\\B1=1.26892852\B2=1.4232963\B3=1.41584706\B4=1.36147953\B5=1.4336
2303\B6=1.42473423\B7=1.40521825\B8=1.37808277\B9=1.40370352\B10=1.377
33201\B11=1.7646714\B12=2.31494918\B13=2.47378428\B14=1.39558361\B15=1
.49353573\B16=1.49562131\B17=1.08017994\B18=1.08261075\B19=1.08204229\
B20=1.08134663\B21=1.08349673\B22=1.08360305\B23=1.0813428\B24=1.08047
805\B25=1.08564199\B26=1.09182875\B27=1.09679532\B28=1.08927065\B29=1.
09732017\B30=1.08933152\B31=1.08786818\A1=121.60396453\A2=121.36571715

\A3=120.65488662\A4=122.07232673\A5=117.97913201\A6=119.0617151\A7=121
.20319712\A8=119.77691763\A9=120.40636737\A10=118.41519953\A11=109.715
40153\A12=90.52486653\A13=90.80616485\A14=96.61713544\A15=96.14673353\
A16=117.90341224\A17=119.54454831\A18=117.73448183\A19=119.9731912\A20
=120.22490842\A21=119.93161528\A22=90.8857581\A23=92.77786775\A24=116.
5523232\A25=109.96884659\A26=109.93719462\A27=112.20704765\A28=108.533
83097\A29=110.90405765\A30=113.47399367\D1=181.78111035\D2=-4.5153358\
D3=2.28671209\D4=-0.25592826\D5=-182.33265673\D6=1.26525361\D7=-0.3110
3294\D8=-0.51365276\D9=181.39888472\D10=-76.87938226\D11=72.2133305\D1
2=51.92915617\D13=171.83166457\D14=-71.2023328\D15=-12.40816773\D16=17
5.93950939\D17=181.66298491\D18=-179.13262293\D19=-180.05256931\D20=-1
79.96187586\D21=-179.37366496\D22=64.70477539\D23=93.35148434\D24=39.9
6698952\D25=157.60273974\D26=-81.32565757\D27=-179.3495072\D28=-61.546
72389\D29=61.10748044\\Version=IA64L-G09RevA.01\State=1-A\HF=-1116.215
9014\RMSD=7.033e-09\RMSF=3.604e-05\Dipole=0.2621944,1.3012182,-0.94836
22\Quadrupole=6.8906663,-4.6735823,-2.217084,1.0337147,-1.5078368,-4.3
73399\PG=C01 [X(C15H15Cl1O1)]\\@

**Toluene ground state**

1\1\ASN_X86_64-DMC24\FOpt\RB3LYP\6-311+G(2d,p)\C15H15Cl1O1\AUBOXA\24-A
ug-2009\1\\# B3LYP/6-311+G(2d,p) opt=(calcfc,noeigen,z-matrix) scrf=(p
cm,solvent=Toluene)\\Allyl Naphthyl Ether TS\\0,1\O\C,1,B1\C,2,B2,1,A1
\C,3,B3,2,A2,1,D1,0\C,4,B4,3,A3,2,D2,0\C,5,B5,4,A4,3,D3,0\C,6,B6,5,A5,
4,D4,0\C,7,B7,6,A6,5,D5,0\C,8,B8,7,A7,6,D6,0\C,9,B9,8,A8,7,D7,0\C,10,B
10,9,A9,8,D8,0\Cl,5,B11,4,A10,3,D9,0\C,1,B12,2,A11,3,D10,0\C,3,B13,2,A
12,1,D11,0\C,13,B14,1,A13,2,D12,0\C,13,B15,1,A14,2,D13,0\C,13,B16,1,A1
5,2,D14,0\H,3,B17,2,A16,1,D15,0\H,4,B18,3,A17,2,D16,0\H,8,B19,7,A18,6,
D17,0\H,11,B20,10,A19,9,D18,0\H,9,B21,8,A20,7,D19,0\H,10,B22,9,A21,8,D
20,0\H,14,B23,3,A22,2,D21,0\H,14,B24,3,A23,2,D22,0\H,15,B25,13,A24,1,D
23,0\H,16,B26,13,A25,1,D24,0\H,16,B27,13,A26,1,D25,0\H,16,B28,13,A27,1
,D26,0\H,17,B29,13,A28,1,D27,0\H,17,B30,13,A29,1,D28,0\H,17,B31,13,A30
,1,D29,0\\B1=1.36071156\B2=1.37581781\B3=1.40991569\B4=1.36515024\B5=1
.42059405\B6=1.43089322\B7=1.41526016\B8=1.37277302\B9=1.40849259\B10=
1.37230636\B11=1.76524516\B12=1.46355587\B13=3.77050068\B14=1.51393387
\B15=1.53520508\B16=1.52865683\B17=1.07788005\B18=1.08199516\B19=1.080
71719\B20=1.08129302\B21=1.08340518\B22=1.08348435\B23=1.08331383\B24=
1.08370403\B25=1.0871857\B26=1.09157594\B27=1.09029789\B28=1.09094389\
B29=1.09091532\B30=1.0908511\B31=1.08871293\A1=126.00303791\A2=120.434
779\A3=120.84567124\A4=121.33536109\A5=117.88721098\A6=118.83725486\A7
=121.04308345\A8=120.1309797\A9=120.36288774\A10=118.57251169\A11=124.
41521274\A12=94.3934171\A13=109.93647476\A14=102.15618339\A15=110.4824
8445\A16=121.00497917\A17=119.18287997\A18=118.49056524\A19=120.127832
\A20=119.99312635\A21=119.85498301\A22=118.11717435\A23=100.74699424\A
24=113.65981577\A25=109.95143905\A26=110.00869802\A27=110.5219021\A28=
110.01378236\A29=109.29363374\A30=112.75711993\D1=181.04101432\D2=0.16
536348\D3=0.06087633\D4=-0.23035449\D5=-179.88036353\D6=-0.05676634\D7
=0.00775013\D8=0.04591517\D9=179.98255825\D10=-11.69219292\D11=-11.528
31943\D12=65.32419661\D13=180.07451558\D14=-62.9695167\D15=-0.09861552
\D16=179.95507825\D17=179.96860349\D18=-180.02704008\D19=-180.02997209
\D20=-179.97660313\D21=-123.60162954\D22=108.22956576\D23=38.4354798\D
24=59.11864668\D25=179.28551858\D26=-60.75713122\D27=-171.03247859\D28
=-52.62743838\D29=68.05918781\\Version=AM64L-G09RevA.01\State=1-A\HF=-
1116.2526579\RMSD=6.153e-09\RMSF=1.542e-05\Dipole=0.5911504,0.183481,-
1.4201002\Quadrupole=7.4825064,-4.256893,-3.2256134,0.5906736,-2.84000
91,-1.1187007\PG=C01 [X(C15H15Cl1O1)]\\@

**<u>Gas-Phase</u> CBS-QB3 allyl napthyl ether optimizations:**

**Transition structure**

```
Temperature=             298.150000 Pressure=
1.000000
 E(ZPE)=                    0.253015 E(Thermal)=
0.268754
 E(SCF)=                -1111.237476 DE(MP2)=                        -
2.841587
 DE(CBS)=                  -0.276123 DE(MP34)=                       -
0.088404
 DE(CCSD)=                 -0.115874 DE(Int)=
0.094587
 DE(Empirical)=            -0.127885
 CBS-QB3 (0 K)=         -1114.339748 CBS-QB3 Energy=             -
1114.324009
 CBS-QB3 Enthalpy=      -1114.323064 CBS-QB3 Free Energy=        -
1114.382531
 1\1\ ASN_ALTIX-ALTIX8\Mixed\CBS-QB3\CBS-QB3\C15H15Cl1O1\AUBOXA\24-Aug-
 2009\0\\# CBS-QB3 opt=(calcfc,TS,noeigen,z-matrix)\\Allyl Naphthyl Eth
 er TS\\0,1\O,0,0.,0.,0.\C,0,0.,0.,1.2667\C,0,1.2155943164,0.,2.0160426
 838\C,0,1.2072689432,-0.0274028373,3.4352537403\C,0,0.0337944753,0.053
 5979821,4.1254326884\C,0,-1.2404602425,0.1056630059,3.4614801394\C,0,-
 1.251794032,0.0818961563,2.035823276\C,0,-2.486023902,0.0718725059,1.3
 605871954\C,0,-3.678607289,0.1186968259,2.0533731277\C,0,-3.6682553953
 ,0.1687317524,3.4576438938\C,0,-2.474471396,0.1576147576,4.1488576274\
 Cl,0,0.0953395196,0.0580521594,5.8857515027\C,0,0.5285114465,2.0606855
 643,-0.7873304363\C,0,1.6553098849,2.3115182448,1.3823804212\C,0,1.668
 3083255,2.1292791151,0.0176828613\C,0,0.658576349,1.6497659284,-2.2208
 358545\C,0,-0.7687487003,2.7116115056,-0.4102207831\H,0,2.1277084949,-
 0.2382862017,1.4874634074\H,0,2.1426410921,-0.1027448471,3.976093952\H
 ,0,-2.4635676871,0.0152500454,0.2796020404\H,0,-2.4733992826,0.1791511
 211,5.2304427029\H,0,-4.6214670053,0.1116849695,1.5181572562\H,0,-4.60
 39421963,0.2049649961,4.0043466033\H,0,2.5821190923,2.3240588561,1.940
 9717442\H,0,0.7869776325,2.6796941647,1.9112048548\H,0,2.6006322253,1.
 802426002,-0.4344470298\H,0,-0.0995380217,0.898326651,-2.4535542749\H,
 0,0.4863806361,2.5081207931,-2.8825475158\H,0,1.6408481512,1.230013054
 9,-2.4387479282\H,0,-0.736896965,3.7664197209,-0.7138467139\H,0,-1.604
 1606843,2.2431207199,-0.9309311063\H,0,-0.9706585388,2.6818554111,0.65
 94838453\\Version=IA64L-G09RevA.01\State=1-A\HF/CbsB3=-1111.237476\E2(
 CBS)/CbsB3=-3.1177102\CBS-Int/CbsB3=0.0945873\OIii/CbsB3=22.0871398\MP
 2/CbsB4=-1113.4721049\MP4(SDQ)/CbsB4=-1113.5605094\MP4(SDQ)/6-31+G(d')
 =-1113.4062459\QCISD(T)/6-31+G(d')=-1113.5221204\CBSQB3=-1114.3397477\
```

**Ground state**

```
Temperature=             298.150000 Pressure=
1.000000
 E(ZPE)=                    0.255424 E(Thermal)=
0.271269
 E(SCF)=                -1111.307427 DE(MP2)=                        -
2.813048
 DE(CBS)=                  -0.273464 DE(MP34)=                       -
0.101212
 DE(CCSD)=                 -0.106524 DE(Int)=
0.093684
```

DE(Empirical)= -0.128798
CBS-QB3 (0 K)= -1114.381365 CBS-QB3 Energy= -1114.365519
CBS-QB3 Enthalpy= -1114.364575 CBS-QB3 Free Energy= -1114.424684
1\1\ ASN_ALTIX-ALTIX7\Mixed\CBS-QB3\CBS-QB3\C15H15Cl1O1\AUBOXA\16-Aug-2009\0\\# CBS-QB3 opt=(calcfc,noeigen,z-matrix)\\Allyl Naphthyl Ether GS\\0,1\O\C,1,1.3635038238\C,2,1.3778714075,1,125.9415159\C,3,1.4117824035,2,120.38674442,1,181.17968325,0\C,4,1.3680137838,3,120.84080908,2,0.19797135,0\C,5,1.4237441509,4,121.31749552,3,0.10241508,0\C,6,1.4326954345,5,117.86820549,4,-0.30970062,0\C,7,1.4169968409,6,118.92389935,5,-179.89217831,0\C,8,1.3742721399,7,120.96185168,6,-0.02678701,0\C,9,1.410347751,8,120.14470683,7,0.00102725,0\C,10,1.3740666501,9,120.41212579,8,0.04029215,0\Cl,5,1.7621083382,4,118.55304047,3,179.98544941,0\C,1,1.4609711537,2,124.45429205,3,-14.53471825,0\C,3,3.7934875996,2,94.33572004,1,-9.21150519,0\C,13,1.5164169889,1,109.99487544,2,66.29246297,0\C,13,1.5370590232,1,102.13765963,2,181.04349662,0\C,13,1.5312577032,1,110.42216904,2,-61.8347039,0\H,3,1.0789161505,2,120.935138,1,-0.11125375,0\H,4,1.0823870006,3,119.34796406,2,179.97948293,0\H,8,1.0811326952,7,118.28561747,6,180.07522296,0\H,11,1.0817775559,10,120.2826156,9,-180.01148775,0\H,9,1.0841140593,8,120.00143887,7,-180.02160887,0\H,10,1.084201805,9,119.82458188,8,-179.9775167,0\H,14,1.0839530157,3,117.87720404,2,-125.08713723,0\H,14,1.0844240312,3,101.67973848,2,106.28838237,0\H,15,1.087826447,13,113.51109861,1,37.44271553,0\H,16,1.0923184935,13,109.85438128,1,59.41598315,0\H,16,1.091320782,13,110.17419764,1,179.82719185,0\H,16,1.0916207444,13,110.32228633,1,-60.12059795,0\H,17,1.0919211456,13,110.22051667,1,-171.39534646,0\H,17,1.0916938097,13,109.18513471,1,-52.84362811,0\H,17,1.089664788,13,112.52485248,1,67.52893811,0\\Version=IA64L-G09RevA.01\State=1-A\HF/CbsB3=-1111.3074268\E2(CBS)/CbsB3=-3.086512\CBS-Int/CbsB3=0.0936843\OIii/CbsB3=22.2448482\MP2/CbsB4=-1113.5114539\MP4(SDQ)/CbsB4=-1113.6126664\MP4(SDQ)/6-31+G(d')=-1113.4575547\QCISD(T)/6-31+G(d')=-1113.5640786\CBSQB3=-1114.3813649\

**Gaussian 09 Reference**

Gaussian 09, Revision A.1, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.

## Appendix II



### Grid Box Dimensions for Docking Simulations

| Ligand | x (Å) | center_x | y(Å) | center_y | z(Å) | center_z |
|---|---|---|---|---|---|---|
| $FMNO^-$ | 18 | 0.497 | 26 | 36.795 | 18 | 40.6 |
| OCS | 18 | 0.883 | 26 | 38.315 | 24 | 44.348 |
| $FMNH_2$-OCS | 18 | 3.529 | 18 | 27.214 | 18 | 38.895 |
| $FMNO^-$-OCS | 14 | 44.156 | 10 | 33.999 | 12 | 29.818 |

**Table 2-1**: Grid box dimensions for all simulations. For the $FMNH_2$/$FMNO^-$-OCS simulations, the flavin ligand was placed in the same location as the $FMNH_2$ and $FMNO^-$ simulations, and octanesulfonate was docked subsequently.

**Flavin Atom Positions**

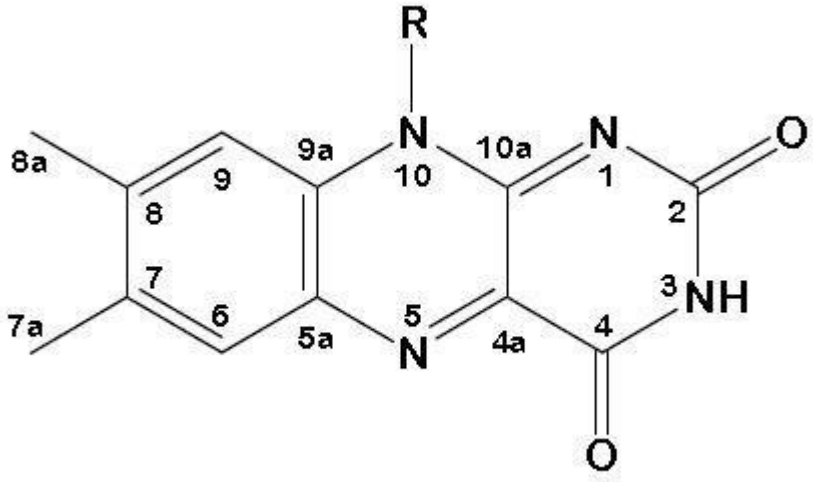



**Figure 2-1**: Flavin isoalloxazine ring numbering[1]

**Distance Calculations**

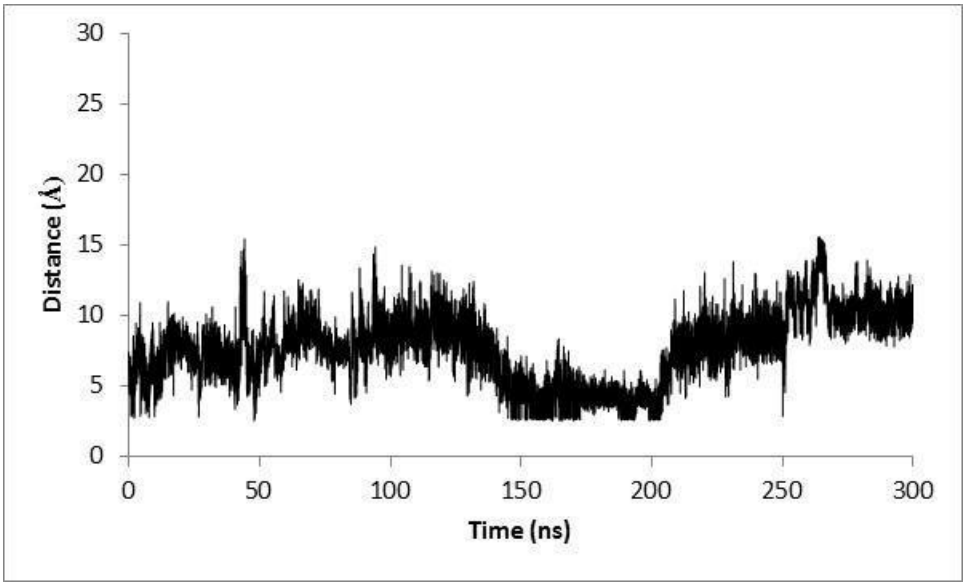



**Figure 2-2**: Distance calculation of an oxygen of Asp111 and nitrogen of Arg297 for the $FMNH_2$-OCS simulation.

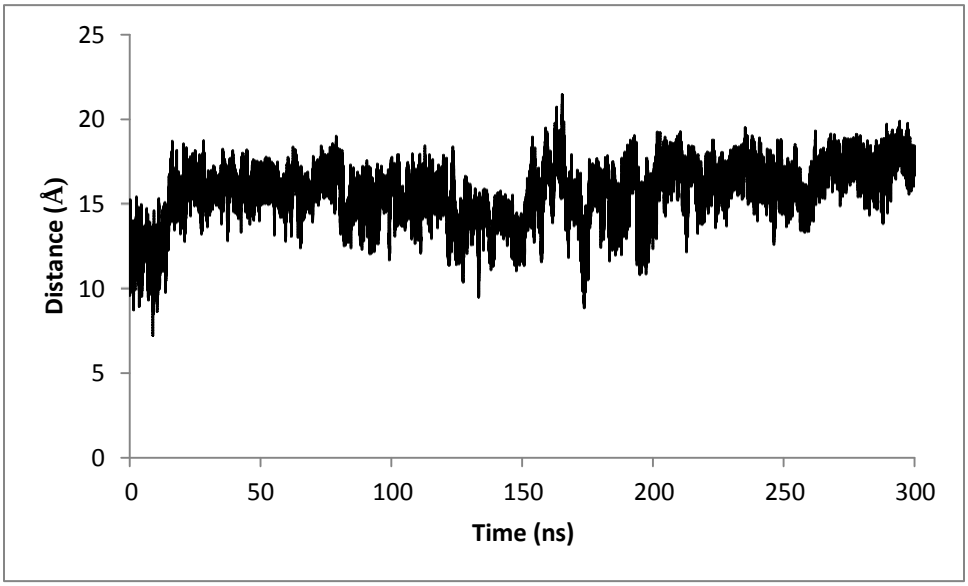



**Figure 2-3**: Distance calculation of the nitrogen of Arg226 with the sulfur of octanesulfonate for the $FMNO^-$-OCS bound SsuD system.

**Clustering Analysis**

| Simulation | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Apo** | 4.9 | 5.4 | 13.7 | 6.3 | 4.4 | 12.7 | 25.4 | 2.0 | 6.2 | 19.1 |
| **$FMNH_2$** | 3.8 | 6.9 | 6.4 | 1.2 | 12.6 | 10.1 | 11.8 | 10.1 | 1.0 | 36.2 |
| **$FMNO^-$** | 1.1 | 2.9 | 3.9 | 34.2 | 6.8 | 1.3 | 5.1 | 11.4 | 23.9 | 9.3 |
| **OCS** | 0.4 | 5.9 | 0.7 | 2.3 | 1.8 | 1.5 | 10.5 | 10.9 | 43.4 | 22.6 |
| **$FMNH_2$-OCS** | 1.0 | 0.7 | 5.4 | 10.7 | 8.0 | 2.5 | 0.6 | 5.2 | 0.4 | 65.6 |
| **$FMNO^-$-OCS** | 1.9 | 2.3 | 11.1 | 37.5 | 0.6 | 0.3 | 9.1 | 9.0 | 27.1 | 1.2 |

**Table 2-2**: Percent occupancy for clusters over the 300ns simulation. All values given as a percentage for each cluster 0-9.

*Structures*

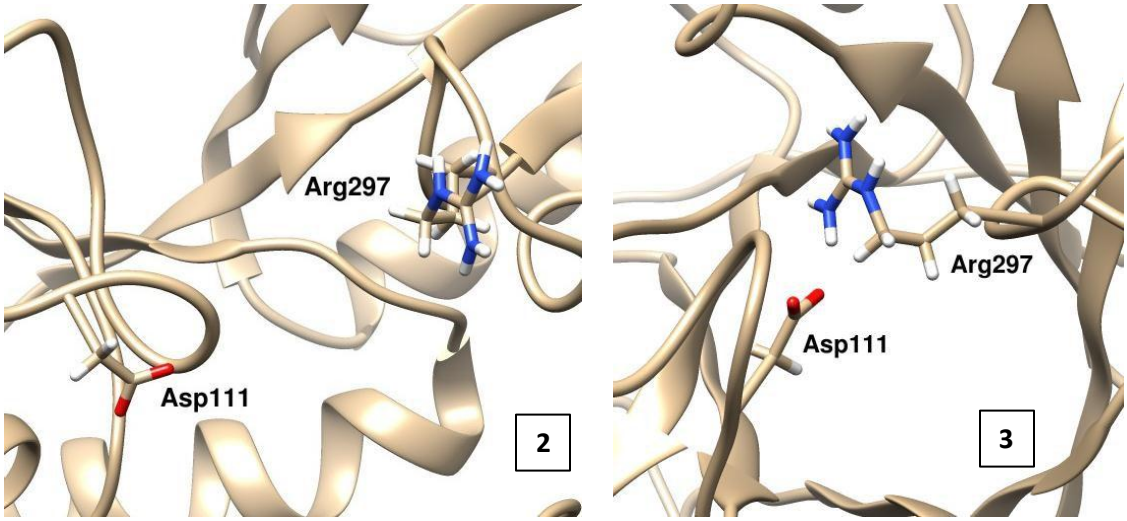



**Figure 2-4:** Apo: 2 represents second cluster structure and 3 represents third cluster structure with the important residues of Asp111 and Arg297.

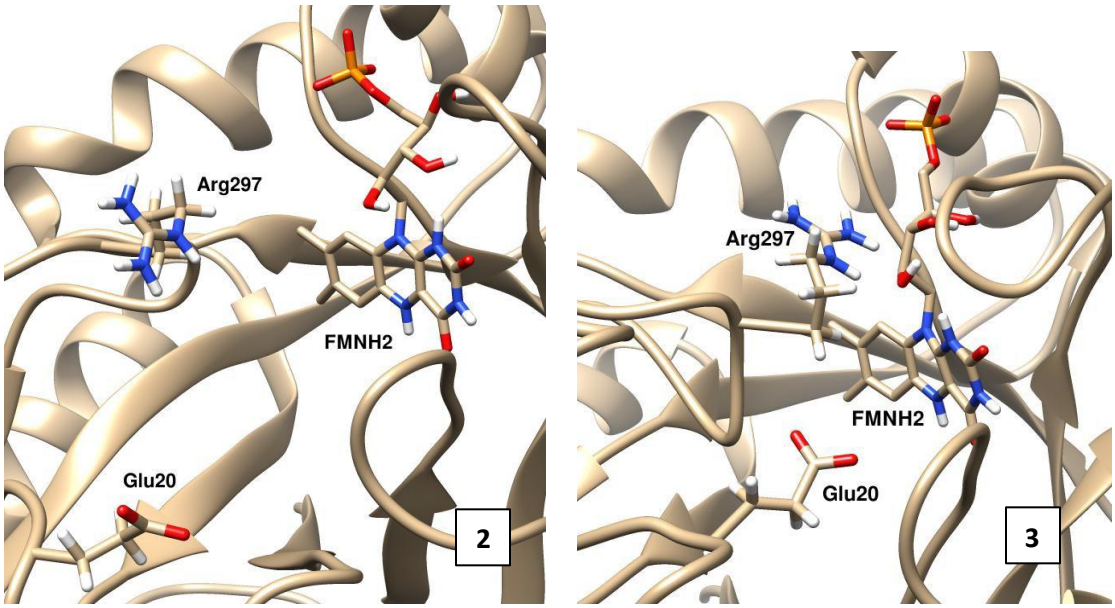



**Figure 2-5:** $FMNH_2$: 2 represents second cluster structure and 3 represents third cluster structure with the important residues of Glu20 and Arg297.

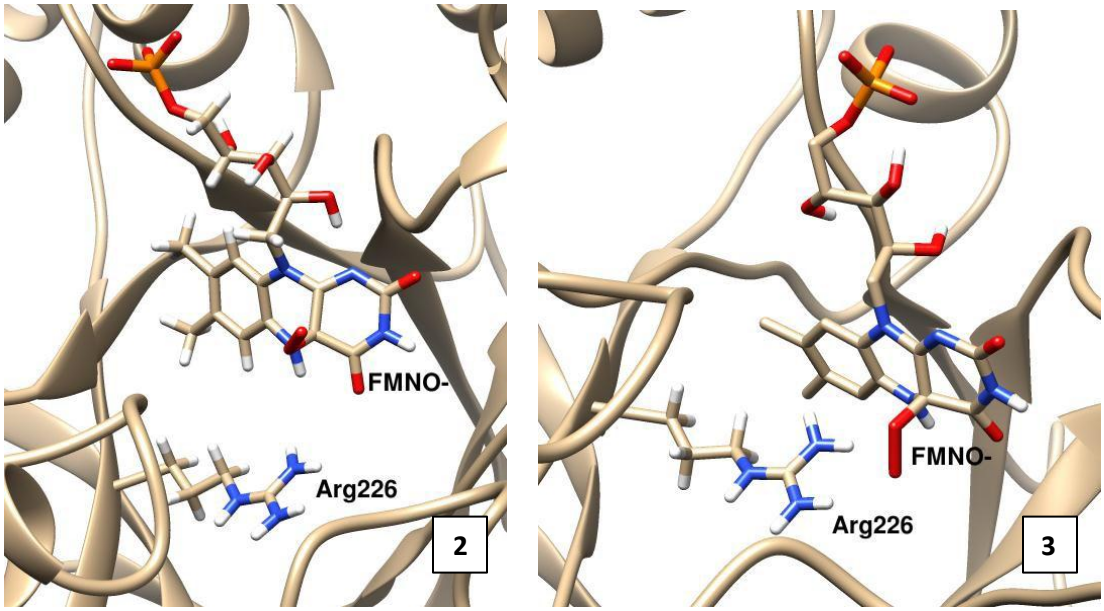



**Figure 2-6:** $FMNO^{-}$: 2 represents second cluster structure and 3 represents third cluster structure with the important residue of Arg226.

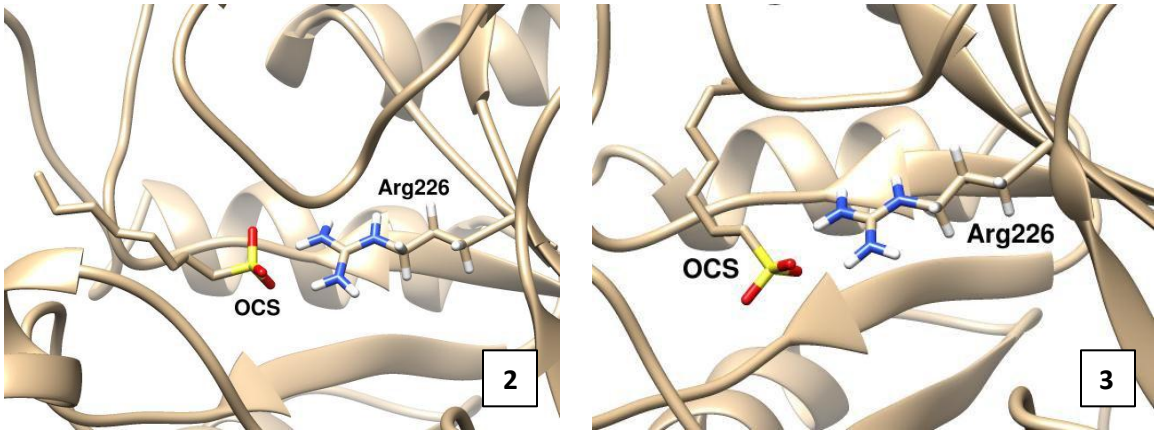



**Figure 2-7:** OCS: 2 represents second cluster structure and 3 represents third cluster structure with the important residue of Arg226.

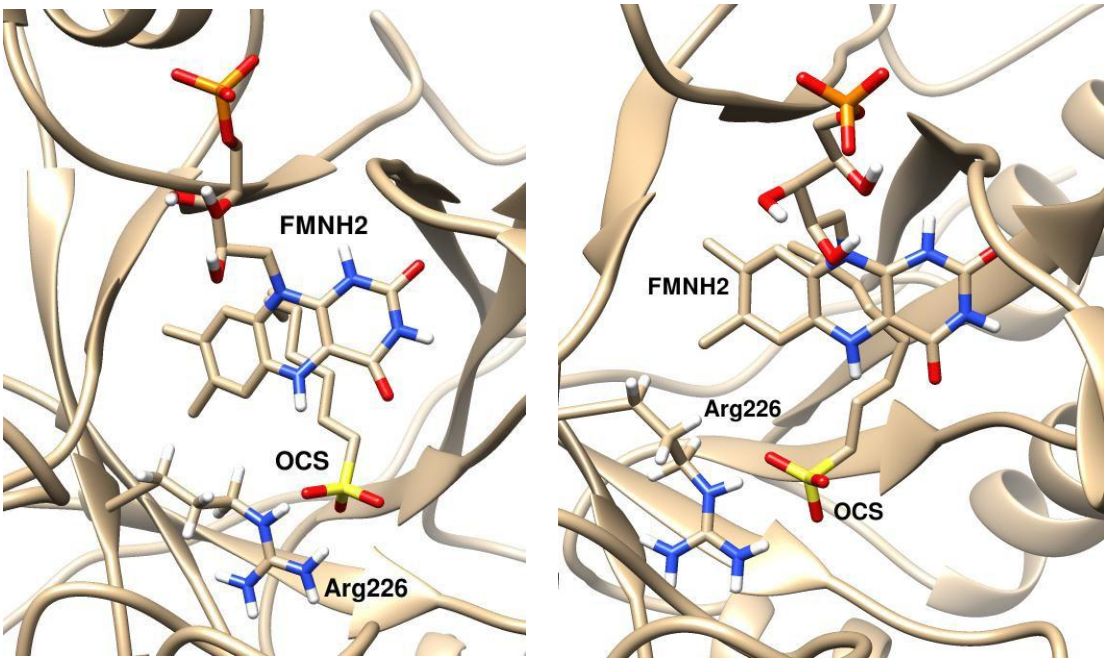



**Figure 2-8:** $FMNH_2$-OCS: 2 represents second cluster structure and 3 represents third cluster structure with the important residue of Arg226.

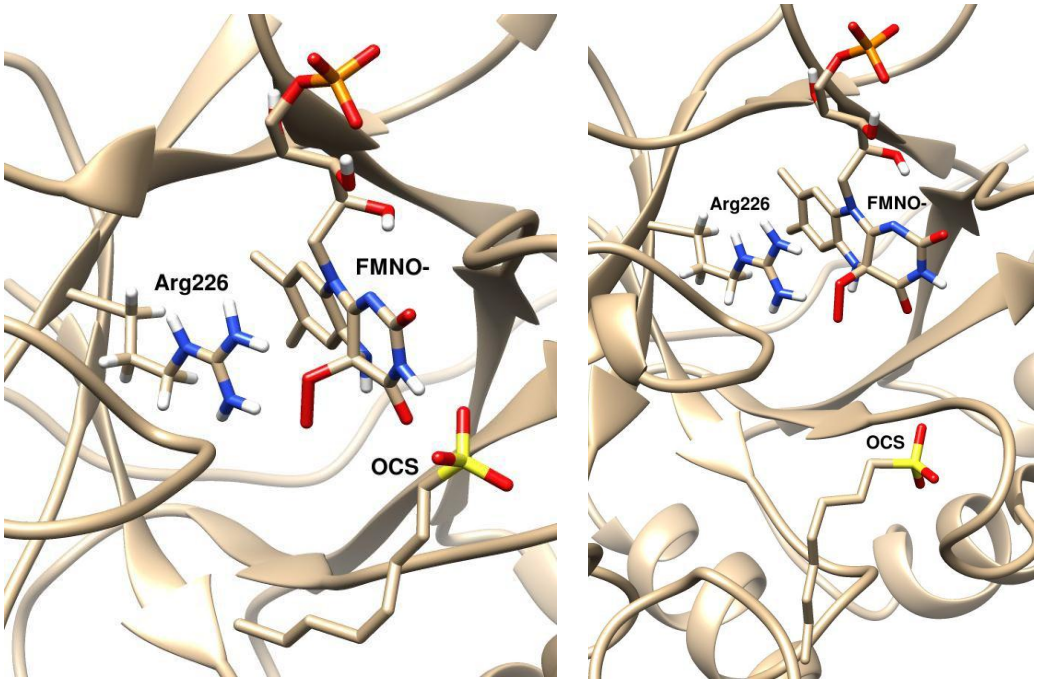



**Figure 2-9:** $FMNO^-$-OCS: 2 represents second cluster structure and 3 represents third cluster structure with the important residue of Arg226.

1. Müller, F., Flavin Radicals: Chemistry and Biochemistry. *Free Radic. Biol. Med.* **1987,** 3, 215-230.