QM/MM/MC Simulations for the Elucidation of Ionic Liquid Solvent Effects upon Organic Reactions

by

Caley R. Allen

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

Keywords: chemistry, organic chemistry, quantum mechanics, molecular mechanics, QM/MM, ionic liquids

Approved by

Orlando Acevedo, Chair, Associate Professor of Chemistry and Biochemistry
Vincent Ortiz, Department of Chemistry Chair, Department of Chemistry and Biochemistry
Peter Livant, Associate Professor of Chemistry and Biochemistry
Konrad Patkowski, Assistant Professor of Chemistry and Biochemistry

Abstract

The research in this dissertation is focused on elucidating and understanding ionic liquid solvent effects on classical organic reactions. Ionic liquids are an exciting class of solvents that have attracted the attention of a growing number of scientists as a reaction medium for an assortment of chemical reactions over the last ten years. The increase in interest is due in part to the classification of ionic liquids as “green”, but more so for their complex reaction environment. Ionic liquids are molten salts, therefore the microenvironment is completely different than that of molecular solvents. Ionic liquids are complex solvent systems capable of many types of intermolecular interactions. Proper description of the physico-chemical properties aspects of solute-solvent interactions is vital to understanding the solvent effects on chemical processes, and reaction rates. Three organic reactions were chosen to investigate the complex behavior of the ionic liquid environment: (1) the base catalyzed β-elimination, (2) the nucleophilic aromatic substitution reaction, and (3) a heterocyclic rearrangement reaction called the Boulton-Katrizky rearrangement. These specific reactions were chosen to elucidate the experimentally observed ionic liquid behavior, e.g., increased favorable electrostatic interactions, increased stabilization through π-π interactions, and an increase in hydrogen bond donor and acceptor capabilities. The empirical effects of ionic liquids on many organic systems have been reportedly, however a theoretical examination of the unique reaction environment and ionic liquid effects is still in its infancy. It is hoped the research

in this dissertation will shed light on the microscopic details of how ionic liquids operate upon and affect chemical reactions.

I dedicate this to my mother and father.

Believe.

Acknowledgments

I would like to express my very great appreciation to Auburn University, and Auburn University's Chemistry and Biochemistry Department. This department comprises individuals with amazing talents all of whom are dedicated to sharing their knowledge and ideas with others so that they too might go on to become successful in their lives. I take immense pleasure in thanking my research advisor, Professor Orlando Acevedo. I am particularly grateful for his assistance, understanding, and advice given. I am highly indebted to him, Auburn University's Chemistry and Biochemistry faculty and staff, and to Auburn University. Finally, yet importantly, I would like to express my sincere thanks and everlasting appreciation to my parents and family for their blessings and enduring support, my friends and classmates for their help and wishes for the successful completion of this chapter in my scholastic journey.

Table of Contents

Chapter 2: Computational Methods

List of Abbreviations

RTIL(s)	room temperature ionic liquid(s)
VOC	volatile organic compounds
E1cb	elimination unimolecular conjugate base
E2	elimination bimolecular
S_NAr	aromatic nucleophilic substitution reaction
MHR	mononuclear heterocyclic rearrangement
Pip	piperidine
Mor	morpholine
Pyr	pyrrolidine
QM	quantum mechanical
QM/MM	hybrid quantum mechanical molecular mechanical
MM	molecular mechanical
DFT	density functional theory
PDDG/PM3	pairwise distance directed Gaussian third parameterization of modified neglect of differential overlap
OPLS-AA	optimized potentials for liquid simulations all atom force field
AA FF	all atom force field
UA FF	united atom force field
AMBER FF	assisted model building with energy refinement force field

CHARMM	chemistry at Harvard macromolecular mechanics
CNDO	complete neglect of differential overlap
INDO	intermediate neglect of differential overlap
NDDO	neglect of diatomic differential overlap
ZDO	zero differential overlap
MINDO/3	modified intermediate neglect of differential overlap
MNDO	modified neglect of differential overlap
AM1	Austin model one
PM3	third parameterization of modified neglect of differential overlap
RH	Roothan-Hall equations
CRF	core repulsion functions
BGE	bond group equivalents
BOSS	biochemical and organic simulation system
MC	Metropolis Monte Carlo
ONIOM	our own *n*-layered integrated molecular orbital and molecular mechanics
CM1	charge model 1
CM3	charge model 3
NVT	canonical ensemble of constant number of molecules, volume and temperature
NPT	isothermal-isobaric ensemble of constant number of molecules, pressure and temperature
MD	molecular dynamics
PBC	periodic boundary conditions
PES	potential energy surface

RDFs	radial distribution functions
FEP	free energy perturbation theory
TI	thermodynamic integration
FF	force field
NDTI	numerical derivative thermodynamic integration
PMF	potentials of mean force
PME	particle mesh Ewald
SPME	smooth particle mesh Ewald
SFG	shifted force gradient
SF3D	shifted force 3rd derivative

List of Figures

Chapter 1

Chapter 2

List of Schemes

List of Tables

Chapter 1

Ionic Liquids and Classical Organic Reactions

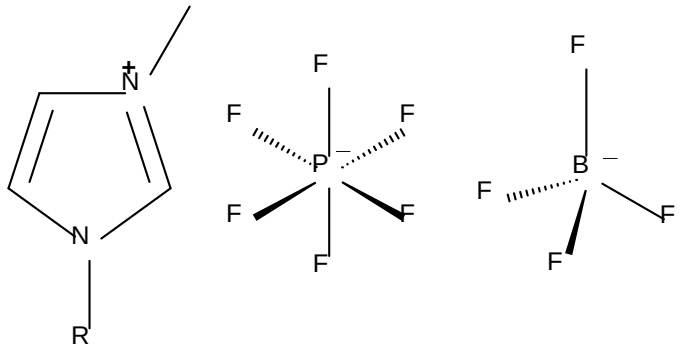
1.1 Introduction

There are two categories of solvents of interest in this work: (1) molecular liquids, and (2) ionic liquids.[1] A solvent can be characterized by physical properties such as melting point, boiling point, density, vapor pressure, surface tension, and heat of vaporization. A solvent can also be thought of from the microscopic point of view, characterized through the dipole moment, electronic polarizability, hydrogen bond donor and acceptor ability, and electron donor and acceptor ability. Proper description of the physico-chemical properties between the solute and solvent is vital to understanding the solvents effects on chemical processes, reaction rates and spectral properties. These intermolecular forces have two general classes: (1) non-directional forces such as induction and dispersion forces, which are non-specific and (2) directional forces such as H-bonding, charge transfer and electron pair donor-acceptor forces that can lead to stereospecific molecular compounds.[1] The complexity of these solvent-solvent and solute-solvent interactions makes the overall structure of a liquid the least known of all aggregation states, most specifically for the class of ionic liquids. As a result, the modeling and theoretical examination of the structure of liquids and the corresponding solvent effects on solutes is among the most difficult tasks in computational chemistry.

1.2 Ionic Liquids

1.2.1 Introduction

Ionic liquids are molten salts making the microenvironment completely different than that of molecular solvents. Ionic liquids generally have a melting point below 100 °C and many are liquid at room temperature,[2] hence the term room temperature ionic liquids (RTILs). Generally the cation has a low degree of symmetry, which reduces the corresponding lattice energy of the crystalline form of the salt, lowering the melting point to below room temperature. Ion components can be fine-tuned through different functional groups to enhance the degree of localized structuring in the liquid phase, which distinguishes ionic liquids from molecular solvents and solutions containing dissociated ions.[3] These different combinations of anions and counter-ions can alter the solvent's properties, such as melting point, viscosity, density, H-bond capabilities, electrophilicity, nucleophilicity and hydrophobicity.[2d] These "designer"[4] solvents are typically composed of a low symmetry organic cation, such as the 1-alkyl-3-methylimidazolium [RMIM] cation (R = M (methyl), E (ethyl), B (butyl), H (hexyl), and O (octyl)), and a weakly coordinating inorganic or organic anion with a diffuse negative charge like hexafluorophosphate PF_6^- or tetrafluoroborate BF_4^-[3, 5] (Figure 1.1).



(R = methyl, ethyl, butyl, hexyl, octyl)

Figure 1.1: Ionic liquids composed of an organic cation, such as the 1-alkyl-3-methylimidazolium [RMIM] cation (R = M (methyl), E (ethyl), B (butyl), H (hexyl), and O (octyl)), and a weakly coordinating inorganic or anion, hexafluorophosphate PF_6^- or tetrafluoroborate BF_4^-.

1.2.2 The Green Aspect

RTILs are an exciting class of solvents that have attracted considerable attention over the last ten years (Figure 1.2).

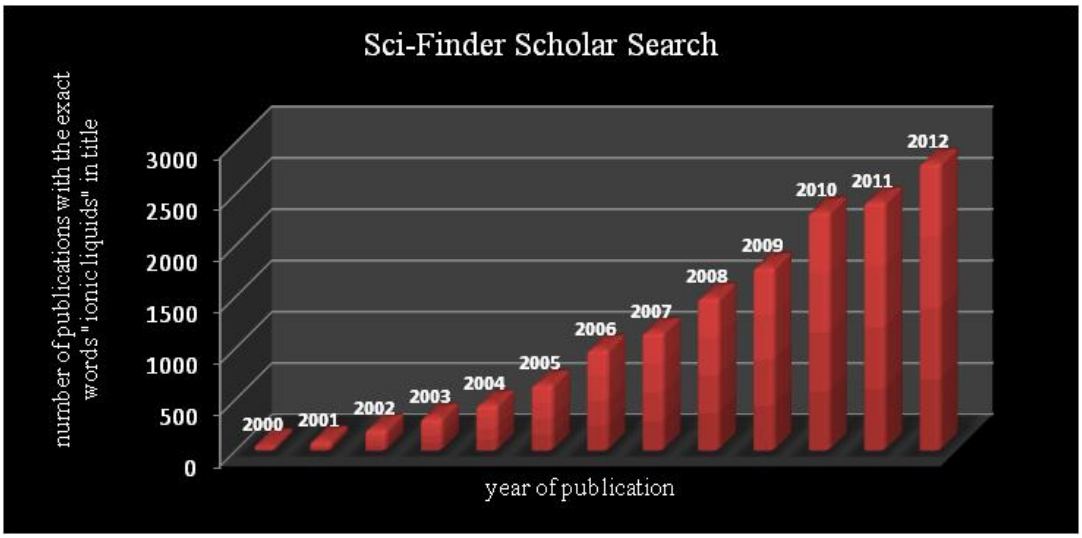


Figure 1.2: Sci-Finder scholar search of "ionic liquids" versus publication year.

The popularity and interest of ionic liquids has increased dramatically in part due to their classification as green solvents. Besides their negligible vapor pressure which prevents solvent evaporation into the atmosphere, two additional properties are of great interest for the advancement of green chemistry. Often a biphasic layer is formed with reacting substances because of the special solubility characteristics of ionic liquids, which allows for effective separation and recycling.[6] For example, some hydrophobic ionic liquids will form three liquid phase systems with water and hexane, offering interesting possibilities for extraction and phase transfer chemistry.[6b] Second, the non-volatile nature of the ionic liquids allows for more effective product isolation by distillation.[6a] The ultimate goal is to reduce or eliminate the related costs of disposal requirements and hazards associated with

the use of volatile organic compounds (VOCs). However, it is important to note that the use of ionic liquids consisting of halogen containing anions can be limited in their greenness. For example, the "work horse ionic liquids"[6a, 7] 1-butyl-3-methyl imidazolium ([bmim]) with hexafluorophosphate (PF_6) or tetrafluoroborate (BF_4) have been shown to hydrolyze after the addition of excess water when the sample was kept at 100°C for eight hours.[6a, 8] The tendency of these specific anions to hydrolyze and form toxic, highly corrosive by-products such as HF, is reduced compared to other ionic liquid anions, e.g., chloroaluminate, but is still problematic. Consequently, the application of such ionic liquids is therefore restricted to those cases where water-free conditions can be utilized at an acceptable cost, and therefore can detract from the overall green appeal of the solvent.[6a]

1.2.3 Solvent Effects

RTILs are complex solvent systems capable of undergoing many types of intermolecular interactions. An effective means to characterize RTILs on the basis of more than one type of solvation interaction would greatly increase our understanding and efficient use of these solvents. Characterizing ionic liquids with a single polarity or polarizability term, as is common for molecular solvents, fails to provide adequate correlation with experimental observations.[9] Additionally, this narrow scope on interpreting ionic liquid solvent effects fails to describe the type and magnitude of these intermolecular interactions that make each designer solvent unique.[9] The ability to fine tune the chemical properties of ionic liquids through the critical selection of the ionic liquid's components allows for their solvent properties to differ considerably from other RTILs, as well as from conventional molecular solvents. RTILs are useful in many

chemical applications, including as solvents in organic synthesis, matrixes in matrix-assisted laser desorption/ionization (MALDI) mass spectroscopy, liquid-liquid extractions, and as stationary phases in gas chromatography.[9a] Figure 1.3 and Table 1.1 below provide data that can be useful in identifying the interactions, properties, and corresponding chemical applications that are important for specific RTILs and comparisons to select molecular solvents.

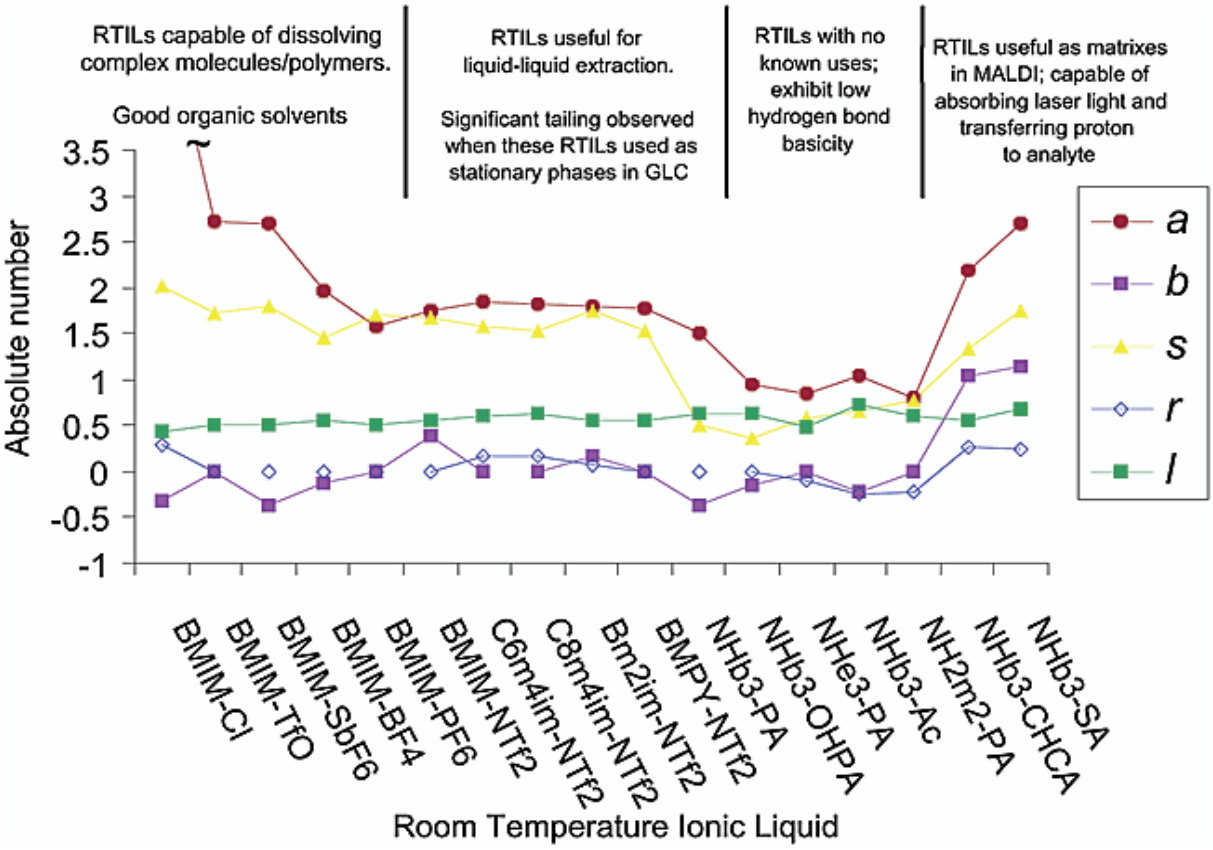


Figure 1.3: [9a] Characterizing ionic liquids on the basis of multiple solvation interactions: *a* – hydrogen bond basicity, *b* – hydrogen bond acidity, *s* – dipolarity/polarizability, *r* – nonbonding interactions, and *l* – dispersion forces.

Solvent	E_T^N (Lit.)	π^* (Lit.)	α (Lit.)	β (Lit.)
[EtNH$_3$][NO$_3$][11]	(0.954)	(1.12)	(1.10[a])	(0.46[b])
[HO(CH$_2$)$_2$-mim][Tf$_2$N][10]	(0.929)			
[CH$_3$O(CH$_2$)$_2$-mim][Tf$_2$N][10]	(0.722)			
[bmim][SbF$_6$]	0.673	1.039	0.639	0.146
[bmim][BF$_4$][8]	0.670 (0.673)	1.047	0.627	0.376
[bmim][PF$_6$][8]	0.669 (0.667)	1.032	0.634	0.207
[bmim][TfO][8]	0.656 (0.667)	1.006	0.625	0.464
[bmim][N(Tf)$_2$][8]	0.644 (0.642)	0.984	0.617	0.243
[bm$_2$im][BF$_4$]	0.576	1.083	0.402	0.363
[bmpy][N(Tf)$_2$]	0.544	0.954	0.427	0.252
[bm$_2$im][N(Tf)$_2$][8]	0.541 (0.552)	1.010	0.381	0.239
Water[3,6,28]	(1.00)	(1.33) (1.13[b])	(1.12) (1.16[b])	(0.14) (0.50[b])
Methanol[3,6,28]	(0.762)	(0.73)	(1.05)	(0.61)
Acetonitrile[28]	0.460	0.799	0.350	0.370
Acetone[28]	0.350	0.704	0.202	0.539
Dichloro-methane[28]	0.309	0.791	0.042	−0.014
Toluene	0.100	0.532	−0.213	0.077
Hexane[3,7]	(0.009)	(−0.12)	(0.07)	(0.04)

Table 1.1:[9b] A range of ionic liquids and select molecular solvents have been investigated using the Kamlet-Taft parameters: E_T^N - solvating ability, π^* - dipolarity/polarizability, α – hydrogen bond acidity, and β – hydrogen bond basicity.

A range of ionic liquids and selected molecular solvents were investigated by Welton and co-workers[9] using the Kamlet-Taft parameters[10] (Table 1.1): were E_T^N stands for the solvating ability of the solvent, π^* is the dipolarity/polarizability of the solvent, α is the hydrogen bond acidity, and β is the hydrogen bond basicity of the solvent. The Kamlet-Taft parameters are determined through the solvatochromic comparison method; for example, the measures of hydrogen bond acceptor strengths, β, are determined from the magnitudes of enhanced solvatochromic shifts.[10a] Solvatochromism is the change in

shape, intensity, and/or position of the UV/vis spectrum of a solute molecule induced by the solvent. The probe solute molecules that were used for the Kamlet-Taft analysis corresponding to Table 1.1 are seen below. [9b]

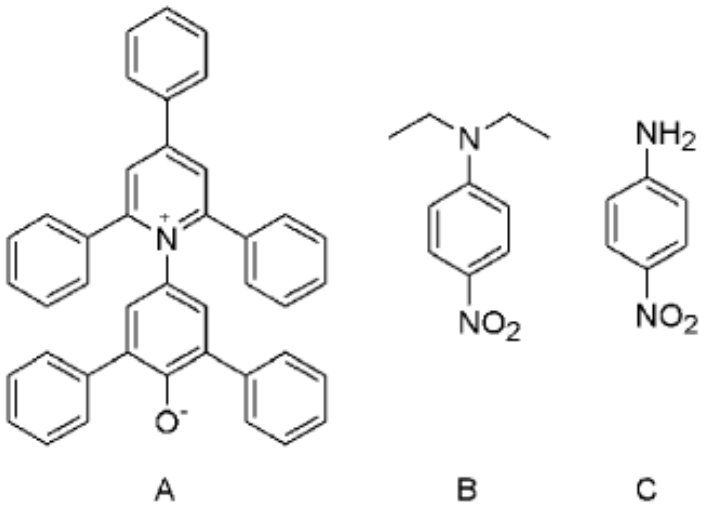


Figure 1.4:[9b] The probe solute molecules used in the Kamlet-Taft analysis of the select solvents seen in Table 1.1: (a) Reichardt's dye, (b) N,N-diethyl-4-nitroaniline, and (c) 4-nitroaniline.

There are many sets of dyes that can be used to determine the Kamlet-Taft solvation parameters. Many studies quote an average of the values obtained from several different sets of dyes. However, Welton and coworkers reported the values obtained using only this single set of dyes[9b] (Figure 1.4). The ultimate goal of the Kamlet-Taft parameters is to link a desired, specific property or feature of the ionic liquid to fundamental properties that can be predicted. For example, the averaged polarity of methanol and [BMIM][BF_4] are very similar.[11] On the other hand, the experimentally measured dipolarity/polarizability (π^*) values for [BMIM][BF_4] and [BMIM][PF_6] of 1.047 and 1.032, respectively, are significantly higher than that of 0.73 for methanol.[9b] Comparison

of the Kamlet-Taft parameters of RTILs with those of molecular solvents leads to three general observations: (1) ionic liquids are polar solvents, (2) ionic liquids have a varying capability to act as a hydrogen bond donor or acceptor, and (3) ionic liquids have uniform polarizability. These trends, when compared directly to the molecular solvents properties, demonstrate how care must be taken when making claims of similarities between the solvents and ultimately their corresponding physico-chemical properties.

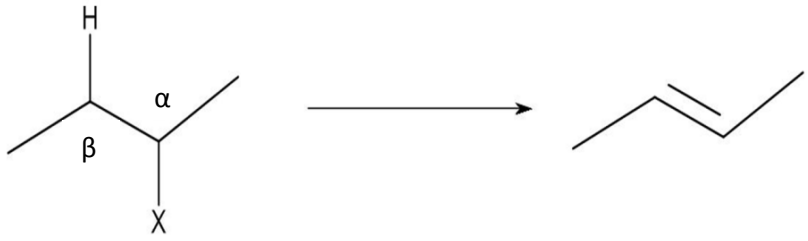
1.3 Organic Reactions

The use of ionic liquids as a reaction medium for chemical reactions has dramatically increased in recent years, due in large part to numerous reported advances in catalysis,[12] separation science,[13] and organic synthesis[14] when employing the unique solvents. For example, the Diels-Alder reaction highlights the advantages provided by ionic liquids as the reaction between cyclopentadiene and methyl acrylate in 1-ethyl-3-methylimidazolium tetrachloroaluminate and heptachlorodialuminate ([EMIM][$AlCl_4$] and [EMIM][Al_2Cl_7], respectively), has been reported to react with rates over 200 times faster and *endo* selectivity 10 times greater than for commonly used reaction conditions.[15] Our hybrid quantum mechanical molecular mechanical (QM/MM) investigation of the same Diels-Alder reaction in the chloroaluminate ionic liquids emphasized the importance of intermolecular interactions on the rate of reaction with excellent $\Delta\Delta G^{\ddagger}$ agreement reported between the solvents.[16] Another example of recent success utilizing our QM/MM method was the Kemp elimination ring-opening of benzisoxazole in 1-butyl-3-methylimidazolium hexafluorophosphate [BMIM][PF_6] using piperidine as the base.[17] Additionally, ionic liquids have been proposed to induce a mechanistic change in the reaction pathway for the fundamentally important base-

induced β-elimination class compared to conventional solvents.[18] The effect of RTILs on many organic systems has been reported experimentally;[2a] however, theoretical calculations have only begun to elucidate the microscopic details on how ionic liquids operate upon and affect chemical reactions.

1.3.1 Introduction to Elimination Reactions

An elimination reaction is when a single compound splits into different species and transforms into an alkene or alkyne. The reaction takes place around as sp^3 – sp^3 carbon to carbon covalent bond with an β-acidic hydrogen atom and a α-leaving group. Often *X* is a halogen atom in Scheme 1.1. When the acidic proton and leaving group are lost from adjacent atoms on the substrate to form the alkene, the reaction is called a β-elimination.



Scheme 1.1: Schematic demonstrating an elimination reaction. α and β atoms are labeled as adjacent atoms on the substrate.

The order in which the starting substrate, or alkyl halide, splits apart determines if the reaction mechanism is an elimination bimolecular mechanism (E2, Scheme 1.2), or an elimination unimolecular conjugate base mechanism (E1cb, Scheme 1.3). Many experimental techniques can be utilized to distinguish between the elimination mechanisms, e.g., deuterium or tritium isotope effects, carbon isotope effects, substituent effects, and leaving group effects. However, experimentally distinguishing between the

irreversible E1cb and E2 mechanisms for dehydrohalogenation reactions can be notoriously difficult.[19]

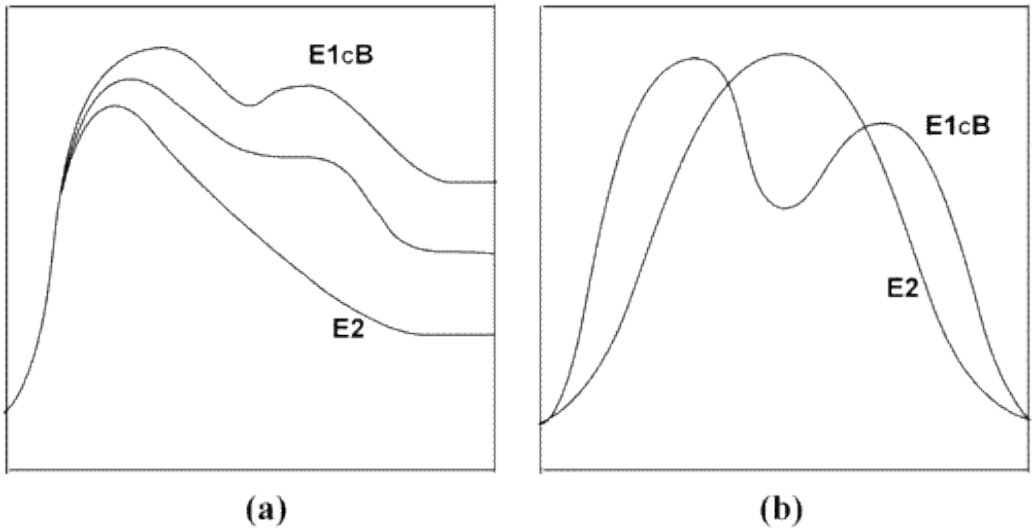
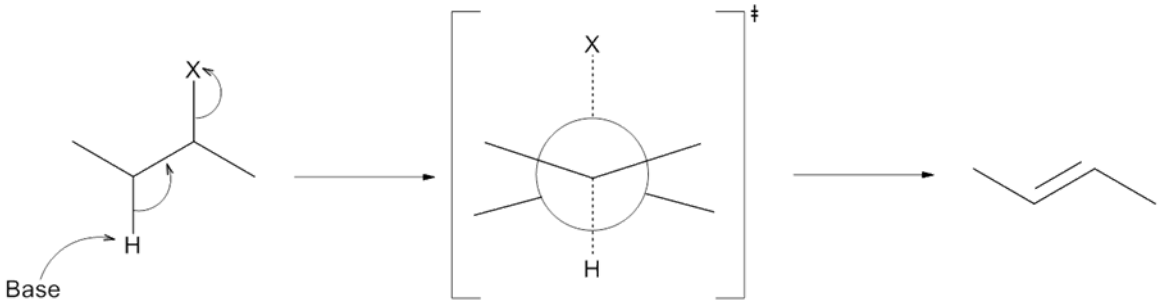


Figure 1.5:[20] Demonstration of the experimental difficulty in determining the difference between an E1cb and E2 reaction path: (a) reaction pathways that merge at the transition state, (b) two reaction pathways that demonstrate nonequivalent transition states.

Figure 1.5 demonstrates how a continuum of mechanism is possible from an E2 to an E1cb reaction pathway and how either the deprotonation or the leaving group departure can be more or less advanced at the transition state.[20] Figure 1.5 (a) illustrates how the carbanion intermediate, indicative of an E1cb mechanism, ceases to exist due to an increase in energy and can ultimately shift to resemble an E2 mechanism.[20] Figure 1.5 (b) is an example when the reaction conditions, e.g. solvent, have been changed such that the E1cb intermediate is stabilized and as a result there is a distinct difference in the kinetics of the reaction mechanisms.[20]

1.3.1.1 Elimination Bimolecular Mechanism (E2)

The concerted, or one step elimination mechanism is called an elimination bimolecular (E2) mechanism, in which the two substituents depart the substrate simultaneously to leave behind a new carbon-carbon π – bond. As seen in Scheme 1.2, a base abstracts the acidic β – hydrogen from the carbon adjacent to the C_α – X bond; a new C – C π bond is formed and the C_α – X bond is broken. The reaction mechanism involves three simultaneous bonding changes while an antiperiplanar transition state is formed.



Scheme 1.2: The E2 in which three bonds change simultaneously and an antiperiplanar transition state is formed (illustrated by a Newman projection).

An E2 mechanism follows second order kinetics, first order with respect to the substrate and first order with respect to the base. Often the E2 mechanism competes with the substitution nucleophilic bimolecular mechanism (S_N2). The difference between the two competing reaction pathways is whether the attacking species acts as a nucleophile and attacks the C_α, or acts as a Bronsted base and abstracts the H_β which leads to the formation of the alkene product. Evidence of an E2 mechanism over an S_N2 mechanism is determined through stereochemical examinations,[21] because the E2 reaction pathway is stereospecific. In the transition state of an E2 mechanism the proton that is abstracted by

the base and the leaving group typically have a dihedral angle of 180° and form an antiperiplanar transition state. This configuration is greatly favored over the syn-eliminations, because the molecule requires less energy to reach the transition state than it does to reach the eclipsed, syn-periplanar transition state (Figure 1.6).

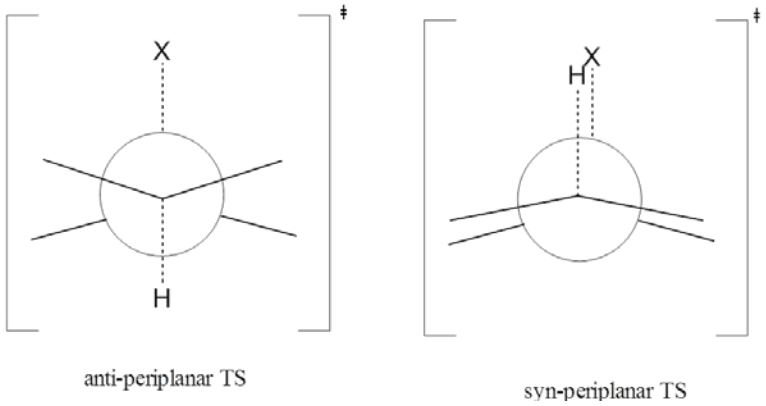
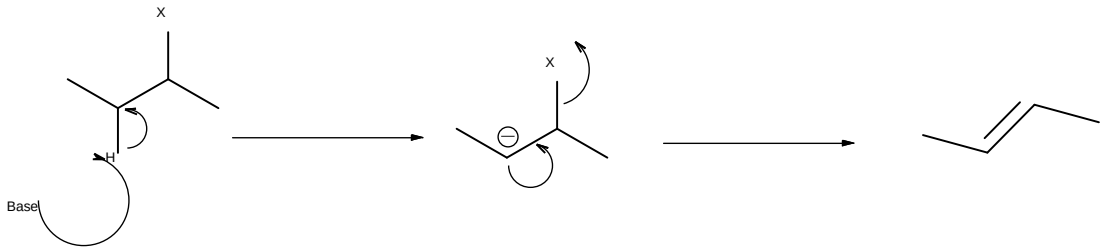


Figure 1.6: The H and the leaving group, *X*, can take the antiperiplanar configuration so that they depart in opposite directions or the syn-periplanar configuration in which they are eclipsed, with a dihedral angle of 0°.

1.3.1.2 Elimination Unimolecular Conjugate Base (E1cb)

The E1cb reaction pathway is a two-step process where the proton is abstracted first, and then the leaving group departs after the formation of a carbanion intermediate.



Scheme 1.3: The general mechanism for an E1cb reaction. Step one is the abstraction of the proton, forming the carbanion intermediate. Step two is the formation of the π – bond in conjunction with departure of the leaving group.

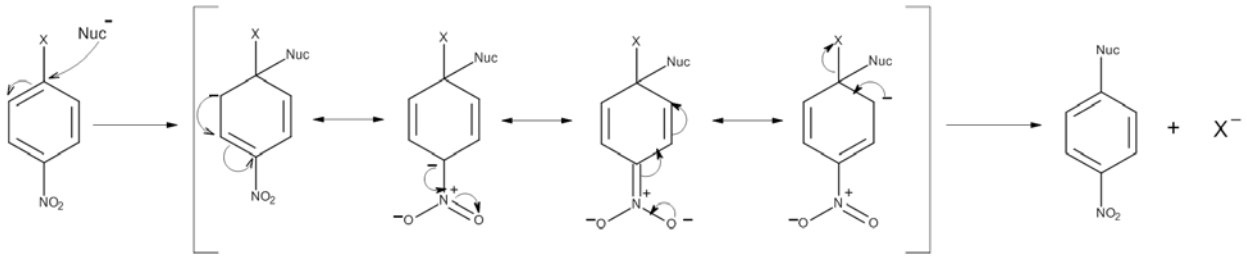
The mechanism is named E1cb because it is the conjugate base of the starting material from which the leaving group is departing. There are three different scenarios. (1) The carbanion is unstable and therefore will quickly reverse to the starting material, making step one reversible and the formation of the product the rate determining step, designated $E1cb_R$; (2) the formation of the carbanion is slow and the formation of the product is fast, making step one rate determining, $E1cb_I$, and lastly (3) the deprotonation step is the fast step and the formation of product is rate determining, because the carbanion is very stable, $E1cb_{irr}$ or irreversible. Kinetic studies to determine orders of reaction are not useful in distinguishing between an E1cb and an E2 mechanism. This is because scenario 1 and 2 are second order, while scenario 3 is first order.[21d, e, 22]

1.3.1.3 E2 vs. E1cb

If a substrate is predicted to undergo an E2 mechanism, structural changes to the substrate can have dramatic effects on the antiperiplanar transition state. For example, an electron withdrawing group to the α – carbon will be increase deprotonation in conjunction with a decrease in leaving group efficiency because the carbanion is better stabilized.[22-23] The same effect has been observed with the addition of β – carbon phenyl substituents.[19e, f] Additionally, mechanistic shifts are observed when the evolution of the carbanion intermediate is stabilized through favorable electrostatic interactions; e.g., through the use of ionic liquids.[18] Overall it is the extent of deprotonation of the substrate in the transition state of the rate determining step that is the true distinguishing feature of these reactions, complete in an E1cb and partial in an E2.

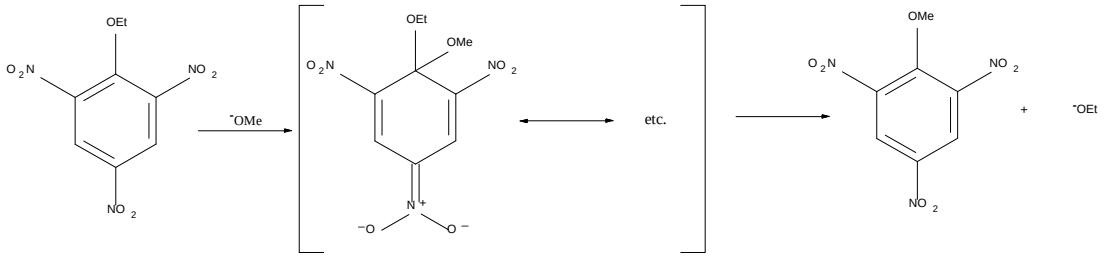
1.3.2 Nucleophilic Aromatic Substitution Reactions (S_NAr)

Nucleophilic aromatic substitution reactions (S_NAr) occur between an aromatic ring and a strong nucleophile. A leaving group is required to be ortho or para to a strong electron withdrawing group, such as a nitro group (Scheme 1.4).



Scheme 1.4: Mechanism for a nucleophilic aromatic substitution, where *Nuc* stands for nucleophile and *X* is the leaving group.

The S_NAr mechanism is a two-step process: step one is nucleophilic attack on the ipso carbon (the carbon with the leaving group) to form the Meisenheimer intermediate complex[24] (MIC), also called Meisenheimer-Jackson salts.[25] Step two is the cleavage of the C-X bond.[21e, 26] The formation of the MIC is usually rate determining and the reaction is almost always first order with respect to the nucleophile and first order with respect to the electrophilic aromatic ring, second order overall.[22] The electron withdrawing group positioned either at the ortho or para position to the leaving group increases the stability of the MIC by an improved delocalization of the charge within the aromatic ring, as illustrated in Scheme 1.4. The MIC was first isolated in 1902 for the reaction of 2,4,6-trinitrophenetole and the methoxide ion (Scheme 1.5).[24, 27] The MIC are commonly stable salts, whose structures have been proved and confirmed through nuclear magnetic resonance[28] (NMR) and X-ray crystallography.[29]



Scheme 1.5:[24] The abbreviated mechanism of the reaction between 2,4,6-trinitrophenole and methoxide ion.

Since the departure of the leaving group occurs in the second step and in most instances is not the rate determining step, different leaving groups will have little effect on the rate of reaction. Identical rates are not expected however, for an increase in the electronegativity of the leaving group will cause a decrease in the electron density at the

ipso carbon, resulting in faster attack by the nucleophile.[21e, 22] However, careful consideration must be taken in the interpretation of leaving group effects since the leaving group is on the carbon undergoing nucleophilic attack.

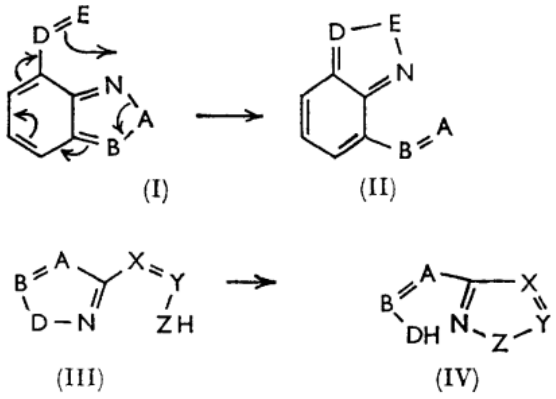
Solvent effects can play a large role in the rate of the S_NAr reaction. For instance, in aprotic solvents there is less extensive solvation of the nucleophile, allowing for an increase in rate when compared to the corresponding reaction in protic solvents.[27] Increased nucleophilicity has also been reported to occur when the nucleophiles are cyclic amines such as piperidine, morpholine and pyrrolidine, in the RTILs 1-butyl-3-methylimidazolium with the counterion hexafluorophosphate or tetrafluoroborate.[30] Furthermore, RTILs have been reported to support the evolution of the transition state in a S_NAr reaction between the ortho- and para-like nitrothiophene isomers and cyclic amines.[30b] A late transition state with an increased charge separation, i.e. more MIC character, is predicted to occur in RTILs; whereas an early TS with less charge separation, i.e., less MIC character, in protic solvents such as methanol. The RTIL effect can be directly attributed to favorable electrostatic interactions with the ions of the RTILs through the formation of a liquid clathrate, in addition to increased hydrogen bond donor and hydrogen bond acceptor abilities.

1.3.3 Mononuclear Heterocyclic Rearrangements (MHRs)

1.3.3.1 Introduction

Mononuclear heterocyclic rearrangement (MHR) reactions are intramolecular nucleophilic substitution reactions which can be categorized into two general classes, see Scheme 1.6.[31] The rearrangement of heterocycles containing suitable sidechains represents an interesting method for the synthesis of heterocyclic derivatives. For

examples, the rearrangement of (I) into (II) can lead to the syntheses of benzotriazoles, benzofurans, anthranils and indazoles.[31-32] The monocyclic rearrangement of (III) into (IV) has several possible combinations for *ABD* and *XYZ* and the associated ring systems: CCN (pyrazole), CNC (imidazole), NCC (imidazole), CCO (isoxazole), CNN (1,2,3-triazole), NNC (1,2,4-triazole), NCN (1,2,4-triazole), CNO (1,2,5-oxadiazole), NCO (1,2,40xadiazole and NNN (tetrazole).[31] There are at least 10^2 possible combinations for this type of intramolecular rearrangement, and if sulfur were included, there are many more possible configurations. However, in 1967 when Boulton and Katritzky recognized and categorized these intramolecular rearrangements, of the 100 possibilities only eight examples from the literature could be collected for the transformation of (III) into (IV), see Table 1.2 for the eight literature cited reactions.[31]



Scheme 1.6:[31] The two general categories for MHR reactions.

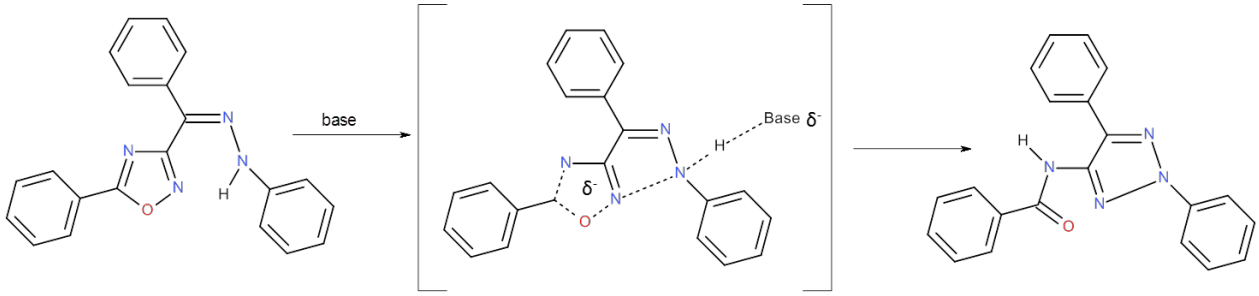
Examples of Monocyclic Rearrangements (III) → (IV)

ABD	*Starting Ring*	*XYZ*	*Finishing Ring*	*Reference*
CNO	1,2,5-Oxadiazole	CNO	1,2,5-Oxadiazole	[33]
CNO	1,2,5-Oxadiazole	CCO	Isoxazole	[34]
NCO	1,2,4-Oxadiazole	CNN	1,2,3-Triazole	[35]
NCO	1,2,4-Oxadiazole	CNO	1,2,5-Oxadiazole	[36]
CCO	Isoxazole	CNN	1,2,3-Triazole	[37]
CCO	Isoxazole	CNO	1,2,5-Oxadiazole	[38]
CCO	Isoxazole	NCN	1,2,4-Triazole	[39]
CCO	Isoxazole	NNN	Tetrazole	[40]

Table 1.2:[31] Examples of MHRs corresponding to the general transformation of (III) into (IV) seen in Scheme 1.6 above.

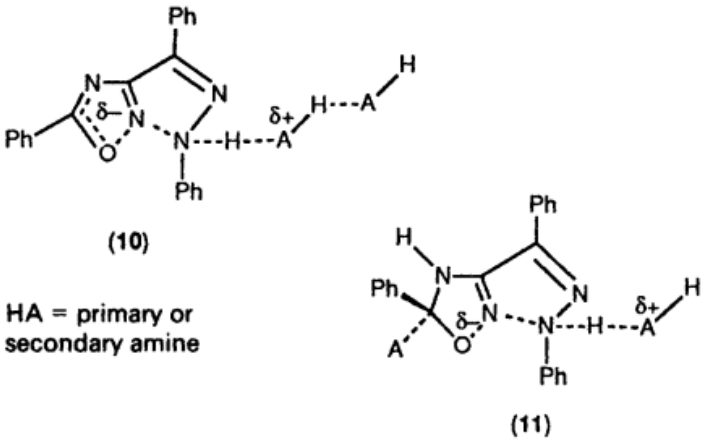
1.3.3.2 The Boulton-Katritzky Rearrangement

Specifically the heterocyclic rearrangement when the starting ring is a 1,2,4-oxadiazole derivative (*ABD* is *NCO)*, into the product ring 1,2,3-triazole derivative (*XYZ* is *CNN*) is of particular interest (Scheme 1.7), and is commonly referred to as the Boulton-Katritzky rearrangement.[30a]



Scheme 1.7: The MHR of the Z-phenylhydrazone: 3-benzoyl-5-phenyl-1,2,4-oxadiazole into the relevant 4-benzoylamino-2,5-diphenyl-1,2,3-triazole induced by amine deprotonation.

The particular intramolecular nucleophilic substitution reaction seen in Scheme 1.7 has been proposed to occur through a bicyclic quasi-aromatic, 10π electron transition state.[41] In the transition state, the N-N bond making and N-O bond breaking occur in a concerted manner. Altering reaction conditions can cause the position of this quasi-aromatic transition state to change along the reaction coordinate. For instance, substituents can affect the acidity of the arylhydrazone hydrogen and the corresponding nucleophilicity of the N atom. An increase in the reactivity of the Z-arylhydrazones results when bonded to strongly electron-donating groups, whereas an increase in N-H deprotonation can occur more readily when the arylhydrazone is bonded to substituents with a low electron-withdrawing effect.[42] Base (in most cases an amine) structure has been varied to analyze the effect on the rate of reaction: primary amines such as butylamine (BuA),[42] the secondary amine piperidine[41b] (Pip) and the tertiary amines triethylamine (TEA)[41b, 42] and 1,4-diazabicyclo[2.2.2]octane (DABCO)[42] have beeb used. In general it was discovered that the tertiary amines TEA and DABCO reacted too slowly to provide sufficient results, whereas the secondary and primary amines produce reactivity, but are strictly dependent on the nature of the solvent (Scheme 1.8 below).[42]



Scheme 1.8:[42] The bicyclic quasi-aromatic, 10π electron transition state of the MHR of Z-phenylhydrazone induced by amine deprotonation. (10) represents the transition state in apolar solvents, and (11) represents the transition state in polar solvents.

In apolar solvents, figure 10 in Scheme 1.8, the transition state is reported to occur via 'catalysis of catalysis' whereas in polar solvents such as acetonitrile or methanol, a significant contribution from the uncatalyzed pathway is observed.[30a, 42] In contrast, an alternative view would be to consider the same reaction in RTILs. The reactivity in RTILs has been reported to be due primarily to their ability to obtain and maintain an organized structure due to the extensive inter- and intramolecular favorable electrostatic interactions. This pre-organized structure is proposed to give the reagents an ordered deposition that is suitable for the reaction to occur, Figure 1.6.[41b]

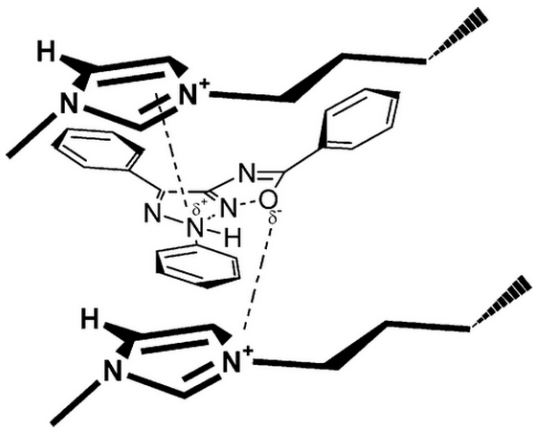


Figure 1.7:[41b] Pre-organized structure of RTILs of the MHR of the Z-phenylhydrazone: 3-benzoyl-5-phenyl-1,2,4-oxadiazole into the relevant 4-benzoylamino-2,5-diphenyl-1,2,3-triazole.

The reaction seen in Scheme 1.7 (Figure 1.6) has been investigated in the RTILs 1-butyl-3-methylimidazolium with tetrafluoroborate and hexafluorophosphate. The MHR had a significant increase in rate of reaction in RTILs when compared to molecular solvents, such as methanol. The effect is attributed to the increased favorable substrate-solvent and amine-solvent interactions; for instance when the amine abstracts the proton from the Z-phenylhydrazone, the amine will develop a positive charge. The generation of this positive charge will cause an increase in the hydrogen bond donor properties of any proton bound to this nitrogen atom and as a consequence a stronger β value of the solvent will preferentially stabilize the transition state more.[41b] This trend is supported by the basicity of the amines used where the reactivity increases in the order TEA < Pip < BuA.[41b, 42]

1.3.4 Summary

RTILs are able to affect the outcome of chemical reactions in a way that is completely different than what is normally observed in conventional organic solvents. In particular higher reactivity in RTILs is commonly observed and cannot be attributed strictly to an increased polarity of the solvent medium, but rather a complex characterization through the use of the Kamlet-Taft parameters is appropriate: E_T^N - solvating ability, π^* - dipolarity/polarizability, α – hydrogen bond acidity, and β – hydrogen bond basicity. RTILs cannot be considered simple solvents but should rather be considered as polymeric supramolecular fluids with regions characterized by "different polar character".[9, 41b]

Chapter 1 References

1. Reichardt, C.; Welton, T., *Solvents and Solvent Effects in Organic Chemistry*. 4 ed.; Wiley-VCH: 2011; p 692.
2. (a) Hallett, J. P.; Welton, T., Room-Temperature Ionic Liquids: Solvents for Synthesis and Catalysis. 2. *Chem. Rev.* **2011,** *111* (5), 3508–3576; (b) Weingaertner, H., Understanding ionic liquids at the molecular level: facts, problems, and controversies. *Angew. Chem. Int. Ed.* **2008,** *47* (4), 654-670; (c) Forsyth, S. A.; Pringle, J. M.; MacFarlane, D. R., Ionic Liquids—An Overview. *Aust. J. Chem.* **2004,** *57*, 113-119; (d) Welton, T., Room-Temperature Ionic Liquids. Solvents for Synthesis and Catalysis. *Chem. Rev.* **1999,** *99*, 2071-2083.
3. (a) Castner, E. W.; Wishart, J. F.; Shirota, H., Intermolecular Dynamics, Interactions, and Solvation in Ionic Liquids. *Acc. Chem. Res.* **2007,** *40* (11), 1217-1227; (b) Iwata, K.; Okajima, H.; Saha, S.; Hamaguchi, H., Local Structure Formation in Alkyl-imidazolium-Based Ionic Liquids as Revealed by Linear and Nonlinear Raman Spectroscopy. *Acc. Chem. Res.* **2007,** *40* (11), 1174-1181.
4. Freemantle, M., *Chem. Eng. News* **1998,** *76*, 32.
5. Bonhote, P.; Dias, A.-P.; Papageorgiou, N.; Kalyanasundaram, K.; Gratzel, M., Hydrophobic, Highly Conductive Ambient-Temperature Molten Salts. *Inorg. Chem.* **1996,** *35* (5), 1168-1178.
6. (a) Wasserscheid, P.; van Hal, R.; Bosmann, A.; EBer, J.; Jess, A. 224th American Chemical Society Meeting, Boston, Mass, Press, O. U., Ed. Boston, Mass, 2002; pp 57-69; (b) Rakita, P. E. 224th American Chemical Society Meeting, Boston, Mass, Press, O. U., Ed. Boston, Mass, 2002; pp 32-56.

7. Fuller, J.; Carlin, R. T.; de Long, H. C.; Haworth, D. J., *J. Chem. Soc. Chem. Commun.* **1994**, 299-300.

8. Wasserscheid, P.; Bosmann, A., *unpublished results.*

9. (a) Anderson, J. L.; Ding, J.; Welton, T.; Armstrong, D., *J. Am. Chem. Soc.* **2002,** *124*, 14247; (b) Crowhurst, L.; Mawdsley, P.; Perez-Alandia, J. M.; Salter, P.; Welton, T., *Phys. Chem. Chem. Phys.* **2003,** *5*, 2790.

10. (a) Kamlet, M. J.; Taft, R. W., *J. Am. Chem. Soc.* **1976,** *98*, 377-383; (b) Taft, R. W.; Kamlet, M. J., *J. Am. Chem. Soc.* **1976,** *98* (2886-2894); (c) Yokoyama, T.; Taft, R. W.; Kamlet, M. J., *J. Am. Chem. Soc.* **1976,** *98*, 3233-3237; (d) Kamlet, M. J.; Abboud, J. L.; Taft, R. W., *J. Am. Chem. Soc.* **1977,** *99*, 6027-6038.

11. Carmichael, A. J.; Seddon, K. R., Polarity study of some 1-alkyl-3-methylimidazolium ambient-temperature ionic liquids with the solvatochromic dye, Nile Red. *J. Phys. Org. Chem.* **2000,** *13* (10), 591–595.

12. (a) Pârvulescu, V. I.; Hardacre, C., Catalysis in Ionic Liquids. *Chem. Rev.* **2007,** *107*, 2615-2665; (b) van Rantwijk, F.; Sheldon, R. A., Biocatalysis in Ionic Liquids. *Chem. Rev.* **2007,** *107*, 2757-2785; (c) Welton, T., Ionic liquids in catalysis. *Coord. Chem. Rev.* **2004,** *248*, 2459-2477.

13. Han, X.; Armstrong, D. W., Ionic Liquids in Separations. *Acc. Chem. Res.* **2007,** *40* (11), 1079-1086.

14. (a) Haumann, M.; Riisager, A., Hydroformylation in Room Temperature Ionic Liquids (RTILs): Catalyst and Process Developments. *Chem. Rev.* **2008,** *108* (4), 1474-1497; (b) Zhang, Z. C., Catalysis in Ionic Liquids. *Advances in Catalysis* **2006,** *49*, 153-237.

15. (a) Lee, C., Diels-Alder reactions in chloroaluminate ionic liquids: acceleration and selectivity enhancement. *Tetra. Lett.* **1999,** *40* (13), 2461-2464; (b) Fischera, T.; Sethia, A.; Weltona, T.; Woolf, J., Diels-Alder reactions in room-temperature ionic liquids. *Tetra. Lett.* **1999,** *40* (4), 793-796; (c) Kumar, A.; Pawar, S. S., Converting exo-Selective Diels−Alder Reaction to endo-Selective in Chloroloaluminate Ionic Liquids. *J. Org. Chem.* **2004,** *69* (4), 1419-1420.

16. (a) 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; (b) Acevedo, O., Determination of Local Effects for Chloroaluminate Ionic Liquids on Diels-Alder Reactions. *J. Mol. Graphics Modell.* **2009,** *28*, 95-101.

17. 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.

18. Allen, C.; Sambasivarao, S. V.; Acevedo, O., *J Am. Chem. Soc.* **2013,** *31*, 1065-1072.

19. (a) Fontana, G.; Frenna, V.; Lamartina, L.; Natoli, R., *J. Phys. Org. Chem.* **2002,** *15*, 108-114; (b) Jia, Z. S.; Rudziński, J.; Paneth, P.; Thibblin, A., Borderline between E1cB and E2 Mechanisms. Chlorine Isotope Effects in Base-Promoted Elimination Reactions. *J. Org. Chem.* **2002,** *67* (1), 177-181; (c) Fontana, G.; Frenna, V.; Gruttadauria, M.; Natoli, M. C.; Noto, R., Kinetic study of base-promoted elimination reactions of some 1,1,1-trihalo-2,2-bis(dimethoxyphenyl)ethanes in alcoholic solutions. *J. Phys. Org. Chem.* **1998,** *11*

(1), 54-58; (d) Fontana, G.; Frenna, V.; Lamartina, L.; Natoli, M. C.; Noto, R., Kinetic study of methoxide-promoted elimination reactions of some 1,1,1-trichloro-2,2-bis(phenyl-substituted)ethanes. *J. Phys. Org. Chem.* **2002,** *15* (2), 108–114; (e) Gandler, J. R.; Storer, J. W.; Ohlberg, D. A. A., The E2-E1cB borderline: elimination reactions of 2-(2,4,6-trinitrophenyl)ethyl halides in aqueous solution. *J. Am. Chem. Soc.* **1990,** *112* (21), 7756–7762; (f) Gandler, J. R.; Jencks, W. P., General base catalysis, structure-reactivity interactions, and merging of mechanisms for elimination reactions of (2-arylethyl)quinuclidinium ions. *J. Am. Chem. Soc.* **1982,** *104* (7), 1937-1951; (g) William H. Saunders, J., *Accounts of Chemical Research* **1976,** *9*, 19-25.

20. Jia, S. Z.; Rudzinski, J.; Paneth, P.; Thibblin, A., *J. Org. Chem.* **2002,** *67*, 177-181.

21. (a) Bartsch, R. A.; Zavada, J., *J. Chem. Rev.* **1980,** *80*, 453; (b) Coke, J. L., *Sel. Org. Transform.* **1972,** *2*, 269; (c) Sicher, J., *Angew. Chem. Int. Ed.* **1972,** *11*, 200; (d) Saunders, J. W. H.; Cockerill, A. F., *Mechanisms of Elimination Reactions*. Wiley: NY, 1973; (e) Smith, M. B.; March, J., *March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*. 6th ed.; Wiley: New Jersey, 2007.

22. Anslyn, E. V.; Dougherty, D. A., *Modern Physical Organic Chemistry*. University Science Books: California, 2006.

23. More O'Ferrall, R. A., *J. Chem. Soc. B.* **1970**, 274.

24. Meisenheimer, J., *Liebigs Ann. Chem.* **1902,** *323*, 205.

25. (a) Jackson, C. L.; Gazzolo, F. H., *Am. Chem. J.* **1900,** *23*, 376; (b) Jackson, C. L.; Earle, R. B., *Am. Chem. J.* **1903,** *29*, 89.

26. Wu, Z.; Glaser, R. J., *J. Am. Chem. Soc.* **2004,** *126*, 10632.

27. Terrier, F., *J. Am. Chem. Soc.* **1982,** *82* (2), 78-147.

28. (a) Crampton, M. R.; Gold, V. J., *J. Chem. Soc. B.* **1966**, 893; (b) Buncel, E.; Crampton, M. R.; Strauss, M. J.; Terrier, F., *Electron Deficient Aromatic- and Heteroaromatic-Base Interactions*. Elsevier: NY, 1984.

29. (a) Destro, R.; Gramaccioli, C. M.; Simonetta, M., *Acta. Crystallogr.* **1968,** *24* (1369); (b) Ueda, H.; Sakabe, M.; Tanaka, J.; Furusaki, A., *Bull. Chem. Soc. Jpn.* **1968,** *41*, 2866; (c) Messmer, G. G.; Palenik, G. J., *Chem. Commun.* **1969**, 470.

30. (a) D'Anna, F.; Frenna, V.; Noto, R.; Pace, V.; Spinelli, D., *J. Org. Chem.* **2005,** *70*, 2828-2831; (b) D'Anna, F.; Frenna, V.; Noto, R.; Pace, V.; Spinelli, D., *J. Org. Chem.* **2006,** *71*, 5144-5150.

31. Boulton, A. J.; Katritzky, A. R.; Hamid, A. M., *J. Chem. Soc. C* **1967**, 2005-2007.

32. (a) Boulton, A. J.; Ghosh, P. B.; Katritzky, A. R., *J. Chem. Soc. B.* **1966**, 1004; (b) Anastas, P. T.; Warner, J. C., *Green Chemistry: Theory and Practice*. Oxford University Press: New York, 1998; p 30.

33. Ponzio, G.; Biglietti, F., *Gazzetta* **1933,** *63*, 159.

34. Cusmano, S.; Giambrone, S., *Gazzetta* **1951,** *81*, 499.

35. Gramantieri, P., *Gazzetta* **1935,** *65*, 102.

36. (a) Ponzio, G.; Avogadro, L., *Gazzetta* **1923,** *53*, 318; (b) Ponzio, G., *Gazzetta* **1931,** *61*, 138.

37. (a) Ajello, T.; Cusmano, S., *Gazzetta* **1940,** *70*, 770; (b) Ajello, T.; Tornetta, B., *Gazzetta* **1947,** *77*, 332.

38. (a) Ajello, T., *Gazzetta* **1937,** *67*, 779; (b) Ajello, T.; Cusmano, S., *Gazzetta* **1938,** *68*, 792; (c) Ajello, T.; Cusmano, S., *Gazzetta* **1939,** *69*, 391.

39. Kano, H.; Yamazaki, E., *Tetrahedron* **1964,** *20*, 159.

40. Kano, H.; Yamazaki, E., *Tetrahedron* **1964,** *20*, 461.

41. (a) Bottoni, A.; Frenna, V.; Lanza, C. Z.; Macaluso, G.; Spinelli, D., *J. Phys. Chem. A.* **2004,** *108*, 1731-1740; (b) D'Anna, F.; Frenna, V.; Noto, R.; Pace, V.; Spinelli, D., *J. Org. Chem.* **2006,** *71*, 9637-9642.

42. Frenna, V., *J. Chem. Soc. Perkin Trans. 2* **1990,** *2*, 1289-1295.

Chapter 2

Computational Methods

2.1 Mixed Quantum and Molecular Mechanics (QM/MM)

In the modeling of chemical systems to elucidate thermodynamics, kinetics, mechanistic and chemical behavior, the number of atoms in the system dramatically influences the computational methods utilized. For example, quantum mechanics (QM) calculates the electronic motion of a system. If the chemical system of interest is composed of an organic solute surrounded by a condensed phase medium, the vast number of particles (electrons) that would have to be considered in order to describe the complete region using QM may be computationally unreasonable. Even if a portion of the electrons are partially treated, as in the semiempirical methods, the large number of particles considered would still be computationally expensive. Therefore a balance of time and computational efficiency must be reached in order to properly and realistically model large systems. The combined quantum mechanical and molecular mechanical (QM/MM) methodology[1] provides a computational approach in which the large chemical system of interest is divided into two subsystems, i.e. QM and MM, (see Figure 2.1) such that the total energy of the chemical system can be divided into three terms (equation 2.1).

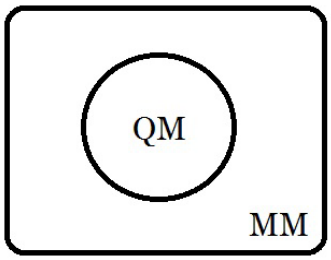


Figure 2.1: Partition of the molecular system of interest within a combined quantum mechanical and molecular mechanical (QM/MM) calculation. The smaller quantum mechanical (QM) region is surrounded by the larger molecular mechanical (MM) region.

$$E_{total} = E_{QM} + E_{MM} + E_{QM/MM} \quad (2.1)$$

The combined quantum mechanical and molecular mechanical (QM/MM) methodology was initially developed by multiple groups, including Warshel, Field, Gao, Merz and others.[1a-h] In the late 1970s, Jorgensen developed the broad-spectrum molecular modeling software *BOSS*, which stands for biochemical and organic simulation system.[2] *BOSS* can perform molecular mechanics (MM) calculations, semiempirical calculations, Monte Carlo (MC) statistical mechanics simulations, and mixed QM/MM calculations. The QM/MM hybrid methodology has been studied extensively on organic and bioorganic reactions,[1a, b, 1i-o] ion solvation,[1p] kinetic isotope effects,[1q] enzymatic reactions,[1a, b, 1r-t] ion-molecule association in solution,[1u] and conformational equilibria.[1v] The combined QM/MM methodology provides a practicable platform for the modeling of structures and the determination of energetics and reaction pathways for complex systems that are in the condensed phase.[1a-g, 1i-l, 1p-am]

As seen in equation 2.1, the total energy of the entire chemical system is determined as the sum of three variables, where E_{QM} is the energy of the QM region; E_{MM} is the energy of the molecular mechanics (MM) region and $E_{QM/MM}$ accounts for the energy of all interactions between the QM and classically described systems. The smaller partitioned region can be described using any QM method, e.g., ab *initio*, density functional theory (DFT), and semiempirical methods; however, our preference is PDDG/PM3 semiempirical method,[3] see equation 2.2, which is to be discussed in the next section. The larger partitioned region is described using a force field, e.g., CHARMM,[4] (chemistry at Harvard macromolecular mechanics), AMBER[5] (assisted model building with energy refinement) and our preference the OPLS-AA[6] (optimized potential for liquid simulations – all atom) force field, see equation 2.3.

$$E_{QM} = (PDDG/PM3) \tag{2.2}$$

$$E_{MM} = (OPLS - AA) \tag{2.3}$$

The total effective Hamiltonian for the chemical system described in Figure 2.1 is determined from the QM, MM and QM/MM parts as described by equation 2.4 below:

$$H_{total} = H_{QM} + H_{MM} + H_{QM/MM} \tag{2.4}$$

where H_{total} is the effective Hamiltonian for the entire chemical system of interest and H_{QM} is the quantum mechanical Hamiltonian which accounts for the full interaction energy of all QM particles with one another. H_{MM} is the molecular mechanical Hamiltonian which accounts for the full interaction energy of all classically described molecules with one another and $H_{QM/MM}$ is the Hamiltonian that accounts for the energy of all interactions between one quantum mechanical particle and one classically described molecule.

The classically treated region typically represents an environment such as that of a macromolecule or hundreds of explicitly simulated solvent molecules. Since molecular mechanics is represented by potential energy functions, it is not an appropriate method for adequately describing chemical reactions, i.e., bond-making or bond-breaking. The QM region, which typically consists of a small number of atoms relative to the molecular mechanics region, is the portion of the molecular system of interest in which significant changes in electronic structure occur along the course of a reaction coordinate. In *BOSS*, all molecules are generally represented with a Z-matrix. A Z-matrix is a coordinate map for the atoms defined by the atomic numbers, bond lengths, bond angles, and dihedral angles. The flexibility of the molecule(s) can be full or limited simply by specifying which internal coordinates are variable in the Z-matrix.

2.2 Optimized Potentials for Liquid Simulations – All Atom Force Field (OPLS-AA FF)

All atom force fields (AA FF) differ from united atom force fields (UA FF) in that AA FF represent all atoms explicitly, whereas UA FF combine some or all of the atoms in a molecule into ‘pseudo-atoms’ or ‘united atoms’. For example, an AA FF model of butane would be represented as fourteen explicit atoms, while in the UA FF model butane would be represented only as a four center model where the carbon atoms in the chain act as the ‘pseudo-atoms’ or ‘united atoms’. Typically in UA FF the only hydrogen atoms that are represented implicitly are those that are attached to aliphatic carbons. Hydrogen atoms that are bonded to aromatic rings and heteroatoms are modeled explicitly.[7] Therefore, the UA FF models have far fewer interaction sites than AA FF models. Since the number of nonbonded interactions scales with the square of the number of interaction sites, the UA FF method can be viewed as being computationally advantageous.

However, decreasing the number of interaction sites can directly affect the nonbonded energetics of the system, especially when the MM system of interest is a solvent. Furthermore an AA FF such as OPLS-AA, allows for more flexibility in the charge distributions within a molecule and also allows for an increase in the angle bending and torsional energetic contributions within the molecules.[7]

The larger MM subsystem described by the OPLS-AA FF ignores electronic motion and thus cannot be used to describe any electronic distribution of a molecule. Therefore the total energy of the solvent, the MM sub-region, is evaluated as a sum of the individual energies of four essential components: bond stretching, angle bending, torsional rotations and nonbonded interactions, see equations below. These four components describe all of the intra- and inter- molecular forces that are occurring within the MM sub-region.

$$E_{total} = E_{bond} + E_{angle} + E_{torsional} + E_{nonbonded} \tag{2.5}$$

$$E_{bond} = \sum_{bonds} K_r\left(r - r_{eq}\right)^2 \tag{2.6}$$

$$E_{angle} = \sum_{angles} K_\theta\left(\theta - \theta_{eq}\right)^2 \tag{2.7}$$

$$E_{torsional} = \sum_i \frac{V_1^i}{2}\left[1 + \cos(\emptyset_i + f_i 1)\right] + \frac{V_2^i}{2}\left[1 - \cos(2\emptyset_i + f_i 2)\right] + \frac{V_3^i}{2}\left[1 + cos(3\emptyset_i + f_i 3)\right] \tag{2.8}$$

$$E_{nonbonded} = \sum_i^{on\ a}\sum_j^{on\ b}\left[\frac{q_i q_j e^2}{r_{ij}} + 4\epsilon_{ij}\left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}}\right)\right] f_{ij} \tag{2.9}$$

The bond stretching, E_{bond}, and angle bending, E_{angle}, energetic contributions to the total energy of the MM sub-region are shown in equations 2.6 and 2.7 respectively. The torsional term, $E_{torsional}$ in equation 2.8, is expressed as a cosine series expansion and describes the energetic barriers of rotation about chemical bonds, i.e. changes in the dihedral angle. Lastly in equation 2.9, $E_{nonbonded}$ is the energy contribution from all nonbonded interactions between all pairs of explicitly modeled OPLS-AA solvent atoms.

A simple harmonic potential is used to describe the bond stretching and angle bending energetic contributions. These energy contributions are characterized by a force constant, K_r and K_θ, and a reference value, r_{eq} and θ_{eq}. The reference value is the value of the bond or angle that would be adopted when all other terms in the OPLS-AA FF are set to zero. The reference bond length/angle is different from an equilibrium bond length/angle in that the equilibrium value is obtained from the minimum energy structure when all other terms in the OPLS-AA FF contribute. As equations 2.6 and 2.7 show, the energy will vary with the square of the shift from this reference bond length and/or reference bond angle. This relationship is known as Hooke's Law, an approximation that states that the restoring force is proportional to the displacement from the equilibrium value. Figure 2.2 shows an example of two masses, i.e. two atoms, connected by a spring described by the appropriate force constant.

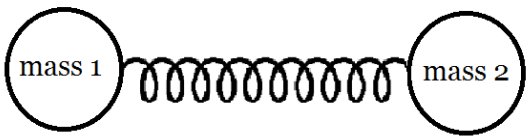


Figure 2.2: Masses 1 and 2 represent atoms described by Hooke's law.

When quantitatively comparing the force constant of bond stretching to the force constant of angle bending, it is chemically intuitive for K_r to be greater than K_θ. The forces between bonded atoms are very strong and considerable energy is needed to cause a bond lenght to deviate significantly from the equilibrium value, whereas less energy is required to distort an angle from an equilibrium value. Therefore the magnitude of the required force to deviate from the equilibrium value is reflected in the respective force constant.

Even though true bond stretching does not follow a harmonic potential, the more realistic Morse potential is not commonly used to model these energetic contributions (Figure 2.3). The harmonic approximation is used because in MM bond lengths are not perturbed very far from their equilibrium value, hence a vibrating ground state molecule will deviate from the equilibrium value only slightly.

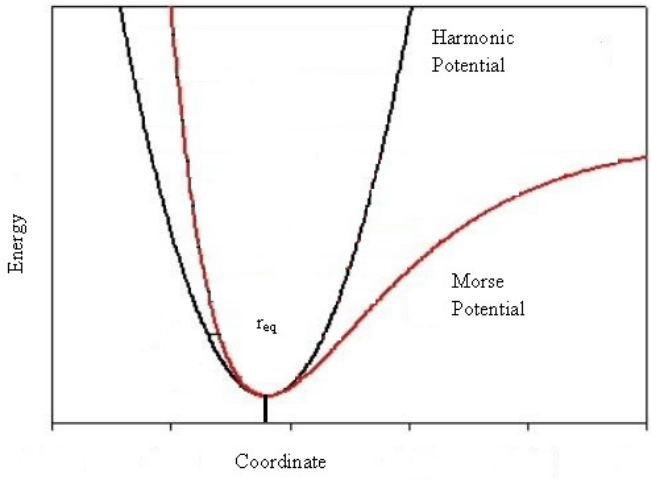


Figure 2.3: A graphical representation of the harmonic potential (black line) and the Morse potential (red line). r_{eq} represents the equilibrium bond length.

As depicted in Figure 2.3, in addition to correctly modeling the energetics at or near the equilibrium bond length, the Morse potential describes very strong attractions which occur during the creation of a bond and the energetics of complete molecular dissociation. This behavior is not modeled in the MM OPLS-AA FF, but requires fewer parameters, and is used over the Morse potential to decrease the computational cost.

The torsional term, $E_{torsional}$ which is seen in equation 2.8, is an intramolecular term. This term describes the energy change as a chemical bond is rotated, thus changing the associated dihedral angle.

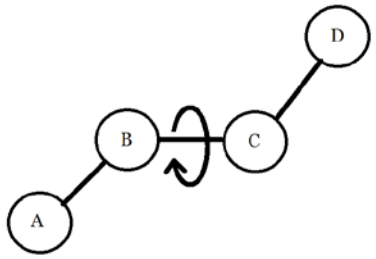


Figure 2.4: A dihedral angle is formed between four adjacent atoms and the corresponding three bonds. The angle between atoms A and D form the dihedral angle. In this example, the dihedral angle A-B-C-D is equal to 180°.

Equation 2.8 is expressed as a cosine series expansion where Φ_i is the dihedral angle of interest and V_1^i, V_2^i and V_3^i are the coefficients in the Fourier series which describe the "barrier height" or energy of rotation about the specific dihedral angle. Whereas f_i1, f_i2 and f_i3 are the phase angles which describe when Φ_i passes through a minimum value. A Fourier series is an expansion of a periodic function, which is a function that repeats its values in regular intervals, in terms of an infinite sum of sine and/or cosine functions. A

Fourier series makes use of the orthogonality relationship of these functions, meaning that no combination of Fourier components can be used to replace another component. This relationship provides an extremely useful way to break up a periodic function into a set of simpler terms. These simplified terms can be solved individually, plugged back into the function and then recombined to obtain the solution to the original problem. The torsional energy is therefore a sum of this series for each dihedral angle that is possible. Barriers of rotation about chemical bonds are fundamental to understanding structural properties of a molecule. Most of the variation in a structure and associated relative energies are due to the complex interplay between torsional energetics and nonbonded contributions.

Lastly, $E_{nonbonded}$ in equation 2.9 describes the intermolecular and intramolecular nonbonded interactions between all pairs of atoms. In 1785 the French physicist Charles-Augustin de Coulomb discovered an inverse relationship of the force between two charged particles, q_i and q_j in equation 2.9, and the square of their distance, r_{ij}, which is now referred to as Coulomb's Law. Today this formalism is used to calculate the electrostatic interactions that are "through-space", meaning intermolecular interactions, which occur in most chemical systems of interest. The same expression is used for intramolecular nonbonded interactions between all pairs of atoms ($i > j$) separated by three or more bonds. However, electrostatics cannot account for all of the nonbonded interactions taking place inside of a chemical system. For example, there are also dispersive interactions and exchange repulsive interactions. These van der Waals interactions are modeled using the Lennard-Jones 12-6 potential[8] to describe the attractive and/or repulsive forces between molecules, or between groups within the same

molecule, other than those due to bond formation or to electrostatic interactions. These potentials also include dipole-dipole, dipole-induced dipole and instantaneous induced dipole-induced dipole interactions.

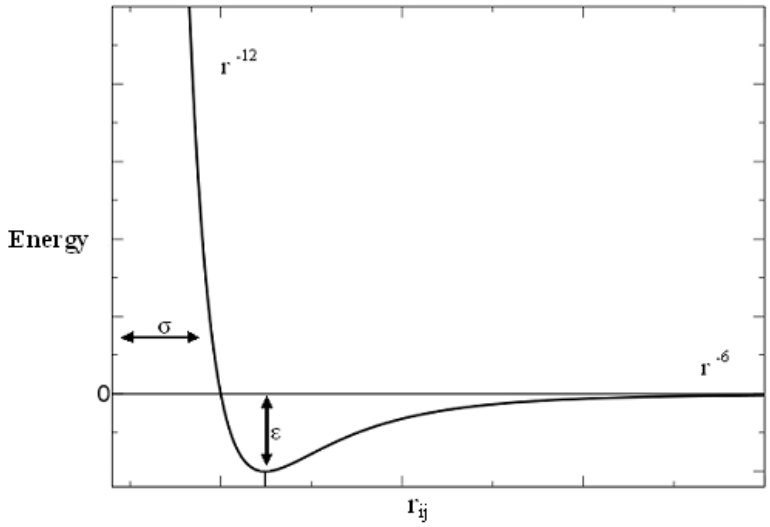


Figure 2.5: The 12-6 Lennard-Jones potential is shown graphically to demonstrate the relationship between two particles over distances, r_{ij}, and the corresponding relative energetics.

A graphical representation of the van der Waals forces is shown in Figure 2.5. As two atoms, q_i and q_j, approach each other, favorable interactions decrease the overall energy. When the atoms become close enough to the point where the repulsive, r^{-12}, and attractive, r^{-6}, forces balance each other out, energy ceases to be released and as a result the atoms reach a minimum energy and a chemical bond is formed. The distance from one nucleus to the other nucleus in a stable bond is known as the equilibrium bond length and this is the r_{ij} value found at the bottom of the well, where ε is the well depth. However, when the atoms get too close together, the energy begins to increase drastically due to repulsive forces between the two nuclei and the two sets of electrons. These short

range repulsive forces are often referred to as “exchange forces” or as “overlap forces”. The overall effect is to reduce the electrostatic repulsion between pairs of electrons, forbidding them to occupy the same region of space. These repulsion forces are a direct result of the Pauli Exclusion Principle, which prohibits any two electrons in a system from having the same set of quantum numbers. Conversely, when the atoms are far apart from each other, large values of r_{ij}, the attractive and repulsive forces are no longer significant and the energy of interaction begins to smooth out, where the collision diameter σ describes the atomic separation for which the energy is zero. At large distances r_{ij} there are dispersion forces which are due to instantaneous dipoles that arise during the fluctuations in the electron clouds. An instantaneous dipole in one molecule can induce a dipole in a neighboring molecule, giving rise to an attractive dispersive effect. In order to use equation 2.9 for both the intramolecular and intermolecular nonbonded interactions described above, for the OPLS-AA FF a scaling factor f_{ij} = ½ was determined to be needed for 1,4 interactions.[6a] Furthermore, equations 2.10 and 2.11 represent the standard combining rules that are used in the OPLS-AA FF.[6a]

$$\epsilon_{ij} = \left(\epsilon_{ii}\epsilon_{jj}\right)^{\frac{1}{2}} \tag{2.10}$$

$$\sigma_{ij} = \left(\sigma_{ii}\sigma_{ij}\right)^{\frac{1}{2}} \tag{2.11}$$

When the OPLS-AA FF was originally created, it differed from other popular force fields, such as the AMBER FF (Assisted Model Building with Energy Refinement force field) developed by Peter A. Kollman and coworkers,[5] in that the charges for the OPLS-AA FF were empirically based. Formerly the charges on the MM atoms were determined from appropriate fitting to reproduce properties of 34 different organic

liquids.[6a] With force fields such as AMBER, the charges for the MM modeled atoms are obtained as needed from fitting to the electrostatic potential surfaces derived from *ab initio* 6-31G(d) calculations.[6a] However, in the most current versions of the OPLS-AA FF, the MM charges are derived in a fashion similar to the AMBER FF. Currently, a notable difference between the OPLS-AA FF and the AMBER FF is that the AMBER FF rotation about a bond is described by a general potential that depends exclusively upon the atom types that comprise the central bond and not upon the atom types of the terminal bonds. Including more than one term in the torsional expansion, such as in the OPLS-AA FF, can potentially describe the energetic barrier more accurately.

Transferability is a key attribute of a successful force field because this allows for the parameters of the force field to be applied to a large range of systems that are of the same family. Even though the OPLS-AA force field cannot accurately describe the electronic motions within a molecule, the fact that the MM region of interest is composed of ground state explicitly represented solvent molecules allows for this methodology to be used.

2.3 Introduction to Semiempirical Quantum Mechanical Approximations

In the hybrid QM/MM methodology as depicted in Figure 2.1.1, the smaller inner sub-region of the chemical system of interest is the QM region. The QM region typically consists of a small number of atoms relative to the larger molecular mechanics region. This smaller region is the portion of the molecular system of interest in which significant changes in electronic structure occur along the course of a reaction coordinate. The QM region is described using the PDDG/PM3[3] method developed by Jorgensen et *al.* in 2002.

Since the QM region may contain an organic substrate upwards of 100 atoms, a more efficient semiempirical methodology is preferred to model the QM region over an *ab initio* method for greater efficiency.

The frameworks of all semiempirical molecular orbital methods are based on the Roothaan-Hall equations. Equations 2.12 – 2.15 are the Roothaan-Hall equations for a closed-shell system with *N* electrons in *N/2* orbitals.

$$\mathbf{FC} = \mathbf{SCE} \tag{2.12}$$

$$F_{\boldsymbol{\mu\upsilon}} = H_{\mu\upsilon}^{core} + \sum_{\lambda=1}^{K}\sum_{\sigma=1}^{K} P_{\lambda\sigma}\left[(\mu\upsilon|\lambda\sigma) - \frac{1}{2}(\mu\lambda|\upsilon\sigma)\right] \tag{2.13}$$

$$P_{\lambda\sigma} = 2\sum_{i=1}^{N/2} c_{\lambda i}c_{\sigma i} \tag{2.14}$$

$$H_{\mu\upsilon}^{core} = \int d\,\upsilon_1 \emptyset_\mu(1)\left[-\frac{1}{2}\nabla^2 - \sum_{A=1}^{M}\frac{Z_A}{|r_1 - R_A|}\right]\emptyset_\upsilon(1) \tag{2.15}$$

During an *ab initio* calculation, all integrals are calculated using the Fock operator (equation 2.13) regardless whether the atomic orbitals (ϕ_μ and ϕ_υ are one of *K* atomic orbitals) are on the same atoms, on atoms bonded to each other, or on atoms that are not formally bonded. The Fock matrix, **F**, is a *K* x *K* symmetric square matrix composed of the Fock operator elements. The elements of the *K* x *K* Fock matrix in equation 2.13 are written as the sum of core, Coulomb and exchange contributions. The core contributions to the Fock operator are described in equation 2.15. Matrix **C**, is composed of the coefficients $c_{\lambda i}$ and $c_{\sigma i}$. $P_{\lambda\sigma}$ (equation 2.14) is the charge density matrix, whose elements are a summation of the **C** matrixs elements. **E** is a diagonal matrix whose elements are the orbital energies and **S** is the overlap integral.

To reduce the computational demand of these QM calculations, semiempirical methods such as PDDG/PM3 only consider the valence electrons and subsume the core electrons into the nuclear core. The reasoning for such an approximation is straightforward; in chemical phenomena such as bonding, only the valence electrons are involved and therefore the contribution from the core electrons can be ignored. Only considering the valence electrons is the first approximation that allows for semiempirical methods to decrease the computational demand when compared to the corresponding *ab initio* calculations. Additionally, a common feature of semiempirical methods is to set the overlap matrix, **S**, equal to the identity matrix, **I**, making all diagonal elements equal to 1 and all off diagonal elements equal to zero. It is important to note that in the calculation of the Fock matrix elements, setting the overlap matrix equal to the identity matrix does not mean that all overlap integrals are set to zero. It is imperative even in the simplest semiempirical methods, to include some of the overlap integrals in order to correctly model chemical phenomena.

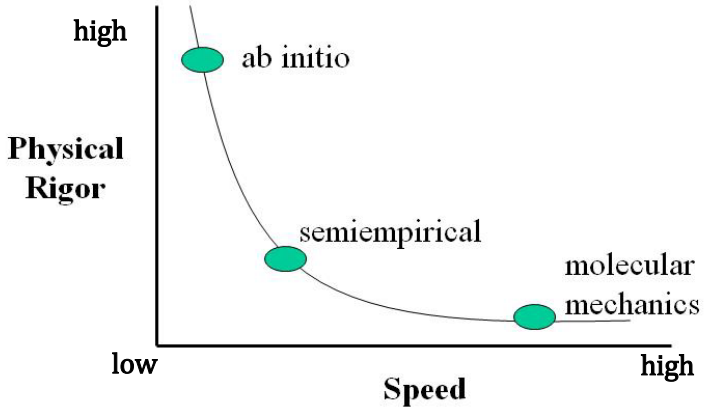


Figure 2.6: The approximations of semiempirical methods not only allow for the system of interest to grow in size, but also can greatly reduce the computational time when compared to strict *ab initio* calculations.

Sir John Anthony Pople developed semiempirical methodologies; he earned his doctoral degree in mathematics in 1964 from Cambridge University and was awarded the Nobel Prize in Chemistry in 1998. As a mathematician, he developed quantum mechanical methods that simplified the Roothaan-Hall equations but were parameterized to match *ab initio* results for a relatively small set of molecules. These semiempirical methodologies are called complete neglect of differential overlap (CNDO),[9] intermediate neglect of differential overlap (INDO)[10] and the neglect of diatomic differential overlap (NDDO).[9e]

2.3.1 Complete Neglect of Differential Overlap (CNDO), Intermediate Neglect of Differential Overlap (INDO) and Neglect of Diatomic Differential Overlap (NDDO)

CNDO[9] was the first method to implement the zero differential overlap (ZDO) approximation in a sensible manner because a major flaw with the ZDO approximation

was that the overlap between pairs of different atomic orbitals was set to zero. Neglecting the overlap with atomic orbitals on different atoms chemically makes little sense, considering that bonding and other chemical phenomena occur with atomic orbitals that are centered on different atoms. The semiempirical QM method CNDO corrects this initial harsh approximation by describing the overlap of the charge densities of different atomic orbitals by a parameter γ_{AB}. γ_{AB} depends only on the nature of atoms A and B and the distance between A and B, not the type of atomic orbitals that are found on atom A or atom B. This parameter can be considered the average electrostatic repulsion between an electron centered on atom A and an electron centered on atom B. Furthermore, if both of the electrons happen to be on the same atom, then this parameter describes the electron-electron repulsion on that atom. CNDO is recognized as the first semiempirical method in the long lineage of semiempirical methods; however CNDO had some important limitations and serious errors. First, CNDO models two neutral atoms as having a high attraction for one another, even when the two atoms are separated by several angstroms. Furthermore, the calculated diatomic equilibrium distances were too short and the corresponding dissociation energies were therefore overestimated. These incorrect modeled behaviors are a result of the electrons on atom A penetrating the valence shell of atom B, resulting in an overall nuclear attraction.[9b, c]

The next major development in semiempirical methodology was the intermediate neglect of differential overlap (INDO).[10] INDO corrected the CNDO approximation which had no allowances for the fact that the interaction between two electrons depended upon the electrons' respective spins. INDO therefore includes monoatomic differential overlap for atomic orbitals that are on the same atom. Including the monoatomic

differential overlap enables the interaction between the two electrons on atom A with parallel spins to have a lower energy than the comparable interaction between electrons with paired spins. As a direct result, the INDO approximations are similar to the approximations of CNDO with the principal difference being that all of the one center exchange integrals are retained.[10] However, the inclusion of these integrals was found to give little to no improvement on calculated equilibrium bond angles, when compared to corresponding CNDO calculations. INDO calculations require a trivial amount of additional computational time when compared to the corresponding CNDO calculation. Yet, INDO can be exploited to provide an improved description of open-shell systems, such as computing the energetics of radicals, which CNDO could not do. Additionally, INDO can be used for the description of singlet and triplet states of small molecules.[10]

The last semiempirical methodology developed by John A. Pople is called neglect of diatomic differential overlap (NDDO).[9e] The NDDO method is less approximate, where only the differential overlap between atomic orbitals on different atoms is neglected. This means that all two electron repulsion integrals which depend on the overlapping of charge densities of different atomic orbitals are equal to zero. All remaining two electron interactions are the two center type, with one electron associated with each atom. When directly compared to CNDO or INDO, there is a remarkable increase in the computational demand of the corresponding NDDO calculation. This is a direct consequence of the increased number of two-electron, two-center integrals that are retained in the NDDO calculation. The number of the integrals retained for an NDDO calculation is increased by a factor of 100 for each pair of heavy atoms in the chemical system when directly compared to the corresponding CNDO or INDO calculations.

2.3.2 Modified Intermediate Neglect of Differential Overlap (MINDO/3), Modified Neglect of Differential Overlap (MNDO), Austin Model 1 (AM1) and Third Parameterization of Modified Neglect of Diatomic Overlap (PM3)

The CNDO, INDO and NDDO semiempirical methodologies developed by Pople and coworkers as originally formulated, are in little use for today's computational problems. However, Pople and coworkers paved the way for Michael J. S. Dewar and colleagues by demonstrating how a logical series of approximations can be used to formulate new methods which perform in a fraction of the time required to solve the full Roothann-Hall equations. Michael Dewar was originally trained as an organic chemist but is probably most famously known for the development of the semiempirical QM methods called the modified intermediate neglect of differential overlap (MINDO/3),[11] the modified neglect of differential overlap (MNDO), [12] and the Austin Model 1 (AM1).[13] Additionally, a former student in the Dewar group, James J. P. Stewart, developed the popular third parameterization of the modified neglect of differential overlap (PM3) method.[14] These methodologies were fashioned with the understanding that previous semiempirical methods were inclined to be problem specific, and a more generalized application of semiempirical QM methods was needed. In addition, the previously mentioned methods required a good starting geometry because the associated geometry optimizations were not very sophisticated. With those limitations in mind, MINDO/3, MNDO, AM1 and PM3 were developed to include parameter sets that would be rigorous and applicable across a wide spectrum of chemical problems, which Dewar chose to be organic molecules.

The semiempirical QM method of modified intermediate neglect of differential overlap (MINDO/3)[11] introduced in 1975 was not so much a significant change in

previous theory, as it is based from Pople and coworkers' INDO method, but rather it differed dramatically in the way it was parameterized making much more use of empirical data. The total energy of a system described by MINDO/3 is defined by five basis terms: the Coulombic inter-electronic repulsions, the electron-core attractions, the core-core repulsion, the one-center exchange terms, and the two-center exchange terms.[11a] In the MINDO/3 approach, some or all of these terms were set equal to parametric functions, which can be adjusted to fit experimental data. In addition, variables that were fixed in CNDO, INDO and NNDO calculations were permitted to vary during the parameterization. Dewar and coworkers made the argument that the MINDO/3 method was a better quantitative guide to chemical behavior; even more so than the corresponding Roothann-Hall calculations (remember this is the early 1970's).[11a] This argument is made predominantly because

> "… the cost of MINDO/3 calculations is in fact less by a factor of at least 100,000, there is really no contest in areas where MINDO/3 can be applied."[11a]

The MINDO/3 method's validity was further investigated by calculating the heats of formation for 193 organic compounds, which on average were within ± 5 kcal/mol of experimental thermochemical data.[11a] Furthermore, over fifty different organic reactions were studied and the calculated energies of activation where within ± 5 kcal/mol of experimental data.[11a] Another notable difference from the previous semiempirical molecular orbital theories is that the MINDO/3 method incorporated geometry optimizations. Since MINDO/3 included a geometry optimization procedure, the method allowed for the initial starting configuration to be unrefined, because MINDO/3 could

derive the associated minimum energy structure.[11a] The '3' in MINDO/3 implies that only after the fourth re-parameterization of Pople and coworkers' INDO method was an acceptable new model obtained. However MINDO/3 had some considerable issues. For example, compounds that contained lone pairs of electrons on adjacent atoms had heats of formation that were too negative and the corresponding bond lengths were too short. Furthermore, the calculated heats of formation for aromatic hydrocarbons were too positive. Even with the errors associated with the MINDO/3 method, Dewar and coworkers explain how they believed

> "…it (is) likely that MINOD/3 represents the limit attainable by an INDO-based semiempirical treatment and we doubt if further efforts in this area will prove rewarding…".[11a]

Therefore, in 1977 Dewar and Thiel proposed the QM semiempirical method modified neglect of differential overlap (MNDO),[12] which was based on NDDO and not INDO. One of the reasons the re-parameterization of the NDDO method was pursued is because the directionality of chemical bonding is dramatically different from that in INDO.[12a] Since the two-center electron-electron repulsions and core electron attractions are included in NDDO, it is reasonable that chemical bonding is modeled more accurately than within the INDO based methods.[12a] For a direct comparison between the MINDO/3 and MNDO methodologies, Dewar and Thiel chose a test set of 138 closed shell "difficult" organic molecules and found that the heats of formation, molecular geometries, ionization potential and dipole moments calculated by MNDO had mean absolute errors of 6.3 kcal/mol compared to 11.0 kcal/mol with MINDO/3.[12b] In addition MNDO was further parameterized to model the chemistry of ground state molecules

containing fluorine,[15] beryllium,[16] boron[17] and third period elements.[18] However, despite the extensive parameterization of the MNDO method, MNDO had serious limitations. MNDO overestimates the repulsion between two atoms that are separated by approximately the sum of their individual van der Waals radii. This results in the heats of formation for the water dimer being drastically low because MNDO could not properly model intermolecular systems that involved hydrogen bonds.

In 1985 Dewar and coworkers developed a semiempirical QM method called the Austin Model 1 (AM1)[13] in an attempt to correct the errors in MNDO. In order to accomplish this goal, Dewar and coworkers decided to modify MNDO's existing core repulsion function[12a] (CRF) by including repulsive and attractive Gaussian functions (see equation 2.16).

$$E_{AB} = E_{MNDO} + \frac{Z_A Z_B}{R_{AB}} \left\{ \sum_i K_{A_i} e^{\left[-L_{A_i}\left(R_{AB} - M_{A_i}\right)^2\right]} + \sum_j K_{B_j} e^{\left[-L_{B_j}\left(R_{AB} - M_{B_j}\right)^2\right]} \right\} \tag{2.16}$$

Equation 2.16 shows the energy of interaction between individual atoms and not pairs of atoms. Z_A and Z_B represent the charge of atom A and B, and R_{AB} is simply the distance between the two atoms, *i.e.* $\frac{Z_A Z_B}{R_{AB}}$ is the Coulombic interaction. The parameter *L* corresponds to the width of the Gaussians and was not optimized for each atom, whereas parameters *K* and *M* were optimized for each atom. Attractive Gaussians were used to compensate for the excessive repulsions and were centered in the regions where these repulsions were greatest. Then repulsive Gaussians where added at small internuclear

distances to reduce the repulsion at larger atomic distances.[13] For the atoms carbon, hydrogen, and nitrogen both attractive and repulsive Gaussian functions needed to be included, while only repulsive Gaussians were needed for oxygen. Furthermore, it was determined that for carbon four terms were required in the Gaussian expansion, whereas only three terms were required for hydrogen and nitrogen and only two terms were needed for oxygen. The inclusion of these Gaussian functions in the CRF significantly increased the number of parameters per atom; in MNDO only seven are required, whereas in AM1 between 13 and 16 parameters where needed per atom. In general Dewar and coworkers were able to demonstrate that AM1 was a very dramatic improvement over MNDO in the modeling of heats of formation for neutral closed-shell molecules, cations, radicals and anions. The energetics and geometries of hydrogen bonds improved when compared to MNDO, however they are still underestimated and "weak". In addition, ionization energies, rational barriers, molecular geometries and activation energies of simple reactions also dramatically improved when compared to MNDO.[13]

Even though AM1 is an improvement over MNDO paired with no increase in the computational demand, one of the authors on the original AM1 paper, James J. P. Stewart developed the third parameterization of MNDO, called PM3.[14] Stewart limited himself to only using two Gaussian functions per atom instead of four Gaussian functions as used in AM1. As a direct result, Stewart was able to use a much larger set of test molecules during the parameterization of H, C, N, O, F, Al, Si, P, S, Cl, Br, and I, because the rate determining step in the parameterization was now the assembling of experimental reference data and not the determination of the parameters.[14] Some parameters have significantly different values in AM1 than in PM3, even though both methods use the

same functional form of MNDO. Furthermore, both methods predict thermodynamic and structural properties to roughly the same level of accuracy.

2.3.3 Pairwise Distance Directed Gaussian Third Parameterization of Modified Neglect of Differential Overlap (PDDG/PM3)

A further improvement in the QM semiempirical methodologies was the addition of the PDDG modifications to the PM3 method, called PDDG/PM3.[3] This semiempirical method employs a pairwise distance directed Gaussian (PDDG) modification to the core repulsion function (CRF) and uses the bond and groups equivalents scheme (BGE) to provide more accurate heats of formation.[3a, 19]

$$\Delta H_f = E_{mol} + \frac{T}{R} + \sum_j a_j n_j \qquad (2.17)$$

In equation 2.17 the total molecular energy, E_{mol}, is converted to the heat of formation using translational, *T*, and rotational, *R*, corrections in addition to the sum of the products of equivalents, a_j, and the number of times that the bonds/groups are used, n_j. The equivalents are determined empirically and there are two types of equivalents. Bond equivalents are the first type and represent the bonding energy of the molecule and include the majority of the energy required to convert molecular energy into enthalpy. For example, for a saturated hydrocarbon there are two different bond equivalents; there is one bond equivalent for each C_{sp^3}-C_{sp^3} bond and a different bond equivalent for each C_{sp^3}-H bond. The second type of equivalent introduced in BGE, is the group equivalent. This accounts for the variations in the chemical environment for a particular functional group; e.g. for saturated hydrocarbons, Me, Iso and Neo represent carbon centers that have 1, 3 and 4 carbon substituents respectively.[19] For a test set of 583 neutral, closed

shell organic molecules, including the BGE scheme reduced the mean absolute errors for the heats of formation with the semiempirical methods MNDO, AM1 and PM3 from 8.2, 6.6, and 4.4 kcal/mol to 3.0, 2.3, and 2.2 kcal/mol respectively.[3a] However the use of this methodology to study transition states and other higher energy structures is not acceptable, for BGE can only be used for molecules at stationary points on the potential energy surface that have conventional bonding. Furthermore, AM1 and PM3 generally give results with poor energetics for diatomics and other small organic molecules; so a significant reevaluation of the eisol parameter or electronic energy of formation, was also required. This parameter is calculated from the restricted single-determinate wave function and in a nonsystematic way, is dependent on the values of all one-center parameters of the semiempirical method, e.g. PM3 or AM1.[3a] Originally, the eisol parameter was parameterized to reproduce experimental properties for molecules that consisted of heavy atoms, and therefore the eisol parameter was re-parameterized for common organic elements.[42]

Since the BGE scheme cannot be applied generally to describe fragments or functional groups, a further modification of the CRF was performed. The modification was an effort to include the fragment and functional group information while not adversely affecting the related molecular energetics.[3a] This modification to the CRF adds Gaussian functions to reduce the core-core repulsions that occur right outside of the bonding distances. This was the first attempt to include atom based functional group information. However, it was difficult to differentiate between the functional groups because the Gaussians only had one distance parameter and therefore were held too closely together. This lead Bernal-Uruchurtu *et* al. to propose an additional correction to

the CRF to improve the energetics and structure of H-bonding, as seen below in equation 2.18.[20]

$$CRF = CRF_{MNDO} + \sum_{A,B}^{inter} \alpha_{AB}\, e^{-\beta_{AB} R_{AB}} + \frac{\chi_{AB}}{R_{AB}^{6}} + \frac{\delta_{AB}}{R_{AB}^{8}} + \frac{\epsilon_{AB}}{R_{AB}^{10}} \tag{2.18}$$

The bond based empirical parameters, α_{AB}, β_{AB}, χ_{AB}, δ_{AB}, and ϵ_{AB}were added to MNDO's CRF to improve the energetics and structures of water dimers and clusters.[3a, 20]

Through all of the above mentioned modifications: adding functional group information into the enthalpy, reevaluating the eisol parameter, and improving the energetics and structure of H-bonding; the PDDG/PM3 QM semiempirical method was created.

$$PDDG(A,B) = \sum_{A}\sum_{B>A} \frac{1}{n_A + n_B} \times \left[\sum_{i=1}^{2}\sum_{j=1}^{2} \left(n_A P_{A_i} + n_B P_{B_j} \right) e^{\left(-10\text{Å}^{-2} \left(R_{AB} - D_{A_i} - D_{B_j} \right)^2 \right)} \right] \tag{2.19}$$

In equation 2.19, the sum is over all distinct atom pairs, A and B, where each element requires a total of four optimizable parameters, two pre-exponential factors, P_{Ai} and P_{Bj}, and two distance terms, D_{Ai} and D_{Aj}. The numbers of valence electrons on atoms A and B, n_A and n_B respectively, are used to avoid problems between atoms pairs with large and small atomic numbers. The PDDG function only uses atom based parameters to generate bond specific Gaussians, by using the difference in the atomic distances to locate the Gaussians and the sum of atomic pre-exponential terms to give the Gaussian its

amplitude. The PDDG modification exhibits the flexibility and shape required to differentiate between various functional groups while only using a small number of atom based parameters.[3a]

A number of considerations went into the parameterization of the PDDG modification in order to have the functional only contribute minimally to the molecular energy and therefore not adversely alter the optimized molecular geometry. Bond and functional group specific parameters were not used in the PDDG functional because the terms in the PDDG expression can differentiate between functional groups based on the molecular geometries. Lastly high energy structures, such as charged species and transition states with nontraditional bonding, can be accurately modeled.[3a] Confidence in the semiempirical method is strengthened by demonstrations of the method's ability to accurately reproduce experimental results. So for instance the calculated mean unsigned errors of the heats of formation for a test set of 622 organic molecules for the semiempirical methods MNDO, AM1, PM3 and PDDG/PM3 was 8.4, 5.2, 4.4 and 3.2 kcal/mol respectively.[3a] Additionally, when the test set was expanded to include 422 halogen containing compounds, the mean unsigned errors of the heats of formation were 14.0, 11.1, 8.1, 6.6 and 5.6 kcal/mol for MNDO, AM1, PM3, PDDG/MNDO and PDDG/PM3 respectively.[3b] The PDDG modification of the semiempirical method PM3 has also shown accurate results for hydrogen bonded complexes and transition states, so much as to be comparable to results obtained from B3LYP/6-311++G(d,p) calculations.[3c] In particular, organic reactions are of specific interest and the use of the PDDG/PM3 methodology to describe the QM region of a QM/MM calculation has shown to be successful in numerous instances; nucleophilic aromatic substitution reactions,[1ac] Kemp

decarboxylations,[1ad] S_N2 reactions,[3c] Diels-Alder reactions,[1ah] the Claisen Rearrangement,[1am] the Menshutkin reaction[1al] and the base catalyzed β-elimination reactions.[21]

2.4 Additional Mixed QM/MM Methodology

2.4.1 Link Atom, Connection Atom and ONIOM

The true distinction between different hybrid QM/MM methods lies in the treatment of the final term, the QM/MM energy of interaction. The interface between the QM and MM region when dealing with for example, large biological macromolecules, involves cutting across chemical bonds and as a result requires special treatment in order to stitch together the distinct partitioned spaces. The link atom[22] methodology was introduced by Singh and Kollman in 1986 in order to properly treat the covalent bonds that exist at the border between the QM and MM regions, see Figure 2.7.

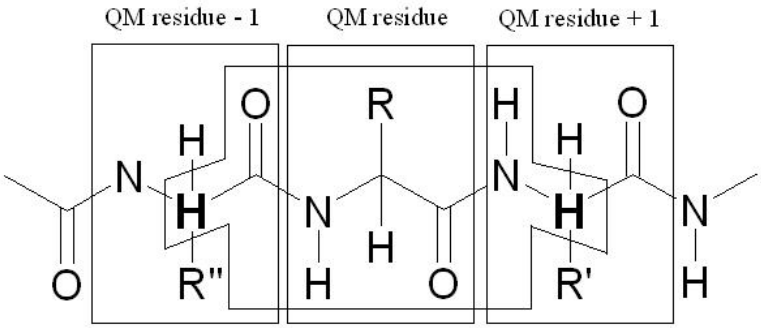


Figure 2.7: Visual representation of the link atom methodology. Hydrogen atoms are added to "cap-off" the chemical bond that lies within the QM region and the MM region (QM residue ± 1) boundary.

Essentially the link atom method caps the QM region with hydrogen atoms in order to fill the free valencies of the QM atoms that are bonded to atoms described by the MM subsystem. These link atoms are treated explicitly during the QM calculation and do not interact with the MM atoms. This approach is non-ideal because including the dummy hydrogens in the QM calculations result in the energy and the gradient being poorly defined. This is simply a result of the inconsistent contributions made by these hydrogen caps. Furthermore there is a debate as to whether or not these link atoms should be involved in Coulombic interactions[23]. Even though the link atom method alters the form of the "real-system" for computational efficiency, as a whole, the "link-atom" method is reasonable when careful consideration of the location of these dummy atoms is taken. In general dummy atoms placed furthest from the QM atoms provide more reliable results.

An alternative to the use of QM hydrogen link atoms is called connection atoms which was developed by Thiel and co-workers.[24] The primary difference from the link atom method is at the boundary between the QM and MM regions where the MM atoms at the border are described as a QM methyl group with a free sp^3 valence (Figure 2.8).

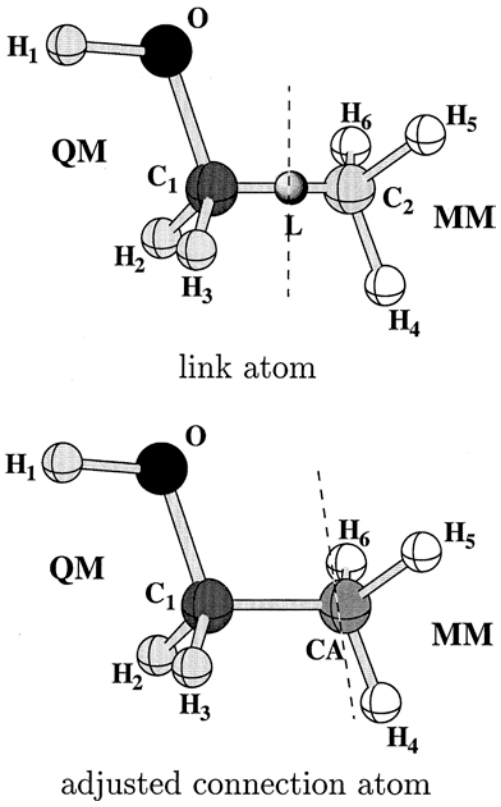


Figure 2.8: The top illustrates the link atom methodology and the bottom figure illustrates the connection atom method.

Another popular method for the special treatment of chemical bonds that cross the QM/MM boundary, for example large bioorganic macromolecules, is the ONIOM[25] method (our own *n*-layered integrated molecular orbital and molecular mechanics) created by Morokuma and coworkers, (Figure 2.9).

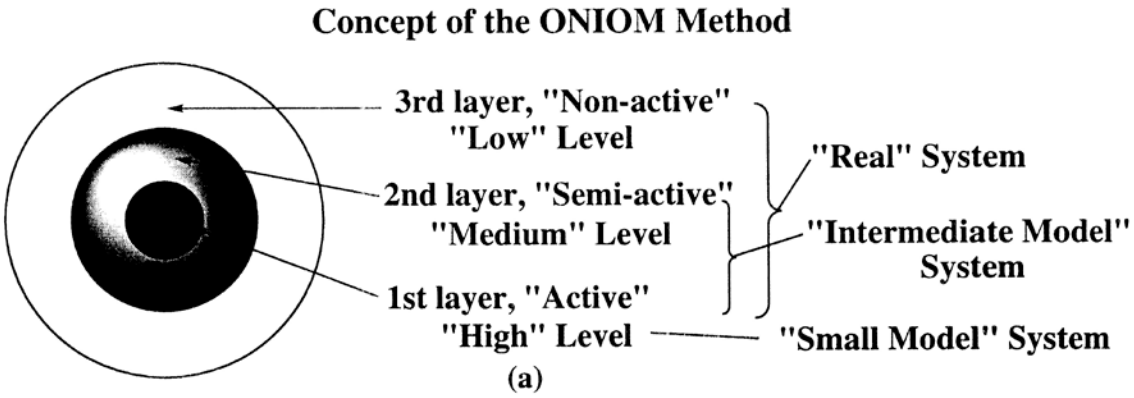


Figure 2.9: A diagram of the ONIOM3 method.

Different levels of computational theory can be used to calculate the overall energetics of the system, due to the "onion-like" multilayered arrangement. The lowest level of theory is applied to the outer most layer; e.g. MM, because it can accurately describe the steric effects of the substituents farthest from the active site. The middle layer is "semi-active" being relatively close to the reacting center, and therefore an intermediate level of theory is needed to accurately describe the electronic effects of the ligands and/or functional groups found here. The active site or core of the onion requires the highest level of theory to adequately describe the energetics and electronic motion of the reaction coordinate of interest. The ultimate goal of Morokuma and coworkers was to develop a computationally effective method in where the modeled system of interest could more closely resemble the "real-world" system without making drastic approximations.

2.4.2 Charge Derivation in QM/MM Region

When the chemical system of interest is a small organic solute molecule in a condensed phase medium, similar to systems in this dissertation, the boundary between the QM and MM region is easily separable and can be strictly described by the

nonbonded interactions occurring between the two regions, i.e., Coulombic and Lennard-Jones contributions. Therefore the energy of interaction between the QM organic solutes and the classically described solvent molecules can be computed as follows:

$$E_{QM/MM} = \sum_{a} E_{as} \tag{2.20}$$

$$E_{as} = \sum_{i}^{on\ s} \sum_{j}^{on\ a} \left[\frac{q_i^* q_J e^2}{r_{ij}} + 4 \in_{ij} \left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}} \right) \right] \tag{2.21}$$

$$q_i^* = \alpha q_i^{CM1A} \tag{2.22}$$

The QM/MM energy of interaction is the double sum over atoms *i* of the solute molecule *s* and atoms *j* of the solvent molecule *a*. The partial charges of the QM solute molecules are represented by q_i^*, as seen in equation 2.21. The AM1 or PDDG/PM3 wave function is used to derive the partial charges for the QM atoms *i* using the charge model 1 (CM1 or CM3) procedure[26] developed by Cramer, Truhlar and coworkers. These partial atomic charges are calculated as a sum of interactions between different parts of the solute molecule. The charges are fixed and there is no polarization, resulting in the charges being restricted to the nuclear centers. The computed atomic partial charges are then scaled by a factor of alpha, α, see equation 2.22. The optimal scaling factor, α, was determined by minimizing the unsigned average error between experimental and calculated free energies of hydration for 25 diverse organic molecules[26c]. For example, it was determined by Marina Udier-Blagovic *et al.* that scaling the CM1 partial atomic charges by 1.14 yielded a mean unsigned error of 1.0 kcal/mol for the free energies of hydration and a mean unsigned error of 0.7 kcal/mol with the OPLS-AA charges. The scaled QM partial atomic charges has been shown to give consistent results for free energies of hydration and medium effects on equilibria and reaction rates.[1j, 1l, 1ab-aj, 1al, am

The partial charges of the MM explicit solvent molecules are represented by q_j and are computed from the OPLS-AA force field.

2.5 Introduction to Statistical Mechanics

Statistical mechanics is a theory which analyzes the fluctuations and behaviors of natural, spontaneous processes. Statistical mechanics is essentially a "bridge" from the microscopic point of view of a system to the corresponding macroscopic phenomena; *e.g.*, experimental measurements. The microscopic phenomena explored in computational chemistry can be described by quantum mechanical and/or molecular mechanical simulations; incorporating statistical mechanics into the calculations provides a link to macroscopic properties such as densities and thermodynamics. From the point of view of the microscopic world, this can be a difficult and tedious procedure. For instance, take a MM-described box of 300 water molecules. To characterize this box of water, the energy is proportional to approximately 10^N, meaning that 10^{300} velocities, positions, trajectories, *etc.* would be required.

A theoretically efficient method to make the jump from observable properties to the corresponding microscopic properties is through the use of ensembles. In the macroscopic world, a system at a constant volume (V), contains a constant number of molecules (N), and is held at a constant temperature (T) is ideal for measuring thermodynamic properties. For example, the average energy of a hydrogen bond between water molecules is approximately 3.6 kcal/mol. These hydrogen bonds are repeatedly and simultaneously made and broken. Therefore the corresponding energies are constantly changing because everything is in motion. Statistically, not every energy of interaction is going to equal the experimental value of 3.6 kcal/mol, some energies of interaction will

be larger, some will be smaller. But since there are more than a million possible configurations, statistically even if a single energy of interaction has an astronomically large value, it will not contribute significantly to the overall average energy of interaction. The two most common ensembles used in order to relate microscopic phenomena to the corresponding macroscopic thermodynamic properties are the canonical (constant N, V, and T) and isothermal-isobaric (constant N, P and T) ensembles.

$$\langle Q \rangle = Q_k + \int Q(\vec{X}) P(\vec{X}) d\vec{X} \tag{2.23}$$

$$P(\vec{X}) = \frac{e^{[-\beta E(\vec{X})]}}{\int e^{[-\beta E(\vec{X})] d\vec{X}}} \tag{2.24}$$

$$\beta = \frac{1}{kT} \tag{2.25}$$

The partition function, *Q*, is the central statistical mechanical function of an ensemble and the equations 2.23 – 2.25 are specifically for the canonical NVT ensemble, where $P(\vec{X})$ is the Boltzmann factor or the configurational probability, $E(\vec{X})$ is the total potential energy of the system, and $\langle Q \rangle$ is the average value for a specific property. The integrals are solved over all possible geometric configurations, $\vec{X}$, of the chemical system of interest. The kinetic energy contribution, Q_k, is separable from the contributions of the configurations as seen in equation 2.23.[27] Likewise the corresponding classical statistical mechanics equations for the isothermal-isobaric NPT ensemble are as follows.

$$\langle Q \rangle = Q_k + \int \int Q(\vec{X}, V) P(\overrightarrow{X}, V) \, d\overrightarrow{X} \, dV \tag{2.26}$$

$$P(\vec{X},V) = \frac{e^{[-\beta H(\vec{X},V)]}}{\int\int e^{[-\beta H(\vec{X},V)]}\, d\vec{X}\, dV} \tag{2.27}$$

$$H(\vec{X},V) = E(\vec{X}) + PV(\vec{X}) \tag{2.28}$$

The integrations now further include the complete volume of the system, *V*, and in equation 2.28 the enthalpy of the system is calculated including the external pressure, *P*, of the system.[27b] For a system of *N* particles, all possible geometric configurations, $\vec{X}$, would have approximately 3N coordinates. Therefore the integrals seen in the classical statistical mechanics equations above are highly dimensional and as a direct result, quickly become impractical for a system containing more than a few particles.

2.5.1 Metropolis Monte Carlo (MC)

Metropolis et *al*.[28] solved the fundamental problem of simplifying the key integrals above in a practical and accurate manner in which

> "instead of choosing configurations randomly, then weighing them with exp(-E/kT), we choose configurations with a probability exp(-E/kT) and weigh them evenly."[28]

With the development of the Metropolis et *al.* sampling algorithm, configurations can now be selected so they occur with a certain probability proportional to their Boltzmann factor. Using the algorithm results in the classical statistical mechanics equations previously described, dramatically simplify from integrals over all configuration space to a sum over discrete configurations.

$$\langle Q \rangle = Q_k + \frac{1}{L} \sum_{i}^{L} Q_i(\vec{X}_M) \tag{2.29}$$

In equation 2.29, the average value of a specific property, $\langle Q \rangle$, is now described by a sum over all snapshots of all possible configurations, in which L is the number of configurations considered and $\vec{X}_M$ indicates a Metropolis-sampled configuration.[27b] The partition function Q_i for the NVT ensemble would be represented as follows:

$$Q(N,V,T) = \sum_{j} e^{\left(\frac{-E_j(N,V)}{kT}\right)} \tag{2.30}$$

In equation 2.30, j represents all of the microstates that a system can occupy, and $E_j(N,V)$ is the corresponding energy for that particular microstate. Manipulation of the partition function allows for a direct way to link microscopic phenomena to macroscopic phenomena by extrapolating the thermodynamic properties of interest. Furthermore, the probability that a system occupies a particular microstate

$$P_i = \frac{e^{(-\beta E_i)}}{\sum_i e^{(-\beta E_i)}} \tag{2.31}$$

is shown in equation 2.31, and the probability is now inversely proportional to the sum of potential microstates. Using the MC algorithm encodes how the probabilities are partitioned among different microstates.

When sampling a chemical system of interest, there is a need to sample extensively in order to reach proper convergence. If the energy of the system is very high for a particular configuration then the probability, P_i , will be very low; whereas the ideal situation is to obtain very high probability states at a very low energetic cost. Repeated

translating, rotating, and/or any internal structural variation on the molecules within the system will eventually allow the system to reach convergence. Metropolis et *al.* simplified the classical statistical mechanics equations into an algorithm, see equation 2.30, which involves the generation of a Markov chain of configurations. The Metropolis sampling algorithm follows two essential conditions; (1) the outcome of each trial depends only upon the preceding trial. (2) Each MC trial belongs to a finite set of possible outcomes.

The Metropolis algorithm involves computing a probability that a random configurational change should be applied to the initial starting configuration, *i*, to create the next new configuration, *j* and so forth. For rigid molecules, a random MC configurational change normally consists of picking one molecule at random, translating it randomly in all three Cartesian directions, see equations 2.32-2.34, and randomly rotating it about its Cartesian axis.[27b] On the other hand, for flexible molecules both the external and internal degrees of freedom are sampled.[27b]

$$x_{new} = x_{old} + (2\zeta_x - 1)\delta r_{max} \quad (2.32)$$

$$y_{new} = y_{old} + (2\zeta_y - 1)\delta r_{max} \quad (2.33)$$

$$z_{new} = z_{old} + (2\zeta_z - 1)\delta r_{max} \quad (2.34)$$

The random number generator produces a number, ζ, in the range 0 to 1 and moves in both the positive and negative directions are possible. A unique random number is generated for each of the three directions x, y and z, where r_{max} is an adjustable parameter that determines the maximum displacement possible in any direction. This parameter governs the size of the perturbation, whose value is chosen so that approximately between 30-50% of sampled configurations are accepted. The external degrees of

freedom are treated the same as in rigid molecules, however the internal degrees of freedom of the molecule are typically represented by a Z-matrix, were the variable entries in the Z-matrix are changed randomly within specified ranges and the molecules are "rebuilt" for the new configuration, j.[27b] The probability that the random configurational change is accepted can be expressed as $p = \pi_i/\pi_j$, where the expressions for π_j for the NVT (equations 2.30 and 2.31) and NPT ensembles are shown below in equation 2.35.[27b]

$$\pi_j = e^{(-\beta H_j)} V_j^N = e^{\left(-\beta(E_j + PV_j - NkT\, lnV_j)\right)} \tag{2.35}$$

After the new configuration is generated, the energy of the new configuration is calculated using the partition function Q_i, and it is compared to the initial configuration's energy. If the probability, i.e. the Boltzmann factor, is greater than or equal to one, $p \geq 1$, the random configurational change is accepted and becomes the new starting configuration, *i*. For the NVT ensemble, when $p \geq 1$, the potential energy of the final configuration is less than the potential energy of the starting configuration, $E_j \leq E_i$, and therefore the new configuration is accepted and becomes the new starting configuration.[27b] If however, the probability is less than one, $p < 1$, the configuration *j* can still be accepted. This configuration's probability is then compared to a random number *x* between 0 and 1, and if $p \geq x$ then the random configurational change is accepted and the configuration *j* becomes the new starting configuration.[27b]

$$random\ number\ x\ (0,1) \leq e^{\left(\frac{-\Delta E_j(N,V)}{kT}\right)} \tag{2.36}$$

However, if $p < x$ then the configurational change is rejected and the initial configuration, *i*, repeats. It is important to realize that the random number generated in equation 2.36, is not truly "random". The same sequences of numbers are generated by

the computer if the initial conditions are the same. This sequence of numbers is altered when different "seeds" are provided at the beginning of every trial, e.g., the time and date of the calculation. Figure 2.10 depicts the general schematic for the Metropolis et *al.* algorithm.

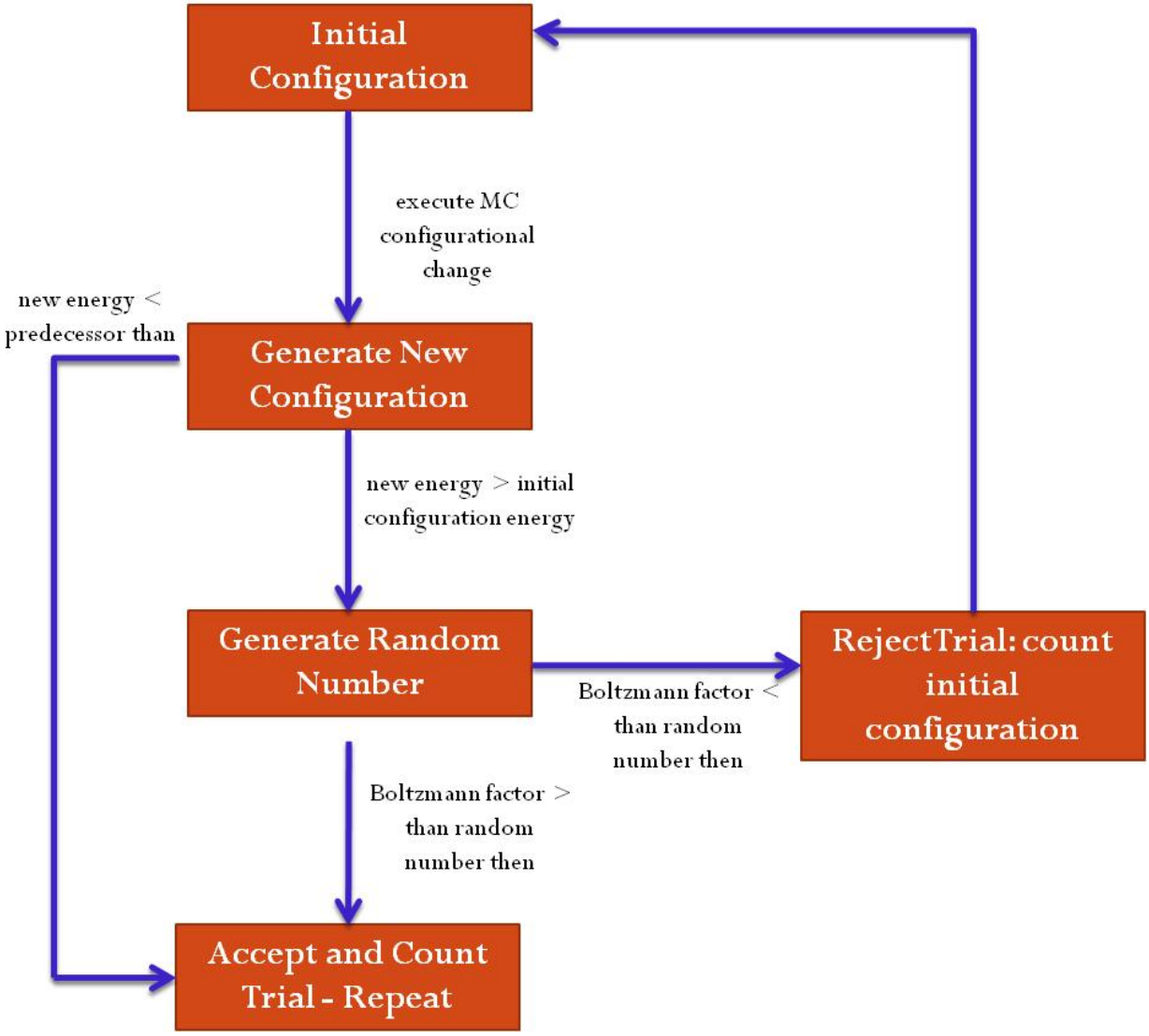


Figure 2.10: The Metropolis MC sampling algorithm flow chart.

The size of the random perturbation to the internal and/or external degrees of freedom is crucial in the determination of the acceptance rate. If the size of the perturbation is too small, then the initial state and the final state will be extremely similar, resulting with the phase space explored being too little. On the other hand, if the size of the move is too large, then many of the Metropolis configurations will be rejected, resulting in a small acceptance rate and very little progression across the potential energy surface (PES). The fact that higher energy configurations can be accepted in the chain of microstates is an important feature of the Metropolis sampling algorithm. Allowing for higher energy states to be accepted results in a greater conformational space to be explored, allowing for a multitude of pathways to be generated along the PES.

2.5.2 Potential Energy Surfaces (PES)

A PES, like the example in Figure 2.11, can be conceptualized as a landscape with a mountainous terrain, containing valleys, ridges and peaks.[29] PESs describe how the energy of a molecule(s) in a particular microstate varies as a function of the molecule's structure. When exploring the PES, there are local minimas and maximas, saddle points and global minima and maxima. The valleys in the landscape represent the reactants, intermediates and products. The ridges represent transition states that typically are either first-order transition states or second-order transition states. A transition state, or saddle point, is a maximum in one direction and a minimum in all others. Furthermore, a first-order saddle point is a peak connecting two valleys and a second-order transition state is a maximum in two directions but a minimum in all other directions.[29] The favored reaction pathway, which generally determines the reaction's energy of activation, is the lowest energy connection between the reactants and products. Even though most

molecules have a multitude of geometric variables, the most important features of a potential energy surface, i.e., the bond making and breaking during a reaction, can be displayed in such a way to generate a surface which accurately depicts the energetic landscape of the chemical reaction of interest.

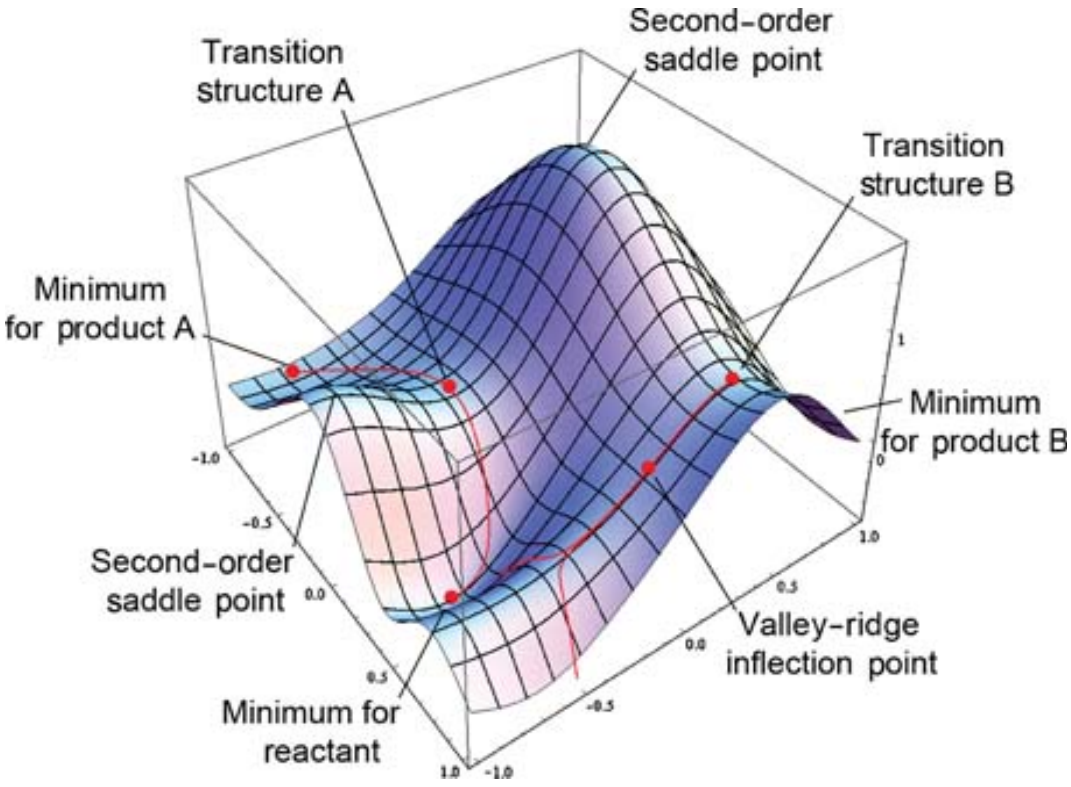


Figure 2.11[29]: An example of a potential energy surface showing products, the reactants, transition structures or saddle points and the corresponding reaction pathways.

2.5.3 Periodic Boundary Conditions (PBC)

In addition to the sampling algorithm, Metropolis *et al*. also developed an efficient method to properly model bulk solvent while utilizing only a small number of particles, *i.e*., periodic boundary conditions (PBC), see Figure 2.12.

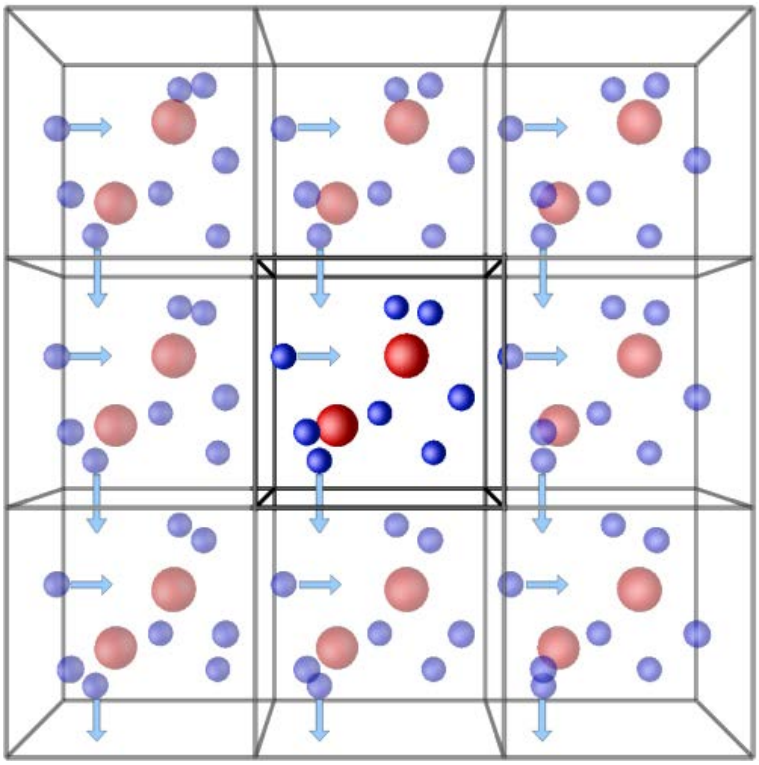


Figure 2.12: Illustration of periodic boundary conditions.

Typically reacting systems of interest that utilize MC and PBC contain upwards of 400-1000 solvent molecules.[27b] Periodic boundary conditions use the original box containing these solvent molecules and replicate it in all directions to give a periodic array of solvent molecules. The locality of the particles in the image boxes can be calculated by adding or subtracting multiples of the length of the central box side. Should a solvent molecule, (represented as a small blue sphere in Figure 2.12) exit a box during the duration of the simulation, it is replaced by an image particle simultaneously from the opposite side of

the box (illustrated as blue arrows). This allows for a constant number of particles to remain in the chemical system. The solute molecule(s), represented as large red sphere(s) in Figure 2.12, are then placed in the central box and therefore replicated in all image boxes. In order to prevent the solute molecule from approaching the "edge of the box", a solute-solvent cutoff interaction distance is implemented. The solute-solvent cutoff distance is utilized because it was determined in early MC simulations that when radial distribution functions (RDFs) or energy pair distributions were calculated, the energy would quickly fall to zero due to the scarcity of the interactions within the solvent box that are beyond the distance equal to half the length of the box, *a*/2, to prevent double counting of molecules.[30] This effect is simply due to the fact that atom pairs would be separated by very large distances and the resulting energy of interaction would contribute very little, if any at all.

2.5.4 Radial Distribution Functions (RDFs)

Radial distribution functions (RDFs) are one of the most important techniques to measure the local structure within a liquid.[27b, 31] RDFs analyze the probability of finding an atom of type *j* at a distance *r* from an atom of type *i* in the solvent of interest, see equation 2.37 and 2.38 below.[27b, 31]

$$\rho_{ij}(r) = \rho_j^0 \, g_{ij}(r) \tag{2.37}$$

$$\rho_j^0 = \frac{N_j}{V} \tag{2.38}$$

Where $\rho_{ij}(r)$ is the density of atom *i* relative to atom *j* at a distance *r*, ρ_j^0 is the original density of atom *j* and $g_{ij}(r)$ is the atom-atom radial distribution function. It is chemically

intuitive that the density, ρ, of atom type j around atom type i should never be uniform. However, if the solvent of interest happens to be structureless then,[27b, 31]

$$\rho_{ij}(r) = \rho_j^0 \; \forall \, r \tag{2.39}$$

$$g_{ij}(r) = 1 \tag{2.40}$$

Furthermore, the average number of atoms j in a solvation shell between distance r and $r + dr$ centralized around atom type i is calculated from equation 2.41,[27b, 31] which can be manipulated to give the appropriate RDFs of particular interest, $g_{ij}(r)$.

$$\langle N_j(r, r + dr) \rangle = 4\pi r^2 \, dr \, \rho_j^0 \, g_{ij}(r) \tag{2.41}$$

The peaks generated from the RDFs calculations are associated with solvation shells of thickness dr, and can be integrated to provide the coordination numbers, c_j, see equation 2.42 below.[27b]

$$c_j = \rho_j^0 \int_0^q 4\pi r^2 \, g_{ij}(r) dr \tag{2.42}$$

The upper integration limit, q, is usually the minimum value allowed in the RDFs, $g_{ij}(r)$, of interest. An example of an RDFs for the reaction to be discussed in Chapter 3 between piperidine and para-methoxynitrothiophene in the solvent methanol is seen below in Figure 2.13.

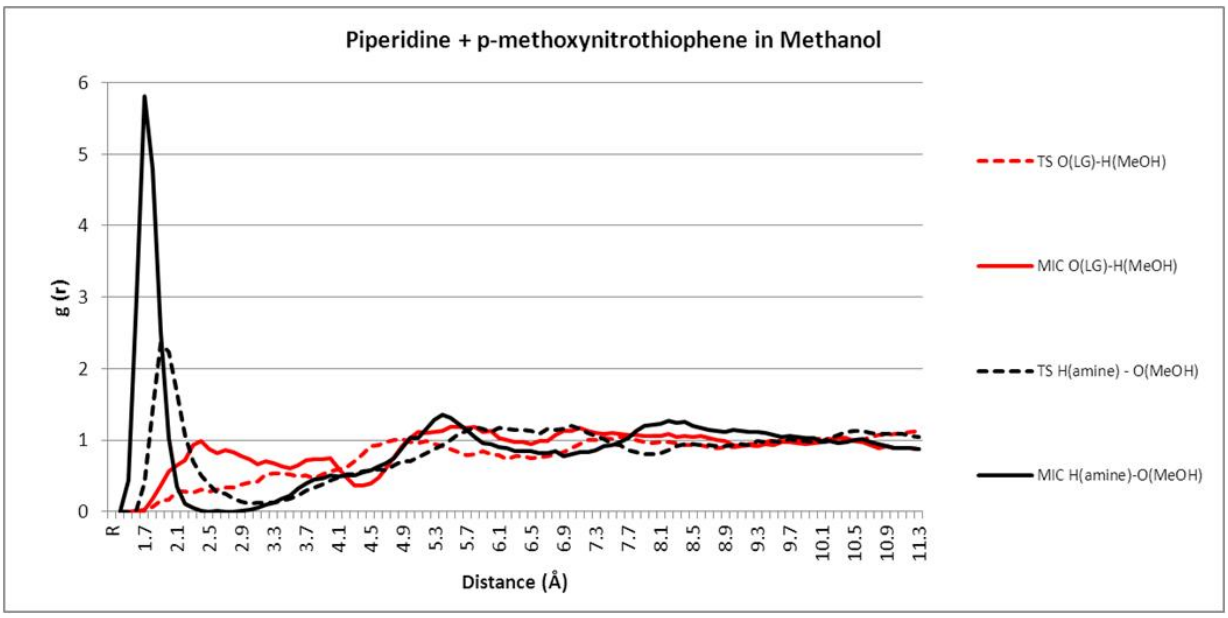


Figure 2.13: Select radial distribution function determined from QM/MM/FEP/MC reaction between piperidine and para-methoxynitrothiophene in methanol for the Meisenheimer Intermediate Complex (MIC) and the transition state (TS). (LG = OCH_3).

For short distances which are less than the atomic radius, the $g_{ij}(r) = 0$, being due to the repulsive nature of the corresponding interactions. Figure 2.13 shows the radial distribution functions for select interactions that occur along the reaction pathway during the substitution nucleophilic aromatic reaction, to be discussed in Chapter 4. The RDFs show an increase in solvation of the solute, piperidine reacting with para-methoxynitrothiophene, when passing from the transition state to the Meisenheimer intermediate complex in methanol; dashed and solid lines respectively. Utilizing MC easily allows for the tracking of interatomic separations and thus obtaining the resulting RDFs of interest for a particular chemical system. RDFs can be measured experimentally

using X-ray diffraction or neutron diffraction,[32] allowing for comparisons with computed RDFs.

It has been shown for solvent boxes containing approximately 200 water or organic solvent molecules, that the energy, enthalpy, volume and radial distribution functions normally converge in 0.5-1.0 million configurations after equilibration has been reached.[33] Nonetheless, other properties that are derived from fluctuations in enthalpy, energy and volume, such as heat capacity and isothermal expansion converge more slowly.[33] Furthermore, MC simulations cannot be used to accurately describe and determine the free energy of a chemical process. These simulations do not adequately sample from the specific regions of phase space that are primarily responsible and contribute the most to the corresponding free energy of a reaction. In order to accurately determine the free energy of a chemical process, which is often considered the most important thermodynamic property, Free Energy Perturbation theory (FEP) must be utilized in conjunction with MC and QM/MM calculations.

2.6 Introduction to Free Energy Calculations

MC simulations sample the important low energy regions of phase space and therefore will never adequately sample or describe the higher energy regions of the surface required to calculate the corresponding free energy of the chemical process. Free energy is a fundamental property for the thermodynamic characterization of chemical reaction pathways. Free energy is usually expressed as the Helmholtz function, *A*, which describes the useful work obtainable from a closed thermodynamic system that is at constant temperature and volume (NVT ensemble) or the Gibbs free energy function, *G*, which measures the amount of useful work obtained from a system that is at constant

temperature and pressure (NPT ensemble). Unfortunately, free energy is a difficult thermodynamic property to calculate for systems such as liquids or large flexible macromolecules. The difficulty is due to these systems having numerous minimum energy configurations, which are separated by low energy barriers. There are three conventional techniques to compute the free energy of a chemical system: (1) thermodynamic perturbation,[1l, 1n, 1ab, 2a, 34] also commonly referred to as Free Energy Perturbation Theory (FEP), (2) thermodynamic integration[34] (TI) and (3) slow growth. Application of any one of the three methods for computing the free energy of a chemical system should give the same energetic result because free energy is a state function and is independent of the path or methodology taken to calculate it.

2.6.1 Free Energy Perturbation (FEP) Theory

Thermodynamic perturbation or Free Energy Perturbation Theory (FEP) is based on Zwanzig's equation,[35] which is shown below.

$$\Delta G = k_B T \, ln \left[e^{\left[\frac{-(E_x - E_y)}{k_B T}\right]} \right] \tag{2.43}$$

The Zwanzig equation relates the free energy difference between an initial state, *X*, and a final state, *Y*, and k_B is the Boltzmann constant. E_x and E_y are the energies for the corresponding state, which describe *N* number of particles interacting with the initial state *X* and the final state *Y*. The application of this formula to the medium effects on the interconversion of two molecular entities, *A* and *B* from medium 1 to medium 2, is shown below in the thermodynamic cycle (Figure 2.14) and the corresponding equation 2.44 below.

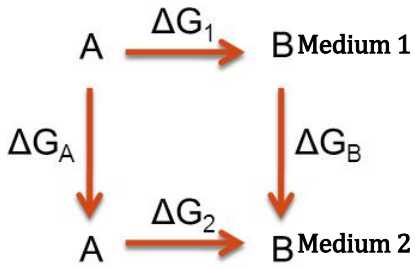


Figure 2.14: The thermodynamic cycle for calculating the state function Gibbs Free Energy, ΔG.

$$\Delta\Delta G(A \rightarrow B) = \Delta G_B - \Delta G_A = \Delta G_2 - \Delta G_1 \qquad (2.44)$$

ΔG_A and ΔG_B correspond to the free energy of solvation of solutes *A* and *B* from medium 1 into medium 2, whereas ΔG_1 and ΔG_2 correspond to the free energy of mutation of solute *A* into solute *B* in constant medium. Since Gibbs free energy is a state function, its value around the complete thermodynamic cycle pictured in Figure 2.14 will equal zero. The reason ΔG_1 and ΔG_2 are calculated and not ΔG_A and ΔG_B are due to the complexities involved in computing ΔG_A and ΔG_B. For example, if solute A is initially in the gas phase (medium 1) and in the final state, solute A is located in a solvent (medium 2); the ΔG_A would correspond to the free energy of solvation on going from gas to solvent. This computation would require extensive reorganization for the entire chemical system to converge energetically.

However, if the medium remains constant, i.e. only medium 1 or medium 2 is used, and the initial solute *A* is transformed into a different final state, solute *B*, the ΔG_1 corresponds to the free energy required for the mutation to occur. The computational demand and cost of such a calculation is practical and is readily preformed, even though the computed ΔG_1 and ΔG_2 do not correspond to any transformation that can be

performed in a wet lab. The free energy difference only depends on the endpoints, allowing for the corresponding Hamiltonians to be manipulated in any way.

2.6.2 Sampling Procedures: forward, backward and double wide

In order for proper convergence of the Zwanzig equation upon the mutation of chemical species *A* into chemical species *B*, geometric and force field parameters such as bond lengths, bond angles, torsions, and the nonbonded interactions have to be perturbed and then adequately sampled. These parameters are perturbed between the two different systems, as previously described, in terms of a coupling parameter, λ, where λ is a value between 0 and 1.[34b] The simplest calculation directly perturbs *A* into *B* to derive free energy. However, when the initial molecular state *A* and the final molecular state *B* are substantially different, a series of simulations for multiple values of λ is necessary. These individual simulations are called "windows" and critical choices have to be made for the individual terms λ for each specific FEP series. The chemical system is first equilibrated for the initial conditions of molecular species *A* with the corresponding initial λ chosen, (λ_i). Then the free energy difference is calculated for each step of $\lambda_i \rightarrow \lambda_{i+1}$. The total free energy for $\lambda = 0$ to $\lambda = 1$ is the sum of the free energy differences for all values of λ_i.

The approach of determining the free energy through a series of calculations from $\lambda_i \rightarrow \lambda_{i+1}$, is known as *forward sampling*.[34b] The FEP calculations could also be performed in the reverse direction, more commonly referred to as *backward sampling*.[34b] For example, if a constant $\Delta\lambda$ of 0.20 is chosen, then a series of five windows would be created to perturb $\lambda = 0.0 \rightarrow 0.20 \rightarrow 0.40 \rightarrow 0.60 \rightarrow 0.80 \rightarrow 1.0$ for *forward sampling*, or $\lambda = 1.0 \rightarrow 0.80 \rightarrow 0.60 \rightarrow 0.40 \rightarrow 0.20 \rightarrow 0.0$ for *backward sampling*. See Figures 2.15 and 2.16 respectively.

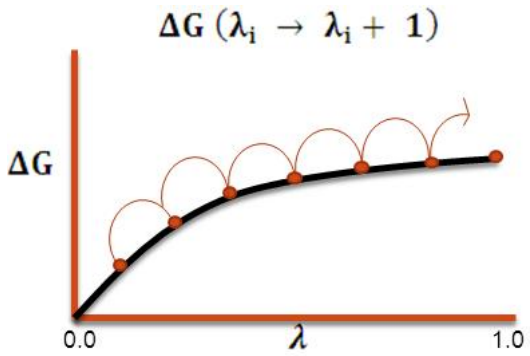


Figure 2.15: Calculation of the free energy difference using forward sampling FEP.

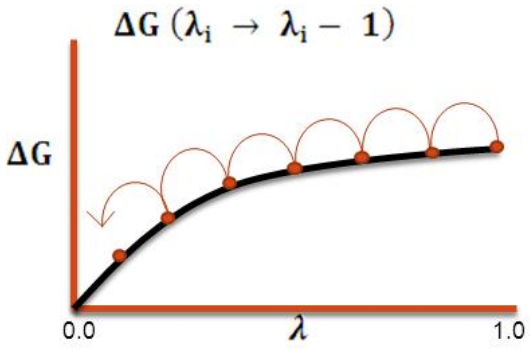


Figure 2.16: Calculation of the free energy difference using backward sampling FEP.

Jorgensen et *al.* modeled the same series of perturbations, using only two simulations to perform the same perturbations previously described.[36] Instead of five individual windows, only two windows are needed; $\lambda = 0.75$ and $\lambda = 0.25$. This approach is called *double-wide sampling*[36] where the forward and reverse perturbations were evaluated simultaneously; $1 \leftarrow 0.75 \rightarrow 0.50$ and $0.50 \leftarrow 0.25 \rightarrow 0.0$.

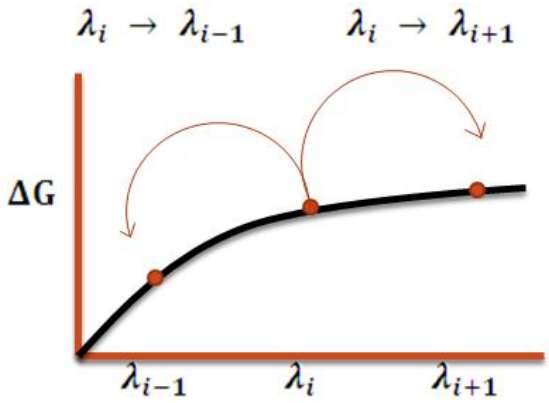


Figure 2.17: Calculation of the free energy difference using double-wide sampling FEP.

In 1985, Jorgensen and Ravimohan performed the first FEP calculations of ethane to methanol in water, using MC and double-wide sampling.[36] The computed uncertainty for this particular mutation was only 0.2 kcal/mol or 3% of the total free energy change.[36] It was concluded in their paper, that double-wide sampling provided a more accurate estimate of the free energy because both the reference perturbation, $\Delta\lambda = 0.20$, and the two perturbed solutes, $\lambda = 0.80$ and $\lambda = 040$, are maintained to evaluate both the forward and reverse free energy changes "on-the-fly".[34b, 36] From the Zwanzig equation previously seen (equation 2.43), the free energy change for $0 \rightarrow 1$ from double-wide sampling is seen in equation 2.45, where *M* represents the intermediate point rather than an endpoint, 0 or 1.

$$\Delta G^{DW} = k_B T \ln \left[\frac{\left(e^{\frac{-(E_0 - E_M)}{k_B T}} \right)_M}{\left(e^{\frac{-(E_1 - E_M)}{k_B T}} \right)_M} \right] \qquad (2.45)$$

2.6.3 Alternatives to FEP

A second alternative to FEP is called thermodynamic integration (TI), see the illustration below in Figure 2.18.

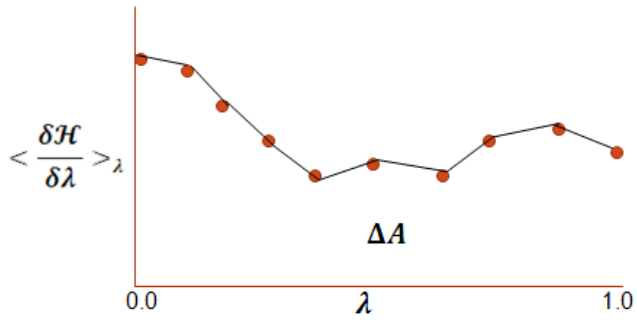


Figure 2.18: Calculation of the free energy difference using thermodynamic integration.

To calculate the free energy difference (ΔA or ΔG) using thermodynamic integration a series of calculations are performed which correspond to discrete values of λ. A few specific points are selected, which can then be integrated to give the corresponding curve. In Figure 2.18, the x-axis represents the values of λ and the y-axis is the change in the corresponding Hamiltonian over the perturbation λ. The total free energy difference (ΔA or ΔG) is then equal to the area under the curve. This approach is commonly referred to as *numerical derivative thermodynamic integration* (NDTI). Similarly to NDTI is the slow growth method, in which all points are simulated and thus integrated. This technique is highly accurate because the resulting Hamiltonian, $H(\lambda_{i+1})$, is practically identical to original Hamiltonian, $H(\lambda_i)$. In general, the smaller the incremented value of λ, regardless of the methodology used, the more accurate the resulting calculation will be. However, as a direct result, the computational demand and cost increase accordingly.

2.6.4 Potentials of Mean Force (PMFs)

The free energy changes that have been considered thus far correspond to a nonphysical mutation of chemical species *A* into a different chemical species *B*. However, chemical reactions are also of interest and unlike the mutations previously described these free energy changes are normally calculated as a function of an inter- or intramolecular reaction coordinate, such as the distance between two atoms or the dihedral angle within a molecule. The corresponding free energy surface along the chosen reaction coordinate is known as a potential of mean force (PMF).[1ab, 2a, 34b] The inter- or intramolecular change between two points on the potential energy surface must be a small incremental change, just as with FEP, to ensure accuracy in the calculation of the relative free energy. Typically these geometric changes are on the order of 2-5° for an angle[2a] and 0.01-0.05Å[1ab, 2a] for bond lengths. Multidimensional PMFs are commonly generated by perturbing between points on the potential energy surface defined by multiple reaction coordinates. For example, the generation of a 2D potential energy surface can be computed via two separate PMFs perturbing the lengths of the reacting bonds during a chemical reaction.

2.7 Introduction to Proper Treatment of Long Range Electrostatic Interactions

Computer simulations of organic reactions in classical fluids, such as common molecular solvents and water, have proven to be successful when the interaction between particles is short ranged. However, when the electrostatic potential is long-ranged the precise method of sampling the particles interactions can have a profound effect on the simulation. In 1978 Adams stated that there are two common approaches to the

summation of charge-charge interactions: either the true periodic nature of the simulation shell is used and infinite-lattice sums are calculated, or the summation is restricted to near neighbor atoms only.[37] The Ewald summation is a popular method for the treatment of these long range charge-charge interactions that retains the infinite array of periodic cells, and allows for particles to interact with all other particles within a particular cutoff distance. Such approximations are required for simulations of ionic liquids. Unlike molecular solvents, ionic liquids are completely composed of anions and cations, creating a microenvironment that is entirely charged. Proper treatment of these long range electrostatic interactions is required for the correct characterization of ionic liquids during computational simulations.

In 1921, a German-born United States crystallographer and physicist, Von Paul Peter Ewald devised a lattice summation to study the energetics of ionic crystals.[38] Ewald developed a method in which a particle interacts with all the other particles in the simulation box and with all of their images in an infinite periodic array. The position of the image boxes can be determined in relation to the central box by specifying a vector whose components are a multiple of the box length; i.e. $(\pm iL,\ \pm jL,\ \pm kL)$ where *i, j* and *k* are equal to a positive integer. The charge-charge contributions to the potential energy, due to all pairs of charges in the central simulation box can be written as follows, see equation 2.46 below.

$$V = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \quad i \neq j \tag{2.46}$$

In equation 2.46, *N* is the total number of charged sites in the system of interest, q_i and q_j are the charge on site *i* and *j* respectively, ε_o is the dielectric constant of the medium and

r_{ij} is the minimal distance between charges i and j. Additionally, the contribution to the potential energy of the system from the charge-charge interactions between the charges, *i* and *j*, in the central box and all images of the particles in the periodic array is:

$$V = \frac{1}{2}\sum_{|\boldsymbol{n}|=0}^{\prime}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{4\pi\varepsilon_0|r_{ij} + \boldsymbol{n}|} \tag{2.47}$$

$$|\boldsymbol{n}| = \left(n_x L, n_y L, n_z L\right) \tag{2.48}$$

This expression, unlike the expression in equation 2.46, is written to include the interactions between all pairs of charges, q_i and q_j, in the central box and all image boxes positioned at a cubic lattice point $|\boldsymbol{n}|$. The prime on the first summation indicates that the series does not include interaction $i = j$ for $n = 0$. There is thus a contribution to the total energy from the interactions in the central box together with the interaction between the central box and all image boxes. One of the major issues with equation 2.47 is that the summation converges extremely slowly due to the fact that it is a conditionally convergent series. A conditionally convergent series is a series that contains a mixture of positive and negative terms, such that alone will form a divergent series; a series without a finite sum. The sum of a conditionally convergent series depends on the order in which the terms are considered and thus adds a level of complexity to the corresponding calculation. Another major issue with equation 2.46 is that the Coulombic term can vary rapidly and dramatically at very small distances, as previously discussed in section 2.1 of this chapter.

2.7.1 The Ewald Summation

The key to calculating the Ewald approximation is to convert the summations into two absolutely convergent series; a short ranged direct sum in Cartesian space and a long

range reciprocal sum in Fourier space, which individually converge more rapidly than the triple summation in equation 2.47.

$$\frac{1}{r} = \frac{f(r)}{r} + \frac{1 - f(r)}{r} \tag{2.49}$$

The mathematical foundation in which the two absolutely convergent Ewald summations are based on is seen in equation 2.49. The goal is to choose an appropriate function, *f(r)*, which can properly handle the rapid variation of *1/r* at small distances of *r* and the slow decay at long distances of *r*. Furthermore, in the Ewald approximation, each charge is surrounded by a neutralizing charge distribution of equal magnitude but of opposite sign, as seen below in Figure 2.19.

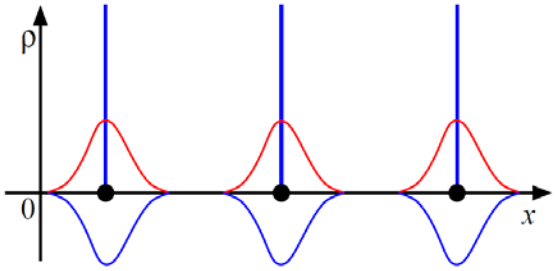


Figure 2.19: The initial set of charges (black dots with the blue vertical line) are surrounded by a Gaussian distribution calculated in real space (the blue Gaussians) to which a cancelling charge distribution is added (red Gaussians), which is calculated in reciprocal space.

The sum over the point charges as seen in equation 2.47, using the fundamental mathematical foundation given in equation 2.49, is now converted to a sum of the interactions between the charges, q_i and q_j, in addition to the neutralizing distributions. This "real space" summation is given by:

$$V = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{|\boldsymbol{n}|=0}^{'}\frac{q_i q_j}{4\pi\varepsilon_0}\frac{erfc\left(\alpha|r_{ij}+\boldsymbol{n}|\right)}{|r_{ij}+\boldsymbol{n}|} \tag{2.50}$$

$$erfc(x) = \frac{2}{\sqrt{\pi}}\int_{x}^{\infty} e^{(-t^2)}\,dt \tag{2.51}$$

The "real space" term, which is calculated in Cartesian space, in the Ewald summation utilizes the complementary error function, (equation 2.51). The critical advantages to using this equation are that it converges very rapidly, and beyond the specified cutoff distance its value can be considered negligible. Importantly, the rate of convergence for the real space term is directly dictated by the width of the cancelling Gaussian distribution, α. The wider the width of the Gaussian, the faster the series will converge. However, the value of α should be chosen so that only interactions with other charges within the specified cutoff are included. Again, the prime on the last summation indicates that the series does not include interactions when $i = j$ and $n = 0$.

The second term in the Ewald summation is the "reciprocal space" term interpolated in Fourier space, in which a second charge distribution is added to the system which exactly counteracts the first charge distribution Gaussians.

$$V = \frac{1}{2}\sum_{k\neq 0}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\pi L^3}\frac{q_i q_j}{4\pi\varepsilon_0}\, e^{\left(-\frac{k^2}{4\alpha^2}\right)}\cos\left(\boldsymbol{k}\cdot\boldsymbol{r_{ij}}\right) \tag{2.52}$$

$$\boldsymbol{k} = \frac{2\pi n}{L} \tag{2.53}$$

This reciprocal space term also converges much faster than the original point charge summation, however the number of terms included in this expression increases as the width of the Gaussian, α, increases. Therefore considerable care must be taken when

choosing an appropriate value for the width of the Gaussians, α. As previously discussed the "real-space" term converges faster with large values of α , whereas the "reciprocal space" term converges faster with a small α value. A balance between the "real space" term and the "reciprocal space" term has to be reached for proper treatment of the long range electrostatic interactions, as well as to maintain the balance of computational accuracy and cost.

The sum of the Gaussian functions in the "real space" term includes self-interaction and therefore, a third term is required in the Ewald summation to counteract this behavior. This term is consequently subtracted from the Ewald summation and is shown below in equation 2.54.

$$V = -\frac{a}{\sqrt{\pi}} \sum_{k=1}^{N} \frac{q^2 k}{4\pi\varepsilon_0} \tag{2.54}$$

The last term in the Ewald summation takes into account if the medium surrounding the simulated box of solvent is a vacuum, meaning the relative permittivity = 1. If this is the case, the fourth term must be included and is commonly referred to as the "correction term".

$$V_{correction} = \frac{2\pi}{3L^3} \left| \sum_{i=1}^{N} \frac{q_i}{4\pi\varepsilon_0} \right|^2 \tag{2.55}$$

The overall and complete Ewald summation is therefore:

$$V = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left\{\sum_{|\boldsymbol{n}|=0}^{\infty}\frac{q_i q_j}{4\pi\varepsilon_0}\frac{erfc\left(\alpha|r_{ij}+\boldsymbol{n}|\right)}{|r_{ij}+\boldsymbol{n}|}\right. \tag{2.56}$$

$$+\sum_{k\neq 0}\frac{1}{\pi L^3}\frac{q_i q_j}{4\pi\varepsilon_0}\, e^{\left(-\frac{k^2}{4\alpha^2}\right)}\cos\left(\boldsymbol{k}\cdot\boldsymbol{r_{ij}}\right)$$

$$\left.-\frac{a}{\sqrt{\pi}}\sum_{k=1}^{N}\frac{q^2 k}{4\pi\varepsilon_0}+\frac{2\pi}{3L^3}\left|\sum_{i=1}^{N}\frac{q_i}{4\pi\varepsilon_0}\right|^2\right\}$$

The first term is the “real space” term in which the original expression for the sum over point charges was converted into a sum of interactions between the neutralizing distributions, incorporating the complementary error function. The second term is the “reciprocal space” summation which exactly counteracts the first neutralizing distribution added with the ‘real space” term. The third expression accounts for the neutralizing “real space” Gaussian interacting with itself, and lastly the fourth term is the correction term which has to be implemented if the environment surrounding the simulation box has a relative permittivity equal to one.

One of the major downsides to performing an Ewald summation is the fact that the simulations will scale approximately as N^2, where *N* is the number of particles in the simulation central box and the same density of reciprocal vectors is used, i.e. constant α. This scaling is a consequence of the balance that has to be established for the width of the neutralizing Gaussians of the “real space” and a “reciprocal space” terms to converge. Additionally, it is the calculation of the “reciprocal space” term that is ultimately rate determining and plays the largest factor in the computationally expensive Ewald calculation.

The Ewald summation has been extensively used to correctly simulate highly charged systems, such as simulations of ionic liquids, lipid bilayers, proteins, and DNA. However, due to the expense of performing an Ewald summation, numerous alternative methods have been developed. Hockney and Eastwood developed the Particle-Mesh Ewald (PME)[39] method in 1988, Essman et *al.* developed the Smooth Particle Mesh Ewald (SPME)[40] method. Both methods scale approximately like *N·log(N)*. More recently, Wolf et *al.* in 1999 and Zahn et *al*. in 2002[41] created a set of potentials that were constructed using two very different, separable, computational tricks (1) shifting through the use of image charges and (2) dampening the electrostatic interactions. These potentials built the foundation for a new set of pairwise summation methods in which the charge neutralization within the cutoff sphere plays a significant role in the energy convergence.[41-42]An additional alternative method to the traditional infinite lattice sums was fashioned by Kale and Herzfeld in 2011 with the Shifted Force Gradient Method (SFG).[42a] The SFG method is an expansion of the Taylor series found in the Shifted Force method developed by Levitt and coworkers.[43] Lastly, and perhaps most significantly for the correct treatment of ionic liquids, there was the development of a new method called the Shifted Force 3rd Derivative (SF3) by McCann and Acevedo.[42c] The SF3 was tested on 59 unique ionic liquid combinations of 1-alkyl-3-methylimidazolium [RMIM] (R = M (methyl), E (ethyl), B (butyl), H (hexyl), and O (octyl)) and N-alkylpyridinium [RPyr] cations, along with Cl^-, PF_6^-, BF_4^-, NO_3^-, $AlCl_4^-$, $Al_2Cl_7^-$, and TfO^- anions in conjunction with QM/MM/MC calculations utilizing the custom OPLS-AA ionic liquid force field.[42c] Multiple cutoff distances and electrostatic damping values were directly compared to the energetics from full Ewald sums. McCann

and Acevedo reported that the SF3 method was accurate and computationally efficient in the treatment of the long range electrostatic interactions during the simulation of ionic liquids.[42c]

2.8 Summary

The hybrid QM/MM methodology is a molecular simulation method that combines the strengths of both QM and MM calculations, ultimately allowing for the study of chemical processes in solution and in macromolecules. An important advantage of QM/MM is the efficiency and lower computational costs relative to full QM. QM offers accuracy to the modeling of chemical systems where electron motion is paramount, whereas MM simulations allow for a large number of atoms to be modeled as the solvent or non-reacting portion of a macromolecule. A recent challenge for the QM/MM approach has been to use a high level *ab initio* description of the QM region, while still being able to perform extensive sampling in order to accurately and properly evaluate free energies of chemical processes. However, with continuing advancement in computer technology and complementary forward thinking in the scientific field, the field of QM/MM simulations will continue to grow and expand to include new challenging chemical problems and higher level, modern methodologies.

Chapter 2: References

1. (a) Aqvist, J. W., A., *Chem. Rev.* **1993,** *93*, 2523; (b) Gao, J., *Acc. Chem. Res.* **1996,** *29*, 298; (c) Field, M. J., *J. Chem. Phys.* **1992,** *96*, 4583; (d) Gao, J.; Xia, X., *Science* **1992,** *258*, 631; (e) Gao, J., *J. Am. Chem. Soc* **1993,** *115*, 2930; (f) Wesolowski, T. A.; Warshel, A., *J. Phys. Chem.* **1993,** *97*, 8050; (g) Stanton, R. V.; Hartsough, D. S.; Merz, K. M. J., *J. Phys. Chem.* **1993,** *97*, 11868; (h) Warshel, A.; Karplus, M., *J. Am. Chem. Soc.* **1972,** *94*, 5612-5625; (i) Thompson, M. A.; Schenter, G. K., *J. Phys. Chem.* **1995,** *99*, 6375; (j) Gao, J., *J. Am. Chem. Soc* **1995,** *117*, 8600-8607; (k) Furlani, T. R.; Gao, J., *J. Org. Chem.* **1996,** *61*, 5492; (l) Kaminski, G. A.; Jorgensen, W. L., *J. Phys. Chem. B* **1998,** *102*, 1787-1796; (m) Acevedo, O.; Jorgensen, W. L., *Ann. Rep. Comput. Chem.* **2006,** *3*, 263-278; (n) Jorgensen, W. L., *Acc. Chem. Res.* **1989,** *22*, 184-189; (o) Jorgensen, W. L.; Blake, J. F.; Lim, D.; Severance, D. L., *J. Chem. Soc., Faraday Trans.* **1994,** *90*, 1727-1732; (p) Kerdcharoen, T.; Liedl, K. R.; Rode, B. M., *Chem. Phys.* **1996,** *211*, 313; (q) Barnes, J. A.; Williams, I. H., *J. Chem. Soc., Chem. Commun.* **1996**, 193; (r) Hartsough, D. S.; Merz, K. M. J., *J. Phys. Chem.* **1995,** *99*, 11266; (s) Lyne, P. D.; Mullholland, A. J.; Richards, W. G., *J. Am. Chem. Soc* **1995,** *117*, 11345; (t) Freindorf, M.; Gao, J., *J. Comput. Chem* **1996,** *17*, 386; (u) Thompson, M. A., *J. Am. Chem. Soc* **1995,** *117*, 11341; (v) Thompson, M. A., *J. Phys. Chem.* **1995,** *99*, 4794; (w) Warshel, A.; Levitt, M., *J. Mol. Biol.* **1976,** *103*, 227-249; (x) Field, M. J.; Bash, P. A.; Karplus, M., *J. Comput. Chem* **1990,** *11*, 700; (y) Gao, J.; Habibollazadeh, D.; Shao, L., *J. Phys. Chem.* **1995,** *99*, 16460;

(z) Wang, J.; Boyd, R. J.; Laaksonen, A., *J. Chem. Phys.* **1996,** *104*, 7261; (aa) Cunningham, M. A.; Ho, L. L.; Nguyen, D. T.; Gillilan, R. E.; Bash, P. A., *Biochemistry* **1997,** *36*, 4800; (ab) Acevedo, O.; Jorgensen, W. L., *Acc. Chem. Res.* **2010,** *43*, 142-151; (ac) Acevedo, O.; Jorgensen, W. L., *Org. Lett.* **2004,** *6*, 2881-2884; (ad) Acevedo, O.; Jorgensen, W. L., *J. Am. Chem. Soc.* **2005,** *127*, 8829-8834; (ae) Pan, D. G.; Pan, Y. K., *J. Org. Chem.* **1999,** *64*, 1151-1159; (af) Acevedo, O.; Jorgensen, W. L., *J. Org. Chem.* **2006,** *71*, 4896-4902; (ag) Acevedo, O.; Jorgensen, W. L., *J. Am. Chem. Soc.* **2006,** *128*, 6141-6146; (ah) Acevedo, O.; Jorgensen, W. L., *J. Chem. Theory Comput.* **2007,** *3*, 1412-1419; (ai) Acevedo, O.; Jorgensen, W. L.; Evanseck, J. D., *J. Chem. Theory Comput.* **2007,** *3*, 132-138; (aj) Acevedo, O.; Squillacote, M. E., *J. Org. Chem.* **2008,** *73*, 912-922; (ak) Chandrasekhar, J.; Shariffskul, S.; Jorgensen, W. L., *J. Chem. Phys. B.* **2002,** *106*, 8078-8085; (al) Acevedo, O.; Jorgensen, W. L., *J. Phys. Chem. B.* **2010,** *114*, 8425-8430; (am) Acevedo, O.; Armacost, K., *J. AM. Chem. Soc.* **2010,** *132*, 1966-1975.

2. (a) Jorgensen, W. L.; Tirado-Rivers, J., *J. Comput. Chem.* **2005,** *26*, 1689-1700; (b) Jorgensen, W. L., *BOSS, Version 4.5; Yale University: New Haven, CT* **2005**.

3. (a) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L., *J. Comput. Chem.* **2002,** *23*, 1601-1622; (b) Tubert-Brohman, I.; Guimaraes, C. R. W.; Repasky, M. P.; Jorgensen, W. L., *J. Comput. Chem* **2003,** *25*, 138-150; (c) Tubert-Brohman, I.; Guimaraes, C. R. W.; Jorgensen, W. L., *J. Chem. Theory Comput.* **2005,** *1*, 817-823.

4. Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; Slater, D. J.; Swaminathan, S.; Karplus, M., *J. Comput. Chem.* **1983,** *4*, 187-217.

5. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M. J.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., *J. Am. Chem. Soc.* **1995,** *117*, 5179-5197.

6. (a) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rivers, J., *J. Am. Chem. Soc.* **1996,** *118*, 11225-11236; (b) Jorgensen, W. L.; Tirado-Rivers, J., *Proc. Natl. Acad. Sci. U.S.A.* **2005,** *102*, 6665-6670; (c) Sambasivarao, S. V.; Acevedo, O., *J. Chem. Theory Comput.* **2009,** *5*, 1038-1050.

7. (a) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J., *J. Am. Chem. Soc* **1984,** *106*, 6638; (b) Jorgensen, W. L., *J. Phys. Chem.* **1986,** *90*, 1276; (c) Jorgensen, W. L., *J. Phys. Chem.* **1986,** *90*, 6379; (d) Briggs, J. M.; Matsui, T.; Jorgensen, W. L., *J. Comput. Chem.* **1990,** *11*, 958; (e) Briggs, J. M.; Nguyen, T. B.; Jorgensen, W. L., *J. Phys. Chem.* **1991,** *95*, 3315; (f) Jorgensen, W. L.; Swenson, C. J., *J. Am. Chem. Soc.* **1985,** *107*, 1489; (g) Jorgensen, W. L.; Severance, D. L., *J. AM. Chem. Soc.* **1990,** *112*, 4768; (h) Jorgensen, W. L.; Nguyen, T. B., *J. Comput. Chem.* **1993,** *14*, 195; (i) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L., *J. Am. Chem. Soc.* **1991,** *113*, 2810; (j) Jorgensen, W. L.; Tirado-Rivers, J., *J. Am. Chem. Soc.* **1988,** *110*, 1657.

8. Jones, J. E., *Proc. R. Soc. Lond. A.* **1924,** *106*, 463-477.

9. (a) Pople, J. A., *J. Chem. Phys.* **1965,** *43*, 2026-2033; (b) Pople, J. A., *J. Chem. Phys.* **1965,** *43*, S136-S149; (c) Pople, J. A., *J. Chem. Phys.* **1966,** *44*, 3289-3296;

(d) Santry, D. P.; Segal, G. A., *J. Chem. Phys.* **1967,** *47*, 158-174; (e) Pople, J. A., *J. Chem. Phys.* **1965,** *43*, S129-S135.

10. Pople, J. A., *J. Chem. Phys.* **1967,** *47*, 2026-2033.

11. (a) Bingham, R. C.; Dewar, M. J. S.; Lo, D. H., *J. Am. Chem. Soc.* **1974,** *97*, 1285-1293; (b) Bingham, R. C.; Dewar, M. J. S.; Lo, D. H., *J. Am. Chem. Soc.* **1974,** *97*, 1294-1301; (c) Bingham, R. C.; Dewar, M. J. S.; Lo, D. H., *J. Am. Chem. Soc.* **1974,** *97*, 1302-1306; (d) Bingham, R. C.; Dewar, M. J. S.; Lo, D. H., *J. Am. Chem. Soc.* **1974,** *97*, 1307-1310.

12. (a) Dewar, M. J. S.; Thiel, W., *J. Am. Chem. Soc.* **1977,** *99*, 4899-4907; (b) Dewar, M. J. S.; Thiel, W., *J. Am. Chem. Soc.* **1977,** *99*, 4907-4917.

13. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P., *J. Am. Chem. Soc.* **1985,** *107*, 3902-3909.

14. (a) Stewart, J. J. P., *J. Comput. Chem.* **1989,** *10*, 209-220; (b) Stewart, J. J. P., *J. Comput. Chem.* **1989,** *10*, 221-264.

15. Dewar, M. J. S.; Rzepa, H. S., *J. Am. Chem. Soc.* **1978,** *100*, 58-67.

16. Dewar, M. J. S.; McKee, M. L., *J. Am. Chem. Soc.* **1978,** *100*, 777-784.

17. Dewar, M. J. S.; Rzepa, H. S., *J. Am. Chem. Soc.* **1977,** *99*, 5231-5241.

18. Dewar, M. J. S.; McKee, M. L.; Rzepa, H. S., *J. Am. Chem. Soc.* **1978,** *100*, 3607-3607.

19. Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L., *J. Comput. Chem.* **2002,** *23*, 498-510.

20. (a) Bernal-Uruchurtu, M. I.; Ruiz-Lopez, M. F., *J. Comput. Chem.* **2000,** *330*, 118; (b) Bernal-Uruchurtu, M. I.; Martins-Costa, M. T. C.; Millot, C.; Ruiz-Lopez, M. F., *J. Comput. Chem.* **2000,** *21*, 572.

21. Allen, C.; Sambasivarao, S. V.; Acevedo, O., *J Am. Chem. Soc.* **2013,** *31*, 1065-1072.

22. Singh, U.; Kollman, P. J., *J. Comput. Chem.* **1986,** *7*, 718.

23. (a) Bakowies, D.; Thiel, W., *J. Comput. Chem.* **1996,** *17*, 87-108; (b) Eurenius, K. P.; Chatfiled, D. C.; Brooks, B. R.; Hodoscek, M., *Int. J. Quantum Chem.* **1996,** *50*, 1189-1200; (c) Bakowies, D.; Thiel, W., *J. Phys. Chem.* **1996,** *100*, 10580-10594.

24. Antes, I., *Ph.D. Thesis, University of Zurich* **1998**.

25. (a) Svensson, M.; Humbel, S.; Froese, R. J.; Matsubara, T.; Sieber, S.; Morokuma, K., *J. Phys. Chem.* **1996,** *100*, 19357-19363; (b) Maseras, F.; Morokuma, K., *J. Comput. Chem.* **1995,** *16*, 1170-1179; (c) Humbel, S.; Sieber, S.; Morokuma, K., *J. Chem. Phys.* **1996,** *105*, 1959; (d) Svensson, M.; Humbel, S.; Morokuma, K., *J. Chem. Phys.* **1996,** *105*, 3654.

26. (a) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G., *J. Comput.-Aided Mol. Des.* **1995,** *9*, 87; (b) Winget, P.; Thompson, J. D.; Xidos, J. D.; Cramer, C. J.; Truhlar, D. G., *J. Phys. Chem. A.* **2002,** *106*, 10707-10717; (c) Blagovic, M. U.; Morales de Tirado, P.; Pearlman, S. A.; Jorgensen, W. L., *J. Comput. Chem.* **2004,** *25*, 1322-1332; (d) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G., *J. Comput. Chem.* **2003,** *24*, 1291-1304; (e) Vilseck, J. Z.; Sambasivarao, S. V.; Acevedo, O., *J. Comput. Chem.* **2011,** *32*, 2836-2842.

27. (a) Jorgensen, W. L., *Theor. Chem. Acc.* **2000,** *103*, 225-227; (b) Jorgensen, W. L., *Encyclopedia of Computational Chemistry* **1998**, 1754-1763.

28. Metropolis, N.; Rosenbluth, A. E.; Rosenbluth, M. N.; Teller, A. H.; Teller, E., *J. Chem. Phys.* **1953,** *21*, 1087-1092.

29. (a) Schlegel, H. B., *Interdisciplinary Reviews: Computational Molecular Science* **2011,** *1*, 790-809; (b) Schlegel, H. B., *Encyclopedia of Computational Chemistry* **1998,** *2*, 1136-1142.

30. Theodorou, D. N.; Suter, U. W., *J. Chem. Phys.* **1985,** *82*, 955-966.

31. Barker, J. A.; Henderson, D., *Rev. Mod. Phys.* **1976,** *48*, 587.

32. (a) Karnicky, J. F.; Pings, C. J., *Adv. Chem. Phys.* **1976,** *34*, 157-202; (b) Enderby, J. E.; Cummings, S.; Herdman, G. J.; Neilson, G. W.; Salmon, P. S.; Skipper, N., *J. Phys. Chem.* **1987,** *91* (5851-5858).

33. Jorgensen, W. L., *J. Chem. Phys.* **1983,** *87*, 5304-5314.

34. (a) Kirkwood, J. G., *J. Chem. Phys.* **1935,** *3*, 300-313; (b) Jorgensen, W. L.; Thomas, L. L., *J. Chem. Theory Comput.* **2008,** *4*, 869-876.

35. Zwanzig, R. W., *J. Chem. Phys.* **1954,** *22*, 1420-1426.

36. Jorgensen, W. L.; Ravimohan, C., *J. Chem. Phys.* **1985,** *83*, 3050-3055.

37. Adams, D., *Chemical Physics Letters* **1979,** *62*, 329-332.

38. Ewald, P., *Ann. Phys.* **1921,** *369*, 253-287.

39. (a) Hockney, R. W.; Eastwood, J. W., Computer Simulations Using Particles. **1988**; (b) Darden, T., *J. Chem. Phys.* **1993,** *98*, 10089-10092.

40. Essmann, U., *J. Chem. Phys.* **1995,** *103*, 8577-8593.

41. Zahn, D., *J. Chem. Phys. B.* **2002,** *106*, 10725.

42. (a) Kale, S.; Herzfeld, J., *J. Chem. Theory Comput.* **2011,** *7*, 3620-3624; (b) Hansen, J. S.; Schroder, T. B.; Dyre, J. C., *J. C. J. Phys. Chem. B* **2012,** *116*, 5738-5743; (c) McCann, B. W.; Acevedo, O., *J. Chem. Theory Comput.* **2012,** *in press*.

43. Levitt, M.; Hirshberg, M.; Sharon, R.; Daggett, V., *Comput. Phys. Commun.* **1995,** *91*, 215-231.

Chapter 3

An Ionic Liquid Dependent Mechanism for Base Catalyzed β-Elimination Reactions from QM/MM Simulations

3.1 Abstract

Ionic liquids have been proposed to induce a mechanistic change in the reaction pathway for the fundamentally important base-induced β-elimination class compared to conventional solvents. The role of the reaction medium in the elimination of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane via two bases, piperidine and pyrrolidine, has been computationally investigated using methanol and the ionic liquids 1-butyl-3-methylimidazolium tetrafluoroborate and hexafluorophosphate [BMIM][BF_4] and [BMIM][PF_6], respectively. QM/MM Monte Carlo simulations utilizing free-energy perturbation theory found the ionic liquids did produce a reaction pathway change from an E1cB-like mechanism in methanol to a pure E2 route that is consistent with experimental observations. The origin of the ionic liquid effect has been found as: (1) a combination of favorable electrostatic interactions, for example, bromine-imidazolium ion, and (2) π–π interactions that enhance the coplanarity between aromatic rings maximizing the electronic effects exerted on the reaction route.

Reprinted with permission from CALEY ALLEN, SOMISETTI V. SAMBASIVARAO, AND ORLANDO ACEVEDO; JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, 2013, *135*, 1065-1072. Copyright 2013 American Chemical Society.

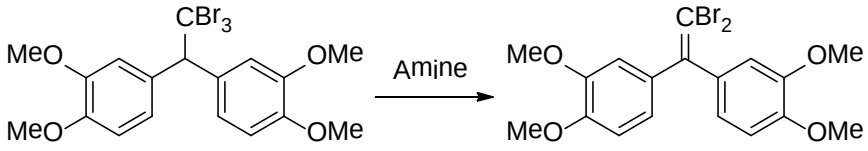
Complete Supporting Information Available at:
http://pubs.acs.org/doi/suppl/10.1021/ja3098614

Solute−solvent interaction energies have been analyzed and show that liquid clathrate solvation of the transition state is primarily responsible for the observed mechanistic changes. This work provides the first theoretical evidence of an ionic liquid dependent mechanism and elucidates the interplay between sterics and electrostatics crucial to understanding the effect of these unique solvents upon chemical reactions.

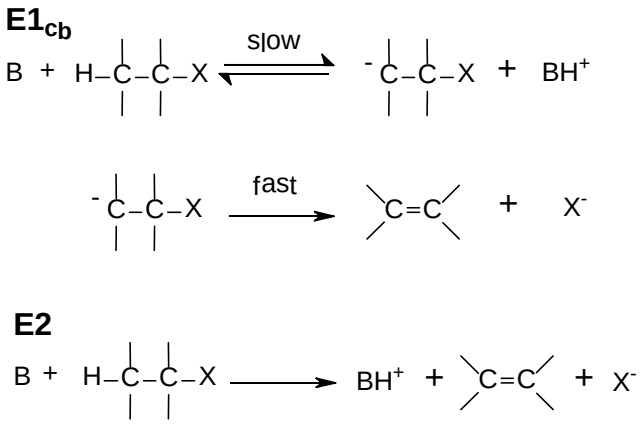
3.2 Introduction

The effect of room temperature ionic liquids (RTILs) upon the β-elimination reaction of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane by cyclic amines (Scheme 1), was reported by D'Anna et al.[1] and an interesting hypothesis was put forth; that a change in the reaction mechanism occurs from an irreversible E1cb route in methanol[2] to an E2 in the ionic liquids. The E2 mechanism is a one-stage process in which the base attacks the β-hydrogen and abstracts it with a concomitant cleavage of the α-C−Br bond; in the E1cb process, the removal of the β-hydrogen is rate limiting and generally reversible, and can be detected by isotopic exchange studies (Scheme 3.2).[3] Despite the obvious differences in the mechanisms, experimentally distinguishing between the irreversible E1cb and E2 mechanisms for dehydrohalogenation reactions can be notoriously difficult.[2,4-7] The present work applied mixed quantum and molecular mechanical (QM/MM) calculations utilizing Monte Carlo sampling and free-energy perturbation theory (MC/FEP) to the β-eliminations in methanol, [BMIM][BF_4], and [BMIM][PF_6] to investigate the proposed mechanism change and its origins. Two secondary cyclic amine bases, pyrrolidine and piperidine, were studied as intriguingly no primary or tertiary amines were found to experimentally induce eliminations. The QM/MM simulations with explicit solvent representation can provide the medium-dependence of the activation barriers and atomic-

level structural detail for characterization of the nature of the ionic liquids. Comparisons are made to density functional theory (DFT) calculations using an implicit continuum model to simulate conventional solvent effects. The results presented provide new insights as to the ionic liquid effects on the reaction pathway and help clarify experimental observations.



Scheme 3.1: β-Elimination Reaction of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane.



Scheme 3.2: General E1cb and E2 elimination mechanisms.

3.3 Computational Methods

QM/MM calculations were carried out on 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with piperidine and pyrrolidine in [BMIM][BF_4], [BMIM][PF_6], and methanol. The solutes were treated with the PDDG/PM3 semiempirical QM method.[8,9] PDDG/PM3 has given excellent results for a wide variety

of organic and enzymatic reactions in the solution-phase.[10-14] Potentials of mean force (PMF) calculations coupled to Metropolis Monte Carlo (MC) statistical mechanics were used to build a free-energy profile for the β-elimination reactions at 25 °C and 1 atm. The starting geometries for the solutes were determined by executing a MC conformational search that resulted in up to 100 unique structures. The top ten most favorable MC structures were then recomputed using full DFT geometry optimizations and the resultant lowest energy structure was used as the starting geometry for the QM/MM calculations. The solvent molecules were represented explicitly using our custom ionic liquid OPLS-AA force field[15] and the united-atom OPLS force field for methanol.[16] The current QM/MM methodology allows simulations of reactions in solution on-the-fly with full sampling and polarization of the solutes by the environment.[17] The systems consisted of the reactants plus 395 solvent molecules for methanol or 188 ion pairs for the ionic liquids. The boxes are periodic and tetragonal with c/a = 1.5 where a is 26.7, 34.3, and 35.5 Å for methanol, [BMIM][BF_4], and [BMIM][PF_6], respectively, with long range electrostatic interactions handled with Ewald summations. The ionic liquid boxes were thoroughly equilibrated by raising the temperature to 1000 °C and carrying out 10 million configurations in the NVT ensemble. The simulations were then equilibrated at 25 °C for 100-200 million MC steps in the NPT ensemble. The heating/NVT and equilibration/NPT simulations on each ionic liquid system were repeated sequentially an average of 4-6 times until the energy and volume of the system no longer changed. Solutes were inserted with the appropriate solute geometry corresponding to each free energy perturbation (FEP) window and re-equilibrated for minimally 100 million MC configurations. The computation of the QM energy and atomic charges was performed for each attempted

move of the solute, which occurred every 100 configurations. For electrostatic contributions to the solute-solvent energy, CM3 charges[18] were obtained for the solute and scaled by 1.14 to reflect the polarization in a condensed-phase environment.[19] In addition, Lennard-Jones interactions between solute and solvent atoms were taken into account using OPLS parameters. This combination is appropriate for a PM3-based method as it minimizes errors in computed free energies of hydration.[20]

The simulations were performed with the BOSS program.[21] All cations were fully flexible, i.e. all bond stretching, angle bending, and torsional motions were sampled. Anions were simulated as rigid molecules. The use of rigid anions in OPLS-AA has been shown to provide an accurate representation of ionic liquid physical properties, including use as a reaction medium for computed QM/MM Diels-Alder[22] and Kemp elimination[15] reaction studies. Solute-solvent and solvent-solvent intermolecular cutoff distances of 12 Å were employed for the tail carbon atom of each side chain (methyl and alkyl), a midpoint carbon on the alkyl chain, and the ring carbon between both nitrogens for imidazolium. Center atoms, e.g. B in BF_4^- and P in PF_6^-, were used for the anions. If any distance is within the cutoff, the entire solvent-solvent interaction was included. Adjustments to the allowed ranges for rotations, translations, and dihedral angle movements led to overall acceptance rates of about 30% 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.

Free energy maps were computed by using a distance, R_{NH} - R_{HC}, for the proton transfer between the nitrogen on piperidine/pyrrolidine and the reacting hydrogen on the solute; R_{NH} + R_{HC} was kept constant at 2.85 Å. The fixed distance of 2.85 Å was

determined to be appropriate from our recent work[13,15] and additional test calculations. Our 5th order polynomial quadrature method was used to provide a 7-fold improvement in speed over traditional potentials of mean force (PMF) methods.[13] A second perturbation was necessary, R_{CBr}, which entailed breaking of the C-Br bond. Combining the R_{NH} - R_{HC} PMF which runs along one reaction coordinate with the R_{CBr} PMF in a second direction produced a two-dimensional (2D) PMF. The resultant free-energy map was used to identify minima and the transition state present in the reaction. The breaking of the C-Br bond was split into increments of 0.025 Å. Each PMF calculation required extensive reorganization of the solvent for the ionic liquid, requiring up to 125 million configurations of equilibration followed by 10 million MC steps of averaging per FEP window; in methanol, 2 and 5 million steps of equilibration and averaging, respectively, sufficed for each reaction. Every solution-phase MC/FEP calculation required over 100 million single point QM calculations per free-energy map in the ionic liquids, demonstrating the need for highly efficient QM methods.

The M06-2X density functional method[23,24] and 6-31+G(d, p) basis set was also used to optimize geometries in vacuum, methanol, and water using Gaussian 09.[25] The effect of solvent was explored by full DFT geometry optimizations using the conductor-like polarizable continuum model (CPCM) with the UFF cavity.[26] Frequency calculations were performed in order to verify all stationary points as minima for ground states or as saddle points for transition structures. All calculations were run on a Linux cluster at Auburn University and on computers located at the Alabama Supercomputer Center.

3.4 Results and Disscusion

3.4.1 Energetics

The QM/MM/MC calculations for the β-elimination of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane in methanol gave computed activation barriers $\Delta G^{\ddagger}$ of 36.5 and 34.6 kcal/mol when using piperidine and pyrrolidine, respectively. The $\Delta\Delta G^{\ddagger}$ values between the secondary amines are consistent with experimental trends, where pyrrolidine is reported to react faster than piperidine based on measurements of second order rate constants.[1] Interestingly, the ring dimension rather than nitrogen basicity appears to control the rate, where the flexibility of the ring is thought to play a major role.[1] Error ranges in the computed free-energy values have been estimated from fluctuations in the ΔG values for each FEP window using the batch means procedure with batch sizes of 0.5 million configurations; computed errors in the free energies imply overall uncertainties in the $\Delta G^{\ddagger}$ of ca. 0.5 kcal/mol. Table 3.1 gives a summary of the activation energies and reacting geometries for the reactions in methanol and the ionic liquids. Errors in the free energies and geometries for the ionic liquids result in overall uncertainties of approximately 1.5 kcal/mol and ± 0.1 Å.

	$\Delta G^{\ddagger}$ (calc)	R(C-Br)	$\Delta G^{\ddagger}$ (exptl)[b]
piperidine			
methanol	36.5	2.43	-
[BMIM][BF_4]	34.8	2.68	24.2
[BMIM][PF_6]	29.2	2.58	23.9
pyrrolidine			
methanol	34.6	2.33	-
[BMIM][BF_4]	36.3	2.63	22.7
[BMIM][PF_6]	28.9	2.60	-

[a] PDDG/PM3 and MC/FEP. [b] Ref [1].

Table 3.1: Free Energy of Activation, $\Delta G^{\ddagger}$ (kcal/mol) and transition structure geometries (Å) at 25 °C for the β-elimination of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane in ionic liquid from QM/MM/MC calculations.[a]

The calculated $\Delta G^{\ddagger}$ overestimation is a systematic error common in many organic reactions when employing a semiempirical method.[27] Dewar also reported mixed energetic agreement with experiment when employing the AM1 method on elimination reactions, but the analogous trends provided excellent results in differentiating between E2, E1cb-like, and SN2 mechanisms.[28,29] In addition, COSMO-AM1 and experimental values of the free activation enthalpy on a set of elimination reactions in water showed large discrepancies; however, reactivity aspects were correctly predicted.[30] 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 dehydrobromination reaction as similar findings have been reported for multiple Diels-Alder reactions,[22,31,32] ene reactions,[14] Claisen rearrangements,[11] and methyl transfer reactions.[33] Conceivably, a straightforward 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. Consequently, there is no difference in leaving the Hamiltonian in its original form or in scaling the energies when one considers the relative solvent effects.

3.4.2 E1cb vs. E2 Mechanism

A representative free energy map from the QM/MM/MC calculations in methanol (Figure 3.1) predicts a concerted E2 transition structure with a large amount of proton transfer characteristic of an E1cb-like mechanism, but not the irreversible stepwise E1cb route proposed in recent publications[1,2,5] as no carbanion intermediate was located. The differences in the results may be rationalized through the disparities in base strengths used, where the base reported MeO^- is an extremely strong base relative to the cyclic amines and could potentially accelerate proton removal but not affect the carbanion decomposition.[3] However, pyrrolidine was reported to induce elimination of the current tribromo-ethane reactant faster than a methoxide/methanol reaction despite the difference in basicity.[1] Experimental studies have proposed that the transition structure should be strikingly similar for E1cb and E2 mechanisms.[5,6] For example, Gandler and Jencks speculated in 1982 about a theoretical potential energy surface for the elimination reactions of (2-arylethyl)quinulidinium ions where the transition state for the E1cb mechanism converts into an E2 mechanism as a substituent change to a β-phenyl group caused the carbanion to become less stable and cease to exist (Figure 3.2).[7] In addition to the β-phenyl groups present in the current reaction, Br as a leaving group generally tends to favor the E2 mechanism, as does the use of a moderate-strength base, such as the cyclic amines piperidine and pyrrolidine, in polar solvents.[34,35] It should be noted that the

banding observed in Figure 3.1 is a consequence of using of 0.05 Å increments in the FEP calculations. To locate the critical points more precisely, the regions surrounding the free-energy minima and maxima from the initial maps in all solvents were explored using final increments of 0.025 Å with increased sampling. This provided the refined results in the energetics and geometries for the β-elimination reactions summarized in Table 3.1.

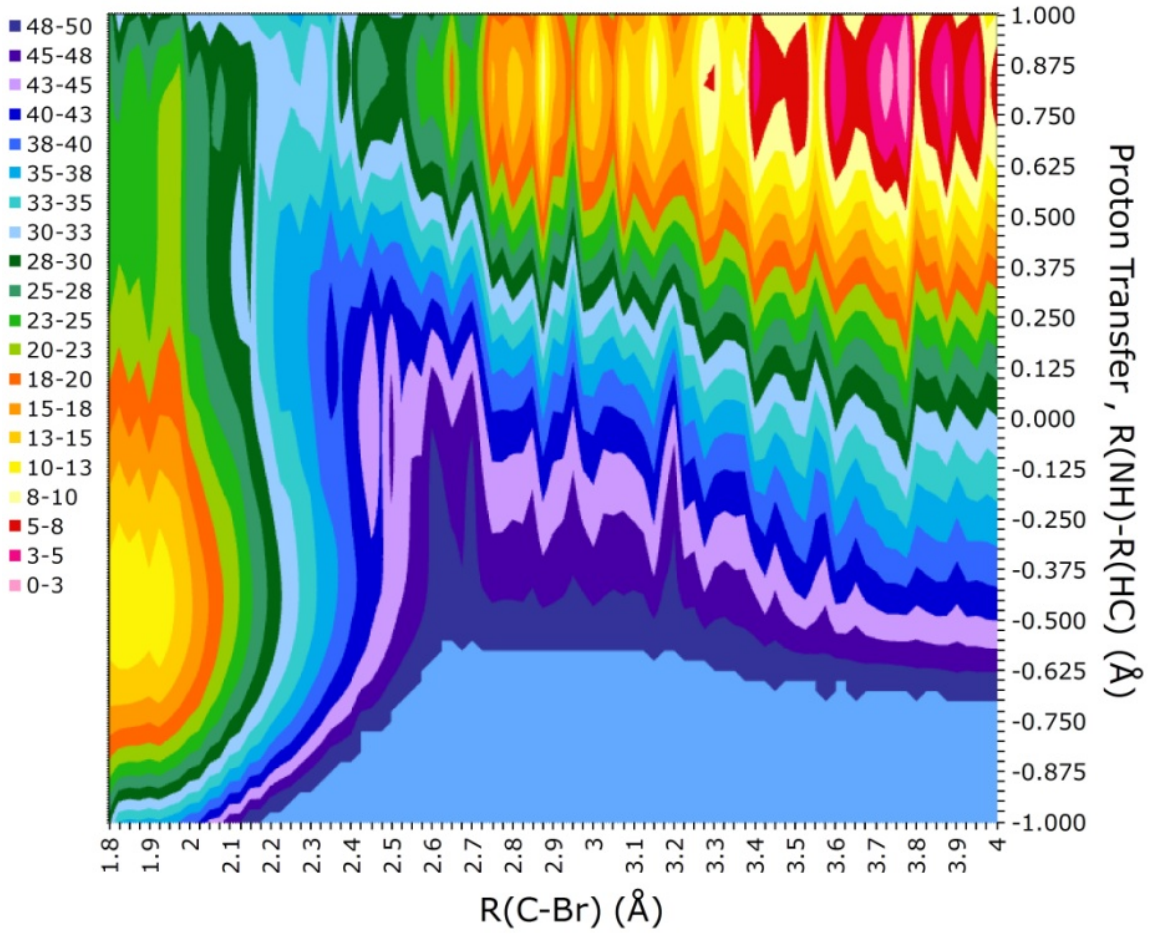


Figure 3.1: Free energy map (kcal/mol) computed for the β-elimination of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with piperidine in methanol from QM/MM/MC simulations. Energy values truncated after 50 kcal/mol for clarity.

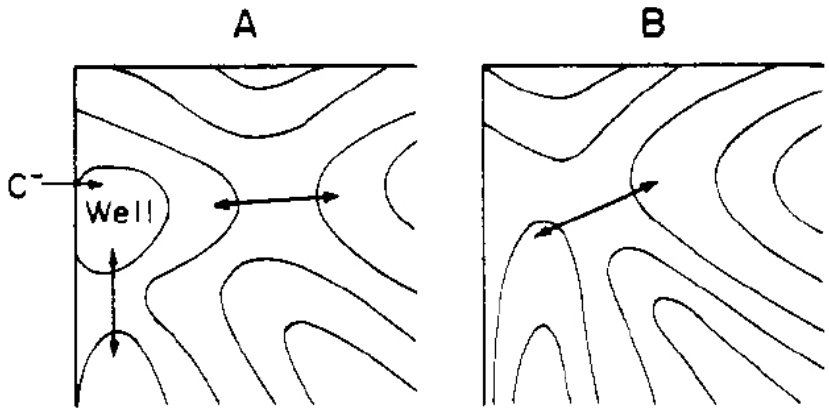


Figure 3.2: Speculative reactive coordinate contour diagram proposed by Gandler and Jencks[125] to illustrate the transition from an E1cb mechanism with a carbanion intermediate (A) to a concerted E2 mechanism when the carbanion no longer exists (B).

	$\Delta G^{\ddagger}$ (calc)	R(C-Br)
piperidine		
gas	27.7	2.31
methanol	19.9	2.16
water	20.2	2.16
pyrrolidine		
gas	28.7	2.35
methanol	19.0	2.15
water	19.2	2.15

Table 3.2: Free Energy of Activation, $\Delta G^{\ddagger}$ (kcal/mol), and transition structure geometries (Å) at 25 °C for the β-elimination of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane from M06-2X/6-31+G(d, p)/CPCM.

As another point of reference, DFT calculations were carried out using the M06-2X/6-31G+(d,p) method in vacuum and in solution (methanol and water) by using the CPCM continuum solvent method (Table 3.2). In the gas-phase, the DFT method predicted a traditional E2 mechanism, where the C-Br making/breaking bond distance is 2.31 – 2.35 Å and is moving in concert with the proton transfer when visualizing the 'imaginary frequency,'[36] e.g, –150.85 cm^{-1} for piperidine. In contrast, the single 'imaginary frequency' of –1166.51 or –1181.18 cm^{-1} for the reaction with piperidine in methanol and water, respectively, showed the proton transfer motion occurring with stationary, but extended C-Br distances of 2.16 and 2.15 Å (Figure 3.3 and Table 3.2) compared to the equilibrium distance of ca. 1.94 Å. The transition state C-Br distance is predicted to be earlier in solution than gas. The DFT calculations are in agreement with the QM/MM/MC simulations in methanol suggesting an E1cb-like mechanism, i.e., E2 with a significant amount of E1cb character, in accord with experiment[2] and with previous studies of borderline cases between E2 and E1cb mechanisms.[37]

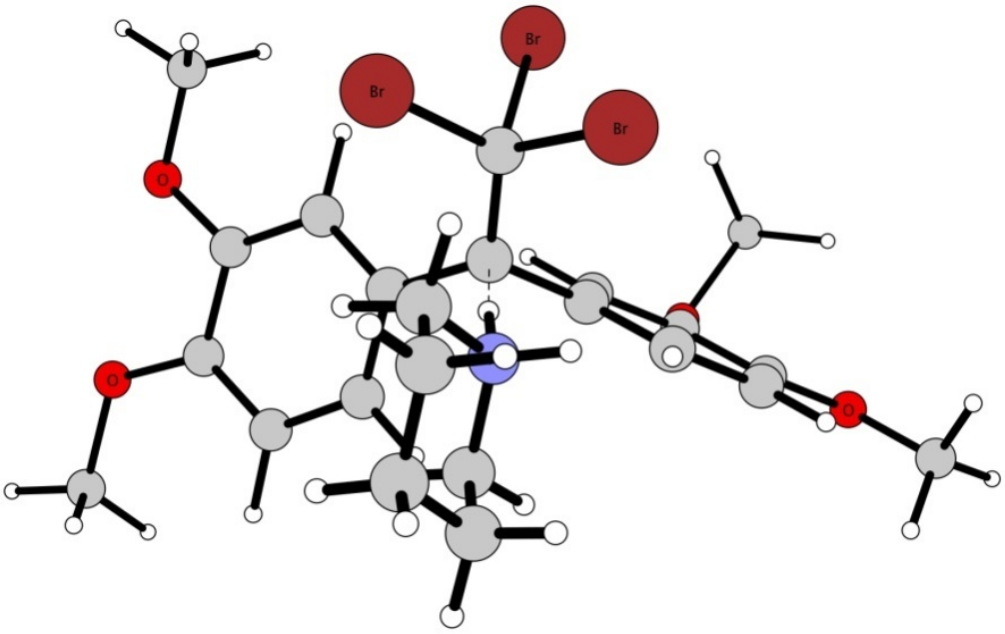


Figure 3.3: Illustration of the optimized transition structure for the β-elimination of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with piperidine in methanol from the M06-2X/6-31+G(d, p)/CPCM calculations.

3.4.3 Charges

The scaled CM3 charges computed for the β-elimination in each solvent can also be used to differentiate between mechanisms. For example, a greater concentration of anionic charge in the antiperiplanar reacting Br at the transition state would be more indicative of a traditional E2 reaction pathway than a distributed negative charge, spread among the Br and ethane carbon atoms, expected from an irreversible E1cb carbanion. In the QM/MM/MC calculations, the leaving Br in the β-elimination transition structures in [BMIM][BF_4] and [BMIM][PF_6] have computed partial charges of -0.94 e compared to -0.54 e in methanol. Figure 3.4 gives selected atomic charges for the transition state with piperidine in the ionic liquids and methanol; Figure 3.5 gives the charges for the

pyrrolidine system. In addition, the partial charges on the ethane carbon responsible for proton transfer are more positive in the ionic liquids than in methanol, i.e., -0.23 and -0.16 e in [BMIM][BF_4] and [BMIM][PF_6], respectively, compared to -0.45 e in CH_3OH. The more developed Br charge in the ionic liquids in concert with the extensive proton transfer at the transition state is more consistent with a pure E2 mechanism. The greater concentration of negative charge on the solute at the transition structure in methanol is more indicative of a reaction with considerable E1cb character. Consistent with the charges, the QM/MM/MC calculations for the dehydrobromination in the [BMIM][BF_4] and [BMIM][PF_6] ionic liquids predicted the C-Br bond cleavage at the transition structure to be noticeably longer, ca. 2.6 – 2.7 Å, compared to the distances of 2.3 – 2.4 Å in methanol for piperidine and pyrrolidine (Table 3.1). The charges on the piperidine N and reacting H are essentially the same in all solvents suggesting that solvation of the leaving group may be primarily responsible for the difference in mechanism.

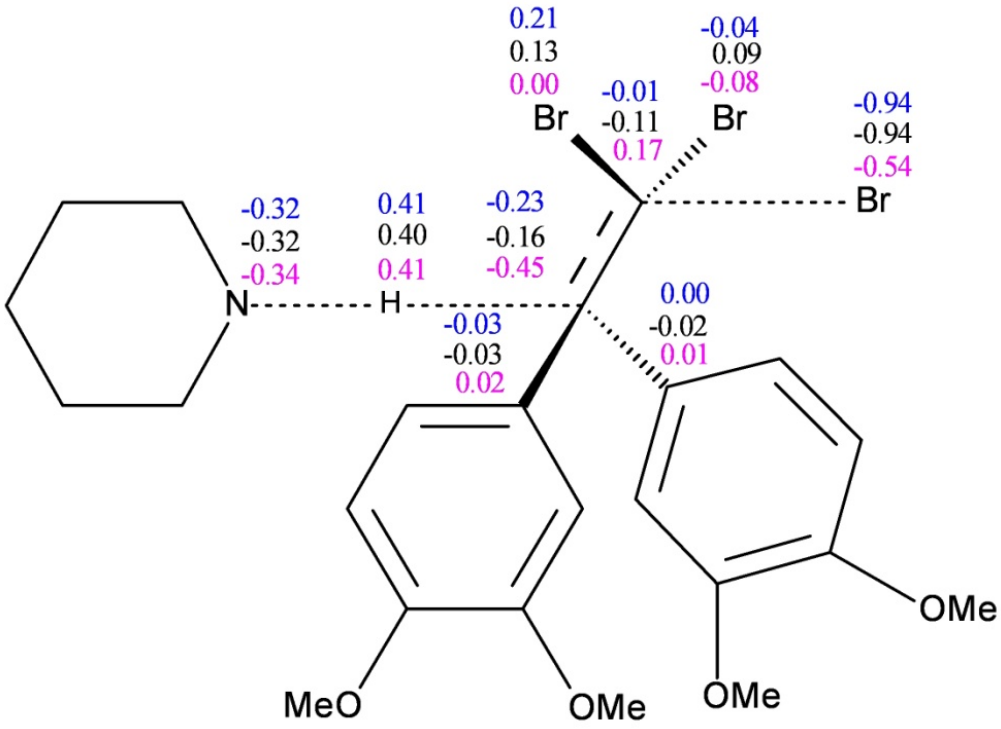


Figure 3.4: Selected atomic charges (e units) for the transition structure in [BMIM][BF_4] (blue), [BMIM][PF_6] (black), and methanol (pink) for the β-elimination of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with piperidine from the QM/MM/MC calculations.

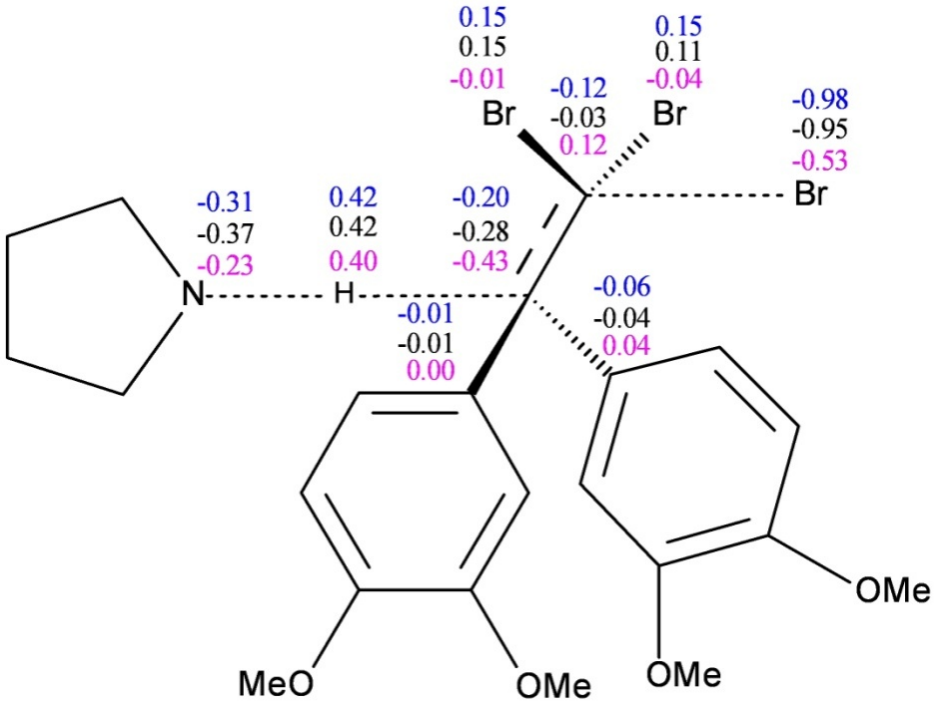


Figure 3.5: Selected CM3 atomic charges (e units) for the transition structure in [BMIM][BF_4] (blue), [BMIM][PF_6] (black), and methanol (pink) for the β-elimination of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with pyrrolidine from the QM/MM/MC calculations.

3.4.4 Solute-Solvent Interactions

To elucidate the differences between the ionic liquids and methanol, the interaction energies for the solvents were quantified by analyzing the solute-solvent energy pair distributions from QM/MM/MC calculations in the representative FEP windows near the reactants and transition state. The distributions record the average number of ions in the ionic liquids or molecules for methanol that interact with the reacting system and their corresponding energies. Highly favorable electrostatic interactions between solute and solvent components are reflected in the left-most region

with energies more attractive than ca. –5 kcal/mol (Figure 3.6 and Figure 3.7). The large band near 0 kcal/mol arises from the many ions in outer shells.

The β-elimination reactions with piperidine and pyrrolidine have weaker energy distributions for the reactants when compared to the transition state in both ionic liquids and in methanol (Figures 3.6 and Figure 3.7). Integration of the distributions from –10.0 to –5.0 (or -3.5) kcal/mol confirms the more favorable interactions for the transition states (Table 3.3). Specifically, the number of solute-solvent interactions increases by 1 – 3 ions in going from the reactants to transition state in the ionic liquids. In addition, there is a shift in the average strength of the most favorable interactions to lower energy particularly for pyrrolidine, which could explain the enhanced rate of reaction reported versus piperidine.[1]

The exact nature of these most favorable solute-ion interactions is of obviously relevant interest. In both reactions, a shift in the average strength of the most favorable interactions to a lower energy in the transition state is consistent with the stabilization of the emerging charge at the reacting Br. Figure 3.8 shows a snapshot of the piperidine-based β-elimination transition structure in [BMIM][BF_4] with nearby ions retained from the QM/MM/MC simulations. The emerging Br anion is stabilized by two BMIM cations forming hydrogen bonds with the more sterically exposed hydrogens on carbons at the 4 and 5 positions and the side chain hydrogens, rather than the most acidic imidazolium proton at the 2 position (pK_a of ca. 21 – 23).[38,39] The BMIM cations are forming a cage-like structure to favorably interact with the Br anion and β-phenyl substituents (Figure 3.9), which agrees with experimental reports of liquid clathrate formation in 1-alkyl-3-methylimdazolium-based ionic liquids with aromatic compounds.[40] In addition, the

proton transfer is facilitated by favorable electrostatic interactions of the emerging positive charge on the base with a BF_4 anion (Figure 3.8). While the averaged polarity of methanol and [BMIM][BF_4] are similar,[41] the experimentally measured dipolarity/polarizability (π^*) values for [BMIM][BF_4] and [BMIM][PF_6] of 1.047 and 1.032, respectively, are significantly higher than that of 0.73 for methanol.[42] This is consistent with more the favorable specific interactions towards the E2 mechanism by the ionic liquids as compared to methanol.

The ground state, represented by a reactant complex between the base and the tribromo-ethane, has computed interactions between the ions and the reactants at greater distances than at the transition structure. For example, in the piperidine reaction in [BMIM][BF_4] the cation interacts with the emerging Br anion with multiple interacting distances as short as 2.7 Å at the transition state, whereas the closest interaction at the ground state is 3.0 Å and it occurs with the less acidic hydrogen atoms bonded to the methyl side-chain. The results are similar for the [PF_6]-based reactions. Detailed hydrogen bonding distances between the ionic liquids ions and the piperidine- and pyrrolidine-based β-eliminations at the transition and ground states are given in the Supporting Information. The structural configuration of the methanol molecules with the transition and ground states are also given Supporting Information and support the computed reduced solute-solvent interaction energies as compared to the ionic liquids.

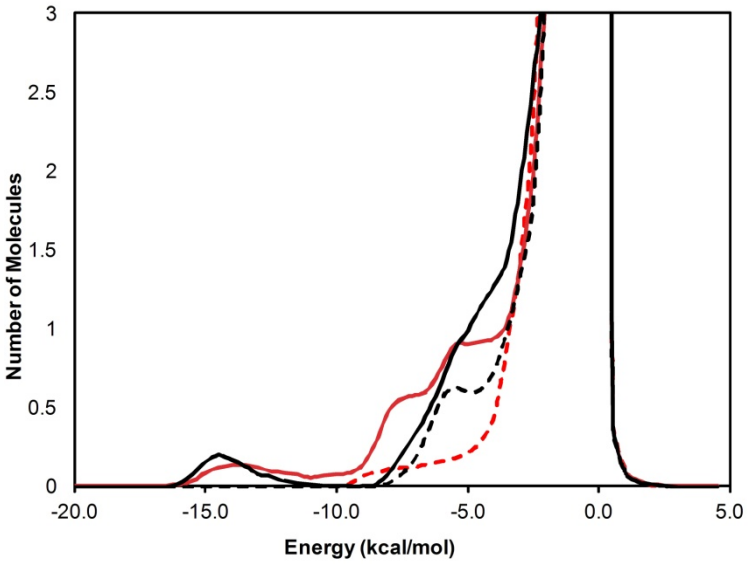


Figure 3.6: Solute-solvent energy pair distributions for β-elimination of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with piperidine (in red) and pyrrolidine (in black) for the reactants (dashed line) and transition state (solid line) in methanol at 25 °C. The ordinate records the number of solvent molecules that interact with the solutes and their interaction energy on the abscissa. Units for ordinate are number of molecules per kcal/mol.

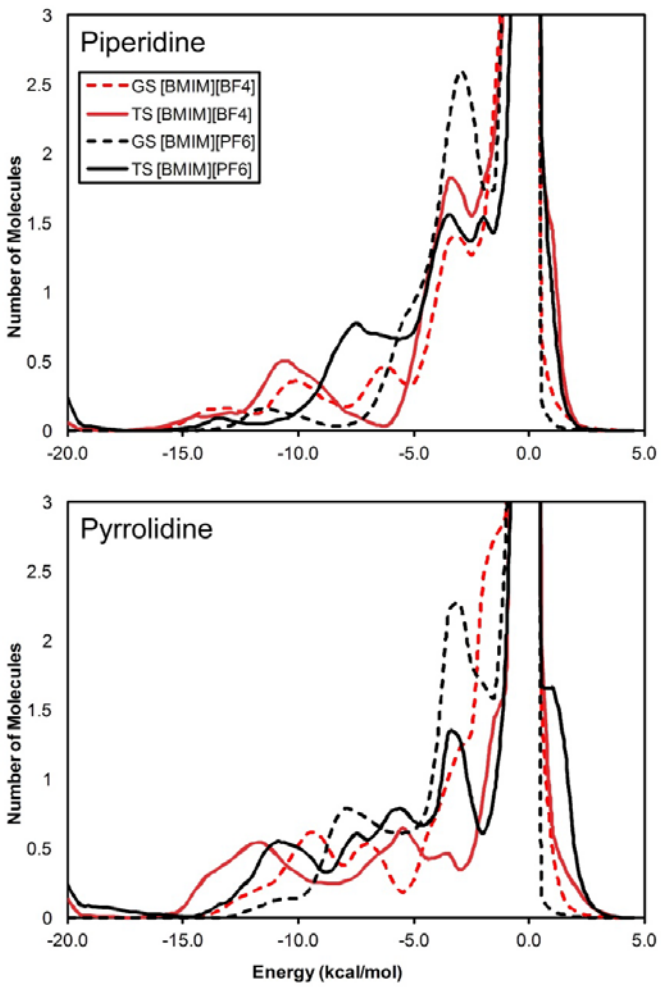


Figure 3.7: Solute-solvent energy pair distributions for β-elimination of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with piperidine (top) and pyrrolidine (bottom) for the reactants (dashed line) and transition state (solid line) in [BMIM][BF_4] and [BMIM][PF_6] at 25 °C. The ordinate records 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.

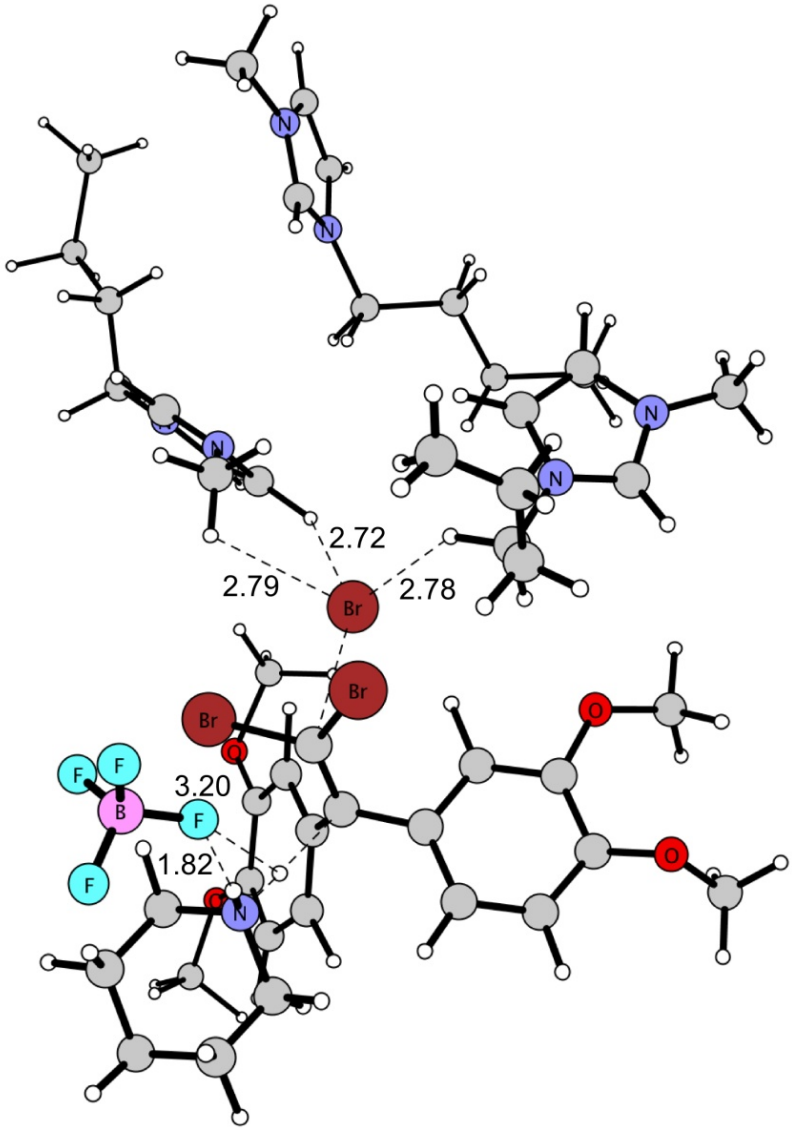


Figure 3.8: Typical snapshot of a transition state for the β-elimination with piperidine in [BMIM][BF_4]. The distances (in Å) are average values over the final 10 million configurations of QM/MM/MC simulations.

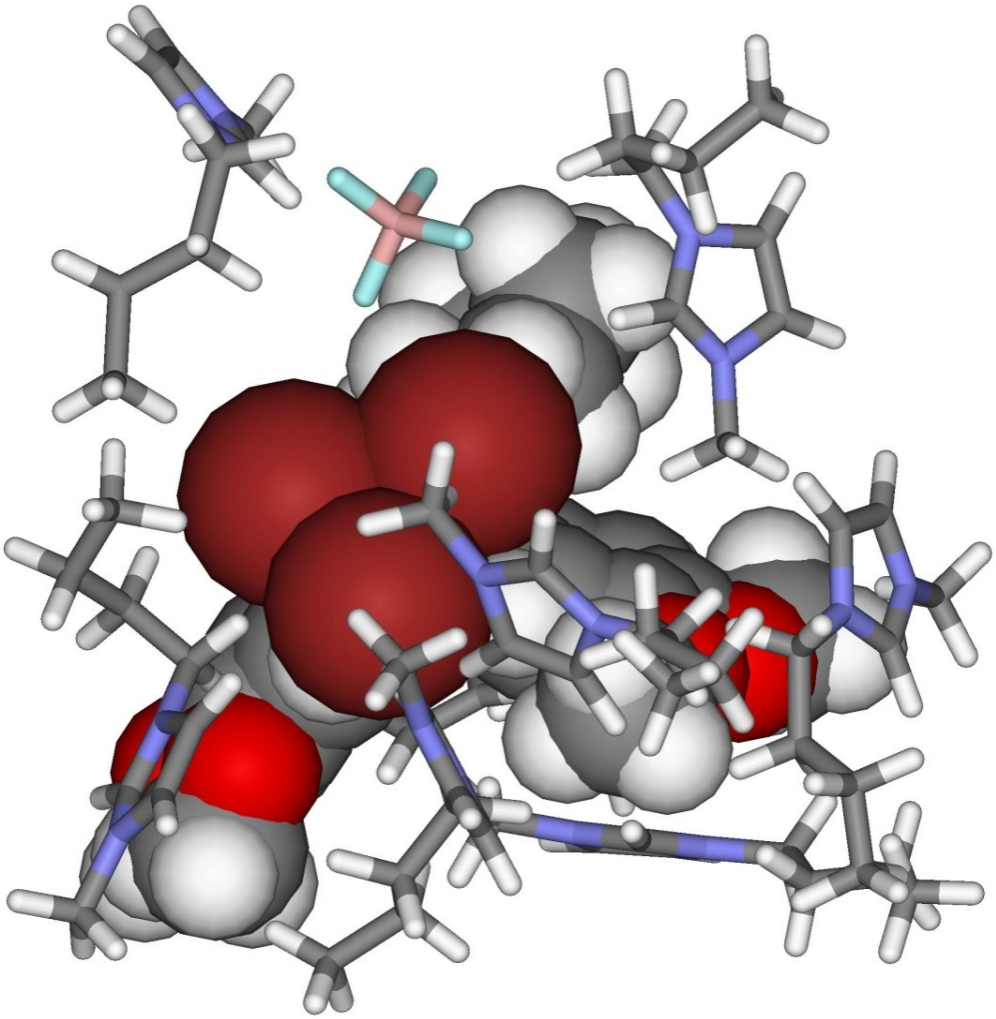


Figure 3.9: Illustration of the encapsulation of the β-elimination solute with piperidine transition state (given as a CPK space-filling model) by nearby ions from [BMIM][BF_4] (shown as sticks).

	Piperidine GS	Piperidine TS	Pyrrolidine GS	Pyrrolidine TS
[BMIM][BF_4]	5.1 (8.0)	4.9 (9.2)	6.1 (8.5)	8.2 (9.6)
[BMIM][PF_6]	3.7 (8.6)	7.1 (11.1)	6.7 (10.9)	9.3 (12.1)
methanol	1.3 (2.9)	6.4 (9.2)	2.5 (4.8)	4.7 (8.6)

[a]From Figures 5 and S2.

Table 3.3: Solute-Solvent Energy Pair Distributions for the β-elimination of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane for the Reactant (GS) and Transition Structure (TS) in [BMIM][BF_4], [BMIM][PF_6], and methanol integrated to -5.0 kcal/mol (and -3.5 kcal/mol in parenthesis).[a]

3.4.5 Aromatic Ring Orientations

A snapshot of the transition state for the elimination reaction using pyrrolidine in [BMIM][PF_6] is given in Figure 3.10 and gives a good representation of the geometry orientation for the reactions in both ionic liquids (Table 3.4). The ionic liquid structures have the aromatic rings generally coplanar at the transition state. For example, torsion angles of -146.4 and 134.5 ° between the two carbons on the aromatic ring and the two carbons on the ethane, defined as Φ_1 and Φ_2 = C1-C2-C3-C4 in Figure 8, for piperidine in [BMIM][BF_4] is a dramatic contrast from the values of -162.6 and 87.1 ° predicted from DFT (Figure 3.3) or -154.4 and 113.5 ° from QM/MM/MC for the same reaction in methanol.

The coplanar orientation of the phenyl rings at the transition state should maximize the electronic effects exerted on the reaction route. Favorable π-π interactions with the ionic liquid cation [BMIM] forces a coplanarity between aromatic rings (Figure

3.9 and Figure 3.11), which is consistent with previous hypotheses.[40,43] It should be noted that the OPLS-AA force field has been reported to yield excellent agreement with experiment for computed benzene dimer interaction energies and geometries in the gas and condensed phase.[44] More recently, Fu and Tian carried out molecular dynamics (MD) simulations for liquid benzene with eight potentials consisting of Lennard-Jones and Coulomb terms and recommended the OPLS-AA as the best model based agreement with high-resolution neutron diffraction data.[45] In addition, Takeuchi also reported that the OPLS-AA force field was more reliable in reproducing the structures of benzene clusters, consisting of up to 30 rings, than MP2 calculations.[46]

Monitoring the average Φ_1 and Φ_2 torsions over the final 10 million configurations of the QM/MM/MC calculations suggests that the reaction spends minimally 70 % of the simulation in the coplanar configuration and 30 % in a t-shaped configuration. In contrast, the transition structures in methanol adopted an approximate t-shaped conformation for nearly 100% of the QM/MM/MC simulations. The aromatic rings of the reaction in water were also predicted to favor a t-shaped orientation from DFT simulations (Table 3.4).

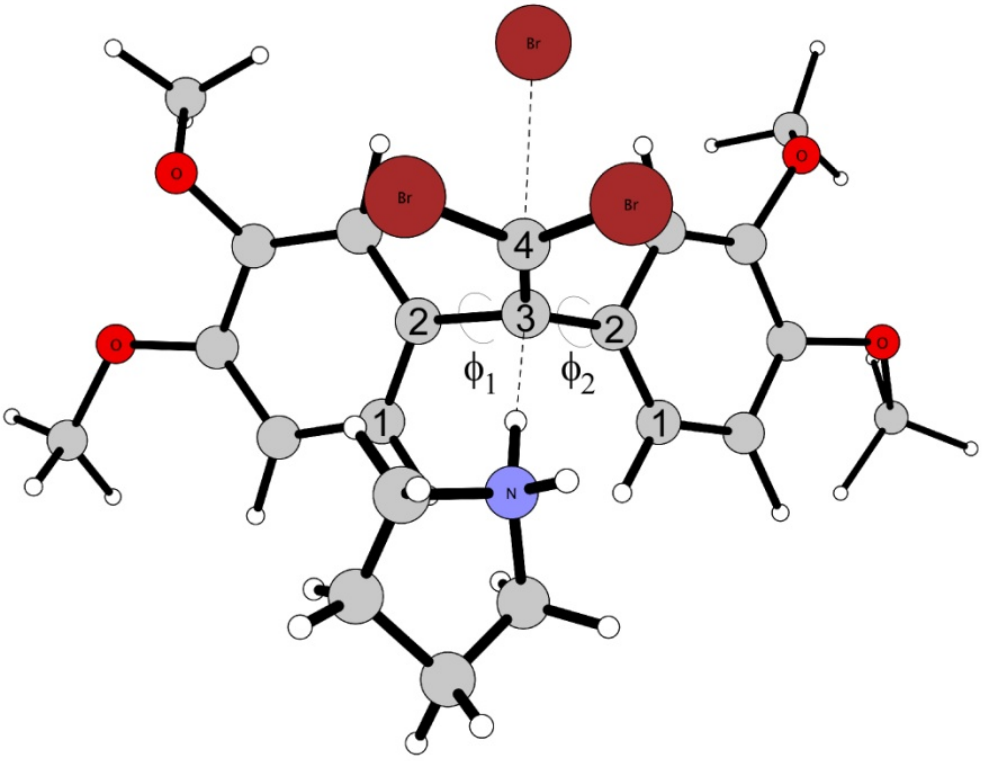


Figure 3.10: Snapshot of a transition state for the β-elimination of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with pyrrolidine in [BMIM][PF_6] from the QM/MM/MC calculations. Φ = C1-C2-C3-C4.

	QM/MM [b]	DFT [c]
piperidine		
[BMIM][BF_4]	-146.4/134.5	-
[BMIM][PF_6]	-105.9/113.2	-
Methanol	-154.4/113.5	-162.6/87.1
Water	-	-162.6/87.1
Gas	-	-154.1/83.7
pyrrolidine		
[BMIM][BF_4]	-106.4/132.7	-
[BMIM][PF_6]	-135.3/129.8	-
Methanol	-173.9/126.2	-157.8/89.1
Water	-	-157.8/89.2

[a] See Figure 3.10 for definitions of Φ_1 and Φ_2. [b] PDDG/PM3 and MC/FEP. Angles averaged over final 10 million configurations. [c] M06-2X/6-31+G(d, p) optimization.

Table 3.4: Dihedral angles Φ_1/Φ_2 (degrees) for the transition structure at 25 °C for the β-elimination of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane.[a]

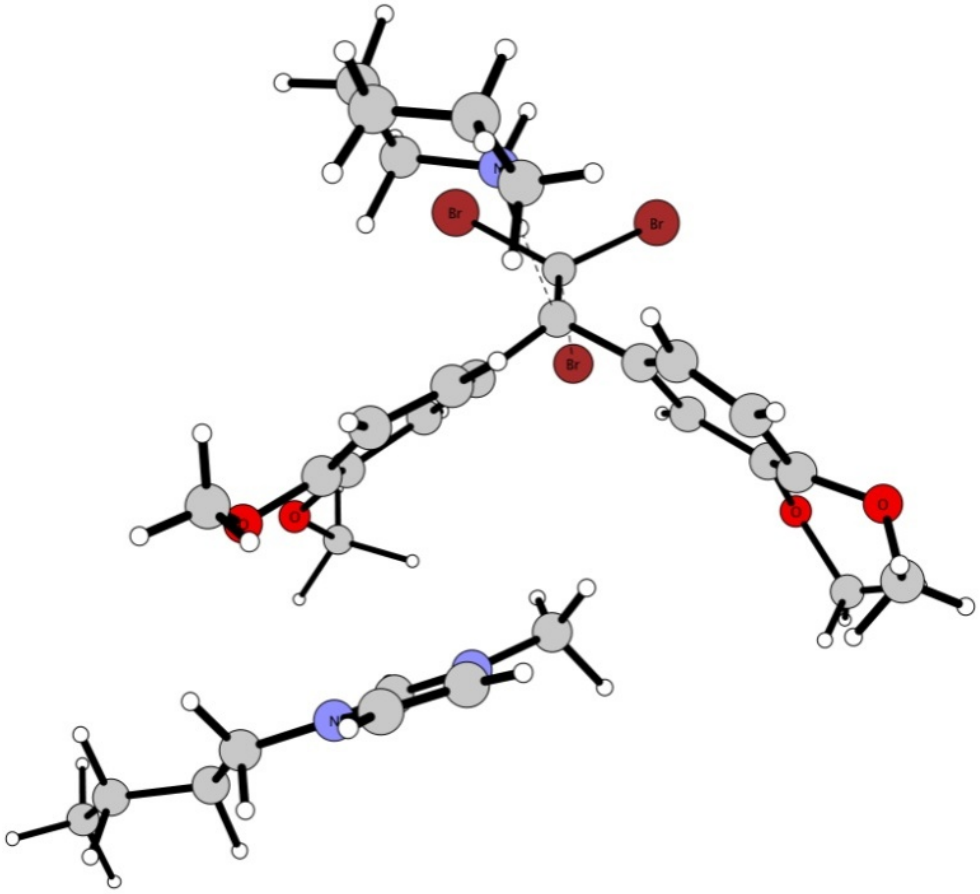


Figure 3.11: Typical snapshot of a transition state for the β-elimination with piperidine in [BMIM][BF_4] from the QM/MM/MC calculations. A single BMIM cation is illustrated to highlight the π-π interaction with the solute.

3.5 Conclusions

QM/MM calculations have been carried out to determine the origin of the ionic liquid effect on a reported mechanism change for the β-elimination between 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane and the cyclic amines piperidine and pyrrolidine. D'Anna et al. proposed the reaction to occur via an irreversible E1cb route in methanol,[2] but as an E2 mechanism in the [BMIM][BF_4] and [BMIM][PF_6] ionic

liquids.[1] Our computed free energy surfaces agree in principle with their hypothesis, with the exception that in methanol the reaction route followed an E1cb-like mechanism, i.e., E2 with a significant amount of E1cb character, as no carbanion intermediate was located. Our results are consistent with previous experimental studies of borderline cases between E2 and E1cb mechanisms.[37] The E1cb-like mechanism in methanol is further verified with additional calculations using an alternative M06-2X/CPCM method. In the case of the two ionic liquids studied, [BMIM][BF_4] and [BMIM][PF_6], the simulations reproduced the full E2 mechanism as suggested. The structural configuration of the ions play a large role, as the observed mechanistic change has been computed as a combination of favorable electrostatic interactions with the leaving Br anion and π-π interactions between the [BMIM] cation and β-phenyl substituents on the tribromo-ethane molecule. Specifically, the number of solute-solvent interactions are computed to increase by 1 – 3 ions in going from the reactants to transition state in the ionic liquids. In addition, there is a shift in the average strength of the most favorable interactions to lower energy, particularly for pyrrolidine, which could explain the enhanced rate of reaction reported versus piperidine.[1] In addition, the ionic liquids form a liquid clathrate structure that enforce a coplanar orientation of the β-phenyl rings at the transition state maximizing the electronic effects exerted on the reaction route.[40,43] Monitoring the average torsions over the final 10 million MC configurations of the QM/MM calculations found the phenyl rings to spend minimally 70 % of the simulation in the coplanar configuration and 30 % in a t-shaped configuration. In contrast, the transition structures in methanol adopted an approximate t-shaped conformation for nearly 100% of the simulation for both the QM/MM and M06-2X /CPCM methods. Deeper insight into the effect of ionic

liquids upon important organic reaction types should allow researchers to exploit this understanding to predict optimal conditions for additional reactions in similar classes.

Acknowledgement. Gratitude is expressed to the National Science Foundation (CHE-1149604) and the Alabama Supercomputer Center for support of this research.

Complete S**upporting Information Available:** Additional figures for charges and transition states in solution; solute-solvent energy pair distributions; hydrogen bond distances between ions/methanol and solute at the transition and ground states; DFT energies, frequencies, and coordinates of structures; and complete Ref. 31. This material is available free of charge via the Internet at http://pubs.acs.org.

3.6 Abbreviated Supplemental Information

page number

[BMIM][BF_4] Reactant Complex Figures

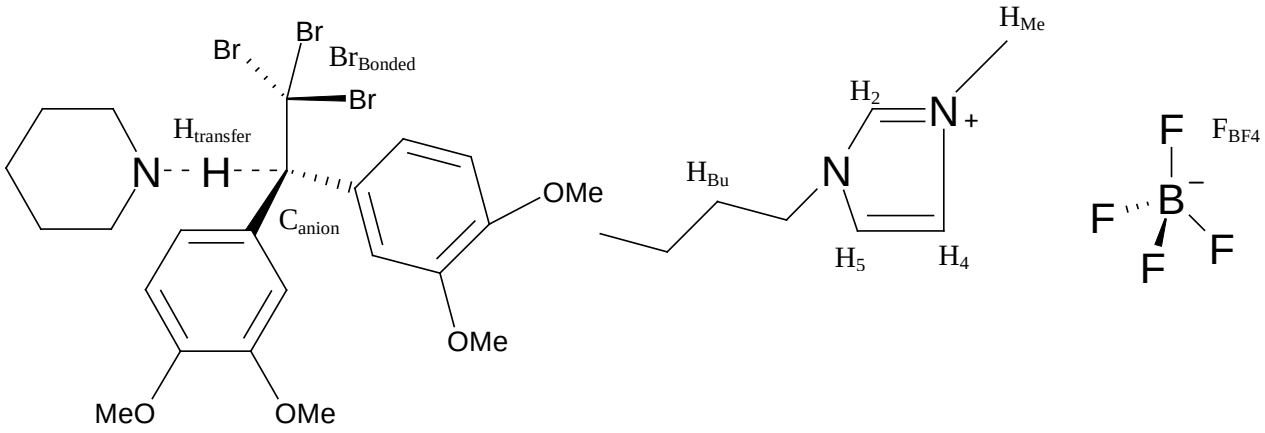


Figure 3S.1 – The reactant complex of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with piperidine in the ionic liquid [BMIM][BF_4]. For solvent effect data see Table 3S.1.

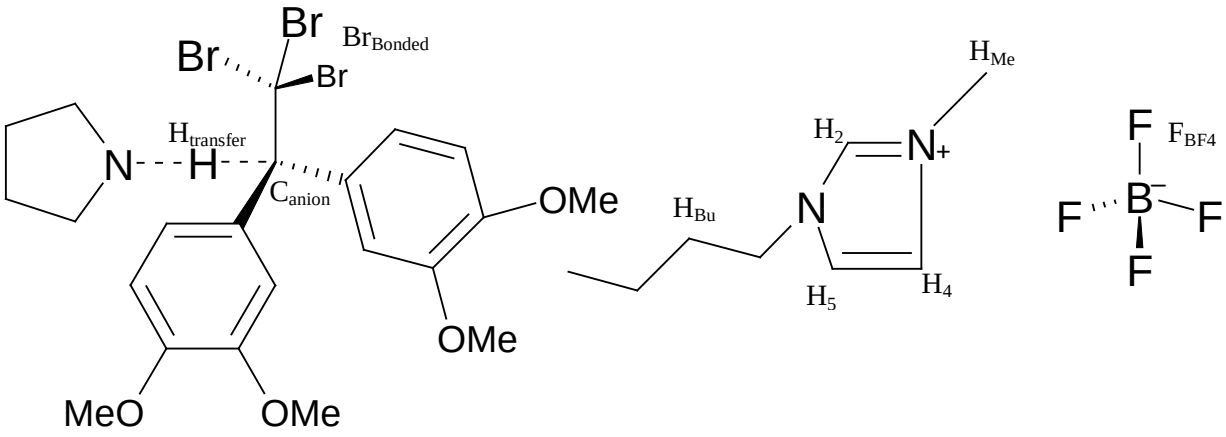


Figure 3S.2 – The reactant complex of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with pyrrolidine in the ionic liquid [BMIM][BF_4]. For solvent effect data see Table 3S.2.

[BMIM][BF_4] Transition Structure Figures

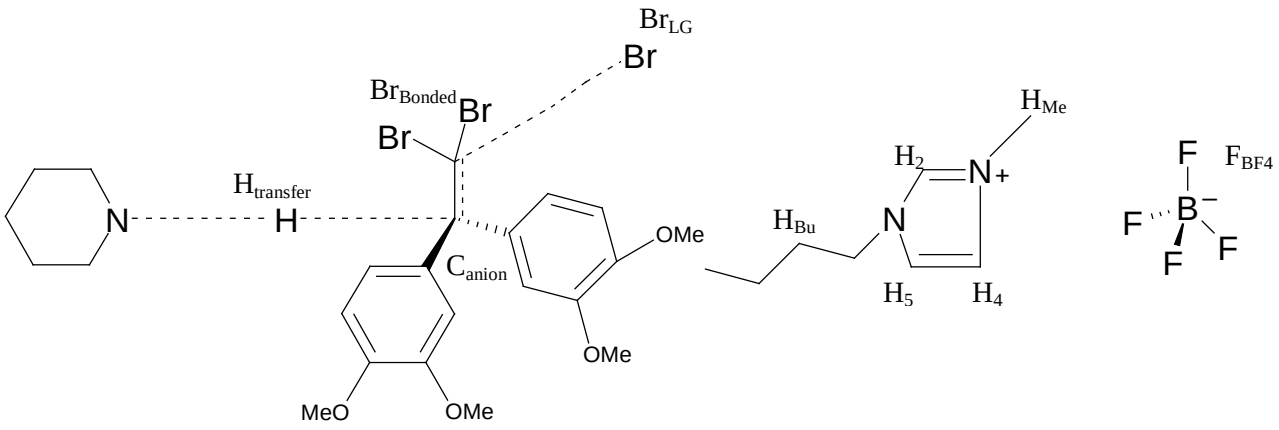


Figure 3S.3 – The transition structure of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with piperidine in the ionic liquid [BMIM][BF_4]. For solvent effect data see Table 3S.3.

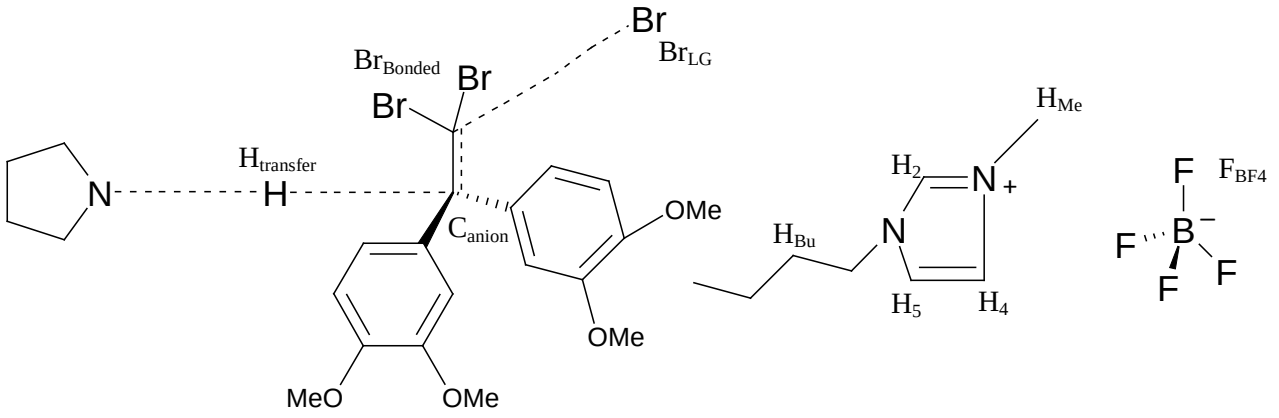


Figure 3S.4 – The transition structure of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with pyrrolidine in the ionic liquid [BMIM][BF_4]. For solvent effect data see Table 3S.4.

Table 3S.1: Piperidine reactant-complex in [BMIM][BF_4] – solvent effect data

Interactions < 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
213	3.75	H_{Me} with C_{anion}
102	3.39	H_4 with Br_{bonded}
102	3.13	H_{Me} with Br_{bonded}
102	3.21	H_{Me} with Br_{bonded}
102	3.47	H_{Me} with Br_{bonded}
102	3.62	H_{Me} with Br_{bonded}
213	2.99	H_{Me} with Br_{bonded}
213	3.55	H_{Me} with Br_{bonded}
213	3.55	H_{Me} with Br_{bonded}
162	3.20	H_{Bu} with Br_{bonded}
162	3.35	H_{Bu} with Br_{bonded}
162	3.54	H_{Bu} with Br_{bonded}
210	3.67	H_{Bu} with Br_{bonded}
213	3.68	H_{Me} with Br_{bonded}

All Interactions < 5 Angstroms

BF_4 Anion	Average Distance (Å)	Interaction
164	3.31	F_{BF4} with $H_{transfer}$

Interactions > 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
102	4.99	H_4 with Br_{bonded}
213	4.76	H_4 with Br_{bonded}
102	5.07	H_4 with Br_{bonded}
210	4.05	H_2 with Br_{bonded}
213	4.24	H_2 with Br_{bonded}
102	4.25	H_{Me} with Br_{bonded}
102	4.55	H_{Me} with Br_{bonded}
102	4.81	H_{Me} with Br_{bonded}
213	4.63	H_{Me} with Br_{bonded}
213	4.69	H_{Me} with Br_{bonded}
213	4.79	H_{Me} with Br_{bonded}
15	3.90	H_{Bu} with Br_{bonded}
15	4.01	H_{Bu} with Br_{bonded}
15	4.79	H_{Bu} with Br_{bonded}
162	4.53	H_{Bu} with Br_{bonded}
162	4.7	H_{Bu} with Br_{bonded}
184	4.62	H_{Bu} with Br_{bonded}
205	4.07	H_{Bu} with Br_{bonded}
205	5.04	H_{Bu} with Br_{bonded}
210	4.07	H_{Bu} with Br_{bonded}
210	4.03	H_{Bu} with Br_{bonded}
210	4.44	H_{Bu} with Br_{bonded}
210	4.59	H_{Bu} with Br_{bonded}
162	4.97	H_{Bu} with C_{anion}
98	4.34	H_{Bu} with C_{anion}
98	4.55	H_{Bu} with C_{anion}
98	4.78	H_{Bu} with C_{anion}
162	4.48	H_{Bu} with C_{anion}
213	4.83	H_{Me} with C_{anion}

Table 3S.2: Pyrrolidine reactant-complex in [BMIM][BF_4] – solvent effect data

Interactions < 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
170	3.72	H_4 with Br_{bonded}
164	3.75	H_4 with Br_{bonded}
170	2.91	H_5 with Br_{bonded}
170	2.91	H_5 with Br_{bonded}
164	3.57	H_5 with Br_{bonded}
101	3.19	H_{Me} with Br_{bonded}
101	3.19	H_{Me} with Br_{bonded}
101	3.29	H_{Me} with Br_{bonded}
101	3.52	H_{Me} with Br_{bonded}
101	3.52	H_{Me} with Br_{bonded}
101	3.61	H_{Me} with Br_{bonded}
170	2.95	H_{Bu} with Br_{bonded}
170	2.95	H_{Bu} with Br_{bonded}
170	3.19	H_{Bu} with Br_{bonded}
170	3.19	H_{Bu} with Br_{bonded}
170	3.51	H_{Bu} with Br_{bonded}
170	3.51	H_{Bu} with Br_{bonded}
164	3.03	H_{Bu} with Br_{bonded}
164	3.62	H_{Bu} with Br_{bonded}

Interactions > 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
101	4.03	H_4 with Br_{bonded}
170	4.31	H_4 with Br_{bonded}
170	4.61	H_4 with Br_{bonded}
164	4.99	H_4 with Br_{bonded}
170	4.27	H_5 with Br_{bonded}
164	4.91	H_2 with Br_{bonded}
170	4.90	H_5 with C_{anion}
101	4.43	H_{Me} with Br_{bonded}
101	4.43	H_{Me} with Br_{bonded}
101	4.72	H_{Me} with Br_{bonded}
170	4.55	H_{Bu} with Br_{bonded}
170	4.56	H_{Bu} with Br_{bonded}
170	4.63	H_{Bu} with Br_{bonded}
170	4.64	H_{Bu} with Br_{bonded}
170	4.71	H_{Bu} with Br_{bonded}
170	4.72	H_{Bu} with Br_{bonded}
170	4.74	H_{Bu} with Br_{bonded}
170	4.74	H_{Bu} with Br_{bonded}
170	4.99	H_{Bu} with Br_{bonded}
170	4.99	H_{Bu} with Br_{bonded}
187	4.82	H_{Me} with C_{anion}
164	4.94	H_{Bu} with C_{anion}
101	4.99	H_{Me} with C_{anion}

All Interactions < 5 Angstroms

BF_4 Anion	Average Distance (Å)	Interaction
227	3.71	F_{BF4} with $H_{transfer}$
227	4.26	F_{BF4} with $H_{transfer}$
227	4.49	F_{BF4} with $H_{transfer}$

Table 3S.3: Piperidine TS in [BMIM][BF_4] – solvent effect data

Interactions < 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
210	3.28	H_2 with Br_{bonded}
102	2.72	H_4 with Br_{LG}
102	2.79	H_{Me} with Br_{LG}
102	3.16	H_{Me} with Br_{LG}
213	2.78	H_{Bu} with Br_{LG}
213	3.12	H_{Bu} with Br_{LG}
213	3.82	H_{Bu} with Br_{LG}
213	3.90	H_{Bu} with Br_{LG}
210	3.38	H_{Me} with Br_{bonded}
210	3.79	H_{Me} with Br_{bonded}
102	3.38	H_{Me} with Br_{bonded}
102	3.57	H_{Me} with Br_{bonded}
213	3.05	H_{Bu} with Br_{bonded}
15	3.48	H_{Bu} with Br_{bonded}
15	3.49	H_{Bu} with Br_{bonded}
15	3.86	H_{Bu} with Br_{bonded}
15	3.87	H_{Bu} with Br_{bonded}
213	3.91	H_{Bu} with Br_{bonded}

Interactions > 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
102	4.99	H_4 with Br_{bonded}
213	4.78	H_5 with Br_{LG}
102	4.33	H_{Me} with Br_{LG}
176	4.46	H_{Me} with Br_{LG}
213	4.30	H_{Bu} with Br_{LG}
213	4.32	H_{Bu} with Br_{LG}
213	4.74	H_{Bu} with Br_{LG}
213	4.91	H_{Bu} with Br_{LG}
213	4.99	H_{Bu} with Br_{LG}
15	4.45	H_{Bu} with Br_{LG}
15	4.74	H_{Bu} with Br_{LG}
210	4.83	H_{Me} with Br_{bonded}
102	4.30	H_{Me} with Br_{bonded}
102	5.14	H_{Me} with Br_{bonded}
15	4.11	H_{Bu} with Br_{bonded}
15	4.72	H_{Bu} with Br_{bonded}
184	4.63	H_{Bu} with Br_{bonded}
184	4.65	H_{Bu} with Br_{bonded}
210	4.89	H_{Bu} with Br_{bonded}
213	4.03	H_{Bu} with Br_{bonded}
213	4.62	H_{Bu} with Br_{bonded}
210	4.29	H_{Me} with C_{anion}
176	4.37	H_{Me} with C_{anion}
176	4.51	H_{Me} with C_{anion}
102	4.33	H_{Me} with C_{anion}
184	4.71	H_{Bu} with C_{anion}

All Interactions < 5 Angstroms

BF_4 Anion	Average Distance (Å)	Interaction
61	3.20	F_{BF4} with $H_{transfer}$
61	4.43	F_{BF4} with $H_{transfer}$
61	4.77	F_{BF4} with $H_{transfer}$

Table 3S.4: Pyrrolidine TS in [BMIM][BF_4] – solvent effect data

Interactions < 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
170	3.60	H_5 with Br_{LG}
164	2.58	H_4 with Br_{bonded}
164	3.94	H_4 with Br_{bonded}
164	3.73	H_4 with Br_{bonded}
102	2.98	H_4 with Br_{bonded}
102	3.29	H_4 with Br_{bonded}
170	2.81	H_{Bu} with Br_{LG}
170	3.04	H_{Bu} with Br_{LG}
170	3.22	H_{Bu} with Br_{LG}
170	3.48	H_{Bu} with Br_{bonded}
170	3.74	H_{Bu} with Br_{bonded}
117	3.33	H_{Bu} with Br_{bonded}
117	3.55	H_{Bu} with Br_{bonded}
117	3.84	H_{Bu} with Br_{bonded}
164	3.53	H_{Me} with Br_{bonded}
164	3.68	H_{Me} with Br_{bonded}
102	3.17	H_{Me} with Br_{bonded}
102	3.72	H_{Me} with Br_{bonded}

All Interactions < 5 Angstroms

BF_4 Anion	Average Distance (Å)	Interaction
166	3.42	F_{BF4} with $H_{transfer}$
166	4.14	F_{BF4} with $H_{transfer}$
166	4.71	F_{BF4} with $H_{transfer}$
122	4.96	F_{BF4} with $H_{transfer}$

Interactions > 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
170	4.58	H_2 with Br_{LG}
170	4.89	H_4 with Br_{LG}
102	4.46	H_4 with Br_{LG}
164	4.5	H_5 with Br_{LG}
102	4.31	H_2 with Br_{bonded}
102	4.59	H_5 with Br_{bonded}
164	4.95	H_4 with Br_{bonded}
170	4.47	H_{Bu} with Br_{LG}
170	4.71	H_{Bu} with Br_{LG}
170	4.97	H_{Bu} with Br_{LG}
170	5.03	H_{Bu} with Br_{LG}
164	4.12	H_{Me} with Br_{LG}
164	4.31	H_{Me} with Br_{LG}
102	4.37	H_{Me} with Br_{LG}
170	4.28	H_{Bu} with Br_{bonded}
170	5.03	H_{Bu} with Br_{bonded}
117	4.46	H_{Bu} with Br_{bonded}
117	4.79	H_{Bu} with Br_{bonded}
117	4.79	H_{Bu} with Br_{bonded}
117	4.98	H_{Bu} with Br_{bonded}
208	4.50	H_{Bu} with Br_{bonded}
102	4.02	H_{Me} with Br_{bonded}
102	4.56	H_{Me} with Br_{bonded}
102	4.73	H_{Me} with Br_{bonded}
102	4.99	H_{Me} with Br_{bonded}
102	5.09	H_{Me} with Br_{bonded}
164	4.37	H_{Me} with C_{anion}
170	4.79	H_{Bu} with C_{anion}
117	4.57	H_{Bu} with C_{anion}
117	4.81	H_{Bu} with C_{anion}

[BMIM][PF_6] Reactant Complex Figures

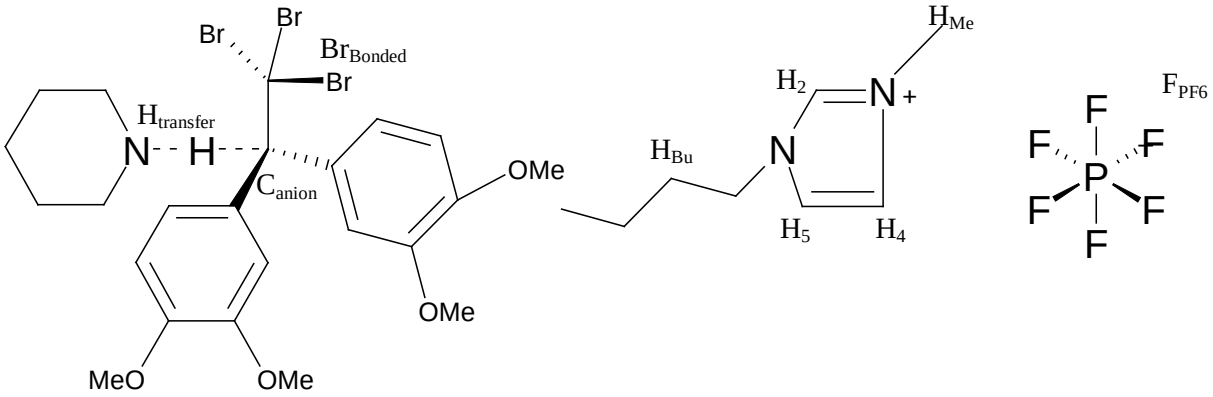


Figure 3S.5 – The reactant complex of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with piperidine in the ionic liquid [BMIM][PF_6]. For solvent effect data see Table 3S.5.

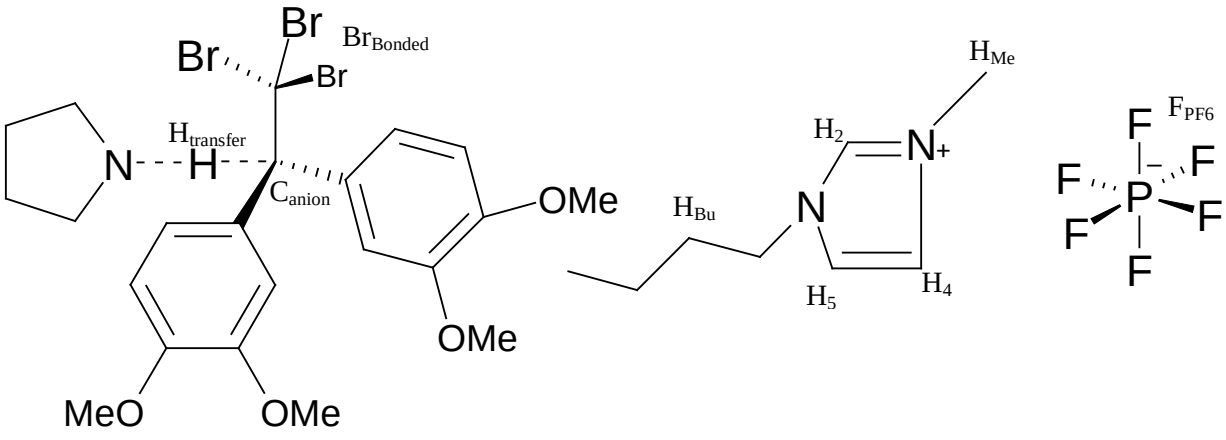


Figure 3S.6 – The reactant complex of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with pyrrolidine in the ionic liquid [BMIM][PF_6]. For solvent effect data see Table 3S.6.

[BMIM][PF_6] Transition Structure Figures

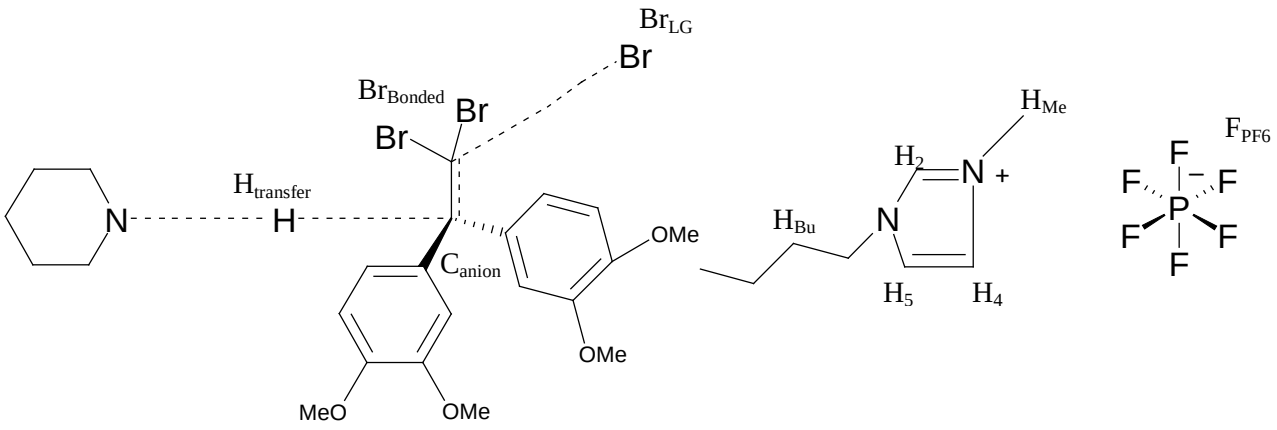


Figure 3S.7 – The transition structure of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with piperidine in the ionic liquid [BMIM][PF_6]. For solvent effect data see Table 3S.7.

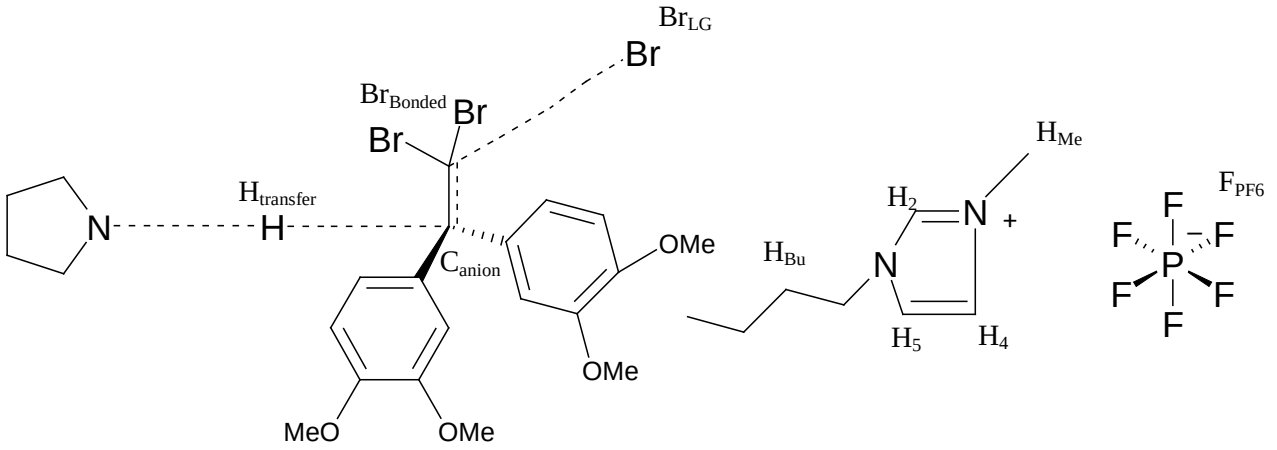


Figure 3S.8 – The transition structure of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with pyrrolidine in the ionic liquid [BMIM][PF_6]. For solvent effect data see Table 3S.8.

Table 3S.5: Piperidine reactant-complex in [BMIM][PF_6] – solvent effect data

Interactions < 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
218	3.35	H_{Me} with Br_{bonded}
218	3.80	H_{Me} with Br_{bonded}
218	3.86	H_{Me} with Br_{bonded}
216	3.57	H_{Bu} with Br_{bonded}
216	3.85	H_{Bu} with Br_{bonded}
102	3.35	H_{Bu} with Br_{bonded}
102	3.58	H_{Bu} with Br_{bonded}
102	3.75	H_{Bu} with Br_{bonded}
102	3.93	H_{Bu} with Br_{bonded}
109	3.48	H_{Bu} with Br_{bonded}

Interactions > 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
216	4.02	H_5 with Br_{bonded}
102	4.16	H_5 with Br_{bonded}
118	4.54	H_5 with Br_{bonded}
218	4.99	H_4 with Br_{bonded}
218	4.11	H_{Me} with Br_{bonded}
218	4.06	H_{Me} with Br_{bonded}
218	4.65	H_{Me} with Br_{bonded}
15	4.79	H_{Bu} with Br_{bonded}
102	4.34	H_{Bu} with Br_{bonded}
102	4.55	H_{Bu} with Br_{bonded}
102	4.62	H_{Bu} with Br_{bonded}
109	4.44	H_{Bu} with Br_{bonded}
109	4.48	H_{Bu} with Br_{bonded}
109	4.51	H_{Bu} with Br_{bonded}
109	4.54	H_{Bu} with Br_{bonded}
109	4.59	H_{Bu} with Br_{bonded}
216	4.17	H_{Bu} with Br_{bonded}
216	4.46	H_{Bu} with Br_{bonded}
216	4.83	H_{Bu} with Br_{bonded}
216	4.96	H_{Bu} with Br_{bonded}
118	4.47	H_{Me} with C_{anion}
118	4.61	H_{Me} with C_{anion}
109	4.50	H_{Bu} with C_{anion}
109	4.49	H_{Bu} with C_{anion}
98	4.85	H_{Bu} with C_{anion}

All Interactions < 5 Angstroms

PF_6 Anion	Average Distance (Å)	Interaction
299	4.51	F_{PF6} with $H_{transfer}$
299	4.67	F_{PF6} with $H_{transfer}$
152	4.70	F_{PF6} with $H_{transfer}$

Table 3S.6: Pyrrolidine reactant-complex in [BMIM][PF_6] – solvent effect data

Interactions < 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
109	3.09	H_2 with Br_{bonded}
216	3.02	H_5 with Br_{bonded}
118	3.00	H_{Me} with Br_{bonded}
118	3.61	H_{Me} with Br_{bonded}
118	3.98	H_{Me} with Br_{bonded}
109	3.13	H_{Me} with Br_{bonded}
102	3.65	H_{Bu} with Br_{bonded}
102	3.66	H_{Bu} with Br_{bonded}
102	3.69	H_{Bu} with Br_{bonded}
102	3.77	H_{Bu} with Br_{bonded}
99	3.68	H_{Bu} with Br_{bonded}
99	3.97	H_{Bu} with Br_{bonded}
213	3.70	H_{Me} with C_{anion}

All Interactions < 5 Angstroms

PF_6 Anion	Average Distance (Å)	Interaction
222	3.54	F_{PF6} with $H_{transfer}$
222	4.05	F_{PF6} with $H_{transfer}$
222	4.78	F_{PF6} with $H_{transfer}$

Interactions > 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
216	4.68	H_4 with Br_{bonded}
216	4.85	H_4 with Br_{bonded}
216	4.14	H_5 with Br_{bonded}
216	4.29	H_5 with Br_{bonded}
118	4.17	H_{Me} with Br_{bonded}
118	4.72	H_{Me} with Br_{bonded}
109	4.48	H_{Me} with Br_{bonded}
109	4.54	H_{Me} with Br_{bonded}
109	4.95	H_{Me} with Br_{bonded}
28	4.78	H_{Me} with Br_{bonded}
99	4.36	H_{Bu} with Br_{bonded}
102	4.14	H_{Bu} with Br_{bonded}
102	4.29	H_{Bu} with Br_{bonded}
102	4.44	H_{Bu} with Br_{bonded}
102	4.55	H_{Bu} with Br_{bonded}
102	4.82	H_{Bu} with Br_{bonded}
109	4.76	H_{Bu} with Br_{bonded}
118	4.44	H_{Bu} with Br_{bonded}
118	4.84	H_{Bu} with Br_{bonded}
216	4.59	H_{Bu} with Br_{bonded}
216	4.78	H_{Bu} with Br_{bonded}
109	4.02	H_{Me} with C_{anion}
216	4.36	H_{Me} with C_{anion}
213	4.39	H_{Me} with C_{anion}
99	4.39	H_{Bu} with C_{anion}
102	4.75	H_{Bu} with C_{anion}

Table 3S.7: Piperidine TS in [BMIM][PF_6] – solvent effect data

Interactions < 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
216	3.37	H_4 with Br_{bonded}
102	3.65	H_4 with Br_{bonded}
102	2.76	H_{Bu} with Br_{LG}
102	3.29	H_{Bu} with Br_{LG}
102	3.54	H_{Bu} with Br_{LG}
15	3.44	H_{Bu} with Br_{LG}
15	3.86	H_{Bu} with Br_{LG}
109	3.95	H_{Bu} with Br_{LG}
216	2.88	H_{Bu} with Br_{LG}
216	3.63	H_{Bu} with Br_{LG}
118	3.87	H_{Me} with Br_{LG}
15	2.98	H_{Bu} with Br_{bonded}
15	3.10	H_{Bu} with Br_{bonded}
15	3.43	H_{Bu} with Br_{bonded}
15	3.99	H_{Bu} with Br_{bonded}
102	3.72	H_{Bu} with Br_{bonded}
216	3.92	H_{Bu} with Br_{bonded}

Interactions > 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
102	4.47	H_4 with Br_{LG}
216	4.08	H_4 with Br_{LG}
15	4.25	H_2 with Br_{bonded}
15	4.17	H_4 with Br_{bonded}
102	4.37	H_{Bu} with Br_{LG}
102	4.64	H_{Bu} with Br_{LG}
102	4.76	H_{Bu} with Br_{LG}
102	4.87	H_{Bu} with Br_{LG}
216	4.33	H_{Bu} with Br_{LG}
216	4.85	H_{Bu} with Br_{LG}
15	4.43	H_{Bu} with Br_{LG}
109	4.77	H_{Bu} with Br_{LG}
109	4.91	H_{Bu} with Br_{LG}
118	4.39	H_{Me} with Br_{LG}
118	4.55	H_{Me} with Br_{LG}
118	4.45	H_{Me} with Br_{bonded}
118	4.67	H_{Me} with Br_{bonded}
118	5.05	H_{Me} with Br_{bonded}
15	4.79	H_{Me} with Br_{bonded}
118	4.26	H_{Me} with C_{anion}
118	4.61	H_{Me} with C_{anion}
109	4.49	H_{Bu} with C_{anion}
109	4.50	H_{Bu} with C_{anion}
109	4.49	H_{Bu} with C_{anion}

All Interactions < 5 Angstroms

PF_6 Anion	Average Distance (Å)	Interaction
152	4.43	F_{PF6} with $H_{transfer}$
152	4.63	F_{PF6} with $H_{transfer}$
152	4.81	F_{PF6} with $H_{transfer}$

Table 3S.8: Pyrrolidine TS in [BMIM][PF_6] – solvent effect data

Interactions < 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
109	3.09	H_2 with Br_{bonded}
216	3.01	H_{Bu} with Br_{LG}
102	3.65	H_{Bu} with Br_{LG}
102	3.66	H_{Bu} with Br_{LG}
102	3.69	H_{Bu} with Br_{bonded}
102	3.77	H_{Bu} with Br_{bonded}
99	3.68	H_{Bu} with Br_{bonded}
99	3.96	H_{Bu} with Br_{bonded}
118	3.00	H_{Me} with Br_{bonded}
109	3.13	H_{Me} with Br_{bonded}
28	3.61	H_{Me} with Br_{bonded}
213	3.70	H_{Me} with C_{anion}
109	4.01	H_{Bu} with C_{anion}

Interactions > 4 Angstroms

BMIM Cation	Average Distance (Å)	Interaction
216	4.14	H_4 with Br_{bonded}
216	4.29	H_4 with Br_{bonded}
216	4.68	H_3 with Br_{bonded}
118	4.44	H_2 with Br_{bonded}
118	4.83	H_2 with Br_{bonded}
102	4.14	H_{Bu} with Br_{LG}
102	4.29	H_{Bu} with Br_{LG}
102	4.55	H_{Bu} with Br_{LG}
216	4.59	H_{Bu} with Br_{LG}
216	4.78	H_{Bu} with Br_{LG}
118	4.72	H_{Me} with Br_{LG}
109	4.94	H_{Me} with Br_{LG}
99	4.36	H_{Bu} with Br_{bonded}
102	4.44	H_{Bu} with Br_{bonded}
109	4.76	H_{Bu} with Br_{bonded}
118	4.78	H_{Bu} with Br_{bonded}

All Interactions < 5 Angstroms

PF_6 Anion	Average Distance (Å)	Interaction
222	3.54	F_{PF6} with $H_{transfer}$
222	4.05	F_{PF6} with $H_{transfer}$
222	4.78	F_{PF6} with $H_{transfer}$
213	4.36	H_{Me} with C_{anion}

102	4.82	H_{Bu} with Br_{bonded}
109	4.48	H_{Me} with Br_{bonded}
109	4.56	H_{Me} with Br_{bonded}
118	4.17	H_{Me} with Br_{bonded}
118	4.78	H_{Me} with Br_{bonded}
102	4.75	H_{Me} with C_{anion}
213	4.36	H_{Me} with C_{anion}
213	4.39	H_{Me} with C_{anion}

Methanol Reactant-Complex Figures

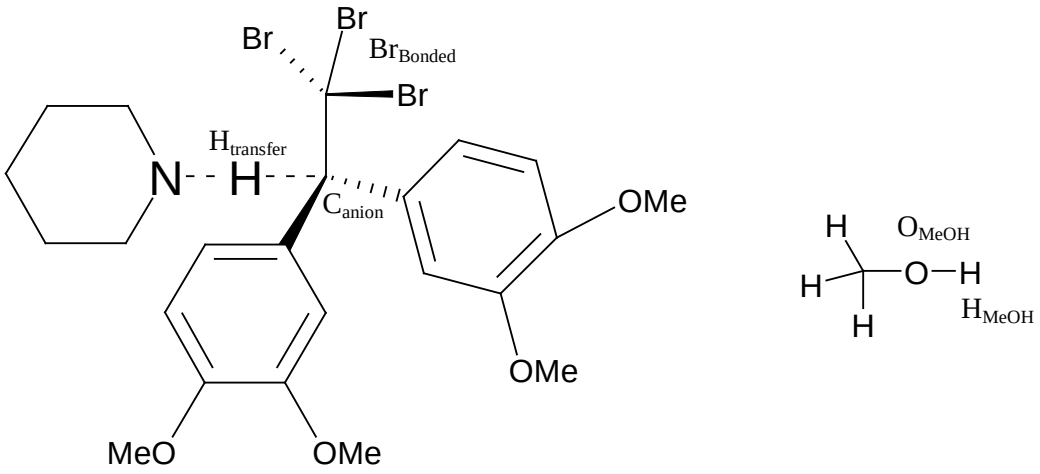


Figure 3S.9 – The reactant complex of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with piperidine in methanol. For solvent effect data see Table 3S.9.

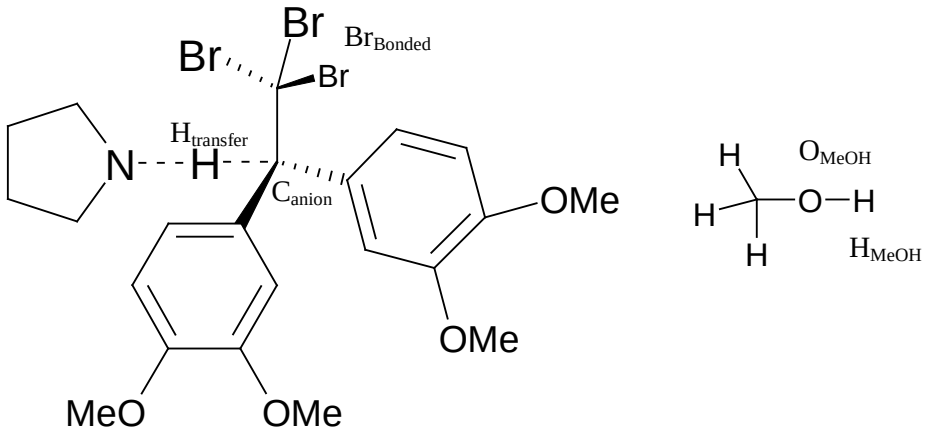


Figure 3S.10 – The reactant complex of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with pyrrolidine in methanol. For solvent effect data see Table 3S.10.

Methanol Transition Structure Figures

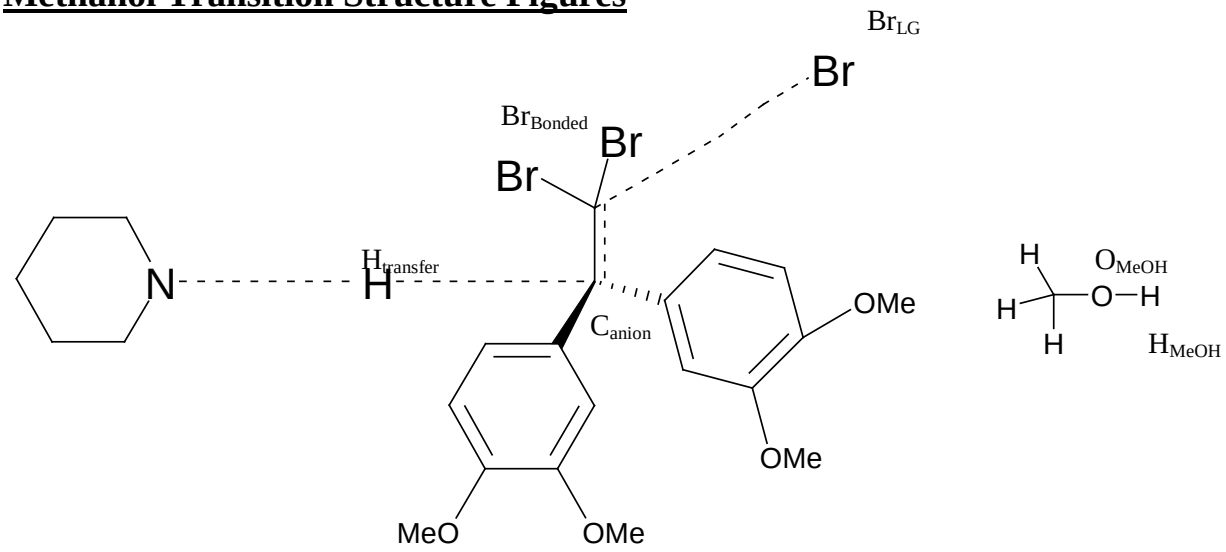


Figure 3S.11 – The transition structure of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with piperidine in methanol. For solvent effect data see Table 3S.11.

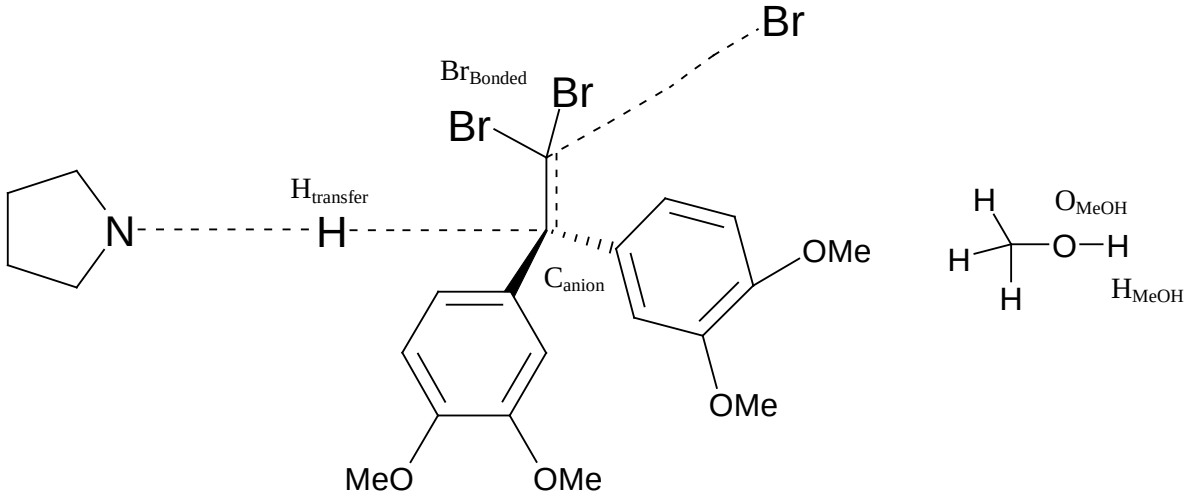


Figure 3S.12 – The transition structure of 1,1,1-tribromo-2,2-bis(3,4-dimethoxyphenyl)ethane with pyrrolidine in methanol. For solvent effect data see Table 3S.12.

Table 3S.9: Piperidine reactant-complex in methanol – solvent effect data

Interactions < 4 Angstroms | **Interactions > 4 Angstroms**

Methanol	Average Distance (Å)	Interaction
71	3.30	O_{MeOH} with $H_{transfer}$

Methanol	Average Distance (Å)	Interaction
90	5.42	O_{MeOH} with $H_{transfer}$
55	5.55	O_{MeOH} with $H_{transfer}$
199	5.91	O_{MeOH} with $H_{transfer}$
99	5.97	O_{MeOH} with $H_{transfer}$
54	4.99	H_{MeOH} with Br_{bonded}
58	5.32	H_{MeOH} with Br_{bonded}
59	4.64	H_{MeOH} with Br_{bonded}
66	5.83	H_{MeOH} with Br_{bonded}
70	5.76	H_{MeOH} with Br_{bonded}
70	5.88	H_{MeOH} with Br_{bonded}
71	4.37	H_{MeOH} with Br_{bonded}
71	4.75	H_{MeOH} with Br_{bonded}
82	4.98	H_{MeOH} with Br_{bonded}
82	5.44	H_{MeOH} with Br_{bonded}
92	5.48	H_{MeOH} with Br_{bonded}
97	4.36	H_{MeOH} with Br_{bonded}
101	5.40	H_{MeOH} with Br_{bonded}
101	5.45	H_{MeOH} with Br_{bonded}
104	4.74	H_{MeOH} with Br_{bonded}
104	5.53	H_{MeOH} with Br_{bonded}
104	5.66	H_{MeOH} with Br_{bonded}
114	5.87	H_{MeOH} with Br_{bonded}
118	5.07	H_{MeOH} with Br_{bonded}
118	5.86	H_{MeOH} with Br_{bonded}
166	5.88	H_{MeOH} with Br_{bonded}
213	4.94	H_{MeOH} with Br_{bonded}
70	4.37	H_{MeOH} with C_{anion}
55	4.46	H_{MeOH} with C_{anion}
199	4.94	H_{MeOH} with C_{anion}
114	5.86	H_{MeOH} with C_{anion}
68	5.88	H_{MeOH} with C_{anion}

Table 3S.10: Pyrrolidine reactant-complex in methanol – solvent effect data

Interactions < 4 Angstroms | **Interactions > 4 Angstroms**

Methanol	Average Distance (Å)	Interaction
79	3.20	O_{MeOH} with $H_{transfer}$

Methanol	Average Distance (Å)	Interaction
20	5.24	O_{MeOH} with $H_{transfer}$
29	5.65	O_{MeOH} with $H_{transfer}$
252	5.71	O_{MeOH} with $H_{transfer}$
20	4.31	H_{MeOH} with Br_{bonded}
20	5.41	H_{MeOH} with Br_{bonded}
36	5.01	H_{MeOH} with Br_{bonded}
36	5.62	H_{MeOH} with Br_{bonded}
56	5.29	H_{MeOH} with Br_{bonded}
56	5.29	H_{MeOH} with Br_{bonded}
56	5.83	H_{MeOH} with Br_{bonded}
62	4.76	H_{MeOH} with Br_{bonded}
62	5.82	H_{MeOH} with Br_{bonded}
69	5.93	H_{MeOH} with Br_{bonded}
79	4.58	H_{MeOH} with Br_{bonded}
79	5.20	H_{MeOH} with Br_{bonded}
81	5.05	H_{MeOH} with Br_{bonded}
81	5.62	H_{MeOH} with Br_{bonded}
84	4.72	H_{MeOH} with Br_{bonded}
103	5.11	H_{MeOH} with Br_{bonded}
104	5.28	H_{MeOH} with Br_{bonded}
213	5.57	H_{MeOH} with Br_{bonded}
20	4.96	H_{MeOH} with C_{anion}
79	4.98	H_{MeOH} with C_{anion}

Table 3S.11: Piperidine TS in methanol – solvent effect data

Interactions < 4 Angstroms

Methanol	Average Distance (Å)	Interaction
105	3.23	O_{MeOH} with $H_{transfer}$
70	2.16	H_{MeOH} with Br_{LG}
181	2.55	H_{MeOH} with Br_{LG}
83	2.65	H_{MeOH} with Br_{LG}
240	3.14	H_{MeOH} with Br_{LG}
240	3.17	H_{MeOH} with Br_{bonded}
83	3.18	H_{MeOH} with Br_{bonded}
141	3.49	H_{MeOH} with Br_{bonded}
116	3.76	H_{MeOH} with Br_{bonded}
105	3.94	H_{MeOH} with Br_{bonded}
105	3.99	H_{MeOH} with Br_{bonded}

Continuation of Right Hand Column

16	4.72	H_{MeOH} with C_{anion}
105	4.77	H_{MeOH} with C_{anion}
38	4.87	H_{MeOH} with C_{anion}
70	5.02	H_{MeOH} with C_{anion}
240	5.32	H_{MeOH} with C_{anion}
79	5.67	H_{MeOH} with C_{anion}
116	5.69	H_{MeOH} with C_{anion}

Interactions > 4 Angstroms

Methanol	Average Distance (Å)	Interaction
79	4.82	O_{MeOH} with $H_{transfer}$
38	5.11	O_{MeOH} with $H_{transfer}$
116	5.30	O_{MeOH} with $H_{transfer}$
16	5.52	O_{MeOH} with $H_{transfer}$
85	5.09	H_{MeOH} with Br_{LG}
141	5.09	H_{MeOH} with Br_{LG}
146	5.17	H_{MeOH} with Br_{LG}
58	5.43	H_{MeOH} with Br_{LG}
137	5.84	H_{MeOH} with Br_{LG}
38	4.06	H_{MeOH} with Br_{bonded}
38	4.06	H_{MeOH} with Br_{bonded}
58	4.57	H_{MeOH} with Br_{bonded}
58	4.57	H_{MeOH} with Br_{bonded}
70	4.13	H_{MeOH} with Br_{bonded}
70	4.23	H_{MeOH} with Br_{bonded}
79	4.05	H_{MeOH} with Br_{bonded}
79	4.98	H_{MeOH} with Br_{bonded}
83	5.28	H_{MeOH} with Br_{bonded}
85	5.74	H_{MeOH} with Br_{bonded}
97	5.16	H_{MeOH} with Br_{bonded}
137	5.57	H_{MeOH} with Br_{bonded}
141	5.75	H_{MeOH} with Br_{bonded}
146	5.06	H_{MeOH} with Br_{bonded}
181	4.67	H_{MeOH} with Br_{bonded}
181	5.04	H_{MeOH} with Br_{bonded}
240	4.72	H_{MeOH} with Br_{bonded}
181	4.12	H_{MeOH} with C_{anion}
83	4.67	H_{MeOH} with C_{anion}

Table 3S.12: Pyrrolidine TS in methanol – solvent effect data

Interactions < 4 Angstroms

Methanol	Average Distance (Å)	Interaction
70	3.11	O_{MeOH} with $H_{transfer}$
103	2.91	H_{MeOH} with Br_{LG}
103	3.60	H_{MeOH} with Br_{bonded}

Interactions > 4 Angstroms

Methanol	Average Distance (Å)	Interaction
194	4.88	O_{MeOH} with $H_{transfer}$
20	5.38	O_{MeOH} with $H_{transfer}$
58	5.88	O_{MeOH} with $H_{transfer}$
107	4.28	H_{MeOH} with Br_{LG}
82	5.61	H_{MeOH} with Br_{LG}
87	5.74	H_{MeOH} with Br_{LG}
58	5.47	H_{MeOH} with Br_{LG}
20	4.47	H_{MeOH} with Br_{bonded}
43	4.82	H_{MeOH} with Br_{bonded}
70	4.04	H_{MeOH} with Br_{bonded}
70	4.25	H_{MeOH} with Br_{bonded}
77	5.99	H_{MeOH} with Br_{bonded}
87	5.36	H_{MeOH} with Br_{bonded}
103	4.19	H_{MeOH} with Br_{bonded}
107	4.37	H_{MeOH} with Br_{bonded}
194	4.92	H_{MeOH} with Br_{bonded}
194	5.68	H_{MeOH} with Br_{bonded}
240	4.85	H_{MeOH} with Br_{bonded}
58	4.07	H_{MeOH} with C_{anion}
70	4.80	H_{MeOH} with C_{anion}
103	4.84	H_{MeOH} with C_{anion}
20	5.09	H_{MeOH} with C_{anion}

Chapter 3 References

(1) D'Anna, F.; Frenna, V.; Pace, V.; Noto, R. *Tetrahedron* **2006**, *62*, 1690-1698.

(2) Fontana, G.; Frenna, V.; Lamartina, L.; Natoli, M. C.; Noto, R. *J. Phys. Org. Chem.* **2002**, *15*, 108–114.

(3) Saunders Jr., W. H. *Acc. Chem. Res.* **1976**, *9*, 19-25.

(4) Jia, Z. S.; Rudziński, J.; Paneth, P.; Thibblin, A. *J. Org. Chem.* **2002**, *67*, 177-181.

(5) Fontana, G.; Frenna, V.; Gruttadauria, M.; Natoli, M. C.; Noto, R. *J. Phys. Org. Chem.* **1998**, *11*, 54-58.

(6) Gandler, J. R.; Storer, J. W.; Ohlberg, D. A. A. *J. Am. Chem. Soc.* **1990**, *112*, 7756–7762.

(7) Gandler, J. R.; Jencks, W. P. *J. Am. Chem. Soc.* **1982**, *104*, 1937-1951.

(8) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. *J. Comput. Chem.* **2002**, *23*, 1601-1622.

(9) Tubert-Brohman, I.; Guimarães, C. R. W.; Repasky, M. P.; Jorgensen, W. L. *J. Comput. Chem.* **2003**, *25*, 138-150.

(10) Acevedo, O.; Jorgensen, W. L. *Acc. Chem. Res.* **2010**, *43*, 142-151.

(11) Acevedo, O.; Armacost, K. *J. Am. Chem. Soc.* **2010**, *132*, 1966-1975.

(12) Acevedo, O.; Jorgensen, W. L. *J. Phys. Chem. B* **2010**, *114*, 8425-8430.

(13) Acevedo, O. *J. Phys. Chem. B* **2009**, *113*, 15372-15381.

(14) Sheppard, A. N.; Acevedo, O. *J. Am. Chem. Soc.* **2009**, *131*, 2530-2540.

(15) Sambasivarao, S. V.; Acevedo, O. *J. Chem. Theory Comput.* **2009**, *5*, 1038-1050.

(16) Jorgensen, W. L. *J. Phys. Chem.* **1986**, *90*, 1276-1284.

(17) Kaminski, G. A.; Jorgensen, W. L. *J. Phys. Chem. B* **1998**, *102*, 1787-1796.

(18) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **2003**, *24*, 1291-1304.

(19) Vilseck, J. Z.; Sambasivarao, S. V.; Acevedo, O. *J. Comput. Chem.* **2011**, *32*, 2836-2842.

(20) Blagović, M. U.; Morales de Tirado, P.; Pearlman, S. A.; Jorgensen, W. L. *J. Comput. Chem.* **2004**, *25*, 1322-1332.

(21) Jorgensen, W. L.; Tirado-Rives, J. *J. Comput. Chem.* **2005**, *26*, 1689-1700.

(22) Acevedo, O.; Jorgensen, W. L.; Evanseck, J. D. *J. Chem. Theory. Comput.* **2007**, *3*, 132-138.

(23) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215-241.

(24) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157-167.

(25) Frisch, M. J.; Revision B.01 ed.: Wallingford, CT, 2009 [Full reference given in Supporting Info.].

(26) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669–681.

(27) Sattelmeyer, K. W.; Tubert-Brohman, I.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2006**, *2*, 413-419.

(28) Dewar, M. J. S.; Yuan, Y.-C. *J. Am. Chem. Soc.* **1990**, *112*, 2088-2094.

(29) Dewar, M. J. S.; Yuan, Y.-C. *J. Am. Chem. Soc.* **1990**, *112*, 2095-2105.

(30) Krop, H. B.; Cheung, C. L.; Govers, H. A. J. *J. Mol. Struct. (THEOCHEM)* **2000**, *505*, 1-10.

(31) Acevedo, O.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2007**, *3*, 1412-1419.

(32) Chandrasekhar, J.; Shariffskul, S.; Jorgensen, W. L. *J. Phys. Chem. B* **2002**, *106*, 8078-8085.

(33) Gunaydin, H.; Acevedo, O.; Jorgensen, W. L.; Houk, K. N. *J. Chem. Theory Comput.* **2007**, *3*, 1028-1035.

(34) Smith, M. B.; March, J. In *March's Advanced Organic Chemistry*; 6 ed.; John Wiley & Sons: Hoboken, New Jersey, 2007; pp 1477-1558.

(35) Reichardt, C.; Welton, T. *Solvents and Solvent Effects in Organic Chemistry*; 4th ed.; Wiley-VCH Verlag GmbH, 2010.

(36) An 'imaginary' frequency is a negative eigenvalue of the Hessian matrix corresponding to the motion of the system along the reaction pathway.

(37) Alunni, S.; Angelis, F. D.; Ottavi, L.; Papavasileiou, M.; Tarantelli, F. *J. Am. Chem. Soc.* **2005**, *127*, 15151-15160.

(38) Amyes, T. L.; Diver, S. T.; Richard, J. P.; Rivas, F. M.; Toth, K. *J. Am. Chem. Soc.* **2004**, *126*, 4366–4374.

(39) Dupont, J.; Spencer, J. *Angew. Chem. Int. Ed.* **2004**, *43*, 5296–5297.

(40) Holbrey, J. D.; Reichert, W. M.; Nieuwenhuyzen, M.; Sheppard, O.; Hardacre, C.; Rogers, R. D. *Chem. Commun.* **2003**, 476-477.

(41) Carmichael, A. J.; Seddon, K. R. *J. Phys. Org. Chem.* **2000**, *13*, 591–595.

(42) Crowhurst, L.; Mawdsley, P. R.; Perez-Arlandis, J. M.; Salter, P. A.; Welton, T. *Phys. Chem. Chem. Phys.* **2003**, *5*, 2790–2794.

(43) Deetlefs, M.; Hardacre, C.; Nieuwenhuyzen, M.; Sheppard, O.; Soper, A. K. *J. Phys. Chem. B* **2005**, *109*, 1593–1598.

(44) Jorgensen, W. L.; Severance, D. L. *J. Am. Chem. Soc.* **1990**, *112*, 4168-4114.

(45) Fu, C.-F.; Tian, S. X. *J. Chem. Theory Comput.* **2011**, *7*, 2240−2252.

(46) Takeuchi, H. *J. Phys. Chem. A* **2012**, *116*, 10172−10181.

Chapter 4

Effects of Ionic Liquid Solvents on a Nucleophilic Aromatic Substitution Reaction from QM/MM/MC Simulations.

4.1 Abstract

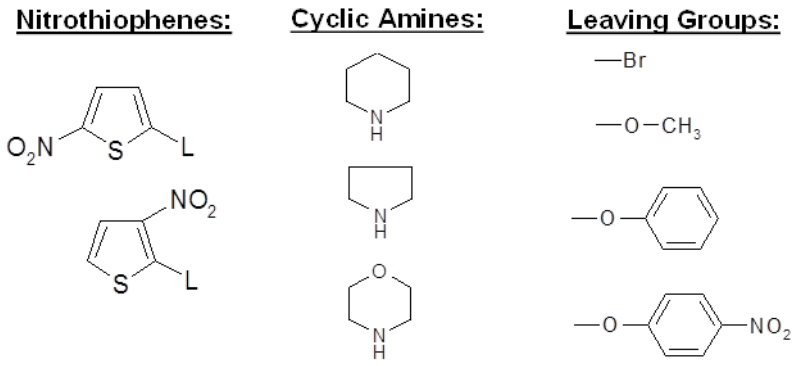
Ionic liquids have been reported to enhance the rates of reaction for the fundamentally important nucleophilic aromatic substitution reaction when compared to the conventional solvent methanol. The role of the reaction medium in the nucleophilic aromatic substitution reaction between three amines, piperidine, pyrrolidine, and morpholine, with 2-L-5- (para-like isomer) and 2-L-3-nitrothiophene (ortho-like isomer) has been computationally investigated using methanol and the room temperature ionic liquids 1-butyl-3-methylimidazolium tetrafluoroborate and hexafluorophosphate [BMIM][BF_4] and [BMIM][PF_6], respectively. QM/MM Monte Carlo simulations utilizing free-energy perturbation theory found the ionic liquids did enhance the rates of reaction which is consistent with experimental observations. The origin of the ionic liquid effect can be attributed to (1) the decrease in solvation of the amine nucleophile, (2) the structural and energetic similarities of the addition step transition state to the Meisenheimer intermediate complex, and (3) the formation of an ionic liquid clathrate. This work elucidates the ionic liquids role upon the accelerated nucleophilic

aromatic substitution reaction rates and describes the critical interplay between sterics and electrostatics crucial to develop a deeper understanding of the effect these unique solvents have upon chemical reactions.

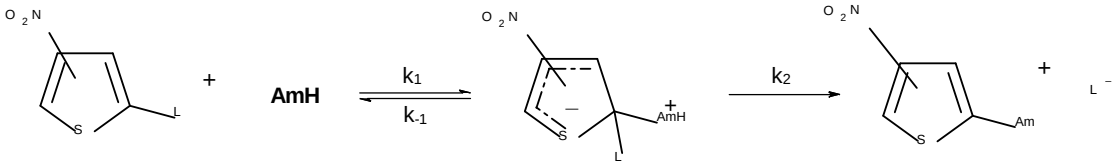
4.2 Introduction

Of current interest are the nucleophilic aromatic substitution reactions (S_NAr) reported by D'Anna *et al.* of three different amines: pyrrolidine (Pyr), piperidine (Pip), and morpholine (Mor) with 2-L-5-nitrothiophene and 2-L-3-nitrothiophene (L = bromine, methoxy, phenoxide and 4-nitrophenoxide) in methanol and the room temperature ionic liquids [BMIM][BF_4], and [BMIM][PF_6].[1] [BMIM] is 1-butyl-3-methylimidazolium and the counterions [BF_4] tetrafluoroborate or [PF_6] hexafluorophosphate (see Scheme 4.1 and Scheme 4.2).[1] The S_NAr reaction undergoes a rate acceleration when the reaction medium is changed from the conventional solvent methanol to RTILs. The amines were chosen because of their differences in nucleophilicity and/or structure that could cause specific interactions with these highly organized solvents. In addition, the ortho-like nitrothiophene isomer possesses an intramolecular interaction between the nitro group and the amine proton in the transition state that can be used to elucidate information about the reaction rate dependence on the medium. The present work applied mixed quantum and molecular mechanical (QM/MM) calculations utilizing Monte Carlo sampling and free-energy perturbation theory (MC/FEP). The QM/MM simulations with explicit solvent representation can provide the medium-dependence of the activation barriers and atomic-level structural detail for characterization of the nature of the ionic liquids. The ability of an ionic liquid to act as a hydrogen bond donor and hydrogen bond acceptor[2] is believed to be the origin of the enhanced rates of reaction. The goal of this

work is to provide new insights into the effect of ionic liquids upon an important organic reaction, which can help predict optimal conditions for other reactions in addition/elimination classes.



Scheme 4.1: Reactants: 2-L-5-nitrothiophene and, 2-L-3-nitrothiophene, three cyclic amine nucleophiles: piperidine, pyrrolidine, and morpholine, and the leaving groups: bromine, methoxy, phenoxide and 4-nitrophenoxide.



Scheme 4.2[1]: Mechanism of the aromatic nucleophilic substitution reaction.

4.3 Computational Methods

The S_NAr reaction between pyrrolidine, piperidine, or morpholine with 2-L-5-nitrothiophene (para-like isomer) and 2-L-3-nitrothiophene (ortho-like isomer) were examined using mixed quantum and molecular mechanical calculations (QM/MM) in conjunction with Metropolis Monte Carlo simulations (MC) and free energy perturbation theory (FEP). Calculations were performed with *BOSS*[3] at 25˚C and 1 atm and the QM

method of choice was PDDG/PM3[4] for the reacting system. Partial charges were obtained from the CM3 charge model.[5] PDDG/PM3 has been previously utilized in solution phase QM/MM of nucleophilic substitution bimolecular reactions (S_N2)[6], nucleophilic aromatic substitution reactions (S_NAr),[7] decarboxylation reactions,[8] Cope eliminations, the Claisen rearrangement,[9] and for gas phase energetics and structures.[4] In addition, PDDG/PM3 has been successful in the study of ionic liquids for the Diels-Alder reaction,[10] the Kemp elimination,[11] and a β-elimination.[2d] The computation of the QM energy and atomic charges was performed for each attempted move of the solute, which occurred every 100 configurations. Lennard-Jones interactions between solute and solvent atoms were taken into account using OPLS parameters.

Periodic boundary conditions were applied to boxes containing the solutes plus 188 ion pairs for the ionic liquids and 390 solvent molecules for methanol. The solvent molecules were represented explicitly using our custom ionic liquid OPLS-AA force field[11] and the united-atom OPLS force field for methanol.[12] All ionic liquid cations were fully flexible, meaning that all bond stretching, angle bending, and torsional motions were sampled with MC. Anions were simulated as rigid molecules. The use of rigid anions in OPLS-AA has been shown to provide an accurate representation of ionic liquid physical properties, including the solvent effects for previous computed QM/MM reaction studies. The periodic boxes are tetragonal with c/a = 1.5 where a is 26.7, 34.3, and 35.5 Å for methanol, [BMIM][BF_4], and [BMIM][PF_6], respectively, with long range electrostatic interactions handled with Ewald summations. These long range electrostatic effects are prevalent in ionic liquids and thus the Ewald summation corrects the inaccuracy resulting from a simple cutoff. Solute-solvent and solvent-solvent

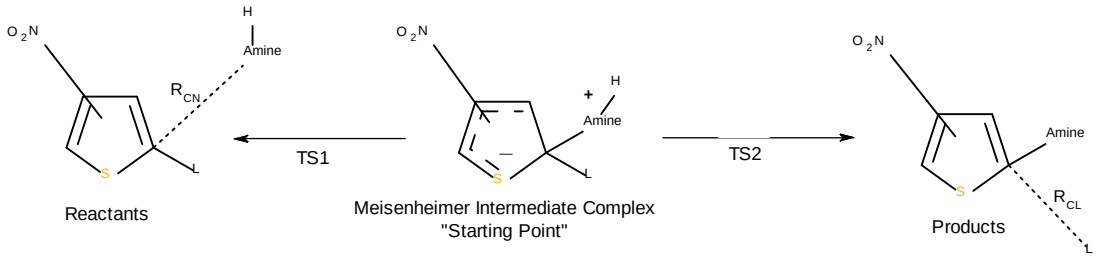
intermolecular cutoff distances of 12 Å were employed for the tail carbon atom of each side chain (methyl and alkyl), a midpoint carbon on the alkyl chain, and the ring carbon between both nitrogens for imidazolium. Center atoms, e.g. B in BF_4^- and P in PF_6^-, were used for the anions. If any distance was within the cutoff, the entire solvent-solvent interaction was included. Adjustments to the allowed ranges for rotations, translations, and dihedral angle movements led to overall acceptance rates of about 69% 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.

Solutes were inserted with the appropriate solute geometry corresponding to each free energy perturbation (FEP) window and re-equilibrated for minimally 160 million MC configurations. The starting geometries for the solutes were determined by executing a MC conformational search that resulted in up to 100 unique structures. The top ten most favorable MC structures were then recomputed using B3LYP/6-31G(d) geometry optimizations and the resultant lowest energy structure was used as the starting geometry for the QM/MM calculation.

4.4 Potential Energy Surfaces

To locate the energy minima and maxima of the nucleophilic aromatic substitution reaction, a one-dimensional potential free energy diagram was constructed. The free energy of activation, $\Delta G^{\ddagger}$, for the addition step was calculated by perturbing the distance R_{CN} between the reacting carbon of the substituted nitrothiophenes and the nitrogen in the nucleophile (piperidine, pyrrolidine or morpholine). The remainder of the molecule was allowed to be variable. The transition state for the elimination step was computed by perturbing the distance R_{CL} between the leaving group (Br or O atom) and

the reacting carbon in the Meisenheimer intermediate complex (MIC). The bond length range for R_{CN} and R_{CL} were 1.40-5.00Å with an increment of 0.01Å (see Figure 4.1 and 4.2). Each potentials of mean force (PMF) calculation in methanol required 2.5 million and 5 million steps of equilibration and averaging, respectively.



Scheme 4.3: PMF procedure for construction of the free energy surface beginning from the Meisenheimer Intermediate Complex (MIC).

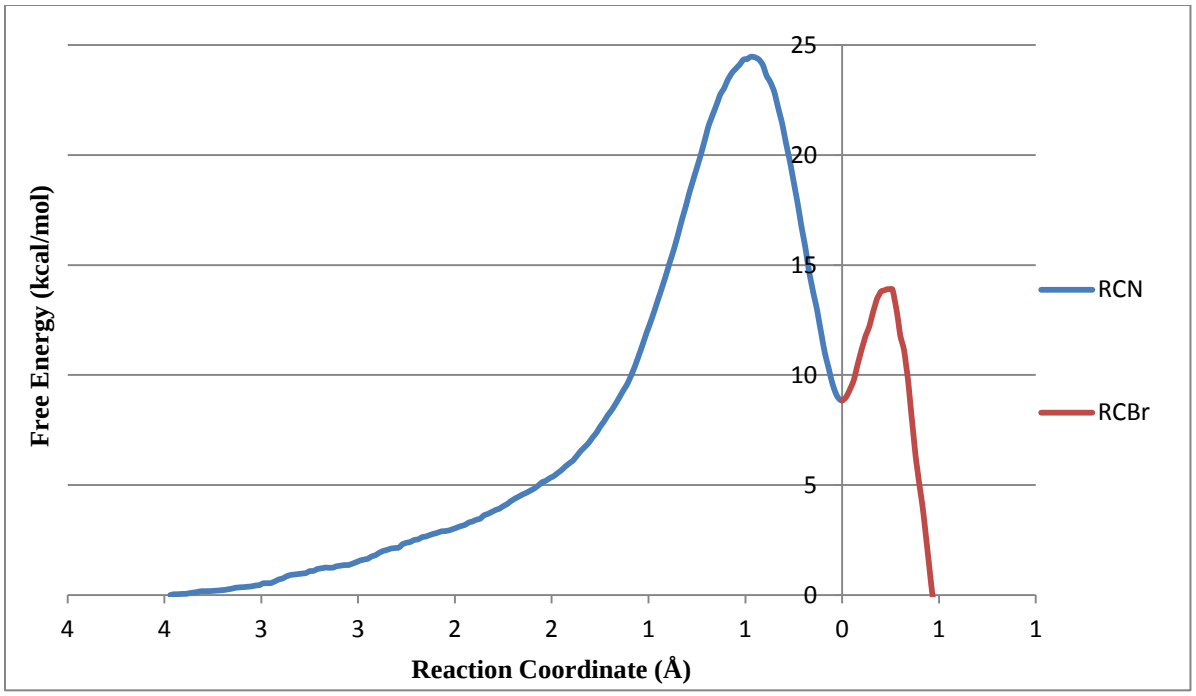


Figure 4.1: The free energy surface of the nucleophilic aromatic substitution reaction between morpholine and 2-bromo-3-nitrothiophene in methanol.

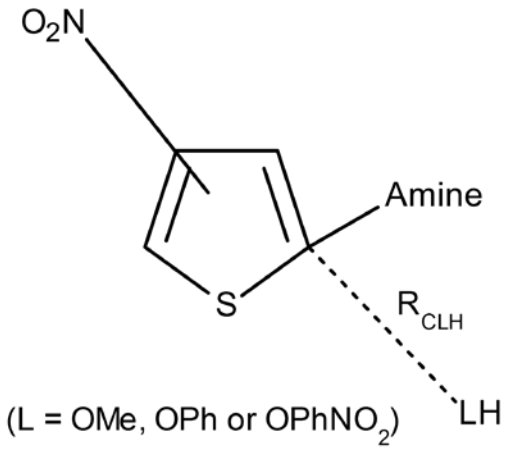
4.4.1 Energetics

The calculated free energies of activation for the addition step in methanol generally agreed well with experimental values (Table 4.1). For instance, the S_NAr reaction between morpholine and 2-bromo-5-nitrothiophene yielded a computed $\Delta G^{\ddagger}$ of 24.6 kcal/mol (exptl. 26.7 kcal/mol[1]) and the reaction between piperidine and 2-methoxy-5-nitrothiophene gave a computed $\Delta G^{\ddagger}$ of 21.8 kcal/mol (exptl. 23.5 kcal/mol[1]). The estimated errors for the precision of the calculations are ±0.4 kcal/mol. When the $\Delta G^{\ddagger}$ for the elimination step were computed for the leaving groups methoxy, phenoxide, and 4-nitrophenoxide, it was found that those corresponding mechanism energies were more complex than in Br due to the proton transfer. Computed energies were predicted in a range from > 50 kcal/mol to no activation barrier depending on the protonation state of

the amine in the MIC (see Figure 4.2). For example, if the proton was transferred from the amine to the leaving group in the MIC neutralizing the charge separation between the nitrothiophene and leaving group, (Scheme 4.4), no free energy barrier for the elimination step was found. Experimental rate constants have a linear dependence on amine concentration, in both the RTILs and methanol.[1] The experimental evidence suggests that the reaction is uncatalyzed and the addition step is rate determining.[1] Therefore for this investigation, we chose to focus on the rate determining addition step, which leads to the formation of the MIC.

The free energy profile for the rate determining addition step was determined using 0.02Å increments for R_{CN} in the RTILs (Table 4.2). Extensive reorganization of the solvent for the ionic liquids resulted in 160 million configurations of equilibration followed by 20 million MC steps of averaging per FEP window. Excellent agreement between the computed and experimental $\Delta G^{\ddagger}$ for the S_NAr reaction was found. For example, pyrrolidine and 2-bromo-5-nitrothiophene in [BMIM][PF_6] and piperidine and 2-bromo-5-nitrothiophene in [BMIM][BF_4] gave computed $\Delta G^{\ddagger}$of 23.9 kcal/mol (exptl. 23.2 kcal/mol[1]) and 25.8 kcal/mol (exptl. 23.0 kcal/mol[1]), respectively. The error bars in the ionic liquids simulations are estimated at ± 1 kcal/mol. It is important to note the overestimation of the absolute $\Delta G^{\ddagger}$ values. The calculated $\Delta G^{\ddagger}$ overestimation is a systematic error common in many organic reactions when using semiempirical QM/MM methods.[13] This overestimation is not limited to the S_NAr reactions in ionic liquids as similar findings have been reported previously for our β – elimination reaction[2d] and as well as for multiple Diels-Alder reactions,[14] ene reactions,[15] Claisen rearrangements,[16] and methyl transfer reactions.[17] When utilizing the AM1 method, Dewar reported mixed

energetics on the elimination class of reactions.[18] In addition, COSMO-AM1 and experimental values of the free activation enthalpy on a set of organic reactions in water showed large discrepancies; however, reactivity aspects were correctly predicted.[19] Conceivably, a straightforward reparameterization of the PDDG/PM3 Hamiltonian by scaling the energies from points along the reaction coordinate could provide more accurate $\Delta G^{\ddagger}$ values; however, the physical reasons for determining structures would be absolutely the same as the original Hamiltonian. Consequently, there is no difference in leaving the Hamiltonian in its original form or in scaling the energies when one considers the relative solvent effects.



Scheme 4.4: An alternative PMF procedure to calculate the free energy for the elimination step. The proton has been transferred from the amine nucleophile to the leaving group, neutralizing the charge separation.

Nucleophile	Thiophene Derivative	Leaving Group	$\Delta G^{\ddagger}$ Methanol (kcal/mol)
Morpholine	ortho	Br	24.5
		OMe	21.3
		OPh	20.9
		$OPhNO_2$	20.9
	para	Br	24.6 (26.7)
		OMe	23.0
		OPh	22.8
		$OPhNO_2$	21.1
Piperidine	ortho	Br	21.5 (24.8)
		OMe	21.8
		OPh	21.5
		$OPhNO_2$	19.7
	para	Br	26.1 (25.9)
		OMe	21.8 (23.5)
		OPh	24.3
		$OPhNO_2$	20.4 (24.1)
Pyrrolidine	ortho	Br	21.9 (24.3)
		OMe	21.3
		OPh	19.1
		$OPhNO_2$	17.4
	para	Br	22.1 (25.6)
		OMe	21.9
		OPh	21.5
		$OPhNO_2$	21.2

Table 4.1: The QM/MM/MC free energies of activation of the addition step for the nucleophilic aromatic substitution reactions between 2-L-3- (ortho isomer) and 2-L-5-nitrothiophene (para isomer) in methanol. Experimental energies are given in parentheses.

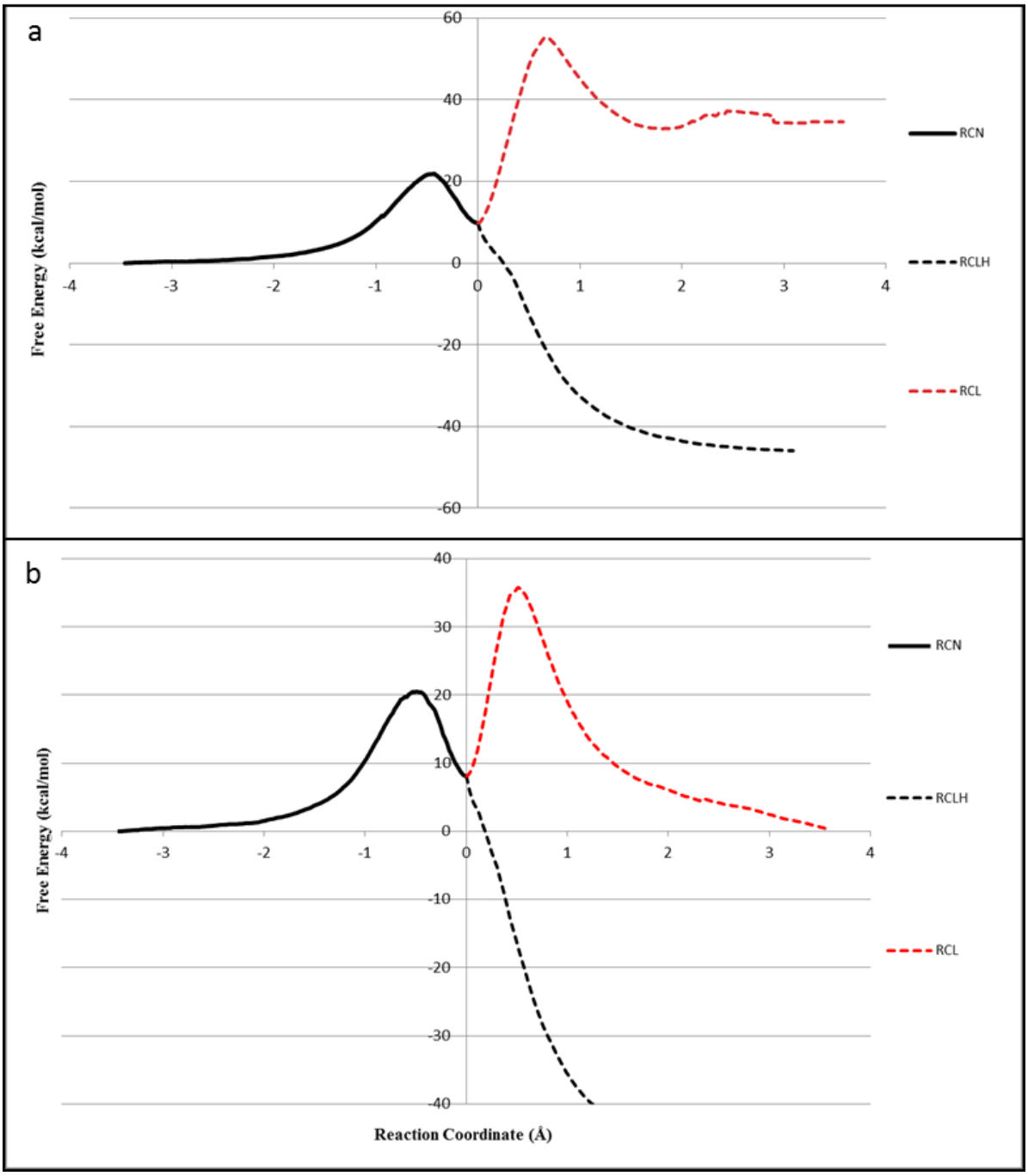


Figure 4.2: The free energy surfaces of the S_NAr reactions between (a) piperidine and 2-methoxy-5-nitrothiophene and (b) piperidine and 2-(4-nitrophenoxide)-5-nitrothiophene in methanol.

Nucleophile	Thiophene Derivative	Leaving Group	$\Delta G^{\ddagger}$ [BMIM][BF_4] (kcal/mol)	$\Delta G^{\ddagger}$ [BMIM][PF_6] (kcal/mol)
Morpholine	ortho	Br	27.9 (23.7)	26.5
		OMe	25.1	25.9
		OPh	28.5	27.8
		$OPhNO_2$	27.8	23.8
	para	Br	25.0 (24.5)	27.5
		OMe	30.0	26.6
		OPh	27.3	28.5
		$OPhNO_2$	28.0	28.6
Piperidine	ortho	Br	23.5 (22.5)	22.8
		OMe	27.3	24.0
		OPh	24.8	24.0
		$OPhNO_2$	25.4	21.1
	para	Br	25.8 (23.0)	25.1
		OMe	27.6 (21.8)	26.2
		OPh	23.8 (22.1)	26.0
		$OPhNO_2$	24.1 (21.5)	29.1
Pyrrolidine	ortho	Br	25.1 (21.7)	27.7
		OMe	23.6	25.6
		OPh	26.7	25.1
		$OPhNO_2$	21.7	22.6
	para	Br	28.6 (22.7)	23.9 (23.2)
		OMe	28.8	26.4
		OPh	26.1	25.5
		$OPhNO_2$	25.1	23.7

Table 4.2: The QM/MM/MC free energies of activation for the addition step of the nucleophilic aromatic substitution reactions in [BMIM][BF_4] and [BMIM][PF_6]. Experimental energies are given in parentheses.

4.5 Results and Discussion

4.5.1 Amine Solvation

As the rate determining step of the reaction is the nucleophilic attack, the effective nucleophilicity of the amine should be directly related to the extent and degree of solvation. The increased rate of reaction on going from methanol to RTILs suggests a potential decrease in solvation of the nucleophile or increase in solvation of the substrate. The structurally similar amines piperidine and pyrrolidine have comparable pK_a values in water, 11.12 – 11.22,[20,21] and 11.27[20] respectively, and morpholine has a pK_a value of 8.33 – 8.36[20,21] in water. It is understood that the amine basicity measured in water is not an adequate description of the effective strength of the nucleophile and that particular interactions between the RTIL and amine will determine different nucleophilicity orders and corresponding strengths in RTILs and methanol.

The energy pair distributions of the cyclic amines in methanol, seen in Figure 4.3, exhibit an increase in the number of solvent molecules coordinating with morpholine at much lower interaction energies than piperidine or pyrrolidine. Integration of the distributions from -10.0 to -5.0 (or -3.5) kcal/mol confirms the more favorable interactions occurring with morpholine than with piperidine or pyrrolidine: 1.97 (4.16), 0.80 (2.25), and 1.70 (3.64) respectively. When the radial distribution functions of the cyclic amines in methanol are analyzed (Figure 4.4), the electrostatic interactions that morpholine undergoes are shorter than for either piperidine or pyrrolidine. For instance Figure 4.4c illustrates how six methanol molecules interact at an average distance of 2.60 Å with the oxygen atom in morpholine. Additionally, an average of five methanol molecules interact with the amine proton at a distance of 2.50 Å, and seven methanol

molecules interact with the amine nitrogen at an average distance of 3.60 Å. On the other hand piperidine (in Figure 4.4a) has an average of only two methanol molecules interacting with the amine nitrogen at an average distance 2.40 Å and 13 methanol molecules interacting with the amine proton, but at a distance of 4.70 Å. These methanol results, in conjunction with the free energy results in Table 4.1, suggest the amine nucleophilicity in methanol to be Pip > Pyr > Mor. A shift to longer, less energetically favorable interactions increases the rate of the S_NAr reaction from methanol to [BMIM][BF4] and [BMIM][PF6], see Table 4.3 and Figure 4.5. Morpholine was found to have an average of 4 ions within 4 Å of the reacting system in [BMIM][PF_6] and 3 ions in [BMIM][BF_4]. In contrast, pyrrolidine had one ion within 4 Å in [BMIM][BF_4] and 2 ions in [BMIM][PF_6]. Table 4.3 shows the average solute-solvent interactions between the amines in the RTILs. Morpholine has the greatest number of interactions that are within 4 Å, the strongest being in the RTIL [BMIM][PF_6] between the acidic proton on the imidazolium cation and the oxygen in morpholine. On the other hand pyrrolidine has the fewest solute-solvent interactions within 4 Å in either RTIL. The shortest interaction occurs between the counterion BF_4^- and the amine proton at a distance of 3.46 Å. In the RTILs [BMIM][BF_4] and [BMIM][PF_6] the amine nucleophilicity is Pyr $\geq$ Pip > Mor. Furthermore, the decrease in solvation of the nucleophile, which results in the increase of the rate of reaction, is directly related to the corresponding Kamlet-Taft parameters of the solvents. The solvent methanol has a hydrogen bond acidity that is nearly double that of [BMIM][BF_4] or [BMIM][PF_6]; 1.05, 0.627, and 0.634 respectively and a hydrogen bond basicity that is over double that of the RTILs, 0.61, 0.376 and 0.207 respectively.[22]

The increase in hydrogen bond acceptor and donor capability increases the number of favorable electrostatic interactions with the nucleophile. These solvent interactions ultimately form a solvent cage around the amine changing the effective nucleophilicity and thus the corresponding rate of reaction (see Figure 4.6 and 4.7). This trend has also been previously reported by Welton and coworkers[23] who performed a kinetic study of substitution reactions between platinum (II) complexes with thioacetate in RTILs and conventional solvents, such as methanol. It was reported that upon the analysis of solvent effects it was important to consider the free energy of solvation of the nucleophile since the reaction was reported to decrease in rate with increasing hydrogen bonding ability or acidity of the solvent. Moreover, as the reacting electrophilic substrate is constant, the decreased rate of reaction was partly attributed to the increased solvation of the nucleophile.[23]

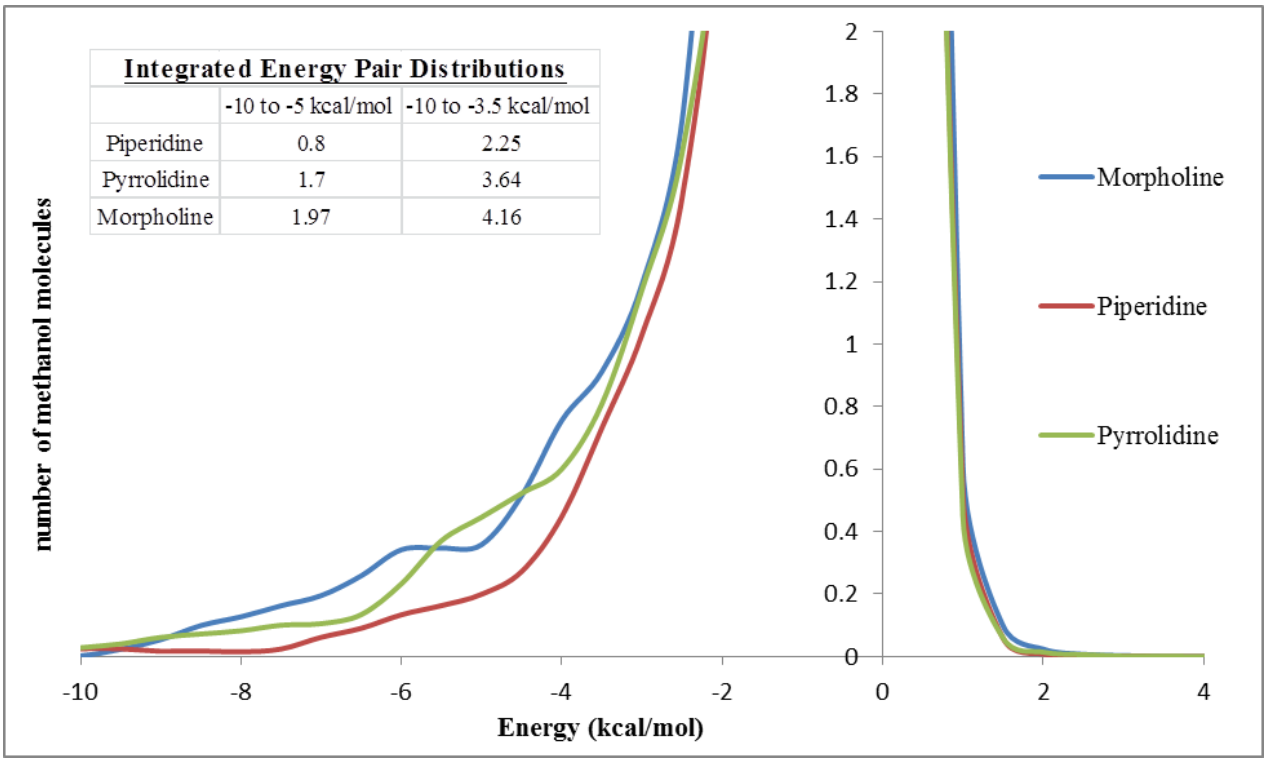


Figure 4.3: The solute-solvent energy pair distributions for morpholine, piperidine, and pyrrolidine in methanol from the QM/MM/MC calculations. The ordinate records the number of solvent molecules that interact with the solute and their interaction energy on the abscissa. Units for ordinate are number of molecules per kcal/mol.

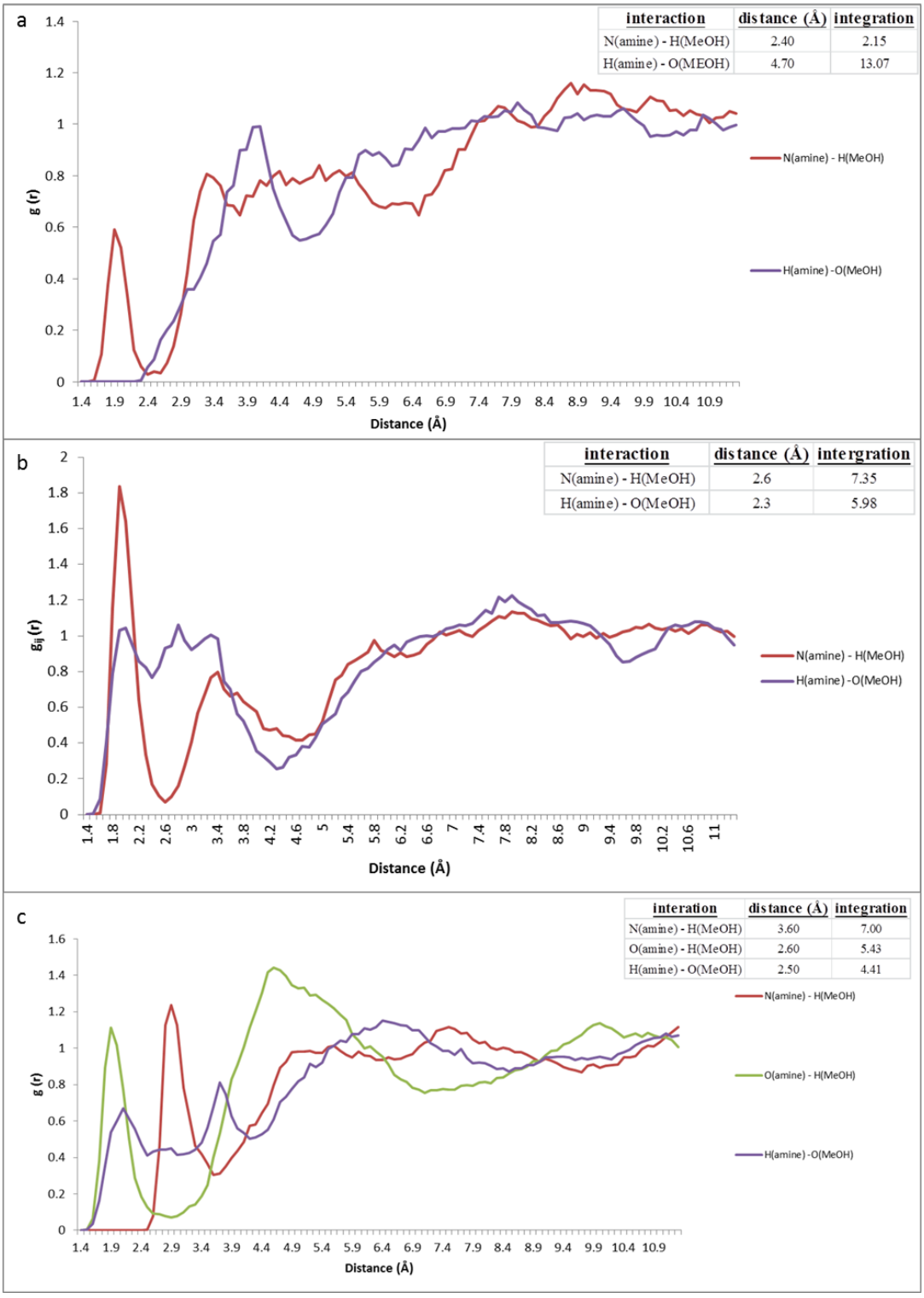
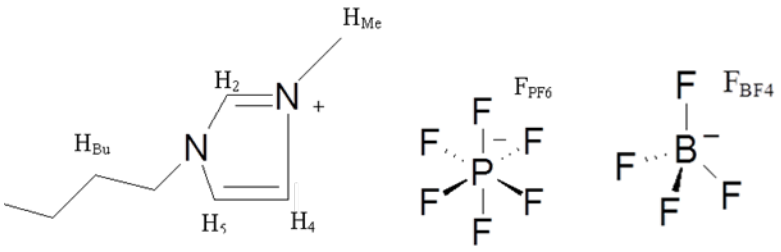


Figure 4.4: The radial distributions of (a) piperidine, (b) pyrrolidine and (c) morpholine in methanol from QM/MM/MC calculations.



Amine	Interaction < 4 Å in [BMIM][PF_6]	Distance (Å)
Morpholine	$F_{PF6} - H_{amine}$	4.09
	$H_4 - N_{amine}$	4.04
	$H_{bu} - N_{amine}$	3.73
	$H_2 - O_{amine}$	2.79
	$H_{bu} - O_{amine}$	3.58
	$H_{bu} - O_{amine}$	3.70
	$H_{bu} - O_{amine}$	3.71
Pyrrolidine	$H_{bu} - N_{amine}$	3.69
	$F_{BF4} - H_{amine}$	4.04
Piperidine	$H_4 - N_{amine}$	3.69
	$H_5 - N_{amine}$	3.42

Amine	Interaction < 4 Å in [BMIM][BF_4]	Distance (Å)
Morpholine	$H_{me} - N_{amine}$	3.49
	$H_{me} - O_{amine}$	3.66
	$H_2 - O_{amine}$	3.57
	$H_2 - O_{amine}$	3.51
	$H_{bu} - O_{amine}$	3.90
Pyrrolidine	$F_{BF4} - H_{amine}$	3.46
Piperidine	$F_{BF4} - H_{amine}$	3.18
	$H_4 - N_{amine}$	2.95
	$H_5 - N_{amine}$	3.63
	$H_{me} - N_{amine}$	3.75
	$H_{bu} - N_{amine}$	3.96

Table 4.3: The average solvent-solute interactions between the cyclic amines (Mor, Pyr, Pip) and [BMIM][PF_6] and [BMIM][BF_4] within ≈ 4 Å from QM/MM/MC calculations.

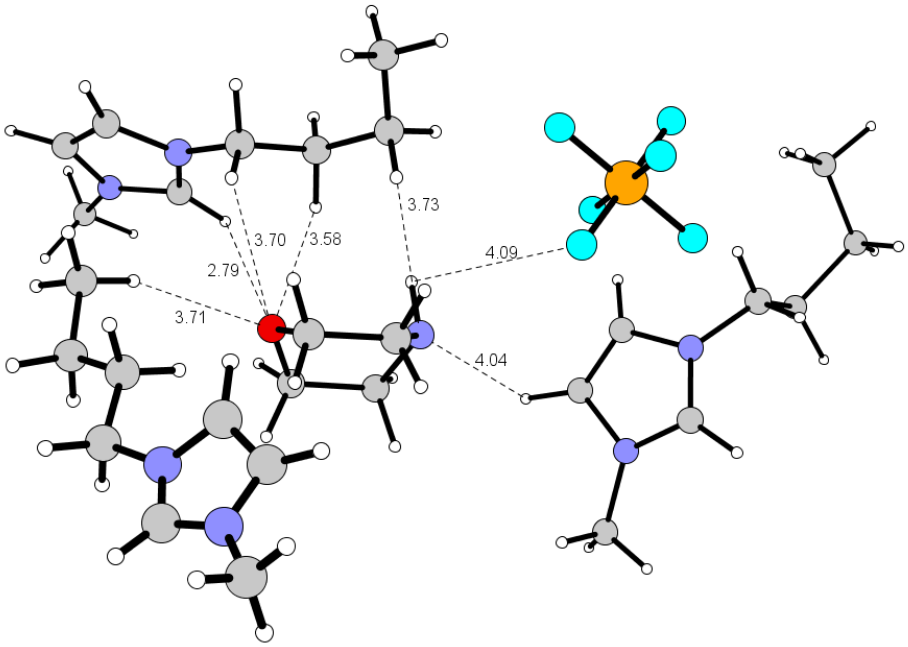


Figure 4.5: The key interactions of morpholine in the RTIL [BMIM][PF_6], from Table 4.3.

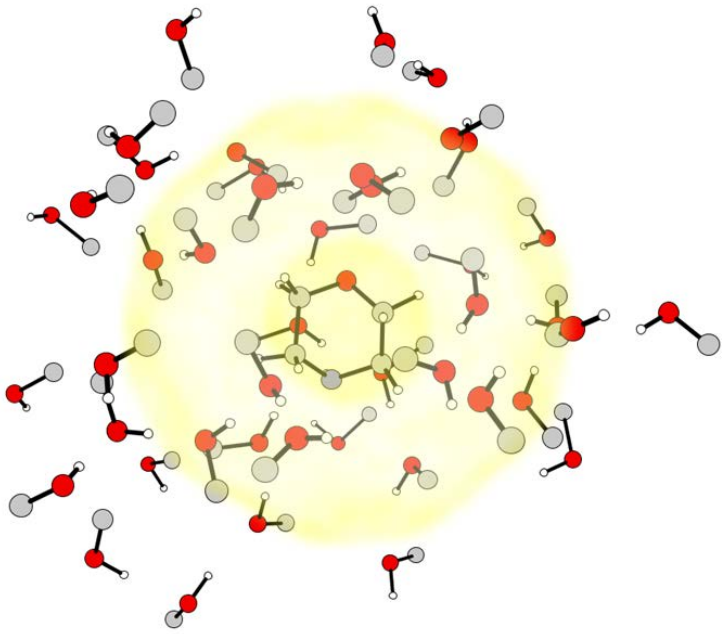


Figure 4.6: The methanol 8 Å solvation shell of morpholine from QM/MM/MC calculations.

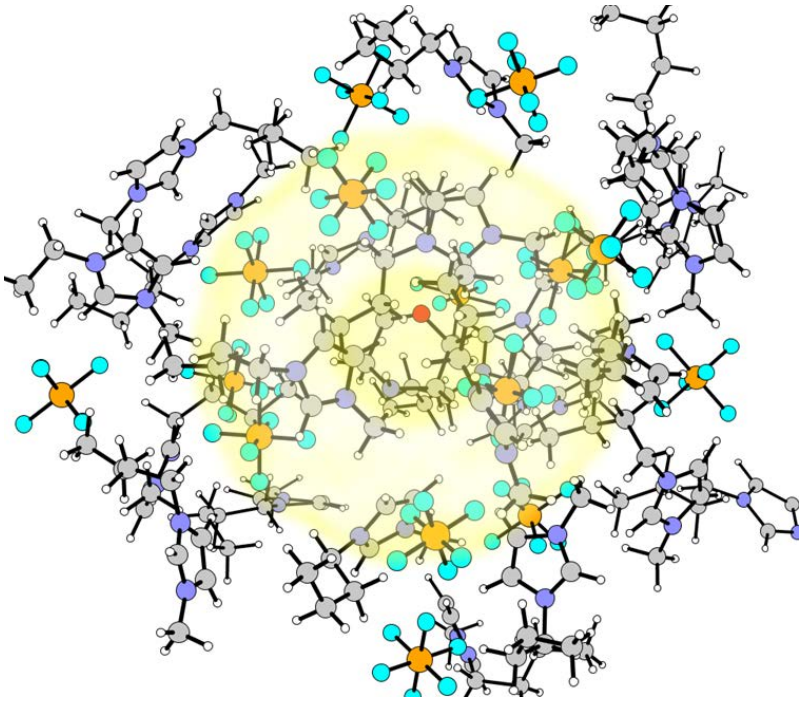


Figure 4.7: The [BMIM][PF_6] 8 Å solvation shell of morpholine from QM/MM/MC calculations.

4.5.2 Leaving Group Effects

Selected atomic charges from the QM/MM/MC calculations (Figure 4.8), show how the higher electronegativity of the oxygen atom from the methoxide, phenoxide, and 4-nitrophenoxide leaving groups, induces a higher partial positive charge density on the reacting carbon than the leaving group bromine. This higher positive charge density favors nucleophilic attack, increasing the rate of the addition step. This trend is observed in methanol and in both RTILs for the 2-L-5- and 2-L-3-nitrothiophenes. The rates of reaction are increased as the partial positive charge increases in the order of OPh ≈ 4-OPh-NO_2 > OMe > Br. Additionally, a coplanar transition state between the substituted aromatic nitrothiophene and the leaving groups: phenoxide and 4-nitrophenoxide, should maximize the electronic effects exerted on the reaction route. These leaving groups have

the capability to distribute the negative charge density more effectively in a planar configuration, due to a favorable conjugation effect. Monitoring the average Φ (defined in Figure 4.10) over the final 20 million configurations of the QM/MM/MC calculations suggests that the addition step transition states in methanol prefer to be exclusively in the gauche or trans configuration. In the RTILs however, on average there is a distribution of the ratio of the gauche to the trans configurations. Take for example the transition state formed from the S_NAr reaction between morpholine and 2-phenoxide-3-nitrothiophene. In methanol the addition transition state has a trans configuration 100% of the time, but in the RTILs the transition state exists in the gauche configuration about 20% of the time. This suggests that the RTILs provide additional stability allowing for a small percentage of the less favorable gauche configuration to be occupied. However both solvents favor a coplanar configuration (Table 4.4) with the 2-L-5- (para-like) substituted nitrothiophene and a twisted configuration with the 2-L-3- (ortho-like) substituted nitrothiophene, owing to the steric hindrance of the nitro group. The coplanar emphasizes conjugation, increasing the electron-withdrawing effect of these oxygen atoms bonded to the heteroaryl moiety, which is supported by the calculated partial atomic charges in Figure 4.8.

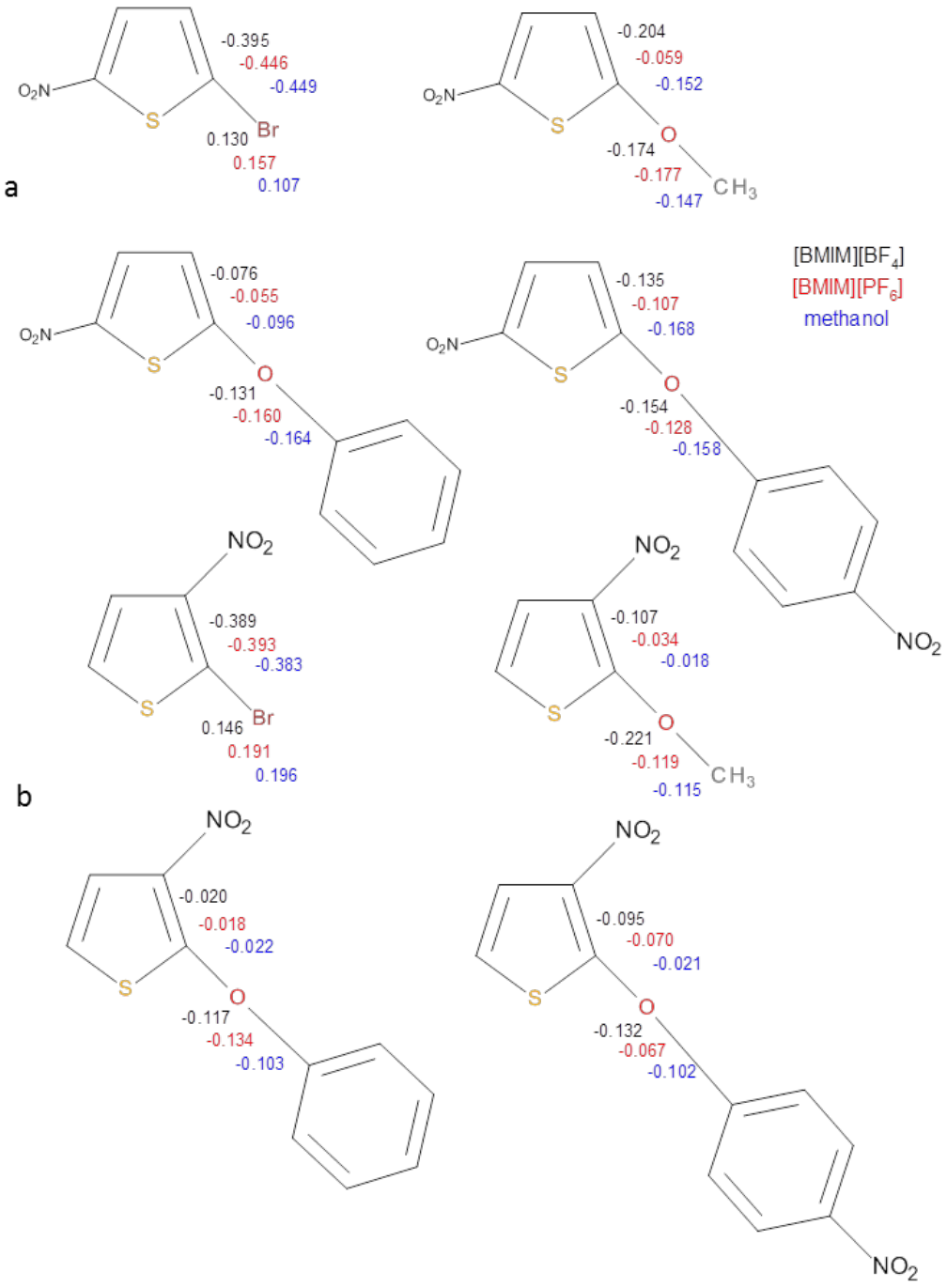


Figure 4.8: Selected atomic charges (e units) for the (a) para substituted nitrothiophenes and the (b) ortho substituted nitrothiophenes in [BMIM][BF_4] (black), [BMIM][PF_6] (red) and methanol (blue) from the QM/MM/MC calculations.

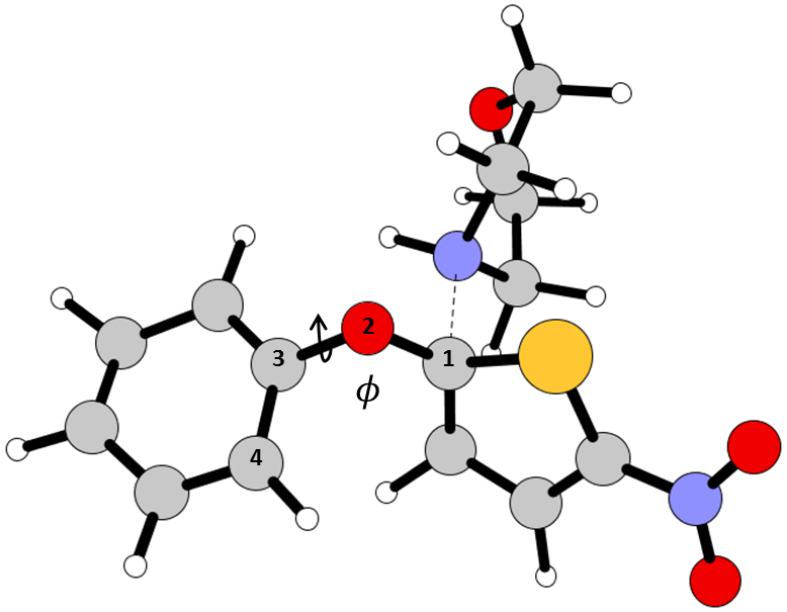


Figure 4.9: Snapshot of the addition step transition state between the substrates morpholine and 2-phenoxide-5-nitrothiophene from the QM/MM/MC calculations. ϕ = C1-O2-C3-C4. $\Phi = (360^{\circ} - \phi)$

Nucleophile	Thiophene Derivative	Leaving Group	Methanol Φ (degrees)	[BMIM][BF_4] Φ (degrees)	[BMIM][PF_6] Φ (degrees)
Piperidine	ortho	OPh	72.0	48.0	46.5
		$OPhNO_2$	37.0	76.9	39.4
	para	OPh	159.5	199.2	207.5
		$OPhNO_2$	166.4	206.4	197.5

Table 4.4: Average dihedral angles (degrees, see Figure 4.10 for definition of Φ) for the addition transition structures of piperidine with 2-L-5- and 2-L-3-nitrothiophene at 25°C in methanol, [BMIM][BF_4], and [BMIM][PF_6]. Angles averaged over final 5 million configurations or 20 million configurations for methanol and the RTIL respectively.

4.5.3 Transition State Evolution

The structure of the activated complex is of great importance. Transition states are transient in nature and therefore generally experimental characterization cannot be done directly. The Hammond Postulate[24] states that the transition state most resembles the adjacent reactant, intermediate, or product that it is closest in energy to. The Hammond postulate incorporates that molecules generally do not undergo rapid, discontinuous structural changes along the reaction coordinate, but rather that the geometrical changes are smooth and continuous. Examining the QM/MM/MC calculated free energy differences, $\Delta G^{\ddagger}_{reverse}$, between the transition state and the intermediate as a result of the addition step, see Table 4.4, it is clear that the transition state in RTILs is considerably closer in energy to the intermediate. Therefore the addition step transition state resembles the structure of the MIC more than for the corresponding reaction in methanol. For instance, the reaction between piperidine and 2-phenoxide-3-nitrothiophene in methanol has a calculated $\Delta G^{\ddagger}_{\text{reverse}}$ of 14.3 kcal/mol, whereas in [BMIM][BF_4] and [BMIM][PF_6] the energies between the transition state and the intermediate are 4.2 and 2.9 kcal/mol respectively. Consequently, the RTIL addition step transition state will have more MIC charge separation characteristics with shorter R_{CN} distances than those bond distances observed in methanol (Table 4.5). For example, the reaction between morpholine and 2-(4-nitrophenoxide)-3-nitrothiophene in methanol has an R_{CN} bond distance of 2.12 Å in the transition state, but an R_{CN} bond distance of 1.92 Å and 1.96 Å is observed in [BMIM][BF_4] and [BMIM][PF_6] correspondingly.

The structural and energetic similarities of the transition state to the MIC in RTILs can be attributed to the RTILs ability to support the evolution of the charged

complex more readily than in methanol. The molecular dipole moment and dielectric constant are not suitable representations to understand the role of solvent molecules in a chemical reaction, because the property of an individual molecule does not necessarily describe the properties and thus behavior of a bulk solvent. The RTILs [BMIM][BF_4] and [BMIM][PF_6] have a dipolarity/polarizability Kamlet-Taft parameter, π^*, of 1.047[22] and 1.032[22] respectively and for methanol the value is 0.73.[22] An increase in the rate of reaction upon going from methanol to the RTILs follows this increase in the medium polarizability which can better stabilize transition structures that are more polar than the starting reactants. Welton et *al.* proposed a similar development when studying an aliphatic nucleophilic substitution reaction in the presence of neutral nucleophiles such as amines.[25] They suggested that the development of a charge separation in the transition state can be stabilized by the ionic liquid ions through favorable electrostatic interactions.

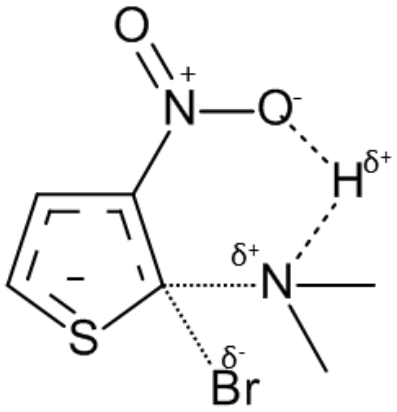
An intramolecular hydrogen bond between the nitro group of the thiophene and the ortho substituted amine (Scheme 4.4) in the transition state increases the stabilization and assists in the deprotonation of the amine. This favorable electrostatic interaction causes the elongation of R_{CL} in the MIC, see Table 4.5. Since the bromine atom induces the poorest electron withdrawing effect (see Figure 4.9) and in general has the smallest rates of reaction (see Table 4.1 and 4.2), this intramolecular hydrogen bond plays a vital role in the development of the MIC for these 2-bromo-3-nitrothiophene S_NAr reactions in methanol.

Nucleophile	Thiophene Derivative	Leaving Group	Methanol $\Delta G^{\ddagger}_{reverse}$ (kcal/mol)	[BMIM][BF_4] $\Delta G^{\ddagger}_{reverse}$ (kcal/mol)	[BMIM][PF_6] $\Delta G^{\ddagger}_{reverse}$ (kcal/mol)
Morpholine	ortho	Br	15.6	9.5	10.3
		OMe	8.2	2.8	1.9
		OPh	12.4	3.2	3.7
		$OPhNO_2$	16.3	5.6	4.2
	para	Br	15.9	10.4	9.6
		OMe	10.7	1.9	1.2
		OPh	10.3	1.4	1.6
		$OPhNO_2$	16.3	0.97	1.97
Piperidine	ortho	Br	18.3	11.0	10.8
		OMe	10.2	3.3	4.4
		OPh	14.3	4.2	2.9
		$OPhNO_2$	17.3	5.2	4.5
	para	Br	18.5	9.2	11.1
		OMe	12.1	5.2	3.0
		OPh	13.9	3.1	4.4
		$OPhNO_2$	12.3	3.1	1.4
Pyrrolidine	ortho	Br	17.7	12.1	11.2
		OMe	16.6	8.6	4.4
		OPh	16.8	8.6	4.2
		$OPhNO_2$	16.9	7.6	6.9
	para	Br	19.9	8.5	10.9
		OMe	12.8	3.3	1.4
		OPh	14.0	5.9	5.7
		$OPhNO_2$	16.3	4.9	7.4

Table 4.5: The QM/MM/MC $\Delta G^{\ddagger}_{reverse}$ of the S_NAr reactions in methanol, [BMIM][BF_4] and [BMIM][PF_6].

Nucleophile	Thiophene Derivative	Leaving Group	Methanol TS R_{CN} (Å)	[BMIM][BF_4] TS R_{CN} (Å)	[BMIM][PF_6] TS R_{CN} (Å)
Morpholine	ortho	Br	2.00	2.00	1.94
		OMe	2.06	1.92	1.84
		OPh	2.06	1.92	1.88
		$OPhNO_2$	2.12	1.92	1.96
	para	Br	2.00	1.92	1.94
		OMe	2.00	1.88	1.80
		OPh	1.96	1.88	1.80
		$OPhNO_2$	2.12	1.76	1.92
Piperidine	ortho	Br	2.07	2.00	2.04
		OMe	2.06	1.94	1.98
		OPh	2.10	1.96	1.92
		$OPhNO_2$	2.10	2.04	1.96
	para	Br	1.99	1.82	1.92
		OMe	1.97	1.96	1.92
		OPh	1.99	1.88	1.94
		$OPhNO_2$	2.00	1.88	1.80
Pyrrolidine	ortho	Br	2.04	2.00	1.94
		OMe	2.06	2.02	1.84
		OPh	2.12	1.92	1.96
		$OPhNO_2$	2.12	2.08	2.02
	para	Br	2.00	1.94	1.92
		OMe	2.03	1.80	1.72
		OPh	2.04	1.98	1.96
		$OPhNO_2$	2.02	1.94	1.96

Table 4.6: The QM/MM/MC calculated R_{CN} bond distances of the addition step's transition state in methanol, [BMIM][BF_4] and [BMIM][PF_6].



Scheme 4.5: The formation of an intramolecular hydrogen bond in the transition state during the reaction of piperidine, pyrrolidine or morpholine and 2-bromo-3-nitrothiophene in methanol.

Nucleophile	Thiophene Derivative	Leaving Group	Methanol MIC R_{CN}/R_{CL} (Å)	[BMIM][BF_4] MIC R_{CN}/R_{CL} (Å)	[BMIM][PF_6] MIC R_{CN}/R_{CL} (Å)
Morpholine	ortho	Br	1.53/1.98	1.54/1.81	1.54/1.88
Piperidine	ortho	Br	1.53/2.90	1.52/1.97	1.52/1.90
Pyrrolidine	ortho	Br	1.52/2.87	1.52/1.91	1.54/1.92

Table 4.7: The QM/MM/MC calculated bond lengths for R_{CN} and R_{CL} of the MIC for the S_NAr reaction between morpholine, piperidine and pyrrolidine with 2-bromo-3-nitrothiophene in methanol, [BMIM][BF_4] and [BMIM][PF_6].

4.5.4 Solute-Solvent Interactions

To further understand the rate accelerations observed for the S_NAr reactions when the environment is changed from the conventional solvent methanol to the RTILs, the interaction energies for the solvents were computed by analyzing the solute-solvent energy pair distributions from QM/MM/MC calculations in the representative FEP window near the corresponding ground state (GS) and transition state. The energy pair distributions determine the average number of methanol molecules, or ions of the RTIL that interact with the solute and the corresponding energies of these interactions. Interactions between the solute and solvent that are highly favorable are reflected in the leftmost region with energies more attractive than -5 kcal/mol (Figures 4.11 – 4.14). The large band at or near 0 kcal/mol is a result of many methanol molecules or RTIL ions in the outermost solvation shell of the system.

The S_NAr reactions generally have flatter energy distributions for the ground state when compared to the corresponding transition state in both methanol and the RTILs (Table 4.7). For instance, in the reaction between piperidine and 2-methoxy-5-nitrothiophene, an integration of the distributions from -20.0 to -5.0 (or -3.5) kcal/mol shows more favorable solute-solvent interactions occurring at the TS: 3.02 (5.11) in methanol, 4.48 (7.29) in [BMIM][BF_4] and 4.36 (5.26) in [BMIM][PF_6], than at the GS: 2.83 (4.61) in methanol, 3.75 (5.86), and 2.96 (5.41) in [BMIM][BF_4] and [BMIM][PF_6] respectively. When the leaving group is phenoxide or 4-nitrophenoxide, the integration of the energy pair distributions shows approximately equivalent solute-solvent interactions occurring at the GS and TS in the RTILs (Table 4.7). This observed behavior could be related to the large resonance interactions occurring between the leaving group and the

nitro group of the thiophene, which would lower the energy of reorganization occurring in the RTIL for the development of the TS.[1] This behavior suggests that the higher reactivity in the ionic liquids can in part be due to π-π interactions between the substrate and imidazolium cation in the RTIL. Additionally, the S_NAr reactions in RTILs have weaker energy distributions for the GS when compared to the transition state. In methanol however, the average energy distributions for the GS and TS are approximately equal in magnitude. This reinforces the findings that the GS cyclic amines are less solvated in RTILs than in methanol and that the evolution of the TS is more robust in the RTILs, contributing to the increased rates of reaction.

Figure 4.14 shows a snapshot of the exact nature of these most favorable solute-ion interactions of the MIC of the S_NAr reaction between piperidine and 2-methoxy-5-nitrothiophene in [BMIM][BF_4]. The amine proton is stabilized by a BF_4 anion, in addition to the leaving group methoxy being stabilized by a hydrogen bond with an acidic proton on the imidazolium ring. Interestingly, the methoxy group interacts with the BMIM cation's more sterically exposed hydrogens on the terminal butyl chain carbons, rather than the most acidic imidazolium proton at the 2 position (pk_a of 21 – 23).[26] The encapsulation of the first transition state and the MIC of the S_NAr reaction between piperidine and 2-methoxy5-nitrothiophene in [BMIM][BF_4] can be seen in Figure 4.15. In the transition state there are 10 ions that are within 3 Å of the substrate and in the MIC there are 14 ions that are within 3 Å of the substrate. Figure 4.15 illustrates how the [BMIM][BF_4] ions are forming a cage-like structure to favorably interact with the TS and the MIC. Furthermore, there is an increased degree of ordering in the cage-like liquid clathrate structure ongoing from the TS to the MIC, which can be attributed to the degree

of charge separation in the TS and MIC. For instance in the MIC, the positive amine is surrounded by a “layer” of BF_4^- counterions, that are further encased by a second “layer" of imidazolium cations. The formation of these structured solvent cages agrees with experimental reports of liquid clathrate formation in 1-alkyl-3-methylimidazolium-based ionic liquids with aromatic compounds.[27]

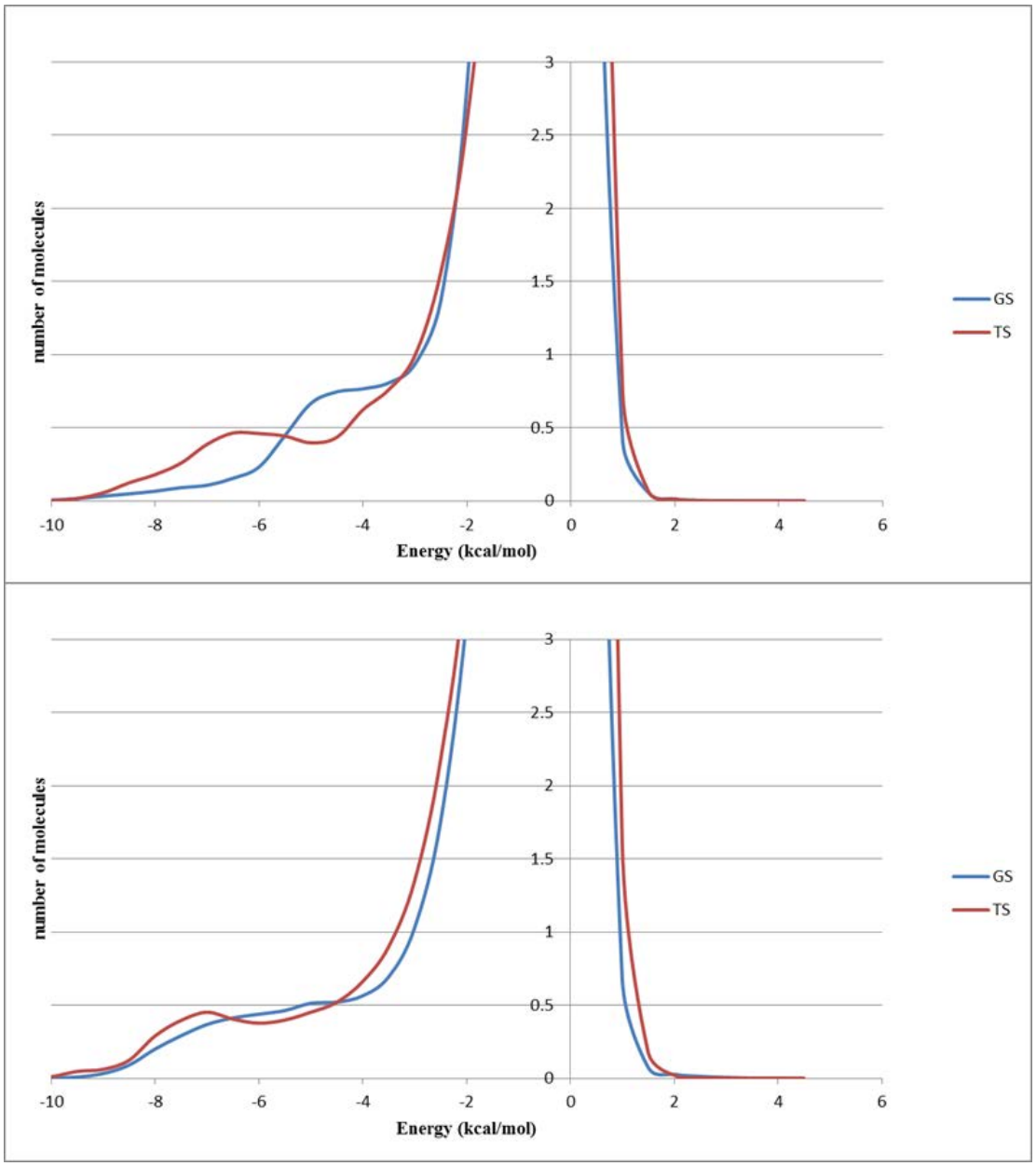


Figure 4.10: The solute-solvent energy pair distributions of the ground state (GS – blue line) and first transition state (TS – red line) for the S_NAr reactions between the nucleophile piperidine and 2-bromo-5-nitrothiophene (top) and 2-methoxy-5-nitrothiophene (bottom) in methanol.

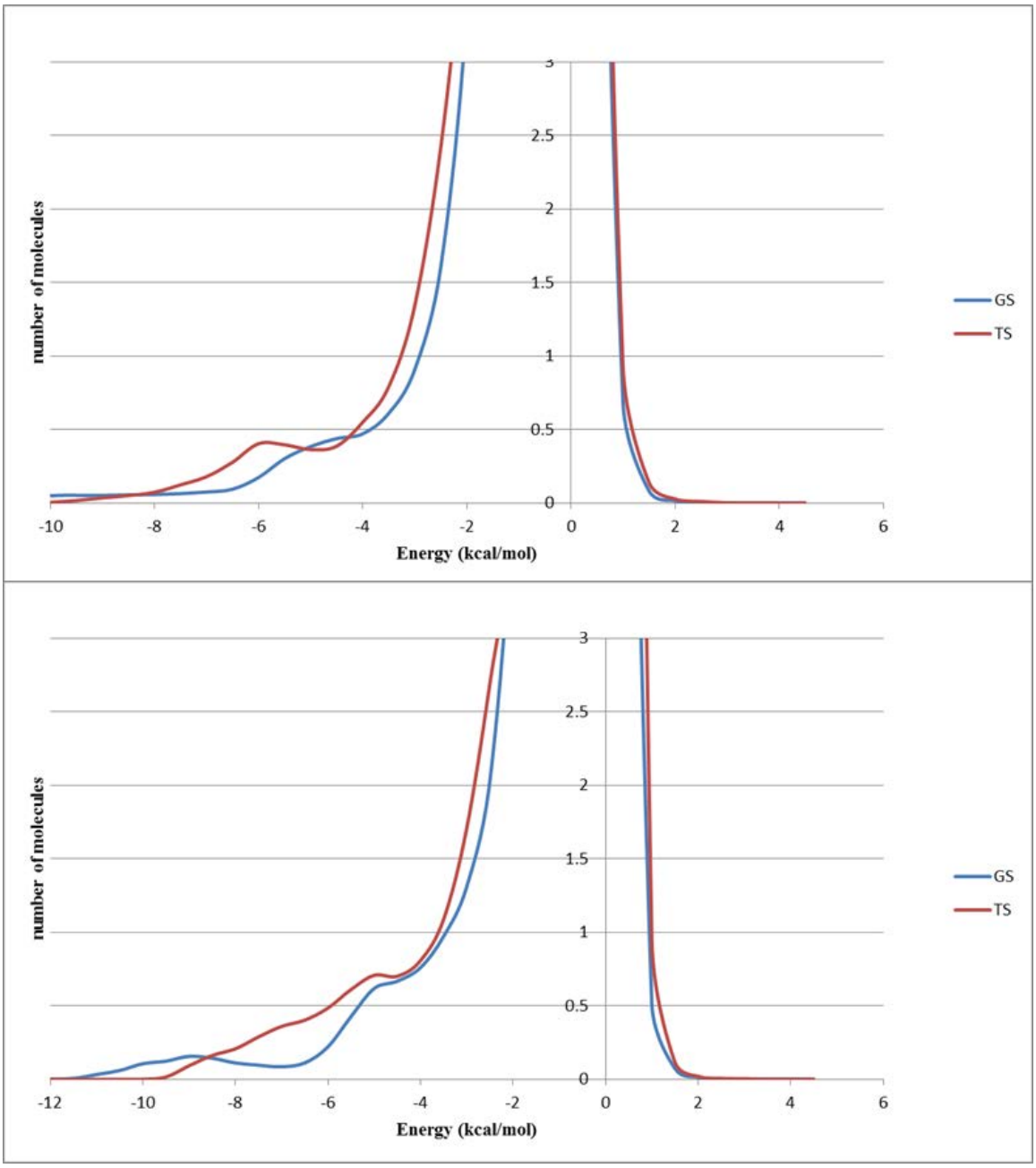


Figure 4.11: The solute-solvent energy pair distributions of the ground state (GS – blue line) and first transition state (TS – red line) for the S_NAr reactions between the nucleophile piperidine and 2-phenoxide-5-nitrothiophene (top) and 2-(4-nitrophenoxide)-5-nitrothiophene (bottom) in methanol.

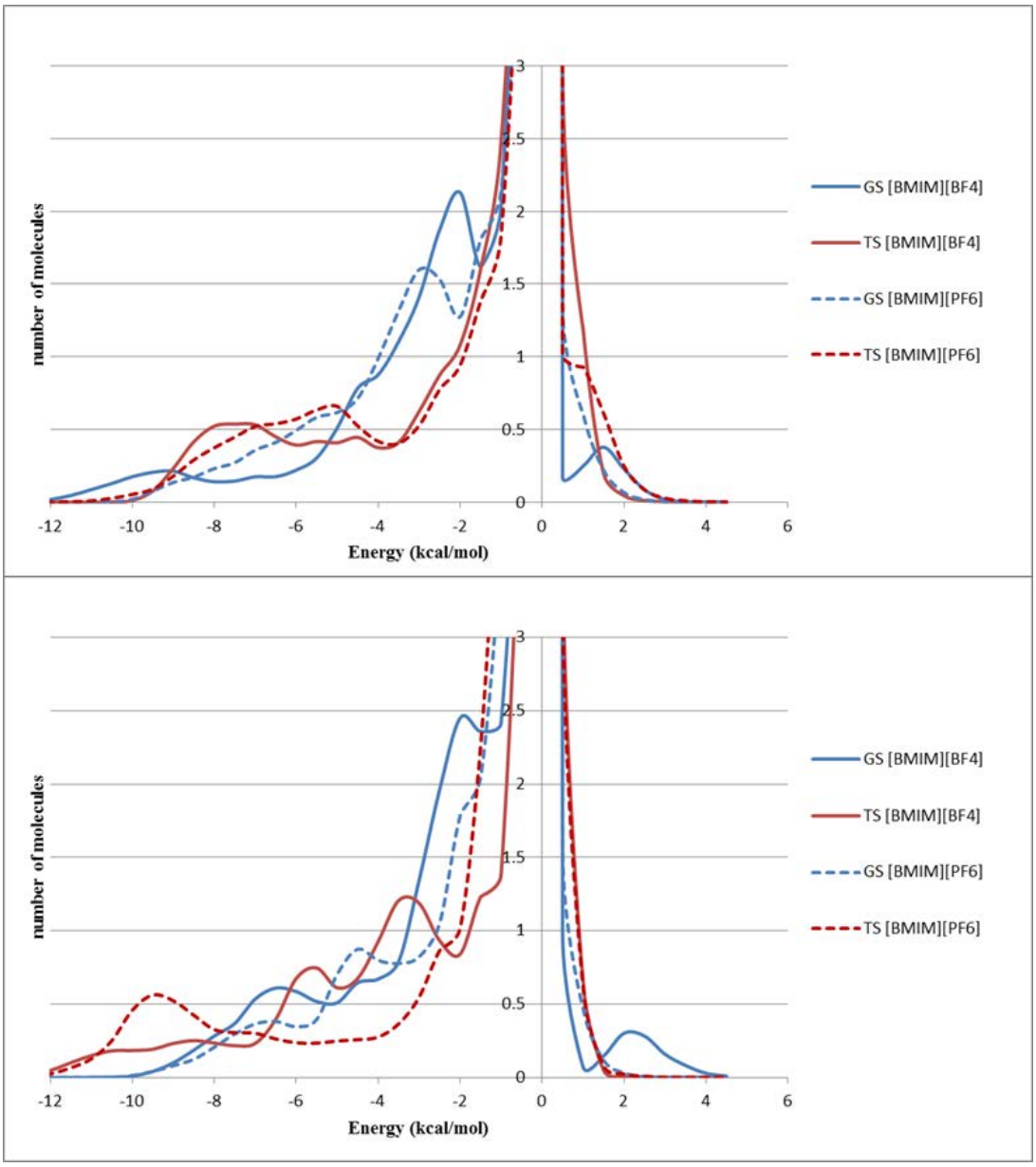


Figure 4.12: The solute-solvent energy pair distributions of the ground state (GS – blue line) and first transition state (TS – red line) for the S_NAr reactions between the nucleophile piperidine and 2-bromo-5-nitrothiophene (top) and 2-methoxy-5-nitrothiophene (bottom) in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

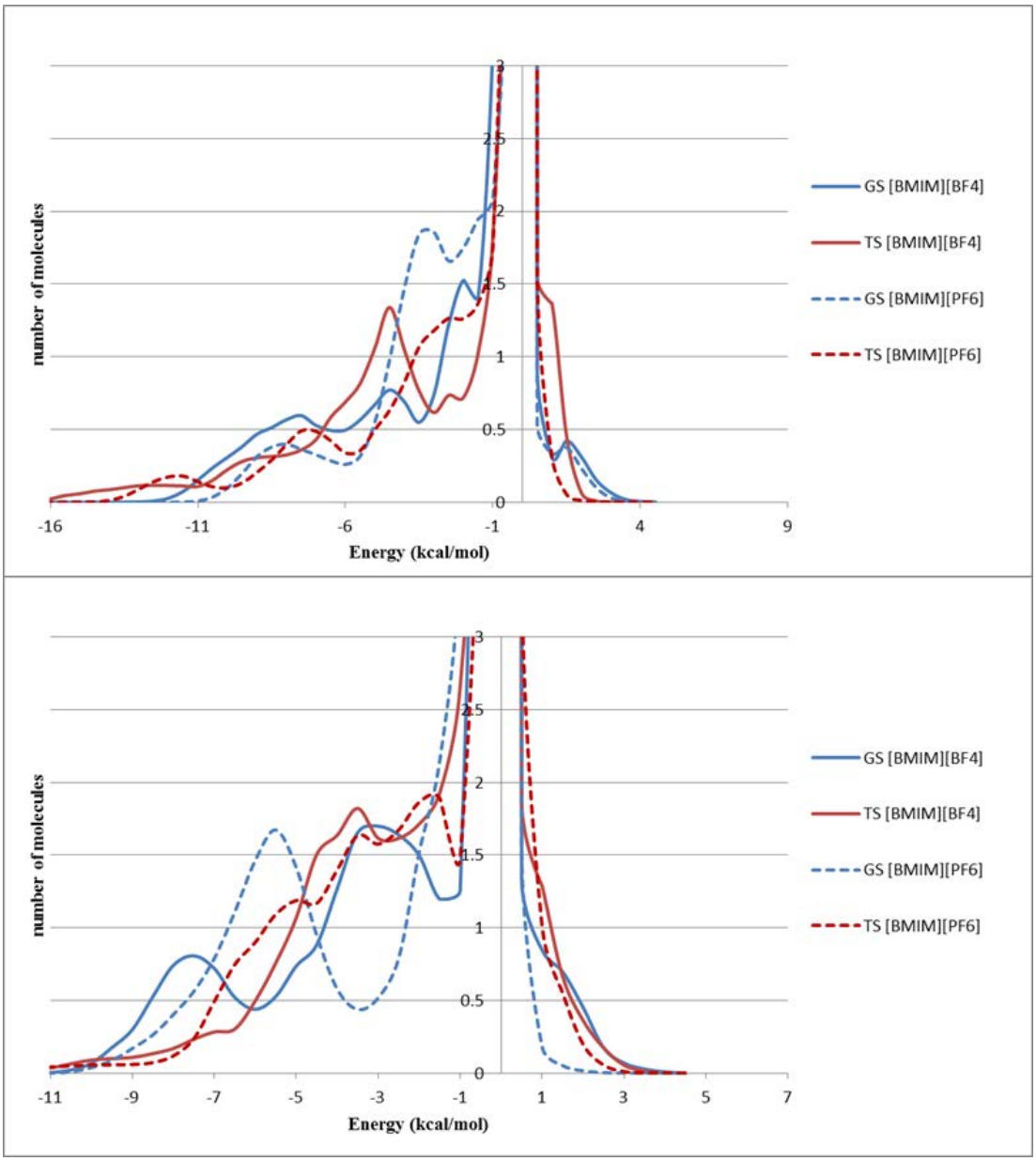


Figure 4.13: The solute-solvent energy pair distributions of the ground state (GS – blue line) and first transition state (TS – red line) for the S_NAr reactions between the nucleophile piperidine and 2-phenoxide-5-nitrothiophene (top) and 2-(4-nitrophenoxide)-5-nitrothiophene (bottom) in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

Solvent	Thiophene Derivative	Leaving Group	G.S.	TS
Methanol	para	Br	1.87 (4.19)	2.79 (4.61)
		OMe	2.83 (4.61)	3.02 (5.11)
		OPh	1.40 (2.92)	1.93 (3.66)
		$OPhNO_2$	2.30 (4.70)	3.34 (5.93)
[BMIM][BF_4]	para	Br	2.71 (5.48)	3.99 (5.22)
		OMe	3.75 (5.86)	4.48 (7.29)
		OPh	6.09 (8.10)	6.49 (9.64)
		$OPhNO_2$	5.59 (9.40)	4.81 (9.77)
[BMIM][PF_6]	para	Br	3.37 (6.39)	4.39 (5.74)
		OMe	2.96 (5.41)	4.36 (5.26)
		OPh	3.53 (7.84)	4.60 (7.13)
		$OPhNO_2$	7.99 (9.95)	5.25 (9.45)

Table 4.8: The solute-solvent energy pair distributions of the reactants (G.S.) and first transition state (TS) for the S_NAr reactions between the nucleophile piperidine and 2-L-5-nitrothiophene in methanol, [BMIM][BF_4], and [BMIM][PF_6] integrated to -5.0 kcal/mol (and -3.5 kcal/mol in parentheses).

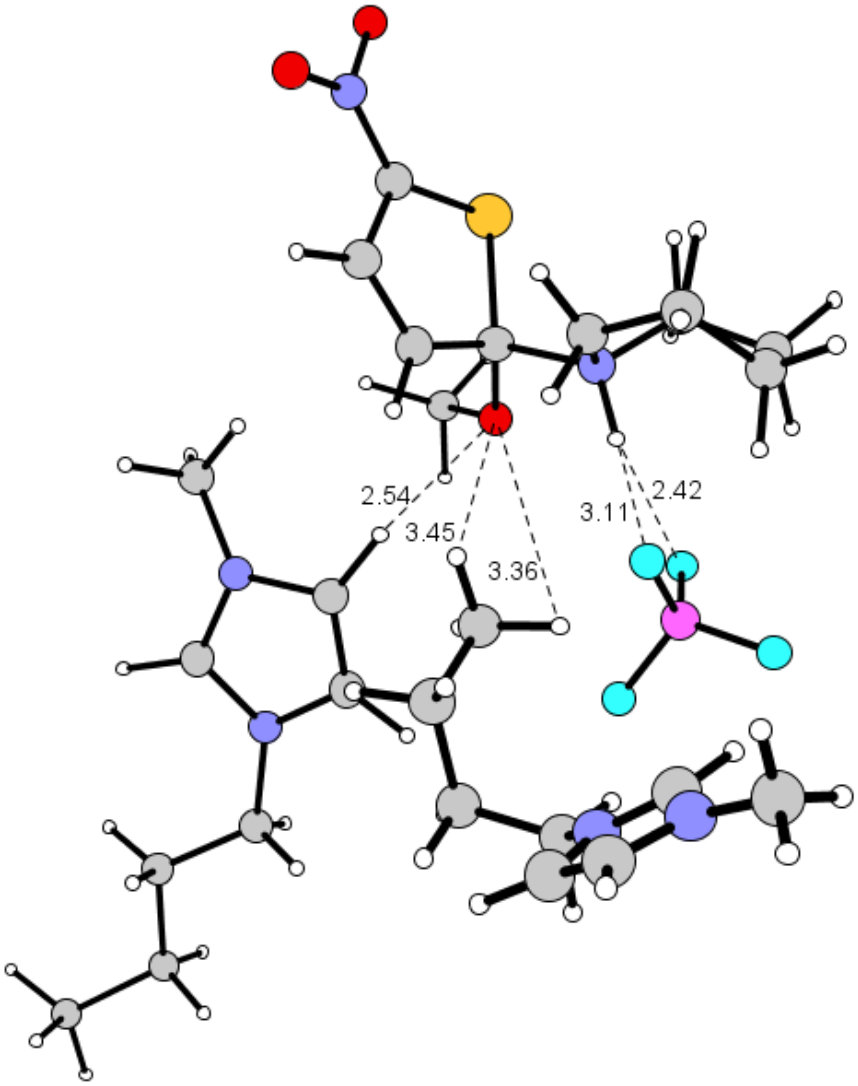


Figure 4.14: Typical snapshot of the MIC for the S_NAr reaction between piperidine and 2-methoxy-5-nitrothiophene in [BMIM][BF_4]. The distances (in Å) are average values over the final 20 million configurations of QM/MM/MC simulations. Only nearby ions are retained for clarity.

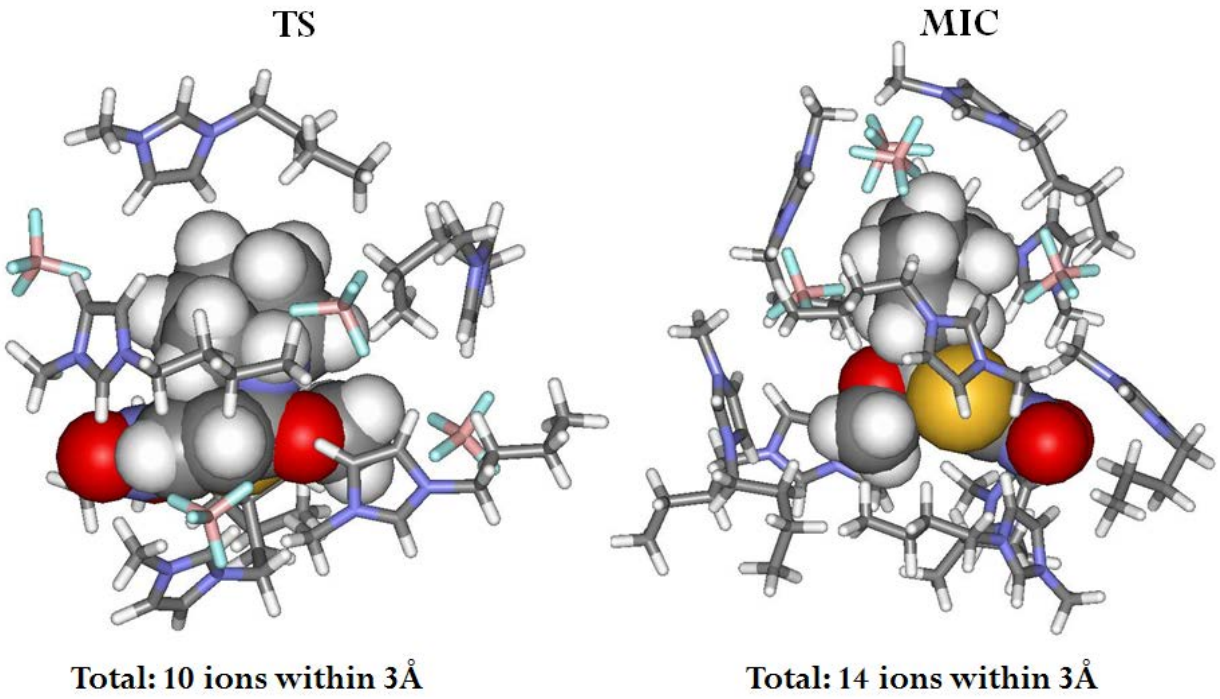


Figure 4.15: Illustration of the [BMIM][BF_4] (shown as sticks) ion encapsulation of the first transition state (left) and the MIC (right) of the S_NAr reaction between piperidine and 2-methoxy-5-nitrothiophene (given as CPK space-filling model).

4.6 Conclusions

QM/MM/MC calculations have been performed to help elucidate the origin of the ionic liquid effect on the reported enhanced rates of reaction between the nucleophiles: piperidine, pyrrolidine, and morpholine with 2-L-3- and 2-L-5-nitrothiophenes (L = bromine, methoxy, phenoxide and 4-nitrophenoxide). The enhanced rates of reaction observed in the RTILs [BMIM][BF_4] and [BMIM][PF_6] can be partly attributed to a less extensive solvation of the starting amine. In the RTILs and methanol the amine nucleophilicity was Pip $\geq$ Pyr > Mor. The decrease in solvation of the secondary amines in the RTILs can be related to the corresponding Kamlet-Taft parameters of the solvents.

The solvent methanol has a hydrogen bond acidity that is nearly double that of [BMIM][BF_4] or [BMIM][PF_6] and a hydrogen bond basicity that is over double that of the RTILs.[22] This was found to increase the number of favorable electrostatic interactions with the nucleophile that can form a solvent cage around the amine changing the effective nucleophilicity and thus decrease the corresponding rate of reaction. Examining the QM/MM/MC calculated free energies of activation for the reverse addition, $\Delta G^{\ddagger}_{reverse}$, it was clear that the transition state in RTILs is considerably closer in energy to the intermediate. The first transition state should therefore occur later along the reaction coordinate in RTILs than in methanol. Hence, the RTIL transition state will have greater charge separation characteristics with shorter R_{CN} distances than those bond distances observed in methanol. Through the analysis of selected atomic charges from the QM/MM/MC calculations, the higher electronegativity of the oxygen atom from the methoxy, phenoxide, and 4-nitrophenoxide leaving groups induced a larger partial positive charge density on the reacting carbon than bromine. The positive charge favors nucleophilic attack. An intramolecular hydrogen bond between the nitro group of the ortho-substituted thiophene and amine in the transition state increases the stabilization and assists in the deprotonation of the amine. This favorable electrostatic interaction causes the elongation of R_{CL} in the MIC. The interaction energies for the solvents were computed by analyzing the solute-solvent energy pair distributions at the ground state (GS) and transition state. The S_NAr reactions were found to have flatter energy distributions for the ground state when compared to the corresponding transition state in both methanol and the RTILs. When the leaving group is phenoxide or 4-nitrophenoxide, there are large resonance interactions occurring between the leaving group and the nitro

group of the thiophene. This behavior suggests that the higher reactivity in the ionic liquids can in part be due to π-π interactions between the substrate and imidazolium cation in the RTIL. The encapsulation of the first transition state and the MIC of the S_NAr reaction between piperidine and 2-methoxy-5-nitrothiophene in [BMIM][BF_4] illustrates how the [BMIM][BF_4] ions are forming a cage-like structure to favorably interact with the TS and the MIC. The increased degree of ordering in the cage-like structure ongoing from the TS to the MIC can be attributed to the degree of charge separation in the TS and MIC. The formation of these structured solvent cages agrees with experimental reports of liquid clathrate formation in 1-alkyl-3-methylimidazolium-based ionic liquids with aromatic compounds.[27] Through the deeper insight into the effect of the ionic liquids and their unique environment upon important organic reaction types, scientists can exploit this understanding to interpret and predict chemical phenomena for similar class reactions in RTILs.

4.7 Supplemental Information: Table of Contents

All Energy Pair Distributions of GS and MIC in:

All Free Energy Surfaces in Methanol

Piperidine + para-LG-nitrothiophene

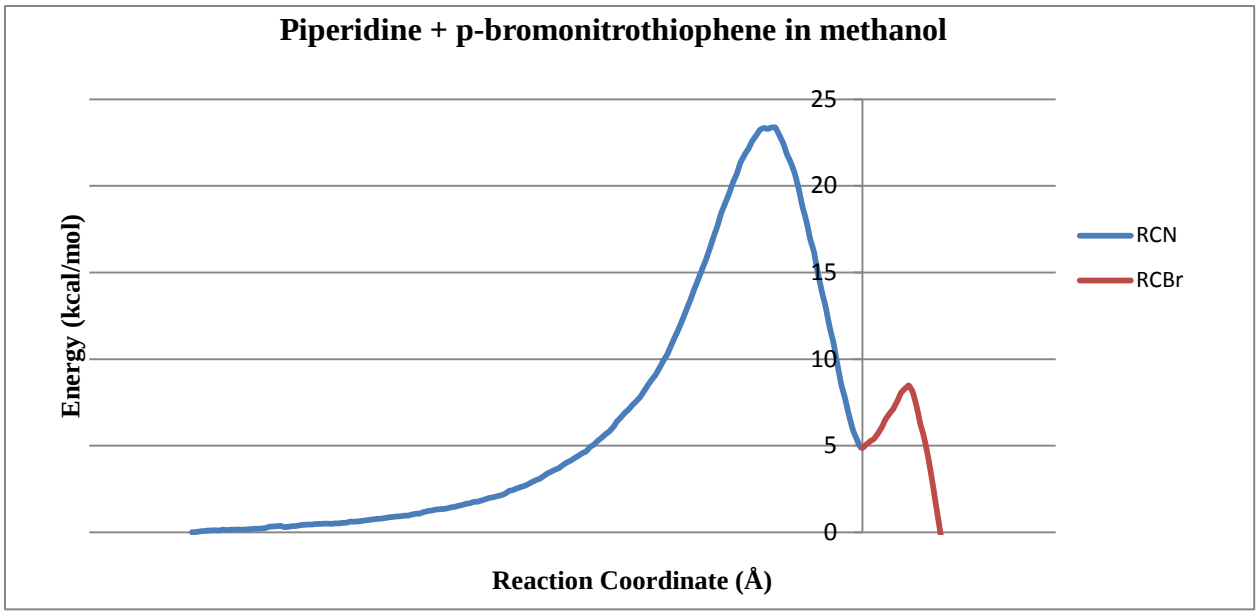


Figure 4S.1: The free energy diagram of the reaction Piperidine with para-bromonitrothiophene in methanol.

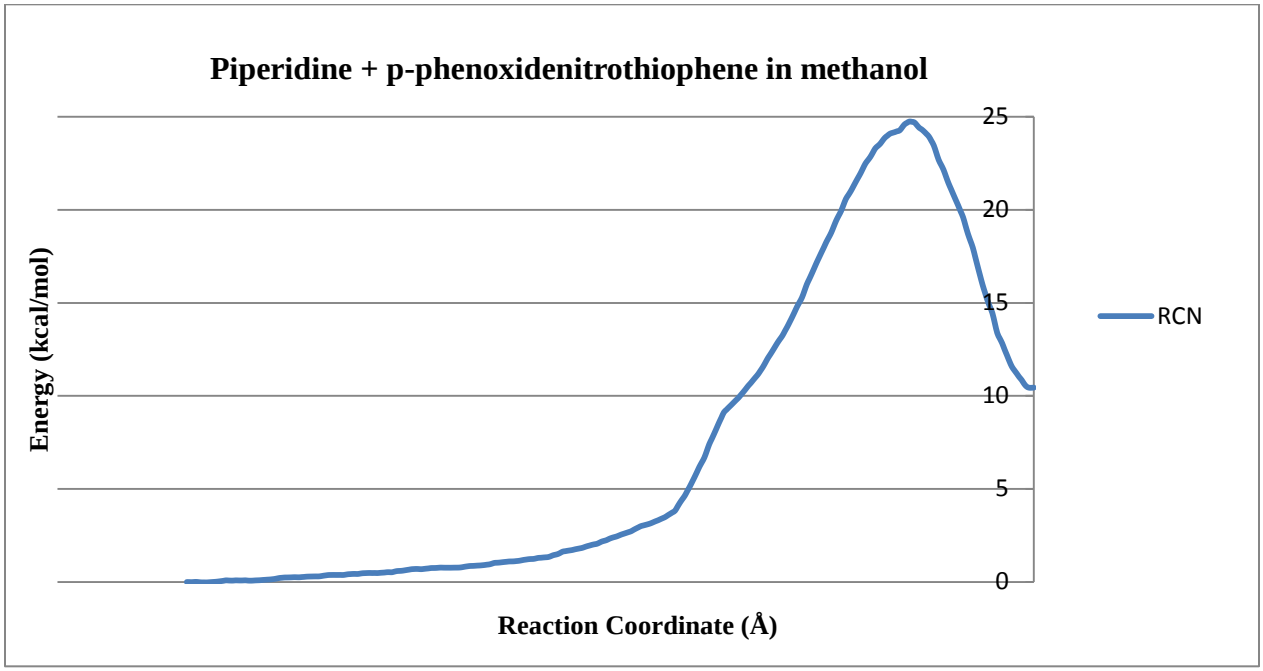


Figure 4S.2: The free energy diagram of the reaction Piperidine with para-phenoxidenitrothiophene in methanol

Piperidine + ortho-LG-nitrothiophene

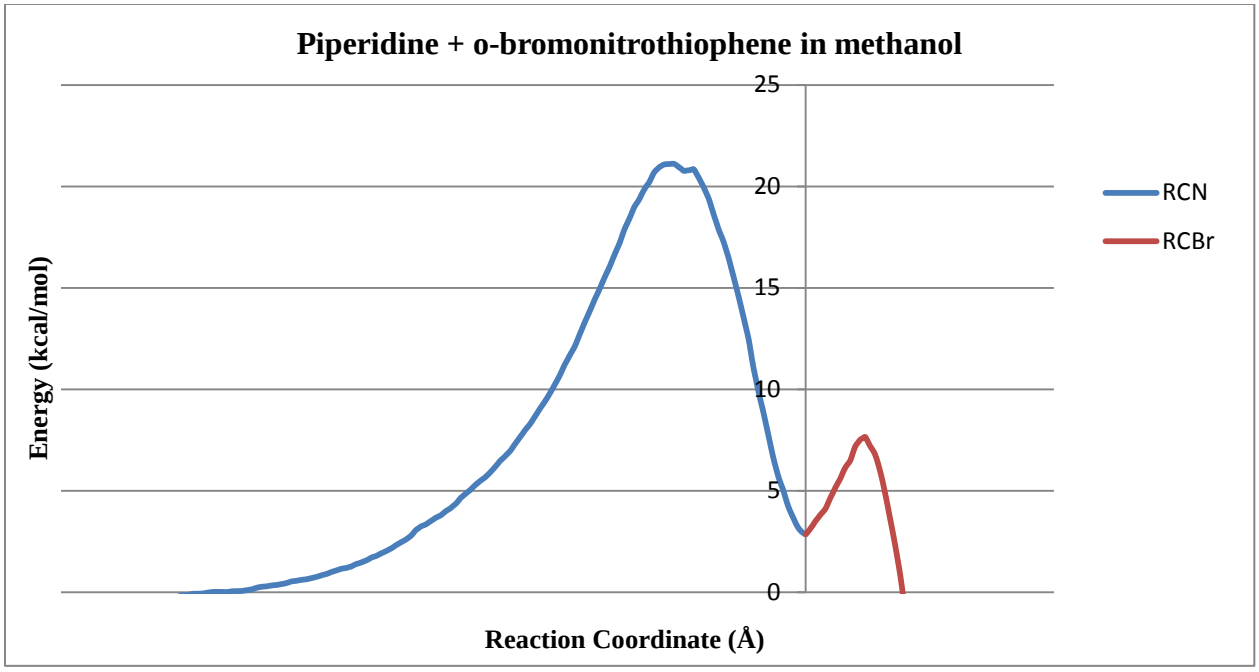


Figure 4S.3: The free energy diagram of the reaction Piperidine with ortho-bromonitrothiophene in methanol

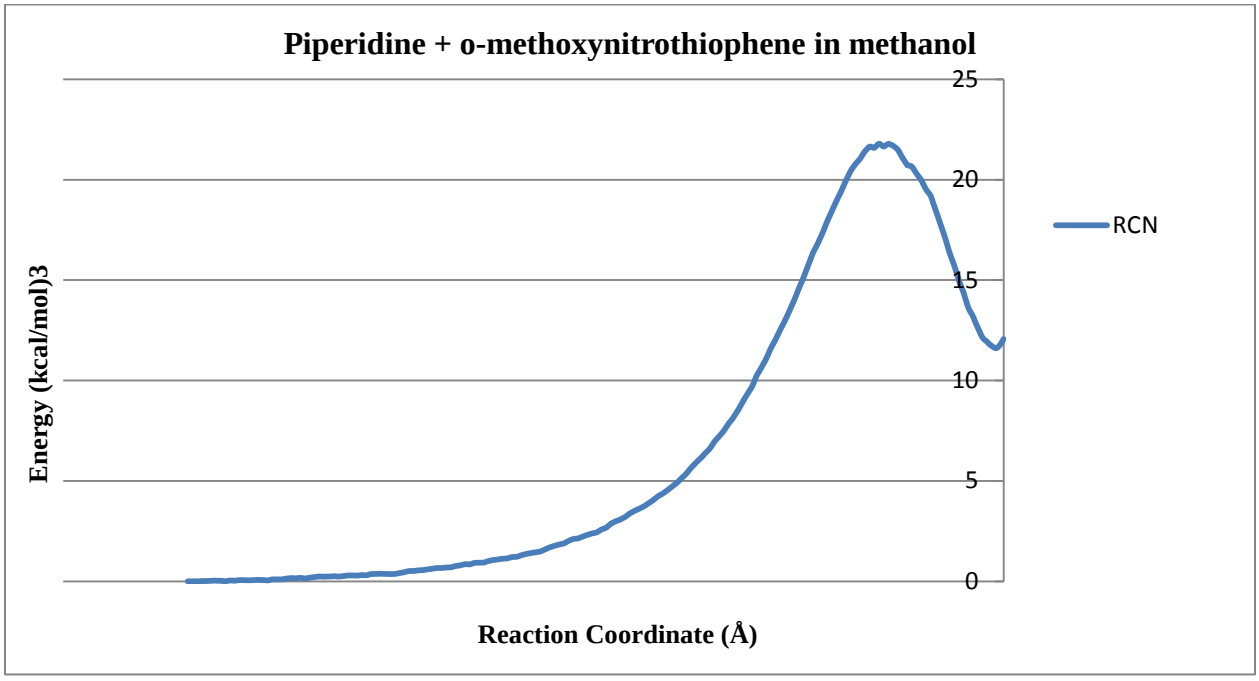


Figure 4S.4: The free energy diagram of the reaction Piperidine with ortho-methoxynitrothiophene in methanol

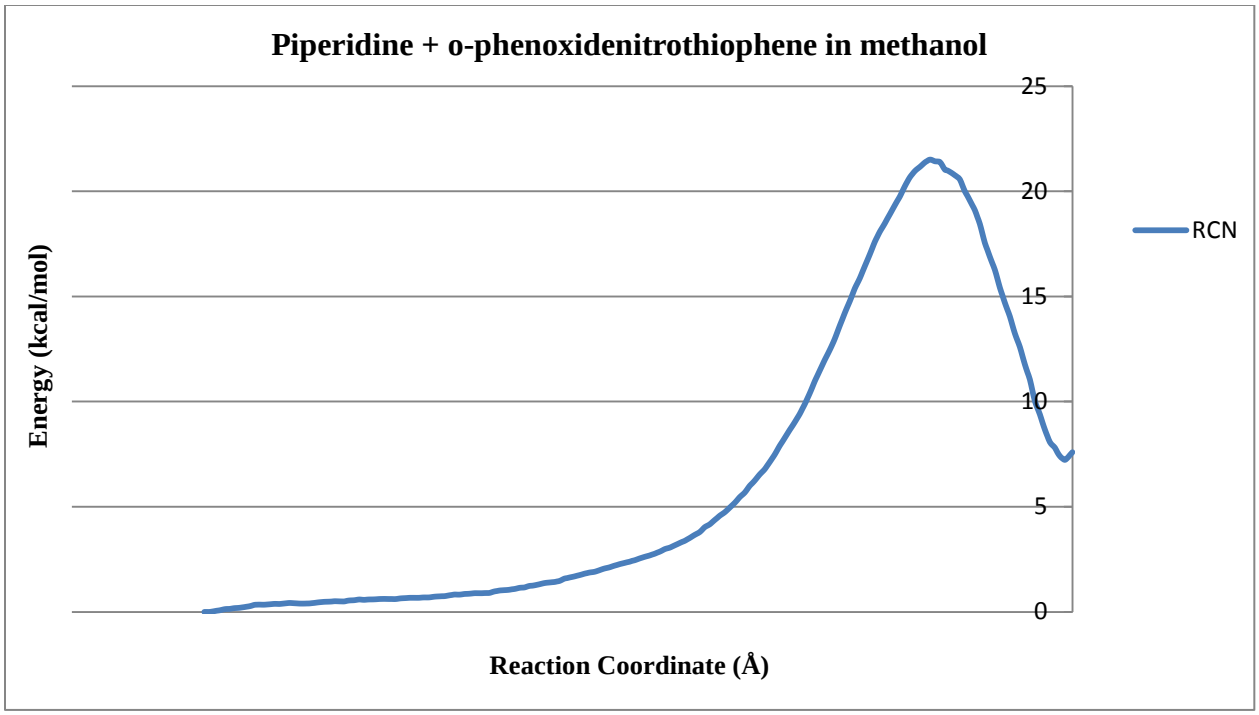


Figure 4S.5: The free energy diagram of the reaction Piperidine with ortho-phenoxidenitrothiophene in methanol

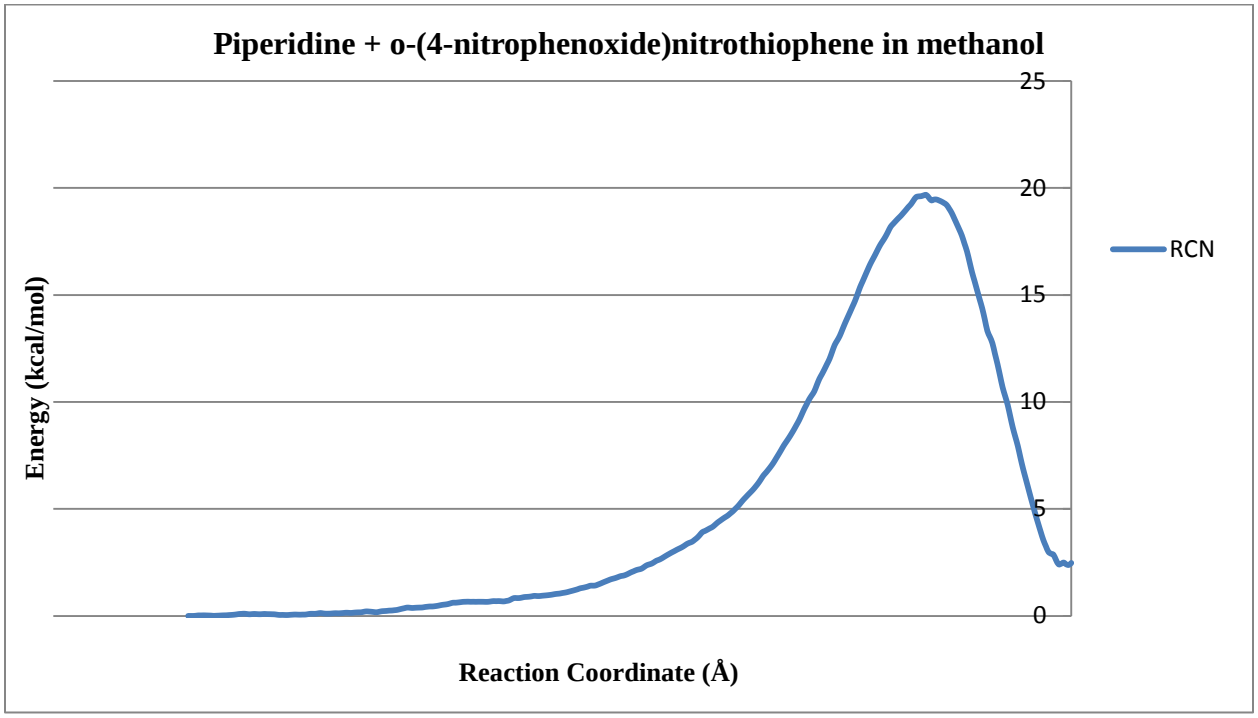


Figure 4S.6: The free energy diagram of the reaction Piperidine with ortho-(4-nitrophenoxide)nitrothiophene in methanol

Morpholine + para-LG-nitrothiophene

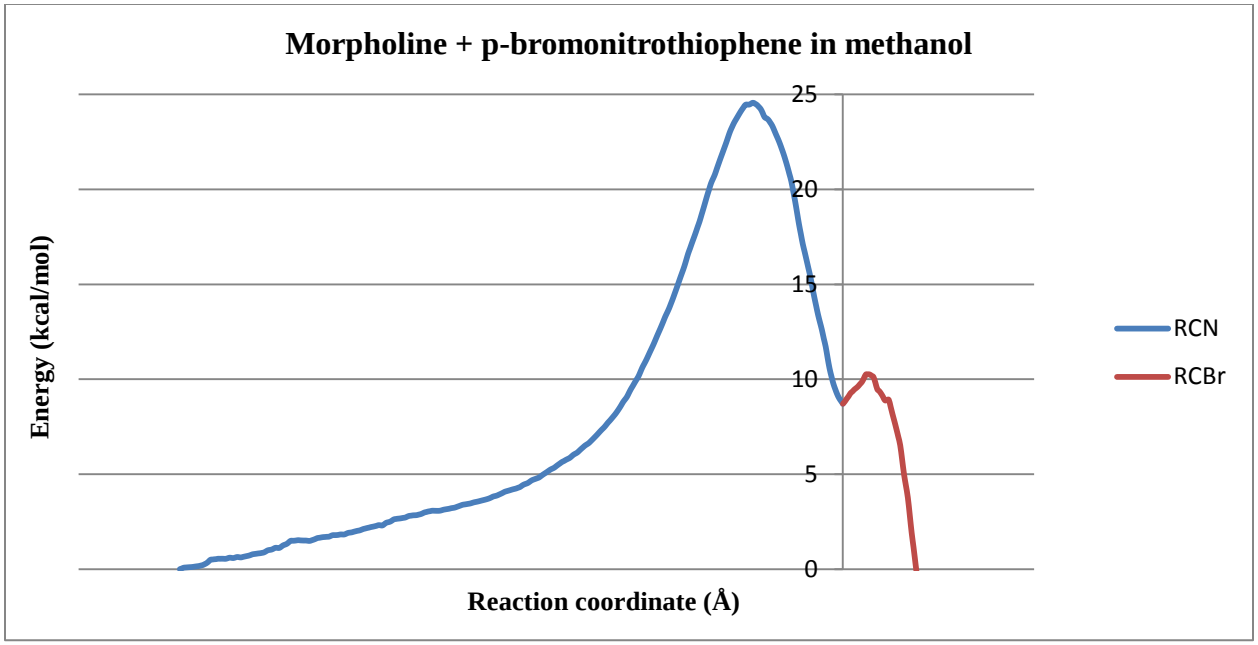


Figure 4S.7: The free energy diagram of the reaction Morpholine with para-bromonitrothiophene in methanol

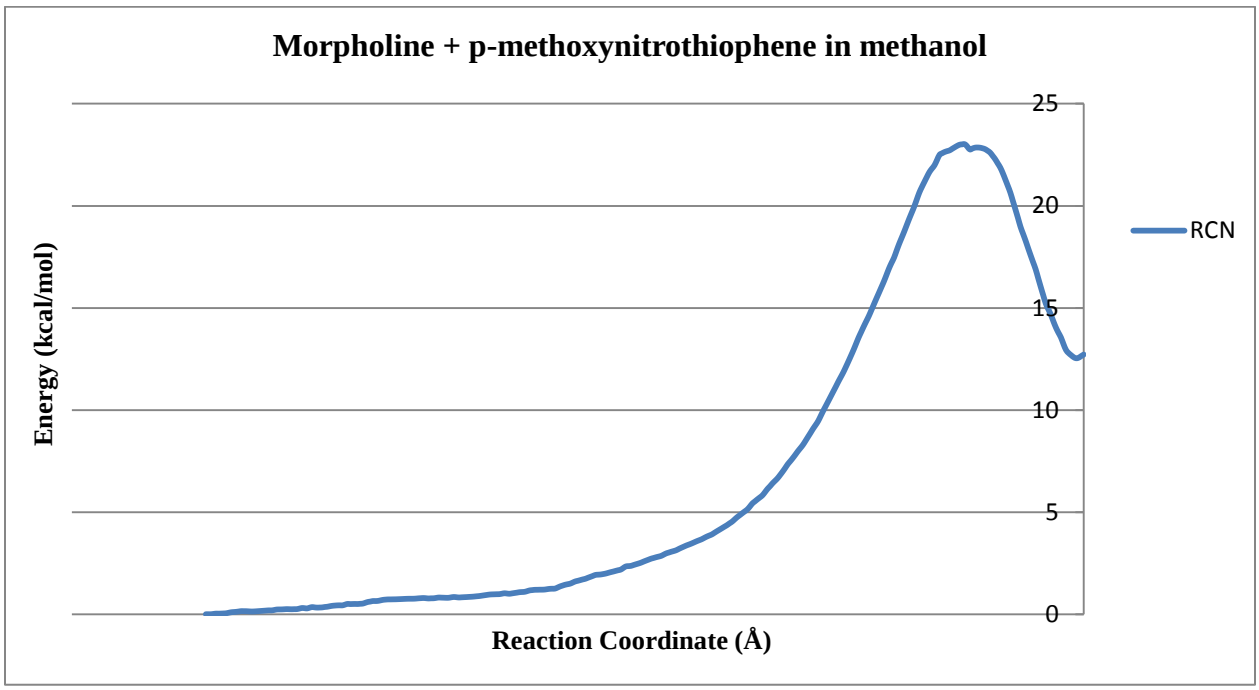


Figure 4S.8: The free energy diagram of the reaction Morpholine with para-methoxynitrothiophene in methanol

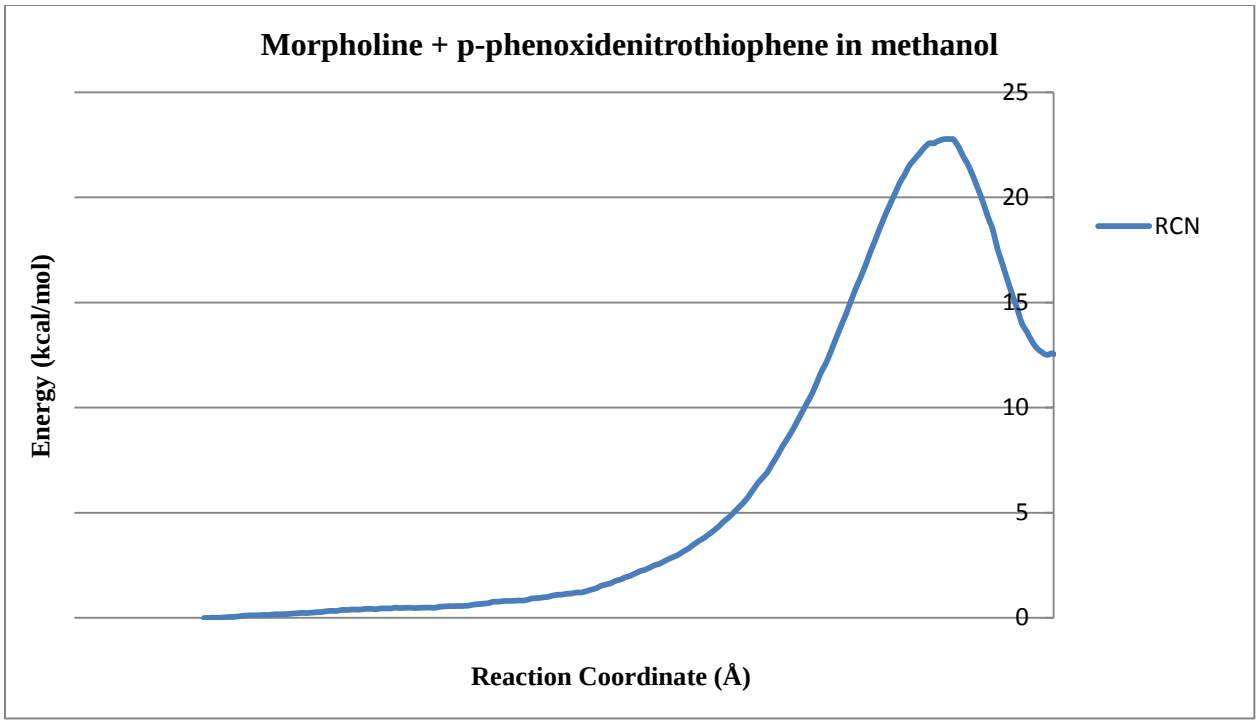


Figure 4S.9: The free energy diagram of the reaction Morpholine with para-phenoxidenitrothiophene in methanol

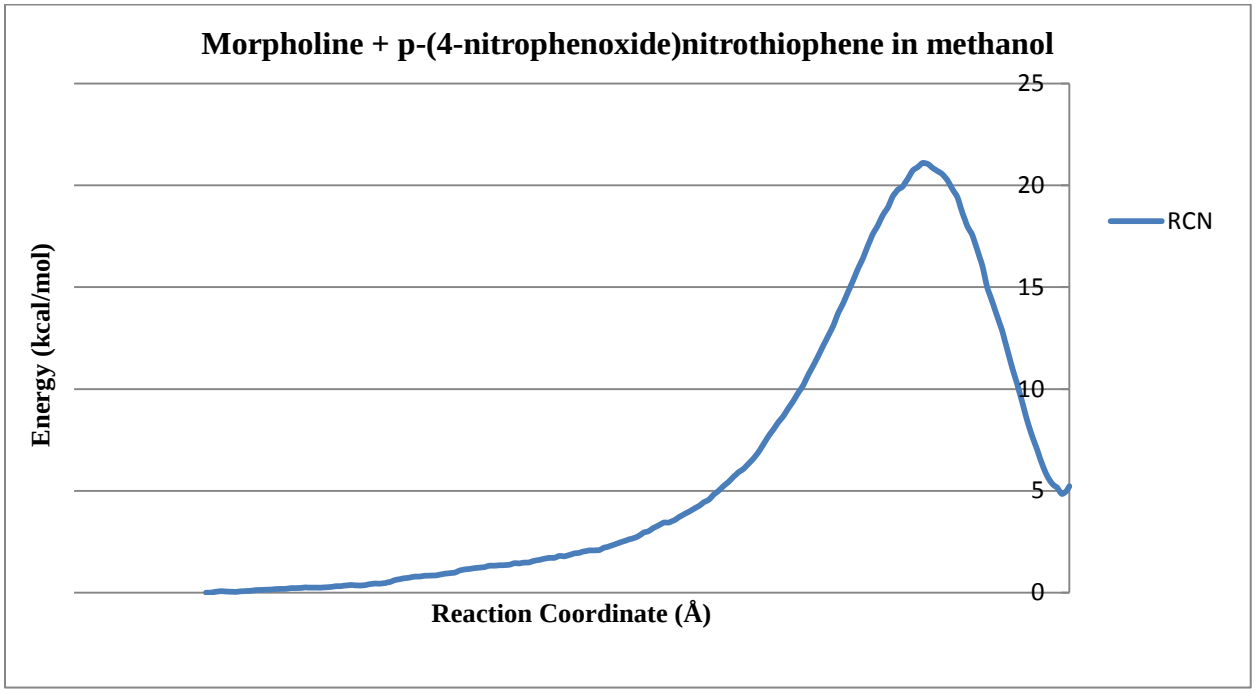


Figure 4S.10: The free energy diagram of the reaction Morpholine with para-(4-nitrophenoxide)nitrothiophene in methanol

Morpholine + ortho-LG-nitrothiophene

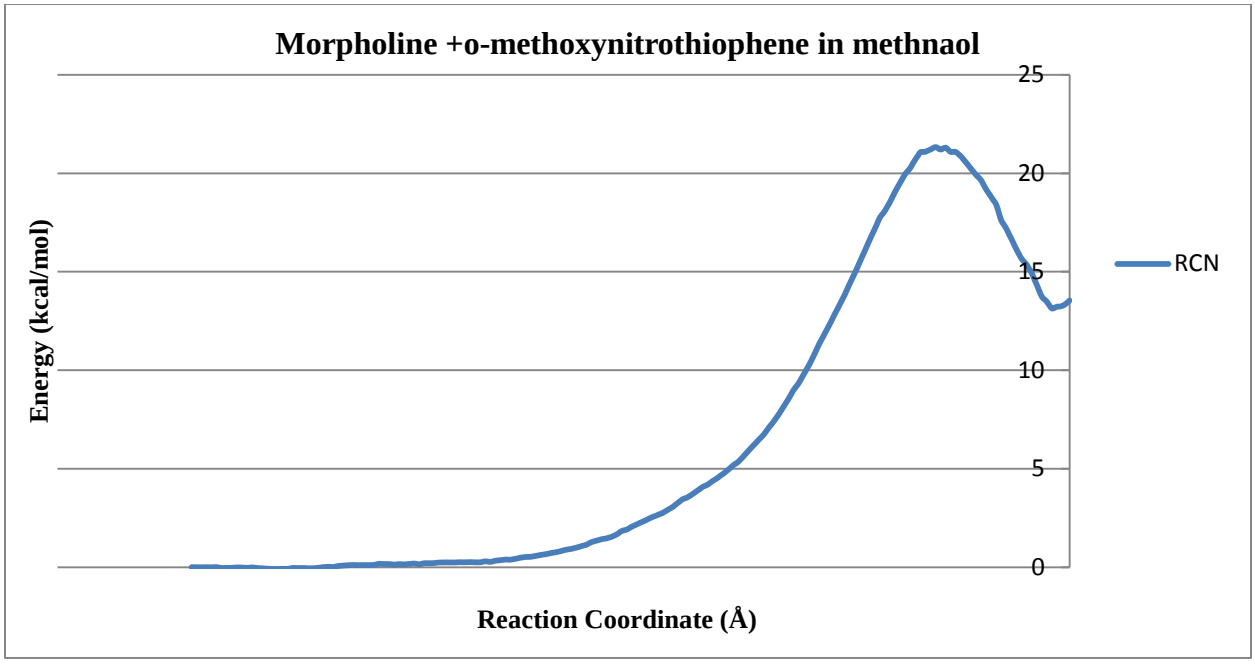


Figure 4S.11: The free energy diagram of the reaction Morpholine with ortho-methoxynitrothiophene in methanol

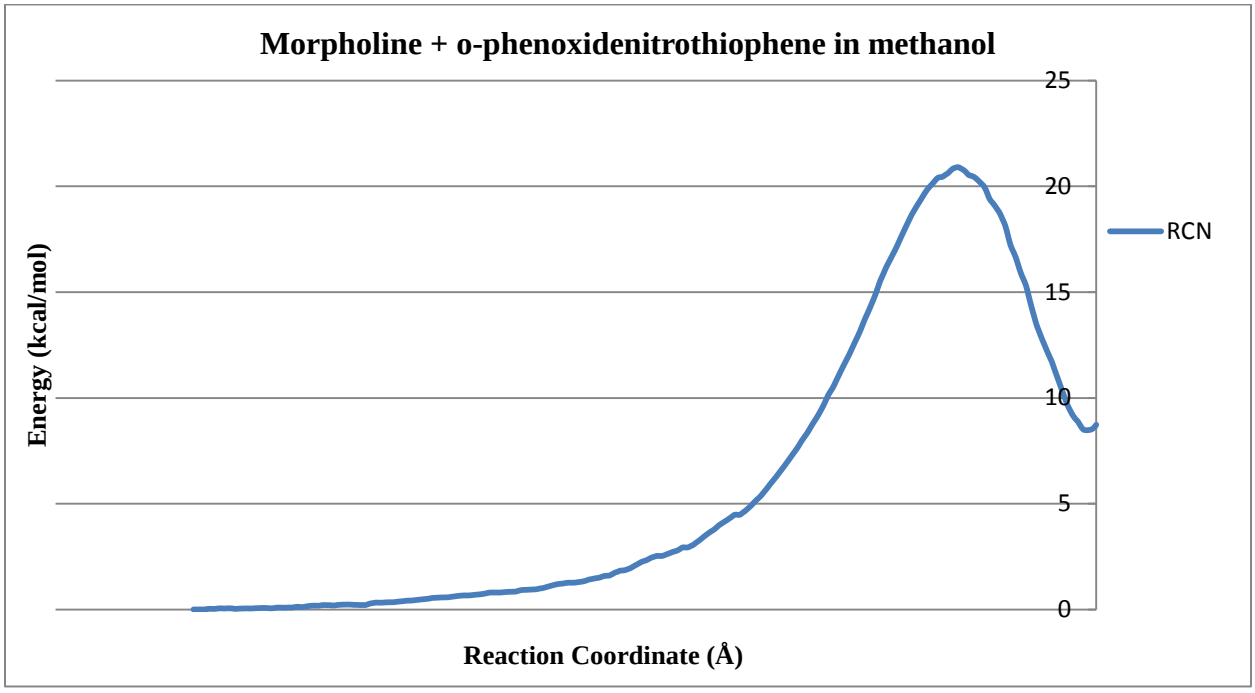


Figure 4S.12: The free energy diagram of the reaction Morpholine with ortho-phenoxidenitrothiophene in methanol

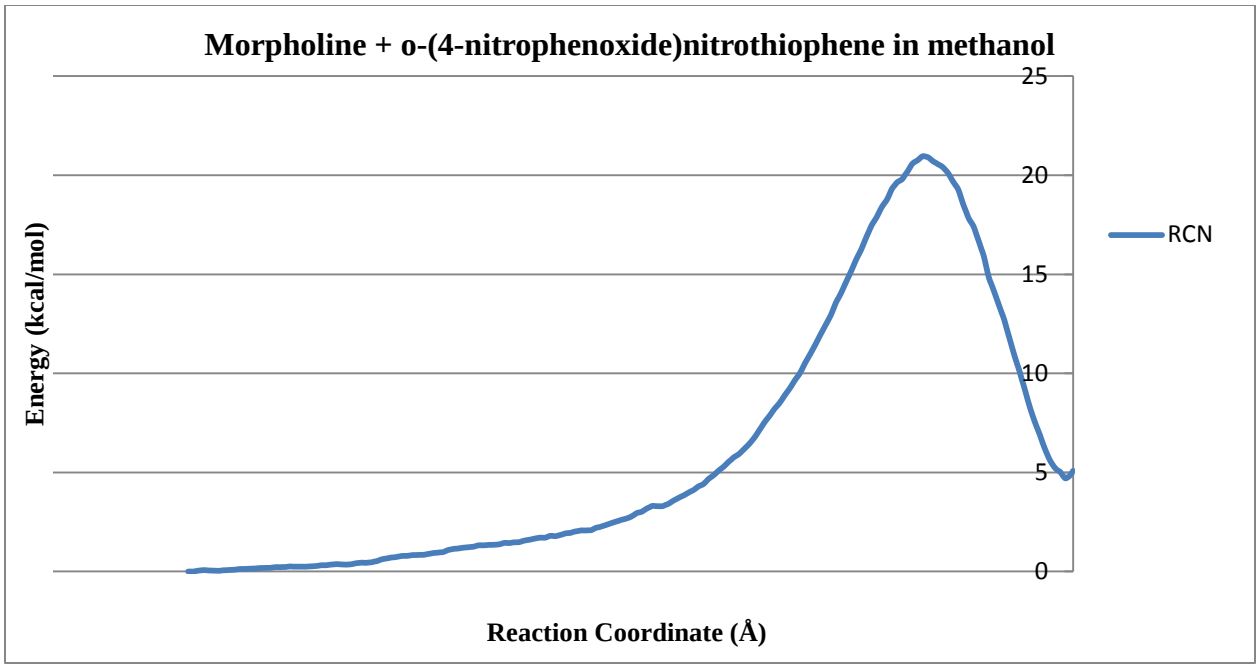


Figure 4S.13: The free energy diagram of the reaction Morpholine with ortho-(4-nitrophenoxide)nitrothiophene in methanol

Pyrrolidine + para-LG_nitrothiophene

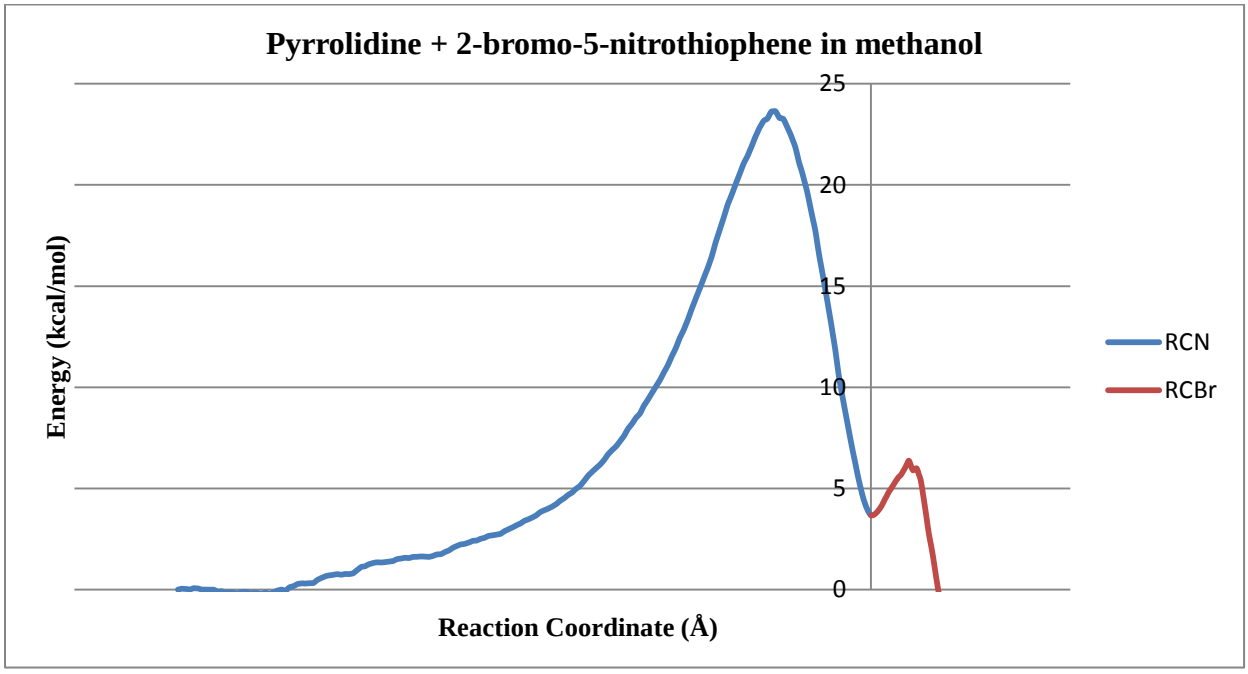


Figure 4S.14: The free energy diagram of the reaction Pyrrolidine with para-bromonitrothiophene in methanol

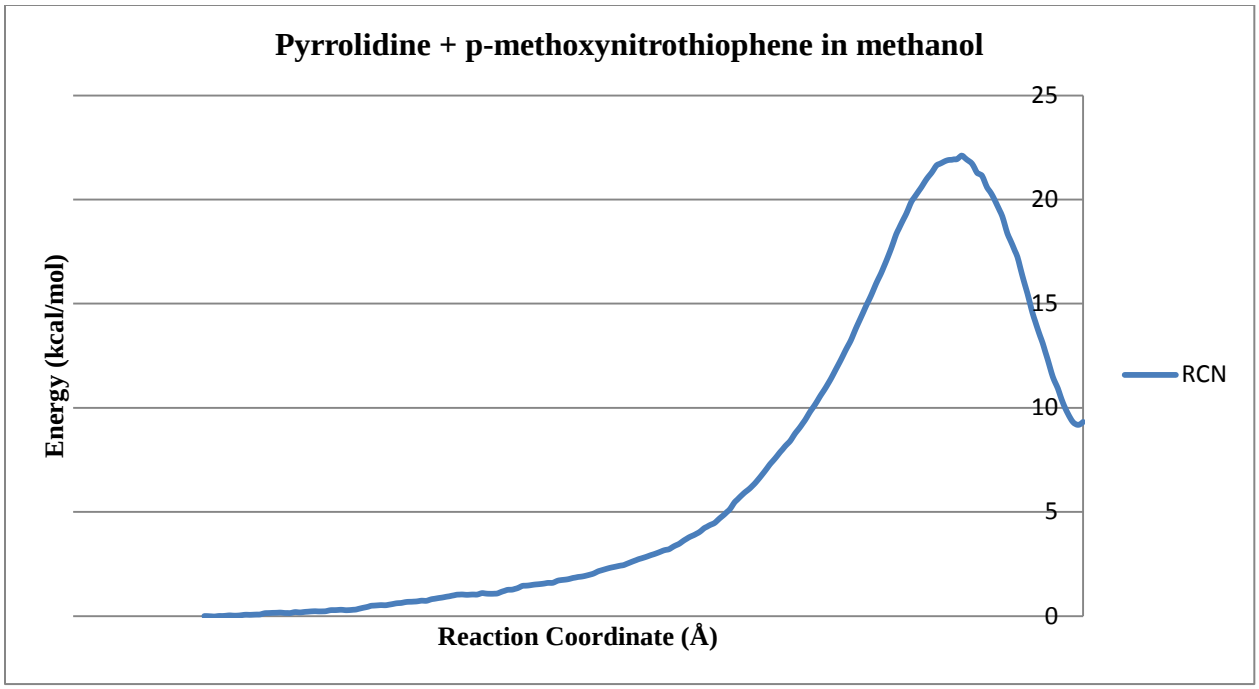


Figure 4S.15: The free energy diagram of the reaction Pyrrolidine with para-methoxynitrothiophene in methanol

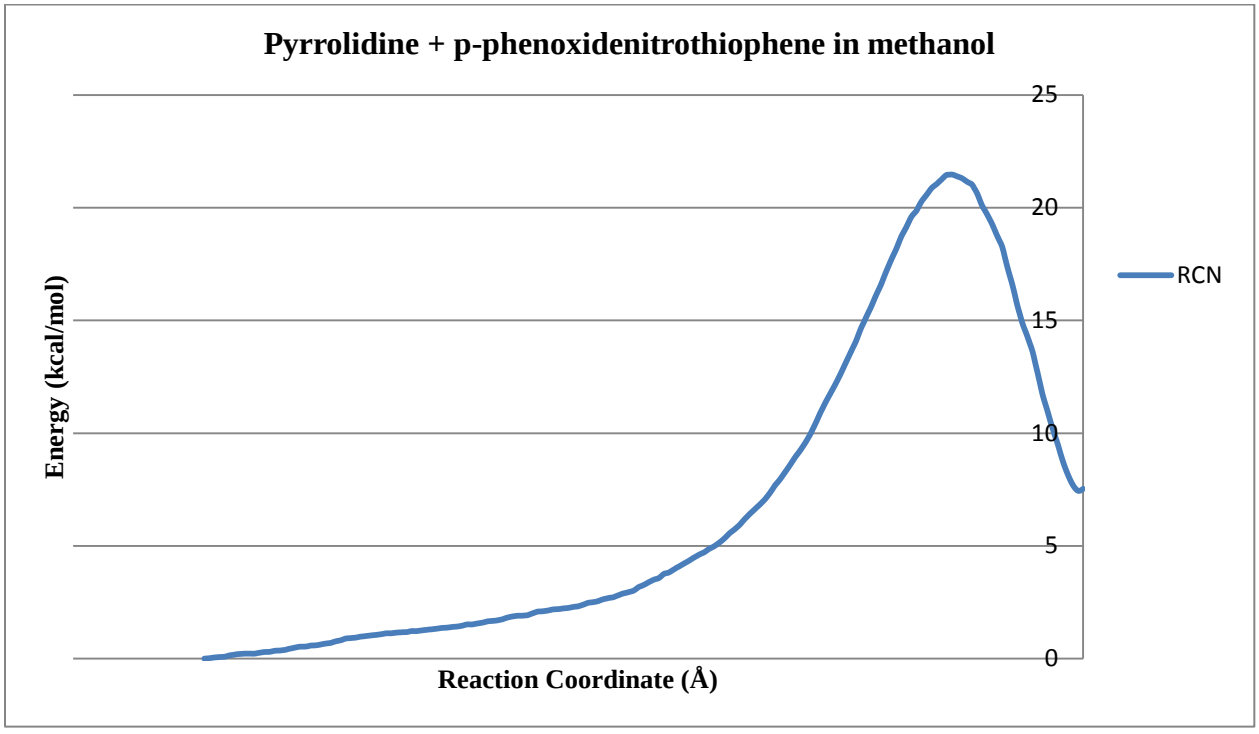


Figure 4S.16: The free energy diagram of the reaction Pyrrolidine with para-phenoxidenitrothiophene in methanol

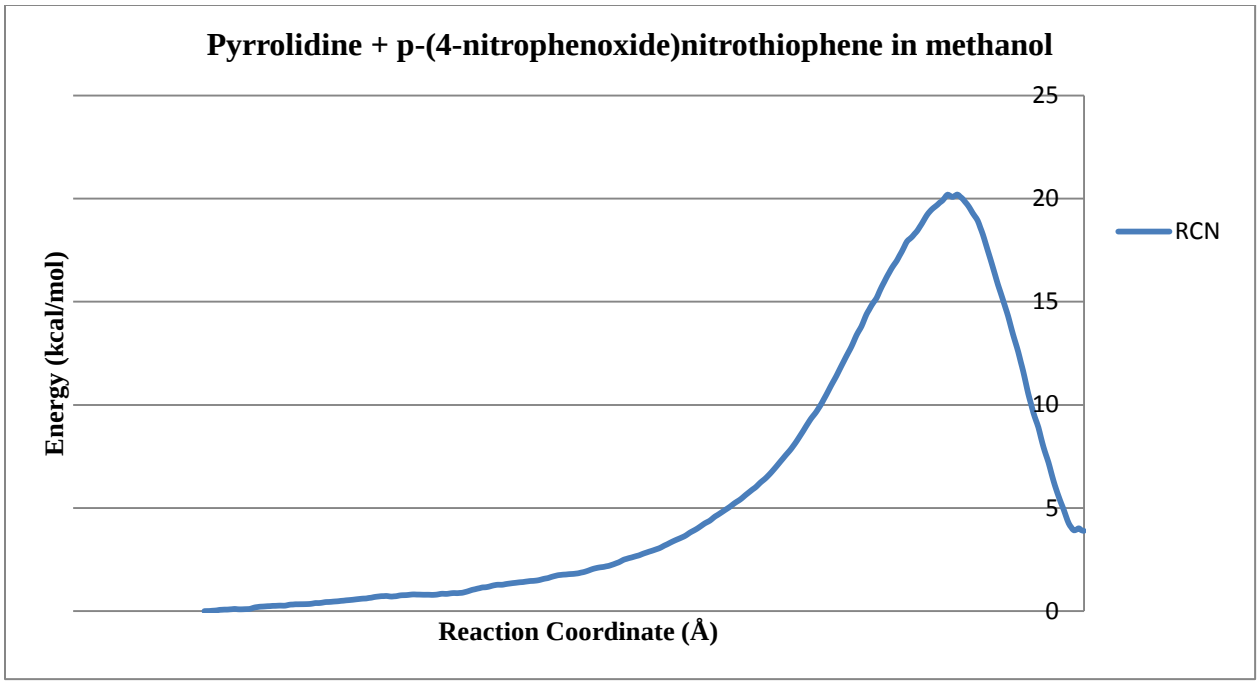


Figure 4S.17: The free energy diagram of the reaction Pyrrolidine with para-(4-nitrophenoxide)nitrothiophene in methanol

Pyrrolidine + ortho-LG-nitrothiophene

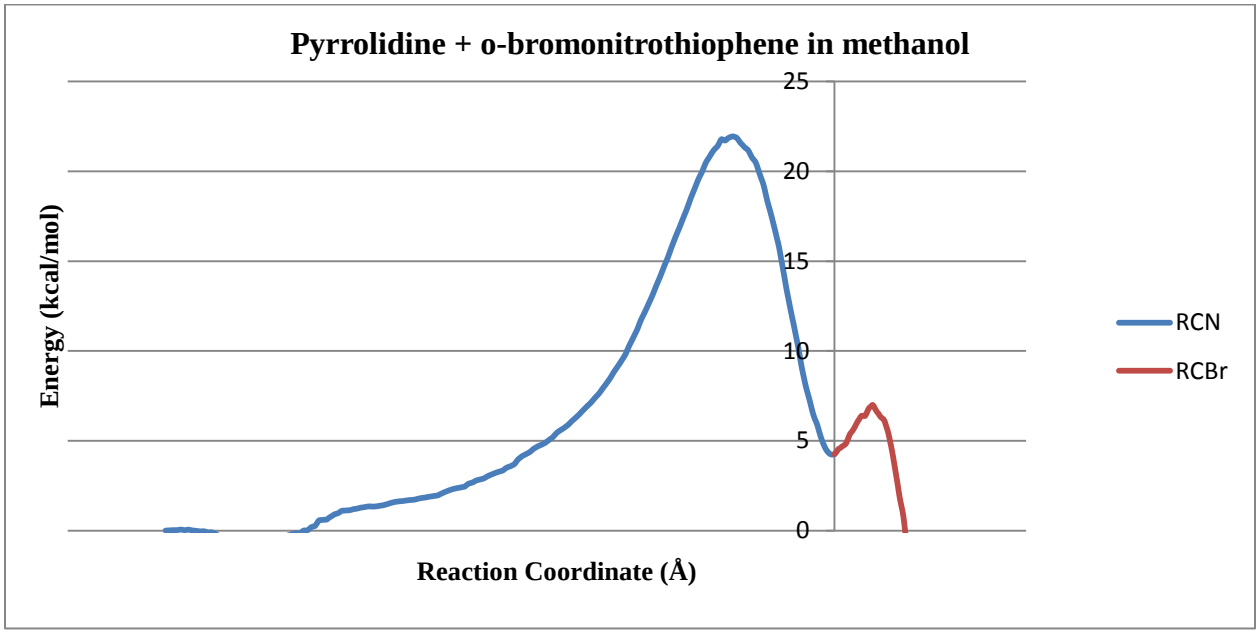


Figure 4S.18: The free energy diagram of the reaction Pyrrolidine with ortho-bromonitrothiophene in methanol

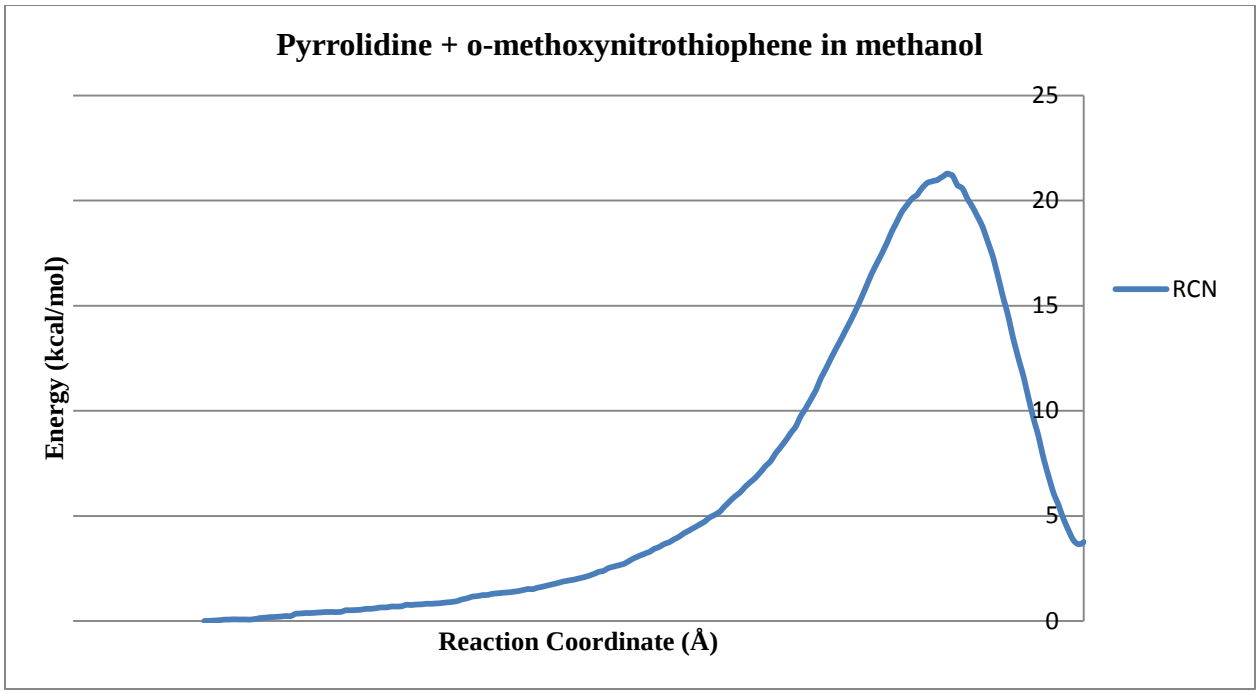


Figure 4S.19: The free energy diagram of the reaction Pyrrolidine with ortho-methoxynitrothiophene in methanol

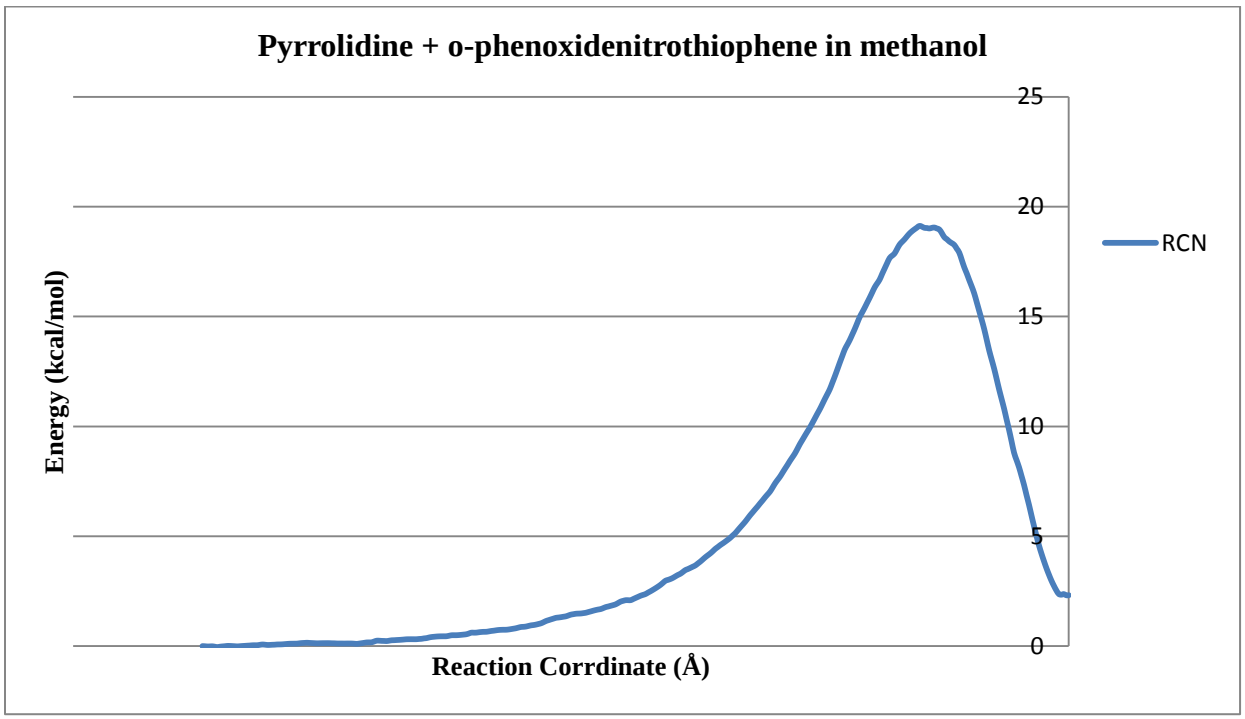


Figure 4S.20: The free energy diagram of the reaction Pyrrolidine with ortho-phenoxidenitrothiophene in methanol

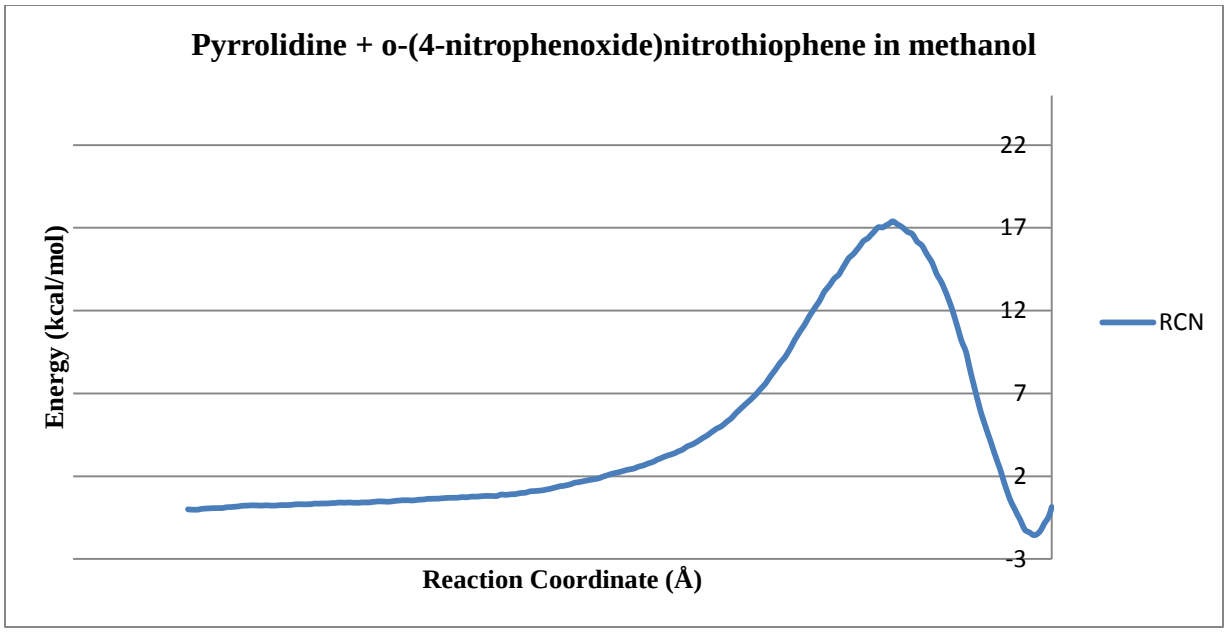


Figure 4S.21: The free energy diagram of the reaction Pyrrolidine with ortho-(4-nitrophenoxide)nitrothiophene in methanol

All Free Energy Surfaces in [BMIM][BF_4]

Piperidine + para-LG-nitrothiophene

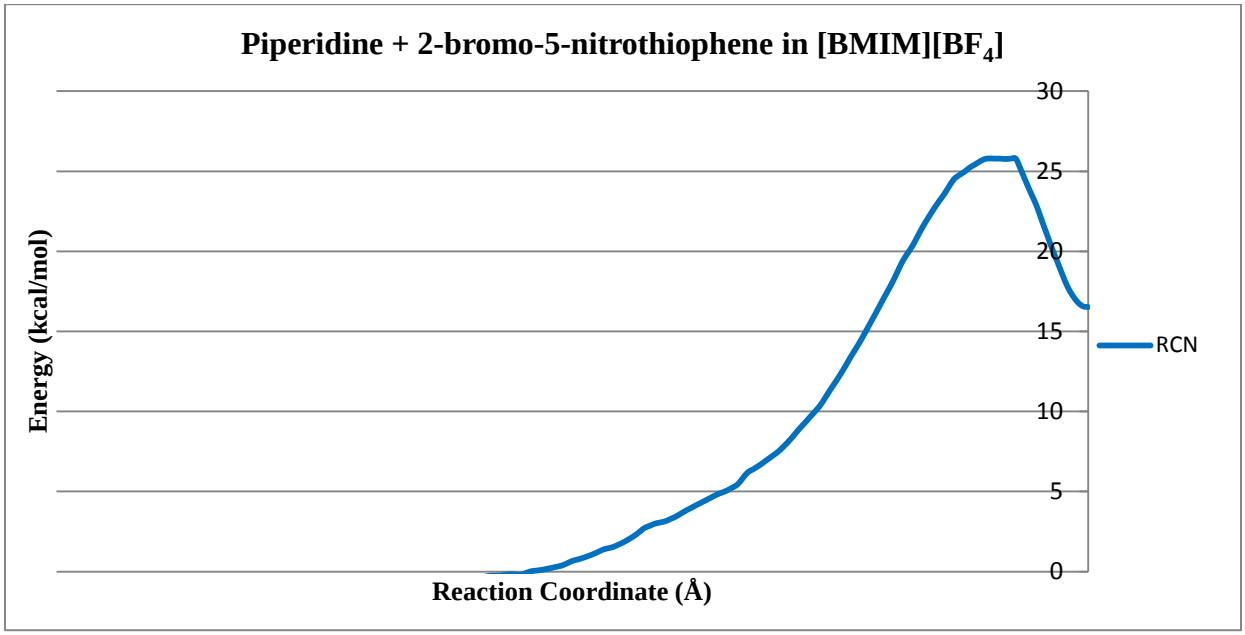


Figure 4S.22: The free energy diagram of the reaction piperidine with para-bromonitrothiophene in [BMIM][BF_4].

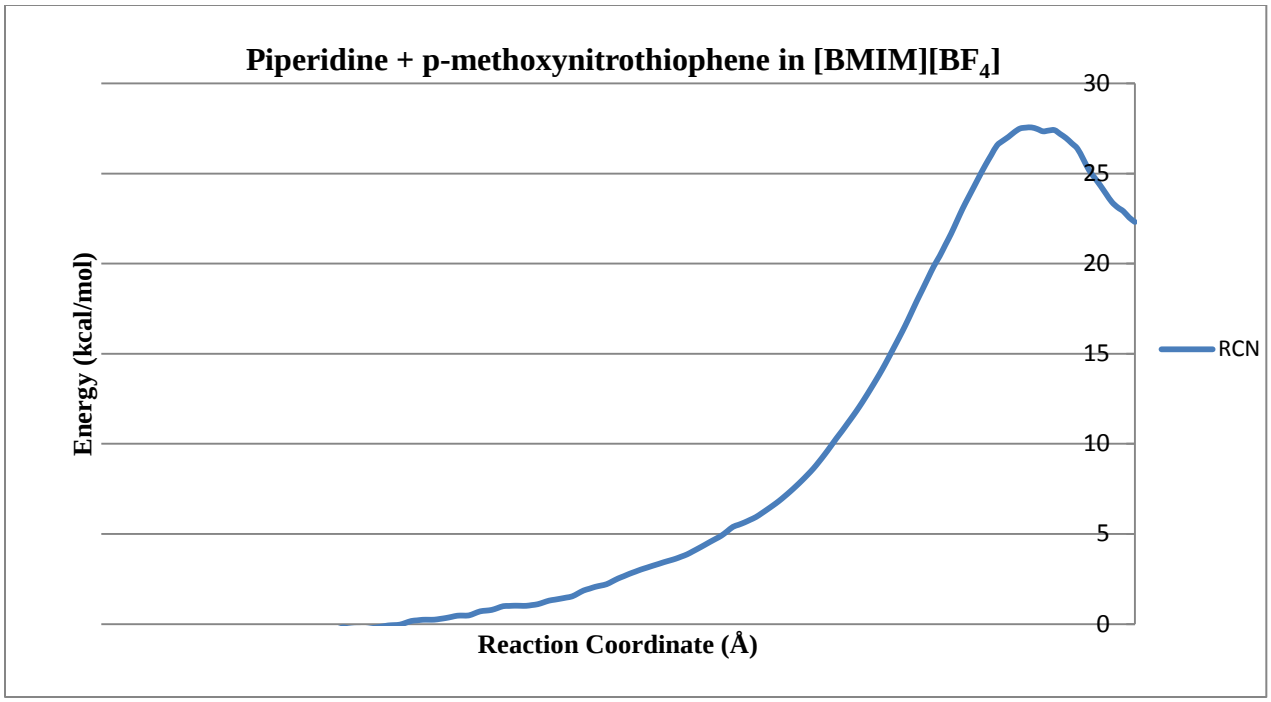


Figure 4S.23: The free energy diagram of the reaction piperidine with para-methoxynitrothiophene in [BMIM][BF_4].

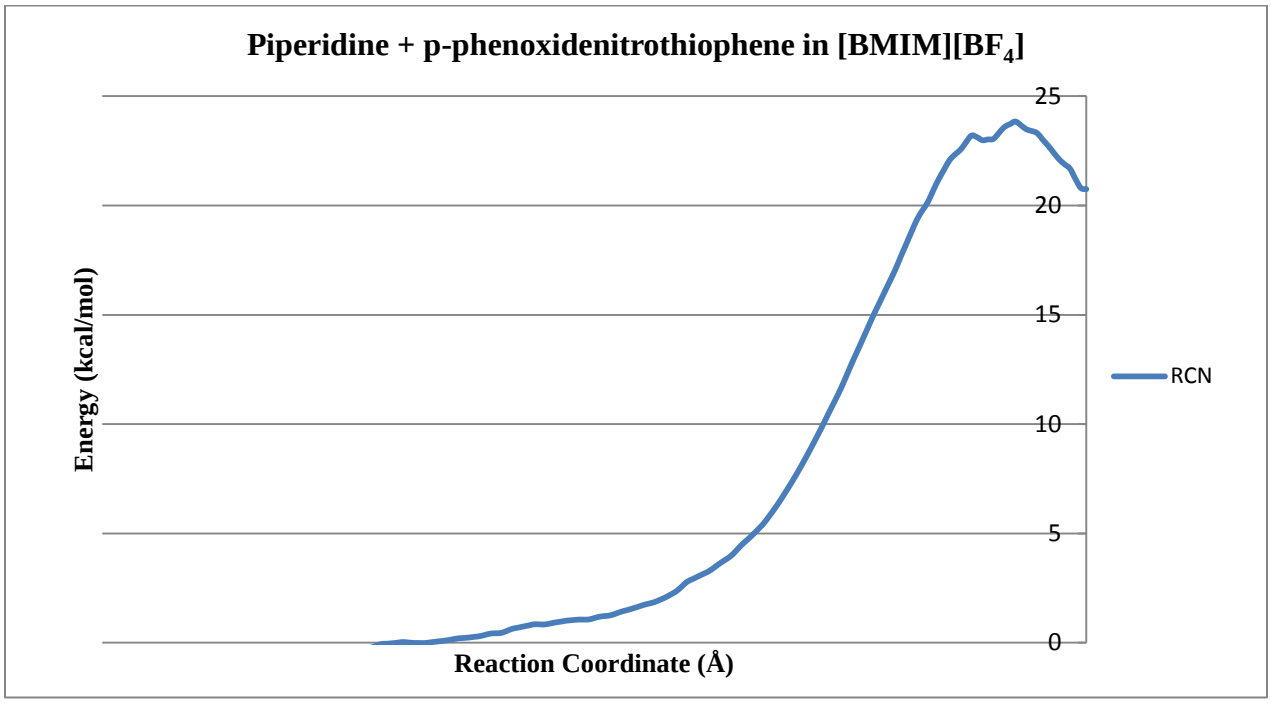


Figure 4S.24: The free energy diagram of the reaction piperidine with para-phenoxidenitrothiophene in [BMIM][BF_4].

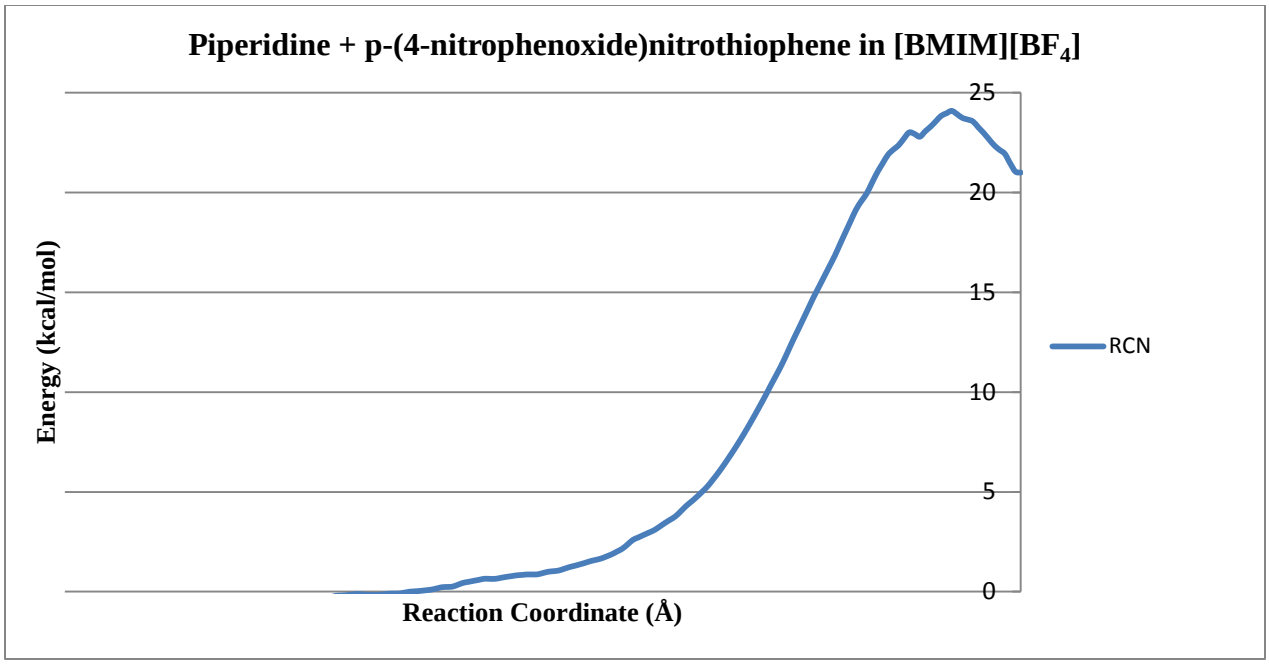


Figure 4S.25: The free energy diagram of the reaction piperidine with para-(4-nitrophenoxide)nitrothiophene in [BMIM][BF_4].

Piperidine + ortho-LG-nitrothiophene

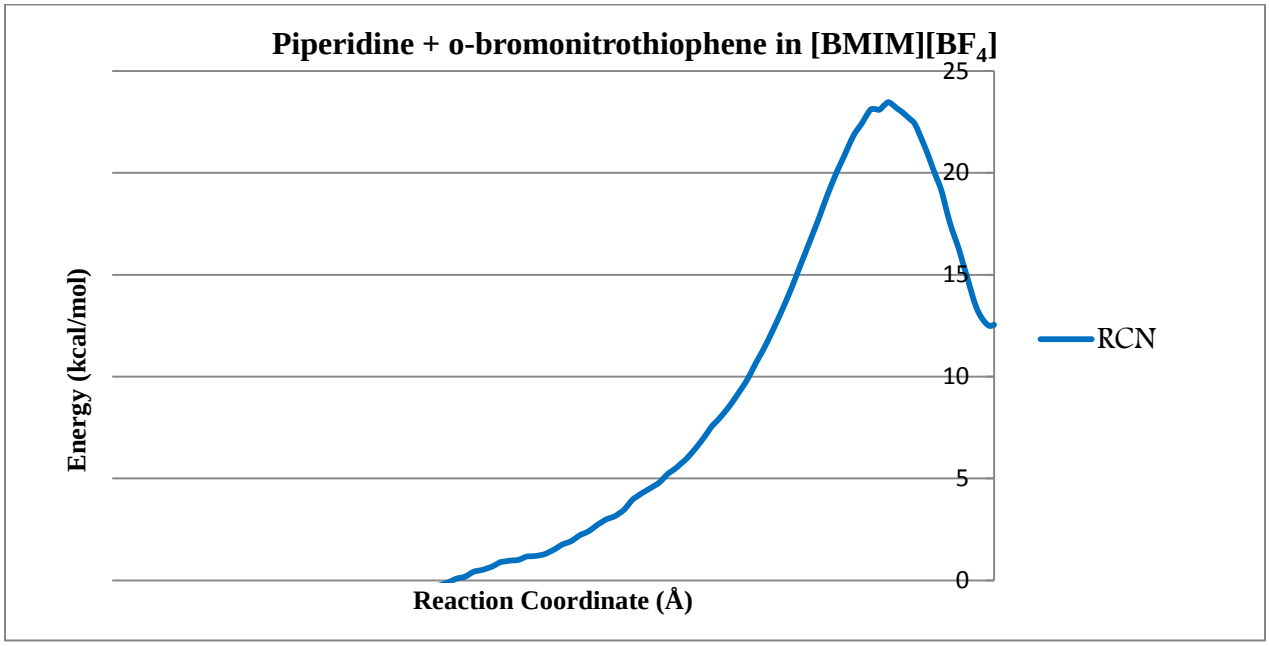


Figure 4S.26: The free energy diagram of the reaction piperidine with ortho-bromonitrothiophene in [BMIM][BF_4].

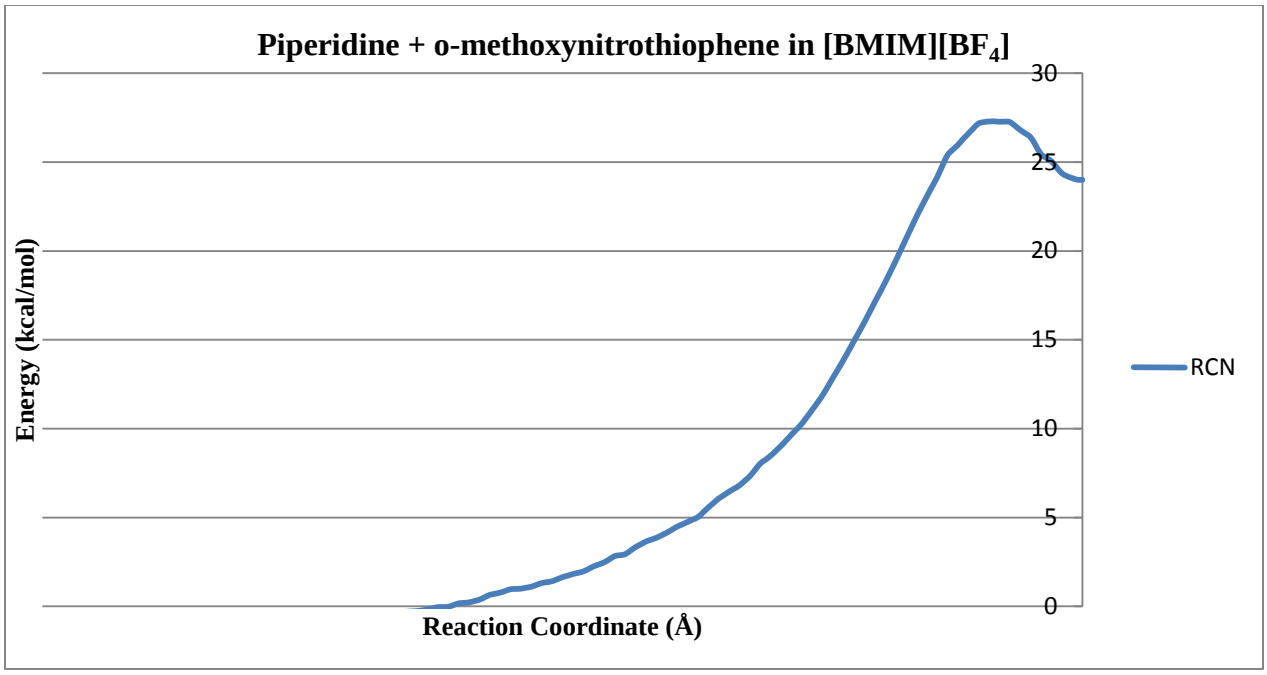


Figure 4S.27: The free energy diagram of the reaction piperidine with ortho-methoxynitrothiophene in [BMIM][BF_4].

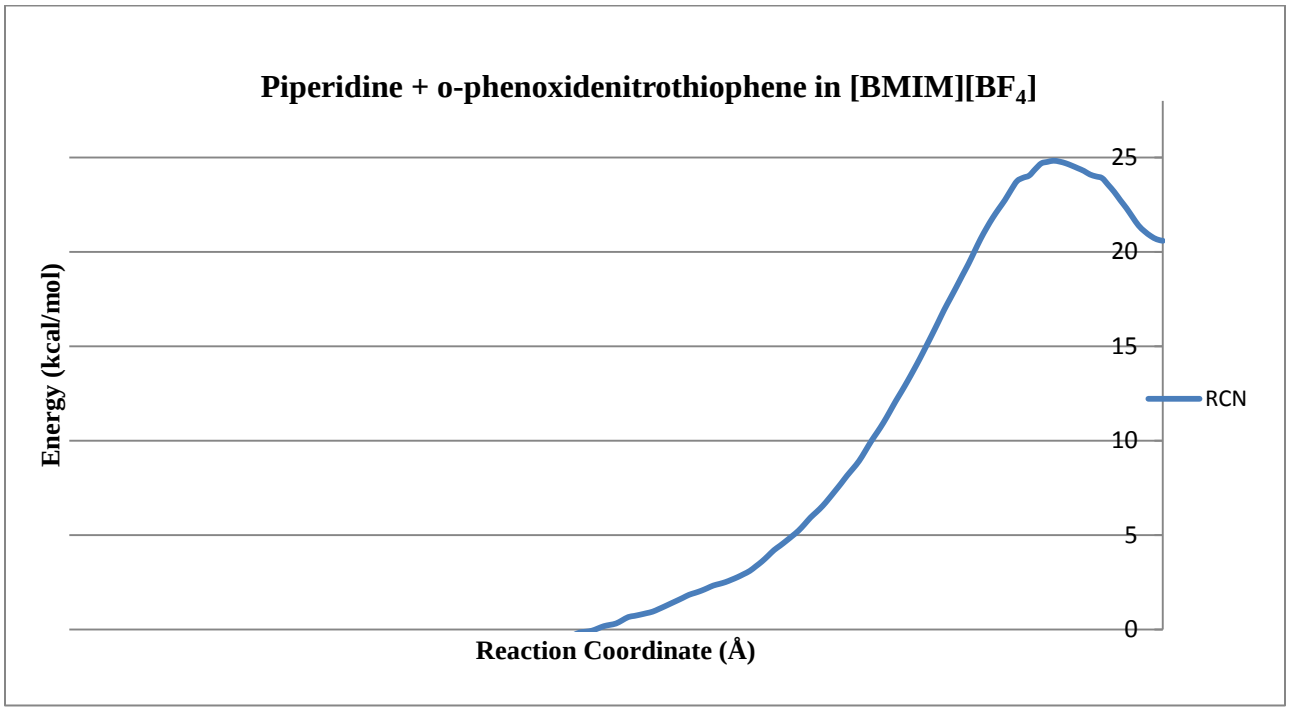


Figure 4S.28: The free energy diagram of the reaction piperidine with ortho-phenoxidenitrothiophene in [BMIM][BF_4].

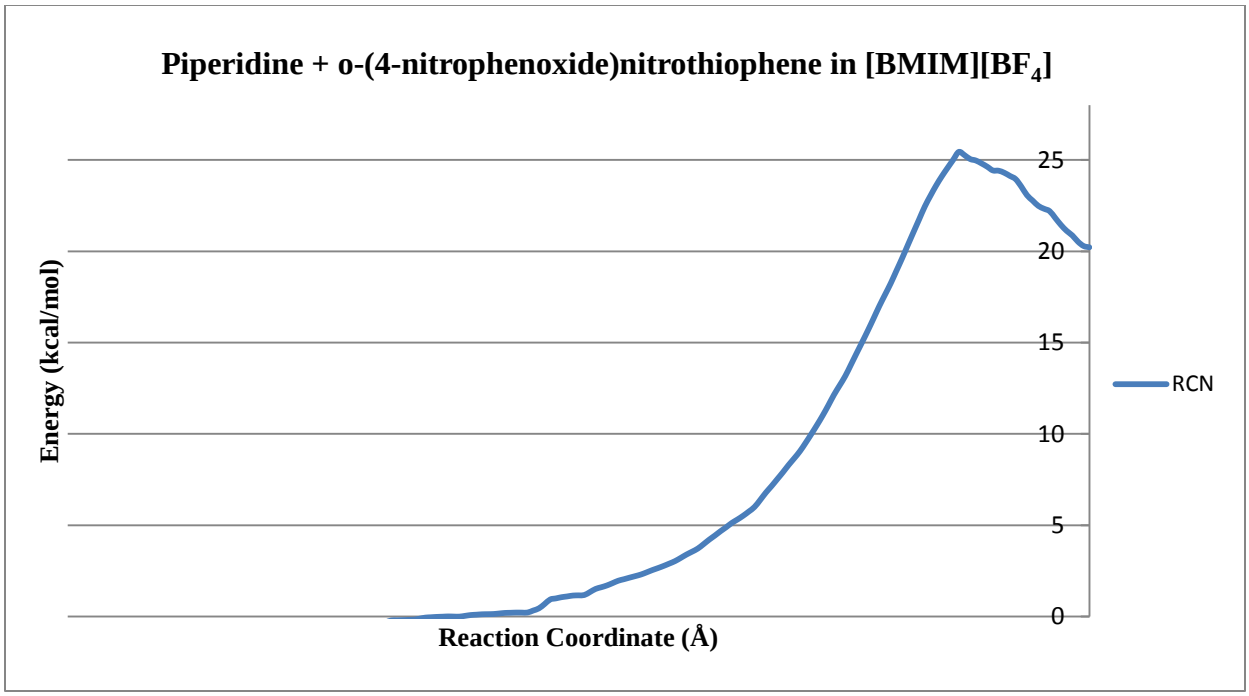


Figure 4S.29 The free energy diagram of the reaction piperidine with ortho-(4-nitrophenoxide)nitrothiophene in [BMIM][BF_4].

Morpholine + para-LG-nitrothiophene

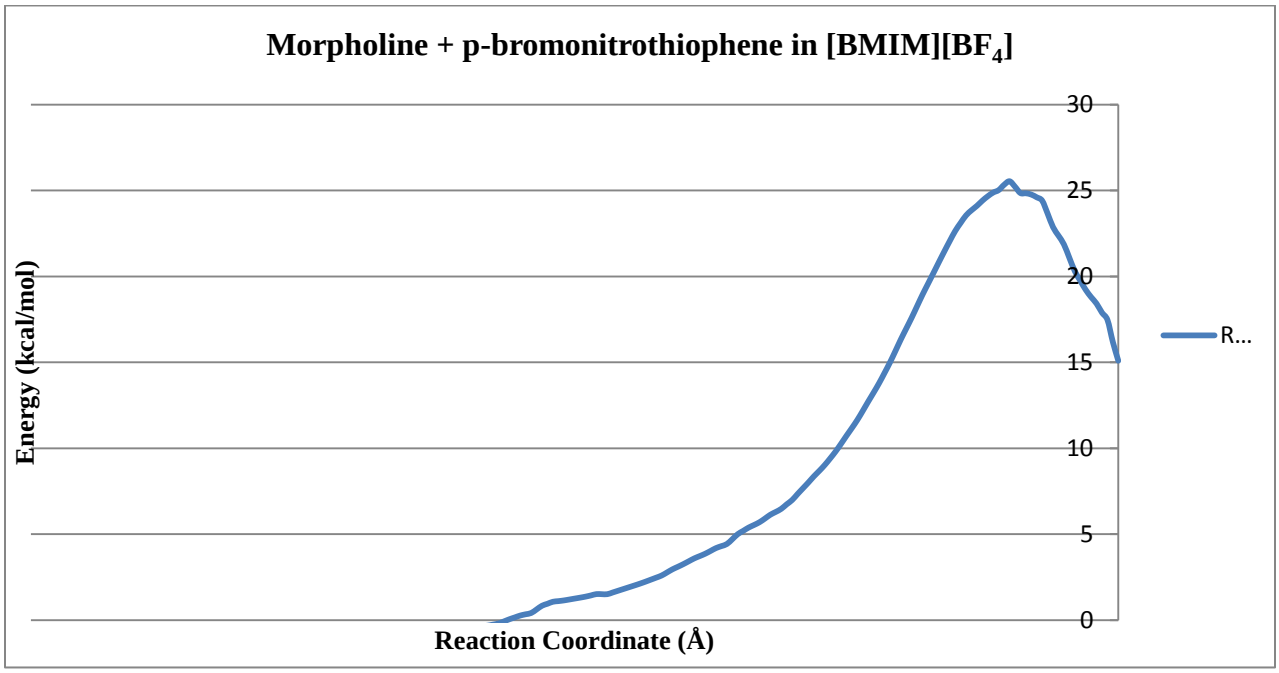


Figure 4S.30: The free energy diagram of the reaction morpholine with para-bromonitrothiophene in [BMIM][BF_4].

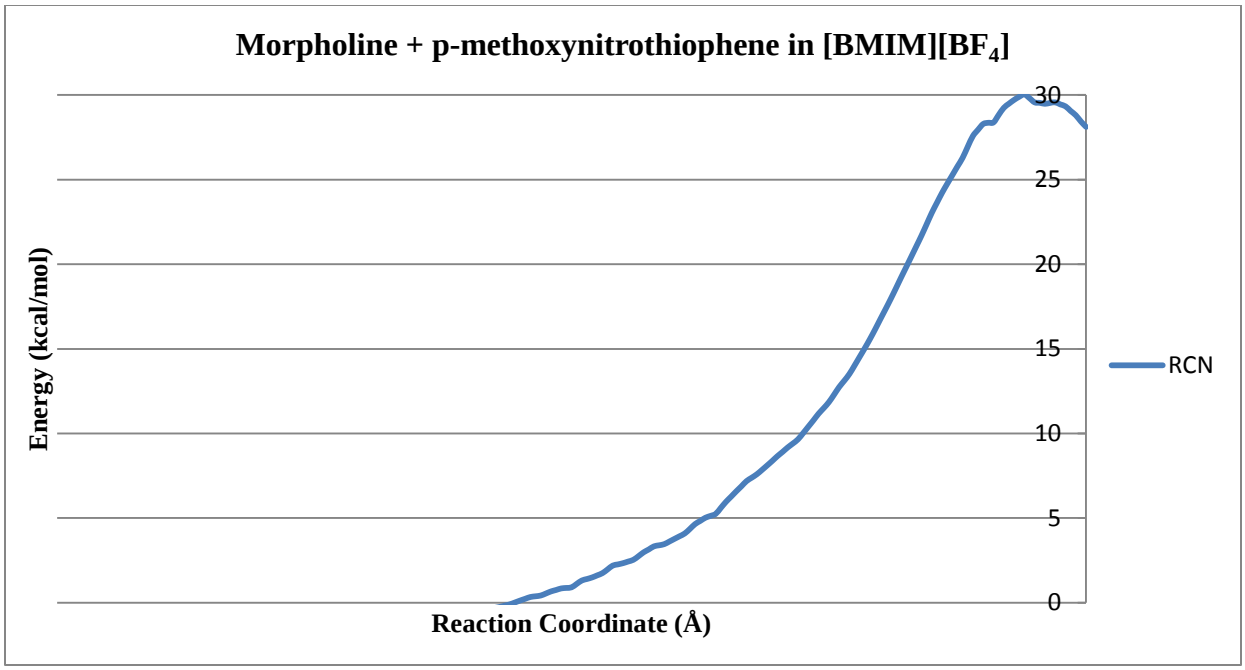


Figure 4S.31: The free energy diagram of the reaction morpholine with para-methoxynitrothiophene in [BMIM][BF_4].

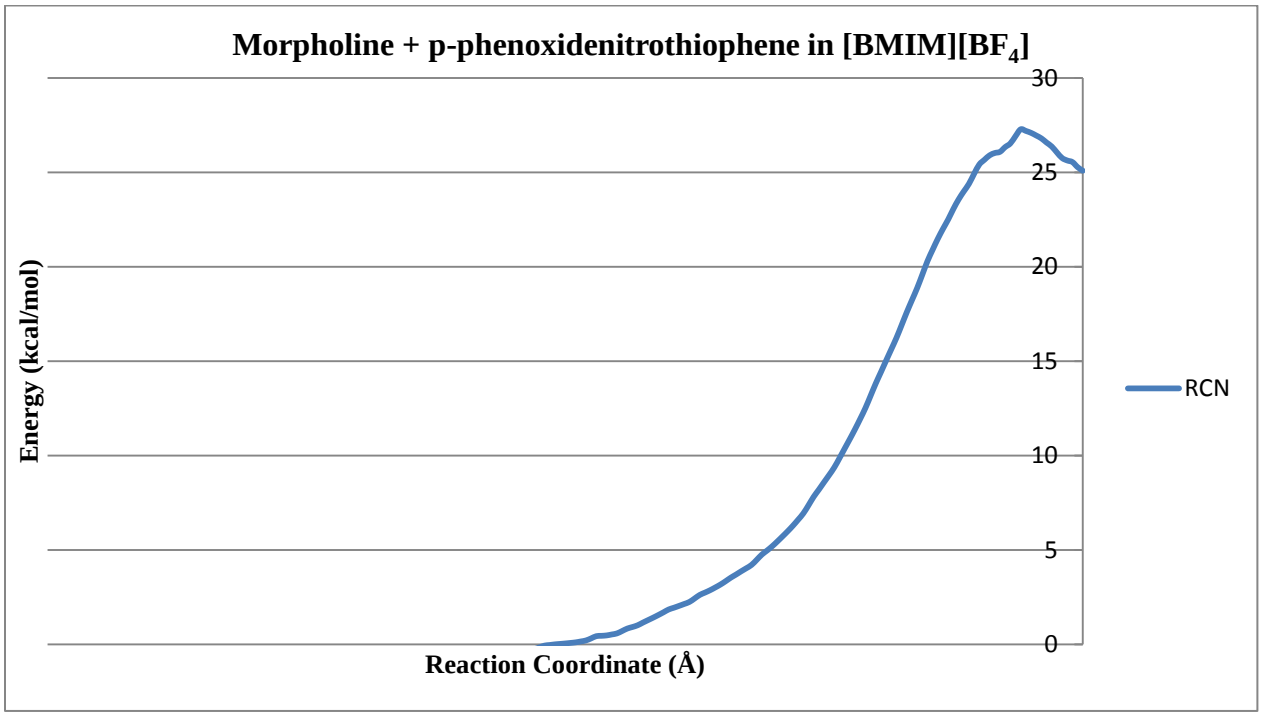


Figure 4S.32: The free energy diagram of the reaction morpholine with para-phenoxidenitrothiophene in [BMIM][BF_4].

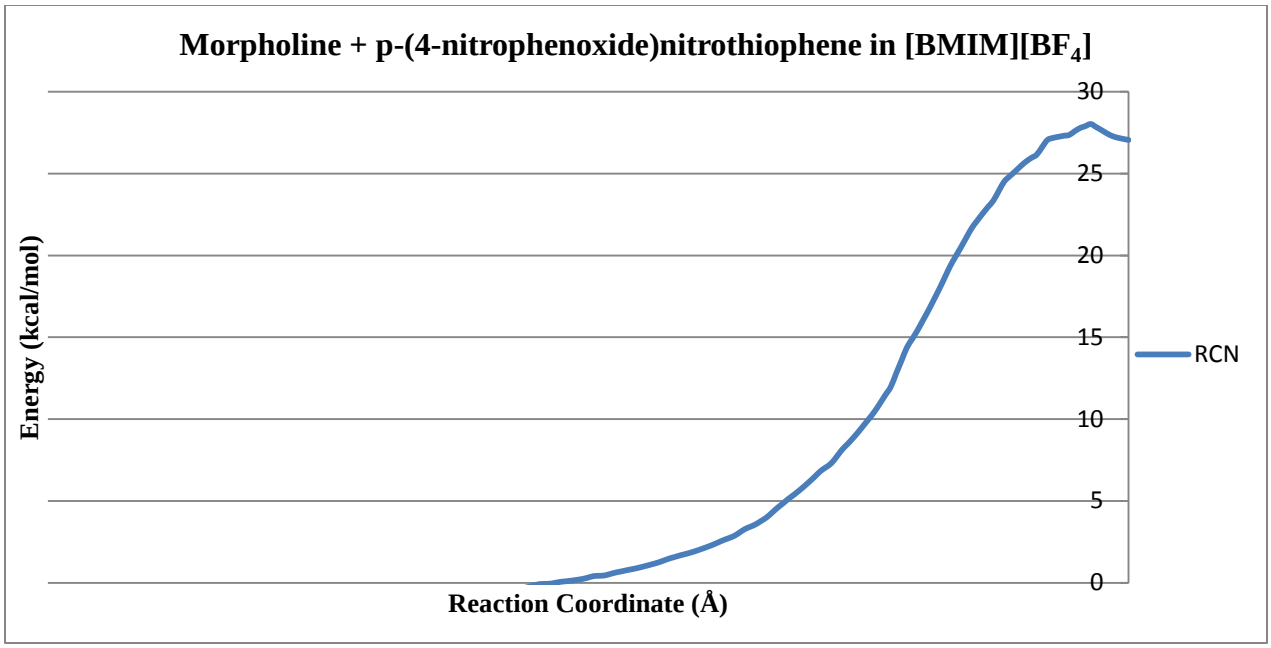


Figure 4S.33: The free energy diagram of the reaction morpholine with para-(4-nitrophenoxide)nitrothiophene in [BMIM][BF_4].

Morpholine + ortho-LG-nitrothiophene

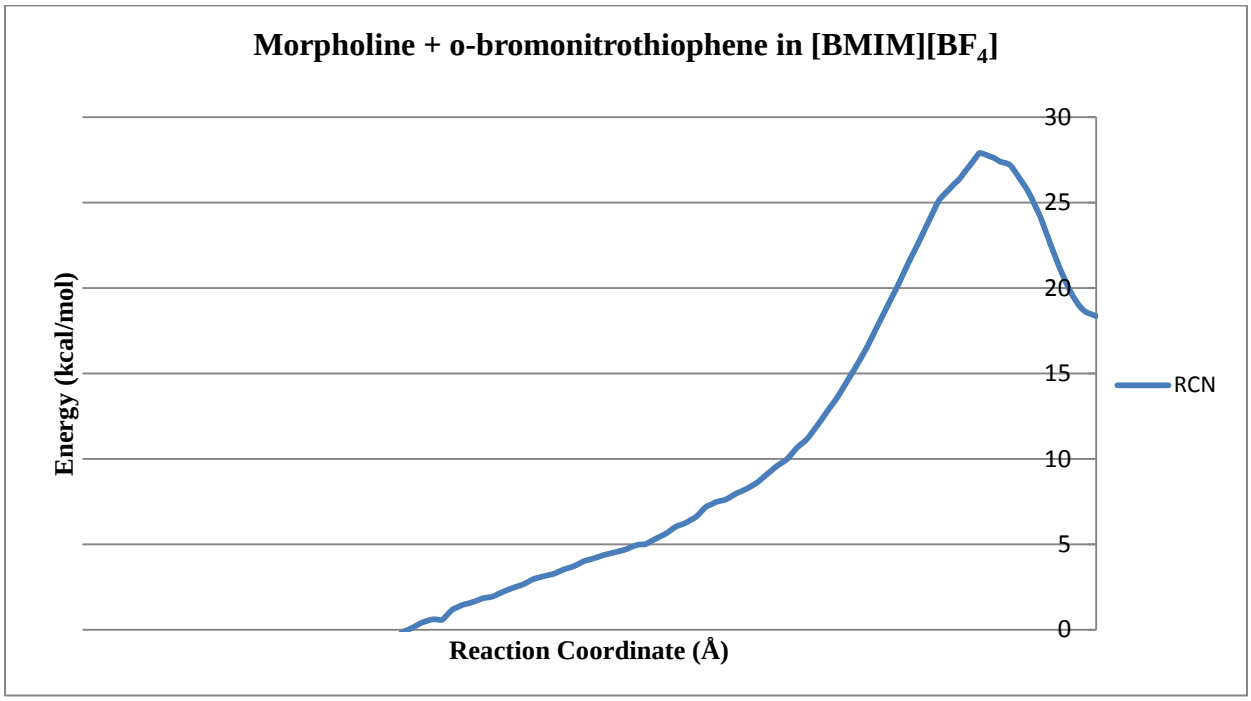


Figure 4S.34: The free energy diagram of the reaction morpholine with ortho-bromonitrothiophene in [BMIM][BF_4].

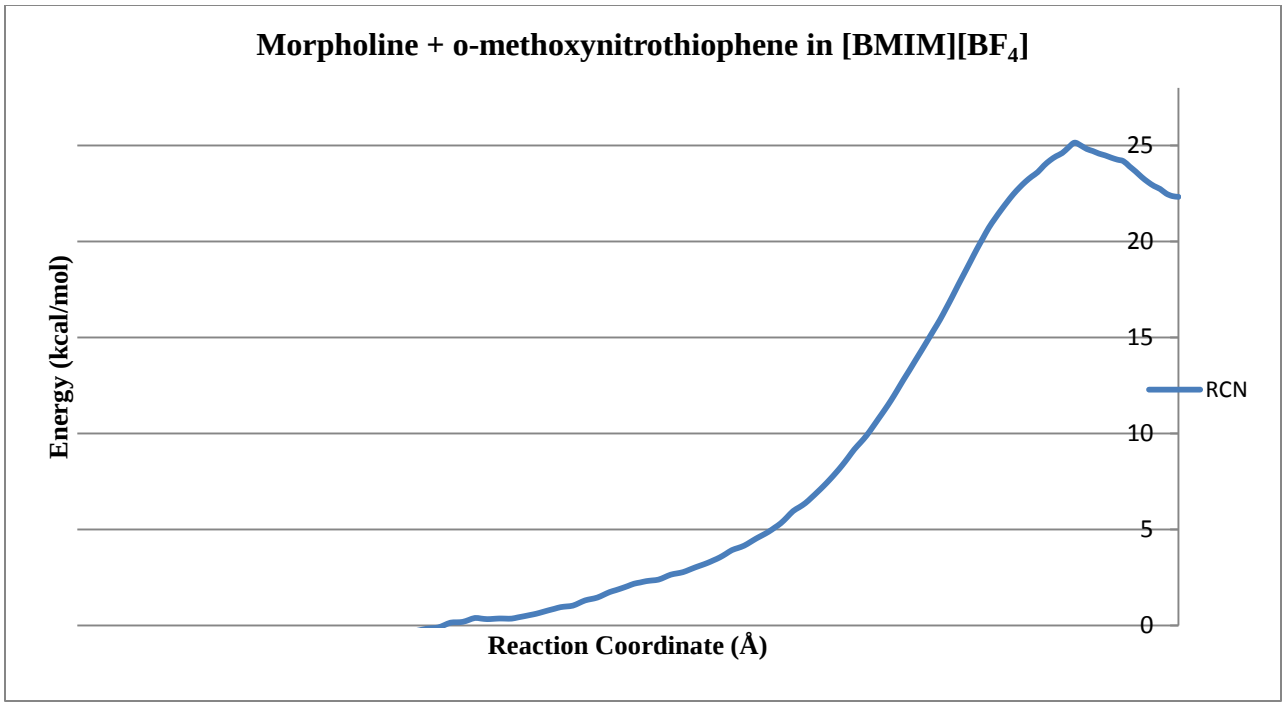


Figure 4S.35: The free energy diagram of the reaction morpholine with ortho-methoxynitrothiophene in [BMIM][BF_4].

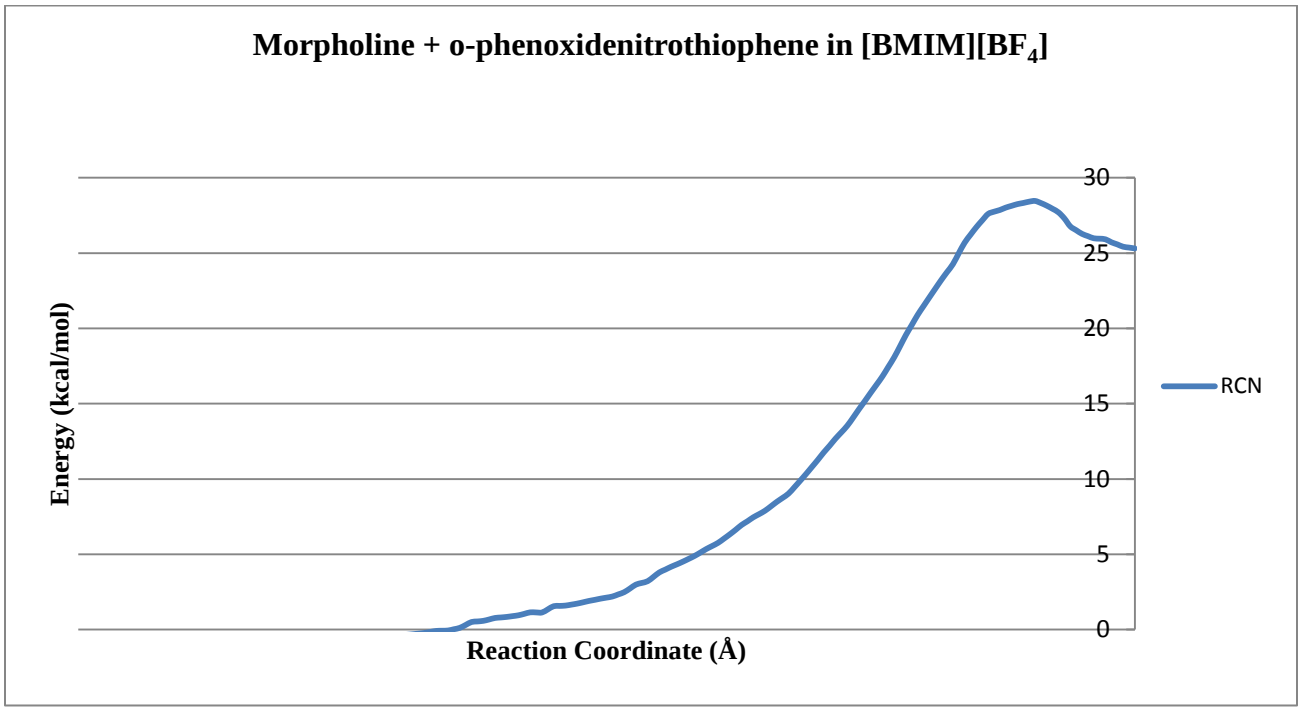


Figure 4S.36: The free energy diagram of the reaction morpholine with ortho-phenoxidenitrothiophene in [BMIM][BF_4].

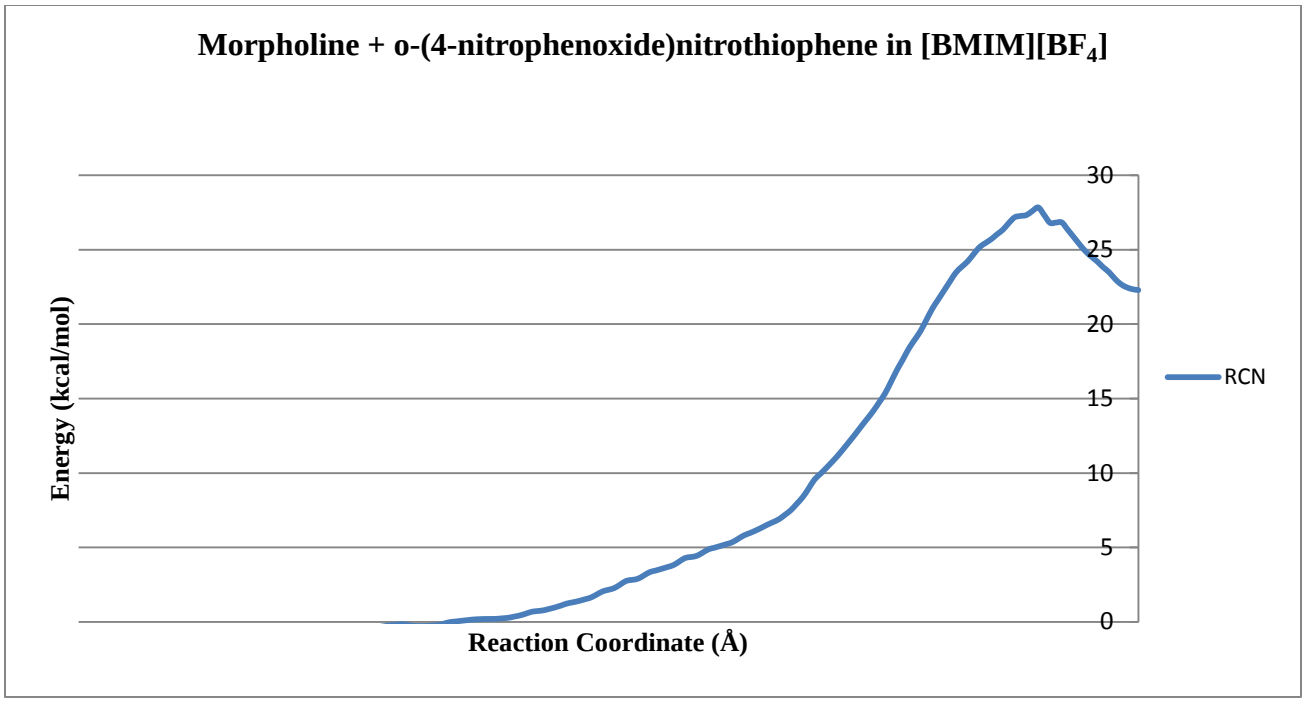


Figure 4S.37: The free energy diagram of the reaction morpholine with ortho-(4-nitrophenoxide)nitrothiophene in [BMIM][BF_4].

Pyrrolidine + para-LG-nitrothiophene

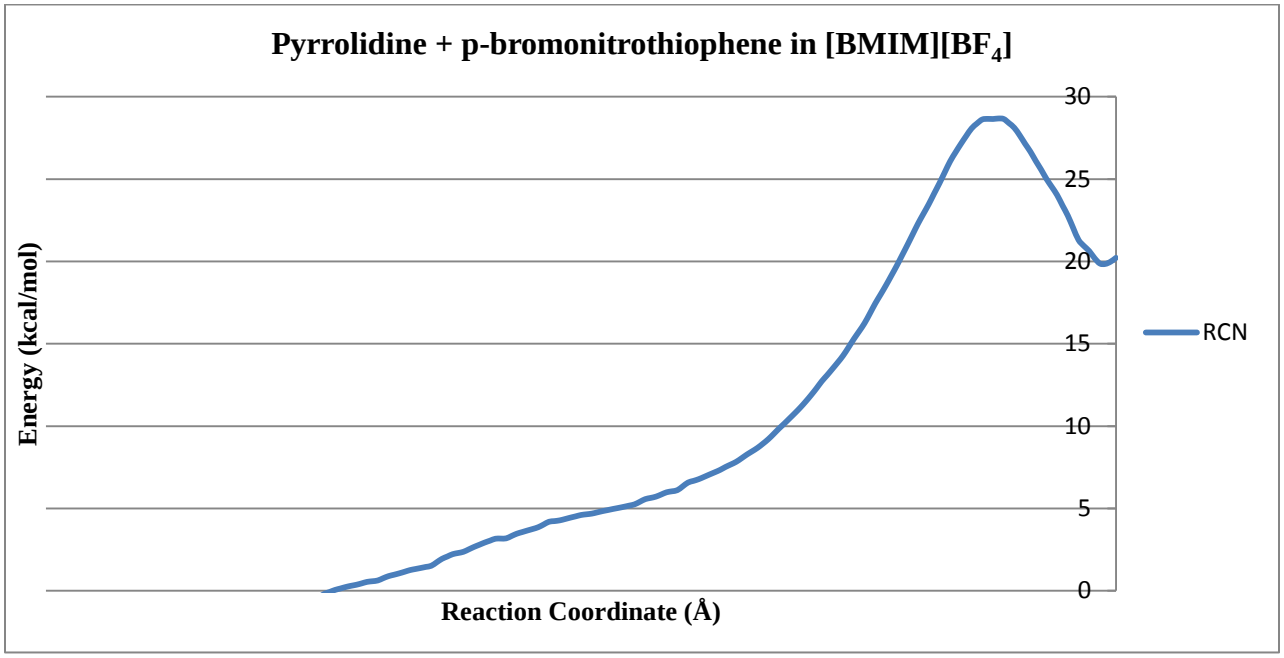


Figure 4S.38: The free energy diagram of the reaction pyrrolidine with para-bromonitrothiophene in [BMIM][BF_4].

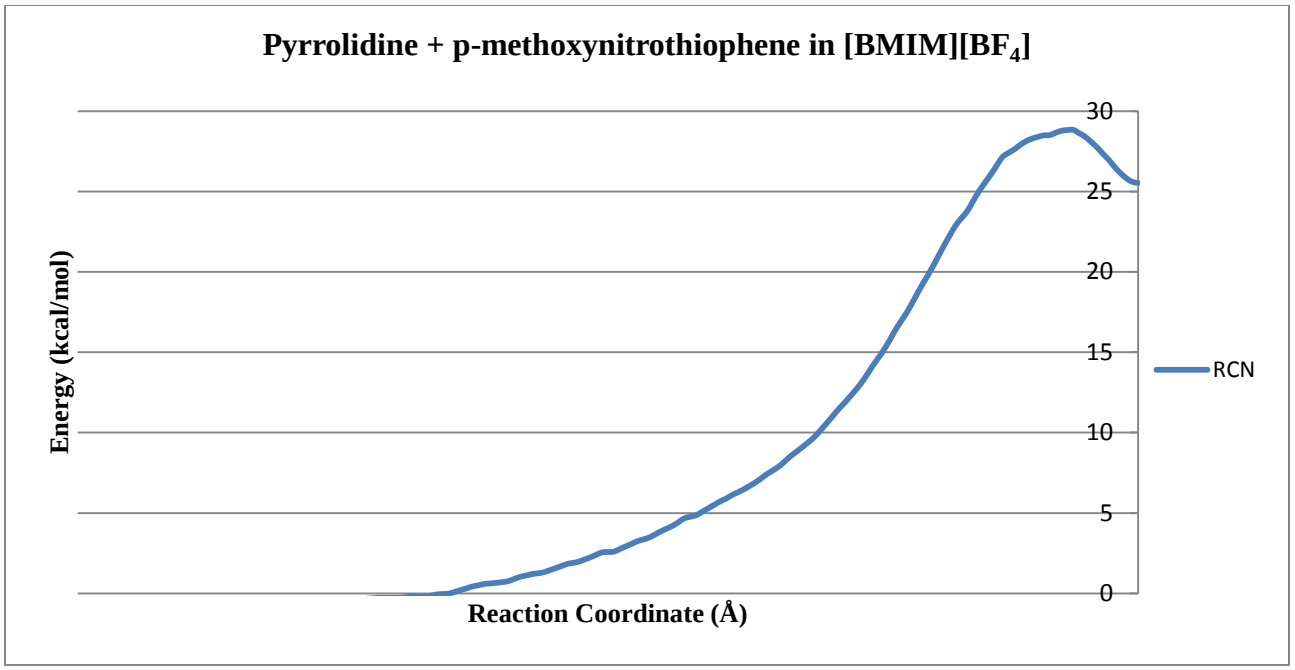


Figure 4S.39: The free energy diagram of the reaction pyrrolidine with para-methoxynitrothiophene in [BMIM][BF_4].

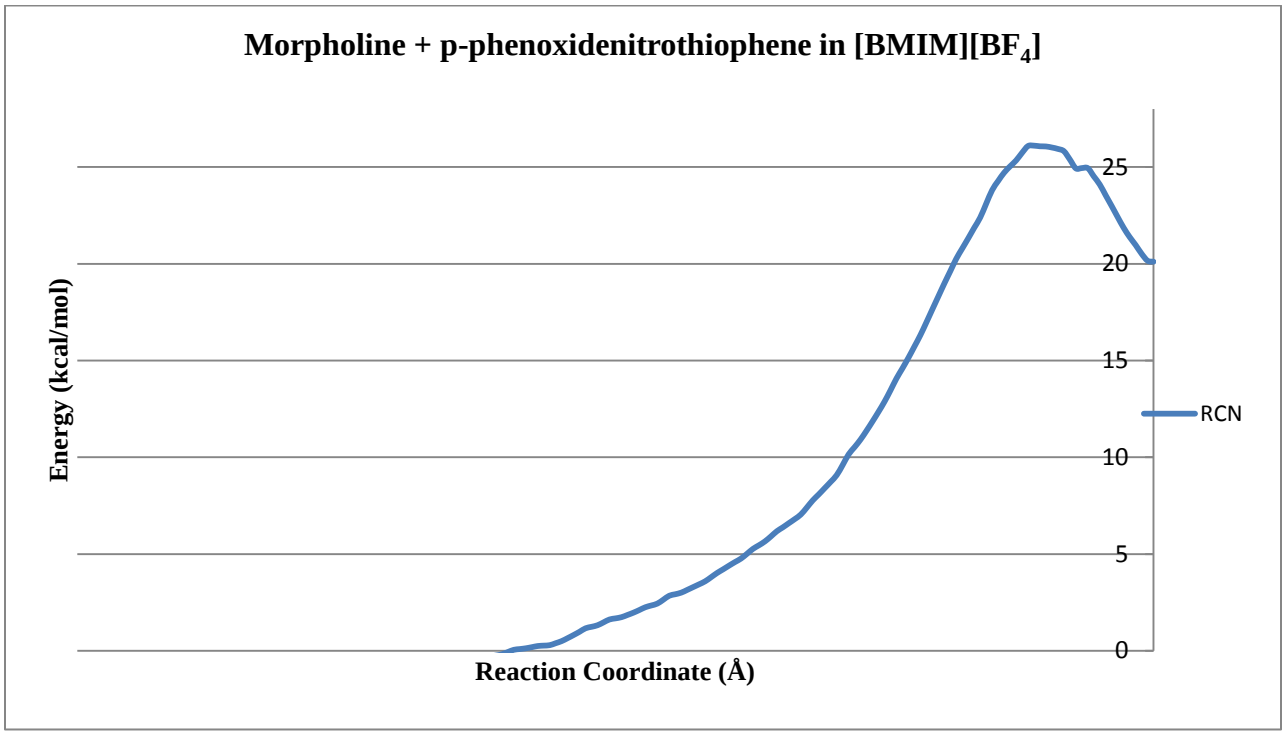


Figure 4S.40: The free energy diagram of the reaction pyrrolidine with para-phenoxidenitrothiophene in [BMIM][BF_4].

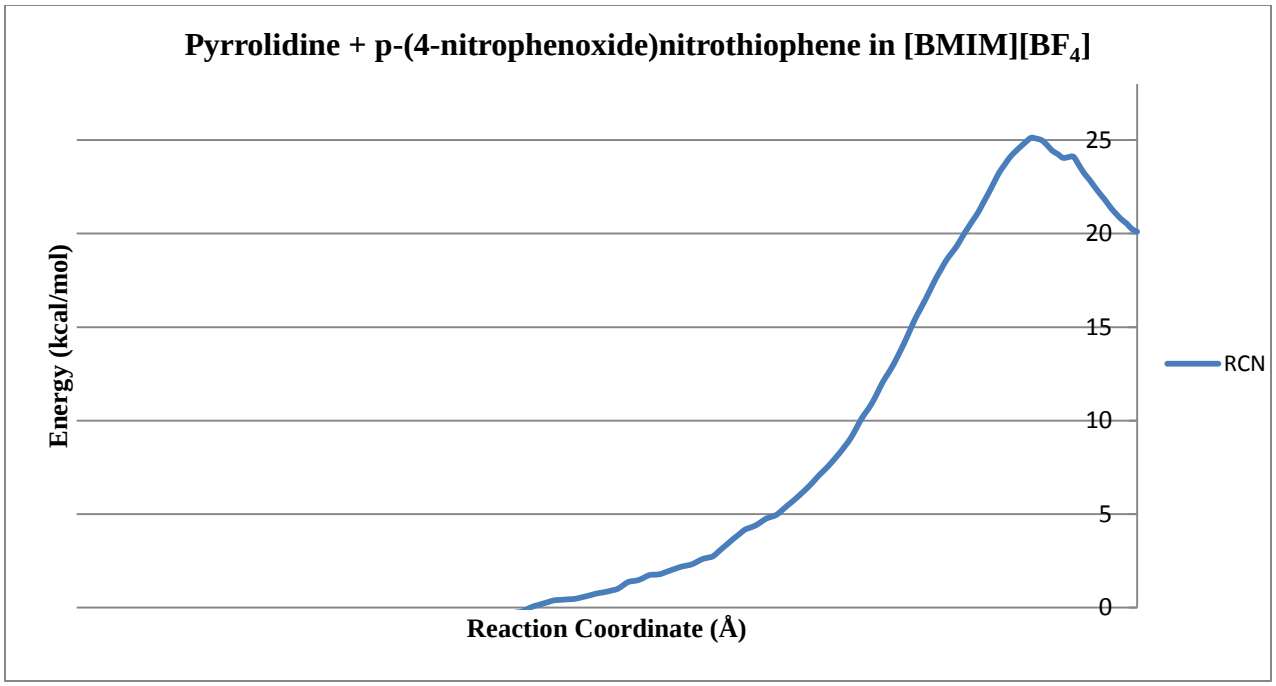


Figure 4S.41: The free energy diagram of the reaction pyrrolidine with para-(4-nitrophenoxide)nitrothiophene in [BMIM][BF_4].

Pyrrolidine + ortho-LG-nitrothiophene

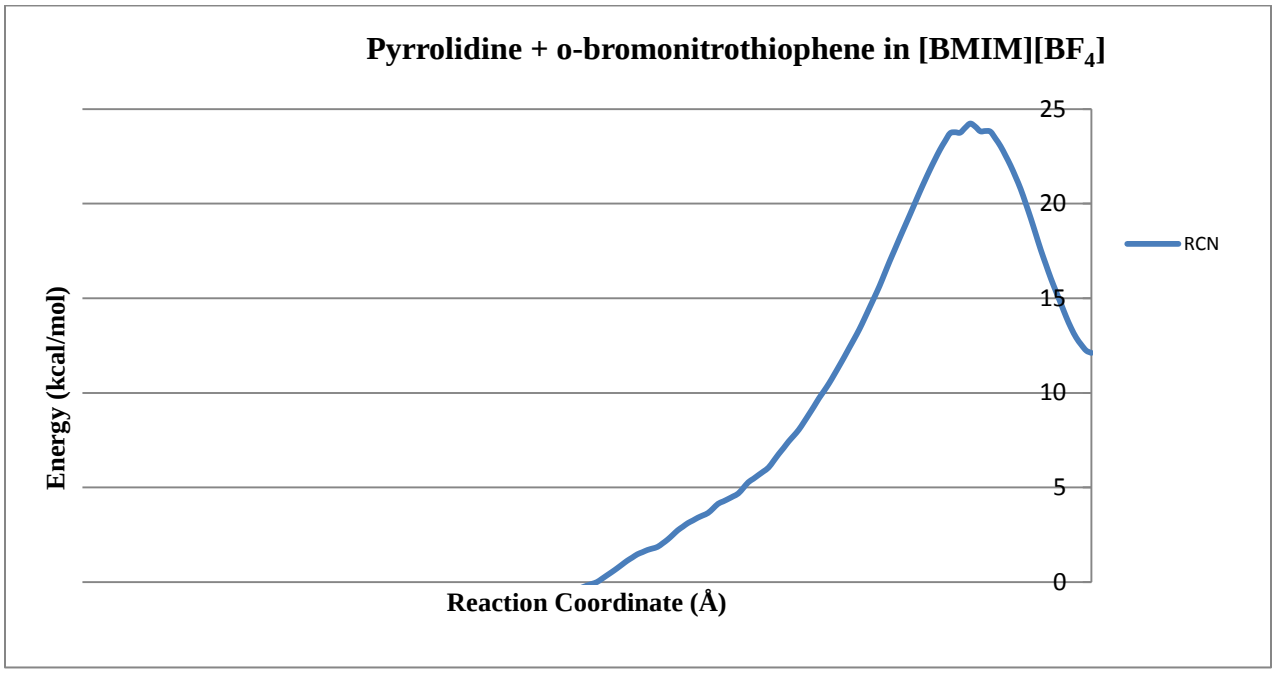


Figure 4S.42: The free energy diagram of the reaction pyrrolidine with ortho-bromonitrothiophene in [BMIM][BF_4].

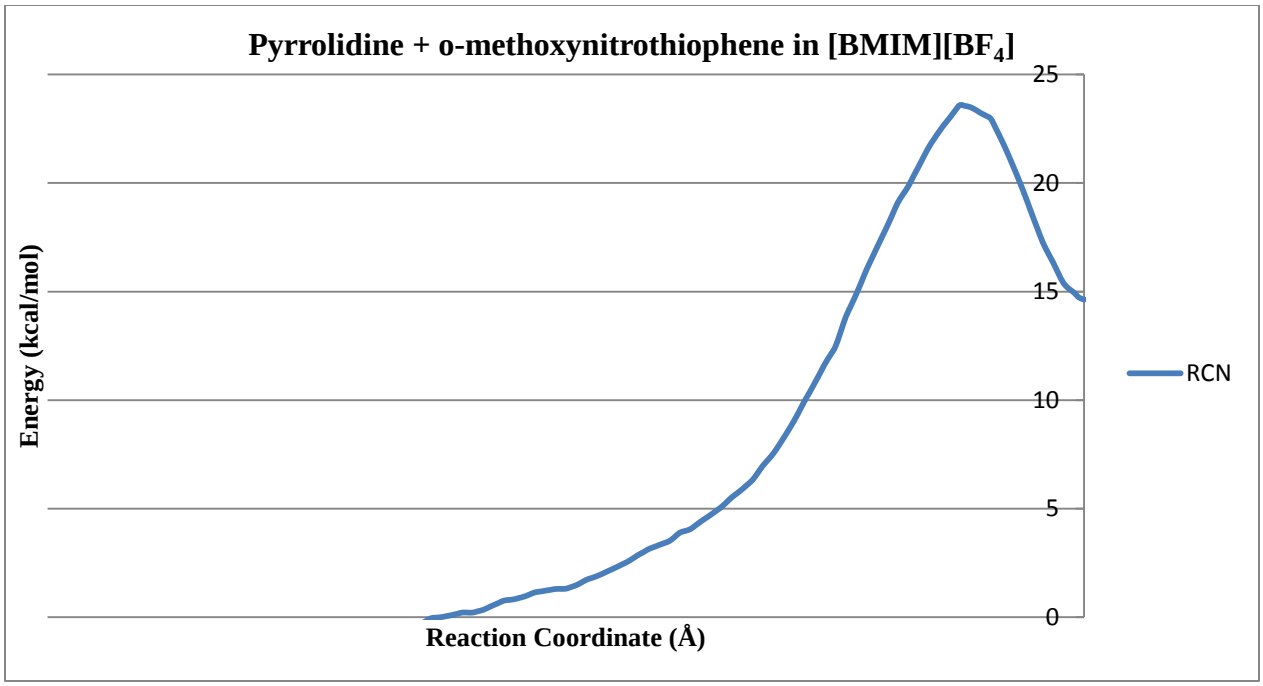


Figure 4S.43: The free energy diagram of the reaction pyrrolidine with ortho-methoxynitrothiophene in [BMIM][BF_4].

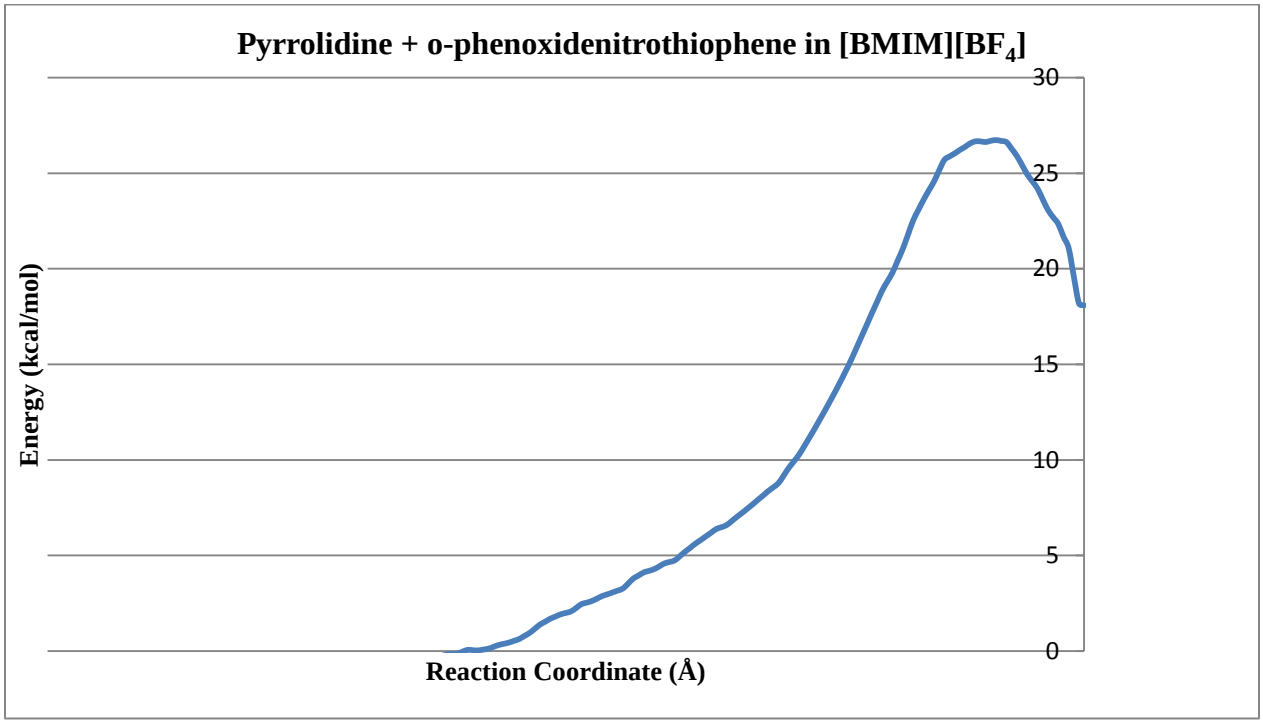


Figure 4S.44: The free energy diagram of the reaction pyrrolidine with ortho-phenoxidenitrothiophene in [BMIM][BF_4].

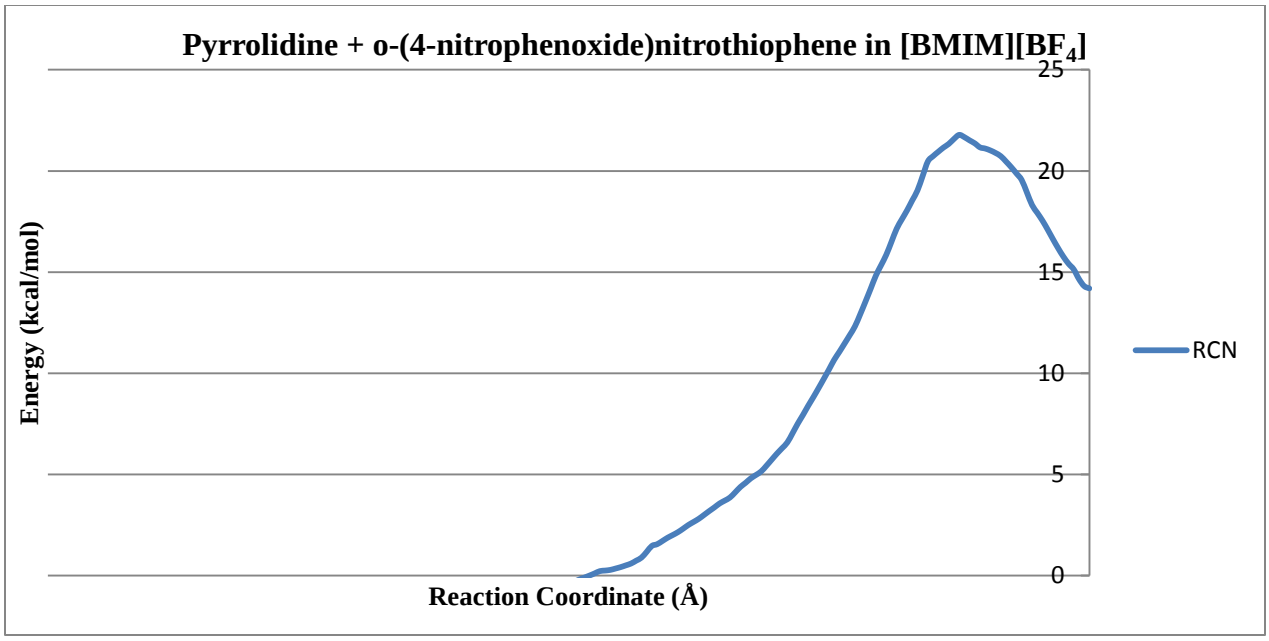


Figure 4S.45: The free energy diagram of the reaction pyrrolidine with ortho-(4-nitrophenoxide)nitrothiophene in [BMIM][BF_4].

Free Energy Surfaces in [BMIM][PF_6]

Piperidine + para-LG-nitrothiophene

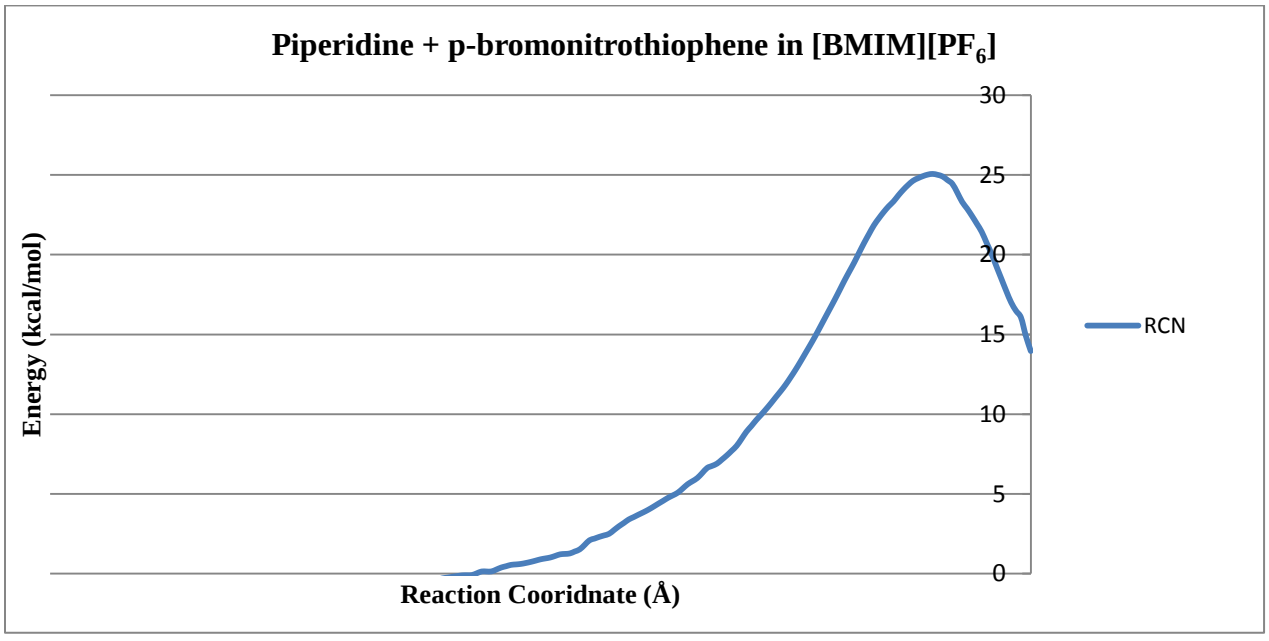


Figure 4S.46: The free energy diagram of the reaction pyrrolidine with para-bromonitrothiophene in [BMIM][PF_6].

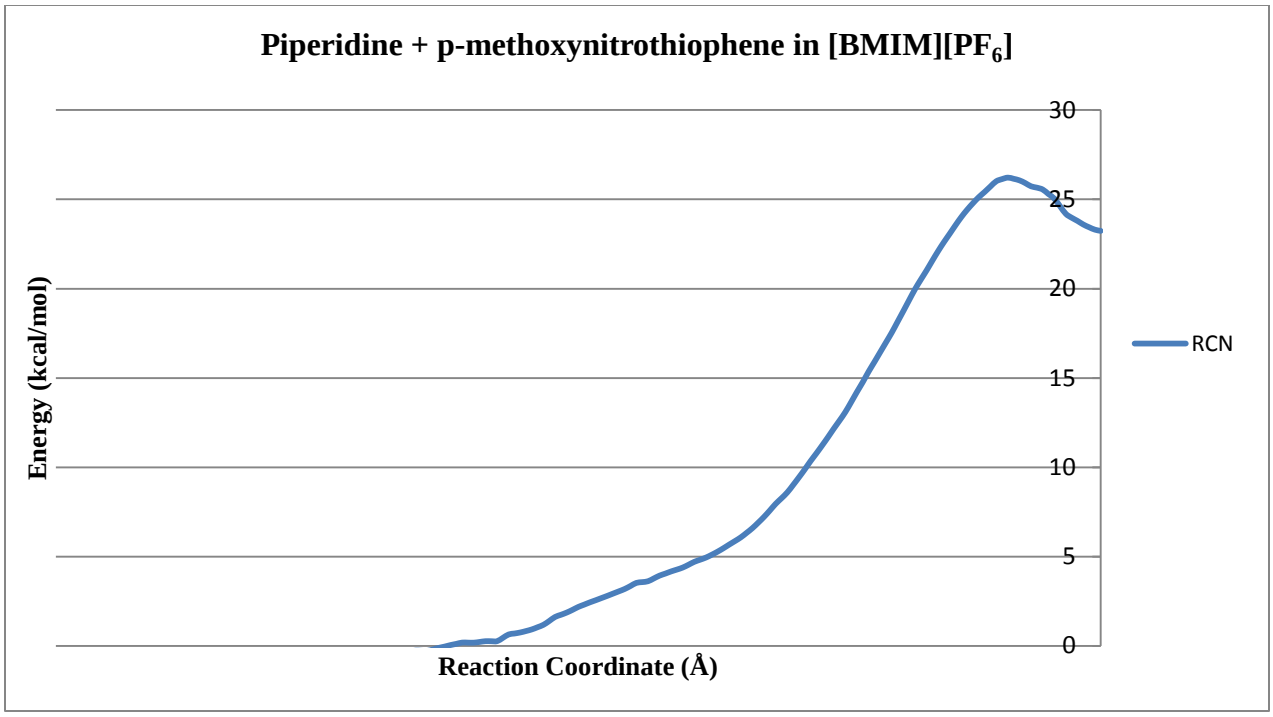


Figure 4S.47: The free energy diagram of the reaction pyrrolidine with para-methoxynitrothiophene in [BMIM][PF_6].

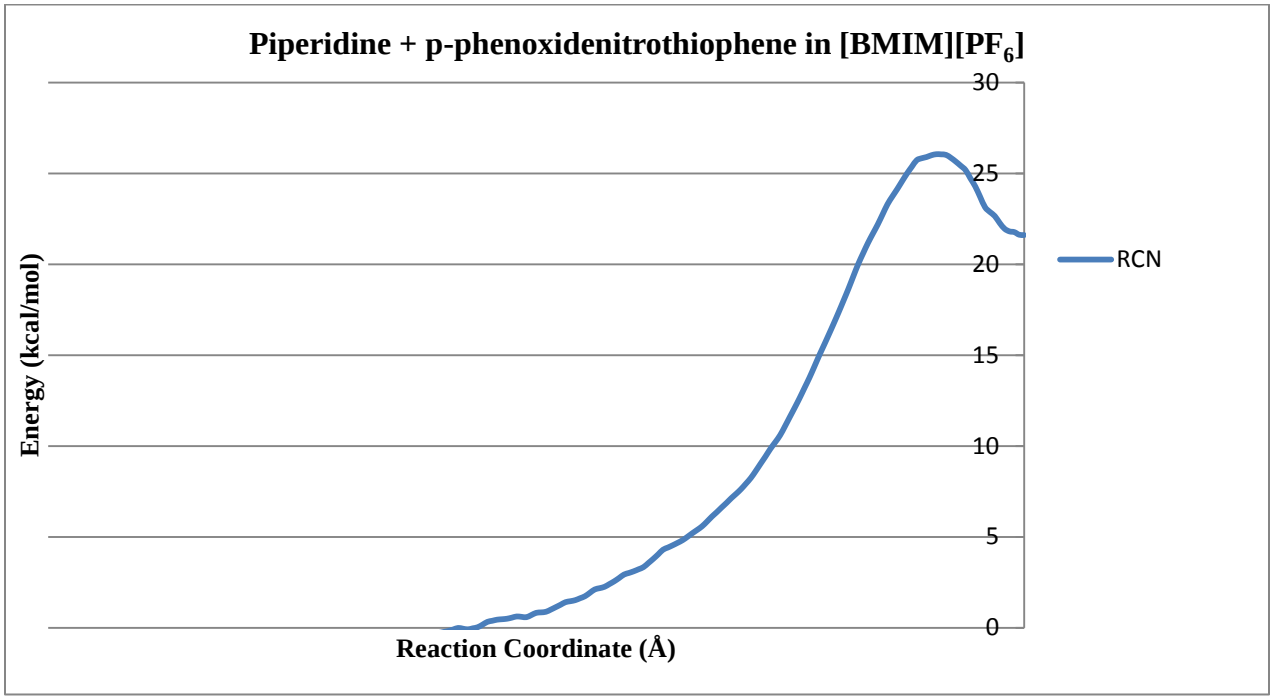


Figure 4S.48: The free energy diagram of the reaction pyrrolidine with para-phenoxidenitrothiophene in [BMIM][PF_6].

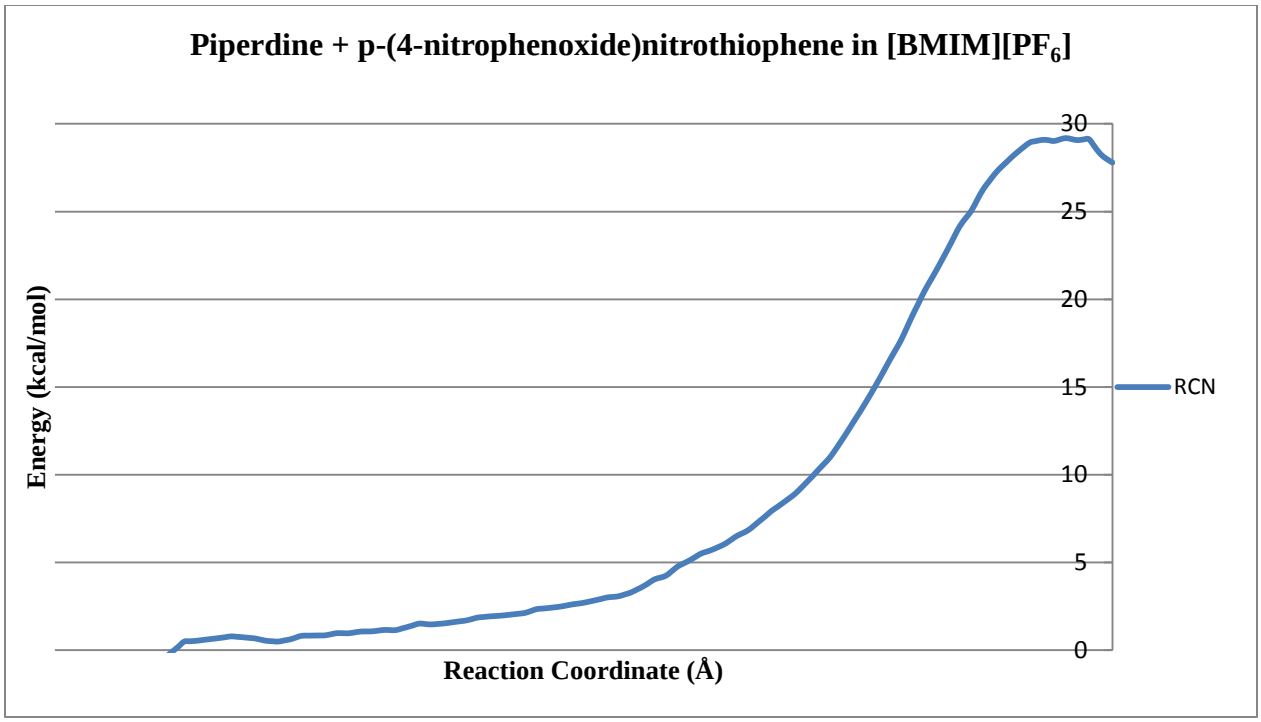


Figure 4S.49: The free energy diagram of the reaction pyrrolidine with para-(4-nitrophenoxide)nitrothiophene in [BMIM][PF_6].

Piperidine + ortho-LG-nitrothiophene

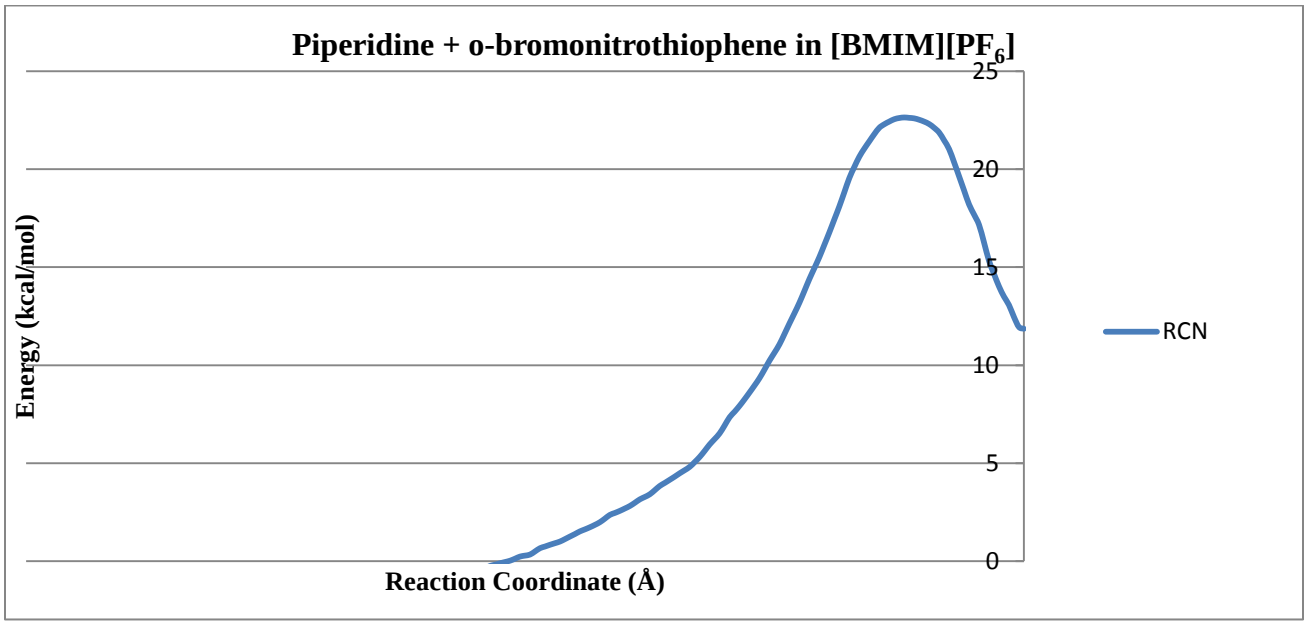


Figure 4S.50: The free energy diagram of the reaction pyrrolidine with ortho-bromonitrothiophene in [BMIM][PF_6].

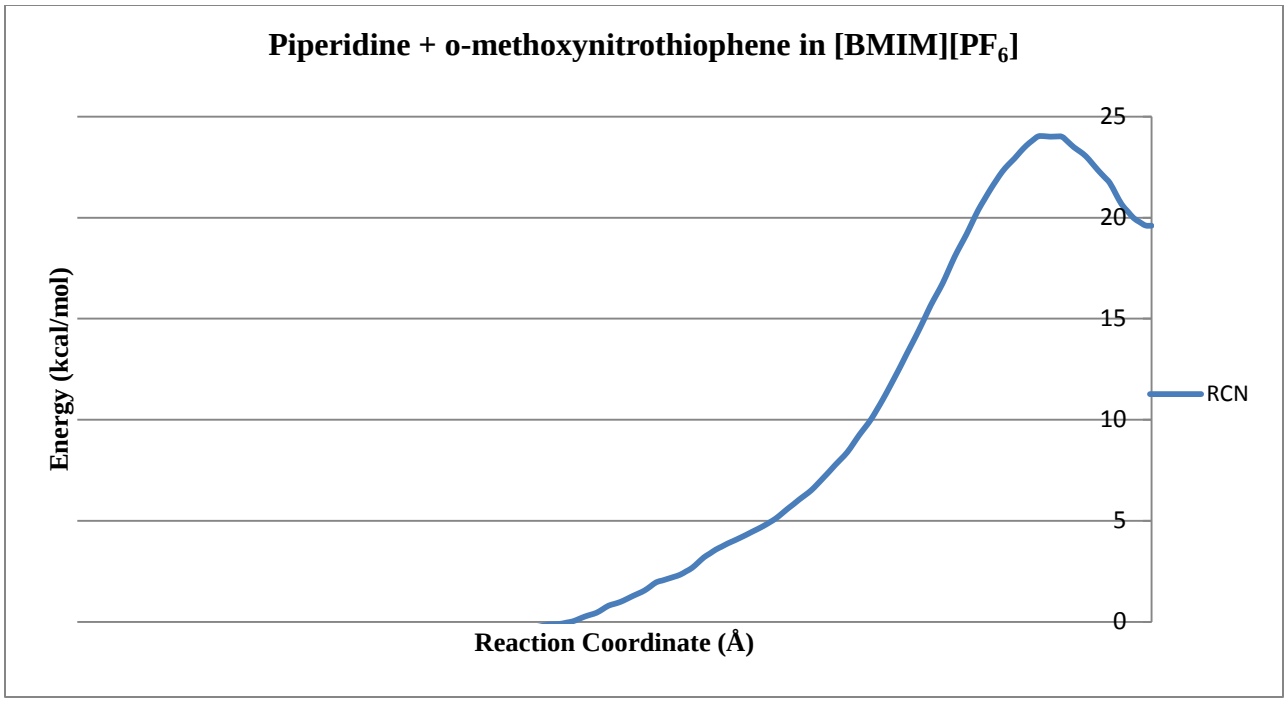


Figure 4S.51: The free energy diagram of the reaction pyrrolidine with ortho-methoxynitrothiophene in [BMIM][PF_6].

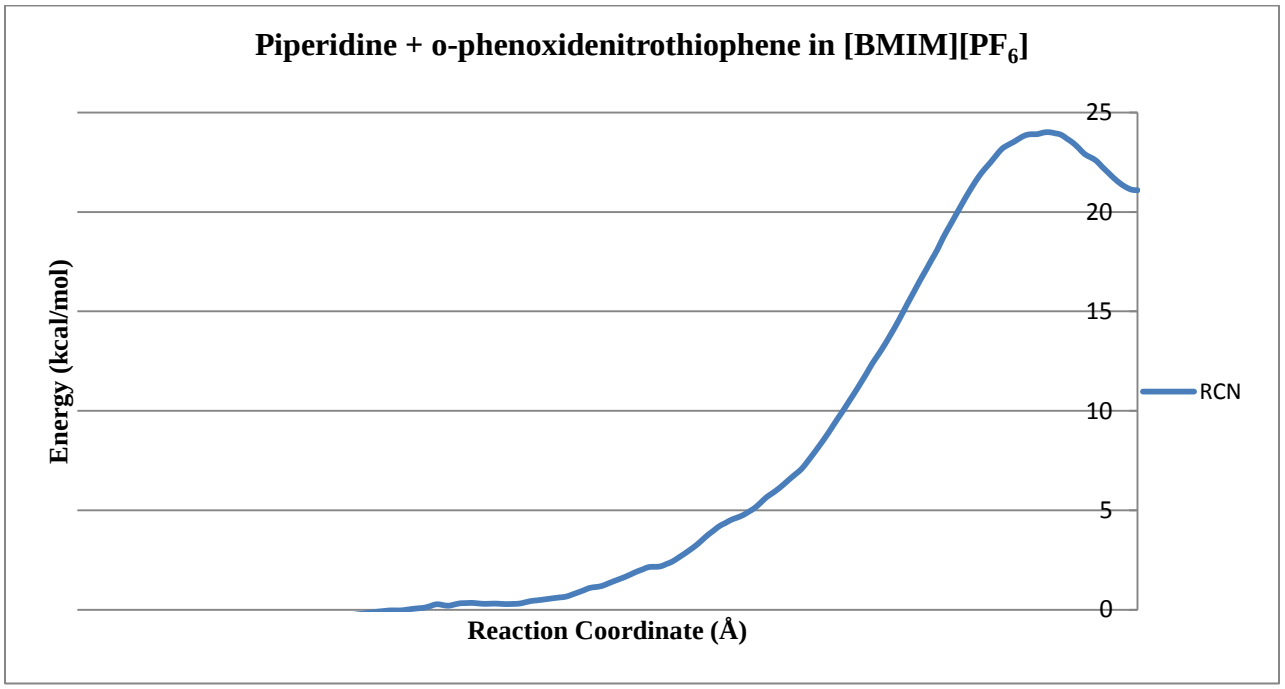


Figure 4S.52: The free energy diagram of the reaction pyrrolidine with ortho-phenoxidenitrothiophene in [BMIM][PF_6].

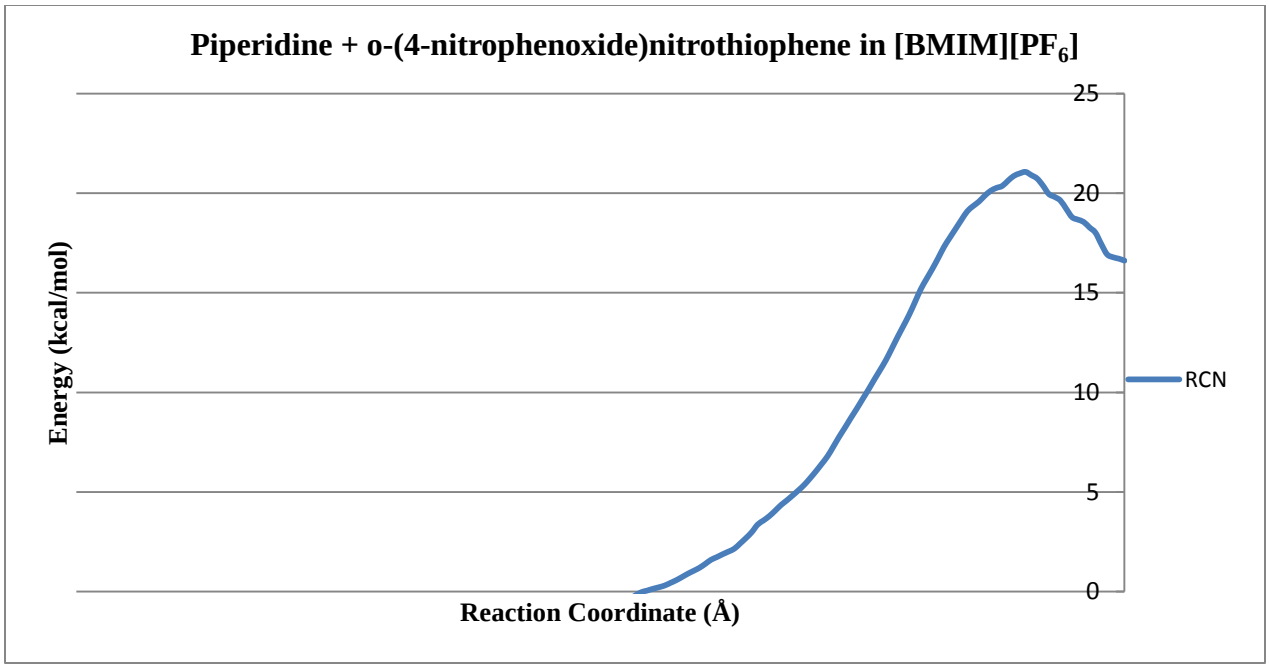


Figure 4S.53: The free energy diagram of the reaction pyrrolidine with ortho-(4-nitrophenoxide)nitrothiophene in [BMIM][PF_6].

Morpholine + para-LG-nitrothiophene

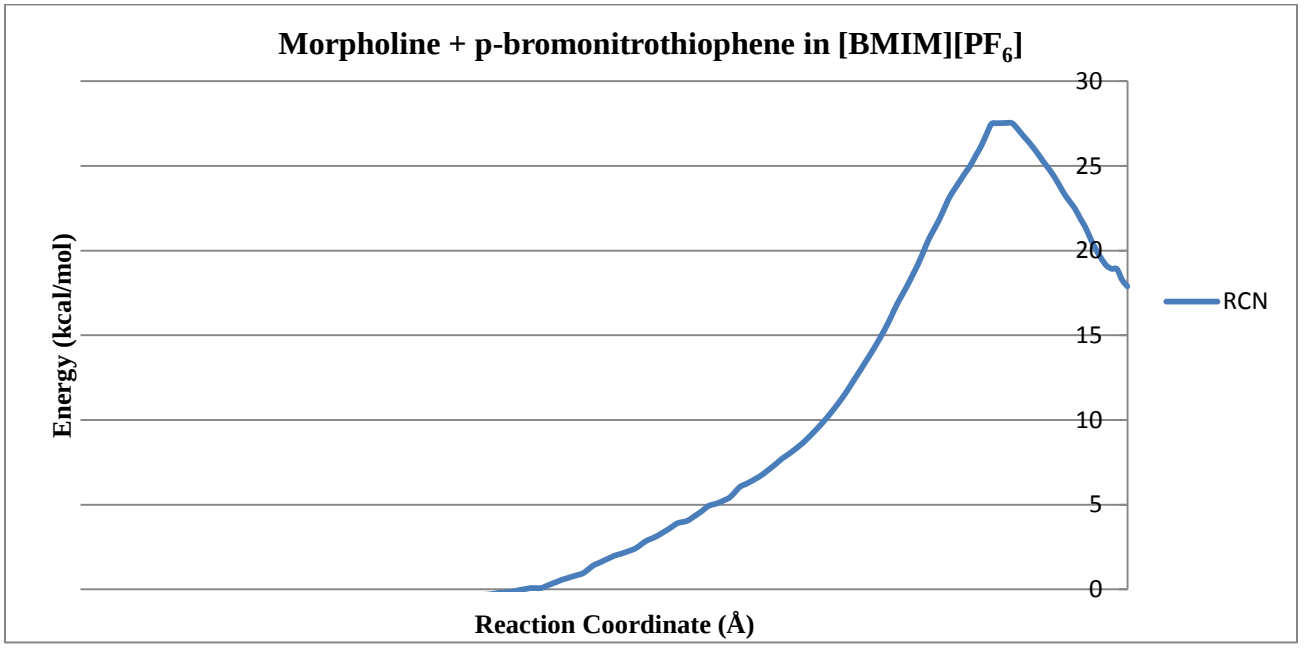


Figure 3S.54: The free energy diagram of the reaction morpholine with para-bromonitrothiophene in [BMIM][PF_6].

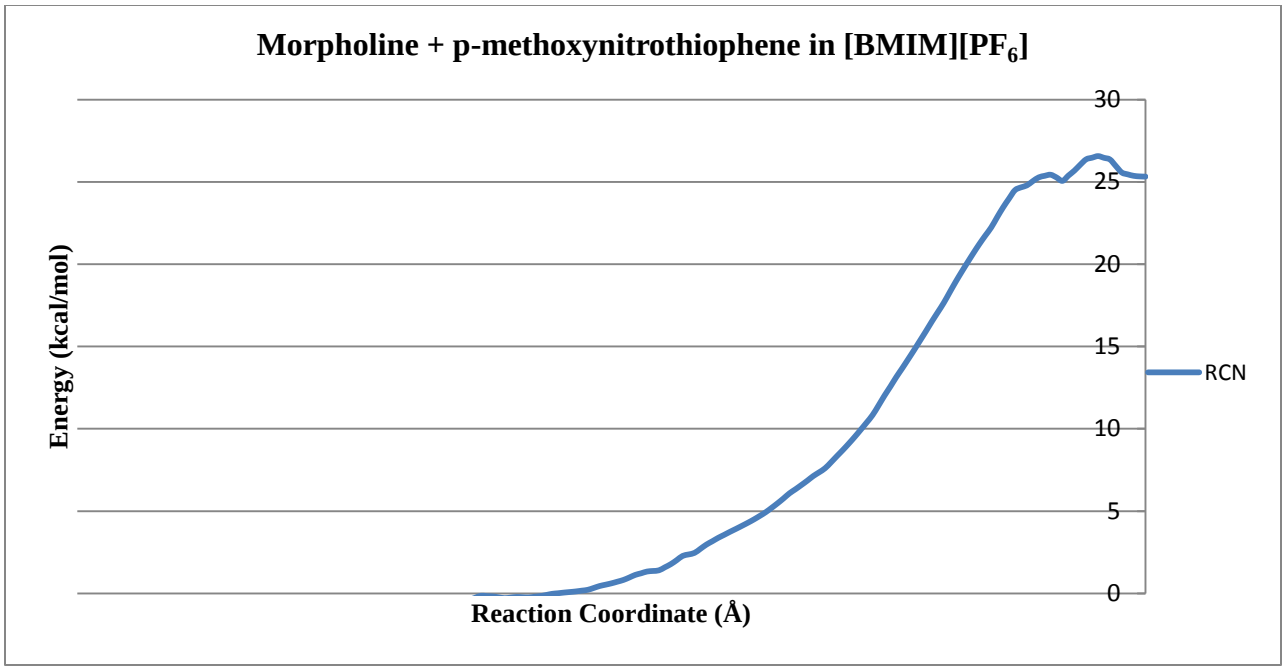


Figure 4S.55: The free energy diagram of the reaction morpholine with para-methoxynitrothiophene in [BMIM][PF_6].

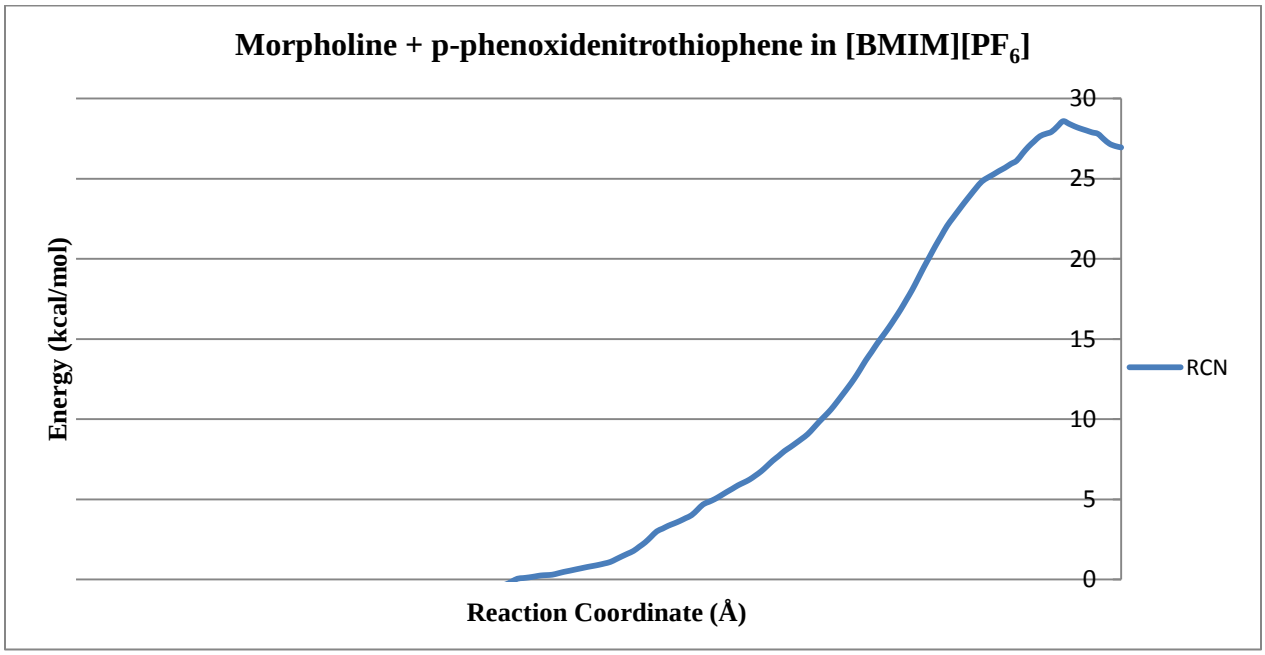


Figure 4S.56: The free energy diagram of the reaction morpholine with para-phenoxidenitrothiophene in [BMIM][PF_6].

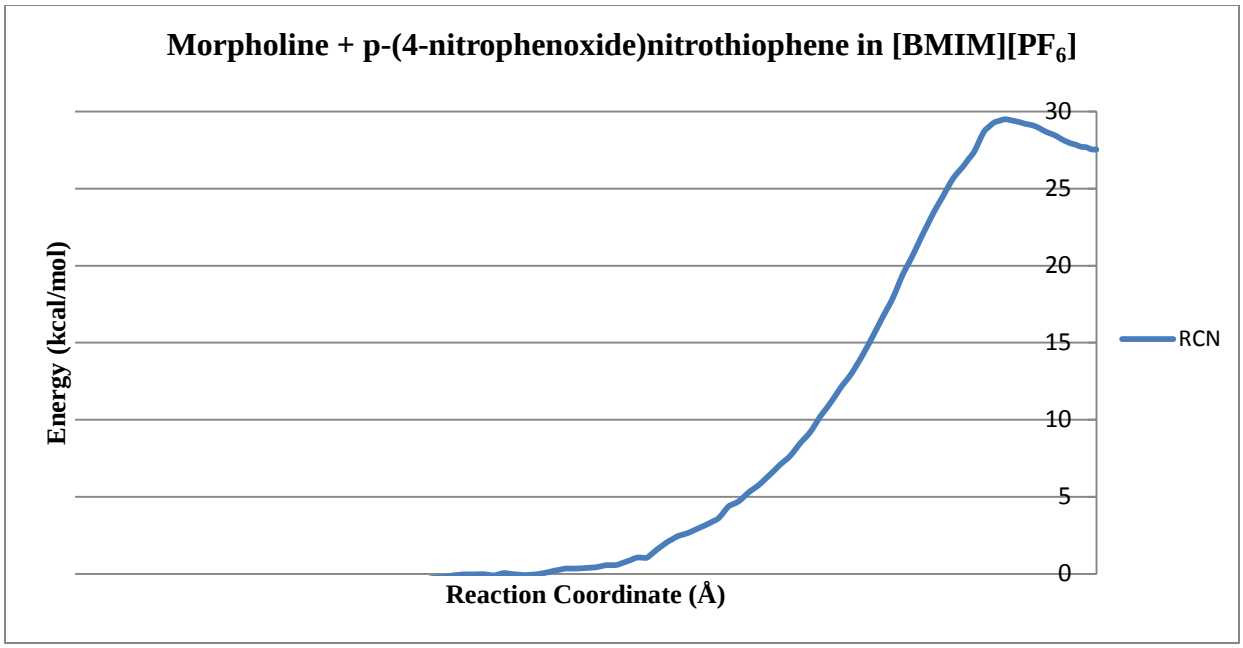


Figure 4S.57: The free energy diagram of the reaction morpholine with para-(4-nitrophenoxide)nitrothiophene in [BMIM][PF_6].

Morpholine + ortho-LG-nitrothiophene

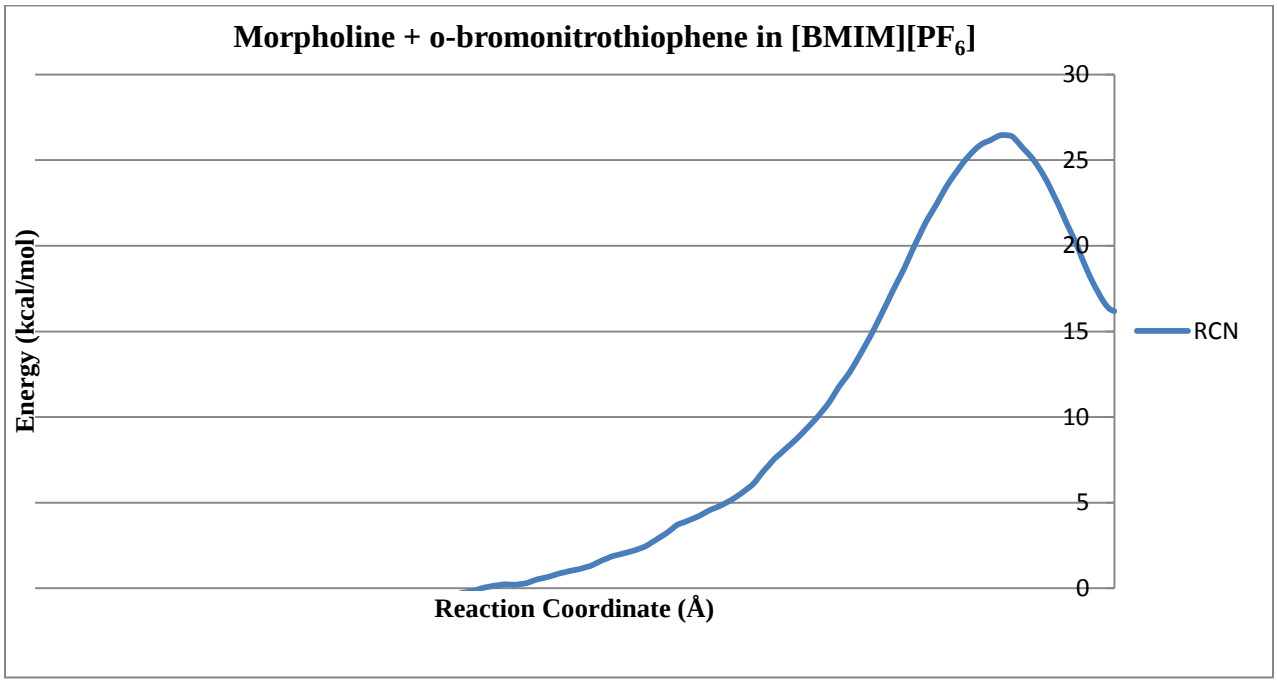


Figure 4S.58: The free energy diagram of the reaction morpholine with ortho-bromonitrothiophene in [BMIM][PF_6].

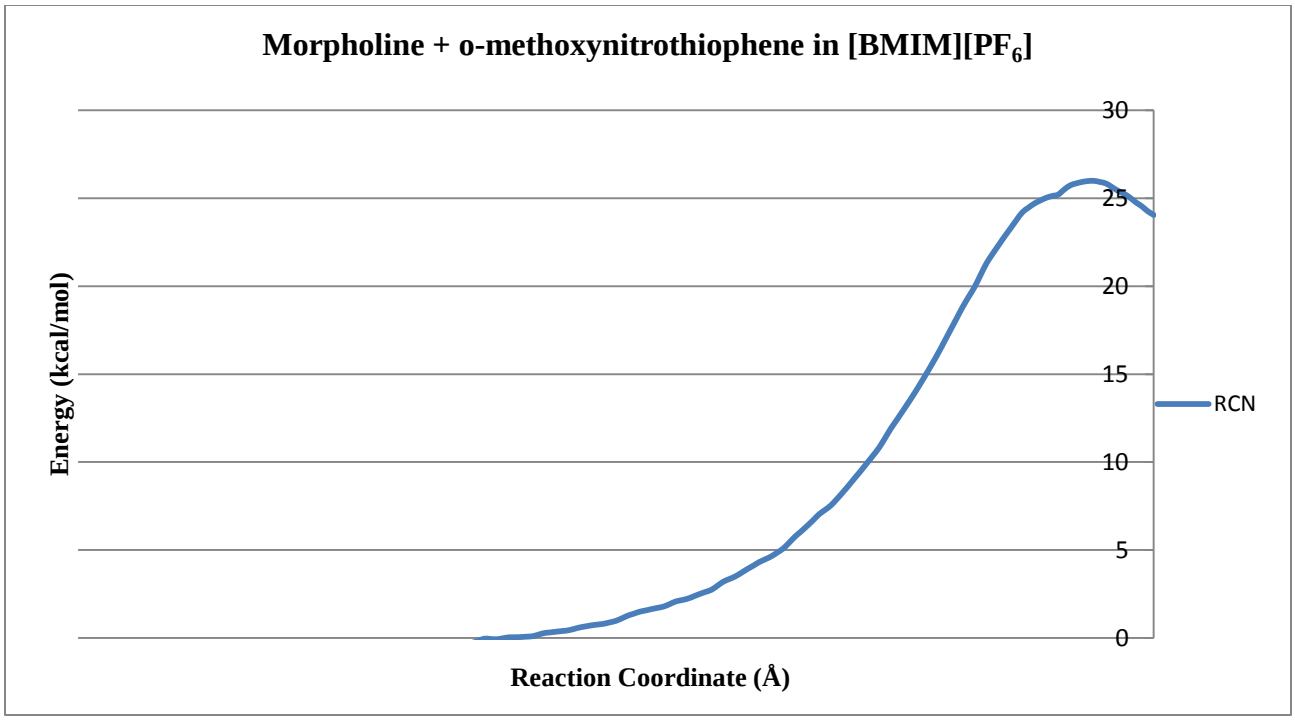


Figure 4S.59: The free energy diagram of the reaction morpholine with ortho-methoxynitrothiophene in [BMIM][PF_6].

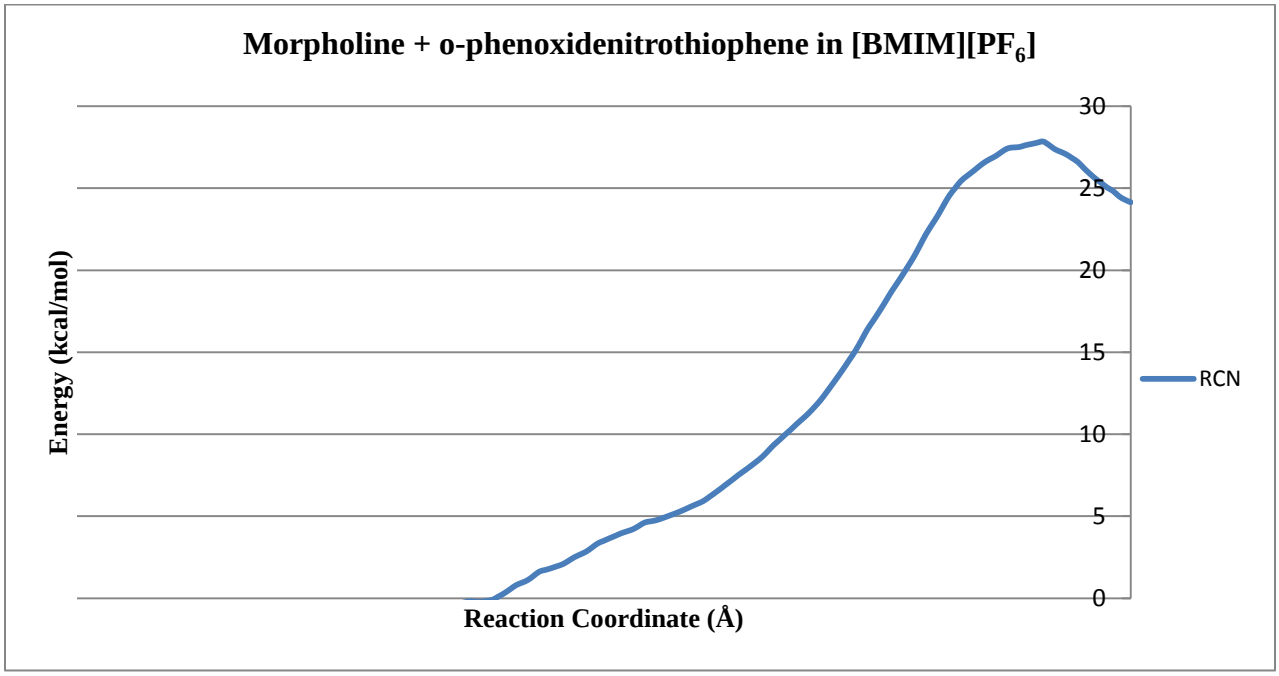


Figure 4S.60: The free energy diagram of the reaction morpholine with ortho-phenoxidenitrothiophene in [BMIM][PF_6].

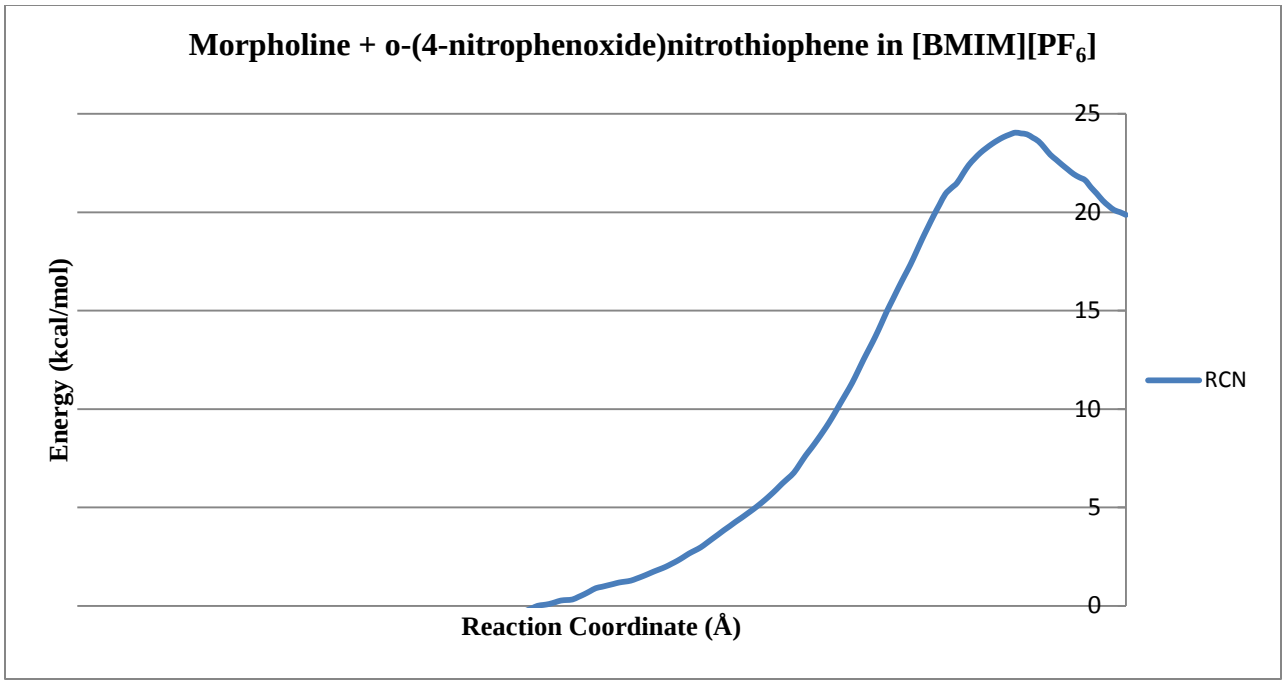


Figure 4S.61: The free energy diagram of the reaction morpholine with ortho-(4-nitrophenoxide)nitrothiophene in [BMIM][PF_6].

Pyrrolidine + para-LG-nitrothiophene

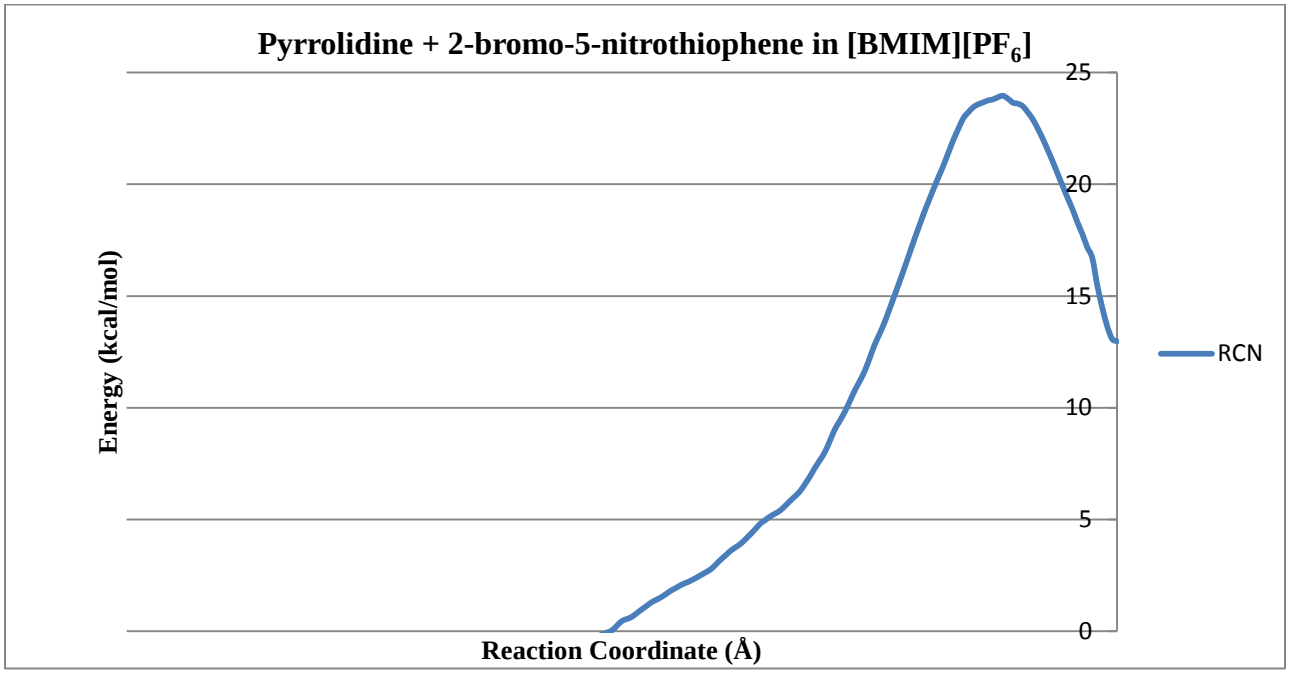


Figure 4S.62: The free energy diagram of the reaction pyrrolidine with para-bromonitrothiophene in [BMIM][PF_6].

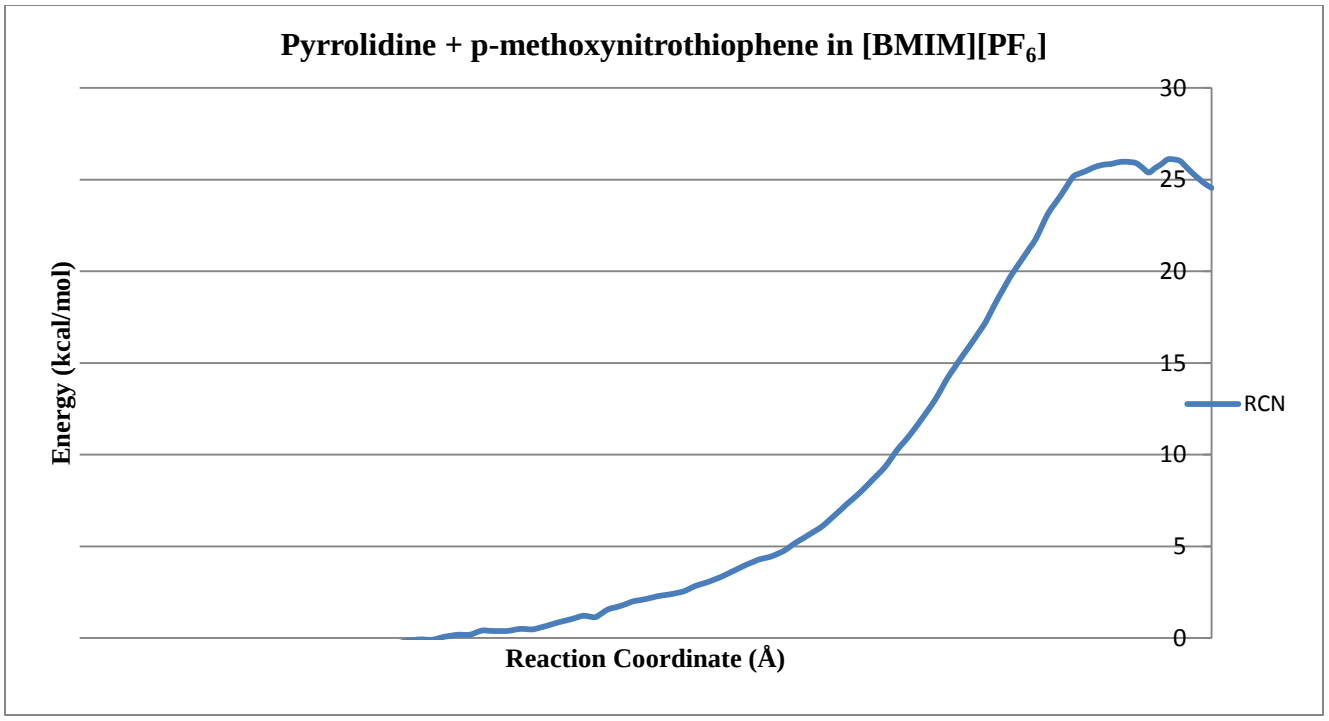


Figure 4S.63: The free energy diagram of the reaction pyrrolidine with para-methoxynitrothiophene in [BMIM][PF_6].

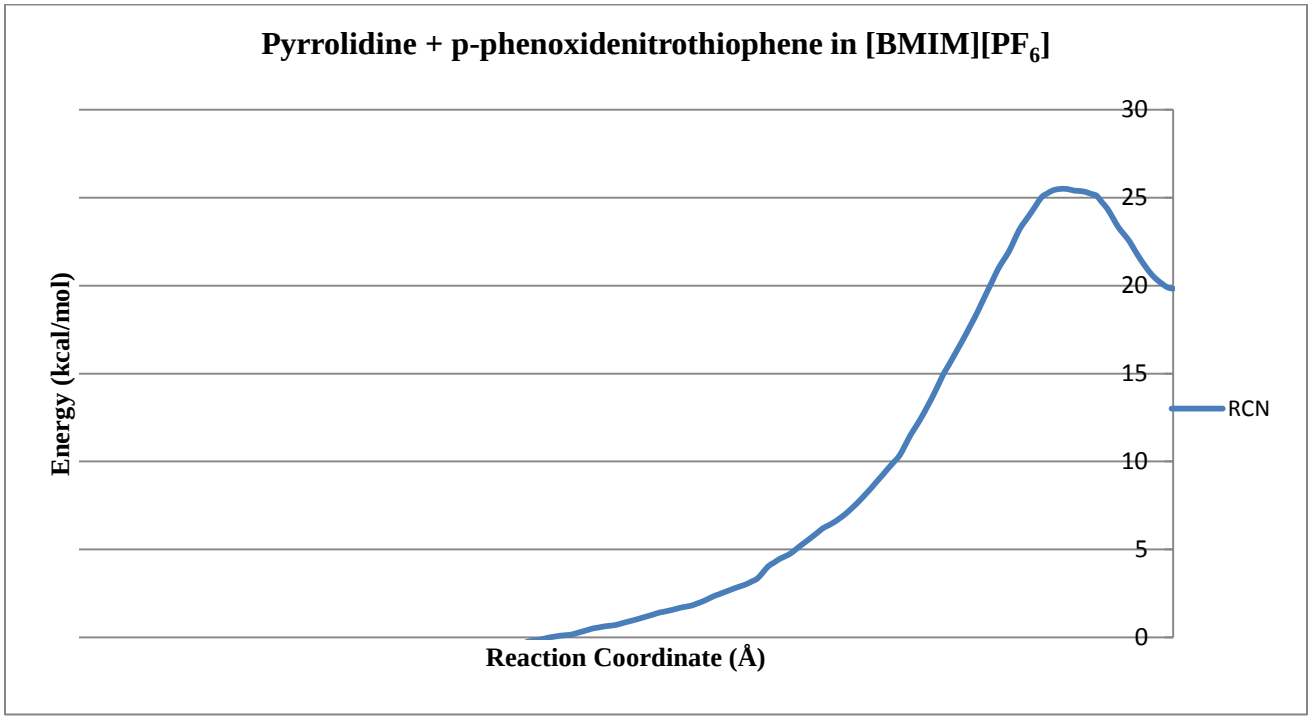


Figure 4S.64: The free energy diagram of the reaction pyrrolidine with para-phenoxidenitrothiophene in [BMIM][PF_6].

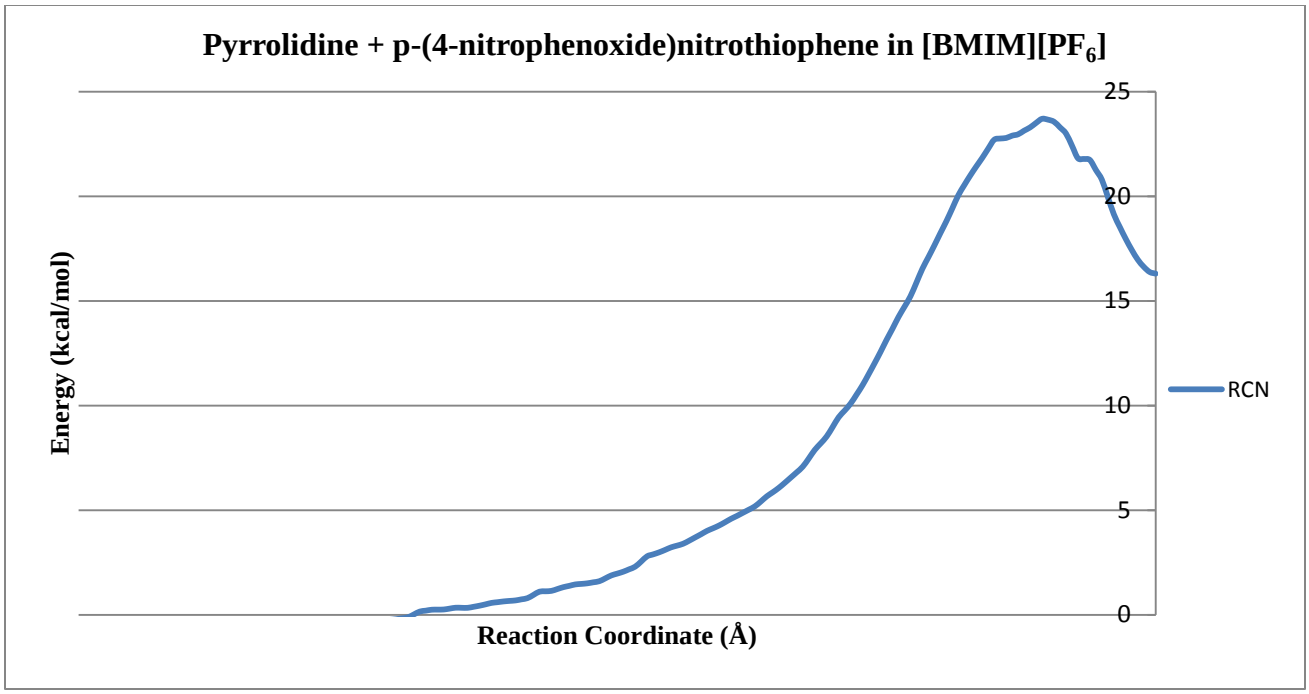


Figure 4S.65: The free energy diagram of the reaction pyrrolidine with para-(4-nitrophenoxide)nitrothiophene in [BMIM][PF_6].

Pyrrolidine + ortho-LG-nitrothiophene

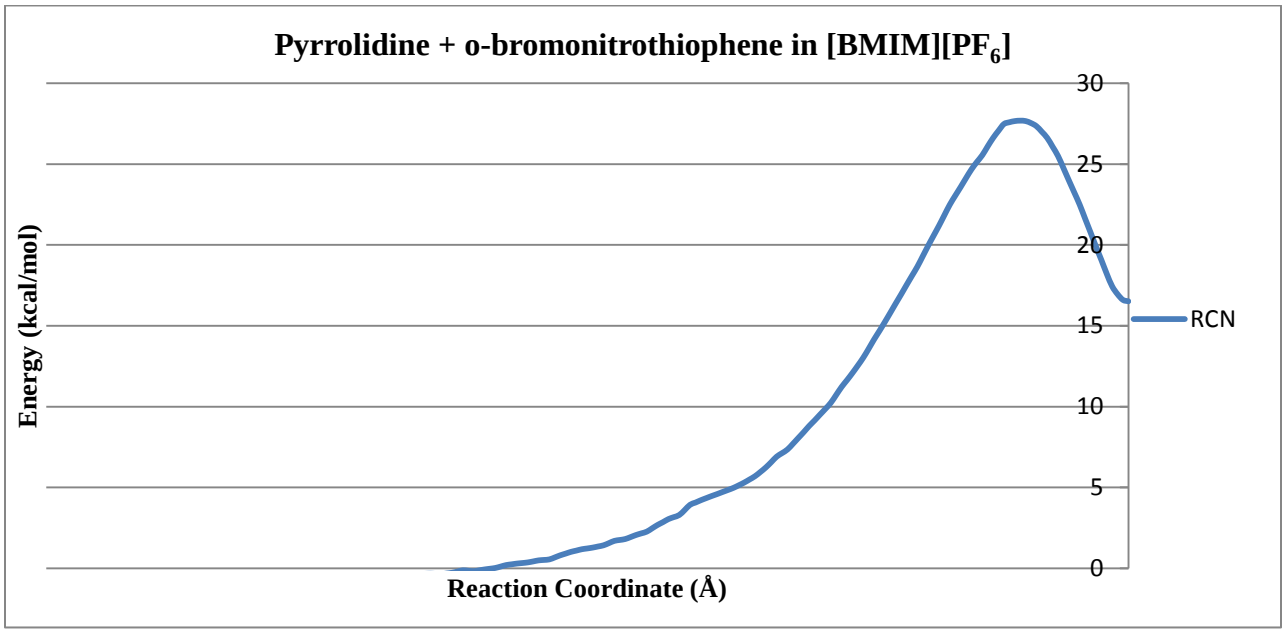


Figure 4S.66: The free energy diagram of the reaction pyrrolidine with ortho-bromonitrothiophene in [BMIM][PF_6].

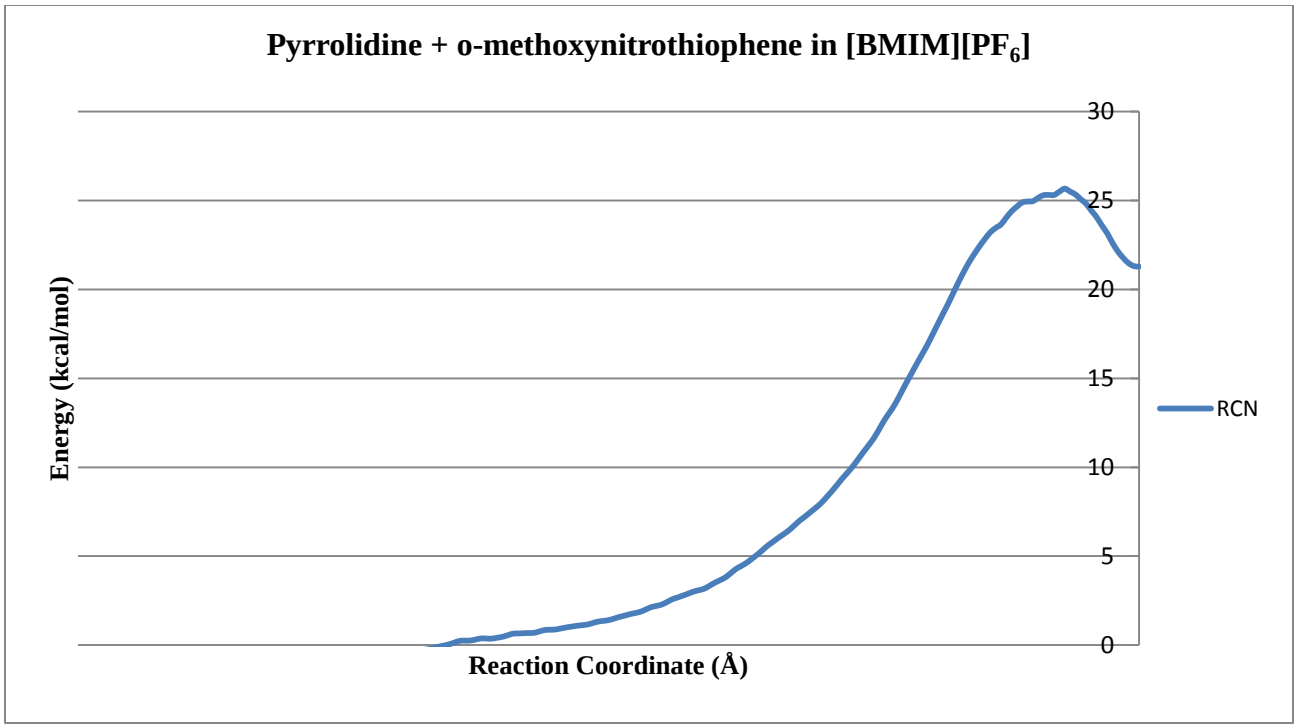


Figure 4S.67: The free energy diagram of the reaction pyrrolidine with ortho-methoxynitrothiophene in [BMIM][PF_6].

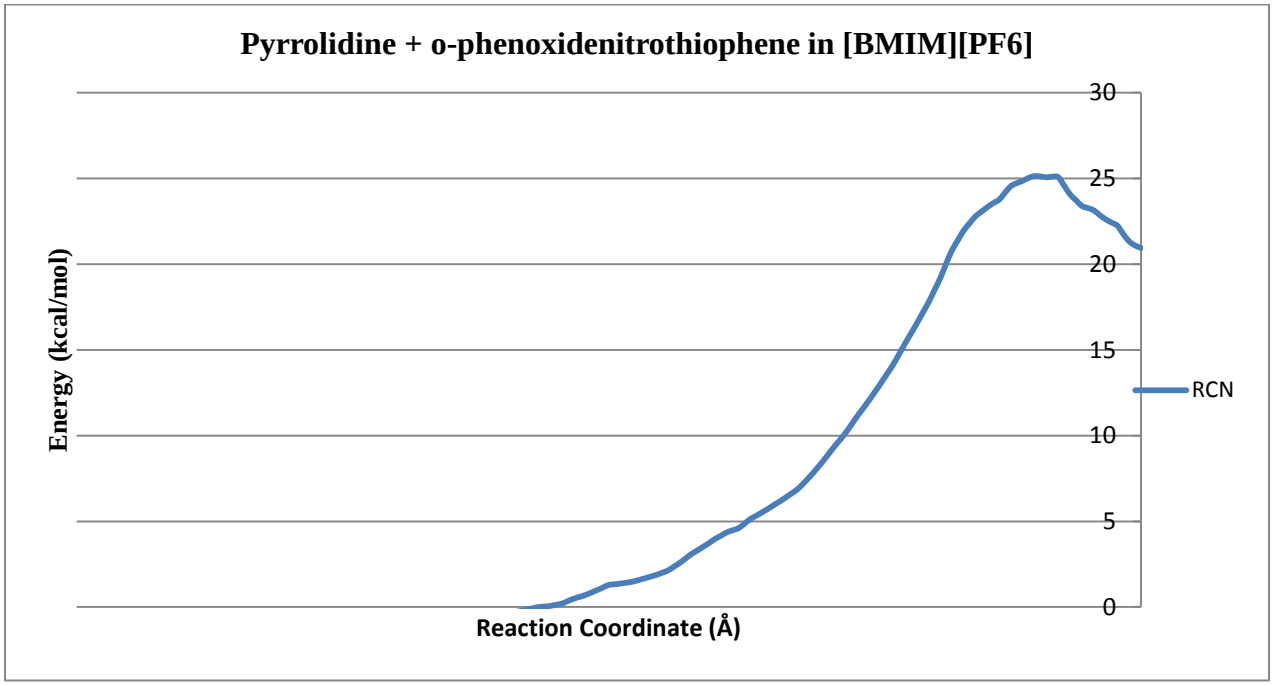


Figure 4S.68: The free energy diagram of the reaction pyrrolidine with ortho-phenoxidenitrothiophene in [BMIM][PF_6].

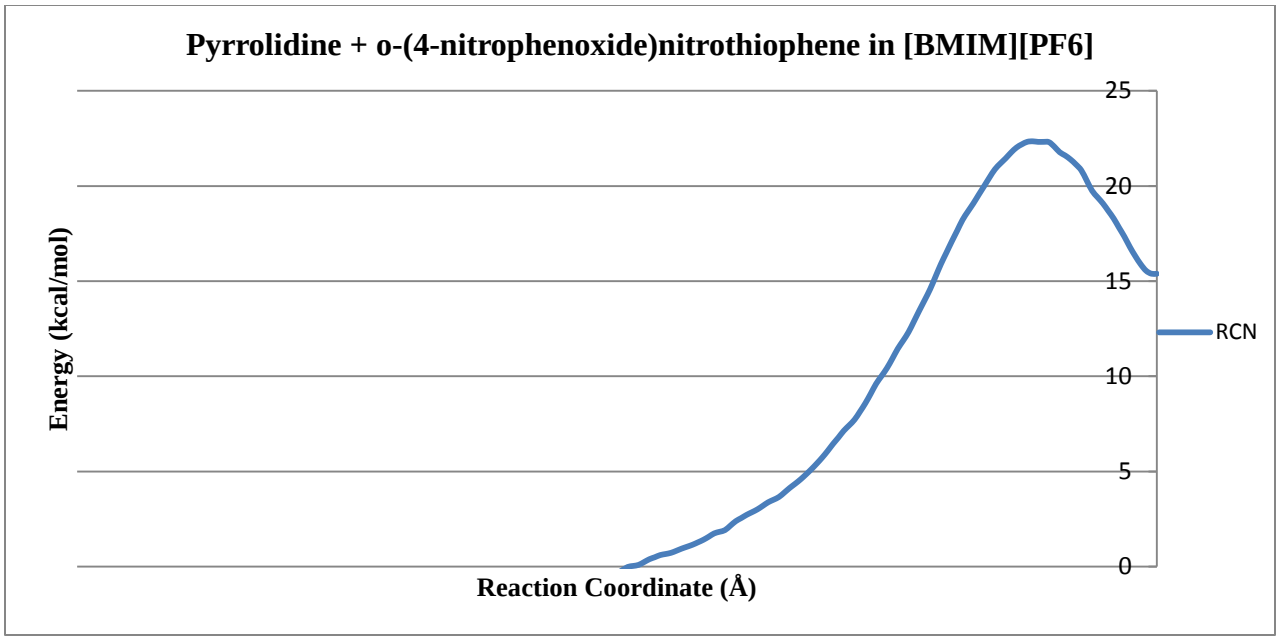


Figure 4S.69: The free energy diagram of the reaction pyrrolidine with ortho-(4-nitrophenoxide)nitrothiophene in [BMIM][PF_6].

Energy Pair Distributions of the GS and TS in methanol

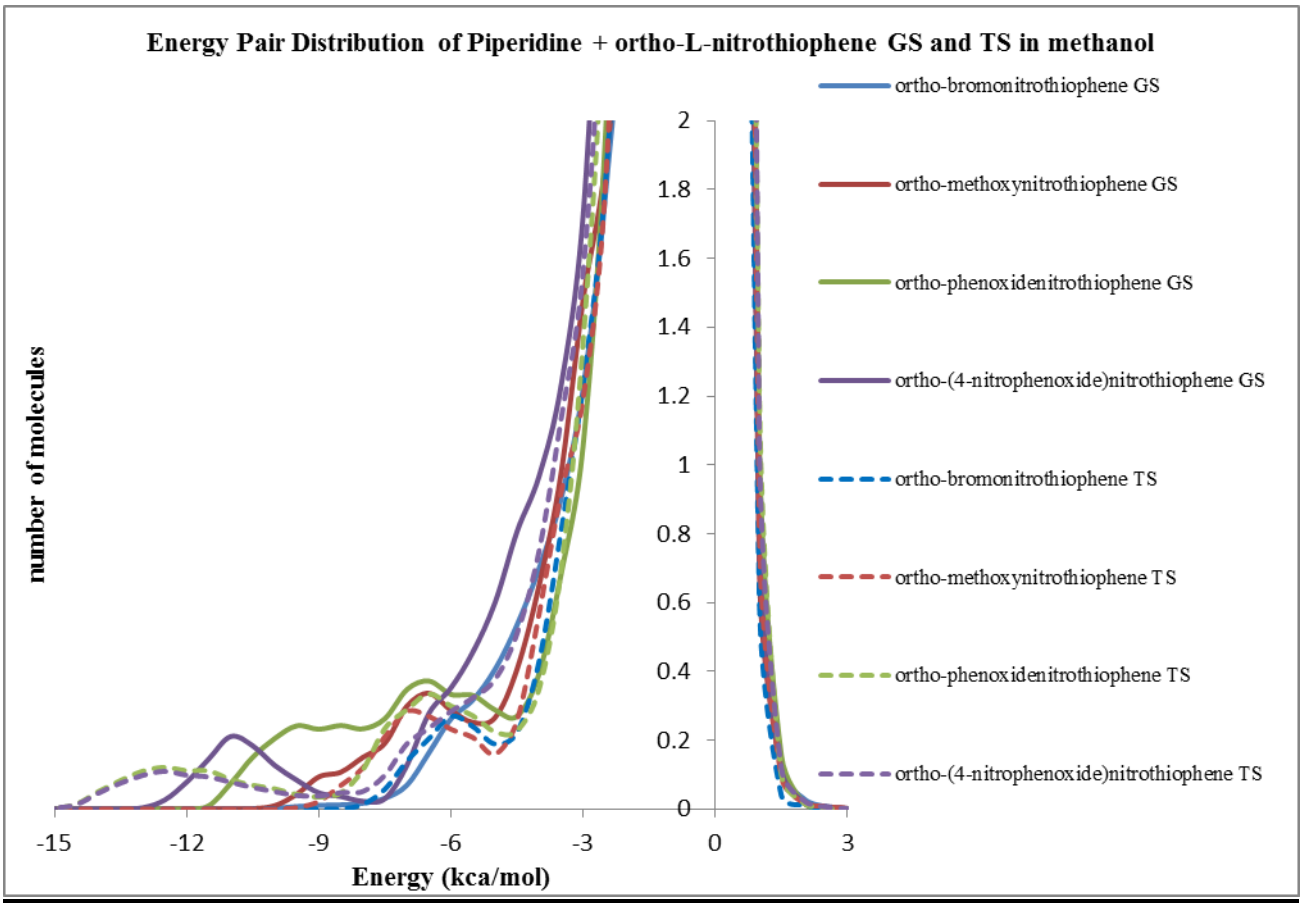


Figure 4S.70: The solute-solvent energy pair distributions of the ground state (GS – solid line) and first transition state (TS – dashed line) for the S_NAr reactions between the nucleophile piperidine and 2-L-3-nitrothiophene in methanol.

	-5 kcal/mol	**-3.5 kcal/mol**
ortho-bromonitrothiophene GS	1.30	3.44
ortho-bromonitrothiophene TS	1.11	2.57
ortho-methoxynitrothiophene GS	2.04	4.07
ortho-methoxynitrothiophene TS	1.58	3.32
ortho-phenoxidenitrothiophene GS	3.32	4.67
ortho-phenoxidenitrothiophene TS	2.71	3.98
ortho-(4-nitrophenoxide)nitrothiophene GS	2.81	5.82
ortho-(4-nitrophenoxide)nitrothiophene TS	2.41	4.81

Table 4S.1: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

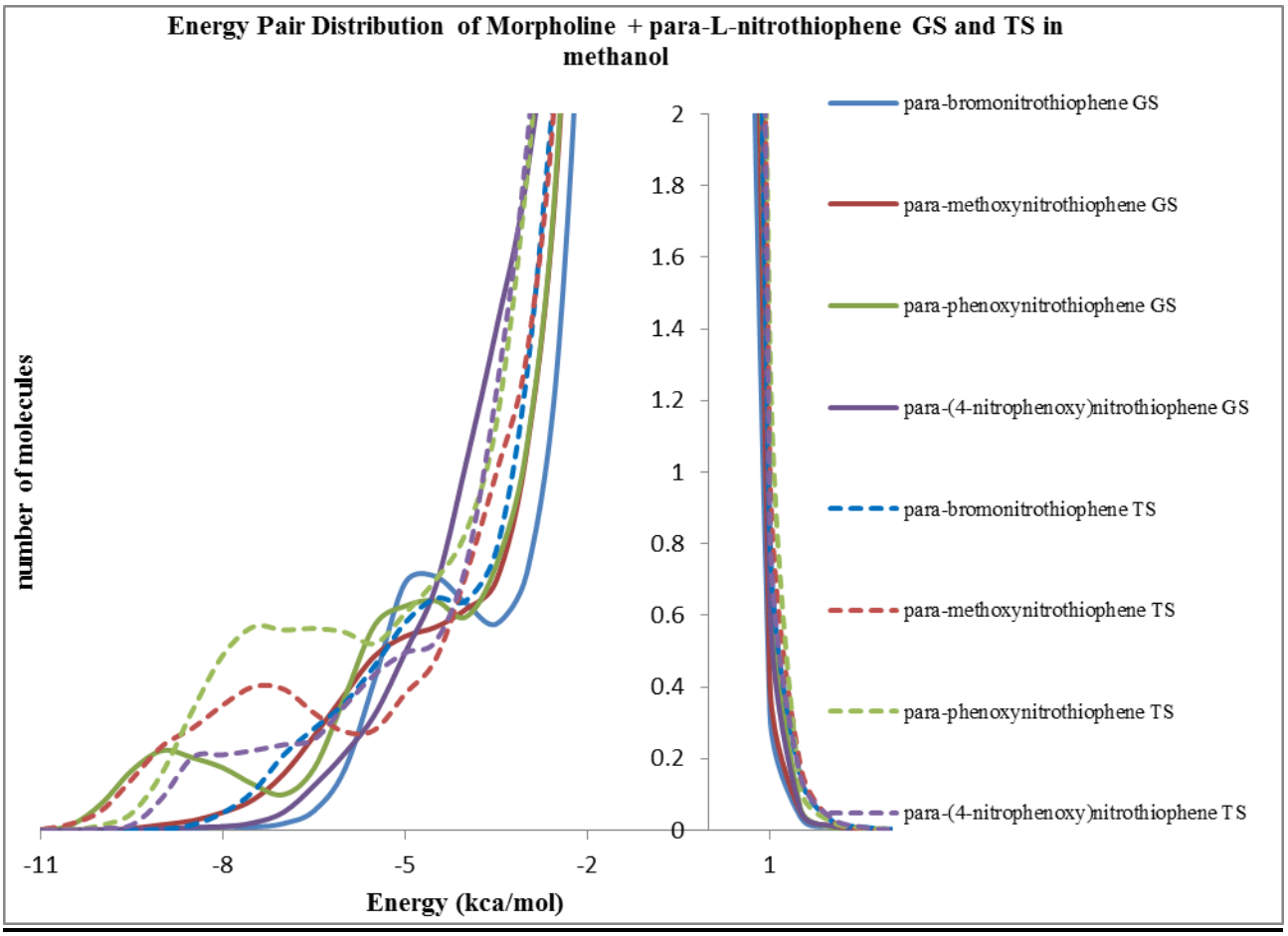


Figure 4S.71: The solute-solvent energy pair distributions of the ground state (GS – solid line) and first transition state (TS – dashed line) for the S_NAr reactions between the nucleophile morpholine and 2-L-5-nitrothiophene in methanol.

	-5 kcal/mol	-3.5 kcal/mol
para-bromonitrothiophene GS	1.38	3.30
para-bromonitrothiophene TS	2.06	4.14
para-methoxynitrothiophene GS	2.01	3.89
para-methoxynitrothiophene TS	3.15	5.34
para-phenoxynitrothiophene GS	2.83	4.80
para-phenoxynitrothiophene TS	4.42	7.09
para-(4-nitrophenoxy)nitrothiophene GS	1.26	4.36
para-(4-nitrophenoxy)nitrothiophene TS	2.52	5.02

Table 4S.2: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

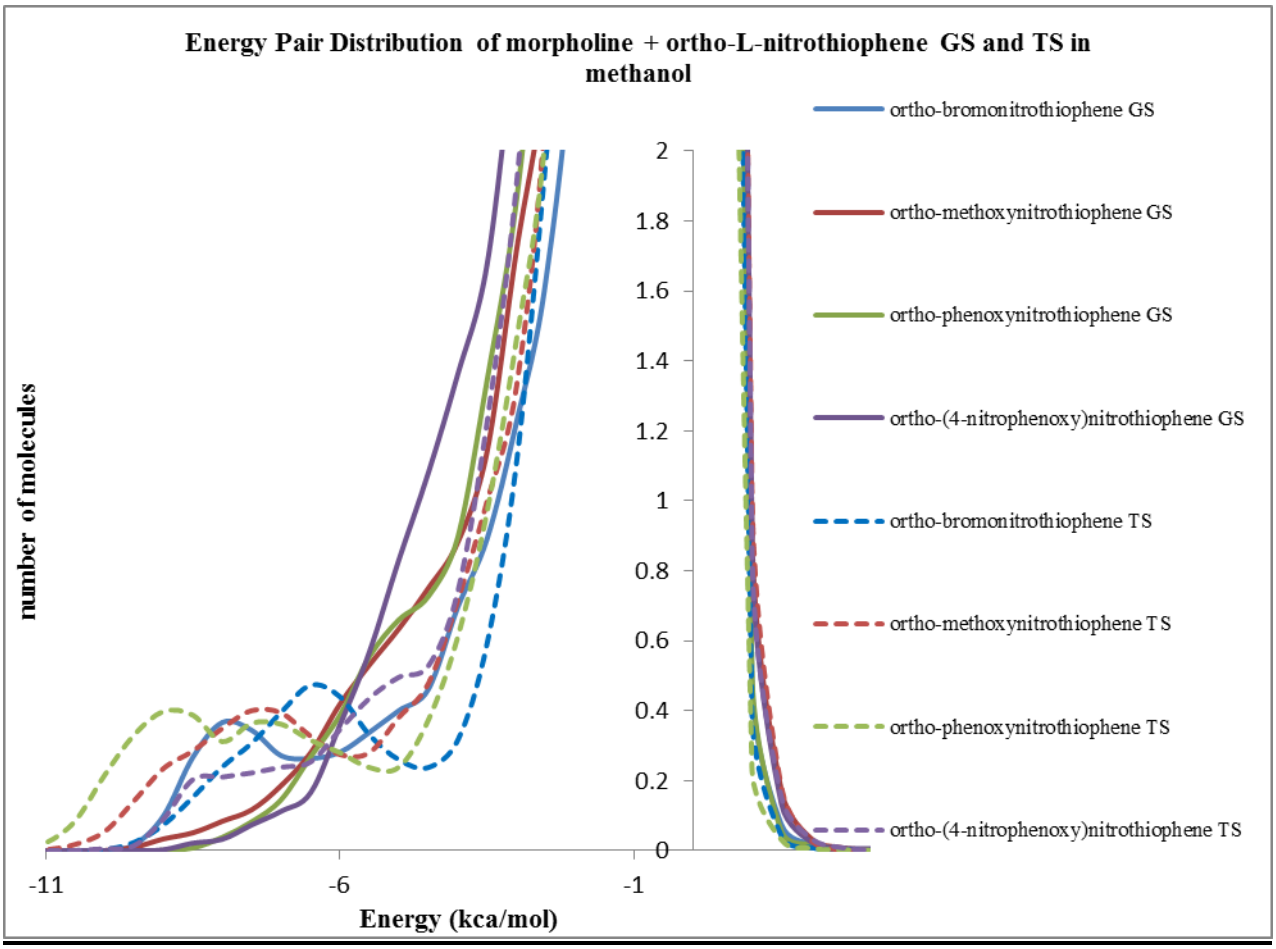


Figure 4S.72: The solute-solvent energy pair distributions of the ground state (GS – solid line) and first transition state (TS – dashed line) for the S_NAr reactions between the nucleophile morpholine and 2-L-3-nitrothiophene in methanol.

	-5 kcal/mol	-3.5 kcal/mol
ortho-bromonitrothiophene GS	2.65	4.70
ortho-bromonitrothiophene TS	2.70	3.85
ortho-methoxynitrothiophene GS	2.35	5.12
ortho-methoxynitrothiophene TS	3.15	5.34
ortho-phenoxynitrothiophene GS	2.15	5.12
ortho-phenoxynitrothiophene TS	3.55	5.51
ortho-(4-nitrophenoxy)nitrothiophene GS	2.17	6.29
ortho-(4-nitrophenoxy)nitrothiophene TS	2.52	5.02

Table 4S.3: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

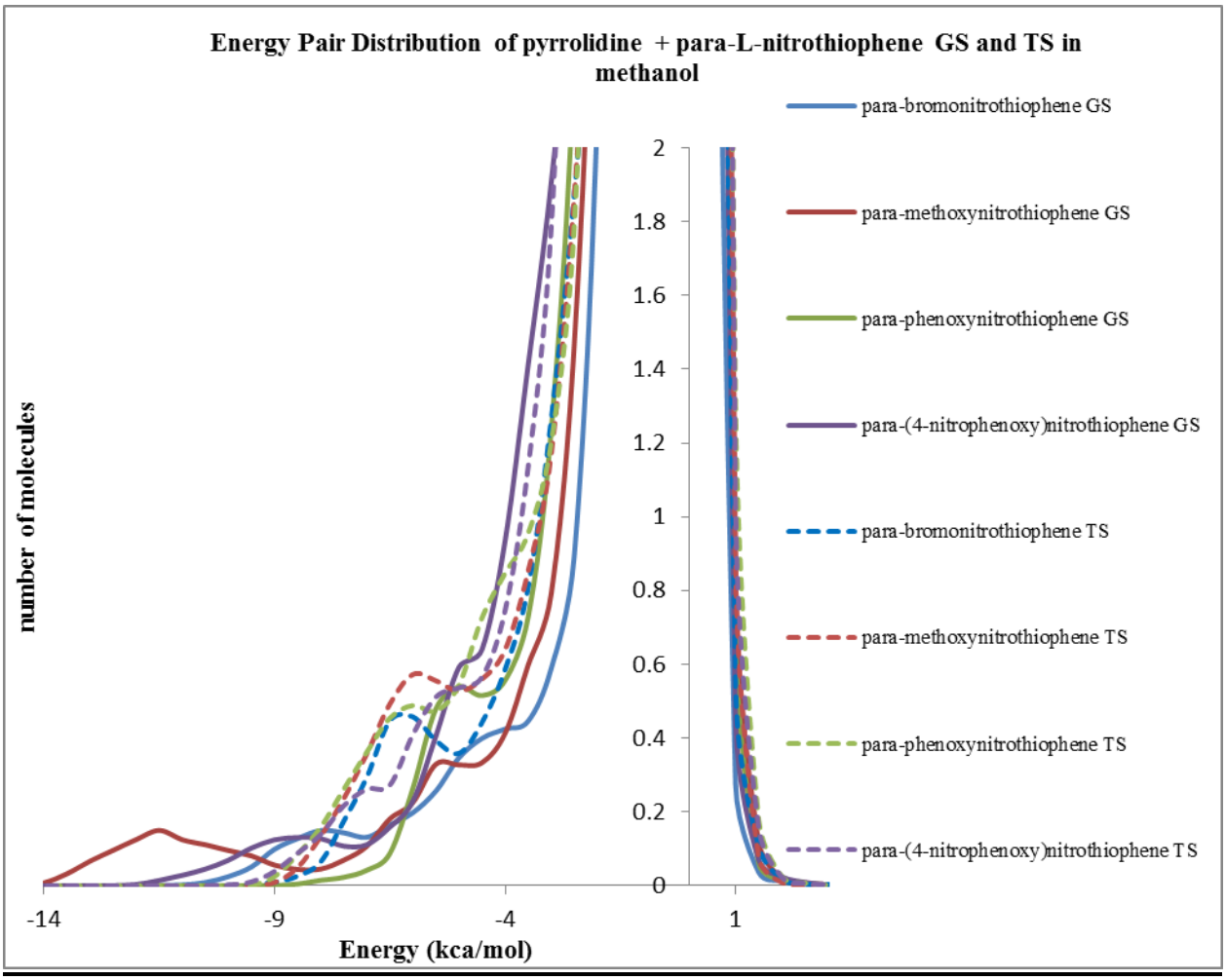


Figure 4S.73: The solute-solvent energy pair distributions of the ground state (GS – solid line) and first transition state (TS – dashed line) for the S_NAr reactions between the nucleophile pyrrolidine and 2-L-5-nitrothiophene in methanol.

	-5 kcal/mol	**-3.5 kcal/mol**
para-bromonitrothiophene GS	1.70	2.97
para-bromonitrothiophene TS	2.23	4.07
para-methoxynitrothiophene GS	2.27	3.61
para-methoxynitrothiophene TS	2.93	4.98
para-phenoxynitrothiophene GS	1.45	3.25
para-phenoxynitrothiophene TS	2.85	5.38
para-(4-nitrophenoxy)nitrothiophene GS	2.25	5.23
para-(4-nitrophenoxy)nitrothiophene TS	2.48	4.93

Table 4S.4: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

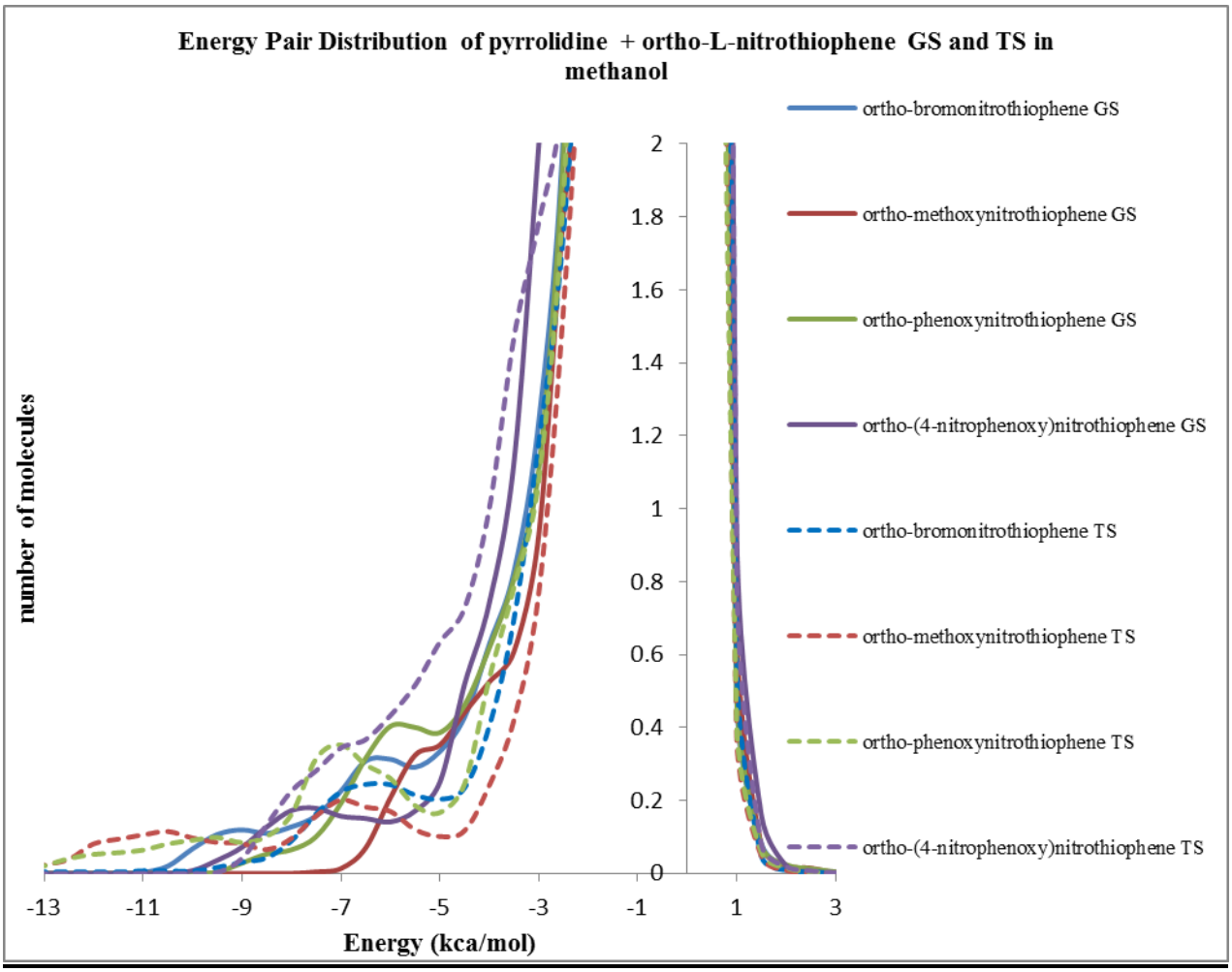


Figure 4S.74: The solute-solvent energy pair distributions of the ground state (GS – solid line) and first transition state (TS – dashed line) for the S_NAr reactions between the nucleophile pyrrolidine and 2-L-3-nitrothiophene in methanol.

	-5 kcal/mol	**-3.5 kcal/mol**
ortho-bromonitrothiophene GS	2.19	4.07
ortho-bromonitrothiophene TS	1.52	2.86
ortho-methoxynitrothiophene GS	0.97	2.55
ortho-methoxynitrothiophene TS	1.82	2.60
ortho-phenoxynitrothiophene GS	1.95	3.84
ortho-phenoxynitrothiophene TS	2.45	4.01
ortho-(4-nitrophenoxy)nitrothiophene GS	1.47	3.86
ortho-(4-nitrophenoxy)nitrothiophene TS	2.98	6.18

Table 4S.5: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

Energy Pair Distributions of the GS and TS in RTILs

Piperidine + ortho-L-nitrothiophene

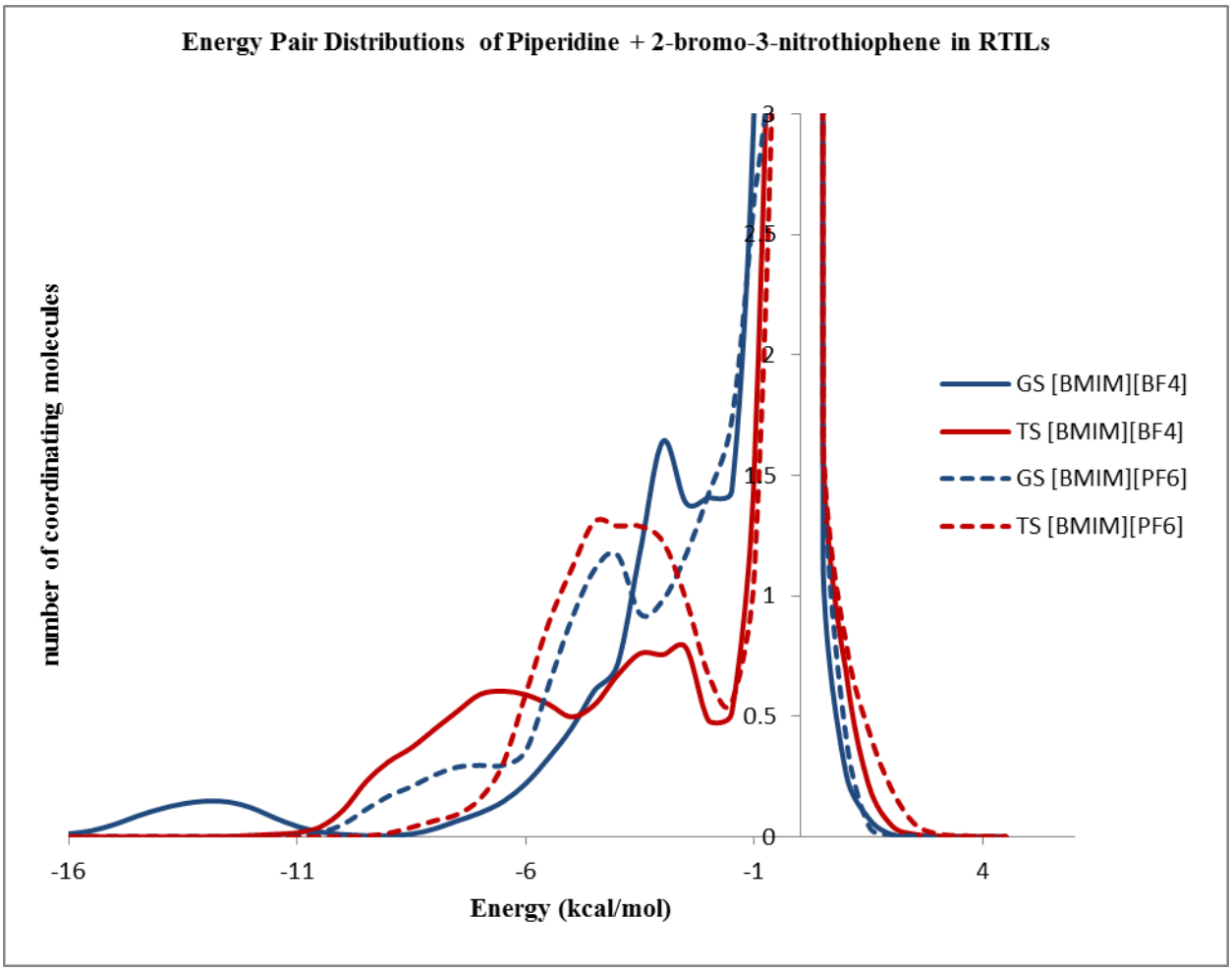


Figure 4S.75: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile piperidine and 2-bromo-3-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	-3.5 kcal/mol
GS [BMIM][BF4]	2.36	4.86
TS [BMIM][BF4]	4.89	6.87
GS [BMIM][PF6]	3.59	6.79
TS [BMIM][PF6]	3.27	7.16

Table 4S.6: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

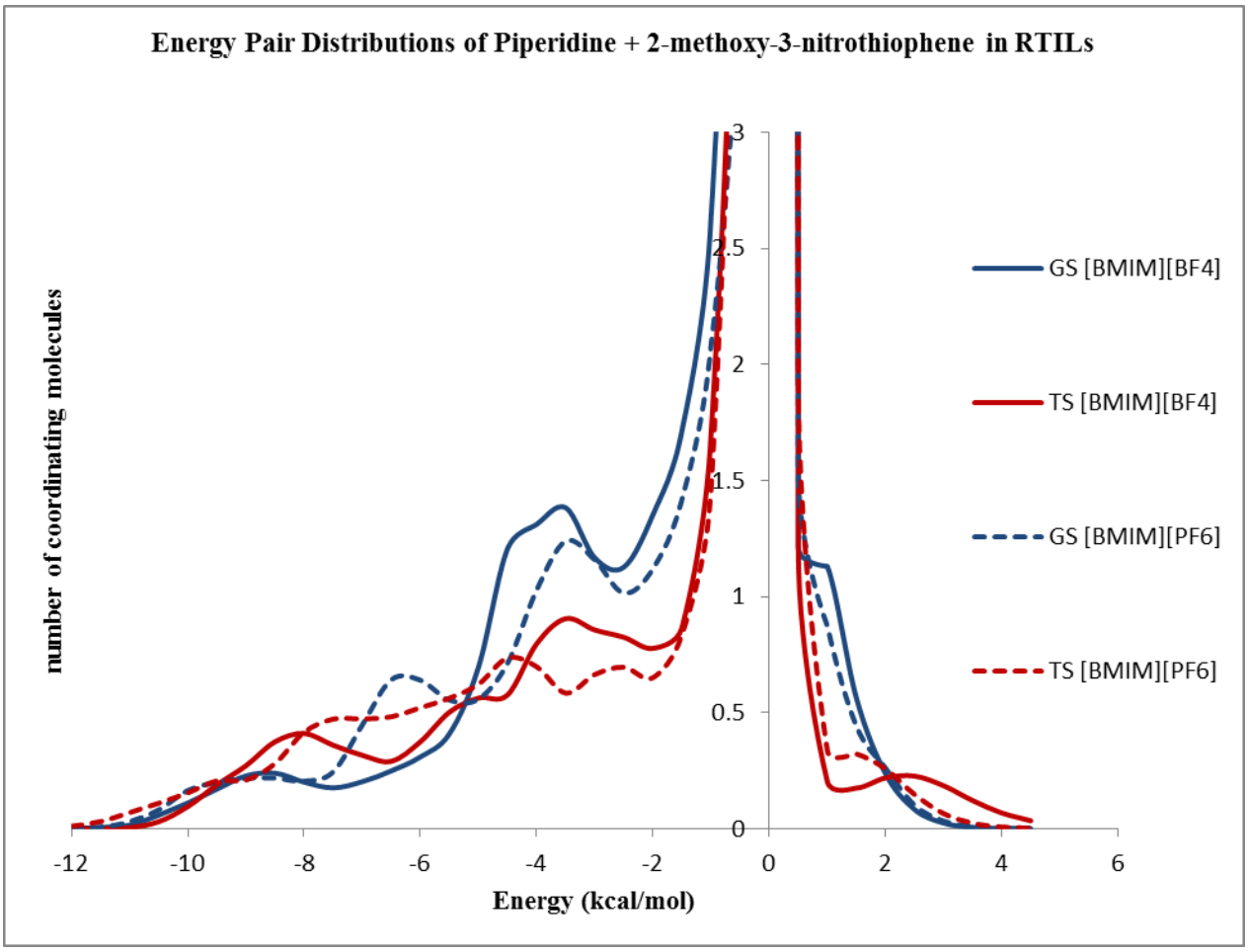


Figure 4S.76: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile piperidine and 2-methoxy-3-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	3.07	6.97
TS [BMIM][BF4]	3.78	6.06
GS [BMIM][PF6]	4.21	7.19
TS [BMIM][PF6]	4.62	6.64

Table 4S.7: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

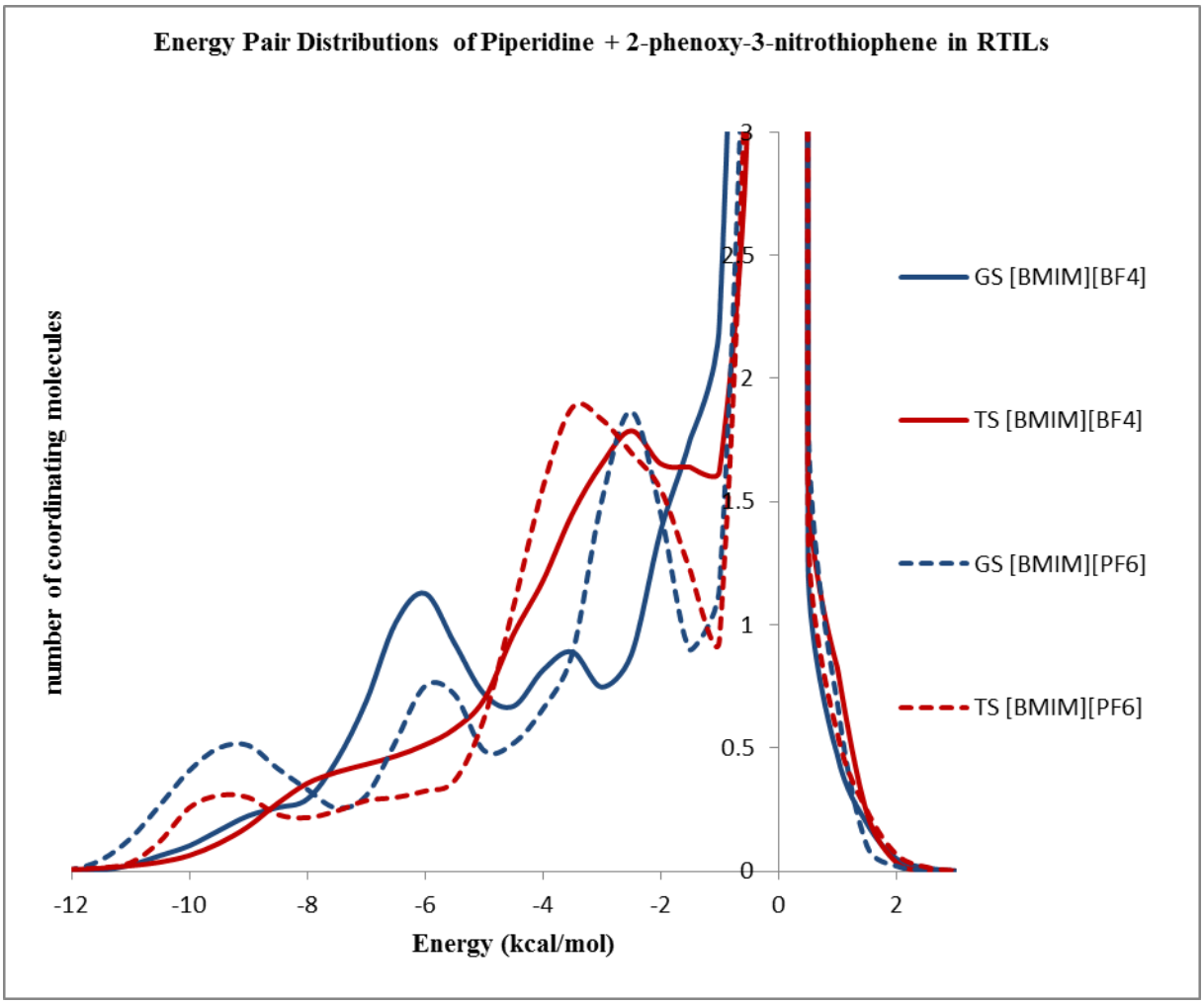


Figure 4S.77: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile piperidine and 2-phenoxy-3-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	-3.5 kcal/mol
GS [BMIM][BF4]	6.06	8.44
TS [BMIM][BF4]	4.17	7.77
GS [BMIM][PF6]	5.68	7.74
TS [BMIM][PF6]	3.65	8.16

Table 4S.8: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

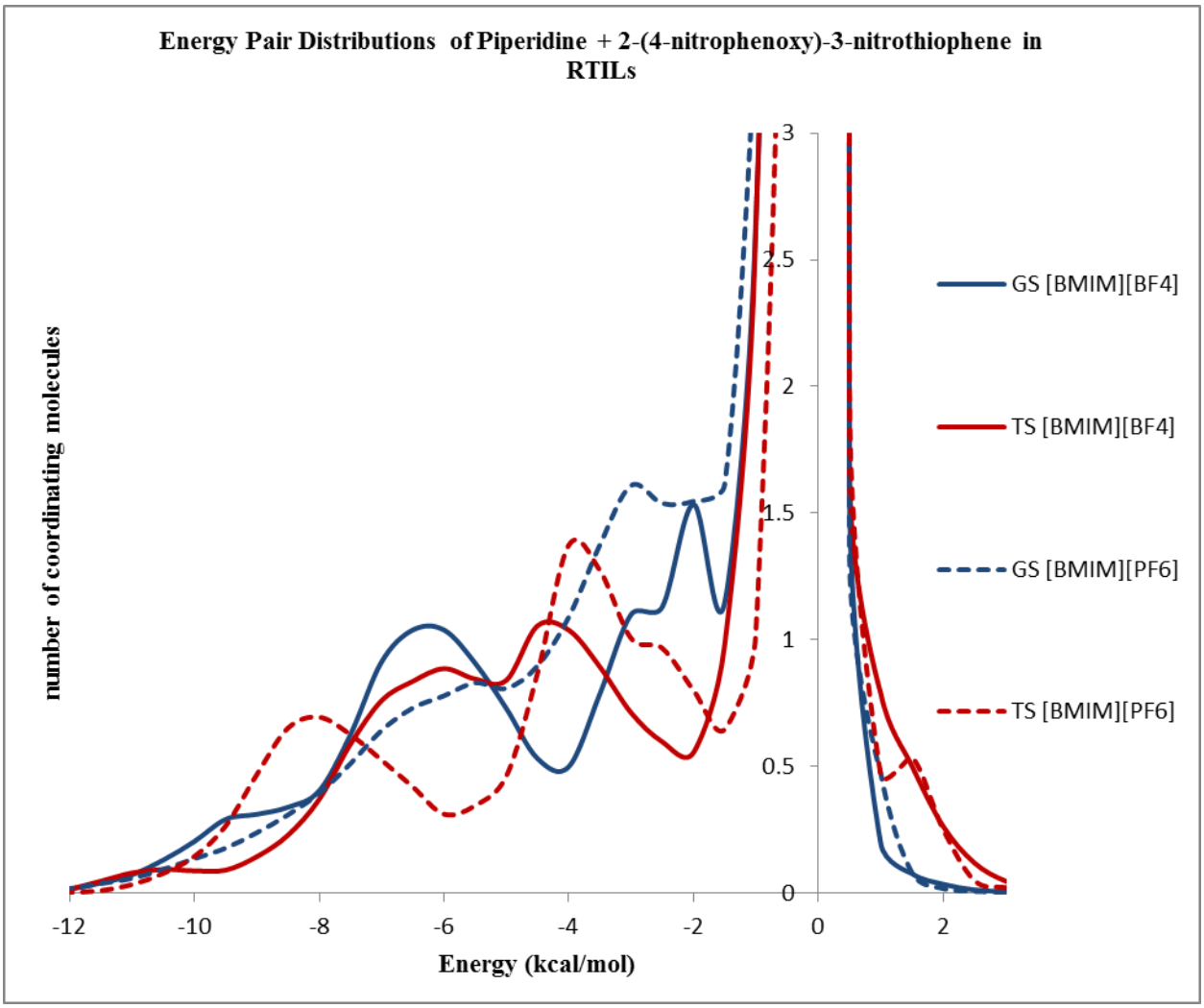


Figure 4S.78: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile piperidine and 2-(4-nitrophenoxy)-3-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	7.07	8.88
TS [BMIM][BF4]	5.92	8.90
GS [BMIM][PF6]	5.77	9.13
TS [BMIM][PF6]	5.03	8.55

Table 4S.9: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

Morpholine + para-L-nitrothiophene

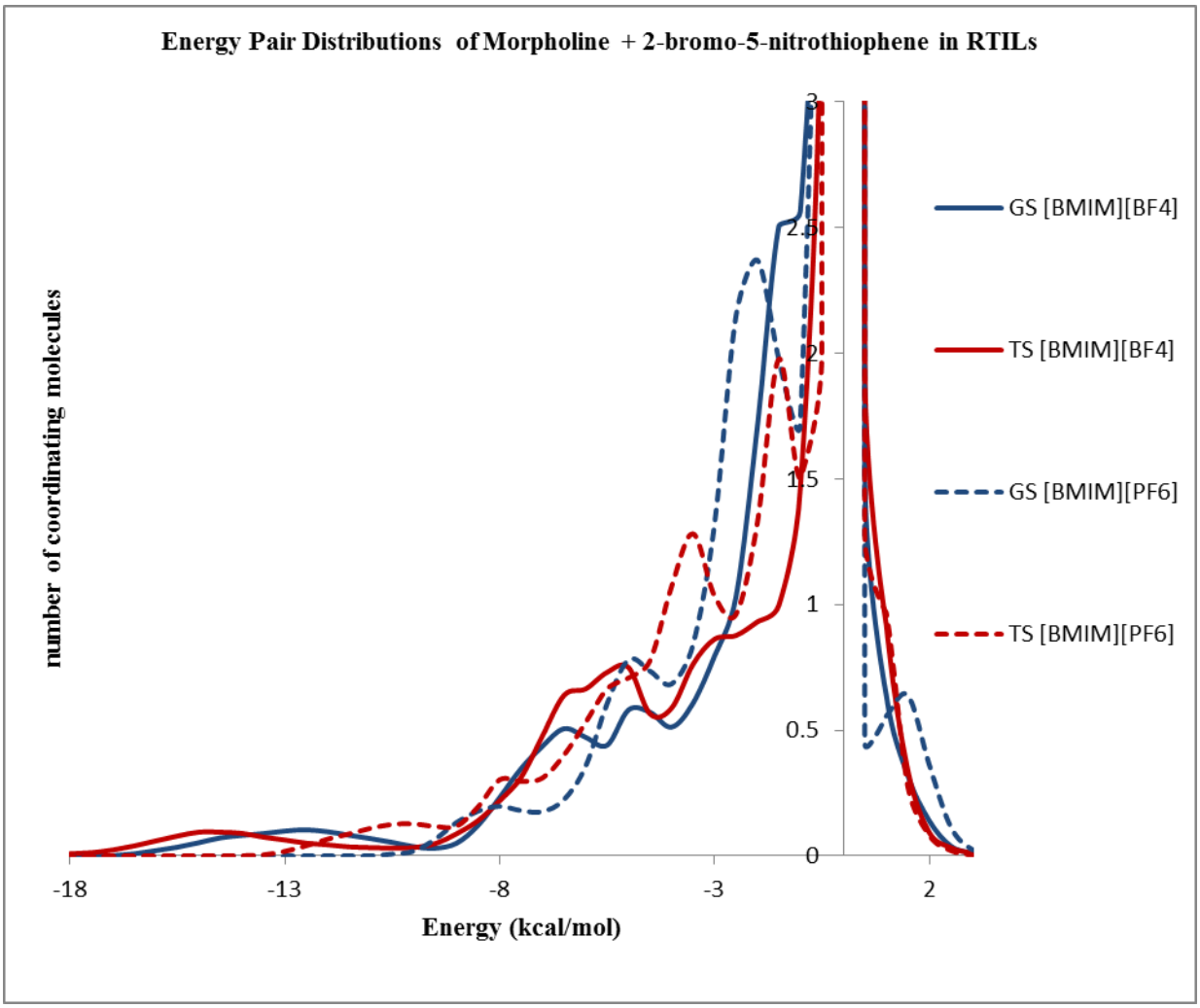


Figure 4S.79: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile morpholine and 2-bromo-5-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	-3.5 kcal/mol
GS [BMIM][BF4]	4.13	5.82
TS [BMIM][BF4]	4.96	6.88
GS [BMIM][PF6]	2.93	5.19
TS [BMIM][PF6]	4.24	7.37

Table 4S.10: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

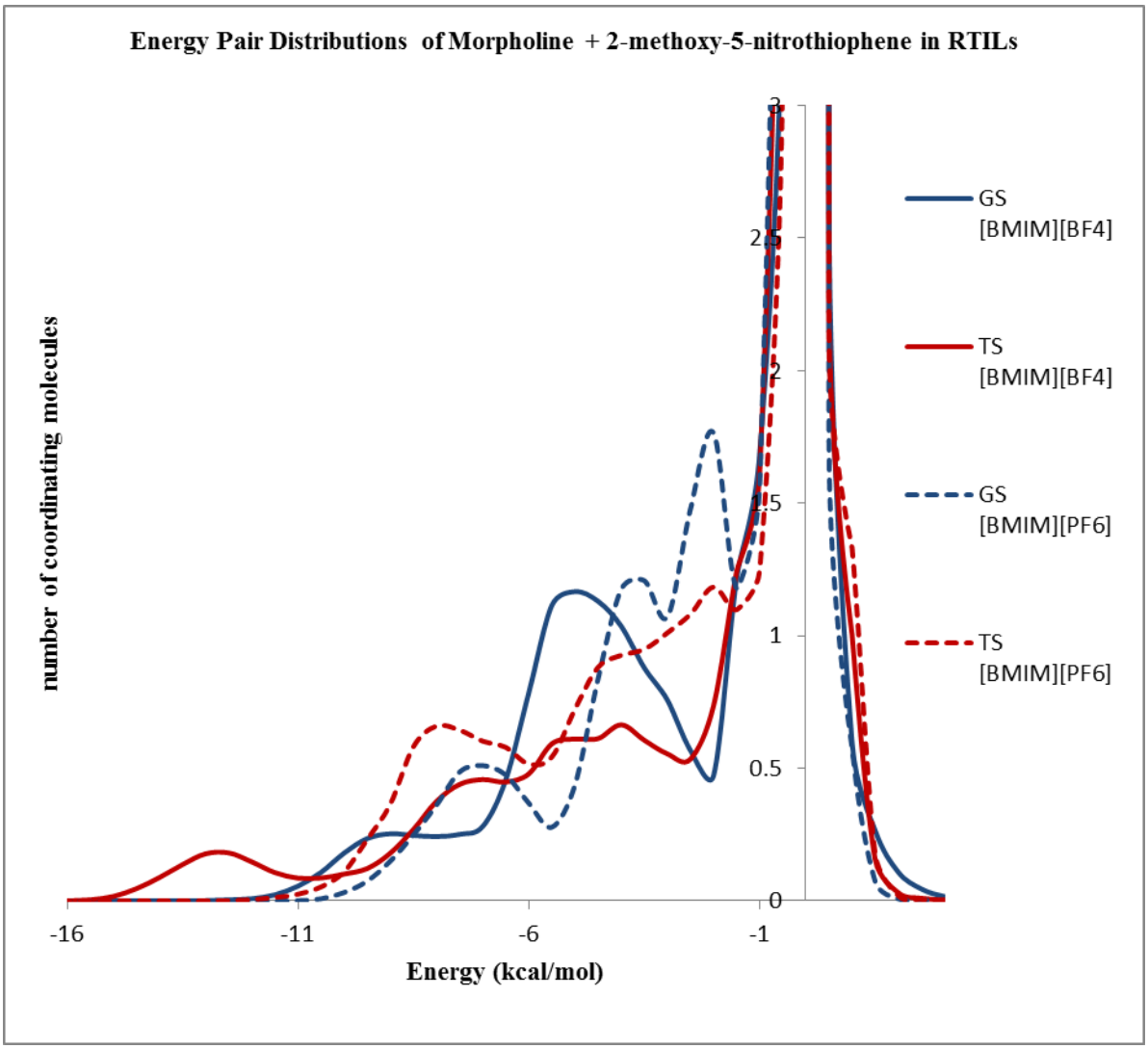


Figure 4S.80: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile morpholine and 2-methoxy-5-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	5.42	8.46
TS [BMIM][BF4]	5.16	7.04
GS [BMIM][PF6]	3.43	6.65
TS [BMIM][PF6]	5.65	8.41

Table 4S.11: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

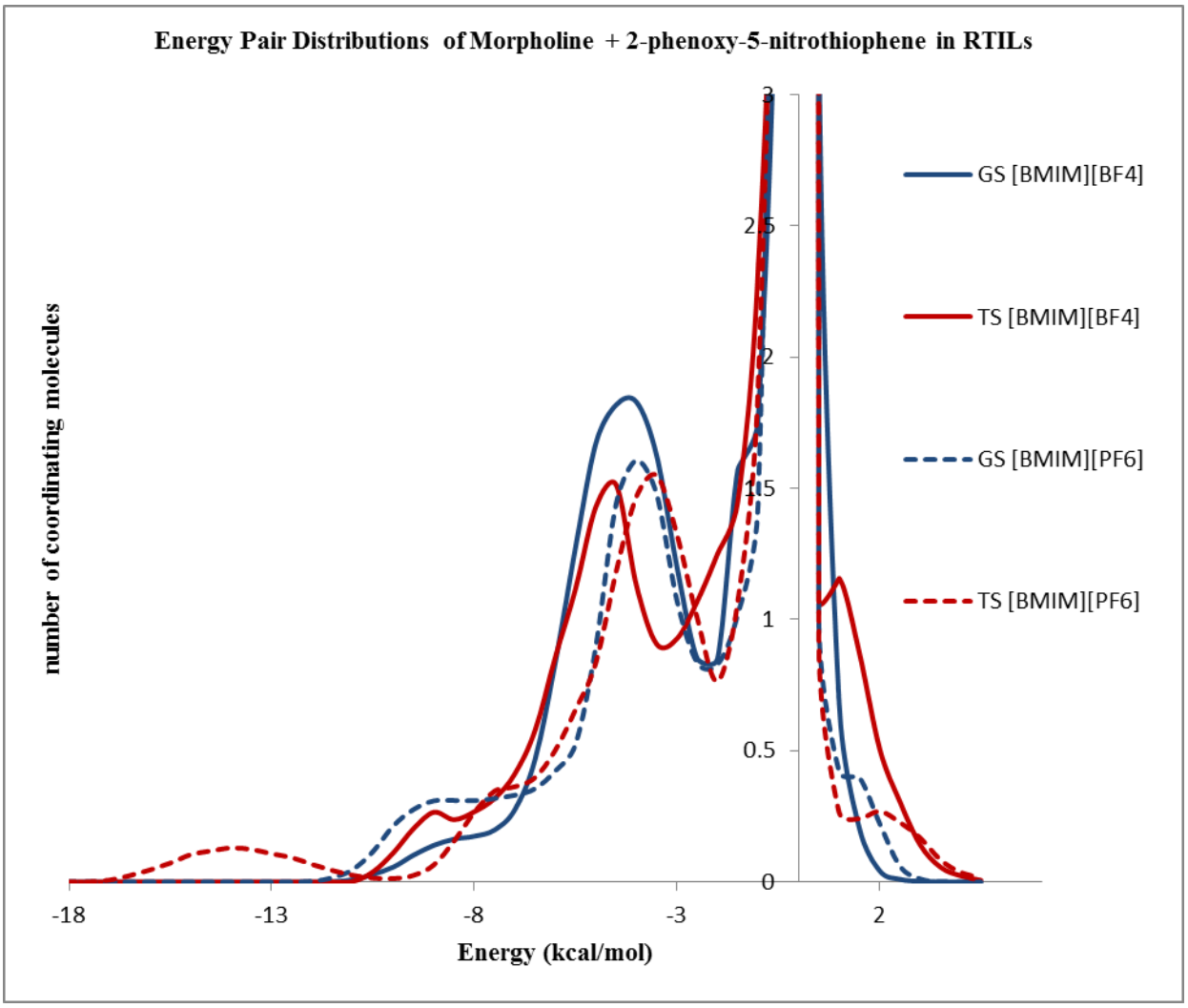


Figure 4S.81: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile morpholine and 2-phenoxy-5-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	5.41	10.68
TS [BMIM][BF4]	5.84	9.39
GS [BMIM][PF6]	4.45	8.98
TS [BMIM][PF6]	4.61	8.81

Table 4S.12: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

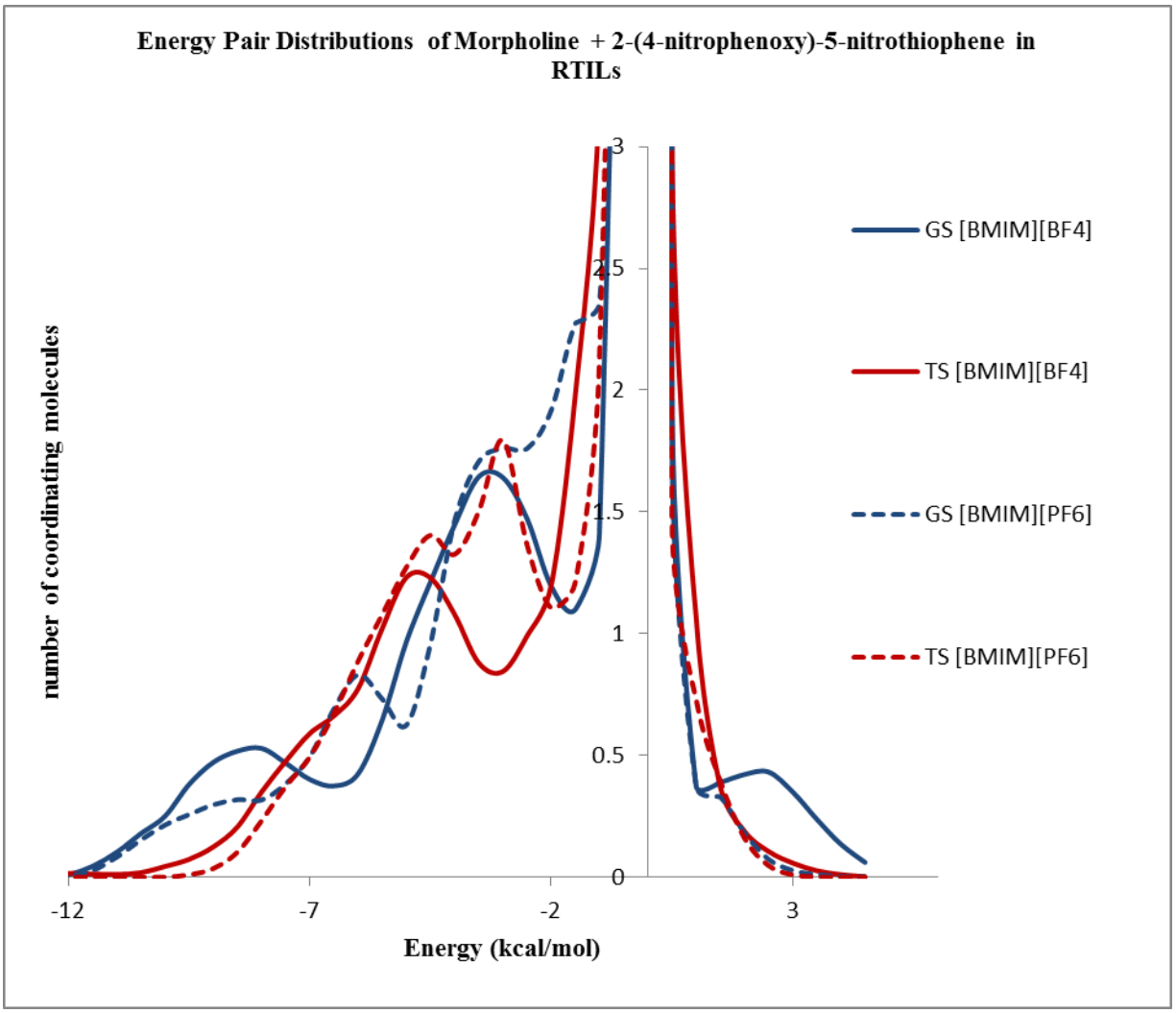


Figure 4S.82: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile morpholine and 2-(4-nitrophenoxy)-5-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	5.77	10.08
TS [BMIM][BF4]	6.58	9.76
GS [BMIM][PF6]	5.44	9.57
TS [BMIM][PF6]	5.16	9.39

Table 4S.13: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

Morpholine + ortho-L-nitrothiophene

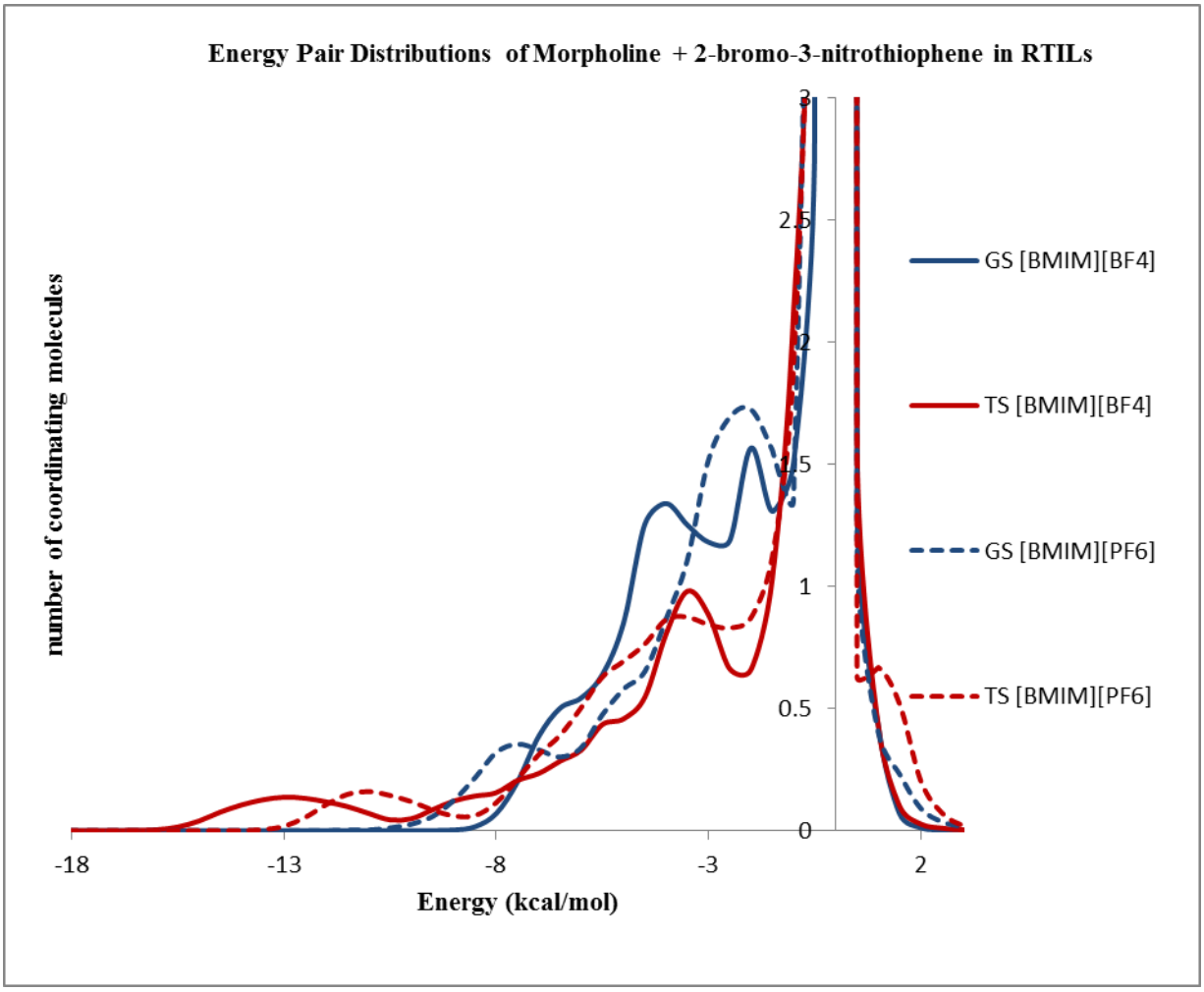


Figure 4S.83: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile morpholine and 2-bromo-3-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	3.21	7.06
TS [BMIM][BF4]	3.46	5.79
GS [BMIM][PF6]	3.14	5.77
TS [BMIM][PF6]	3.81	6.32

Table 4S.14: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

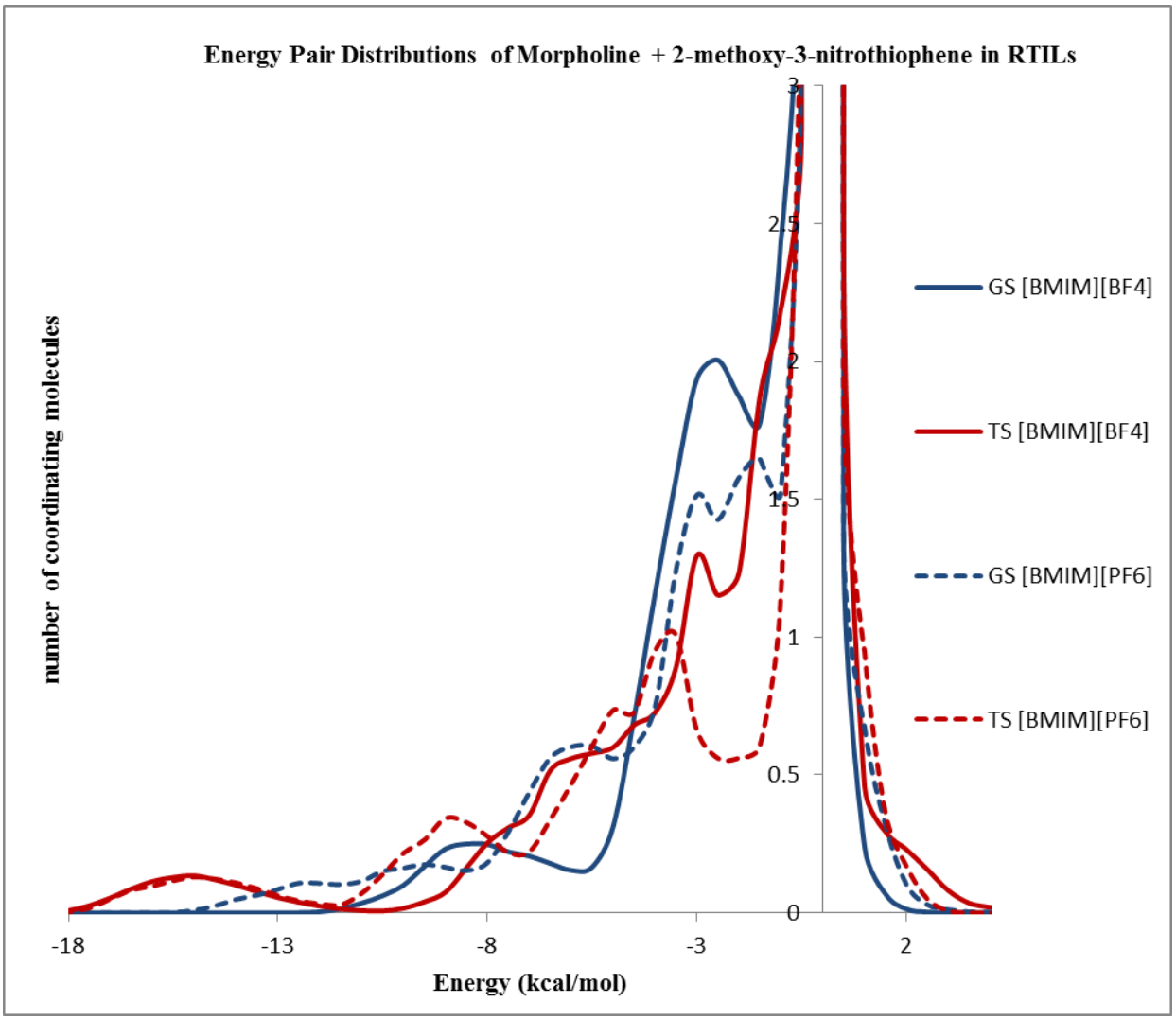


Figure 4S.84: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile morpholine and 2-methoxy-3-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	2.35	5.80
TS [BMIM][BF4]	4.46	6.76
GS [BMIM][PF6]	4.71	7.30
TS [BMIM][PF6]	5.21	7.89

Table 4S.15: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

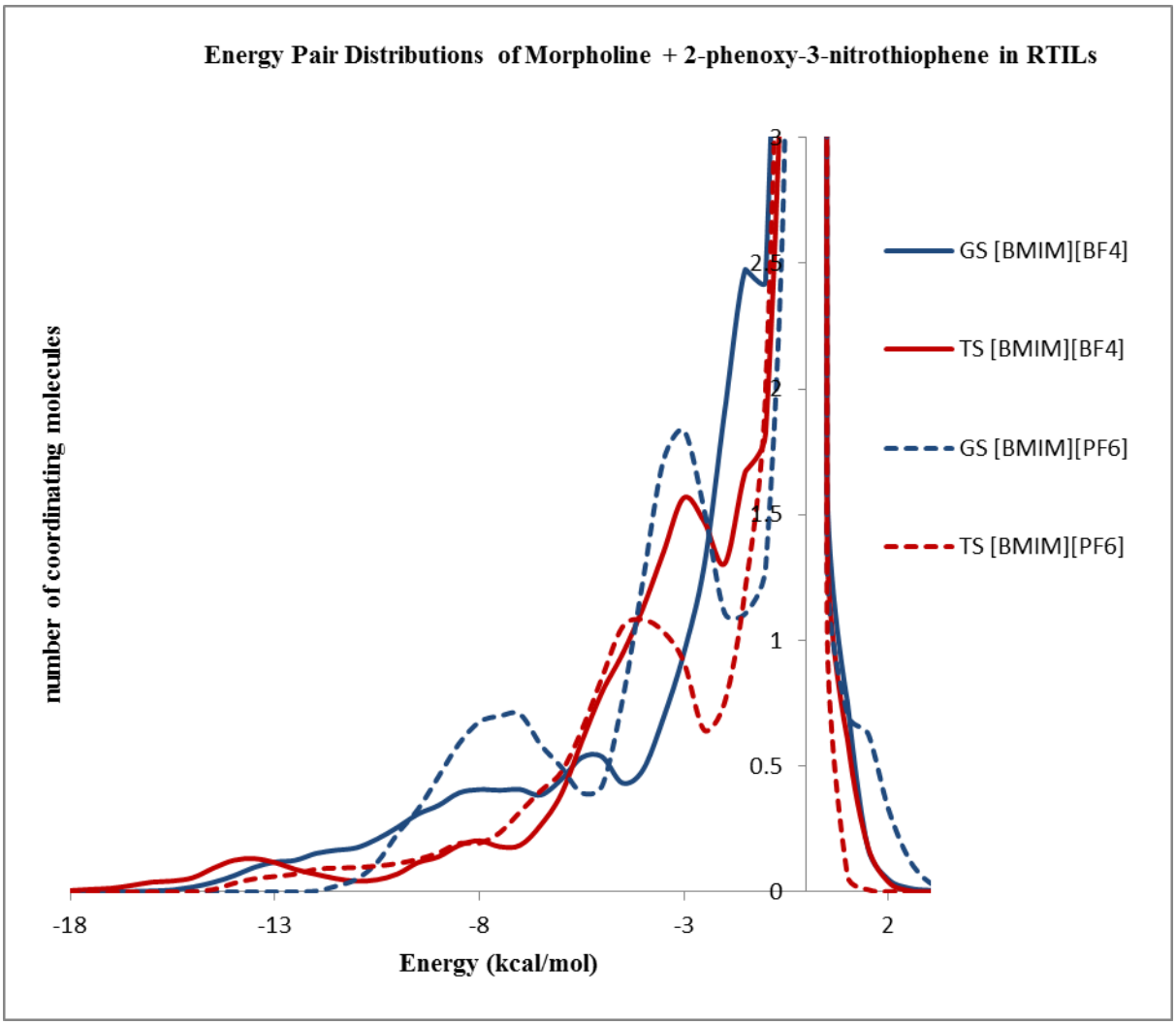


Figure 4S.85: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile morpholine and 2-phenoxy-3-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	5.60	7.21
TS [BMIM][BF4]	4.09	7.51
GS [BMIM][PF6]	5.78	9.49
TS [BMIM][PF6]	4.32	7.49

Table 4S.16: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

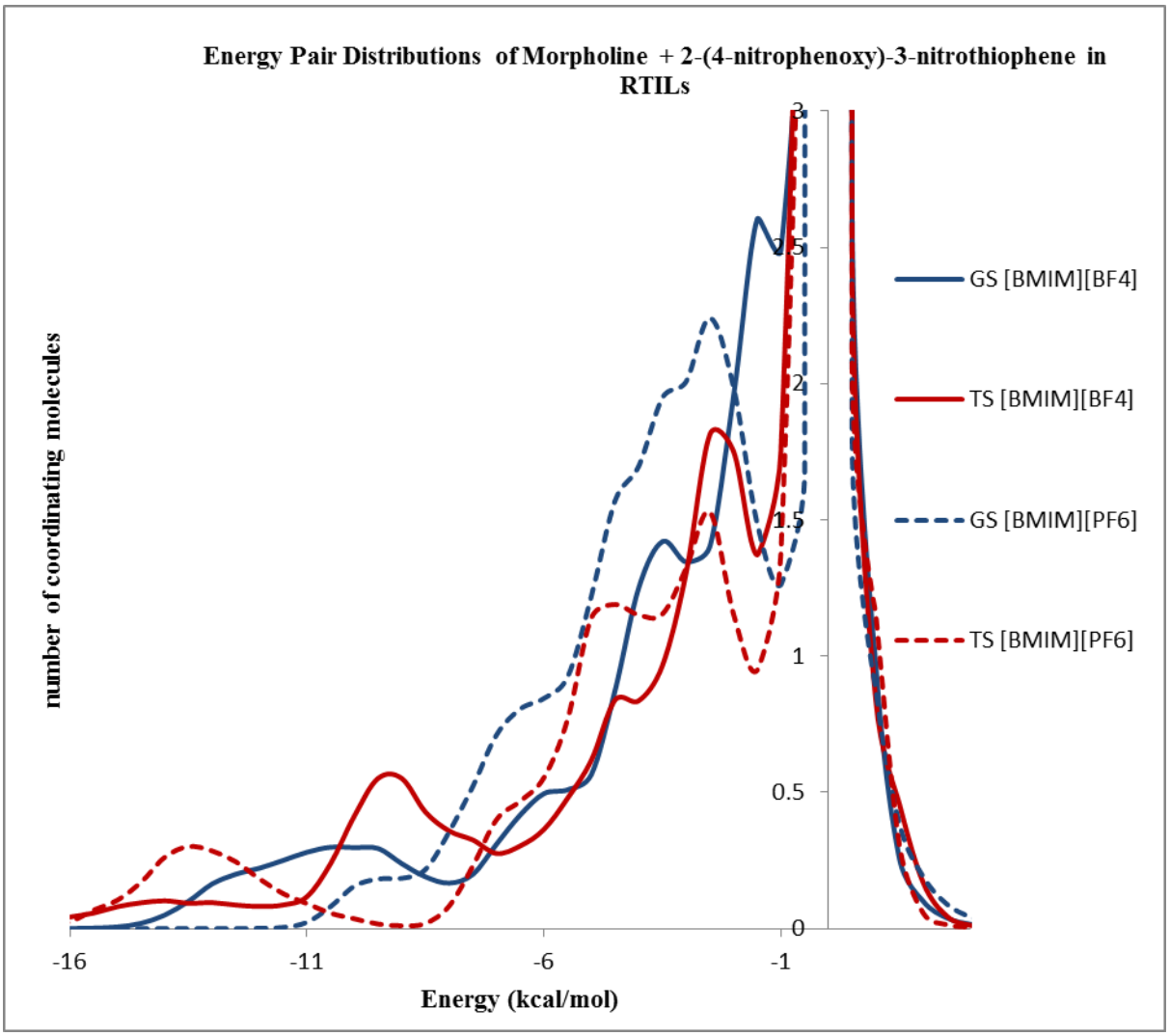


Figure 4S.86: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile morpholine and 2-(4-nitrophenoxy)-3-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	5.28	8.82
TS [BMIM][BF4]	5.91	8.55
GS [BMIM][PF6]	6.23	11.45
TS [BMIM][PF6]	5.66	9.16

Table 4S.17: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

Pyrrolidine + para-L-nitrothiophene

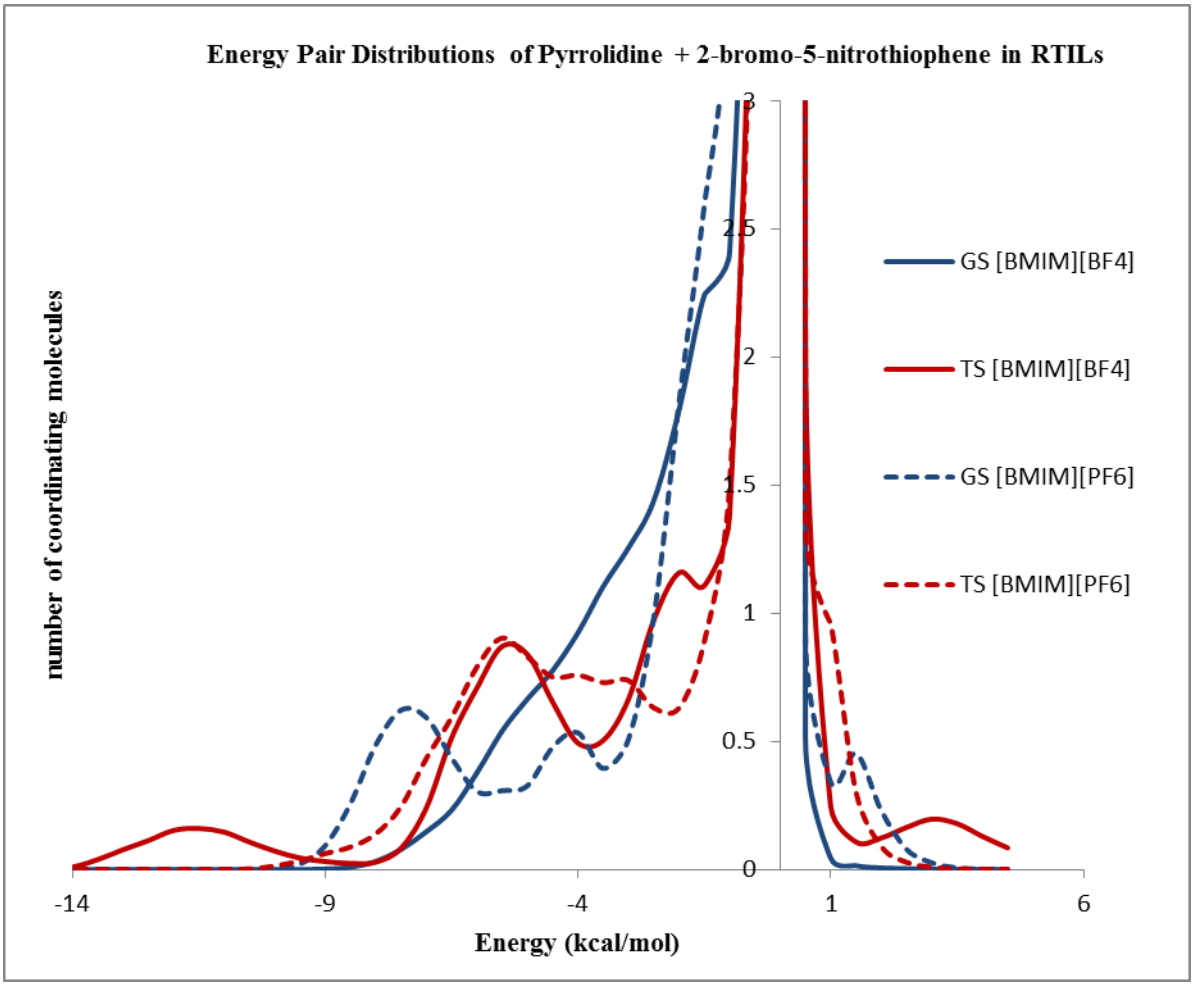


Figure 4S.87: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile pyrrolidine and 2-bromo-5-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	2.08	4.89
TS [BMIM][BF4]	4.26	5.91
GS [BMIM][PF6]	3.45	4.85
TS [BMIM][PF6]	4.13	6.37

Table 4S.18: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

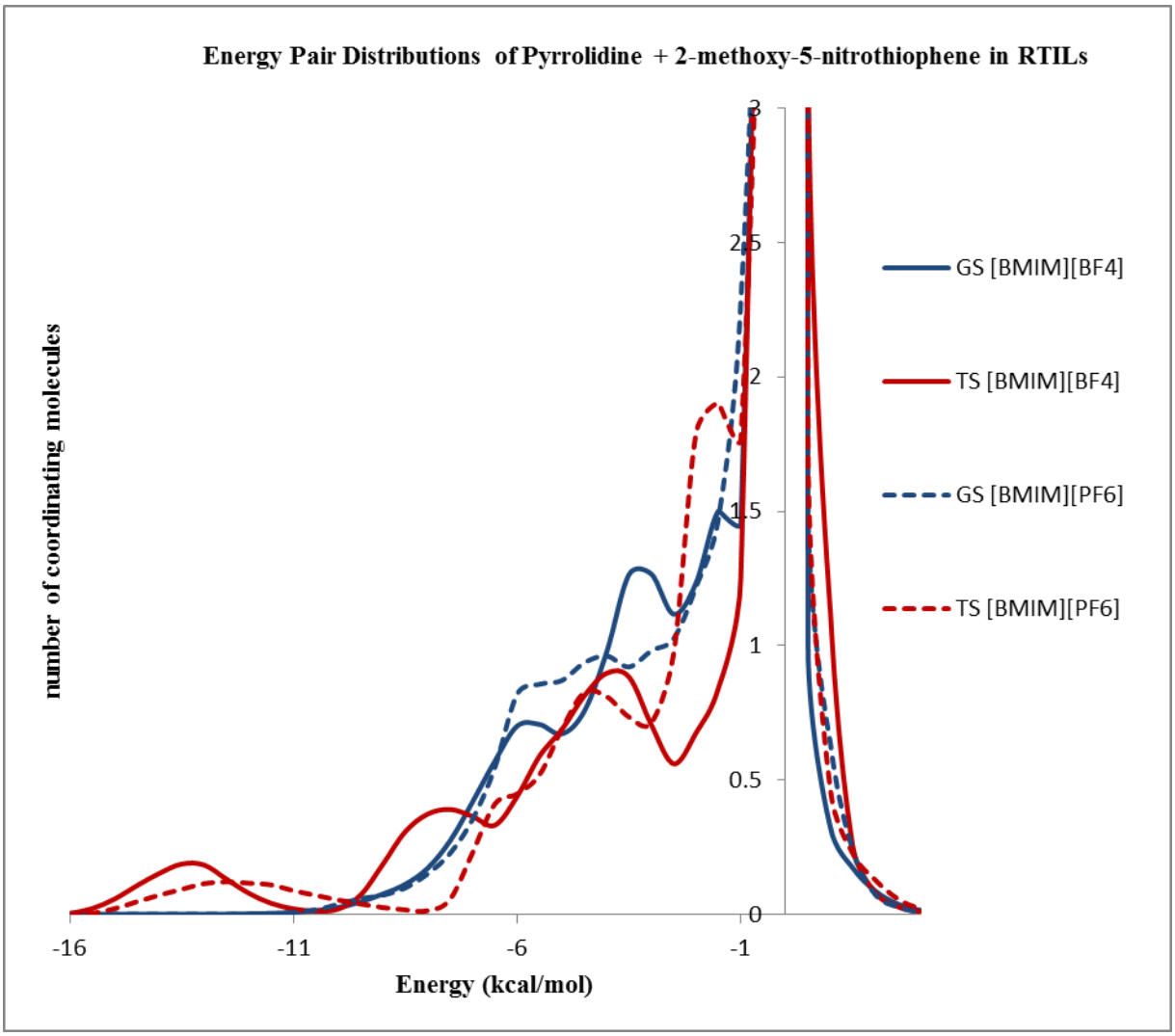


Figure 4S.88: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile pyrrolidine and 2-methoxy-5-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	3.80	6.80
TS [BMIM][BF4]	4.76	7.35
GS [BMIM][PF6]	4.11	6.93
TS [BMIM][PF6]	3.35	5.72

Table 4S.19: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

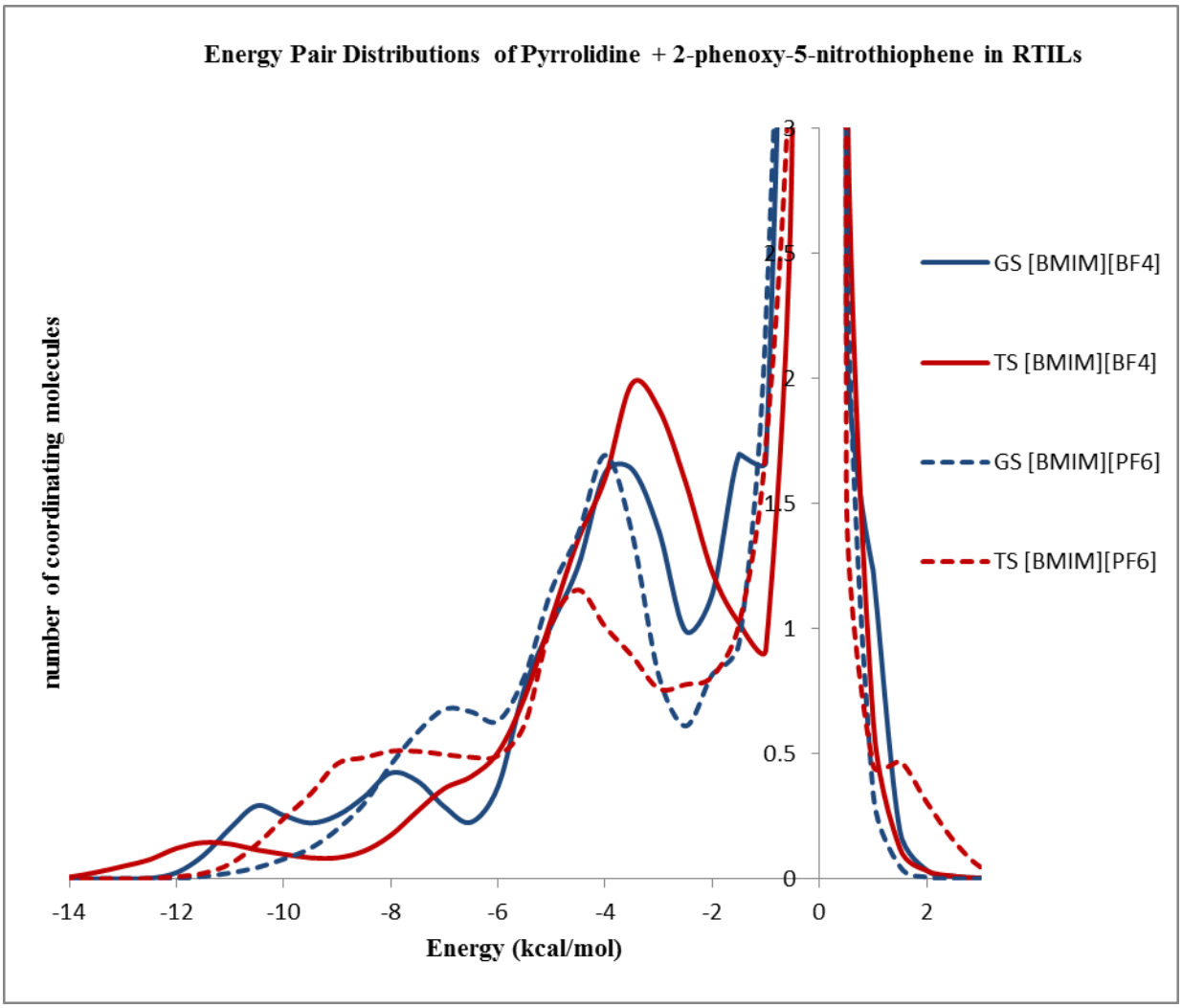


Figure 4S.89: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile pyrrolidine and 2-phenoxy-5-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	5.14	9.65
TS [BMIM][BF4]	4.54	9.48
GS [BMIM][PF6]	5.76	10.21
TS [BMIM][PF6]	5.88	8.93

Table 4S.20: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

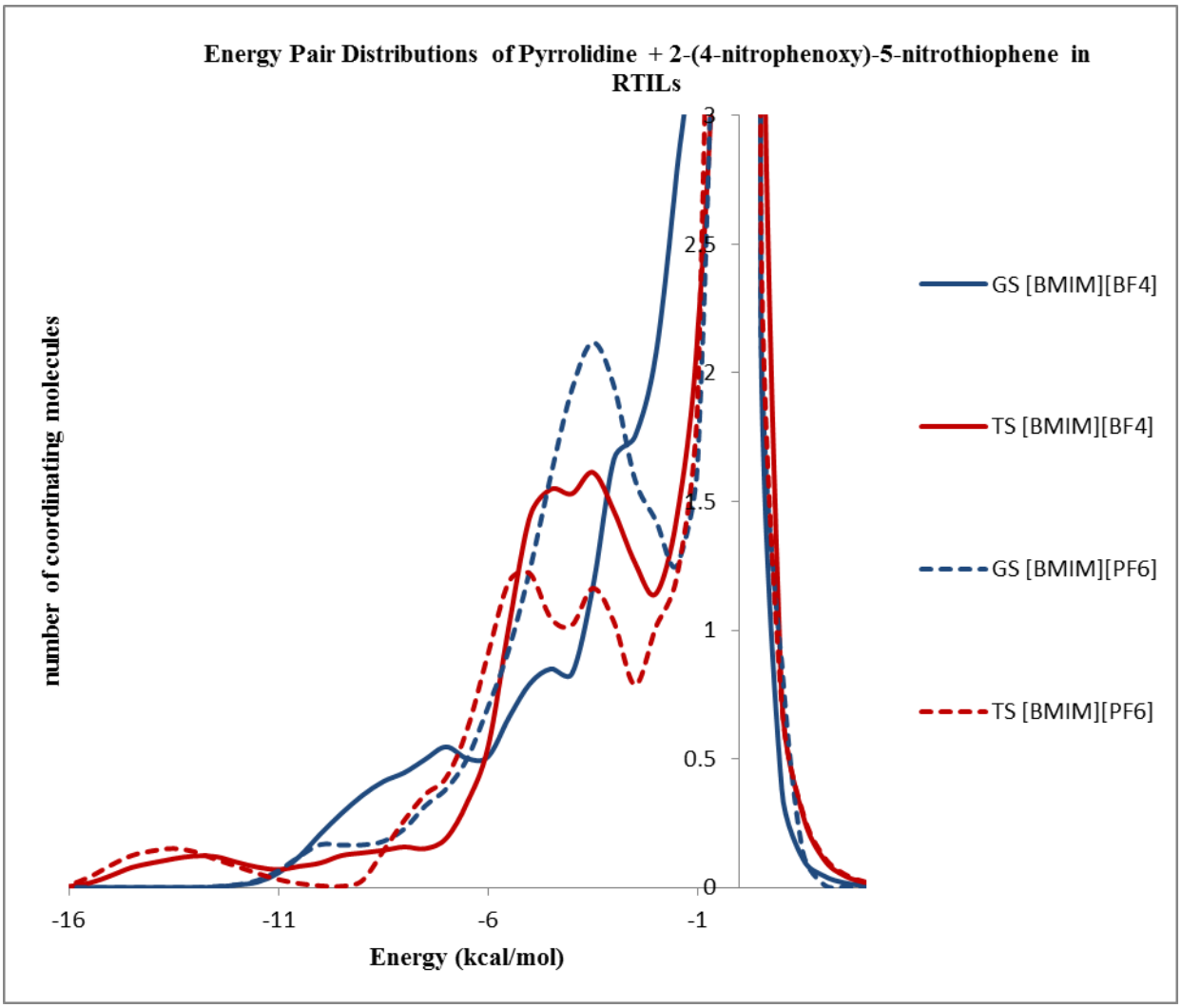


Figure 4S.90: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile pyrrolidine and 2-phenoxy-5-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	5.43	8.29
TS [BMIM][BF4]	5.27	9.96
GS [BMIM][PF6]	5.21	10.87
TS [BMIM][PF6]	6.15	9.38

Table 4S.21: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

Pyrrolidine + ortho-L-nitrothiophene

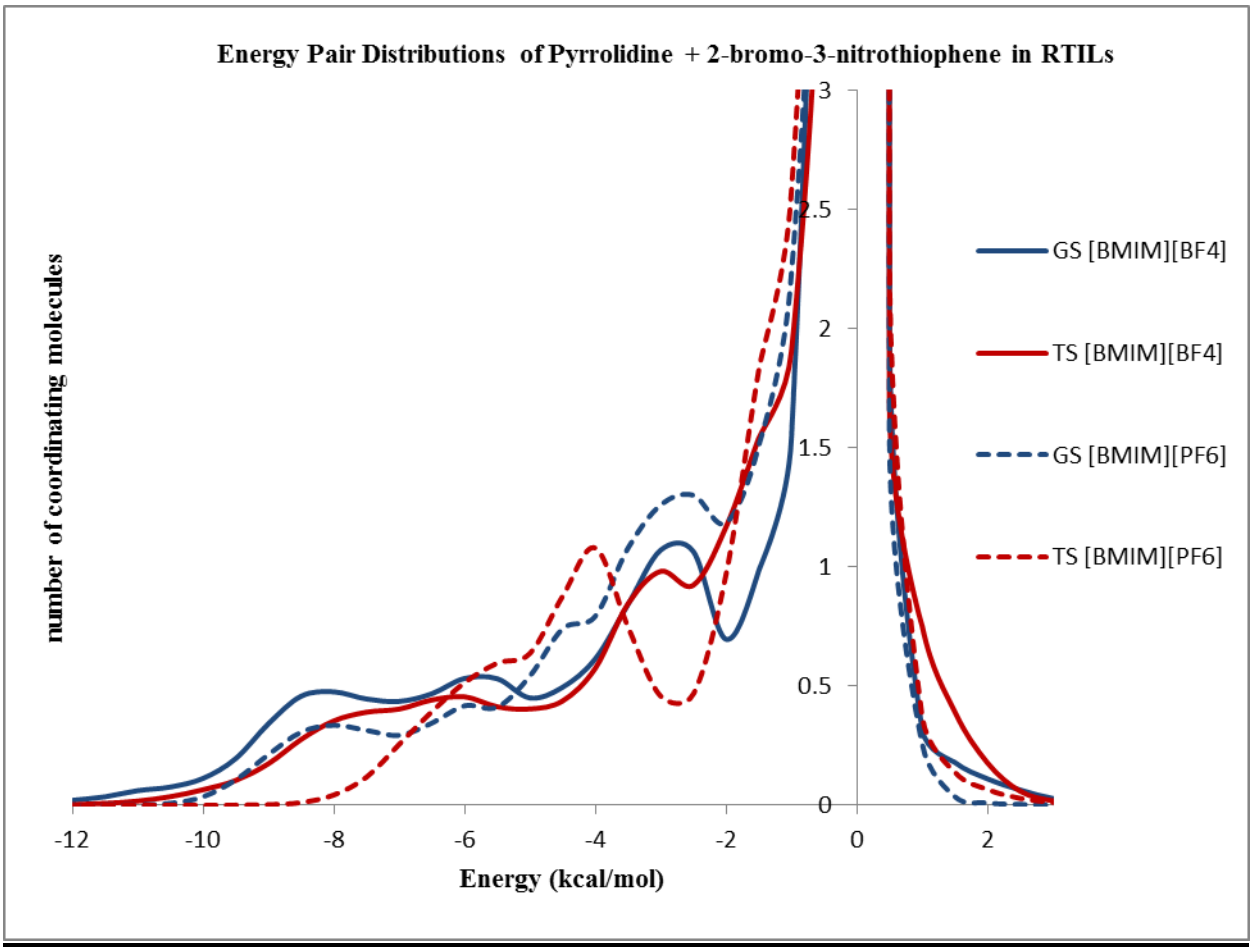


Figure 4S.91: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile pyrrolidine and 2-bromo-3-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	4.65	6.60
TS [BMIM][BF4]	3.54	5.40
GS [BMIM][PF6]	3.33	5.94
TS [BMIM][PF6]	2.57	5.28

Table 4S.22: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

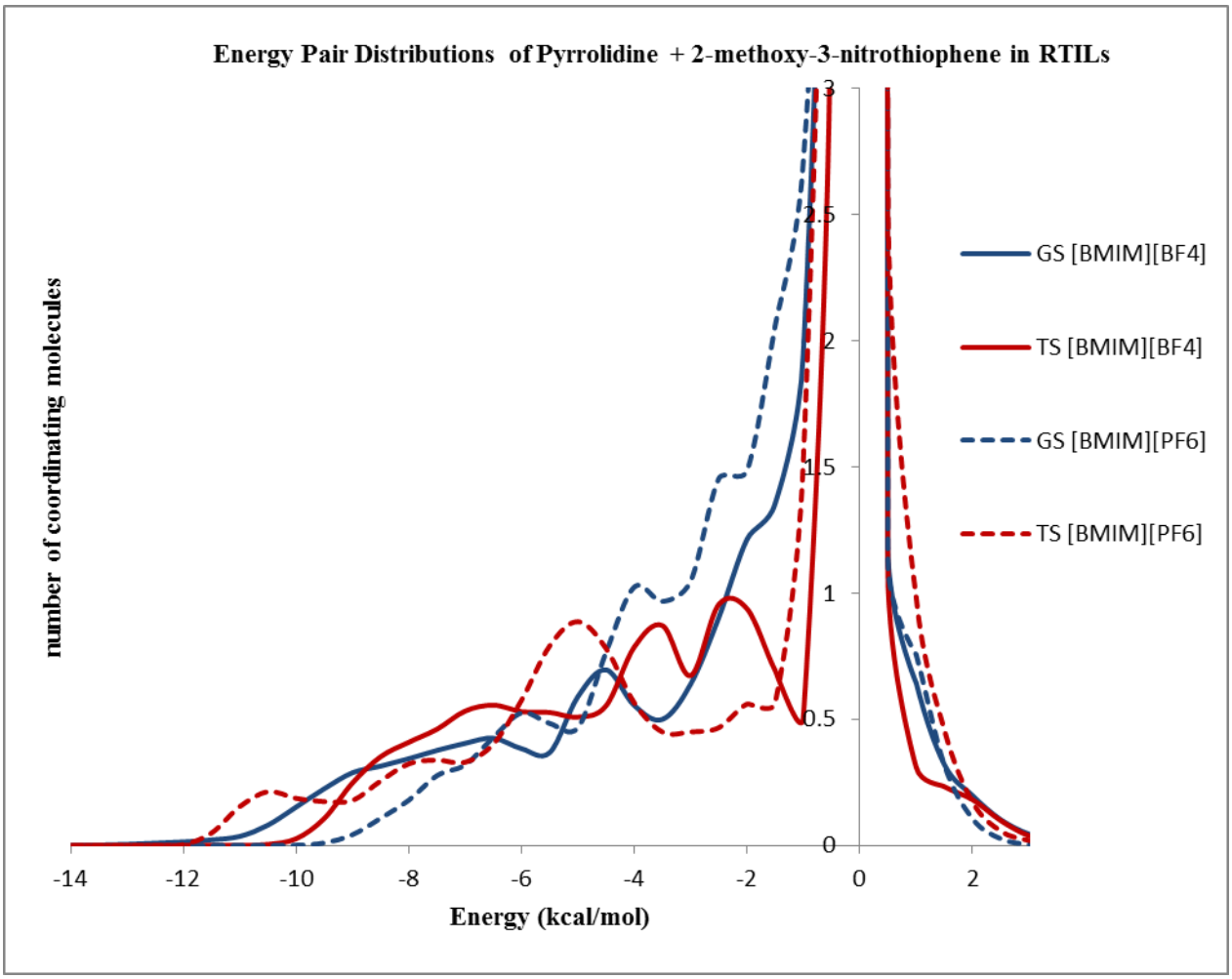


Figure 4S.92: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile pyrrolidine and 2-methoxy-3-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	-3.5 kcal/mol
GS [BMIM][BF4]	4.04	5.80
TS [BMIM][BF4]	4.27	6.48
GS [BMIM][PF6]	2.84	5.60
TS [BMIM][PF6]	4.87	6.66

Table 4S.23: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

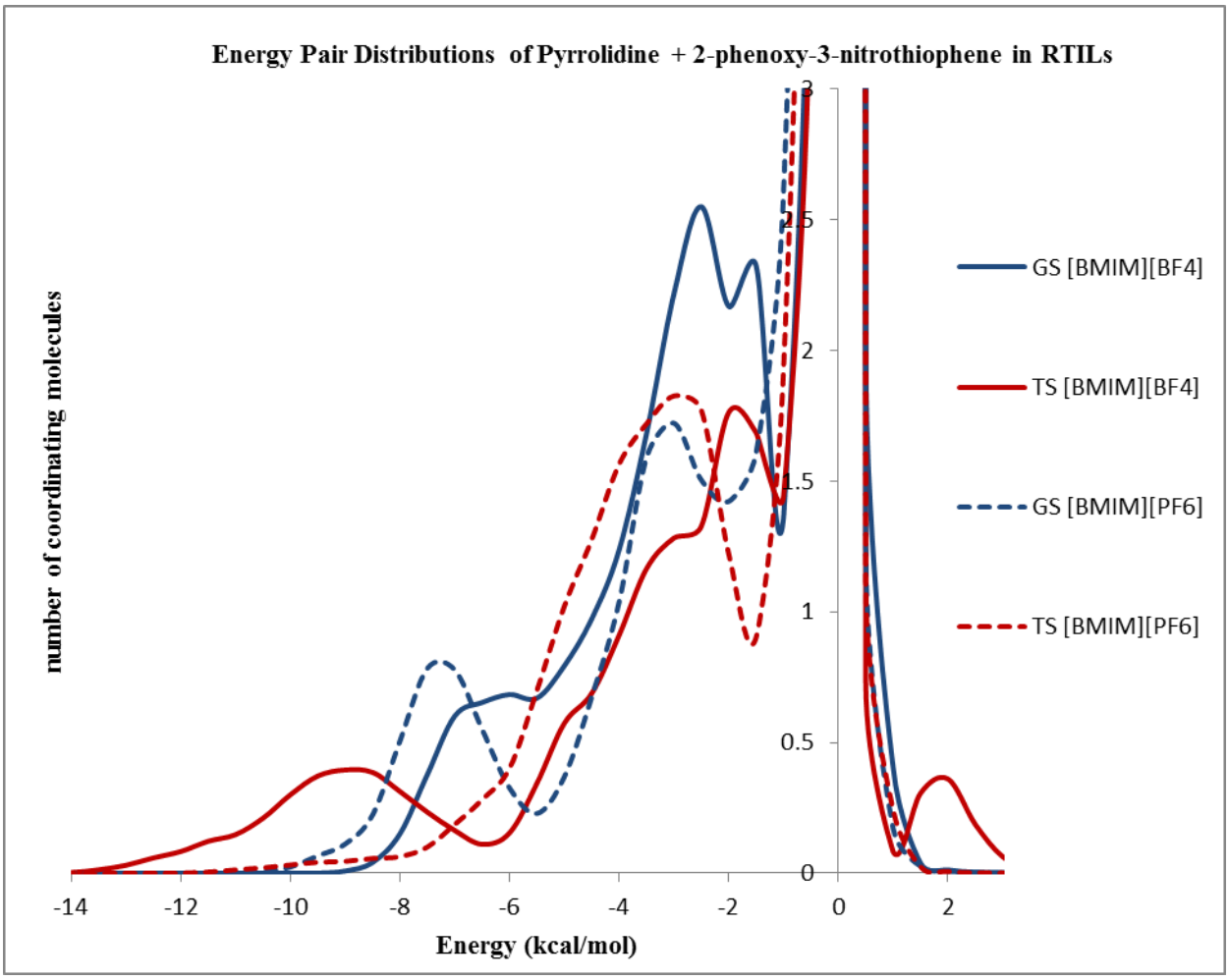


Figure 4S.93: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile pyrrolidine and 2-phenoxy-3-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	3.98	7.87
TS [BMIM][BF4]	4.01	6.78
GS [BMIM][PF6]	3.97	7.28
TS [BMIM][PF6]	2.97	7.53

Table 4S.24: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

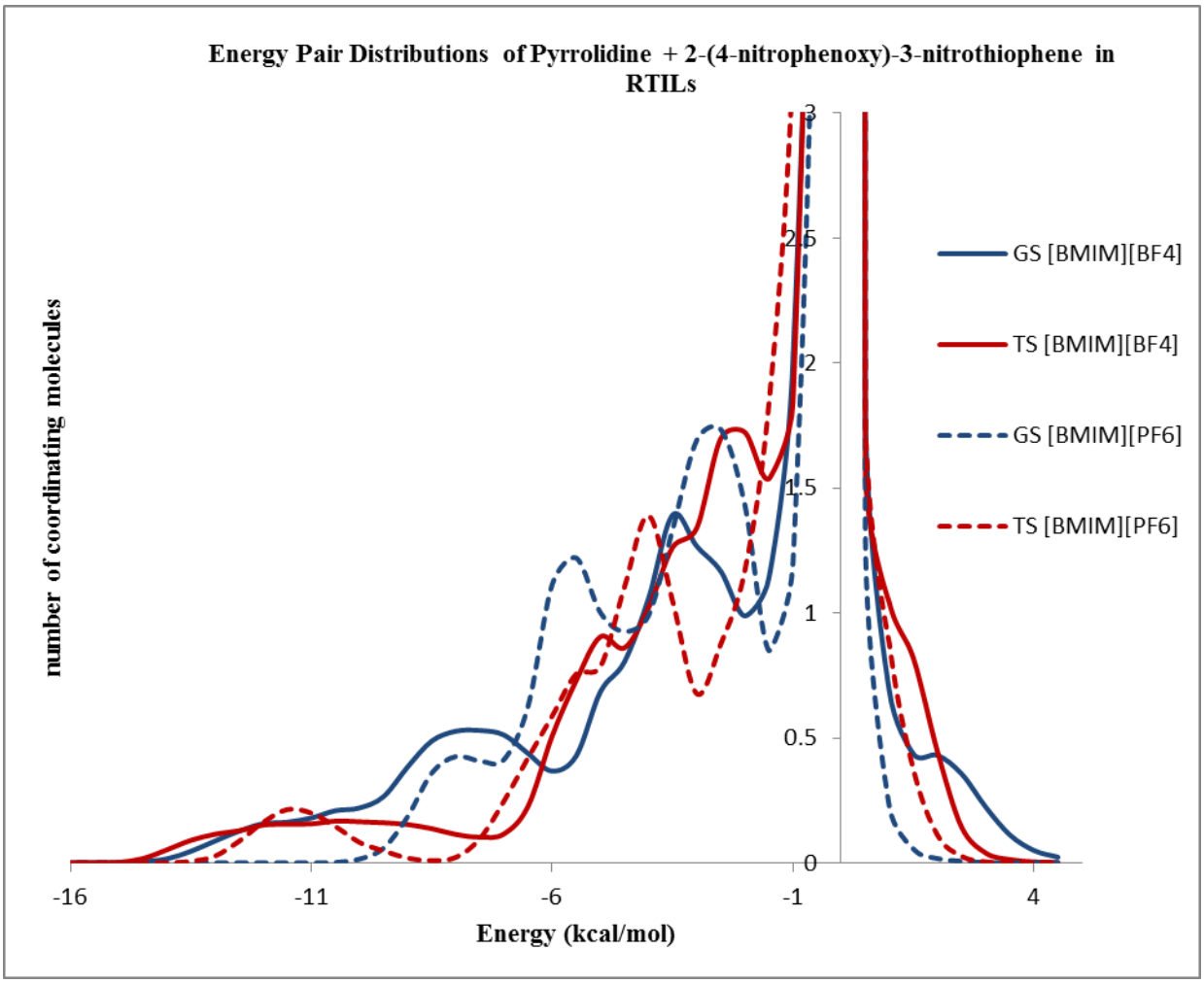


Figure 4S.94: The solute-solvent energy pair distributions of the ground state (GS – blue) and first transition state (TS – red) for the S_NAr reactions between the nucleophile pyrrolidine and 2-(4-nitrophenoxy)-3-nitrothiophene in [BMIM][BF_4] (solid line) and [BMIM][PF_6] (dashed line).

	-5 kcal/mol	**-3.5 kcal/mol**
GS [BMIM][BF4]	5.84	9.09
TS [BMIM][BF4]	4.35	7.50
GS [BMIM][PF6]	5.83	9.08
TS [BMIM][PF6]	3.92	7.46

Table 4S.25: The solute-solvent energy pair distributions of the reactants (GS) and first transition state (TS) for the S_NAr reactions, integrated to -5.0 kcal/mol and -3.5 kcal/mol.

Nucleophile	Thiophene Derivative	Leaving Group	Methanol Φ (degrees)	[BMIM][BF_4] Φ (degrees)	[BMIM][PF_6] Φ (degrees)
Morpholine	ortho	OPh	43.6	43.4	56.4
		$OPhNO_2$	9.0	66.0	10.7
	para	OPh	113.4	127.5	120.7
		$OPhNO_2$	176.8	158.8	136.8
Pyrrolidine	ortho	OPh	39.0	65.9	66.6
		$OPhNO_2$	55.6	8.1	68.2
	para	OPh	123.0	93.3	132.7
		$OPhNO_2$	139.0	138.3	82.4

Table 4S.26: Average dihedral angles (degrees, see Figure 4.10 for definition of Φ) for the addition transition structures of morpholine (or pyrrolidine) with 2-L-5- and 2-L-3-nitrothiophene at 25°C in methanol, [BMIM][BF_4], and [BMIM][PF_6]. Angles averaged over final 5 million configurations or 20 million configurations for methanol and the RTIL respectively.

Chapter 4 References

1. D'Anna, F.; Frenna, V.; Noto, R.; Pace, V.; Spinelli, D., *J. Org. Chem.* **2006,** *71*, 5144-5150.

2. (a) Vidis, A.; Olin, C. A.; Laurenczy, G.; Kusters, E.; Seldelmeier, G.; Dyson, P. J., *Adv. Synth. Catal.* **2005,** *347*, 266-274; (b) Aggarwal, A.; Lancaster, N. L.; Sethi, A. R.; Welton, T., *Green Chem.* **2002,** *4*, 517-520; (c) Welton, T., *Chem. Rev.* **1999,** *99*, 2071; (d) Allen, C.; Sambasivarao, S. V.; Acevedo, O., *J. Am. Chem. Soc.* **2013,** *31*, 1065-1072.

3. Jorgensen, W. L.; Tirado-Rivers, J., *J. Comput. Chem.* **2005,** *25*, 1689-1700.

4. (a) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L., *J. Comput. Chem.* **2002,** *23*, 1601-1622; (b) Tubert-Brohman, I.; Guimaraes, C. R. W.; Repasky, M. P.; Jorgensen, W. L., *J. Comput. Chem.* **2003,** *25*, 138-150; (c) Tubert-Brohman, I.; Guimaraes, C. R. W.; Jorgensen, W. L., *J. Chem. Theory Comput.* **2005,** *1*, 817-823.

5. (a) Storer, J. W.; Gisen, D. J.; Cramer, C. J.; Truhlar, D. G., *J. Comput. - Aided Mol. Des.* **1995,** *9*, 87; (b) Winget, P.; thompson, J. D.; Xidos, J. D.; Cramer, C. J.; Truhlar, D. G., *J. Phys. Chem. A.* **2002,** *106*, 10707-10717; (c) Blagovic, M. U.; Morales de Tirado, P.; Pearlman, S. A.; Jorgensen, W. L., *J. Comput. Chem.* **2004,** *25*, 1322-1332; (d) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G., *J. Comput. Chem.* **2003,** *24*, 1291-1304; (e) Vilseck, J. Z.; Sambasivarao, S. V.; Acevedo, O., *J. Comput. Chem.* **2011,** *32*, 2836-2842.

6. Vayner, G.; Houk, K. N.; Jorgensen, W. L.; Brauman, J. I., *J. Am. Chem. Soc.* **2004,** *126*, 9054-9058.

7. Acevedo, O.; Jorgensen, W. L., *Org. Lett.* **2004,** *6*, 2881-2884.

8. (a) Acevedo, O.; Jorgensen, W. L., *J. Am. Chem. Soc.* **2005,** *127*, 8829-8834; (b) Acevedo, O.; Jorgensen, W. L., *J. Am. Chem. Soc.* **2006,** *71*, 4896-4902.

9. Acevedo, O.; Armacost, K., *J. Am. Chem. Soc.* **2010,** *132*, 1966-1975.

10. Acevedo, O.; Jorgensen, W. L., *J. Chem. Theory Comput.* **2007,** *3*, 132-138.

11. Sambasivarao, S. V.; Acevedo, O., *J. Chem. Theory Comput.* **209,** *5*, 1038-1050.

12. Jorgensen, W. L., *J. Phys. Chem.* **1986,** *90*, 1276-1284.

13. 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.

14. (a) 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; (b) 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; (c) 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.

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

16. 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.

17. 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.

18. (a) Dewar, M. J. S.; Yuan, Y.-C., AM1 Studies of E2 Reactions. 1. Mechanism and Leaving Group Effects. *J. Am. Chem. Soc.* **1990,** *112*, 2088-2094; (b) Dewar, M. J. S.; Yuan, Y.-C., AM 1 Studies of E2 Reactions. 2. Regioselectivity, Stereochemistry, Kinetic Isotope Effects, and Competition with SN2 Reactions. *J. Am. Chem. Soc.* **1990,** *112*, 2095-2105.

19. Krop, H. B.; Cheung, C. L.; Govers, H. A. J., Correlation between the experimental and calculated (COSMO-AM1) activation free enthalpy of the hydroxide-induced elimination reaction of small haloalkanes in water. *J. Mol. Struct. (THEOCHEM)* **2000,** *505*, 1-10.

20. Hall, H. K., Jr, *J. Am. Chem. Soc.* **1957,** *79*, 5441-5444.

21. D'Anna, F.; Frenna, V.; Pace, V.; Noto, R., *Tetrahedron* **2006,** *62*, 1690-1698.

22. (a) Crowhurst, L.; Mawdsley, P.; Perez-Alandia, J. M.; Salter, P.; Welton, T., *Phys. Chem. Chem. Phys.* **2003,** *5*, 2790; (b) Kamlet, M. J.; Abboud, J.-L. M.; Abraham, M. H.; Taft, R. W., *J. Org. Chem.* **1983,** *48*, 2877.

23. Correia, I.; Welton, T., *Dalton Trans.* **2009**, 4115-4121.

24. Hammond, G. S., *J. Am. Chem. Soc.* **1955,** *77*, 334-338.

25. Crowhurst, L.; Lancaster, N. L.; Perez-Alandia, J. M.; Welton, T., *J. Am. Chem. Soc.* **2004,** *126*, 11549-11555.

26. (a) Amyes, T. L.; Diver, S. T.; Richard, J. P.; Rivas, F. M.; Toth, K., *J. Am. Chem. Soc.* **2004,** *126*, 4366-4374; (b) Dupont, J.; Spencer, J., *Angew. Chem. Int. Ed.* **2004,** *43*, 5296-5297.

27. Hollbrey, J. D.; Reichert, W. M.; Nieuwenhuyzen, M.; Sheppard, O.; Hardacre, C.; Rogers, R. D., *Chem. Commun.* **2003**, 476-477.

Chapter 5

The QM/MM/MC Study of the Mononuclear Heterocyclic Boulton – Katritzky Rearrangement of the Z-phenylhydrazone of 3-benzoyl-5-phenyl-1,2,4-oxadiazole into 4-benzoylamino-2,5-diphenyl-1,2,3-triazole in Room Temperature Ionic Liquids.

5.1 Introduction

Mononuclear heterocyclic rearrangements (MHR) are a common and efficient way to synthesize five-membered heteroaromatics[1] and dihydroheteraromatics.[2] The reaction mechanisms do not require π system reorganizations, but rather the making and breaking of σ-bonds.[1c, 1g, 3] Of specific interest is the Boulton-Katritky reaction[1b] of the Z-phenylhydrazone of 3-benzoyl-5-phenyl-1,2,4-oxadiazole into 4-benzoylamino-2,5-diphenyl-1,2,3-triazole (Scheme 5.1, compounds 1 and 2 respectively) in methanol and two room temperature ionic liquids (RTILs): 1-butyl-3-methylimidazolium tetrafluoroborate [BMIM][BF_4] and 1-butyl-3-methylimidazolium hexafluorophosphate [BMIM][PF_6].[4] Different bases, the amines piperidine, triethylamine, and butylamine, have been studied experimentally[4-5] in RTILs, and conventional solvents. Piperidine (Pip) was chosen to study the MHR reaction due to reported amine reactivity in the [BMIM] based RTILs. Experimental $\Delta H^{\ddagger}$ for the primary and tertiary amines were significantly higher than that for piperidine.[4]

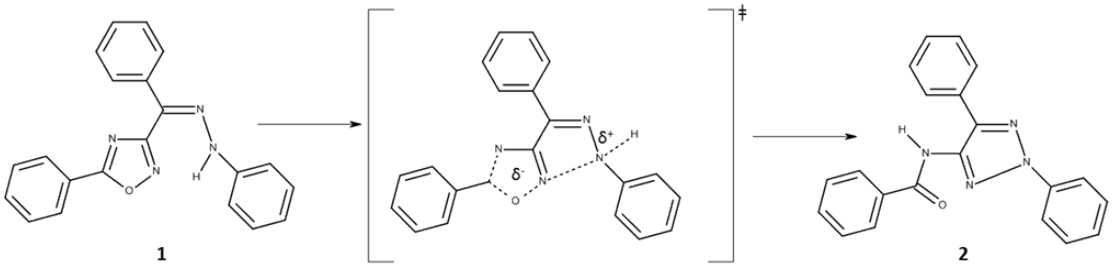
The MHR has attracted limited attention from the computational chemistry community. Some simple systems such as the acylamino derivatives of isoxazole have

been studied using semiempirical and *ab initio* calculations, however these results are in relatively poor agreement with experimental data.[6] In a 2004 paper published by Bottoni *et al.*[3b] they reported consistent results between experimental and DFT studies of a MHR of Z-hydrazones of 3-acyl-1,2,4-oxadiazoles. A mechanistic study of the MHR uncatalyzed pathway included one or two water molecules to describe the direct involvement of the solvent as a base. They concluded that the effect of the solvent is fundamental in determining the energy profile of the favored pathway.[3b] Furthermore, under experimental conditions they concluded that the rearrangement should proceed through a concerted, asynchronous path where nucleophilic attack and proton transfer occur in the same step but are not simultaneous.[3b]

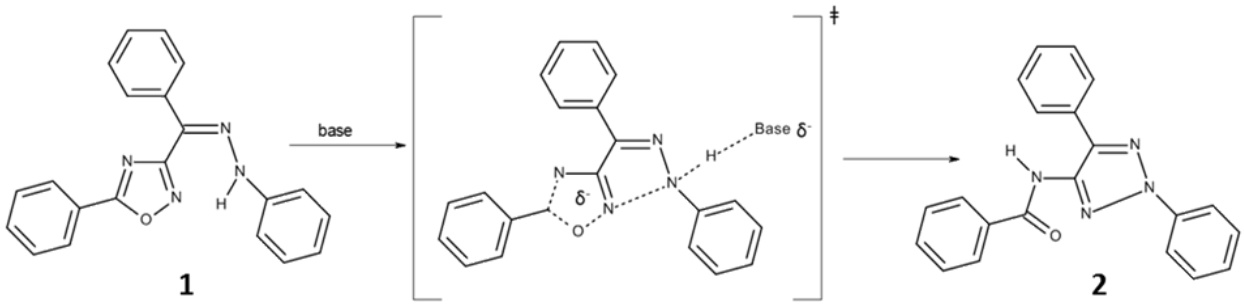
Experimental data for MHRs in multiple solvents; e.g., RTILs,[4] acetonitrile,[4, 5b, 7] benzene,[4, 5b, 7] dioxane/water,[4, 5b] methanol,[4, 7] and ethyl acetate,[4, 7] provide kinetic evidence for the occurrence of two different reaction pathways: (1) a proton-independent, uncatalyzed pathway (Scheme 5.1) and (2) a proton-dependent, base-catalyzed pathway (Scheme 5.2) that requires general or specific base catalysis.[3b] In protic solvents with high permittivity and basicity, such as methanol or dioxane/water, as well as in dipolar aprotic solvents such as acetonitrile, the uncatalyzed pathway has been found.[8] In nonpolar solvents such as benzene, the uncatalyzed pathway has not been observed.[8] On the other hand, the occurrence of a specific acid-catalyzed pathway has been proposed for MHRs when there is a basic center present in the attacked heterocycle.[3b, 9] Due to the unusual solvent environment of ionic liquids, RTILs have been shown to alter reaction mechanisms[10] as well as effect reactivity in diverse ways.[8] Since it is very difficult and often inaccurate to describe the ionic liquid effect by means of a single solvent property[11]

such as polarity, it has been difficult to determine the correct MHR reaction mechanism in RTILs.

The RTILs provide a significant increase in the rate of MHR when compared to methanol.[4] The accelerated rates of rearrangement have been attributed to the unique ionic liquid effects, however substrate-solvent and base-solvent interactions have been reported to have scarce relevance.[4-5] Conversely, their solvent effects are based on their partially pre-organized structure.[4, 12] Ionic liquids affect substrates that have structural requirements, such as Z-phenylhydrazones, and may allow the reactant to adopt a pre-ordered geometric arrangement that may enhance the reaction rate.[4, 13] Boulton-Katritzky rearrangements are an intramolecular nucleophilic substitution reaction, which is reported to occur through a bicyclic 10π quasi-aromatic transition state, and thus represent a special case of S_N2 displacement reactions.[3b, 4] Previous studies propose a concerted but asynchronous pathway in which the bond making and breaking occur simultaneously, but are influenced differently by the viscosities of the solvents.[3b, 4] As viscosity increases; e.g., from methanol (υ = 0.544cp[14]) to [BMIM][BF4] (υ = 233cp[14]) to [BMIM][PF_6] (υ = 450cp[14]), bond making is favored while bond breaking is disfavored, altering the position of the transition state along the reaction coordinate.[15] Moreover, with these particular rearrangements, the RTILs provide favorable π-π interactions due to the conjugation and aromatic nature of the substrates.[4] For these reasons, RTILs may function as "entropic drivers",[16] that can be described as polymeric supramolecular fluids[4] that have regions described by different polarity character.[17]



Scheme 5.1: The uncatalyzed MHR of the Z-phenylhydrazone of 3-benzoyl-5-phenyl-1,2,4-oxadiazole (compound 1) into 4-benzoylamino-2,5-diphenyl-1,2,3-triazole (compound 2).



Scheme 5.2: The base-catalyzed MHR of the Z-phenylhydrazone of 3-benzoyl-5-phenyl-1,2,4-oxadiazole (compound 1) into 4-benzoylamino-2,5-diphenyl-1,2,3-triazole (compound 2).

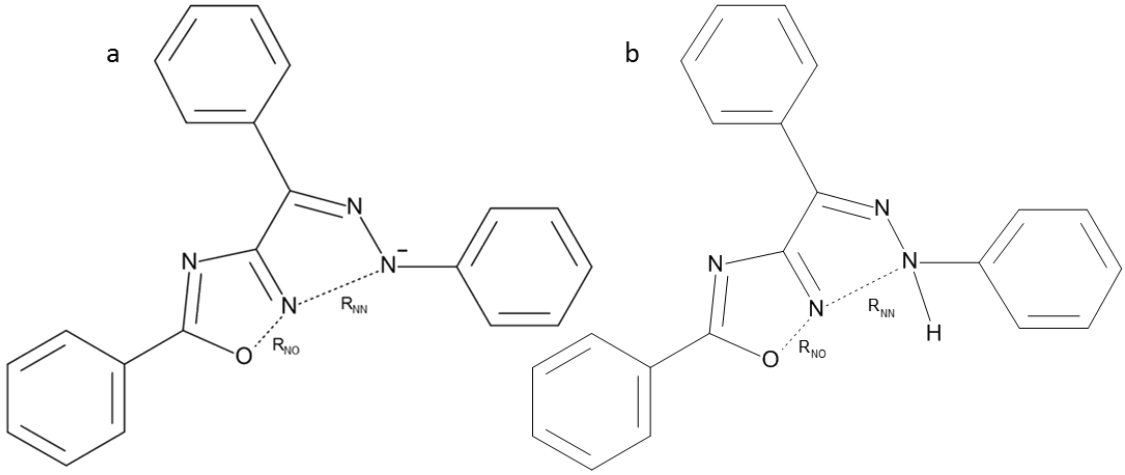
5.2 Computational Methods

Mixed quantum mechanical and molecular mechanical (QM/MM) calculations were carried out on the Boulton-Katritzky MHR of the Z-phenylhydrazone of 3-benzoyl-5-phenyl-1,2,4-oxadiazole (compound 1) into 4-benzoylamino-2,5-diphenyl-1,2,3-triazole (compound 2) in [BMIM][BF_4], [BMIM][PF_6], and methanol. The starting geometry for the solute was determined by executing a Monte Carlo (MC) conformational search that resulted in up to 100 unique structures. The top ten most

favorable MC structures were then recomputed using M06-2X/6-311+G(d,p) geometry optimizations and the resultant lowest energy structure was used as the starting geometry for the QM/MM calculations. A free-energy profile for the MHR reaction at 25°C and 1 atm was created by potentials of mean force (PMF) calculations coupled to MC statistical mechanics performed with *BOSS*.[18] The semiempirical method PDDG/PM3 was used to treat the solute in the QM region. The semiempirical QM method PDDG/PM3 has given excellent results for a wide variety of organic reactions in the solution phase.[19] The RTILs were represented explicitly using the custom OPLS-AA force field[20] and methanol was represented using the untied atom OPLS force field.[21] The systems consisted of the reactants plus 390 solvent molecules for methanol or 188 ion pairs for the ionic liquids in boxes that are periodic and tetragonal with c/a = 1.5 where a is 26.7, 34.3, and 35.5 Å for methanol, [BMIM][BF_4], and [BMIM][PF_6], respectively. Long range electrostatic interactions were handled with the Ewald summations[22] and the shifted force 3rd derivative (SF3) which was previously tested on 59 unique ionic liquid combinations of 1-alkyl-3-methylimidazolium [RMIM] (R = M (methyl), E (ethyl), B (butyl), H (hexyl), and O (octyl)) and N-alkylpyridinium [RPyr] cations, along with Cl^-, PF_6^-, BF_4^-, NO_3^-, $AlCl_4^-$, $Al_2Cl_7^-$, and TfO^- anions.[23] Solutes were inserted with the appropriate solute geometry corresponding to each free energy perturbation (FEP) window and re-equilibrated.

Free energy maps were computed for the catalyzed and uncatalyzed pathways to identify the minima and the transition states present in the reaction. The free energy maps were created by perturbing the bond distance between the nucleophilic nitrogen atom and the nitrogen in the oxadiazole ring denoted R_{NN} from 1.30 Å to 2.80 Å in 0.05 Å

increments (see Schemes 5.3). A second perturbation was necessary, R_{NO}, which entailed the breaking of the oxadiazole ring at the N – O bond, from 1.40 Å to 2.50 Å in 0.005 Å increments. Combining the R_{NN} PMF which runs along one reaction coordinate with the R_{NO} PMF in a second direction produced a two-dimensional (2D) PMF. Each PMF calculation will require extensive reorganization of the solvent for the ionic liquids, and may require up to 160 million configurations of equilibration followed by 20 million MC steps of averaging per FEP window, as previously reported in our β-elimination[10] and S_NAr work (see unpublished results in Chapter 4). However in methanol, only 2 and 5 million steps of equilibration and averaging were required for the title MHR reaction.



Scheme 5.3: The PMF procedure to calculate the free energy of activation for (a) the base-catalyzed and (b) the uncatalyzed MHR of the Z-phenylhydrazone of 3-benzoyl-5-phenyl-1,2,4-oxadiazole into 4-benzoylamino-2,5-diphenyl-1,2,3-triazole.

5.3 Methanol Energetics

The QM/MM/MC calculations for the Boulton-Katritzky MHR of the Z-phenylhydrazone of 3-benzoyl-5-phenyl-1,2,4-oxadiazole (compound 1) into 4-benzoylamino-2,5-diphenyl-1,2,3-triazole (compound 2) for the catalyzed pathway in methanol gave a computed activation barrier, $\Delta G^{\ddagger}$, of 22.0 kcal/mol (exptl. 23.2 kcal/mol[4]), see Table 5.1. The calculated free energy of activation is in good agreement with the experimental value. A free-energy profile for the catalyzed MHR reaction is shown in Figure 5.2. To locate the critical points more precisely, the regions surrounding the free-energy minima and maxima from the initial maps were explored using final increments of 0.01 Å with increased sampling. This provided the refined results in the energetics and geometries. The resultant geometries of the reactant are R_{NN} of 2.69 Å and R_{NO} = 1.38 Å, and the transition state geometries are R_{NN} of 2.01 Å and R_{NO} = 1.75 Å (Table 5.1). It has been proposed that as the viscosity of the solvent increases bond making (R_{NN}) becomes favored and bond breaking (R_{NO}) becomes disfavored, ultimately affecting the transition state position on the reaction coordinate.[15] As seen in Table 5.1, only the QM/MM/MC values for the free energy of activation for the methanol catalyzed reaction are provided. The uncatalyzed pathway calculations (and RTILs calculations) are currently underway.

	$\Delta G^{\ddagger}$ (calc.)	R_{NN}	R_{NO}	$\Delta G^{\ddagger}$
catalyzed	22.0			23.2
Methanol GS	-	2.69	1.38	-
Methanol TS	-	2.01	1.75	-
Δ bond	-	0.68	0.37	-

[a] PDDG/PM3 and MC/FEP. [b] Ref 4.

Table 5.1. Free energy of activation, $\Delta G^{\ddagger}$ (kcal/mol), and geometries (Å) at 25 °C for the title MHR from QM/MM/MC calculations.[a]

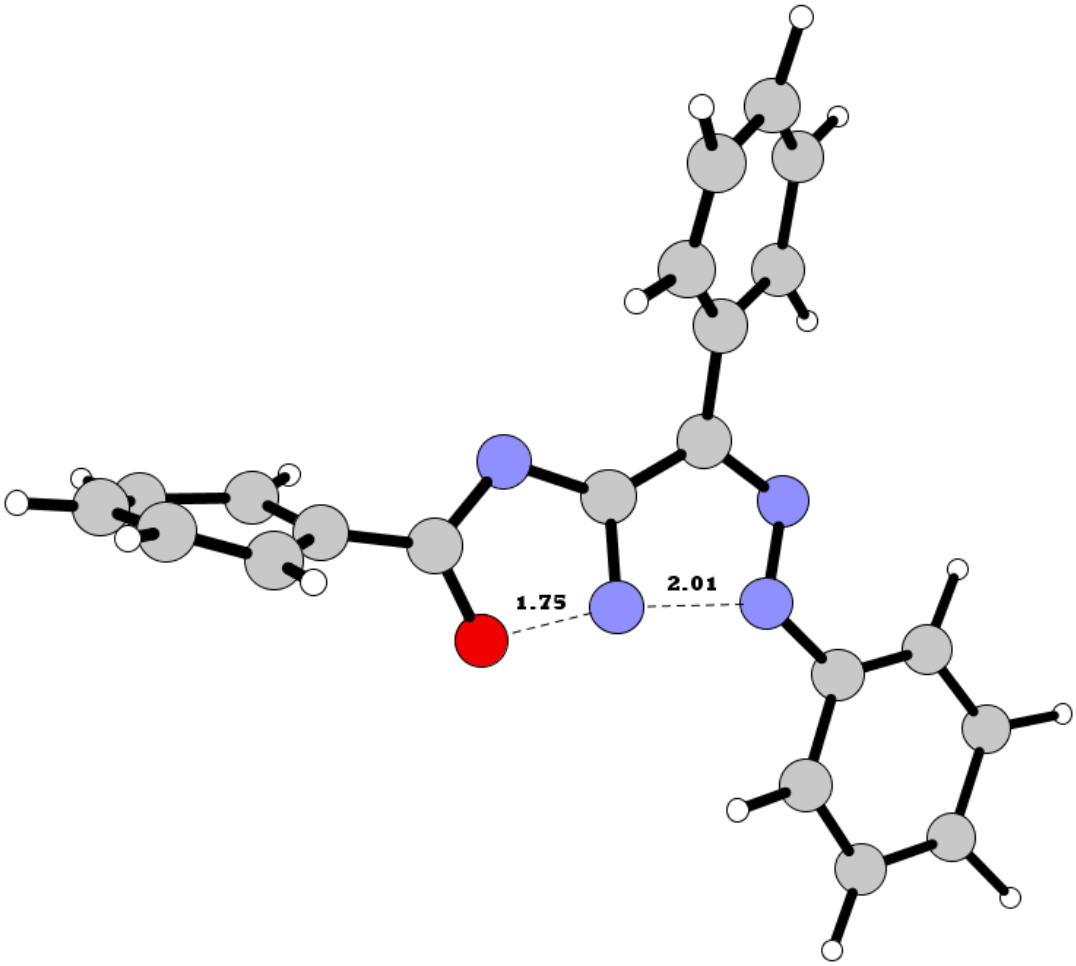


Figure 5.1: A snapshot of the 10π quasi-aromatic transition state observed for the MHR reaction in methanol. Bond lengths are given in Å.

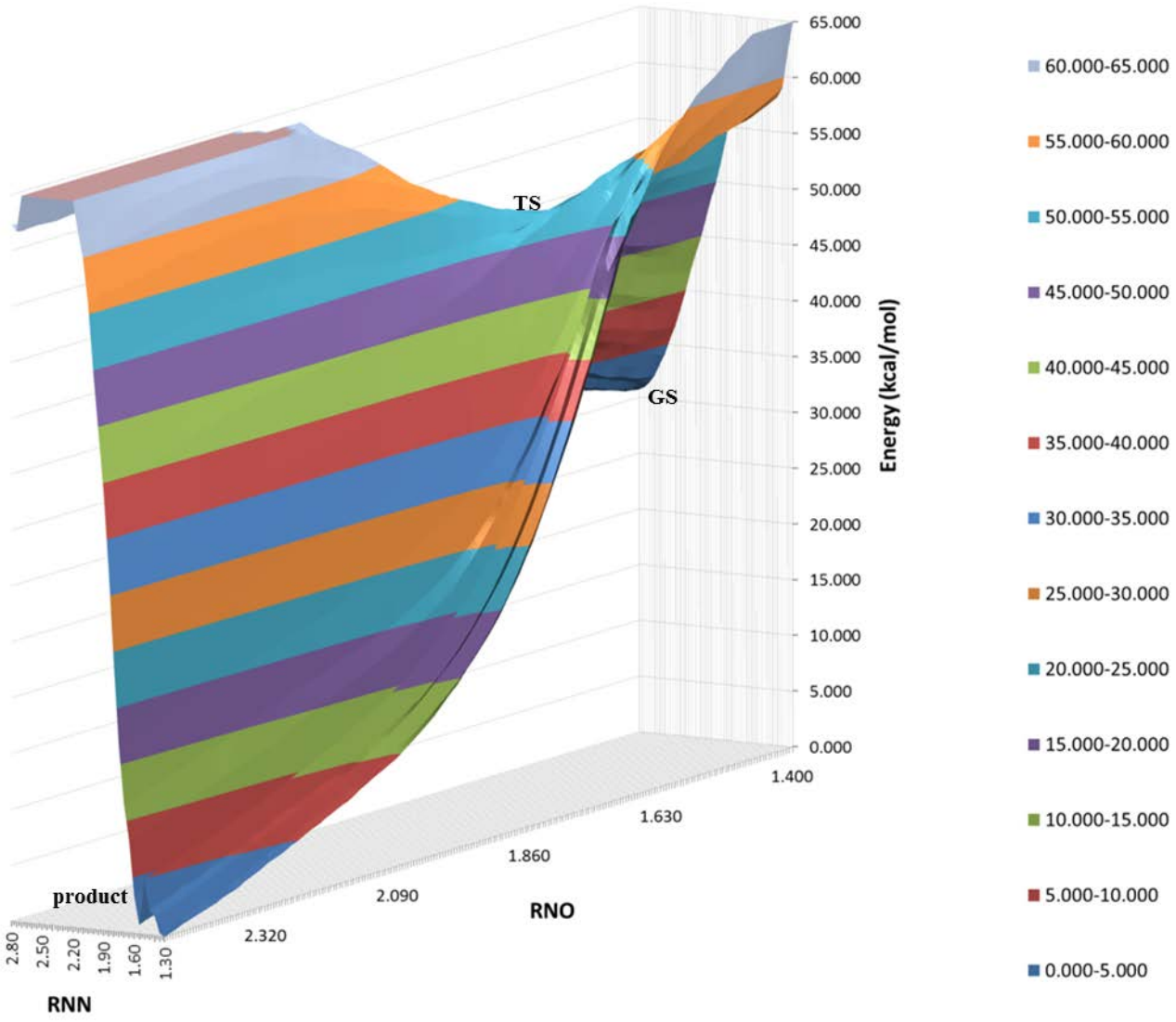


Figure 5.2: The free energy map (kcal/mol) computed for the MHR of the Z-phenylhydrazone of 3-benzoyl-5-phenyl-1,2,4-oxadiazole into 4-benzoylamino-2,5-diphenyl-1,2,3-triazole in methanol from QM/MM/MC simulations. Energy values truncated after 65 kcal/mol for clarity.

5.4 Solvent Effects

To elucidate the differences between the solute-solvent interactions occurring in the ground state (GS) and transition state (TS) for the catalyzed pathway, the energy pair

distributions from QM/MM/MC calculations in the representative FEP window near the corresponding geometry were analyzed (Figure 5.3). The distributions record the average number of methanol molecules that interact with the reacting system and their corresponding energies. Highly favorable electrostatic interactions between solute and solvent components are reflected in the left-most region with energies more attractive than ca. –5 kcal/mol. The large band near 0 kcal/mol arises from the many ions in outer shells. Integration of the distributions in Figure 5.3 from –13.5 to –5.0 (or -3.5) kcal/mol shows that the GS of the MHR of the Z-phenylhydrazone of 3-benzoyl-5-phenyl-1,2,4-oxadiazole into 4-benzoylamino-2,5-diphenyl-1,2,3-triazole has 8.7 (12.5) favorable interactions, which is greater than at the TS, 5.8 (10.6). The number of solute-solvent interactions decreases by 2 – 3 solvent molecules in going from the reactants to transition state in methanol, suggesting a destabilization of the transition state.

A planar orientation of the phenyl rings at the ground state should maximize the electronic effects exerted on the reaction pathway. It is proposed that favorable π-π interactions with the ionic liquid cation [BMIM] forces a coplanarity between aromatic rings, which is consistent with previous hypotheses.[10, 24] Figure 5.4 illustrates the three dihedral angles used to describe the planarity of the phenyl ring substituents. Table 5.2 shows that for the catalyzed pathway in methanol, the TS has the phenyl substituents in a more planar, less twisted configuration (Φ_1 = 162.2, Φ_2 = 169.4, and Φ_3 = 198.6 degrees) with the reacting center than the GS (Φ_1 = 137.0, Φ_2 = 183.2, and Φ_3 = 147.0 degrees), suggesting that methanol does not fulfill the structural requirements of the GS to enhance the rates of reaction.

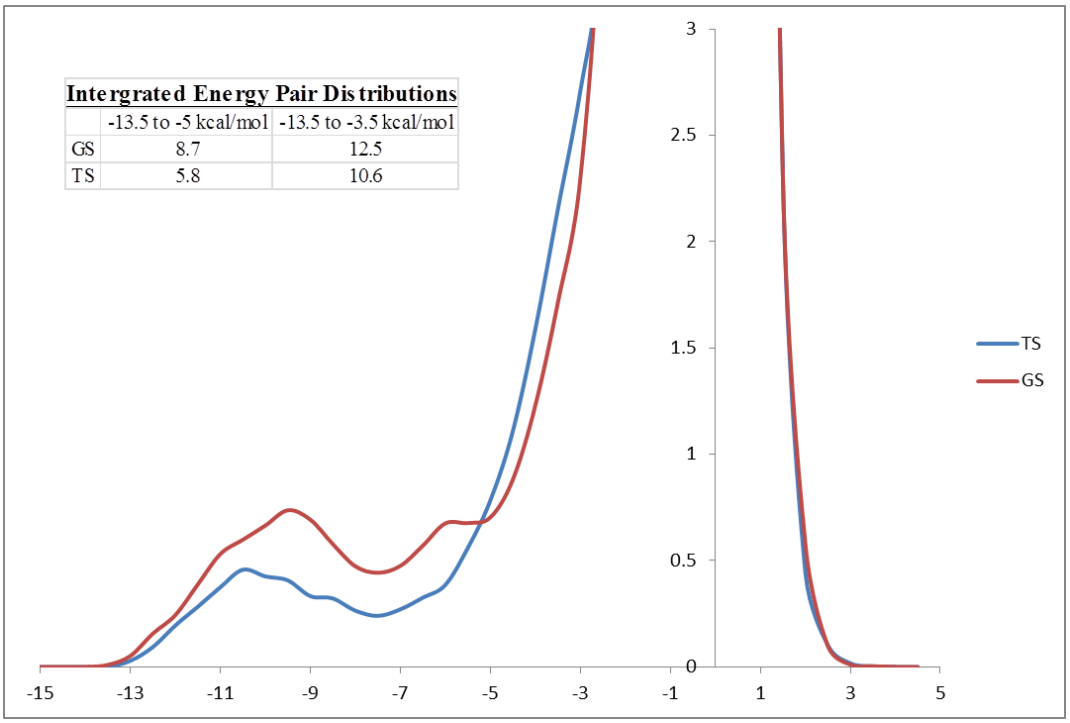


Figure 5.3: The solute-solvent energy pair distributions of the ground state (GS) and transition state (TS) for the catalyzed MHR of the Z-phenylhydrazone of 3-benzoyl-5-phenyl-1,2,4-oxadiazole into 4-benzoylamino-2,5-diphenyl-1,2,3-triazole in methanol from QM/MM/MC simulations. The ordinate records the number of solvent molecules that interact with the solute and their interaction energy on the abscissa. Units for ordinate are number of molecules per kcal/mol.

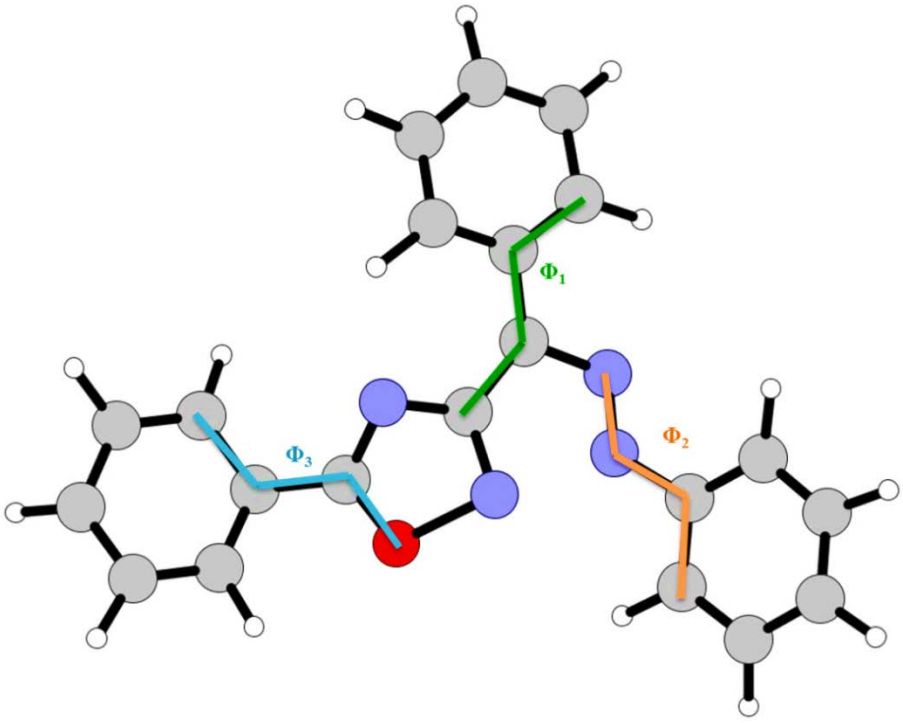


Figure 5.4: Snapshot to illustrate the three dihedrals of interest used to determine the planarity of the system, Φ_1 = C-C-C-C, Φ_2 = N-N-C-C, and Φ_3 = O-C-C-C.

	Φ_1	Φ_2	Φ_3
Catalyzed MeOH			
GS	137.0	183.2	147.0
TS	162.2	169.4	198.6

Table 5.2: Average dihedral angles (degrees, see Figure 5.4 for definition of Φ_1, Φ_2, and Φ_3) for the MHR at the ground state (GS) and transition state (TS) in methanol.

5.5 Conclusions

QM/MM/MC calculations have been performed for the Boulton-Katritzky MHR of the Z-phenylhydrazone of 3-benzoyl-5-phenyl-1,2,4-oxadiazole (compound 1) into 4-benzoylamino-2,5-diphenyl-1,2,3-triazole (compound 2) for the catalyzed pathway in methanol and gave a computed activation barrier, $\Delta G^{\ddagger}$, of 22.0 kcal/mol (exptl. 23.2 kcal/mol[4]). It has been proposed that as the viscosity of the solvent increases: methanol < [BMIM][BF4] < [BMIM][PF_6], bond making becomes favored and bond breaking becomes disfavored, ultimately affecting the transition state position on the reaction coordinate.[15] For the catalyzed pathway in methanol the geometries of the reactant were R_{NN} of 2.69 Å and R_{NO} = 1.38 Å, and the transition state geometries were R_{NN} of 2.01 Å and R_{NO} = 1.75 Å. The solute-solvent interactions occurring in the ground state and transition state for the catalyzed pathway were determined by energy pair distributions. The ground state of the MHR had 8.7 (12.5) favorable interactions in methanol, which was greater than the number of favorable interaction occurring at the transition state in methanol, 5.8 (10.6) (Figure 5.3). The number of solute-solvent interactions decreased by 2 – 3 solvent molecules in going from the reactants to transition state in methanol, suggesting a destabilization of the transition state. Additionally, it is proposed that favorable π-π interactions with the ionic liquid cation [BMIM] force a coplanarity between aromatic rings in the ground state maximizing the electronic effects exerted on the reaction pathway. Three dihedral angles are used to describe the planarity of the phenyl ring substituents and it was found that for the catalyzed pathway in methanol, the transition state had the phenyl substituents in a more planar, less twisted configuration with the reacting center than the ground state (Table 5.2). This suggests that methanol

cannot satisfy the structural requirements of the ground state to enhance the rates of reaction.

The Boulton-Katritzky rearrangements are an intramolecular nucleophilic substitution reaction, which is reported to occur through a bicyclic 10π quasi-aromatic transition state, and thus represent a special case of S_N2 displacement reactions. Ultimately the purpose of exploring this particular mechanism is to elucidate how the RTILs provide a significant increase in the rate of MHR when compared to methanol. The accelerated rates of rearrangement have been attributed to their partially pre-organized structure, such that the ionic liquids can affect substrates that have structural requirements through favorable π-π interactions with the [BMIM] cation. This pre-organized structure and favorable interactions will allow the reactant to adopt a pre-ordered geometric arrangement that may enhance the reaction rate. The uncatalyzed pathway calculations in methanol and the RTILs calculations are currently underway.

Chapter 5 References

1. (a) Boulton, A. J., *Lectures in Heterocyclic Chmistry; Hetero-Corporation*. Provo, UT., 1973; (b) Boulton, A. J.; Katritzky, A. R.; Majid-Hamid, A. J., *J. Chem. Soc. C* **1967**, 2005-2007; (c) Katritzky, A. R.; Gordev, M. F., *Heterocycles* **1993,** *35*, 483-518; (d) Ghosh, B. P., *J. Chem. Soc. B* **1968**, 334-338; (e) Van der Plas, H. C., *Ring Transformations of Heterocycles*. Academic Press: London, 1973; Vol. 1 and 2; (f) L'abbe, G. J., *Heterocycl. Chem.* **1984,** *21*, 627-638; (g) Ruccia, M.; Vivona, N.; Buscemi, S.; Frenna, V.; Cudmano, G., *Adv. Heterocycl. Chem.* **1993,** *56*, 49-154.
2. Korbonits, D.; Bako, E. M.; Horvath, K., *J. Chem. Res., Synop.* **1979,** (64-65).
3. (a) Ruccia, M.; Vivona, N.; Spinelli, D., *Adv. Heterocycl. Chem.* **1981,** *29*, 141-169; (b) Bottoni, A.; Frenna, V.; Lanza, C. Z.; Macaluso, G.; Spinelli, D., *J. Phys. Chem. A* **2004,** *108*, 1731-1740.
4. D'Anna, F.; Frenna, V.; Noto, R.; Pace, V.; Spinelli, D., *J. Org. Chem.* **2006,** *71*, 9637-9642.
5. (a) D'Anna, F.; Frenna, V.; Lanza, C. Z.; Macaluso, G.; Spinelli, D., *J. Org. Chem.* **2005,** *70*, 2828-2831; (b) Frenna, V., *J. Chem. Soc. Perkin Trans.* **1990,** *2*, 1289-1295.
6. (a) La Manna, G.; Buscemi, S.; Frenna, V.; Vivona, N.; Spinelli, D., *Heterocycles* **1991,** *32*, 1547-1557; (b) La Manna, G.; Buscemi, S.; Vivona, N., *J. Mol. Struct. (THEOCHEM)* **1998,** *452*, 67-74; (c) Andrianov, V. G.; Makushenkov, S. V.; Eremeev, A. V., *Mendeleev Commun.* **1992**, 129-130.

7. Frenna, V.; Vivona, N.; Consiglio, G.; Spinelli, D., *J. Chem. Soc., Perkin. Trans.* **1983,** *2,* 1199-1202.

8. Reichardt, C.; Welton, T., *Solvents and Solvent Effects in Organic Chemistry*. 4 ed.; Wiley-VCH: Weinheim, Germany, 2011.

9. Cosimelli, B.; Frenna, V.; Guernelli, S.; Lanza, C. Z.; Macaluso, G.; Petrillo, G.; Spinelli, D., *J. Org. Chem.* **2002,** *67*, 8010-8018.

10. Allen, C.; Sambasivarao, S. V.; Acevedo, O., *J. Am. Chem. Soc.* **2013,** *31*, 1065-1072.

11. Hallet, J. P.; Welton, T., *Chem. Rev.* **2011,** *111*, 3508-3576.

12. (a) Consorti, C. S.; Suarez, P. A. Z.; de Souza, R. F.; Burrow, R. A.; Farrar, D. H.; Lough, A. J.; Loh, W.; da Silva, L. H. M.; Dupont, J., *J. Phys. Chem. B* **2005,** *109*, 4341-4349; (b) Canongia, J. N.; Padua, A. A. H., *J. Phys. Chem. B* **2006,** *110*, 3330-3335.

13. Weiss, R. G., *Tetrahedron* **1988,** *44*, 3413-3474.

14. (a) Huddleston, J. G.; Visser, A. E.; Reichert, W. M.; Willauer, H. D.; Broker, G. A.; Rogers, R. D., *Green Chem.* **2001,** *3*, 156-164; (b) Pringle, J. M.; Golding, J.; Baranyai, K.; Forsyth, C. M.; Deacon, G. B.; Scott, J. L.; MacFarlane, D. R., *New J. Chem.* **2003,** *27*, 1504-1510.

15. Swiss, K. A.; Firestone, R. A., *J. Phys. Chem. A* **1990,** *103*, 5369-5372.

16. Antonietti, M.; Kuang, D.; Smarsly, B.; Zhou, Y., *Angew. Chem., Int. Ed.* **2004,** *43*, 4988-4992.

17. Schroder, U.; Wadhawan, J. D.; Compton, R. G.; Marken, F.; Suarez, P. A. Z.; Consorti, C. S.; de Souza, R. F.; Dupont, J., *New J. Chem.* **2000,** *24*, 1009-1015.

18. Jorgensen, W. L.; Tirado-Rivers, J., *J. Comput. Chem.* **2005,** *25*, 1689-1700.

19. (a) 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; (b) 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; (c) Acevedo, O.; Jorgensen, W. L., Exploring Solvent Effects upon the Menshutkin Reaction Using a Polarizable Force Field. *J. Phys. Chem. B* **2010,** *114* (25), 8425-8430; (d) 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; (e) Sheppard, A. N.; Acevedo, O., Multidimensional Exploration of Valley-Ridge Inflection Points on Potential Energy Surfaces. *J. Am. Chem. Soc.* **2009,** *131*, 2530-2540.

20. 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.

21. Jorgensen, W. L., Optimized intermolecular potential functions for liquid alcohols. *J. Phys. Chem.* **1986,** *90* (7), 1276-1284.

22. Ewald, P., *Ann. Phys.* **1921,** *369*, 253-287.

23. McCann, B. W.; Acevedo, O., *J. Chem. Theory Comput.* **2012,** *in press*.

24. (a) Deetlefs, M.; Hardacre, C.; Nieuwenhuyzen, M.; Sheppard, O.; Soper, A. K., Structure of Ionic Liquid−Benzene Mixtures. *J. Phys. Chem. B* **2005,** *109* (4), 1593–1598; (b) Holbrey, J. D.; Reichert, W. M.; Nieuwenhuyzen, M.; Sheppard,

O.; Hardacre, C.; Rogers, R. D., Liquid clathrate formation in ionic liquid-aromatic mixtures. *Chem. Commun.* **2003**, 476-477.