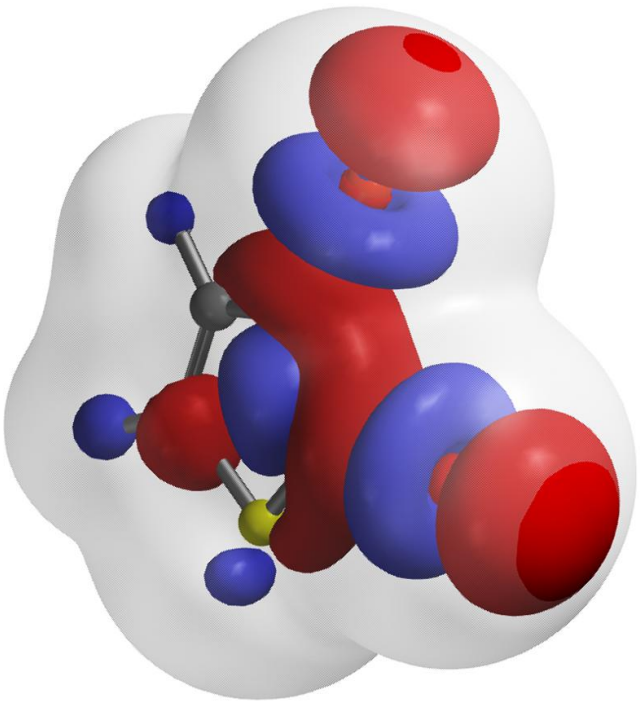
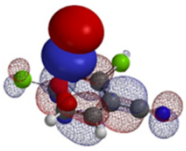
WuXi AppTec

Quantum Mechanics for Organic Chemists

An Experimentalist Approach

Fourth Edition





WuXi AppTec

Edited by

Liting Dong, Qiuyue Wang, Yongsheng Chen, John S. Wai

Quantum Mechanics for Organic Chemists

An Experimentalist Approach

Fourth Edition

ISBN 978-1-7354471-5-5 (eBook)

Front cover: 2,3-Dibromothiophene LUMO+1/Density Map. On why lithiation of it is C2 selective.

Table of Contents

Foreword

Science is human's continuous pursuit for better understanding of the physical world. In organic chemistry, we use structural formula and arrow-pushing extensively to rationalize intuitively reactivity observed. This is proven to be a powerful approach, yet it also leaves us with a lot of observations incomprehensible.

Quantum Mechanics, Molecular Orbital Theory in particular, enables us to understand organic reactions at deeper levels. It is a powerful retrospective analysis tool for computational chemists. Building upon our efforts on retrospective analyses of chemistry with Quantum Mechanics, we have successfully incorporated these calculations in our daily prospective retrosynthetic analyses.

Our laboratory chemists initiated a writing project to share with the chemistry community what we have learned with this "Magical Power", focusing on parameters that are useful in correlating with chemical reactivity and reaction mechanisms of interest. This fourth edition of our eBook collated the first fifty articles from their efforts. Furthermore, these relevant Quantum Mechanics parameters highlighted are essential for creating high performing Machine Learning chemistry tools and are keys for our scientists to understand their output. Incorporation of these relevant tools into our daily workflow enables our scientists to develop deeper insights to reaction mechanisms, make predictions and decisions with relevant data, accelerate discoveries and innovations, etc.

Please join us in this pursuit and enjoy the Eureka moments ahead.

John S. Wai & Tao Guo

July 2024

Summary of QM calculations

Dong Pan, Wenfeng Liu, Liting Dong, Tommy Lai, Chen Yongsheng

Structure determination	**Calculate ^{13}C NMR for all possible isomers（Chapter 11）**
	Compare with experimental ^{13}C NMR (in $CDCl_3$). The structure with smallest differences is most likely.
Atropisomerism	**Torsional scan and rotation energy barrier（Chapters 14, 40）**
	Class 1 (ΔE_{rot} < 20 kcal/mol), no need of SFC separation; Class 2 (20 kcal/mol < ΔE_{rot} < 30 kcal/mol), possible interconversion after SFC separation; Class 3 (ΔE_{rot} > 30 kcal/mol), SFC separation into individual isomers. Calculate equilibrium ratio of atropisomers based on the energy difference of low-energy conformations.
Relative acidity, deprotonation site	**Calculate Electrostatic Potential（Chapters 3, 43）**
	The site with highest ESP Max value ("bluest") is most acidic and preferred deprotonation site.
Relative basicity, protonation site	**Calculate Electrostatic Potential or Natural charge（Chapter 32）**
	The site with lowest ESP Max value ("reddest") is most basic or with most negative natural charge.
Nucleophilic reactivity	**Calculate HOMO (or HOMO-1, HOMO-2) energy（Chapters 25, 27, 30, 46）**
	Higher HOMO energy means higher nucleophilic reactivity. Smaller energy difference between electrophile LUMO and nucleophile HOMO means lower activation energy. Calculate HOMO-1 or HOMO-2 energy instead when there is no lobe on HOMO.
Electrophilic reactivity	**Calculate LUMO (or LUMO+1, LUMO+2) energy（Chapters 7, 25, 27, 29, 43, 44, 45, 47, 48）**
	Lower LUMO energy means higher electrophilic reactivity. Smaller energy difference between electrophile LUMO and nucleophile HOMO means easier reaction. Inspect LUMO+1 or LUMO+2 energy if there is no suitable lobe on LUMO.
Stereoselectivity of carbanion attack	**Calculate HOMO map and Inaccessible Markers for most dominant conformation of carbanion（Chapter 4）**
	Attack to electrophile occurs on the more accessible side, that is, with fewer inaccessible markers.
Reaction site in Chan-Lam	**Calculate HOMO of substrate anion（Chapters 34, 43, 47）**
	Reaction preferentially occurs at the N anion with bigger HOMO lobe.
Regioselectivity of oxidative addition	**Calculate LUMO or LUMO+1 and IR（Chapters 5, 47）**
	Reaction preferentially occurs at the site with bigger LUMO (or +1) lobe. If inconclusive, reaction at the C-X bond with lower IR stretching frequency is favored.
Regioselectivity of electrophilic aromatic halogenation	**Calculate substrate HOMO and ^{13}C NMR; substrate-X^+ relative energy and activation energy（Chapters 2, 6, 17, 18, 23, 27, 30）**
	Reaction preferentially occurs at the site with bigger HOMO lobe. If inconclusive, reaction at the carbon with more upfield chemical shift in ^{13}C NMR is favored, highly selective when the difference of chemical shift is > 6 ppm.
	From calculation of equilibrium geometry and relative energy of possible substrate-X^+ halonium ions, the lowest energy one is preferred. For estimation of product ratio, calculate activation energy.
Regioselectivity of nucleophilic substitution	**Calculate LUMO/LUMO+1, LUMO map, and activation energy（Chapters 1, 7, 10, 18, 19, 27, 29, 35, 36, 37）**
	Reaction preferentially occurs at the site with bigger LUMO lobe. If the energy difference between LUMO+1 and LUMO is small, consider LUMO+1 lobe too.
	If inconclusive by LUMO lobe, from LUMO map with inaccessible markers, the "bluest" and most accessible site is preferred. For estimation of product ratio, calculate activation energy.
Product ratio or reaction mechanism	**Calculate difference in activation energy from reaction energy profile of competing pathways（Chapters 9, 26: pyrazole N-alkylation; Chapter 10: S_NAr; Chapters 11, 15, 16, 18, 22, 23, 24, 28, 31, 36, 37, 50）**
	Simplify substrate structure; consider all possible bonding pathways and calculate reaction energy profile; identify transition state and validate by iFreq; calculate product ratio based on difference in activation energy.
	Activation energy < 10 kcal/mol, reaction occurs at low or room temp; 10-20 kcal/mol, reaction needs heating; 20-40 kcal/mol, reaction needs heating and longer time; > 40 kcal/mol, reaction is unlikely to occur.
Regioselectivity of metal-halogen exchange	**Calculate LUMO or LUMO+n, relative energy of carbanion（Chapters 39, 47）**
	Reaction preferentially occurs at the site with bigger string-of-pearl shaped lobe, which might be on LUMO or LUMO+1/2/3. If inconclusive, reaction occurs at the carbon with lowest relative energy of corresponding carbanion.

Chapter 1 Application of LUMO Analysis in Nucleophilic Reactions

Qiuyue Wang, Guqin Shi, John S. Wai

Would you like to predict organic reaction outcomes accurately? Design synthetic sequences with higher rate of success?
Would you prefer to learn and rationalize organic chemistry in a data-driven manner?
Please join us to unleash the magical power of Quantum Mechanics (QM).

First Example

We learned that nucleophilic substitution of 2,4-dichloro-5-methylpyrimidine with a nucleophile, such as piperidine, proceeds preferentially at C-4 position (Figure 1). Intuitively it makes more sense for the substitution to occur faster at C-2, which has two electron withdrawing aromatic nitrogen atoms next to it. Why is it not the case? Furthermore, when the nucleophile is alkyllithium, only C-6 addition was observed. Why?

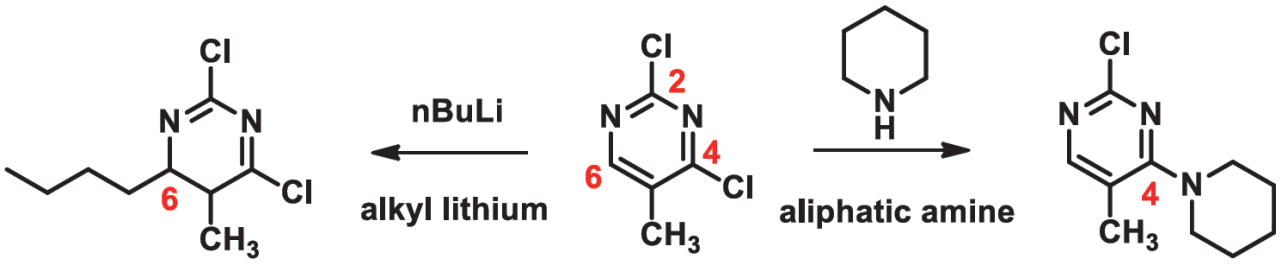


Figure 1. Nucleophilic reaction of 2,4-dichloro-5-methylpyrimidine with alkyllithium or aliphatic amine

The question becomes “What could be a better scientific method to rationalize, to learn these chemistry?” In our synthetic organic chemistry laboratories, we found QM to be the most valid and efficient tool for these purposes. In this article, we will discuss how we use LUMO and LUMO map to analyze nucleophilic aromatic substitution reactions. You will see that it is relatively easy to understand and apply such analyses in your laboratories.

Frontier Molecular Orbital Theory (FMO)

First let's review the basics of Molecular Orbital (MO) Theory, which proposes that when atoms come together to constitute a molecule, their atomic orbitals interact with each other and form new molecular orbitals. These orbitals have their own energy levels and unique spatial distribution. For organic reaction, we could focus on the use of Frontier Molecular Orbital (FMO) Theory, which emphasizes the analyses with outer orbitals, i.e., the Highest Occupied Molecular Orbital (HOMO) and the Lowest Unoccupied Molecular Orbital (LUMO).

Based on FMO theory, an organic reaction is electron transfer from nucleophilic part (HOMO) to the electrophilic part (LUMO) of the reacting system. It enables chemists to understand how orbitals interact with one another throughout the course of reactions, and to account for chemical reactivity & selectivity observed, etc.

Now let's use FMO theory to rationalize the aforementioned nucleophilic reactions with the dichloropyrimidine (Figure 1).

Using Spartan for Our QM calculations

There are a number of powerful QM computing tools available to us. Our laboratory chemists prefer to use Spartan, which has a user-friendly graphic interface, making calculations relatively easy to set up and the results easier to interpret. We use Density Functional Theory (DFT) ωB97X-D/basis set 6-31G* for our calculations unless otherwise stated.

First, let's calculate the LUMO and its associated orbital energies of 2,4-dichloro-5-methylpyrimidine (the electrophile in the reaction) with Spartan (Figure 2).

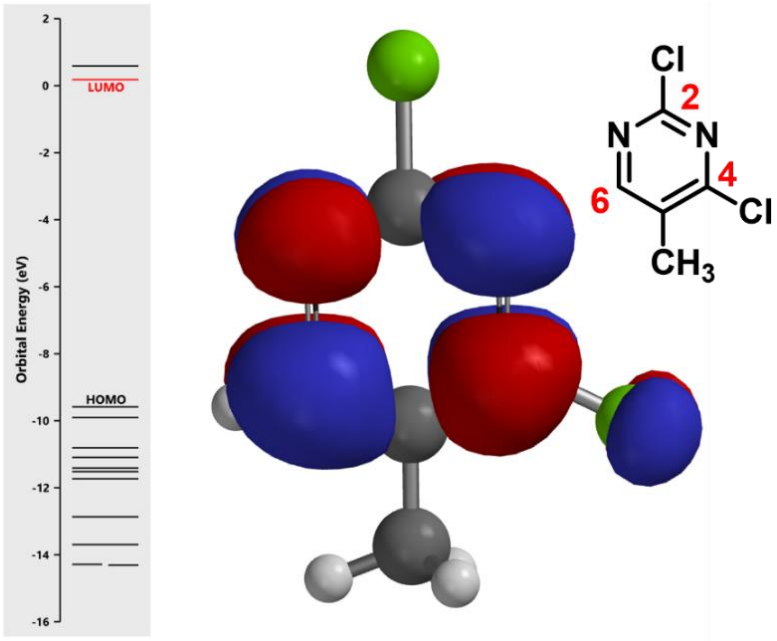


Figure 2. LUMO of 2,4-dichloro-5-methylpyrimidine
(Left: Molecular orbital energy level, red line represents the LUMO energy level.
Middle: LUMO superimposed onto ball & stick model.)

Notice that there is no LUMO lobe over C-2, as such, accounting for no displacement reaction observed at this carbon. Whereas at C-4 and C-6, both have LUMO lobes associated with them, as such these LUMO lobes will be able to interact with incoming nucleophiles. Moreover, addition of piperidine at C-6 site is reversible, while addition at C-4 site will lead to displacement of the chloro group, an irreversible event. That leads to the overall transformation we observed (Figure 3).

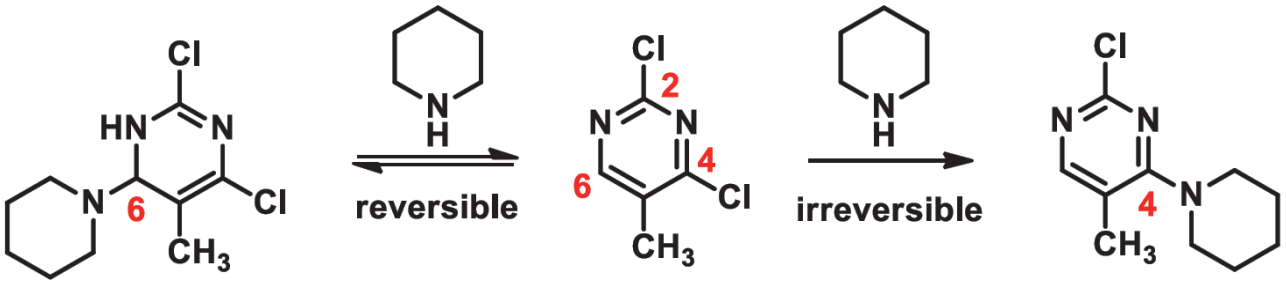


Figure 3. Reversible and irreversible interaction of piperidine and 2,4-dichloro-5-methylpyrimidine at C-6 and C-4 positions

Then why treatment of 2,4-dichloro-5-methylpyrimidine with n-butyllithium affords only the C-6 addition product? It is the same substrate, same LUMO.

To understand this, we calculated for the LUMO map, which is an overlay of LUMO and Electron density surface of the molecule. This provides us a quantitative comparison of which part of the molecule is more accessible to attack by an incoming nucleophile. In the default settings of Spartan, colors near red indicate that the LUMO value at this location is close to zero, i.e., a low probability of being attacked. On the contrary, colors near blue indicate a much higher level, suggesting that the LUMO lobe at this location is highly accessible.

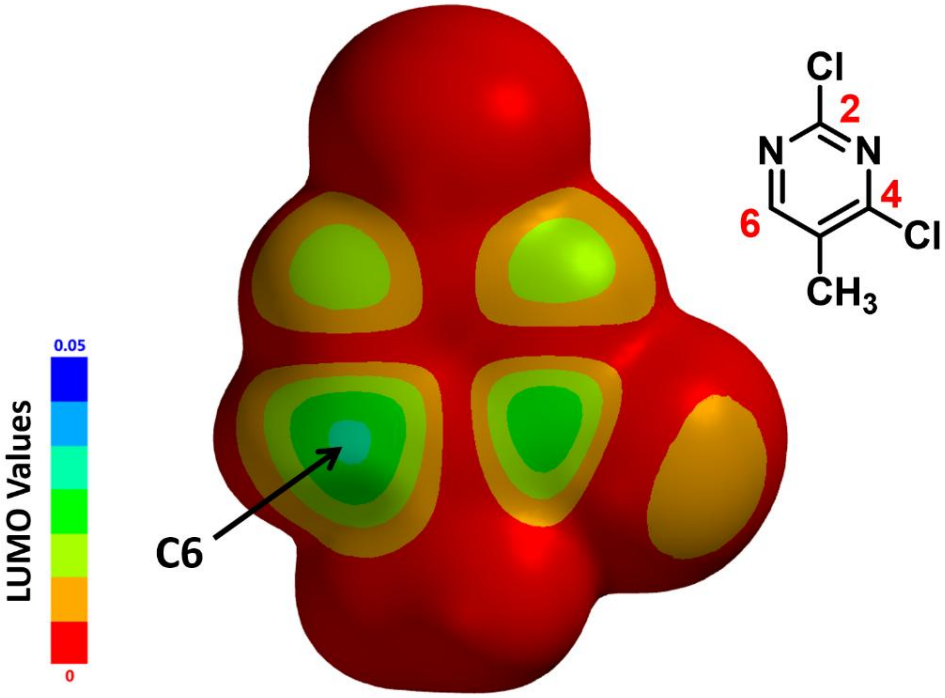


Figure 4. LUMO map of 2,4-dichloro-5-methylpyrimidine
Note: The LUMO map "paints" the absolute value of the LUMO onto an electron density surface

With the LUMO map of 2,4-dichloro-5-methylpyrimidine (Figure 4), it becomes obvious that the C-6 position is more accessible to nucleophilic attack than the C-4 carbon. Furthermore, addition of n-butyllithium at C-6 site is irreversible, as such, we could only observe the C-6 addition product. This is a highly selective addition reaction under kinetic control (Figure 5).

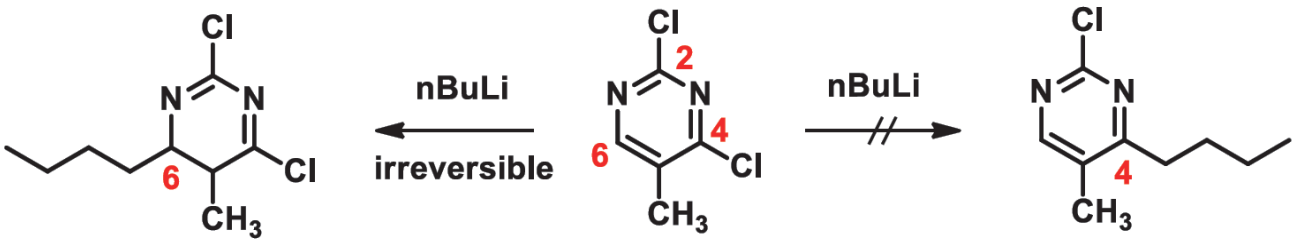


Figure 5. Treatment of 2,4-dichloro-5-methylpyrimidine with n-butyllithium

In summary, we integrate QM (calculated LUMO, LUMO map), our knowledge of the inherent properties of different nucleophiles, and the stability of their addition products, to rationalize this set of pyrimidine chemistry. And we learned that knowledge gained from retrospective analyses of our chemistry with QM enables us to transform it into a powerful prospective tool.

Building on What We Just Learned

If the methyl group in the substrate (2,4-dichloro-5-methylpyrimidine) is replaced with Cl, Br, CF_3, and CCl_3, which position do you think an aliphatic amine will attack, respectively? What will be the rationale? (Hint: Look for differences in LUMOs of the CF_3 and CCl_3 analogs.)

Reaction Site	5-CH_3	5-Cl	5-Br	5-CF_3	5-CCl_3
C2	✘				
C4	✓				

Return to Table of Contents

References:

[1] W. Hehre and S. Ohlinger, *A Guide to Molecular Mechanics and Quantum Chemical Calculations.* Irvine, CA, USA: Wavefunction, Inc., **2003**.

[2] J.A. Joule & K. Mills, "The Diazines: Pyridazine, Pyrimidine, and Pyrazine: Reactions and Synthesis." *Heterocyclic Chemistry 4th Ed.* Malden, MA, USA: Blackwell Publishing Ltd., **2000**.

Chapter 2 Application of HOMO Analysis in Electrophilic Reactions

Qiuyue Wang, Guqin Shi, Jian Wang, John S. Wai

Electrophilic aromatic substitution (S_EAr) is a very important class of reactions for chemists to introduce functional groups onto heterocycles.

Regioselectivity Prediction of S_EAr

We all learned that substituents on aromatic rings affect regioselectivity of electrophilic substitutions. With mono-substituted phenyl analogs, if the substituent is an electron-donating group, such as a methoxy group, we usually observe reactions at *ortho*- and *para*-positions. On the other hand, when the substituent is an electron-withdrawing group, such as a nitro group, *meta*-substituted products are usually produced. Halogen substituents are unique, they are electron-withdrawing, yet *ortho*-and *para*-directing (Figure 1).

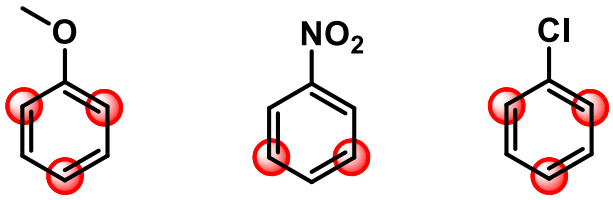


Figure 1. Regioselectivity of S_EAr on different mono-substituted benzenes

We also learned that predicting regioselectivity of S_EAr on complex multi-substituted heterocyclic systems could be more challenging.

For instance, with substrate **A** (Figure 2), which heterocycle would be halogenated first, the pyridazine or the pyrazole? For substrate **B**, would halogenation be on the fused ring or on the benzene ring, and at which position? For substrate **C**, it is quite clear that the imidazole would be halogenated. Will we have selective halogenation at C-2 or C-3, or a mixture of products? Intuitively, these are not easy to predict with high confidence.

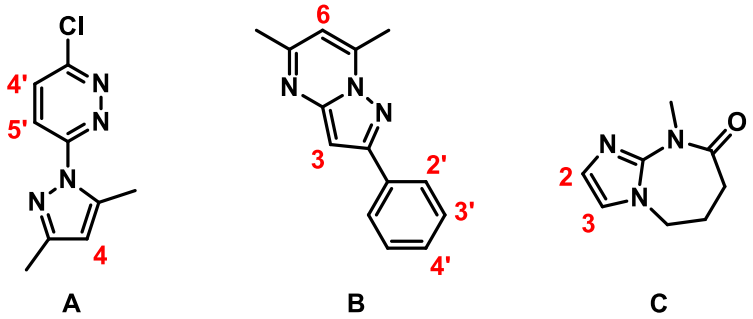


Figure 2. Potential sites on substrates for aromatic halogenation

In this article, we will discuss how we use calculated HOMO and ^{13}C NMR data to make predictions on the regioselectivity of S_EAr on heterocyclic aromatic systems.

Application of HOMO

Frontier Molecular Orbital (FMO) Theory emphasizes the use of the Highest Occupied Molecular Orbital (HOMO) for analyses of nucleophiles and the Lowest Unoccupied Molecular Orbital (LUMO) for electrophiles.

In Chapter 1, we showed how we utilize LUMO and LUMO map to analyze nucleophilic aromatic substitutions. Here we will be looking at HOMO of three heterocyclic substrates (Figure 3) to rationalize the regioselectivity observed with S_EAr

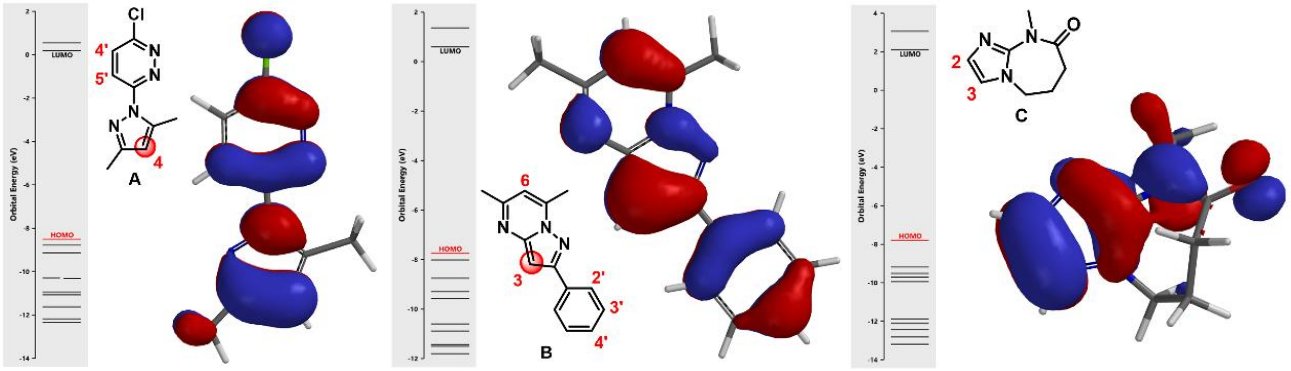


Figure 3. HOMO of substrates **A**, **B**, and **C**
In each diagram, on the left side shows the orbital energy level of the molecule, and the red line represents the HOMO energy level. On the right side shows the HOMO superimposed onto molecule in tube model.

For substrate **A**, among three potential sites, there are HOMO lobes on C-4 and C-5', which suggests only these two positions are more likely to be interacting with incoming electrophiles. The HOMO lobe over C-4 is much larger than the one over C-5' indicating that electrophiles would react selectively at C-4. Indeed we observed only the C-4 halogenated product experimentally. Similar for substrate **B**, the C-3 carbon on the fused ring is associated with the largest HOMO lobe among all potential sites of reaction, resulting in selective halogenation at C-3.

Things get slightly complicated for substrate **C**. The HOMO lobes covering C-2 and C-3 are similar in size on both atoms. In this case, will we observe a mixture of substitution products? What else have we learned to predict the regioselectivity in such circumstances?

Application of Calculated ^{13}C NMR

We found QM-calculated ^{13}C NMR chemical shift numbers of these carbons correlate quite well with regioselectivity of S_EAr.

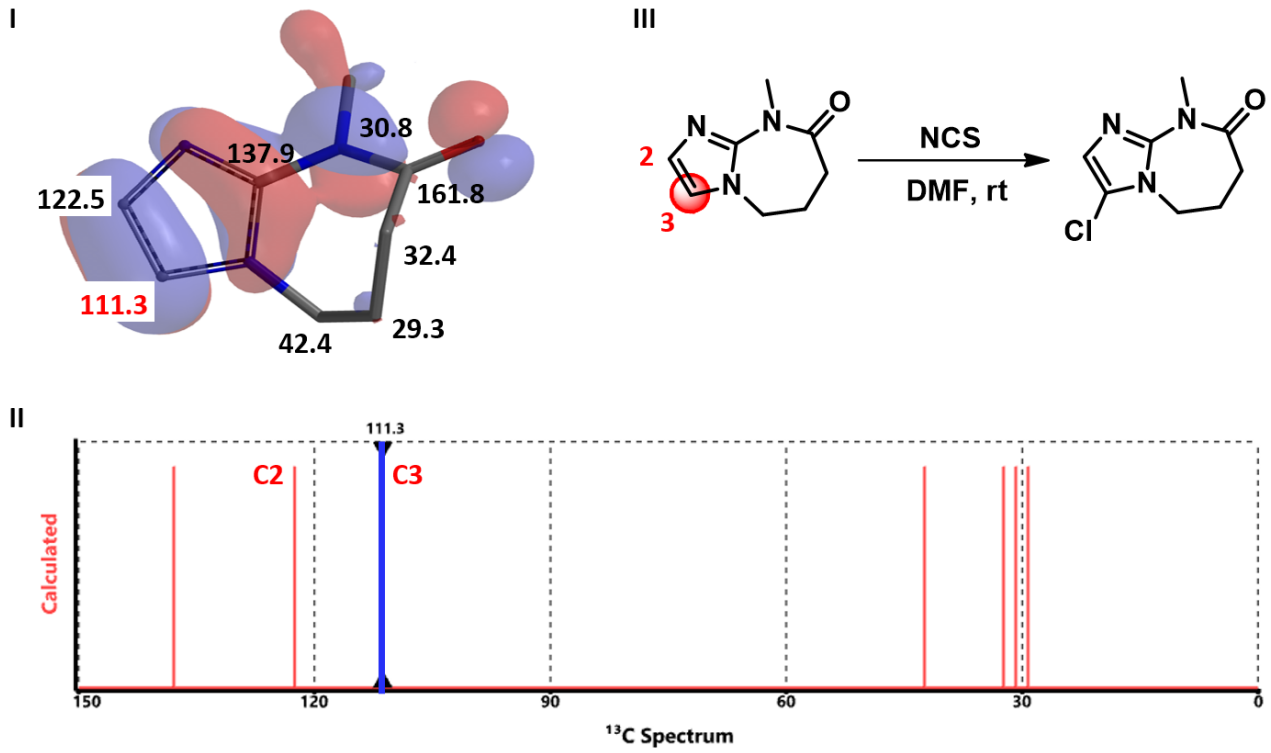


Figure 4. I: Both HOMO lobes and ^{13}C NMR chemical shift values are superimposed on substrate **C** in tube model II: Calculated NMR spectrum. III: Selective chlorination of substrate **C**

Calculated ^{13}C NMR of substrate **C** (Figure 4) shows a significant difference in chemical shift values of C-2 (122.5 ppm) *versus* C-3 (111.3 ppm). A lower chemical shift value on C-3 suggests larger electron density on the carbon and preference for C-3 halogenation. Indeed the reaction is observed to be highly selective at C-3.

In summary, by integrating QM-calculated HOMO and ^{13}C NMR data of the substrates, we could rationalize the electrophilic reactions and then predict regioselectivity reliably. This becomes part of the standard QM evaluations we do routinely in our retrosynthetic planning.

Building on What We Just Learned

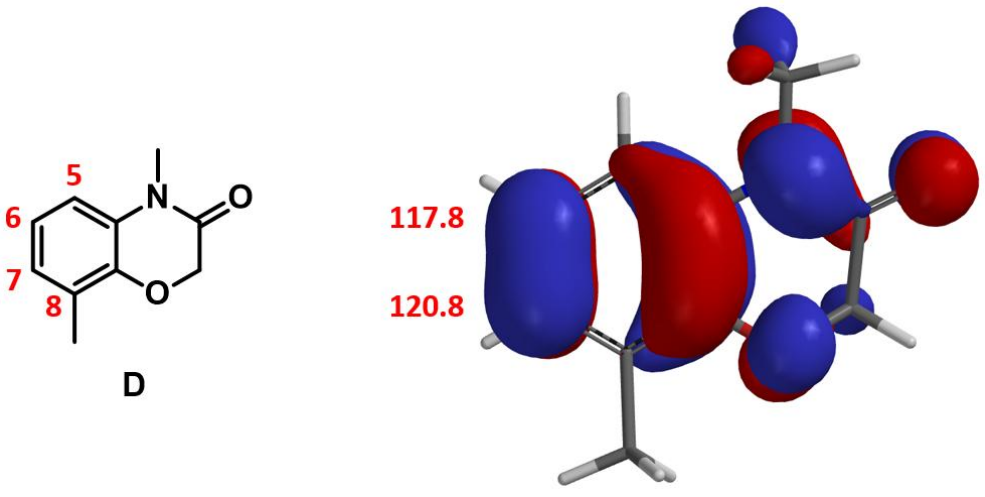


Figure 5. Compound **D**, HOMO and ^{13}C NMR chemical shift of C-6 and C-7.

On substrate **D**, there are three potential sites that could be brominated. They are C-5, C-6, and C-7. Since there is no HOMO lobe associated with C-5, we only need to evaluate relative reactivity at C-6 and C-7. However, differences in HOMO lobe volumes and the chemical shift values of C-6 and C-7 are not significant. Would C-8 methyl group exert any steric effect? What other QM parameters shall we calculate?

Return to Table of Contents

References:

[1] W. Hehre and S. Ohlinger, *A Guide to Molecular Mechanics and Quantum Chemical Calculations*. Irvine, CA, USA: Wavefunction, Inc., **2003**.
[2] F.A. Carey & R.J. Sundberg, *Advanced Organic Chemistry Part A: Structure and Mechanisms.* New York, NY, USA: Springer Science+Business Media, LLC., **2007**; pp 771-779.
[3] M. Kruszyk, M. Jessing, J.L. Kristensen, M. Jorgensen, *J. Org. Chem.* **2016**, *81*, 5128.

Chapter 3 Application of Electrostatic Potential Map in Acidity Estimation

Renwei Zhang, Guqin Shi, Tommy Lai, John S. Wai

In organic synthesis, we often need to estimate acidity of our substrates to choose a proper base for deprotonation. For simple compounds we refer to the p*K*a table. But for complex heterocyclic molecules, such as substrate (**1**) shown in Figure 1, in which two acidic hydrogens, H-1 and H-2, are susceptible to be deprotonated, it could be challenging to predict accurately.

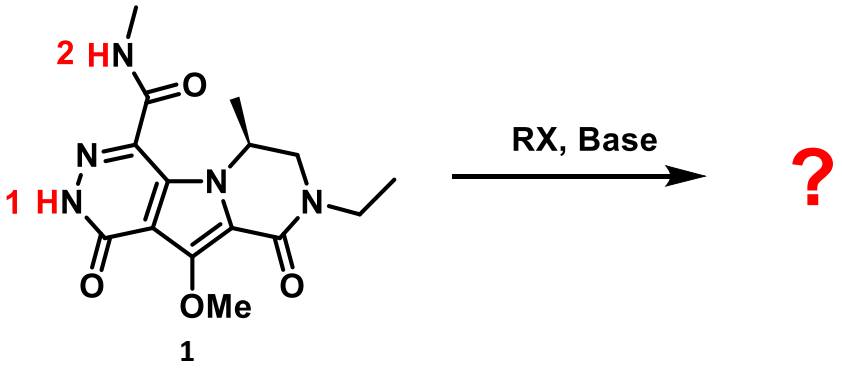


Figure 1. Selective deprotonation of **1** for subsequent alkylation

In this article, we would like to introduce the use of Electrostatic potential map (ESP map) to correlate with acidity and then use relative differences in electrostatic potential values to predict regioselectivity of deprotonation.

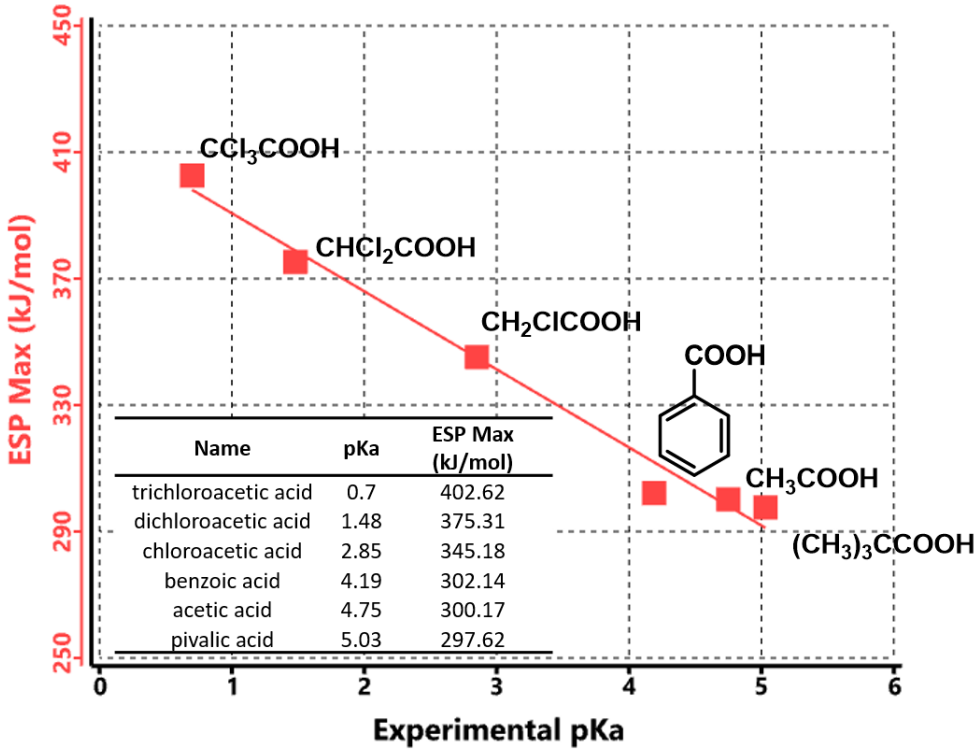
Electrostatic Potential

Electrostatic potential is the interaction energy of a point positive charge with the nuclei and electrons of a molecule. The value of electrostatic potential depends on the location of the point positive charge. If the point positive charge is placed in a region of excess positive charge (an electron-deficient region), the point charge-molecule interaction is repulsive and the electrostatic potential will be positive. On the contrary, if the point positive charge is placed in a region of excess negative charge (an electron-rich region), the interaction is attractive and the electrostatic potential will be negative.[2] By mapping the electrostatic potential onto an electron density surface (which represents the molecular shape), we could assess the charge distributions on molecules and identify the most electron deficient region(s).

Prediction on Acid Strength

The maximum values of the electrostatic potential in the vicinity of the acidic hydrogens are known to correlate with their p*K*a values.[3] Figure 2 shows experimental p*K*a values of a set

of carboxylic acids plotted along with the calculated maximum values of the electrostatic potential (ESP Max). The ESP Max values could be determined from the ESP map where the acidic hydrogen atoms are located, as they are the most electron deficient regions within these molecules.



Name	pKa	ESP Max (kJ/mol)
trichloroacetic acid	0.7	402.62
dichloroacetic acid	1.48	375.31
chloroacetic acid	2.85	345.18
benzoic acid	4.19	302.14
acetic acid	4.75	300.17
pivalic acid	5.03	297.62

Figure 2. Correlations between p*K*a and ESP Max values of carboxylic acids

These ESP Max values correlate almost linearly with reported experimental p*K*a values of the carboxylic acids. The more positive of the ESP Max value, the lower the p*K*a, the more acidic it is.

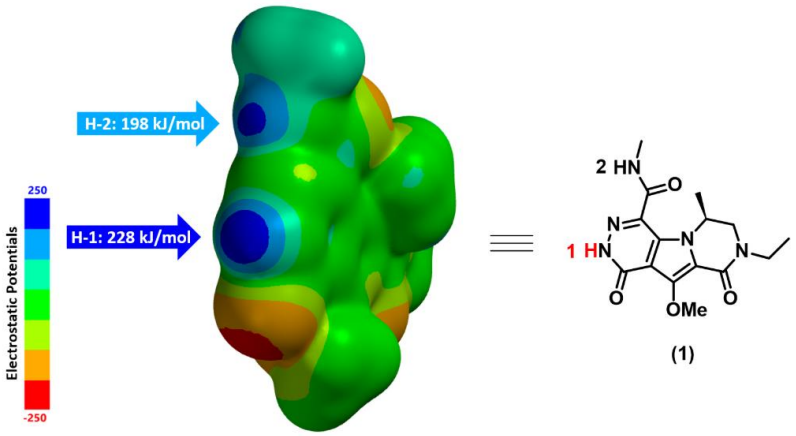
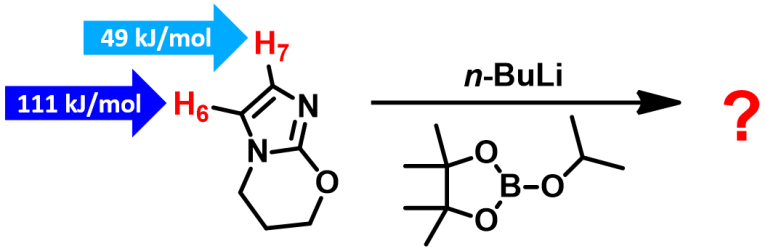


Figure 3. Electrostatic potential map (ESP map, electrostatic potential of a molecule overlay onto its electron density surface) of substrate **1**

Figure 3 displays the calculated ESP map of substrate **1**. By default, red color depicts regions of more negative electrostatic potentials (electron-rich) whereas blue color depicts regions of more positive electrostatic potentials (electron-deficient). The ESP Max value at H-1 site (228 kJ/mol) is much larger than the value at H-2 site (198 kJ/mol), suggesting that H-1 is significantly more acidic than H-2. We therefore predict that H-1 could be selectively deprotonated. Indeed we observed only the anticipated alkylated product experimentally.

In summary, QM-calculated ESP map enables us to estimate the relative acidity of deprotonable hydrogens, and therefore to predict their selective deprotonation with high confidence. This method to estimate relative acidity is routinely being used in our retrosynthetic planning.

Building on What We Just Learned



2

In substrate **2**, both H-6 and H-7 are susceptible to deprotonation with n-butyl lithium. With the calculated ESP Max values shown in the figure, we could readily predict which hydrogen will be deprotonated, and as such, the structure of the resultant boronate ester. It is quite reliable.

Return to Table of Contents

References:

[1] W. Hehre and S. Ohlinger, *A Guide to Molecular Mechanics and Quantum Chemical Calculations*. Irvine, CA, USA: Wavefunction, Inc., **2003**.
[2] *Spartan'18 Tutorial and User's Guide.* Irvine, CA, USA: Wavefunction, Inc. **2019**; pp 116-118, 546-551.
[3] S. Liu & L.G. Pedersen, *J. Phys. Chem. A.* **2009**, *113*, 3648.

Chapter 4 Application of Conformational Analyses & HOMO Map for Stereoselectivity Prediction

Guqin Shi, Qiuyue Wang, Dong Pan, Tommy Lai, John S. Wai

In Chapters 1 & 2, we discuss how LUMO and HOMO analyses are applied for predicting regioselectivity on heterocyclic aromatic substrates. In this chapter we will be looking at cases in which saturated ring systems are involved.

For the piperidine substrate **1** in Figure 1, deprotonation of the α-carbon of the ester group followed by alkylation results in a quaternary chiral center in the product. Would the reaction be stereoselective?

1

Figure 1. Alkylation of (2*R*, 4*S*)-dimethyl 2-methylpiperidine-1,4-dicarboxylate (**1**)

A conformational analysis is often needed in such cases. Substituents on the saturated ring system could potentially exert substantial steric effects, biasing for one stereoisomer or the other.

In this article, we will present how we combine conformational analyses along with HOMO map to provide a solid and straightforward way for stereoselectivity prediction.

Identify Substrate Dominant Conformation

The "Conformer Distribution" task in Spartan is programmed to generate all reasonable conformations for the target and identify the most dominant one with the lowest energy. Figure 2-a shows the preferred conformation of the piperidine substrate **1**. Both 1-*N*-methyl carbamate and 4-methyl ester groups are at equatorial positions as we generally recognize, whereas the 2-methyl group adopts an axial position.

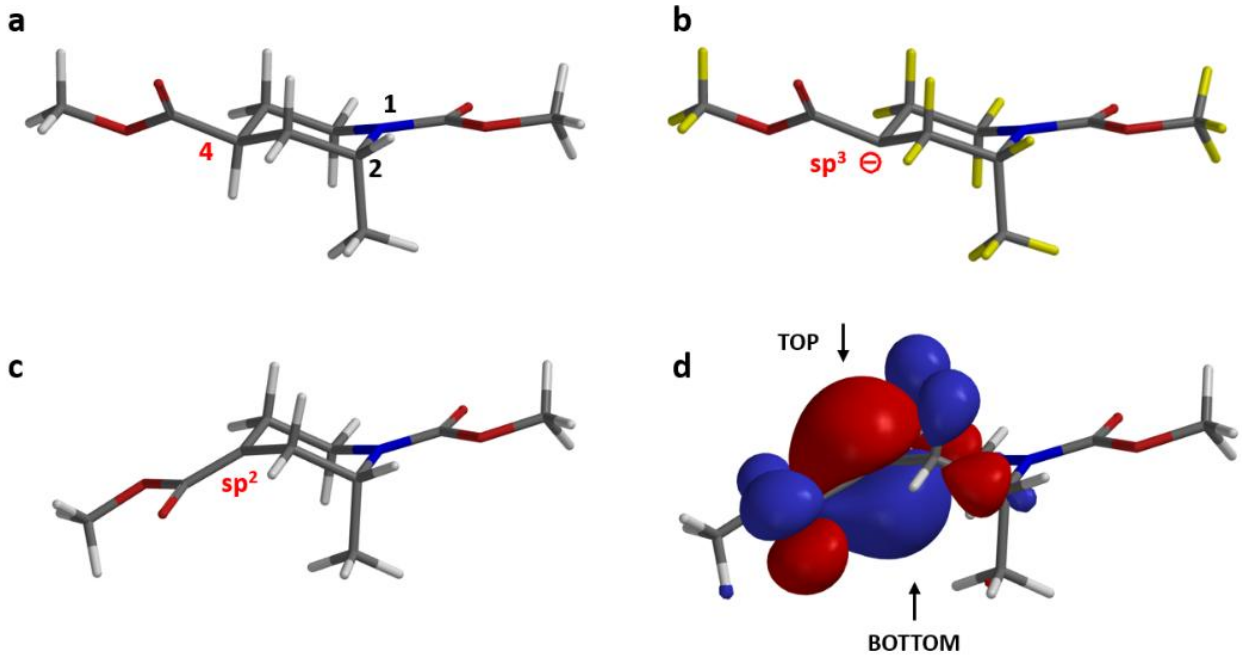


Figure 2. a. Dominant conformation of substrate **1**. b. Substrate **1** is deprotonated at the α-carbon of methyl ester yielding a carbanion intermediate c. QM-minimized carbanion intermediate adopts the enolate form d. HOMO of the enolate intermediate

Calculate Intermediate Equilibrium Conformation

The reaction proceeds in two steps. LDA deprotonates the α-carbon of the ester group, generating a carbanion intermediate, followed by alkylation with methyl iodide. Here we remove the hydrogen atom from the α-carbon at 4-position (Figure 2-b) to generate the chemical structure for QM calculations.

The QM-minimized carbanion intermediate structure is shown in Figure 2-c. The power of QM calculation is that it is capable of recognizing the electronic environment surrounding the carbanion and successfully correcting the hybridization status of the atom. The deprotonated α-carbon is sp^2 hybridized, and the resultant enolate adopts a co-planar conformation.

Evaluate HOMO Lobes Accessibility

Figure 2-d shows the HOMO of the QM-minimized enolate intermediate. Clearly the largest HOMO lobes are on top and bottom of the plane of the enolate anion. It is anticipated that the 2-methyl group could limit access of the HOMO lobe from the bottom side through steric hindrance (1,3-diaxial interactions, a specific class of *syn*-pentane interactions). The question is: will it be completely or partially blocked?

Here, we would like to introduce the use of HOMO map, along with “Inaccessible Markers”, a surface property provided in Spartan, to assist the analysis.

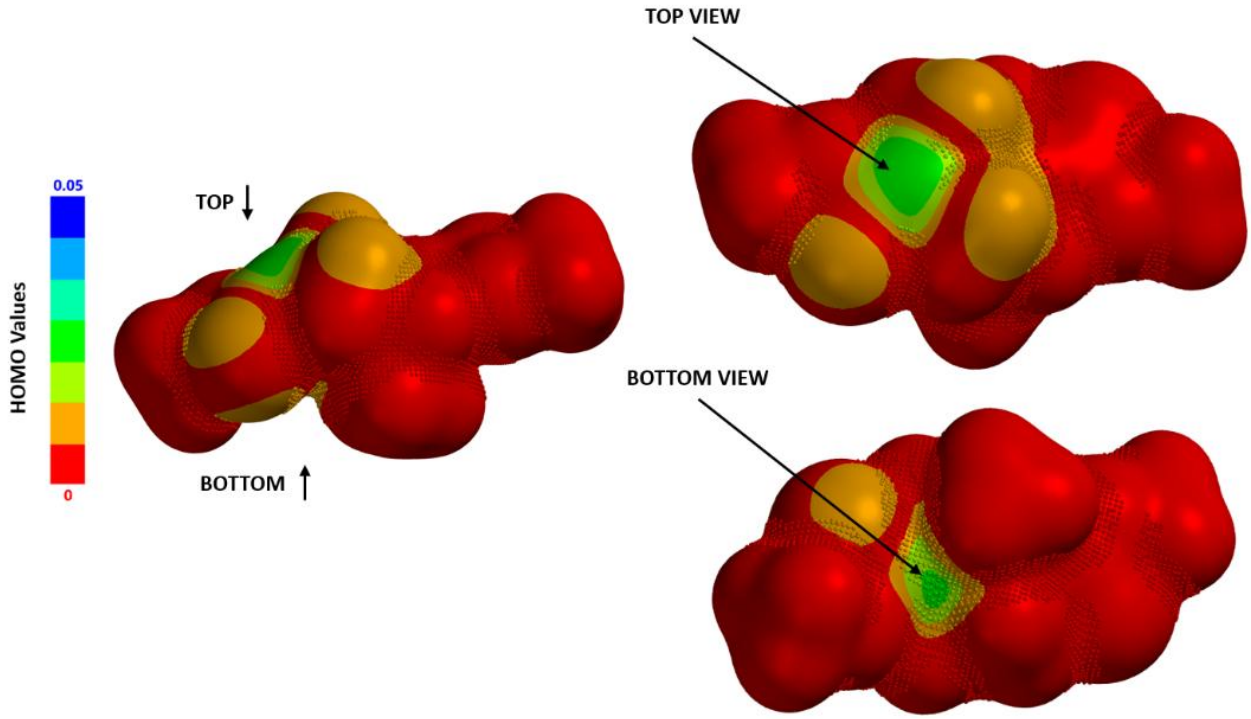


Figure 3. HOMO map of the enolate intermediate

Note: HOMO map is an overlay of the absolute value of the HOMO on electron density surface

Inaccessible Markers is a display indicator for inaccessible surface area

By mapping HOMO onto electron density surface, we could generate the HOMO map of the enolate intermediate, shown in Figure 3. Color near red indicates a low HOMO value with a low probability being attacked by electrophiles, whereas color near green-to-blue indicates higher HOMO value, suggesting the HOMO lobe is available to interact with electrophilic reagent. Spartan also provides a useful tool, "Inaccessible Markers", represented as dots distributed on surface, to distinguish regions that are not accessible.

Although HOMO values on top and bottom of the enolate anion are identical, as they are both in green colors, the bottom side is completely covered with inaccessible markers, which explicitly confirms what we anticipated: 2-methyl group brings in substantial 1,3-diaxial interactions, completely blocking the bottom side from being attacked. The methyl iodide can only approach from the top side, leading to 4-*R* stereoisomer as the sole product observed (Figure 4).

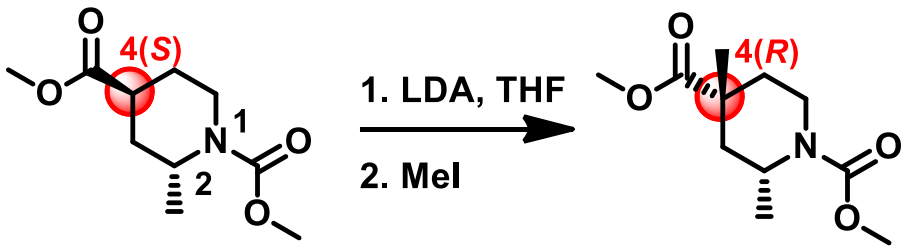


Figure 4. Alkylation of substrate **1** gives selectivity 4(*R*) isomer

To summarize, for reactions involving saturated ring systems, we could use modeling tools to identify the dominant low-energy conformation. Subsequent QM calculations could further

provide more accurate structural information along with frontier orbital distributions to locate where the reaction might occur. By taking advantage of molecular surface properties, we can then visually and objectively evaluate the steric effects and predict the stereoselectivity with high confidence.

Building on What We Just Learned

Substrate **2** is a bicyclic trans-decalin with multiple substituents. Will deprotonation of carbon α to the nitrile group followed by methylation provide the required quaternary chiral center (Figure 5-a)?

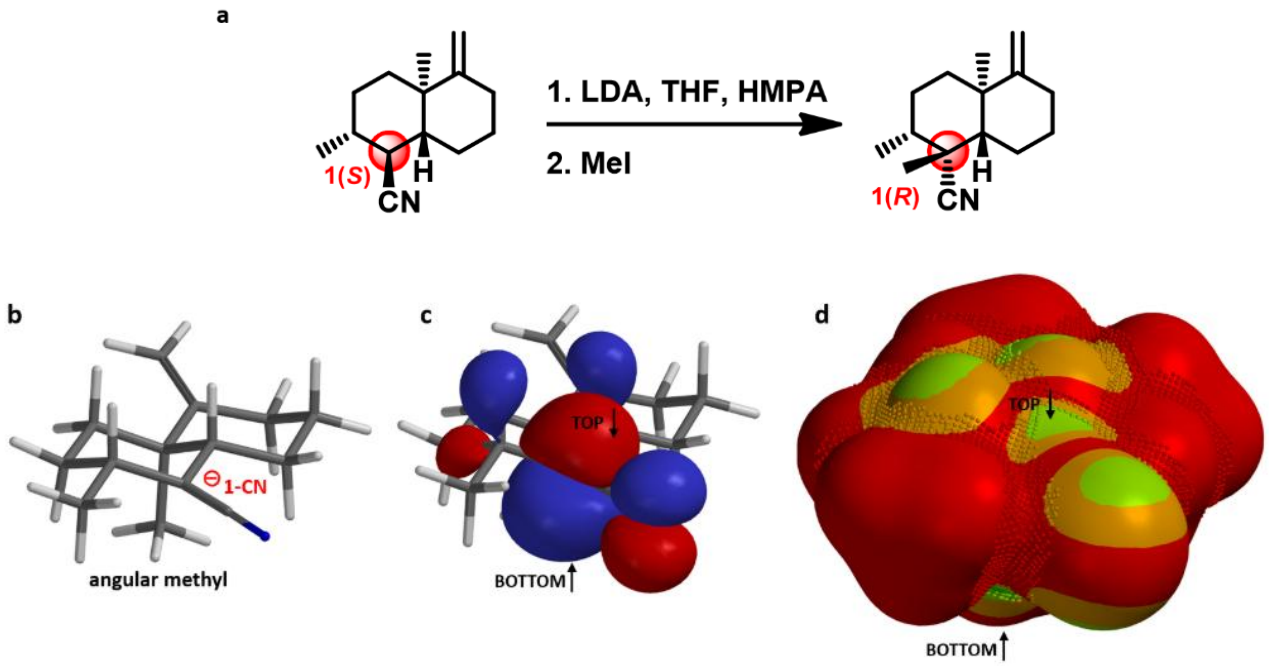


Figure 5. a. Alkylation reaction of substrate **2**. b, c, and d. QM-minimized carbanion intermediate, the corresponding HOMO lobes, and the HOMO map of substrate (2), respectively.

QM energy minimized carbanion intermediate structure (Figure 5-b) suggests a close-to-planar conformation at reaction center, in conjugation with the nitrile group. HOMO lobes on top and bottom of carbanion are in similar sizes (Figure 5-c). However, HOMO map (Figure 5-d) reveals that the bottom side is completely covered with inaccessible markers, with access totally blocked by the angular methyl group. Experimentally only the 1-*R* methylated product is produced.

Return to Table of Contents

References:

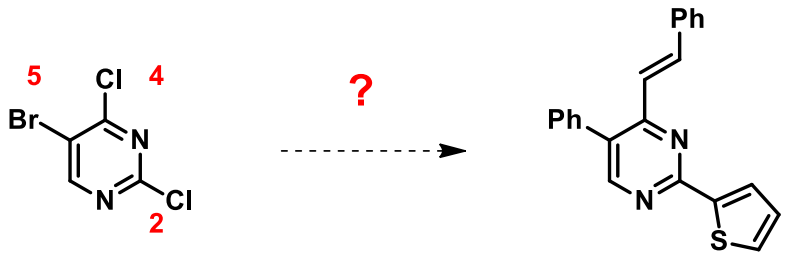
[1] W. Hehre and S. Ohlinger, *A Guide to Molecular Mechanics and Quantum Chemical Calculations.* Irvine, CA, USA: Wavefunction, Inc., **2003**.
[2] *Spartan'18 Tutorial and User's Guide.* Irvine, CA, USA: Wavefunction, Inc., **2019**.
[3] R.W .Hoffmann, *Angew. Chem. Int. Ed. Engl.*, **1992**, *31*, 1124.

Chapter 5 Sequential Cross Coupling Reactions of Polyhalogenated Heterocycles

Qiuyue Wang, Guqin Shi, Zhong Zheng, Jian Wang, John S. Wai

In the previous QM chapters, we learned the concepts of LUMO and LUMO map, and their application in analyses of organic chemical reactions. In this chapter, we will apply these concepts on prospective analyses of a sequence of Cross Coupling Reactions, starting with a polyhalogenated pyrimidine.

Synthesis of Target Molecule (4) from Polyhalogenated Pyrimidine (1)



(1) (4)

For a polyhalogenated benzene ring, reactivity towards oxidative addition usually follows the order of I > Br > Cl. With polyhalogenated heterocyclic aromatic systems, experience taught us that it may not be that straightforward.

For 5-bromo-2,4-dichloropyrimidine **(1)**, halogenated at position 2-, 4-, and 5-, which position is most susceptible towards oxidative addition? In palladium catalyzed cross-coupling reactions, we could consider Pd[0] as a nucleophile, as such, we first calculate for LUMO of **1** for prospective analysis.

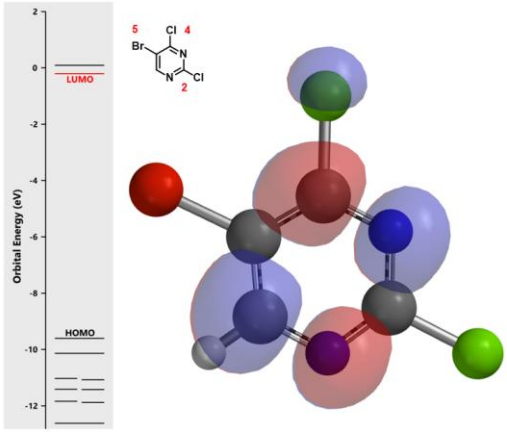
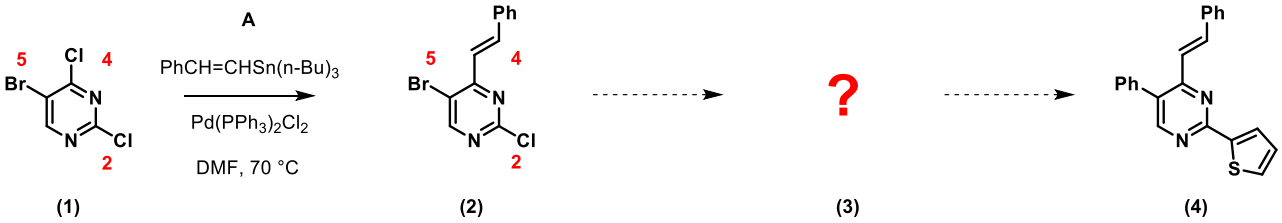


Figure 1. LUMO and associated energy of 5-bromo-2,4-dichloropyrimidine (**1**)

Calculated LUMO of starting material **1** reveals only the carbon at position 4, connecting to a chloro group, has a significant LUMO lobe associated with it (Figure 1). As such, under well controlled conditions, with reaction temperature maintained at 70 °C, pyrimidine **1** underwent selective cross coupling reaction with tin reagent **A** to provide C-4 vinyl intermediate 2 selectively.[2]



For the next step, we calculate again the LUMO of intermediate **2** for analysis (Figure 2):

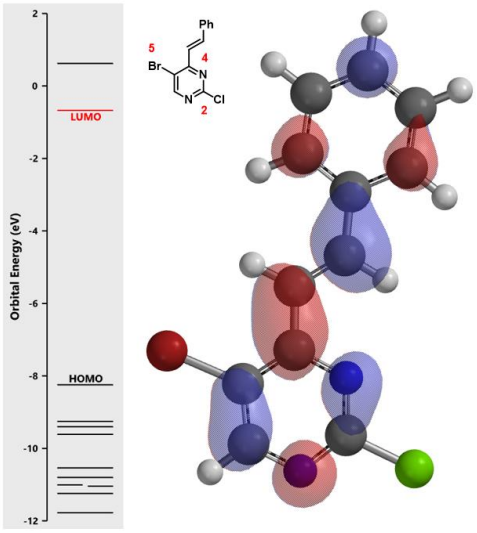


Figure 2. LUMO and associated energy of intermediate **2**

There are no obvious LUMO lobes on either C-2 or C-5 with a chloro and bromo substituents, respectively. In such situation, we learned to inspect the second lowest unoccupied molecular orbital, LUMO+1 (Figure 3).

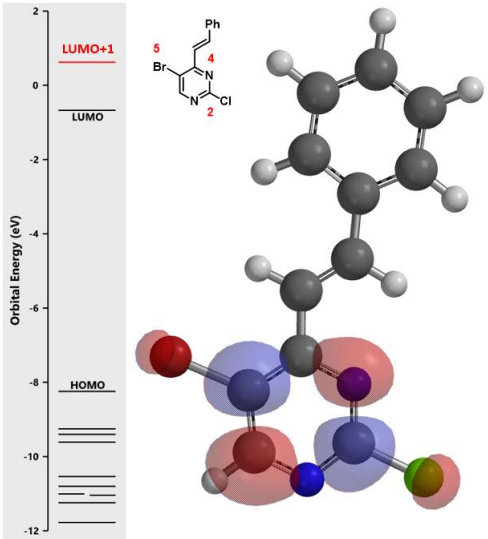


Figure 3. LUMO+1 and associated energy of intermediate **2**

LUMO+1 of pyrimidine intermediate **2** shows LUMO lobes at C-2 and C-5 positions, but no significant difference in their relative sizes. Does that mean there is no intrinsic bias for the next oxidative addition? We know the answer is no. Then what will be the next trick up our sleeves?

Application of QM Calculated IR Frequency

Here we would like to introduce the use of QM calculated C-X bond infrared stretching vibrational frequencies to differentiate their relative reactivity toward oxidative addition. The weaker the C-X bond is, the lower the wave number it has [3], and the easier it is for the bond to undergo reaction.

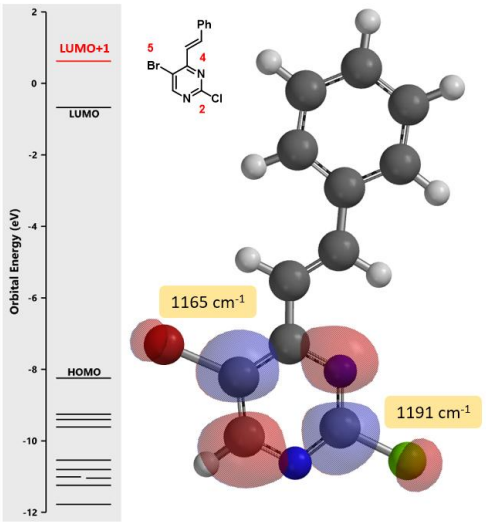
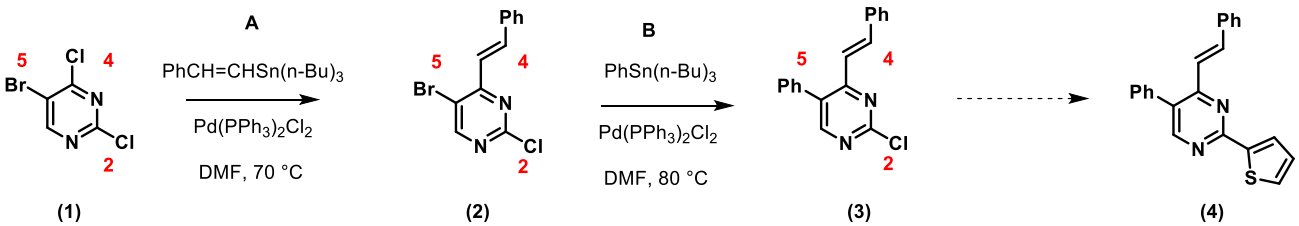


Figure 4. LUMO+1 and IR stretching frequencies of intermediate **2**

QM calculation results indicated that the C-Br bond (1165 cm^{-1}) is weaker than the C-Cl bond (1191 cm^{-1}) (Figure 4), and therefore the oxidative addition reaction is more likely to occur, consistent with experimental observation. It is worth noting that the lowest unoccupied orbital energy of starting material **1** is -0.21 eV, and the energy of LUMO+1 of intermediate **2** is 0.62 eV. This suggests that we need to raise the reaction temperature for the second cross coupling reaction to proceed. Indeed, reaction of compound **2** with tin reagent **B** proceeded smoothly at 80 °C, with the same $Cl_2Pd(PPh_3)_2$ catalyst as in previous step, to provide the phenylated intermediate **3** as the single product.



In the last step, intermediate **3** does not have LUMO lobe on C-2, yet a very clear one with the higher energy LUMO+1.

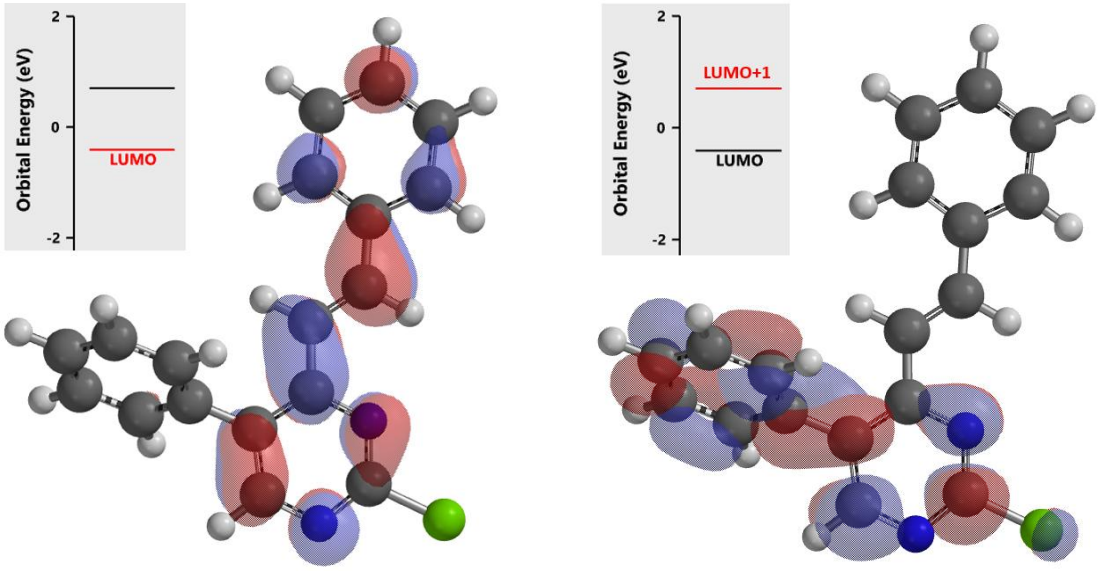
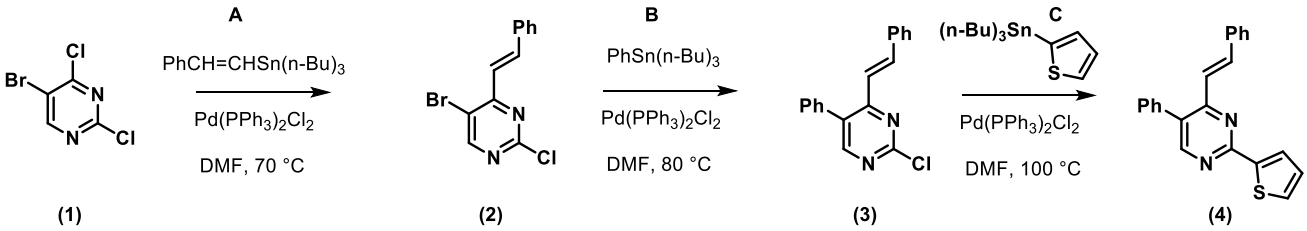


Figure 5. Schematic of LUMO (left) and LUMO+1 (right) for compound **3**

With intermediate **3** (Figure 5), since its LUMO+1 orbital energy (0.70 eV) is higher than that of intermediate **2** (0.62 eV) the reaction temperature was raised further to 100 °C for the final cross coupling reaction to proceed.



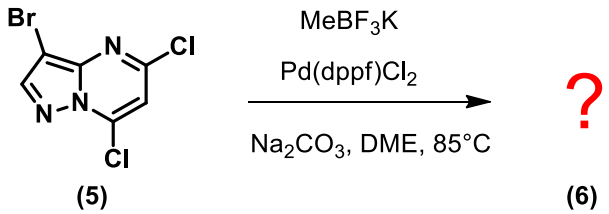
With the above sequence of Cross Coupling reactions, using the same catalyst and solvent, one could envision running the whole sequence in one pot by adding corresponding tin reagents at different stages and raising the reaction temperature accordingly.

QM has proven to be a very powerful tool for organic chemists to analyze their synthetic schemes prospectively, and to evaluate objectively each of the reactions in the sequence.

In summary, for sequential Cross Coupling reactions of polyhalogenated heterocycles, we first look at relative sizes of LUMO or LUMO+1 lobes on the carbons with halogen connecting to them. If the relative sizes of the lobes are similar, we compare the calculated C-X bond infrared stretching vibrational frequencies. These parameters enable us to sequence the order of reactions in a highly predictable manner.

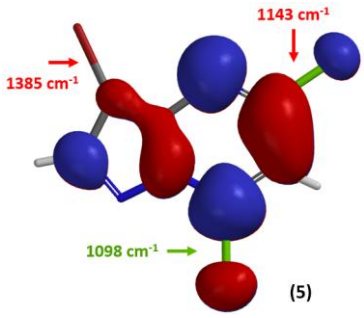
Building on What We Just Learned

Let's apply what we learned above in 3-bromo-5,7-dichloropyrazolo[1,5-a]pyrimidine **5** shown below.

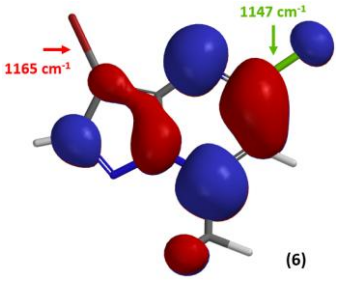


(5) (6)

Based on the LUMO and calculated C-X bond stretching info, where will the first oxidative addition occur?



Yes, on the C-7 carbon – chloride bond. How about the next oxidative addition?



Return to Table of Contents

References:

[1] W. Hehre and S. Ohlinger, *A Guide to Molecular Mechanics and Quantum Chemical Calculations.* Irvine, CA, USA: Wavefunction, Inc., **2003**.

[2] J.A. Joule & K. Mills, *Heterocyclic Chemistry 4th Ed.* Malden, MA, USA: Blackwell Publishing Ltd., **2000**; pp 202-203.

[3] R.M. Badger, *J. Chem. Phys.* **1934**, *2*, 128.

Chapter 6 Stepwise Multiple Halogenation & Coupling Reaction Strategies

Dong Pan, Guqin Shi, Qiuyue Wang, Tommy Lai, John S. Wai

In Chapter 5, we introduced "Sequential Cross Coupling Reactions of Polyhalogenated Heterocycles" to design synthetic routes. In this article, we will discuss the use of a complimentary strategy - Stepwise Multiple Halogenation and Coupling Reactions!

Synthesis of Polysubstituted Aromatic Compounds

Polysubstituted aromatic targets are very common targets in organic synthesis. In general, a functional group can be introduced by halogenation and then metal-catalyzed cross-coupling reaction. The resultant intermediates could be treated with a few cycles of halogenation and coupling reaction to provide the final products. However, regioselectivity of halogenation on complex aromatic substrates is sensitive to the nature and position of the array of their substituents, making it difficult to predict their outcomes accurately, and rendering these syntheses “trial and error” processes. In our laboratories, incorporation of QM analyses in planning of these synthetic sequences has proven to be invaluable.

Retrosynthetic Analysis

Shown on Figure 1 is the retrosynthetic analysis of the target molecule. Starting with compound **A**, we need three halogenations and two cross-coupling reactions to obtain the highly functionalized target molecules. Will each halogenation reaction occur in the required position?

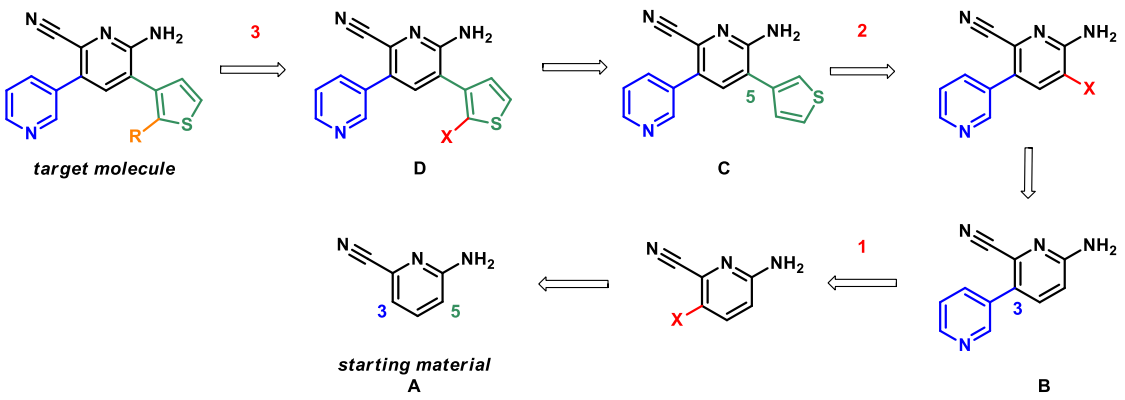


Figure 1. Retrosynthetic analysis

HOMO and ^{13}C NMR Calculation

We discussed in Chapter 2 the use of HOMO and calculated ^{13}C NMR to predict the selectivity of electrophilic aromatic substitutions[1]. These parameters are calculated for starting material **A**, intermediates **B** and **C** to assess their chemical properties and feasibility of the proposed synthetic sequence.

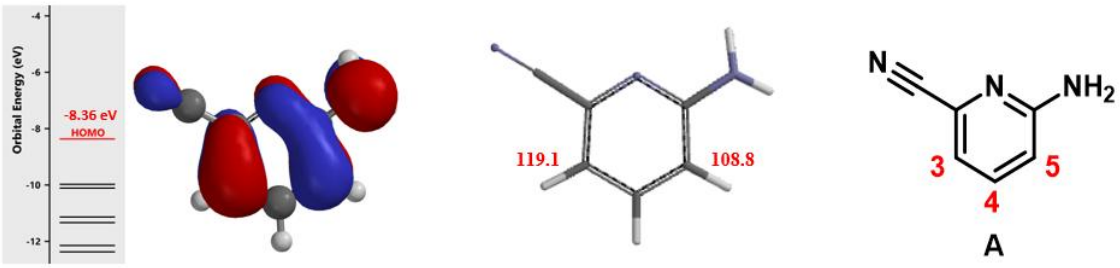


Figure 2. HOMO, associated energy and carbon spectrum of starting material **A**

First, how do the cyano and amine groups affect regioselectivity in bromination of pyridine **A**? With HOMO (Figure 2, left), the 4-position has almost no HOMO lobe distribution, while the HOMO lobes at C-3 and C-5 are similar in size. Calculated ^{13}C NMR chemical shift values are 119.1 ppm at C-3 and 108.8 ppm at C-5, (Figure 2, middle), suggesting that C-5 will be preferentially halogenated, different from what we need. To overcome this intrinsic preference, we planned to use n-Bu_4NBr_3, which is known to selectively brominate *para* to amino groups, for reagent control (see chapter 23)[2]. Being the first step of the synthetic sequence, this level of uncertainty is acceptable.

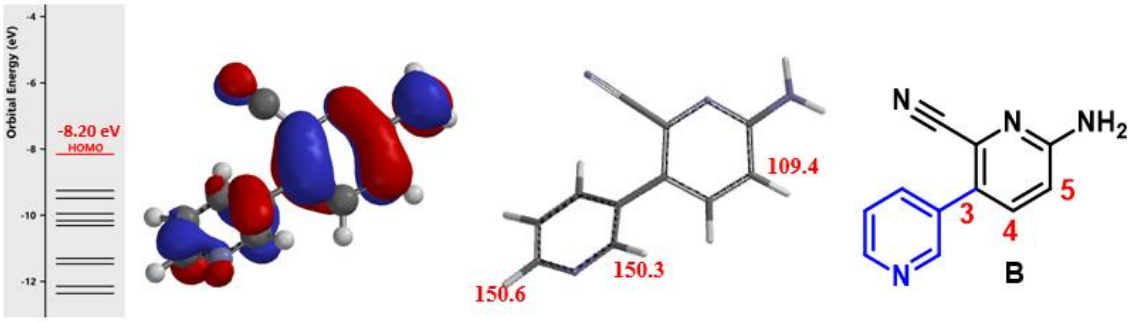


Figure 3. HOMO, associated energy and carbon spectrum of intermediate **B**

Intermediate **B** contains two pyridine rings, which one is more nucleophilic? The same QM analysis was carried out. HOMO lobe at the C-5 of the pyridine a ring is the largest among all the substitutable carbons (Figure 3, left), and its calculated chemical shift is 109.4 ppm, lower than those calculated for the other two carbons on the pyridine b ring (Figure 3, middle). These suggest that halogenation of **B** will occur selectively at C-5.

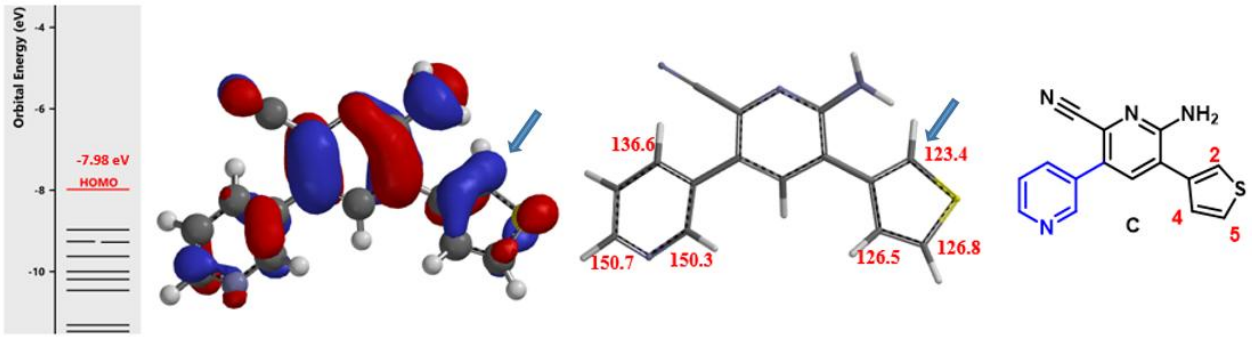


Figure 4. HOMO, associated energy, and carbon spectrum of intermediate **C**

For the third halogenation, an increasingly more complicate intermediate **C**, QM calculations show that C-2 of thiophene has the largest HOMO lobe (Figure 4, left) and a chemical shift value of 123.4 ppm, the lowest among all substitutable carbons with significant HOMO lobes (Figure 4, middle). Together they enable us to predict with confidence: highly selective C-2 halogenation of the thiophene ring.

Next is to find out how well this sequence works in the labs.

Experimental Results Align with QM Calculations

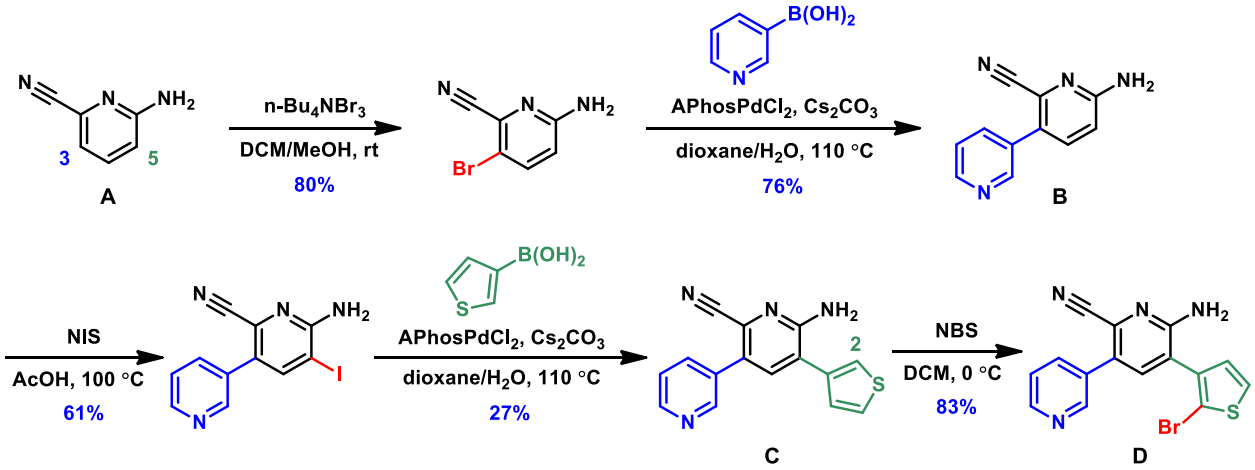


Figure 5. Stepwise multiple halogenation and coupling synthesis of halogenated tetrasubstituted pyridine compounds

The experimental results of this synthetic route are shown in Figure 5. Reaction of the starting material **A** with n-Bu_4NBr_3 indeed allowed us to overcome the intrinsic preference for reaction at C-5, and provided us selectively the required C-3 *para* bromination product[2]. It is remarkable that QM analysis alerted us to seek for a *para* selective reagent for the reaction.

The following iodination of intermediate **B** with NIS and bromination of intermediate **C** with NBS proceeded selectively as calculated QM data suggested. Stitching these cycles of halogenation and cross coupling reactions in the proper order provided the key bromo intermediate **D** for further elaborations.

In summary, this chapter introduces the use of stepwise multiple halogenation and coupling strategies to design synthetic routes for target molecules. Success of this strategy hinges on

generating the key mono-halo intermediate selectively in each cycle. While in Chapter 5, we obtained the target through the selective C-X coupling of multi-halogenated substrates, requiring us to know in each reactions which C-X bond is more reactive toward oxidative addition.

Our examples highlighted the use of QM to understand and utilize the inherent chemical properties of starting materials, intermediates, and reagents, to evaluate feasibilities and risks associated with diverse synthetic strategies conceived, to learn chemistry in a more fundamental manner, at a deeper level.

Return to Table of Contents

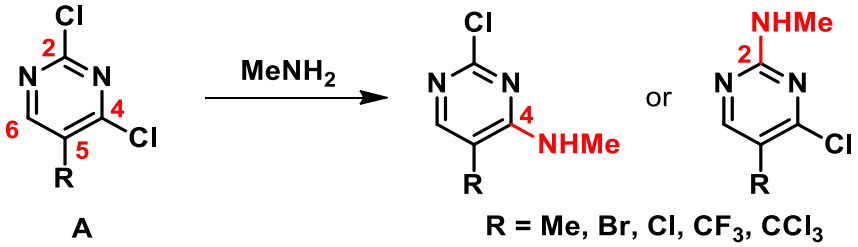
References:

[1] *Spartan'18 Tutorial and User's Guide*. Irvine, CA, USA: Wavefunction, Inc., **2019**
[2] n-Bu_4NBr_3: a) I. Macsári, K. Paulsen, L. Rakos, *Compounds and Their Use for Treatment of Amyloid Beta-Related Diseases* [P]: US, 2012/122843, **2012**, A1, 50. b) J. Berthelot, C. Guette, M. Essayegh, P.L. Desbene, J.J. Basselier, *Synth. Commun.* **1986**, *16*, 1641. c) Z. Chen & K. Zhou, *University Chemistry* **1995**, *10*, 10. https://doi.org/10.3866/PKU.DXHX19950403.

Chapter 7 Application of LUMO Analysis in Nucleophilic Reactions (Part II)

Zhong Zheng, Guqin Shi, Dong Pan, Qiuyue Wang, John S. Wai

Our introductory chapter (Chapter 1) on LUMO leaves us with an intriguing question, that is, when the 5-CH_3 group in dichloropyrimidine **A** is replaced with Cl, Br, CF_3, and CCl_3, which chloro group will be displaced by an aliphatic amine? In this chapter, we will discuss the orbital analyses of these pyrimidines.



Reaction Site	5-Me	5-Br	5-Cl	5-CF_3	5-CCl_3
2	✘	✘	✘	✓	✓
4	✓	✓	✓	✓	✘

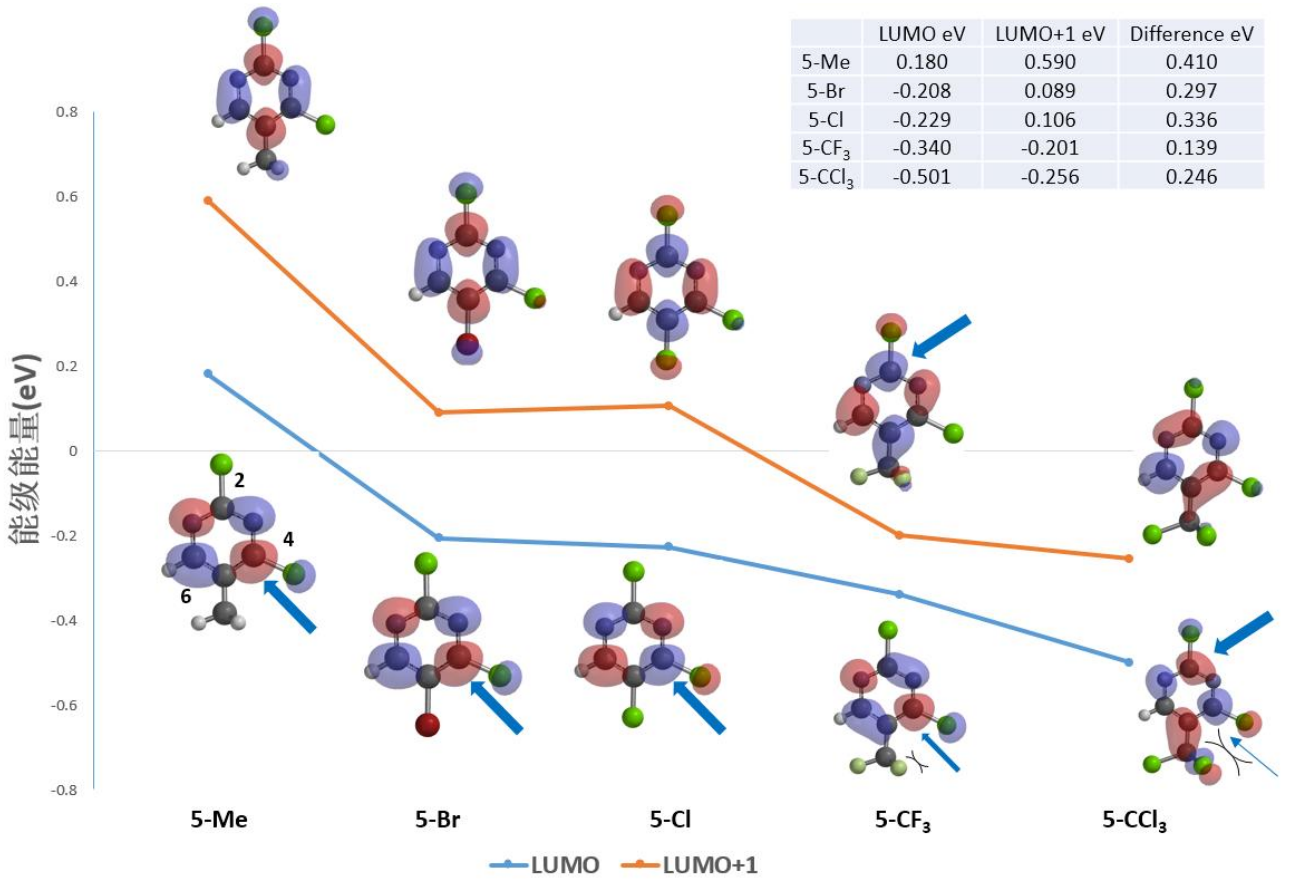
Table 1. Nucleophilic substitution of substrate **A** with methylamine

LUMO and LUMO+1 in Nucleophilic Substitutions

First, we would like to introduce LUMO+1, which is the second Lowest Unoccupied Molecular Orbital, one level higher in energy than LUMO. Its involvement in reaction could become significant when energy gap between LUMO and LUMO+1 is small.

Shown on Figure 1 are LUMO, LUMO+1 orbitals and their corresponding energy levels of five dichloropyrimidine substrates. In the 5-CH_3 substituent, LUMO lobes are located at C-4 and C-6 positions, accounting for their interactions with incoming nucleophilic reagents discussed in Chapter 1. Addition of methylamine at C-6 is reversible, whereas addition at C-4 is irreversible, leading to selective displacement of the chloro group observed. Energy gap between LUMO and LUMO+1 for this substrate is big enough that LUMO lobe distribution is enough to account for the observation.

Similar orbital and energy patterns can be observed for the 5-Br and 5-Cl substituents, accounting for the C-4 selective displacement observed. No surprise.



	LUMO eV	LUMO+1 eV	Difference eV
5-Me	0.180	0.590	0.410
5-Br	-0.208	0.089	0.297
5-Cl	-0.229	0.106	0.336
5-CF_3	-0.340	-0.201	0.139
5-CCl_3	-0.501	-0.256	0.246

Figure 1. LUMO and LUMO+1 orbitals and corresponding energy levels of substrate **A** with various 5-substitutents

When it comes to the 5-CF_3 substituent, the distribution of LUMO lobes has changed. Now the C-2 position is partially covered with LUMO lobe, enabling the substrate to potentially interact with incoming nucleophiles at this position. In addition, the energy levels between LUMO and LUMO+1 have become closer together, and the 5-CF_3 group imposes higher degree of steric hindrance to incoming nucleophile, accounting for a 1:1 mixture of C-2 and C-4 substitution products observed.

With the 5-CCl_3 dichloro-pyrimidine, the energy gap between LUMO and LUMO+1 is back to the normal range, we could focus our analysis on LUMO. This substrate has LUMO lobes on both C-2 and C-4, yet 5-CCl_3 is substantially larger in size than 5-CF_3, limiting accessibility to reaction at C-4, and leading to selective displacement at C-2.

Plausible Reaction Mechanism Based on LUMO Analysis

5,6-Pyrimidyne was postulated as intermediate in the reaction of 4-methyl-5-bromopyrimidine (**B**) with sodamide (Figure 2), starting with dehydrobromination, followed by deprotonation of the methyl group, and then amide addition to the resultant benzyne anion exclusively at C-6.[1]

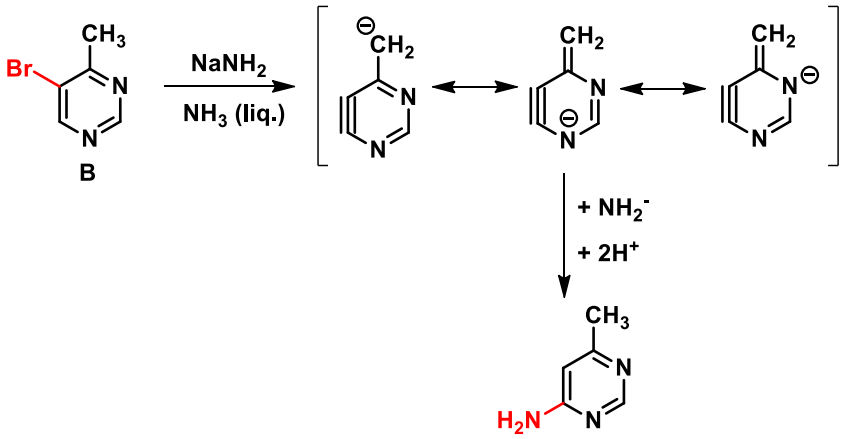


Figure 2. Amination of compound **B**: formation of a benzyne-like intermediate

We reasoned that it is not likely to have two anions reacting with each other and performed a QM analysis of the bromopyrimidine **B**.

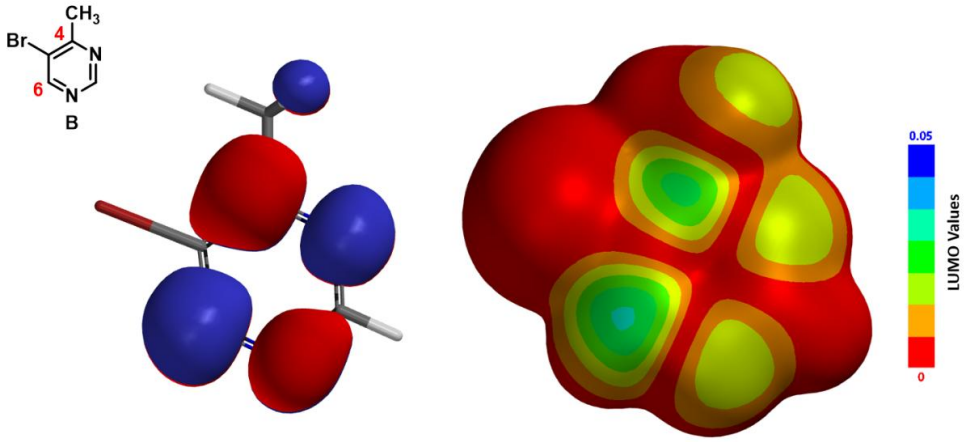


Figure 3. LUMO and LUMO map of compound **B**

Shown on Figure 3 are the calculated LUMO and LUMO map for bromopyrimidine **B**. There is little LUMO lobe associated with C-5, explaining why a S_NAr substitution of the bromide with an amino group at C-5 was not observed. Both C-4 and C-6 have sizable lobes, yet LUMO map indicates that C-6 is more accessible to nucleophilic attack (light blue). We reasoned that an amide anion is first added at C-6, leading to an intermediate anion shown in Figure 4, followed by an irreversible elimination of hydrogen bromide to generate 4-amino-6-methylpyrimidine.

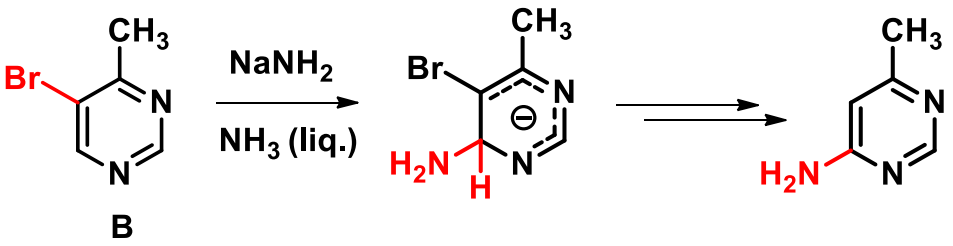
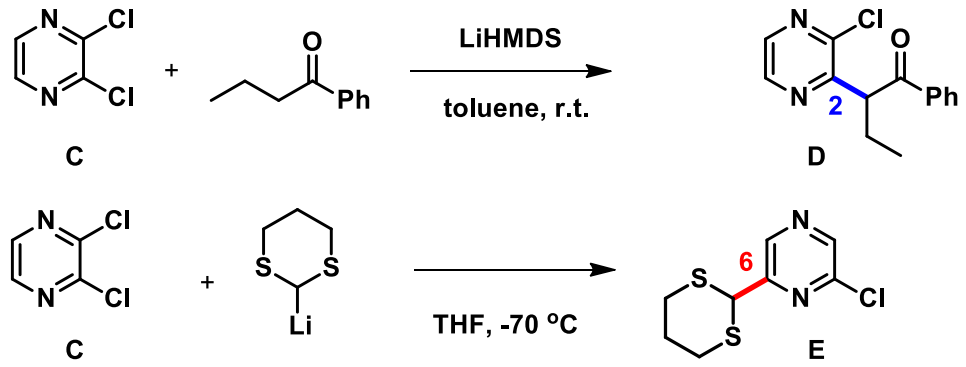


Figure 4. Amination of compound **B**: formation of an amino anion intermediate

To summarize, for nucleophilic substitution reactions on complex heterocycles, LUMO and LUMO map have been very useful for us to rationalize and predict reaction outcomes. When the energy gap between LUMO and LUMO+1 is small, it is important to consider both orbitals.

Inclusion of steric consideration in such analysis provides further insight to the reaction of interest.

Building on What We Just Learned



Reaction of compound **C** with lithium enolate of phenylbutanone is known to provide C-2 substituted product **D**. However, in 2006 McDonald *et al*.[2] reported that treatment of **C** with lithio dithiane afforded a *tele*-substitution product **E**, that is, addition of dithiane moiety at C-6 followed by elimination of the chloro group across the ring at C-3. Shown below in Figure 5 are the LUMO/LUMO+1 and their corresponding maps calculated for compound **C**. We hope you will have as much fun as we did in solving this decade old chemistry puzzle.

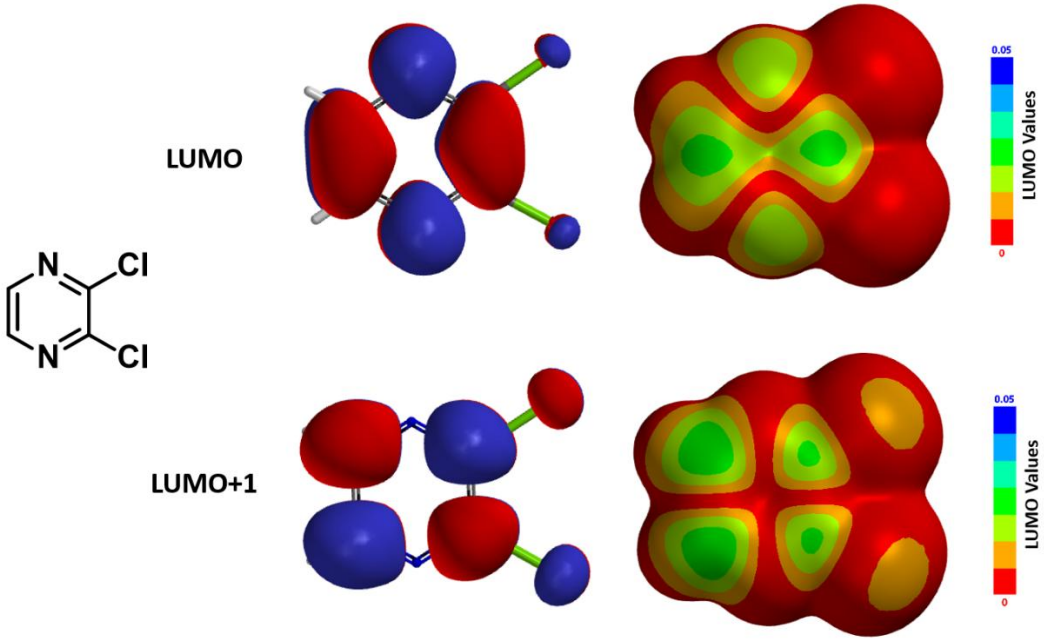


Figure 5. LUMO, LUMO+1 lobe and corresponding LUMO map, LUMO+1 map of compound **C**

Return to Table of Contents

References:

[1] J.A. Joule & G.F. Smith, *Heterocyclic Chemistry* 2nd Edition. Wokingham, Berkshire, England: Van Nostrand Reinhold Company Ltd., **1972**; pp 131.

[2] J.E. Torr, J.M. Large, P.N. Horton, M.B. Hursthouse, E. McDonald, *Tetrahedron Lett.* **2006**, *47*, 31.

Chapter 8 Standing on the Shoulders of Giants: Hammond's Postulate

Guqin Shi, Shouliang Wang, Qiuyue Wang, John S. Wai

In previous chapters, we have introduced the application of HOMO/LUMO analysis of heterocycle substrates to rationalize reaction outcome, make quick prediction, and facilitate synthetic planning.

In our laboratory practice, we found that there are occasions in which further calculations for reaction energy profile and transition state are necessary to provide quantitative estimate for selectivity in reaction, and to gain deeper insights on reaction mechanisms.

Today, we would like to review three key theories which laid the foundation of reaction energy calculations.

1889 • Arrhenius Equation

How can a chemical reaction take place? What controls the reaction rate and determines the product selectivity? In 1889, Svante Arrhenius first proposed an empirical formula that summarizes the relationship between reaction rate and temperature, which is now well-recognized as Arrhenius Equation (Figure 1).

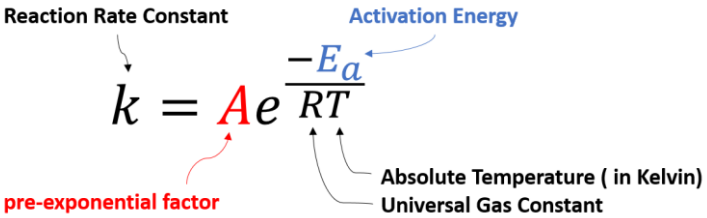


Figure 1. Empirical Arrhenius equation

It states that for reactants to transform into products, they must first **acquire a minimum amount of energy, activation energy E*a*.** When a reaction has a rate constant that obeys Arrhenius Equation, a plot of *ln* k versus T^{-1} gives a straight line (Arrhenius plot). By measuring a set of reaction rates under different temperatures, one can determine the activation energy E*a* and pre-exponential factor A from the slope and the intercept, respectively.

1935 • Transition State Theory

In 1935, benefiting from significant development of statistical thermodynamics, Henry Eyring *et al.* proposed the Transition State Theory, along with the Eyring Equation (Figure 2), provided a more sophisticated formula, and addressed the physical interpretations of two parameters in Arrhenius Law:

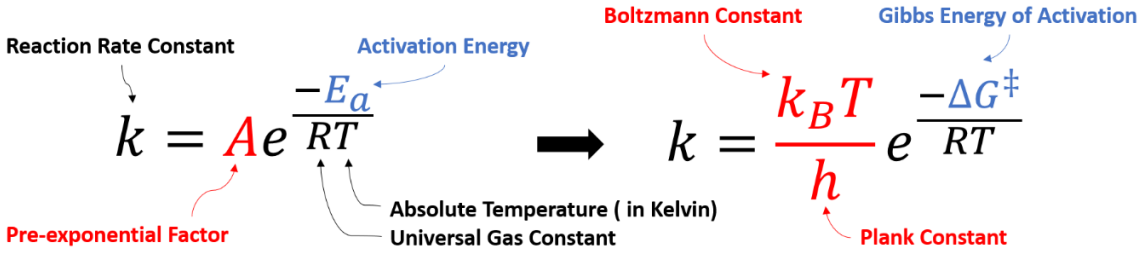


Figure 2. Arrhenius equation versus Eyring equation

1. **Activation Energy**: Eyring *et al*. argued that the reaction does not occur by simple collisions, instead, it requires the formation of an activated transition state complex from reactant molecules. **The energy required for reactant molecules to form the transition state complexes is the activation energy**. The Gibbs Free Energy of Activation, ΔG‡, was introduced to account for both enthalpy and entropy changes. For multi-step complex reactions, each transition state should have a corresponding ΔG‡.

2. **Pre-exponential Factor**: It is a function of temperature, representing the frequency with which the activated transition state complexes are converted into products.

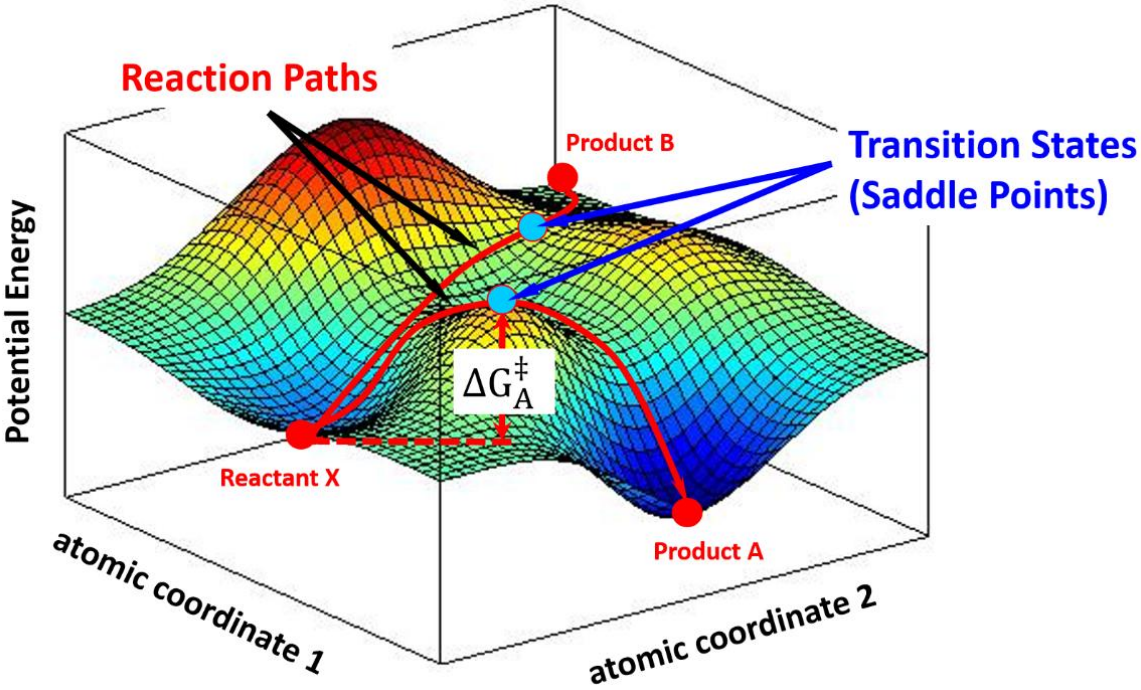


Figure 3. Different reaction paths over a potential energy surface

In the progress of forming a transition state complex, the potential energy of the entire system changes correspondingly. A chemical reaction could then be described as a moving point over the potential energy surface with coordinates in atomic momenta and distances. Transition State Theory states that the saddle point, i.e. the minimax point, on a potential energy surface, is the transition state of the reaction. For example, as shown in Figure 3, a stable reactant X sits in a local minimum of the potential energy surface. By following different energy paths, distinct transition state complexes are formed at saddle points, leading to products **A** or **B**.

In principle, by constructing a potential energy surface from microscopic quantities, one can calculate the absolute reaction rate constants for the reaction; although in practice, it is never easy to obtain the precise knowledge of a potential energy surface. Nevertheless, the concept of "activated transition state complex" has been well accepted.

1955 • Hammond's Postulate

In 1955, George Hammond proposed a hypothesis which is of profound significance:

"If two states, as, for example, a transition state and an unstable intermediate, occur consecutively during a reaction process and have nearly the same energy content, their inter-conversion will involve only a small reorganization of the molecular structures."

For a one-step reaction, it could be further interpreted as "the structure of a transition state resembles the side of the reaction that is higher in energy." For an exothermic reaction (Figure 4, left), the transition state is close to the reactant in terms of energy, therefore the structure of the transition state is similar to the reactant. Whereas in an endothermic reaction (Figure 4, right), the transition state structure resembles the product as now they are closer in energy.

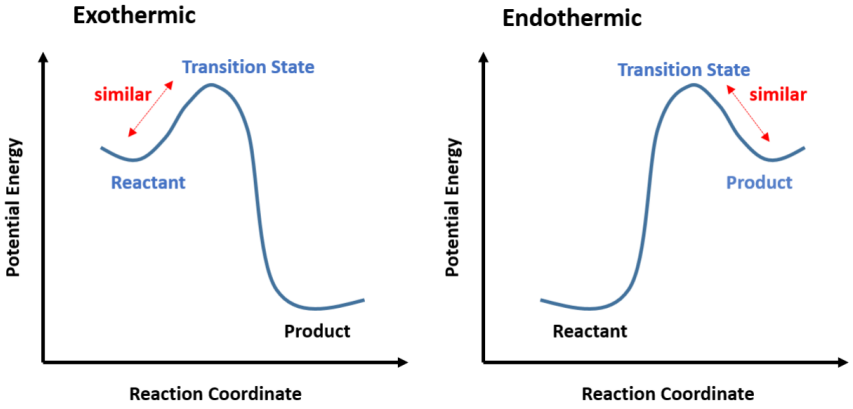


Figure 4. Hammond's postulate interpretation in one-step reactions

Transition states are extremely short-lived, roughly at picosecond level, rendering great experimental challenges to capture or to observe. Thus, the postulate provides a geometric model for the transition state complex, enabling chemists to predict its properties by characterizing the known molecules, and modeling the reaction progress.

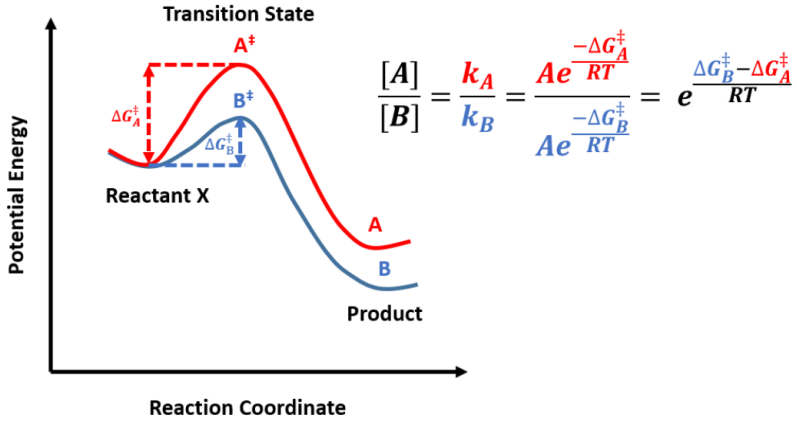


Figure 5. Calculate product ratio from activation energy difference

In kinetic-controlled reactions, product concentration is proportional to reaction rate. The product ratio can be calculated from activation energy difference according to Arrhenius Law, as shown in Figure 5, which is the core equation we will use, for product selectivity predictions in following chapters. Whereas in thermodynamic-controlled reactions, the product ratio is related to the relative energy difference between products.

Standing on the Shoulders of Giants

Transition states often involve bond breaking and formation. While Quantum Mechanics is down to the fundamentals of electron behaviors, not limited by arbitrary bond definitions, making it a perfect tool for transition state modeling.

We appreciate those efforts and dedications contributed by generations of scientists into the development of theories for understanding chemical reactions. We also realize standing on the shoulders of giants enable us to see further.

Combining Frontier Molecular Orbital Theory and Hammond's Postulate, with increasingly faster computing tools, we now could model the transition state complexes at molecular level. These retrospective analyses provide us with deeper mechanistic insight of the reactions and enable us to analyze chemistry prospectively in a more fundamental manner.

Return to Table of Contents

References:

[1] Fu Xiancai *et al*. *Physical Chemistry*. Beijing, P.R. China: Higher Education Press, **2006**.
[2] Y. Wang. *First principles process planning for computer-aided nanomanufacturing* in J.G. Michopoulos, C.J. J. Paredis, D.W. Rose, J.M. Vance (Ed.), *Advances in Computers and Information in Engineering Research, Volume 1*. New York, NY, USA: ASME Press, **2014**.
[3] H. Eyring, *J. Chem. Phys*. **1935**, *3*, 107.
[4] G. S. Hammond, *J. Am. Chem. Soc.* **1955**, *77*, 334.
[5] W.J. Hehre, A.J. Shusterman, W.W. Huang, *A Laboratory Book of Computational Organic Chemistry.* Irvine, CA, USA: Wavefunction, Inc., **1998**.

Chapter 9 Activation Energy Estimation for Alkylation of Pyrazole

Shouliang Wang, Jian Wang, Guqin Shi, Qiuyue Wang, Dong Pan, John S. Wai

In Chapter 8, we discussed the theories leading to the modeling of chemical reactions, calculation of activation energy of reaction, and structure of transition state, etc. In this article, we will discuss how these routine calculations are used to guide our synthetic planning.

Let's start with two practical questions for the following alkylation of pyrazole **1** (Figure 1). Will the alkylation be selective on N-1 or N-2? If it is not selective, will there be a reliable way to estimate the product ratio?

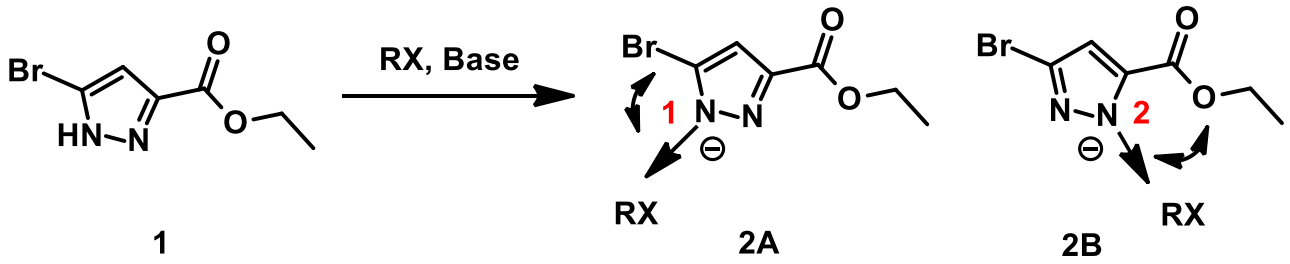


Figure 1. Regioselectivity of N-alkylation on pyrazole **1**

First, we calculate for the Electrostatic potential map (discussed in Chapter 3) of the two tautomeric forms of pyrazole **1** to estimate the relative acidity of the NH group (Figure 2, **A** and **B**). The electrostatic potential of the NH group in tautomer **A** (281 kJ/mol) is found to be slightly higher than that in tautomer **B** (266 kJ/mol). However this difference may not be consequential, since after deprotonation, charge is delocalized within the pyrazole molecule. Indeed, HOMO calculated for the resultant pyrazole anion shows similar lobe size on N-1 and N-2 (Figure 2, **C**). These couldn't help us to differentiate the two potential alkylation pathways.

Reaction Energy Profile

QM calculation enables us to model stepwise the bond forming & breaking processes in a chemical reaction, establish the transition state structures, estimate the associated activation energy, compare the relative difference of competing paths, and calculate for the potential product ratio.

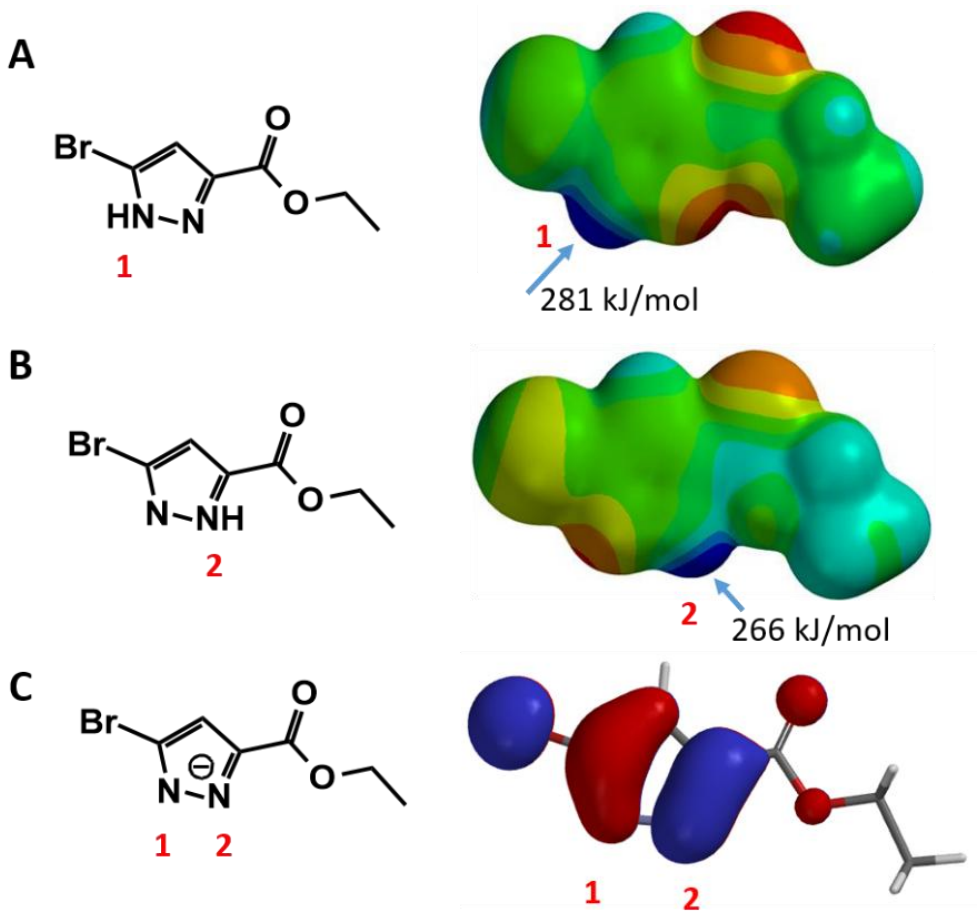


Figure 2. A and B: Electrostatic potential maps of the two tautomeric form of pyrazole **1**. C: HOMO of pyrazole **1** anion

Establish the Reaction Model

For the pyrazole alkylation in consideration, we assume a S_N2 mechanism, with pyrazole N approaching the alkyl halide gradually, forming the C-N bond and breaking the C-X bond in a concerted manner.

We set up two calculations, using CH_3Br as our generic alkylating reagent, whereas one approaches N1 of pyrazole 1, and the other interacts with N2, starting with a C-N distance of 3.8 Å, and calculate for the system energy with a step size of 0.1 Å (Figure 3).

Energy of the system starts to rise when the distance between C-N atoms is shortened to 3.0 Å, indicating that the reactant molecules start to interact. When the bond length is shortened to about 2.1 Å, the energy level reaches the peak and reaction reaches the transition state. The energy difference between these two points is our estimated activation energy.

Estimation of Product Ratio

Activation energy calculated for N-1 and N-2 alkylation is 10.77 and 9.24 kcal/mol, respectively (Figure 3), suggesting that the reaction will occur more readily at N-2.

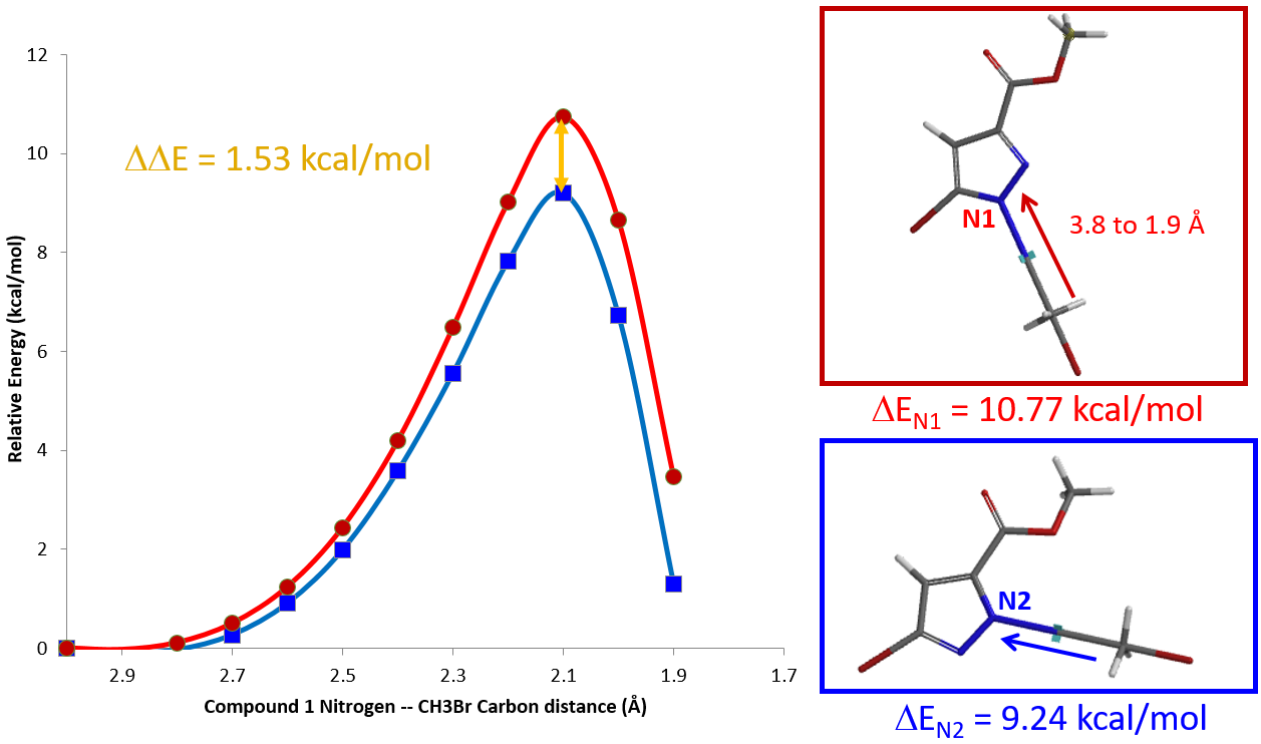


Figure 3. Reaction energy profiles of N1- (**red**) and N2- (**blue**) alkylation paths

Plugging into Arrhenius Equation the estimated difference in activation energies (ΔE_{N1} -ΔE_{N2}, 1.53 kcal/mol), an estimated product ratio of around 13:1 at 25 °C is obtained (See Chapter 8). Experimentally, when we alkylated pyrazole **1** at 25 °C, an N-2 to N-1 product ratio of 10:1 was obtained, close to the estimation!

$$k = Ae^{\frac{-E_a}{RT}}$$

$$\frac{[N2]}{[N1]} = \frac{k_{N2}}{k_{N1}} = \frac{Ae^{\frac{-\Delta E_{N2}}{RT}}}{Ae^{\frac{-\Delta E_{N1}}{RT}}} = e^{\frac{\Delta E_{N1} - \Delta E_{N2}}{RT}}$$

Figure 4. Calculating product ratio with Arrhenius equation

Relative Difference in Activation Energy Controls Product Ratio

Table 1 illustrates how changes in relative difference in activation energy (ΔE1-ΔE2) and reaction temperature affect product ratio, an important consideration in our synthetic planning and execution.

ΔE1-ΔE2 (kcal·mol^{-1})	(ΔE1-ΔE2)/RT	N2/N1 product ratio	T (°C)
0.0	1.000	1:1	25
1.4	2.364	11:1	25
1.4	1.793	6:1	125
2.8	4.728	113:1	25
4.2	7.092	1201:1	25

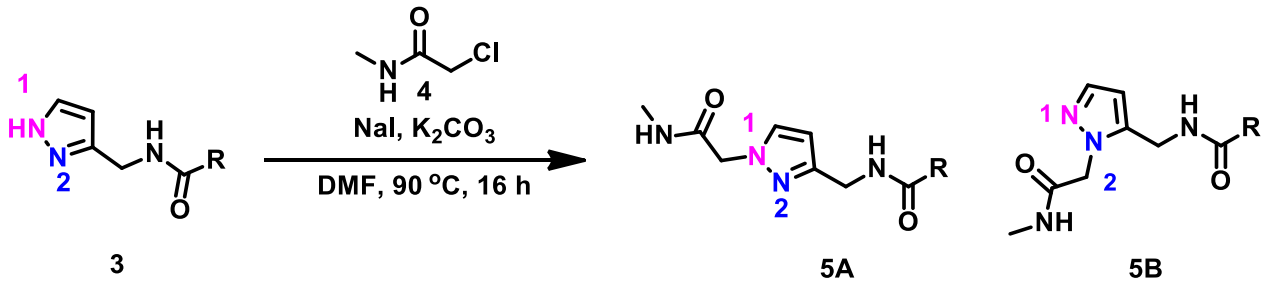
Table 1. Product ratio obtained with different ΔΔE and temperature

For a reaction with two competing pathways, with a difference of 1.4 kcal/mol in activation energy, the product ratio will be 11:1 with reaction run at 25 °C; the ratio will decrease to 6:1 when the reaction temperature is raised to 125 °C.

We found in practice that when the activation energy difference is around or below 1.4 kcal/mol, we can empirically adjust reaction temperature and solvent to optimize the ratio. On the other hand, if the calculated activation energy difference is above 3 kcal/mol, it is difficult to reverse the intrinsic bias.

In summary, through QM calculations of competing alkylation paths of pyrazole, we could calculate the difference in their activation energies, estimate the ratio of products, and use the result to guide our synthetic planning and execution.

Building on What We Just Learned



3 5A 5B

Alkylation of pyrazole **3** with simple alkyl halide is N-1 selective. However, reaction with chloroacetamide **4** provides only N-2 alkylation product. What could have possibly caused the reversal in selectivity?

Return to Table of Contents

References:

[1] W. Hehre and S. Ohlinger, *A Guide to Molecular Mechanics and Quantum Chemical Calculations.* Irvine, CA, USA: Wavefunction, Inc., **2003**.
[2] *Spartan'18 Tutorial and User's Guide.* Irvine, CA, USA: Wavefunction, Inc., **2019**; pp 116-118.
[3] S. Liu & L.G. Pedersen, *J. Phys. Chem. A.* **2009**, *113*, 3648.
[4] 无机化学 第四版 （杨宏孝 王建辉 著） 高等教育出版社第七章 270-279 页。

Chapter 10 S_NAr Reaction of Polyhalogenated Heterocycles

Xiaocang Wei, Qiuyue Wang, Guqin Shi, John S. Wai

We discussed the use of QM calculated activation energy to analyze regioselective alkylation of pyrazole in Chapter 9. In this chapter we will discuss its application on S_NAr reactions of polyhalogenated heterocyclic compounds.

For the dichloropyrazolopyrimidine carboxylate **1** in Figure 1, there are three potential sites for nucleophilic attack, C1, C2 on the pyrimidine ring, and C3 on the ester moiety. What will happen when we treat compound **1** with aqueous sodium hydroxide in tetrahydrofuran? Is there a reliable way to predict reaction outcome?

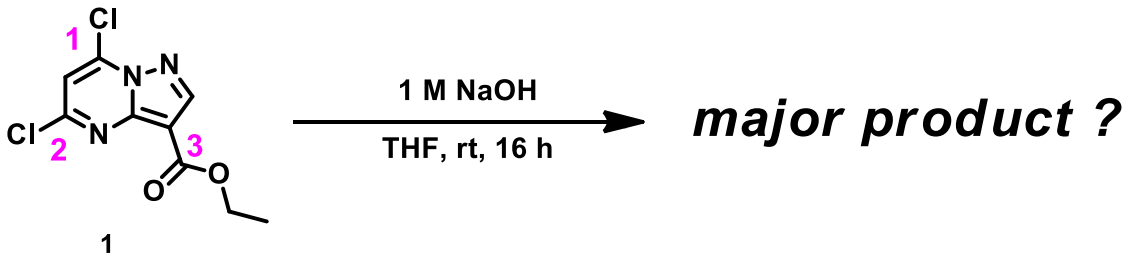


Figure 1. Hydrolysis of ethyl 5,7-dichloropyrazolo[1,5-a]pyrimidine-3-carboxylate (**1**)

LUMO Analysis of Compound 1

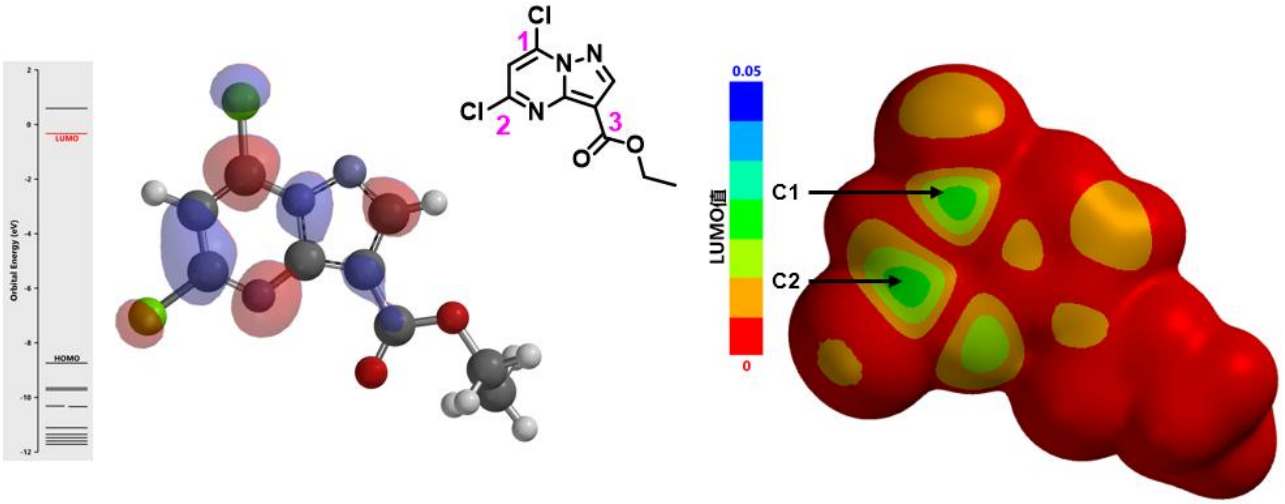


Figure 2. LUMO (left) and LUMO map (right) of compound **1**

Since compound 1 is an electrophile, we first look at its LUMO (Figure 2). There is close to none orbital lobe on C3, nucleophilic attack of the ester group is not likely to happen. There are significant LUMO lobes on C1 and C2, however they are similar in size, and their accessibility could not be differentiated with the calculated LUMO map (Figure 2). Under

these circumstances, we will proceed to calculate and compare the activation energy difference of these two potentially competing paths.

Calculation Model and Estimation of Product Ratio

We learned from experience that using bromide as a surrogate nucleophile to model this kind of reaction could provide a more reliable estimate. This will avoid overestimation of hydrogen bonding effects occasionally observed between the hydroxide anion and the relatively electron-deficient hydrogen on the aromatic system.

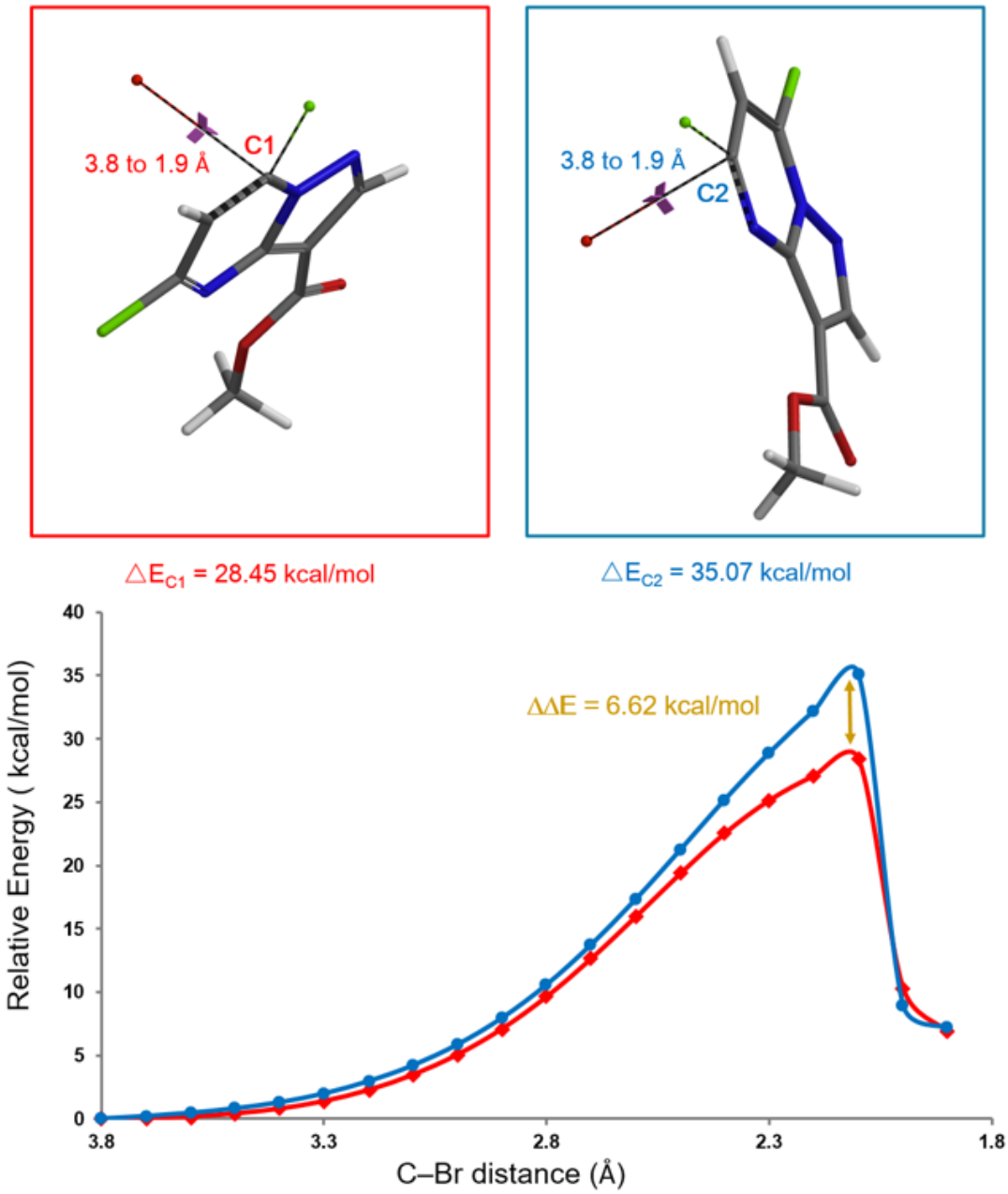


Figure 3. Activation energies of competing S_N2 reactions at C1 (**red**) and C2 (**blue**)

Shown on Figure 3 are the two reaction energy profiles calculated. Activation energies required for reactions at C1 and C2 are 28.45 and 35.07 kcal/mol, respectively, a difference

in 6.62 kcal/mol, in favor of nucleophilic attack at C1. Referring to Table 1 (discussed in Chapter 9), it becomes obvious that the hydrolysis will be highly selective at C1.

ΔE2-ΔE1 (kcal·mol^{-1})	(ΔE2-ΔE1)/RT	C1/C2 product ratio	T (°C)
0.0	1.0	1:1	25
1.4	2.4	11:1	25
2.8	4.7	113:1	25
4.2	7.1	1201:1	25

Table 1. Product ratio obtained with different ΔΔE and temperature

Experimental Results

After we treated compound **1** with aqueous sodium hydroxide in tetrahydrofuran at room temperature, we saw preferentially reaction at C1 position to provide compound **2**, the only product[1,2] (Figure 4). Neither nucleophilic substitution of the C2 chloro group nor hydrolysis of the ester group was observed.

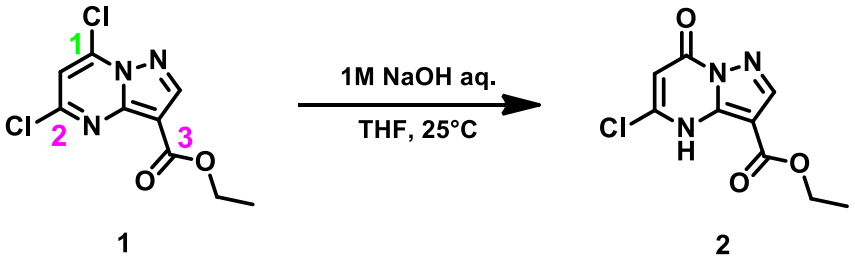


Figure 4. Selective hydrolysis of 5,7-dichloropyrazolo[1,5-a]pyrimidinecarboxylate (**1**)

In summary, for QM analysis of S_NAr reaction, we first calculate LUMO and LUMO map of the substrate to look for differences in LUMO lobes and their accessibility for nucleophilic interactions. When further differentiation is needed or an estimation of the product ratio is required, we will calculate for and compare the activation energies of competing reaction paths.

Building on What We Just Learned

For the chloropyridines below, they have a unique order of reactivity for nucleophilic attack[3].

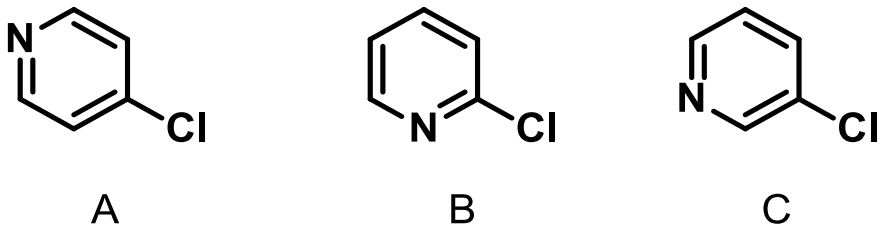


Figure 5. 4-chloro (**A**), 2-chloro (**B**), and 3-chloropyridines (**C**)

The relative differences in calculated activation energies correlate well with the reactivity pattern observed.

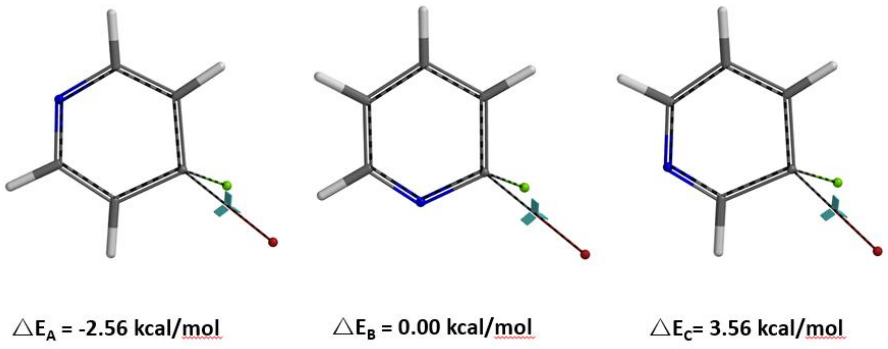


Figure 6. Relative differences in calculated activation energies of S_NAr reactions on 4-, 2-, and 3-chloropyridines

Then we arrive at an intriguing question. Why 4-chloropyridine is more reactive than 2-chloropyridine? The aromatic nitrogen is further away in the 4-chloro analog[4]. Why is 3-chloropyridine the least reactive?

Return to Table of Contents

References:

[1] R.K. Robins, G.R. Revankar, *et al*. *J. Heterocyclic Chem,* **1985***, 22*, 601.
[2] J.R. Zbieg, P.P. Beroza, J.J. Crawford, WO2019232216
[3] J.A. Joule & K. Mills. *Heterocyclic Chemistry 5th Ed*. Chichester, West Sussex, UK: Blackwell Publishing Ltd., **2010**; pp 117-118.
[4] Hint: Compare and contrast LUMO and LUMO+1 of the three chloropyridines.

Chapter 11 QM Assisted Structure Assignment with ^{13}C NMR Calculation

Dajin Tan, Yongsheng Chen, Guqin Shi, Qiuyue Wang, John S. Wai

For isomeric reaction products, 2D NMR spectroscopy is a tremendous tool for their structure assignment. Yet, there are still many instances that we could not do so conclusively. In this chapter, we will discuss the use of Quantum Mechanics calculated ^{13}C NMR data to assist structural assignment of these compounds.

Structure Assignment of Regioisomers

From a Chan-Lam coupling reaction of a chloropyrazolopyridine, we obtained a mixture of regioisomers (Figure 1). 2D NMR spectroscopy did not provide us with the necessary C-H or H-H correlation to differentiate them conclusively.

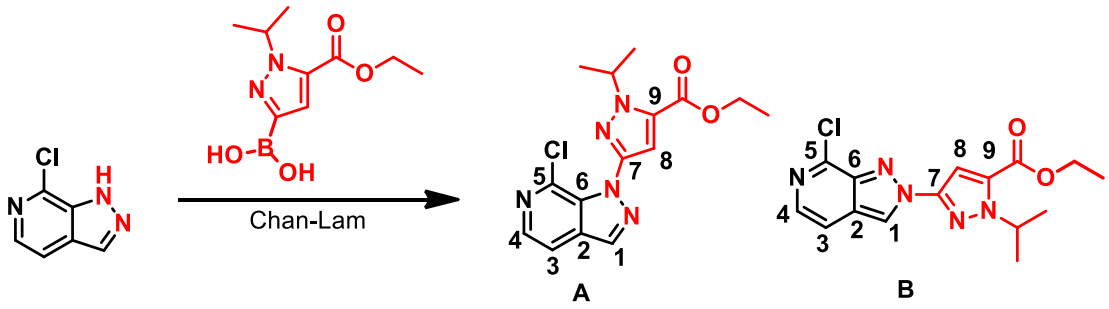


Figure 1. Reaction products from Chan-Lam reaction with pyrazolopyridine

We did ^{13}C NMR calculations for structures **A** and **B** with Spartan software[1] and compared with the spectroscopic data obtained experimentally. Tabulated in Table 1 are the experimental ^{13}C NMR data of product **1** and the calculated values for structures **A** and **B**. Based on the differences, product **1** is more likely to have structure **A**.

C#	(1) Experimental ^{13}C Chemical Shift, $CDCl_3$	Structure A Calculated δ	|Difference|	Structure B Calculated δ	|Difference|
1	134.7	134.0	0.7	123.0	11.7
2	130.3	131.3	1.0	125.8	4.5
3	114.0	114.2	0.2	112.8	1.2
4	138.7	139.0	0.3	138.8	0.1
5	134.3	134.9	0.6	146.4	12.1
6	133.9	135.9	2.0	144.9	11.0
7	132.1	145.5	13.4	148.1	16.0
8	108.5	114.7	6.2	105.3	3.2
9	144.2	134.3	9.9	134.1	10.1

Table 1. Comparison of experimental and calculated ^{13}C NMR data between product **1** and structures **A** and **B**, respectively (Spartan'16). Note: Green indicates $\Delta < 2$ ppm; blue: $2 \leq \Delta \leq 10$ ppm; red $\Delta > 10$ ppm. |Difference| represents the absolute value of the difference between the experimental and calculated values. Same for following Tables.

Similarly, we compared the experimental data of product **2** with the calculated values (Table 2). Product **2** aligned with structure **B** much better.

C#	(2) Experimental ^{13}C Chemical Shift, $CDCl_3$	Structure A Calculated δ	|Difference|	Structure B Calculated δ	|Difference|
1	121.1	134.0	12.9	123.0	1.9
2	125.0	131.3	6.3	125.8	0.8
3	113.2	114.2	1.0	112.8	0.4
4	137.2	139.0	1.8	138.8	1.6
5	143.1	134.9	8.2	146.4	3.3
6	143.3	135.9	7.4	144.9	1.6
7	133.1	145.5	12.4	148.1	15.0
8	102.3	114.7	12.4	105.3	3.0
9	146.7	134.3	12.4	134.1	12.6

Table 2. Comparison of experimental and calculated ^{13}C NMR data between product **2** and structures **A** and **B**, respectively (Spartan'16)

This allows us to tentatively assign the structures to products **1** and **2** with confidence. Subsequent experimental results substantiate the assignments.

We were curious to know whether this could be applied to natural products with more complex structures. Below is one of the examples we looked at.

Structure Assignment of Complex Natural Product

In 2010, Yongsheng Che *et al*. reported the discovery of Cytosporolides **A - C** with antibacterial activity in *Org. Lett.*[2]. Proposed structure for Cytosporolides **A** and its absolute stereo-chemistry assignment were based on 2D NMR spectroscopy, single crystal X-ray diffraction, data in the literature, etc. The antibacterial properties of Cytosporolides and their unique structures attracted significant interest among synthetic chemists. In 2011, Australian scientists Jonathan H. George *et al*., based on chemical synthesis and data analysis, proposed

an alternate six-membered aryl ether structure for Cytosporolides **A** (Figure 2), and published their results in *Org. Lett.*[3].

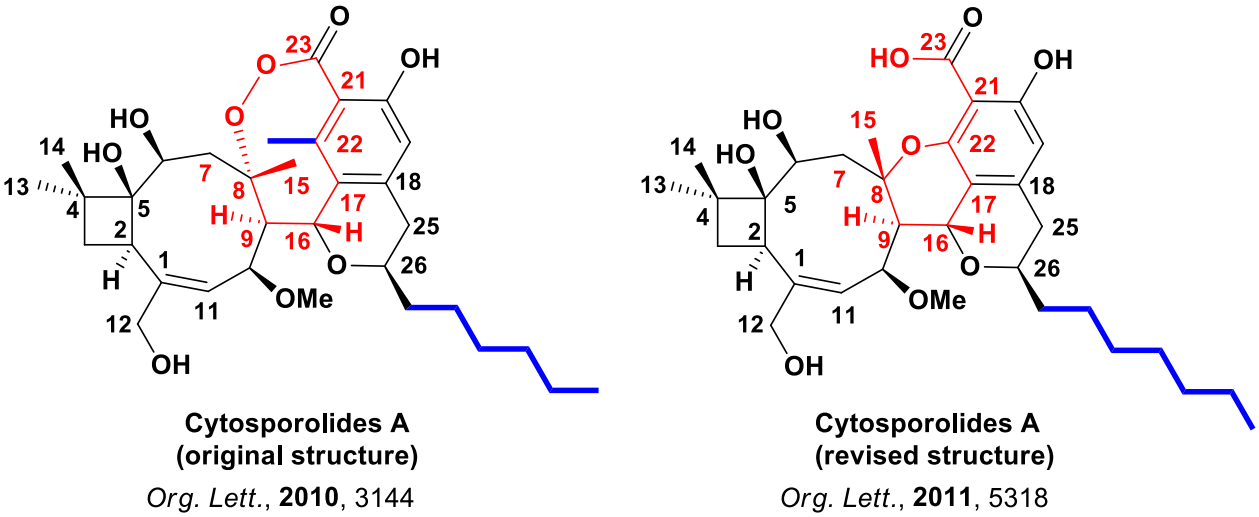


Figure 2. Proposed structures for Cytosporolides **A**

The key structural differences of the two proposed structures are highlighted in red and blue (Figure 2). Calculated ^{13}C NMR data of the proposed structures for the key signals for C7-C9, C17-C23, and C25 were compared with experimental ones (Table 3). The data are supportive of the structure proposed by George *et al*.

C#	Experimental ^{13}C Chemical Shift, $(CD_3)_2C=O$	Original 2010_Structure Calculated δ	|Difference|	Revised 2011_Structure Calculated δ	|Difference|
C7	46.2	51.9	5.7	45.8	0.4
C8	87.5	96.1	8.6	87.7	0.2
C9	49.6	66.3	16.7	46.9	2.7
C17	113.5	132.2	18.7	111.7	1.8
C18	145.3	138.6	6.7	143.5	1.8
C19	109.4	111.3	1.9	107.3	2.1
C20	163.8	156.4	7.4	166.9	3.1
C21	99.2	117.5	18.3	99.0	0.2
C22	151.0	150.5	0.5	153.6	2.6
C23	171.9	170.6	1.3	173.7	1.8
C25	34.9	30.0	4.9	34.2	0.7

Table 3. Comparison of experimental and QM-calculated ^{13}C NMR data with the two proposed structures of Cytosporolides **A**. Tabulated are the core differences. (Spartan'18)

The above cases highlighted the application of QM NMR calculation to assist structure assignments, from relatively simple structures to more complicated natural products. In our laboratories, we set up the NMR experiments and calculations concurrently to enable proper structure assignments in a timely manner.

An Important Tip

NMR algorithm in the Spartan software was optimized with experimental data obtained with $CDCl_3$ as solvent. It is prudent to obtain experimental data in $CDCl_3$ for proper comparison.

Return to Table of Contents

References:

[1] a. W. Hehre and S. Ohlinger, *Spartan'18 Tutorial and User's Guide*, Irvine, CA, USA: Wavefunction, Inc. **2019**; *Calculating NMR Spectra*, pp 525-529, b. W. Hehre, P. Klunzinger, B. Deppmeier, A, Driessen, N. Uchida, M. Hashimoto, E. Fukushi, Y. Takata, *J. Nat. Prod.*, **2019**, *82*, 2299.
[2] Y. Li, S.B. Niu, B.D, Sun, S.C. Liu, X.Z. Liu, Y.S. Che, *Org. Lett.*, **2010**, *12*, 3144.
[3] J.T.J. Spence & J.H. George, *Org. Lett.*, **2011**, *13*, 5318.

Chapter 12 LUMO-Assisted Design of Disulfide Linkers

Zijin Xu, Jinhua Chen, Dong Pan, Guqin Shi, John S. Wai

Antibody-Drug Conjugates

Antibody-Drug Conjugates (ADCs) are complex molecules composed of biologically active small molecule linked to antibodies (mAb) through a cleavable or non-cleavable linker. The mAb acts as a specific carrier to transport the small molecule into the target cells. Linkers are usually based on chemical motifs such as hydrazones, peptides, and disulfides. In this chapter, we will discuss the optimization of the disulfide linker for ADC drug discovery.

Design of Disulfide Linkers

Conjugation makes use of the sulfhydryl group of cysteines (mAb-SH) on monoclonal antibody to undergo a S_N2 type reaction with the disulfide bond on the linker[1] to generate the target ADC (Figure 1, product **A**). A specific antibody is chosen for target delivery of the ADC into tumor cells. Glutathione, which is high in concentration in tumor cells, will reduce the disulfide bonds, release the highly active small molecule drugs for cell kill.

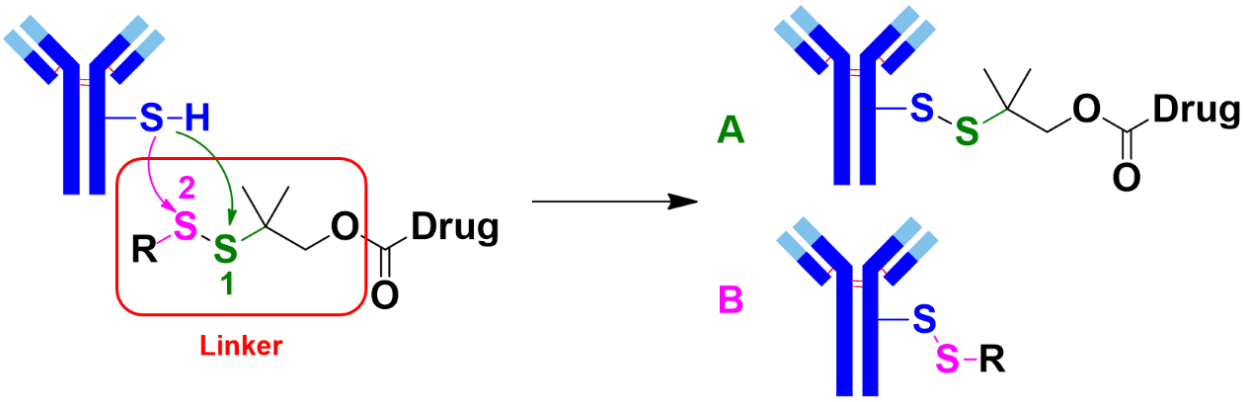


Figure 1. Preparation of ADC by thiol-disulfide exchange reaction

To improve ADC *in vivo* stability, two methyl groups are usually added at the α position of the central sulfur atom S1[2] to minimize undesired dissociation of ADC before it reaches the target cells. However, this modification leads to mAb-SH attacking at the undesired S2 during conjugation, resulting in by-product **B** (Figure 1).

As such, we need to optimize the linker for selective generation of product **A**, improving the Drug/Antibody Ratio (DAR) value of ADC target.

In our case, the antibody contains two free cysteine thiols (mAb-SH), giving a maximum DAR value of 2. We need to design better linkers to make full use of each mAb internalized into the cancer cells.

General organic chemistry taught us that the lower the p*K*a value of the conjugated acid of the leaving group R-S- (Figure 1), the better it is as a leaving group. Then in principle, conjugation with linker-drugs derived therefrom would provide more target products **A**, i.e., a higher DAR value. However, such correlation is not observed with our results (Figure 2).

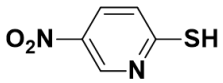
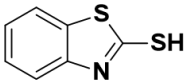
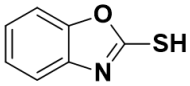
R–SH			
pKa	**5.39**	**6.12**	**5.64**
DAR	**0.1**	**0.2**	**0.5**

Figure 2. p*K*a of conjugated acids (leaving group) and corresponding DAR values[3]

Relationship between LUMO and DAR Values

We decided to analyze the above results with Quantum Mechanics.

For thiol exchange to proceed, HOMO of mAb-SH needs to interact with LUMO on either **S1** or **S2**. Selectivity will depend on relative size of the LUMO lobes. Shown on Figure 3 are LUMO of the three disulfide linkers above. LUMO lobe on the sulfur atom **S1** gradually increases from nitropyridine to benzoxazole, as the corresponding DAR value improves (Figure 3). This could be a useful QM parameter to guide our linker design.

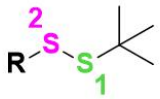
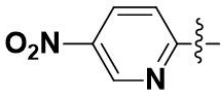
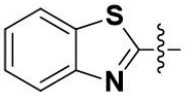
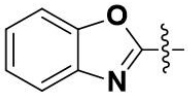
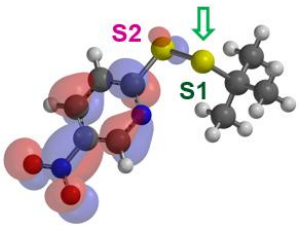
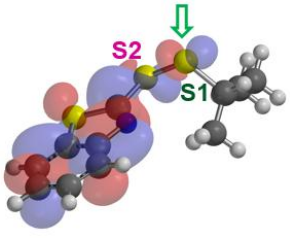
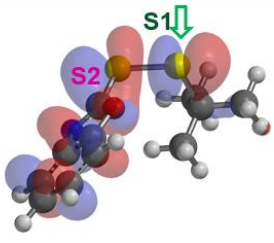
			
LUMO			
DAR	**0.1**	**0.2**	**0.5**

Figure 3. LUMO of linkers and corresponding DAR values

LUMO-Assisted Design of Novel Disulfide Linkers

Based on the above analysis, we evaluated the relative size of their LUMO lobes on sulfur atoms **S1** and **S2** of a series of linkers designed. Isothiourea and methyl sulfone (MTS) derived linkers were identified with extensive LUMO lobes over the desired **S1** (Figure 4). These suggest that mAb-SH could attack selectively at **S1** and provide excellent conjugation DAR. To avoid the production of toxic thiourea, we focused our exploration efforts on MTS.

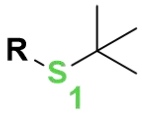
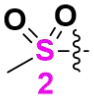
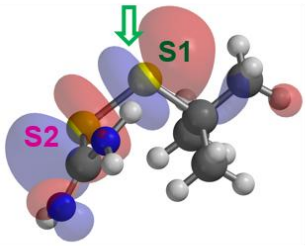
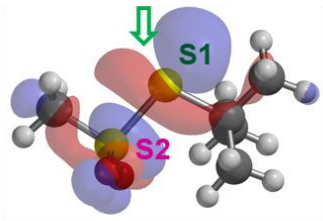
	 isothiourea	 MTS
LUMO		
DAR	**1.8**	**1.9**

Figure 4. LUMO of isothiourea and MTS linkers and corresponding DAR values

Even though synthesis of the key linker for conjugation with mAb-SH encountered significant difficulties, LUMO analysis provides us with compelling reasons to persist and prepare the MTS linker-drug successfully. More importantly, a DAR of 1.9 was obtained in conjugation!

In summary, we established the use of LUMO to guide the design of disulfide linkers, discovered MTS linker, and improved DAR by nearly 20 folds over the nitropyridine. A major hurdle for mAb thiol exchange is resolved[4].

Return to Table of Contents

References:

[1] R.D. Bach, O. Dmitrenko, C. Thorpe., *J. Org. Chem*. **2008**, *73*, 12.
[2] R.V.J. Chari, M.L. Miller, W.C. Widdison, *Angew. Chem. Int. Ed.* **2014**, *53*, 3796.
[3] See Chapter 3, "Application of Electrostatic Potential Map in Acidity Estimation".
[4] J.D. Sadowsky, T.H. Pillow, J. Chen, H. Yao, Z. Xu, J.S. Wai, *et al., Bioconjugate Chem*. **2017**, *28*, 2086.

Chapter 13 The Mysterious Factor X that Confers Stereoselectivity

Qiuyue Wang, Guqin Shi, John S. Wai

In synthetic planning of complex organic molecules, π-facial selectivity is often an important element to consider. The origin of such stereoselectivity has been attributed to a number of factors such as electrostatics, steric, hyperconjugation effects, and a mysterious Factor X[1,2]. In this chapter we will discuss what is Factor X, how we could use QM to model, analyze, quantify, and visualize this important controlling element in stereoselectivity.

Norbornene Epoxidation

It is well-known that peracid epoxidation of norbornene results in highly selective *exo* product, as shown in Figure 1.

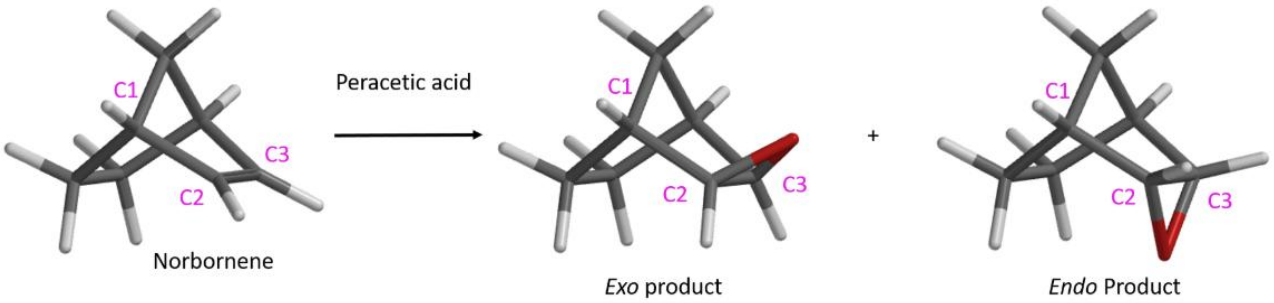


Figure 1. Epoxidation of norbornene

We modeled this Prilezhaev reaction with peracetic acid attacking from both faces of norbornene. Shown on Figure 2 is the approach from bottom face, leading to the formation of *endo* product. Geometry and energy were calculated as O-C2 distance varies from 3.2 Å to 1.9 Å, with a step size of 0.1 Å, assuming synchronized O-C2 and O-C3 bond formations.

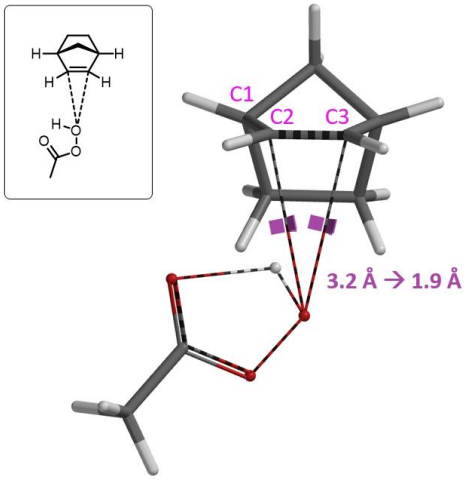


Figure 2. Model of epoxidation of norbornene with peracetic acid attacking from bottom face

Reaction energy profile calculations (Figure 3) suggested activation energies ΔE of 26.95 and 22.33 kcal/mol for peracetic acid attacking from bottom and top faces, respectively, to provide the corresponding *endo* and *exo* products. A difference (ΔΔE) of 4.62 kcal/mol indicates a highly selective reaction, producing mainly the *exo* product.

What causes the energy difference of 4.62 kcal/mol?

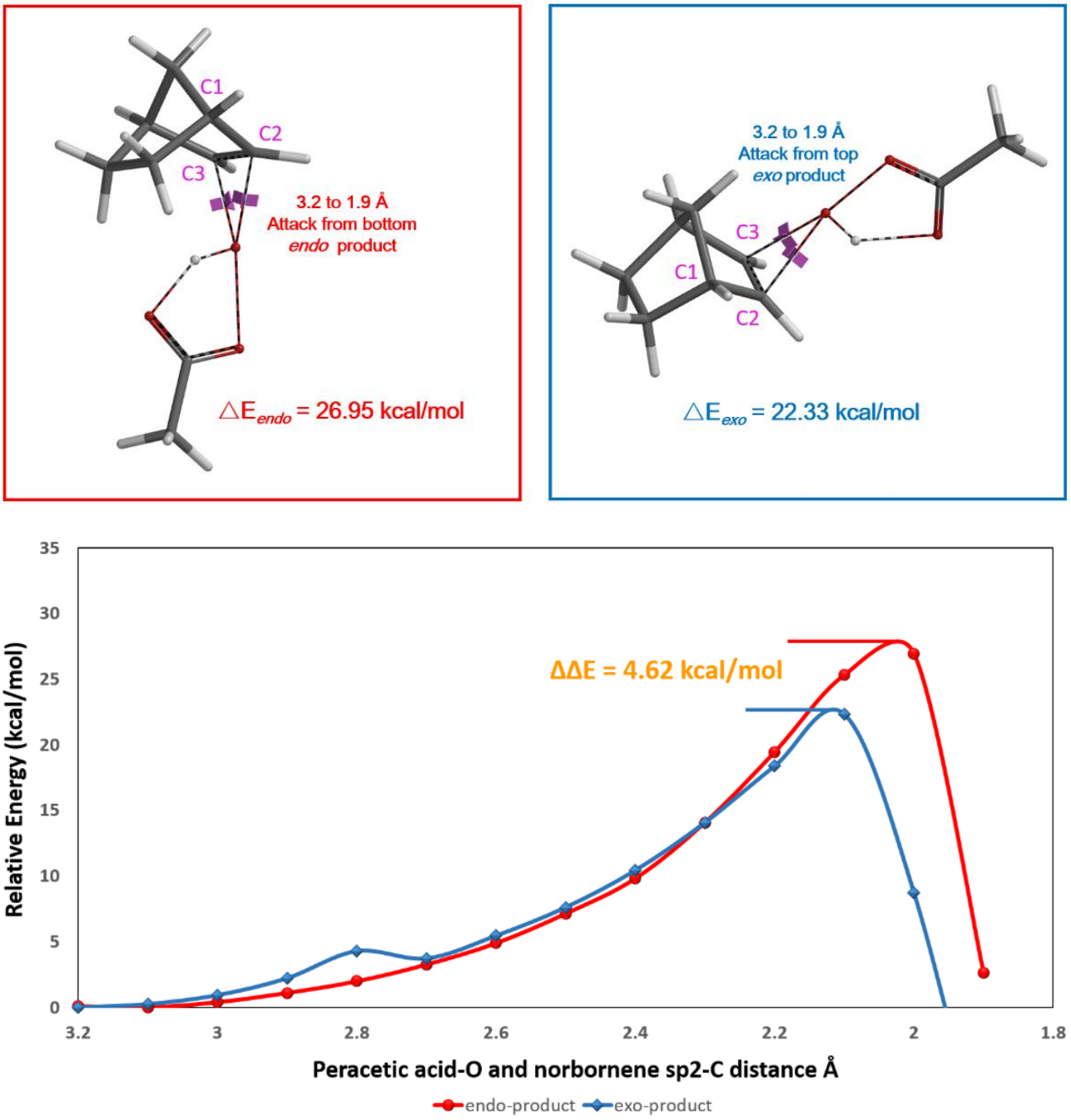


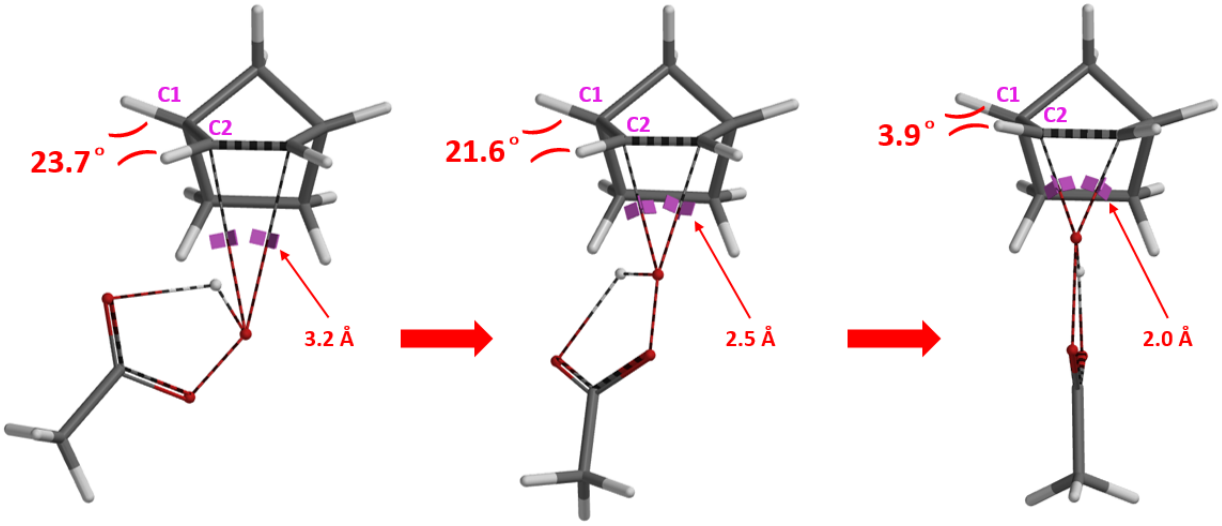
Figure 3. Reaction energy profiles of peracetic acid attacking from bottom/top faces giving *endo*/*exo* products.

Torsional Strain -- The Factor X

In stereochemistry, every three non-collinear atoms of a molecule determine a plane. When two such planes intersect, the angle between them is defined as a dihedral angle, or a torsion angle.

Careful inspection on the geometries at initial states (with O-C bond length at 3.2 Å), mid-way (2.5 Å), and the transition states (2.0 Å or 2.1 Å), showed that as epoxidation proceeds, C1-H and C2-H bonds approach each other and build up steric congestion when peracid attacks from bottom face (Figure 4). In contrast, C1-H and C2-H move away from each other and relieve steric stress when being attacked from top face.

Attack from bottom face providing endo product:



Attack from top face providing exo product:

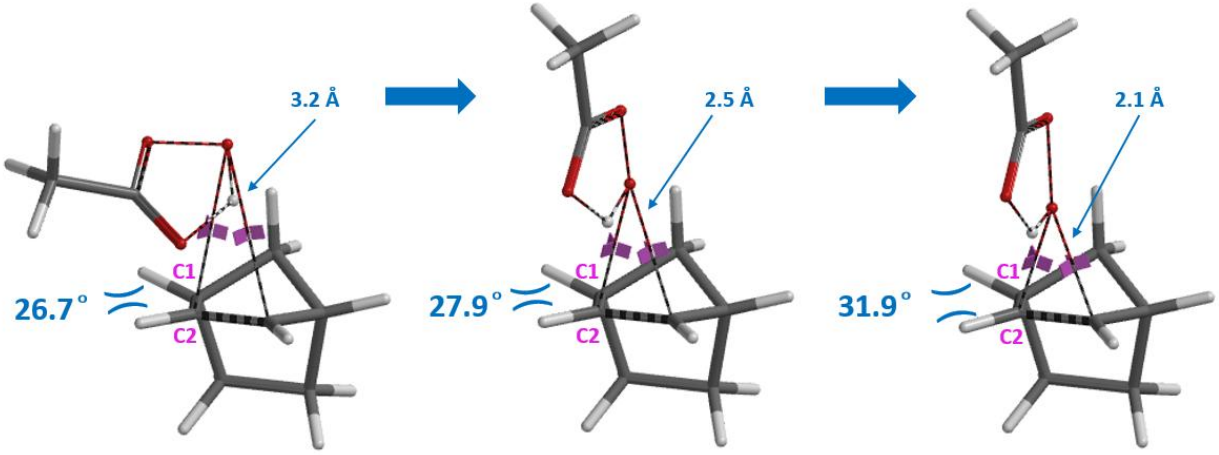


Figure 4. Dihedral angle changes during Prilezhaev reaction progress

In summary, it is the difference in torsional strain that controls the selectivity observed in peracetic epoxidation of norbornene. For what was once considered to be a mystery (the Factor X), QM analysis shows us that it is comprehensible, predictable, and quantifiable.

Building on What We Just Learned

For the following nucleophilic opening of epoxide **1**, if we apply the above torsional strain analysis, shall we expect **2** or **3** to be the major product?

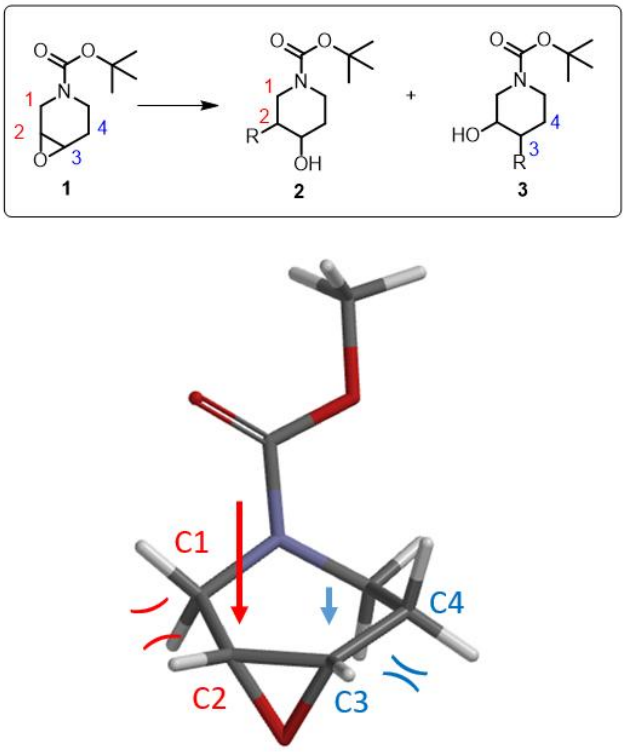


Figure 5. Ring-opening reaction of epoxide **1**

Return to Table of Contents

References:

[1] a) R. Huisgen, P.H.J. Ooms, M. Mingin, N.L. Allinger, *J. Am. Chem. Soc.* **1980**, *102*, 3951. b) R. Huisgen, *Pure Appl. Chem.,* **1981**, *53*, 171.
[2] H. Wang & K.N. Houk, *Chem. Sci*. **2014**, *5*, 462.
[3] R.D. Bach, C. Canepa, J.E. Winter, P.E. Blanchette, *J. Org. Chem.* **1997**, *62*, 5191.

Chapter 14 Application of Torsion Scan in Medicinal Chemistry

Zhengquan Zhou, Guqin Shi, Liting Dong, Dong Pan, and John S. Wai

We learned about dihedral angle and torsional strain in Chapter 13. In this chapter, we will discuss torsion scan, a valuable tool for medicinal chemists to evaluate what are the likely active conformations of their drug molecules and to assess potential atropisomerism.

Application of Torsion Scan in Drug Design

Shown on Figure 1 are 2D structure of two drug molecules, in which naphthyridine amide **1** has excellent inhibitory activity, while quinoline amide **2** exhibits much lower biological activity. These two chemical structures differ only at position 6 with one having a nitrogen atom and the other having a carbon atom, respectively.

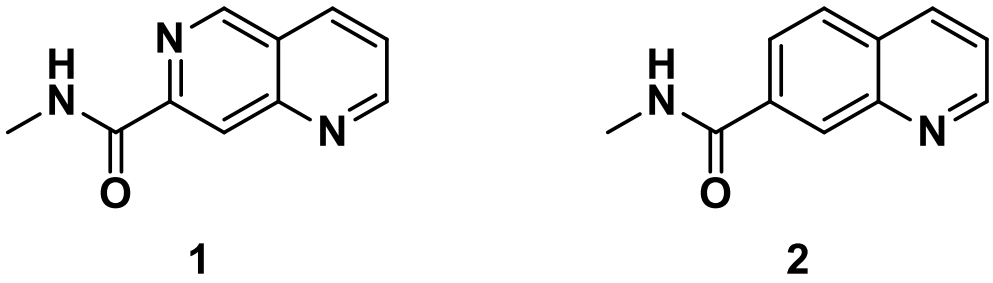


Figure 1. Naphthyridine amide **1** and quinoline amide **2**

In this situation, we conduct a torsion scan with Density Functional Theory calculation, rotating the dihedral bond in 20° increment from 0° to 360° between the amide plane and the aromatic core for these two inhibitors, to compare their energy profiles.

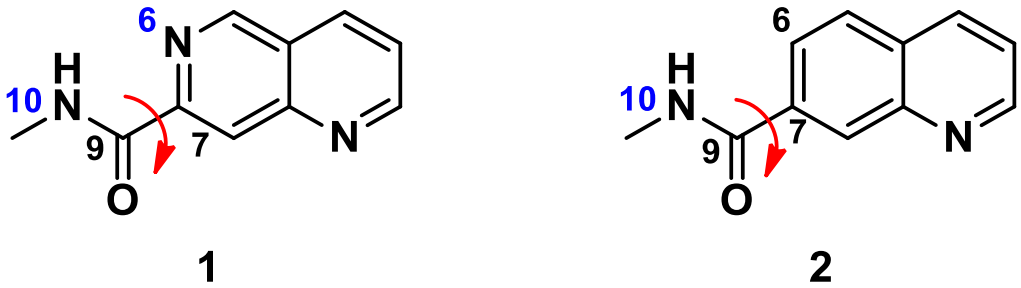


Figure 2. Torsion scan of naphthyridine amide **1** and quinoline amide **2**

The torsion scan energy profile of naphthyridine **1** is shown in Figure 3, with **1A** being the lowest energy conformation. In this conformation, the amide group is coplanar with the naphthyridine ring and its N10-H is hydrogen bonding with the N6 of the naphthyridine.

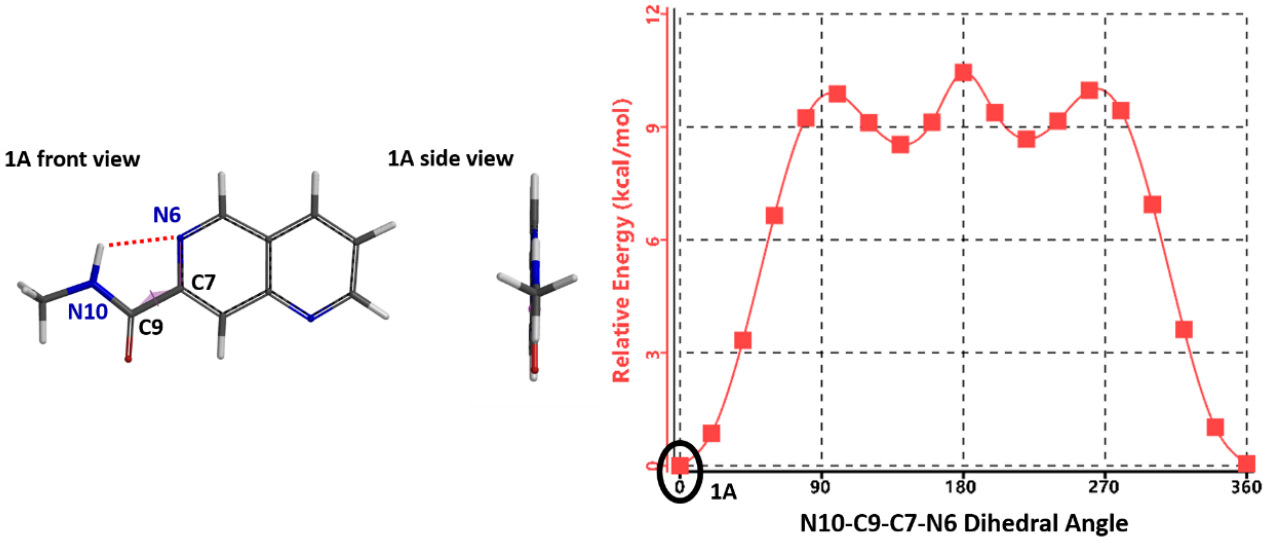


Figure 3. Torsion scan energy profile of naphthyridine amide **1**

Quinoline amide **2**, with the N6 atom on the naphthyridine ring replaced with a C atom, loses the intramolecular hydrogen bonding, as shown in Figure 4, conformation **2B**. Furthermore **2B** encounters steric congestion between the N10-H and C6-H. As such, the lowest energy conformation of quinoline amide **2A** adopts a N10-C9-C7-C6 dihedral angle at 200° with the amide group being non-coplanar to the quinolone. To adopt the ground state conformation as naphthyridine **1A**, the quinoline conformation **2B** will be 2 kcal/mol higher in energy than its ground state **2A**.

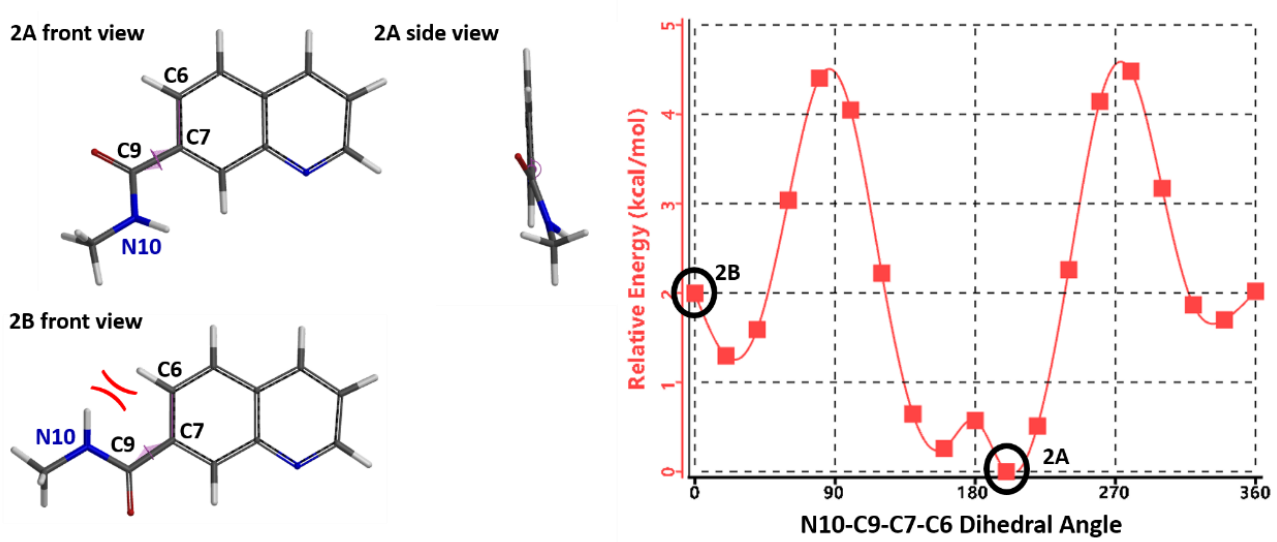


Figure 4. Torsion scan energy profile of quinoline amide **2**

Since naphthyridine **1** is more biologically active, we can speculate that its coplanar structure is likely to be the dominant conformation for target binding, and that the orientation of the carbonyl group and/or the amino group in the amide might also be biased for specific hydrogen bonding with the target. This provides important conformational information for design of drug molecules.

On top of assisting us in analyzing the low-energy conformations of drug molecules, we found torsion scan to be very useful in analyzing the rotational isomers (Atropisomerism), providing

guidance for their separation, and enabling us to assess whether this could pose significant challenges for further development of these molecules.[1]

Application of Torsion Scan in Separation and Characterization of Drug Targets

^{1}H NMR spectroscopy revealed that in our preparation of pyrazolyl triazole derivative **3**, a mixture of two compounds was obtained (Figure 5). Could they be regio-isomers? After this mixture was subjected to Supercritical Fluid Chromatography (SFC) on chiral stationary phase, we found that the enriched fractions quickly equilibrated back to same ratio. This makes more sense if the product is actually a mixture of rotational isomers, with hindered rotation of the C-N single bond across the pyrazole and triazole planes, instead of being a mixture of regio-isomers.

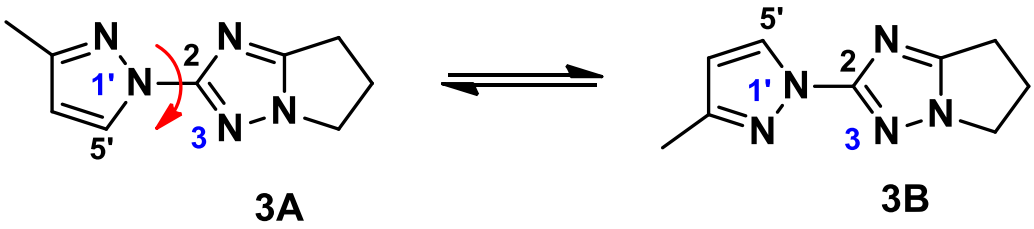


Figure 5. Rotamers of pyrazolyl triazole derivative **3**

We have been able to further substantiate the conclusion with a QM torsion scan calculation of compound **3**. Shown on Figure 6 is the energy profile of the torsion scan on dihedral angle N3-C2-N1'-C5' from 0° to 360°, in 19 steps.

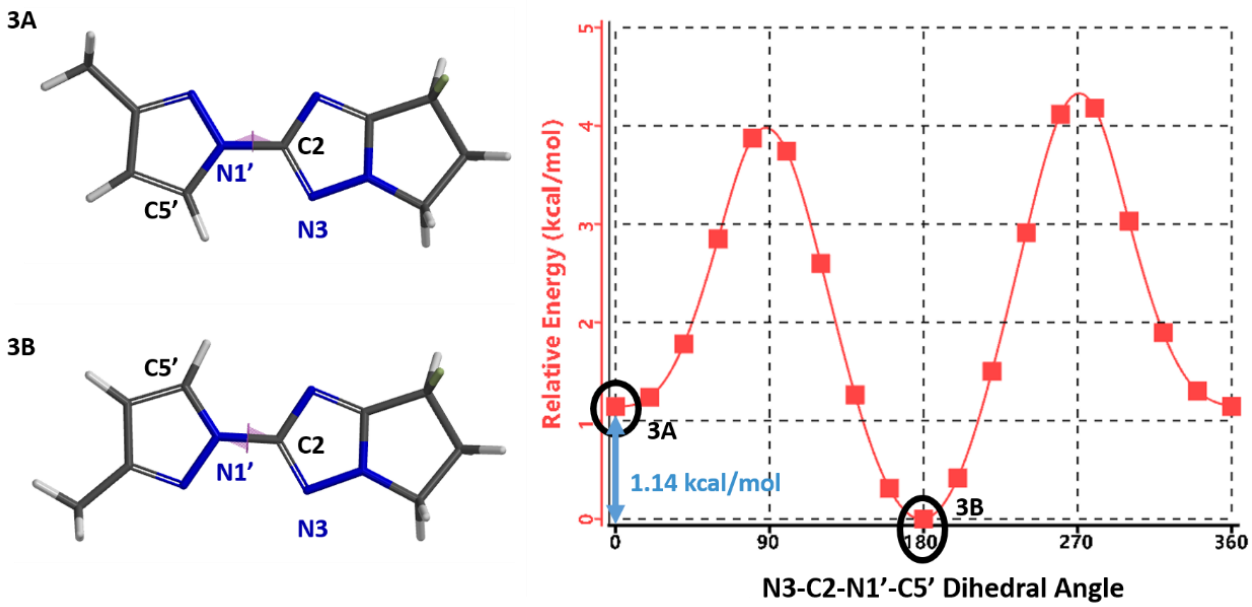


Figure 6. Torsion scan energy profile of pyrazolyl triazole derivative **3**

Calculation shows an energy difference of around 1.14 kcal/mol for the two low-energy conformations, **3A** and **3B**. Based on the established relationship between Gibbs Free Energy

change and equilibrium constant, $\Delta G = -RT\ln K_{eq}$（with $K_{eq} = \frac{[3B]}{[3A]}$），the energy difference translates to an equilibrium ratio of **3A** : **3B** ≈ 1:7 at 25 °C, consistent with experimental observation.

Furthermore, the calculation indicates that the energy barriers for the conversion of conformations are 2.73 (**3A** to **3B**) and 3.87 kcal/mol (**3B** to **3A**), respectively (Figure 7). These level of energy barriers are enough to slow down their interconversion for discrete ^{1}H NMR signals of each rotamer and to have them partially separated with SFC. This is consistent with the report that an energy barrier higher than 23 kcal/mol (kinetic barrier[1]) is needed to prevent the purified rotamers from equilibrating back to an equilibrium mixture, accounting for our observation that partially separated mixtures of pyrazolyl triazole rotamers reverse back to its equilibrium ratio.

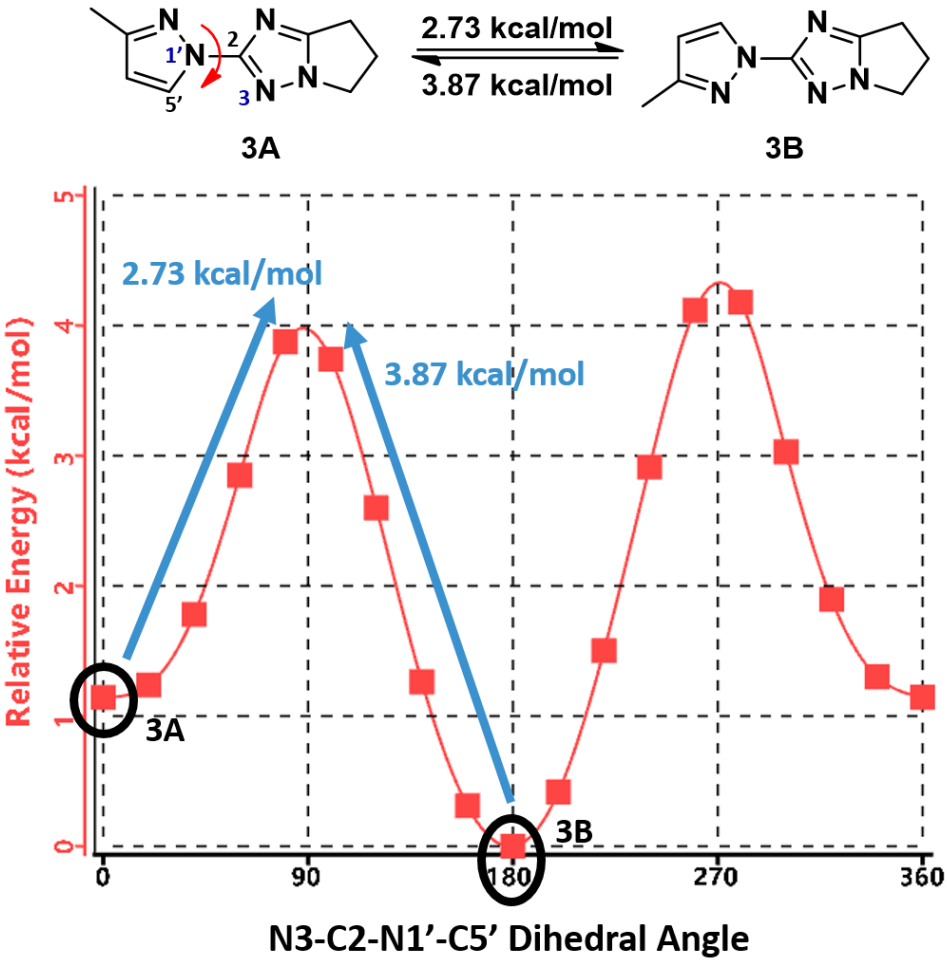


Figure 7. Interconversion energy barriers for pyrazolyl triazole conformers **3A** and **3B**

In summary, torsion scan is a very powerful calculation tool for medicinal chemists to establish the likely active conformation of their inhibitors, to assess whether their targets of interest could exhibit atropisomerism, and to estimate how quickly they could interconvert.

Building on What We Just Learned

Shown on Figure 8 is the torsion scan energy profile calculated for the phenylated pyrazolyl imidazole analog 4. What will be the equilibrium ratio of these two rotamers (**4A** and **4B**)? Is it likely we could separate and purify each of the rotamers for biological assays?

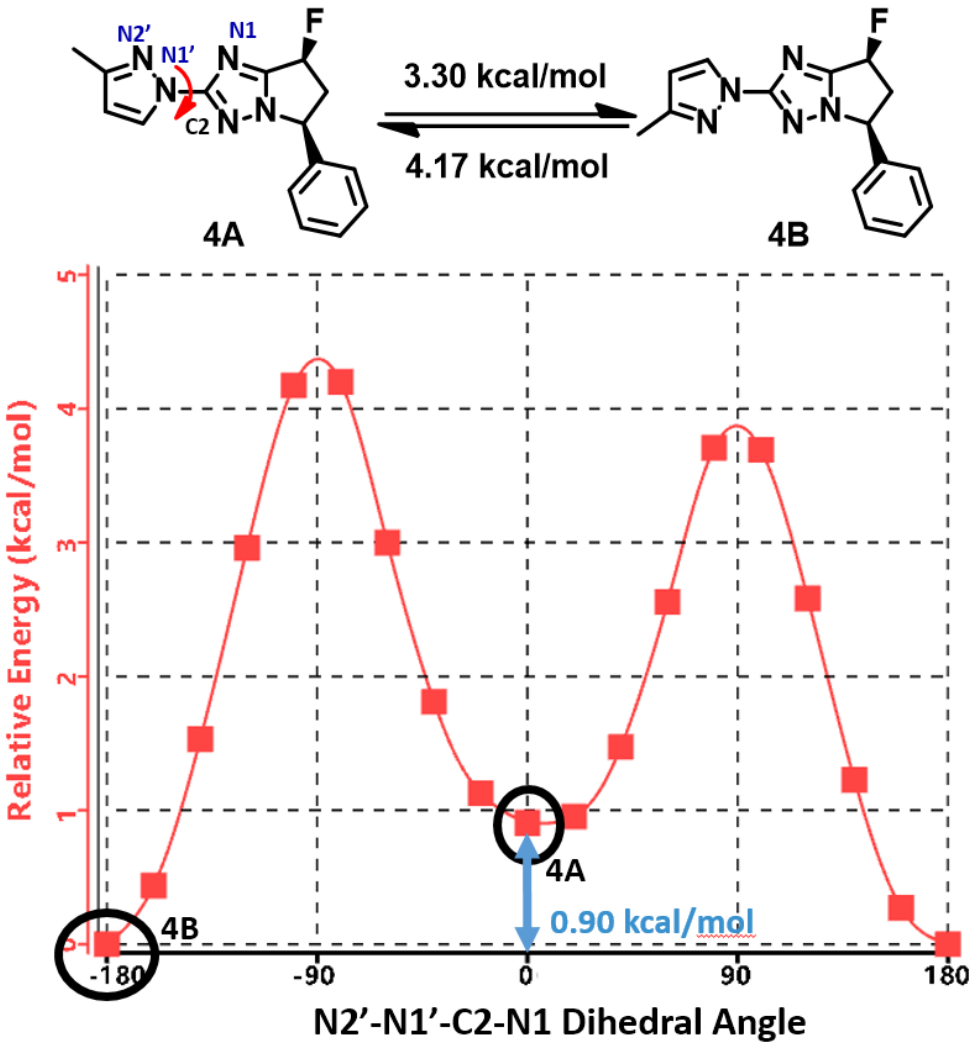


Figure 8. Torsion scan energy profile for pyrazolyl imidazole analog **4**

Return to Table of Contents

References:

[1] a) A.M. Belostotskii. *Conformational Concept for Synthetic Chemist's Use: Principles and in Lab Exploitation*. Hackensack, NJ, USA: World Scientific Publishing, **2015**. b) Kinetic Barrier: Atropisomers with low energy barriers usually exist as equilibrium mixtures due to their rapid interconversion. For them to be effectively purified/separated, the kinetic barrier generally has to be ≥ 23 kcal/mol, with a t½ at around 33 days, at 25 °C. The interconversion rate is also affected by temperature and solvents. c) S.R. LaPlante, P.J. Edwards, L.D. Fader, A. Jakalian, *ChemMedChem* **2011**, *6*, 505.

Chapter 15 Alcohol Speeds up Boc Protection of Primary Amines

Renwei Zhang, Guqin Shi, Dong Pan, and John S. Wai

Protection of nucleophilic amino groups is an important strategy in organic synthesis. Among the protecting groups, Boc offers excellent stability towards a variety of nucleophilic and electrophilic reaction conditions, yet readily undergoes deprotection under mild acidic conditions. Reaction of Boc anhydride with aliphatic amines is generally quite fast, however protection of weakly nucleophilic aromatic amines will take longer, may require heating or even addition of strong bases.

In 2006, Tirayut Vilaivan reported that alcoholic solvents greatly enhance the rate of Boc protection of aromatic amines at room temperature, and do not need to include any base in the reaction mixture (Figure 1).[2] Kinetics study with NMR spectroscopy showed that the reaction between *p*-toluidine and Boc anhydride in CD_3OD is 70 fold faster than in $CDCl_3$. We are curious in how alcohol catalyzes this reaction and decide to analyze this reaction further with QM calculations.

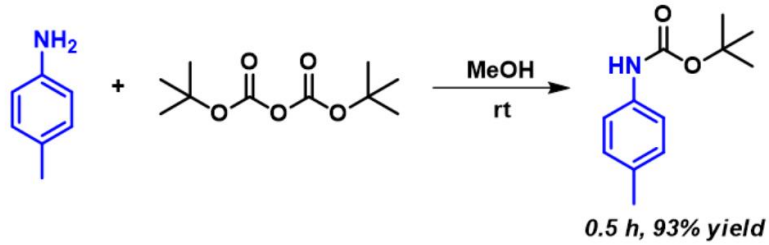


Figure 1. Methanol enhances the rate of Boc protection of aromatic amines

Aniline Boc Protection Modeling

QM analyses quickly revealed potentially the presence of hydrogen bond between N-H of aniline and the carbonyl oxygen O1' in Boc anhydride (Figure 2A, blue dotted line) that could lead to a 6-membered transition state. As such, a reaction energy profile calculation with the N-C2 distance to change from 3.2 Å to 1.4 Å (Figure 2A, magenta dotted line), with a step size of 0.1 Å, and enabling hydrogen bonding interactions between N-H of aniline and O1' of Boc anhydride (Figure 2B) was set up and provided a calculated activation energy (ΔE) of 16.22 kcal/mol (Figure 4, red line).

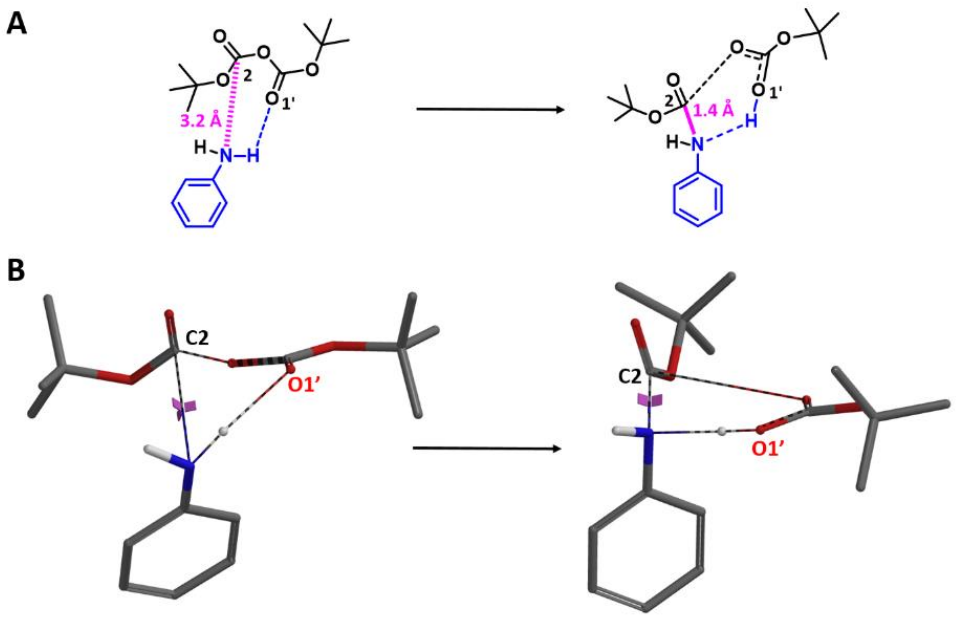


Figure 2. Aniline Boc protection modeling

Include Methanol in Modeling of Aniline Boc Protection

For modeling of the reaction in the presence of methanol, we positioned it for potential non-covalent interactions with both the carbonyl oxygen O2' and the second N-H of aniline (Figure 3A, red dotted line). QM provided a calculated activation energy (ΔE) of 13.94 kcal/mol, 2.28 kcal/mol lower than in the absence of methanol (Figure 4, blue line), accounting for the experimental observation that the rate of aniline Boc protection proceeds faster in methanol.

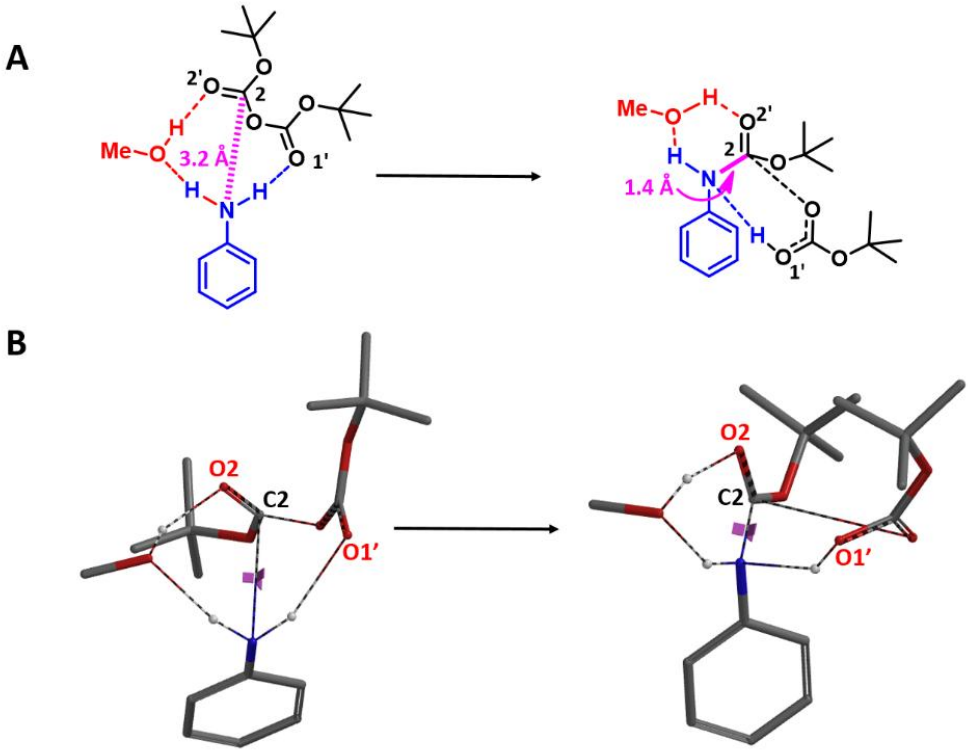


Figure 3. Including methanol in modeling of aniline Boc protection

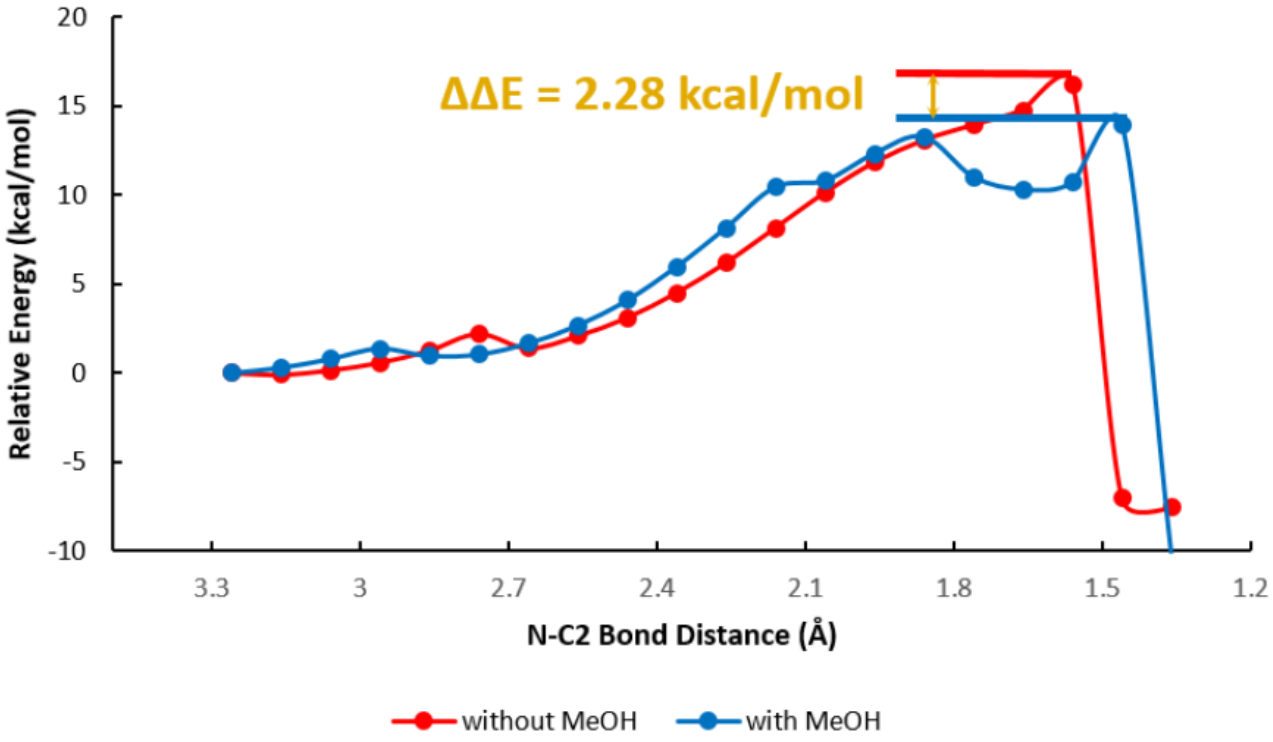


Figure 4. Reaction energy profiles of Boc protection of aniline in the absence (**red**) and in the presence (**blue**) of methanol

Stabilization of Transition States with Alcohol

Let's examine further the transition state structures **TS1** and **TS2** resulted from the calculations. In the transition state **TS1** (Figure 5, without methanol), the distance between N-C2 is 1.56 Å, which is close to the bonding distance, meanwhile, the distance between N-H of aniline and the departing carbonyl oxygen O1' in Boc anhydride is 1.96 Å. This is within the range of hydrogen bonding and will have a stabilizing effect on transition state **TS1**.

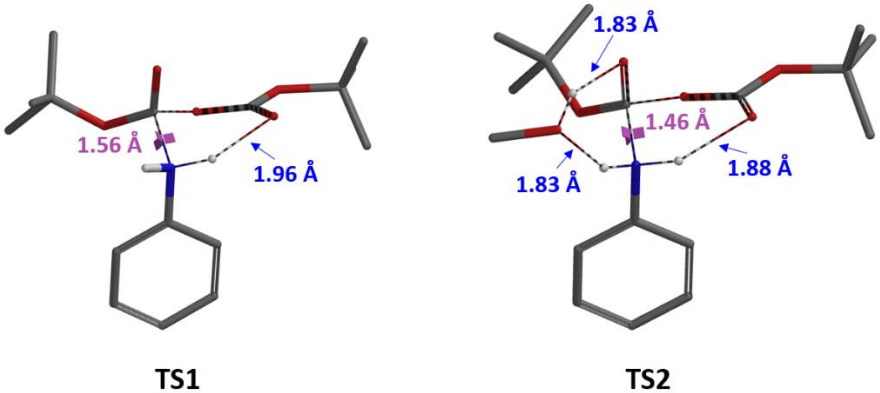


Figure 5. Transition states without methanol (**TS1**) and with methanol (**TS2**) in Boc protection of aniline

In the transition state TS2 (Figure 5, with methanol), the distance between N-C2 is 1.46 Å and already bonded, and there is the same hydrogen bond between the N-H of aniline and the departing carbonyl oxygen O1' in Boc anhydride (1.88 Å). Furthermore, the included methanol, acts like a bridge, establishes hydrogen bonding to both the carbonyl oxygen O2' in Boc anhydride and the second N-H of aniline, with similar calculated bond distances of 1.83 Å.

These supports our hypothesis that methanol introduces two additional hydrogen bonds to the transition state, further stabilizes it, lowers the activation energy, and speeds up the reaction.

Boc Protection of Secondary Aromatic Amines

In the same paper, Vilaivan reported that in contrast to primary amines, with reactions complete from < 1 minutes to 2 hr, secondary aromatic amines react much slower, and will take 24 to 48 hr to go to completion in the same alcoholic solvent (Figure 6A). This observation could be accounted for with our proposed transition state (Figure 6B, **TS3**) losing the stabilization from the hydrogen bond between NH and alcohol.

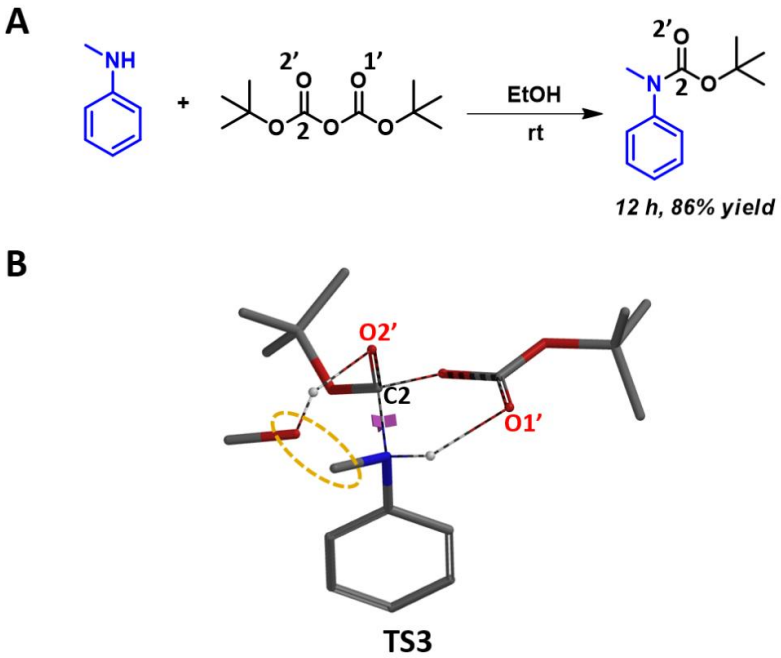


Figure 6. Boc protection of secondary aromatic amine

In summary, our QM calculations arrived with a useful model in Boc protection of primary amines, indicative that alcohol stabilizes the transition state with two additional hydrogen bonds, bridging the two reactants with a six-membered system, increasing the rate of reaction. The model could also account for the lack of acceleration with secondary amines. [3]

Building on What We Just Learned

In organic synthesis, we learned to lactamize γ-amino esters in alcohol with a catalytic amount of acetic acid (Figure 7). Why do we add acetic acid? Will it protonate the amino group, making it less nucleophilic?

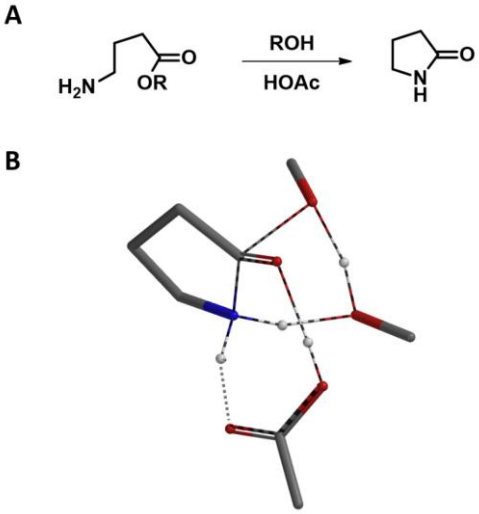


Figure 7. γ -aminoester lactamization

Return to Table of Contents

References:

[1] W. Hehre and S. Ohlinger, *A Guide to Molecular Mechanics and Quantum Chemical Calculations*. Irvine, CA, USA: Wavefunction, Inc., **2003**.
[2] T. Vilaivan, *Tetrahedron Lett.* **2006**, *47*, 6739.
[3] Our model is also consistent with the use of water or aqueous organic solvent for primary amine Boc protection, see S.V. Chankeshwara & A.K. Chakraborti, *Org. Lett.* **2006**, *8*, 3259.

Chapter 16 Ketalization with Tetramethoxysilane: QM Mechanism Investigation

Dong Pan, Zhengquan Zhou, Wenfeng Liu, Guqin Shi, Liting Dong, Tommy Lai, and John S. Wai

In organic synthesis, ketalization is a very useful strategy for protection of ketones. This is usually accomplished by treatment of the ketone with a diol in the presence of an acid catalyst. Protection of sterically congested ketones could be challenging, which usually requires high reaction temperature, long reaction time, yet still stalls at partial conversion. Even though ketalization with tetramethoxysilane was reported in the literature,[1] experimental conditions are lacking. We found that treatment of these ketones in neat tetramethoxysilane in the presence of *p*-toluenesulfonic acid could provide almost quantitatively the corresponding dimethoxy ketal (Figure 1).

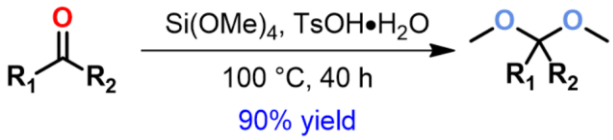


Figure 1. Synthesis of dimethoxy ketal in $Si(OMe)_4$/TsOH

Plausible Mechanism of Reaction

Since mechanism of this reaction has not been reported so far, we decided to study this with QM calculations. We reasoned that the reaction proceeds with the following mechanism (Figure 2). First, tetramethoxysilane transfers a methoxy group to the carbonyl carbon of ketone *via* a four-membered ring transition state **I**, resembling mechanism of a Wittig reaction, to give intermediate **II**, which is then protonated to provide **III**. Cleavage of the C-O bond of **III** generates oxonium intermediate **IV** and $(MeO)_3SiOH$. Then a second methoxy group is transferred from the silyl reagent to **IV** to provide the ketal product **C** and dimethyl silicate **D**.

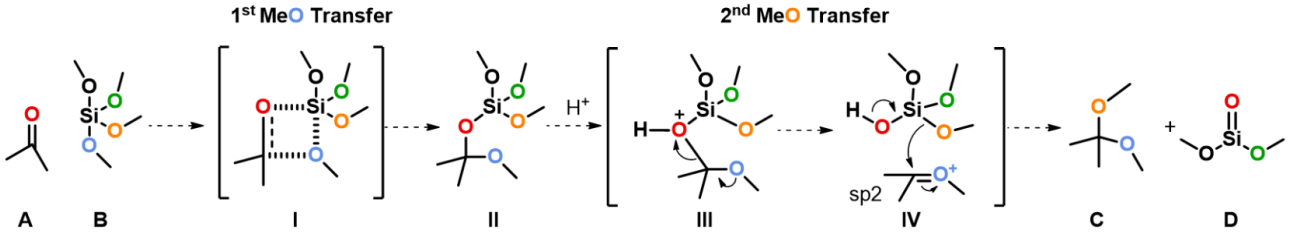


Figure 2. Plausible mechanism of ketalization with tetramethoxysilane

Modeling of the First Methoxy Group Transfer

We modeled reaction energy profile of the first methoxy group transfer from tetramethoxysilane to ketone, with Si approaching carbonyl O from 3.6 Å to 1.7 Å and silyl O approaching carbonyl C from 3.3 Å to 1.4 Å, concurrently with a step size of 0.1 Å, *via* a four-membered ring transition state **I** as shown in Figure 3[2]. The energy maximum of the reaction energy profile is approximated as the transition state of the reaction, translating to an activation energy of 20.7 kcal/mol[3,4].

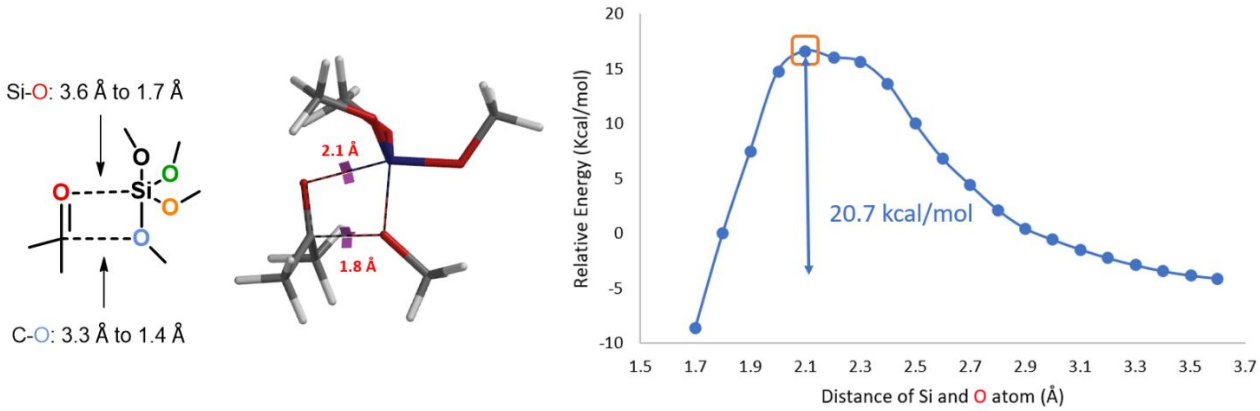


Figure 3. Transition state **I** and reaction energy profile of the first methoxy group transfer

Modeling of the Second Methoxy Group Transfer

For the next step, electrostatic potential map calculations showed that $(MeO)_3SiOH$ is more acidic than methanol (Figure 4, right), and as such $(MeO)_3SiO$ is a better leaving group. This suggests that the reaction will proceed through C-O cleavage of the O-protonated intermediate **III** (Figure 4, left).[5]

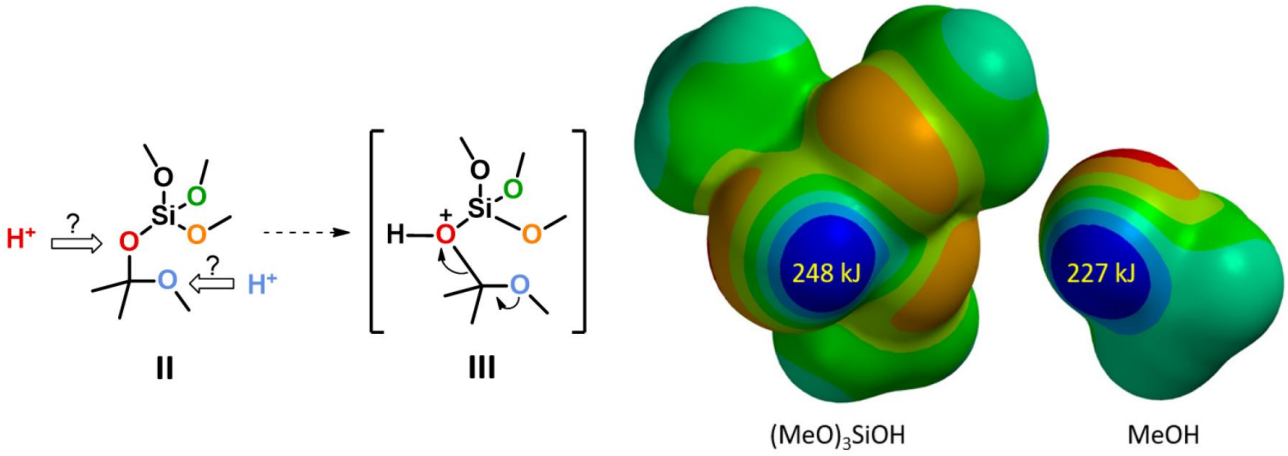


Figure 4. Electrostatic potential maps of $(MeO)_3SiOH$ and MeOH

The following step will be the transfer of the second methoxy group (Figure 5). But where does the methanol come from? Could it be from decomposition of $(MeO)_3SiOH$?

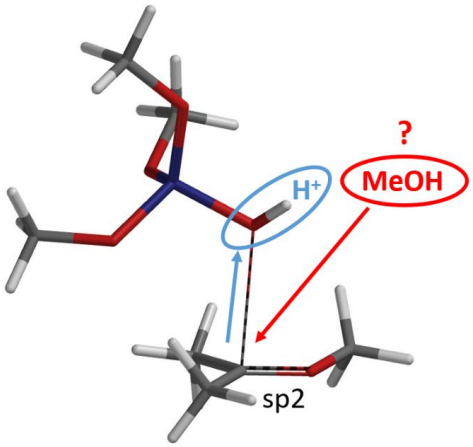


Figure 5. Second methoxy transfer

Reaction energy profile calculation showed that decomposition of $(MeO)_3SiOH$ to dimethyl silicate and methanol could have a very high energy barrier of 64 kcal/mol that is not likely to be happening under our reaction conditions (Figure 6).

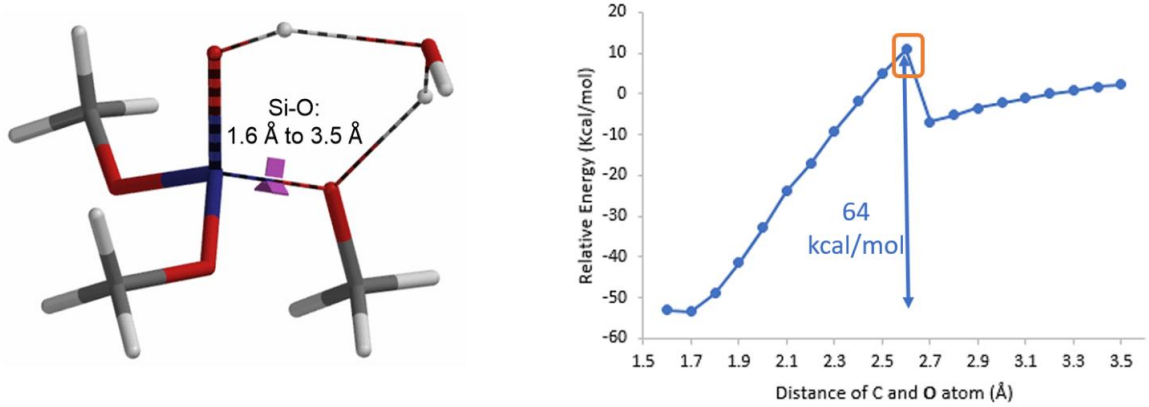


Figure 6. Transition state and reaction energy profile of $(MeO)_3SiOH$ Dissociation

So where could the methanol come from? It becomes obvious that methanol could come from reaction of tetramethoxysilane with water in the hygroscopic *p*-toluenesulfonic acid added to the reaction mixture (Figure 7).[1b]

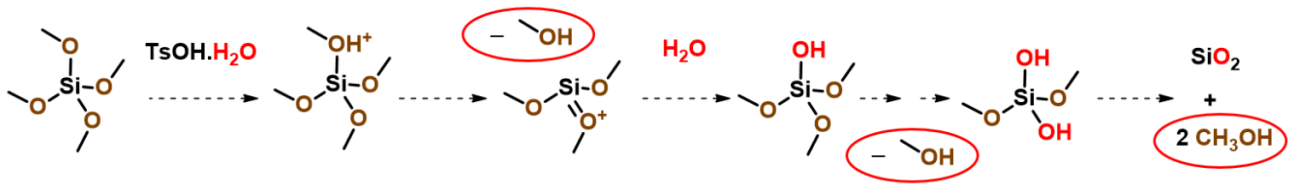


Figure 7. $TsOH \bullet H_2O$ hydrolysis of tetramethoxysilane to methanol

With the source of second MeO group settled, we are ready to model the next methoxy transfer step. We reasoned that a concerted mechanism, in which cleavage of the C-O and formation of the C-O bonds proceed concurrently, is more likely. Indeed, modeling of this reaction shows a relatively low activation energy of 8.93 kcal/mol[3] (Figure 8), indicative of a very facile transfer process.

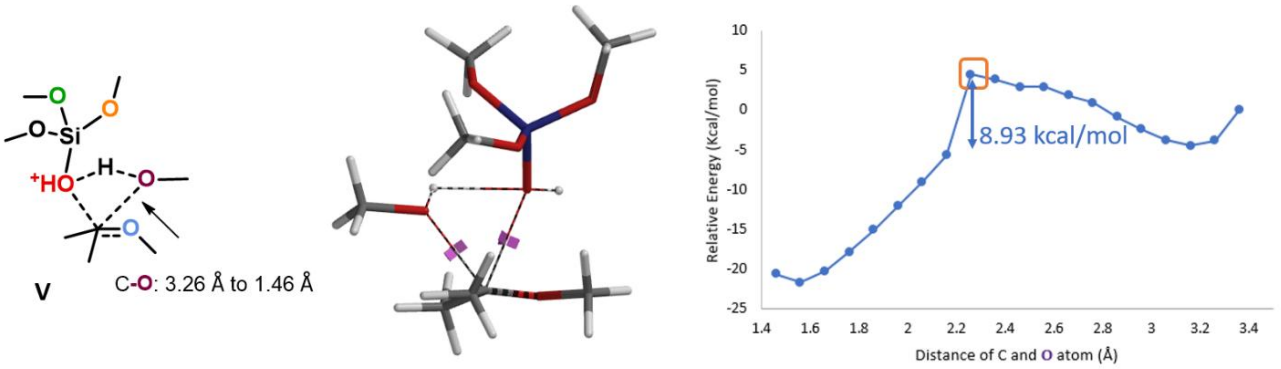


Figure 8. Concerted mechanism: transition state **V** & reaction energy profile of the second methoxy transfer

Based on the above calculation results, we updated our proposed mechanism as shown on Figure 9. Ketone **A** and tetramethoxysilane **B** react, through a four-membered ring transition state **I**, to give intermediate **II**, with an activation energy of ~20 kcal/mol. Reaction of protonated **II** and methanol *via* transition state **V** with a concerted mechanism produces ketal **C**, with a relatively low activation energy of ~9 kcal/mol. Hygroscopic *p*-toluenesulfonic acid monohydrate catalyzes the second methoxy transfer step and also provides the source of methanol, through hydrolysis of tetramethoxysilane.

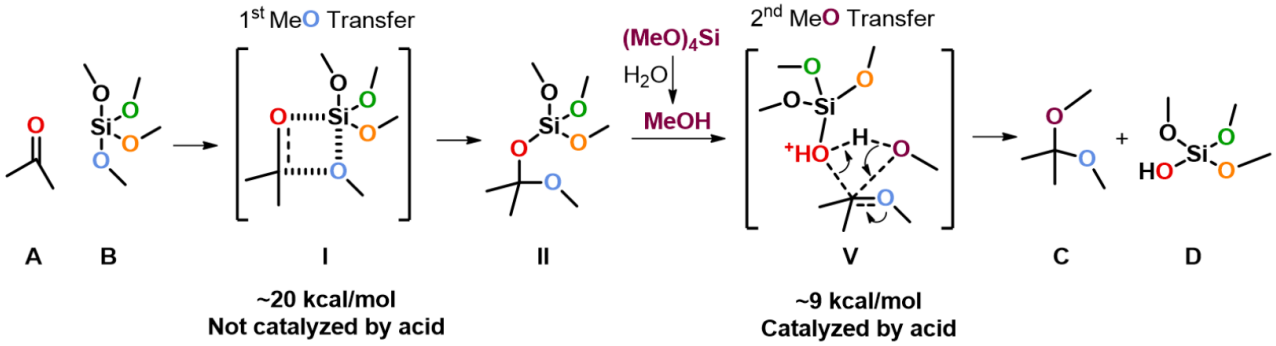


Figure 9. Mechanism for ketalization with tetramethoxysilane: QM analysis summary

Experimental Procedure

Reaction conditions for ketalization with tetramethoxysilane is not available from the literature, below are conditions we identified:

A mixture of ketone **A** (1.0 mmol), tetramethoxysilane (2.5 mL), and TsOH • H_2O (0.04 mmol) was stirred under an atmosphere of nitrogen for 40 h at 100 °C. The product mixture was diluted with ethyl acetate and washed with brine (3 times). The organic layer was dried and concentrated under vacuum. The residue was purified by column chromatography on silica gel to provide the ketal product **C**.

The resultant ketals could be readily hydrolyzed to corresponding ketones under very mild conditions (TFA/$CHCl_3$/H_2O, 15/15/1 volume, rt, 1 h).

Return to Table of Contents

References:

[1] a) Nasarow *et al. Zhurnal Obshchei Khimii,* **1959**, *29*, 106. b) Sakurai, H. and Love, J.A., "Silicon (IV) methoxide" in *Encyclopedia of Reagents for Organic Synthesis*, **2009**, John Wiley & Sons.
[2] *Spartan '18 Tutorial and User's Guide* (**2019**). Irvine, CA, USA: Wavefunction, Inc.
[3] Reactions with activation energies less than 10 kcal/mol can usually proceed at room temperature or lower. For those with activation energies between 10 and 20 kcal/mol, they generally require heating. Reaction with activation energy >20 kcal/mol usually requires higher reaction temperature and longer reaction time.
[4] Attempts to model the first methoxy transfer with ketone carbonyl oxygen protonated were futile.
[5] Since $(MeO)_3SiO$ is a better leaving group than MeO, we reasoned that this step proceeds through protonated intermediate **III**. This is analogous to what we observed with aminal.

Chapter 17 Predicting Regioselectivity of Electrophilic Halogenation Reactions

Zhou Xu, Guqin Shi, Yongsheng Chen, John S. Wai

Chapter 2, "Application of HOMO in Electrophilic Reactions", left us with a unique substrate with HOMO lobes very close in size at C6 and C7 positions, and no significant difference in the ^{13}C chemical shift. This is consistent with the electron donating characteristic of the amide and ether functionalities on the phenyl ring, leading us to expect a mixture of C6 and C7 brominating products. But experimentally only C7 bromide was obtained. How could we account for this?

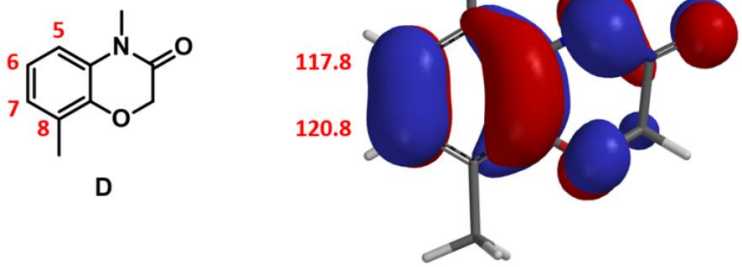


Figure 1: Schematic of HOMO & ^{13}C NMR of compound **D**

Predicting Regioselectivity of Electrophilic Halogenation by Calculation of Reaction Activation Energy

From Chapter 9, "Activation Energy Estimation", we learned the use of calculated activation energy to compare two competing reaction paths. Let's do this first.

To save computational time, we substituted NCS for NBS. The activation energies calculated for halogenation at C6 and C7 positions are 36.21 and 34.08 kcal/mol, respectively (Figure 2), suggesting that the reaction occurs more readily at C7 position.

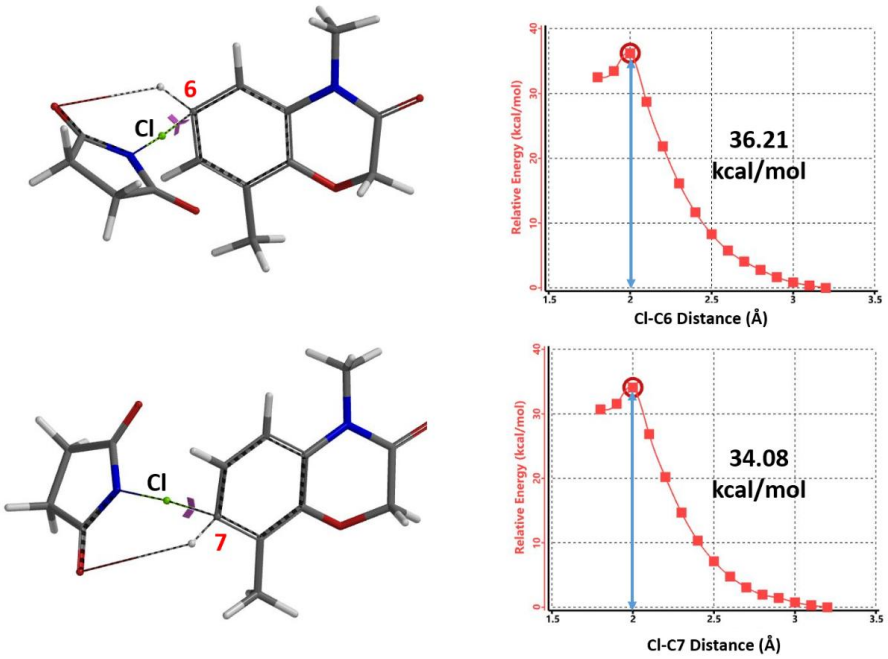


Figure 2: Calculation of halogenation activation energy at C6 and C7

Based on Arrhenius equation (Figure 3), an activation energy difference of 2.13 kcal/mol translates to a ratio of 37:1 at 25 °C, accounting for why only C7 brominated product is observed.

$$k = Ae^{\frac{-E_a}{RT}}$$

$$\frac{[C7]}{[C6]} = \frac{k_{C7}}{k_{C6}} = \frac{Ae^{\frac{-\Delta E_{C7}}{RT}}}{Ae^{\frac{-\Delta E_{C6}}{RT}}} = e^{\frac{\Delta E_{C6} - \Delta E_{C7}}{RT}}$$

Figure 3: Calculating product ratio with Arrhenius equation

In general, activation energy calculations require relative accurate model of the reaction and longer computation time[1,2]. Do we have other more convenient, faster methods to predict accurately regioselectivity of such halogenation reactions?

Predicting Regioselectivity of Electrophilic Halogenation by Calculation of Relative Energy of Halonium ion

In 2018, Jørgensen *et al*. reported in *Chem. Sci.*[3] the development of RegioSQM method for rapid prediction of the regioselectivity of electrophilic aromatic substitution reactions. RegioSQM works by protonating all aromatic C–H carbon atoms and identifying those with the lowest free energies in chloroform using the PM3 semiempirical method as the most nucleophilic center (Figure 4). When RegioSQM was used to predict regioselectivity for compound **D** (Figure 1), it predicted a mixture of halogenation products.

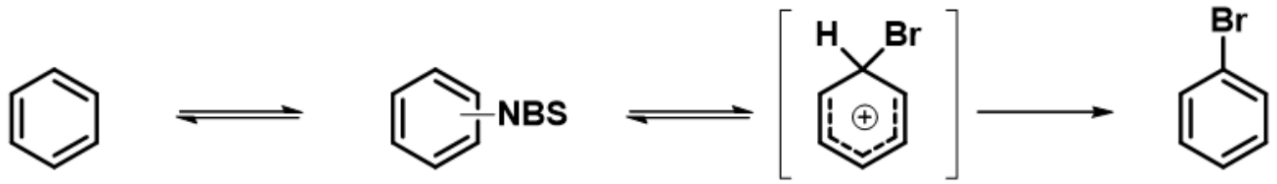
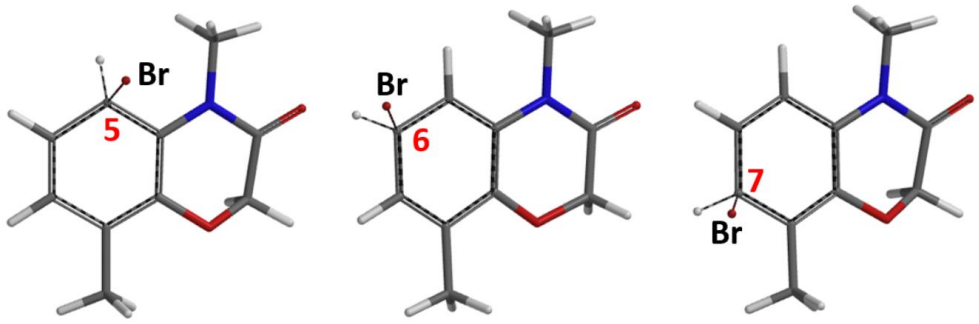


Figure 4: Classical halogenation reaction mechanism

We reasoned that Jørgensen's method could be improved by calculating Equilibrium Geometry of the relevant halonium ions with a more accurate QM method (with DFT ωB97X-D 6-31G* [4]) and comparing their relative energies to predict selectivity of the halogenation.

Relative energies from Equilibrium Geometry calculation of the C5, C6, and C7 brominium ions reveal that the C7 one has the lowest energy, in both polar and nonpolar solvents, suggesting that the brominated product at position C7 would be the major product (Figure 5). Consistent with the observation.

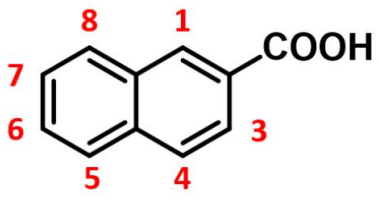


Relative Energy (kcal/mol)	5-	6-	7-
nonpolar solvent	0.00	-4.21	-5.64
polar solvent	0.00	-4.60	-5.79

Figure 5: Equilibrium geometry and relative energy of brominium ion at positions C5, C6, and C7

In summary, for electrophilic halogenation, we could use activation energy calculations to find transition states, and predict regioselectivity and product ratio. However, such calculations are time-consuming. A faster alternative, with comparable accuracy, is calculating for Equilibrium Geometry and comparing the relative energies of halonium ions for prediction.

Building on What We Just Learned



A

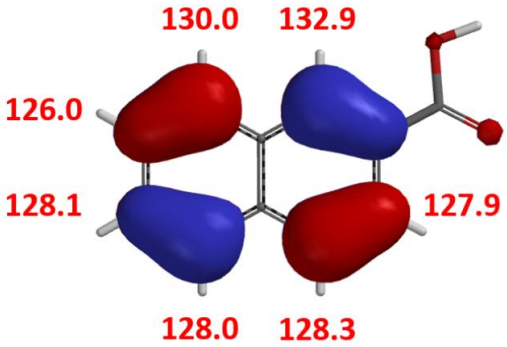
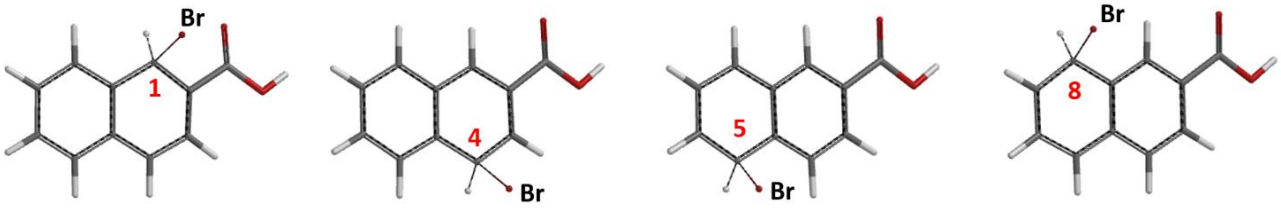


Figure 6: Calculated HOMO and ^{13}C NMR of 2-naphthoic acid (**A**)

For predicting regioselectivity of the bromination of 2-naphthoic acid (**A**), we found that the HOMO lobe sizes of four out of seven sites are very similar, and the chemical shifts of the ^{13}C NMR are not significantly different. Relative energies from Equilibrium Geometry calculation of the bromonium structures are tabulated below. The major product is……?



Relative Energy (kcal/mol)	1-	4-	5-	8-
nonpolar Solvent	4.03	1.37	0.00	1.22
polar solvent	4.27	1.62	0.00	1.21

Figure 7: Equilibrium conformation and relative energy of brominated cations at C1, C4, C5, and C8

Return to Table of Contents

References:

[1] R. Hoffmann, J.P. Malrieu, *Angew Chem Int Ed Engl*., **2020**, *59*, 12590 (Simulation *vs*. Understanding, "*To predict is not to explain*.")

[2] Structurally, halogenating reagent is an integral part of the transition state of halogenation. The successful use of relative energy difference between halonium ion to predict regioselectivity of halogenation does not necessary mean these simplified structures are the transition states.

[3] J.C. Kromann, J.H. Jensen, M. Kruszyk, M. Jessing, M. Jørgensen, *Chem. Sci*., **2018**, *9*, 660. RegioSQM http://www.regiosqm.org

[4] *Spartan'18 Tutorial and User's Guide* (**2019**). Irvine, CA, USA: Wavefunction, Inc.

Chapter 18 Integrating QM HOMO/LUMO/^{13}C NMR Calculations

Chunrui Wu, Guqin Shi, Qiuyue Wang, Dong Pan, Yongsheng Chen, John S. Wai

In previous chapters, we discussed the application of HOMO and LUMO for predicting regioselectivity of S_EAr/S_NAr reactions of heterocyclic aromatic substrates, and calculated ^{13}C NMR for structure assignments. Below is an example on how these QM calculations are integrated into our routine chemistry workflow[1].

For the synthesis of target molecule **4**, we envisioned a C-2 selective bromination of starting material imidazo[1,2-a]pyrazine **1** to provide the bromo intermediate **2** (Figure 1), followed by a bromide-Grignard exchange and addition of acetaldehyde to afford intermediate compound **3**, then a selective S_NAr reaction of **3** with methyl amine at C-6. Next, we evaluated feasibility of the sequence with QM.

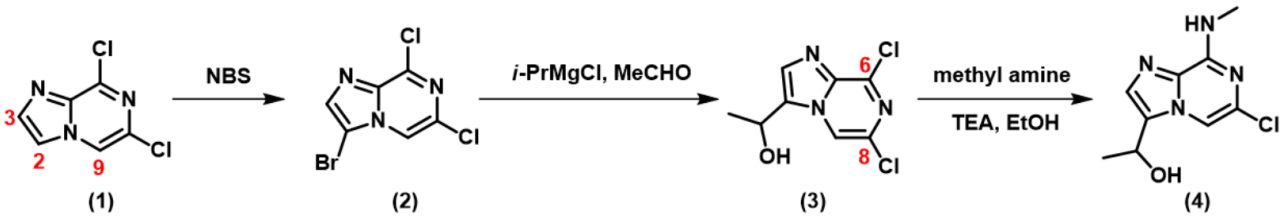


Figure 1: Designed synthetic route of target molecule **4**

Predicting Regioselectivity of Bromination

From Chapter 2 "Application of HOMO in Electrophilic Reactions", we learned that regioselectivity of S_EAr halogenation[2,3] could be reliably predicted using calculated HOMO, and ^{13}C NMR.

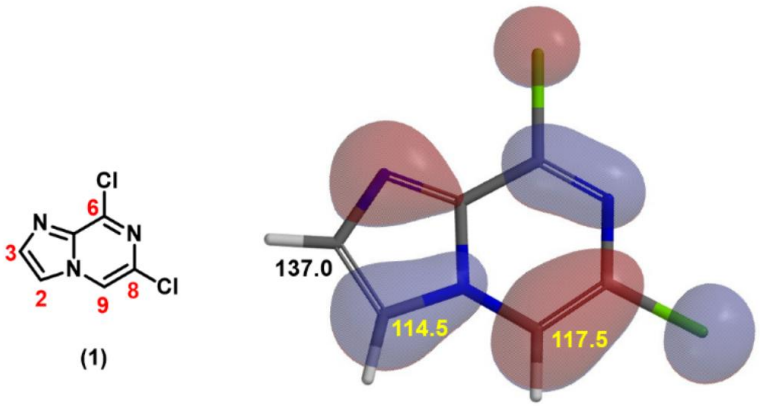
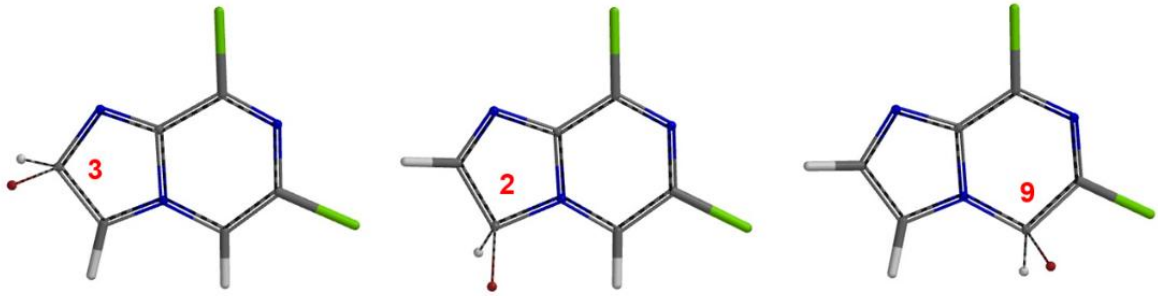


Figure 2: HOMO and calculated ^{13}C chemical shift of compound **1**

HOMO of compound **1** shows no HOMO lobe at C-3 position, while a big one over C-2 position (Figure 1), suggesting a preference for C-2 regioselectivity of halogenation on the imidazole ring. Furthermore, calculated ^{13}C chemical shift of C-2 (114.5 ppm) and C-3 (137.0 ppm) showed significant difference, with a lower chemical shift on C-2, indicating a higher electron density on the carbon and C-2 halogenation to dominate. Both HOMO and ^{13}C NMR point to selective reaction at C-2 over C-3.

On the other hand, HOMO lobe on C-9 position of compound **1** is comparable in size to C-2 of the imidazole ring, and ^{13}C NMR chemical shifts only differ by 3 ppm – will we run the risk of having competing halogenation on C-2 and C-9? In general, as imidazole is electron-rich while pyrazine is electron-deficient, electrophilic substitution reactions are more likely to proceed on the more electron-rich imidazole. Do we have additional QM calculation tools that confer us with a higher level of confidence on the desired regio-selectivity before we start chemistry in the lab?

In Chapter 17 "Predicting Regioselectivity of Electrophilic Halogenation Reactions", we introduced the use of relative energy of halonium ion as the third QM parameter for such prediction. For substrate **1**, calculated relative energy of brominium ion (Equilibrium Geometry) for the three potential substitutable sites, C-2, C-3 and C-9, are tabulated in Figure 3. The relative energy of **2B** is much lower than that of **2C** and **2A**, suggesting the bromination should occur only at C-2 position. And indeed, this is what we observed. Although intermediate **2** has multiple halogen atoms, we reasoned that the rate of Br-Mg exchange is much faster than that of Cl-Mg exchange, and the formation of intermediate **3** will be highly selective.



2A 2B 2C

Relative Energy (kcal/mol)	3-	2-	9-
nonpolar solvent	0.00	-21.09	-7.96
polar solvent	0.00	-20.75	-7.10

Figure 3: Equilibrium geometry and relative energy of C-3, C-2, C-9 brominium ions

Prediction of nucleophilic substitution reaction sites using LUMO

The next question is on the selectivity of the S_NAr reaction of intermediate **3**. Which chloride will be preferentially displaced by the amine? In Chapter 1 "Application of LUMO in Nucleophilic Reactions", we learned that nucleophile will preferentially attack the carbon with larger LUMO lobe

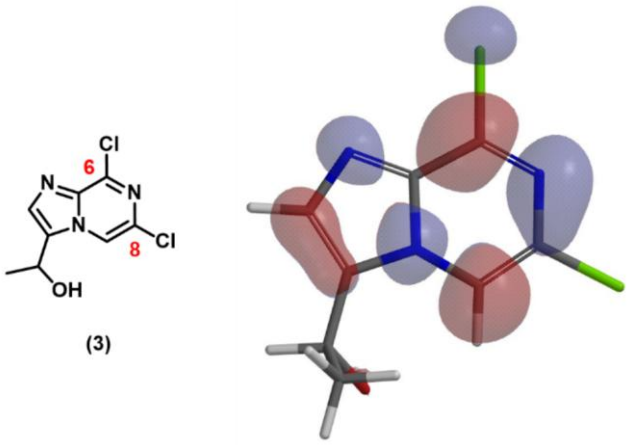
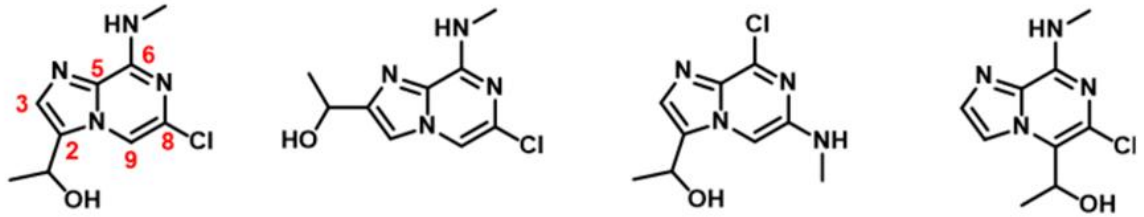


Figure 4: LUMO of intermediate **3**

LUMO of intermediate **3** shows significant lobe over C-6 and close to none over C-8, suggesting that the nucleophile would react preferentially at C-6 position. This is what we observed experimentally.

QM Assisted Structure Assignment

In Chapter 11 "QM Assisted Structure Assignment with ^{13}C NMR Calculation", we learned to use Quantum Mechanics calculated *versus* experimental ^{13}C NMR chemical shift for structure assignment. In this example, we applied such analysis to be assured that we indeed obtained target **4** (Figure 5), not any of the other potential isomers, **4A**, **4B** or **4C**[4,5]. The calculated ^{13}C NMR data of structure **4** matched the experimental data the best *versus* the other isomeric structures, enabling us to assign the structure of compound **4** with high confidence.



	^{13}C Chemical Shift (ppm)				
	Experimental	Calculated (4)	Calculated (4A)	Calculated (4B)	Calculated (4C)
C2	128	128	112	128	117
C3	132	131	150	133	133
C5	136	134	133	138	134
C6	150	150	148	145	149
C8	133	136	137	146	134
C9	107	109	108	97	119

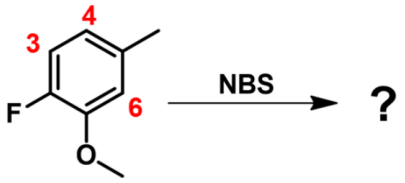
Figure 5: Experimental ^{13}C NMR of the final product 4 and the calculated data of structural isomers (Spartan'18)

(Note: black indicates $\Delta \leq 2$ ppm; blue: $2 < \Delta \leq 10$ ppm; red: $\Delta > 10$ ppm.)

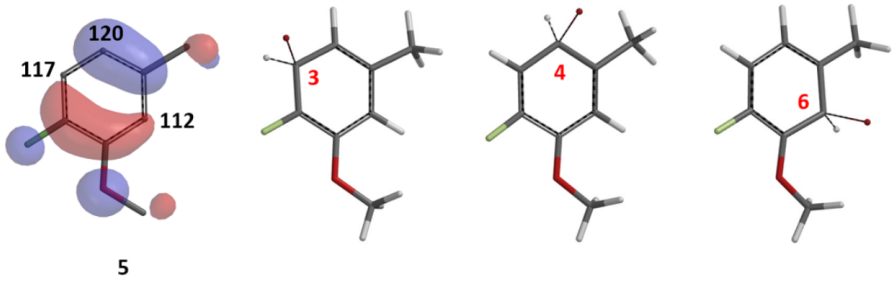
In summary, we integrate multiple QM applications to evaluate feasibility of synthetic routes and compare calculated *versus* experimental ^{13}C NMR for accurate structure assignment of the final targets. **Quantum Mechanics give us Insight and Numbers**, changing the ways we analyze chemistry retrospectively and prospectively.[6]

Building on What We Just Learned

For bromination of compound **5** with NBS, let's look at QM calculated HOMO, ^{13}C NMR and relative energy of the three potential brominium ions (Equilibrium Geometry) on Figure 6. We could predict with high confidence which isomer will be the only product produced.



(5)



5

Relative Energy (kcal/mol)	3-	4-	6-
nonpolar solvent	10.61	-7.77	0.00
polar solvent	10.48	-7.91	0.00

Figure 6: HOMO, ^{13}C NMR of compound **5** and eqm geometry-relative energies of C-3, C-4, and C-6 brominium ions.

Return to Table of Contents

References:

[1] W. Hehre and S. Ohlinger, *A Guide to Molecular Mechanics and Quantum Chemical Calculations*. Irvine, CA, USA: Wavefunction, Inc., **2003**.

[2] F.A. Carey & R.J. Sundberg. *Advanced Organic Chemistry Part A: Structure and Mechanisms*. New York, NY, USA: Springer Science + Business Media, LLC., **2000**; pp 551-557.

[3] M. Kruszyk, M. Jessing, J.L. Kristensen, M. Jorgensen, *J. Org. Chem.* **2016**, *81*, 5128.

[4] M. Lodewyk, M. Siebert, D. Tantillo, *Chem. Rev*. **2012**, *112*, 1839.

[5] W. Hehre, P. Klunzinger, B. Deppmeier, A. Driessen, N. Uchida, M. Hashimoto, E. Fukushi, Y. Takata, *J. Nat. Prod*. **2019**, *82*, 2299.

[6] F. Neese, M. Atanasov, G. Bistoni, D. Maganas, S. Ye, *J. Am. Chem. Soc*., **2019**, *141*, 2814.

Chapter 19 *Tele*-Substitution of 2,3-Dichloropyrazine

Liting Dong, Zhengquan Zhou, Tommy Lai, Dong Pan, Guqin Shi, Yongsheng Chen, John S. Wai

Chapter 7, " Application of LUMO Analysis in Nucleophilic Reactions (Part II)", left us with a surprising observation reported by McDonald *et al*. in 2006[1]: 2,3-dichloro-pyrazine reacts with lithium dithiane to provide a single product in which the hydrogen atom at C-6 of pyrazine is substituted by the nucleophile and the chloro group at C-3 is replaced by a hydrogen atom. In contrast to the selective C-2 chloride displacement usually observed for reactions between 2,3-dichloropyrazine and other nucleophiles (e.g. amines, enolate anions, etc.) (Figure 1).[2]

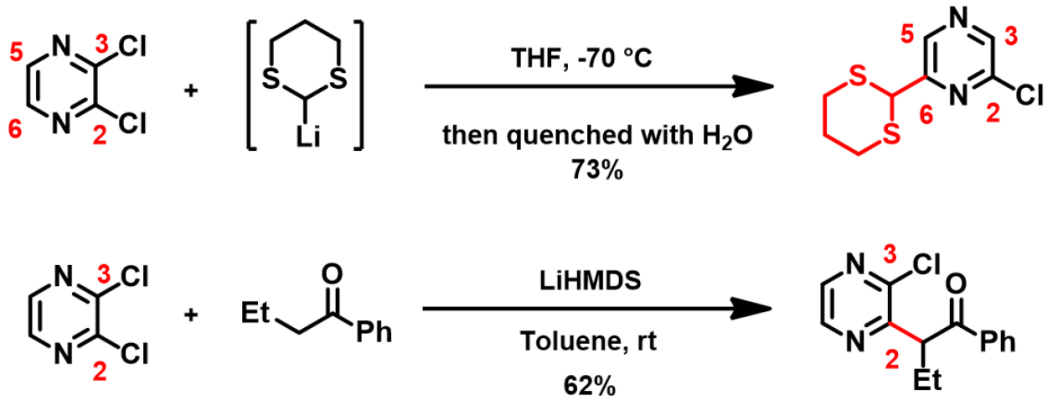


Figure 1. Nucleophilic substitution of 2, 3-dichloropyrazine at C2 and C6

Why does lithiated dithiane add to 2,3-dichloropyrazine at C-6 position? Why would the chloro group at C-3 be selectively displaced?

LUMO & LUMO Map Analyses

Since 2,3-dichloropyrazine is the electrophile in the reaction, let's look at its LUMO and LUMO map first (Figure 2). This dichloropyrazine has its LUMO lobes centered mainly between C-2 and C-3, C-5 and C-6 (colored in red), not on the carbons themselves, while LUMO+1, 0.65 eV higher in energy than LUMO, has significant lobes centered on C-2, C-3, C-5 & C-6, and are roughly equal in sizes.

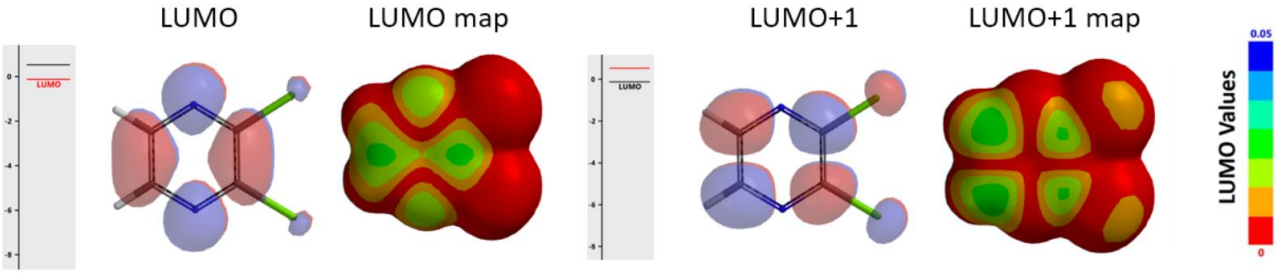


Figure 2. LUMO/LUMO map (left) and LUMO+1/LUMO+1 map (right) of 2,3-dichloropyrazine

LUMO+1 map reveals that the LUMO+1 lobes on C-5 and C-6 are more exposed beyond the surface of the molecule than the lobes on C-2 and C-3 (due to the attached chloro groups), and LUMO ones, suggesting a higher accessibility for nucleophilic attack at C-5 and C-6. It is more likely that 2,3-dichloropyrazine uses LUMO+1 for the addition.

Reversibility and Irreversibility of Interactions

Based on the LUMO+1/LUMO+1 map analyses, we reason that C-6 of 2,3-dichloropyrazine is more accessible than C-2 for nucleophilic interaction with lithium enolate (Figure 3). However, such reaction at the C-6 is reversible, while addition at the C-2 position could lead to displacement of the chloro group, an irreversible process.

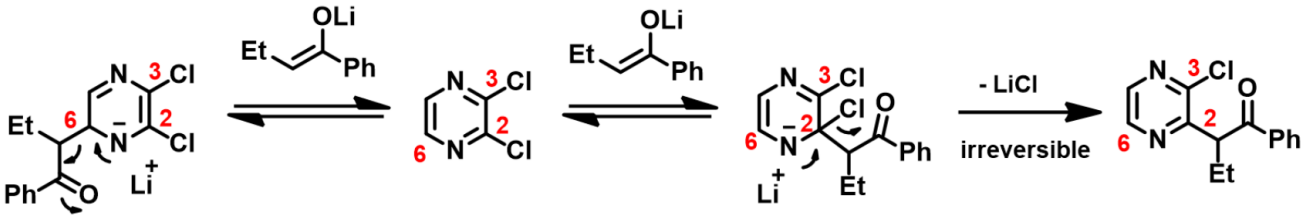


Figure 3. Nucleophilic substitution of 2, 3-dichloropyrazine with lithium enolate

When lithiated dithiane is used in the reaction (Figure 4), since C-6 is more accessible than C-2 and formation of this carbon-carbon bond is irreversible, preferential addition of dithiane occurs at the C-6.

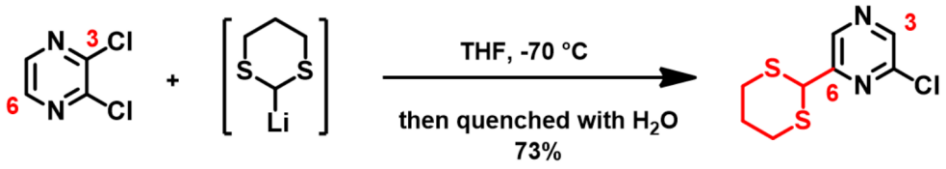


Figure 4. Nucleophilic substitution of 2,3-dichloropyrazine with lithiated dithiane

Next question, why is the chloro group at C-3 being selectively displaced?

Mechanism for Chloride Elimination

We propose the reaction mechanism shown in Figure 5. First, 2,3-dichloropyrazine reacts with lithiated dithiane irreversibly to provide intermediate **I** (along with its resonant structures **I'** and **I"**); after quenching with water, the anionic intermediate acquires a proton to form compound **B**, which potentially is an equilibrium mixture of **B1**, **B2**, **B3** and **B4** ; Intermediate **B** eliminates a molecule of HCl, rearomatizes to provide product **C**, analogous to the mechanism proposed by McDonald *et al* for 2,6-dichloropyrazine *tele*-substitution.[1]

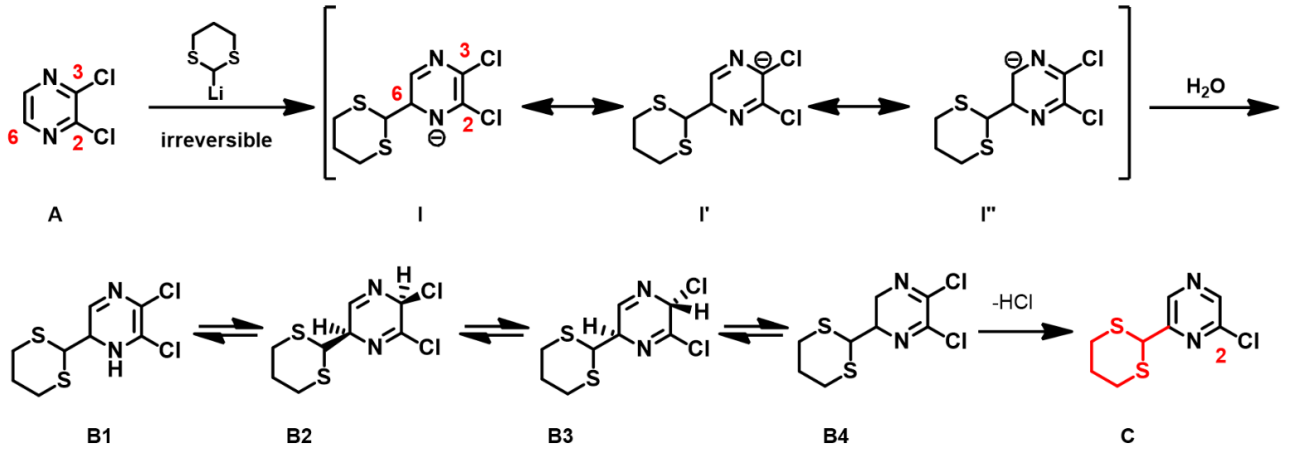
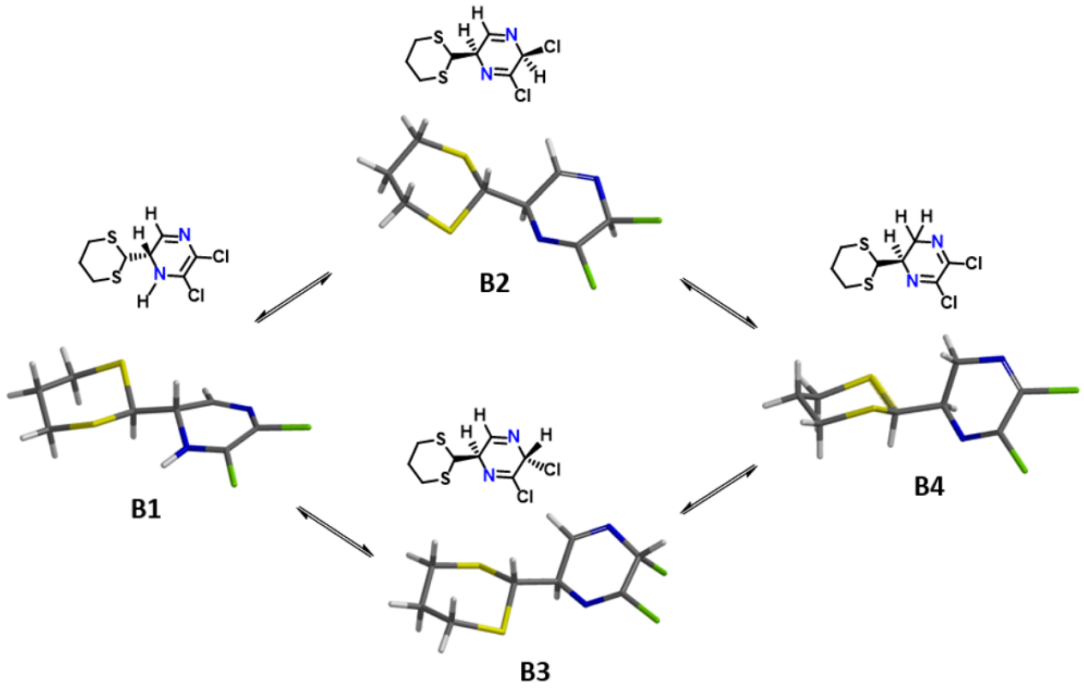


Figure 5. Proposed mechanism: Reaction of 2,3-dichloropyrazine with lithiated dithiane

Next question becomes "Elimination of HCl proceeds via **B1**, **B2**, **B3** or **B4**?" Density Functional Theory equilibrium geometry calculations (polar solvent) suggests an equilibrium distribution in the ratio of 0 : 0.8 : 7.2 : 91.9% for **B1**, **B2**, **B3** and **B4**, respectively (Figure 6). **B1** has the highest energy and its distribution is almost zero. Isomer **B4** has the lowest energy of -6.53 kcal/mol and the highest distribution percentage, presumably due to stabilization from the conjugated N=C-C=N system. But since both of its chloro groups are attached to sp^2 hybridized carbon atoms, they are not amendable to undergo HCl elimination. **B2** and **B3**, with their C=N bonds not in conjugation, are higher in energy than **B4**. How could we evaluate them for subsequent elimination?



Intermediate	Relative Energy (kcal/mol)	Boltzmann Weights[3]
B1	0.00	0.000
B2	-3.72	0.008
B3	-5.02	0.072
B4	-6.63	0.919

Figure 6. The four possible structures of **B**, their relative energies and distribution.

Reaction Energy Profile Calculation

When water is added to quench the reaction, hydroxide is generated. Shown on Figure 7 is the reaction energy profile we modeled for **B-3** HCl elimination with hydroxide. The C-H bond distance at C-6 increased from 1.096 to 3.096 Å with a step size of 0.2 Å. Modeling result suggests a concerted departure of the pseudo-axial proton at C-6 and the pseudo-axial chloro group at C-3, and the chloride interacts with the water molecule formed. C-3 and C-6 change from sp^3 to sp^2 hybridization, and the molecule aromatizes to provide compound **C**. Calculation also indicates that this elimination has no activation energy barrier to overcome, and should be a very fast process.

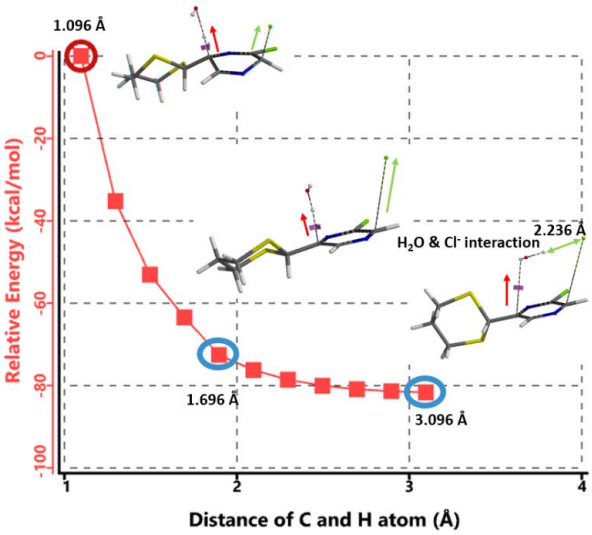


Figure 7. Modeling of **B3** HCl elimination with hydroxide anion

For intermediate **B2**, with the departing chloro in pseudo-equatorial position (Figure 8), reaction energy profile calculation does not arrive at elimination product and suggests a high activation barrier, making this pathway not likely.

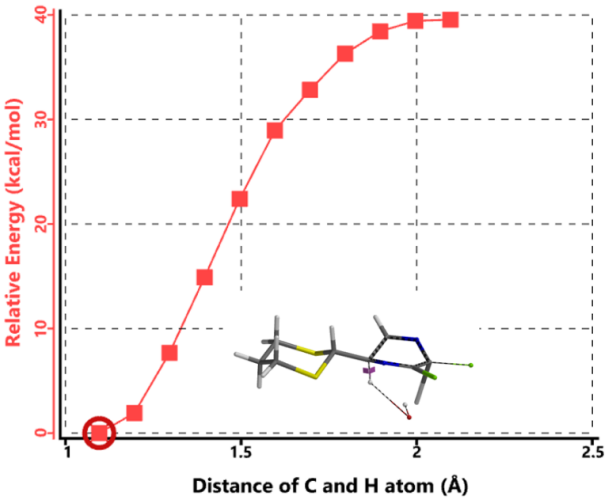
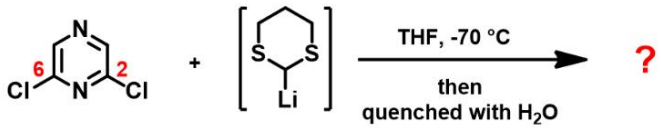


Figure 8. Modeling of B2 HCl elimination with hydroxide anion

In summary, LUMO/LUMO map analyses of 2,3-dichloropyrazine, irreversible nature of lithiated dithiane addition, and reaction energy profile calculation of the HCl elimination step, provide us with satisfactory explanations to the "surprising" *tele*-substitution product observed by McDonald *et al*.

Building on What We Just Learned

In the same paper, McDonald *et al* also reported that treatment of 2,6-dichloropyrazine with lithiated dithiane led to another *tele*-substitution product.



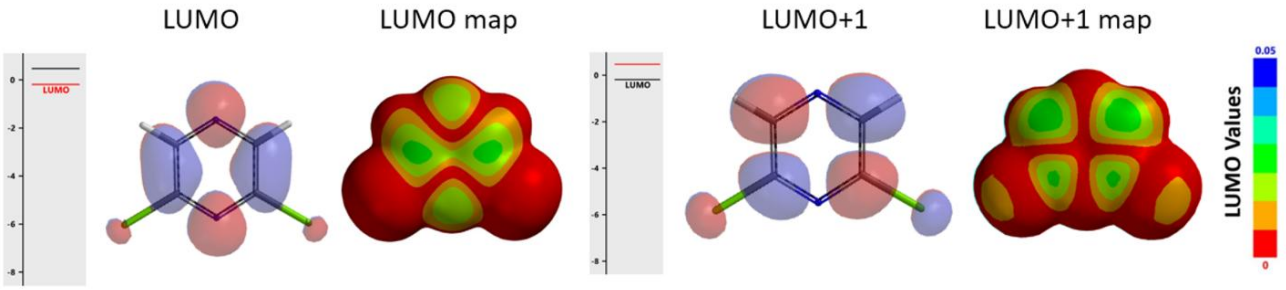


Figure 9. LUMO/LUMO+1 and LUMO/LUMO+1 map of 2,6-dichloropyrazine

Shown in Figure 9 are LUMO/LUMO+1 and LUMO/LUMO+1 map of 2,6-dichloropyrazine. How will the addition and elimination proceed? What could be the structure of the product identified?

Return to Table of Contents

References:

[1] J.E. Torr, J.M. Large, P.N. Horton, M.B Hursthouse, E. McDonald, *Tetrahedron Lett*. **2006**, *47*, 31.
[2] J.A. Joule & K. Mills. *Heterocyclic Chemistry 5th Ed.* Chichester, West Sussex, UK: Blackwell Publishing Ltd., **2010**; pp 256-7.
[3] Boltzmann Weights: Distribution of equilibrating isomers at thermodynamic equilibrium, based on difference in their energies.

Chapter 20 Analysis of Ambident Reactivity of Dimethyl Carbonate

Shouliang Wang, Jian Wang, Qiuyue Wang, Zhong Zheng, Guqin Shi, Yongsheng Chen, John S. Wai

Dimethyl carbonate (DMC) is a low toxicity, fast biodegradability chemical feedstock and important organic reagent. DMC, structurally consisting of a carbonyl and two methoxy functional groups, is an excellent substitute for highly toxic phosgene and dimethyl sulfate in acylation and methylation, respectively.

Hard-Soft-Acid-Base (HSAB) principle was introduced by Pearson in 1963[1] and is being widely used to explain stability of reaction products and mechanisms, where "hard" refers to higher charge density and smaller radius species (ions, atoms, molecules) that are weakly polarizable with greater polarity, while "soft" means lower charge density and larger radius species that are strongly polarizable with lower polarity. HSAB principle qualitatively explains reactivity patterns with "hard likes hard, soft likes soft" preference.

In 2005, Tundo *et al.* published their observations on how ambident DMC reacts differently with phenols, thiophenol, and aniline under various reaction conditions[2]. However, their application of the HSAB principle failed to explain the experimental results in a consistent manner, creating more confusion. In 2011, Mayr *et al.* advocated "Farewell to the HSAB Treatment of Ambident Reactivity"[3] and revealed that thermodynamic/kinetic control could better account for reactivity of organic ambident reagents observed (Figure 1).

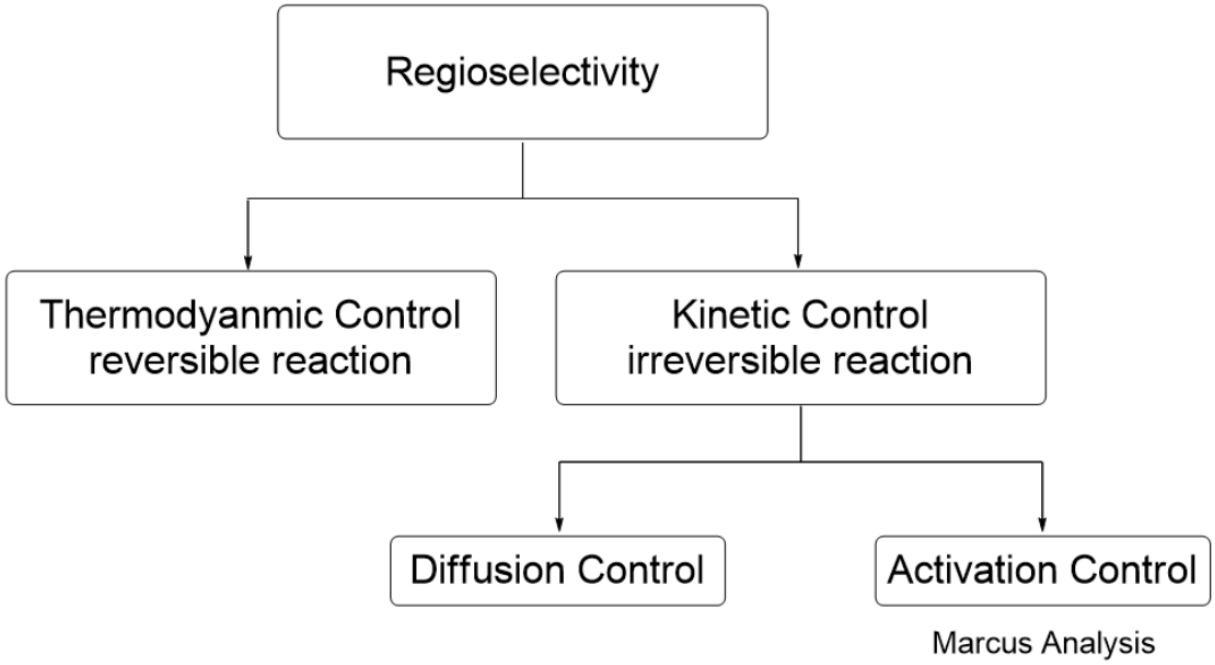


Figure 1: A systematic approach to ambident reactivity[3]

Inspired by Mayr *et al*., we found that correlating the LUMO & LUMO+1 characteristics of DMC and p*K*a values of the substrates (phenols, thiophenol, and aniline) is indeed sufficient to understand the complex ambident reactivity patterns of DMC, avoiding the unnecessary confusion with HSAB analysis[4].

As an ambident electrophilic reagent, DMC can react differently with nucleophilic reagents to produce the carboxymethylation or methylation product or both (Figure 2). When the carbonyl carbon of DMC is attacked by the nucleophile, the acyl-oxygen bond is broken, forming an acylation product, accompanied by the formation of CH_3O^-; when the methoxy carbon of DMC is attacked by the nucleophile, the methyl-oxygen bond is broken, forming a methylation product, accompanied by the formation of CH_3O^- and CO_2. Below is our analysis of DMC reactivity patterns with various phenols, thiophenol, and aniline.

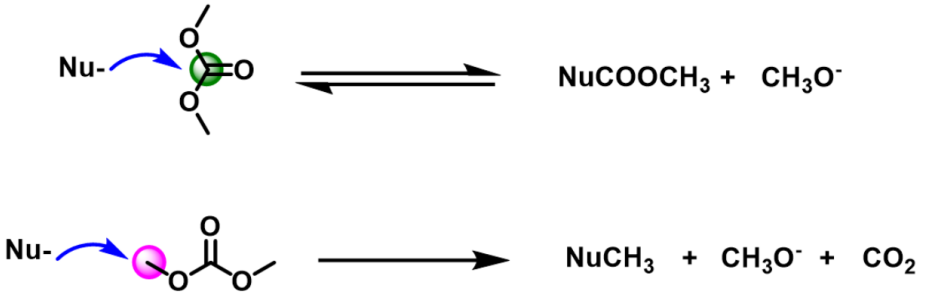


Figure 2: Ambident reactivity of dimethyl carbonate

Reaction of Phenolic Compounds with DMC

Shown on Table 1 are the results reported by Tundo *et al*. on how various phenols react with DMC to give different relative percentage of methylation *vs.* carboxymethylation products, along with p*K*a values of various phenols we tabulated from the Reich and Bordwell p*K*a compilation[5]. For *p*-methoxyphenol, acylation product is predominant; while for *p*-cyanophenol, only the methylation product is formed.

Reaction of Phenols with DMC in the presence of K_2CO_3[a]

Entry	Phenol	pKa [5]	% Conversion	% Methylation	% Carboxy-methylation
1	*p-Methoxyphenol*	19.1	7.0	23	77
2	*p-Cresol*	18.9	36.3	39	61
3	Phenol	18.0	27.7	72	28
4	*p-Chlorophenol*	16.7	62.0	98	2
5	*p-Cyanophenol*[b]	13.2	100	100	0

[a]DMC reflux temp, 90 °C; molar ratio phenol/DMC/K_2CO_3 = 1.0/40/1.2. Conversions after 53 h.

[b]After 30 h, 82% conversion, methylation only.

Table 1: Reaction of phenols with different substituents with DMC[2]

Since DMC is an electrophile, we calculated for its LUMO and LUMO map first[6]. See Chapter 1 for LUMO analysis of electrophiles.

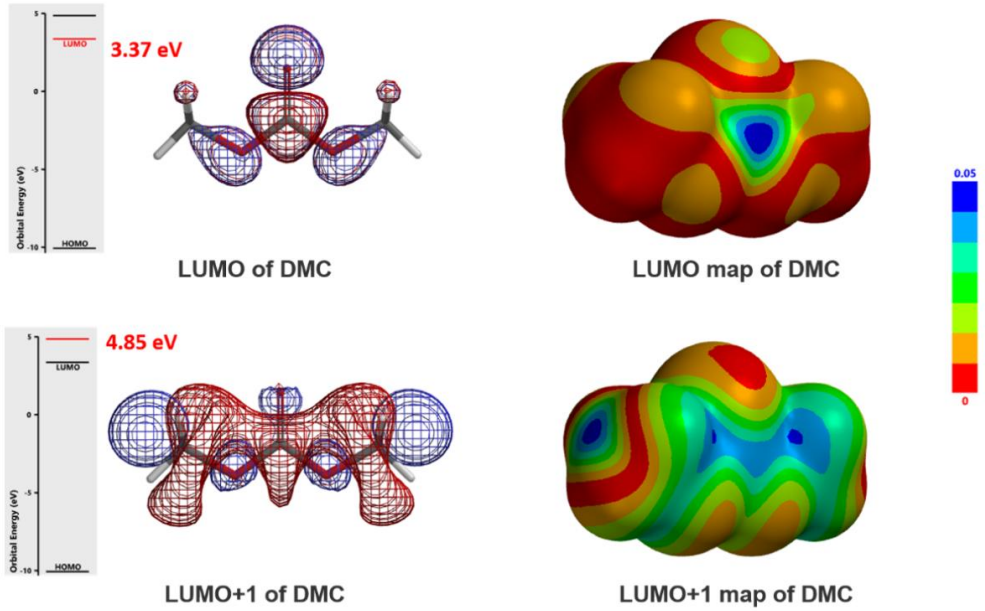


Figure 3: DMC's LUMO/LUMO+1 and LUMO/LUMO+1 map

LUMO (3.37 eV) and LUMO map of DMC show accessible reaction site on carbonyl carbon, while LUMO+1 (4.85 eV) and LUMO+1 map show the accessible reaction site on methoxy carbon (Figure 3), with a significant orbital energy difference of 1.48 eV. These suggest that HOMO of phenols would prefer to interact with LUMO lobe on the carbonyl carbon of DMC *via* path A to provide the corresponding phenyl carbonates (Figure 4), and the interaction with LUMO+1 lobe on the methoxy carbon of DMC *via* path B to produce the corresponding O-methyl ethers will encounter higher energy barriers.

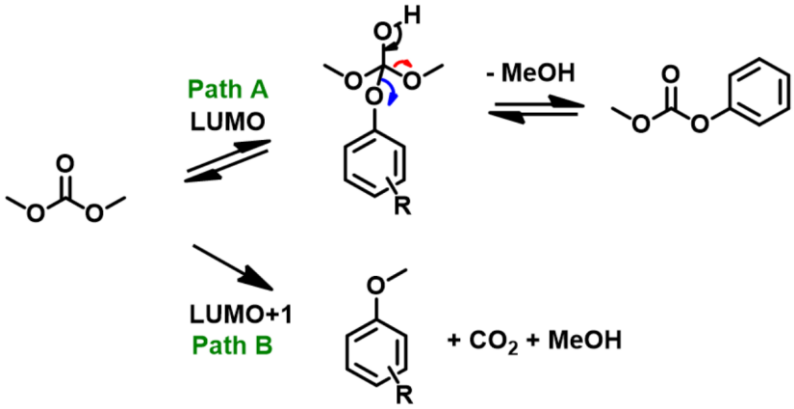


Figure 4: Two potential reaction paths between different phenols and DMC

p-Methoxyphenol, which is the least acidic phenol in the set, reacts with DMC to provide a 77:23 mixture of carboxymethylation and methylation products *via* path A and the higher energy path B, respectively, with the highest percentage of carboxymethylation reported in the study. In contrast, reaction with *p*-cyanophenol, which is the most acidic phenol in the paper, provides quantitatively and selectively, only the methylation product (Figures 4, 5). We reasoned that the tetrahedral dimethoxy(aryloxy)methanol intermediate formed with the most acidic *p*-cyanophenol reverts most readily back to starting phenol, which is then methylated irreversibly *via* path B, a higher energy process. The overall observed ratios of

carboxymethylation *vs* methylation products are related to the reaction temperature, reaction time, p*K*a of the phenol itself (Figure 5), the reversibility and irreversibility of the competing reactions, and the amount of methanol remaining in the reacting mixture. Figure 5 enables us to predict the ratio of methylation *vs* carboxymethylation based on p*K*a of phenols of interest with the reported reaction conditions.

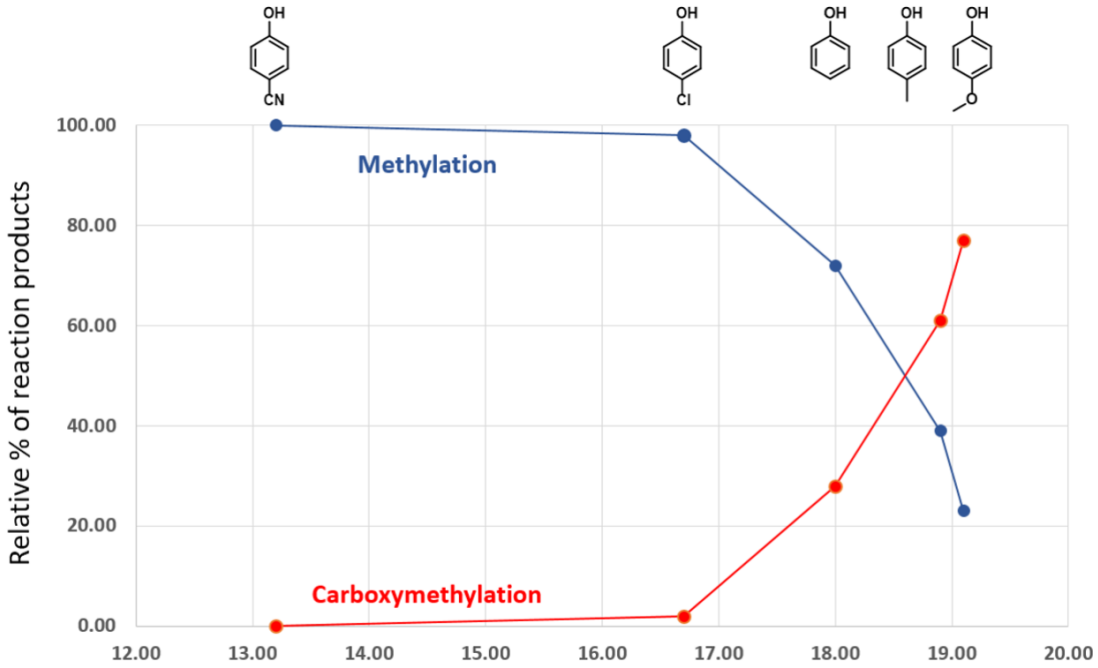


Figure 5: Plot of p*K*a of various phenols versus relative % of reaction products

Reaction of Thiols and 1-Octanol with DMC

Next, let's look at the reactions of thiols and 1-octanol with DMC (Table 2).

Reaction of Thiols and 1-Octanol with DMC in the Presence and Absence of K_2CO_3[a]

Entry	Thiol	pKa [5]	Base	Time (h)	% Methylation	% Carboxyl-methylation
1	Thiophenol	10.3	None	1	27	0
2	Thiophenol	10.3	K_2CO_3	4	100	0
3	1-Octanthiol	17.0	K_2CO_3	24	34	0
4	1-Octanol	30.0	K_2CO_3	7	0	65

[a]DMC reflux temperature, 90 °C; molar ratio thiol/DMC/K_2CO_3 = 1.0/40/1.2.

Table 2: Reaction of thiols and 1-octanol with DMC [2]

Thiols are more nucleophilic than alcohols, readily react with DMC, even in the absence of base. Thiol and thiophenol are significantly more acidic than the corresponding alcohol and phenol. With the tetrahedral intermediates formed by addition of thiophenol or thiol to the

carbonyl group of DMC, thiol is a better leaving group than alcohol. Reverting of the tetrahedral intermediates back to dimethyl carbonate is more favorable than generating the thioesters, which are more reactive than DMC; as such, carboxymethylation of thiols was not observed (Figure 6). 1-Octanol interacts with DMC to reach an equilibrium mixture of DMC and transesterification products (methyl octyl and dioctyl carbonates). Since 1-octanol (p*K*a of 30.0) is not acidic enough to be deprotonated with potassium carbonate, no methylation product was observed.

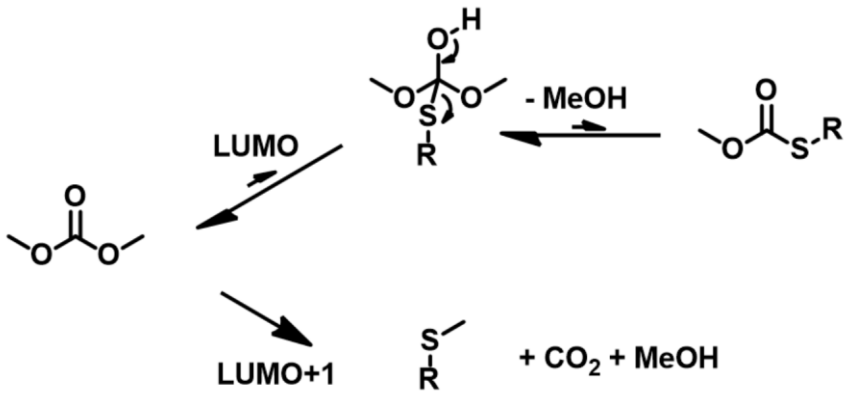


Figure 6: Potential reaction paths between thiols and DMC

Reaction of aniline compound with DMC

Finally, let's look at the reactions of aniline with DMC (Table 3).

Reaction of Aniline with DMC in the Absence and Presence of Base[a]

Entry	Base	Time (min)	% MNM	% DNM	% Carbamate	% NMcarbamate
1	None	360	0	0	0	0
2	CH_3ONa	210	0	0	67	2
3	$(CH_3)_3COK$	1	0	0	100	0
4	$(CH_3)_3COK$	180	0	0	60	40

[a]DMC reflux temperature, 90 ° C; molar ratio aniline/DMC/base = 1.0/40/1.2.

MNM = mono N-methylation; DNM = di N-methylation; NMcarbamate = N-methyl carbamate.

Table 3: Reaction of aniline with DMC[2]

Reaction of aniline and DMC exhibited very different patterns as compared with those with phenols and thiols. Aniline does not react with DMC in the absence of base. When CH_3ONa is used as the base, aniline reacts rapidly with the carbonyl carbon of DMC to form a relatively more stable carbamate and the reaction is irreversible (Figure 7, Path A-I). Methylation of aniline with DMC (Path B-I) will require using LUMO+1 of DMC with a higher energy barrier, not competitive with the irreversible acylation. The carbamate **a** might be further acylated to form intermediate **b** *via* Path A-II (LUMO), but due to its high reactivity towards methanol generated in the reaction, it'll readily revert back to **a** and is not observed. Furthermore, since

the carbamate NH is more acidic than that of NH_2 on aniline, **a** will undergo deprotonation, then methylation to give **c**, *via* Path B-II (LUMO+1), an irreversible step. With $(CH_3)_3COK$, a stronger base, and extended reaction time, **c** continued to accumulate, until the reaction was stopped.

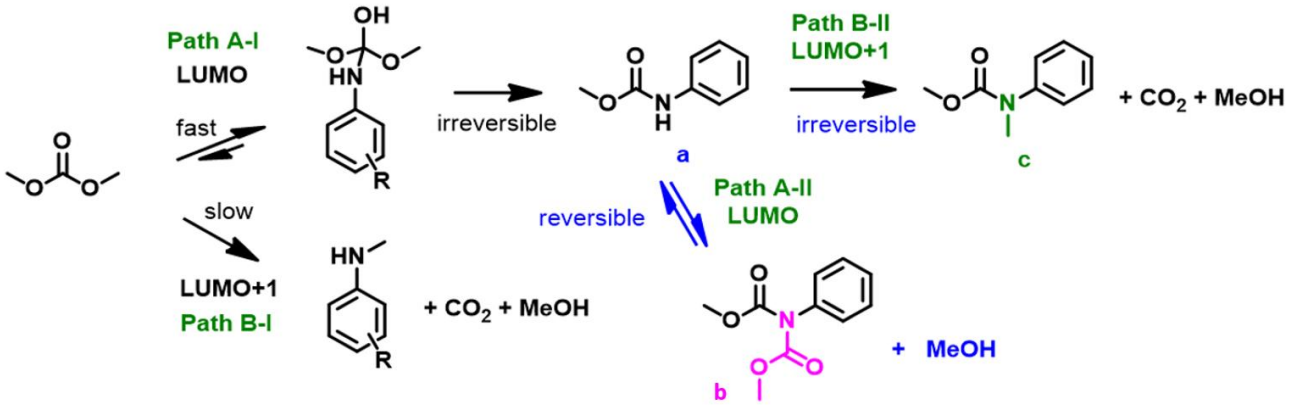


Figure 7: Potential reaction paths between aniline and DMC

It is easier to understand the complex ambident reactivity pattern of dimethyl carbonate from the perspective of LUMO, p*K*a, reversibility and irreversibility of the reactions. Our insights enable us to predict how DMC will react under specific conditions. If you are interested in p*K*a estimation and LUMO analyses of electrophiles, please refer to chapters 3 and 12.

Building on What We Just Learned

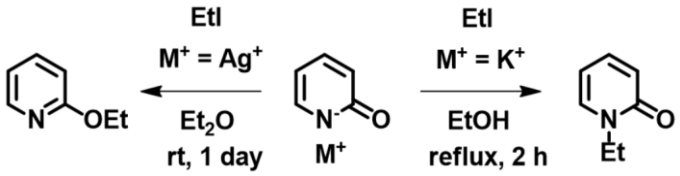


Figure 8: Alkylation of pyridone

Let us now consider the experimental results of the pyridone ethylation shown in Figure 8. In 2010, Mayr *et al.* pointed out alkylation results under silver salt conditions could not be accounted for with the S_N1 carbonium ion mechanism[7]. Based on the calculated electron density map[8] below, what will you think a reasonable explanation could be? We also encourage you to read Mayr's excellent review article[3], say "Farewell to the HSAB concept", and analyze reaction of ambident reagents from thermodynamics and kinetics perspectives.

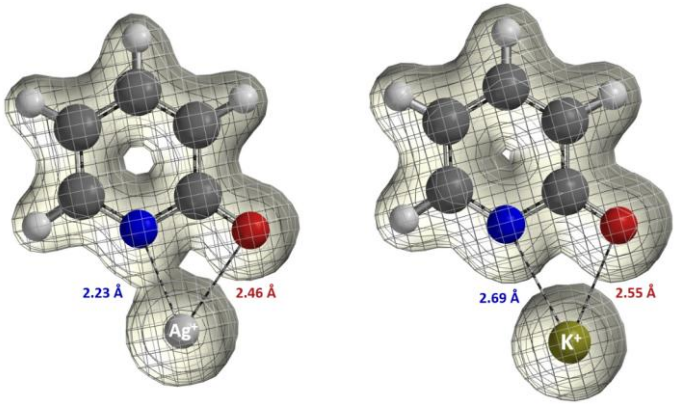


Figure 9: Calculated electron density of pyridone silver and potassium Salts
(Isovalue: Ag+ 0.046 e/au3, 92.0%; K+ 0.030 e/au3, 92.0%) [6,8]

Return to Table of Contents

References:

[1] R.G. Pearson, *J. Am. Chem. Soc.* **1963**, 3533.
[2] P. Tundo, L. Rossi, A. Loris, *J. Org. Chem.* **2005**, *70*, 2219.
[3] H. Mayr, M. Breugst, A.R. Ofial, *Angew. Chem. Int. Ed. Engl*. **2011**, 6470.
"Farewell to the HSAB Treatment of Ambident Reactivity: We have analyzed literature reports on the regio-selectivities of these and other ambident nucleophiles and electrophiles, and we have realized that the number of cases where the HSAB principle and the concept of charge- and orbital-controlled reactions give correct predictions approximate the number of cases where they fail. For that reason, we suggest abandoning these concepts as guides for predicting ambident reactivity."
[4] R.S. Drago, *J. Chem. Educ.* **1974**, *51*, 300 and references 6-9 cited in H. Mayr, M. Breugst, A. Ofial, R. *Angew. Chem. Int. Ed. Engl.* **2010**, 6470.
[5] A: https://organicchemistrydata.org/hansreich/resources/pka/pka_data/pka-compilation-reich-bordwell.pdf
B: https://organicchemistrydata.org/hansreich/resources/pka/#kaalcohol
C: https://organicchemistrydata.org/hansreich/resources/pka/#kathiol
[6] https://www.wavefun.com/spartan-documentation
[7] M. Breugst, H. Mayr, *J. Am. Chem. Soc.* **2010**, 15380.
[8] E.R. Johnson, S. Keinan, P. Mori-Sánchez, J. Contreras-García, J. Aron, A.J. Cohen, W. Yang, *J. Am. Chem. Soc.* **2010**, 6498.

Chapter 21 A Comprehensive Overview of Quantum Mechanics-Free Organic Chemistry

Qiuyue Wang, Yongsheng Chen, Zhong Zheng, and John S. Wai

Inspired by the 2014 Nature Chemistry online publication "A comprehensive overview of chemical-free consumer products"[1], we have examined and subjected an exhaustive number of organic reactions[2-6] to Quantum Mechanics analysis. Herein are described all the organic reactions, to our knowledge, that do not involve Quantum Mechanics.

References:

1. A.F.G. Golberg & C.J. Chemjobber, *Nature Chemistry*, **2014**, "A comprehensive overview of chemical-free consumer products"
http://blogs.nature.com/thescepticalchymist/files/2014/06/nchem_-Chemical-Free.pdf
2. Organic Chemistry: https://en.wikipedia.org/wiki/Organic_chemistry
3. "Quantum Mechanics" Free Molecular Mechanics:
https://en.wikipedia.org/wiki/Molecular_mechanics
4. "Quantum Mechanics" contaminants in Semi Empirical methods:
https://en.wikipedia.org/wiki/Semi-empirical_quantum_chemistry_method
5. Quantum Mechanics Free Particle: https://en.wikipedia.org/wiki/Free_particle
6. Nesse, F. M. Atanasov, G. Bistoni, D. Maganas, S. Ye, *J. Am. Chem. Soc.* **2019**, *141*, 2814. Chemistry and Quantum Mechanics in 2019: Give Us Insight *and* Numbers.

Acknowledgment

Authors appreciate Nature Chemistry for pioneering word-free publications.

Author Statements

Richard Feynman (Physics Nobel prize, 1965) noted in **The Character of Physical Law**, "I think I can safely say that nobody understands quantum mechanics." MIT Press, ISBN 987-0262533416. Analogously, authors acknowledge that we might not fully understand why the QM analyses came out as what they are as described in previous 20 chapters and future ones, yet it will not stop us from learning from it. QM doesn't speak, it is an excellent teacher we cherish.

Return to Table of Contents

Chapter 22 Exploring the Reaction Mechanism of Menke Nitration

Xiaoli Shen, Tommy Lai, Dong Pan, Yongsheng Chen, John S. Wai

Nitration is a common reaction in organic synthesis, and a large variety of nitration reagents and reaction conditions are reported in the literature. Menke Nitration, first reported by Menke in 1925[1], refers to the nitration of aromatic compounds by reaction with copper nitrate and acetic anhydride. A reaction mechanism proposed in the current literature is shown below[2].

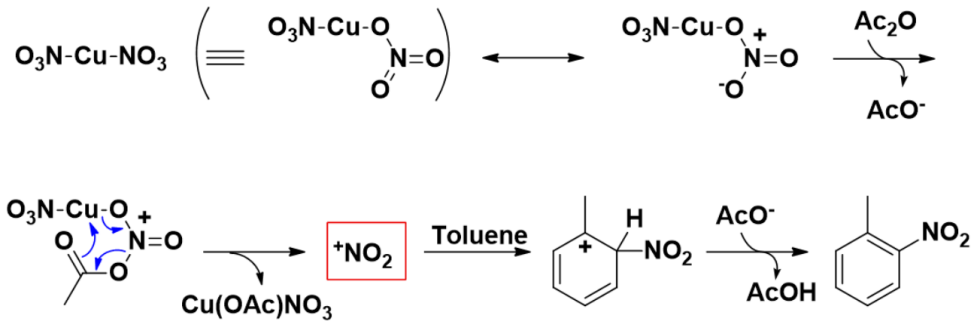


Figure 1. Proposed Menke nitration reaction mechanism

In 1960 F.G. Bordwell *et al*.[3] found that olefins can react with $AcONO_2$ in the nitration system of nitric acid-acetic anhydride, and suggested that protonated acetyl nitrate $(AcOHNO_2)^+$ may be the active species for nitration. Inspired by this, we speculate that it is $AcONO_2$ rather than NO_2^+ that is involved in Menke Nitration, and decide to use QM calculations to evaluate our hypothesis.

Reaction Modeling

For Menke nitration, copper nitrate reacts with acetic anhydride to form the intermediate $AcONO_2$. We reasoned that the following two reaction pathways are possible (Figure 2). Path 1: the electrophilic substitution reaction of the intermediate $AcONO_2$ with benzene ring, then departure of H^+ to form the product and acetic acid. Path 2: nitration proceeds *via* a six-membered ring transition state, with an intramolecular transfer of the "aromatic" H to the acetate group in a concerted manner.

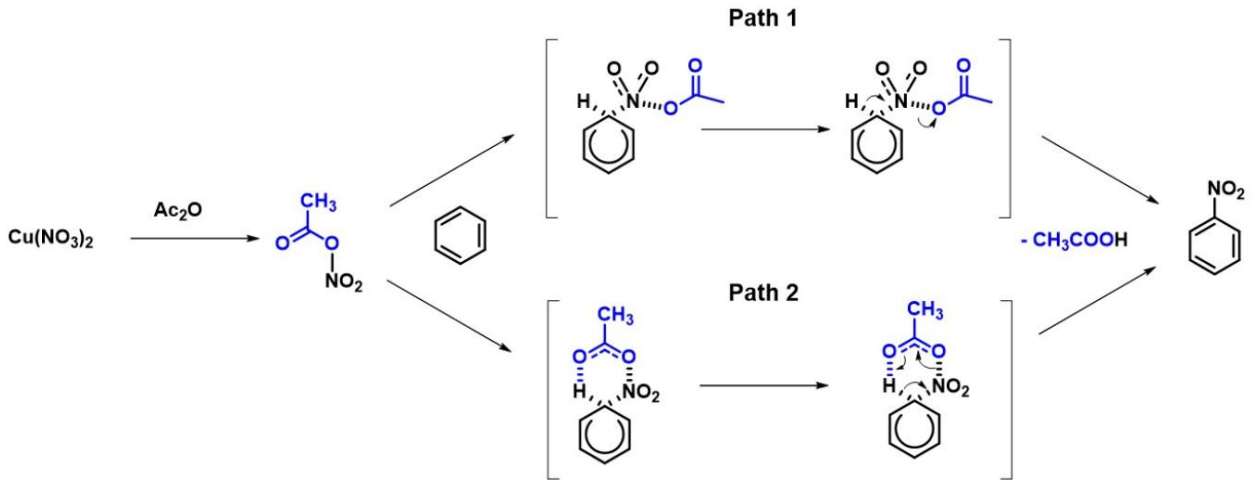


Figure 2. Two possible reaction pathways of Menke nitration

To evaluate which potential reaction pathway is energetically more favorable, we calculated for both reaction energy profiles, with the aromatic C to nitro group N distance changing from 3.0 Å to 1.5 Å over 16 steps. These calculations suggest Path 1 and 2 will have activation energy of ~ 60 and ~50 kcal/mol, respectively. Path 2 is ~10 kcal/mol lower than that with Path 1 (Figure 3), suggestive of reaction mechanism with a six-membered ring transition state.

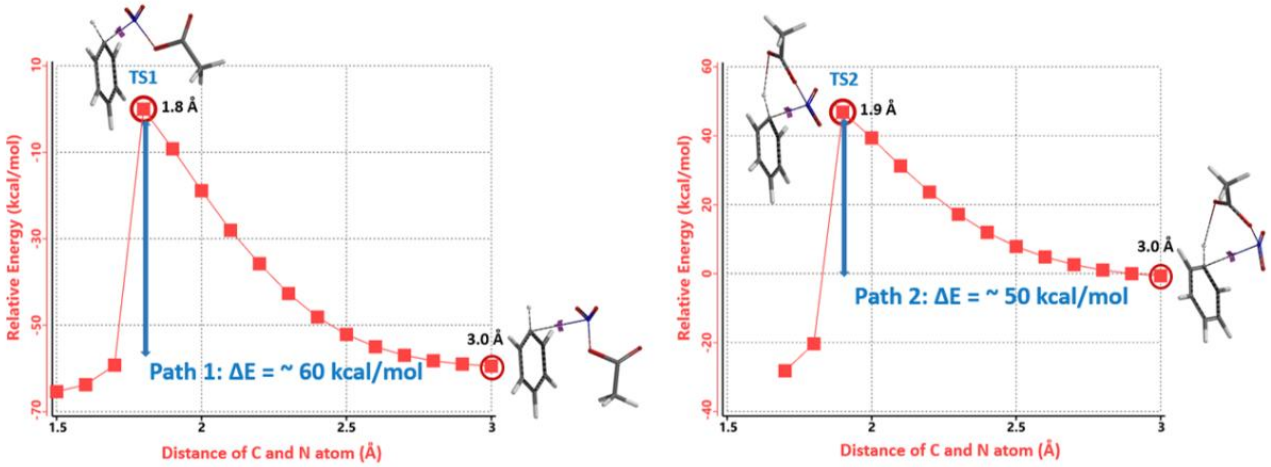
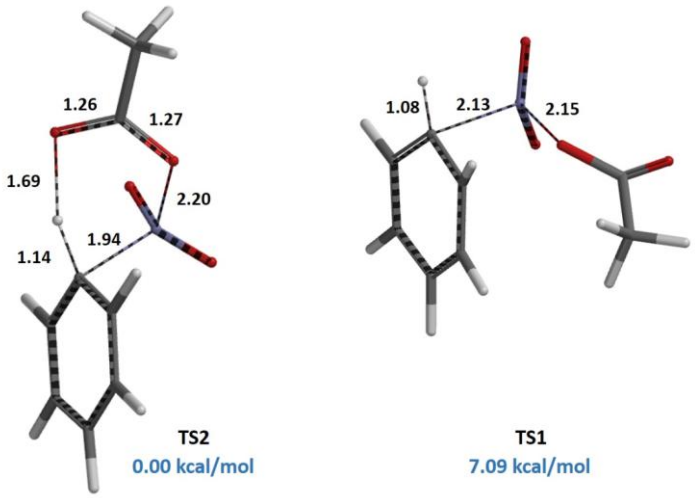


Figure 3. Reaction energy profiles for the two possible reaction pathways

The structures **TS1** and **TS2**, at the highest points in the two reaction energy profiles, were then used to calculate for more accurate "Transition State Geometry" and to determine their relative energy difference. Energy of **TS2** is found to be 7.09 kcal/mol lower than that of **TS1**, with a Boltzmann distribution favoring **TS2** by 100% (Figure 4), presumably from the intramolecular hydrogen bond stabilization in **TS2**.

Infrared vibration calculation of structure **TS2** showed one and only one imaginary frequency at i412 cm^{-1} (Figure 5)[5]. Intrinsic Reaction Coordinate (IRC) calculation of **TS2** reaches both the starting material and nitration product. These results provide further support that **TS2** could be a viable transition state for Menke Nitration.



ΔE (kcal/mol)	Boltzmann [4]
-7.09	100:0

Figure 4. Relative energy and proportion of **TS1** and **TS2**

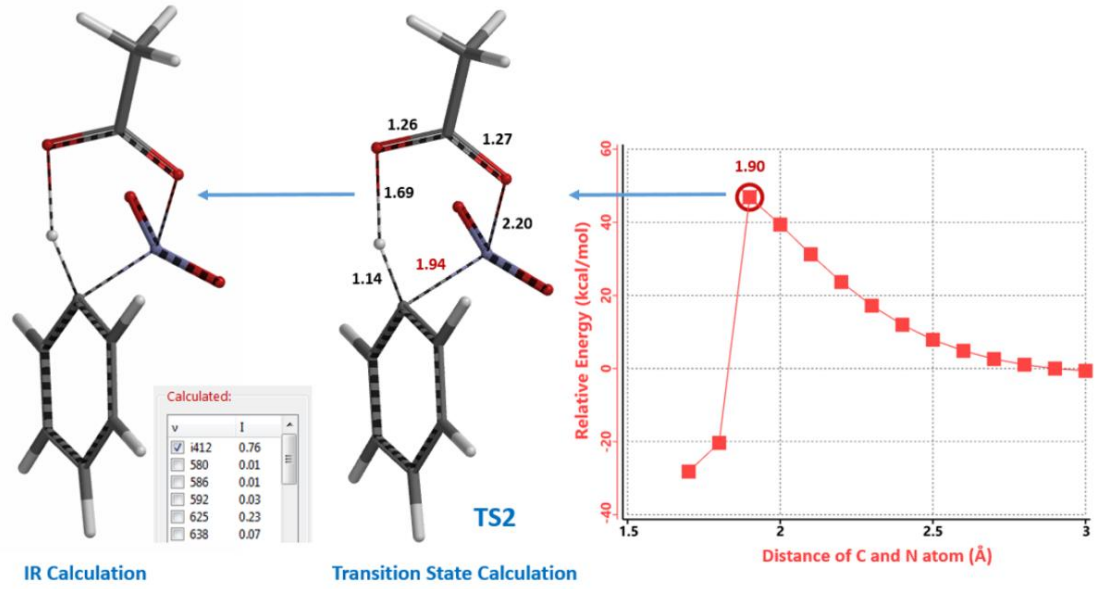


Figure 5. Reaction energy profile, **TS2** Transition State and IR Calculation Summary

Refining and Validating the Menke Nitration Model

Menke nitration of chroman 1 provided a 3 : 1 mixture of 6- and 8-nitro products, respectively. We will use this example to illustrate how we use calculated HOMO and ^{13}C NMR data of a substrate (see Chapter 2) along with the above Menke nitration model to **predict** and **quantify** the regio-selectivity observed.

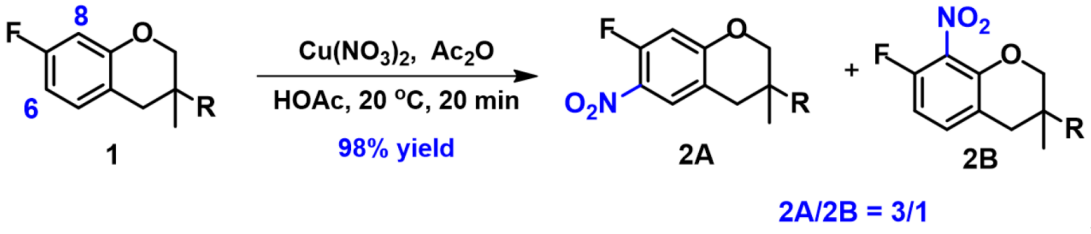


Figure 6. Menke nitration of chroman **1**

The R-group of chroman **1** was simplified to a methyl group for faster calculations. Energy difference between HOMO and HOMO-1 is 0.66 eV (Figure 7). HOMO lobe on C-6 is significantly larger than that on C-8, while the HOMO-1 lobe on C-8 is significantly larger than that on C-6. Calculated ^{13}C NMR showed a chemical shift difference of < 6 ppm between C-6 (108.2 ppm) and C-8 (104.8 ppm). These calculated parameters could account for the observation of a mixture of **2A** and **2B** nitration products, with **2A** as the major one.

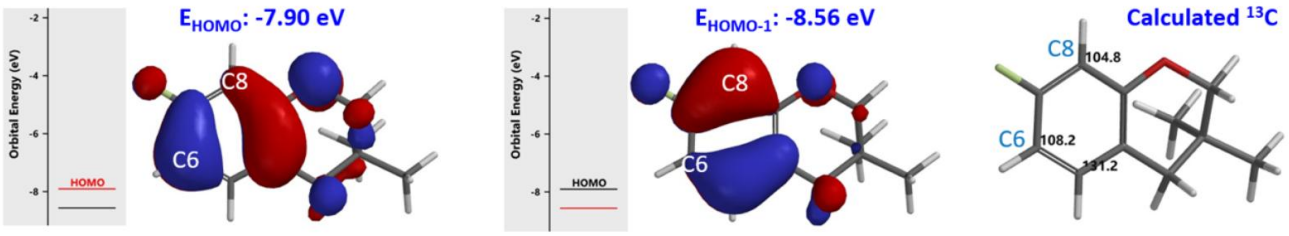


Figure 7. Scheme of HOMO, HOMO-1 & ^{13}C NMR of compound **1**

To refine the Menke Nitration reaction model proposed above, we first calculated for the reaction energy profiles for nitration at C-6 and C-8. The estimated activation energies are found to be ~38.7 kcal/mol and ~42.6 kcal/mol, respectively, favoring C-6 nitration.

Then we selected the structures that correspond to the highest energy point of the two energy profiles to calculate for "Transition State Geometry" of **TS6** and **TS8** (Figure 8) and their relative energy differences with various dielectric constants settings (CPCM model). Results are tabulated in Table 1.

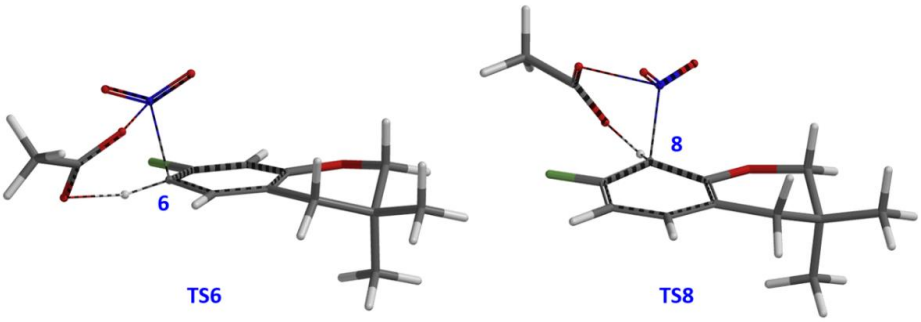


Figure 8. Transition state structures of C-6 & C-8 nitration

Options	TS6 kcal/mol	TS8 kcal/mol	ΔΔE kcal/mol	ratio at 25 °C
Gas	0.00	4.31	-4.31	218:1
Nonpolar solvent	0.00	1.42	-1.42	11:1
Polar solvent	0.00	0.97	-0.97	5.2:1
Water	0.00	0.90	-0.90	4.6:1

Table 1. Relative energy difference and product ratio calculated under different dielectric constants

"Polar solvent" (dielectric constant of 37), consistent with acetic acid being used as solvent in the reaction, shows the best correlation with the experimental result of 3:1. This further substantiates the hypothesis that $AcONO_2$ is the active nitrating species in Menke Nitration, the reaction proceeds *via* a six-membered ring transition state, and the model is useful to predict and quantify regio-selectivity of Menke Nitration.

This chapter highlights the use of QM calculation to evaluate possible mechanisms of a reaction, the use of transition state geometry and ifrequency calculations to further assess the models, and the use of experimental results to refine necessary dielectric settings used in calculations to better quantify regio-selectivity observed. Validated QM reaction models are powerful prospective analysis tools in synthetic planning, and their refinement with experimental data will lead to better mechanistic understanding of reactions of interest.

Building on What We Just Learned

HSO_3ONO_2 (LUMO energy of -1.00 eV, Figure 9) is much more electrophilic than $AcONO_2$ (LUMO energy -0.29 eV), could be generated by the reaction of anhydrous HNO_3 with SO_3 in fuming sulfuric acid, and its use patented in a remarkable continuous and closed-loop process for nitration of dinitrotoluene to TNT, at a relatively low temperature of 70-90 °C[6]. Based on the QM analysis of Menke nitration discussed above, the reaction energy profile of HSO_3ONO_2 nitration of benzene (Figure 10), the transition state structure, and an imaginary frequency (i194 $^{cm-1}$), an obvious question emerges: mechanistically could both $AcONO_2$ and HSO_3ONO_2 nitration proceed through analogous six-membered ring transition states?

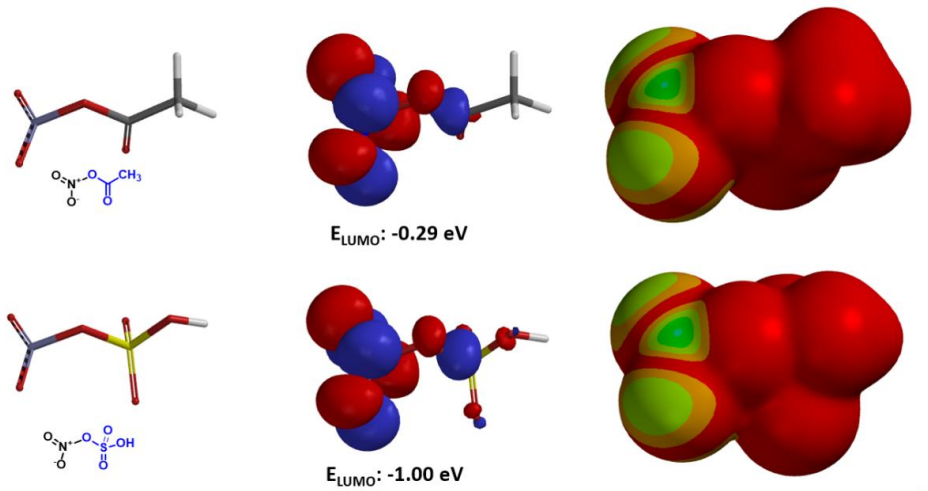


Figure 9. LUMO and LUMO maps calculated for $AcONO_2$ and HSO_3ONO_2

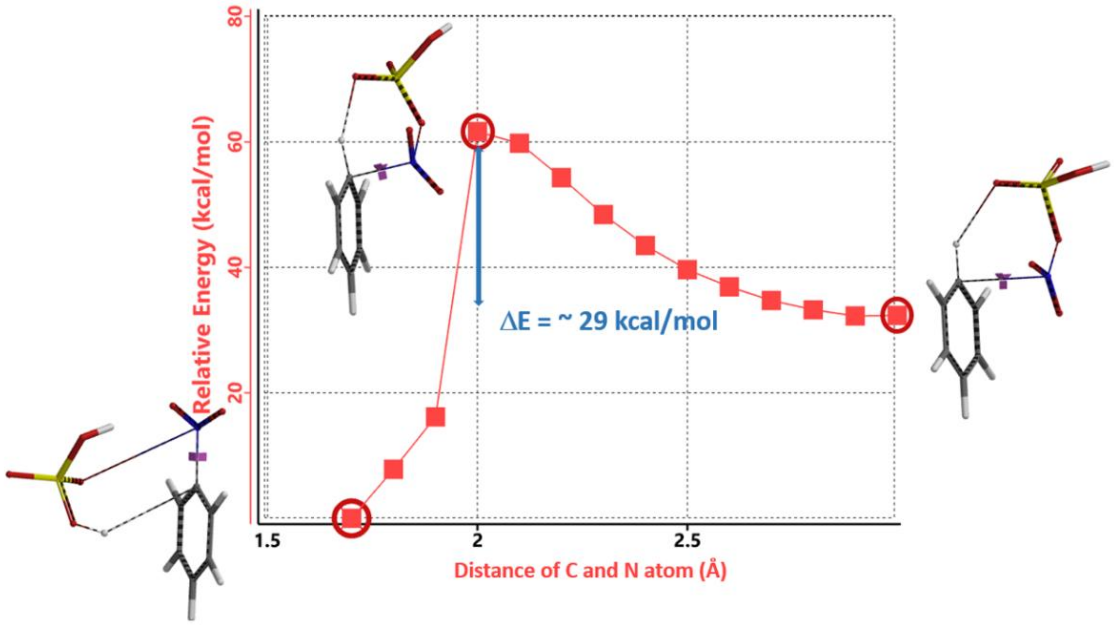


Figure 10. Reaction energy profile of HSO_3ONO_2 nitration of benzene

References:

[1] (a) J.B. Menke, *Recl. Trav. Chim. Pays-Bas,* **1925**, *44*, 141. (b) J.B. Menke, *Recl. Trav. Chim. Pays-Bas*, **1925**, *44*, 269. (c) J.B. Menke, *Recl. Trav. Chim. Pays-Bas*, **1925**, *44*, 270.
[2] Z. Wang, *Comprehensive Organic Name Reactions and Reagents,* volume 2. John Wiley & Sons, Inc., Hoboken, New Jersey, US, **2009**, 1893.
[3] F.G. Bordwell, E.W. Garbisch Jr., *J. Am. Chem. Soc.,* **1960**, *82*, 3588.
[4] The Boltzmann distribution function describes the probability distribution of molecules according to the level of energy under thermal equilibrium.
[5] *Spartan'18 Tutorial and User's Guide*. Irvine, CA, USA: Wavefunction, Inc., **2019**; pp 152, 495, & 613. In computational chemistry, molecular vibrations at the transition state can be described approximately with a harmonic oscillator model. As shown in the equation below, the vibration frequency is proportional to the square root of the ratio of force constant *k* and reduced mass *μ* (a combination of the masses of atoms involved in motion along that coordinate). The force constant *k* is the curvature of the potential energy surface along a specific coordinate (defined by the second derivative of the energy relative to the geometric coordinate). Transition State Theory and the transition state is located at the saddle point on the potential energy surface (see Chapter 8). The curvature of the transition state on the reaction path is negative and the reduced mass is positive, so the frequency calculated is an imaginary number. Having one and only one imaginary frequency that corresponds to vibration of the bonds being made/broken is a crucial criteria to satisfy for the calculated structure to be considered as a transition state.

$$\Delta E = h\nu = \frac{h}{2\pi}\sqrt{\frac{k}{\mu}}; \quad k = \text{force constant}; \quad \mu = \text{reduced mass.}$$

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}}; \quad k < 0 \rightarrow \nu \text{ is an imaginary frequency}$$

[6] M.E., Hill, W.E. Tolber, G.J. McDonald, US Patent 3,799,993, March 26, 1974.

Chapter 23 QM Study of the *para* Regioselectivity of $TBABr_3$ Bromination

Dong Yu, Dong Pan, Xiaoli Shen, Tommy Lai, Yongsheng Chen, John S. Wai

In Chapter 6, "Stepwise Multiple Halogenation Coupling Reaction Strategies", we discussed the use of tetrabutylammonium tribromide ($TBABr_3$) for a highly *para* selective bromination of 2-amino-6-cyanopyridine **A** (Figure 1)[1], overriding the intrinsic bias for *ortho* substitution, suggestive from its HOMO and calculated ^{13}C chemical shifts. We are interested in understanding the mechanism of this reaction.

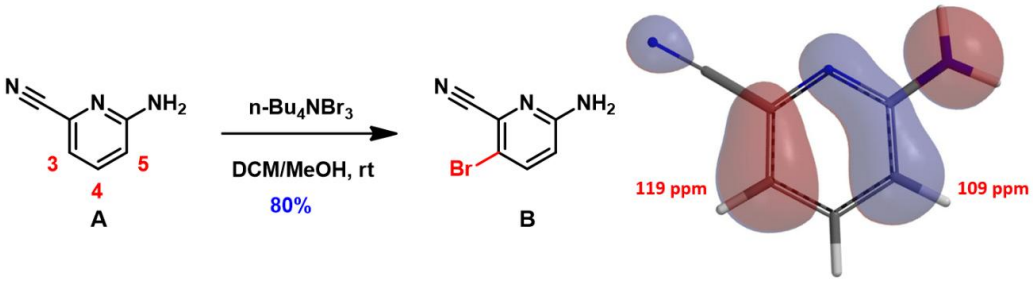


Figure 1. Bromination of compound **A**, HOMO and calculated ^{13}C chemical shifts

Mechanism of $TBABr_3$ reaction reported in the literature

Shown in Figure 2 (left) is the reaction mechanism proposed by Chen & Zhou[2a] for $TBABr_3$ bromination of phenol. It was suggested that free Br_3 anion is the brominating species. The reaction involves proton transfer from the phenolic hydroxyl group to tribromide and a series of electron transfers to provide the *para* bromide selectively. We are intrigued to explore how an electron-rich tribromide anion could react with an electron-rich phenol[2] and calculated for the LUMO of the Br_3 anion (Figure 2, right). The LUMO orbital energy is 3.07 eV, too high to be considered as an active brominating species for a room temperature reaction.

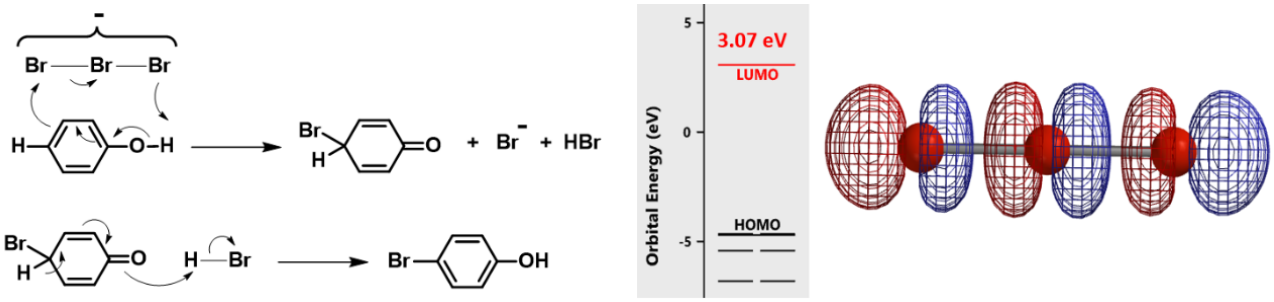


Figure 2. Literature proposed mechanism[2a], LUMO and LUMO energy of Br_3 anion

In parallel, we calculated for the reaction energy profile for this proposed mechanism (Figure 3). The C – Br bond distance changes stepwise from 3.8 Å to 1.9 Å; after a small dip at 2.1 Å,

energy level continues to increase, casting further doubts that the tribromide anion is the active brominating species.

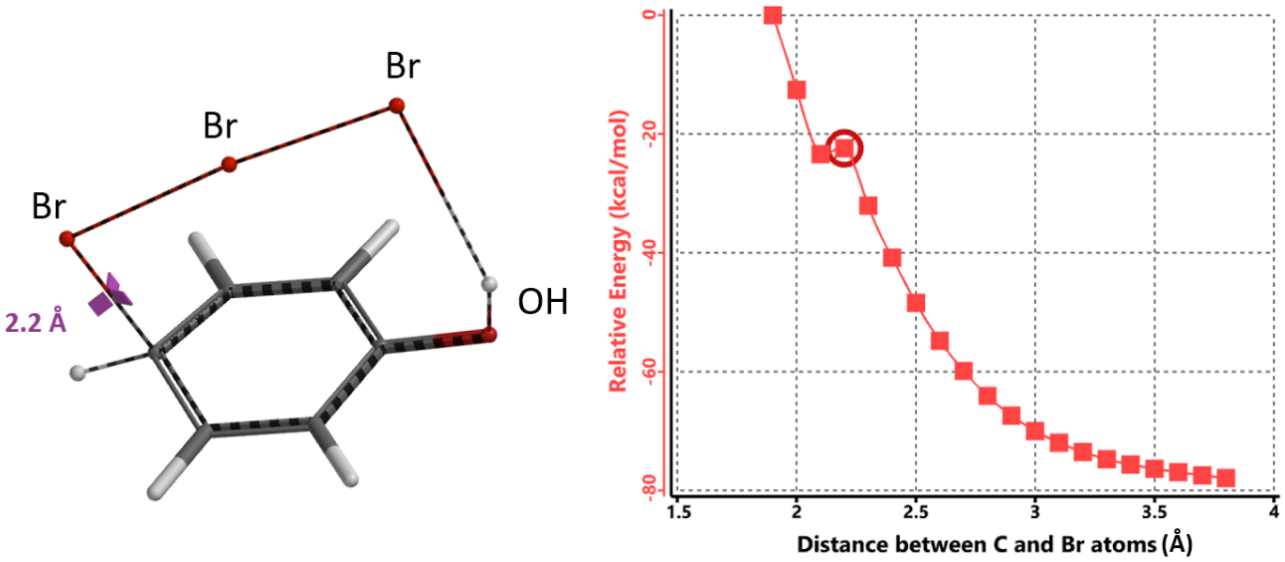


Figure 3. Reaction energy profile for the literature proposed mechanism[2a]

QM Modeling of the Active Bromination Species

We reasoned that the undissociated $TBABr_3$ is the active bromination species, not the free tribromide anion, and there are two possible ways for the tetrabutylammonium cation and the tribromide anion to be associated. We simplified the tetrabutylammonium cation as tetramethylammonium (TMA) for calculations. LUMO energies of structures with TMA^+ attached to Br-1 and Br-2 are 0.41 and -0.34 eV, respectively (Figure 4). The lower the LUMO energy of an electrophile, the more reactive it is expected to be. The calculated equilibrium structure with TMA^+ attached to Br-2 is remarkable. The N of TMA^+ lies between Br-1 and Br-2, presumably energetically more stable bridging the two bromo atoms, and is actually closer to Br-1 than Br-2 (distances of 3.96 Å and 4.19 Å, respectively). In addition, this TMA–Br-1–Br-2 cyclic structure is 13.42 kcal/mol lower in relative energy than that of the TMA–Br-1 linear structure. Together, these parameters suggest that the cyclic structure is more likely to be the active bromination species.

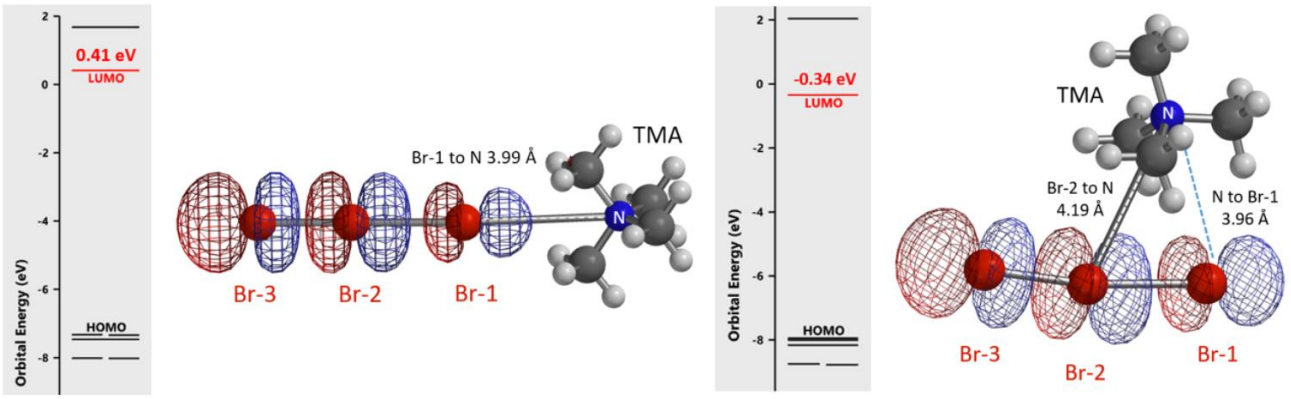


Figure 4. Possible structures of $TMABr_3$, their LUMO, and LUMO energies

QM Modeling of Reaction Energy Profiles

Next we modeled the bromination with TMA^+ connecting to the tribromide anion at all three possible positions (Figure 5).

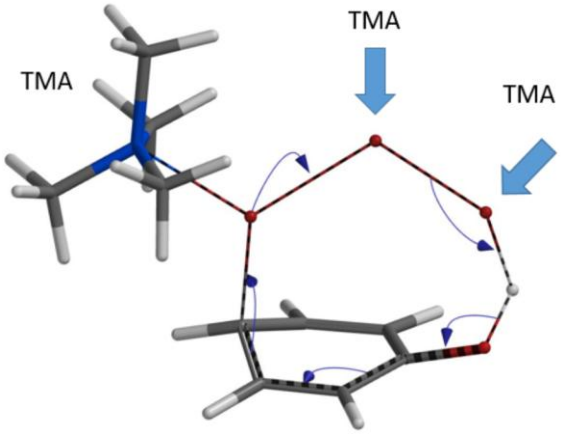


Figure 5. Three possible connections

Remarkably, all three calculations found an identical transition structure, suggesting that the connection between TMA^+ and tribromide anion is dynamic. Shown on Figure 6 is the reaction energy profile calculated with TMA^+ attached to Br-2 and the C – Br-3 distance changing from 3.5 Å to 2.2 Å. At a distance of 2.4 Å, the relative energy reaches a maximum, providing an estimated activation energy of ~ 25 kcal/mol.

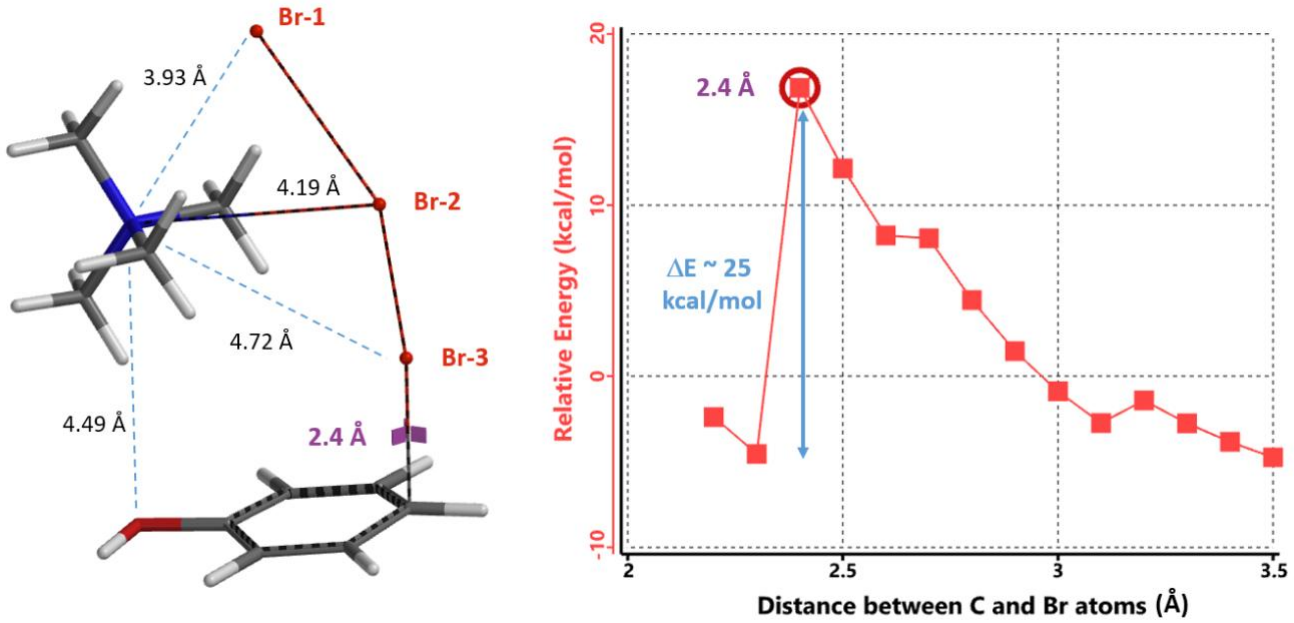


Figure 6. Reaction energy profile modeling for $TMABr_3$ bromination of phenol

This calculated transition state structure does not involve hydrogen bonding between the phenolic hydrogen and tribromide[2a], and is more consistent with the observation that $TBABr_3$ selectively *para* brominates N,N-diethylaniline (with no NH proton)[3]. The $TMABr_3$ moiety in the transition state structure (Figure 6) closely resembles the ground state $TMABr_3$ structure modeled (Figure 4, right), supportive of our proposal that the whole $TBABr_3$ entity, not the tribromide anion, is the active brominating species. Furthermore, calculated structure shows that TMA^+ is spatially close to the electron-rich phenol oxygen (N-O distance of 4.49

Å), suggestive of an attraction interaction between them, anchoring $TMABr_3$ for its Br-3 to have favorable contact (Figure 6).

The structures from Reaction energy profile calculation also show that Br-3 of $TMABr_3$ is approaching the *para* carbon at close to 90° angle to the plane of phenol. This enables more optimal orbital interaction, as exemplified with the structures and orbitals in Figure 7, with C – Br-3 distance at 2.6 Å, approaching the transition state (Figure 7).

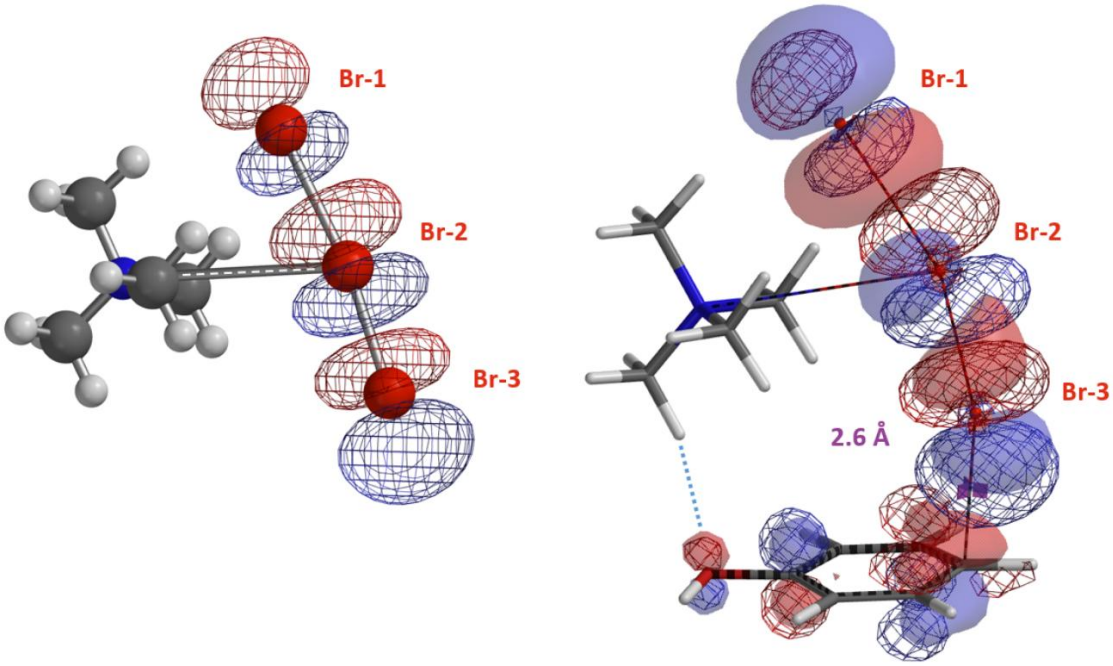


Figure 7. $TMABr_3$ LUMO and $TMABr_3$-Phenol HOMO (solid)/LUMO (mesh) at a C – Br distance of 2.6 Å

Furthermore the electrostatic potential energy map of the reacting complex at a C – Br-3 distance of 2.6 Å shows that the methyl group on the tetramethylammonium ion is highly electron deficient (Figure 8 left, dark blue)[4]. The electron density map reveals a noncovalent interaction between the hydroxyl oxygen and the hydrogen on the nearest methyl group of TMA^+ [5], forming a cyclic $TMABr_3$ – PhOH reacting complex (Figure 8 right), stabilizing the transition state.

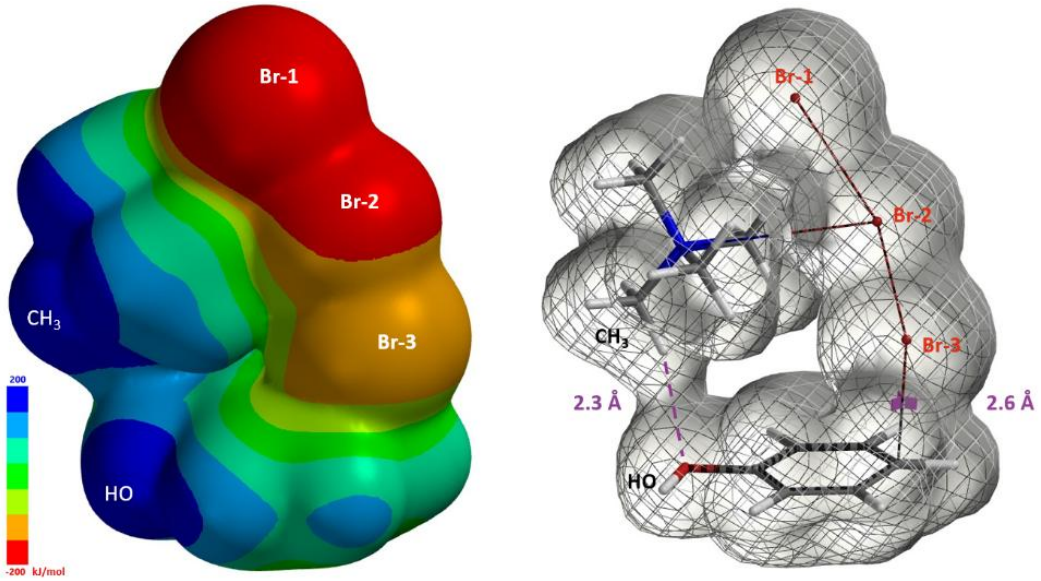


Figure 8. Electrostatic potential energy map (Isovalue 0.002 e/au3, 99.4%) and Electron density map (Isovalue 0.00744 e/au3, 98.0%) of $TMABr_3$ – PhOH reaction complex with C – Br distance at 2.6 Å

Modeling ortho bromination for comparison

Next we calculated for the Reaction energy profile of *ortho* bromination of phenol for comparison (Figure 9). The activation energy of *ortho* bromination is about 45 kcal/mol, 20 kcal/mol higher than that of the *para* substitution, accounting for the high *para* selectivity observed.

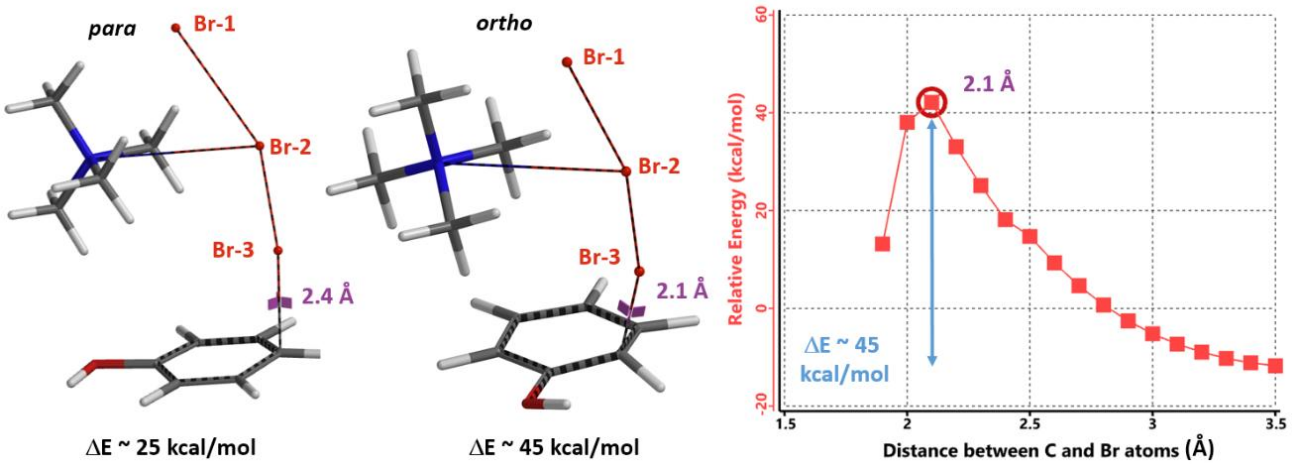


Figure 9. $TMABr_3$ *para* and *ortho* bromination TS structures, and Reaction energy profile of *ortho* bromination

In summary, our QM analyses offer an alternative reaction mechanism to account for *para* selective bromination of phenol with $TBABr_3$. Reaction energy profile calculation suggests that Br-3 in $TMABr_3$ is approaching the *para* carbon at close to 90° angle to the plane of phenol to enable optimal orbital interaction. The electrostatic potential energy and electron density maps reveal the presence of a non-covalent interaction between phenol oxygen and the hydrogen on the nearest methyl group on TMA^+ cation, anchoring the reagent for selective bromination at the *para* position. Activation energy estimated for *ortho* bromination reaction is about 20 kcal/mol higher than that of the *para* bromination, accounting for the high *para* selectivity observed with $TBABr_3$.

Building on What We Just Learned

Selective *para* chlorination of phenols has also been reported[6]. In 2019, El-Hiti *et al* reported that sulfonyl chloride, in the presence of catalytic amount of tetrahydrothiopyran and $AlCl_3$, chlorinates phenol selectively at the *para* position (Figure 10)[6e].

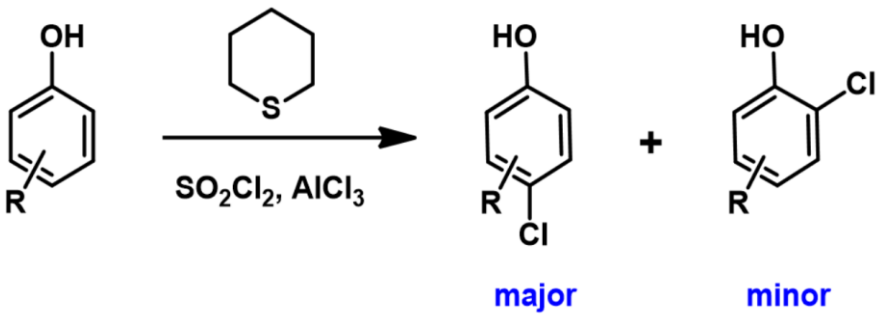


Figure 10. Selective *para* chlorination of phenols

Assuming tetrahydrothiopyran-Cl^+ $AlCl_4^-$ complex to be the chlorinating species, we calculated for the reaction energy profiles for *para* and *ortho* substitution, and the LUMO for the resultant transition state structures (Figure 11). We encourage our readers to compare and contrast them with those calculated above with $TMABr_3$.

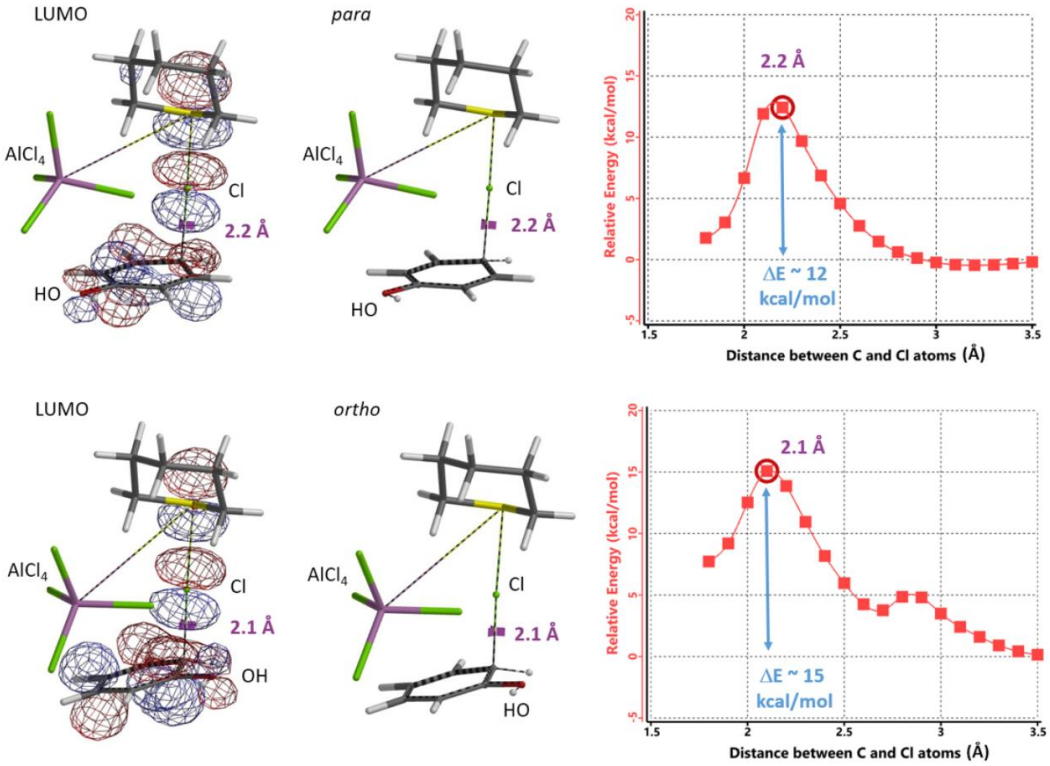


Figure 11. Tetrahydrothiopyran-Cl^+ $AlCl_4^-$ chlorination,
calculated transition state structure with LUMO and reaction energy profiles.

Return to Table of Contents

References:

[1] I. Macsari, K. Paulsen, L. Rakos, *et al* Compounds and Their Use for Treatment of Amyloid Beta-Related Diseases: US2012122843, 2012, A1, 50.

[2] (a) Z. Chen & K. Zhou, $TBABr_3$ in Organic Synthesis. *University Chemistry* **1995**, *10*, 10. https://doi.org/10.3866/PKU.DXHX19950403 (b) R. Karmaker, N.B. Kuotsu, A. Ganguly, N. Guchhait, U.B. Sinha, *Chem. Phys. Lett.* **2018**, 118. (c) J. Berthelot, C. Guette, P.-L. Desbene, J.-J. Basselier, P. Chaquin, D. Masure, *Can. J. Chem.* **1990**, *68*, 464.

[3] J. Berthelot, C. Guette, M. Essayegh, P.-L. Desbene, J.-J. Basselier, *Synth. Commun*. **1986**, *16*, 1641.

[4] Chapter 3: Application of Electrostatic Potential Map in Acidity Estimation.

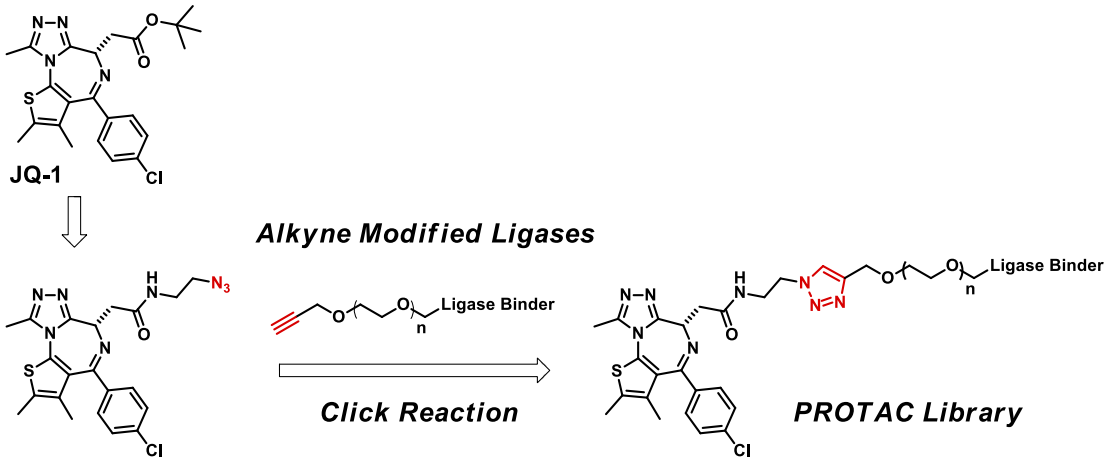
[5] E.R. Johnson, S. Keinan, P. Mori-Sánchez, J. Contreras-García, A.J. Cohen, W. Yang, *J. Am. Chem. Soc.* **2010**, *132*, 6498.

[6] (a) W.D. Watson, *J. Org. Chem. 1985*, **50**, 2145. (b) S.M. Maddox, A.N. Dinh, F. Armenta, J. Um, J.L. Gustafson, *Org. Lett.* **2016**, *21*, 5476. (c) Y. Xiong, H. Duan, X. Meng, Z. Ding, W. Feng, *J. Chem*., **2016**, Article ID 2960414. https://doi.org/10.1155/2016/2960414. (d) A. Jassim. & A.*J.* Zuhairi, *Kerbala University*, **2014**, *12*, 230. (e) K. Smith, D. Williams, G.A. El-Hiti, *J. Sulfur Chem*. **2019**, *40*, 528.

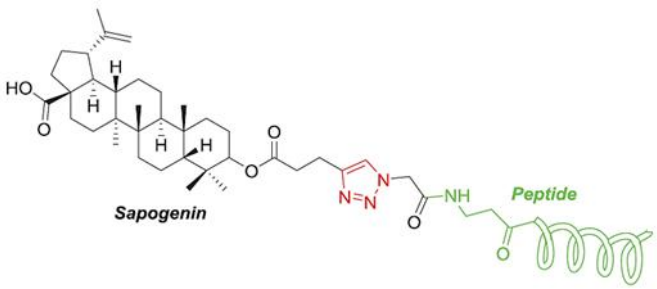
Chapter 24 QM Study of the Reaction Mechanism of Diazotransfer

Zhong Zheng, Qiuyue Wang, Jian Wang, Dong Pan, Yongsheng Chen, John S. Wai

1,2,3-Triazole is a common heterocycle in medicinal chemistry, providing us novel structures with a wide range of biological activities. 1,4-Disubstituted 1,2,3-triazoles are versatile (Figure 1): they have been used as part of the pharmacophores; as bioequivalent for specific functional groups; as peptidomimetic to improve metabolic stability against proteases; and as linkers for multispecific drugs, such as PROTAC®, drug conjugates, and so on[1].



Click-Chemistry-mediated rapid synthesis of bispecific molecules for inducing protein degradation.



Click-Chemistry-mediated conjugation of a nonspecific antiviral Sapogenin with a HIV fusion inhibitor.

Figure 1. Examples of the use of 1,2,3-triazole moiety in drug design

1,2,3-triazoles are usually constructed with [3 + 2] cycloaddition reaction, including the Cu(I) catalyzed Click Reaction (Figure 2) developed by Professor Sharpless. The required azido compounds are highly reactive, yet usually associated with toxic and explosive properties. As such, safer methods for their generation and handling are highly desirable in MedChem discovery efforts.

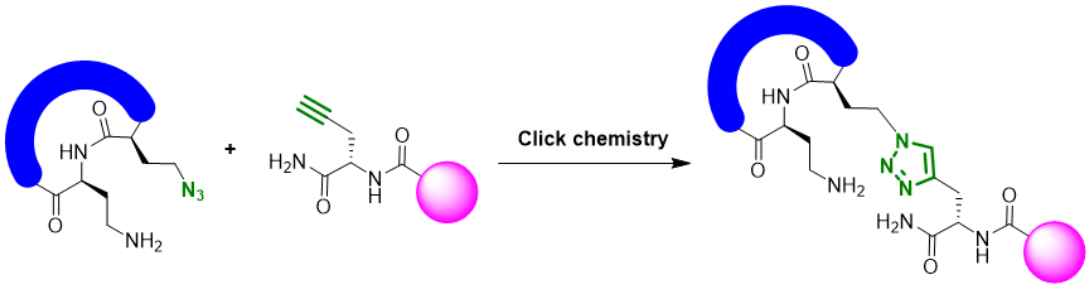


Figure 2. Example of 1, 2, 3-triazole skeleton construction

In 2019, Jiajia Dong of Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, and Barry Sharpless published in Nature the use of fluorosulfuryl azide as a safer and more efficient reagent for the preparation of various azides and their use in the Click reaction (aka Double Click) to construct a triazole library with more than a thousand compounds (Figure 3).[2]

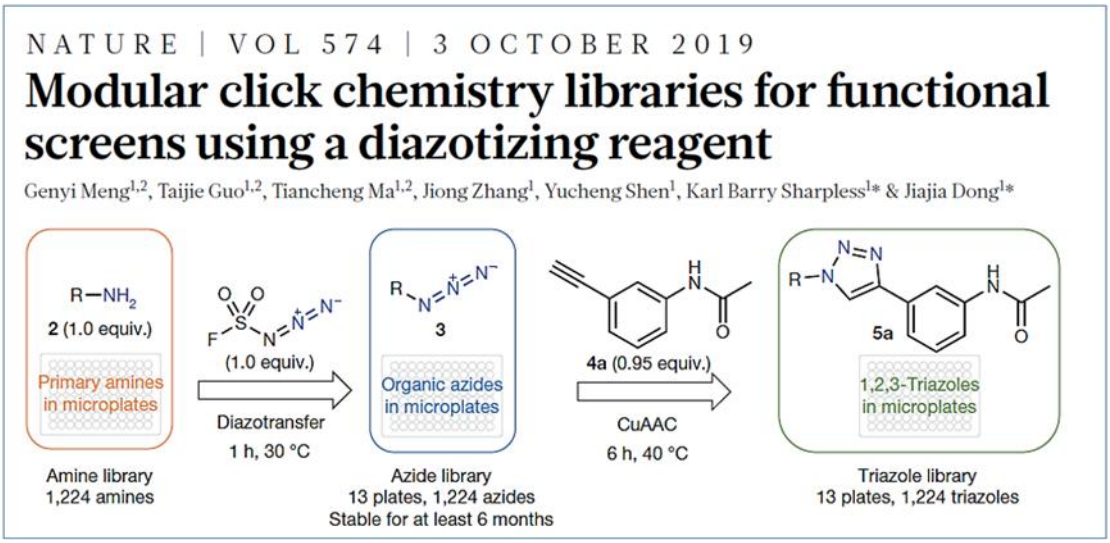
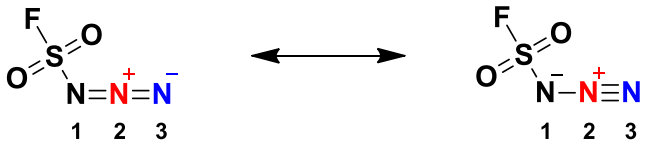


Figure 3. Diazo transfer reaction published in Nature, 2019

Fluorosulfuryl azide, FSO_2N_3, introduced in this paper shows superior efficiency in diazo transfer to primary amines. Compared with the traditional diazo transfer reagents, it presents many advantages: the reactions are carried out at room temperature, 1:1 stoichiometric, fast and efficient conversion, suitable for all primary amines, including tertiary alkyl amines (e.g. *t*-$BuNH_2$), near quantitative conversion. In this chapter, we'll discuss QM analysis of the reaction mechanism of this useful diazo transfer reagent.

Discussion on reaction mechanism

Fluorosulfuryl azide is highly electrophilic. Among the three azido nitrogen, we need to determine which nitrogen atom is the most electrophilic one to interact with the incoming amine (Fig. 4).



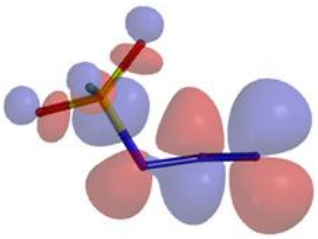


Figure 4. Resonance structures of fluorosulfuryl azide and its LUMO

The LUMO and LUMO map of fluorosulfuryl azide were calculated first (see Chapter 1). Results are shown on Figure 5, with LUMO lobe on N-3 > N-2 > N-1, and LUMO map clearly shows that N-3 is more accessible than N-2, suggesting that nucleophilic attack is more likely to occur on N-3. This is consistent with results from a ^{15}N isotope labeling experiment reported by Samuelson *et al*. in 2014[3].

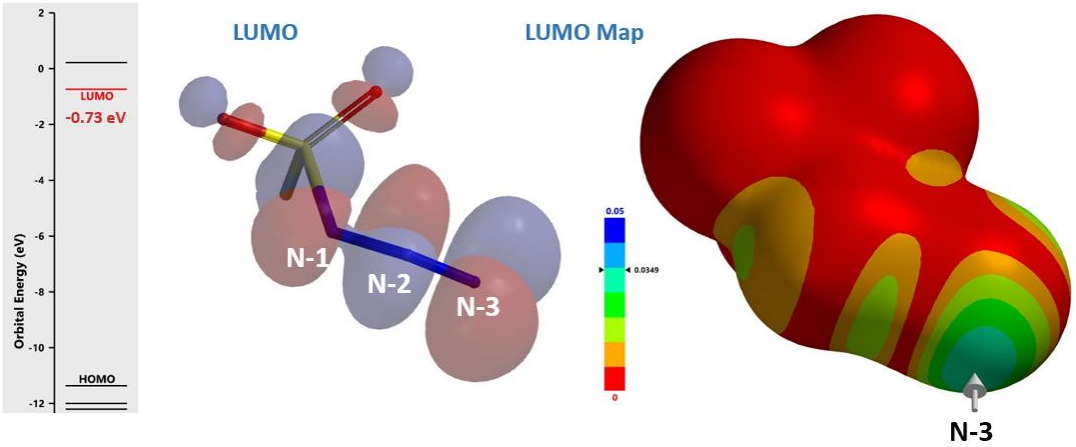


Figure 5. LUMO and LUMO map of fluorosulfuryl azide

Modeling of the First Step

The Nature paper pointed out that water is crucial for fluorosulfuryl azide to transfer a diazo group to amines efficiently. Our preliminary calculations indicated that nucleophilic attack of amine groups is a stepwise mechanism, not a one-step concerted mechanism (Figure 6), and shall include a water molecule in the reaction modeling, i.e. the amine attacks N-3 of fluorosulfuryl azide first, followed by transfer of a hydrogen atom from the amino group *via* a water molecule to N-1 of the reagent.

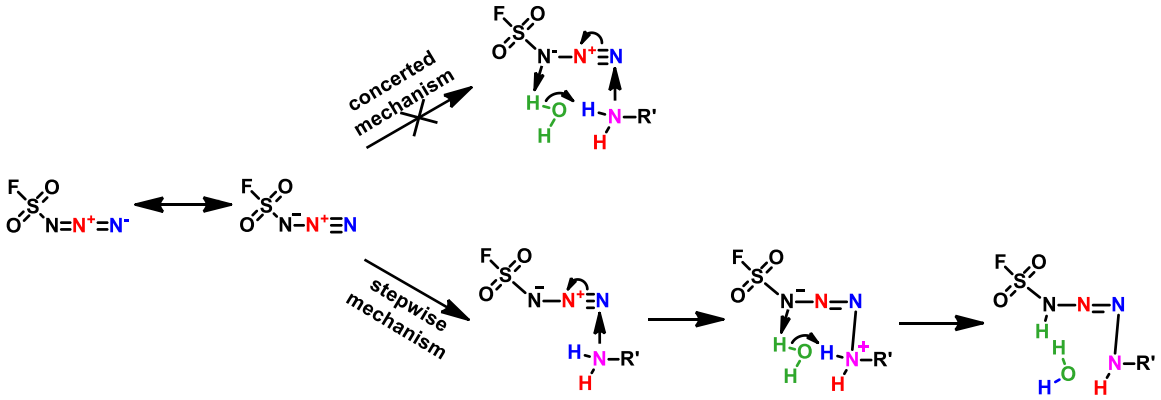


Figure 6. Possible mechanism for nucleophilic attack of amine on FSO_2N_3.

Shown on Figure 7-1 is the reaction energy profile calculated for the addition step, with the distance of the amine N—N-3 bond changing from 2.4 Å to 1.5 Å over 10 steps and an energy barrier less than 5 kcal/mol. Next, we used the final structure in the addition step as the starting point to calculate for its reaction energy profile of the hydrogen atom transfer step. The results show that the activation energy of this step is about 15 kcal/mol, an energy barrier that could readily be overcome under room temperature conditions (Figure 7-2)[4]. The hydrogen transfer is relayed *via* a water molecule, with a seven-membered ring transition state.

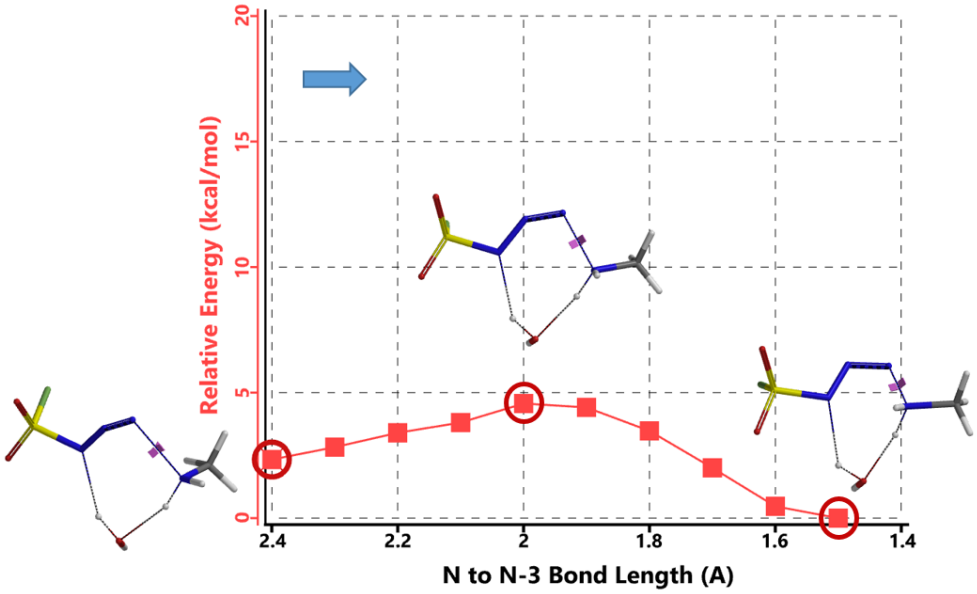


Figure 7-1. Reaction energy profile of amine addition to FSO_2N_3

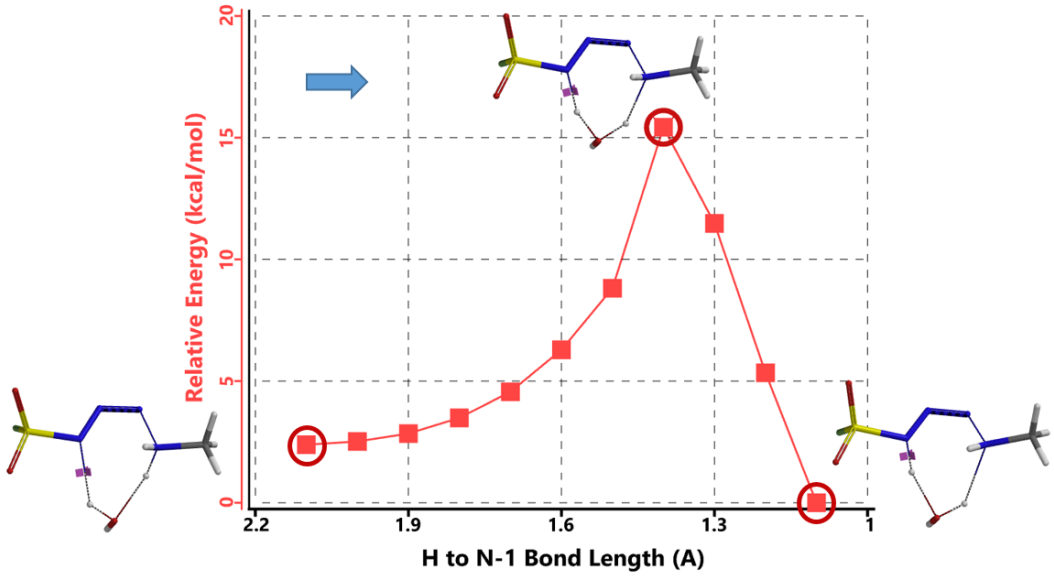


Figure 7-2. Reaction energy profile of hydrogen transfer *via* a water molecule

We selected the structures at the highest energy point from the two reaction energy profiles for more accurate Transition State and imaginary frequency calculations[5]. Resulting transition state structures are shown on Figure 8, with their respective imaginary frequency of 136 and 1280 cm^{-1} that are corresponding to the bonds being made and broken, satisfying the criteria for them to be considered as transition states.

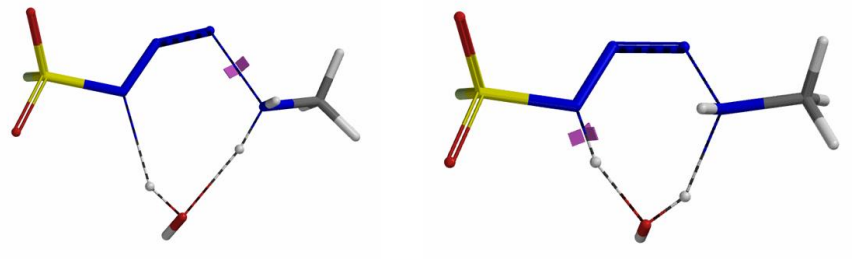


Figure 8. Transition state structures for amine addition to N-3 of FSO_2N_3 and hydrogen atom transfer steps. Imaginary frequencies 136 cm^{-1} (left) & 1280 cm^{-1} (right)

Modeling of the Second Step

Next, we calculated for the reaction energy profile of the N-1—N-2 bond cleavage of the amine-FSO_2N_3 addition complex and the transfer of the second hydrogen atom from amine to N-1 step, *via* a water molecule. Result is supportive of a concerted mechanism with an activation energy of about 20 kcal/mol (Figure 9). This step has the highest energy barrier for the whole diazo transfer reaction, indicating that this is the Rate Limiting Step of the reaction. Calculated transition state has a single ifrequency at 239 cm^{-1}, offering further support to our water catalysis model.

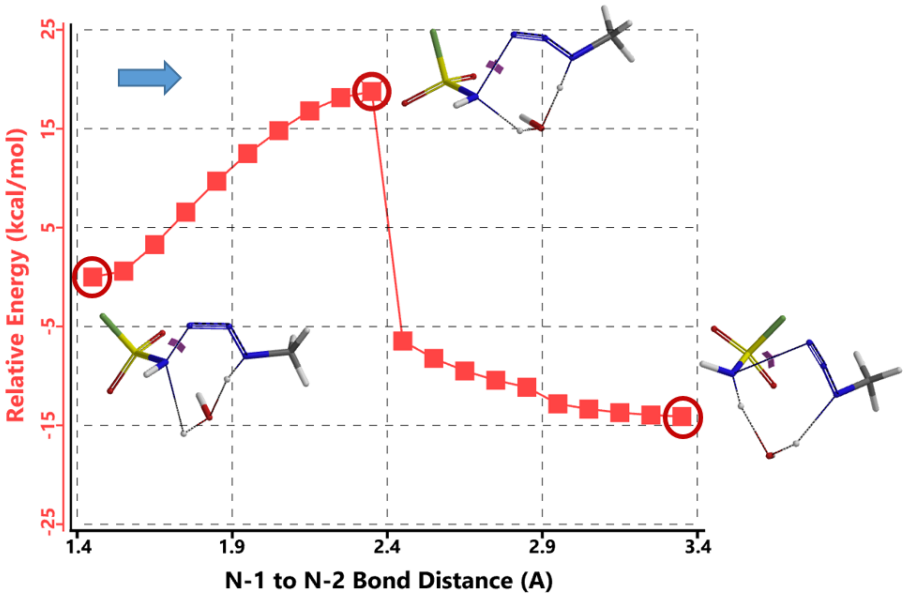


Figure 9. Reaction energy profile of step 2: N-1—N-2 bond of cleavage and second hydrogen transfer *via* water

Shown on Figure 10 is an overall reaction energy profile of the reaction. We set the relative energy of the starting species as zero, and included the key reaction intermediates and transition states in the energy plot. The energy barrier of each step of the reaction is relatively low, consistent with room temperature reaction conditions. The reaction is exothermic, ~22 kcal/mol. Reversing the reaction from the products side will encounter an energy barrier of >30 kcal/mol, consistent with the observation that this diazotransfer reaction is not reversible at room temperature.

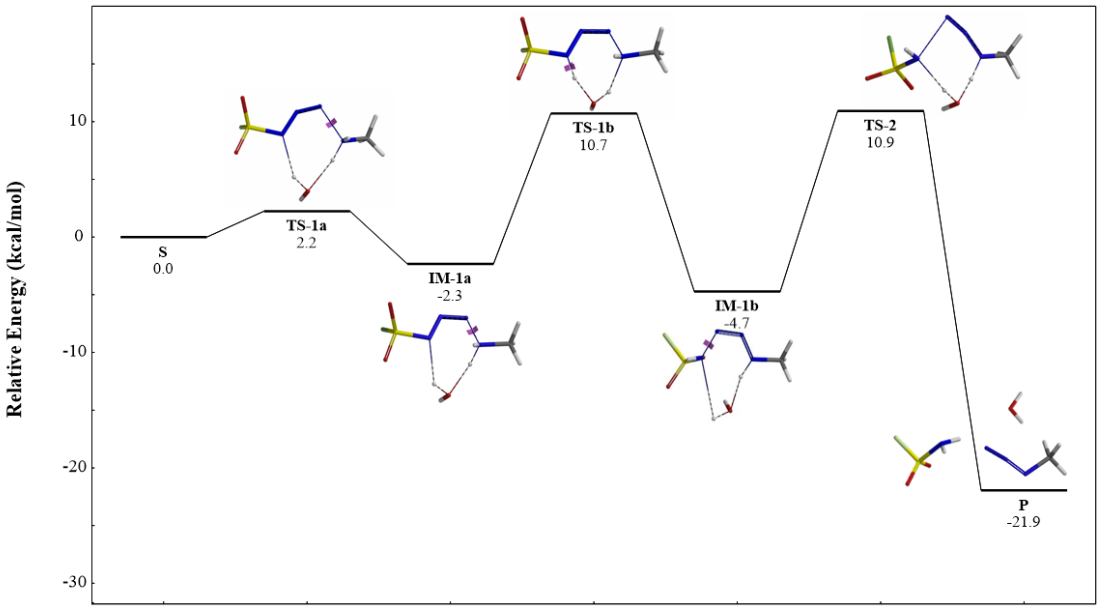


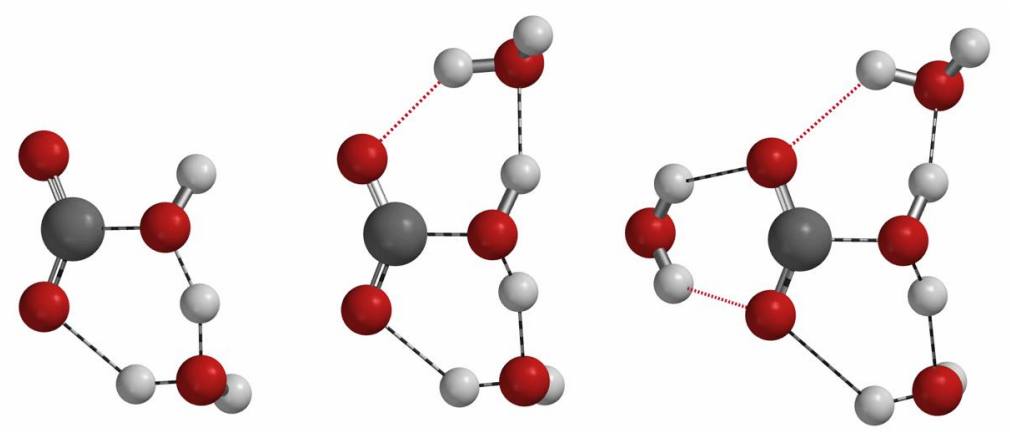
Figure 10. Overall reaction energy profile of FSO_2N_3 diazotransfer reaction

In summary, to establish our current QM model of this FSO_2N_3 diazotransfer reaction, we calculated reaction energy profiles for various possible mechanisms that could account for the key experimental observations. We conducted further transition state and imaginary frequency calculations to ensure that they meet the essential criteria. A useful protocol for qualitative QM study of reaction mechanism of interest.

Building on What We Just Learned

$$n\ H_2O + CO_2 \rightarrow n\text{-}1\ H_2O + H_2CO_3?$$

Reaction of carbon dioxide with water to form carbonic acid is a process of important geological and biological significances. Naturally, we are curious to find out how many water molecules are involved in the formation of a molecule of carbonic acid[6,7,8]. One, two, three, or four?



iFreq 732 cm-1 iFreq 1032 cm-1 iFreq 39 cm-1 (Irrelevant vibration)

Return to Table of Contents

References:

[1] X.Y. Jiang, X. Hao, L.L Jing, G.C Wu, D.W. Kang, X.Y. Liu, P. Zhan, *Expert Opin Drug Discov*, **2019**, *14*, 779.
[2] G.Y. Meng, T.J. Guo, T.C. Ma, J. Zhang, Y.C. Shen, K.B. Sharpless, J.J. Dong, *Nature*, **2019**, *574*, 86.
[3] A.K. Pandiakumar, S.P. Sarma, A.G. Samuelson, *Tetrahedron Lett*. **2014**, *55*, 2917.
[4] D.C. Young, *Computational Chemistry: A Practical Guide for Applying Techniques to Real-World Problems*, Wiley, New York (**2001**).
[5] See Chapter 22, reference 5 and footnote.
[6] M.T. Nguyen, G. Raspoet, L.G. Vanquickenborne, *J. Phys. Chem. A* **1997**, *101*, 7379 (Molecular Orbital Model: water self catalysis, n = 3).
[7] R.K. Lam, A.H. England, A.T. Sheardy, O. Shih, J.W. Smith, A.M. Rizzuto, D. Prendergast, R.J. Saykally, *Chem. Phys. Lett.,* **2014**, *614*, 282. (First Principles Molecular Dynamic Model and X-Ray Absorption Spectroscopy: hydration number = 3.17)
[8] S. Lindskog, *Pharmacol. Ther.,* **1997**, *74*, 1. (Zn^{++} Catalysis, n = 1)

Chapter 25 Assessing Reactivity with LUMO and HOMO Energy Gap

Dong Pan, Wenfeng Liu, Tommy Lai, Tao Wang, Yongsheng Chen, John S. Wai

In organic synthesis, we need to pair substrates and reagents of proper reactivity for the reactions to proceed smoothly. For example, halogenation of electron-rich aromatic substrates with N-chlorosuccinimide (NCS) usually works, while for electron-deficient ones, we'll need to use dichlorohydantoin (DCH) or even trichloroisocyanuric acid (TCCA). DCH and TCCA are stronger chlorination reagents than NCS. How much stronger? Is there a way to quantify the differences?

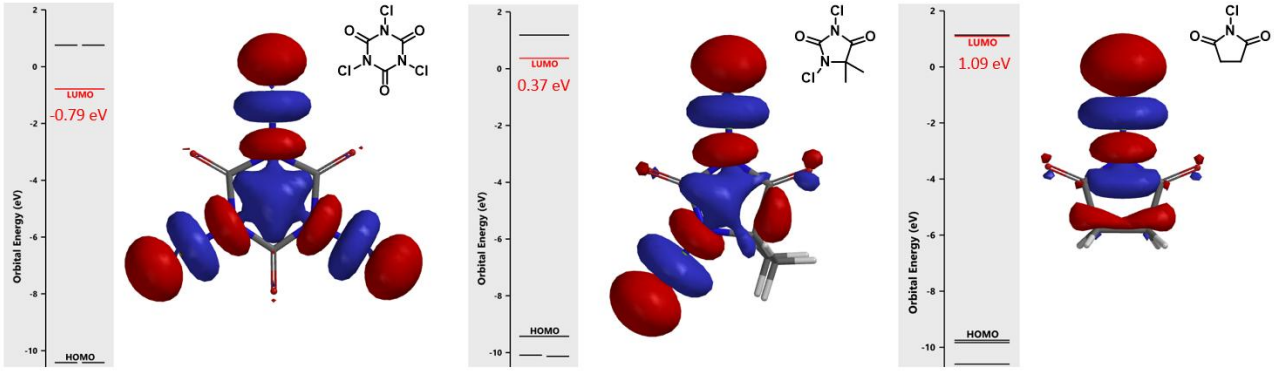


Figure 1. LUMO orbitals and energies of TCCA, DCH and NCS

The calculated LUMO orbitals and energies of TCCA, DCH, and NCS are shown in Figure 1. The LUMO lobes of these reagents are distributed around the chlorine atoms, not on the carbonyl carbons, consistent with the fact that they are halogenation reagents rather than acylation reagents. TCCA has the lowest LUMO energy of -0.79 eV, DCH has an intermediate energy of 0.37 eV, and NCS has the highest energy of 1.09 eV. These calculated LUMO energy numbers correlate with chemical reactivity of the reagents: the lower the LUMO energy, the more reactive it is as halogenation reagent.

Correlation between LUMO-HOMO energy difference and reactivity

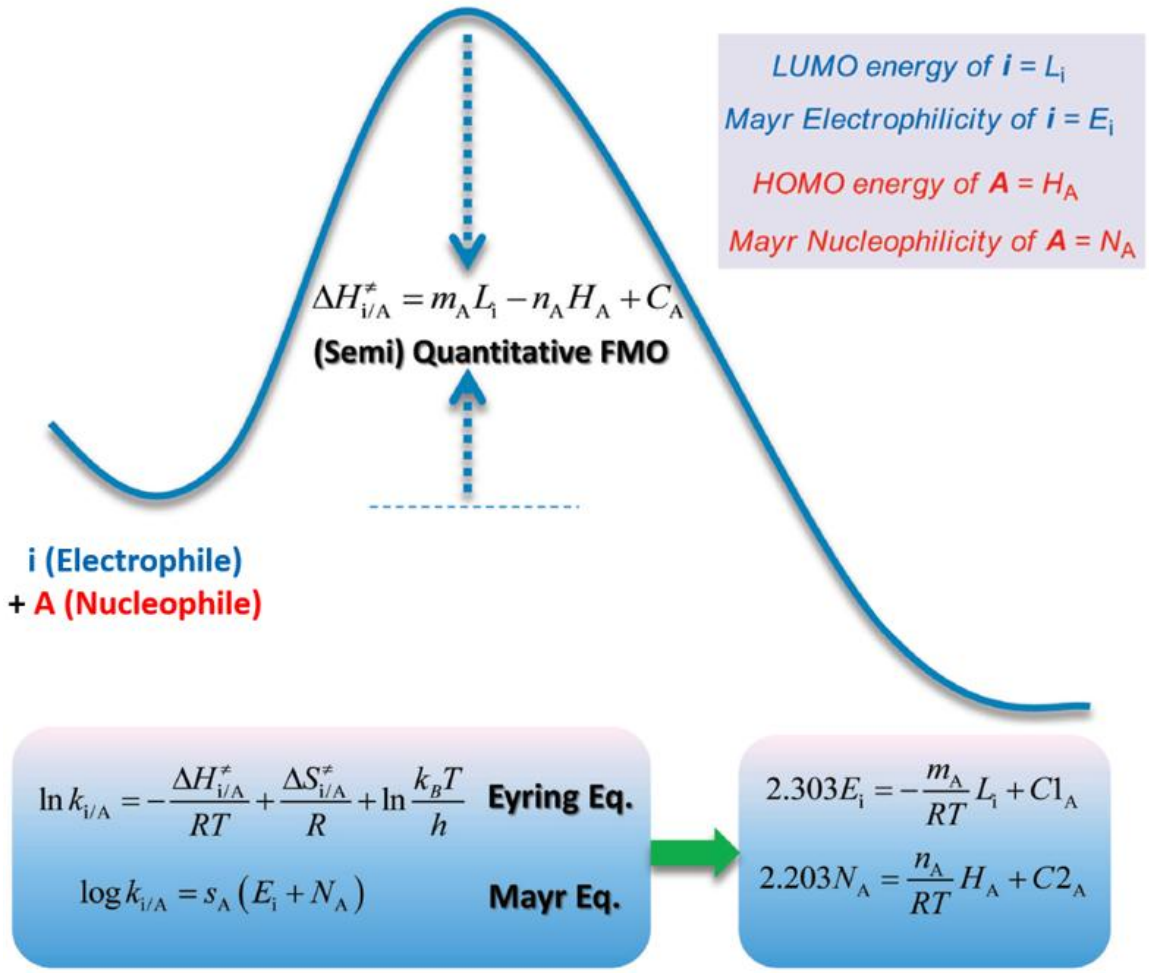


Figure 2. A Frontier Molecular Orbital Theory approach to understanding the Mayr Equation

Why LUMO energy could be used to gauge chemical reactivity? Zhi-Xiang Yu's group provided a theoretical understanding of the Mayr equation based on frontier molecular orbital theory (FMO) and the Eyring equation of the transition state theory, showing that the nucleophilicity of a molecule is related to the energy of this molecule's highest occupied molecular orbital (HOMO), while the electrophilicity is related to the energy of the lowest unoccupied molecular orbital (LUMO) of the electrophile (Figure 2)[1]. In other words, the nucleophilicity and electrophilicity of substrate/reagent can be assessed by their HOMO and LUMO energies, and whether the reaction will proceed or not could be judged by the difference in energy of the orbitals involved in the reactions.

Next, we'll make use of this LUMO-HOMO energy difference analysis to account for reactivity pattern observed with a set of substrates for Pictet-Spengler reaction.

Analysis of a set of substrates for Pictet-Spengler reaction

Pictet-Spengler reaction is the condensation of β-(hetero)aryl ethylamines with aldehydes or ketones under acidic conditions, followed by addition of the nucleophilic aromatic moiety of the molecule onto the resultant electrophile imine to provide the cyclization product (Figure 3, top).

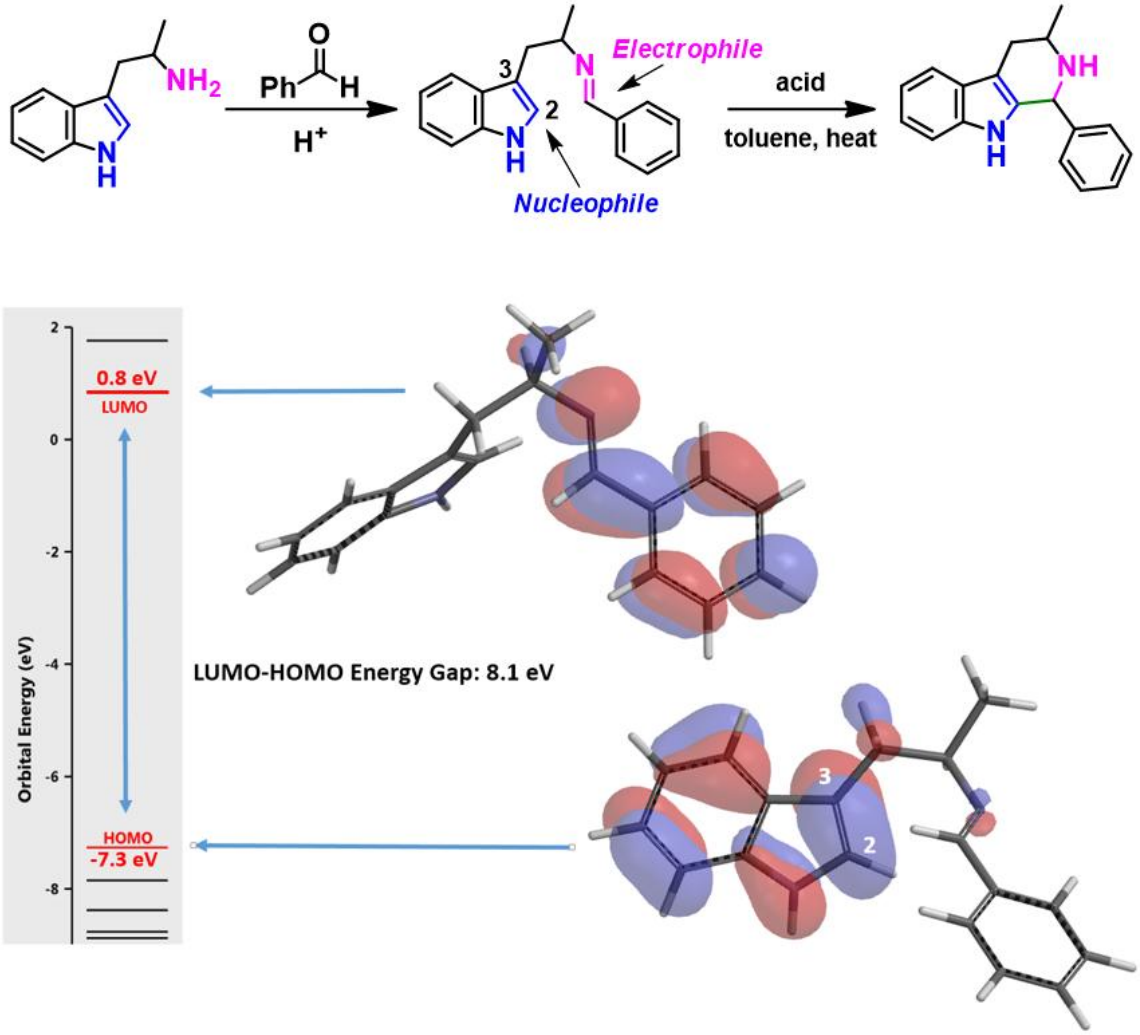


Figure 3. Pictet-Spengler reaction (top) and LUMO orbital, HOMO orbital and energy of imine (bottom)

As illustrated in Figure 3, there are HOMO lobe distributed between C2 and C3 of indole, and LUMO lobe on the imine carbon on the condensation intermediate. Interactions of these orbitals lead to the cyclization observed. Reactivity can be gauged by the energy difference between the LUMO and HOMO orbitals, which is 8.1 eV for the structure in Figure 3. If there is no lobe distributed on HOMO or LUMO at the reaction sites, HOMO-1/HOMO-2 or LUMO+1/LUMO+2 with appropriate lobe distribution should be considered for the energy gap estimation.

Next, we change the substitution position on the indole ring or replace the indole with various aromatic and hetero-aromatic systems, while holding the phenyl imino moiety unchanged. The LUMO and HOMO energies and their differences of this series of imine intermediates are shown in Table 1 (see Figure 4 for their structures).

Entry	E_{HOMO} (eV)	E_{LUMO} (eV)	E_{LUMO}-E_{HOMO} (eV)
1	-6.99	0.43	7.42
2	-7.26	0.84	8.10
3	-7.25	0.86	8.11
4	-7.57	0.80	8.37
5	-7.64	0.74	8.38
6	-7.68	0.80	8.48
7	-7.76	0.80	8.56
8	-7.82	0.90	8.72
9	-8.35	0.74	9.09
10	-8.32	0.83	9.15

Table 1. LUMO and HOMO orbital energies and their differences for a series of imine intermediates in Pictet-Spengler reaction

As shown in Figure 4, as the energy gap between LUMO and HOMO gradually increases, higher reaction temperature and/or stronger acid are required for the cyclization to occur[2]. When the orbital energy difference exceeds the threshold of about 9 eV, the reaction does not proceed under the conditions investigated.

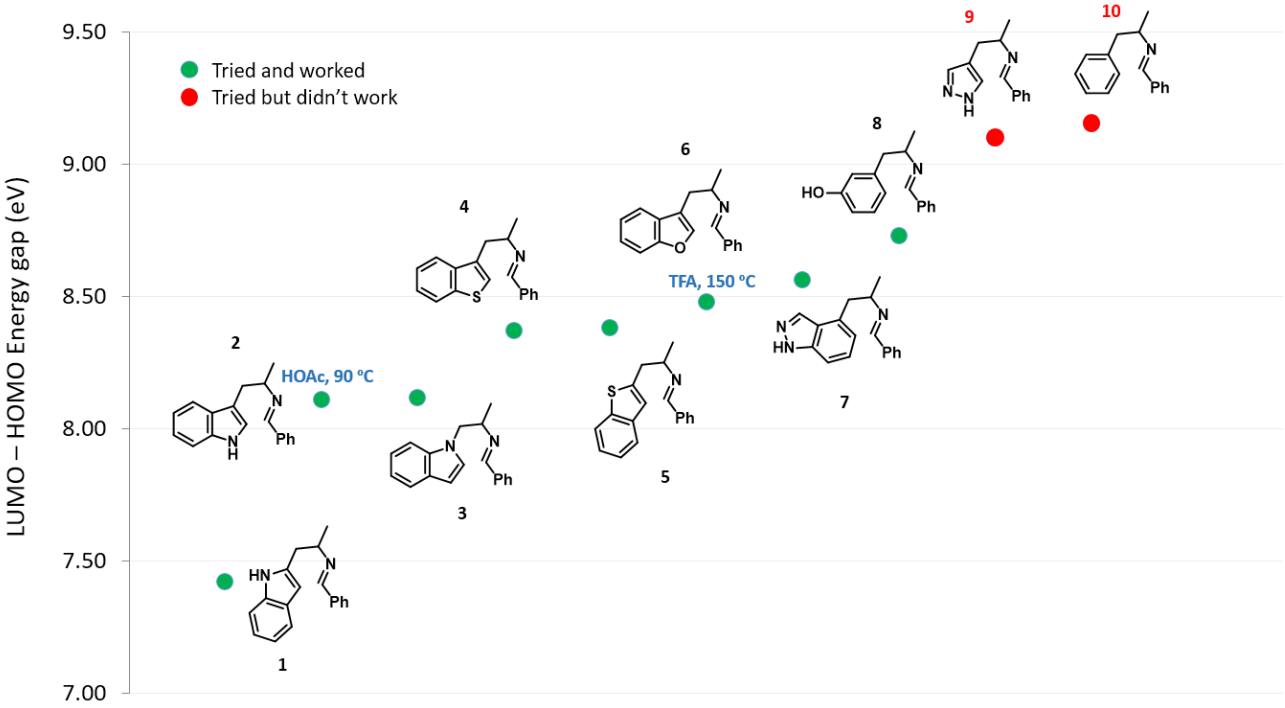


Figure 4. Correlation of LUMO-HOMO energy difference with reactivity in Pictet-Spengler reaction

The above correlation enables us to predict whether a substrate could undergo the Pictet-Spengler reaction or not, simply by comparing the energy gap calculated *versus* the threshold of 9.09 eV, established with the pyrazole analog **9** and phenyl analog **10**, and what cyclization conditions to use by referencing to examples on the Table/Figure.

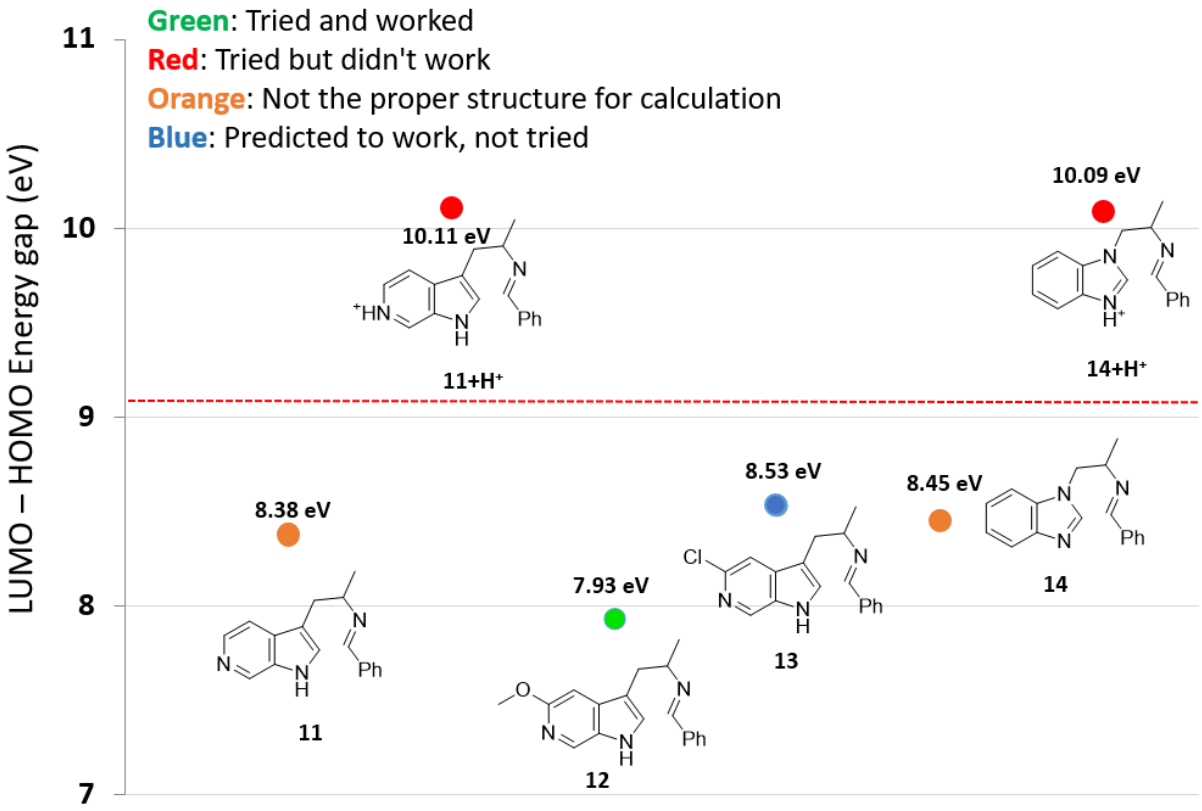


Figure 5. The LUMO-HOMO energy difference was used to predict the reaction activity of different substrates (Energy differences: **11, 12, 13** are based on LUMO and HOMO; **11+H⁺** is based on LUMO+1 and HOMO-2; **14** is based on LUMO and HOMO-1; **14+H⁺** is based on LUMO+2 and HOMO-3)

For substrates containing basic nitrogen, we also need to consider the p*K*a of the substrate. For 5-azaindole **11**, the LUMO-HOMO energy difference is 8.38 eV, which is within the range where the reaction can occur. However, the cyclization did not proceed and literature search did not reveal any successful example. We reasoned that in the TFA acidic reaction conditions, the pyridyl nitrogen is protonated (Fig. 5, pyridine p*K*a ~ 5.2), and the LUMO-HOMO energy difference for **11+H⁺** is actually 10.11 eV.

This challenge was readily resolved by introducing a methoxy or chloro substituent adjacent to aromatic N to lower its basicity (p*K*a of 3.2 and 0.7 for 2-MeO- and 2-Cl pyridines, respectively)[3]. The LUMO-HOMO energy differences of MeO-substituted substrate **12** and Cl-substituted substrate **13** are estimated to be 7.93 eV and 8.53 eV, respectively. We decided to prioritize the synthesis of the methoxy product first because the LUMO-HOMO energy difference of 7.93 eV is further away from the threshold of 9.09 eV, and successfully synthesized the methoxy Pictet-Spengler reaction product. The methoxy group was converted into triflate and hydrogenated to provide the 6-azaindole product[4]. The triflate intermediate also enables us to establish further SAR in that part of the inhibitors.

A similar hurdle was anticipated for the benzimidazole substrate **14+H⁺** with an estimated orbital energy gap of 10.09 eV. This alerted us to design an alternative synthetic route for the target[4].

In this chapter, we introduce a simple method for assessing feasibility of a reaction, by calculating for the energy difference between LUMO and HOMO of the reactant and/or substrate. Compared to the calculation of transition state, this method is simpler and faster,

and especially suitable for reactions that have a useful number of examples for correlation. It is worth noting that this method is very useful when the electronic effect is the main driving factor for the reaction, and requires us to choose the proper orbitals, with lobe at the reaction sites, for energy gap estimation. Potential steric interactions may require special attention.

Building on What We Just Learned

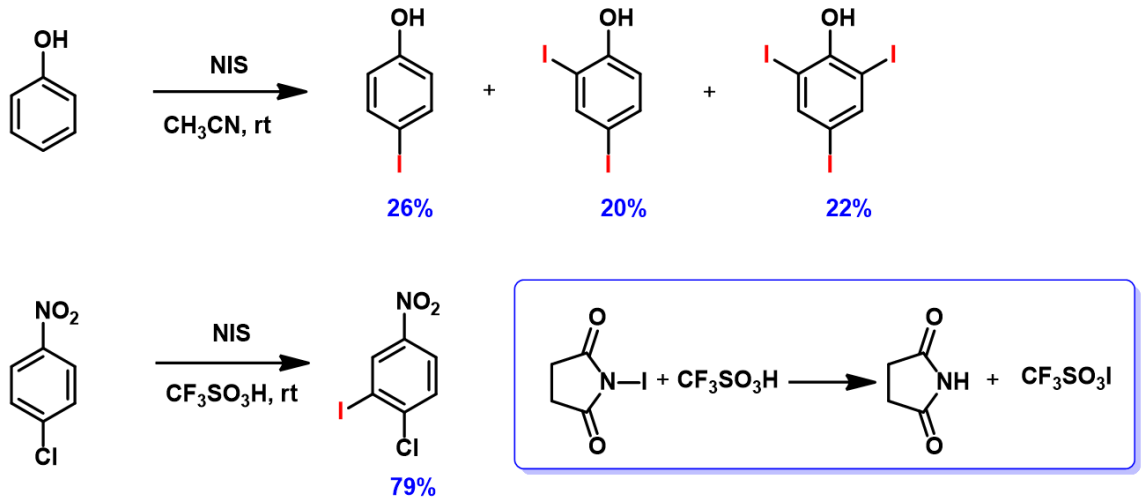
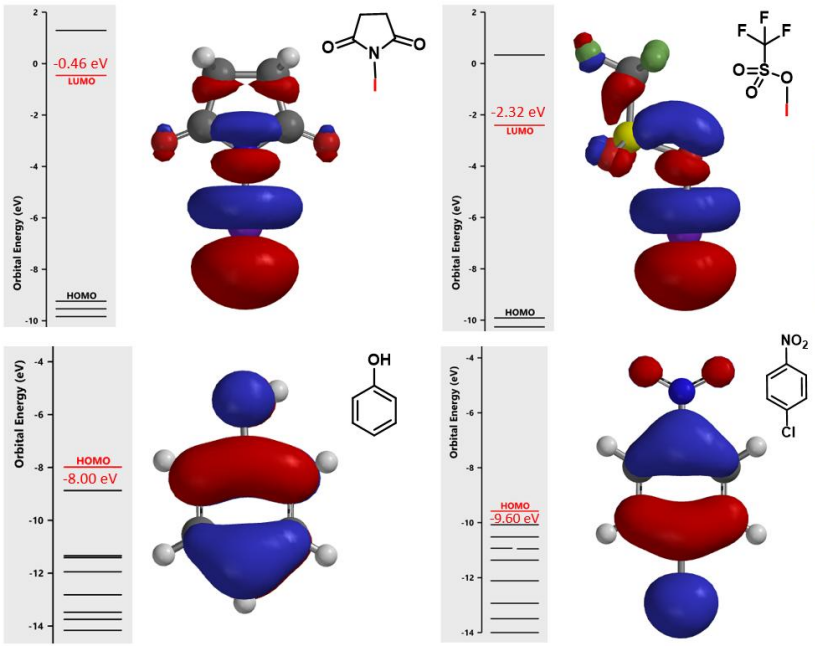


Figure 6. Iodination of phenol and 4-chloronitrobenzene

Reaction of phenol with N-iodosuccinimide (NIS) readily proceeds to provide a mixture of mono-, di- and tri-iodinated products[5]. However, iodination of 4-chloronitrobenzene with NIS proceeded only in the presence of trifluoromethanesulfonic acid[6]. It is reasoned that trifluoromethanesulfonic acid reacts with NIS to generate iodine(I) trifluoromethane-sulfonate, which is the active species for the halogenation (Figure 7). What reagent will you choose to iodinate 1,2,3-trifluoro-4-nitrobenzene? And will halogenation occur at C-5 or C-6 position? [7]



Reagent	Substrate	E_{LUMO}-E_{HOMO} (eV)
NIS	PhOH	7.54
NIS	$C_6H_4NO_2Cl$	9.14
CF_3SO_2I	$C_6H_4NO_2Cl$	7.28

Figure 7. LUMO orbitals of NIS and CF_3SO_3I and HOMO orbitals of phenol and 4-chloronitrobenzene with corresponding energies

Return to Table of Contents

References:

[1] L.G. Zhuo, W. Liao, Z.X. Yu, *Asian J. Org. Chem.* **2012**, *1*, 336.
[2] A. Yokoyama, T. Ohwada, K. Shudo, *J. Org. Chem.* **1999**, *64*, 611.
[3] https://www.chem.wisc.edu/areas/reich/pkatable/pKa_compilation-1-Williams.pdf
[4] M.K. Jackl, I. Kreituss, J.W. Bode, *Org. Lett.* **2016**, *18*, 1713.
[5] P. Bovonsombat, J. Leykajarakul, C. Khan, K. Pla-on, M.M. Krause, P. Khanthapura, R. Ali, N. Doowa, *Tetrahedron Lett.* **2009**, *50*, 2664.
[6] G.A. Olah, Q. Wang, G. Sandford, G.K. Surya Prakash, *J. Org. Chem.* **1993**, *58* , 3194.
[7] Tips: Calculate for HOMO map and HOMO-1 map of 1,2,3-trifluoro-4-nitrobenzene.

Chapter 26 Activation Energy Estimation for Alkylation of Pyrazole (Part II)

Guijun Liu, Qiuyue Wang, Zhong Zheng, Shouliang Wang, Yongsheng Chen, John S. Wai

In Chapter 9, we discussed the application of QM to calculate for the difference in activation energy of competing alkylation paths of pyrazole, estimate product ratio, and use the result to guide our synthetic planning and execution.

For assessment of competing alkylation of pyrazole **1** (Figure 1), we usually set up two calculations in parallel using CH_3Br as generic alkylating reagent to speed up the process: one approaches **N1**, and the other interacts with **N2**, to estimate their activation energies for comparison. In this case, the calculated activation energies are 6.4 kcal/mol for **N1** and 9.4 kcal/mol for **N2** methylation, in favor of **N1** alkylation (Figure 2).

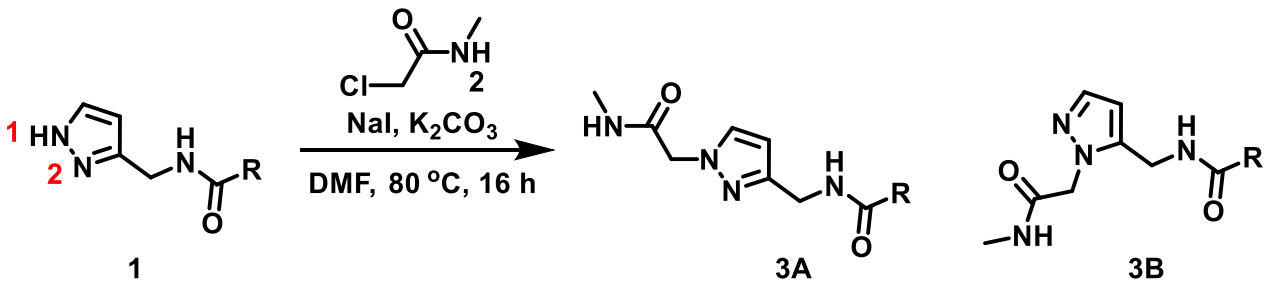


Figure 1. Alkylation of pyrazole **1**

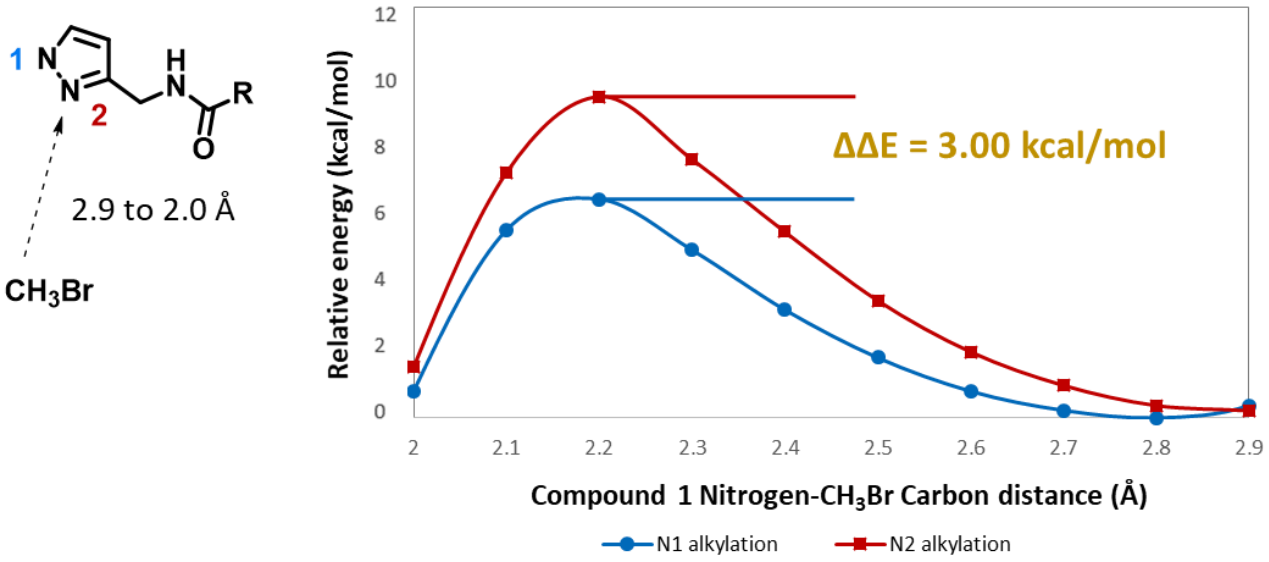
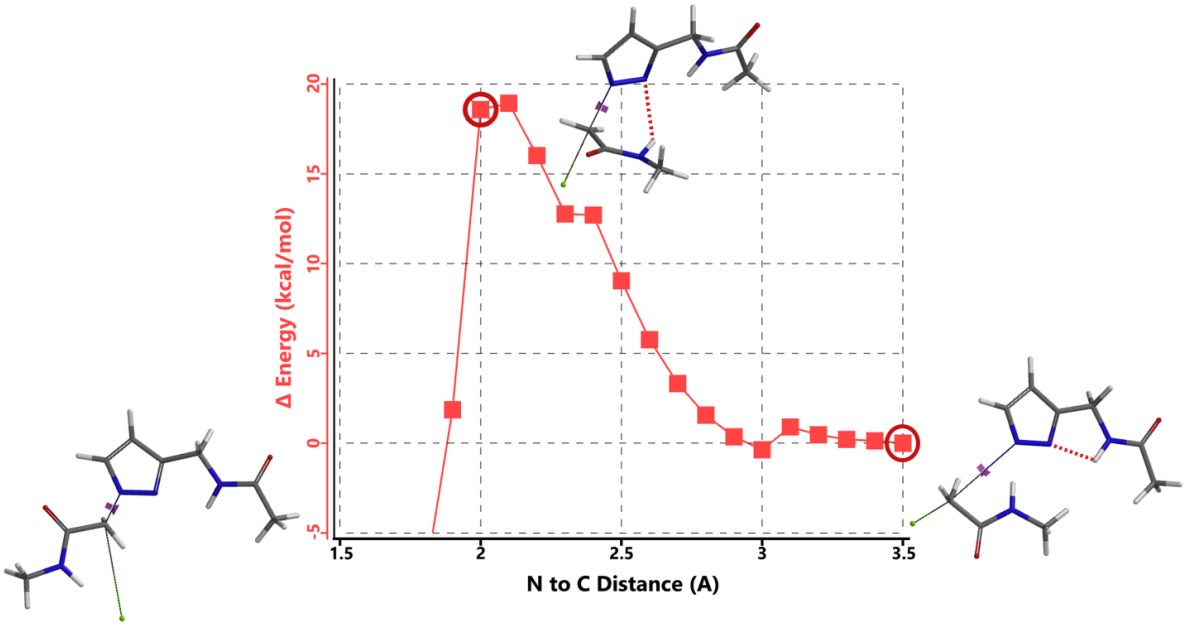


Figure 2. Reaction energy profiles of **N1**, **N2** alkylation of pyrazole **1** using CH_3Br as generic electrophile

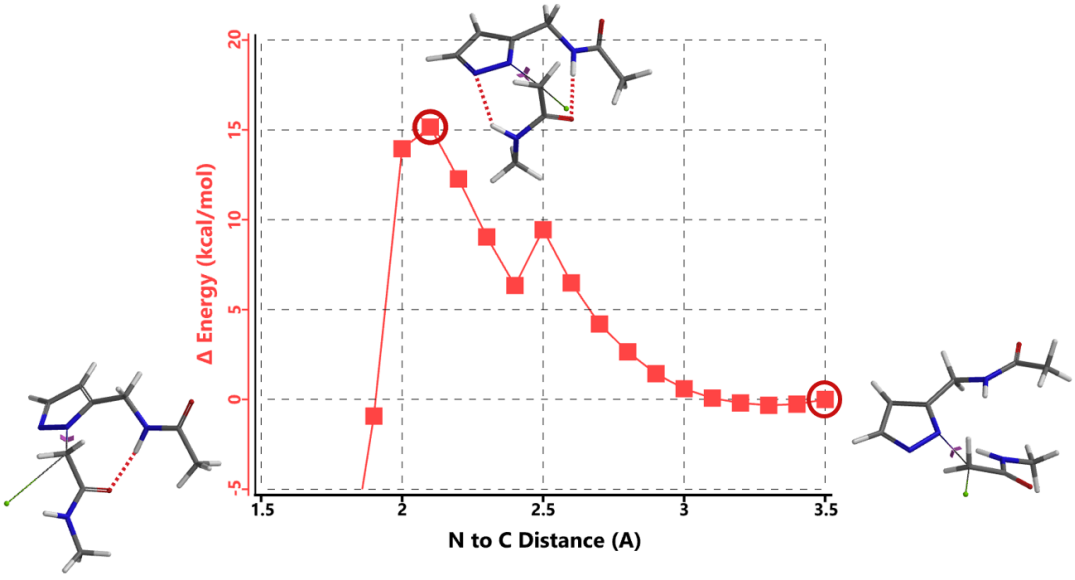
What could have caused a change in N1 *vs* N2 selectivity?

With N-methyl chloroacetamide used as alkylating reagent in our synthesis, only **N2** alkylation product **3B** was observed, different from what the calculations with methyl bromide suggested.

When we repeated the calculations with the chloroacetamide, the activation energy estimated is 18.0 kcal/mol for **N1** methylation and 15.0 kcal/mol for **N2** alkylation, suggesting a reversal in **N1** *vs* **N2** selectivity, different from the one modeled with methyl bromide. The 3 kcal/mol energy difference translates to an anticipated **3A**:**3B** product ratio of ~1:70 at the reaction temperature of 80 °C, consistent with experimental observations (Figure 3). What could have caused a change in **N1** *vs* **N2** selectivity?



Alkylation at N1



Alkylation at N2

Figure 3. Reaction energy profiles of **N1**, **N2** alkylation of pyrazole **1** with N-methyl chloroacetamide (.....Hydrogen bond)

In the calculated transition state structure of **N1** alkylation, a hydrogen bond (2.42 Å) between chloroacetamide NH and the pyrazole **N2**[2] is observed. While for the transition state structure of **N2** alkylation, in addition to the comparable hydrogen bond (2.41 Å) between chloroacetamide NH and the pyrazole **N1,** a second, stronger, hydrogen bond (1.91 Å) is observed between the side chain amide NH and the carbonyl oxygen on the alkylating reagent (Figures 4 and 5). The latter hydrogen bond offers further stabilization to the transition state for **N2** relative to **N1** alkylation, accounting for the **N2** selectivity.[3] It is worth noting that both reaction energy profiles show characteristic bumps at N-C distance ~2.5 to 2.3 Å, due to change in Cl-C-C-O dihedral angle of the incoming chloroacetamide, *i.e.* the chloro group changing from being close to coplanar relative to the amide to close to perpendicular in transition state.[4]

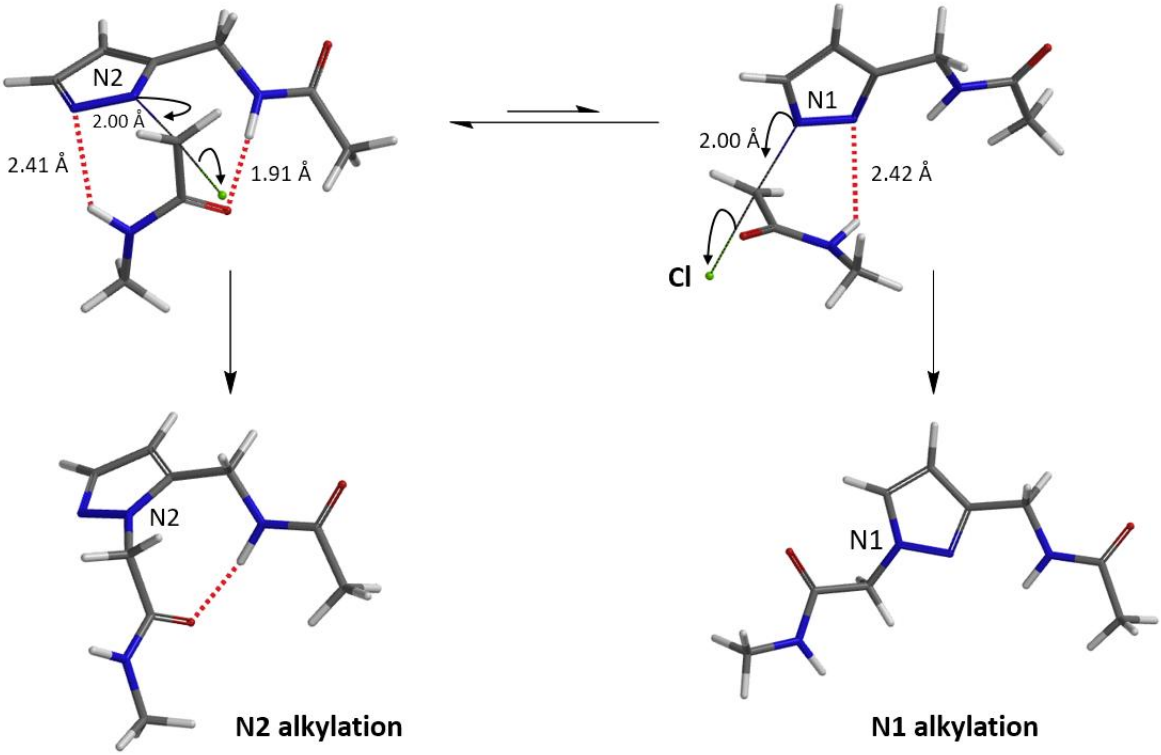


Figure 4. Hydrogen bonding introduced by chloroacetamide as alkylation reagent (.....Hydrogen bond)

Transition state (Figure 5) calculations showed one and only one imaginary frequency which corresponds to the bonds being made and broken, supporting the use of these structures to account for the selectivity observed.

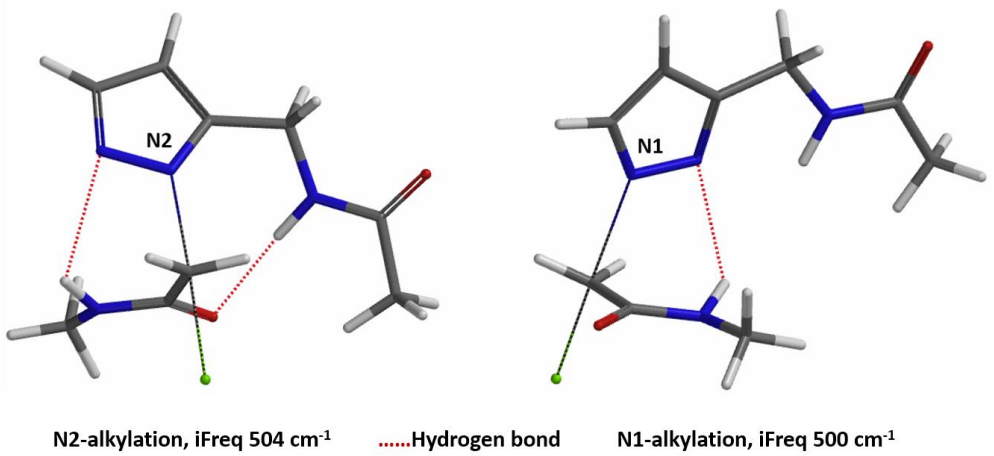


Figure 5. Transition state structures of **N2** & **N1** alkylation, and imaginary frequencies

To further evaluate the hypothesis that these hydrogen bonds are indeed key to the **N2** selectivity observed with N-methyl chloroacetamide, we replaced it with methyl chloroacetate for calculations. Reaction energy profiles showed little hydrogen bond interaction in the transition states, with estimated activation energy of 9.5 kcal/mol for **N1** and 11.6 kcal/mol for **N2** alkylation, contrasting to the selectivity observed with chloroacetamide, substantiating our explanation.

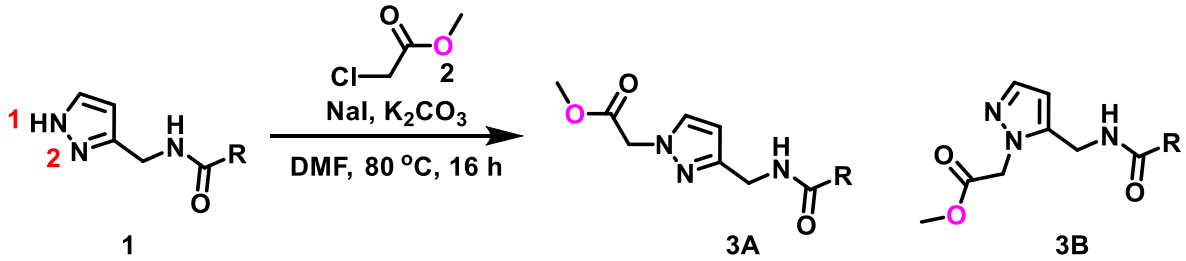
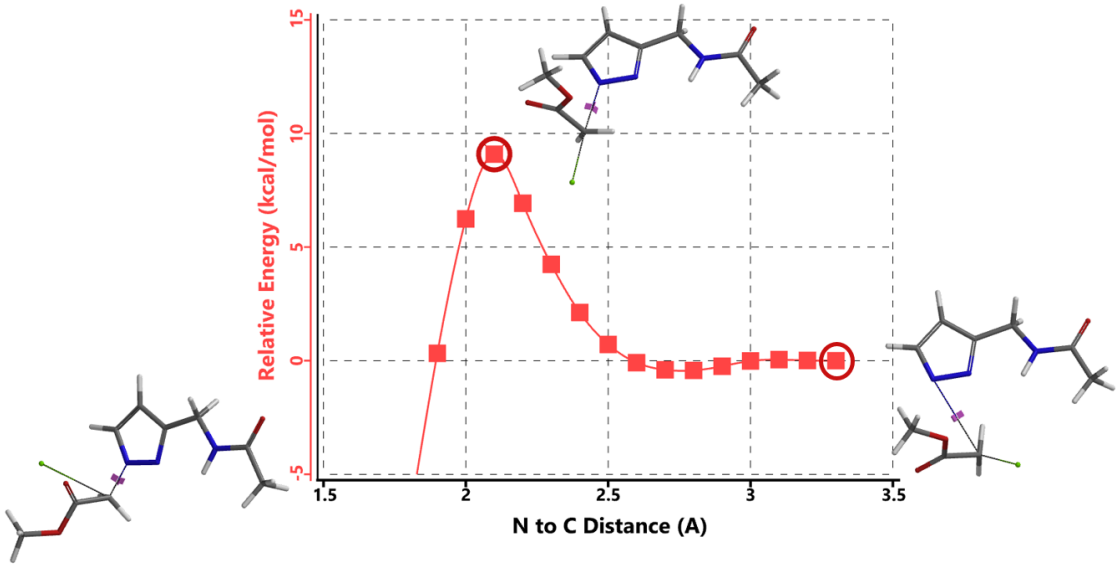
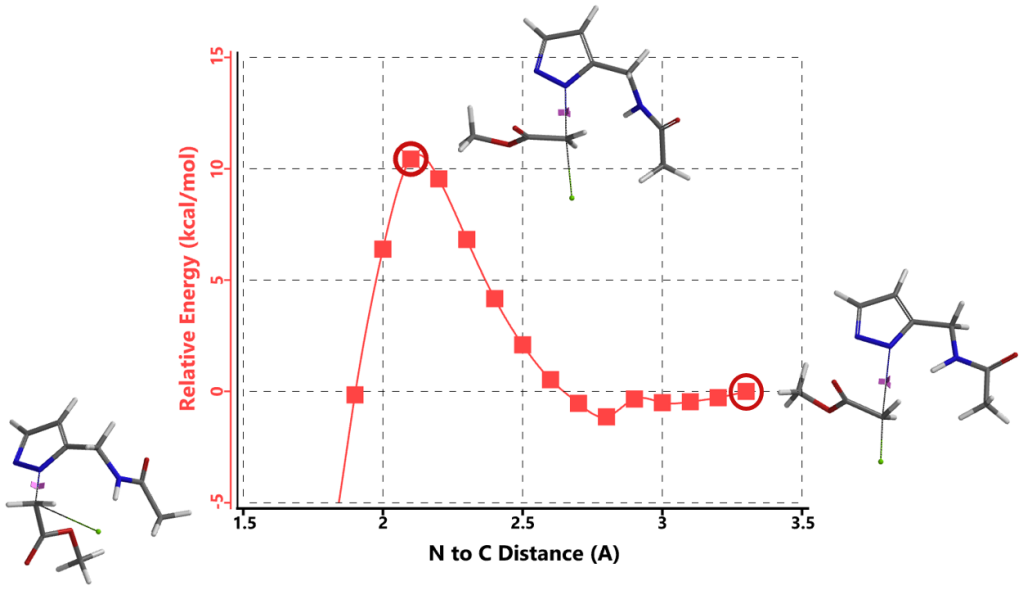


Figure 6. Methyl chloroacetate as alkylation reagent



Alkylation at N1



Alkylation at N2

Figure 7. Reaction energy profiles of **N1**, **N2** alkylation of pyrazole **1** with methyl chloroacetate

Importance of relevant models

The application of relevant computational models for prediction of experimental outcome requires careful consideration of many factors: are surrogate structures representative of the reactions, will there be any intramolecular or intermolecular interactions that are crucial in determining the outcome, do we need to include explicit solvent molecules in the calculations (Chapters 15 and 24), etc. Chemical insights do not automatically present themselves as the results of a calculation, but require additional human effort. It is very important that our retrospective QM calculations could account for the experimental results, before they could be considered for prospective analyses.

Building on What We Just Learned

We are interested in understanding the differences in reactivity between thiol and N,N-dimethyl chloroacetamide *versus* N-methyl chloroacetamide. LUMO and LUMO maps of two amides are shown below in Figure 8. Which one will you expect to react faster with thiol? Hydrogen bond interaction could play a role too? Will it speed up or slow down the reaction?

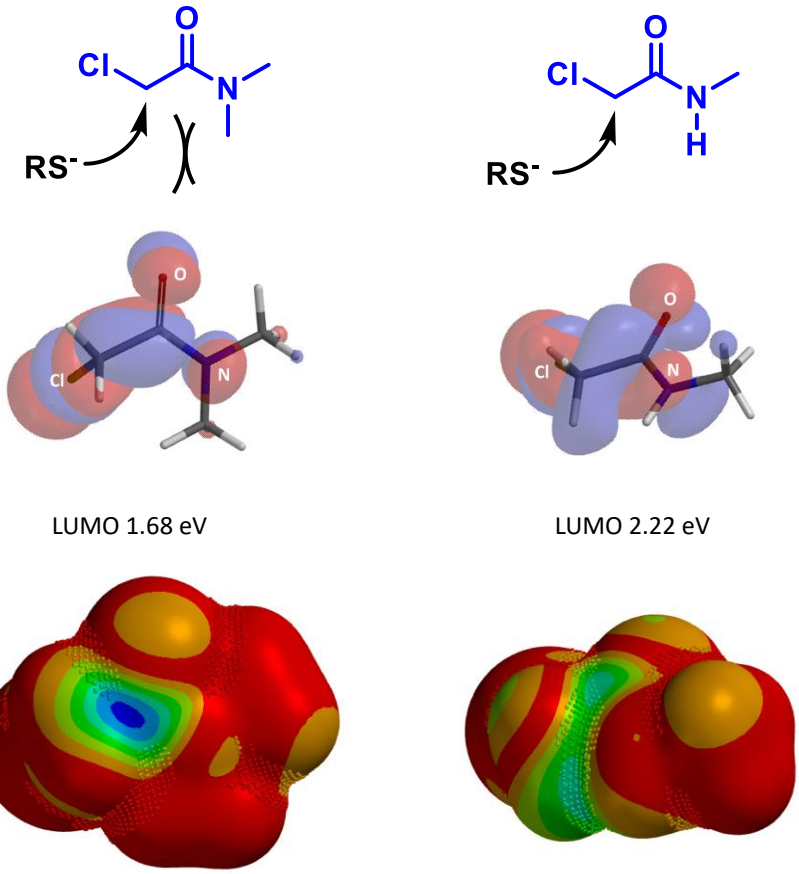


Figure 8. LUMO and LUMO map of N,N-dimethyl chloroacetamide and N-methyl chloroacetamide

Return to Table of Contents

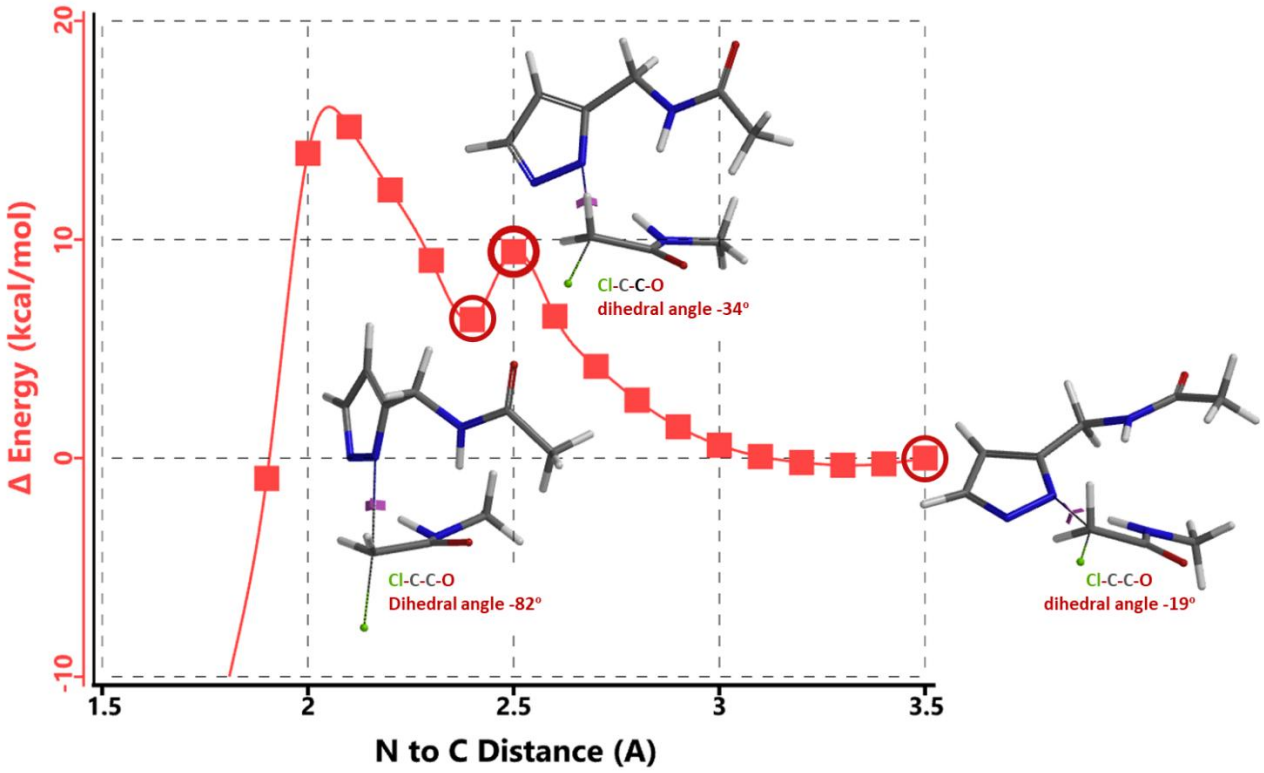
References:

[1] W. Hehre and S. Ohlinger, *A Guide to Molecular Mechanics and Quantum Chemical Calculations*. Irvine, CA, USA: Wavefunction, Inc., **2003**.

[2] Spartan '20 *Tutorial and User's Guide*. Irvine, CA, USA: Wavefunction, Inc. **2021**; p256.

[3] D.C. Young, *Computational Chemistry: A Practical Guide for Applying Techniques to Real-World Problems*, Wiley, New York (2001).

[4] Changes in Cl-C-C-O dihedral angle of the incoming N-methyl chloroacetamide



Chapter 27 Correlating Reactivity Trends with Frontier Molecular Orbitals

Zhengquan Zhou, Dong Pan, Liting Dong, Tommy Lai, Yongsheng Chen, John S. Wai

In chapter 19 on *tele*-substitution of 2,3-dichloropyrazine, we learned that dichloropyrazine uses LUMO+1, not LUMO, for its nucleophilic substitution. In this chapter we will have further examples on choosing the appropriate LUMO or LUMO+1 and HOMO or HOMO-1 to correlate with electrophilic and nucleophilic reactivity trends observed in heterocyclic chemistry[1].

Correlating Electrophilic Reactivity with LUMO Energy

Shown in Figure **1** is the relative reactivities in nucleophilic substitution described in *Heterocyclic Chemistry*[2] for chlorodiazines and chloropyridines.

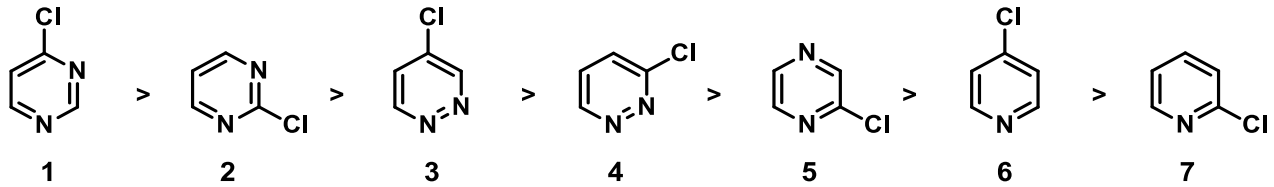


Fig. 1. Relative reactivities for chlorodiazines and chloropyridines

Selecting the appropriate LUMO/LUMO+1 for comparison is the **key** in correlating reactivity of observed. As shown in Figure **2**, 4-chloropyrimidine **1** and 4-chloropyridine **6** have LUMO lobes centered on the C-Cl carbons, little to none in LUMO+1; as such their LUMO are used for correlation. In contrast, for 2-chloropyrimidine **2**, 3-chloropyridazine **4**, 2-chloropyrazine **5**, and 2-chloropyridine **7**, their LUMO have almost no lobe centered on the C-Cl carbons, while the LUMO+1 have significant ones; therefore their LUMO+1 are selected.

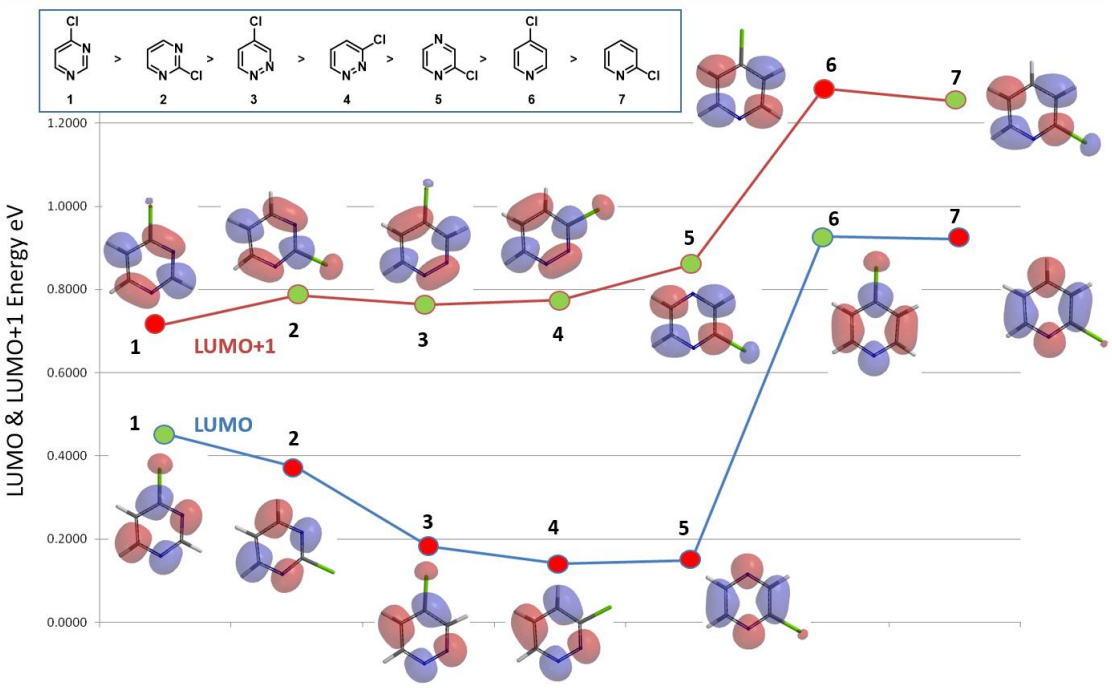


Figure 2. LUMO/LUMO+1 and associated energies of selected chlorodiazines and pyridines

4-Chloropyridazine **3** is unique: its LUMO and LUMO+1 have energy gap of ~ 0.6 eV, and both have lobes slightly off-center on the C-Cl carbon. How do we choose the proper one to correlate?

1) **Observed experimental difference**: 4-Chloropyridazine **3** is less reactive than 2-chloropyrimidine **2**; as such its LUMO+1, of comparable energy, is selected for correlation.

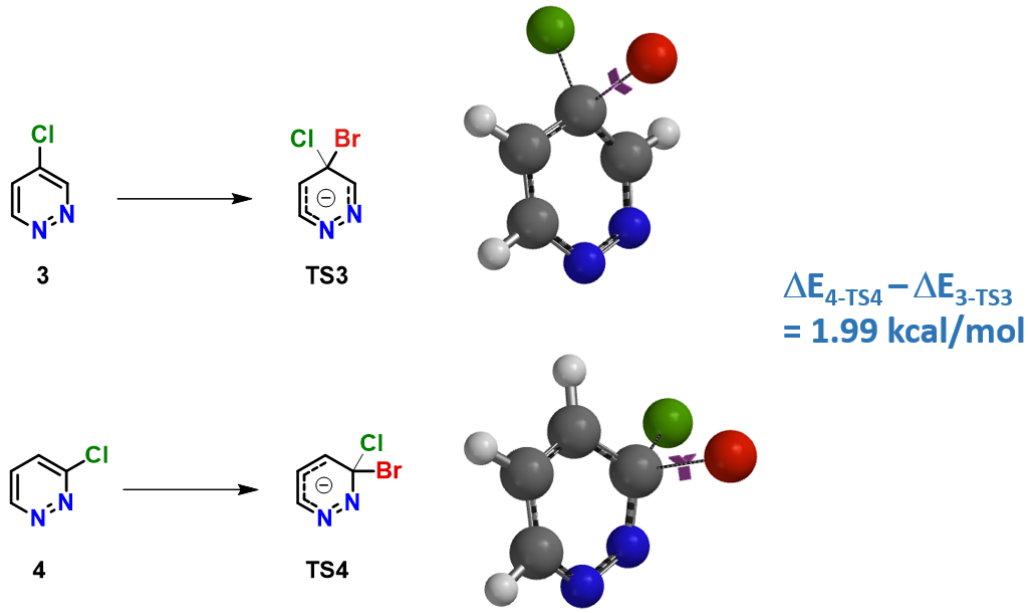


Figure 3. TS of nucleophilic attack of 4- & 3-chloropyridazine **3** & **4** with bromide and their energy difference

2) **Activation energy difference**: Activation energies for bromide substitution of 4- & 3-chloropyridazine **3** & **4** are calculated (Figure **3**). $\Delta E_{4\text{-}TS4}$ is found to be higher than $\Delta E_{3\text{-}TS3}$ by 1.99 kcal/mol, consistent with the observed difference in reactivity, and supportive of picking LUMO+1 of pyridazine **3** for reactivity correlation.

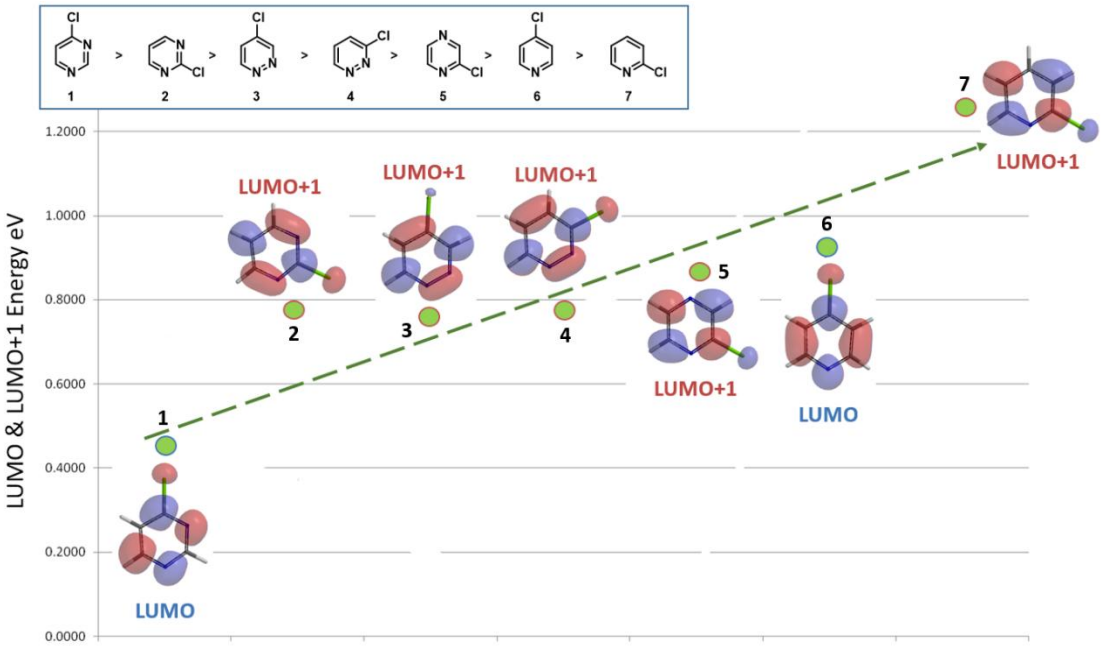


Figure 4. Correlation of reactivities with proper LUMO and LUMO+1 of chlorodiazines and pyridines

With proper LUMO/LUMO+1 selection of the above seven chlorodiazines and pyridines, a direct correlation with their reactivity becomes obvious (Figure **4**). However, a new puzzle emerges: "For 4-chloropyridazine **3**, LUMO lobe for nucleophilic substitution is available, why does the reaction use LUMO+1 instead?"

Correlating Nucleophilic Reactivity with HOMO Energy

Next, correlating reactivity of eight common 5-membered and 6-membered heterocycles with their HOMO/HOMO-1 in electrophilic substitution reaction (Figure **5**).

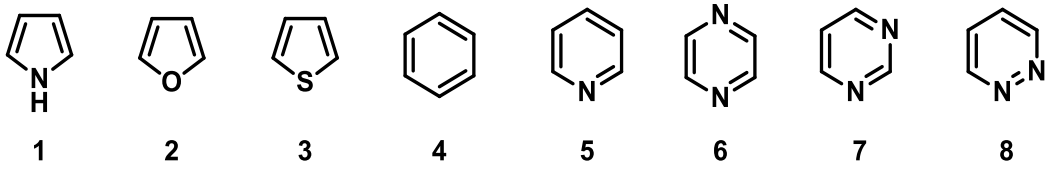


Figure 5. Eight common 5- and 6-membered heterocyclic compounds

As shown in Figure **6**, all the five-membered heterocyclic compounds, pyrrole **1**, furan **2**, and thiophene **3**, have obvious HOMO lobes for electrophilic substitution. Moreover, lobes on C2 are significantly larger than on C3, consistent with higher reactivity observed at C2. Benzene **4** is C_6 symmetrical: HOMO and HOMO-1 have the same energy (-8.78 eV). Each of its HOMO lobes is distributed over two carbon atoms, while the HOMO-1 ones are concentrated over one carbon atom.

For pyridine **5**, pyrazine **6**, pyrimidine **7**, and pyridazine **8**, there are significant HOMO lobes on the nitrogen atoms, while the HOMO-1 lobes are mainly distributed on the carbon atoms, with little or no distribution on the nitrogen atoms. As such their HOMO-1 shall be selected for the comparison for electrophilic reactivity occurring on the carbons.

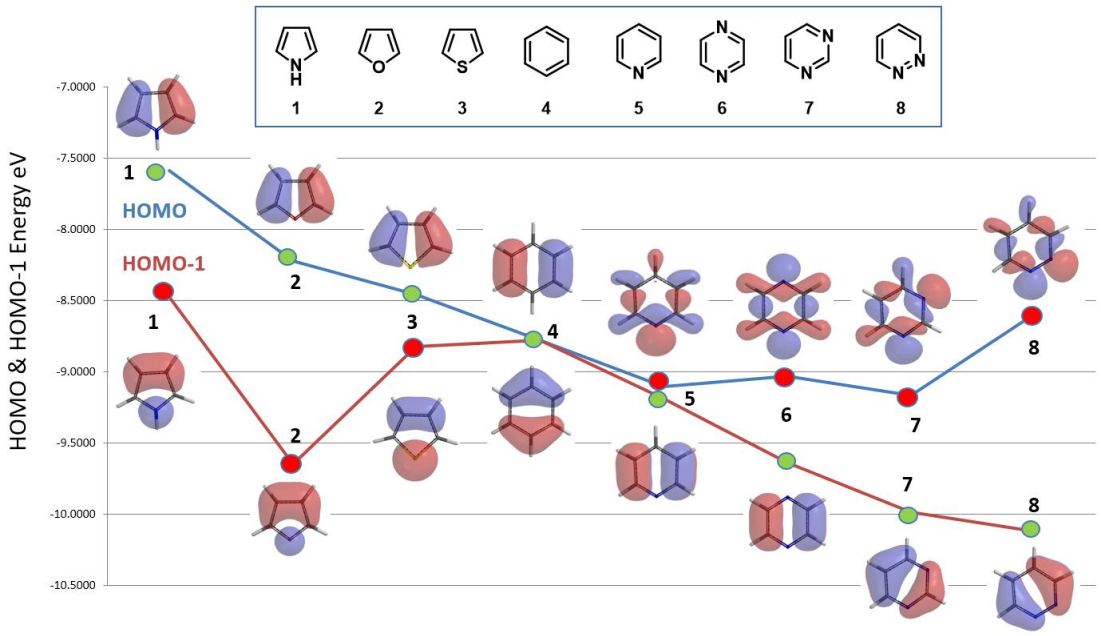
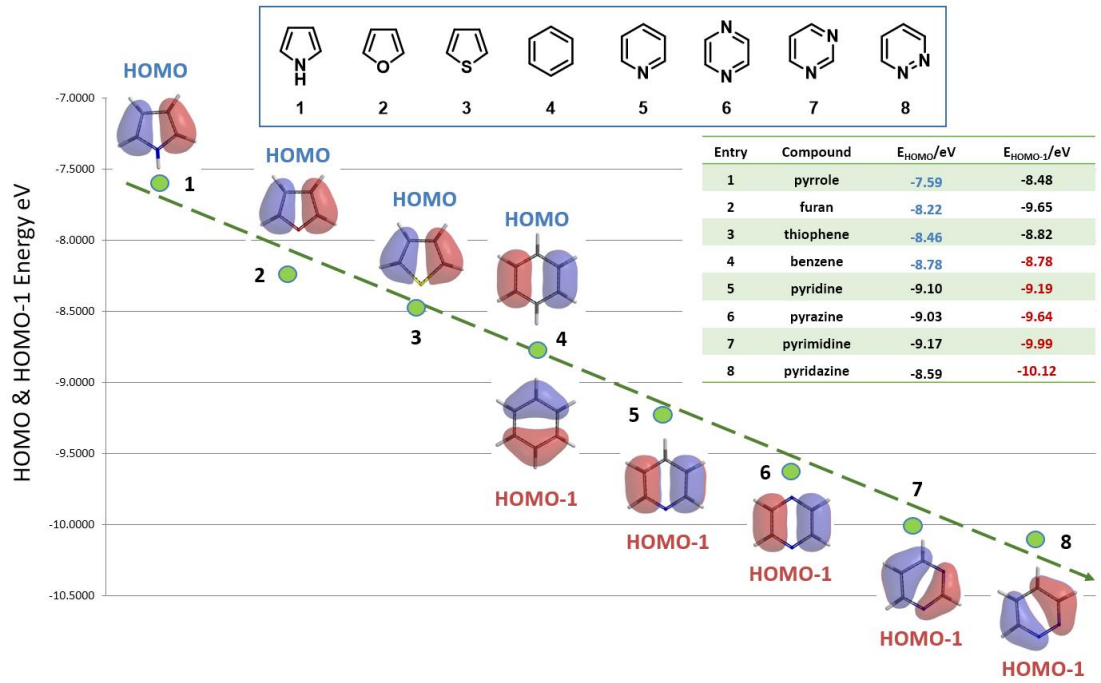


Figure 6. HOMO and HOMO-1 and associated energies of selected 5- & 6-membered heterocycles

As shown in Figure **7**, reactivities of the eight 5- & 6-membered heterocycles correlate well with their proper HOMO/HOMO-1 energies. The five-membered heterocyclic compounds are significantly more reactive, having higher HOMO energies than six-membered heterocyclic compounds. Benzene **4** stands out as unique transition from electron rich to electron deficient heterocycles.



Entry	Compound	E_{HOMO}/eV	E_{HOMO-1}/eV
1	pyrrole	-7.59	-8.48
2	furan	-8.22	-9.65
3	thiophene	-8.46	-8.82
4	benzene	-8.78	-8.78
5	pyridine	-9.10	-9.19
6	pyrazine	-9.03	-9.64
7	pyrimidine	-9.17	-9.99
8	pyridazine	-8.59	-10.12

Figure 7. Correlation of reactivities with corresponding HOMO and HOMO-1 of heterocyclic compounds

In summary, two sets of heterocyclic compounds are used to illustrate the use of FMO analysis to correlate with their nucleophilic or electrophilic reactivities. For halogenated heterocyclic compounds, when they have LUMO lobes centered on C-X carbons, their LUMO shall be selected; when there is little/no LUMO lobes on C-X carbons, LUMO+1 shall be considered;

when their LUMO and LUMO+1 lobes are comparable, we'll need to compare calculated activation energies. Similar criteria apply for HOMO/HOMO-1 analyses.

Building on What We Just Learned

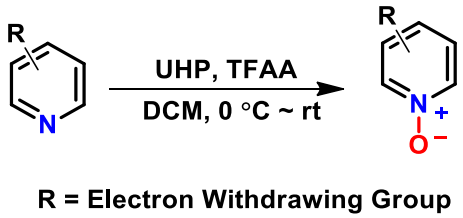


Figure 8. *N*-Oxidation of pyridines under UHP and TFAA conditions

Pyridines, with their HOMO lobes centered on the aromatic nitrogen, could be oxidized with peracid (such as peracetic acid, *m*-CPBA, etc.) to the corresponding *N*-oxides. For oxidation of electron-deficient pyridines, Caron *et al*.[3] reported that oxidation with urea-hydrogen peroxide complex (UHP) and trifluoroacetic anhydride (TFAA) is more efficient. What could be the reasons?

Return to Table of Contents

References:

[1] L.G. Zhuo, W. Liao, Z.X. Yu, *Asian J. Org. Chem.* **2012**, *1*, 336.
[2] J.A. Joule & K. Mills. *Heterocyclic Chemistry 5th Ed.* Chichester, West Sussex, UK: Blackwell Publishing Ltd., **2010**; p 256.
[3] S. Caron, N.M. Do, J.E. Sieser, *Tetrahedron Lett.* **2009**, *41*, 2299.

Chapter 28 A QM Study of the Hydroaminomethylation of Olefins

Yexia Zhang, Liting Dong, Tommy Lai, Guqin Shi, Yongsheng Chen, John S. Wai

For formation of C(sp^3)-N bonds, direct addition of nitrogen-containing fragments to olefins is complementary to classical S_N2 reactions. Metal-catalyzed direct hydroamination of alkenes has been extensively studied[1]; in contrast, hydroaminomethylation, the addition of a hydrogen and a (methyleneamino) unit across an olefin, has received significantly less attention.

Recently, Maulide's group reported the use of Eschenmoser salt for an acid-catalyzed hydro-aminomethylation of unactivated alkenes and alkynes[2]. A broadly applicable approach with readily available substrates and reagents, excellent functional group compatibility and high selectivity. Compared to the classical S_N2 approach, this reaction eliminates potential risks of formation of dimethyl or quaternary ammonium by-products. Reaction is assumed to proceed *via* an electrophilic addition followed by a [1,5]-H shift, a two steps mechanism:

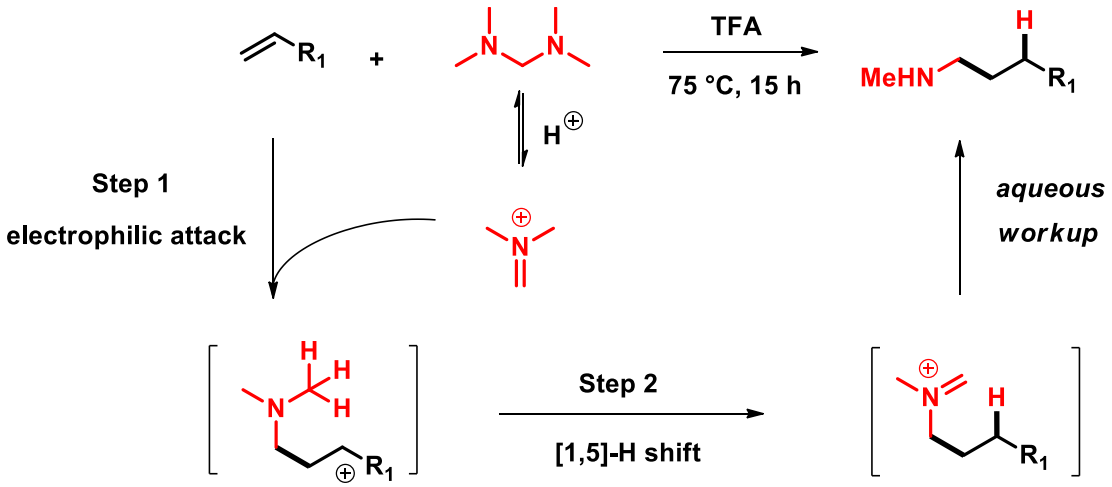
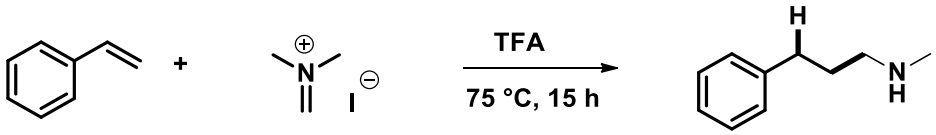


Figure 1: Reaction mechanism suggested in the article

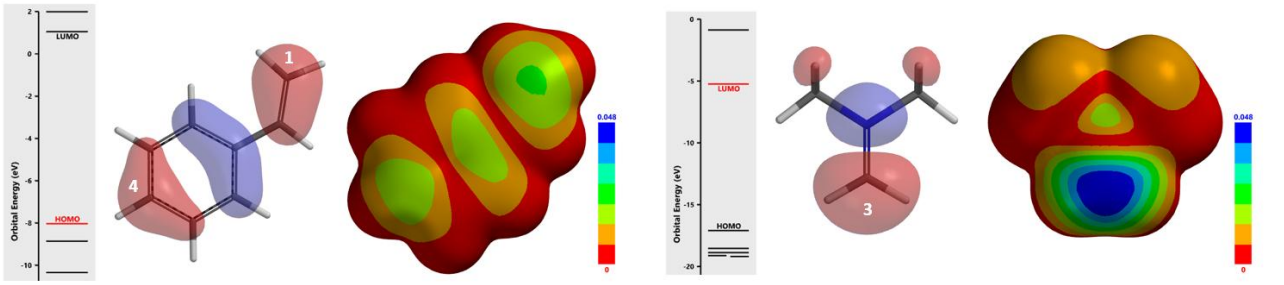
Exploration of Reaction Mechanism

We selected the reaction of styrene and Eschenmoser salt, reported in the publication, for our computational studies.



Based on the reaction mechanism proposed in the paper, the first step is electrophilic addition of Eschenmoser salt to styrene. Site of the reaction is at the terminal carbon of the

olefin. As such, we calculated HOMO and HOMO Map for styrene and LUMO and LUMO Map for iminium ion, as shown in figure 2[3].



HOMO -8.03 eV | HOMO Map (IsoVal 0.0020 e/au^3; 99.05%) | LUMO -5.24 eV | LUMO Map (IsoVal 0.0020 e/au^3; 98.9%)

Figure 2: HOMO and HOMO map for styrene and LUMO and LUMO map for iminium ion

HOMO and LUMO energy difference between styrene (-8.03 eV) and the iminium ion (-5.24 eV) is only 2.79 eV, suggesting that the reaction will proceed readily[4]. HOMO of styrene concentrates on substitutable carbon at C-1 and C-4 positions. Its HOMO map indicates that the lobe on C-1 is more exposed, suggesting that C-1 will interact selectively with electrophilic agents. Moreover, LUMO of the iminium ion is mainly distributed on the methylene carbon atom, the reaction center. Results of the above frontier orbital analyses are consistent with sites of reaction reported in the article.

Step 1: Reaction Energy Profile of the Electrophilic Addition Reaction

Next, we calculated for the reaction energy profile for the electrophilic addition step, with C-1 of styrene set 3.5 Å away from C-3 of iminium ion and moving closer to one another with a step size of 0.1 Å, as shown in figure 3 below. Activation energy is estimated to be about 13 kcal/mol, consistent with experimental conditions. Transition state and IR calculations of the peak energy structure **B** showed one and only one imaginary frequency at i299 cm^{-1}, supportive of the mechanism postulated[5].

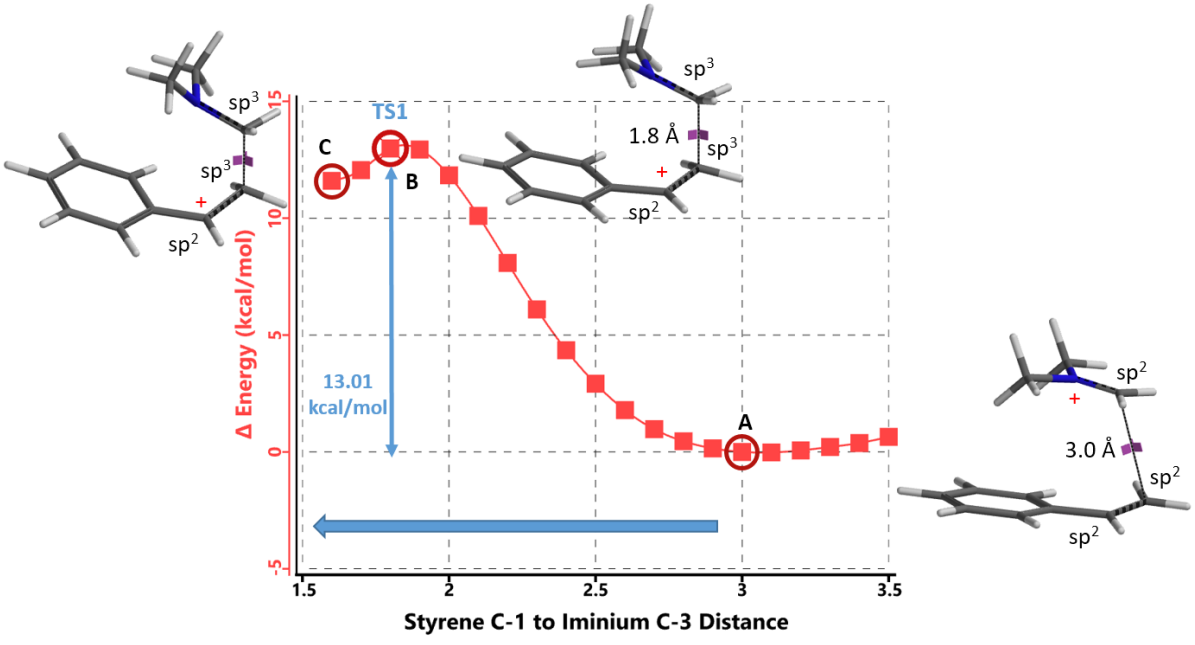


Figure 3: Step 1 Reaction Energy Profile

Step 2: Intrinsic Reaction Coordinate of the [1, 5]-H Shift

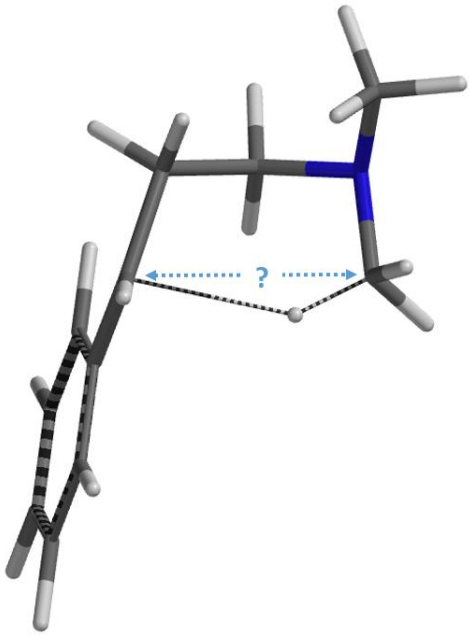


Figure 4: Step 2, conformer **D**

Step 2, for this intramolecular [1,5]-H shift, we did not know how far to put the C-1 and C-5 carbons apart to set up for a reaction energy profile calculation (Figure 4). We used an alternative computational strategy: calculate for the transition state of the H-shift first, then use the transition state structure for intrinsic reaction coordinate (IRC) calculation to establish the energy profile. IRC is an important concept and tool for quantum chemical calculation by starting from the transition state, bifurcating, and connecting reactants and products, to calculate for the reaction potential energy curve (Figure 5)[6].

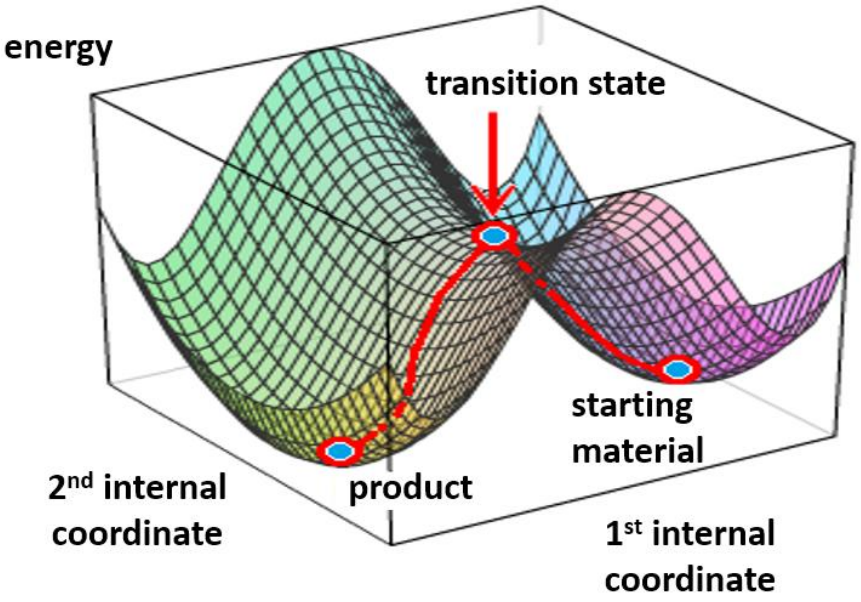


Figure 5: Schematic representation of Intrinsic Reaction Coordinate (IRC)

For this example we varied the C-1 to C-5 distance for a few transition state calculations, cumulating to the identification of the transition state structure **E** (Figure 6), with one and

only one imaginary frequency at i387 cm^{-1}, arriving at a reasonable transition state structure[5].

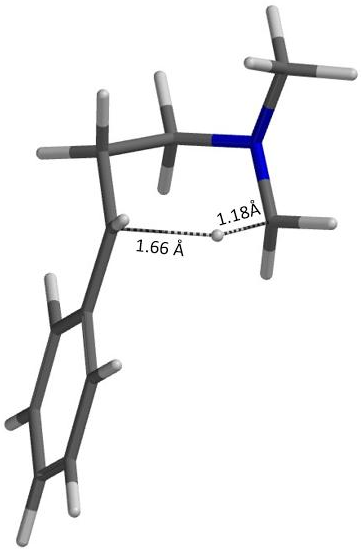


Figure 6: Step 2 transition state **E** (TS2) ifrequency i387 cm^{-1}

Starting from this transition state structure **E**, we calculated for IRC (Figure 7)[6], reached back to starting and arrived at product structures of the [1,5]-H shift step, and obtained the energy profile (Figure 7). The reaction starts from structure **D** and arrives at product structure **F** *via* transition state **E**. The activation energy required for the reaction is ~0.83 kcal/mol.

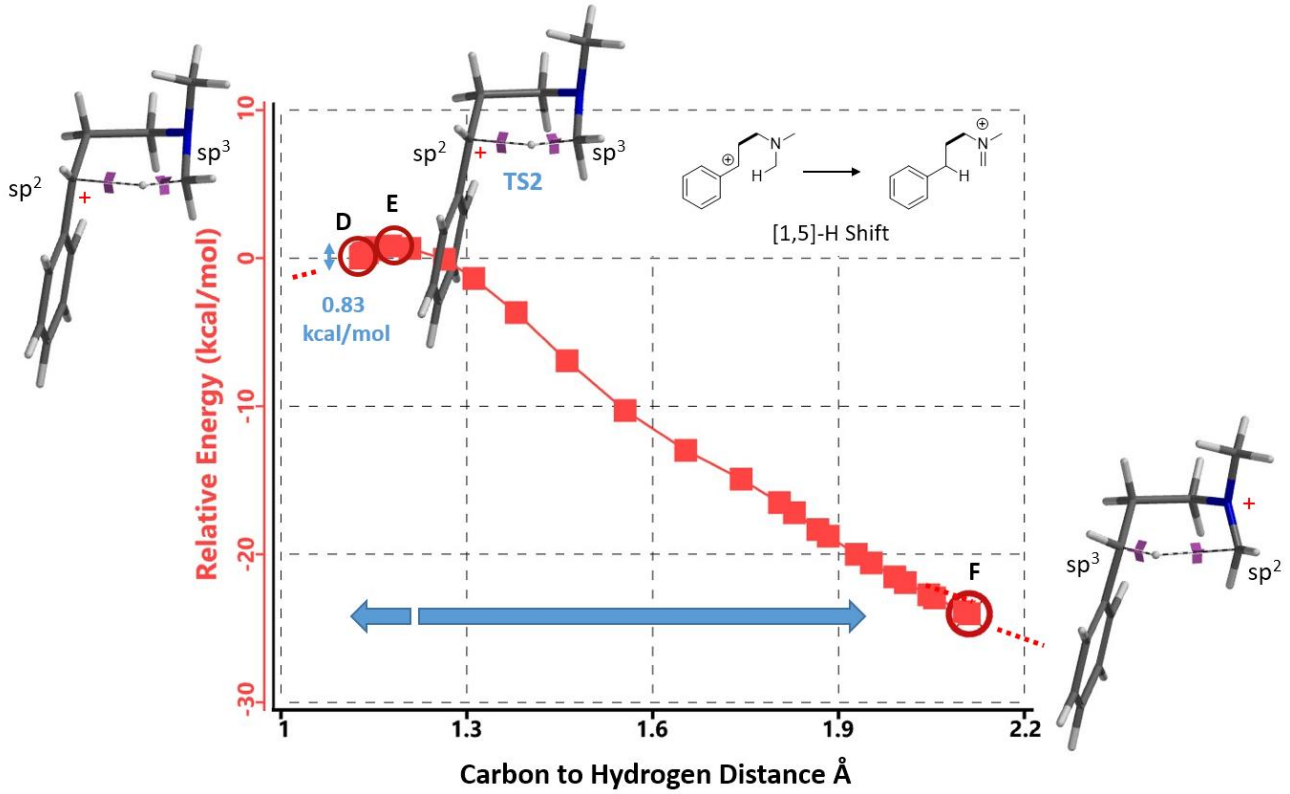
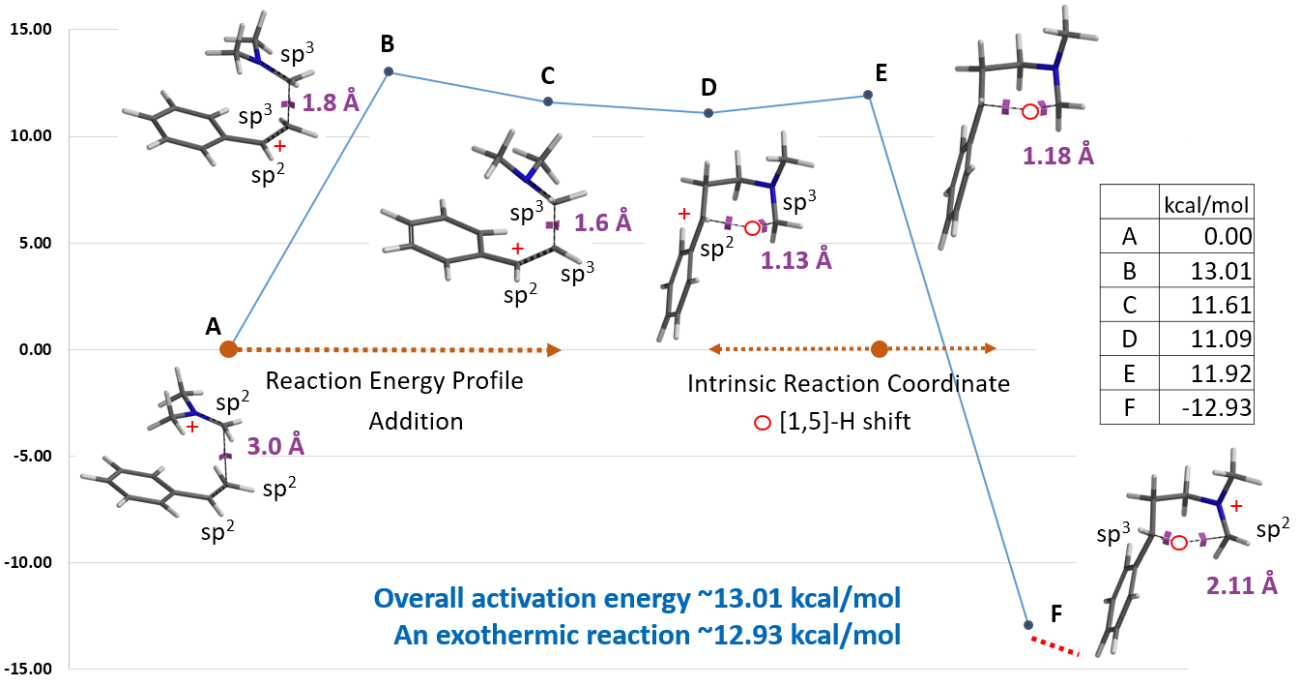


Figure 7: [1,5]-H shift IRC. See footnote [7] for changes in electron density during H shift

Connecting Step 1 and Step 2



	kcal/mol
A	0.00
B	13.01
C	11.61
D	11.09
E	11.92
F	-12.93

Figure 8. Overall reaction energy profile of styrene hydroaminomethylation

Next is to connect the energy profiles of step 1 and 2 together. **C** and **D** are the same intermediate structure, differing in their conformation. QM relative energy calculation of them indicates that conformer **C** is 0.52 kcal/mol higher in energy than conformer **D**.

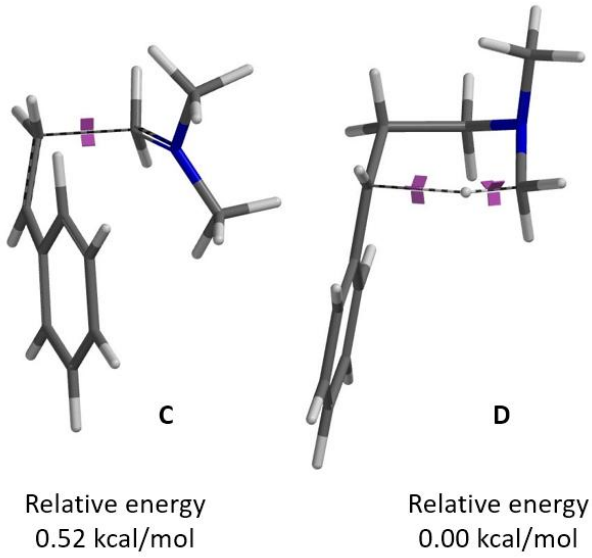


Figure 9. Relative energy of conformers **C** and **D**

Overall, hydroaminomethylation of styrene is exothermic (-12.93 kcal/mol) with step 1 as the rate limiting step having an activation energy of 13.01 kcal/mol. And the product **F** is an iminium ion, ready for further functionalization.

In summary, our FMO analyses of the styrene and iminium ion for step 1 are consistent with sites of reaction. Activation energy and transition state calculated for the electrophilic

addition correlates with experimental conditions. Results from direct transition state and IRC calculations of step 2, [1,5]-H shift, support the mechanism proposed in the literature.

Building on What We Just Learned

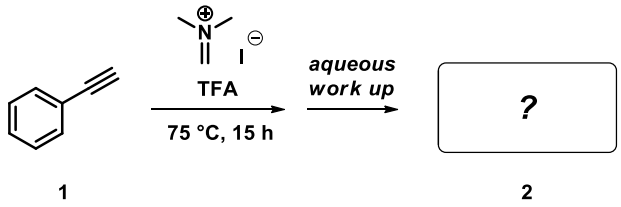


Figure 10: Reaction of phenylacetylene with iminium ion

In addition to alkenes, alkynes can undergo hydroaminomethylation under the same conditions. HOMO (-8.30 eV) and HOMO map of phenylacetylene are shown in Figure 11 below. What is the likely product of the reaction? Will the resultant product have a *cis* or a *trans* geometry?

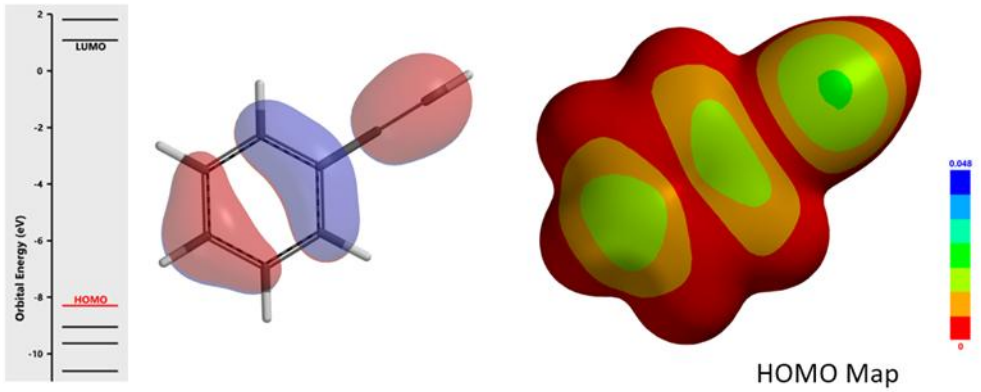


Figure 11: HOMO (-8.30 eV) and HOMO Map for Phenylacetylene (IsoVal 0.0020 e/au^3; 99.05%)

Return to Table of Contents

References and Footnotes:

[1] C. Michon, M.-A. Abadie, F. Medina, F. Agbossou-Niedercorn, *J. Organomet. Chem.*, **2017**, *847*, 13. J.P. Huo, G.Z. He, W.L. Chen, X.H. Hu, Q.J. Deng, D.C. Chen, *BMC Chemistry*, **2019**, *13*, article 89.
[2] D. Kaiser, V. Tona, C.R. Goncalves, S. Shaaban, A. Oppedisano, N. Maulide. *Angew. Chem. Int. Ed*. **2019**, 58, 14639.
[3] *Spartan'20 Tutorial and User's Guide* (**2020**). Irvine, CA, USA: Wavefunction, Inc. pp. 60 & 365.
LUMO map is an overlay of the absolute value of the lowest unoccupied molecular orbital onto an electron density surface. By convention, colors near blue indicate high concentration of the LUMO, while colors near red indicate low concentration. Given that the LUMO designates space available for a pair of electrons, a LUMO map indicates where nucleophilic attack would likely occur. Similar for HOMO map.

[4] a) L.G. Zhuo, W. Liao, Z.X. Yu, *Asian J. Org. Chem*. **2012**, *1*, 336 b) See Chapter 25, *Assessing Reactivity with LUMO and HOMO Energy Gap*.
[5] a) *Spartan'20 Tutorial and User's Guide* (**2020**). Irvine, CA, USA: Wavefunction, Inc. pp. 158, 442, 459 & 536 b) See Chapters 22 and 24, *Exploring the Reaction Mechanism of Menke Nitration* and *A QM Study of the Reaction Mechanism of Diazotransfer*.
[6] K. Fukui, *Acc. Chem. Res.* **1981**, *14*, 368. S. Maeda, Y. Harabuchi, Y. Ono, T. Taketsugu, K. Morokuma, *Int. J. Quantum Chem.* **2015**, *115*, 258.
Intrinsic reaction coordinate was proposed by Fukui in 1970 as a path of chemical reactions. It is the mass-weighted steepest descent paths on the potential energy surface, starting from the transition structure, that is, first-order saddle point. IRC has been used extensively in analysis and prediction of mechanisms of a variety of chemical reactions.
[7] Shown on Figure 12 is electron density of transition state **E** in [1,5]-H shift. Electron density is associated with the hydrogen as it migrates from the amino methyl carbon to the benzylic carbon. (Other than the migrating hydrogen, all the other hydrogens on the molecule are hidden).

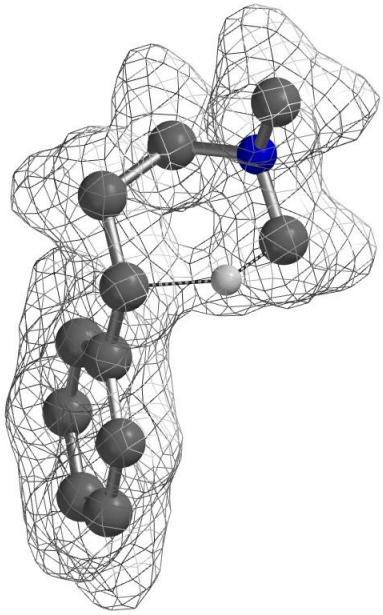


Figure 12. Electron density map of the TS **E** in [1,5]-H shift from **D** to **F**
(IsoVal 0.0457 e/au^3; 86%)

Chapter 29 Understanding the highly variable Regioselectivity in S_NAr reaction of Dichloropyrimidines

Renwei Zhang, Liting Dong, Tommy Lai, Yongsheng Chen, John S. Wai

Pyrimidines are extensively used in drug discovery due to their unique chemical properties. For medicinal chemists, functionalization of the pyrimidine ring with nucleophilic aromatic substitution reaction (S_NAr) of properly substituted 2,4-dichloropyrimidines is a very useful synthetic strategy. We learned that nucleophilic substitution of 2,4-dichloropyrimidines is usually C-4 selective. However, often a time, we also encountered exceptions[1][2][3], that is, C-2 selective displacement or a mixture of products. In this chapter, we continue our QM analyses on why regioselectivity of S_NAr reactions observed with 2,4-dichloropyrimidines is so sensitive to other substitutions on the ring.

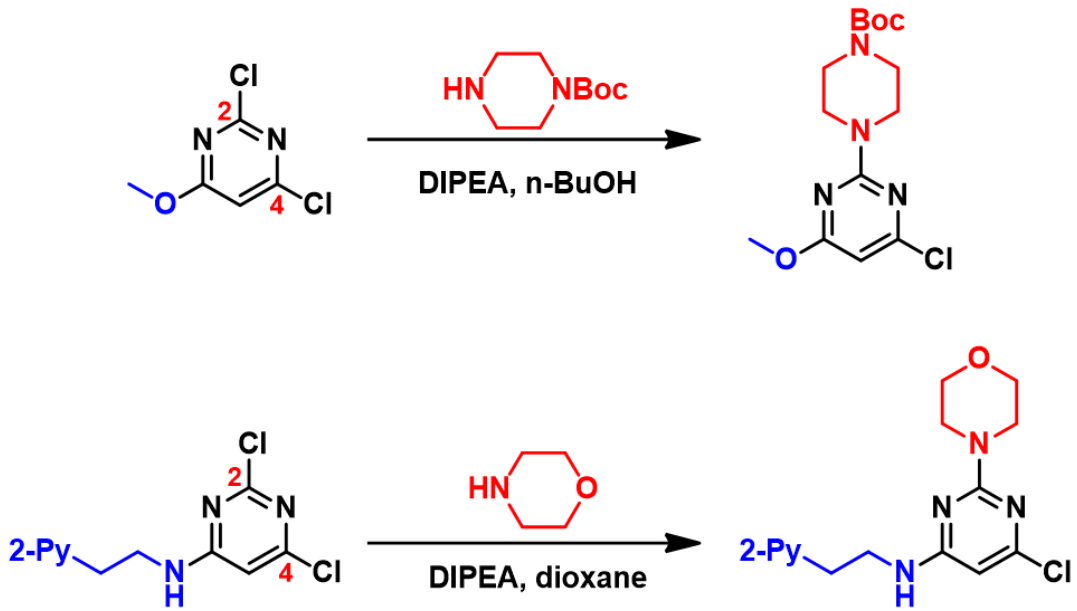


Figure 1: Regioselective C-2 S_NAr reaction of 2,4-dichloropyrimidine with C-6 electron-donating substitution

We selected the above two reactions for QM analysis (Fig. 1). For 2,4-dichloropyrimidine with an electron-donating substituent at C-6, its S_NAr reactions proceed preferably at C-2, instead of C-4. Authors of the second reaction verified the structure of the product by hydrogenating off the chlorine group. Two distinct doublets were observed in the aromatic region of the ^1H NMR spectrum, thus enabling accurate structural assignment of the product[2]. Could we use QM analysis to account for such reversal in selectivity?

LUMO and LUMO Map Analysis

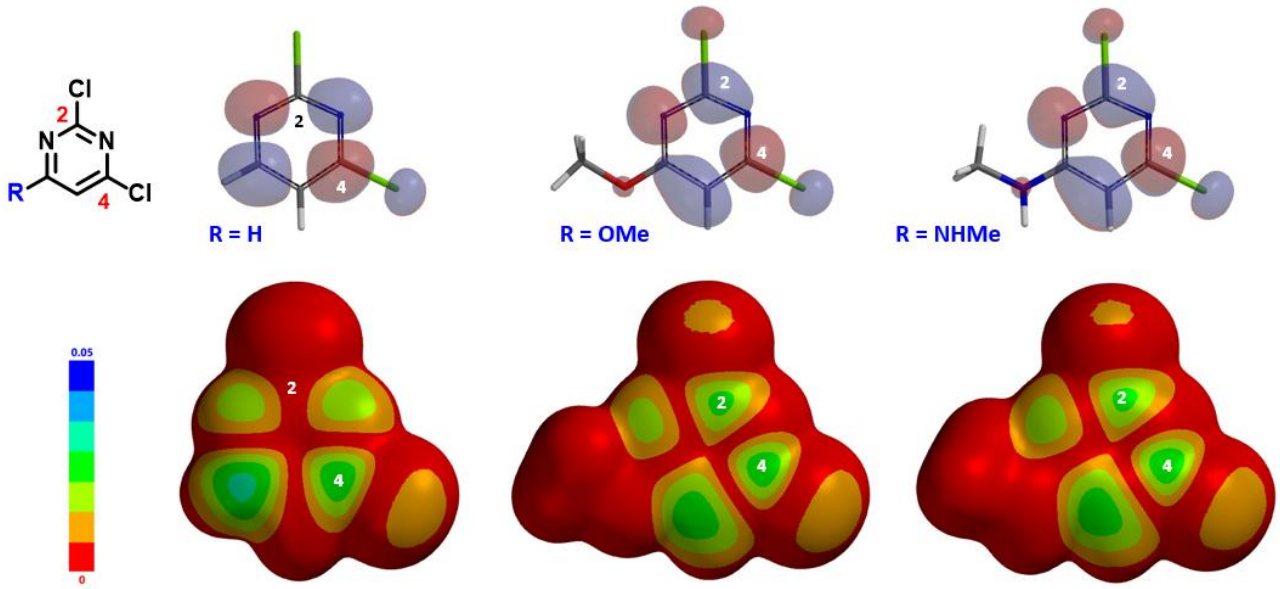


Figure 2: LUMO and LUMO map of 2,4-dichloropyrimidine analogs (R = H, OMe, NHMe)
(IsoVal 0.002 e/au^3; 99.3%)

Shown in Figure 2 are LUMOs and LUMO maps of three 2,4-dichloropyrimidines (C-6 R = H, OMe, NHMe) obtained from DFT calculations (Spartan'20, ωB97X-D/6-31G*). When R is H, LUMO is mainly distributed at C-4, and almost none at C-2, consistent with what we learned from general heterocyclic chemistry textbooks, accounting for C-4 selectivity observed. On the other hand, when R is OMe or NHMe, LUMOs of the pyrimidines are quite different. LUMO lobes at C-2 and C-4 are similar in size, and on LOMO maps, they have similar concentration, suggesting that we could potentially obtain a mixture of C-2 and C-4 substitution products. However, highly selective C-2 substitution is observed. LUMO and LOMO map analyses could account for increased C-2 reactivity but not the selectivity.

Reaction Energy Profile Calculation

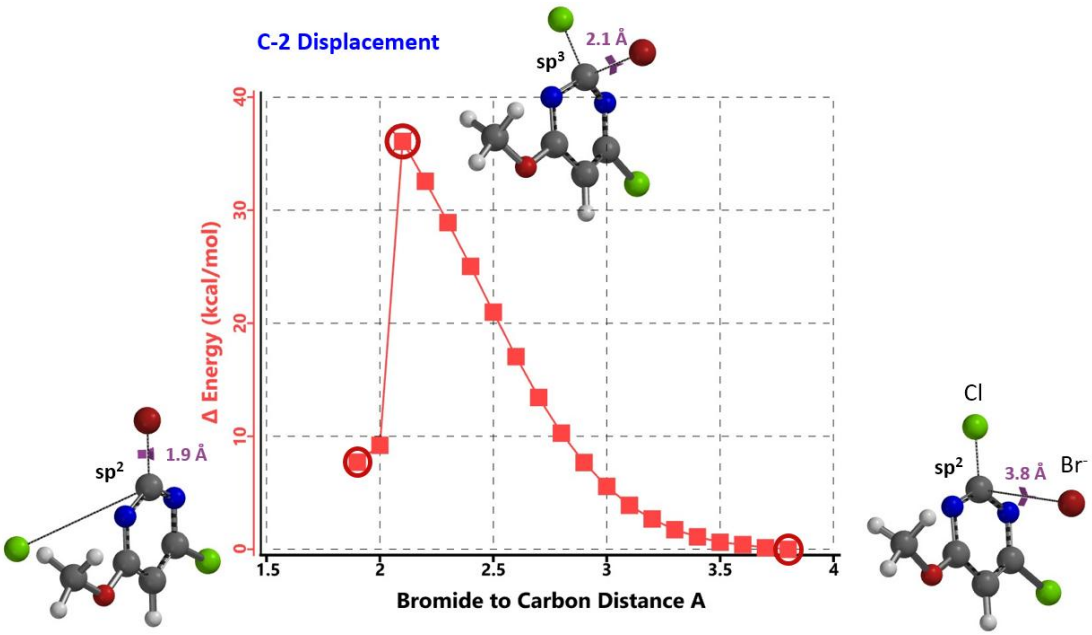


Figure 3: Reaction energy profile for C-2 substitution (with bromide as surrogate nucleophile)

In this case, we calculated for the reaction energy profiles for the C-2 and C-4 substitutions (Figure 3; similar for C-4, not shown), and used bromide as surrogate nucleophile to speed up the calculations[4].

Transition State Calculations

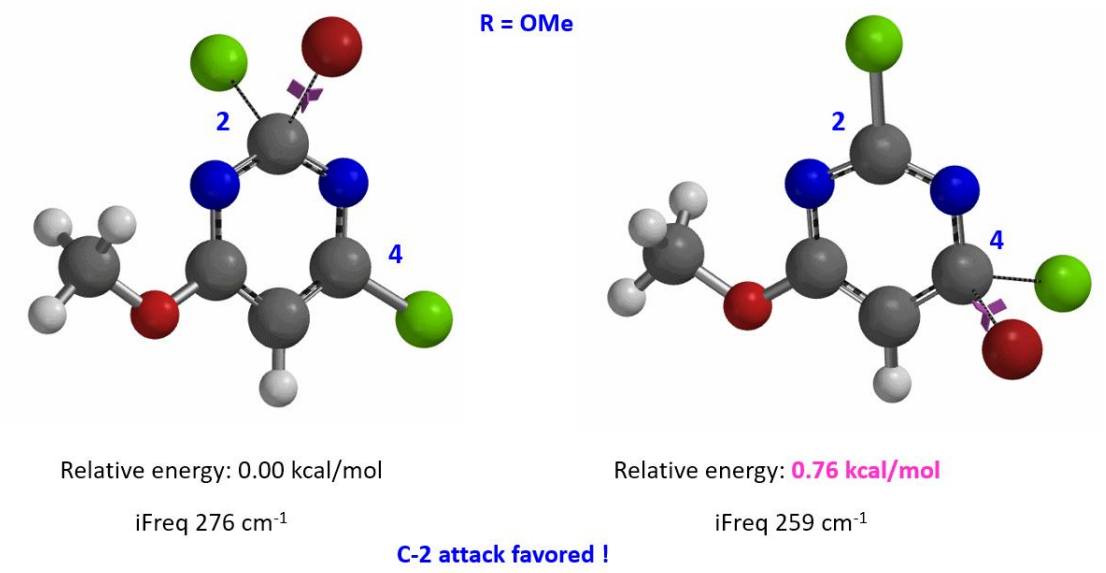


Figure 4: Relative difference in transition state energy for C-2 and C-4 substitution with R = OMe

Next we took the structures at the highest points from the two reaction energy profiles to do more accurate "Transition State Geometry" calculations and determine their relative energy difference. Energy of C-4 TS is found to be 0.76 kcal/mol higher than that of C-2 (Figure 4), suggesting that the reaction will be C-2 selective, consistent with experimental observations.

Infrared vibration calculation showed one and only one imaginary frequency at 276 and 259 cm^{-1} for C-2 and C-4 transition states, respectively[5].

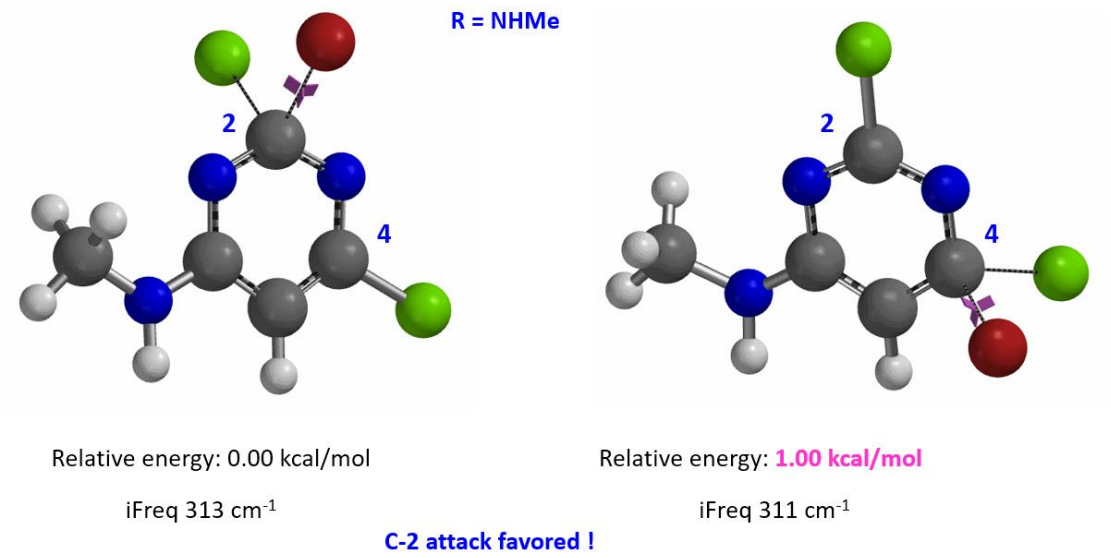


Figure 5: Relative difference in transition state energy for C-2 and C-4 substitution with R = NHMe

We used the same calculation workflow for R = NHMe to determine their relative difference in transition state. Energy of C-4 TS is 1.00 kcal/mol higher than that of C-2, suggesting that the reaction at C-2 will also be more favorable (Figure 5).

When we do LUMO analyses for regioselectivity of these competing reactions, we evoked the Hammond hypothesis, that is, in exothermic reaction, the structure of the transition state is similar to the substrates. This has its limitations. When orbital lobes are similar in size, there could be other factors affecting the relative energy of the transition states and altering the product ratio. Under such circumstance, we need to calculate for reaction energy profiles and transition states for proper comparison.

LUMO+1

In Chapter 7, Application of LUMO Analysis in Nucleophilic Reactions (Part II), we mentioned that if the energy difference between LUMO and LUMO+1 is small, we need to consider both orbitals, along with potential steric interaction from neighboring substituents.

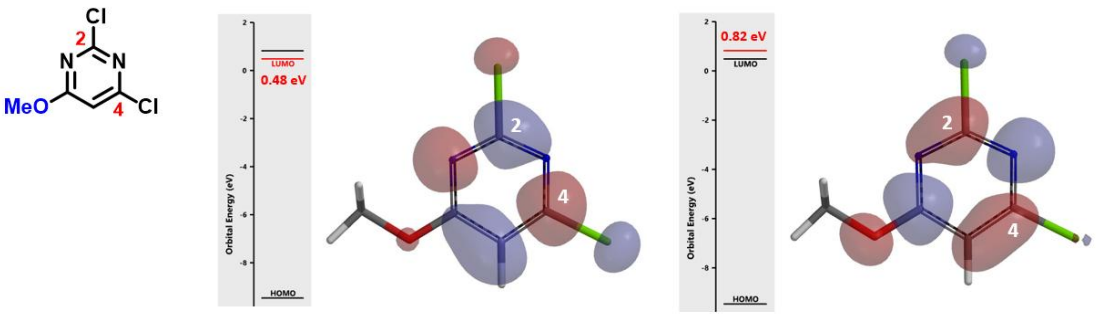


Figure 6: LUMO and LUMO + 1 for 2,4-Dichloro-6-methoxypyrimidine

In the case of 2,4-dichloro-6-methoxypyrimidine (Figure 6), the energy difference between LUMO and LUMO+1 is 0.34 eV. With a big enough energy gap, LUMO+1 is not likely to participate in the reaction.

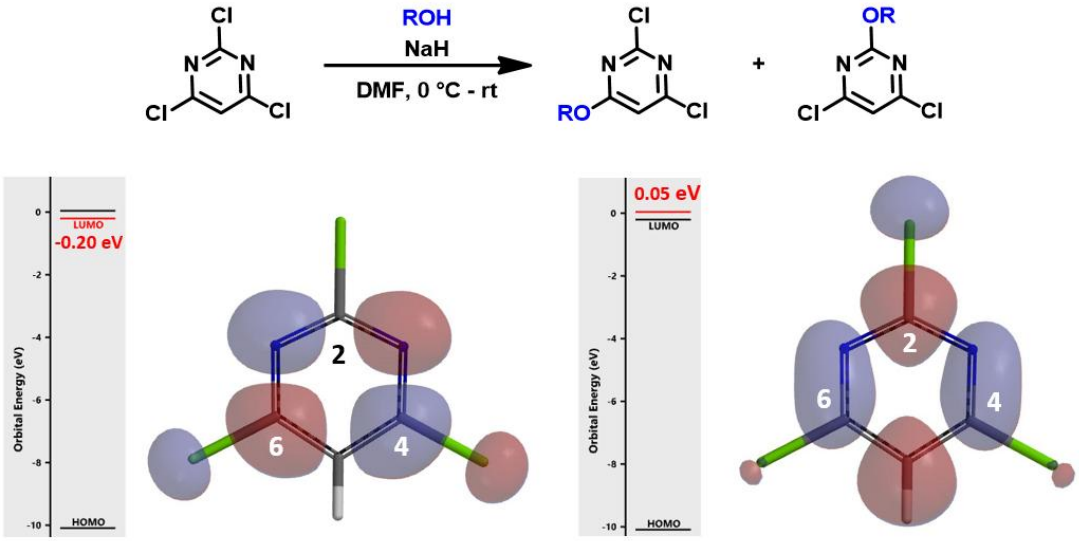


Figure 7: LUMO and LUMO+1 of 2,4,6-trichloropyrimidine

Replacing C-6 methoxy with one more chloro group, we have trichloropyrimidine, which has LUMO lobes on both C-4 and C-6, but no LUMO distribution at C-2 position. Moreover, the energy gap between LUMO and LUMO+1 is 0.25 eV, requiring us to consider LUMO+1 as well. Since there is significant LUMO+1 lobe at C-2, we need to evaluate relative difference in transition states energy. Calculation suggested a mixture of products, consistent with the experimental results.

To summarize what we have learned from chapter 7[3] and this chapter, 1) For S_NAr reaction with 2,4-dichloropyrimidine analogs, when there are strong electron donating or electron withdrawing groups at positions C-5 or C-6, the reaction may not always occur selectively at the C-4 position. We could potentially obtain a mixture of C-4 and C-2 substitution products, or even C-2 displacement product only; 2) If energy gap between LUMO and LUMO+1 is $\leqslant$ 0.25 eV, it is necessary to consider contribution from LUMO+1; 3) When C-5 position is a sterically bulky substituent, it could affect C4/C2 selectivity[3]; 4) Regioselectivity of 2,4-dichloropyrimidine analogs' S_NAr reaction is very sensitive to electronic and steric effects, yet quite predictable with QM calculation tools.

Building on What We Just Learned

Shown on Figure 8 is 2,4-dichloro-5-trimethylsilylpyrimidine and its LUMO and LUMO+1. Energy difference between LUMO and LUMO+1 is 0.14 eV. Only C-2 nucleophilic substitution product could be observed with its S_NAr reactions. What could be the reason?

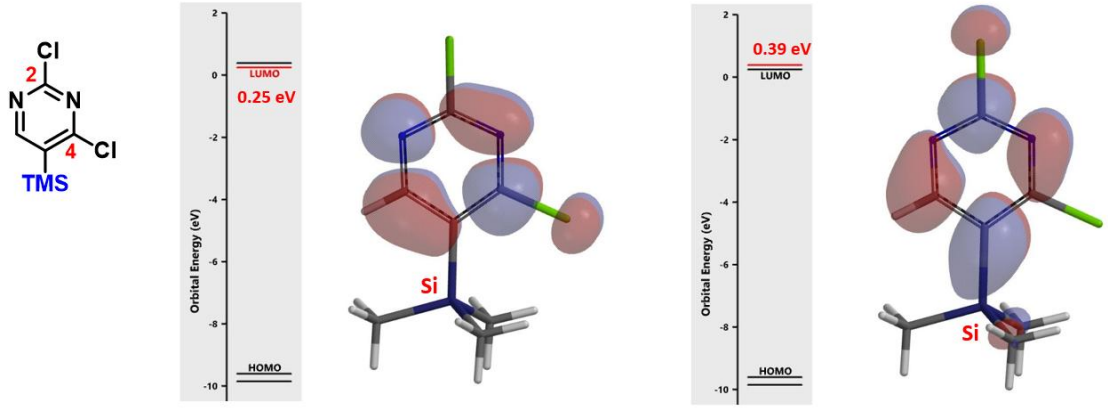


Figure 8: LUMO and LUMO + 1 for 2,4-dichloro-5-trimethylsilylpyrimidine

Return to Table of Contents

References:

[1] L.D. Arnold, H. Maag, W.W. Turner, Jr. WO 2016/168619 A1.

[2] J.M. Large, J.E. Torr, F.I. Raynaud, P.A. Clarke, A. Hayes, F. Stefano, F. Urban, S.J. Shuttleworth, N. Saghir, P. Sheldrake, P. Workman, E. McDonald, *Bioorg. Med. Chem.* **2011**, *19*, 836 – 851.

[3] See Chapter 7, Application of LUMO Analysis in Nucleophilic Reactions (Part II).

[4] See Chapter 10, S_NAr Reaction of Polyhalogenated Heterocycles.

[5] a) *Spartan'20 Tutorial and User's Guide* (**2020**). Irvine, CA, USA: Wavefunction, Inc. pp. 158, 442, 459 & 536. b) See Chapters 22 and 24 for transition state and imaginary frequency calculations.

Chapter 30 QM analyses of Electrophilic Aromatic Iodination

Jiahui Lu, Liting Dong, Zhaowu Xu, Qiuyue Wang, Yongsheng Chen, John S. Wai

Aromatic iodo compounds, with higher reactivity as compared to their bromo and chloro analogs, play an important role in synthetic organic chemistry. In addition to functional group interconversion, iodo intermediates could also be used for preparation of radio labeled analogs. For example, in the preparation of tritium (^{3}H)-labeled HIV-1 integrase inhibitor MK-0248, it is first iodinated, then reduced with sodium borotritide (NaB^{3}H4) for extensive biological evaluations and DMPK profiling (Figure 1).[1]

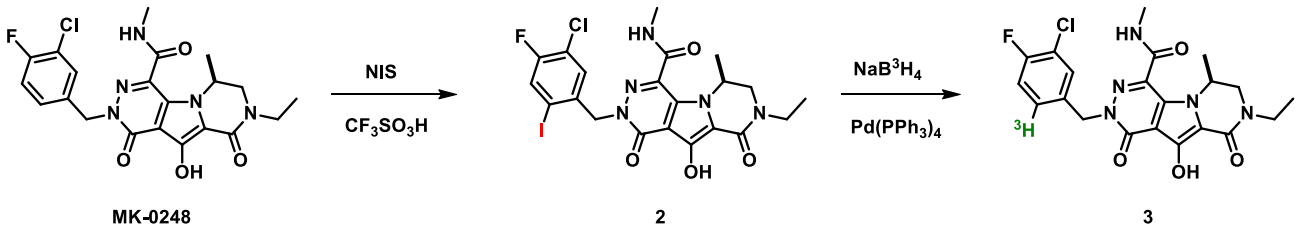


Figure 1. Tritium labeling of HIV-1 Integrase Inhibitor MK-0248

In chapter 25 we discussed the use of HOMO (nucleophile) and LUMO (electrophile) energy gap to assess how readily they will react with one another, as exemplified with halogenation reagents and a set of substrates for Pictet-Spengler reaction. We also described that phenol readily reacts with N-iodosuccinimide (NIS) to provide a mixture of mono-, di- and tri-iodinated products. However, iodination of electron deficient 4-chloronitrobenzene with NIS proceeds only in the presence of trifluoromethanesulfonic acid. It was reasoned that trifluoromethanesulfonic acid reacts with NIS to generate iodine(I) trifluoromethane-sulfonate,[2] which is the active species for the iodination (Figure 1). At the end of the chapter, we were asked to decide whether we shall use NIS or this super-electrophilic reagent to iodinate 1,2,3-trifluoro-4-nitrobenzene substrate, and predict whether reaction will occur at C-5 or C-6 position (Figure 2).

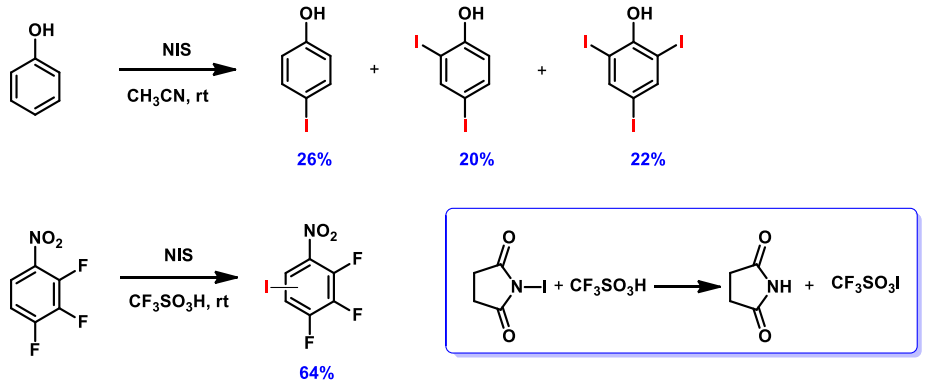


Figure 2. Iodination of phenol and 1,2,3-trifluoro-4-nitrobenzene

Regioselectivity and reagent choice are two essential considerations. In this chapter, we will describe how QM analyses could guide us prospectively.

(I) Regioselectivity of iodination reaction

In chapters 2, 17, and 18, we discussed the use of calculated HOMO, HOMO map, and ^{13}C NMR data to predict regioselectivity of NCS chlorination and NBS bromination of various heterocyclic aromatic systems.

These electrophilic reactions usually occur preferentially at substitutable positions with the largest HOMO lobe of the substrates. As such, we first calculate for the frontier orbitals of 1,2,3-trifluoro 4-nitrobenzene and their energies. Since the energy difference between HOMO and HOMO-1 is relatively small, 0.2 eV, we learned from experience to consider both HOMO and HOMO-1 (Figure 3). For HOMO, the lobe is mainly distributed at substitutable C-5, while for HOMO-1, it is mostly on C-6. The differences are not significant enough to enable any confident prediction. Moreover, calculated ^{13}C NMR of the substrate (Figure 3) showed C-5 and C-6 with chemical shift of 123 and 112 ppm, respectively, a difference in 11 ppm, suggesting that electrophilic aromatic reaction will preferentially occur at C-6.

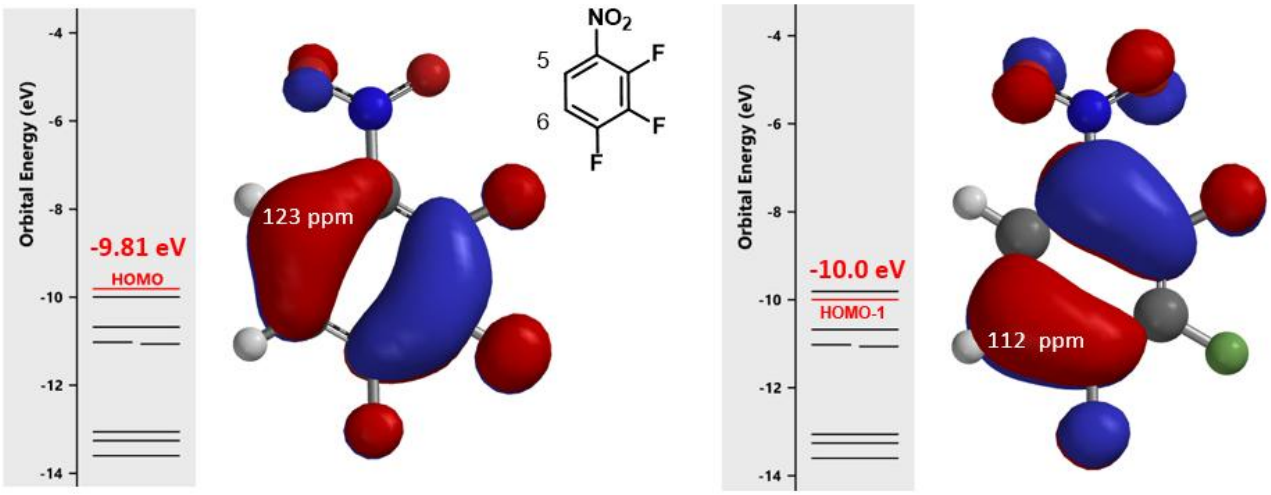


Figure 3. HOMO and HOMO-1 of 1,2,3-trifluoro-4-nitrobenzene; ^{13}C NMR chemical shift for C-5 & C-6

Further comparison of the HOMO and HOMO-1 maps (Figure 4) revealed that the HOMO-1 density at C-6 is significantly higher than that of HOMO at C-5, as indicated by more blue color at C-6 *vs* C-5, augmenting the prediction from calculated ^{13}C NMR; that is, C-6 position is more likely to undergo selective electrophilic substitution.

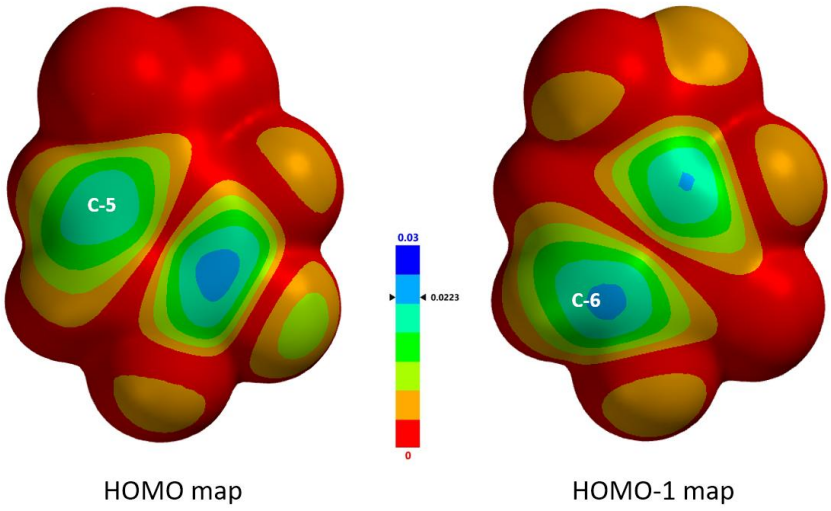
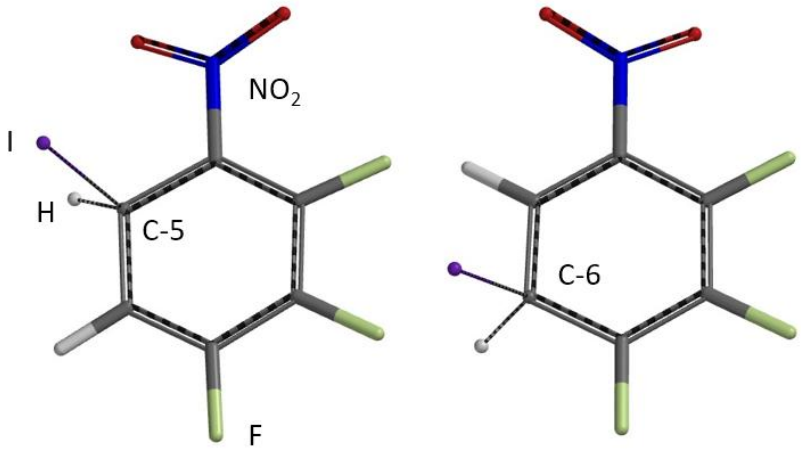


Figure 4. HOMO and HOMO-1 map of 1,2,3-trifluoro-4-nitrobenzene

Extending the "Equilibrium Geometry and Relative Energy" comparison method described in chapter 17 to assess regioselectivity in bromination to iodination, the relative energy calculated for equilibrium geometry of C-6 is found to be significantly lower than that of C-5 (Figure 5). We have enough data to make a prediction that the iodination reaction is likely to occur selectively at C-6 position, and this is consistent with experimental result.[2,3]



Relative Energy (kcal/mol)	C-5	C-6
Nonpolar solvent	5.88	0.00
Polar solvent	5.96	0.00

Figure 5. Equilibrium geometry and relative energy of C-5 vs C-6 iodinium ions

Summarizing key results from the sequence of calculations described above:

1. HOMO and HOMO-1 energies differ by only 0.2 eV, requiring both HOMO and HOMO-1 to be considered. HOMO suggests iodination will occur at C-5, whereas HOMO-1 suggests that it will occur at C-6.
2. Calculated ^{13}C NMR data predicts that the reaction will occur selectively at C-6.
3. HOMO map and HOMO-1 map comparison predicts that iodination will occur selectively at C-6.

4. Calculated equilibrium geometry and relative energy of C-5 *vs* C-6 iodinium ions predict that the reaction will also occur selectively at C-6.

Overall, calculations suggest that the iodination will occur preferentially at C-6, and is consistent with experimental result.[2,3]

(II) Selection of iodination reagent

After determining regioselectivity for iodination of the substrate, we have to select appropriate iodinating reagent (and reaction conditions).

In chapter 25, we discussed that the lower the LUMO energy of a chlorinating reagent, the stronger its reactivity in electrophilic aromatic reaction (Figure 6). For different substrates, their reactivity could be assessed by calculating the energy difference between LUMO of the chlorinating reagent and HOMO of the substrate. When the energy gap exceeds a specific threshold, the reaction will not proceed.

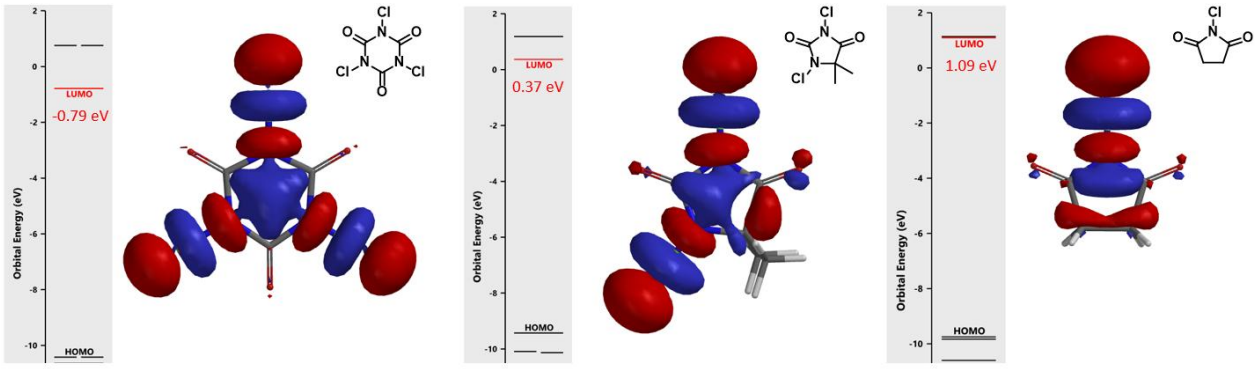


Figure 6. LUMO and energy of trichloroisocyanuric acid, dichlorohydantoin, and NCS

This method is also applicable for assessing the reactivity of the iodination reagents and their reactions.

In iodination of phenol with NIS, the energy difference between LUMO of NIS and HOMO of phenol is 7.52 eV, and the reaction proceeds readily (Table 1). However, under the same conditions, no reaction is observed between 1,2,3-trifluoro-4-nitrobenzene and NIS. The energy gap between HOMO-1 and LUMO of this reacting pair is 9.54 eV, presumably above the threshold. With CF_3SO_3I as iodinating reagent, the LUMO and substrate HOMO-1 energy gap is lowered to 7.68 eV, and the reaction proceeds smoothly.

Reagent	LUMO (eV)	Substrate	HOMO used (eV)	LUMO-HOMO gap (eV)
NIS	-0.46	PhOH	-7.98	7.52
NIS	-0.46	$C_6H_2F_3NO_2$	-10.00	**9.54**
CF_3SO_3I	-2.32	$C_6H_2F_3NO_2$	-10.00	7.68

Table 1. Reagents LUMO and substrates HOMO energy and their energy gaps

LUMO energies of a few common iodination reagents are calculated and shown on the scale below for comparison (Figure 7). For electron rich substrates, reagents with relatively higher LUMO energy will be sufficient and shall be used to avoid over iodination, whereas for electron deficient substrates, reagents with lower LUMO energy, such as CF_3SO_3I and CH_3SO_3I, will be necessary.[4]

Iodination reagent	LUMO energy (eV)
CF_3SO_3I	-2.32
CH_3SO_3I	-1.68
1,3-Diiodo-5,5-dimethylhydantoin (DDH)	-0.92
N-Iodosaccharin (NISac)	-0.82
N-Iodophthalimide (NIP)	-0.56
N-Iodosuccinimide (NIS)	-0.46

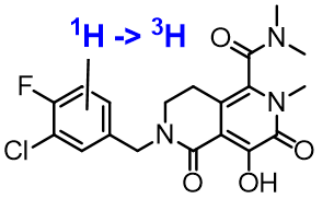
Figure 7. LUMO energies for different iodinated reagents

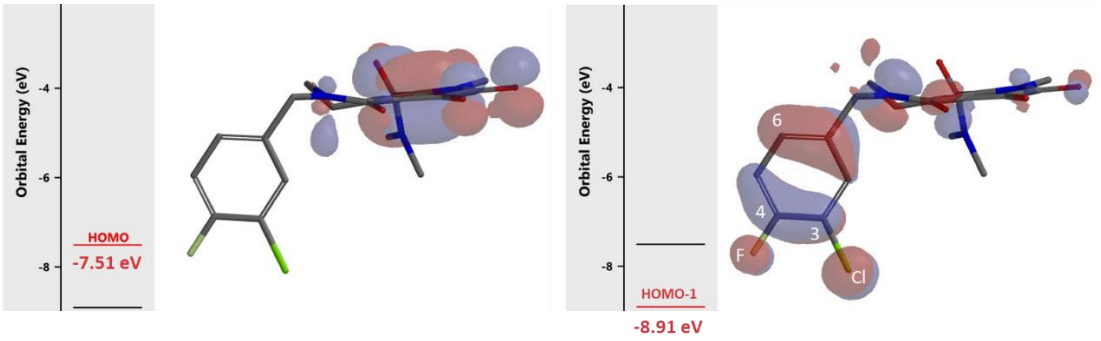
Summary

For electrophilic aromatic iodination, HOMO analysis, calculated ^{13}C NMR data, HOMO map, calculated equilibrium geometries and their relative energy are useful in predicting the site of reaction. Whereas energy difference between LUMO of the iodination reagents and HOMO of the substrates enables assessment of the feasibility of the reaction. For comparison, a few common iodination reagents are ranked by their LUMO energy to serve as a quick guide to their relative reactivity.

Building on What We Just Learned

We need to prepare iodo derivative of the inhibitor below (Figure 8) for tritium labeling. Iodination will occur only on the substitutable carbons with suitable HOMO/HOMO-1 Lobes on the phenyl group. Which carbon will be iodinated and which reagent shall we use?





Substrate	Energy (eV)	
HOMO		
HOMO-1	-8.91	
Iodination reagent	**LUMO (eV)**	**LUMO – HOMO-1 Energy Gap (eV)**
CF_3SO_3I	-2.32	6.59
CF_3CO_2I	-1.90	7.01
CH_3SO_3I	-1.68	7.23
1,3-Diiodo-5,5-dimethylhydantoin	-0.92	7.99
N-Iodosaccharin	-0.82	8.09
N-Iodophthalimide	-0.56	8.35
N-Iodosuccinimide	-0.46	8.45

Figure 8. HOMO/HOMO-1 of inhibitor, LUMO of selected iodination reagents, and their energy gaps

Return to Table of Contents

References:

[1] (a) R. Voges, J.R. Heys, T. Moenius, *Preparation of Compounds Labeled with Tritium and Carbon-14*, John Wiley & Sons Ltd, 2009. (b) T. Nagasaki, K. Sakai, M. Segawa, Y. Katsuyama, N. Haga, M. Koike, K. Kawada, S. Takechi, *J. Labelled Cpd. Radiopharm.* **2001**, *44*, 993.

[2] G.A. Olah, Q. Wang, G. Sandford, G.K. Surya Prakash, *J. Org. Chem.* **1993**, *58*, 3194.

[3] Spartan'20 QM NMR calculation algorithm is not applicable for iodine containing compounds.

[4] X.S. Xue, Y. Wang, M. Li, J.P. Cheng, *J. Org. Chem*. **2016**, *81*, 4280. Fluorinated reagent FPD values (Figure 9).

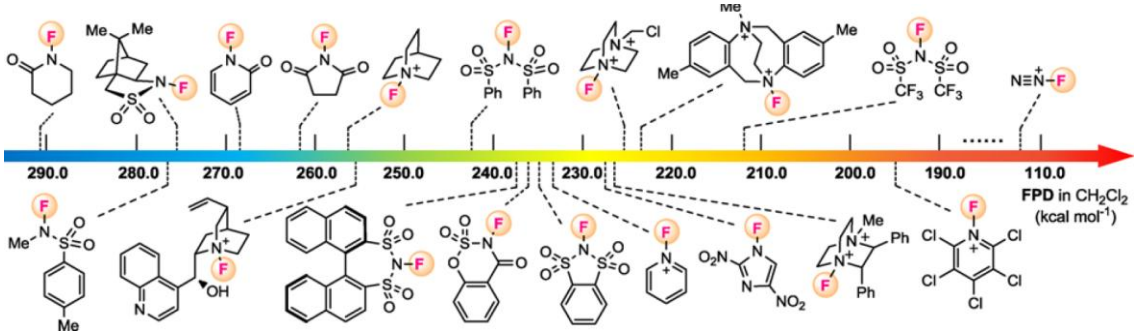


Figure 9. Fluoro FPD Values

Chapter 31 Benzofuran Synthesis *via* Acid Catalyzed Cyclization of Acetal

Qiuyue Wang, Zhong Zheng, Shouliang Wang, Yongsheng Chen, John S. Wai

Benzofuran is an important class of heterocyclic compounds and a common motif in pharmaceuticals. Its synthesis has a wide range of applications in drug discoveries. For example, Amiodarone, an antiarrhythmic drug developed by Pfizer, Darifenacin hydrobromide, a genitourinary drug marketed by Novartis, and Fruquintinib, an antineoplastic drug developed by Hutchison Whampoa (Figure 1).[1]

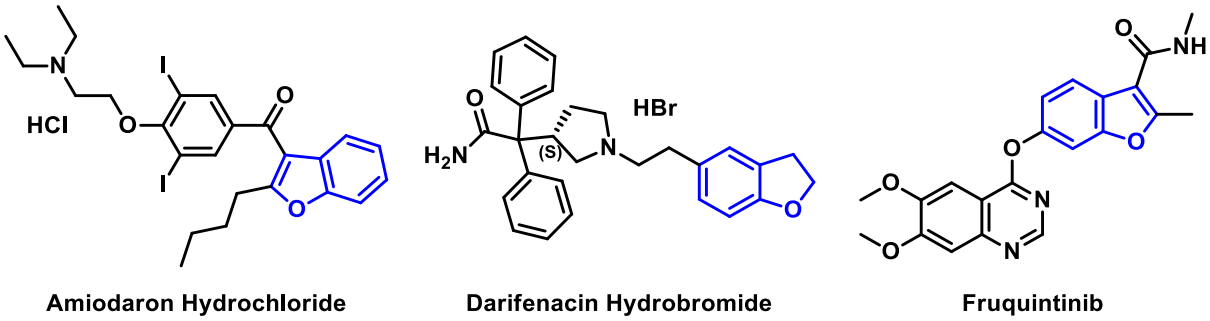


Figure 1. Marketed drugs with benzofuran motif

There are many ways to construct the benzofuran scaffold[2]; among them, cyclization of aryl ether substrates is the most popular synthetic method used in our laboratory (Figure 2). In this chapter, we'll discuss the use of molecular orbital analysis to predict regioselectivity of acid-catalyzed cyclization of the acetal substrates.

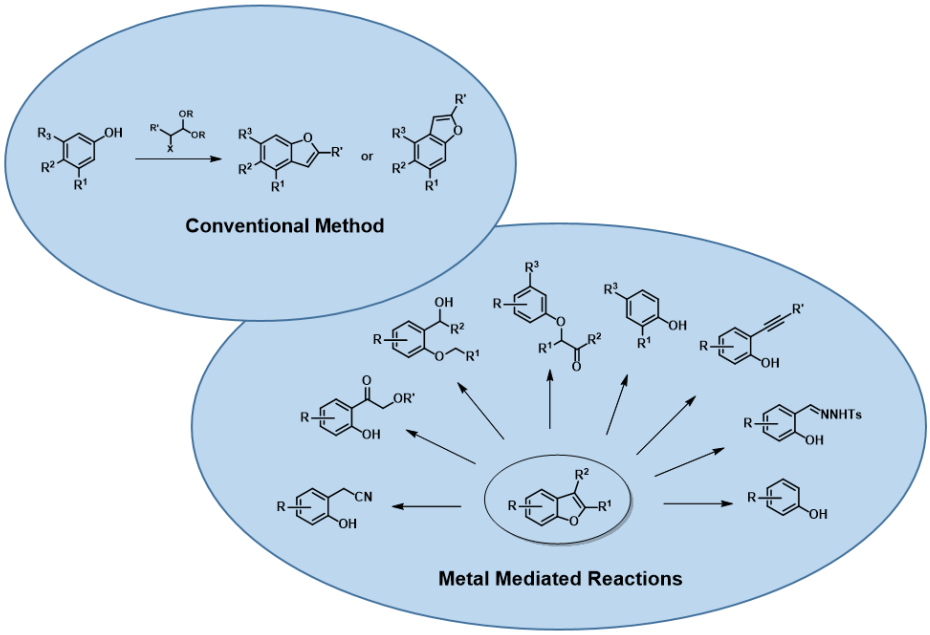


Figure 2. General synthetic methods for benzofuran

Regioselectivity of Cyclization

In our laboratory, we synthesized benzofuran core **2a** (Figure 3) *via* polyphosphoric acid (PPA) catalyzed cyclization of acetal **1**. QM analyses (HOMO and ^{13}C NMR calculation) of **1** suggested that the reaction will provide a mixture of regioisomers **2a** and **2b**, with **2a** being the major product. Experimentally, a 1:5 mixture is observed, with **2b** as the major product (Figure 4), requiring us to reassess QM analysis of this type of reaction.

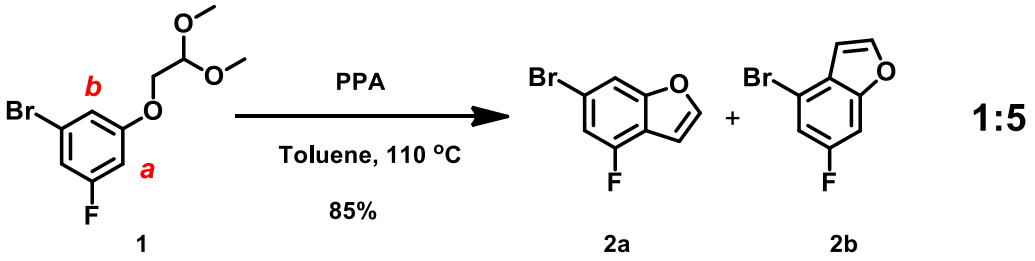


Figure 3. Experimental product ratio, **2a**: **2b** = 1:5

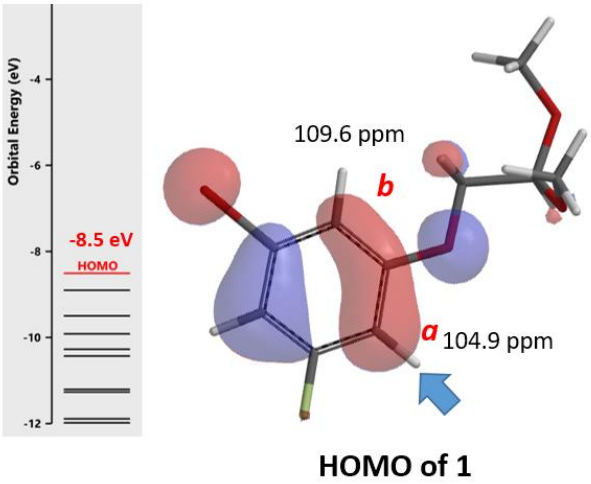


Figure 4. Schematic of HOMO and ^{13}C NMR data for substrate **1**

Mechanistic Insight

Mechanistically, substrate **1** is first protonated under acidic conditions, followed by elimination of MeOH to form oxonium ion **A**. The phenyl ring underwent nucleophilic addition, at sites *a* and *b*, to provide **B** and **C**, followed by a second MeOH elimination to provide a mixture of products, **2a** and **2b** (Figure 5).

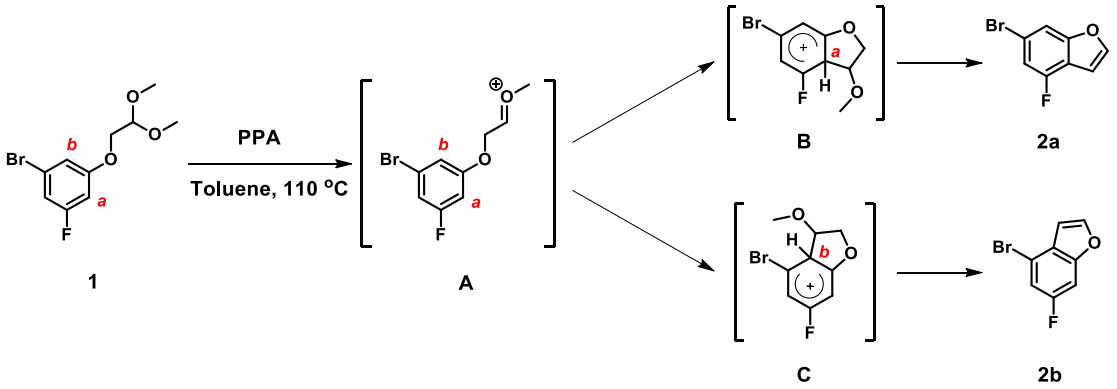


Figure 5. Reaction mechanism for formation of benzofuran under acidic conditions

As such, it is more relevant to look at properties of oxonium ion **A** to assess regioselectivity of the cyclization.

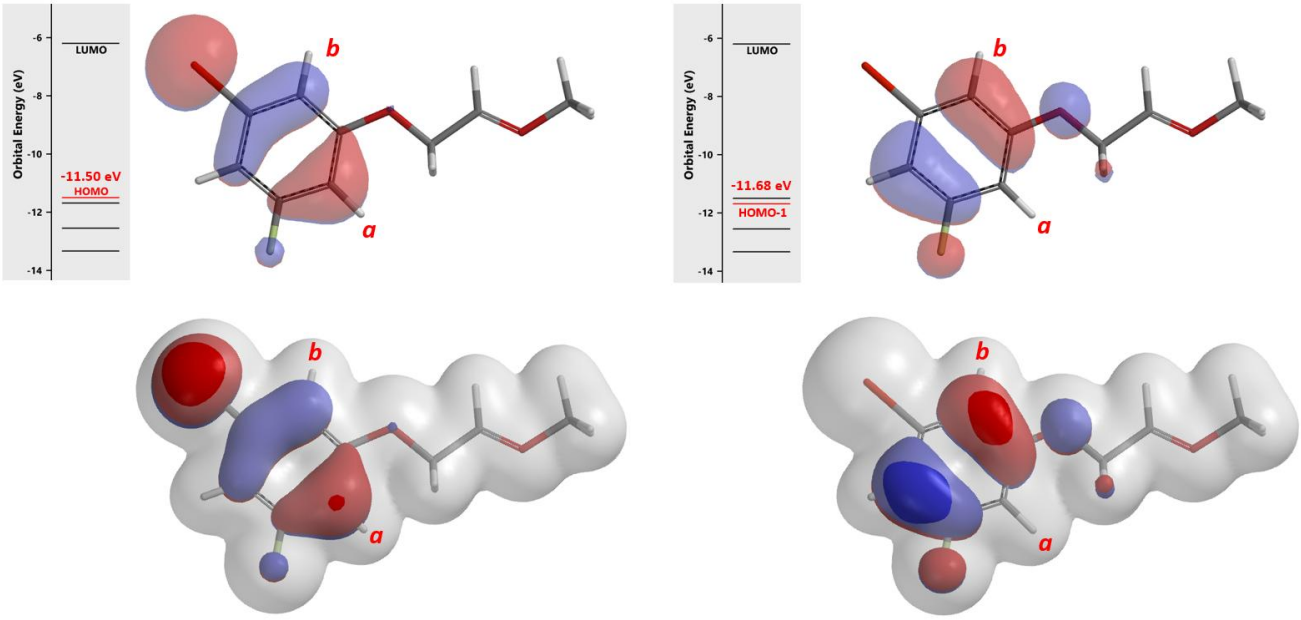


Figure 6. Oxonium ion **A** HOMO/HOMO-1 and overlays with Electron Density Surface (IsoVal 0.0065 e/au^3, 98.20%)

Shown in Figure 6 are the HOMO (-11.50 eV) and HOMO-1 (-11.68 eV) calculated for oxonium ion **A**. Since energy difference between the two orbitals is small, 0.18 eV, we need to consider both of them for subsequent cyclization. In this situation, we learned to superimpose HOMO and HOMO-1 onto Electron Density Surface (EDS)[3] of **A** for further evaluation. The electron density isosurface-orbital overlays revealed that HOMO-1 lobe at site *b* is a lot more accessible than HOMO lobe at site *a*, suggesting that **2b** could be the major product.

Concurrently, we calculated for the reaction energy profiles for the two competition pathways. As shown in Figure 7, activation energy required for the cyclization reaction at site *b* is ~9.22 kcal/mol, which is 0.94 kcal/mol lower than the 10.16 kcal/mol required for reaction at site *a*, substantiating the orbital-EDS analyses above.

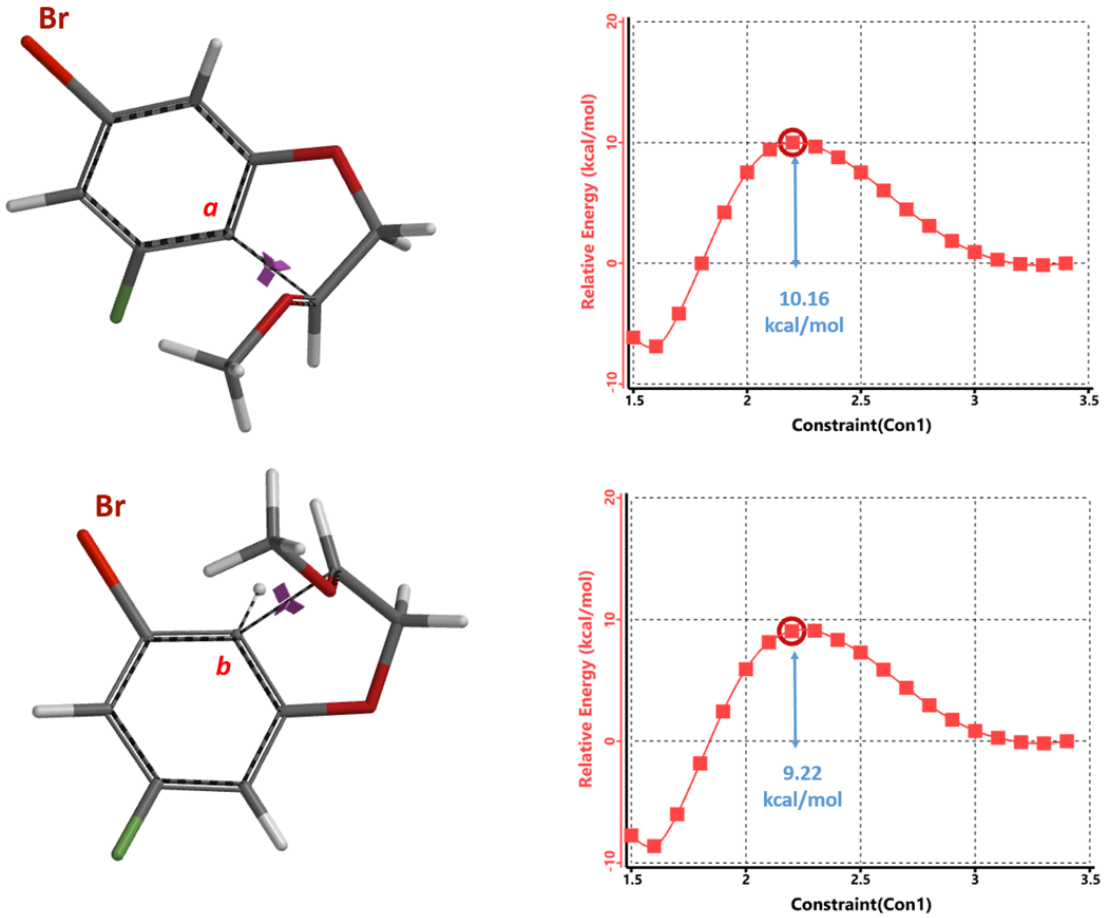


Figure 7. Reaction energy profiles of cyclization leading to **2a** and **2b**

Based on the Arrhenius equation, a 0.94 kcal/mol difference in activation energy at 110 °C translates to a **2a**: **2b** product ratio of ~1:3.44, consistent with the observed ratio of 1:5.

T (°C)	E_a-E_b (kcal mol^{-1})	(E_a-E_b)/RT	2a/2b ratio
110	0.94	1.24	1 : 3.44

Table 1. Calculated product ratio

It is worth noting that even though activation energy of the cyclization reaction is ~10 kcal/mol, a reaction temperature of 110 °C is required, presumably for conversion of substrate **1** to oxonium ion **A**.

Conclusion

In organic chemistry, mechanistic insights focus our QM analyses on crucial structures/factors that determine reactivities observed. **What to calculate, How to calculate, How to interpret/question the results.** "To ask the right question is already half the solution of a problem."-Carl Jung

Building on What We Just Learned

Acid catalyzed cyclization of acetal **3** is reported to be highly regioselective[4]. Since HOMO and HOMO-1 of corresponding oxonium ion have a significant energy gap of 0.66 eV, we could focus QM analysis on the overlay of HOMO and EDS of the oxonium ion (Figure 8). Is **4a** or **4b** the major product observed?

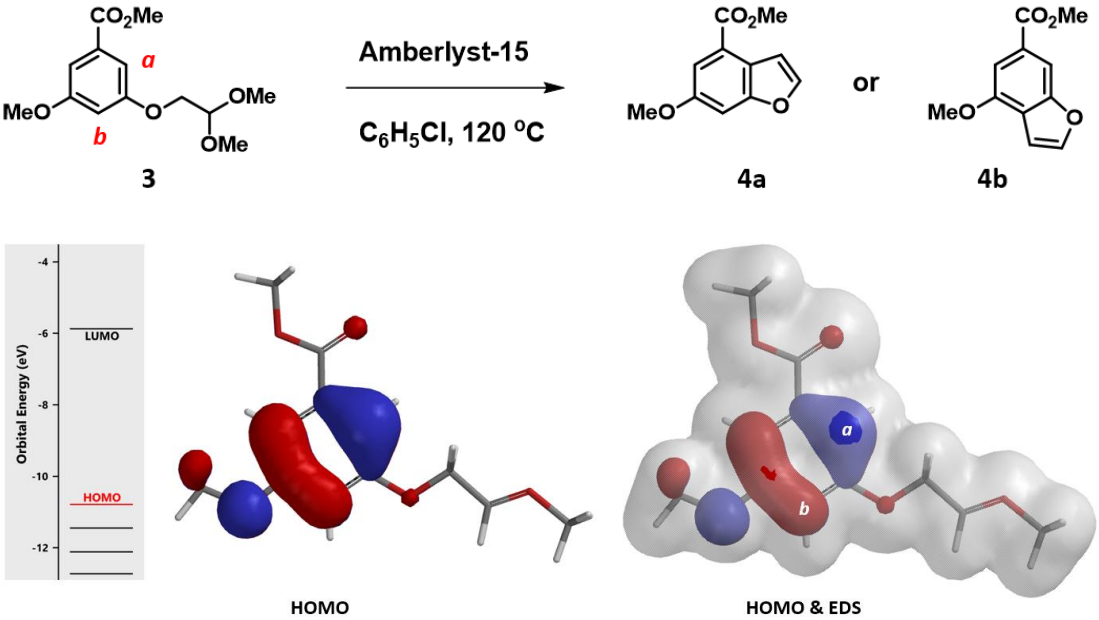


Figure 8. Regioselective conversion of acetal **3** to benzofuran **4**. HOMO and its overlay with EDS of the oxonium ion.

Return to Table of Contents

References:

[1] R.J. Nevagia, S.N. Dighe, S.N. Dighe, *Eur. J. Med. Chem*. **2015**, *97*, 561.

[2] A. Lavanya, K. Narasimhan, V. Padmini, *Mini-Rev. Org. Chem.* **2019**, *17*, 224.

[3] *Spartan'20 Tutorial and User's Guide* (2020). Irvine, CA, USA: Wavefunction, Inc. p368.

Electron Density Surface (EDS) that contains >99% of a molecule's electrons roughly corresponds to a space-filling model, that is, a van der Waals surface. EDS reveals overall molecular size and shape and demarks the steric barrier seen by encroaching molecules, and could be overlaid with molecular orbitals and mapped with electrostatic potential to rationalize physical and chemical properties of molecules.

[4] J-T. Liu, T.J. Do, C.J. Simmons, J.C. Lynch, W. Gu, Z.X. Ma, W. Xu, W.P. Tang, *Org. Biomol. Chem.* **2016**, *14*, 8927.

Chapter 32 One versatile intermediate, three different heterocyclic products

Qiuyue Wang, Shouliang Wang, Yongsheng Chen, John S. Wai

In organic synthesis, we encounter interesting substrates which under different reaction conditions provide different products. For example, reaction of the common intermediate **2**, formed by reaction of acyl hydrazide **1** with isothiocyanate or by coupling of carboxylic acid **3** with hydrazinecarbothioamide (Figure 1), under acidic, basic, and neutral conditions, provided thiadiazole **4**, triazole **5**, and oxadiazole **6**, respectively[1,2,3]. In this chapter, we'll analyze these reactions with QM.

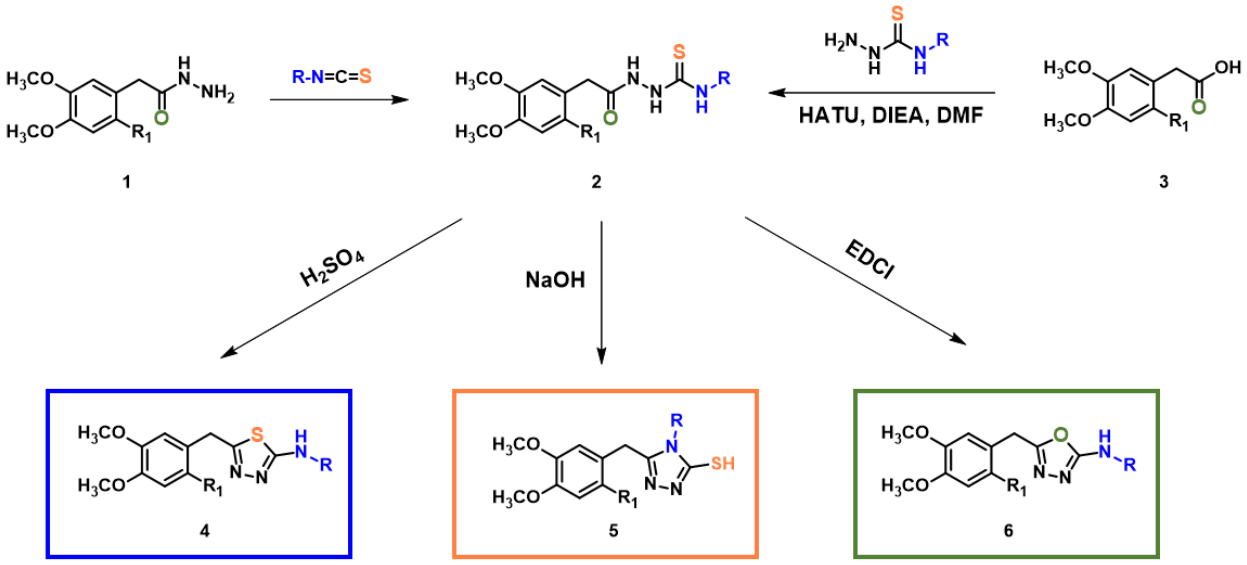


Figure 1. Intermediate **2** provided thiadiazole, triazole, and oxadiazole under different reaction conditions.

In organic chemistry, mechanistic insights are crucial in guiding our QM analyses: what are the structures/factors that could determine reactivities observed, what to calculate, how to calculate.

Formation of thiadiazole under acidic conditions

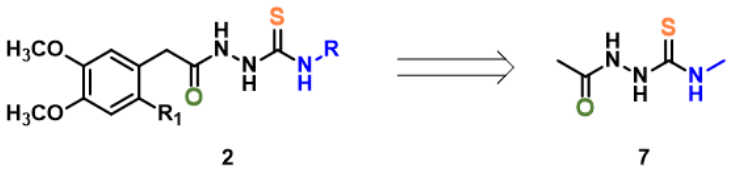


Figure 2. Simplified intermediate **2** as 2-acetyl-*N*-methylhydrazinecarbothiamide **7** for calculations.

For computational efficiency, we simplified intermediate **2** as 2-acetyl-*N*-methylhydrazine-carbothiamide **7** (Figure 2). For acid mediated reactions, we need to find out which atom on **7** could be selectively protonated. We chose Electrostatic Potential Map and Natural Charge for the analysis[4]. With electrostatic potential map, by convention, colors toward red correspond to negative potential (stabilizing interaction between the molecule and a positive charge, eg. proton), while colors toward blue correspond to positive potential.

For structure **7**, the most electronegative area (-176.7 kJ/mol) is on the carbonyl oxygen (Figure 3), consistent with the more negative -0.655 unit Natural Charge calculated for it. Electrostatic potential could also be displayed as two dimensional contour plot (slice). As such, preferential protonation leads to intermediate **8**. Calculation indicated that its HOMO mainly distributed on the sulfur atom, the nucleophilic center of the intermediate (Figure 3).

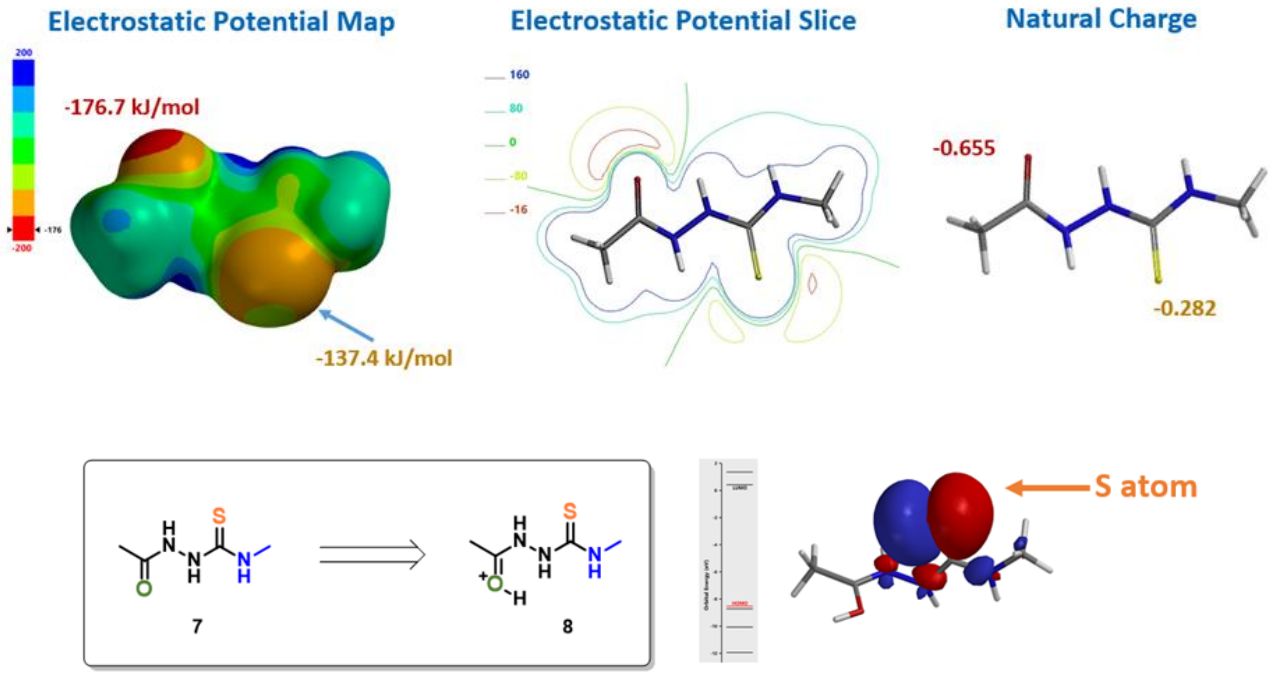


Figure 3. Electrostatic potential map, electrostatic potential slice, and natural charge of structure **7** (above); HOMO of structure **8** (below).

Reaction energy profile calculated for the cyclization step showed an activation energy of 10 kcal/mol (Figure 4), consistent with the observation that the reaction proceeds at 0 °C. Further Transition State calculation with structure **8b** showed only one imaginary frequency, i243 cm^{-1}, which corresponds to the vibration of the C-S bond forming between the carbonyl carbon and sulfur atoms.

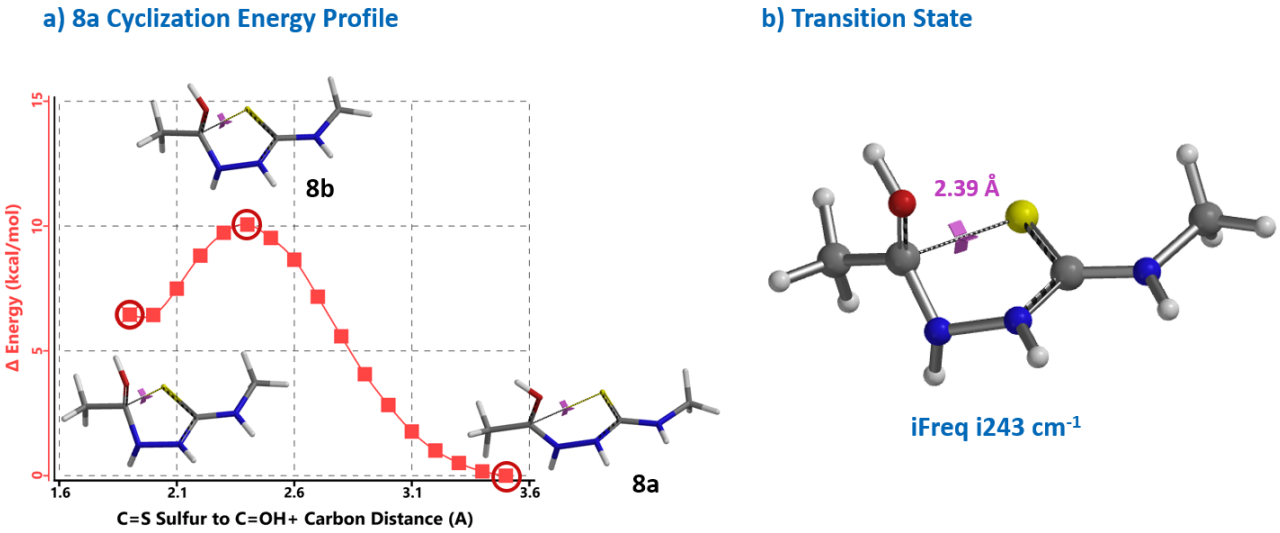


Figure 4. Reaction energy profile & transition state calculations for thiadiazole synthesis.

Formation of triazole under basic conditions

Electrostatic Potential Map revealed that the thioamide N**H** on structure **7** is the most acidic proton. Base deprotonation leads to anion **9**, with the negative charge delocalized among the N-C-S atoms (Figure 5).

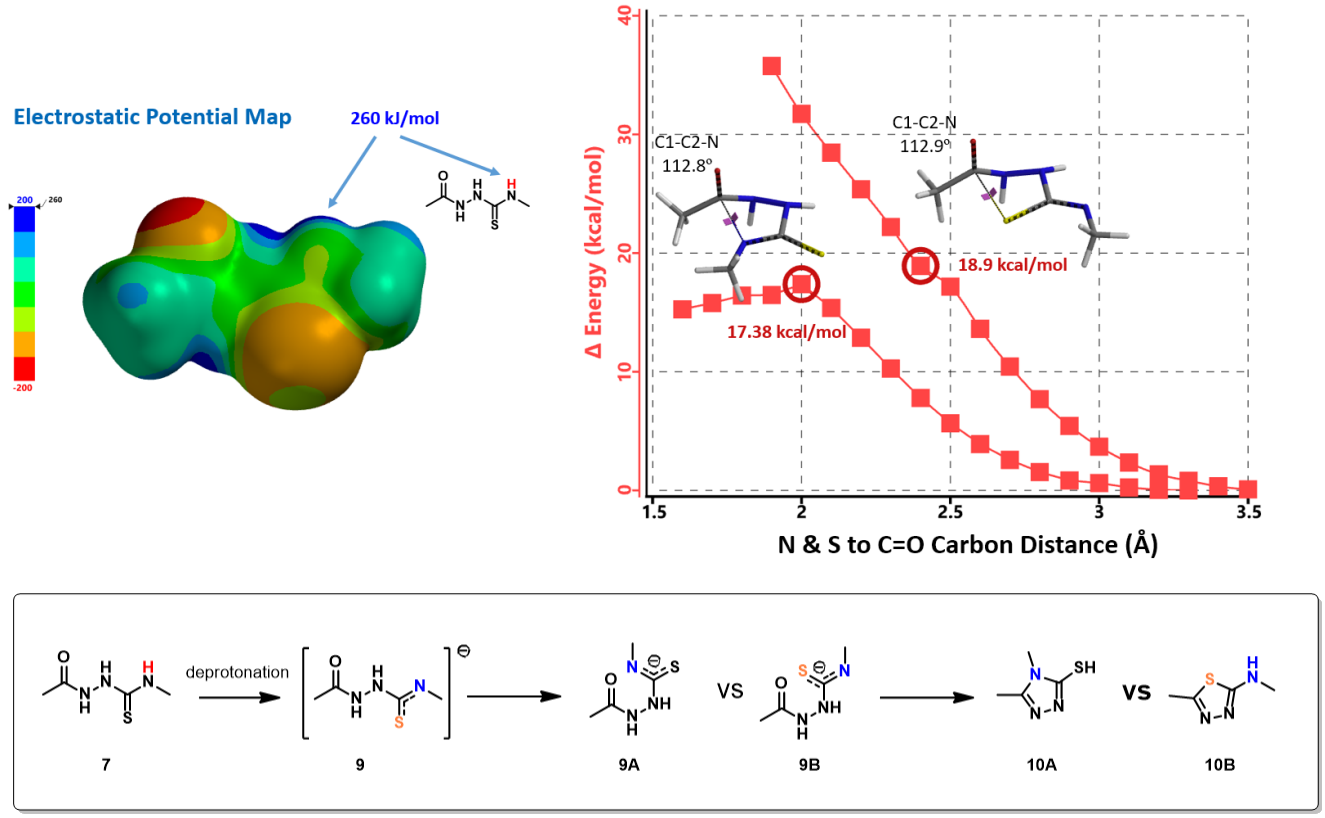


Figure 5. Compound **7** electrostatic potential map (top left); Reaction energy profiles for formation of triazole vs thiadiazole (top right). Reaction scheme under basic conditions (bottom).

Reaction energy profile calculated for formation of triazole **10A** suggested an activation energy of 17.38 kcal/mol, consistent with reaction temperature of 50-100 °C required. For potential competing cyclization to thiadiazole **10B**, calculated reaction profile is significantly

higher in energy which continues to rise. This is consistent with our observations that only the triazole cyclization product **10A** was formed.

Formation of oxadiazole under neutral conditions

Several neutral conditions were reported for cyclization of structure **7** to oxadiazole. In this chapter, we'll focus our discussion with the use of 1-(3-dimethylaminopropyl)-3-ethylcarbo-diimide (EDCI). We reasoned that the reaction is a two-steps process, hydrogen sulfide transfer from **7** to EDCI, and cyclization of resultant diimide **11** to oxadiazole **12**. (Figure 6).

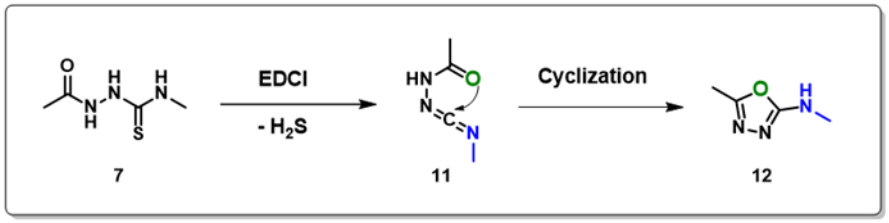


Figure 6. Reaction scheme for formation of oxadiazole

Calculations indicated that the hydrogen sulfide transfer step is exothermic (-6.38 kcal/mol), and activation energy for the cyclization step is 7.92 kcal/mol. Both readily proceed.

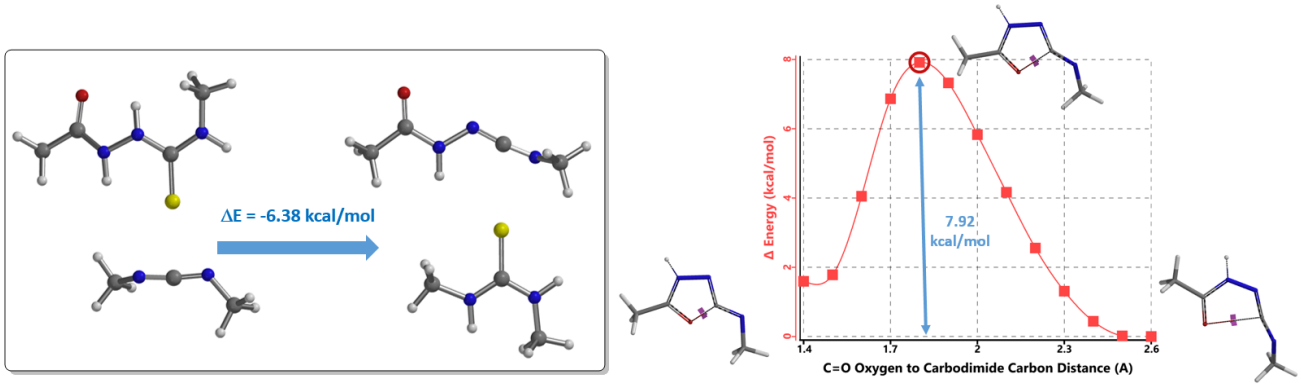


Figure 7. H_2S transfer step (Left); Cyclization step, reaction energy profile (Right)

Conclusion

QM analyses enable us to learn organic chemistry reactions at deeper levels. It requires us to be aware of the chemical properties of the substrates, reagents, reaction conditions, and reaction mechanisms, to find out what relevant QM parameters to calculate for to correlate with experimental results, to incorporate what we learned into our workflow, to turn QM into prospective tools. All these build on one another, require persistent efforts, to learn and understand the reactions of our interest, one at a time. Enjoy.

Building on What We Just Learned

In the hydrogen sulfide transfer step discussed above, for the conversion of **13** to **11**, (Figure 8), which mechanism is more consistent with QM calculation results? A concerted process or stepwise one?

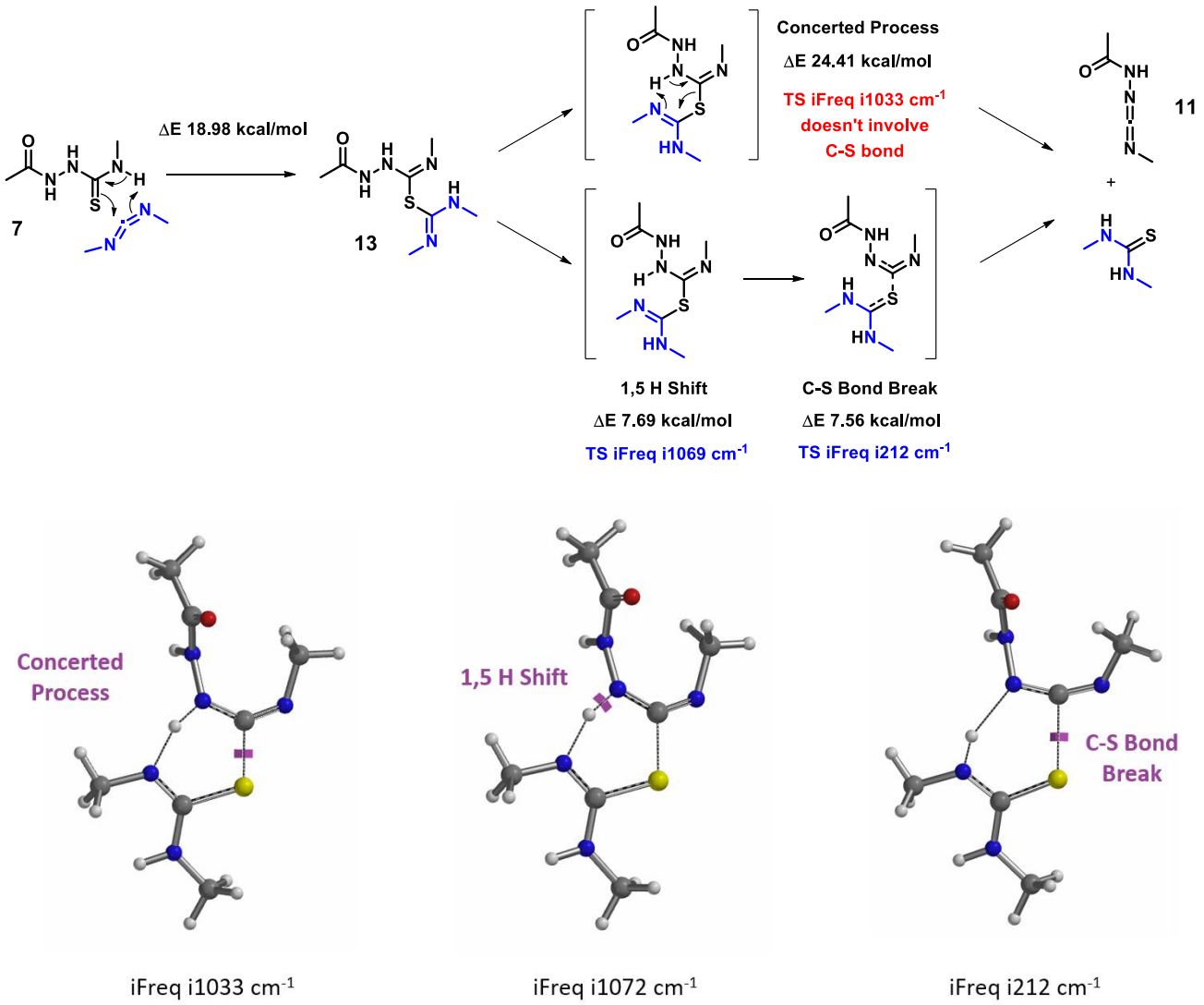


Figure 8. Hydrogen sulfide transfer step under EDCI neutral conditions.

Return to Table of Contents

References:

[1] P. Zoumpoulakis, C. Camoutsis, G. Pairas, M. Sokovic, J. Glamoclija, C. Potamitis, A. Pitas, *Bioorg. Med. Chem. Lett*, **2012**, *20*, 1569.

[2] I. Saramet, A. Banciu, L. Socea, C. Draghici, M.D. Banciu, *Heterocyclic Communications*, **2003**, *9*, 653.

[3] F. Naaz, F. Ahmad, B.A. Lone, Y.R. Pokharel, N.K. Fuloria, S. Fuloria, M. Ravichandran, L. Pattabhiraman, S. Shaf, M.S. Yar, *Bioorganic Chemistry*, **2020**, *95*, 103519.

[4] *Spartan'20 Tutorial and User's Guide* (2020). Irvine, CA, USA: Wavefunction, Inc. p599; A.E. Reed, R.B. Weinstock, F. Weinhold, *J. Chem. Phys.* **1985**, *83*, 735; T. Lu T, F.W. Chen, *Acta Phys.-Chim. Sin.* **2012**, *28*, 1.

Chapter 33 QM analyses for Cysteine – Chloroacetamide Bioconjugation

Shouliang Wang, Zhong Zheng, Qiuyue Wang, Jian Wang, Yongsheng Chen, John S. Wai

Bioconjugation of proteins with thio modification of cysteine residues usually exhibits excellent chemoselectivity as free cysteines rarely occur on protein surface. Electrophiles, such as iodoacetamide, maleimide, and disulfide derivatives are usually used. Shown on Figure 1 is an excellent example in which the iodoacetamide and maleimide strategies are being used sequentially on XTEN to achieve high Drug-to-antibody ratio (DAR) [1]. A recent report showed that chloroacetamide exhibits even greater specificity than iodoacetamide for cysteine residues in bioconjugation[2]. In this chapter we'll discuss the difference observed in reactivity between *N,N*-dimethyl and *N*-methyl chloroacetamide in thiol conjugation and results from our QM analyses.

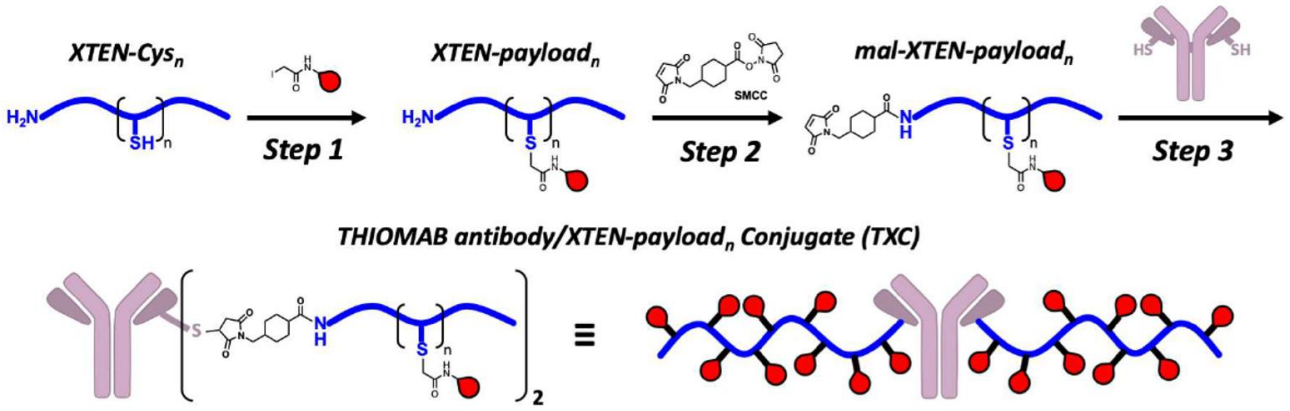


Figure 1. A thiol bioconjugation strategy for high Drug-to-antibody ratio (DAR).

N,N-Dimethyl *versus N*-Methyl chloroacetamide

Shown in Figure 2 is the difference in reactivity observed between *N,N*-dimethyl and *N*-methyl chloroacetamides with a thiol. Intuitively, we were expecting the opposite, that is, the incoming thio will face more steric hindrance with the tertiary amide. To understand why, we first examined LUMO and LUMO map of the two acetamides (Figure 3).

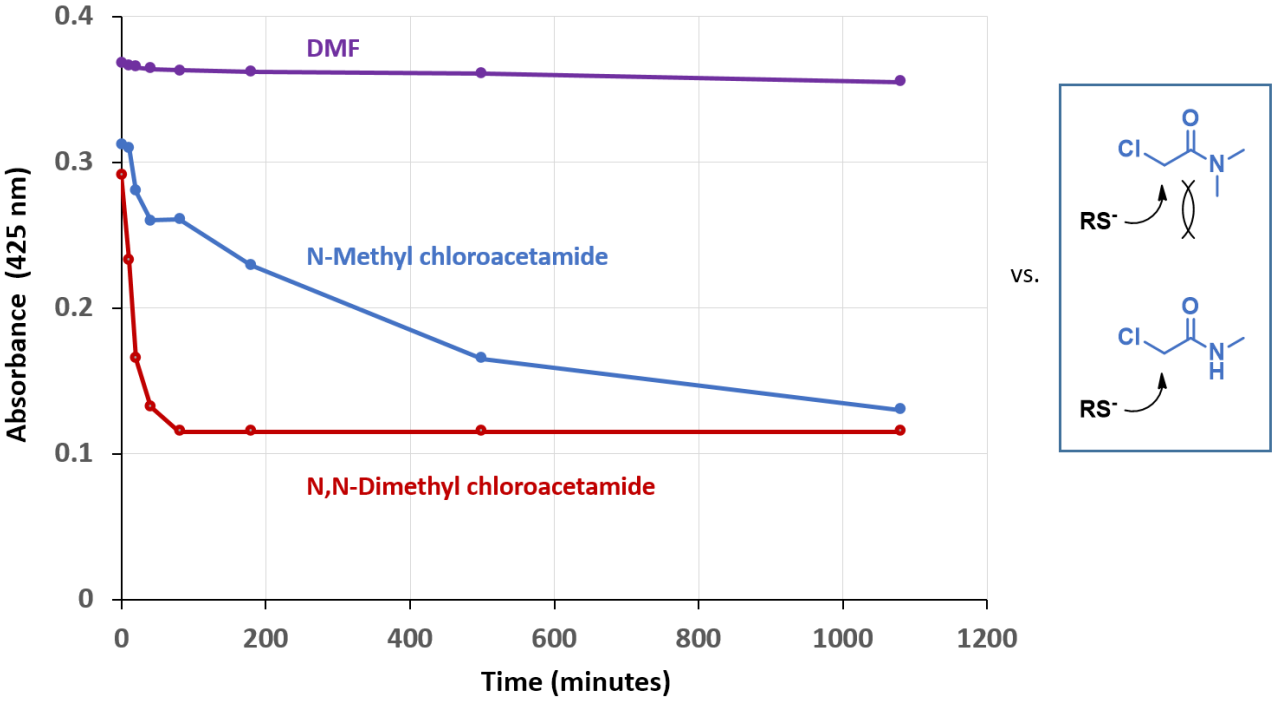


Figure 2. Difference in reactivity observed between *N,N*-dimethyl and *N*-methyl chloroacetamides with a thiol *vs* expectation.

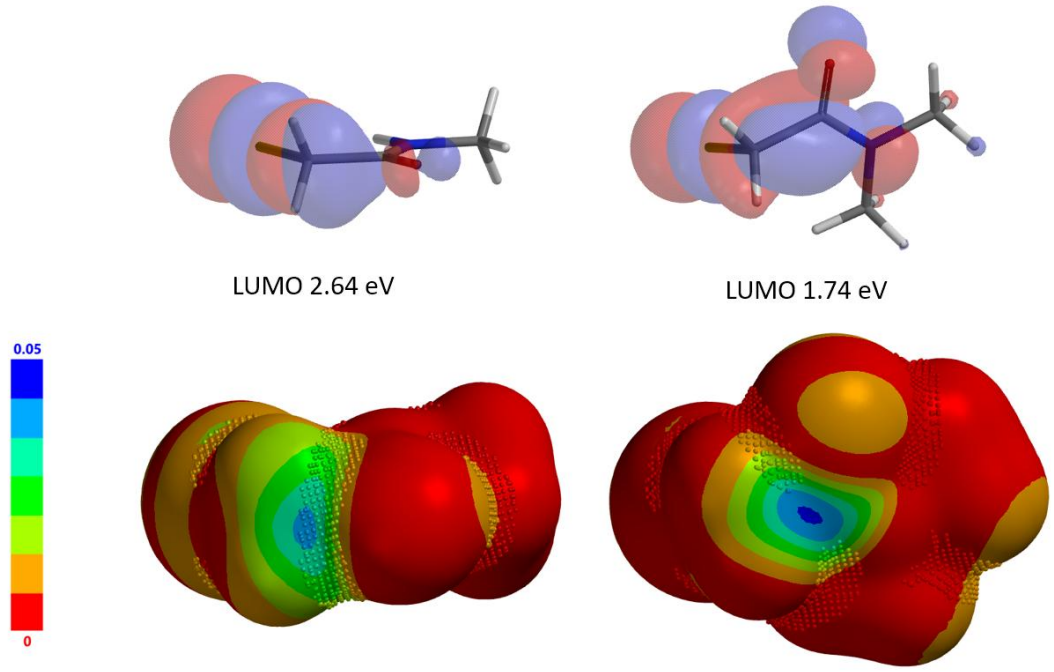


Figure 3. LUMO and LUMO map of *N,N*-dimethyl and *N*-methyl chloroacetamides (IsoVal 0.002 e/au^3, 99.2%).

LUMO energy of *N*-methyl chloroacetamide is 2.64 eV. Its LUMO map and inaccessibility marker suggest that access to the LUMO lobe across on the C-Cl carbon is limited. On the other hand, LUMO energy of *N,N*-dimethyl chloroacetamide is lower, 1.74 eV, and the LUMO lobe across on the C-Cl carbon is more exposed and accessible, suggesting that it shall be more reactive, consistent with experimental observations. Done? We reasoned that there should be additional reasons and decided to conduct further QM analyses of this pair of reactions.

Reaction energy profile and associated conformational changes

First, we modeled for the reaction energy profile of *N,N*-dimethyl chloroacetamide, using methyl thiolate as cysteine surrogate to speed up our calculations (DFT B3LYP-D3/6+31G*, the same below).

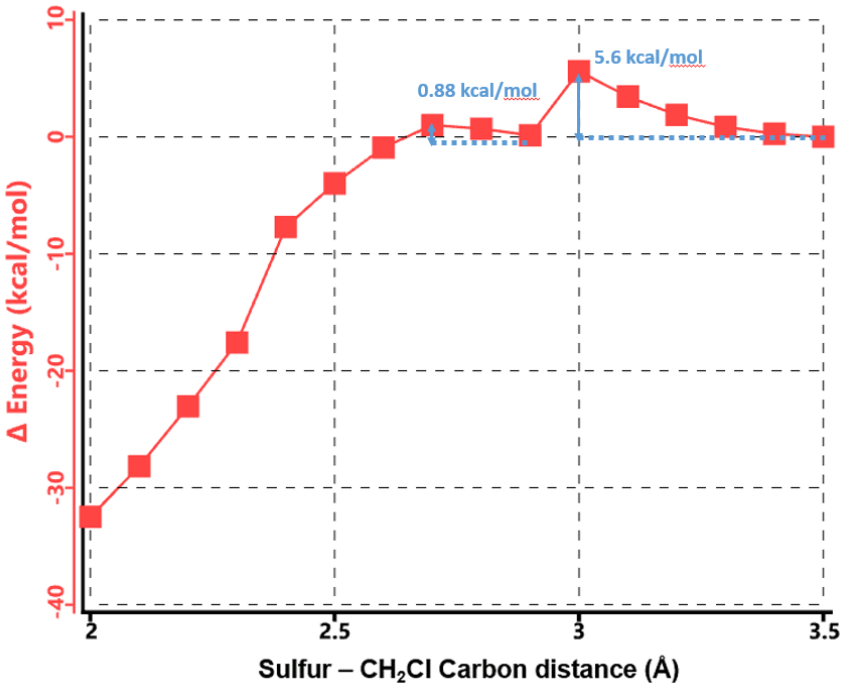


Figure 4. Reaction energy profile of *N,N*-dimethyl chloroacetamide and methyl thiolate.

As shown in Figure 4, there are two relatively low energy barriers, 5.60 and 0.88 kcal/mol, associated with the reaction, peaking at sulfur to CH_2Cl carbon distance of 3.0 Å and 2.6 Å, respectively, while *N,N*-dimethyl chloroacetamide undergoes remarkable conformational changes. When the sulfur and carbon are 3.5 Å apart, the carbon-chlorine bond is almost coplanar with the amide group. Thio displacement of the chloride requires the nucleophile to approach from the backside, and faces steric hindrance from the amide methyl group as energy peaks at 3.0 Å. The barrier is overcome at 2.9 Å, when the carbon-chlorine bond rotates perpendicular to the amide plane and lines up perfectly for the nucleophilic substitution reaction, with a second energy barrier of only 0.88 kcal/mol.

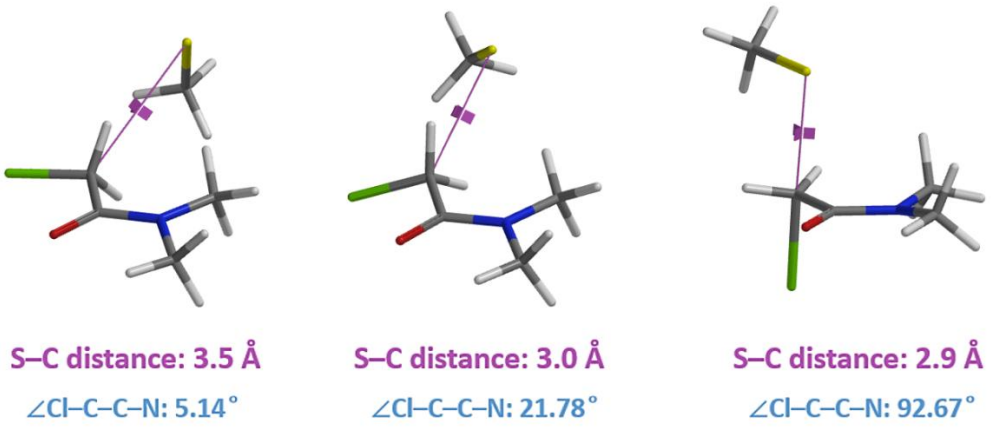


Figure 5. Conformational and dihedral angle Cl–C–C–N changes in reaction with *N,N*-dimethyl chloroacetamide.

Similarly, we calculated for the reaction energy profile of *N*-methyl chloroacetamide (Figure 6), which exhibits a single energy peak. At ~ 10 kcal/mol, it is significantly higher than that of *N,N*-dimethyl chloroacetamide. Both chloroacetamides undergo the same remarkable conformational changes in the process, that is, the carbon-chlorine bond rotates from almost coplanar to perpendicular to the amide group (Figure 7, top).

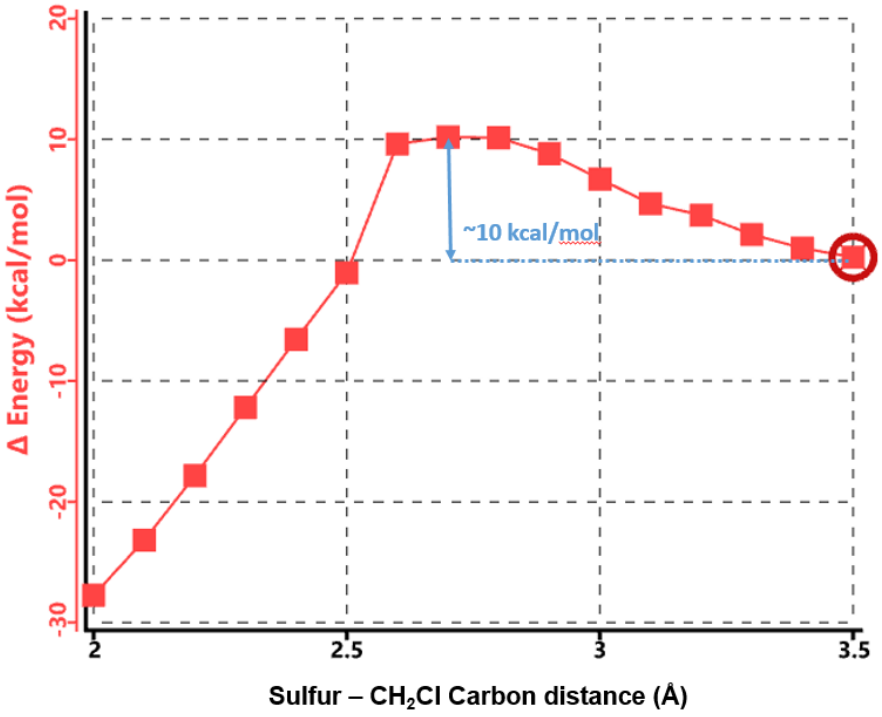


Figure 6. Reaction energy profile of *N*-methyl chloroacetamide and methyl thiolate.

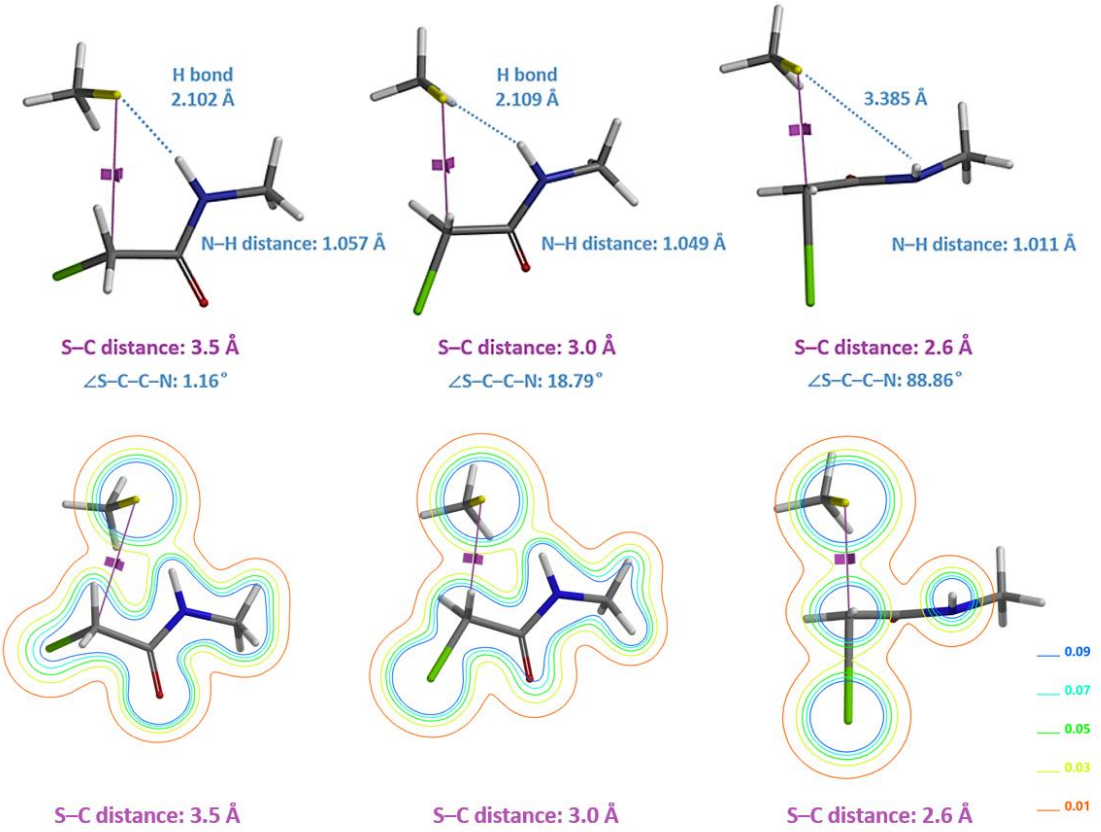


Figure 7. Conformational, dihedral angle Cl–C–C–N, and electron density changes in reaction with *N*-methyl chloroacetamide.

However, for *N*-methyl chloroacetamide, the incoming thiolate hydrogen bonds with the amide NH (Figure 7, S-C distance 3.5 and 3.0 Å), as indicated by the S···H–N bond distances and electron density slices calculated. For the nucleophilic substitution to proceed further, this hydrogen bond needs to be broken, incurring energy cost to reach transition state (Figure 7, bottom, S-C distance 2.6 Å), consistent with lower reactivity observed with *N*-methyl chloroacetamide.

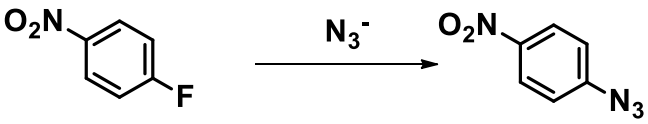
We learned from the example in chapter 26 that hydrogen bond interactions could stabilize transition state, accelerating the rate of the reaction. In this chapter, we learned that when the hydrogen bond stabilizes the reacting complex at early stages and needs to be broken to reach transition state, it could slow down the reaction. For chloroacetamide bioconjugation on cysteine residues, this turns out to be a good thing.

Conclusion

For organic chemistry, we first consider the steric and electronic effect of reactants and reagents at ground state, then intuitively evaluate potential non-covalent interactions that could have a significant impact on the reaction process. QM enables us to quantify such interactions to account for our observations, improves our intuition, and guides our chemistry.

Building on What We Just Learned

S_NAr reactions usually proceed more readily in polar aprotic solvents (DMSO, MeCN, DMF, etc) than in polar protic ones, as shown for S_NAr reactions of 1-fluoro-4-nitrobenzene with azide (Table 1)[4].



Solvent	-Log *k*
DMSO	3.28
MeCN	3.26
DMF	2.76
Water	7.40
MeOH	7.20

Table 1: Reaction rate constants (log k) of *n*-Bu_4N_3 with 1-fluoro-4-nitrobenzene in different solvents at 25 °C[4]

In our laboratories, we encountered remarkable S_NAr reactions that did not proceed at all in DMSO, NMP, MeCN, but proceeded readily in alcohols, as exemplified by the reaction of primary amines with 2-fluoropyridine. Since primary amines are highly solvated in polar protic solvents, their reactions in alcohol were anticipated to be slower than in polar aprotic ones. Why are we observing the opposite?

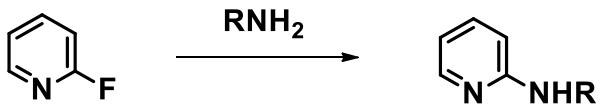


Figure 8. SNAr Reaction of 2-fluoropyridine with primary amines

Return to Table of Contents

References:

[1] https://www.researchsquare.com/article/rs-356231/v1

[2] J. Kalia, R. T. Raines, *Curr. Org. Chem.* **2010**, *14*, 138.

[3] P. Atkins, J. de Paula, J, Keeler, *Atkins' Physical Chemistry* (11th ed.). Oxford University Press. Oxford, 2018: pp 584

[4] B.G. Cox, A.J. Parker, *J. Am. Chem. Soc*. **1973**, *95*, 408.

Chapter 34 QM Analyses of Regioselectivity in Chan-Lam Reaction

Xiaocang Wei, Qiuyue Wang, Jian Wang, Yongsheng Chen, John S. Wai

A Tianjin site colleague raised a very practical question, how to use orbital analysis to predict regioselectivity of Chan-Lam reaction between indazole **1** and phenylboronic acid **2** (Figure 1).

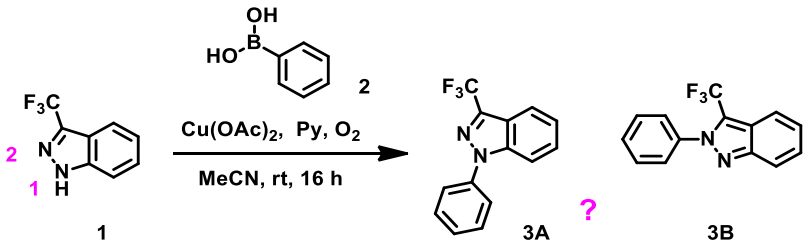


Figure 1. Chan-Lam reaction between indazole **1** and phenylboronic acid **2**.

Mechanism of Chan-Lam Reaction

First question to ourselves became "shall we calculate for HOMO of the neutral form of the molecule or its anion for analysis?" Shown on Figure 2 is the proposed mechanism for Chan-Lam reaction[1]. Arylboronic acid undergoes transmetalation with divalent copper complex **E** to form intermediate **A**, which coordinates with deprotonated heterocycles to form Cu(II) complex **B**, which undergoes air oxidation to form Cu(III) Intermediate **C**. Intermediate **C** undergoes reductive elimination to form coupling product and Cu(I) complex **D**, which is further oxidized to regenerate Cu(II) complex **E**. As such, deprotonation of the heterocycles and addition of the corresponding anion to form Cu(II) **B** are the crucial steps in the catalytic cycle. To understand selectivity of the reaction, we should look at HOMO of the anion, not HOMO of the neutral molecule.

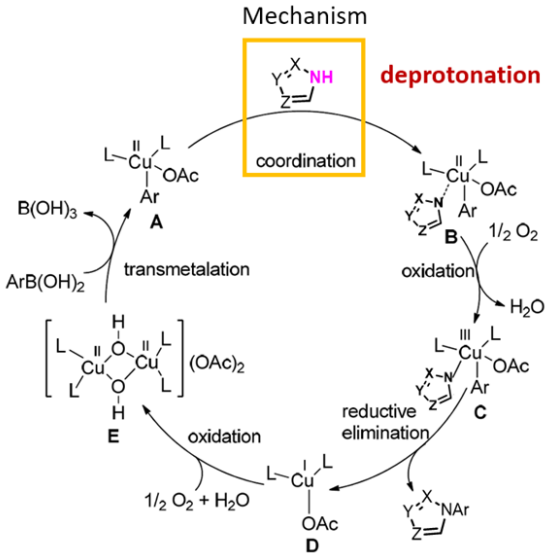


Figure 2. Possible mechanism of Chan-Lam reaction[1]

HOMO analysis

Shown on Figure 3 is HOMO of indazole **1** anion. There is significant HOMO lobe on N1, but almost nothing on N2, suggesting that Chan-Lam reaction will occur preferentially at N1. The determining factor in this reaction is electronic effect, rather than the potential steric interaction from the CF_3 group.

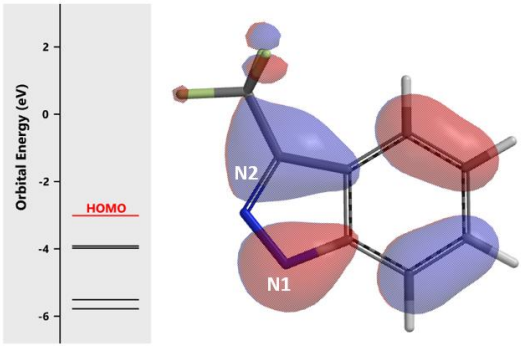


Figure 3. HOMO of indazole **1** anion

Experimentally, Tianjin colleague observed formation of N1 phenylated product **3A** only (Figure 4), lending support for the use of HOMO anion for prospective regioselectivity analysis of the reaction.

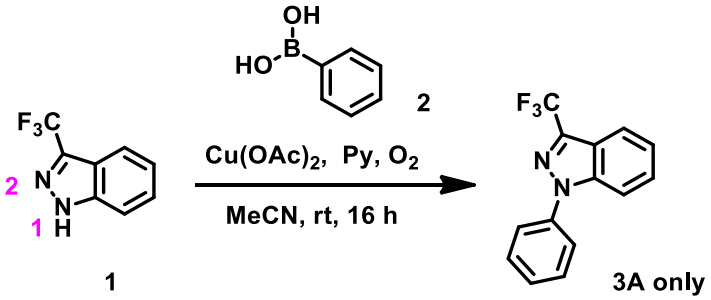


Figure 4. Chan-Lam reaction between indazole **1** and phenylboronic acid **2**.

Further Testing of the HOMO anion Analysis

Next we collected additional examples from the literature for analysis[2]. Pyrazole (Table 1, entry 1), imidazole (entry 2), and benzimidazole (entry 7) do not require analyses for regioselectivity[2a]. HOMO of 1,2,3-triazole anion (Table 1 & Figure 5, entry 3) suggests Chan-Lam reaction will proceed predominately on N2, to a lesser extent on N1, in contrast to reported N1 arylation product only [2a]. **Perplexing!**

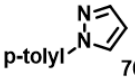
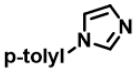
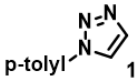
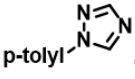
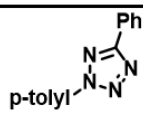
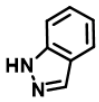
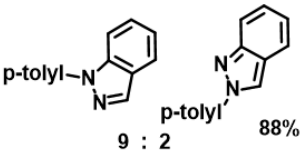
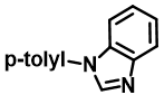
Entry	Heteroarene Substrate	Product and isolated yield
1		 76%
2		 72%
3		 11%
4		 6%
5		 26%
6		 9 : 2 88%
7		 67%

Table 1. Chan-Lam reaction between various heterocycles and *p*-tolylboronic acid[2a].

Moving down the Table, HOMO of 1,2,4-triazole anion (entry 4) suggests reaction to be on N1 or N2, consistent with the experimental result. For phenyl tetrazole (entry 5), HOMO lobes at four nitrogens are similar in size, while C5 phenyl group will hinder coordination with Cu(II) at N1 and N4 position. We expected reaction to occur mainly at N2 or N3 position, aligned with reported result. For indazole (entry 6) selective N1 phenylation was expected and observed. Regioselectivity for all other examples in the literature[2] are accounted for with HOMO anion analysis; the only exception is 1,2,3-triazole (entry 3) discussed above. We were curious and decided to repeat the experiment.

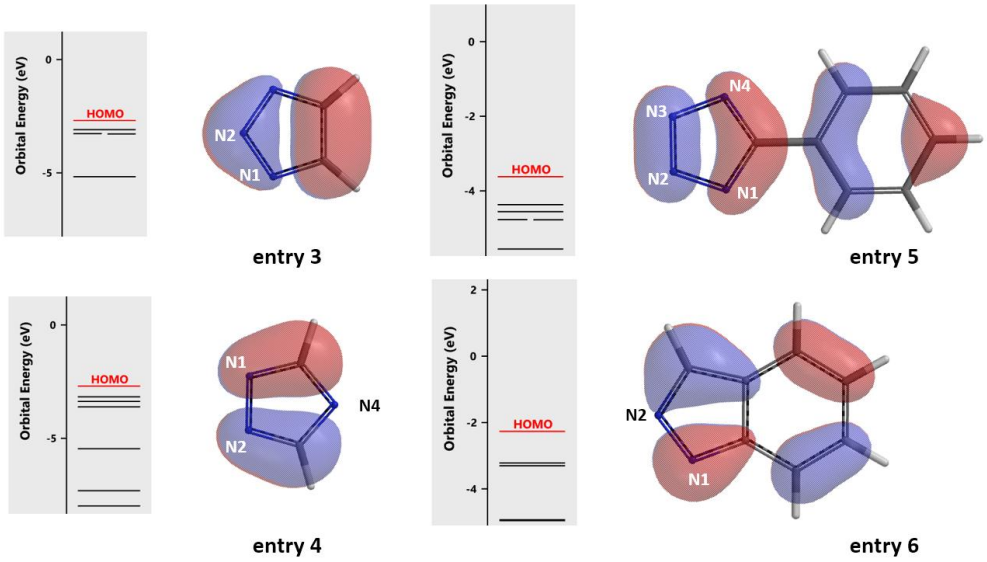


Figure 5. HOMO anion of entries **3 - 6**.

Before running the reaction, we also calculated dipole moments for the two isomeric products and the two potential by-products (aryl ether and aryl acetate) as a measure of their polarities. The dipole moments of the N2 and N1 arylation products are 0.49 and 4.88 Debye, respectively (Figure 6). The N2 product is significantly less polar than N1 one, and is even lower in polarity than aryl ether and aryl acetate.

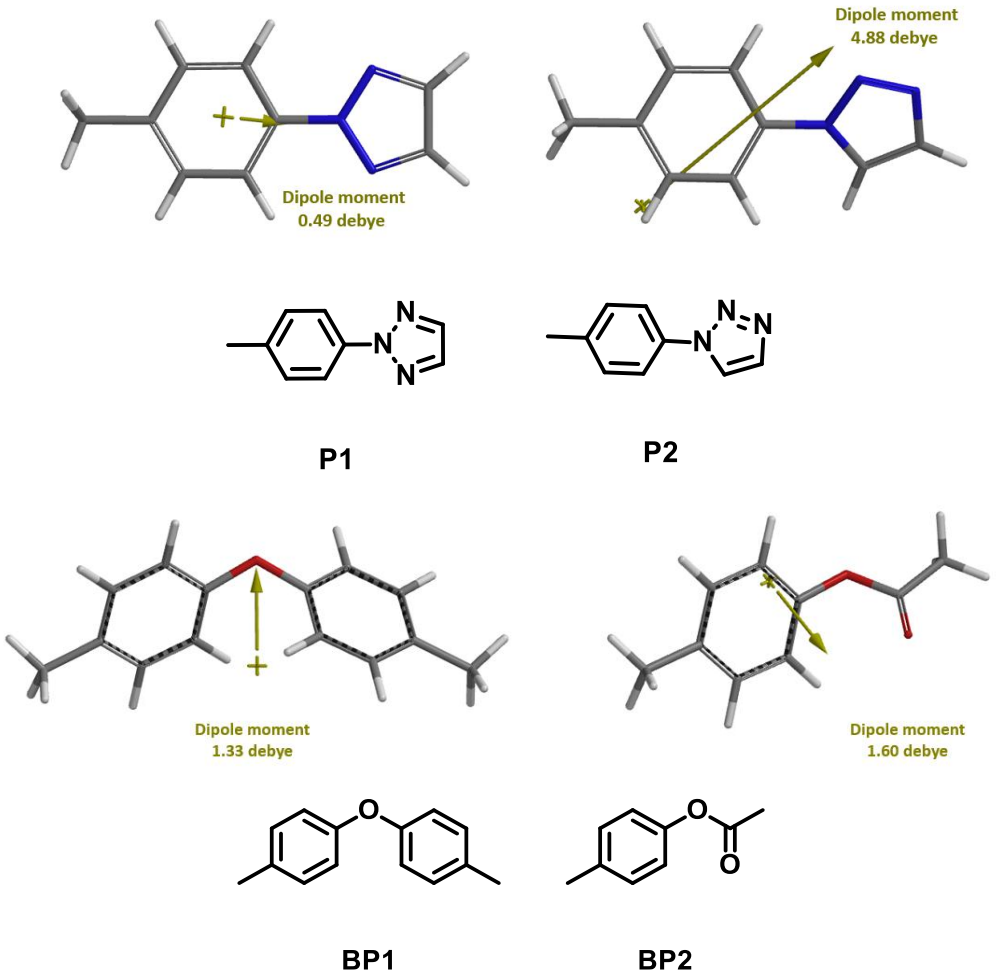


Figure 6. Dipole moment calculated for **P1, P2, BP1, BP2**.

Experimental Validation

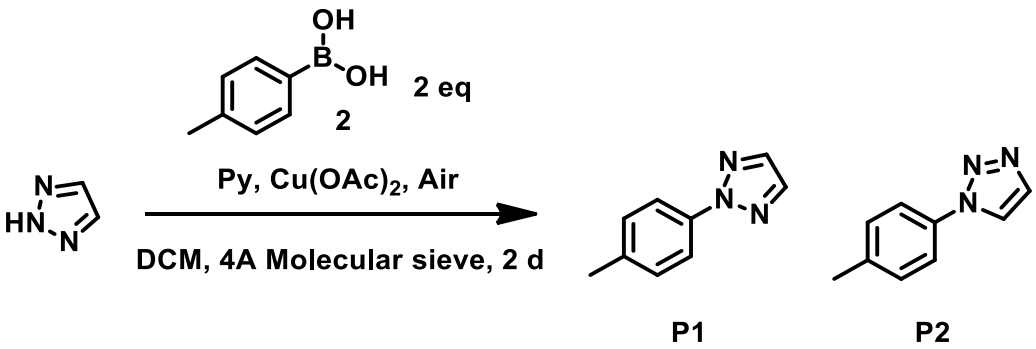


Figure 7. Chan-Lam reaction of 1,2,3-triazole and *p*-tolylboronic acid.

Reaction between 1,2,3-triazole and *p*-tolylboronic acid was repeated and products carefully separated and purified by column chromatography. We observed the formation of two products P1, P2 (~2 : 1), and two by-products BP1, BP2. ^{13}C NMR of P1, major product, has six sets of signals, corresponding to N2 arylation, while P2, minor product, exhibits seven sets of signals, indicating arylation is at N1 position. These are further confirmed by comparing ^{13}C and ^{1}H NMR shifts of both isomers reported in the literature [3],[4], further validating the prediction with HOMO analysis for 1,2,3-triazole anion, resolving the puzzle (Figure 8).

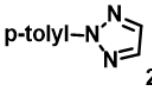
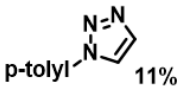
Entry	Heteroarene Substrate	Product and isolated yield
3		 20% major product  11% minor product

Figure 8. Major and minor products from Chan-Lam reaction between 1,2,3-triazole and *p*-tolylboronic acid

During purification, we found that the polarity of P1 is indeed much lower than that of P2, and lower than that of the two by-products BP1 and BP2, consistent with the calculated dipole moments. These parameters provided useful guidance in locating and purifying all four reaction products (Figure 6).

Conclusion

For Chan-Lam reaction, HOMO of substrate anion is a useful model for prospective analysis of regioselectivity. Chan-Lam reaction of 1,2,3-triazoles provides ~2 : 1 N2 *vs* N1 arylation products, consistent with the QM analysis.

Building on What We Just Learned

Chan-Lam Reaction is sensitive to the nature of heterocyclic substrates[2a]. For example, pyrrole and indole failed to react with *p*-tolylboronic acid, and the more acidic triazole and tetrazole did not react well. We learned that deprotonation of the NH proton on the heterocyclic ring and the coordination of the resultant anion onto Cu(II) **A** are crucial steps in the reaction. Tabulated on Table 2 are *pKa* of heterocyclic substrates and isolated yields for the reaction under standard conditions (pyridine as base and ligand at 25^{o}C). From your experience, what *pKa* range the heterocyclic substrates need to fall within for Chan-Lam reaction to be feasible?

	pKa[5]	Isolated Yield[2a]
Pyrrole	**23.0**	**<3%**
Indole	**21.0**	**<3%**
Pyrazole	**19.8**	**76%**
Imidazole	**18.6**	**72%**
Benzimidazole	**16.4**	**67%**
1,2,3-Triazole	**13.9**	**31%***
1,2,4-Triazole	**14.8**	**6%?**
Phenyl**tetrazole**	**8.2**	**26%**

Table 2. Isolation Yield [2a] and *pKa* of Heterocyclic Substrates [5]

* 31% (overall yield of P1 and P2)

Return to Table of Contents

References:

[1] X.P. Ma, F.P. Liu, D. L. Mo, *Chinese J. Org. Chem.,* **2017**, *37*, 1069.

[2] (a) P.Y.S. Lam, C.G. Clark, S. Saubernt, J. Adamst, M.P. Winters, D.M.T. Chan, A. Combst *Tetrahedron Lett*. **1998**, *39*, 2941; (b) D.M.T. Chan, P.Y.S. Lam, in *Boronic Acids*, Ed.: Hall, D. G., Wiley-VCH, Weinheim, 2005, Chapter 5, p205 and references therein cited; (c) J.X. Qiao, P.Y.S. Lam, in *Boronic Acids*, Ed.: Hall, D. G., Wiley-VCH, Weinheim, 2011, Chapter 6, p315 and references cited therein.

[3] Y.W Liu, C.M. Han, X.Y. Ma, J.H. Yang, X.P. Feng, Y.B. Jiang *Tetrahedron Lett.* **2018**, *59*, 650.

[4] C.Y. Chen, X.W. Lu, M.C. Holland, S.C Lv, X.B. Ji, W. Liu, J. Liu, D. Depre, P. Westerduin, *Eur. J. Org. Chem.,* **2020**, 548.

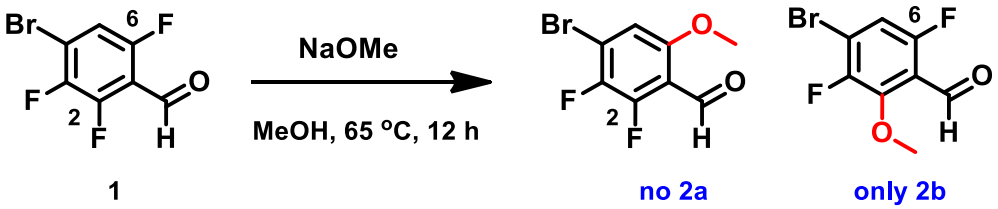
[5] https://organicchemistrydata.org/hansreich/resources/pka/pka_data/pka-compilation-reich-bordwell.pdf

Chapter 35 Highly regioselective S_NAr of a polyhalogenated benzaldehyde

Dajin Tan, Yongsheng Chen, Liting Dong, John S. Wai

In chapters 7 and 10[1], we discussed the use of LUMO and reaction energy profile for regioselectivity analyses of S_NAr reactions of polyhalogenated aromatic compounds in general. In this chapter, we will discuss what we learned with polyhalogenated benzaldehyde substrates specifically.

Shown in Figure 1 is the reaction of sodium methoxide with 2,3,6-trifluoro-4-bromo-benzaldehyde (**1**) with four potentially displaceable halogens. We learned from organic chemistry the following pattern F > Cl ≈ Br > I for leaving ability in S_NAr reaction. We reasoned that the fluoro groups at C2 and C6 will be more labile than the bromo group on C4[2], and the aldehyde group enhances their reactivity. It is not obvious to us whether the reaction will be selective at C2 or C6, or provide a mixture of both products. Experimentally, we observed only C2 displacement. Why is the reaction so selective? Benzaldehyde **1** is quite electron deficient, why do we need to warm up the reaction to 65 °C for it to proceed?



1

Figure 1. Nucleophilic substitution reaction of 2,3,6-trifluoro-4-bromobenzaldehyde with sodium methoxide

LUMO Analysis

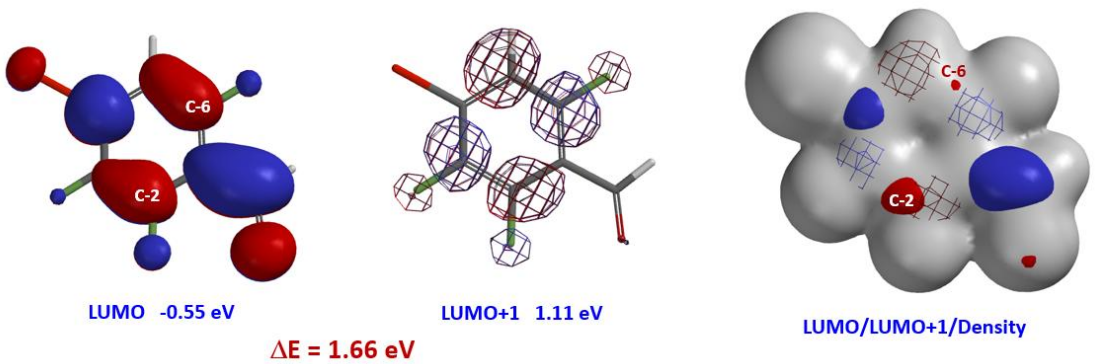


Figure 2. 2,3,6-trifluoro-4-bromobenzaldehyde: LUMO, LUMO+1, and their overlay with Electron Density Map

Shown in Figure 2 are LUMO (-0.55 eV) and LUMO+1 (1.11 eV) of benzaldehyde **1**. Since the energy difference between them, ΔE = 1.66 eV, is relatively large (Figure 2), we could focus our analysis on LUMO. Its lobes on C2 and C6 are similar in size, yet after we overlaid Electron Density Map[3] on the orbital, it became obvious that the LUMO lobe at C2 is more accessible than the one at C6, accounting for the selectivity observed.

Reaction temperature required is higher than expected

Next, we calculated for the reaction energy profile of the addition step at C2 and C6 positions toward formation of the corresponding Meisenheimer complexes.

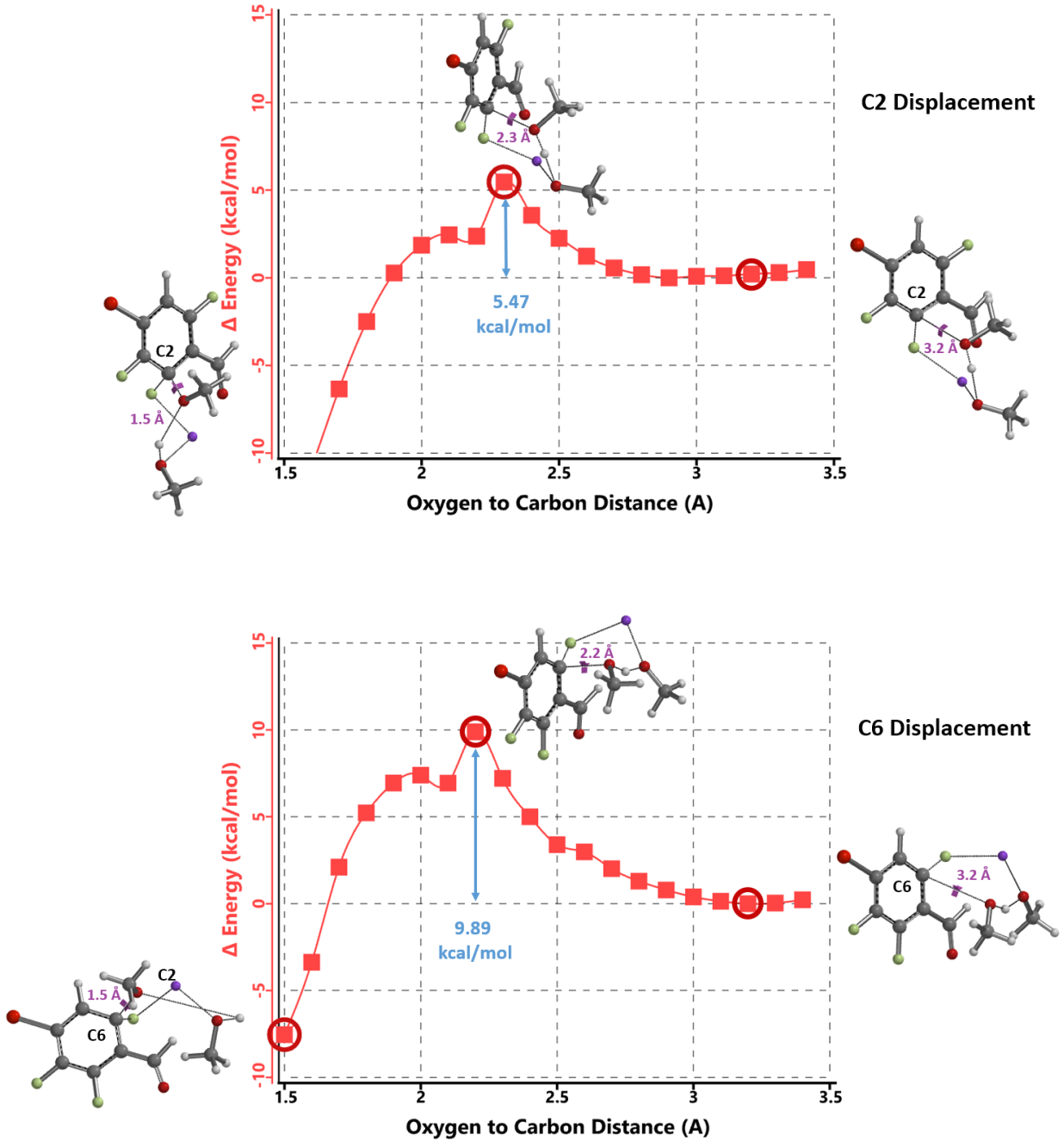


Figure 3. Reaction energy profiles for nucleophilic attack at C2 and C6

The results are shown in Figure 3. The activation energy required for addition at C6 position is 4.42 kcal/mol higher than that required for C2, suggesting that the reaction is more likely to occur at C2 position, consistent with the experimental results. The calculated activation energy of is relatively low at 5.47 kcal/mol, suggesting that the reaction could proceed at room temperature or lower, in contrast to the required experimental conditions (65 ^{o}C, 12 hours).

Reviewing LUMO of substrate **1** (Figure 2), we found LUMO lobe on the aldehyde carbon is significantly larger than that on C2 and C6. Methoxide nucleophilic addition is more likely to occur preferentially at the aldehyde group first. Above assumed reaction mechanism of direct displacement could be an oversimplification.

What will be a better mechanistic explanation?

We reasoned that the addition of sodium methoxide to the aldehyde group precedes nucleophilic attack at the C-F carbons. Reaction energy profile calculated for the formation of the hemiacetal sodium salt (Figure 4) suggests that addition is exothermic and proceeds with no energy barrier, supporting the hypothesis that the hemiacetal sodium salt is the "actual" substrate for subsequent nucleophilic aromatic substitution.

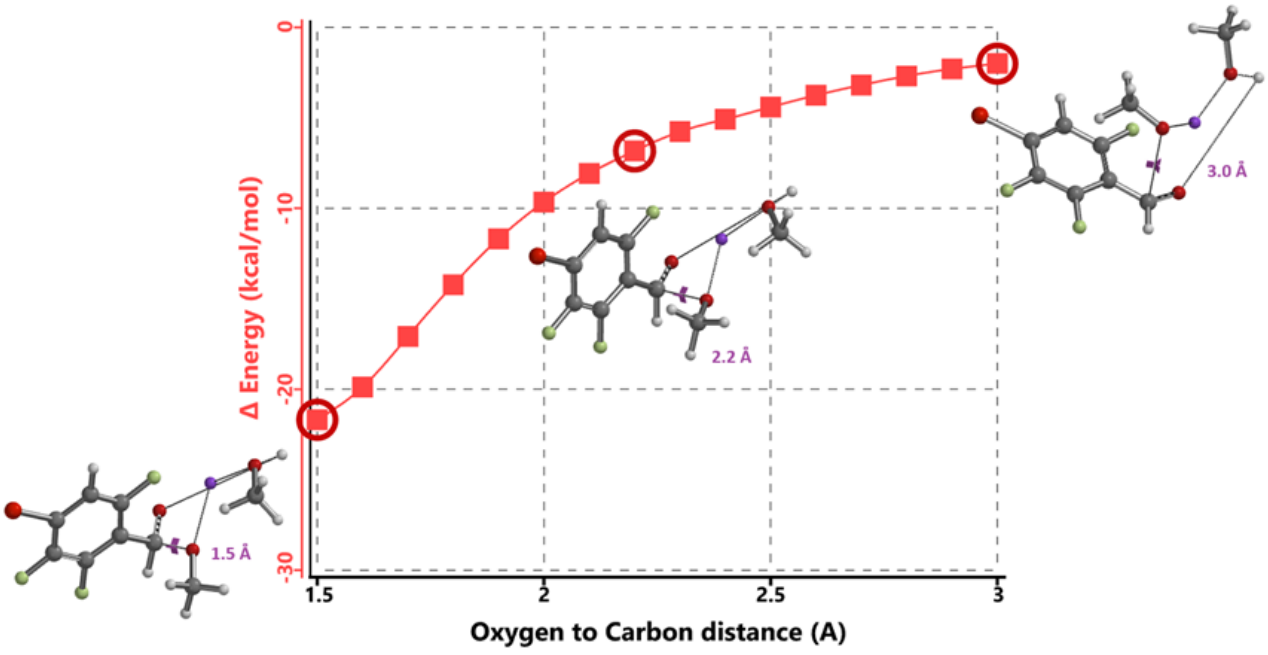
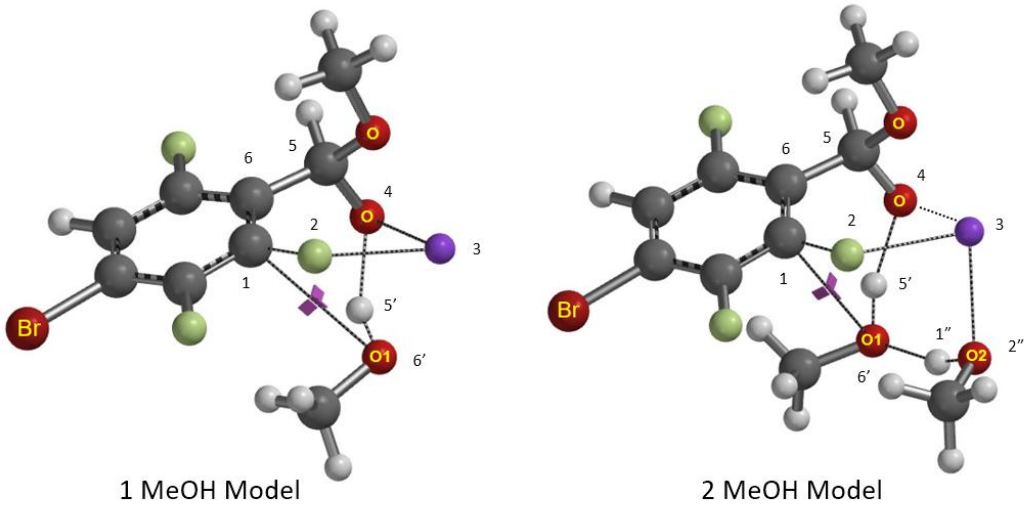


Figure 4. Reaction energy profile for addition of sodium methoxide to benzaldehyde **1**

Next, we calculated for reaction energy profile and transition state of the nucleophilic substitution. Based on structural characteristics of the hemiacetal salt, we evaluated two options, 1 MeOH and 2 MeOH models.



SnAr Displacement	Activation Energy (kcal/mol)	
	1 MeOH Model	2 MeOH Model
2-F	26.67	21.45
6-F	28.07	23.64

Figure 5. Activation energies of 1 and 2 MeOH models for S_NAr of benzaldehyde-methoxide adduct

Activation energy for methoxide addition at C2 of benzaldehyde-methoxide adduct is 1.40 and 2.19 kcal/mol lower than attack at C6 for the 1 MeOH and 2 MeOH models, respectively (Figure 5), accounting for the selective C2 substitution. Reaction energy profile calculated with the 2 MeOH model exhibits unique gentle drop in energy between 2.0 to 1.5 Å (Figure 6, C2-F carbon to incoming methoxy oxygen)[2c] corresponding to an energy barrier of 21.45 kcal/mol, in better agreement with experimental conditions (65 °C, 12 hours).

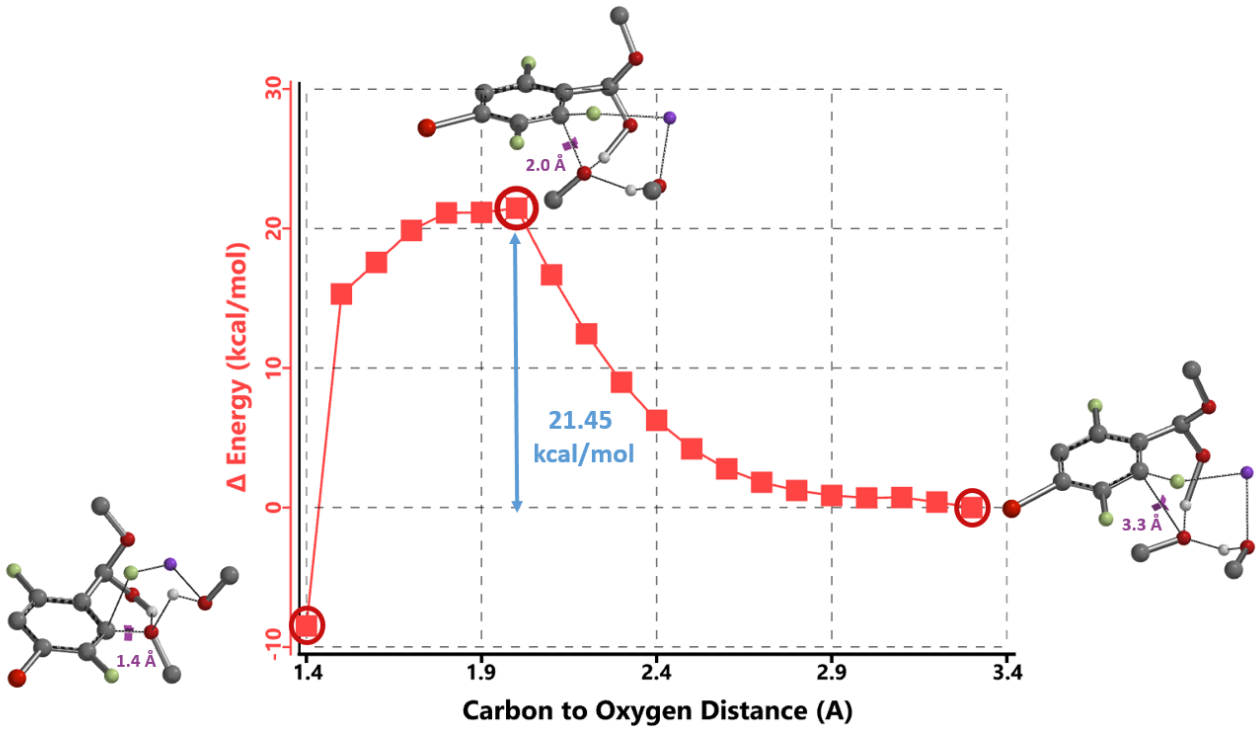


Figure 6. Reaction energy Profile calculated for S_NAr at C2 with a 2 MeOH model

Transition state (iFreq i350 cm^{-1}) and intrinsic reaction coordinate calculated (not shown) are supportive of the above sodium hemiacetal salt 2 MeOH reaction model. We assume upon aqueous work up, the hemiacetal on the S_NAr product hydrolyzes back to the aldehyde (Figure 7).

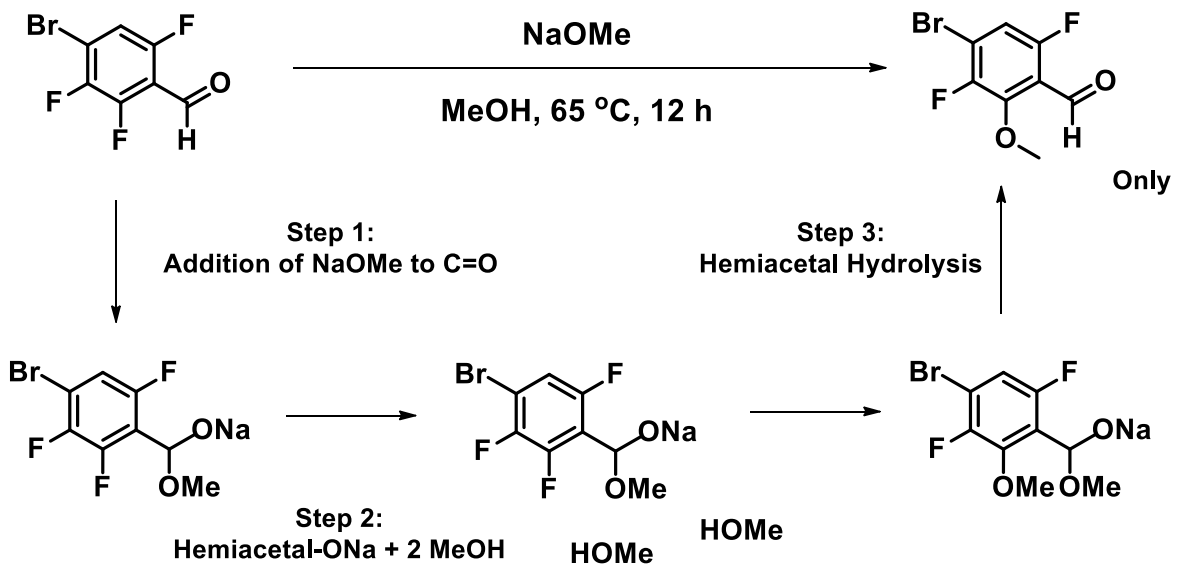


Figure 7. Mechanism for S_NAr reaction between 2,3,6-trifluoro-4-bromobenzaldehyde and sodium methoxide

Conclusion

QM calculations suggest that S_NAr reaction between 2,3,6-trifluoro-4-bromobenzaldehyde (**1**) and sodium methoxide proceeds *via* their corresponding hemiacetal product. Methanol solvent molecules are an integral part of the transition state. Both 1 MeOH and 2 MeOH reaction models could account for preferential formation of C2 displacement product **2b**. Calculated activation energy of 21.45 kcal/mol with the 2 MeOH model is in better agreement with experimental conditions.

Building on What We Just Learned

For the S_NAr reaction of 2-fluoropyridine analogs, electron withdrawing groups, such as CN and CHO, at C3 position, were anticipated to increase the electrophilicity of the substrates and accelerate the reaction (Figure 8). However, less than 5% conversions were observed, and most of the starting materials were recovered. On the other hand, S_NAr of the bromo analog proceeded smoothly with an isolated yield of 80%. How could we account for these? Below are overlays of LUMO and LUMO+1 of the pyridine analogs with their electron density maps.

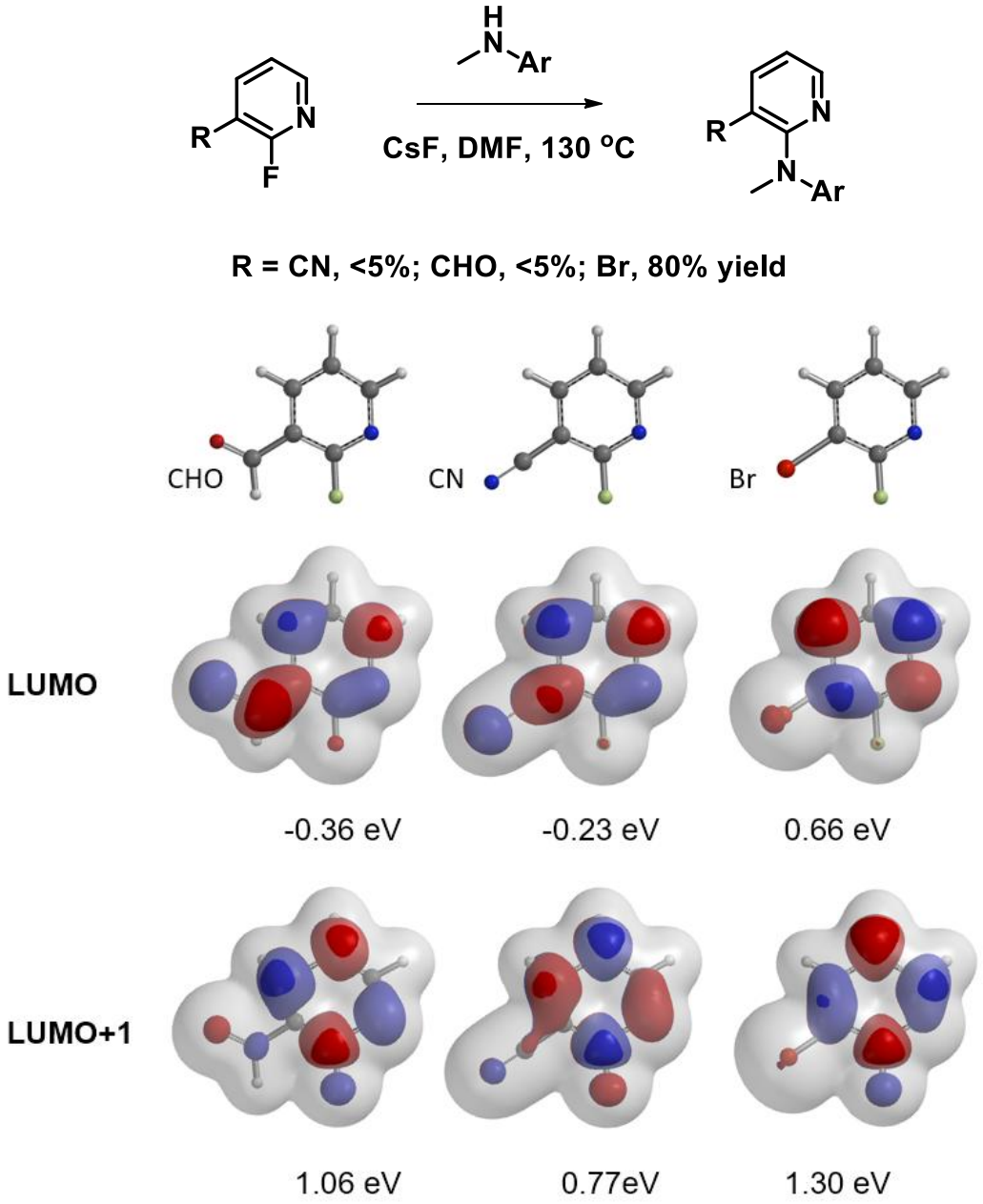


Figure 8. Reaction of 2-fluoropyridine with secondary amine (top);
Overlay of 2-fluoropyridine analogs LUMO and LUMO+1 with electron density map (bottom)

Return to Table of Contents

References:

[1] (a) QM Chapter 7: Application of LUMO Analysis in Nucleophilic Reactions (Part II). (b) QM Chapter 10: S_NAr Reaction of Polyhalogenated Heterocycles

[2] (a) S. Rohrbach, A.J. Smith, J.H. Pang, D.L. Poole, T. Tuttle, S. Chiba S, J.A. Murphy, *Angew. Chem. Int. Ed.*, **2019**, *58*, 16368. (b) O. Acevedo, W.L. Jorgensen, *Org. Lett.*, **2004**, *6*, 2881. (c) E.E. Kwan, Y. Zeng, H.A. Besser, E.N. Jacobsen, *Nature Chemistry*, **2018**, *10*, 917.

[3] *Spartan'20 Tutorial and User's Guide* (2020). Irvine, CA, USA: Wavefunction, Inc. pp362-368.

Chapter 36 Intriguing observations with S_NAr reaction of 4-nitrobenzonitrile

Yu Wang, Yongsheng Chen, John S. Wai

In chapters 10 and 35, we discussed the use of QM analyses to rationalize regioselectivity observed in nucleophilic aromatic substitution (S_NAr) of polyhalogenated substrates[1]. This type of reaction is usually mechanistically simple, operationally straightforward, and quite reproducible. Yet occasionally we do run into cases that are very sensitive to subtle variations, as exemplified by the reaction between 4-nitrobenzonitrile and sodium methoxide described by Ma in ChemRxiv[2] (Figure 1). He found that

1. Addition of 1 eq of MeONa gave no conversion of **1** to **2**;
2. Conversion improved with excess of MeONa; with 10 eq of MeONa, reaction did not go to completion, with~30% of **1** remained;
3. Addition of a small amount of methanol drastically improved conversion;
4. But conversion dropped sharply with more than 40 eq of methanol;
5. All reactions gave a clean mixture of **1** and **2**, with no other detectable byproducts.

In this chapter, we will use QM analyses to understand this interesting reaction, to account for the observations reported.

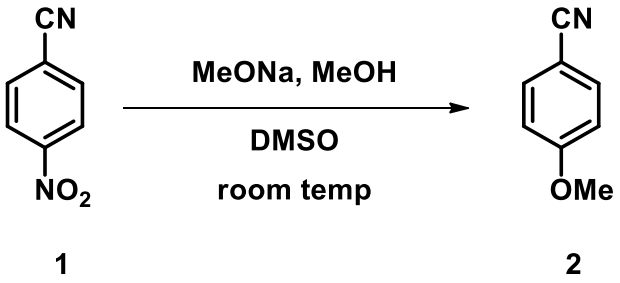


Figure 1. S_NAr reaction between 4-nitrobenzonitrile and sodium methoxide

Initial Calculations

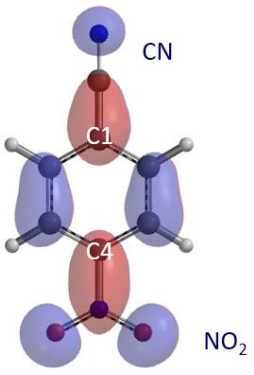


Figure 2. LUMO of 4-nitrobenzonitrile

Shown in Figure 2 is LUMO of **1** with significant lobes on C4 and C1, reaction centers for nucleophilic substitution. Reaction energy profile calculated for nucleophilic attack at C4 shows i) As methoxide approaches C4 carbon from 3.4 Å to 2.5 Å, energy of the reacting system drops by 6.06 kcal/mol; ii) The nucleophilic addition step has a relatively low energy barrier of 2.59 kcal/mol, suggesting that the reaction will proceed smoothly and thoroughly with one equivalent of sodium methoxide, which is **inconsistent** with observation 1.

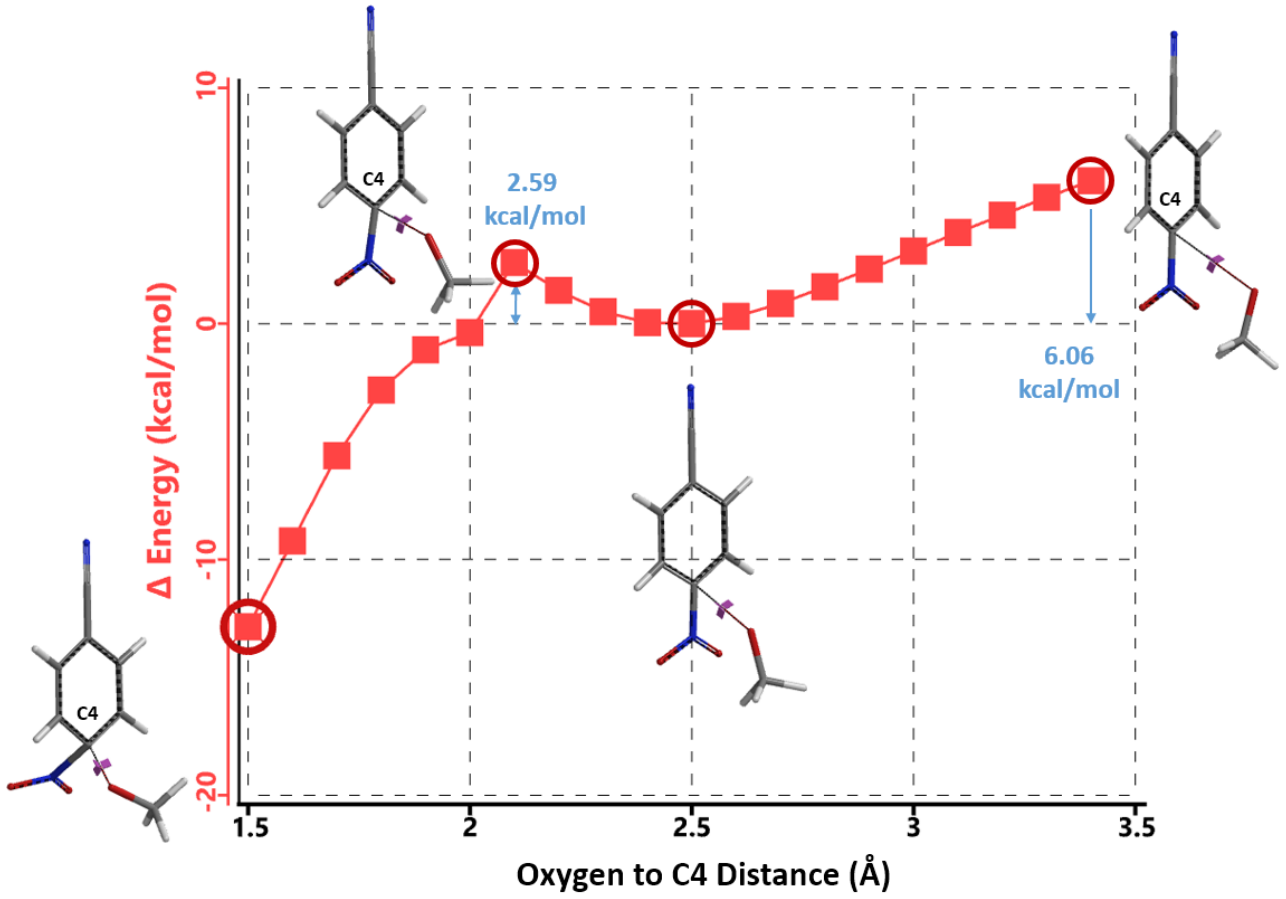


Figure 3. Reaction energy profile calculated for nucleophilic attack at C4

We then calculated for reaction energy profile for nucleophilic attack at C1 (Figure 4). As methoxide approaches C1 (3.4 Å to 2.4 Å), it gets close enough to the nitrile carbon (3.25 Å to 1.48 Å), to be suggestive of formation of a covalent bond. Furthermore, the bond angle between cyanide group and C1 carbon also changes from 176.5° to 127.2°, indicating that the carbon-nitrogen triple bond of cyanide group is converting to the carbon-nitrogen double bond of an imidate. These changes are associated with an energy drop of 18.87 kcal/mol. Subsequent conversion from the imidate to 4-methoxybenzonitrile **2** (2.4 Å to 1.5 Å) will face a very high energy barrier of 35.93 kcal/mol, not consistent with a room temperature reaction.

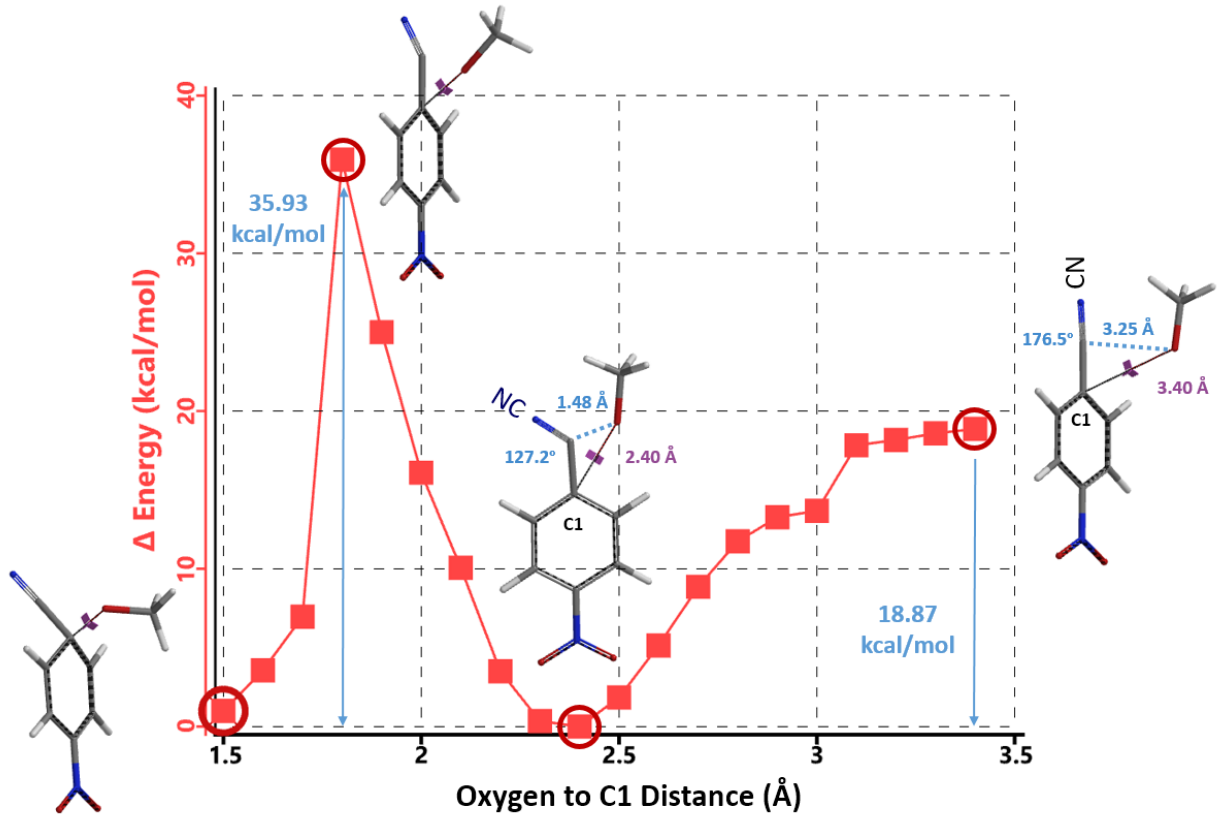


Figure 4. Reaction energy profile calculated for nucleophilic attack at C1

Based on the above calculation results, we reasoned that the imidate is likely to be an intermediate in this reaction. Direct reaction energy profile calculation for its formation suggests an even lower energy barrier of 0.78 kcal/mol (Figure 5) *versus* 2.59 kcal/mol for reaction at C4.

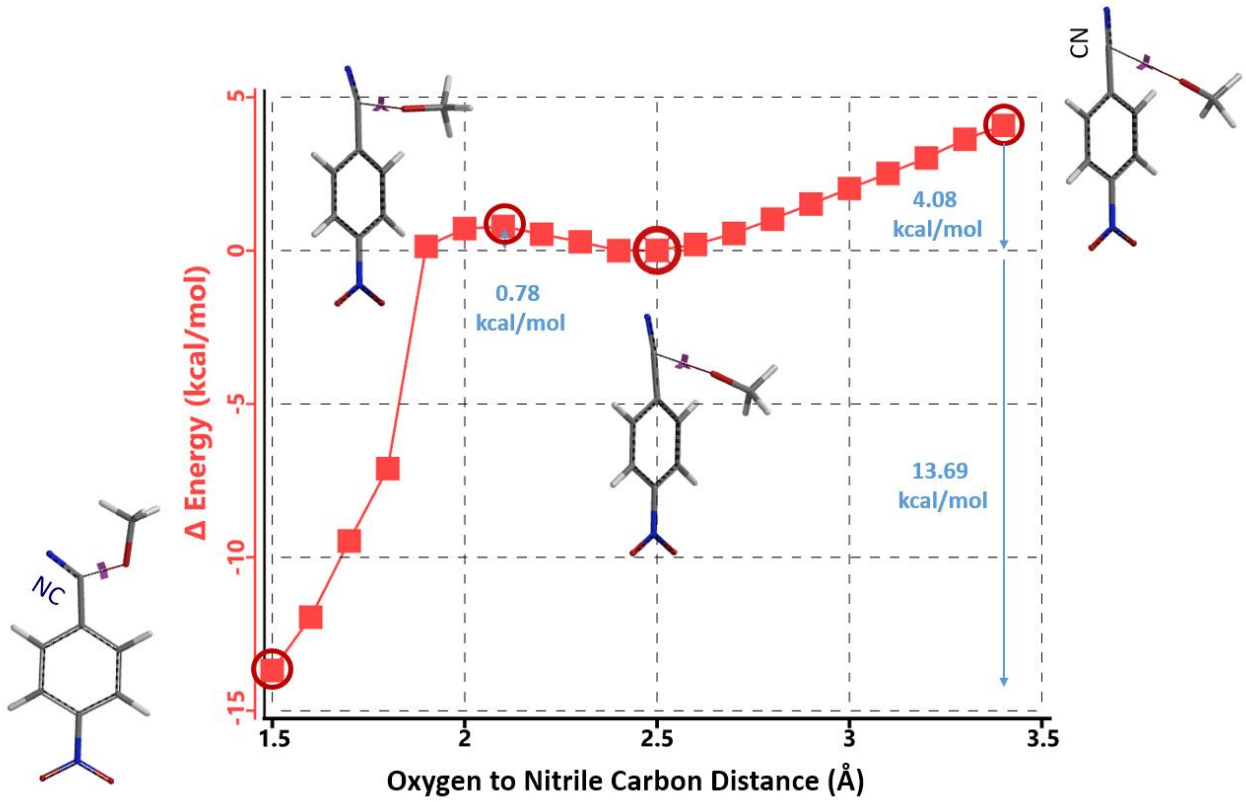


Figure 5. Reaction energy profile calculated for nucleophilic attack at nitrile group

This initial set of calculations enables us to conclude that the first equivalent of sodium methoxide will add on to the nitrile group of **1** to form intermediate **3**, instead of displacing the nitro group to provide **2**. **This is consistent with observation 1, that is, when only one equivalent of sodium methoxide is added, no conversion of 1 to 2 could be detected.**

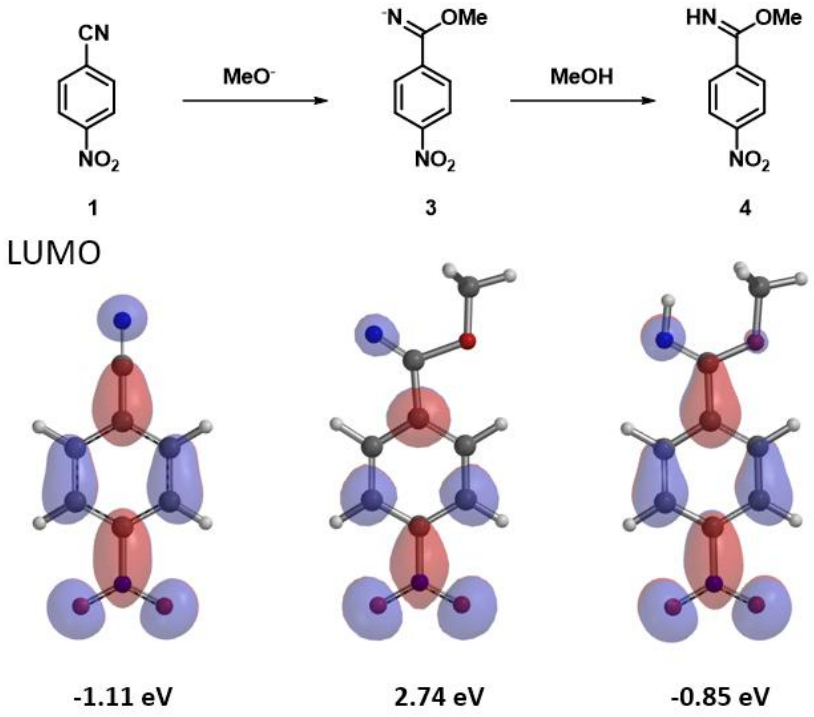


Figure 6. LUMO of 4-nitrobenzonitrile and corresponding imidates

LUMO energy calculated for the resultant imidate anion **3** is 2.74 eV, indicating that the electrophilicity of this intermediate is low (Figure 6). We reasoned that when a small amount of methanol is added, proton exchange will occur to provide the neutral imidate **4**, with LUMO energy of -0.85 eV, electrophilic enough for further nucleophilic reaction. **This could account for observation 3, that is, addition of a small amount of methanol drastically improved conversion.**

Second Nucleophilic Addition

Next we compared reaction energy profiles calculated for methoxide nucleophilic displacement of nitro group at C4 (Figure 7) *versus* addition to imidate group at C1 (Figure 8) with intermediate **4**. Activation energies were found to be 4.54 kcal/mol *versus* 3.26 kcal/mol, respectively, with a difference of 1.28 kcal/mol, in favor of the addition to imidate group.

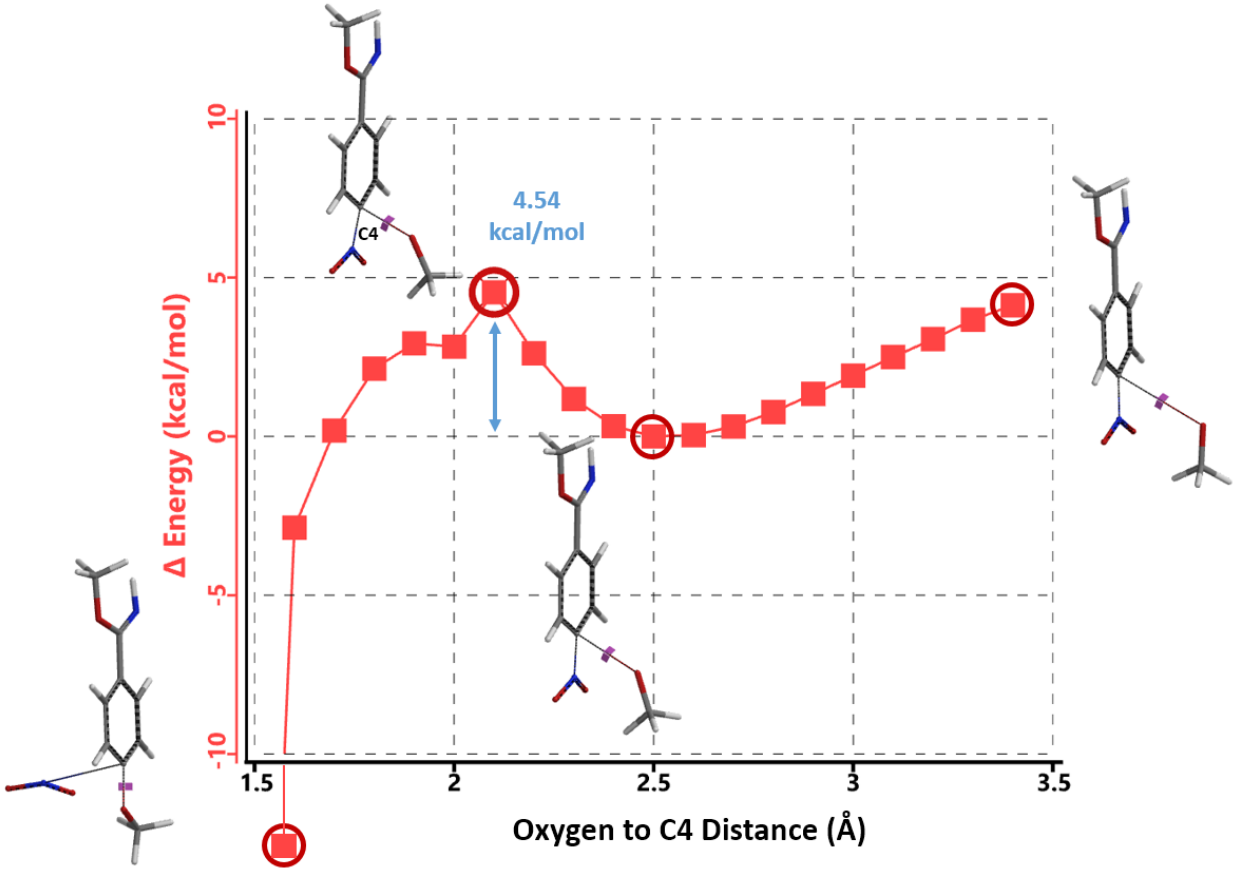


Figure 7. Reaction energy profile calculated for nucleophilic attack at C4

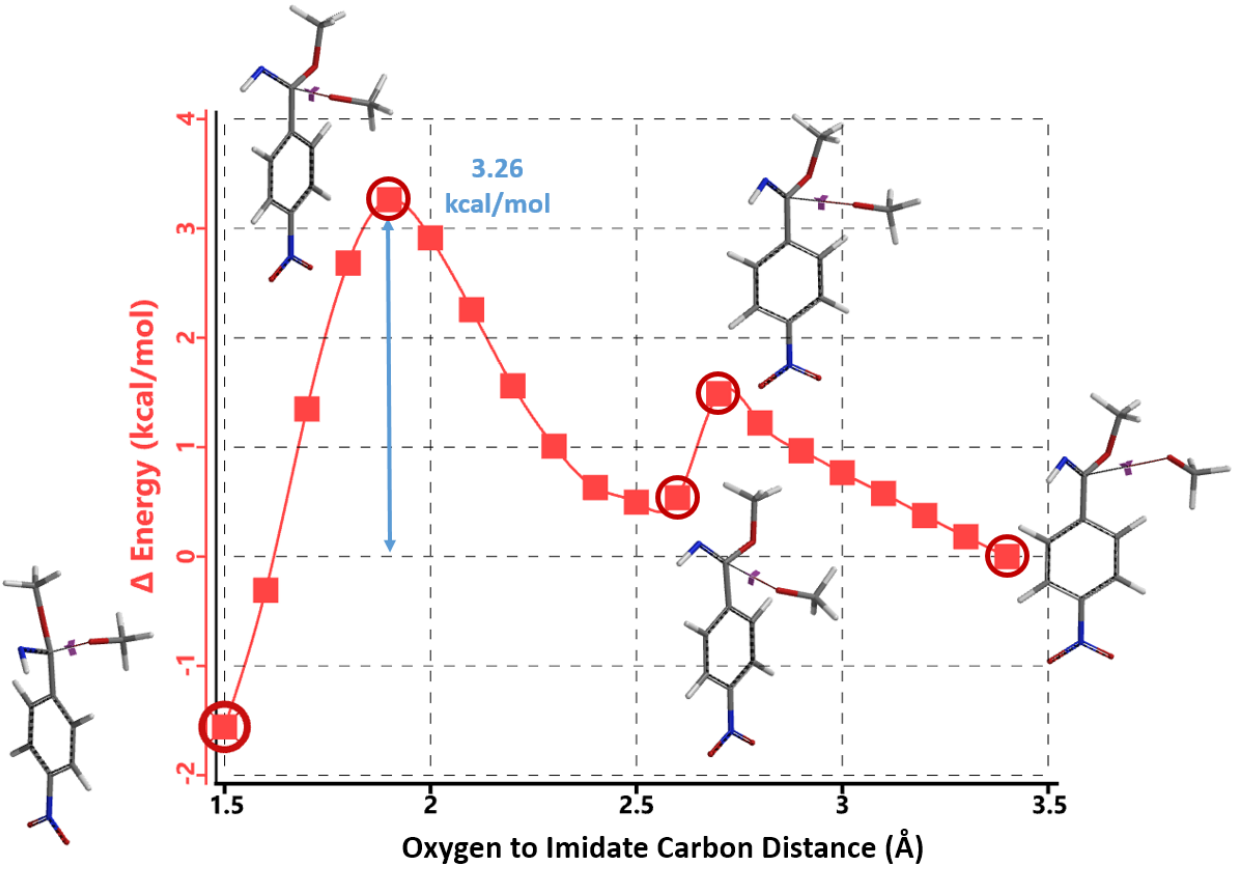


Figure 8. Reaction energy profile calculated for addition to imidate group

Similar to the conversion of intermediate **3** to **4**, we reasoned that the resultant intermediate **5** will undergo proton exchange with methanol to form the neutral intermediate **6** (Figure 9), with a LUMO energy of -0.67 eV, electrophilic enough for reaction with the third equivalent of methoxide, for displacement of the nitro group.

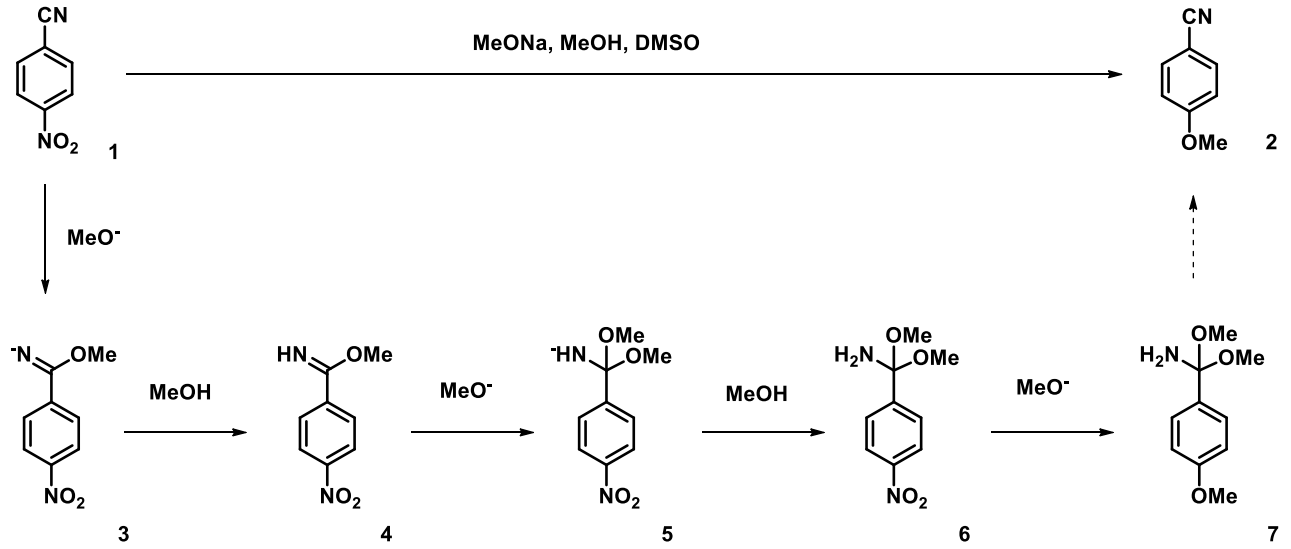


Figure 9. Proposed reaction mechanism with multiple nucleophilic reactions

Third Nucleophilic Addition

With intermediate **6**, we have only one electrophilic center to consider. Remarkably, the reaction energy profile calculated for methoxide displacement of the C4 nitro group exhibits two energy peaks (Figure 10), with oxygen to C4 distance of 2.1 Å and 1.8 Å, respectively. Follow-up calculation for the 2.4 Å to 1.6 Å portion of the energy profile at higher resolution, with a step size of 0.05 Å (Figure 11), shows clearly the presence of two energy peaks, one at 2.15 Å and one at 1.85 Å.

Transition state calculation of both structures revealed that only the structure at 1.85 Å has an imaginary frequency of i331 cm^{-1} which corresponds to the bonds being made and broken, satisfying key computational chemistry criteria as transition state (Figure 12). And the result provides an estimated activation energy of 5.59 kcal/mol for this rate limiting step, consistent with a room temperature reaction[3].

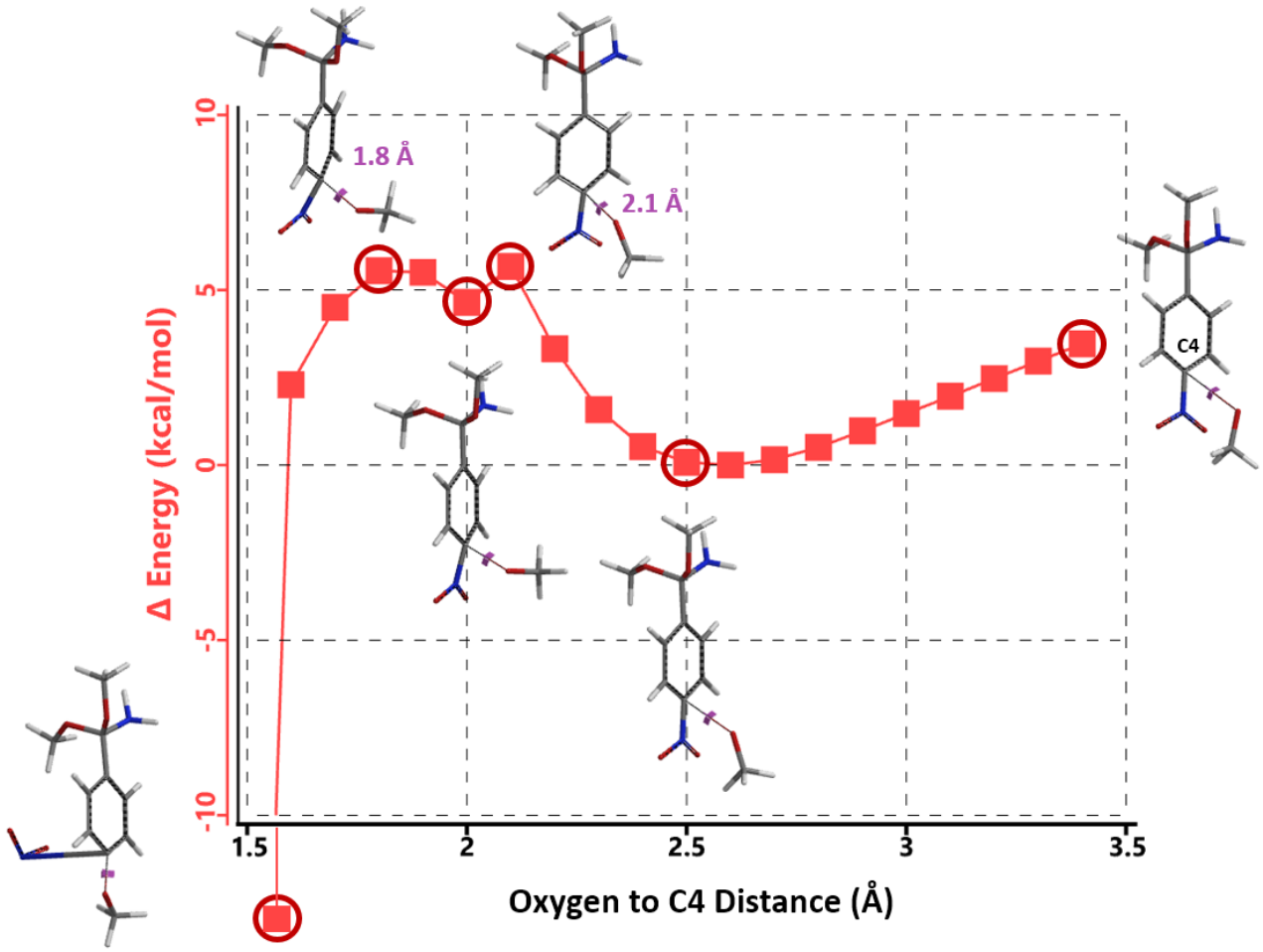


Figure 10. Reaction energy profile calculated for nucleophilic attack of intermediate **6** at C4

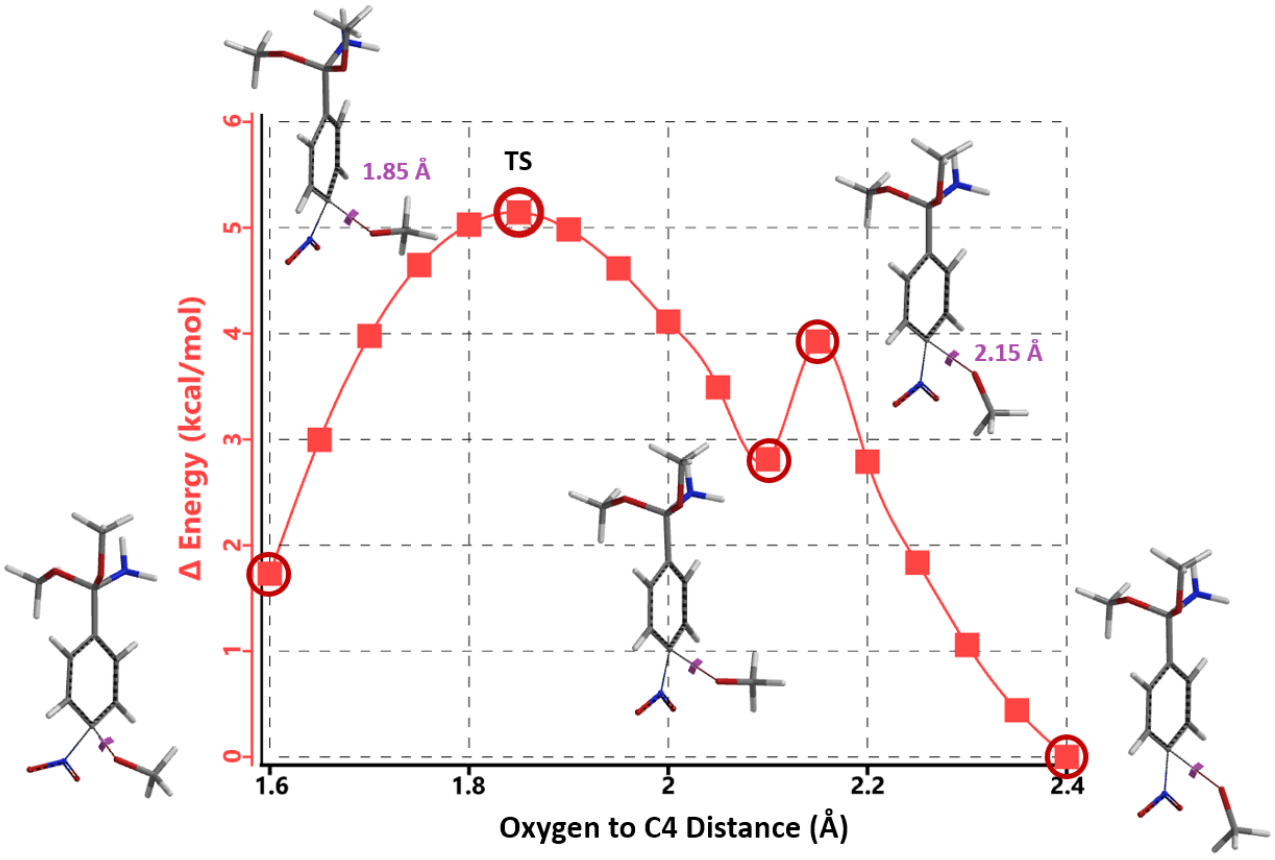


Figure 11. Reaction energy profile between 2.4 Å and 1.6 Å with a step size of 0.05 Å

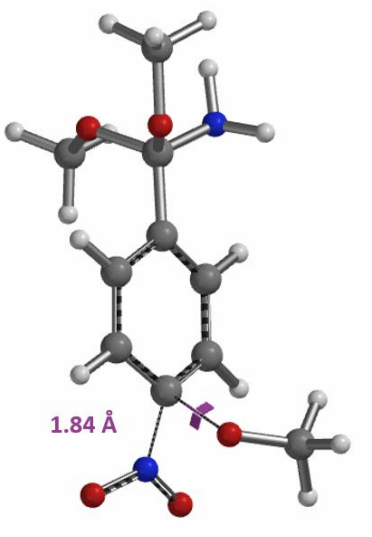


Figure 12. Transition State for third methoxide addition with an imaginary frequency of i331 cm^{-1}

Conclusion

Results from our QM analyses are supportive of a reaction mechanism with three consecutive additions of methoxide, first to 4-nitrobenzonitrile **1**, then to the resultant imidate intermediate **4**, and finally to the intermediate **6**, with twice proton exchanges with methanol in between to restore electrophilicity (Figure 13). The proposed mechanism could account for the intriguing observations reported. Since the difference in activation energy for formation of **5** *versus* **8** in the second methoxide addition step is relatively small ($\Delta\Delta E$ of 1.28 kcal/mol), reaction could go through **8** as a very minor pathway.

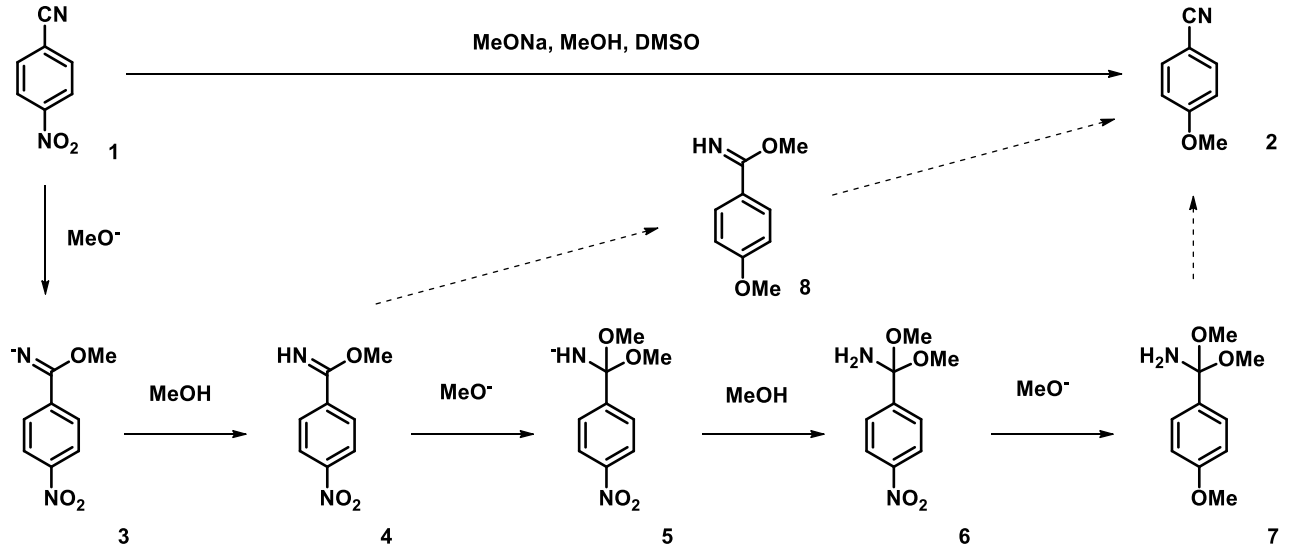


Figure 13. Proposed reaction mechanism for SnAr reaction of 4-nitrobenzonitrile with methoxide

Nucleophilic aromatic substitution (S_NAr) is usually mechanistically simple, operationally straightforward, and quite reproducible. Yet there are examples which are very sensitive to subtle variations, as exemplified by the above reaction between 4-nitrobenzonitrile with

methoxide. Detailed and accurate records are crucial for understanding and optimizing these reactions. Precise execution of experimental conditions as recorded is very important when repeating such reactions!

Building on What We Just Learned

For transformation of intermediate **7** to 4-methoxybenzonitrile (**2**), what are likely mechanisms? This happens during reaction with the excess sodium methoxide, or during aqueous work-up[4]?

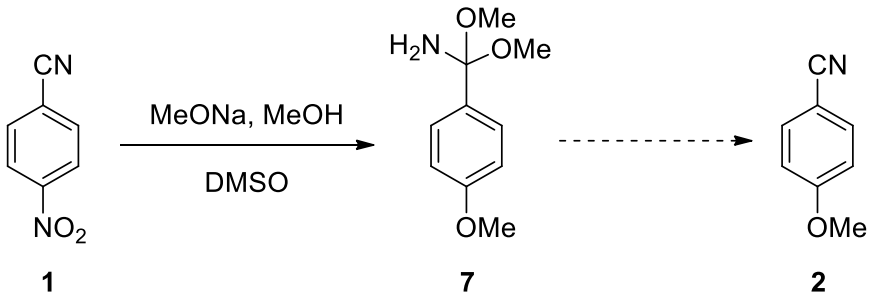


Figure 14. The last step in the reaction mechanism for conversion of **1** to **2**

Return to Table of Contents

References:

[1] (a) QM chapter 10: S_NAr reaction of polyhalogenated heterocycles. (b) QM chapter 35: Highly regioselective SnAr of a polyhalogenated benzaldehyde.

[2] Ma Y. Solvent Molecules Play a Role in an SNAr Reaction. ChemRxiv. Cambridge: Cambridge Open Engage; 2019; 10.26434/chemrxiv. 7782797. V1

[3] *Spartan'20 Tutorial and User's Guide* (2020). Irvine, CA, USA: Wavefunction, Inc. pp. 158, 442, 459 & 536 and Chapters 22 and 24 of the QM Organic Chemistry Class on the calculation of transition states and imaginary frequencies.

[4] Additional information to consider: 1) work-up: the reaction mixture was quenched with aqueous $NaHCO_3$ and extracted with dichloromethane; 2) benzimidate reacts with water to form benzamide; 3) clean conversion of 4-nitrobenzonitrile to 4-methoxybenzonitrile, with no detectable byproducts.

Chapter 37 Is S_N2 substitution at sp^2 reaction centers feasible?

Liting Dong, Tommy Lai, Yongsheng Chen, John S. Wai

Alkyl and vinyl halides are important reagents and intermediates in organic synthesis. Shown in figure 1 are LUMO and LUMO+1 of ethyl and vinyl chlorides. It is noteworthy that LUMO of ethyl chloride is very similar to LUMO+1 of vinyl chloride, suggesting that they could react in similar manners. This is consistent with the observations that both of them react with magnesium to form the corresponding Grignard reagents, and undergo dehydrohalogenation to generate ethylene or acetylene. With ethyl chloride, interaction of the LUMO lobe on the C-Cl carbon with HOMO of incoming nucleophile leads to nucleophilic substitution reaction at the sp^3 carbon with inversion of configuration (S_N2 reaction). Intuitively, this leads to the question, can vinyl chloride undergo a similar in-plane S_N2 reaction?

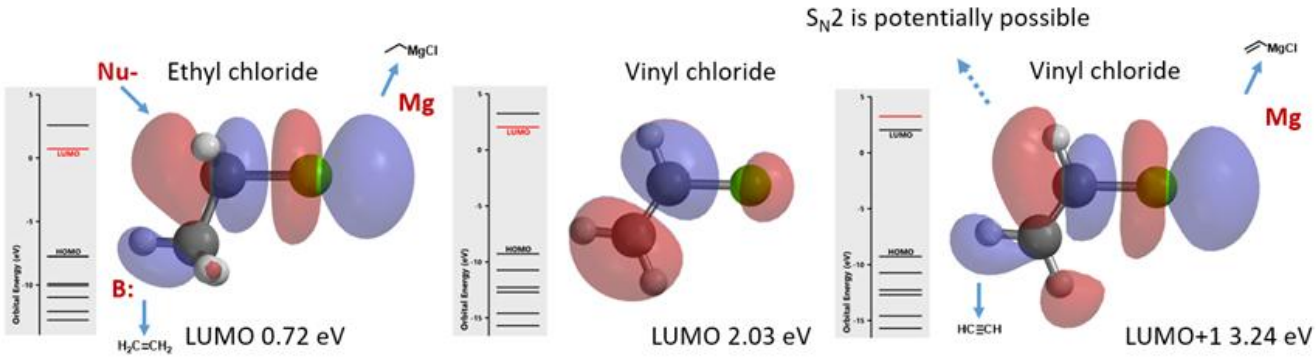


Figure 1. LUMO of ethyl chloride and LUMO, LUMO+1 of vinyl chloride

S_N2 reaction of sp^2 carbon atom ($S_NV\sigma$)

In 1994, Pross and Radom *et al.* predicted by *ab initio* calculations that in-plane S_N2 reaction with configuration inversion of unactivated vinyl chloride is possible in both gas phase and solution[1]. In 2004, Ando and Narasaka *et al.* demonstrated the first in-plane S_N2 reaction of vinyl chloride *E*-**1** with DFT calculations, then confirmed it experimentally (Figure 2)[2]. Substrate is configuration specific for the reaction. When *E*-**1** is treated with NaH in DMF solvent, the reaction proceeds at room temperature to provide benzofuran **2** in 95 % isolated yield. On the other hand, the isomeric *Z*-**1** was recovered quantitatively after heating at 110 °C.

Transition state and intrinsic reaction coordinates (IRC) calculations

For calculation of the transition state (Figure 2), Ando and Narasaka *et al.* used the B3LYP/6-31+G* (Onsager continuum, DMF ε=37) model. Since the exact coordinate data was not reported, we repeated the calculation with ωB97X-D DFT. The calculated transition state has

a carbon-oxygen distance of 2.16 Å, close to the value reported, and an imaginary frequency of i556 cm^{-1}, corresponding to the bonds being made and broken (Figure 3).

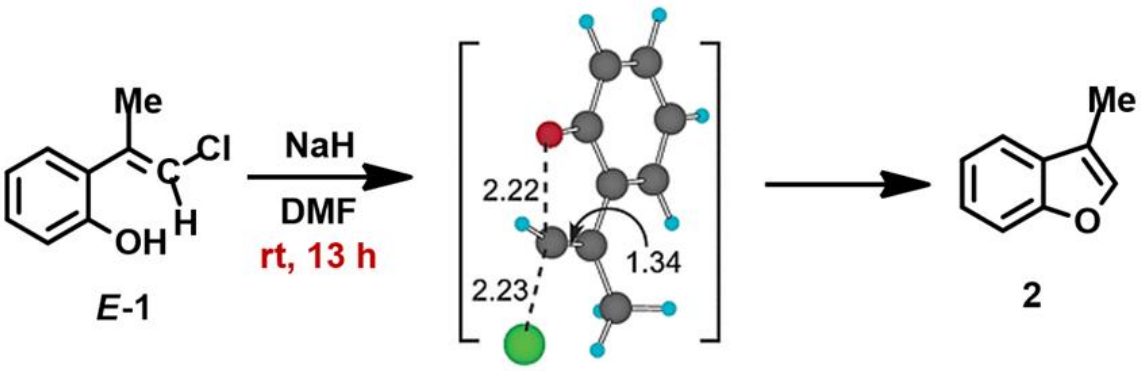


Figure 2. *E*-**1** Intramolecular S_N2 reaction and reported transition state structure

Figure 3. Transition state structure of *E*-**1** S_N2 reaction with an imaginary frequency of i556 cm^{-1}

With the coordinate of transition state, we calculated for the intrinsic reaction coordinate (IRC)[3] of the reaction (Figure 4). This shows that as the oxyanion approaches the sp^2 carbon, the vinylic hydrogen gradually flips from the inner side to the outer side, in the plane defined by Cl, C, O atoms, while the chloro group is being displaced. The carbon-chlorine bond breaks and the carbon-oxygen bond forms in a concerted manner, analogous to the umbrella-like inversion at sp^3 carbon in S_N2 reaction.

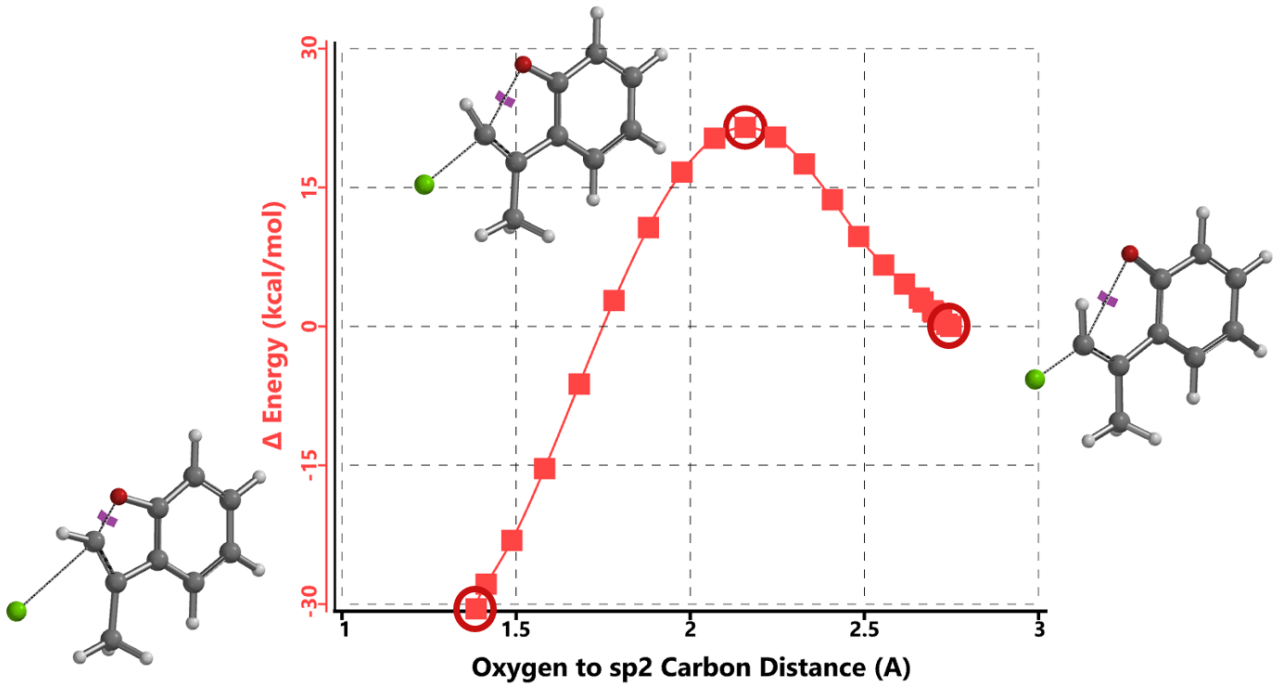


Figure 4. Intrinsic reaction coordinate of *E*-**1** sp^2 S_N2 reaction

Shown in Figure 5 are changes in electron density contour across the plane of the molecule as reaction proceeds. Electron density between the oxyanion and the vinyl hydrogen gradually increases toward the transition state, indicative of hydrogen bond interaction between them. When the distance between the oxygen and sp^2 carbon is approximately 2.4 Å, the hydrogen atom begins to flip to the other side. The electron density gradually shifts from between oxyanion and vinyl hydrogen to between oxygen and sp^2 carbon. These account for the remarkable sp^2 S_N2 reactivity observed with the *E*-**1** isomer.

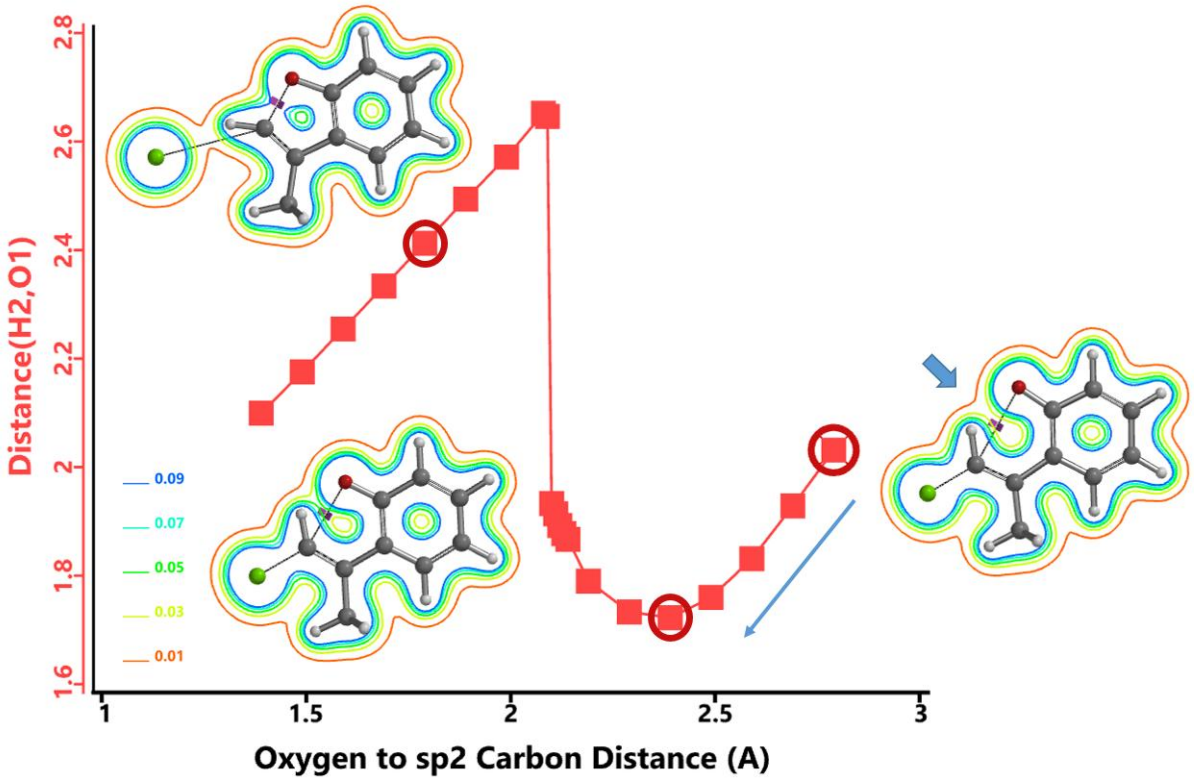
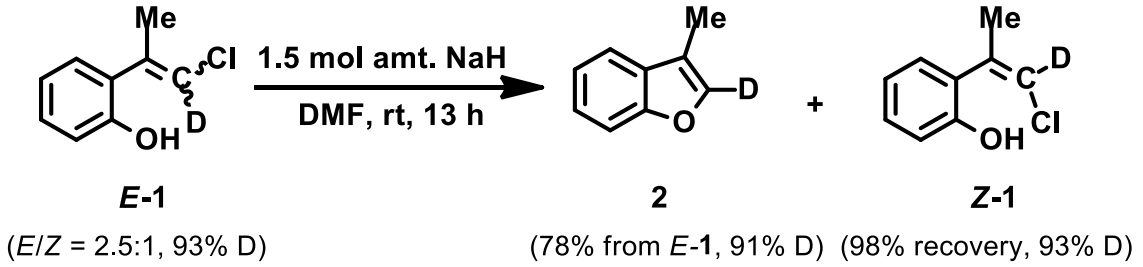


Figure 5. Changes in electron density contour as *E*-**1** sp^2 S_N2 reaction proceeds

Deuterium isotope experiment



***E*-1** (*E*/*Z* = 2.5:1, 93% D) **2** (78% from *E*-**1**, 91% D) **Z-1** (98% recovery, 93% D)

Figure 6. Reaction with a mixture of deuterium labeled *E*-**1** and *Z*-**1**

When a mixture of *E*-**1** and *Z*-**1** deuterium labeled substrate was treated with NaH in DMF at room temperature, *E*-**1** was consumed, benzofuran **2** with deuterium label at C-2 was isolated, while *Z*-**1** was almost quantitatively recovered (Figure 6). These results provide further support for the sp^2 S_N2 mechanism discovered with DFT calculations, and exclude possibility of the alternative allylic isomerization and carbene insertion mechanisms (Figure 7).[4]

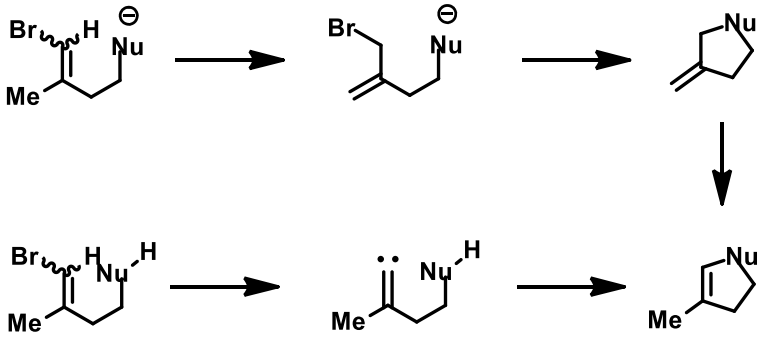
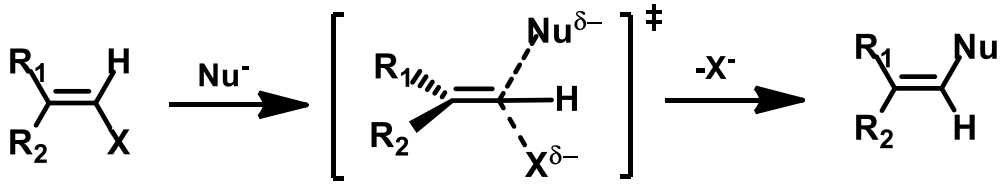


Figure 7. Excluded possible allylic isomerization and carbene insertion mechanism

Further examples of in plane sp^2 S_N2 reactions were reported, along with computational studies.[4]

Substrate scope

Figure 8 summarizes substrate scope for sp^2 S_N2 reaction with carbon reaction center:
1. Nucleophile could be oxygen, nitrogen, carbon, or sulfur based[4]
2. Leaving group X could be Cl, Br



Nu^- = TsN(-)R, NC(RSO_2)C(-)R, RS(-), etc.

Figure 8. Substrate scope for sp^2 S_N2 reaction with carbon reaction center

Is S_N2 reaction of sp^2 nitrogen atom feasible as well?

In 2020, Chen *et al.* reported an efficient method for the synthesis of 2-aryl-1,2,3-triazoles *via* intramolecular nitrogen-nitrogen bond formation (Figure 9)[5]. Selective activation of dimethyl(phenylhydrazono)ethylidenehydrazines with methyl iodide provided hydrazonium intermediates **3**, which readily cyclized in the presence of base to provide a variety of 2-aryl-1,2,3-triazoles **5** in good to excellent yields. The authors suggest that the reaction proceeds *via* **4b**, with an intramolecular S_N2 on the sp^2 nitrogen, displacement of trimethylamine, and formation of a new *N–N* bond. This is substantiated with DFT computational studies.

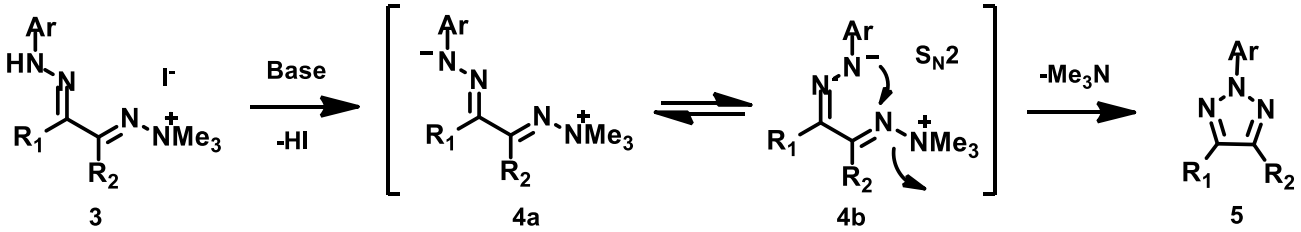


Figure 9. Proposed intramolecular S_N2 reaction mechanism for synthesis of 2-aryl-1,2,3-triazoles

Intrigued by their findings, we calculated the transition state (Figure 10, iFreq i458 cm^{-1}) and then intrinsic reaction coordinate (IRC) of the reaction (Figure 11). The latter illustrated in detail the in-plane sp^2 nitrogen S_N2 concerted process, with relatively lower energy barrier than the sp^2 vinyl halide S_N2 reaction described above, presumably due to the absence of the vinylic hydrogen. This reaction is an excellent application of the sp^2 S_N2 mechanism. Arylation of triazole with Chan-Lam reaction will provide a mixture of *N*-1 and *N*-2 arylation products in low yield[6].

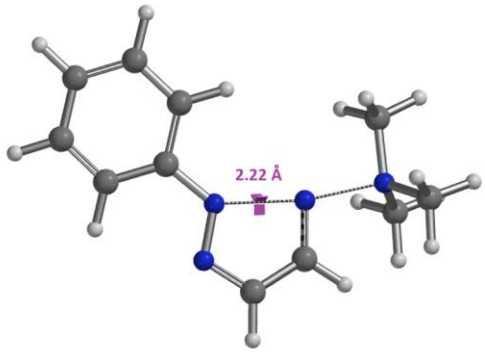


Figure 10. Transition state, with imaginary frequency of i458 cm^{-1}, in the formation of 2-aryl-1,2,3-triazole

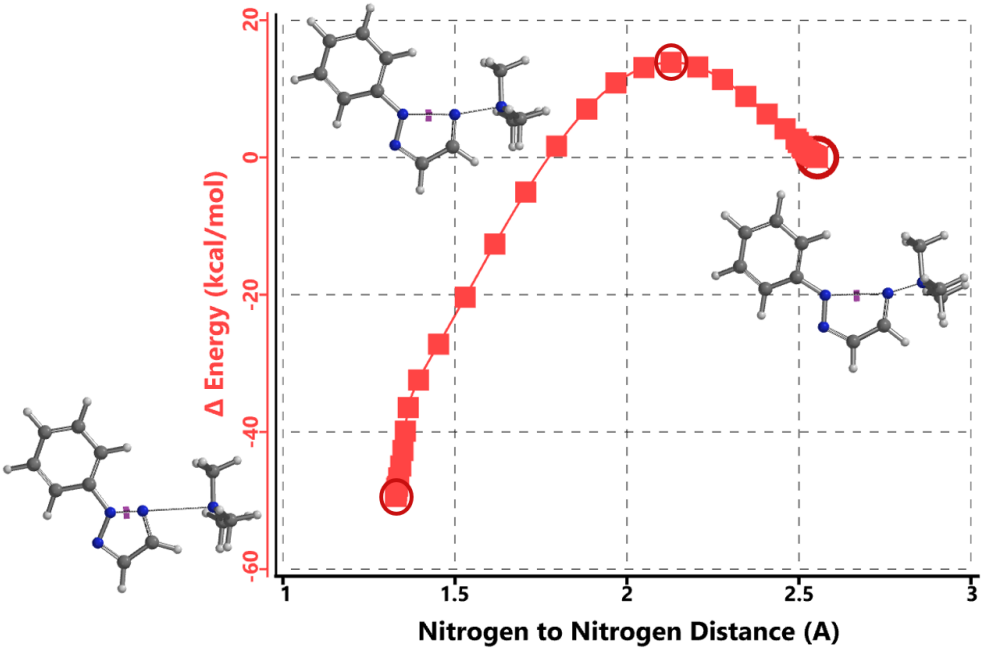


Figure 11. Intrinsic reaction coordinate for formation of 2-aryl-1,2,3-triazole

Conclusion

In-plane sp^2 S_N2 reactions are considered unlikely on the basis of the perception that S_N2 attack with inversion at vinylic centers is a high-energy process[1,7]. Quantum mechanics give organic chemists data and insight to objectively reassess the feasibility of such reactions [1, 2]. In this chapter we learned:

1) S_N2 reactions at sp^2 carbon and nitrogen ($S_NV\sigma$) can occur and are predictable

2) Accessibility of the sp^2 reaction center to nucleophilic attack determines feasibility of such sp^2 S_N2 reaction

3) Potential application/occurrence of these sp^2 S_N2 reactions in retrosynthetic planning[4,5]

4) Quantitative QM analysis can provide excellent guidance on their feasibilities.

Building on What We Just Learned

Based on the LUMO, LUMO/Electron Density overlay of chlorobenzene (Figure 12), a nucleophilic substitution reaction of it *via* a sp^2 S_N2 mechanism will be sterically impossible, would you agree?

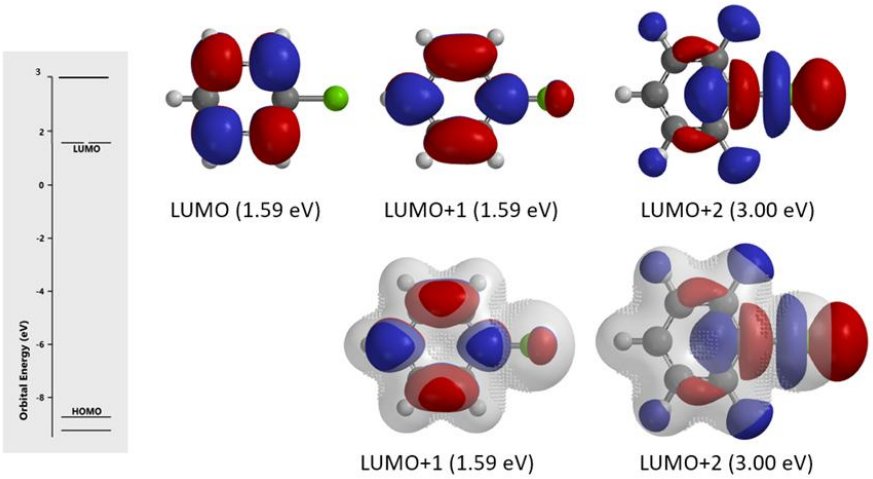


Figure 12. LUMO and LUMO/Electron Density overlay (with inaccessibility markers) of chlorobenzene

Return to Table of Contents

References:

[1] M. N. Glukhovtsev, A. Pross, L. Radom *J. Am. Chem. Soc.,* **1994**, *116*, 5961. In-plane S_N2 was considered unlikely on the basis of the perception that S_N2 attack with inversion at vinylic centers is a high-energy pathway than out-of-plane π-attack. Standard *ab initio* molecular orbital calculations showed that in-plane σ-type S_N2 substitution for Cl^- + CH_2=CHCl with inversion at unactivated sp^2 carbon (136.5 kJ/mol), is actually energetically preferred to the out-of-plane π-pathway (178.9 kJ/mol) by 42.4 kJ/mol.

[2] K. Ando, M. Kitamura, K. Miura, K. Narasaka *Org. Lett.*, **2004**, *6*, 2461.

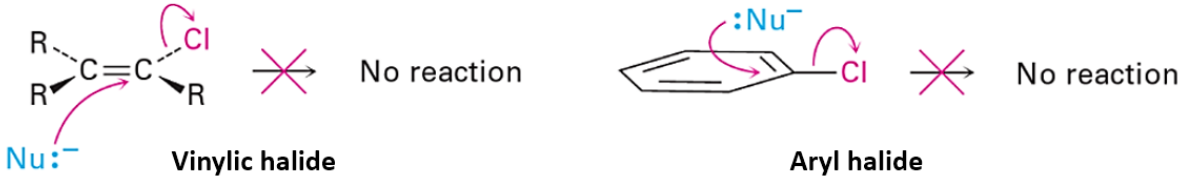
[3] QM Organic Chemistry Classroom Chapter 28 "A QM Study of the Hydroaminomethylation of Olefins".

[4] H. Miyauchi, S. Chiba, K. Fukamizu, K. Andob, K. Narasaka, *Tetrahedron,* **2007**, 63, 5940.

[5] C.Y. Chen, X. Lu, M. C. Holland, S. Lv, X. Ji, W. Liu, J. Liu, D. Depre, P. Westerduin, *Eur. J. Org. Chem.,* **2020**, 548.

[6] QM Organic Chemistry Classroom Chapter 34 "QM Analyses of Regioselectivity in Chan-Lam Reaction".

[7] J. McMurry, *Organic Chemistry*, Eighth Edition, Boston, MA, USA: Cengage Learning, 2011; pp 379-380. "Vinylic halides and aryl halides are….unreactive towards S_N2 displacement. This lack of reactivity is due to steric factors: the incoming nucleophile would have to approach in the plane of the carbon-carbon double bond and burrow through part of the molecule to carry out a backside displacement."



Chapter 38 Frequently Asked Questions

Yongsheng Chen, John S. Wai

Over the years, we received a lot of questions from our readers, below are the frequently asked ones.

1. What software to use for QM calculations?

Gaussian, Gamess, and ORCA are popular ones for organic chemists. WuXi AppTec provides us with Spartan, which uses Q-Chem as the calculation engine.

Figure 1. Highly cited QM calculation engines in the scientific literature[1]

2. Which DFT and basis set do you use as standard method?

Currently we use Density Functional Theory ωB97X-D and basis set 6-31G* as our standard method. It provides a balance of speed and accuracy our organic chemists need.

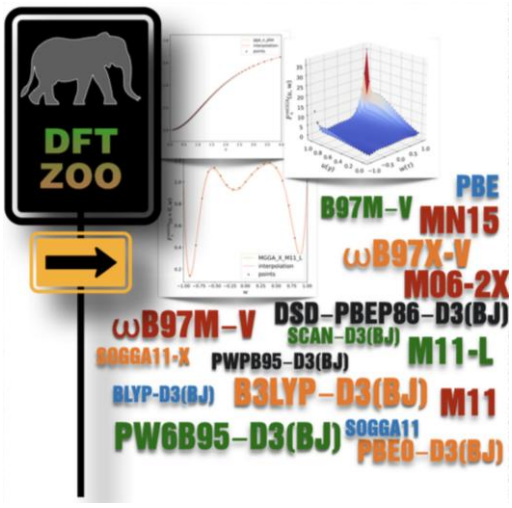


Figure 2. Too many DFT choices[2]

3. Any book will you recommend learning practical QM?

We searched for practical QM textbooks which could be useful for organic chemists. There is none. As such, our colleagues felt the need to share what they learned in articles published in https://rcs.wuxiapptec.com/qm/. They are collated into eBook you could download for free from the same links.

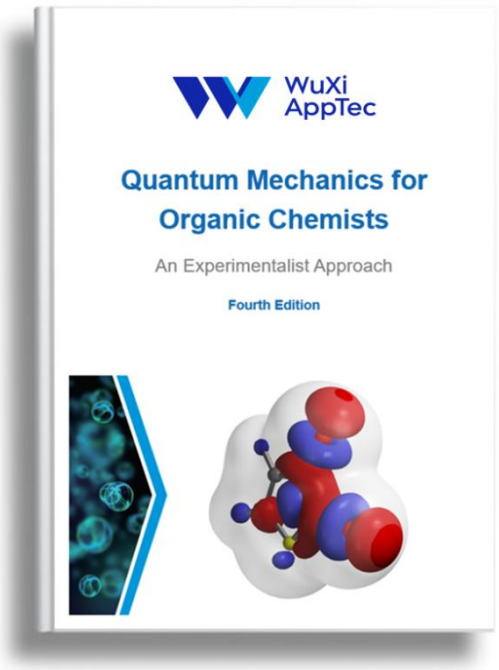


Figure 3. Our practical QM book for organic chemists

4. How do WuXi AppTec chemists learn QM?

We started from retrospective analyses of reactions we run every day, especially the challenging ones we couldn't predict properly intuitively. We optimized our methods, developed SOP, and turned QM into a prospective tool. We discussed and taught one another what we learned. It is a continuous team effort.

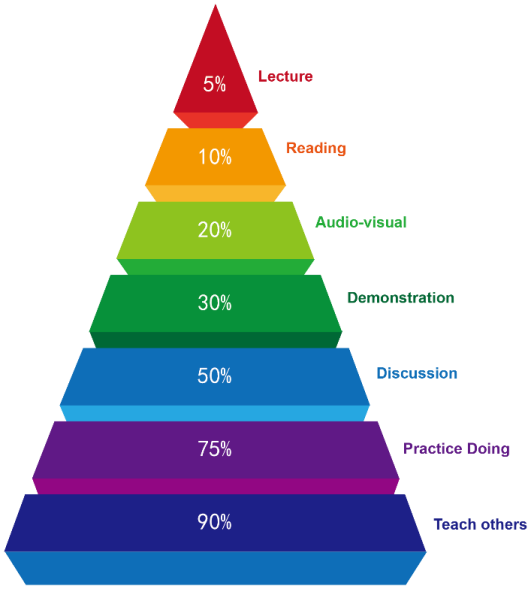


Figure 4. Learning Pyramid

5. What is the difference between QM and ML for reactivity predictions?

QM focuses on physical reality, could pin point uniqueness of each substrate, each reagent. ML focuses on data, statistical average. Currently QM offers better guidance when it is not intuitive easy to make correct prospective analyses and is more useful in retrospective analyses for divergence in reactivity. As QM parameters are incorporated into Machine Learning, we anticipate chemistry ML tools to continue to improve.

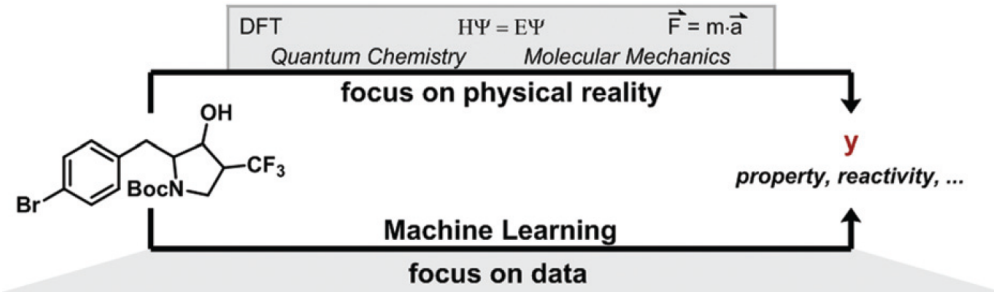
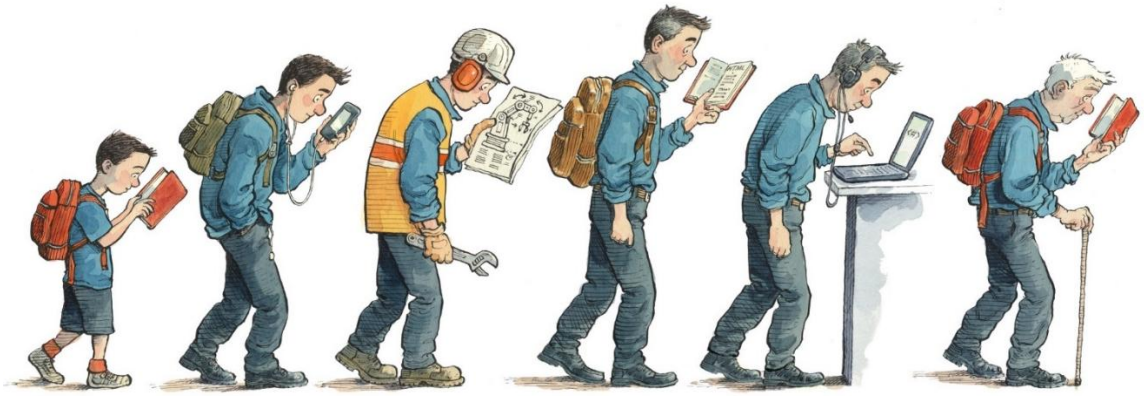


Figure 5. Relevant data driven decision. Differences in nature of data[3]

6. What is the lead time to become proficient in QM?

Literature calculations that are useful for practical analyses are very limited. It is a continuous efforts in WuXi AppTec to develop these methods. We delegate our QM analyses to chemists with appropriate levels of proficiency and encourage everyone to learn. As we invest our efforts to analyze reactions of interest, we become more proficient. QM calculation methods and computational power are evolving very fast, new reactions continue to be discovered, etc. Over time, we learn "What to calculate, How to calculate, How to interpret the results, and How to improve the methods." This is lifelong learning.



David Parkins[4]

Figure 6. Proficiency is relative. Quantum Mechanics is lifelong learning

7. QM analysis is relatively slow, how your teams do them fast enough for prospective analysis for synthetic planning?

We established SOP for different analyses, fit for purpose ones. From orbital to electrostatic potential map analyses, which take a few minutes, to reaction energy profiles, transition state

calculations which could take a few hours. Accessible QM Databases with properties of a large number of organic molecules pre-calculated are tremendously useful.

Figure 7. Fit for purpose calculations are fast enough for prospective analyses

8. What computer system do you use for QM calculations? Are they affordable to synthetic laboratories?

QM calculations are CPU intensive, good to run them on desktops with ≥ 8 cores processors, which have become more affordable. The balance between computational *vs* time+labor+ reagent+disposal costs has changed significantly. QM calculation is becoming a valuable Resource Optimization/Green Chemistry tool.

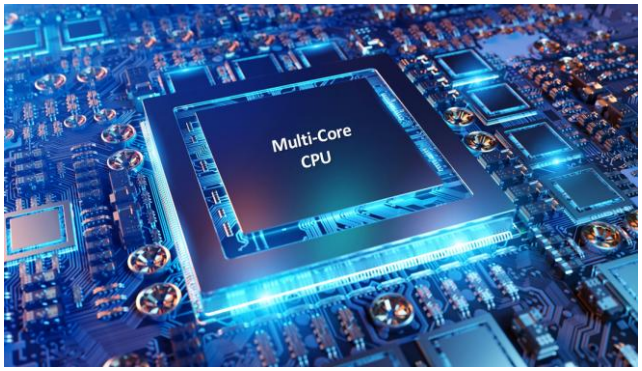


Figure 8. Multiple-core processors speed up routine QM calculations for synthetic planning

Return to Table of Contents

References:

[1] L. Talirz, L.M. Ghiringhelli, B. Smit, *Living Journal of Computational Molecular Science,* **2021**, *3(1),* 1483.

[2] a) R. Peverati, *Int. J. Quantum Chem.* **2021**;121:e26379 b) L. Goerigk, N. Mehta, *Aust. J. Chem.* **2019**, *72*, 563.

[3] a) F. Strieth-Kalthoff, F. Sandfort, M.H.S. Segler, F. Glorius, *Chem. Soc. Rev.,* **2020***, 49,* 6154. b) Y.F. Guan, C.W. Coley, H.Y. Wu, A. Ranasunghe, E. Heid, T.J. Struble, L. Pattanaik, W.H. Green, K.F. Jensen, *Chem. Sci.,* **2021**, *12*, 2198.

[4] https://www.davidparkins.com

[5] a) Spartan Spectra and Physical Properties Database: https://www.wavefun.com b) QM7-X: J. Hoja, L.M. Sandonas, B.G. Ernst, A. Vazquez-Mayagoitia, R.A. DiStasio Jr, A. Tkatchenko, *Sci Data,* **2021**, *8*, 43.

Chapter 39 Unraveling Divergence in Haloselectivity of S_NAr and Cross Coupling *vs* Halogen-Metal Exchange Reactions

Dong Pan, Jun Liu, Wenfeng Liu, Tommy Lai, Yongsheng Chen, John S. Wai

Divergence in haloselectivity of polyhalogenated substrates is a long-standing puzzle in organic chemistry. For example, nucleophilic aromatic substitutions of 2,5-dibromopyridine (see Chapters 1, 7, 10, and 19 for analyses of observed selectivity) and palladium-catalyzed cross coupling reactions occur preferentially at C-2 [1,2] (see Chapter 5 for analysis), while halogen-metal exchange reactions occur preferentially at C-5 (Figure 1)[3,4].

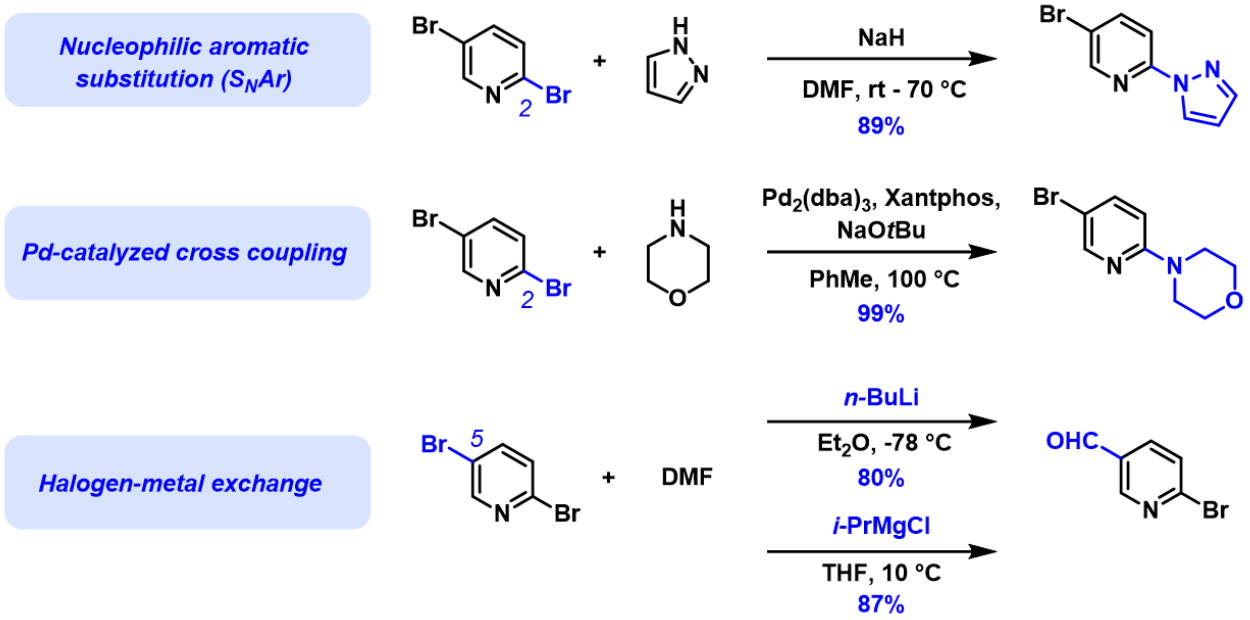


Figure 1. Haloselectivity: S_NAr and cross coupling *vs* halogen-metal exchange reactions of 2,5-dibromopyridine.

How to account for this and other similar divergent examples reported in the literature[5,6]? What is the quantum origin of these divergences in reactivity?

String-of-Pearls: Four characteristic LUMO lobes along the C-X Bond

Let's start by looking at the LUMO of bromoethane (Figure 2). Its lobe on the opposite side of the C-Br σ-bond interacts with HOMO of incoming nucleophile for S_N2 reaction. The small lobe on the C-H bond at the C2 enables interaction with base for elimination reaction. Then the distinctive four LUMO lobes along the C-Br bond, with the outer lobe extending further in space, resemble a **String-of-Pearls**. We reasoned that organic halides use unoccupied orbitals with this characteristic feature for halogen-metal exchange reactions and metalation, and

examined orbitals with such feature to account for/predict the distinctive haloselectivity observed in halogen-metal exchange reactions.

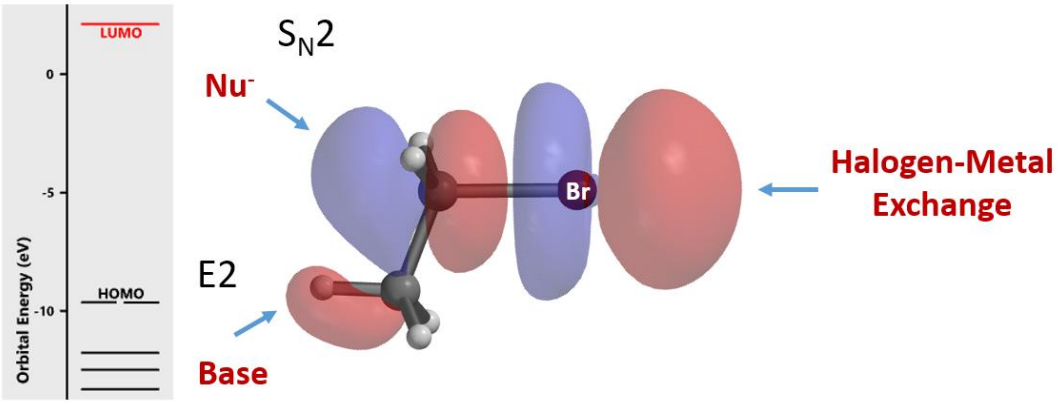


Figure 2. LUMO of bromoethane and chemical reactions associated with its specific lobes[7]

Correlation of Halogen-Metal Exchange Reaction Selectivity with String-of-Pearls shaped LUMO lobes

First, let's see whether we can account for the selectivity observed with the bromine-metal exchange reaction of 2,5-dibromopyridine. Its LUMO, LUMO+1, and LUMO+2 are shown in Figure 3. The "String-of-Pearls" shaped lobe distribution along the C-Br bond is only observed with LUMO+2, not with LUMO nor LUMO+1. The LUMO+2 terminal lobe on C5-Br bond is large and protrudes beyond its Electron Density isosurface (0.002 e/au^3; 99.45%). On the other hand, such a distinctive pattern is not obvious along the C2-Br bond. As such, we reasoned that C5 bromide can preferentially interact with metal reagents, accounting for the C5 haloselectively observed with the bromine-metal exchange reaction of 2,5-dibromopyridine.

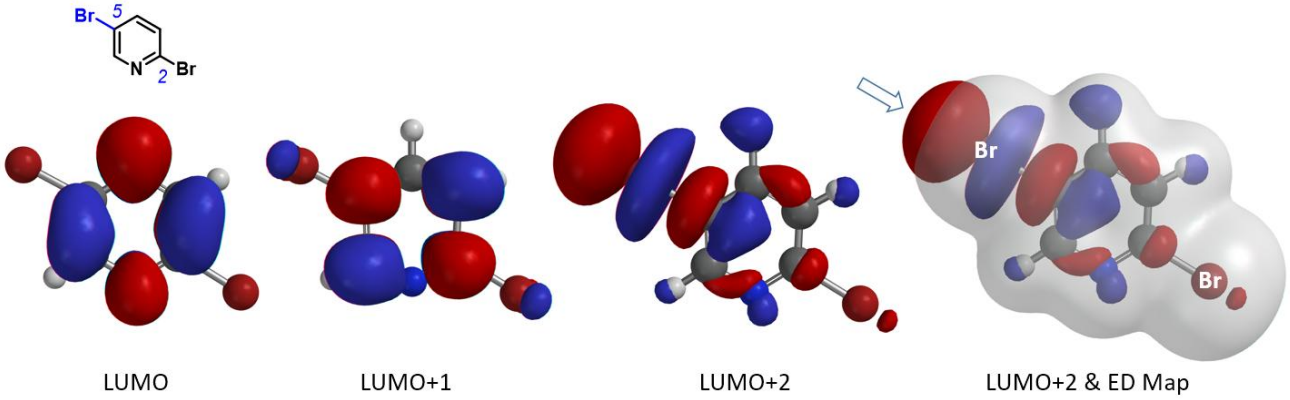


Figure 3. 2,5-Dibromopyridine: LUMO, LUMO+1, LUMO+2, and overlay of LUMO+2 with Electron Density Map[7]

Next, let's apply the analysis to additional examples (Figure 4). With 1,2,4-tribromobenzene, its LUMO+2 has three Strings-of-Pearls, with the largest set along the C2-Br bond, and its terminal lobe extends beyond the isosurface, consistent with C2 selective metalation observed. Similarly for 2,3-dibromothiophene, its LUMO+1 lobes along the C2-Br bond are larger than those on the C3-Br bond, accounting for its selective C2 functionalization. For *N*-methyl-3,5-dibromo-pyrazole and *N*-methyl-5,7-dibromoindole, the characteristic LUMO

lobes are significantly larger at C5 *vs* C3 bromide and C7 *vs* C5 bromide, respectively, aligned with the haloselectivity observed. These examples also exemplify a unique characteristic of halogen-metal exchange reactions, that is, due to the unique unoccupied orbitals involved, with their terminal lobes extending beyond Electron Density isosurfaces, they are not sensitive to neighboring group's steric hindrance.

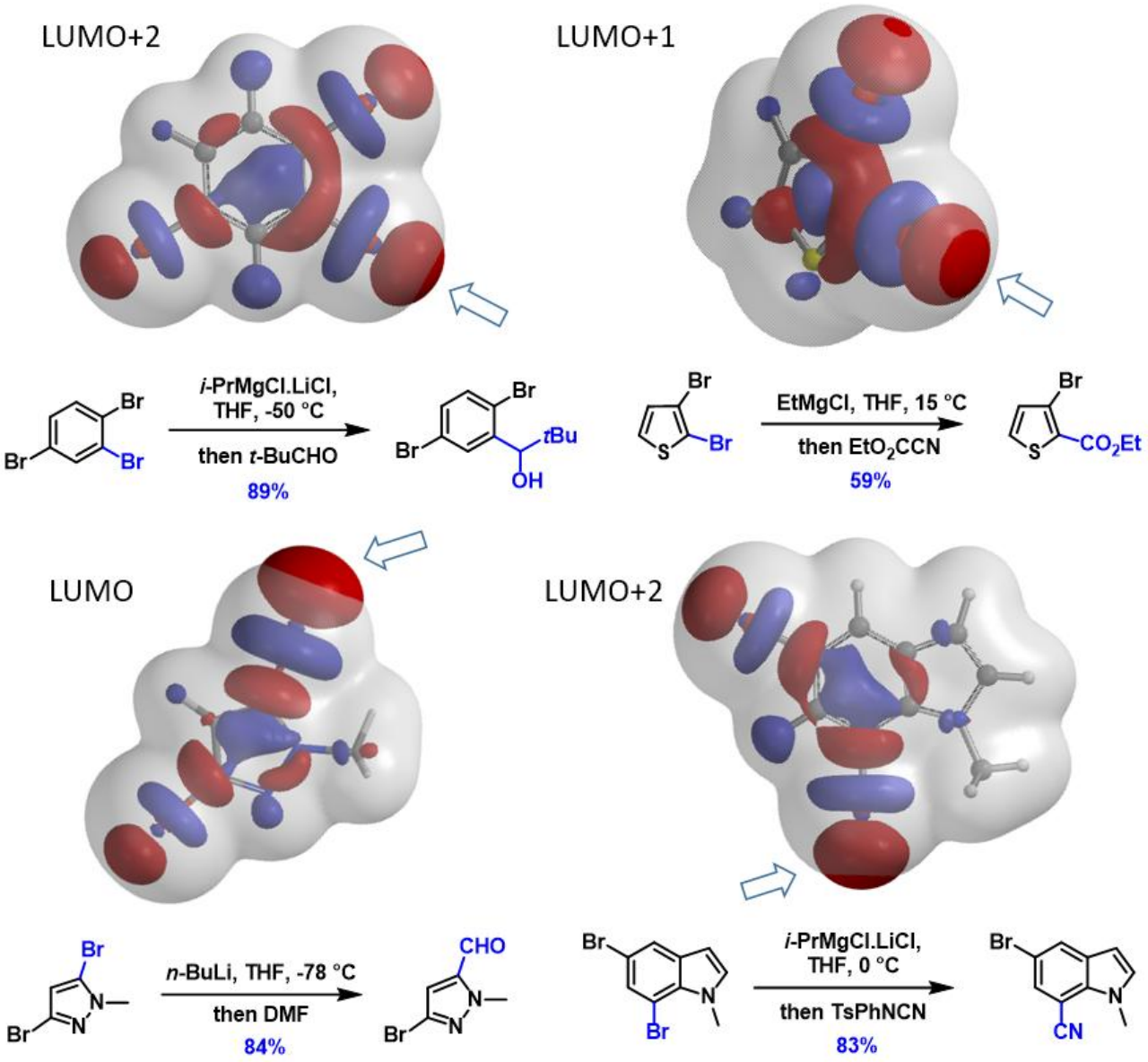


Figure 4. "Pearl-of-String" shaped LUMO or LUMO+n lobes for predicting halogen-metal exchange reactions[7a-d]

This method can also be extended to direct metalation reactions (Figure 5). With 2,4-dichloro-5-fluoropyrimidine, there are two chloro groups. The terminal lobe on the C4-Cl bond extends beyond the isosurface of the molecule while the lobe along the C2-Cl bond does not. This enables selective formation of the corresponding C4-Zn reagent, which reacts with acetic acid, and leads to the selective dechlorination observed.

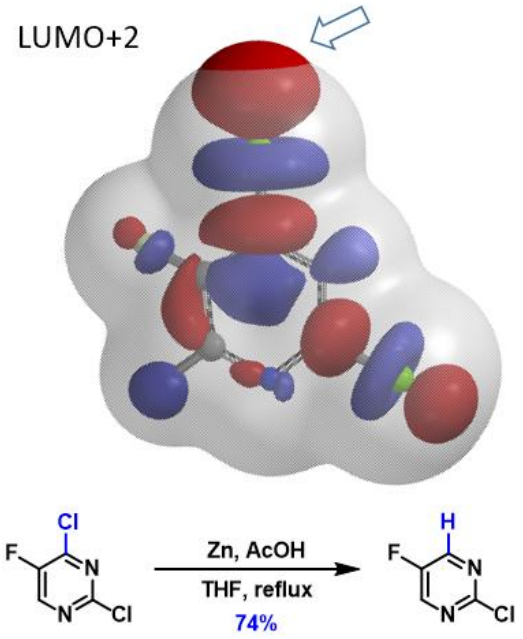


Figure 5. Dechlorination of 2,4-dichloro-5-fluoropyrimidine (LUMO+2 and Overlay with Electron Density Map)[7e]

This LUMO analysis is further validated with >100 examples (reaction solvent THF or Et_2O) reported in the literature, enabling us to predict with high accuracy the haloselectivity in halogen-metal exchange reactions.

With LUMO+2 and LUMO+3 usually at significantly higher energy levels than LUMO and LUMO+1, one may wonder how they can be involved in these reactions. For feasibility of reactions, the key determining factor is the HOMO-LUMO energy gap of the reactants, not their absolute numbers (see QM Chapter 25)[8]. Lithiation and Grignard reagents have high HOMO energy, the resulting energy gaps are within the range where these reactions will proceed. Lithiation reagents usually have higher HOMO energy than Grignard reagents, consistent with the observations that lithium-halide exchange will proceed at lower temperature than the Grignard exchange reactions.

Prediction of Halogen-Metal Exchange Reaction Selectivity with Relative Energy of Carbanions

When the characteristic LUMO or LUMO+n lobes are similar in size and accessibility, we found that relative energies calculated for regioisomers of corresponding carbanions correlate well with the selectivity observed. The regioisomer with lower carbanion energy will have a higher proportion in the product mixture[9]. This is a convenient approximation for prediction. The halogen-metal exchange product should exist with carbon-metal bonds, and their stabilities are affected by various factors such as solvent, temperature, and directing groups.

This can be exemplified with 2,3-dibromobenzothiophene (Figure 6) with two sets of similarly sized LUMO+1 String-of Pearls. Calculated relative energies of C2 *vs* C3 carbanion are 0.00 *vs* 11.66 kcal/mol, respectively, consistent with the observed selective C2 functionalization[10].

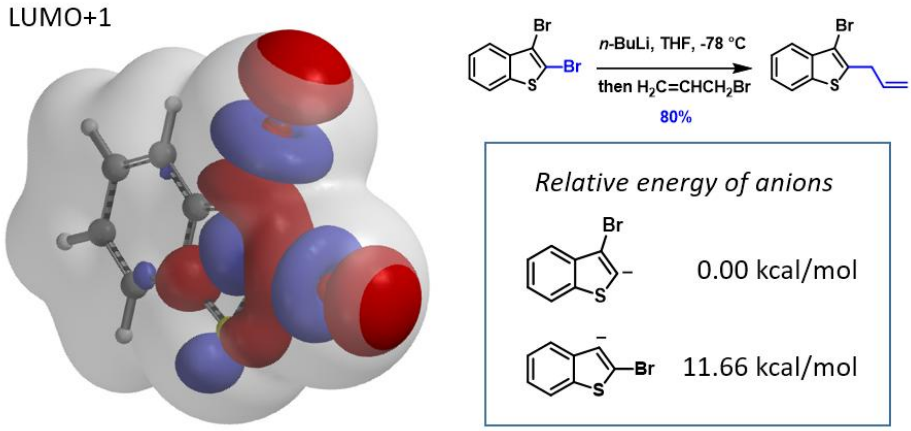


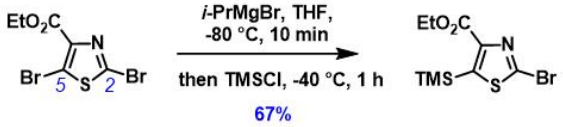
Figure 6. Prediction of haloselectivity of 2,3-dibromobenzothiophene with relative energy of carbanion[10]

In summary, Halogen-Metal Exchange reactions involve unique "String-of-Pearls" shaped LUMO lobes along carbon-halogen bonds of the substrates. Comparison of the size of these unique orbitals/lobes enables us to predict the regioselectivity of the exchange. These unique lobes may be on LUMO, LUMO+1, LUMO+2, or LUMO+3. When the LUMO lobes are similar in size, the relative energy of the corresponding carbanion can be used for differentiation. The differences in haloselectivity of Oxidative Addition and S_NAr *vs* Halogen-Metal Exchange arise from the differences in quantum origin of these reactions, i.e. the former uses LUMO lobe centered over the carbon of the C-X bonds, while the latter utilizes the "String-of-Pearls" large terminal lobes that protrude beyond the Electron Density isosurface.

Expanding our analyses to include LUMO+n enabled us to appreciate why haloselectivity of Halogen-Metal Exchange could be so different. A decades-old chemistry puzzle is solved.

Building on What We Just Learned

Shown in Figure 7 is ethyl 2,5-dibromothiazole-4-carboxylate with its two low energy conformations. LUMO+1 "String-of-Pearls" analyses of them suggested opposite halo-selectivity. Experimentally, the bromine-Grignard exchange is C5 selective[11]. What could be the controlling factor for this transformation?



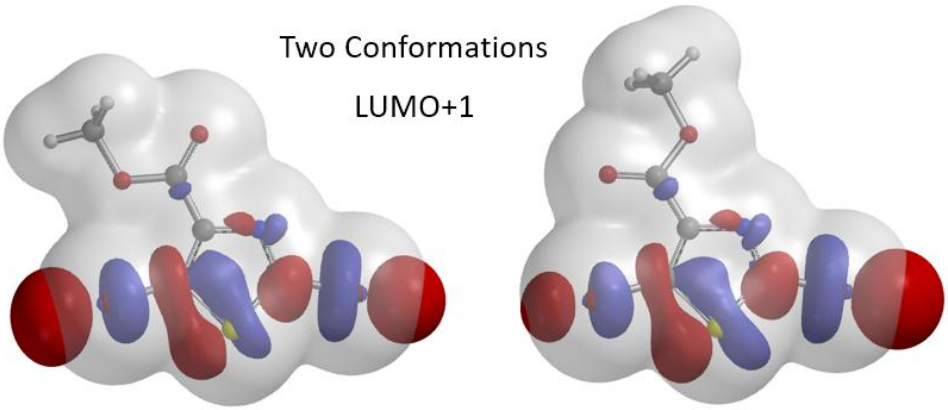


Figure 7. Two conformations of ethyl 2,5-dibromothiazole-4-carboxylate: LUMO+1 and ED map overlay[11]

Return to Table of Contents

References:

[1] C.S. Rani, A.G. Reddy, E. Susithra, K.K. Mak, M.R. Pichika, S. Reddymasu, M.V.B Rao. *Med. Chem. Res*., **2021**, 30, 74.

[2] J.G. Ji, T. Li, W.H. Bunnelle, *Org. Lett*., **2003**, *5*, 4611.

[3] P.N.W. Baxter. *Chem. Eur. J*., **2003**, *9*, 5011.

[4] L. Peng, J.D. Wen, C.G. Pan, X.W. Pan, Y. Xie, W. He, Q.F. Lin, X.Y. Zhang (Chongqing Sansheng Industrial Co., Ltd., China), CN 112479991 A, **2021**.

[5] J.A. Joule & K. Mills, *Heterocyclic Chemistry 4th Ed.* Malden, MA, USA: Blackwell Publishing Ltd., **2000**; pp 26-45.

[6] https://www.scripps.edu/baran/images/grpmtgpdf/Gutekunst_Apr_10.pdf

Divergences in haloselectivity highlighted with 3,6-dibromoindole (pp 2), 2,3,5-tribromo thiophene (pp 3), 2,5-dibromopyridine (pp 6), in palladium catalyzed cross coupling reactions *vs* halide metal exchange reactions.

[7] LUMO, LUMO+1, LUMO+2 and Electron Density Map calculated with Spartan'20 DFT-ωB97X, 6-31G*. (a) A. Krasovskiy, P. Knochel, *Angew. Chem. Int. Ed.,* **2004**, *43*, 3333. (b) C. Christophersen, M. Begtrup, S. Ebdrup, H. Petersen, P. Vedsø. *J. Org. Chem.,* **2003**, *68*, 9513. (c) Y. Yin, C.J. Chen, R.N. Yu, L. Shu, T.T. Zhang, D.Y. Zhang, *Bioorg. Med. Chem*., **2019**, *27*, 1562. (d) P. Anbarasan, H. Neumann, M. Beller, *Chem. Eur. J*. **2011**, *17*, 4217. (e) C. De Savi, A. Pape, J.G. Cumming, A. Ting, P.D. Smith, J.N. Burrows, M. Mills, C. Davies, S. Lamont, D. Milne, C. Cook, P. Moore, Y. Sawyer, S. Gerhardt, *Bioorganic Med. Chem. Lett*., **2011**, *21*, 1376.

[8] L.G. Zhuo, W. Liao, Z.X. Yu, *Asian J. Org. Chem.* **2012**, *1*, 336.

[9] (a) H.J.S. Winkler, H. Winkler, *J. Am. Chem. Soc.,* **1966**, *88*, 964. (b) H.J.S. Winkler, H. Winkler, *J. Am. Chem. Soc*., **1966**, *88*, 969. (c) H.R. Rogers, J. Houk, *J. Am. Chem. Soc.,* **1982**, *104*, 522. (d) K.B. Kenneth, S. Sklenak, F.B. William, *J. Org. Chem*., **2000**, *65*, 2014.

[10] Relative energy of anion calculated with Spartan'20 DFT-ωB97X, 6-31G*, nonpolar solvent. T.P. Sura, D.W.H. MacDowell, *J. Org. Chem*., **1993**, *58*, 4360.

[11] M. Abarbi, J. Thibonnet, L. Bérillon, F. Dehmel, M. Rottländer, P. Knochel, *J. Org. Chem*., **2000**, *65*, 4618.

Chapter 40 QM Torsion Scan for Analysis of Atropisomers

Haowei Wang, Xiaoli Shen, Dong Pan, Tommy Lai, Yongsheng Chen, John S. Wai

In chapters 13 and 14, we discussed how torsional strain controls the selectivity observed in peracid epoxidation of norbornene, and the application of torsion scan to assess whether compounds of interest could exhibit atropisomerism, respectively. In this chapter, we'll discuss further the application of QM torsion scan to guide drug discoveries of atropisomers.

Atropisomers are stereoisomers arising because of hindered rotation about a single bond. This could be exemplified with the biphenyl structure shown in Figure 1. The four substituents A, A', B, and B' on both sides of the single bond of the biphenyl analog exert steric hindrance for rotation. When the energy barrier of rotation is large enough, the pair of atropisomers cannot rapidly interconvert at room temperature and could be separated by chiral chromatography.

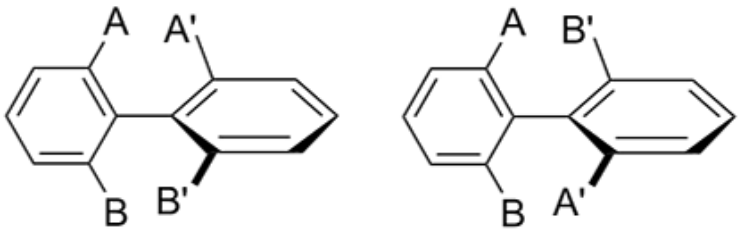


Figure 1. Atropisomer [1]

Classification of Atropisomers

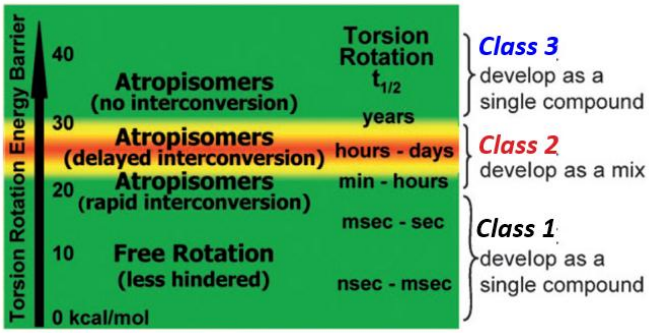
According to statistics, 80% of kinase inhibitors approved by FDA have axial chirality [2]. It was reported in the literature [3] that axial chiral molecules could have different half-lives ($T_{1/2}$) for racemization, depending on the difference in energy barriers for rotation. Atropisomerism can be divided into three classes.

Class 1: ΔE_{Rot} <20 kcal/mol, half-life ($T_{1/2}$) in seconds. Atropisomers can rotate freely and quickly, behaving as single compounds.

Class 2: 20 kcal/mol < ΔE_{Rot} < 30 kcal/mol, $T_{1/2}$ between hours and days, separable with chiral SFC.

Class 3: ΔE_{Rot} > 30 kcal/mol, $T_{1/2}$ in years, the atropisomer is very stable, separable with chiral SFC, could be developed as a single compound.

It has been demonstrated that QM torsion scan analyses could be used to evaluate the energy barrier of the axial rotation, estimate their interconversion rate, guide us to increase or decrease the energy barriers in target designs, provide guidance whether chiral SFC separations are necessary, shorten cycle time during lead optimization, and adopt proper development strategies [4].



Class	$T_{1/2}$	$\Delta E_{rotation}$
3	Years	>> 30 kcal/mol
2	Hours –days	20 - 30 kcal/mol
1	Free to rapid interconversion	< 20 kcal/mol

Figure 2. Classification of Atropisomers [3]

QM Torsion Scan

Next, we'll use Afloqualone to illustrate the use of QM torsion scan to estimate energy barriers of atropisomer interconversion. Shown in Figure 3 is the overall procedure. First, the low-energy conformation of Afloqualone was calculated using MMFF (Spartan'20). Then from the menu, choose Geometry, select Measure Dihedral, click on the **C1-N2-C3-C4** atoms in order, and select the dihedral angle to rotate. The torsion scan was then set to calculate from a dihedral angle of 15° to 345°, with an increment of 15°, in 23 steps (excluding the conformers with dihedral angle of 0° to 360° which will have the methyl and fluoromethyl group crashing into one another) using Hartree-Fock 6-31G**method (gas phase) [3].

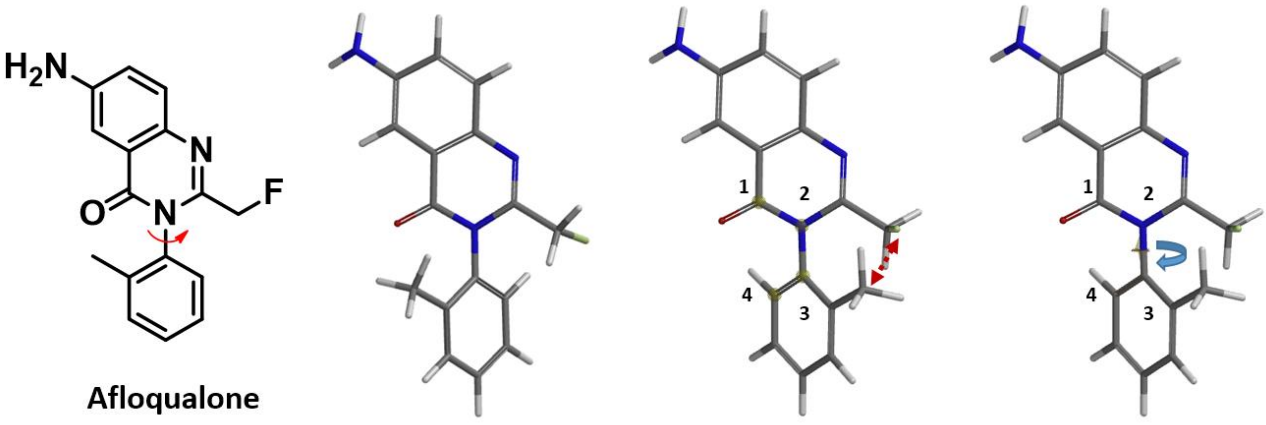
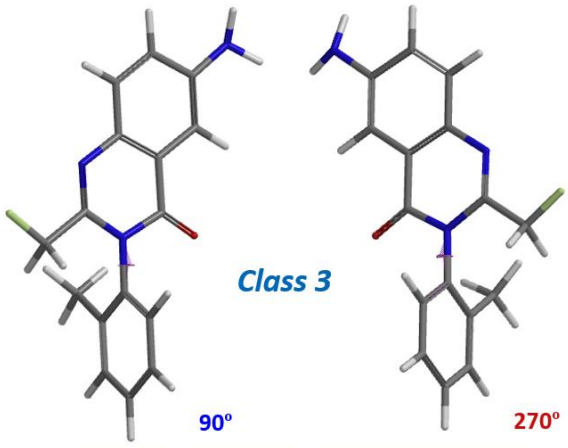


Figure 3. Torsion scan calculation procedure as exemplified with Afloqualone [4]

The calculation results are shown in Figure 4. The two lowest energy conformations have dihedral angles of 90° and 270°, respectively, both with a relative energy of 0 kcal/mol. Their interconversion will require the molecules to go through the conformation with an 180° dihedral angle, with the methyl and the carbonyl group coplanar to one another, and an energy barrier of 34.96 kcal/mol. As such, Afloqualone belongs to class 3 atropisomers,

separable with chiral SFC, having no significant interconversion at room temperature, and can be developed as single chiral compounds.



Class	$T_{1/2}$	$\Delta E_{rotation}$
3	Years	>> 30 kcal/mol

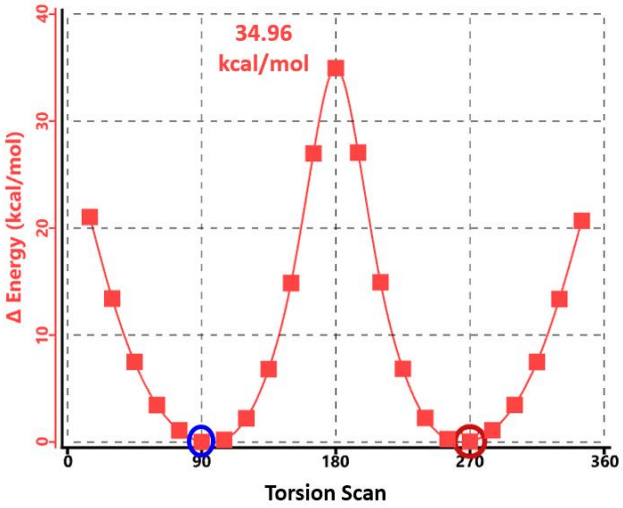
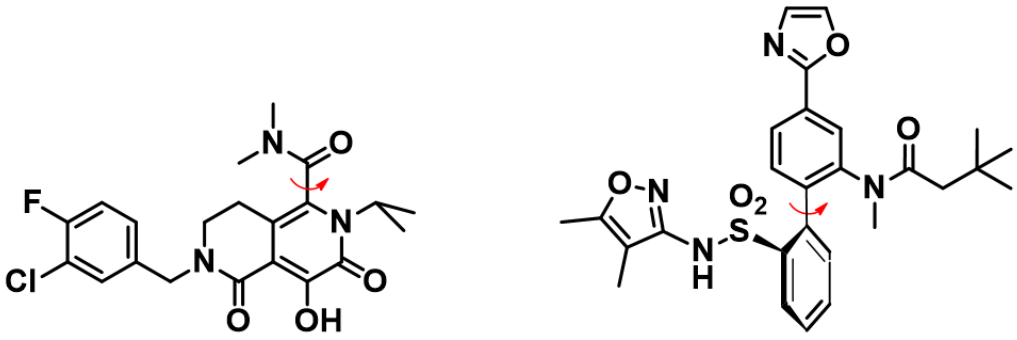


Figure 4. Energy profile from torsion scan of Afloqualone

Drug design can be guided by calculating the interconversion energy barriers of atropisomers, ahead of synthesis. If targeted molecules are predicted to belong to class 2, one could consider increasing the steric hindrance to increase the interconversion energy barrier and transform the compound to class 3 or remove some groups to reduce the steric hindrance and convert it to class 1. And the potential option of eliminating atropisomerism altogether by using the same substituents to make the molecules symmetrical.

Plasma/Serum Could Accelerate Interconversion of Atropisomers

Atropisomers could have different conversion rates in different environments. Atropisomers of HIV-1 integrase inhibitor MK-0536 [5] and endothelin type A antagonist BMS-207940 [6] have been reported to have accelerated interconversion in plasma/serum (Figure 5).



MK-0536 BMS-207940

Figure 5. Molecular structure of MK-0536 and BMS-207940

MK-0536 atropisomerism comes from restricted rotation of the C-C bond marked with the red arrow (Figure 5). QM torsion scan shows that when the two lowest conformations have

dihedral angle of 90° and 270°, respectively, the energy barrier for their interconversion is 27.29 kcal/mol (Figure 6, the *N*-benzyl group is simplified to *N*-methyl group for calculation). As such, MK-0536 belongs to class 2 atropisomer. Under normal conditions, half-life of interconversion will be in days.

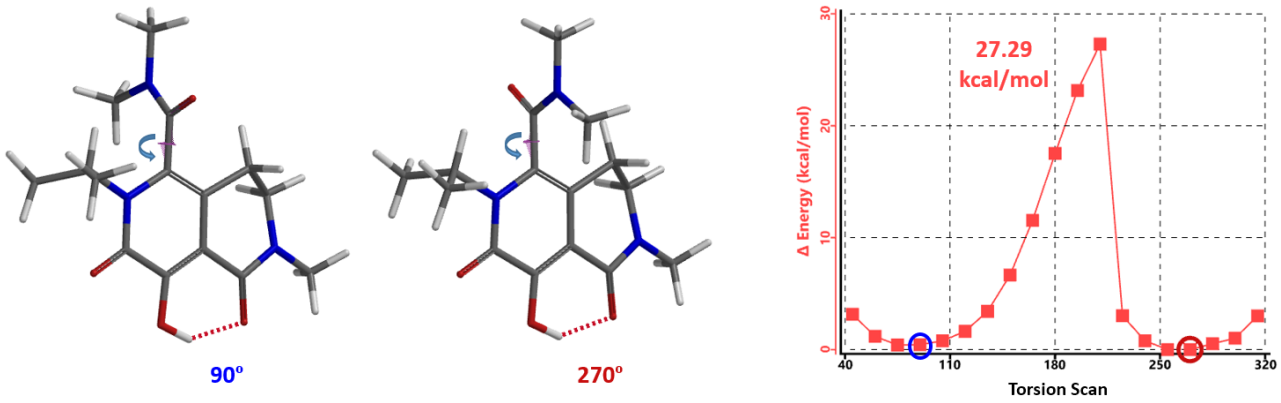


Figure 6. Torsion scan energy profile of MK-0536

However, when MK-0536 was incubated in human plasma at a concentration of 9 μM, the interconversion was completed in <30 minutes, and the $T_{1/2}$ of the racemization is too fast to be measured by standard methods [5]. The rapid interconversion was also observed in dog, rat, and rhesus monkey plasma. With the rapid rate of interconversion in plasma, separation of the atropisomers is not necessary for clinical development of MK-0536.

The above observations remind us that for class 2 atropisomers, it is prudent that we shall determine their interconversion rate in plasma/serum to make informed decisions for their development strategies.

Summary

In this chapter, we discussed the use of QM torsion scan to calculate for energy barrier for interconversion of atropisomers, to guide medicinal chemists to design of drug molecules and establish development strategies. We also learned that plasma/serum could accelerate interconversion of class 2 atropisomers.

Building on What We Just Learned

Shown in Figure 7 is the asymmetrical energy profile obtained from torsion scan of compound **A**. Why is it asymmetrical? Based on the calculated energy barriers, shall we classify it as class 2 or 1? Is chiral SFC purification necessary?

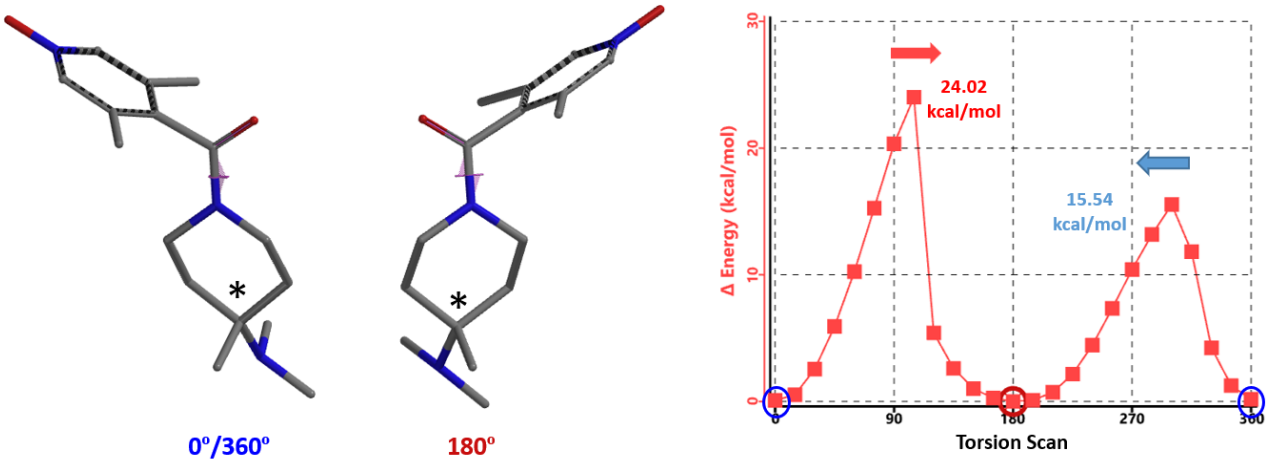


Figure 7. Torsion scan energy profile of compound **A**

Return to Table of Contents

References:

[1] https://cen.acs.org/pharmaceuticals/drug-development/Giving-atropisomers-another-chance/96/i33.

[2] S.T. Toenjes, J.L. Gustafson, *Future Med. Chem*. **2018**, *10*, 409. doi.org/10.4155/fmc-2017-0152.

[3] S.R. LaPlante, L.D. Fader, K.R. Fandrick, D.R. Fandrick, O. Hucke, R. Kemper, S.P.F. Miller, P.J. Edwards, *J. Med. Chem.* **2011**, *54*, 7005. S.R. LaPlante, P.J. Edwards, L.D. Fader, A. Jakalian, O. Hucke, *ChemMedChem* **2011**, *6*, 505.

[4] J. Chandrasekhar, R. Dick, J. Van Veldhuizen, D. Koditek, E-I. Lepist, M.E. McGrath, L. Patel, G. Phillips, K. Sedillo, J. R. Somoza, J. Therrien, N. A. Till, J. Treiberg, A. G. Villaseñor, Y. Zherebina, S. Perreault, *J. Med. Chem.* **2018**, *61*, 6858. L. Xing, B. Devadas, R. V. Devraj, S. R. Selness, H. Shieh, J. K. Walker, M. Mao, D. Messing, B. Samas, J. Z. Yang, G. D. Anderson, E. G. Webb, J. B. Monahan, *ChemMedChem* **2012**, *7*, 273.

[5] M.S. Egbertson, J.S. Wai, M. Cameron, R.S. Hoerrner. Discovery of MK-0536: A Potential Second-Generation HIV-1 Integrase Strand Transfer Inhibitor with a High Barrier to Mutation in *Antiviral Drugs: From Basic Discovery through Clinical Trials*, W.M. Kazmierski, Ed.; Wiley: Hoboken, New Jersey, 2011, Chapter 12, pp 167-8.

[6] Y. Shi, M.H. Huang, J.E. Macorb, D.E. Hughes, *J. Chromatogr. A,* **2005**, *67*, 1078.

Chapter 41 QM-ML tool for Deoxyfluorination

Wenfeng Liu, Chen Yu, Dong Pan, Tommy Lai, Yongsheng Chen, John S. Wai

Introducing fluorine into organic compounds can significantly alter their physical and biological properties, by changing their conformation, non-covalent interaction with target proteins, lipophilicity, acidity/basicity, metabolic stability, permeability, etc.[1] Deoxy-fluorination is an important method to convert alcohols to corresponding aliphatic fluorides. A variety of deoxyfluorination reagents have been developed, among which diethylamino-sulfur trifluoride (DAST) is the most used. But DAST has poor thermal stability with explosive potential. In recent years, safer deoxyfluorinating reagents have been developed, such as PyFluor[2], AlkylFluor[3] etc. Yet, the choice of base is also very important with deoxy-fluorination. For example, PyFluor and DBU are reported to be the best combination of acyclic secondary alcohol deoxyfluorination reaction, which can minimize the generation of by-products[2].

For MedChem synthesis, we need to quickly identify better choices of reagent and base for deoxyfluorination of the substrates of our interest. In 2018, Professor Abigail G. Doyle of Princeton University reported the development of a Quantum Mechanics (QM) and Machine Learning (ML) tool for reliable prediction of the optimal reaction conditions of untested substrates for deoxyfluorination[4a].

Creation of the QM-ML Model for Deoxyfluorination

By now, we are familiar with QM analyses of reactions. But what is machine learning (ML)?

ML is a field of inquiry devoted to understanding and building methods that 'learn'. Its learning algorithms build models based on sample data, known as training data, in order to make predictions without being explicitly programmed to do so.

Figure 1 illustrates the process of building the QM-ML model for deoxyfluorination discussed in this chapter[4a, 5]. First, relevant data were collected for 640 screening reactions, including experimental yields, characteristic data of alcohols, sulfonyl fluorides, and bases used in the reactions. This includes QM calculated molecular and atomic parameters of the alcohols, e.g., LUMO and HOMO energies, electrostatic charge, carbon NMR data, dipole moment, etc. and their structural classification, e.g., primary, secondary, tertiary, cyclic, benzylic, allylic, homobenzylic, homoallylic, etc. Then 70% of the data was used as a training set for machine learning with Random Forest algorithm. The resultant model was then evaluated with the remaining 30% of the data. The validated QM-ML model could be used for prediction of optimal conditions for deoxyfluorination for alcohol substrates that are not in the data set.

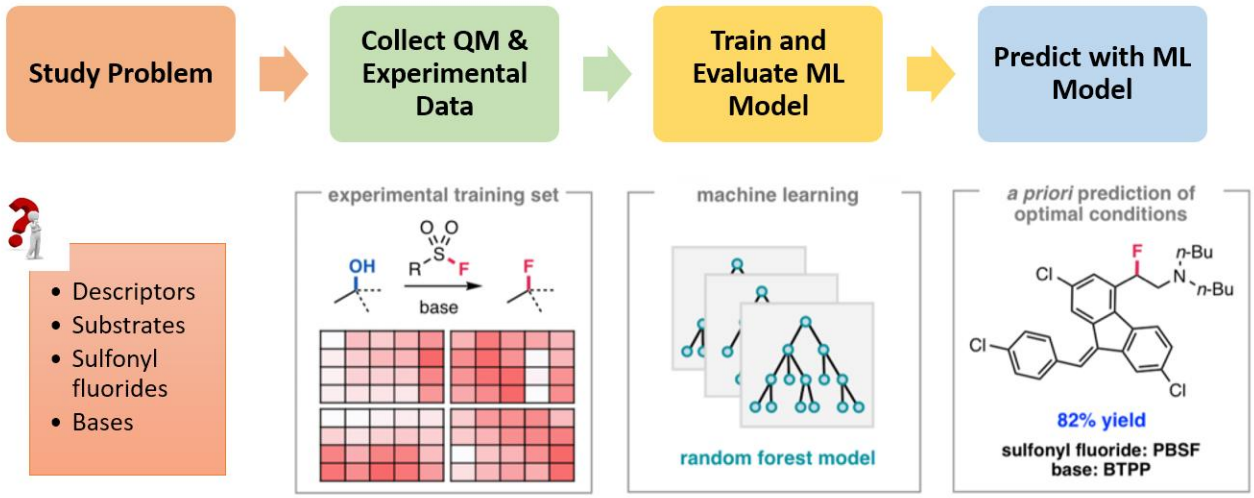


Figure 1. Process for building the QM-ML model for deoxyfluorination[4a, 5]

Figure 2 is the flow chart of using the QM-ML model for predicting reaction conditions of deoxyfluorination. The HOMO and LUMO energies of the alcohol reaction substrate, as well as the electrostatic charge and exposed area of the carbon attached to the hydroxyl group were first calculated with QM. It took a few minutes to create an input file with the above four parameters as well as the classification information of alcohols, and a few seconds for the QM-ML model to generate predicted yields for the 20 reaction conditions. A relatively simple process!

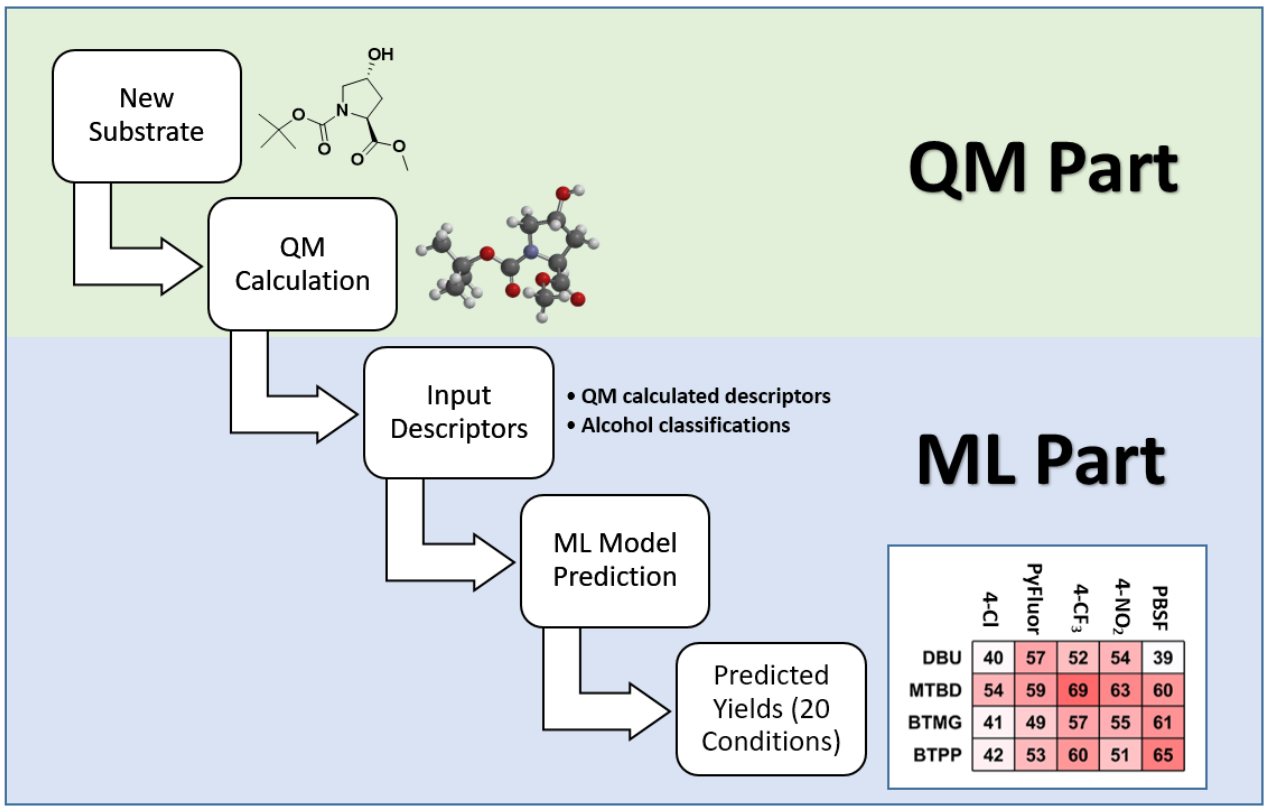


Figure 2. Workflow for using the QM-ML deoxyfluorination model for prediction[4b]

Application of QM-ML Model for Deoxyfluorination

When the QM-ML paper was published, we were converting the precious alcohol **1** to the corresponding mesylate, and then using TBAF displacement to provide the desired fluoride **2**, with an isolated yield of 43%, losing half of the optically pure alcohol **1**.

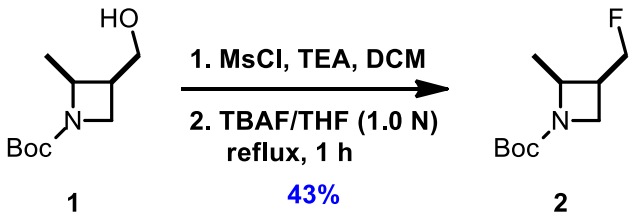


Figure 3. Two step procedure for deoxyfluorination of alcohol **1**

Our prior knowledge in the use of QM for reaction analyses enabled us to quickly deploy Doyle's deoxyfluorination model, which suggested the combination of perfluorobutansulfonyl fluoride (PBSF) and (*tert*-butylimino)tris(pyrrolidino)phosphorane (BTPP) as optimal conditions, with a predicted yield of 42.3%.

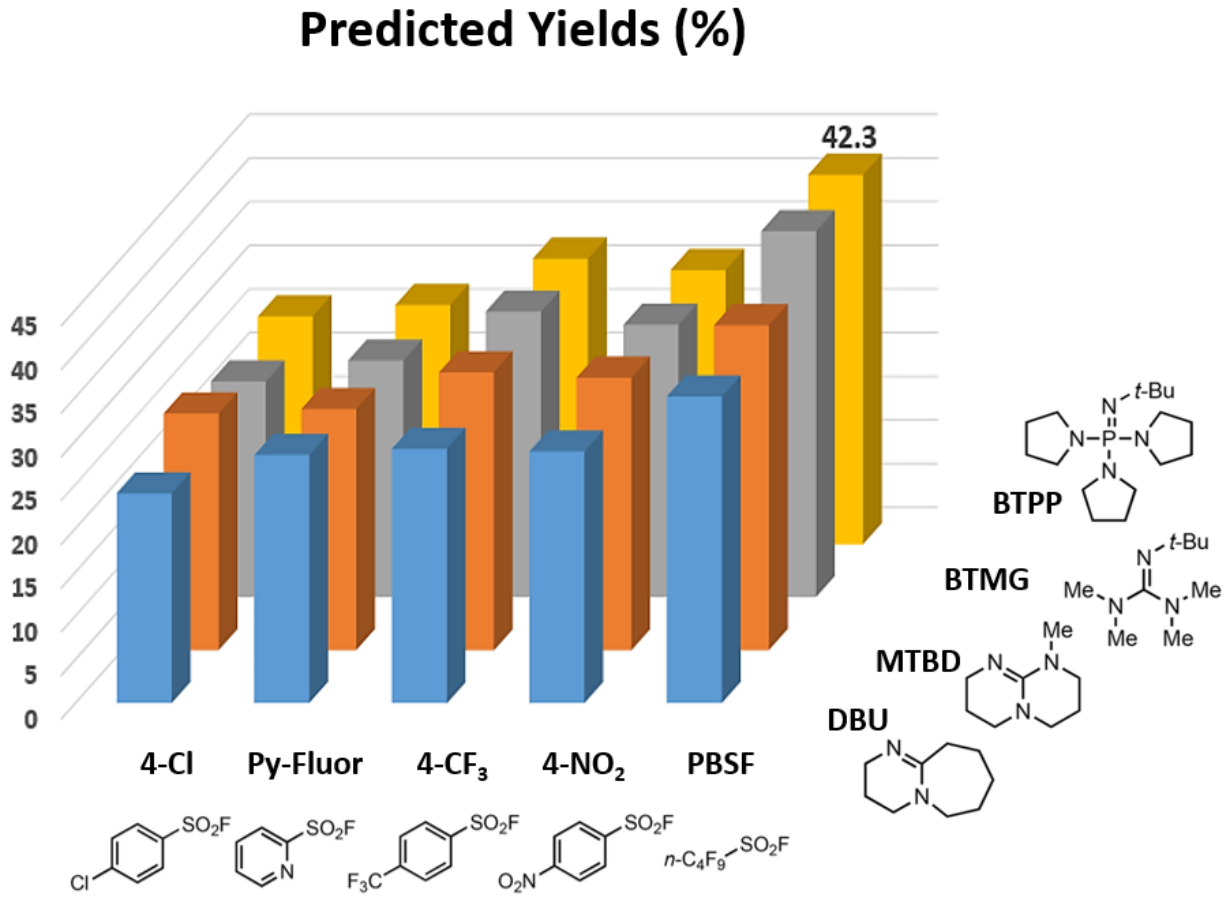


Figure 4. Output from QM-ML deoxyfluorination model for alcohol **1**[4b]

Using the recommended PBSF and BTPP conditions, an isolated yield of 83% was achieved, nearly doubled in yield compared to the earlier two step procedure. Our first application of the QM-ML deoxyfluorination tool in our laboratory is quite successful.

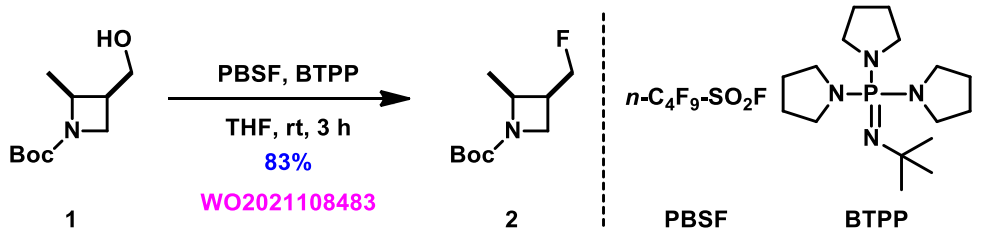


Figure 5. Deoxyfluorination result with suggested reaction conditions from QM-ML tool

With the QM-ML deoxyfluorination model, Doyle's group found optimal reaction conditions for more than 30 alcohols, including active alcohols (benzylic alcohols, allylic alcohols), inactive alcohols, acyclic alcohols, cyclic alcohols, etc. However, for acyclic secondary alcohol **3**, even though the optimal conditions predicted with QM-ML tool was the combination of 4-CF_3PhSO_2F and MTBD with a yield of 69%, experimentally the optimal reaction condition found to be the combination of PyFluor and DBU[4,2].

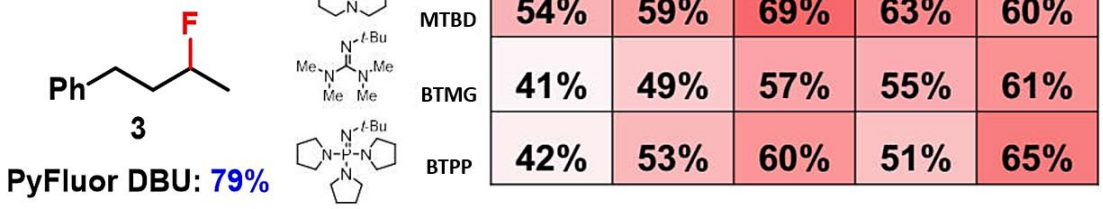


Figure 6. Deoxyfluorination of alcohol **3**[4b]

In summary, combination of QM and ML has generated a very useful model for accurate prediction of optimal conditions for deoxyfluorination and isolated yields. We look forward to development of similar QM-ML reaction conditions predictive tools for other reactions.

We offer the above Deoxyfluorination Prediction as service (Figure 7). If your chemistry requires such analysis, please contact us by email at chemistry_service@wuxiapptec.com.

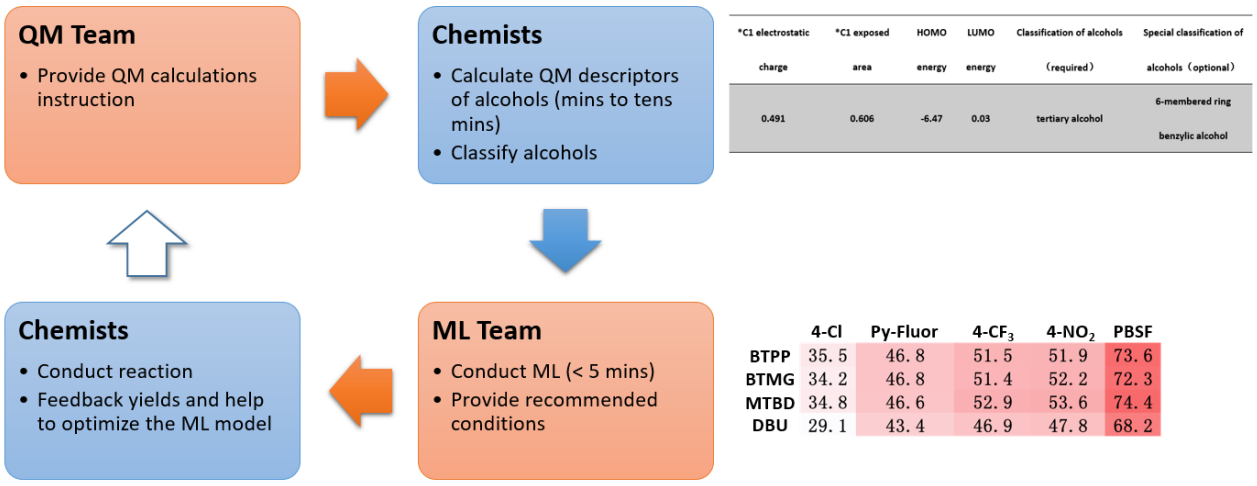


Figure 7. Workflow of our Deoxyfluorination Prediction Service

Building on What We Just Learned

In general, the predicted yields from the QM-ML model have excellent correlation with experimental results (Figure 8). There are some notable exceptions.

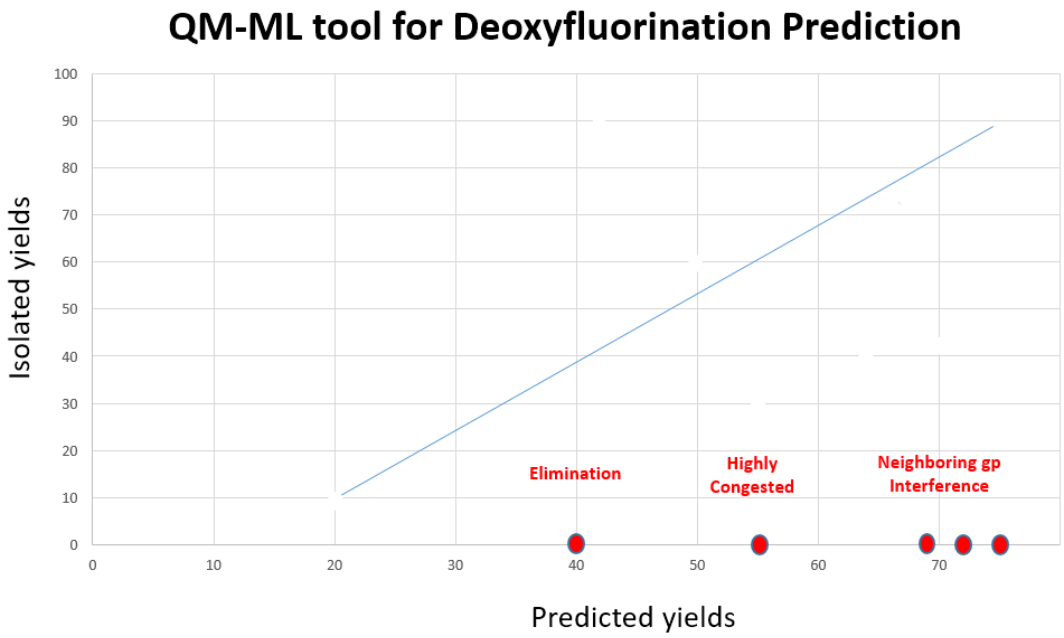


Figure 8. Comparison of QM-ML predicted and isolated yields with our deoxyfluorination

For highly active and easily eliminating alcohols and tertiary alcohols, the above QM-ML model does not provide useful deoxyfluorination conditions. Moreover, we found Sulfox-Fluor[6] works well for alcohols with high activity and easy elimination; Fluolead[7] and XtalFluor-E[8] are useful for tertiary alcohols. If you are aware of other excellent reagents, please share them with us.

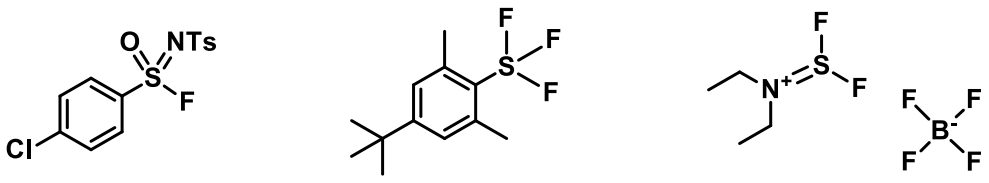


Figure 9. Other useful deoxyfluorination reagents

Return to Table of Contents

References:

[1] D.E. Yerien, S. Bonesi, A. Postigo, *Org. Biomol. Chem.* **2016**, *14*, 8398.

[2] M.K. Nielsen, C.R. Ugaz, W. Li, A.G. Doyle, *J. Am. Chem. Soc.* **2015**, *137*, 9571.

[3] N.W. Goldberg, X. Shen, J. Li, T. Ritter, *Org. Lett.* **2016**, *18*, 6102.

[4] a. M.K. Nielsen, D.T. Ahneman, O. Riera, A.G. Doyle, *J. Am. Chem. Soc.* **2018**, *140*, 5004. b. The three substituted (Cl, CF_3, and NO_2) phenylsulfonyl fluorides are *para* substituted instead of *meta* substituted as described on p 5004 of this paper. They were prepared from commercially available *para* sulfonyl chloride. See paper's supporting information S-21.
[5] F. Strieth-Kalthoff, F. Sandfort, M.H.S. Segler, F. Glorius, *Chem. Soc. Rev.* **2020**, *49*, 6154.
[6] J.K. Guo, C.W. Kuang, J. Rong, L.C. Li, C.F. Ni, J.B. Hu, *Chem. Eur. J.* **2019**, *25*, 7259
[7] T. Umemoto, R.P. Singh, Y. Xu, N. Saito. *J. Am. Chem. Soc.* **2010**, *132*, 18199.
[8] S. Kalidindi, A.S. Gangu, S. Kuppusamy, S. Sathasivam, V. Shekarappa, S. Murugan, S. Bondigela, M. Kandasamy, K. Ghanta, A. Vinodini, A. Shrikant, R. Ramachandran, W.P. Gallagher, N. Kopp, F. González-Bobes, M.D. Eastgate, R. Vaidyanathan, *Org. Process Res. Dev.* **2021**, *25*, 1556.

Chapter 42 Visualizing Hydrogen Bonds Using Electron Density Maps

Dong Pan, Xin Huang, Wenfeng Liu, Tommy Lai, Yongsheng Chen, John S. Wai

Hydrogen Bond (HB) is a common intramolecular and intermolecular interaction. From DNA to structural proteins to enzymes to ligands, HB plays important roles in their structural and biological functions. It is one of the key non-covalent interactions for optimization in drug and new materials design [1].

In general, hydrogen bond can be expressed as "XH····YZ", where "····" represents hydrogen bond, with the hydrogen connected to the electronegative atom X as a hydrogen bond donor, and electron-rich atom Y as the hydrogen bond acceptor, via electrostatic interaction. The most widely used method for judging hydrogen bonds is the bond length. The shorter the hydrogen bond, the stronger the strength. Recently with better understanding of hydrogen bonds and broadening of its scope (e.g. the hydrogen bond acceptor Y is not limited to lone pair electrons; it can also be anions, or π electrons), we could use better method to visualize and even quantify these interactions [2,3].

Visualizing Hydrogen Bonds Using Electron Density Maps

In quantum chemistry, electron density is a measure of the probability that an electron exists in a unit of space around a given point. A surface connecting points of equal value, that is, an isosurface, can be displayed as an array of dots, a mesh, solid or translucent surface. The electron density between atoms reveals the presence of interactions and correlates with the relative strength of these "bonds" [4,5].

As shown in the electron density map of acetic acid dimer in Figure 1, we could see electron density between O and H of the two hydrogen bonds and the three-dimensional boundary of electron density between the hydrogen bonds [2]. This varies along the hydrogen bond, connecting the atoms like a bridge. In addition, we could also visualize electron density in the form of contour (slice) map. In recent years, scientists discovered that there is no clear boundary between hydrogen bonds and covalent bonds, and the essence of short and strong hydrogen bonds can be described as a covalent bond [6].

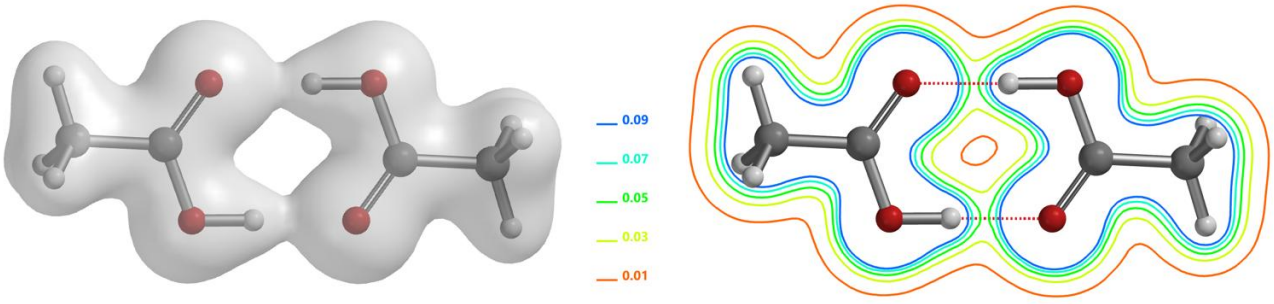


Figure 1. Electron density surface (IsoValue 0.06 e/au^3, 91.13%) and slice contour map of acetate dimer[4]

Determining Hydrogen Bond Strength in Molecules Using Electron Density Maps

For molecules with multiple hydrogen bonds, it is quite straightforward to evaluate relative strength of these intramolecular interactions with QM calculated electron density map. This could be exemplified with 8-hydroxy-*N*-methyl-1,6-naphthyridine-7-carboxamide (Figure 2) with O2-H1 distance of 1.69 Å and O2-H1-O1 bond angle of 148°. This satisfies most modeling software's criteria of hydrogen bond: An N or O -linked hydrogen between 1.6 Å and 2.1 Å from another O or N, and the XH····Y bond angle is >120° [4]. From the electron density contour map, we could see the hydrogen bond between O2-H1 is relatively stronger and the one between N1-H2 is relatively weaker. Note that the latter (N1-H2 distance of 2.19 Å and the N1-H2-N2 bond angle of 108°) does not meet the narrow criteria of "hydrogen bond" [4].

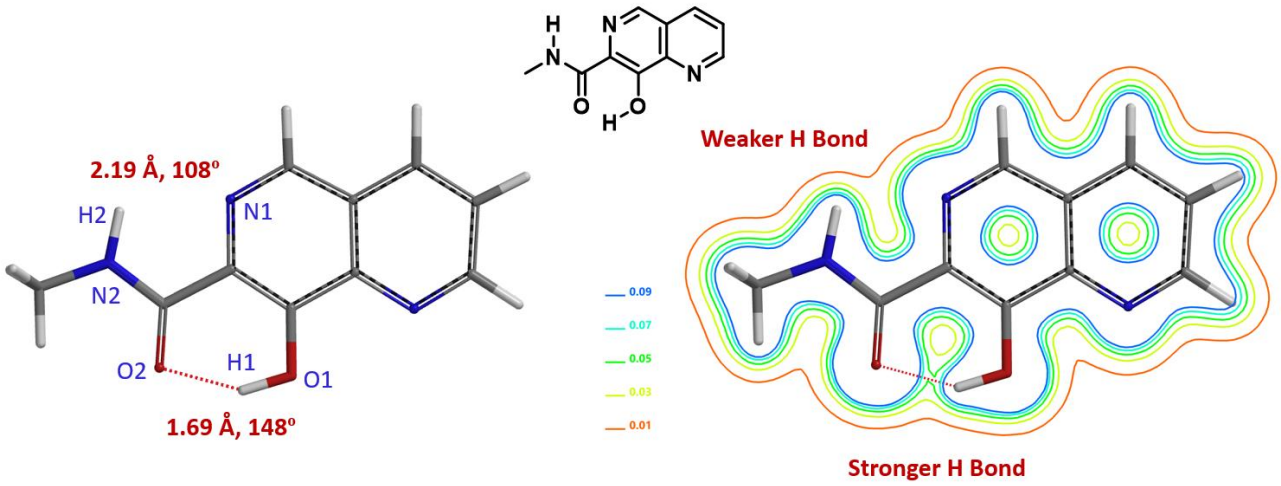


Figure 2. Electron density slice contour plot of 8-hydroxy-*N*-methyl-1,6-naphthyridine-7-carboxamide

The weaker hydrogen bonding interaction between N1-H2 can also be visualized with semi-transparent surface of the electron density map (Fig. 3). When IsoValue is set to 0.0200 e/au^3, it can be clearly seen that there is bridging electron density between O2-H1 and N1-H2 and their difference. We could also set the IsoValue to a larger threshold of 0.0389 e/au^3 to visualize the difference [5].

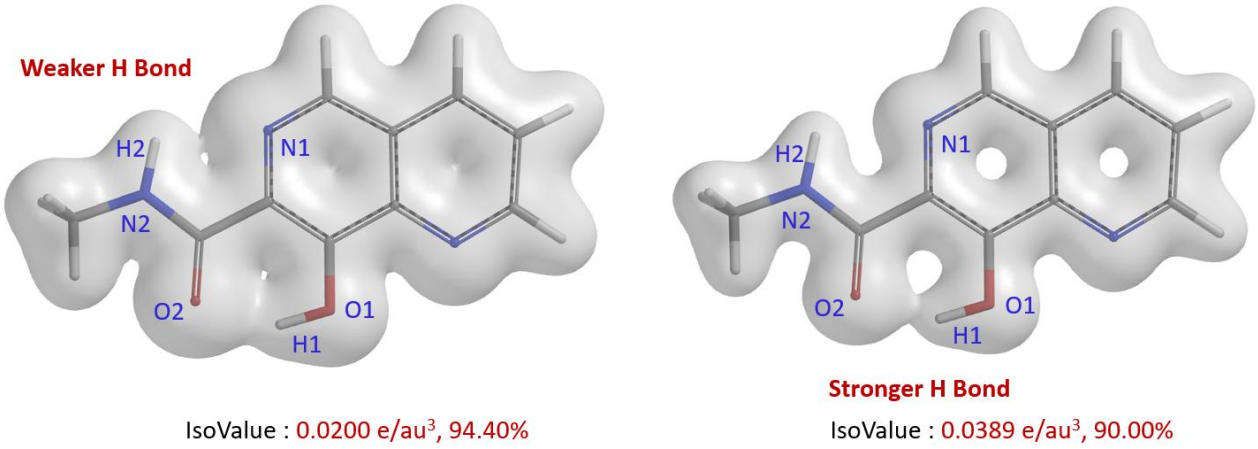


Figure 3. Electron density surface of different IsoValues of 8-hydroxy-*N*-methyl-1,6-naphthyridine-7-carboxamide

Hydrogen bonding is of great significance to biological macromolecules. It is an important factor in determining the secondary, tertiary, and quaternary structures of proteins and nucleic acids. Hydrogen bonding interactions in complex structures can be visualized with QM calculated electron density map. This could be exemplified with the three-dimensional structure of a tetrapeptide which folds into a hairpin-like configuration through hydrogen bonding [7]. Since the three hydrogen bonds are not coplanar, it is easier to visualize them and compare their relative strength with the mesh representation of its electron density surface (Figure 4).

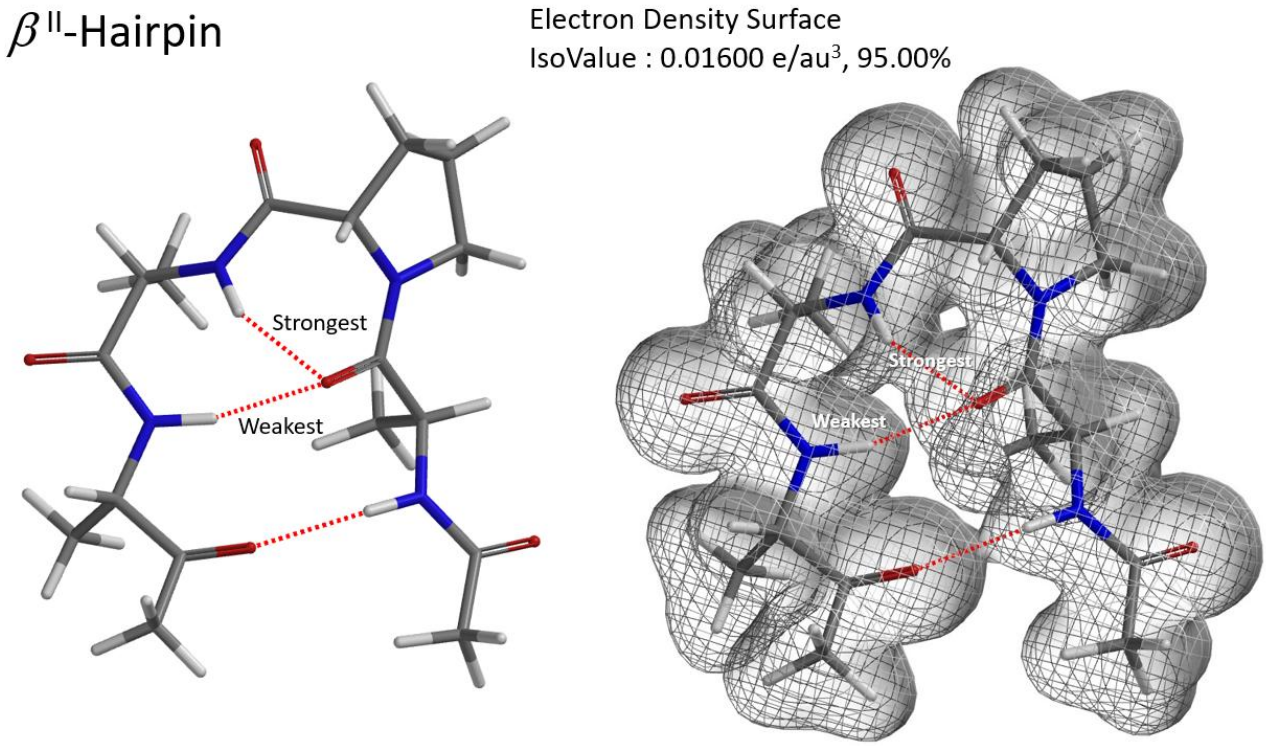


Figure 4. Hydrogen Bonds and Electron density surface of a β-hairpin tetrapeptide

Effect of Hydrogen Bonding on Reactivity

We learned that transition state stabilized by hydrogen bonding could speed up the reaction by lowering the activation energy [8]. However, when the reactant is stabilized by hydrogen bonding, this will need to be broken for the reaction to proceed, raising the energy barrier.

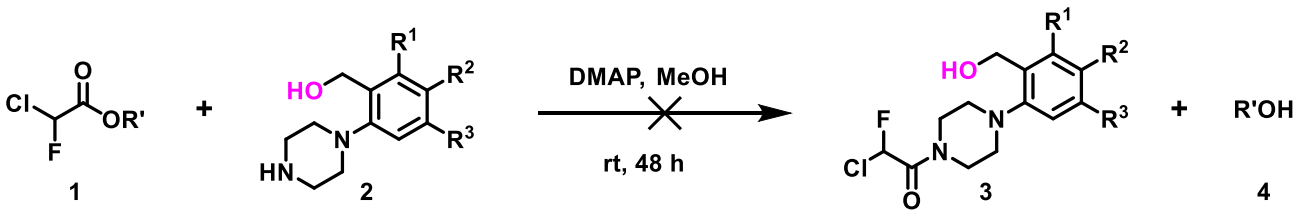


Figure 5. Effect of hydroxyl groups on amine transesterification

In the ester amidation shown in Figure 5, the des-hydroxy analog of substrate **2** proceeded efficiently to the corresponding amide. In the presence of the hydroxy group, the reaction does not occur at all under the same reaction conditions. QM analyses suggest that substrate **2** could form a hydrogen bond complex with the methanol solvent, involving two hydrogen bond interactions between the proton of OH in compound **2** and the N in NH, with an energy drop of 5.59 kcal/mol (Figure 6). We reasoned that for the desired nucleophilic attack to occur, the reaction needs to break both hydrogen bonds first, resulting in an increase in the activation energy, and thus the reaction does not occur. Since methanol solvent is essential for amidation, we could avoid formation of the above-mentioned complexes with silyl protection of the hydroxyl group prior to acylation of piperidine **2**.

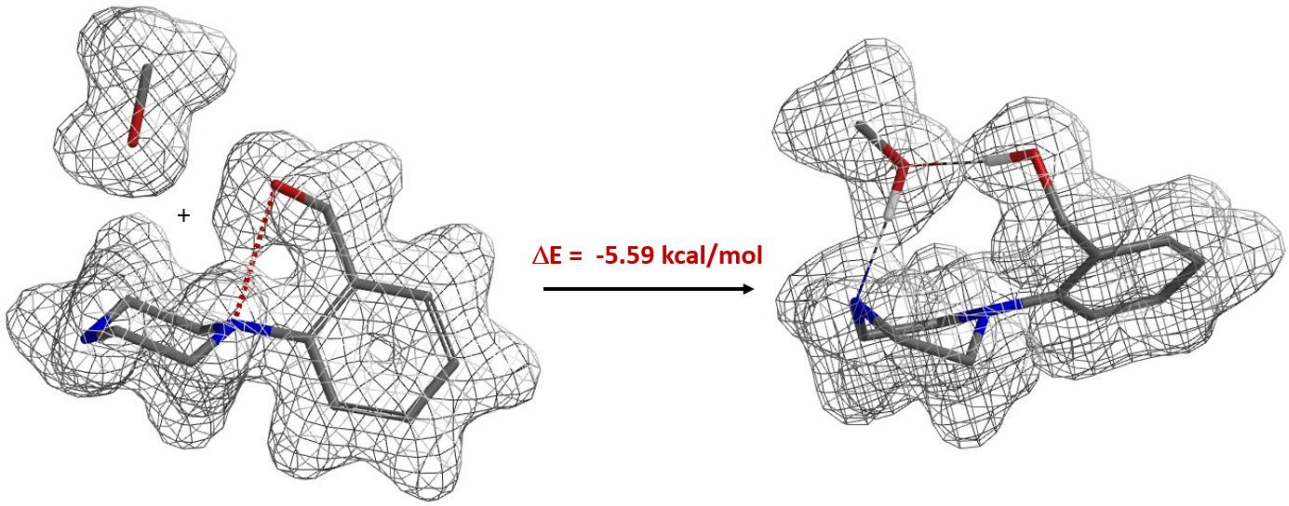


Figure 6. Electron density surface of compound **2**, methanol (left) and their hydrogen bonding complex (right)

Similarly, electron density maps also enable us to visualize and compare various non-covalent interactions in drug discovery [9], such as halogen bond inhibitor binding pocket (Figure 7) [10], lone pair-π interaction in biological systems (Fig. 8) [11], proline H-pi interaction (Fig. 9) [12], dihydrogen bond (Fig. 10) [13], etc.

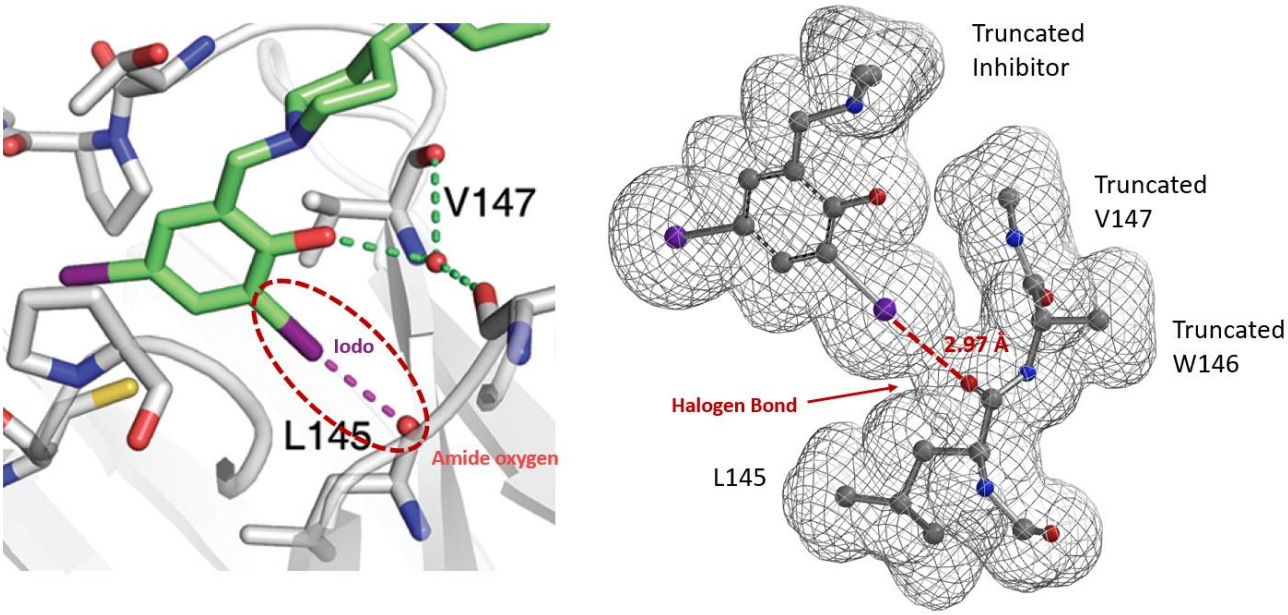


Figure 7. Crystal Structure and Electron Density Surface of a diiodine inhibitor in the p53 mutant Y220C binding pocket

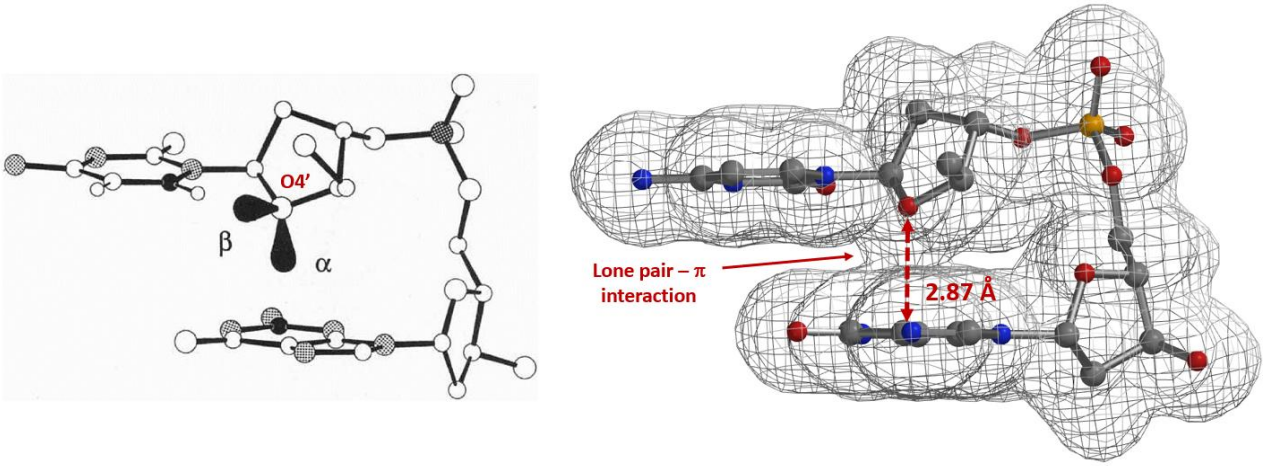


Figure 8. Lone pair-π interaction: Electron-density surface of cytidine deoxyribose-guanine base

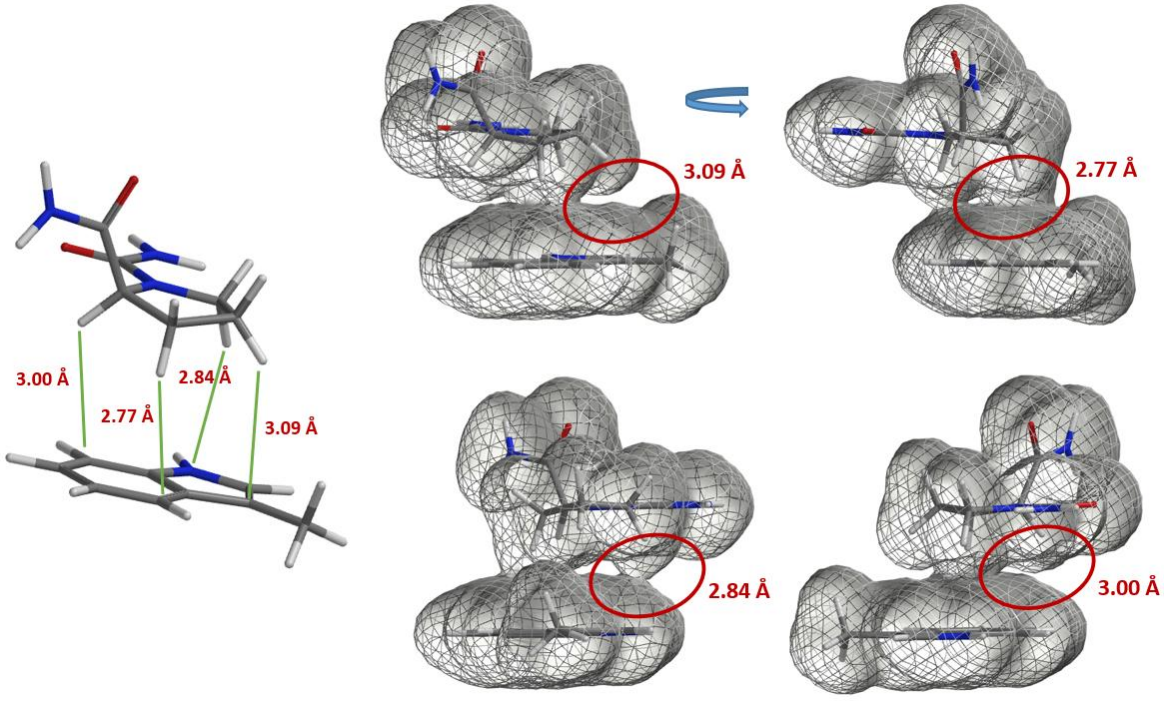


Figure 9. Proline H – pi interactions in proline rich domains

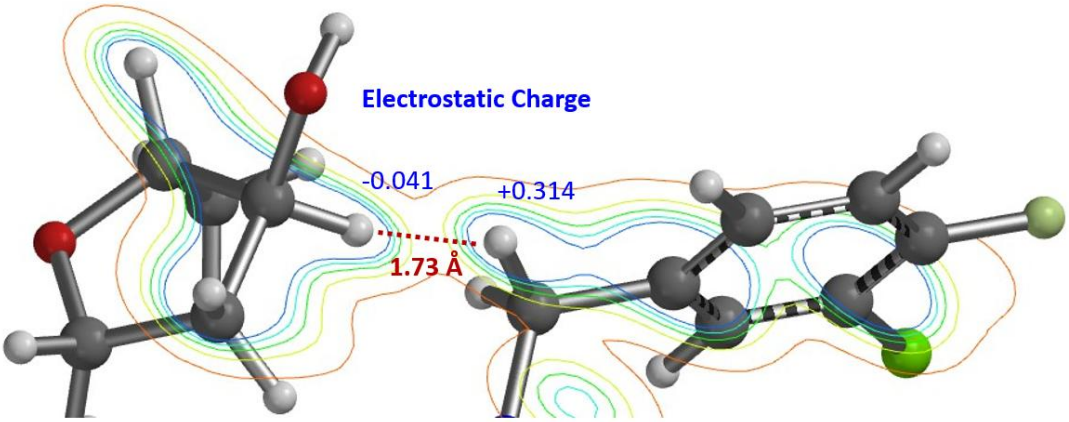


Figure 10. Dihydrogen bond between adenine ribose C3' H and benzylic H of HIV-1 integrase inhibitor MK-2048

In summary, electron density map enables us to visualize and compare the relative strength of intermolecular and intramolecular hydrogen bonds. When there are multiple hydrogen bonds that are not in the same plane, the mesh representation of the electron density surface map could be very useful. The same method can also be applied to other non-covalent interactions in analysis of chemical reactions and biological systems, providing us with better understandings to these interactions, crucial in drugs, materials, or ligands designs. Visualizing these non-covalent interactions is conceptually simple yet very powerful. What matters more is proper analyses of these density maps to generate crucial insights to guide drug designs and accelerate discoveries.

Building on What We Just Learned

A. Feigenbaum reported that on silica gel TLC plate, polarity of *o*-nitrophenol is lower than that of *p*-nitrophenol and *m*-nitrophenol. With DCM-pentane as mobile phase, the polarity is even lower than that of nitrobenzene [14]. Polarity of *p*-nitrophenol and *m*-nitrophenol is close to that of phenol. Are the differences due to acidity, or hydrogen bonding? Based on the pKa and electron density surface maps of the three molecules shown in Figure 11, what will be your conclusions?

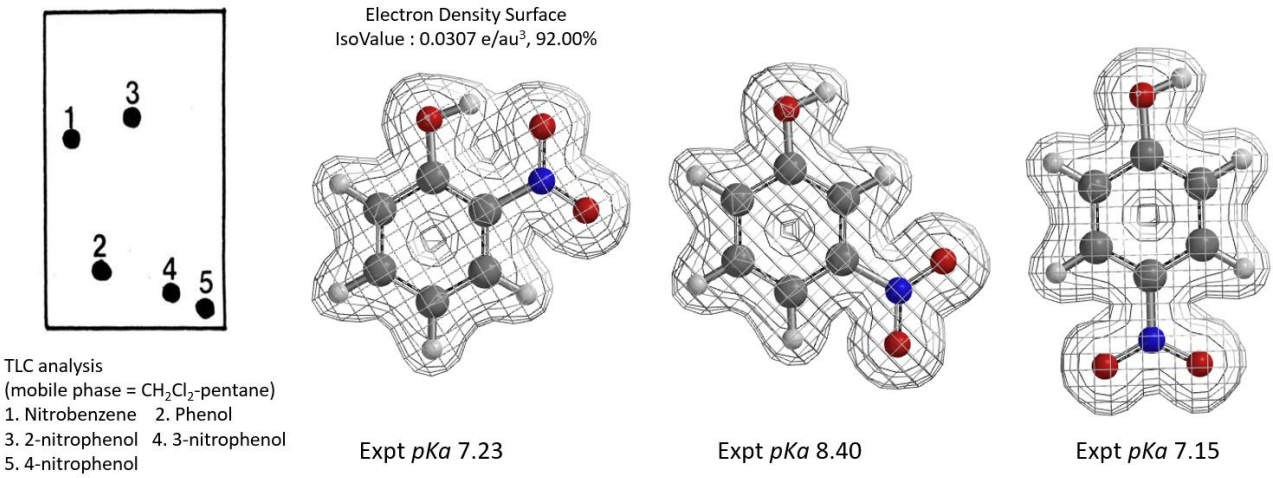


Figure 11. pKa and Electron Density Maps of *ortho*-, *meta*-, and *para*-nitrophenols

Return to Table of Contents

References:

[1] L.J. Karas, C.H. Wu, R. Das, J.I.C. Wu, *WIREs Comput. Mol. Sci*., **2020**, *10*, 1477.

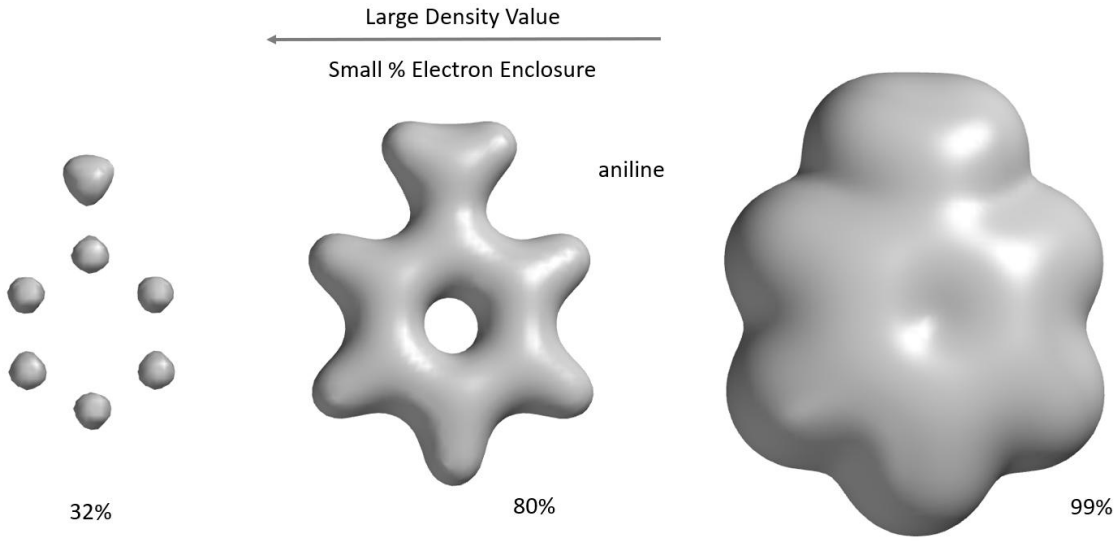
[2] Definition of the hydrogen bond (IUPAC Recommendations 2011): E. Arunan, G. Desiraju, R. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. Clary, R. Crabtree, J. Dannenberg, P. Hobza, H. Kjaergaard, A. Legon, B. Mennucci, D. Nesbitt, *Pure Appl. Chem.*, **2011**, *83*, 1637.

[3] QM Chapter 23: A QM Study of the para Regioselectivity of $TBABr_3$ Bromination.

[4] Spartan'20 *Tutorial and User's Guide*. Irvine, CA, USA: Wavefunction, Inc. **2021**; p 256, 362, 368.

[5] https://en.wikipedia.org/wiki/Electron_density

A space-filling model formed by an electron-dense surface containing the majority (>99%) of electrons in a molecule, a van der Waals surface. The electron-dense surface, which contains about 80% of the electrons, reveals the connections ("bonds") between atoms. The highest electron density is distributed around the atoms. Different electron density surfaces can reflect different volumes and shapes of molecules.



[6] B. Dereka, Q. Yu, N.H.C. Lewis, W.B. Carpenter, J.M. Bowman, A. Tokmakoff, *Science*, **2021**, 371, 160.

[7] R.W. Hoffmann, *Angew. Chem. Int. Ed*., **2000**, *39*, 2054.

[8] Examples in QM chapters 15, 23, 26 and 33 illustrated the effect of hydrogen bondings on activation energies.

[9] E.R. Johnson, S. Keinan, P. Mori-Sanchez, J. Contreras-Garcia, A.J. Cohen, W.T. Yang, *J. Am. Chem. Soc.,* **2010**, *132*, 6498. Non-covalent interactions have a unique signature, and their presence can be revealed solely from the electron density. They are highly non-local and manifest in real space as low-gradient isosurfaces with low densities.

[10] G. Cavallo, P. Metrangolo, R. Milani, T. Pilati, A. Priimagi, G. Resnati, G. Terraneo, *Chem. Rev.* **2016**, *116*, 2478.

[11] J. Novotný, S. Bazzi, R. Marek, J. Kozelka, *Phys. Chem. Chem. Phys*., **2016**, *18*, 19472.

[12] T. Ozawa, K. Okazaki, K. Kitaura, *Journal of Computational Chemistry,* **2011***, 32,* 2774.

[13] S. Hare, A.M. Vos, R.F. Clayton, J.W. Thuring, M.D. Cummings, P. Cherepanov, *Proc Natl Acad Sci USA*, **2010**, 107, 20057.

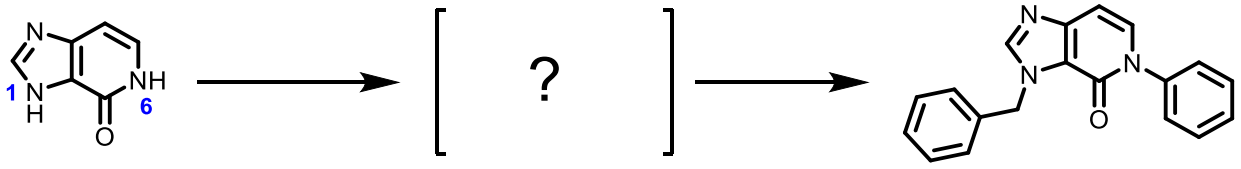
[14] A. Feigenbaum, *J. Chem. Edu*., **1986**, *63*, 815.

Chapter 43 QM Quiz

Dong Pan, Hailong Ren, Qiuyue Wang, Zhong Zheng, Tommy Lai, Yongsheng Chen, John S. Wai

This chapter covers four questions we had in an open-book QM examination in WuXi AppTec. For setting up QM calculations and interpreting results, participants have access to our QM e-book [1] and Spartan manual [2]. They were requested to refrain from doing literature search for answers during the 2-hour exam.

Question 1



Which nitrogen shall be selectively functionalized first to attain the final compound? What will be the governing factor? What QM parameter to calculate for analysis? Please provide QM calculation results and interpretation.

Answer (Dong Pan)

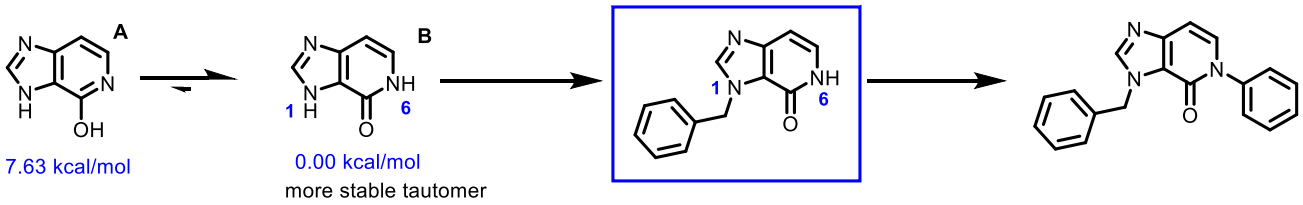


Figure 1. Relative energies of the two major tautomers of the substrate

The imidazopyridinone substrate could exist in several tautomeric forms, among which are tautomers **A** and **B** with relatively lower energy. The relative energy difference between the **A** enolic *vs* the **B** amide form is 7.63 *vs* 0.00 kcal/mol (Figure 1). As such, the substrate is 100% in the amide form and being used for the following analysis. The relative acidity of the two NH groups is the key factor in deciding for the proper order of reactions. This will require us to calculate for the electrostatic potential (ESP) map of structure **B** for comparison [4].

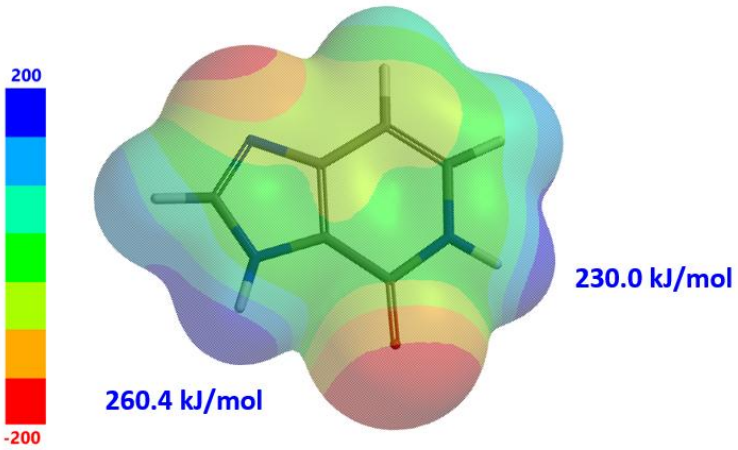
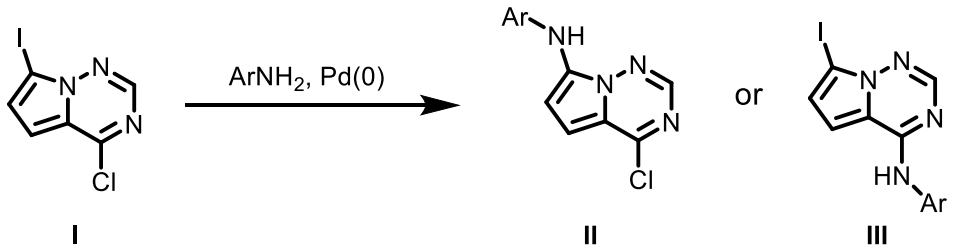


Figure 2. Electrostatic potential diagram of tautomer **B**

ESP map reveals N1-H and N6-H electric potential are 260.4 and 230.0 kJ/mol, respectively (Fig. 2). The larger the value, the stronger the acidity. With electrostatic potential of N1-H significantly higher than that of N6-H by 30 kJ/mol, we could selectively deprotonate N1-H for benzylation first [5].

Question 2



I **II** **III**

For the above Buchwald amination, will the reaction form a single product or mixture of products? If it is a single product, which one, II or III? What are the QM parameters to calculate for analysis? Please provide QM calculation results and interpretation.

Answer (Hailong Ren)

For the Buchwald amination of 4-chloro-7-iodopyrrolotriazine (**I**), we first examine LUMO of the substrate [6] (Figure 3). The LUMO lobe on C-Cl carbon is significantly larger than that on C-I carbon with the iodine carbon atom on the pyrrole, suggestive of selective oxidative addition (OA) and SnAr reaction at the C-Cl carbon [7]. LUMO+1 of **I** has a unique" string of pearls" lobes along the C-I bond, suggestive of selective metal halogen exchange reaction at the C-I carbon [8]. LUMO+1 does not have orbital lobe characteristic for OA and SnAr reactions, we do not need to consider it for Buchwald amination.

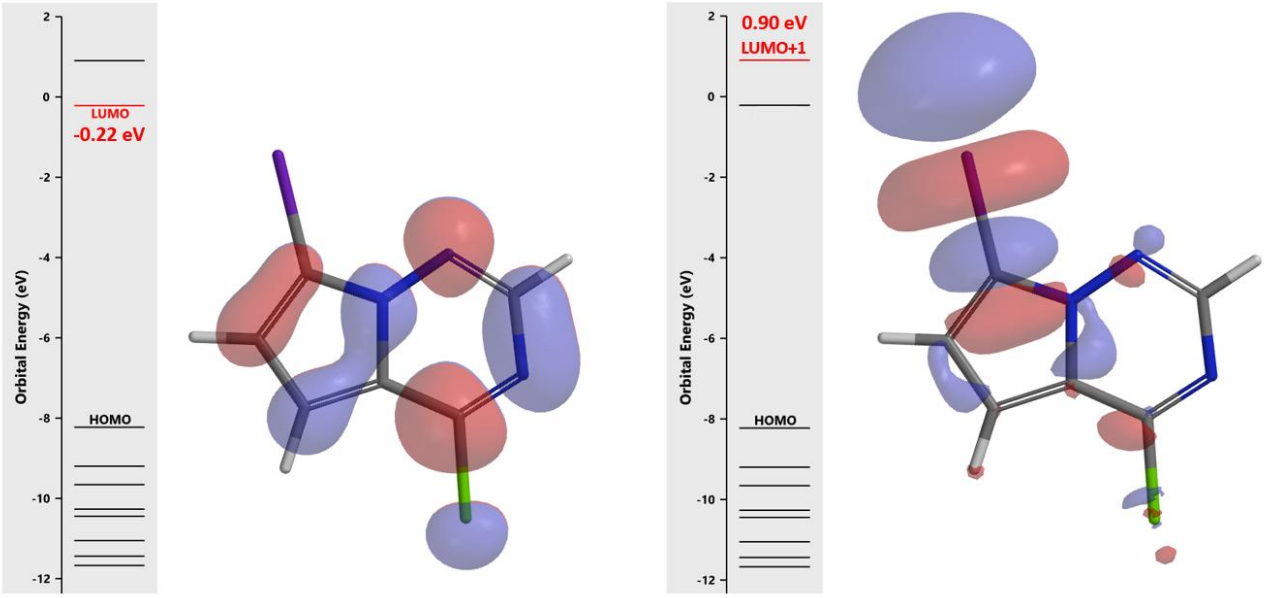


Figure 3. LUMO (left) and LUMO+1 (right) of pyrrolotriazine **I**

The prediction of C-Cl is more reactive than C-I towards OA is further supported with LUMO Map (Figure 4) and infrared (IR) stretching vibration frequencies of CX bonds. For substrate **I**, wavenumber calculated for C-Cl bond is 1191 cm^{-1} *vs* 1361 cm^{-1} for the C-I bond. The C-Cl bond is significantly weaker than the C-I bond. As such, results from LUMO and IR analyses are consistent, pointing towards a selective Buchwald amination at the C-Cl carbon.

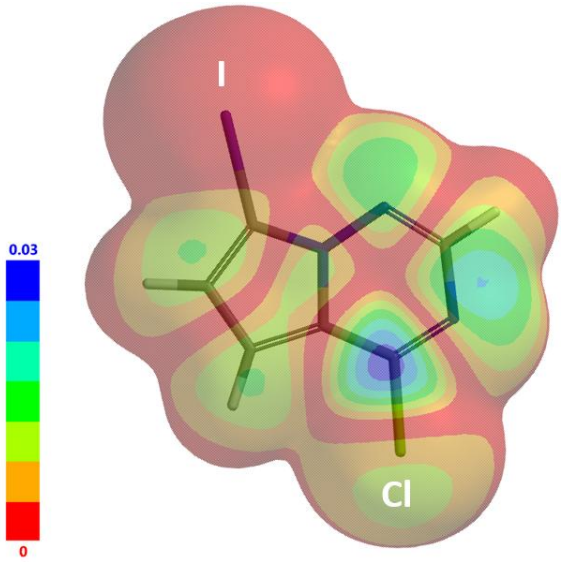
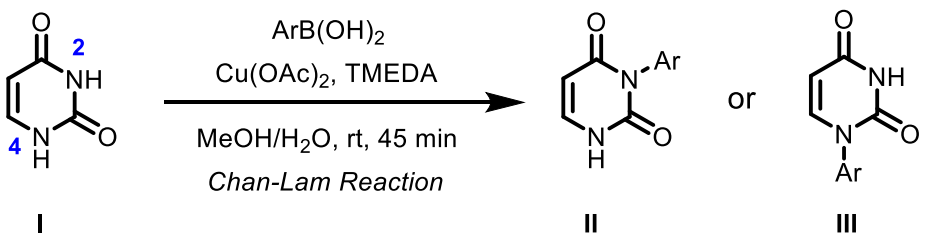


Figure 4. LUMO Map of substrate **I**

Question 3



Will the above Chan-Lam reaction form a single product or mixture of products? If it is a single product, which one, II or III? What is the QM parameter to calculate for analysis? Please provide QM calculation results and interpretation.

Answer (Qiuyue Wang)

Quantum Mechanical Organic Chemistry chapter 34[10] showed us how to predict regio-selectivity for Chan-Lam reaction. We'll find out whether N2-H or N4-H on uracil (**I**) is more acidic, and then examine HOMO of the deprotonated intermediate.

Electrostatic Potential Map of uracil (Figure 5, left) reveals that N4-H is more acidic than N2-H[4] (ESP at N4-H: 272.0 kJ/mol >> N2-H: 200.9 kJ/mol). N4-H will be selectively deprotonated under the basic conditions in Chan-Lam reaction. HOMO of the anionic intermediate (Figure 5, right) has HOMO lobe at N4, none at N2, enabling the prediction that Chan-Lam reaction of uracil (**I**) will occur preferentially at N-4 to provide a single product **III**.

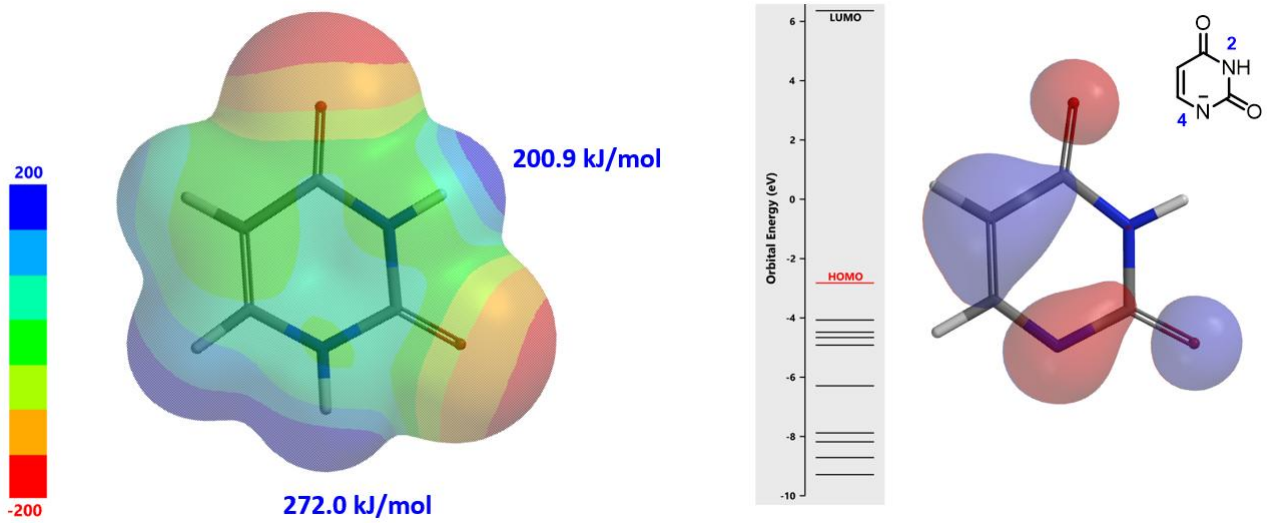
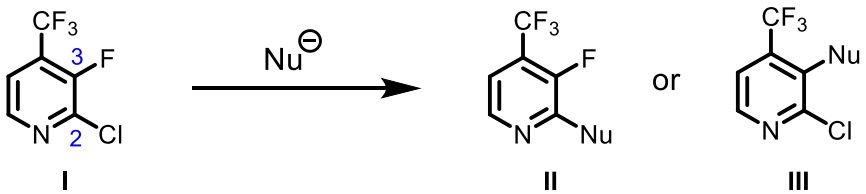


Figure 5. The ESP map of uracil (left) and HOMO of N4 deprotonated intermediate (right)

Question 4



I **II** **III**

Will the above SNAr displacement reaction form a single product or mixture of products? If it is a single product, which one, II or III? What are the QM parameters to calculate for analysis? Please provide QM calculation results and interpretation.

Answer (Zhong Zheng)

For nucleophilic substitution reaction, we'll look at unoccupied orbitals of pyridine **I** [6]. With LUMO, the lobe on C2 and C3 is mainly distributed between them, on the carbon-carbon bond (Figure 6, left). With LUMO+1, we have what we are looking for (Figure 6, right), lobes that are centered on the C-X carbons for further analysis.

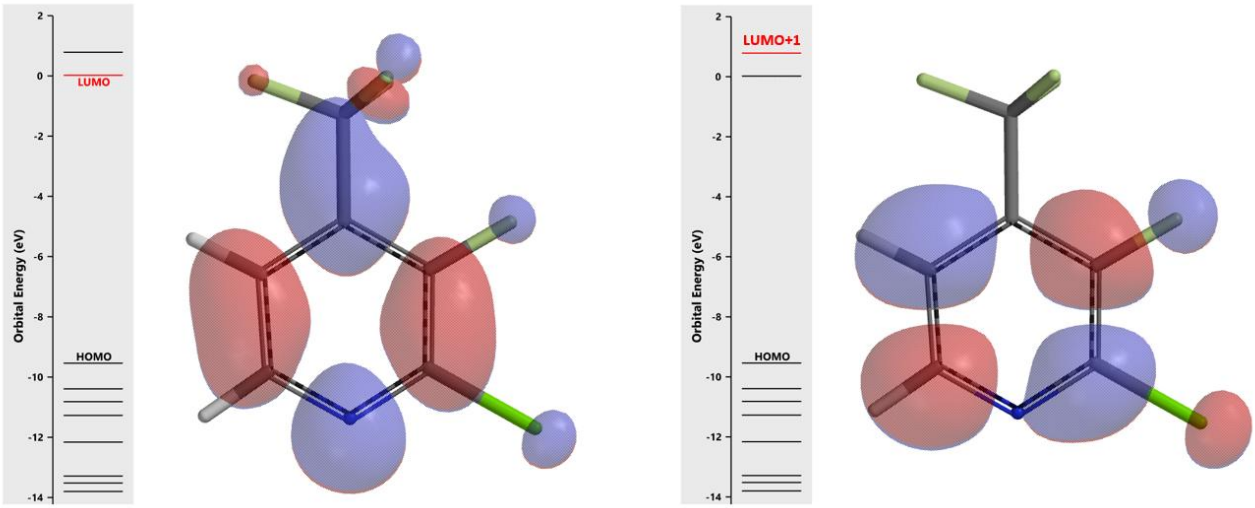


Figure 6. LUMO and LUMO+1 of compound **I**

LUMO+1 map (Figure 7, left) and LUMO+1/Electron density map (Figure 7, right) reveal that the LUMO lobe on C3-F carbon is more exposed and accessible than C2-Cl carbon for nucleophilic substitution reaction, suggestive of potential selective formation of product **III**.

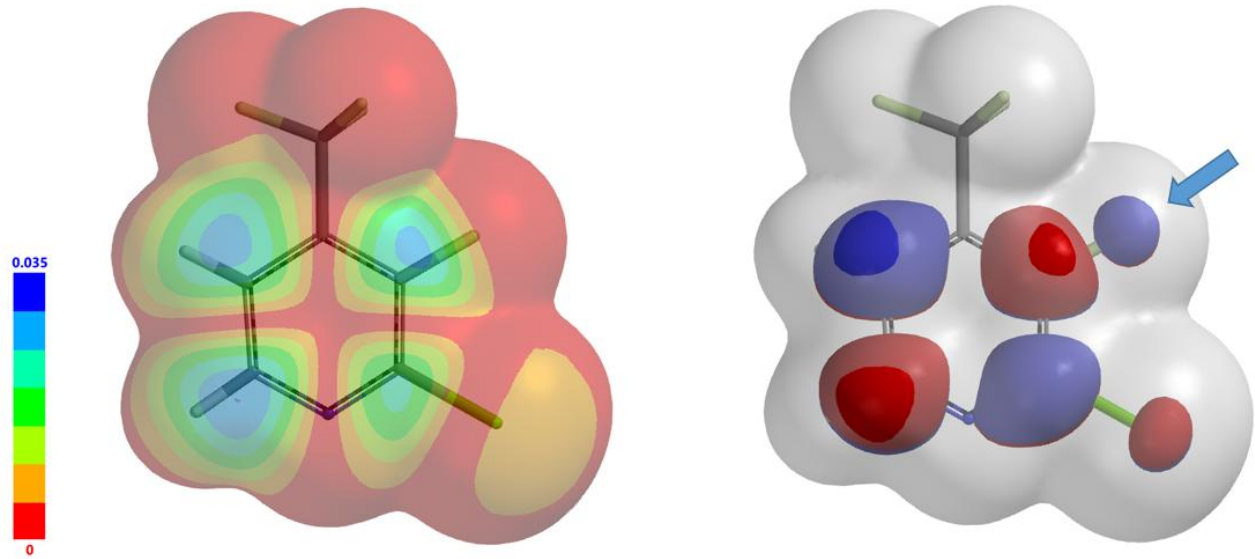


Figure 7. LUMO+1 map and LUMO+1/Electron Density map of compound **I**

Return to Table of Contents

References:

[1] Quantum Mechanics for Organic Chemists: An Experimentalist Approach. *https://wuxibiology.com/resource/quantum-mechanics-for-organic-chemists/*

[2] Spartan Tutorial and User's Guide. Accessible within Spartan software and from https://www.wavefun.com/spartan-documentation

[3] W.W.K.R. Mederski, M. Lefort, M. Germann, D. Kux, *Tetrahedron*, **1999**, *55*, 12757.

[4] Reference 1, chapter 3: Application of Electrostatic Potential Map in Acidity Estimation.

[5] WO2019/147782, 2019, A1 p191, 194; US2012/77814, 2012, A1 p25; WO2011/25940, 2011, A1 p. 92.

[6] Reference 1, chapter 1: Application of LUMO Analysis in Nucleophilic Reactions.

[7] Reference 1, chapter 5: Sequential Cross Coupling Reactions of Polyhalogenated Heterocycles.

[8] QM Chapter 39: Unraveling Divergence in Haloselectivity of SNAr and Cross Coupling vs Halogen-Metal Exchange Reactions. https://wuxibiology.com/unraveling-divergence-in-haloselectivity-of-snar-and-cross-coupling-vs-halogen-metal-exchange-reactions/

[9] L. Tao, Y. Yue, J. Zhang, S.Y. Chen, X.Q. Yu, *Helvetica Chimica Acta*, **2008**, *91*, 1008.

[10] Reference 1, chapter 34: QM Analyses of Regioselectivity in Chan-Lam Reaction.

[11] L.C. Campeau, Q. Chen, D. Gauvreau, M. Girardin, K. Belyk, P. Maligres, G. Zhou, C.Z. Gu, W. Zhang, L.S. Tan, P. D. O'Shea, *Org. Process Res. Dev.* **2016**, *20*, 1476.

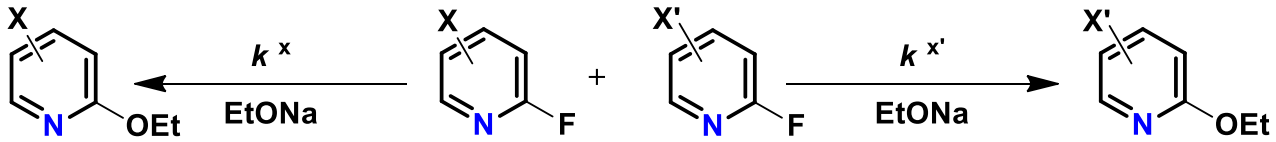
Chapter 44 Correlating Activation Energy with Relevant Orbital Energy

Ying Lu, Zhong Zheng, Yongsheng Chen, John S. Wai

In Chapter 25 "Assessing Reactivity with LUMO and HOMO Energy Gap", we discussed nucleophilicity and electrophilicity of substrate/reagent can be evaluated by their HOMO and LUMO energies, respectively, and the correlation between reactivity and LUMO and HOMO Energy Gap. In this chapter, we will analyze further such correlations between orbital energies and activation energies in nucleophilic aromatic substitution reactions.

SnAr Reaction of Substituted 2-Fluoropyridines

2-Fluoropyridine undergoes SnAr reaction with sodium ethoxide in EtOH about 250 times faster than 2-chloropyridine [1]. The activating effects of various substituents on the displacement of 2-fluoropyridine will be of great interest to medicinal and organic chemists. Schlosser *et al* [2] reported that by mixing equal amount of 2-fluoropyridine and its substituted analogs and reacting the mixture with sub-stoichiometric amount of sodium ethoxide in ethanol, monitoring the disappearance of the substrates and appearance of products, the relative reaction rate of each reaction substrate relative to 2-fluoropyridine and the relative activation free energy ΔΔG could be calculated (Table 1).



2a - 2i **1a - 1i** **2a - 2i**

Substrate	Substituent	k^{rel}	k_f^{rel}	ΔΔG (kcal/mol)	LUMO (eV)	LUMO+1 (eV)
1a	6-EtO	0.078	0.078	+1.5	1.43	2.05
1b	5-F	0.67	0.67	+0.24	0.70	1.68
1c	H	1.0	1.0	0.00	1.06	1.69
1d	6-F	56	28	-2.0	0.92	1.56
1e	4-F	86	43	-2.2	1.03	1.46
1f	3-F	50	50	-2.3	0.90	1.51
1g	4-Cl	69	69	-2.5	0.66	1.29
1h	3-Cl	75	75	-2.6	0.68	1.31
1i	5-CF_3	3100	3100	-4.8	0.54	0.89

Table 1. Fluoride/ethoxide displacement rates relative to 2-fluoropyridine (**1c**), k^{rel} = relative reaction rate, k_f^{rel} = relative reaction rate after statistical correction, differences in the free activation energies, and LUMO/LUMO+1 energies of various fluoropyridines

LUMO and LUMO+1 energies of the fluoropyridine substrates, calculated with DFT ωB97X-D/6-31G*, are tabulated. Linear regression analysis on the LUMO orbital energy level *versus* the relative activation energy obtained from the experiment plot shows poor linear correlation, with a coefficient of determination (R^2) of 0.5379 (Figure 1).

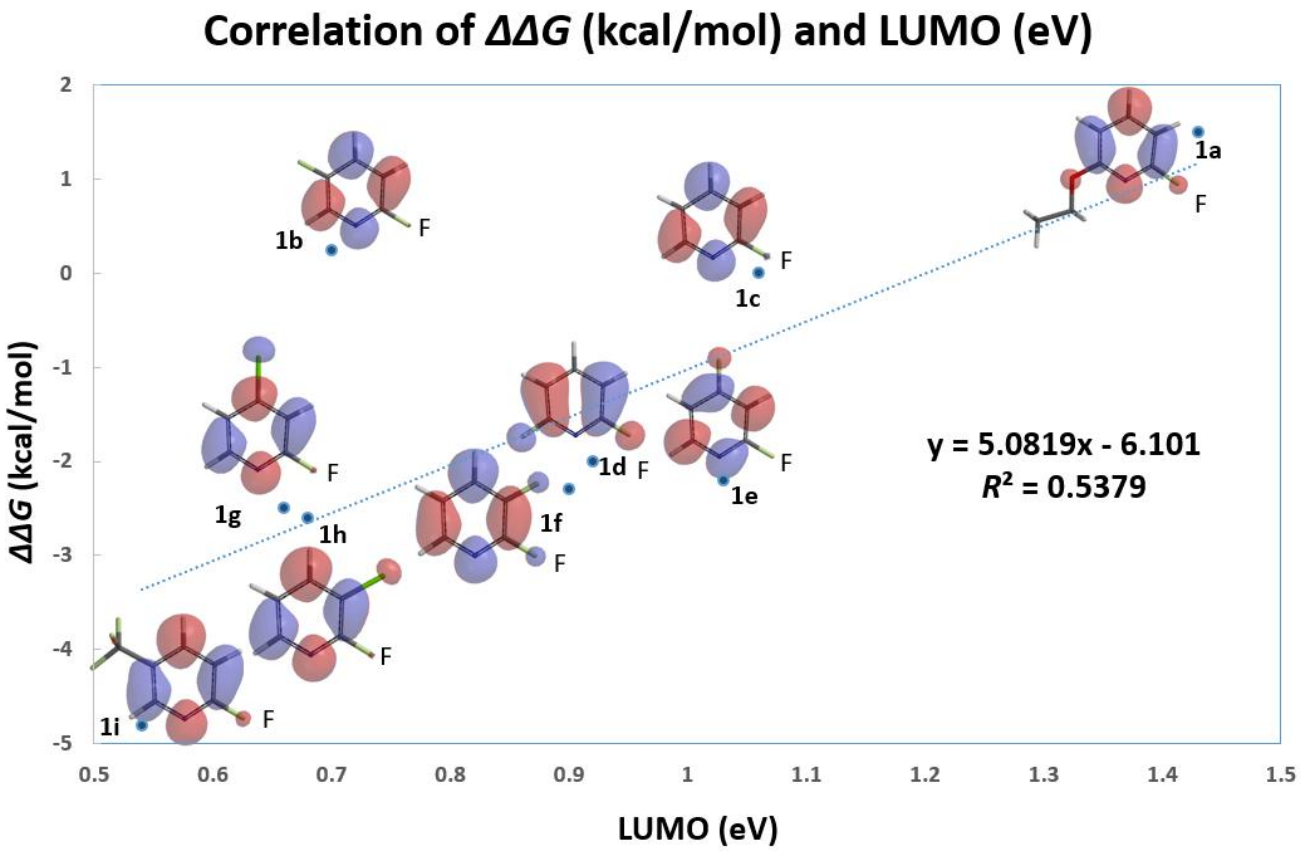


Figure 1. Correlation analysis of relative activation energy and LUMO energy level of various substituted 2-fluoropyridines

We noticed little to none LUMO lobe centered on the C-F carbon with these 2-fluoropyridines, indicating that the associated LUMO energy data may not be relevant for correlation analysis of this set of fluoride displacement reactions. Galabov *et al.* also reported a similar weak correlation in electrophilic aromatic substitution reactions [3].

On the other hand, with LUMO+1, there are significant lobes centered on C-F carbon of this set of 2-fluoropyridines (Figure 2). The linear correlation of their LUMO+1 energies and the activation energies is excellent, with a coefficient of determination (R^2) of 0.9175, significant improvement as compared to the use of LUMO. This shows that it is necessary to select only the relevant frontier orbitals, and the associated orbital energy values, for reactivity analyses. These are crucial prerequisites for successful Quantum Mechanics-Machine Learning of chemical reactivity and regioselectivity.

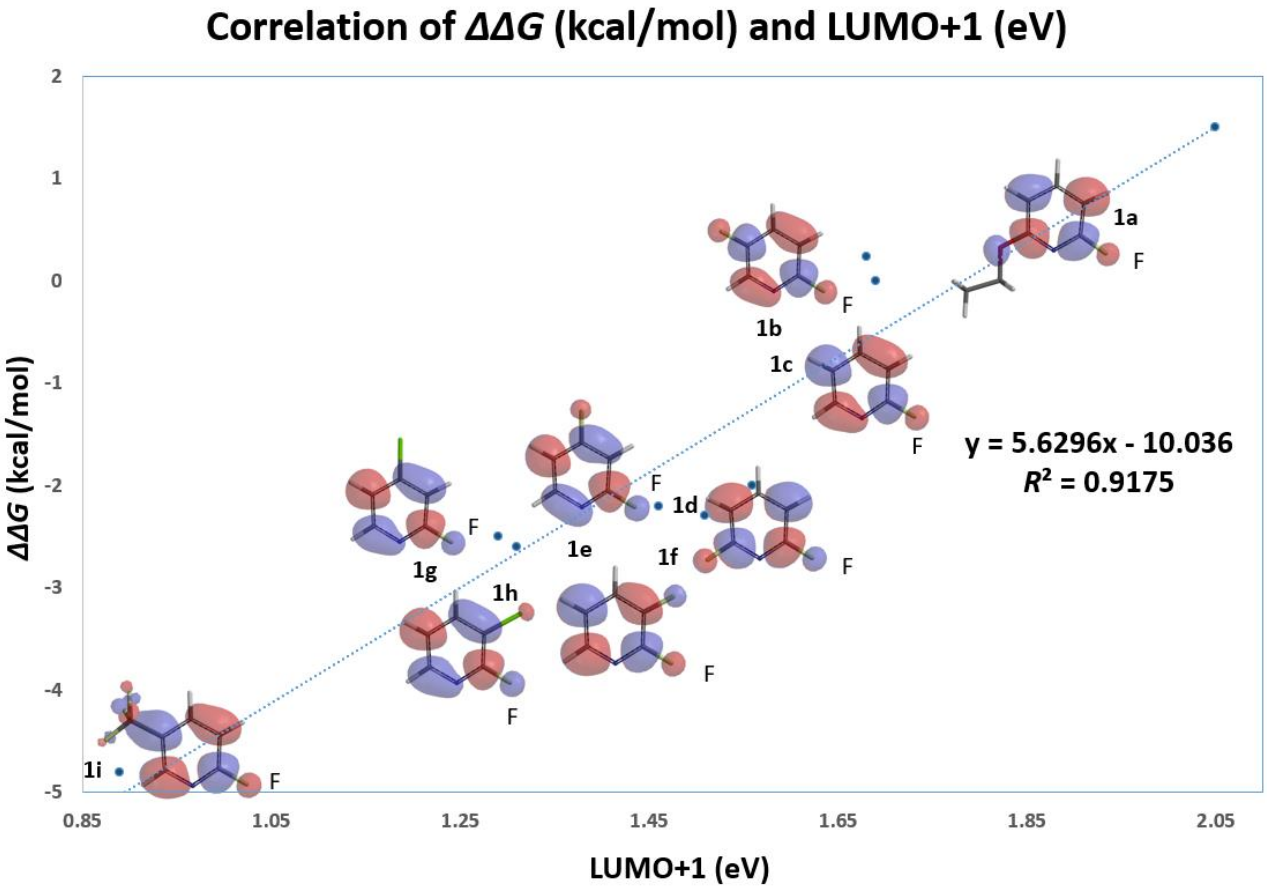


Figure 2. Correlation analysis of relative activation energy and LUMO+1 energy level of various substituted 2-fluoropyridines

This lesson could be further exemplified with the SnAr reaction of various substituted 2-chloropyridines. Excellent linear correlation (R^2 of 0.9995) of reactivity with relevant LUMO or LUMO+1 orbital energy, chosen with an orbital lobe centered on the C-Cl carbon (Figure 3, Table 2). Relevant orbitals for correlation of a set of electrophiles with reactivity could be a mix of LUMO and LUMO+n. We need to choose the appropriate ones first.

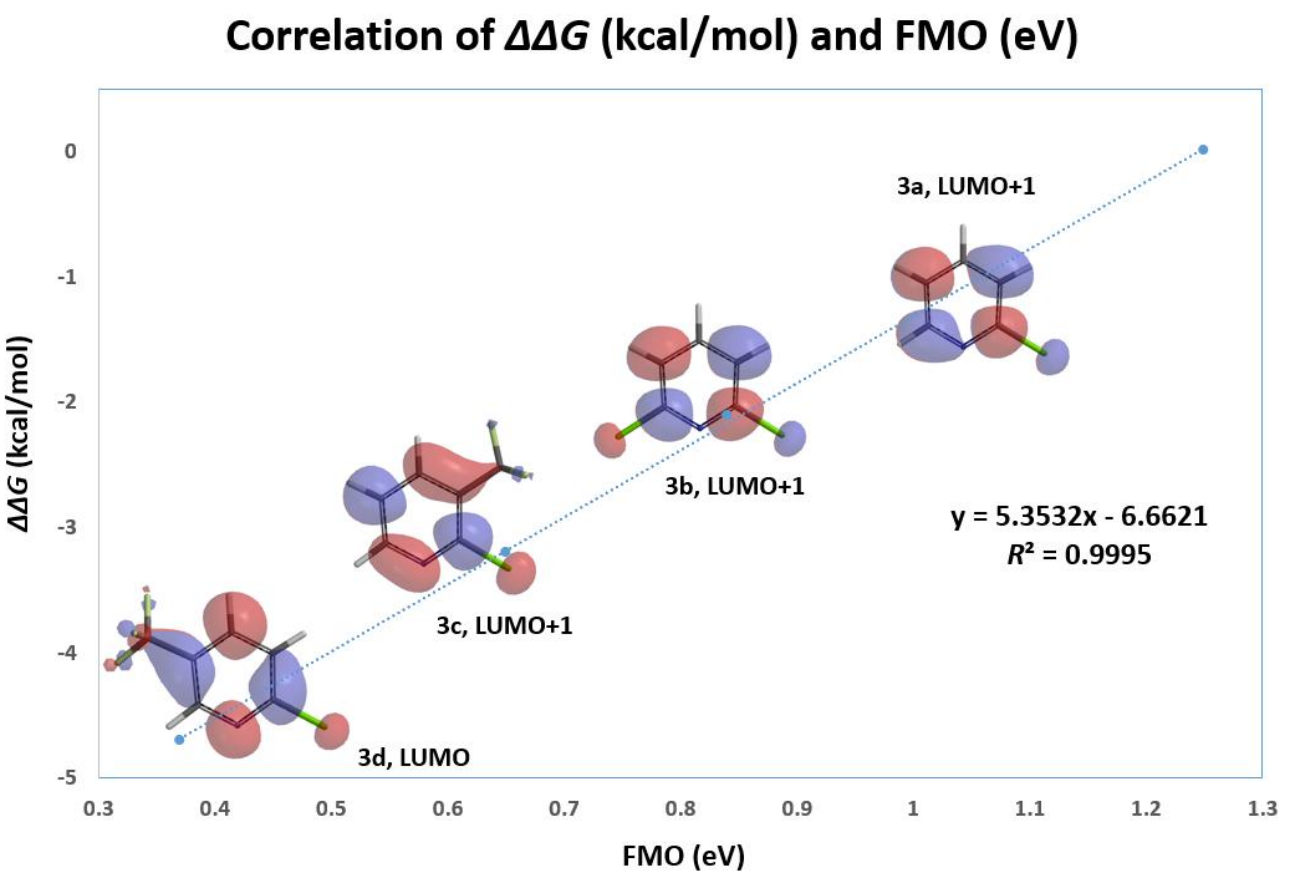
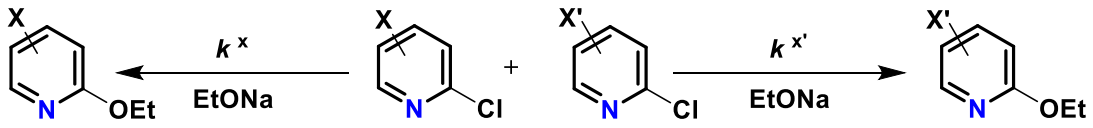


Figure 3. Correlation analysis of relative activation energy and FMO energy level of various substituted 2-chloropyridines



3a -3d

Substrate	Substituent	k_f^{rel}	ΔΔG (kcal/mol)	LUMO (eV)	LUMO+1 (eV)
3a	H	1	0.00	0.92	1.25
3b	6-Cl	65	-2.1	0.56	0.84
3c	3-CF_3	220	-3.2	0.41	0.65
3d	5-CF_3	2800	-4.7	0.37	0.58

Table 2. Chloride/ethoxide displacement rates relative to 2-chloropyridine (2a), k^{rel} = relative reaction rate, differences in the free activation energies, and LUMO/LUMO+1 energies of various chloropyridines

The resultant linear regression equations obtained from above correlations could be used to infer reactivity of other 2-chloro/fluoropyridine analogs (Figures 2 & 3). For example, with 2,6-dichloro-3-trifluoromethylpyridine (Figure 4), LUMO has a lobe centered on C6-Cl carbon, while LUMO+1 has a lobe centered on C2-Cl carbon. Substituting the energy value of LUMO (0.11 eV) and LUMO+1 (0.26 eV) into the linear regression equation in Figure 3 provided ΔΔG

of -6.07 kcal/mol for reaction at C-6 *vs* -5.27 kcal/mol for reaction at C2, respectively (the negative sign indicates their ΔG are smaller than that for 2-chloropyridine). As such, chloride displacement at C6 will be much faster than at C2, consistent with experimental numbers of ΔΔG -6.5 kcal/mol and -5.4 kcal/mol, respectively, reported [2]. Steric hindrance effect of trifluoromethyl group is anticipated to slow down nucleophilic attack at C2 further, resulting in C6 displacement as the dominant reaction path.

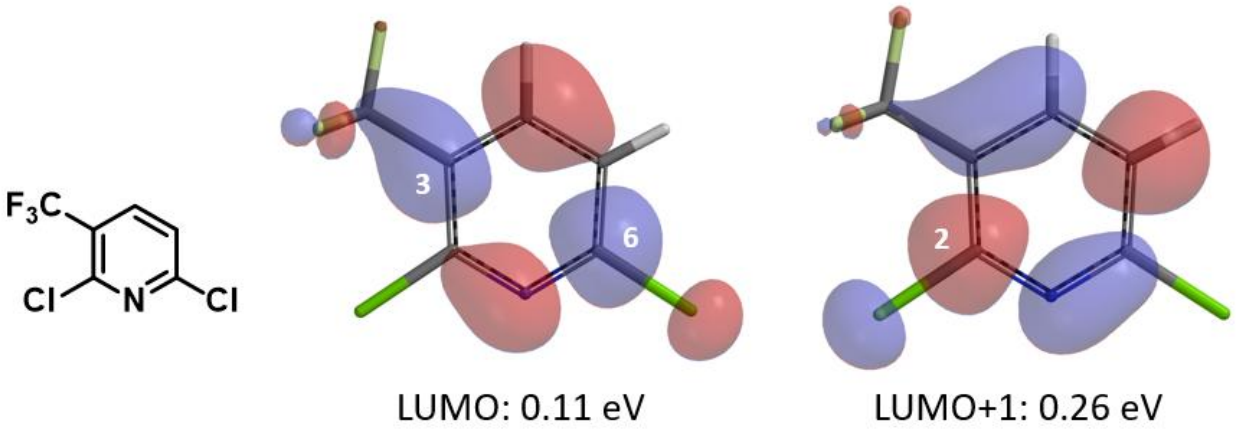


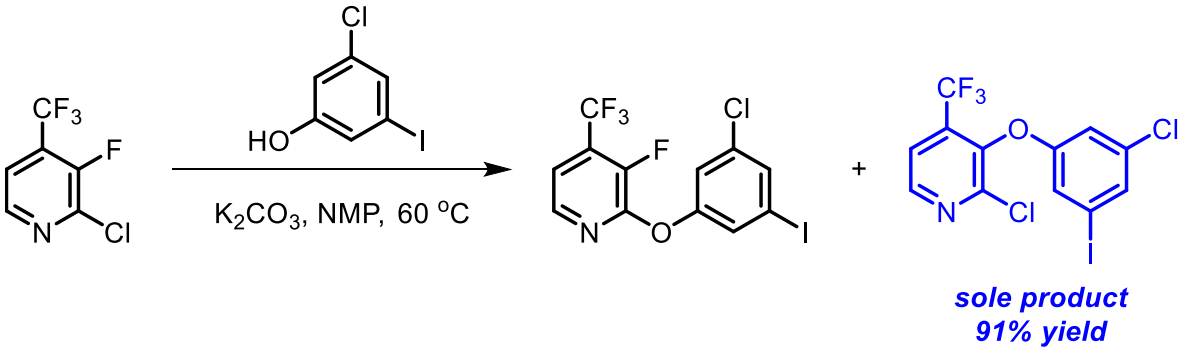
Figure 4. LUMO/LUMO+1 lobe of 2,6-dichloro-5-trifluoromethylpyridine

Conclusion

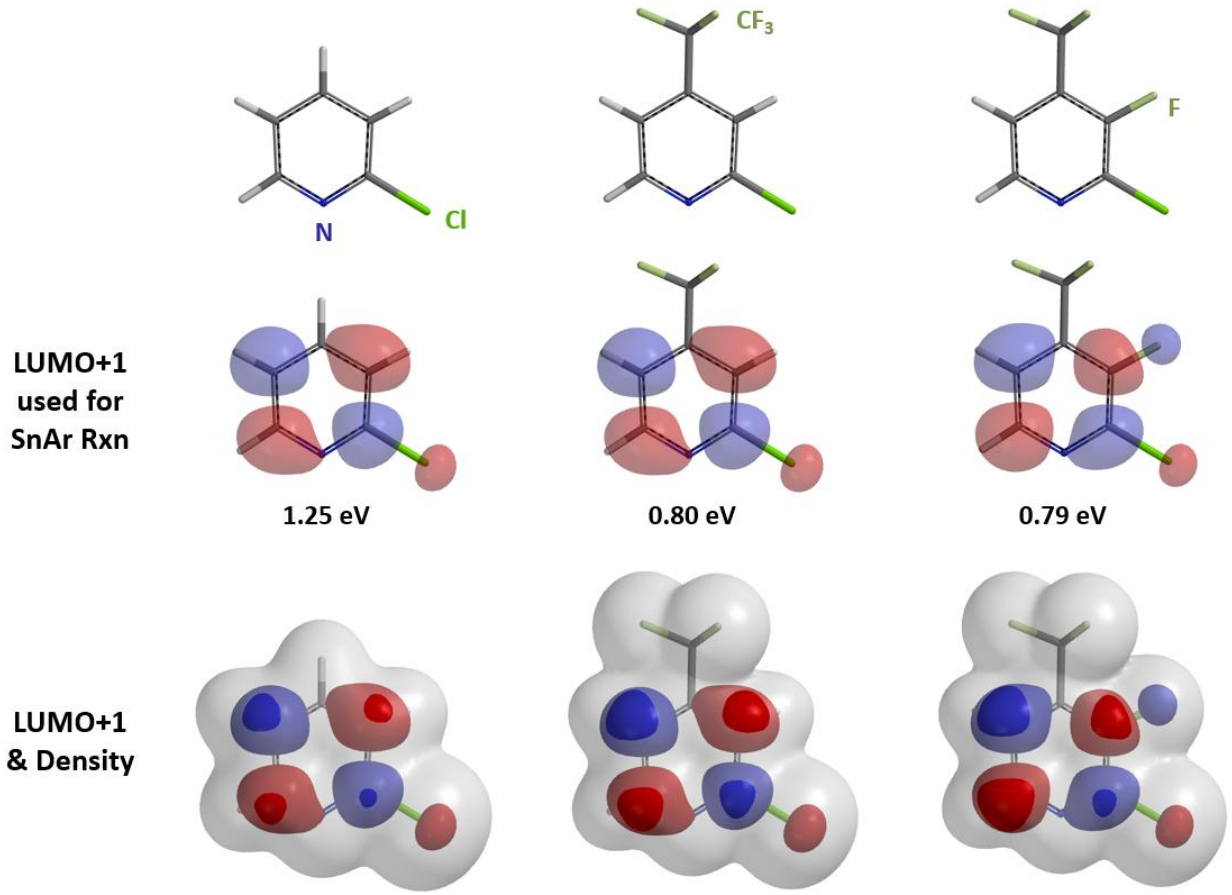
For nucleophilic aromatic substitution reactions, choosing the relevant frontier orbitals is critical for proper correlation of reactivity with LUMO energy. This is an important prerequisite for successful Quantum Mechanics-Machine Learning of chemical reactivity and regioselectivity [4].

Building on What We Just Learned

In the textbook "Heterocyclic Chemistry"[5], the authors taught us that nucleophilic substitution of 3-halopyridine is too slow for practical use. However, in the synthesis of Doravirine, a non-nucleoside reverse transcriptase inhibitor (NNRTI) for the treatment of HIV-1 infection, reaction of 3-chloro-5-iodophenol with 2-chloro-3-fluoro-4-trifluoromethyl-pyridine provides selectively the C3 substituted product [6], no displacement of the "potentially" more reactive C2 chloride is observed.



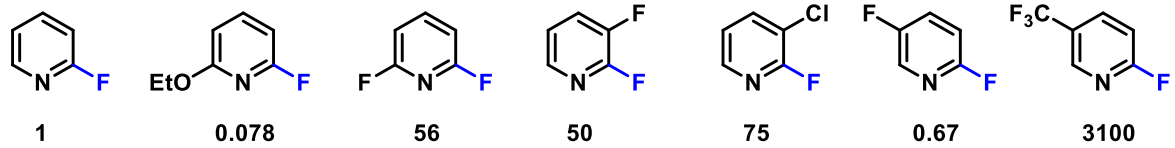
Shown below are LUMO+1 of three 2-chloropyridine analogs and their overlays with respective electron cloud density maps. Note the effects of sequential introduction of trifluoromethyl and fluoro groups on LUMO+1 energy and lobe distribution, in particular at C2 and C3 carbons. It should become obvious why SnAr of 2-chloro-3-fluoro-4-trifluoromethylpyridine proceeds selectively at C3, not at C2.



Return to Table of Contents

References:

[1] J. A. Joule & K. Mills. *Heterocyclic Chemistry 5th Ed.* Chichester, West Sussex, UK: Blackwell Publishing Ltd., **2010**; pp 25.



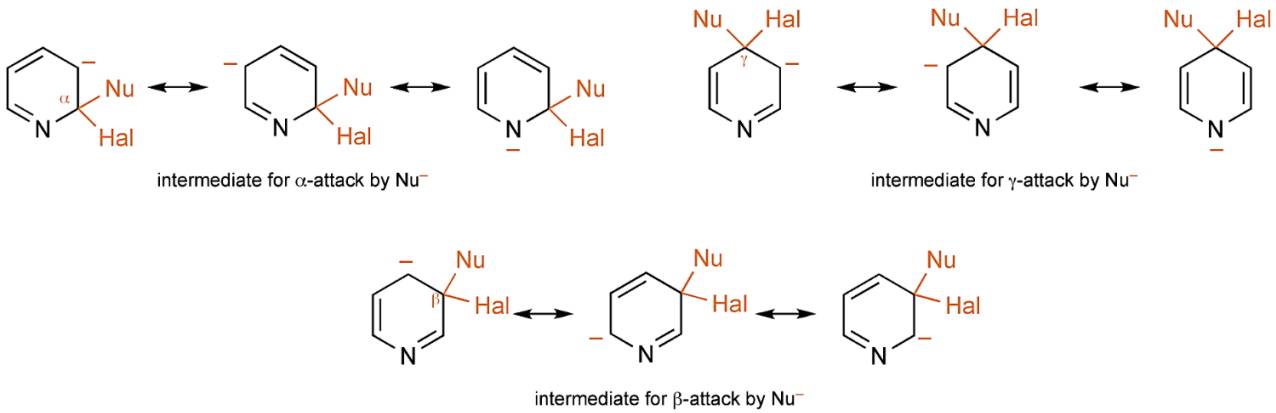
Relative rates of displacement of pyridine-2-fluoride by ethoxide in ethanol

[2] M. Schlosser, T. Rausis, *Helv. Chim. Acta.* **2005**, *88*, 1240.

[3] G. Koleva, B. Galabov, J. I. Wu, H.F. Schaefer III, P. R. Schleyer, *J. Am. Chem. Soc.* **2009**, *131*, 14722.

[4] (a) T. Stuyver, C. W. Coley, J. Chem. Phys. 2022, 156, 084104. (b) Y. Guan, C. W. Coley, H. Wu, D. Ranasinghe, E. Heid, T. J. Struble, L. Pattanaik, W. H. Green, K. F. Jensen, *Chem. Sci.* **2021**, *12*, 2198.

[5] J. A. Joule & K. Mills. *Heterocyclic Chemistry 5th Ed.* Chichester, West Sussex, UK: Blackwell Publishing Ltd., **2010**; pp 118. *Such substitutions follow the same mechanistic route as the displacement of halide from 2- and 4-halonitro-benzenes, i.e., the nucleophile first adds and then the halide departs. By analogy with the benzenoid situation, the addition is facilitated by: (i) the electron - deficiency at α- and γ-carbons, further increased by the halogen substituent, and (ii) the ability of the heteroatom to accommodate negative charge in the intermediate thus produced. A comparison of the three possible intermediates makes it immediately plain that this latter is not available for attack at a* β*-position, and thus* β *nucleophilic displacements are very much slower – for practical purposes they do not occur.*

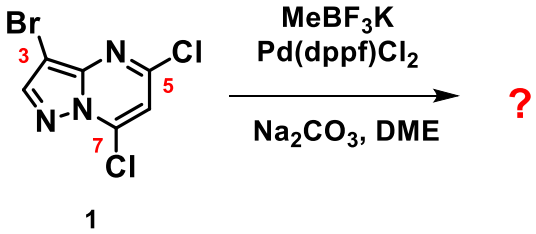


[6] L. C. Campeau, Q. H. Chen, D. Gauvreau, M. Girardin, K. Belyk, P. Maligres, G. Y. Zhou, C. Z. Gu, W. Zhang, L. S. Tan, P. D. O'Shea, *J. Org. Process Res. Dev.* **2016**, *20*, 1476.

Chapter 45 Unusual Reactivities of Polyhalogenated Heteroaromatic Substrates are Predictable

Zhong Zheng, Qiuyue Wang, Yongsheng Chen, John S. Wai

In Chapter 5 "Sequential Cross Coupling Reactions of Polyhalogenated Heterocycles", we discussed the use of QM calculated LUMO/LUMO map and IR carbon-halogen bond stretching wavenumbers to predict the order of haloselectivity for a sequence of palladium-catalyzed cross coupling reactions on monocyclic 2,4-dichloro-5-bromopyrimidine. We left a related question on predicting haloselectivity of palladium catalyzed cross coupling reaction for the bicyclic 3-bromo-5,7-dichloropyrazolopyrimidine (**1**). What will be the order of haloselectivity (Figure 1)?



1

Figure 1. Suzuki Cross Coupling reaction of compound **1**

Shown in Figure 2 are the calculated LUMO and IR stretching vibration wavenumbers of the three carbon-halogen bonds of compound **1**.

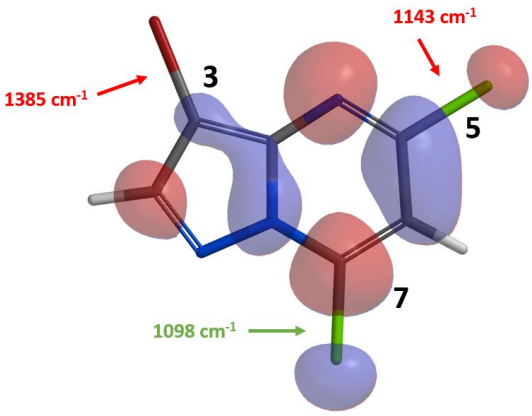


Figure 2. LUMO of compound **1** and IR stretching vibration wavenumbers of the C-X bonds

There is little LUMO lobe on the C3-Br carbon relative to the C5-Cl and C7-Cl carbons on compound **1**. IR calculation shows that the stretching vibration wavenumbers of the C7-Cl and C5-Cl bonds are 1098 and 1143 cm^{-1}, respectively, suggesting that the C7-Cl bond is weaker and will be more reactive toward Pd(0) oxidative addition. Our experimental result is

consistent with the above prospective analysis, Suzuki cross coupling reaction of compound **1** with potassium methyl-trifluoroborate provides compound **2** selectively [1]. An obvious follow-up question is haloselectivity of intermediate **2**.

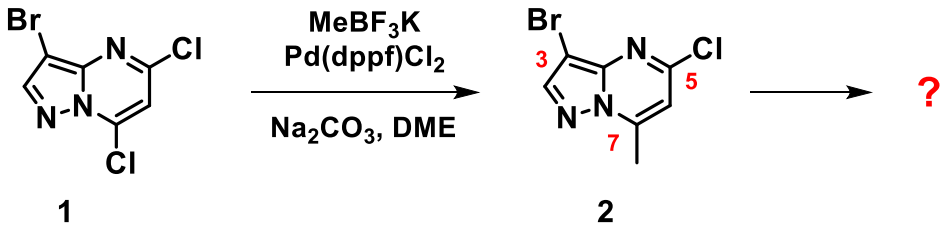


Figure 3. Suzuki Cross Coupling reaction of intermediate **1**

With intermediate **2**, there is little LUMO lobe on the C3-Br carbon relative to the C5-Cl one. IR calculation shows that the stretching vibration wavenumber of the C3-Br and C5-Cl bonds are 1168 and 1147 cm^{-1}, respectively, suggesting that the C5-Cl bond is more reactive toward oxidative addition. Both LUMO and IR analyses are suggestive of selective reaction at C5-Cl, consistent with experimental results. Here we can see that the haloselectivity of metal-catalyzed coupling reactions of compounds **1** and **2** diverges from the general reactivity we learned from general organic chemistry for C-X bonds, i.e., X = I > Br ~ OTf >> Cl >> F, yet fully accountable with relevant QM calculated parameters.

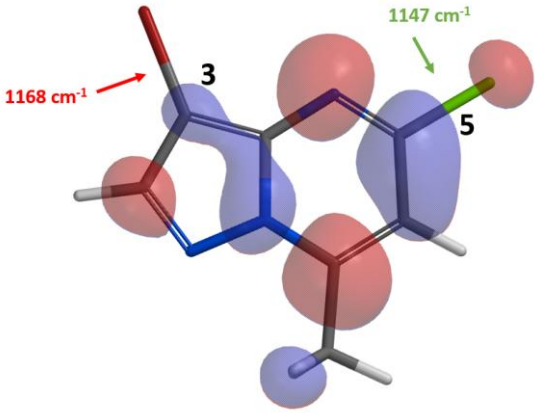


Figure 4. LUMO of intermediate **2** and IR stretching vibration wavenumbers of the C-X bonds

How about Polyhalogenated Bicyclic Heterocycles with Iodo group?

From the previous example, we learned that for certain bicyclic heterocyclic systems, C-Cl bonds could undergo oxidative addition with Pd(0) selectively over C-Br bonds. What have we learned about iodo containing polyhalogenated bicyclic systems? Below are two examples.

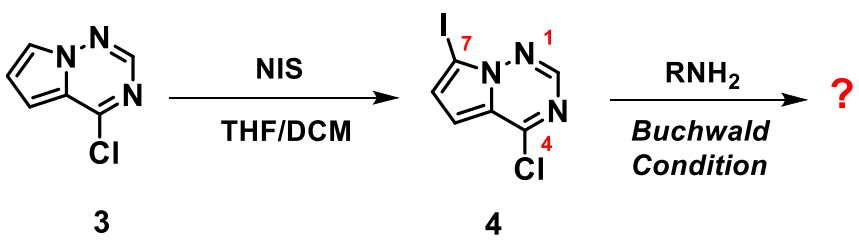


Figure 5. Buchwald reaction of aromatic heterocyclic compound **4**

Shown in Figure 5 is 4-chloro-7-iodopyrrolotriazine **4**, prepared by NIS iodination of commercially available chloropyrrolotriazine **3**. Intuitively, we anticipated the C7-iodide bond to be more susceptible to oxidative addition to give selectively a C-7 amination product. However, only C-4 amination product was obtained under various Buchwald amination conditions. With compound **4**, there is little LUMO lobe on the C7-I carbon relative to the C4-Cl one. IR calculation shows that the stretching vibration wavenumber of the C7-I and C4-Cl bonds are 873 and 823 cm^{-1}, respectively, suggesting that the C4-Cl bond to be more reactive toward oxidative addition. Both LUMO and IR analyses are suggestive of selective reaction at C4-Cl, consistent with experimental observations. We learned to integrate these analyses to our retrosynthetic planning.

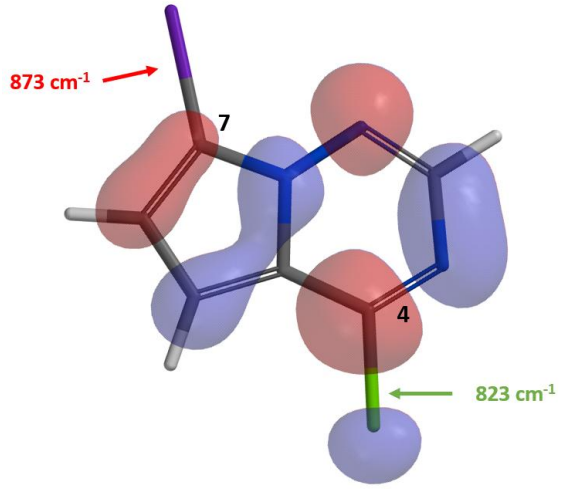


Figure 6. LUMO of compound **4** and IR stretching vibration wavenumbers of the C-X bonds

Second example comes from an OPRD paper published by Pfizer[2] (Figure 7). For nucleophilic aromatic substitution of 2,4-dichloro-5-iodopyrrolopyrimidine **5** with a secondary alcohol compound, we also calculated for its LUMO and C-X bond stretching wavenumbers. As shown in Figure 8, there are LUMO lobes on all three carbon-halogen bonds, yet with the lobe on C4-Cl carbon significantly larger than the ones on C5-I and C2-Cl carbons. IR calculation shows that the stretching vibration wavenumber of the C2-Cl, C4-Cl, and C5-I bonds are 981, 837, and 924 cm^{-1}, respectively, suggesting that the C4-Cl bond is weakest and most reactive toward nucleophilic substitution, accounting for the high regioselectivity reported.

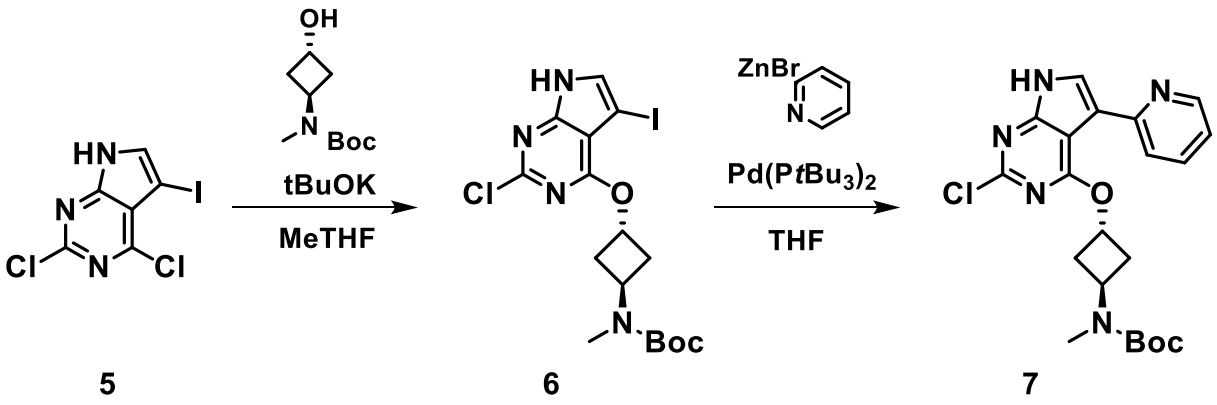


Figure 7. SnAr and Negishi reactions of polyhalogenated pyrrolopyrimidine **5**

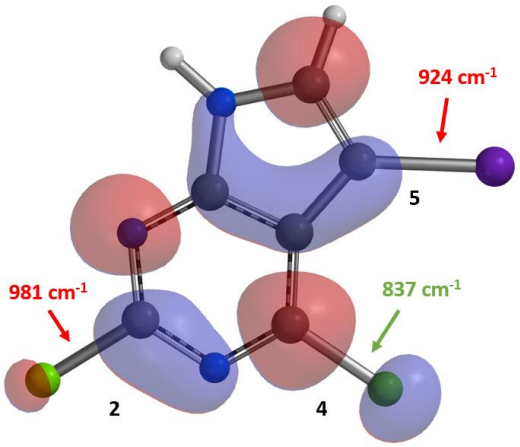


Figure 8. LUMO of compound **5** and IR stretching vibration wavenumbers of the C-X bonds

Conclusion

With polyhalogenated heterocycles, the C-X bonds' relative reactivity towards metal-catalyzed coupling reactions does not necessarily follow the general reactivity order of X = I > Br ~ OTf >> Cl >> F. In the case of bicyclic heteroaromatic systems, if the C-Cl bond is on the electron-deficient part of the system, while the C-I or C-Br bonds are on the electron-rich part, we shall calculate for LUMO and IR C-X bonds' stretching vibration wavenumbers to sequence properly the order of reactions. These intrinsic changes in halo reactivity are predictable.

Building on What We Just Learned

How about polyhalogenated tricyclic heteroaromatic substrates? Let's analyze 5-bromo-9-chloropyridopyrropyrimidine (**8**). Based on the calculated LUMO and C-X bond stretching wavenumbers shown in Figure 9, will palladium-catalyzed reactions be selective on the C-Br or C-Cl bond? Which factor controls the selectivity observed for this substrate [3]?

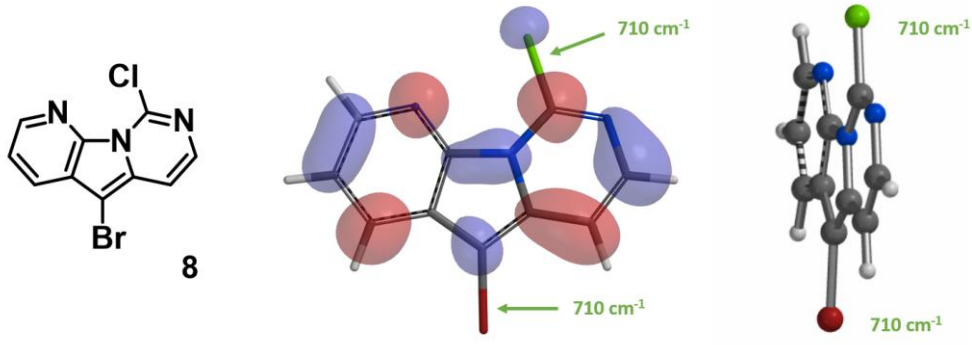


Figure 9. LUMO of compound **8** and IR stretching vibration wavenumbers of the C-X bonds

Return to Table of Contents

References:

[1] Z. Xu, Y. Lou, Fused heterocyclic compound, preparation method therefor, pharmaceutical composition, and uses therefore. US 20160244432 A1, August 25, 2016, page 28 on related Suzuki reaction with Compound **1**.

[2] Y. Tao, N.F. Keene, K.E. Wiglesworth, B. Sitter, J.C. McWilliams, *Org. Process Res. Dev.* **2019**, *23*, 382.

[3] A. Baeza, J. Mendiola, C. Burgos, J. Alvarez-Builla, J. J. Vaquero, *Eur. J. Org. Chem*. **2010**, *29*, 5607.

Chapter 46 Dichotomy in Regioselectivity of Electrophilic Reaction of 3-Acetyl-5-hydroxylindole

Jun Liu, Jianhui Qiao, Qiuyue Wang, Liting Dong, Yongsheng Chen, John S. Wai

Bromination of 3-acetyl-5-hydroxyindole **1** provided selectively only the C6 brominated product **2**, while Mannich reaction of it generated only the C4 substitution product **3** [1]. How to account for this dichotomy?

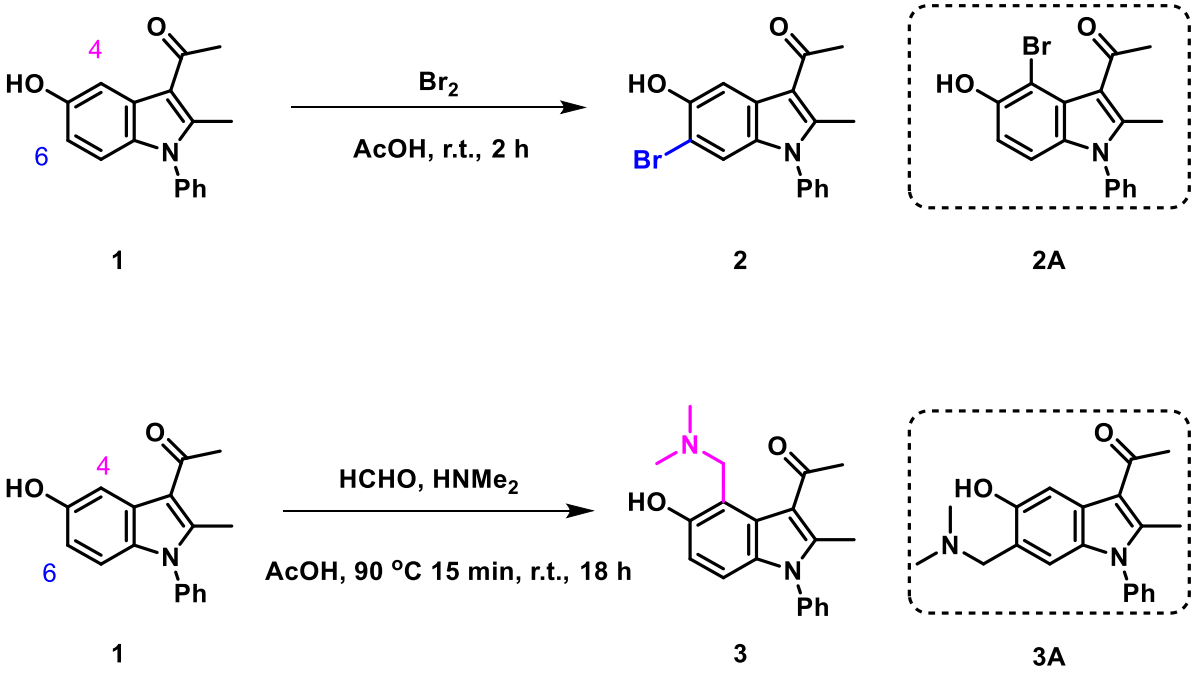


Figure 1. Bromination and Mannich reaction of 3-acetyl-5-hydroxyindole **1**

QM Analysis of the Bromination

In chapter 2 "Application of HOMO Analyses in Electrophilic Reaction", we used the comparative difference in size of **HOMO** lobes and calculated ^{13}C NMR for predicting regioselectivity of halogenation. As shown in Figure 2, energy gap between HOMO (-7.46 eV) and HOMO-1 (-7.57 eV) of substrate **1** is small, and the HOMO lobe at C4 is similar in size as the HOMO-1 lobe at C6, suggesting the reaction will provide a mixture of C4 and C6 products. Calculated **^{13}C NMR** chemical shift value is 104 ppm for C4 *vs* 111 pm for C6, suggesting that C4 bromide will be the major product. Obviously, the above QM parameters derived from equilibrium geometry of indole **1** *fail* to account for the experimental result.

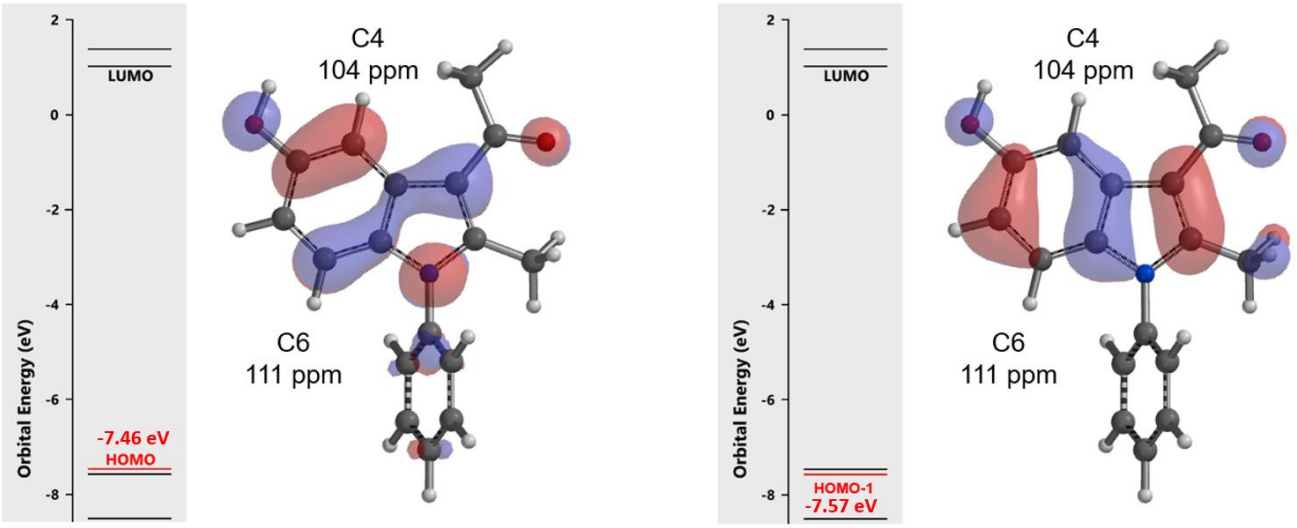
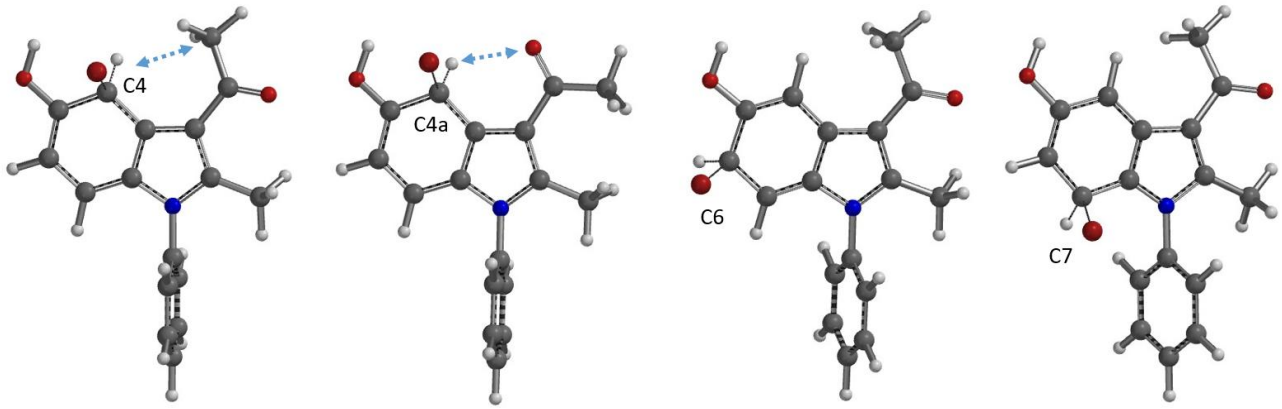


Figure 2. HOMO (left), HOMO-1 (right), calculated C4 & C6 ^{13}C NMR shifts of indole **1**

In Chapter 17 "Predicting Regioselectivity of Electrophilic Halogenation Reactions", we discussed the use of QM calculated **Equilibrium Geometry** of Wheland intermediates' relative energy difference for predicting regioselectivity. The calculated relative energies of the **C4**, **C4a**, and **C6** Wheland intermediates are significantly lower than the **C7** one, with the **C4a** one slightly lower than the **C6**, suggesting the bromination will provide a mixture of **C4** and **C6** substitution products, with **C4** bromide being the major one. This well-established model for electrophilic aromatic reaction *fail* to account for the experimental result as well.



Relative Energy (kcal/mol)	C4	C4a	C6	C7
Nonpolar solvent	1.61	0.00	0.71	14.28
Polar solvent	1.37	0.00	0.46	13.74

Figure 3. Equilibrium geometry and relative energy of bromo Wheland intermediates **C4**, **C4a**, **C6**, and **C7**

Why all these basic methods, differences in relative size of HOMO lobes, in ^{13}C NMR shift, difference in relative energy of the hypothetical Wheland intermediates, fail to predict regioselectivity of bromination of indole **1**? Shown in Figure 4 is a general mechanism for electrophilic aromatic substitution. This involves interaction of positively charged electrophile (E^+) with the aromatic substrate to form the positively charged Wheland intermediate [2], followed by loss of proton (H^+) or rearomatization to provide the bromination product. Formation of the Wheland intermediate is assumed to be the rate-determining step;

the subsequent rearomatization is presumed to be a much faster process. We reasoned that with indole **1** steric hindrance may have slowed down the rearomatization step, rendering it the rate limiting step of the reaction. The established prediction methods are not applicable for indole **1**. The question becomes, how to establish a more relevant model for this substrate?

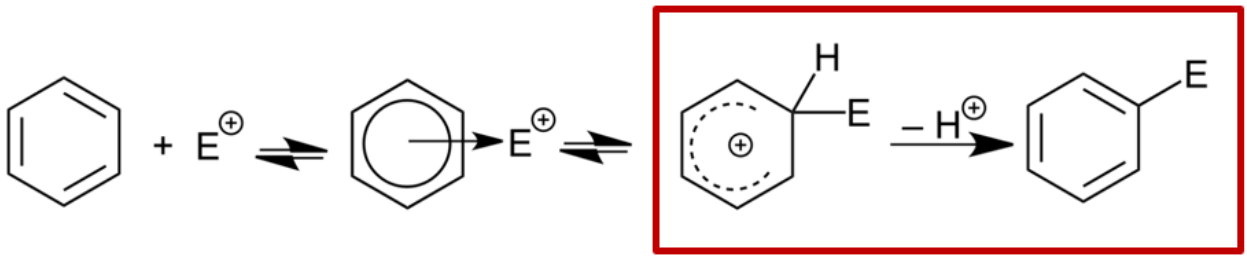


Figure 4. General mechanism for electrophilic aromatic substitution[2]

Since bromine was used as the brominating reagent in the reaction, Br_2 will be the most appropriate for use to model its reaction with indole **1**. Shown in Figure 5 is the calculated energy plot for C4 bromination. From Br_2 and indole **1** (**Reactants**) to transition state 1 (**TS1**) towards formation of the transient Wheland Complex like structure (**WC**) requires an activation energy of **4.92 kcal/mol**. The distal bromine then interacts with both the C-5 hydroxy group and the reaction center's sp^3 CHBr hydrogen to form a unique hydrogen bond stabilized Bromide-Wheland Complex (**Bromide-WC**), which is 4.05 kcal/mol lower in energy than the **Reactants**. The bromide in the 6-membered hydrogen bond ring of the **Bromide-WC** could deprotonate the sp^3 CHBr hydrogen,[3] form HBr, and rearomatize to provide C4 bromination product. However, in transition state 2 (**TS2**), the sp^3 CHBr bromide faces steric hindrance from the 3-acetyl group, as indicated with the calculated energy barrier of **11.59 kcal/mol**, higher than the 8.92 kcal/mol needed to reverse back to the **Reactants**. This is likely to be the case as no C4 bromination was observed.

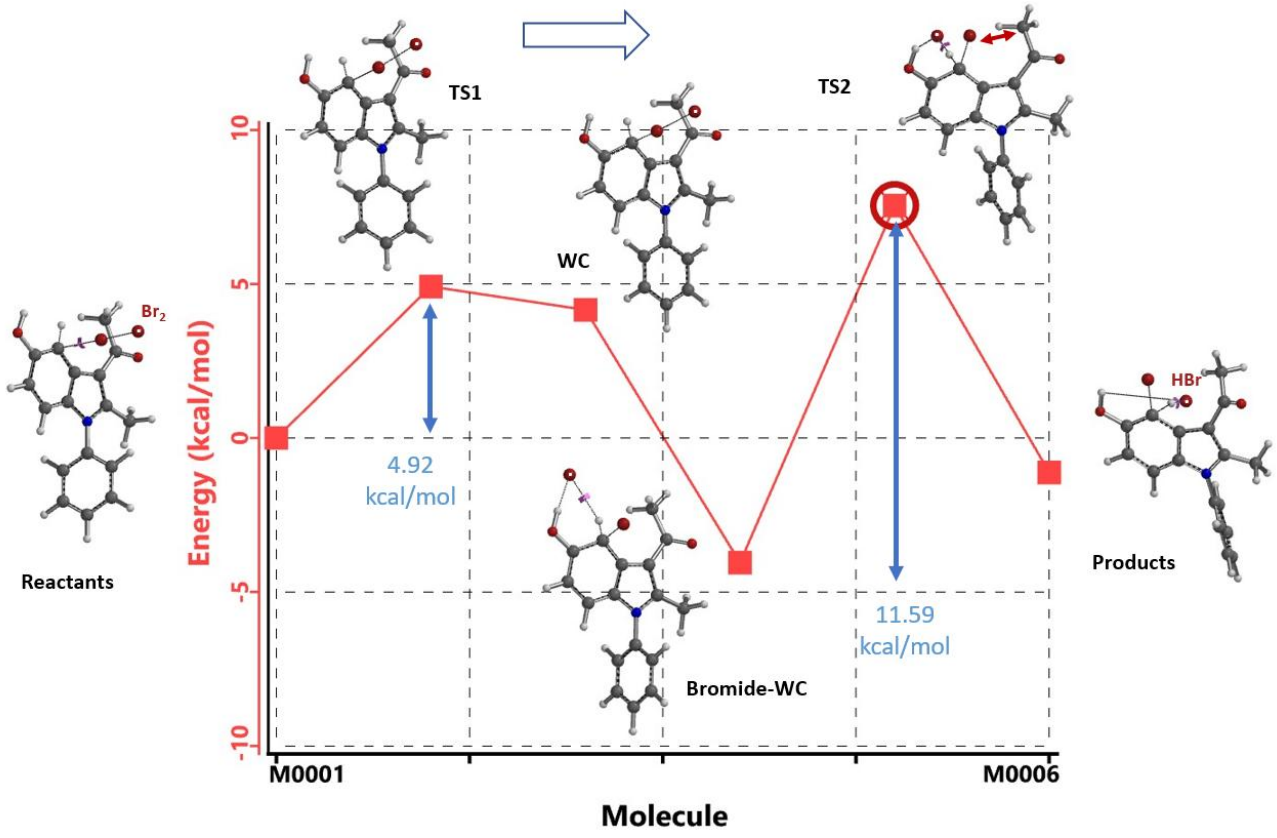


Figure 5. Calculated energy profile of indole **1** C4 bromination

Shown in Figure 6 is the calculated energy profile for C6-bromination. Br_2 and indole **1** (**Reactants**) to **TS1** towards formation of the transient **WC** requires an activation energy of **6.21 kcal/mol**, significantly higher than **4.92 kcal/mol** calculated for C4 bromination. The distal bromine then interacts with both the C-5 hydroxy group and the reaction center's sp^3 CHBr hydrogen to form a hydrogen bond stabilized **Bromide-WC**, which is 2.07 kcal/mol lower in energy than the **Reactants**. The bromide in the 6-membered hydrogen bond ring of the **Bromide-WC** deprotonates the sp^3 CHBr hydrogen, forms HBr, and rearomatizes to provide the C6 bromination product. With no significant steric barrier in the process, this results in a significantly lower calculated energy barrier of 6.71 kcal/mol. It is the difference in energy barrier in this elimination step, 11.59 at C4 *vs* 6.71 at C6, a difference of 4.88 kcal/mol, that differentiates the two competing paths, leads to the highly selective C6 bromination observed.

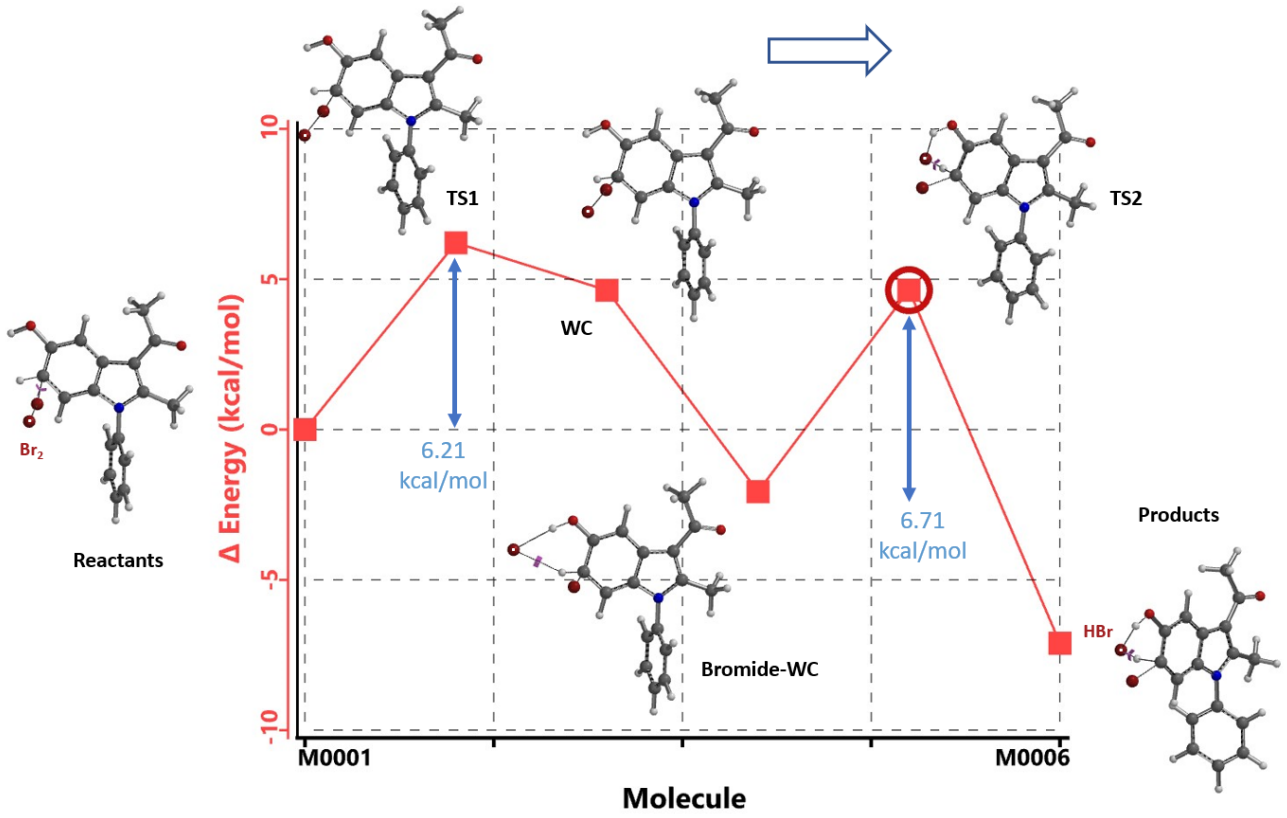
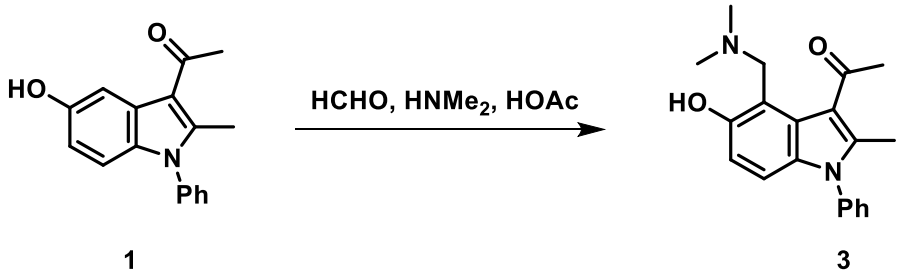


Figure 6. Calculated energy profile of indole **1** C6 bromination

In summary, calculation and comparison of the energy profiles of the potential pathways with full structure of the bromination reagent provide better insight on possible mechanism of the reaction and correlation with experimental observation. Reaction of indole **1** with *N*-bromo-succinamide (NBS) was reported to provide C6 selectively as well. Similar energy profiles could be generated by replacing the distal Br with succinimide in the calculations. For bromination in which the addition is the RLS, the positive charged, hypothetical, Wheland Complexes will be decent approximations; their relative energy usually correlates with experimental observations. While for bromination in which the elimination is the RLS, as shown in Figures 5 and 6, comparing the relative energy of the C4 *vs* C6 bromination products, 5.66 *vs* 0.00 kcal/mol respectively, may also be adequate for prospective regioselectivity analysis.

QM Analysis of the Mannich Reaction

Mannich reaction of indole **1**, in contrast to the highly selective bromination at C6 discussed above, provides only the C4 substitution product **3**[1]. Both are electrophilic aromatic substitutions, shall regioselectivity of one be predictive of the other[2]?



1 **3**

Figure 7. Mannich reaction of indole **1**

Formaldehyde and dimethyl amine react to form an iminium intermediate $[H_2C{=}N(CH_3)_2]^+$, a highly reactive species with LUMO energy of -5.24 eV. LUMO lobe on the iminium carbon is significantly larger than that of the nitrogen, accounting for selective nucleophilic attack on the iminium carbon (Figure 8, left). QM calculation reveals an accurate description of electron distribution of $[H_2C{=}N(CH_3)_2]^+$ (Figure 8, center), i.e. methyl H on iminium ion is positively charged, while the iminium nitrogen is negatively charged, instead of carrying a formal positive one as indicated in most organic chemistry textbooks[4]. Natural charge on nitrogen changes from -0.698 in dimethylamine to -0.275 in $[H_2C{=}N(CH_3)_2]^+$; the N becomes less negatively charged, not positively charged. Similar to tetramethylammonium ion described in QM chapter 23 (Figure 8, right)[5].

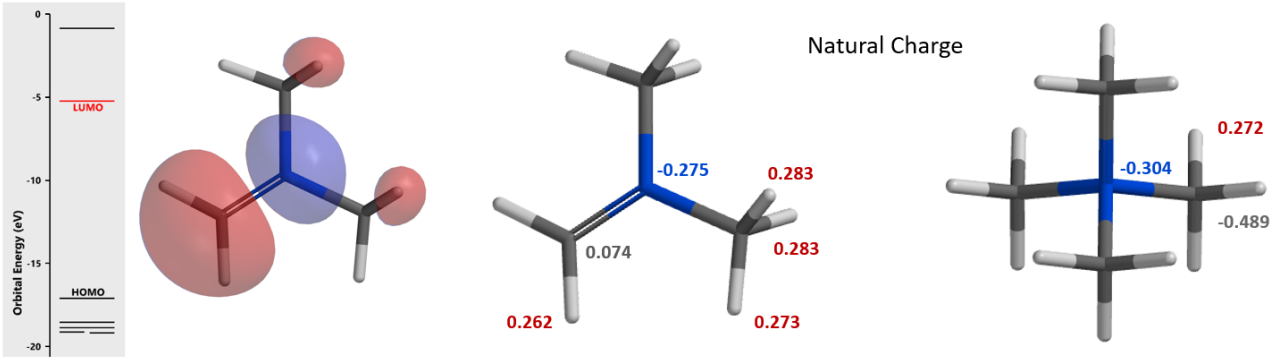


Figure 8. LUMO (left) and natural charge of $[H_2C{=}N(CH_3)_2]^+$ (center); natural charge of tetramethylammonium ion

Shown in Figure 9 (left) is transition state structure calculated for the Mannich reaction proceeding at C4 position, revealing a stabilizing Non-Covalent Interaction (NCI) between the C=O oxygen on 3-acetyl group of indole **1** and the methyl H on $[H_2C{=}N(CH_3)_2]^+$. On the other hand, this is not observed with corresponding transition state with C6 attack (Figure 9, right), leading to an energy difference of 1.74 kcal/mol between the two of them, favoring C4 attack, accounting for the selectivity observed.

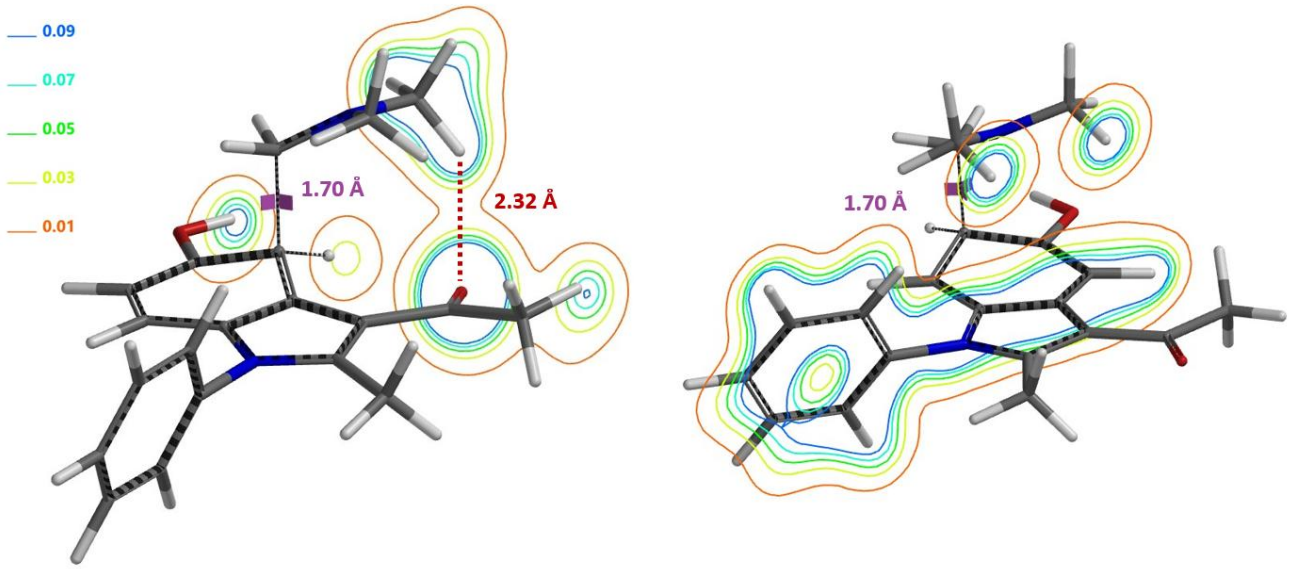


Figure 9. Calculated transition state structures: C4 attack shows 3-acetyl group C=O non-covalent interaction with $[H_2C{=}N(CH_3)_2]^+$ methyl hydrogen (left with electron density contour); C4 attack does not have such NCI stabilization

Conclusion

Differences in regioselectivity between electrophilic aromatic substitutions are not common. For the indole substrate **1**, bromination at C4 faces steric interaction with the 3-acetyl group, while the Mannich reaction at C4 encounters electronic stabilization effect between the acetyl group and incoming iminium ion, accounting for the divergence. QM analyses require us to examine carefully structures of the substrates and electrophiles for potential neighboring group interactions and choose relevant calculation models that are consistent with experimental conditions. This also involves critical examination of the basic reaction mechanisms and recognizing Wheland Complex could be just another convenient conceptual approximation we learned.

Building on What We Just Learned

Bromination of 8-methoxyquinoline with NBS provides only the C-5 product and could be accounted for with QM calculated HOMO, ^{13}C chemical shifts, and relative energy of C-5 *vs* C-7 equilibrium geometry of the brominium ions. However, NBS bromination of analogous 8-hydroxyquinoline, with almost the same calculated QM parameters, generates a mixture of C-5 and C-7 bromides, with C-7 substituted one being the major product.

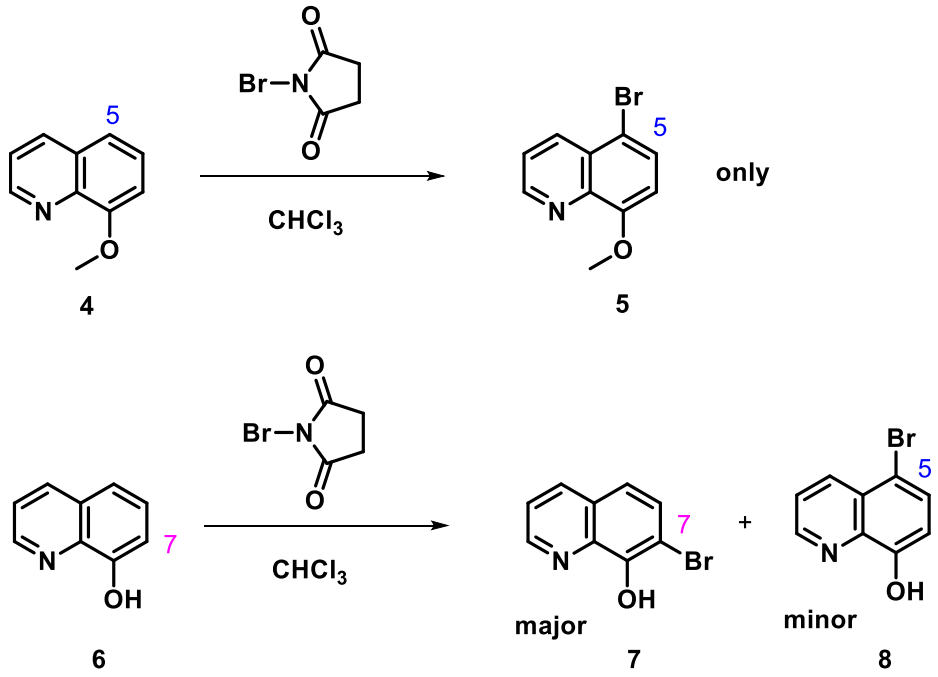


Figure 10. NBS Bromination of 8-methoxyquinoline and 8-hydroxyquinoline

We looked for potential hydrogen bonding interaction between the OH group and the brominating reagent to account for the divergence. Extensive QM analyses led to our current hypothesis that relative overall rate of C-5 *vs* C-7 bromination is controlled by the rate of deprotonation of the C-5 and C-7 Wheland intermediates which are in reversible equilibrium as discussed above[2].

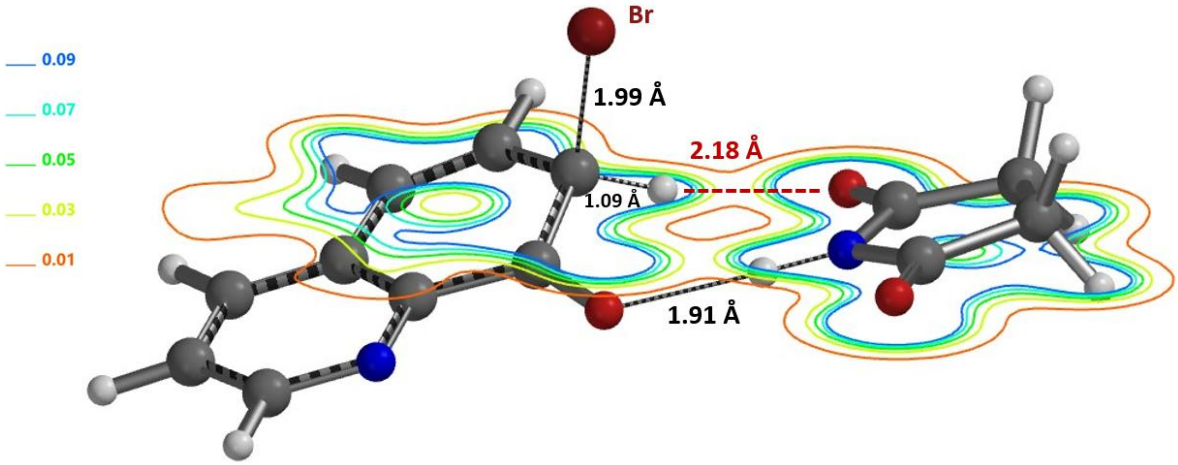


Figure 11. Equilibrium Geometry of 8-hydroxyquinoline-NBS complex with Electron Density Contour on the HO=CN plane.

Obvious questions arise. What is the role of succinimide anion (Figure 11)? Shall we expect a difference in C-7 *vs* C-5 product ratio under different reaction conditions, e.g., NBS *vs* bromine with different counter anion, under different reaction concentration, etc.

Return to Table of Contents

References:

[1] G.S. Gadaginamath, A. G. Kamat, B. G. Pujar, *Revue Roumaine de Chimie*, **1995**, *40*, 265. Both Br_2 and NBS brominating reagents reacts with indole **1** to provide selectively the C-6 bromide.
[2] https://en.wikipedia.org/wiki/Electrophilic_aromatic_substitution.
[3] W.J. Hehre, A.J. Shusterman, J.E. Nelson (1998) The Molecular Modeling Workbook for Organic Chemistry. Irvine, CA, USA: Wavefunction, Inc.; pp 86-7. Similarities between deprotonation and Sn2 reaction.

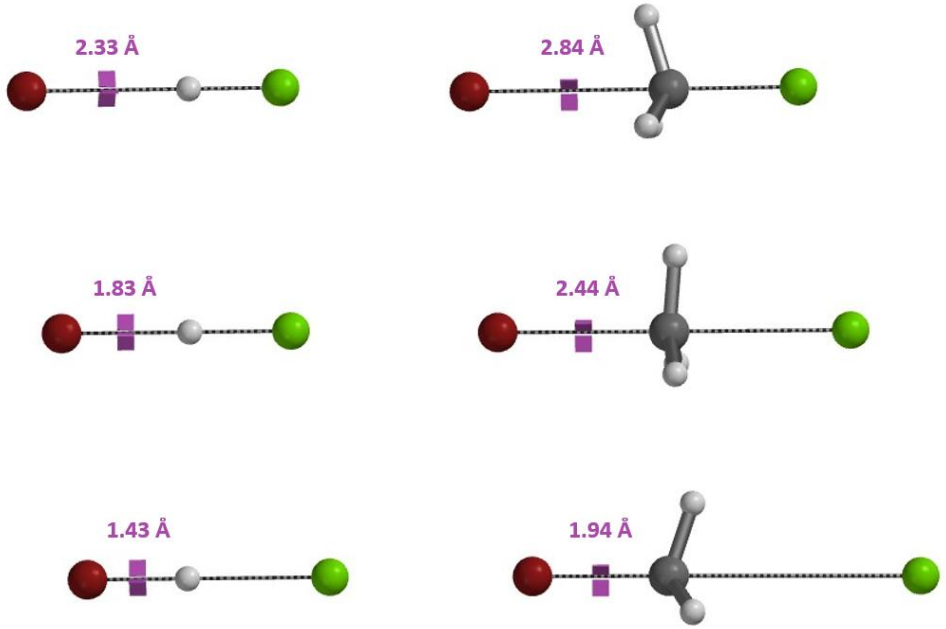


Figure 12. Br^- approaches the hydrogen in HCl (left) and the carbon in CH_3Cl (right) from the "backside", displacing Cl^-. Charge is transferred from Br^- to Cl^-

[4] E.V. Anslyn, D.A. Dougherty (2006) Modern Physical Organic Chemistry. New York, NY, USA: University Science Books; page 7 correctly points out the misrepresentation in most organic chemistry textbooks.
[5] QM Chapter 23 "A QM Study of the para Regioselectivity of $TBABr_3$ Bromination" https://wuxibiology.com/a-qm-study-of-the-para-regioselectivity-of-tbabr3-bromination/

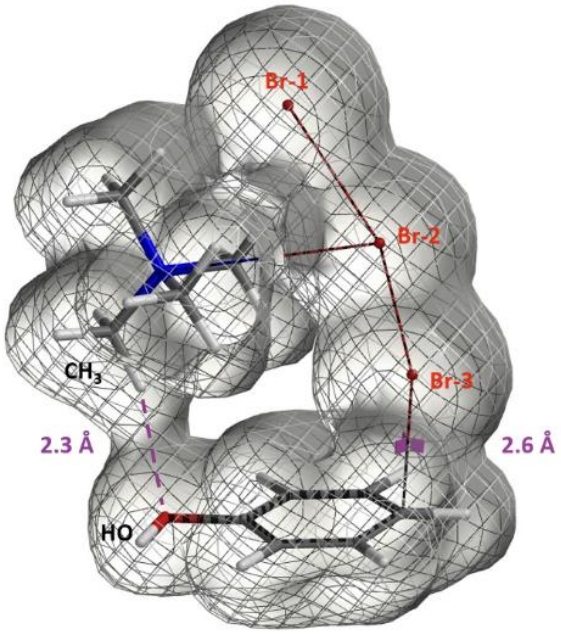


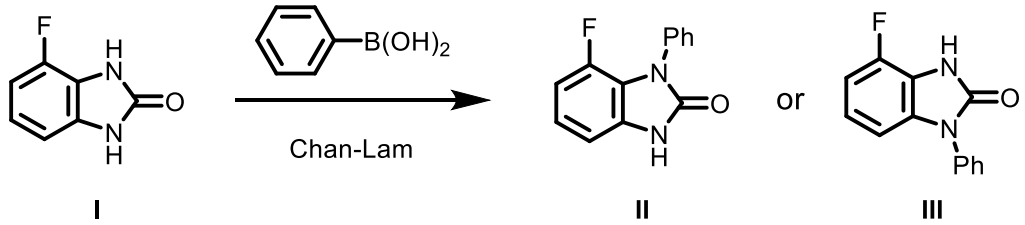
Figure 13. Electron density map (Isovalue 0.00744 e/au3, 98.0%) of $TMABr_3$ – PhOH reaction complex with C – Br distance at 2.6 Å

Chapter 47 QM Quiz II

Zhifa Pu, Qiuyue Wang, Hailong Ren, Wenfeng Liu, Tommy Lai, Yongsheng Chen, John S. Wai

This chapter covers four questions we had in an open-book QM examination in WuXi AppTec. For setting up QM calculations and interpreting results, participants have access to our QM e-Book [1] and Spartan manual [2]. They were requested to refrain from doing literature search for answers during the 2-hour exam. The analyses below are provided by our selected participants.

Question 1



I II III

The above Chan-Lam reaction formed a single product, II or III? What are the QM parameters to calculate for analysis? Please provide and interpret QM calculation results.

Answer (Zhifa Pu)

For the Chan-Lam reaction[3], first we need to find out which NH is more acidic. Based on the calculated Electrostatic Potential Map (ESP Map) of substrate **I**, N1-H is more acidic (N1-H: ESP 250.6 kJ/mol > N3-H: 236.1 kJ/mol) and will be selectively deprotonated. Intuitively, one would have expected N3-H, next to the fluoro group, to be more acidic.

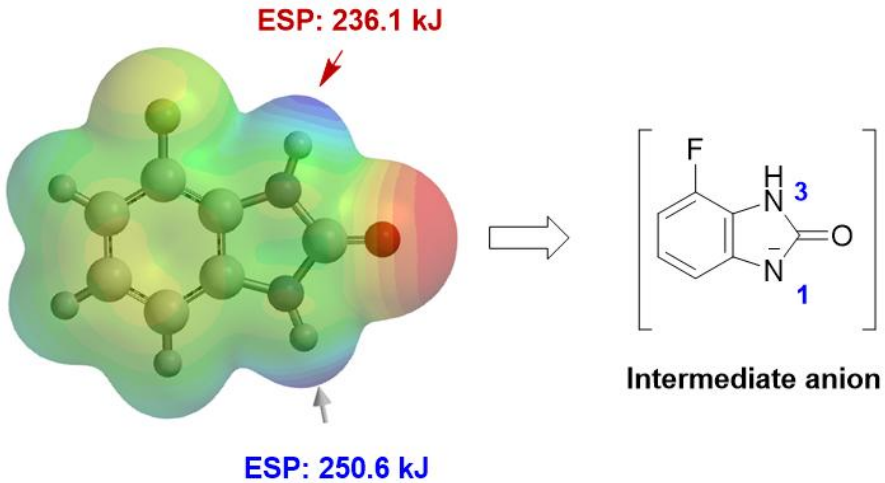


Figure 1. Chan-Lam reaction regioselectivity: step 1, ESP calculation

Next, we calculate for HOMO of the resultant anion, which shows the necessary lobe for N1 phenylation. As such, Chan-Lam reaction of compound **I** should occur preferentially at N1 to generate a single product **III**.

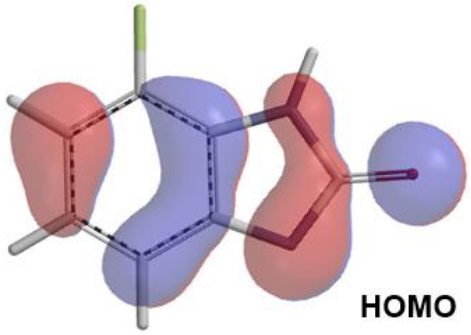
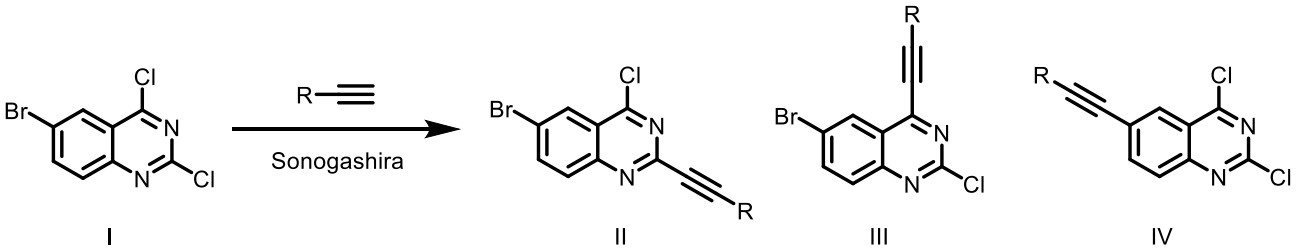


Figure 2. Chan-Lam reaction regioselectivity: step 2, check HOMO

Question 2



The above Sonogashira reaction formed a single product, it is II, III or IV[4]? What are the QM parameters to calculate for analysis? Please provide and interpret QM calculation results.

Answer (Qiuyue Wang)

For Sonogashira reaction, like other metal-catalyzed cross coupling reactions, regioselectivity is controlled by the oxidative addition step. Shown in figure 3 are the LUMO and LUMO Map of substrate **I**, where only C4 has an obvious LUMO lobe centered on the C-halogen carbon. LUMO and LUMO+1 energy gap is relatively large (0.98 eV), so we do not need to consider LUMO+1. We also use QM calculated C-X bond infrared stretching vibrational wavenumbers to differentiate their relative reactivity toward oxidative addition. The weaker the C-X bond is, the lower the wave number it has, and the easier it is for the bond to undergo oxidative addition. From the calculation, C2-Cl, C4-Cl and C6-Br bond stretching wavenumbers are 1242, 825, and 1048 cm^{-1}, respectively, with the C4-Cl one being the weakest bond. Both methods predict that oxidative addition will proceed selectively at C4 and lead to a single product **III** [5].

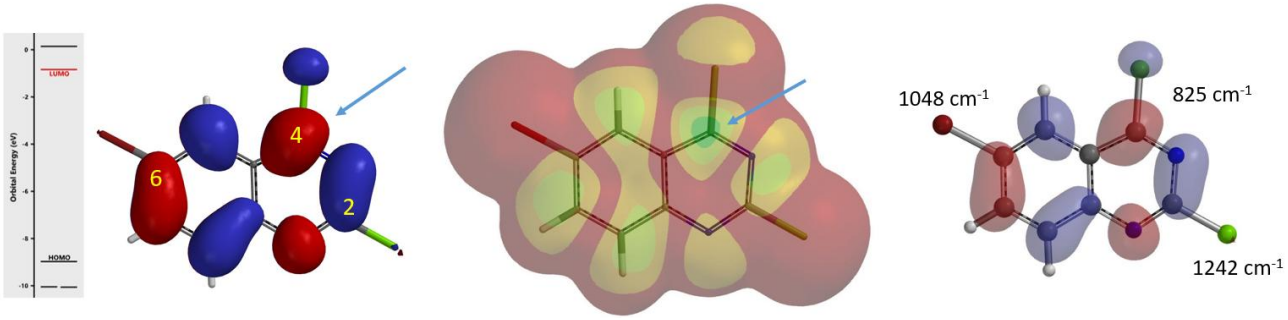
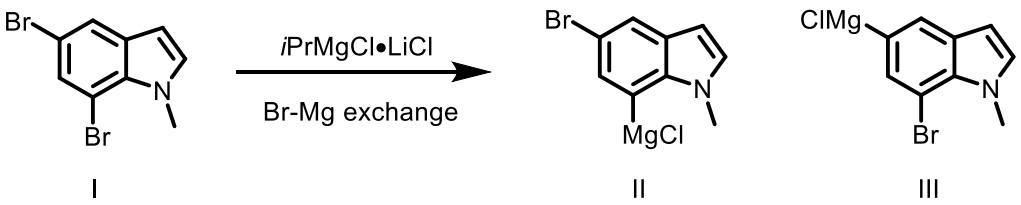


Figure 3. LUMO, LUMO Map, and IR calculations of substrate **I**

Question 3



I II III

The above bromide-magnesium exchange reaction formed a single product, II or III[6]? What are the QM parameters to calculate for analysis? Please provide and interpret QM calculation results.

Answer (Hailong Ren)

For polyhalogenated substrates, we look at LUMO and LUMO Map, compare the calculated stretching wavenumber of the carbon-halogen bonds to correlate with regioselectivity of SnAr and oxidative addition reactions. With halide-lithium exchange or halide-magnesium exchange reactions we discovered that it is more appropriate to compare appropriate LUMO+n with unique "String-of-Pearls" shaped lobes along the carbon-halogen bonds of the substrates [7].

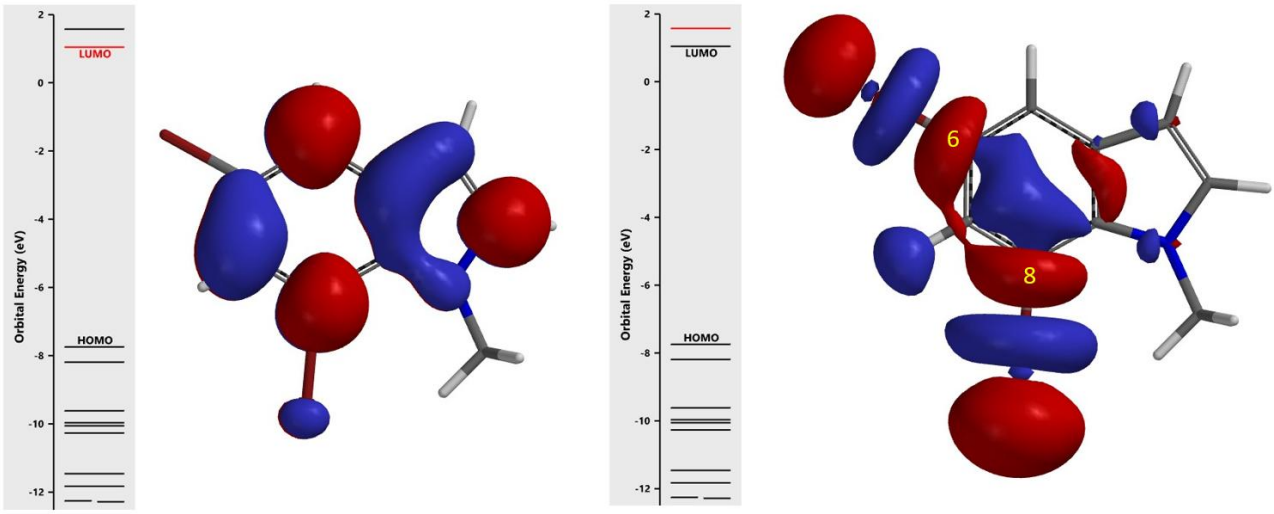


Figure 4. LUMO (left), LUMO+1 (right) calculation of substrate I

As shown in Figure 4, LUMO+1 of substrate **I** has the unique "String-of-Pearls" lobes distribution along the two C-Br bonds. Overlaying it with Electron Density Map (Figure 5) reveals that the C7-Br Lobe is more accessible for Grignard exchange; as such, the C7 bromo group will react preferentially to provide selectively product **II**. This example exemplifies a unique characteristic of halogen-metal exchange reaction, that is, due to the unique unoccupied orbitals required, with its terminal lobe extending beyond Electron Density isosurfaces, it is not sensitive to neighboring group's steric hindrance[7].

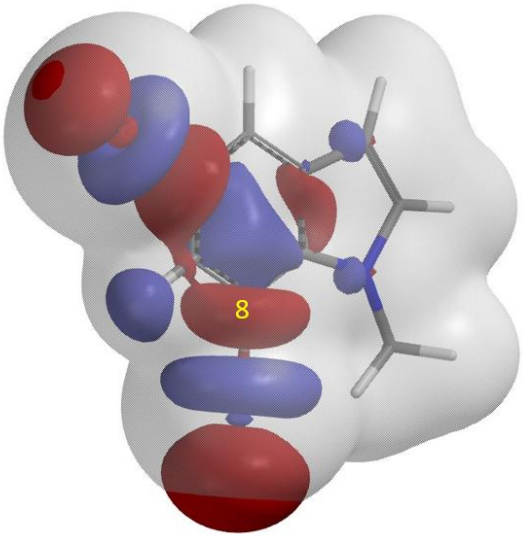
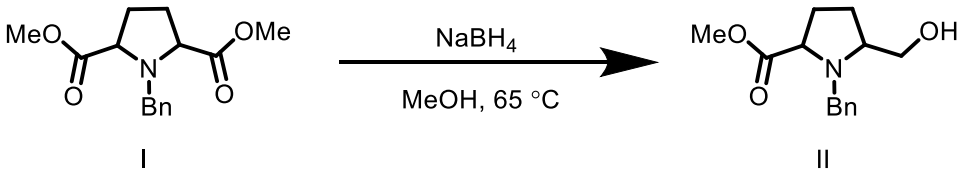


Figure 5. Substrate **I** LUMO+1 and Electron Density Map overlay

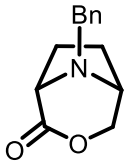
Question 4



I **II**

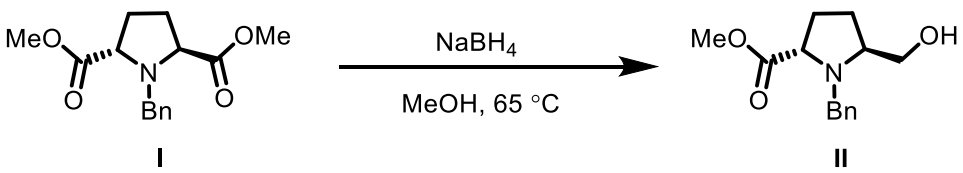
Reduction of diester **I** with sodium borohydride provides exclusively **II** as the only product[8]. What are the QM parameters to calculate for analysis to account for the selectivity? Please provide QM calculation results and interpret.

Answer (Wenfeng Liu)



III

When one ester group is reduced first, only a single product **II** is obtained. Lactone **III** as shown in the figure above is not observed, indicating it is likely that the ester groups are *trans* to one another.



I **II**

Figure 6. Monoester reduction of *trans* diester **I**

Based on this assumption we calculated for the LUMO map of substrate **I** and added to it inaccessibility markers. The blue to green coloration on the LUMO map at the carbonyl carbons, on both sides of the molecule, represents decent values of the orbital, yet only one of them and only from one side is accessible for reduction to provide compound **II**.

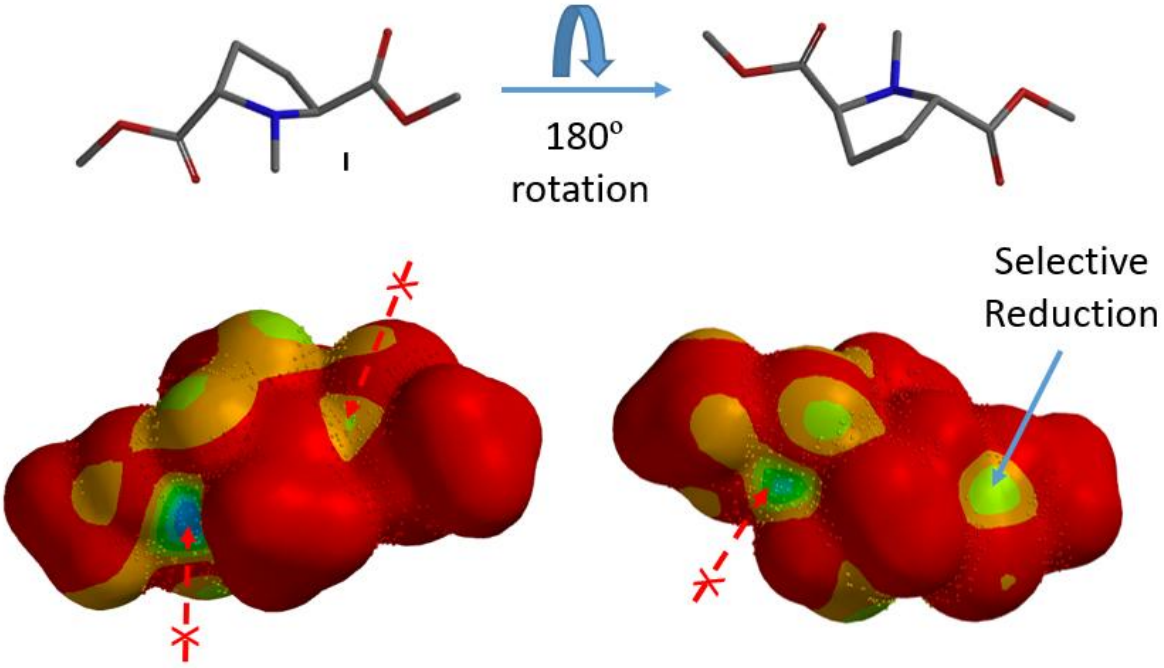


Figure 7. LUMO Map with Inaccessible Markers for substrate **I**

LUMO Map calculated for structure **II** reveals that the remaining carbonyl carbon is covered with inaccessibility markers, indicating that this carbonyl is not readily accessible for reaction, accounting for why product **II** cannot be further reduced under mild reduction conditions.

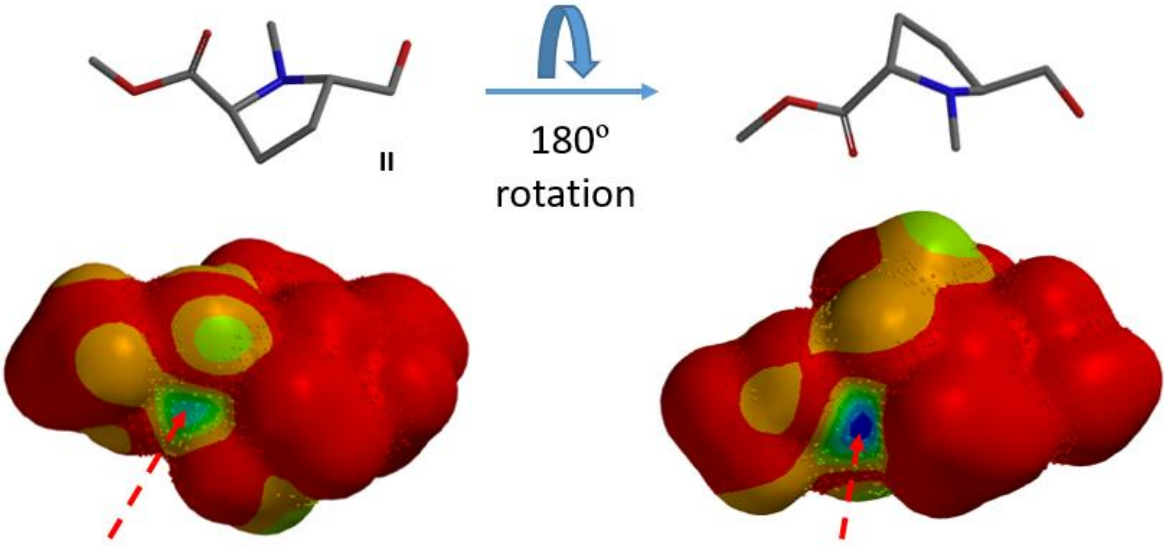


Figure 8. LUMO Map with Inaccessible Markers of product **II**

Return to Table of Contents

References:

[1] Quantum Mechanics for Organic Chemists: An Experimentalist Approach. *https://wuxibiology.com/resource/quantum-mechanics-for-organic-chemists/*

[2] Spartan Tutorial and User's Guide. Accessible within Spartan software and from https://www.wavefun.com/spartan-documentation

[3] QM chapter 34 "QM Analyses of Regioselectivity in Chan-Lam Reaction" https://wuxibiology.com/qm-analyses-of-regioselectivity-in-chan-lam-reaction/

[4] I. Mangalagiu, T. Benneche, K. Undheim, *Acta Chem. Scand.* **1996**, *50*, 914

[5] QM chapter 45 "Unusual Reactivities of Polyhalogenated Heteroaromatic Substrates are Predictable" https://wuxibiology.com/unusual-reactivities-of-polyhalogenated-heteroaromatic-substrates-are-predictable/

[6] P. Anbarasan, H. Neumann, M. Beller, *Chem. Eur. J.* **2011**, *17*, 4217

[7] QM chapter 39 "Unraveling Divergence in Haloselectivity of S_NAr and Cross Coupling *vs* Halogen-Metal Exchange Reactions" https://wuxibiology.com/unraveling-divergence-in-haloselectivity-of-snar-and-cross-coupling-vs-halogen-metal-exchange-reactions/

[8] K. W. Yang, L. Feng, S. K. Yang, M. Aitha, A. E. LaCuran, P. Oelschlaeger, M. W. Crowder, *Bioorg. Med. Chem. Lett.*, **2013**, *23*, 5855

Chapter 48 Dichotomy in Regioselectivity of SnAr Reactions with 2-$MeSO_2$-4-Chloropyrimidine

Ying Lu, Wenfeng Liu, Qiuyue Wang, Yongsheng Chen, John S. Wai

Regioselectivity for SnAr reactions of halo pyrimidines is sensitive to the presence of other substituents on the ring. Intuitively, they are not easy to understand, yet fully accountable for with proper QM analyses (See QM chapters 1, 7, 29). In this chapter, we'll discuss the longstanding puzzle of dichotomy in regioselectivity observed for SnAr reactions with 2-$MeSO_2$-4-chloropyrimidine. Similar to 2,4-dichloropyrimidine, SnAr reactions of 2-$MeSO_2$-4-chloropyrimidine with amines and Stille coupling occur selectively at C-4, but substitution reactions with alkoxides and formamide anions occur selectively at C-2 (Figure 1).

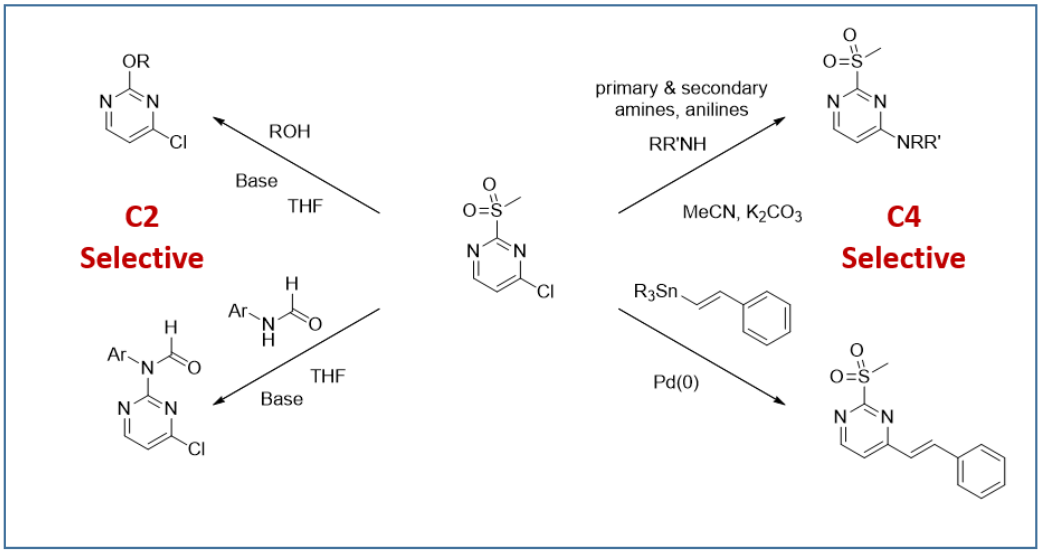


Figure 1. Dichotomy in regioselectivity of various reaction types with 2-$MeSO_2$-4-chloropyrimidine

C-2 Selective SnAr Reaction with Alkoxide

In 1993, SnAr of 2-$MeSO_2$-4-chloropyrimidine with alkoxide was reported to be exclusively at C-2 and proceed at surprisingly low temperature (-78°C)[1]. Even though this unique selectivity has been utilized quite often in MedChem syntheses[2] , there is no further analysis on why it is C-2 selective. Below are results from our QM analyses. LUMO (-0.23 eV) calculated for the pyrimidine shows a lobe centered on C4. LUMO+1 (-0.14 eV) is only 0.09 eV higher in energy and has a lobe centered on C-2 (Figure 2). Together, they suggest that we could have a mixture of C-4 and C-2 substitution products. Obviously, the above LUMO analysis, based on the equilibrium geometry of the pyrimidine, fails to account for the high C-2 selectivity observed.

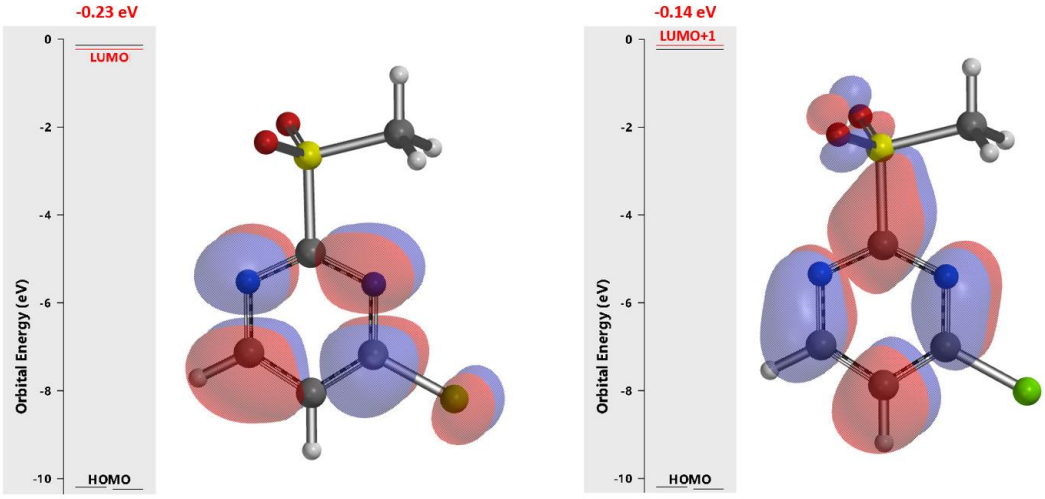


Figure 2. LUMO and LUMO+1 of 2-$MeSO_2$-4-chloropyrimidine

We reasoned that 2-$MeSO_2$-4-Cl-pyrimidine and alkoxide form a hydrogen bond complex. Equilibrium geometry calculated for the complex shows a distance of 1.86 Å between the alkoxide oxygen and the methyl hydrogen of the $MeSO_2$ group (Figure 3, left). There is significant electron density between the oxygen and hydrogen atoms (Figure 3, middle and right), supportive of our hypothesis. It is likely that this is a key factor in directing the SnAr reaction to proceed at the C-2 *vs* C4 positions.

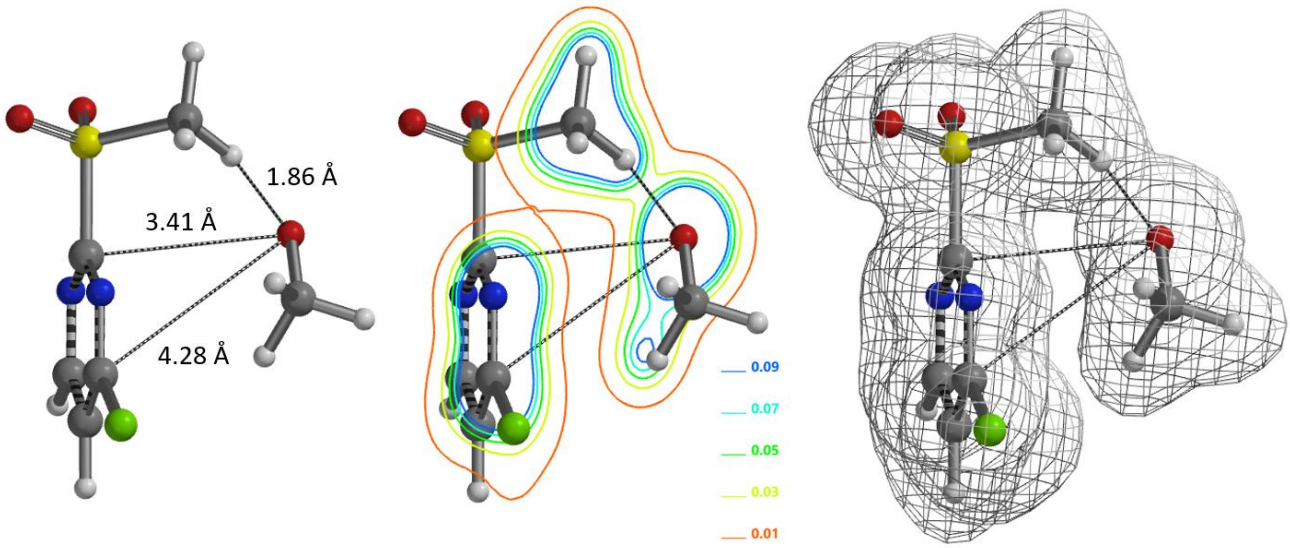


Figure 3. Equilibrium geometry, ED contour and density maps of pyrimidine-alkoxide HB complex (Calculation level: ωB97X-D/6-31+G*)

Next, we modeled the SnAr at C-2 with C-2 to O distance changing from 3.41 Å to 1.51 Å with step size of 0.05 Å. Calculated energy barrier is about 0.25 kcal/mol, accounting for the -78°C low reaction temperature (Figure 4, left hand side). The hydrogen bond between $MeSO_2$ H and alkoxide oxygen in the hydrogen bonded complex is maintained, stabilizing the reaction transition state. There is no energy barrier for subsequent elimination of the $MeSO_2$ group (not shown) to form the C-2 substitution product.

Similarly, calculation of the reaction energy profile for the C-4 substitution (Figure 4, right hand side) shows an energy barrier of ~4.11 kcal/mol, 3.86 kcal/mol higher than that at the

C-2 position, accounting for why the reaction occurs exclusively at the C2 position. The higher barrier corresponds to energy cost in breaking of the $MeSO_2$ H and alkoxide oxygen hydrogen bond.

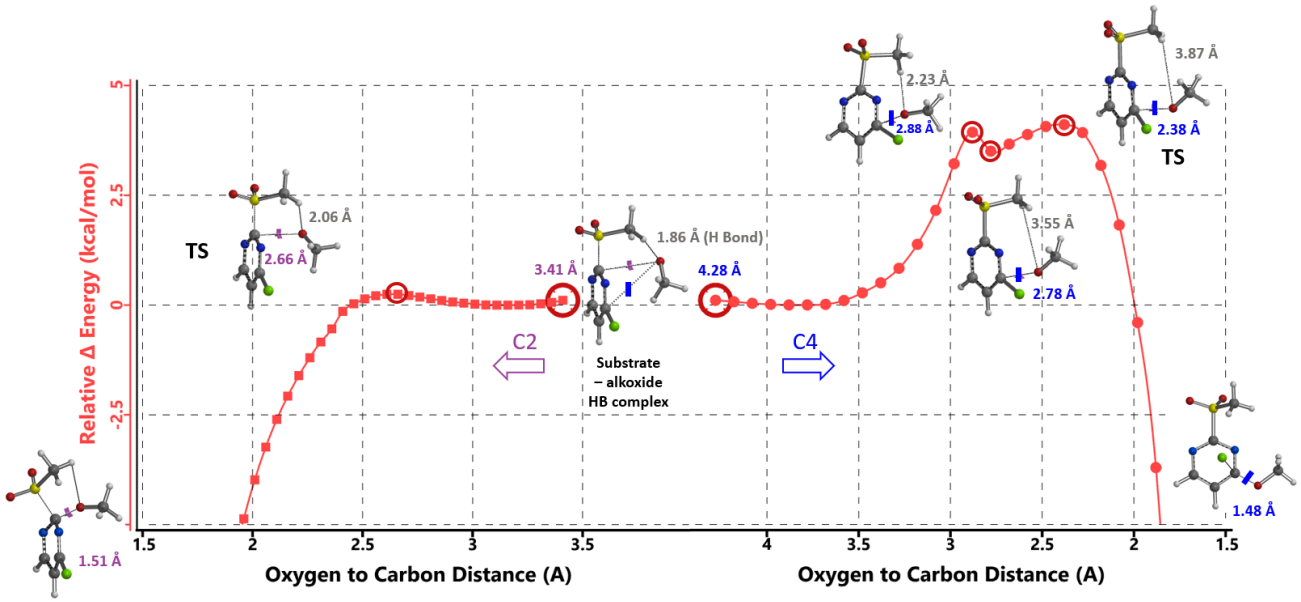


Figure 4. Reaction Energy Profiles of SnAr reaction between 2-$MeSO_2$-4-Cl-pyrimidine and alkoxide: C-2 *vs* C-4

Subsequent transition state geometry calculation reveals a relatively long carbon to oxygen distance of 2.64 A, single and relatively low iFreq of 88 cm^{-1} for SnAr at C-2, suggestive of a relative early transition state. The C-4 transition state is 3.87 kcal/mol higher in energy (Figure 5)[3], in accordance with results from REP calculations above.

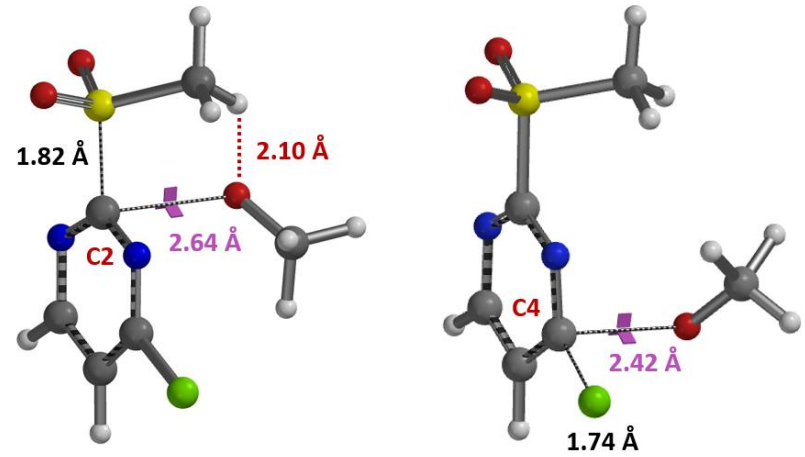


Figure 5. Transition state geometries and imaginary frequencies for SnAr at C-2 and C-4 positions

C-2 Selective SnAr Reaction with Aromatic Formamides

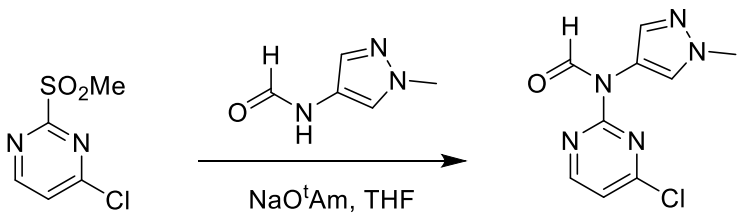


Figure 6. Reaction with aromatic formamide compound

SnAr reaction between 2-$MeSO_2$-4-Cl-pyrimidine and aromatic formamide also occurs selectively at C-2 position [4,5]. Based on what we have learned above, we first calculated for the equilibrium geometry of the pyrimidine and methylpyrazole formamide anion hydrogen bond complex (Figure 7). Results are supportive of the presence of a hydrogen bond (2.24 Å) between the H of the $MeSO_2$ group and the N of the formamide (Figure 7, left and middle) and reveal the presence of a *pi-pi* interaction between the two aromatic rings (Figure 7, right).

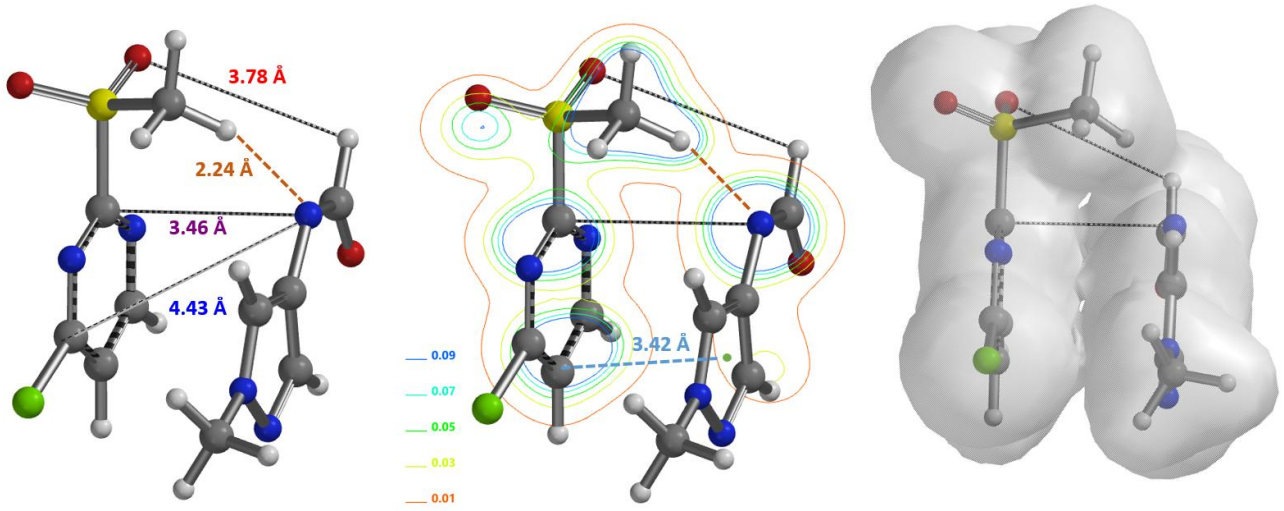


Figure 7. Equilibrium geometry, ED contour and density maps of pyrimidine-formamide anion complex

Modeling of the SnAr reaction at C-2 provided a calculated energy barrier of ~8.75 kcal/mol, consistent with reported reaction temperature of 0°C. The hydrogen bond between $MeSO_2$ H and formamide N in the hydrogen bonded complex is preserved in transition state (Figure 8, left hand side). Similarly, calculation of the reaction energy profile for C-4 substitution (Figure 8, right hand side) shows a higher energy barrier of ~14.71 kcal/mol resulting from the loss of the hydrogen bond. The difference in energy barrier of 5.96 kcal/mol accounts for the highly selective C-2 substitution observed.

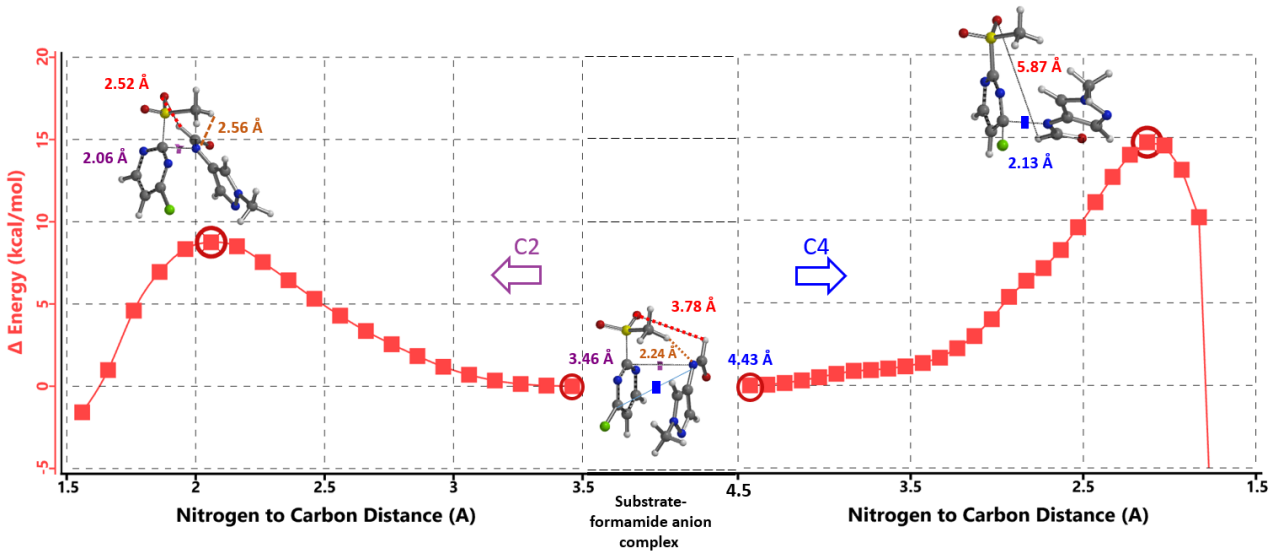


Figure 8. REP of SnAr reaction between 2-$MeSO_2$-4-Cl-pyrimidine and formamide anion: C-2 *vs* C-4

Shown in figure 9 are the calculated transition state geometries and each of their single imaginary frequencies. Energy of the C-4 transition state is 4.12 kcal/mol higher than that at the C-2 position.

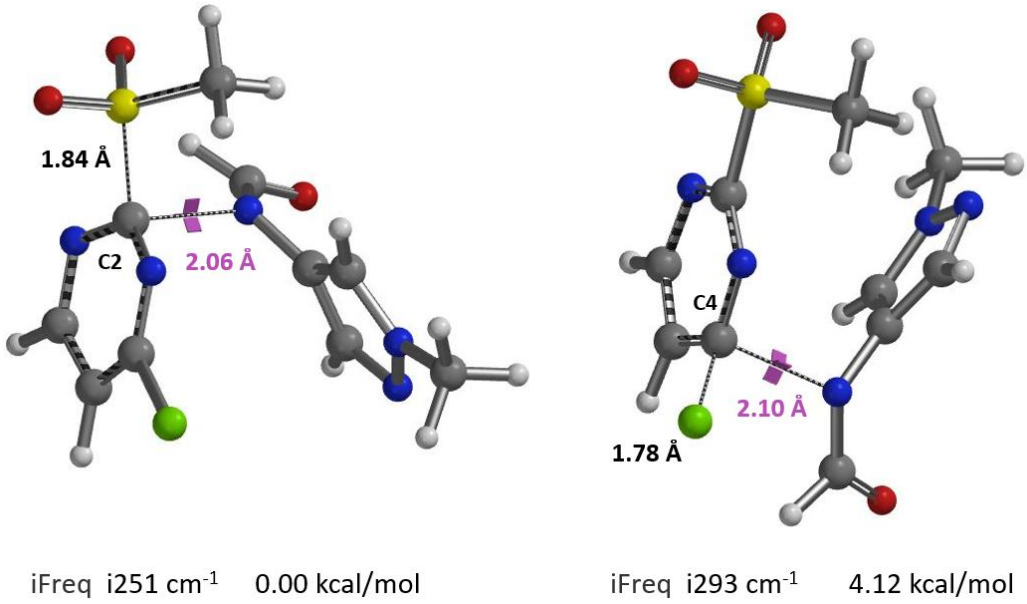


Figure 9. Transition state geometries and iFrequencies calculated for SnAr reaction at C-2 & C-4 positions

The formyl group enables deprotonation of the formamide under mild conditions. Electron density contour maps of the C-2 transition state structure (Figure 10) reveal two important interactions that lead to the high C-2 regioselectivity observed for the SnAr reaction:
1. Hydrogen bond (**2.58 Å**) between SO_2CH_3 H and formamide anion N, similar to the alkoxide discussed above; 2. A second hydrogen bond formed between the formamide H and one of the two $MeSO_2$ oxygens (**2.53 Å**), further lowering its energy. Together, they account better for why the formyl group is so crucial for the high efficiency and selectivity observed for the reaction.

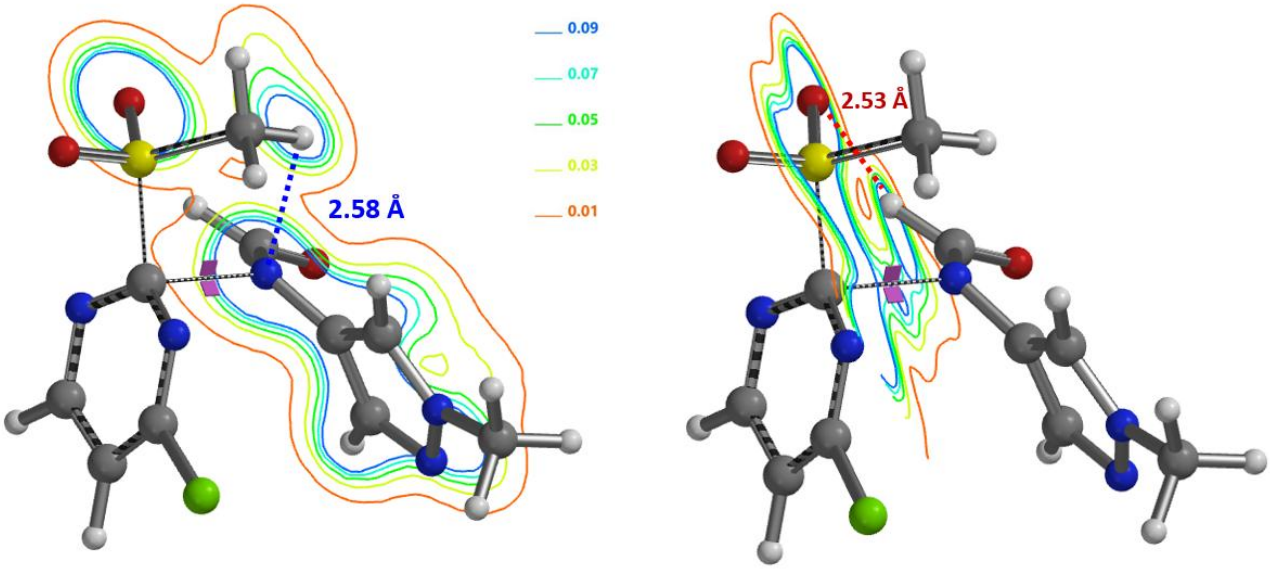


Figure 10. Electron density contour maps of the C-2 transition state revealing two hydrogen bonds

From QM analyses for the two reactions above, we could see why they are highly C-2 selective
1. Non-covalent interaction between the alkoxide/formamide anion with the acidic proton of the methanesulfonyl group, forming hydrogen bond complexes;
2. The nucleophiles need to be anionic for formation of the hydrogen bond complexes;

3. The hydrogen bonds direct the nucleophile to attack at C-2. They are preserved in the transition state, lowering their energies relative to C-4 attack.

C-4 Selective Metal-catalyzed Cross Coupling reactions

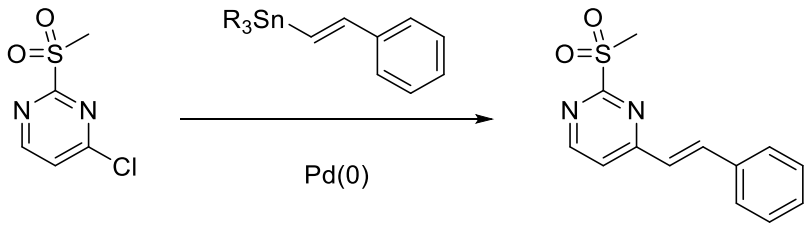


Figure 11. C-4 selectivity for Stille reaction

So why does the oxidative addition reaction occur at the C4 position? Recently, Chatelain [6] *et al.* reported that methylsulfones display intermediate level of reactivity between that of nitroarenes and aryl halides in Pd(0)-catalyzed Suzuki cross coupling reaction (Figure 12), with trifluoromethyl sulfone group providing the best conversion. Similarly, this reported trend in reactivity could account for faster oxidative addition at C-Cl *vs* C-O_2SMe bonds, accounting for the C-4 selectivity observed with the Stille coupling reaction.

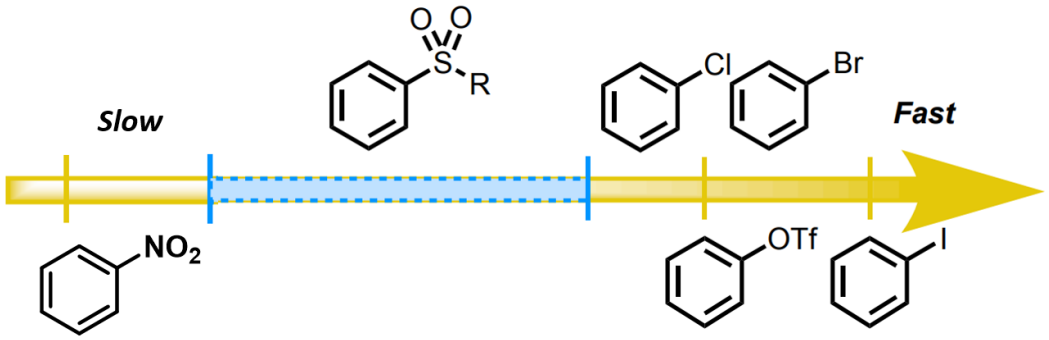


Figure 12. Relative reactivity of aryl electrophiles in Suzuki cross coupling reaction

Conclusion

2-$MeSO_2$-4-Cl-pyrimidine forms hydrogen bond complexes with alkoxide and formamide anion *via* its acidic $MeSO_2$ proton. This hydrogen bond directs the SnAr reaction at C-2 and lowers the energy of the transition state. On the other hand, C-4 attack will require breaking of the hydrogen bonding, increasing the energy barrier. When the reaction involves hydrogen bond, non-covalent interaction (NCI)[8], steric or other electronic interactions, etc. (see chapters 9, 23, 26, 33, 42), their relative contributions are intuitively difficult to compare. QM calculations enable us to evaluate them in a more objective manner. In the process, we learn to go beyond knowing reactions of interest to understanding them better. It is crucial that we build our computational models based on experimental conditions and observations. If you come across other interesting chemistry, please let us know, let's analyze and learn together.

Building on What We Just Learned

The reaction of 2-$MeSO_2$-4-Cl-pyrimidine with secondary amine only occurs selectively at C-4. How to rationalize it?

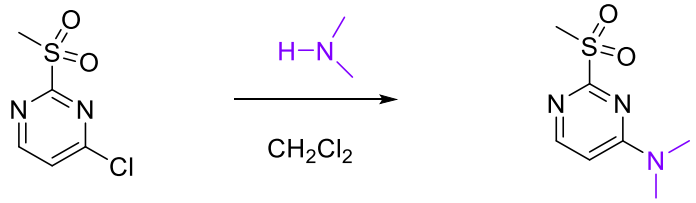


Figure 13. Reaction of 2-$MeSO_2$-4-Cl-pyrimidine with secondary amine

Secondary amines, with lone pair electrons and hydrogen on the N atom, have various potential ways to interact with the $MeSO_2$ group and sp^2 N of the pyrimidine. Calculation shows that they form a complex with a ΔE of -7.61 kcal/mol (structure shown in Figure 14). Electron density contour map reveals:

1. Strong non-covalent interaction (**2.98 Å**) between the secondary amine nitrogen (lone pair electrons) and the electrophilic pyrimidine C-2 carbon.
2. A hydrogen bond (**2.55 Å**) between the amine nitrogen and $MeSO_2$ group hydrogen; complex with the structure shown is lower in energy (> 1 kcal/mol) than other ones being evaluated in which the amine H hydrogen bond to the $MeSO_2$ oxygen.
3. Two charge-charge interactions between the hydrogens on the secondary amine methyl groups and the pyrimidine N-1 and N-3 (**Natural charges**).

How would you analyze this reaction further? [9]

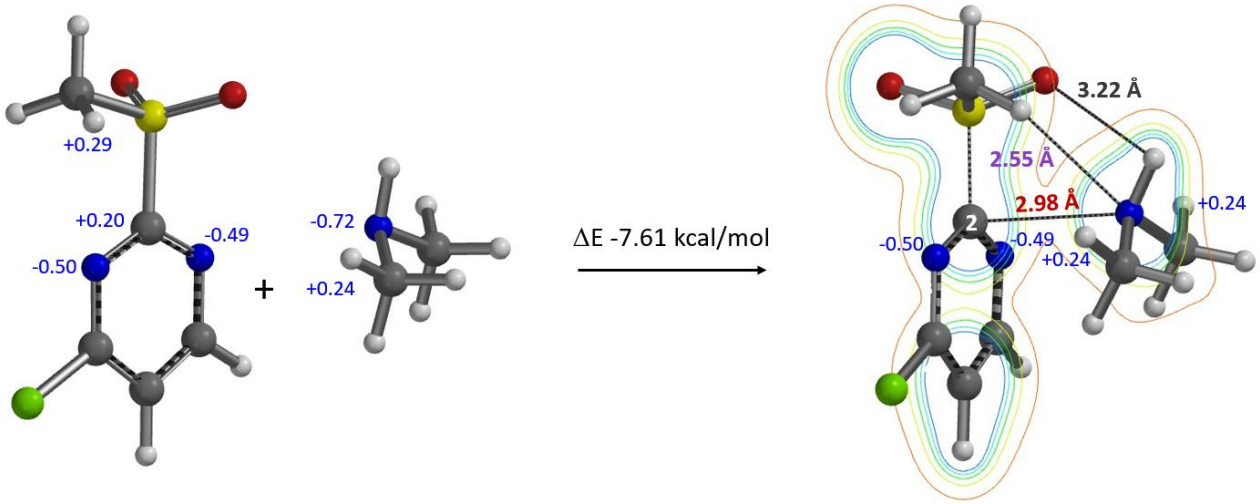


Figure 14. Formation of 2-$MeSO_2$-4-Cl-pyrimidine-secondary amine complex. Electron density contour map and natural charges (blue marker) of the resultant complex

Return to Table of Contents

References:

[1] M. L. Falck-Pedersen, B. Tore, U. Kjell, *Acta Chem. Scand.* **1993**, *47*, 72.

[2] C. M. Serrano, R. E. Looper, *Org. Lett.* **2011**, *13*, 5000; A. S. Carlson, H. R. Cui, A. Divakaran, J. A. Johnson, R. M. Brunner, W. C. K. Pomerantz, J. J. Topczewski, *ACS Med. Chem. Lett.* **2019**, *10*, 1296.

[3] F. M. Bickelhaupt, K. N. Houk, *Angew. Chem. Int. Ed.* **2017**, *56*, 10070. Small HOMO and LUMO energy gap and little structure distortion from reactant to transition state leads to low activation energy.

[4] C. M. Li, F. Haeffner, S. J. Wang, C. C. Yuan, D. Shang, X. L. Shi, B. Ma, B. T. Hopkins, E. M. O'Brien, *Org. Process Res. Dev.* **2022**, *26*, 137.

[5] B. Barlaam, R. Ducray, C. Lambert-van der Brempt, P. Plé, C. Bardelle, N. Brooks, T. Coleman, D. Cross, J. G. Kettle, J. Read, *Bioorg. Med. Chem. Lett.* **2011**, *21*, 2207.

[6] P. Chatelain, A. Sau, C. N. Rowley, J. Moran, *Angew. Chem. Int. Ed.* **2019**, *58*, 14959.

[7] B. L. Mylari, P. J. Oates, W. J. Zembrowski, D. A. Beebe, E. L. Conn, J. B. Coutcher, M. T. O'Gorman, M. C. Linhares, G. J. Withbroe, *J. Med. Chem.* **2002**, *45*, 4398.

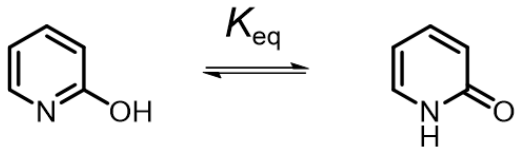
[8] QM Chapter 42 https://wuxibiology.com/visualizing-hydrogen-bonds-using-electron-density-maps/

[9] To continue with the analysis, we'll use the (global minimum energy) structure obtained for 2-$MeSO_2$-4-Cl-pyrimidine-secondary amine complex (shown in figure 14) to set up C-2 and C-4 SnAr REP calculations. Then it will become obvious that the energy barrier of the C-4 substitution is lower than that for C-2. To understand why, transition state calculation, followed by activation strain (distortion) analyses[3] will be very informative. They reveal that the C2-transition state is associated with a larger strain energy, the key reason why SnAr with secondary amine is C-4 selective.

Chapter 49 How about Tautomers?

John S. Wai, Qiuyue Wang, Liting Dong, Yongsheng Chen

A compound exhibits tautomerism if it can be represented by two structures that are related by an intramolecular movement of hydrogen from one atom to another. The different tautomers of a molecule usually have different molecular fingerprints, hydrophobicities, electrostatic properties, acidity, and even 3D shapes [1]. One of the most well studied tautomerizations is the interconversion of 2-hydroxypyridine and 2-pyridone. The equilibrium between these two tautomeric forms is sensitive to properties of solvent. 2-Hydroxypyridine has an OH group capable of hydrogen bonding, yet it is 2-pyridone that is better stabilized in the high polarity solvents. The equilibrium constants in different solvents are shown in Table 1 [2]. Intuitively, this is not easy to rationalize. Could QM help us to understand this dichotomy?



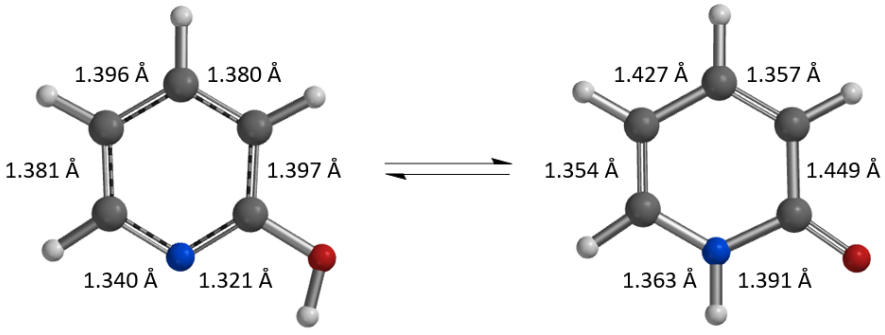
Tautomerization

Solvent	K_{eq}
Cyclohexane	1.7
Chloroform	6.0
Acetonitrile	148
Water	910

Figure 1. Solvent effect in 2-hydroxypyridine and 2-pyridone tautomers equilibria

Tautomerization between 2-hydroxypyridine and 2-pyridone

First, we calculated for the equilibrium geometry of 2-hydroxypyridine and 2-pyridone. DFT ωB97X-V and a bigger basis set 6-311+G(2df,2p) are used for all the calculations discussed in this chapter [3].

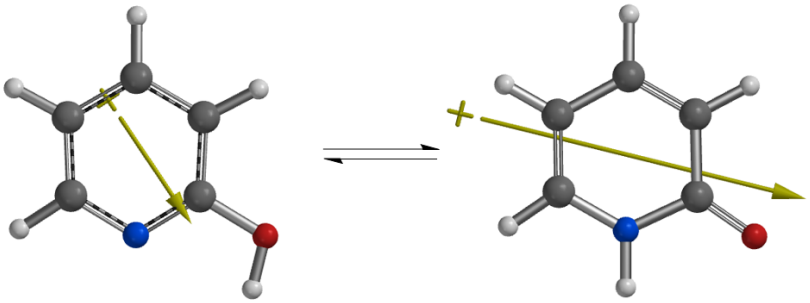


2-Hydroxypyridine is aromatic. The bonding in 2-pyridone is localized.

Figure 2. Equilibrium geometry of 2-hydroxypyridine and 2-pyridone,
calculated with DFT ωB97X-V/6-311+G(2df,2p), C-PCM model, dielectric constant 2.02, cyclohexane

As shown in Figure 2, the calculated carbon to carbon distances in **2-hydroxypyridine** are 1.381 Å, 1.396 Å, 1.380 Å, 1.397 Å, suggestive of **aromatic resonance**. While the carbon to carbon distances in 2-pyridone alternate short and long: 1.354 Å, 1.427 Å, 1.357 Å, 1.449 Å, more consistent with localized CC single and double bonds, and the carbon to nitrogen distances in 2-pyridone are 1.363 Å and 1.391 Å, much longer than the CN distances in 2-hydroxypyridine, 1.340 Å, 1.321 Å, more consistent with CN single bonds in pyridone. These bond distance data suggest **2-pyridone is not aromatic**. Experimental equilibrium constant (K_{eq}) between 2-hydroxypyridine and 2-pyridone is 1.7 in cyclohexane (25 °C), indicative of a relatively small ~0.32 kcal/mol energy difference, with a small preference for 2-pyridone.

To understand the shift in equilibrium between these two tautomeric forms from cyclohexane to chloroform, we compared the calculated Dipole Moment of the tautomers (Figure 3).



Solvent	Expt K_{eq}	Dipole Moment; enol form (Debye)	Dipole Moment; keto form (Debye)	Dielectric Constant
Cyclohexane	1.7	1.65	5.29	2.02
Chloroform	6.0	1.83	5.97	4.71

Figure 3. Equilibria between 2-hydroxypyridine and 2-pyridone and their calculated dipole moments

Calculated dipole moment[4] of 2-pyridone, 5.29 and 5.97 Debye in cyclohexane and chloroform, respectively, are higher than that of 2-hydroxypyridine, 1.65 and 1.83 Debye, respectively, as such 2-pyridone is relatively more stable than 2-hydroxypyridine in chloroform, with a higher dielectric constant than cyclohexane, accounting for the observed K_{eq} shift from 1.7 to 6.0. The location, orientation, and size of the dipole moment vector, where the + sign refers to the positive end of the dipole, correlates with the charge distribution and degree of charge separation on the tautomers. Difference in dipole moment could also be visualized and compared with calculated Electrostatic Potential Maps as shown in Figure 4. 2-Pyridone exhibits significantly larger charge separation than 2-hydroxypyridine.

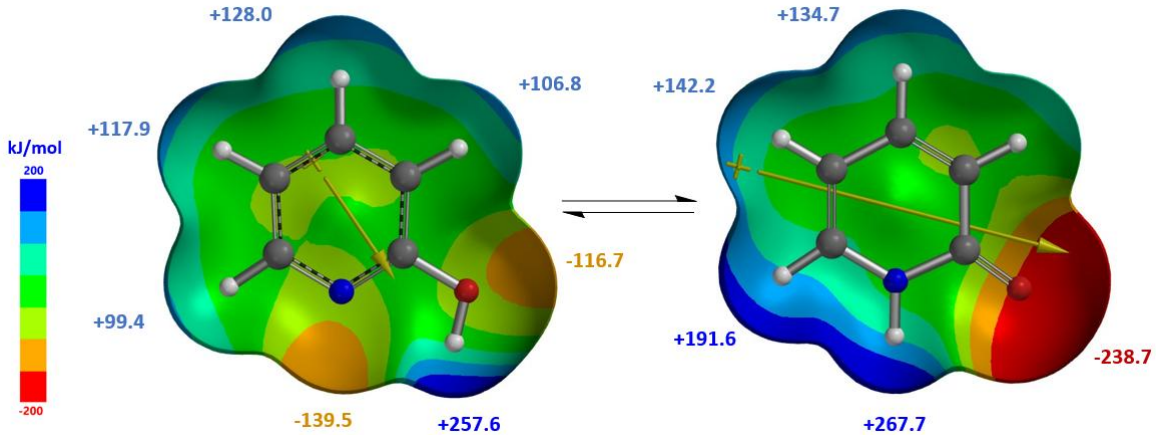


Figure 4. 2-Hydroxypyridine and 2-pyridone (Diploe/Natural Charges/Electrostatic Potential Map)

Moving from chloroform to acetonitrile and water, the changes in K_{eq} are larger than what will be expected from differences in changes in solvents' dielectric constant. How to account for it? We decided to focus further analyses of the 2-hydroxypyridine/2-pyridone tautomerization in the presence of water.

Solvent	Expt K_{eq}	Dipole Moment; enol form (Debye)	Dipole Moment; keto form (Debye)	Dielectric Constant
Cyclohexane	1.7	1.65	5.29	2.02
Chloroform	6.0	1.83	5.97	4.71
Acetonitrile	148	1.98	6.53	37.5
Water	910	1.99	6.57	78.3

Figure 5. Equilibria between 2-hydroxypyridine and 2-pyridone and their calculated dipole moments

For the following analyses, it is important to understand:

a. how solvation will affect the **energy barrier** for tautomerization between 2-hydroxy-pyridine and 2-pyridone?

b. how water could solvate 2-pyridone, change their **relative energy**, shifting the tautomeric equilibrium from the hydroxy to the keto form?

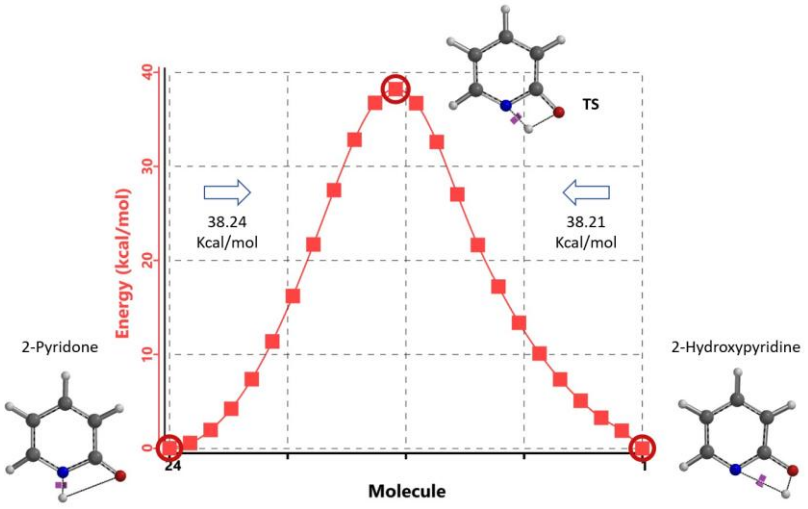


Figure 6. Intrinsic Reaction Coordinate between non-solvated 2-hydroxypyridine and 2-pyridone

Our DFT Transition State [5] and Intrinsic Reaction Coordinate calculations show that for non-solvated equilibration between 2-hydroxypyridine and 2-pyridone, the energy barriers are 38.21/38.24 kcal/mol. Very high energy hurdles (Figure 6). This value is significantly reduced to 12.12/16.74 kcal/mol in the presence of a single water molecule bridging the N-H and C=O part of 2-pyridone (Figure 7). Further change to about 13.17/17.51 kcal/mol for the doubly solvated form (Figure 8). The presence of water molecules strongly reduces the tautomerization barrier between the tautomeric forms, and 2-pyridone water complexes are more stable than the 2-hydroxypyridine ones.

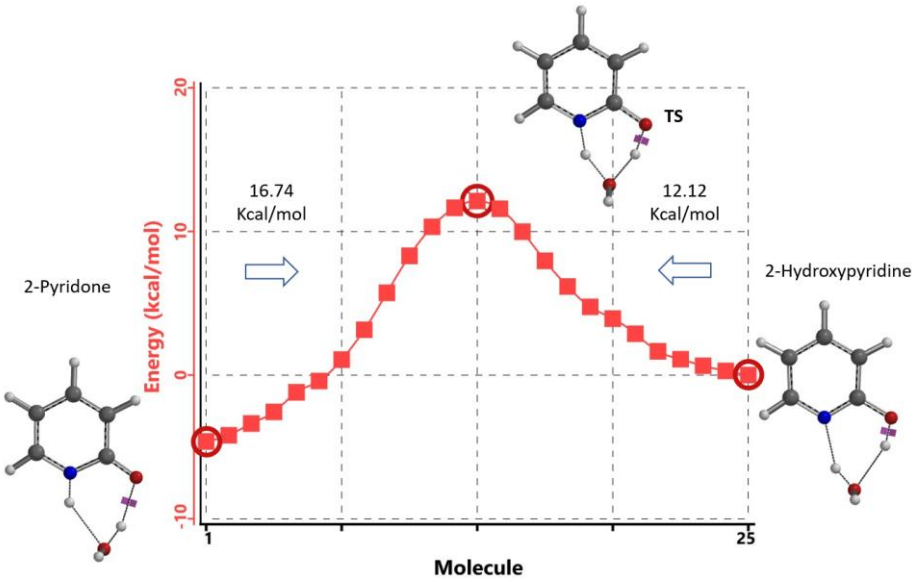


Figure 7. Intrinsic Reaction Coordinate between 2-hydroxypyridine and 2-pyridone monohydrate

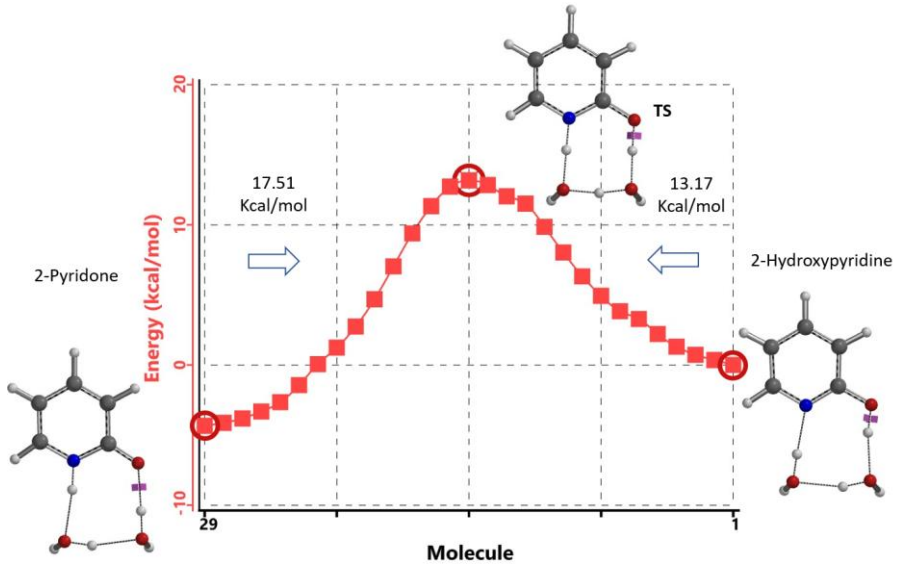


Figure 8. Intrinsic Reaction Coordinate between 2-hydroxypyridine and 2-pyridone dihydrate

Water solvates 2-pyridone *via* formation of hydrogen bonds and shifts the tautomeric equilibrium from the hydroxy to the keto form. Calculations with a microsolvation model show that the solvation is an energetically highly favorable process, -5.08 and -6.69 kcal/mol, respectively with incorporation of the first and second water molecule (Figure 9). The presence of these 2-hydroxypyridine and 2-pyridone hydrated complexes with one and two water molecules and the equilibrations catalyzed by water are supported experimentally with results from Molecular Beam Fourier Transform Microwave Spectroscopy [6].

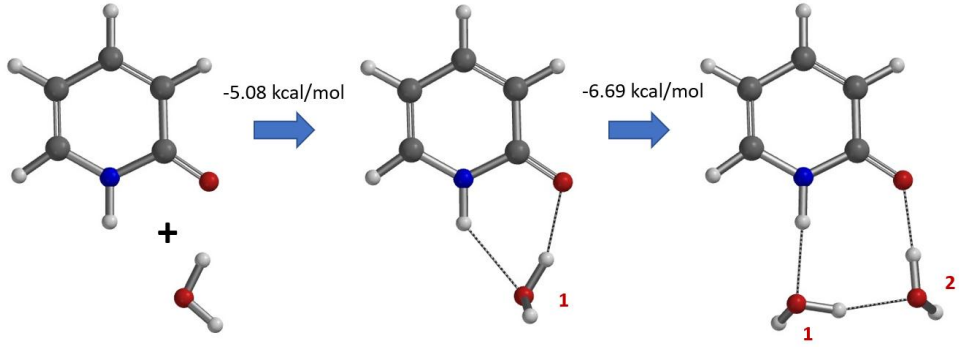


Figure 9. Microsolvation models of 2-pyridone and water

The above microsolvation models provide us semi-quantitatively a picture on how hydrogen bonding shifts K_{eq} of 6 in $CHCl_3$ to 910 in water, a more than 150-folds change, and the variations with K_{eq} values reported for polar aprotic solvents, and how small amount of water complicates things. More realistic models shall have a cluster of water molecules around 2-pyridone as shown in Figure 10 for more precise evaluation of energy change in solvation.

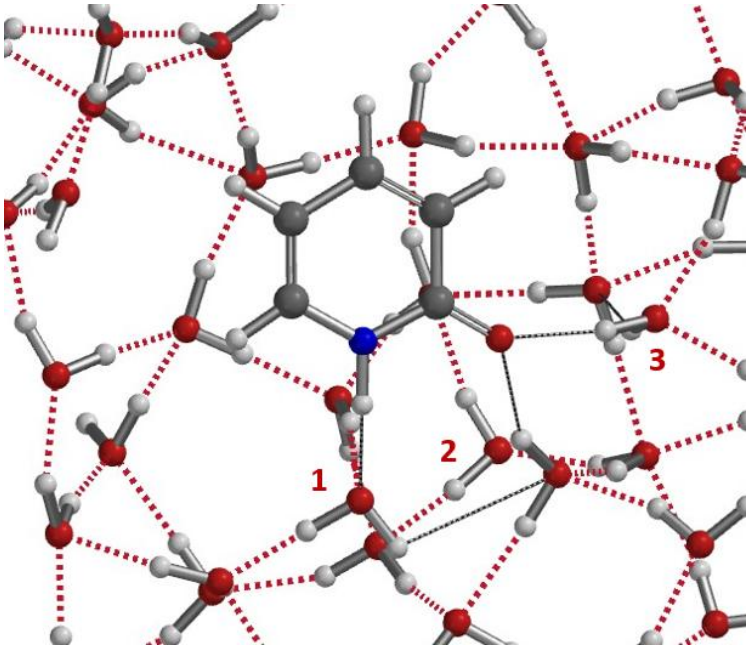


Figure 10. Water cluster model. Very likely more than two water molecules in the solvation complex.

For equilibration in the absence of water/protic solvent, our calculation revealed a **concerted** double proton transfer. Interestingly, the transition state (with single iFreq of 889 cm^{-1}) calculated suggests an **asynchronous** process, with the N-H bond distances of 1.433 Å *vs* 1.206 Å and the O-H bond distances of 1.283 Å *vs* 1.089 Å (Figure 11) [7].

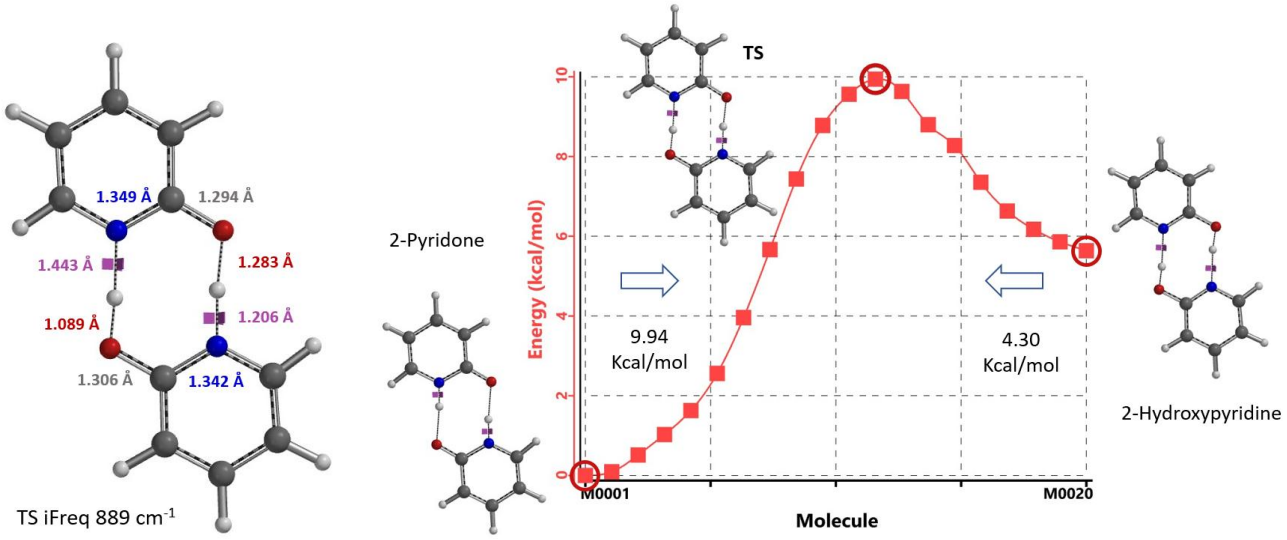


Figure 11. Transition State and IRC for equilibration in the absence of water/protic solvent

DFT analyses help us to understand why K_{eq} of 2-hydroxypyridine and 2-pyridone is highly solvent dependent, and why a little water complicates analyses in polar aprotic solvents.

Tautomerism of 10-hydroxy-1,7-phenanthrolines

10-Hydroxy-1,7-phenanthrolines, structurally related to 4-pyridone, exhibit unique tautomerism features (Figure 12). The system was studied by Katritzky *et al.* extensively in the nineties [8, 9, 10], found to be less sensitive to change in polarity of solvents, and merits revisit with further QM analyses. "Generate Tautomers" tool [11] added tautomer **C** to the list of two, **A** & **B**, reported earlier for consideration.

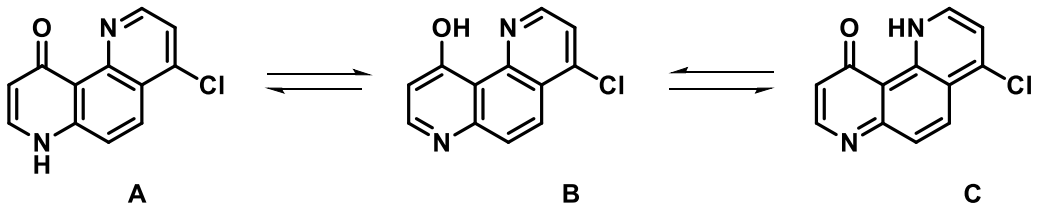


Figure 12. 4-Chloro-10-hydroxy-1,7-phenanthroline tautomers

Our DFT calculated electron density contour maps of 4-chloro-10-hydroxy-1,7-phenanthroline show strong intramolecular hydrogen bonding in enol form **B** and keto form **C**. There is significantly less electron density between the C=O oxygen and pyridine N in keto form **A**. Infrared red spectra in nujol, CCl_4, or $CHCl_3$ exhibit broad absorption ca. 2300–3500 cm^{-1}, indicative of strong hydrogen bonding, suggesting that tautomer **A** is a minor species [8].

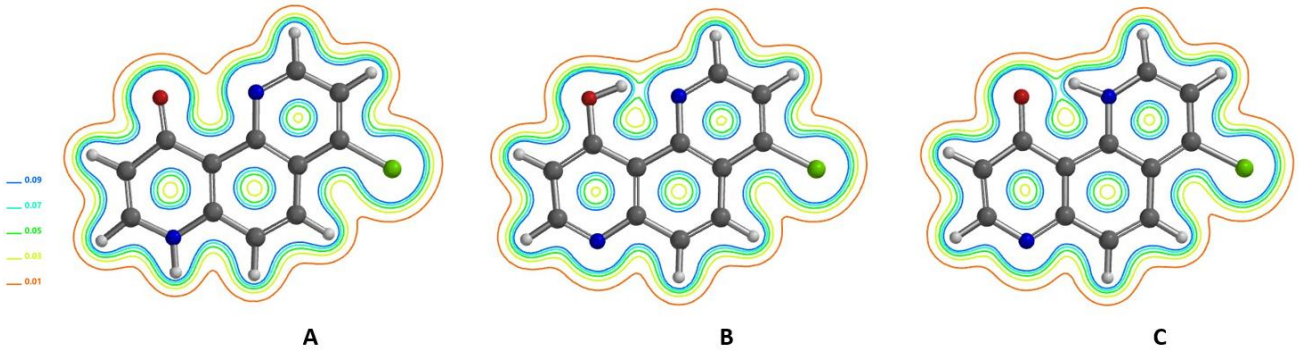


Figure 13. Electron density contour maps of 4-chloro-10-hydroxy-1,7-phenanthroline tautomers **A**, **B**, **C**

DP4[12] analysis of QM calculated ^{13}C NMR of 4-chloro-10-hydroxy-1,7-phenanthroline **A**, **B**, **C** tautomers *versus* experimental data[10] (Figure 14) provides DP4 scores of 0%, 100%, 0%, respectively; and DFT calculated equilibrium geometry relative energy of 6.33, 0.00, 5.83 kcal/mol, respectively, provide strong support that the intramolecularly hydrogen bonded enol **B** is the predominate tautomer. Among the two keto forms, tautomer **C**, with an intramolecular hydrogen bond, is lower in energy than tautomer **A**.

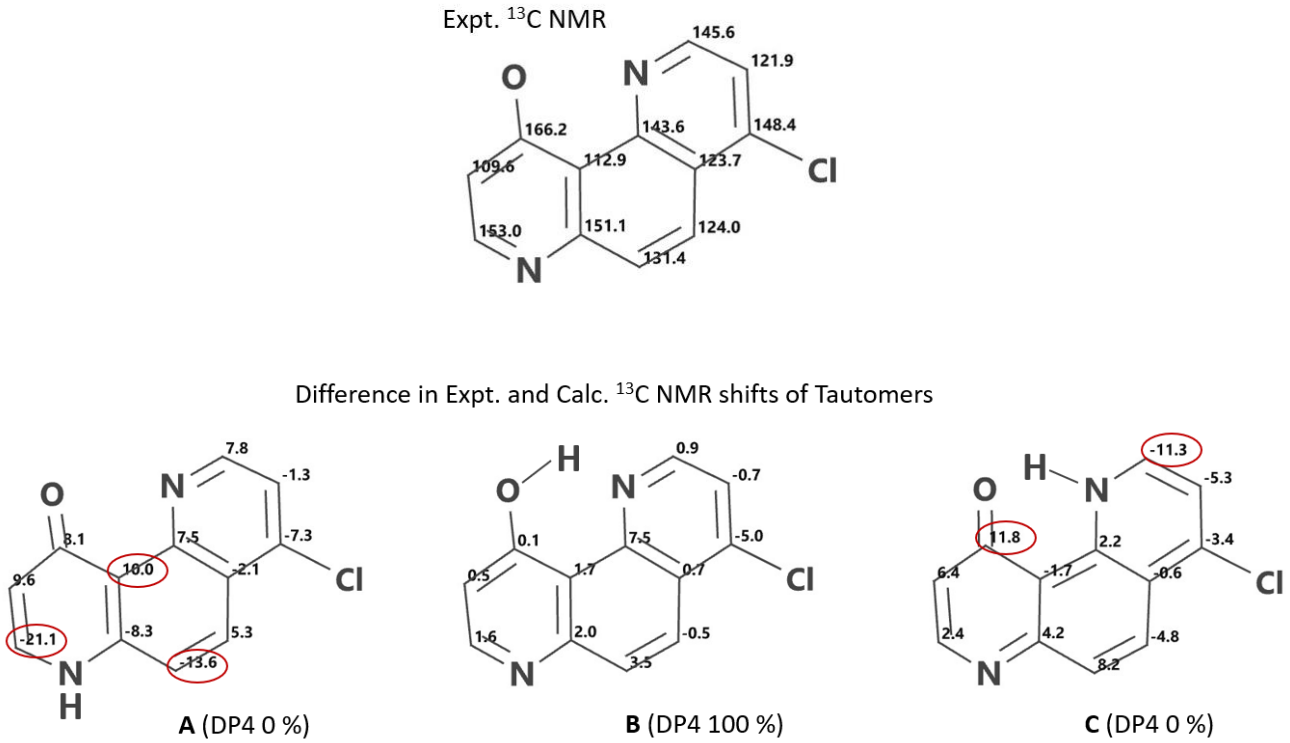


Figure 14. Expt ^{13}C NMR of 4-chloro-10-hydroxy-1,7-phenanthroline
Difference in Expt. and Calc. ^{13}C NMR shifts of tautomers **A**, **B**, **C**, and their DP4 scores

Energy barriers calculated for equilibration are ~ 0.50 kcal/mol from keto **C** to enol **B** and 6.31 kcal/mol in reverse, suggesting that equilibration will proceed rapidly at room temperature.

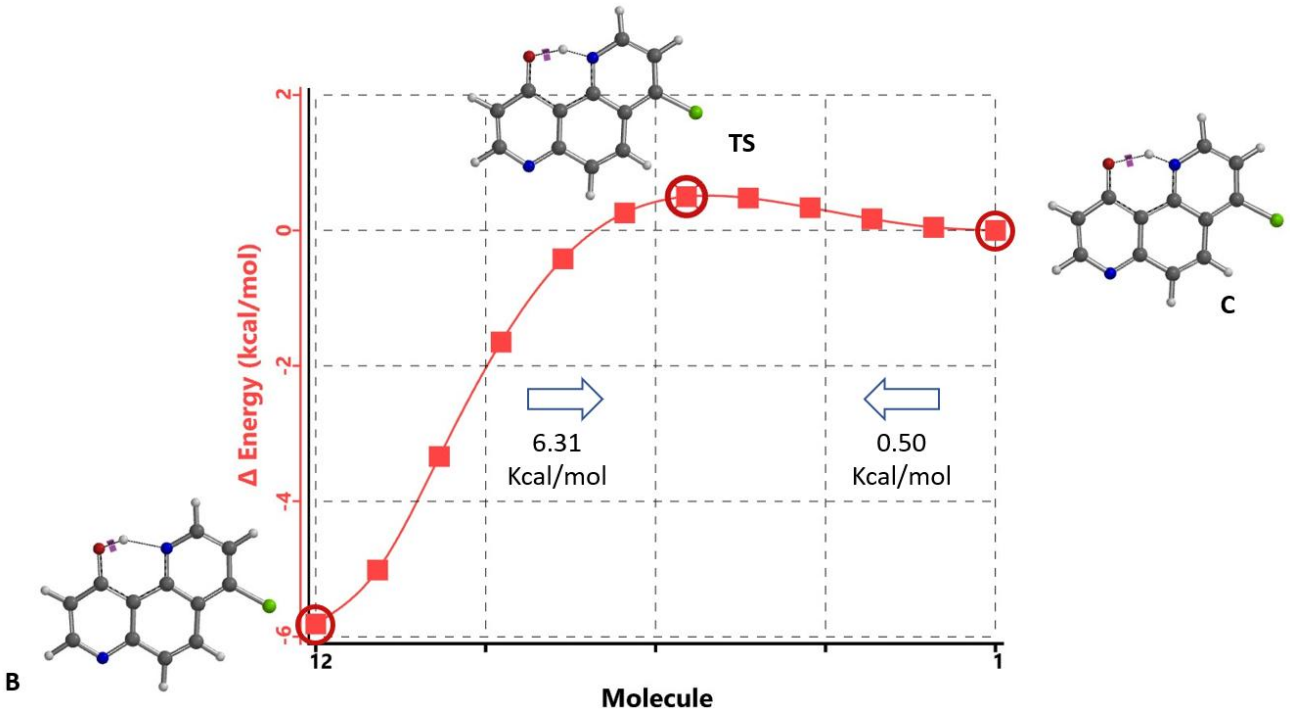


Figure 15. Intrinsic Reaction Coordinate between tautomer **B** and **C**

It is note-worthy to point out that, unlike the tautomerism between 2-hydroxypyridine and 2-pyridone, for tautomers with strong intramolecular hydrogen bond, in general, the K_{eq} are less sensitive to change in polarity of solvents [10].

Retrospect

Tautomer identification is especially important in dealing with heterocyclic compounds where two or more different tautomers may exist in equilibrium, and the identity of the dominant species may not be intuitively obvious. Precise experimental determination and QM analyses of tautomerization equilibria are significant challenges. Complexity and uncertainty multiply with tautomeric structures in MedChem design. Local charged or polar groups in the binding pockets can shift the tautomer ratios and binding to the higher energy tautomer in free solution. *All these make tautomerization easy to ignore but overwhelming to consider*[1]. QM may not be able to precisely quantify the small energy difference between tautomers in solution, yet it is highly reliable in differentiating tautomers with ≥10 kcal/mol difference. It is unlikely that 10 kcal/mol could be compensated by protein binding to arrive at low nano molar potency candidates suitable for clinical development, while a tautomeric energy difference of a few kcal/mol could be more readily compensated. Binding with tautomeric forms which are significantly different in energy could benefit from Fragment Molecular Orbital analyses [13, 14, 15]. QM tools for tautomerization analyses are readily accessible and shall be used routinely for target design and prioritization.

Building on What We Just Learned

γ-Hydroxybutenolide is an important structural core found in many bioactive marine natural products such as Palauolide [16], Cacospongionolide A, Dysidiolide, Manoalide [17], etc., and considered to be the key pharmacophore for this class of compounds with diverse biological properties.

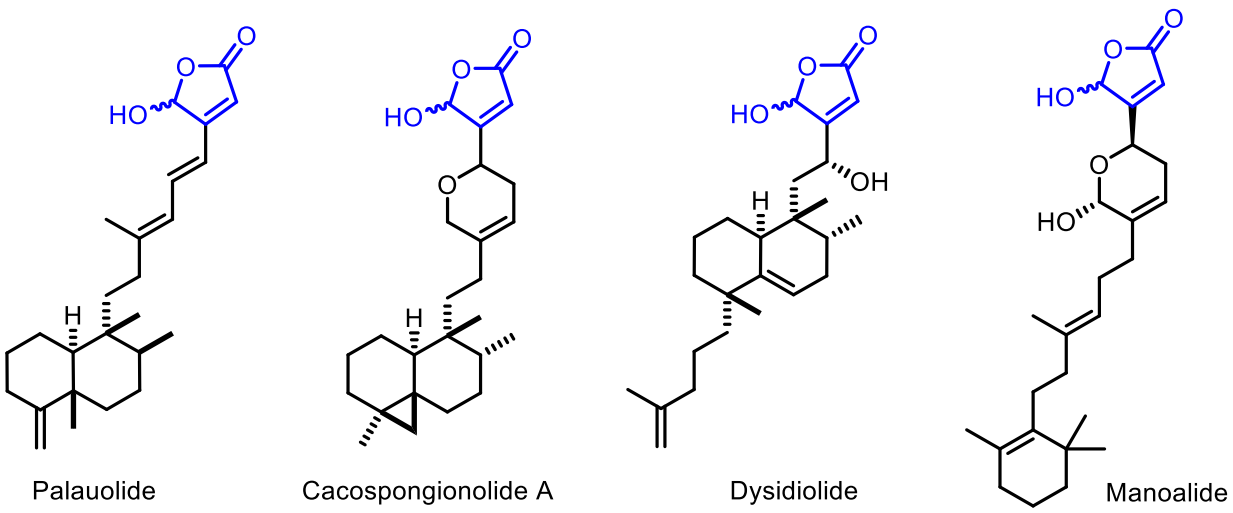
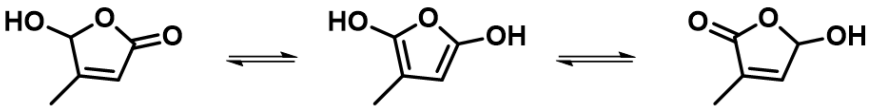


Figure 16. Selected natural products with γ-hydroxybutenolide moiety

Looking at the structures of these important natural products, obvious questions arise.

a. Why the hydroxyl group is always on the more substituted side of the butenolide?

b. Based on the relative energy of three tautomeric forms shown below, will γ-hydroxy-butenolide equilibrate to other tautomeric forms?



0.00 kcal/mol 23.50 kcal/mol 1.62 kcal/mol

Figure 17. Potential γ-hydroxybutenolide tautomerization mechanism

Return to Table of Contents

References:

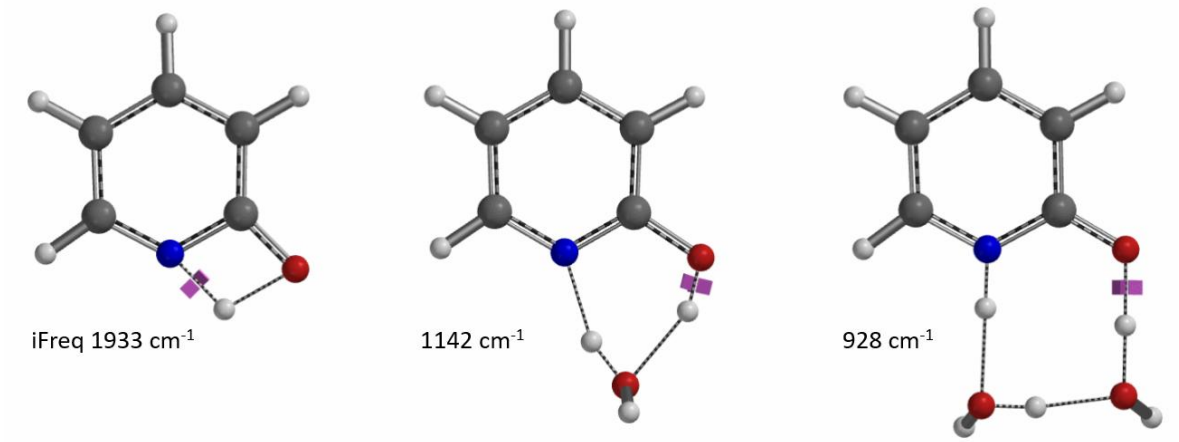
[1] Y. C. Martin *J. Comput. Aided Mol. Des.*, *2009*, *23*, 693.

[2] E.V. Anslyn, D.A. Dougherty (2006) *Modern Physical Organic Chemistry*. New York, NY, USA: University Science Books; page 155.

[3] For DFT calculation of tautomers, ωB97X-V/6-311+G(2df,2p) is a more accurate model than the standard ωB97X-D/6-31G* one we routinely used for other QM calculations.

[4] *Spartan'20 Tutorial and User's Guide* (2020). Irvine, CA, USA: Wavefunction, Inc. page 50.

[5] Transition States calculated for tautomeric equilibration between 2-hydroxypyridine and 2-pyridone in the presence of zero, one, and two water molecules.



[6] S. Mata, V. Cortijo, W. Caminati, J. L. Alonso, M. Eugenia Sanz, J. C. López, S. Blanco *J. Phys. Chem. A* **2010**, *114*, 11393.

[7] A. Fu, H. L. Li, D. M. Du, Z. Y. Zhou *J. Phys. Chem. A* **2005**, *109*, 1468. This paper studied the 2-pyridone dimer tautomer equilibration with B3LYP and BH-LYP hybrid density functional methods at the 6-311++G (2d, 2p) basis set level. Results suggest the process to be **concerted** and **synchronous**. E. Torres, G. A. DiLabio, *J. Phys. Chem. Lett.* **2012**, *3*, 1738 noted that B3LYP is not able to reliably model systems in which noncovalent interactions are important (hydrogen bonding, π-stacking, etc.).

[8] I. Shcherbakova, J. Elguero, A. R. Katritzky *Advances in Heterocyclic Chemistry,* **2000**, 51.

[9] A. R. Katritzky, M. Karelson, P. A. Harris *Heterocycles* **1991**, *32*, 329.

[10] A.R. Katritzky, I. Ghiviriga, P. Leeming, F. Soti *Magnetic Resonance in Chemistry*, **1996**, *34*, 518.

[11] Tautomer Generators

a. *Spartan'20 Tutorial and User's Guide* (2020). Irvine, CA, USA: Wavefunction, Inc. page 269. Spartan'20 tautomer generating tool is limited to tautomers involving nitrogen, oxygen, phosphorus, and sulfur. Tautomers involving carbon have intentionally been excluded. And as such Spartan'20 "Generate Tautomer" does not alert the possibility of the keto form as shown below.

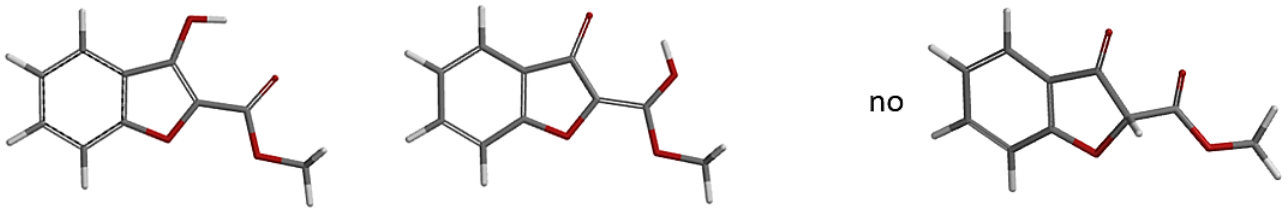


Figure 18. Spartan'20 Tautomer Generation Tool

b. D. K. Dhaked, W.-D. Ihlenfeldt, H. Patel, V. Delannée, M. C. Nicklaus *J. Chem. Inf. Model.* **2020**, *60*, 1253. Tautomerizer is a more thorough tautomer generator. This web tool is available at https://cactus.nci.nih.gov/tautomerizer/

[12] S. G. Smith, J. M. Goodman *J. Am. Chem. Soc,* **2020,** *132*, 12946. T.-H. Duong, M. A. Beniddir, J. Boustie, K.-P.-P. Nguyen, W. Chavasiri, G. Bernadat, P. Le Pogam *Molecules*, **2019**, *24*, 1527.

[13] M. Wieder, J. Fass, J. D. Chodera *Chem. Sci.* **2021**, *12*, 11364.

[14] H. Brandstetter. F. Grams, D. Glitz, A. Lang, R. Huber. W. Bode, H-W. Krell, R. A. Engh *J. Biol. Chem.* **2001**, *276*, 17405. Based on a 1.8-Å crystal structure of human neutrophil collagenase (MMP-8) in complex with an active site-directed inhibitor (RO200-1770; https://www.rcsb.org/structure/1jj9), authors concluded that the inhibitor binds in the enol form, which is 17.14 kcal/mol higher in energy than the keto form.

[15] For recent reviews on Fragment Molecular Orbital Analyses, see Alexander Heifetz editor Quantum Mechanics in Drug Discovery, 2020, Springer Science+Business Media, LLC, New York, NY 10004, U.S.A. Guiding Medicinal Chemistry with Fragment Molecular Orbital (FMO) Method. pp. 37-48, Analyzing Interactions with the Fragment Molecular Orbital Method pp. 49-73.

[16] T. Mai, J. Toullec, S. Van Wynsberge, M. Besson, S. Soulet, S. Petek, E. Aliotti, M. Ekins, K. Hall, D. Erpenbeck, D. Lecchini, M. A. Beniddir, D. Saulnier, C. Debitus *Fish. Aquatic Sci.* **2019**, *22*, 30.

[17] S. Tilvi, S. Khan, M.S. Majik *Current Org. Chem.* **2019**, *23*, 2436.

Chapter 50 Mechanism of a highly selective N2 alkylation of indazole

Liting Dong, Qiuyue Wang, Yongsheng Chen, John S. Wai

Pfizer chemists reported in 2022 a highly selective N2 alkylation of indazoles with primary, secondary, and tertiary alkyl 2,2,2-trichloroacetimidate in *Synthesis*[1]. Diverse 1H-indazoles and azaindazoles are selectively alkylated at N2 and no N1 isomer was observed. The reaction mechanism proposed by the authors, as shown in Figure 1, involves the alkyl 2,2,2-trichloroacetimidate undergoing protonation on the imide under acidic conditions. This activation facilitates nucleophilic displacement by the N2-indazole nitrogen, leading to the desired product after re-aromatization. We conducted quantum mechanical (QM) analyses to understand the high selectivity for N2 alkylation.

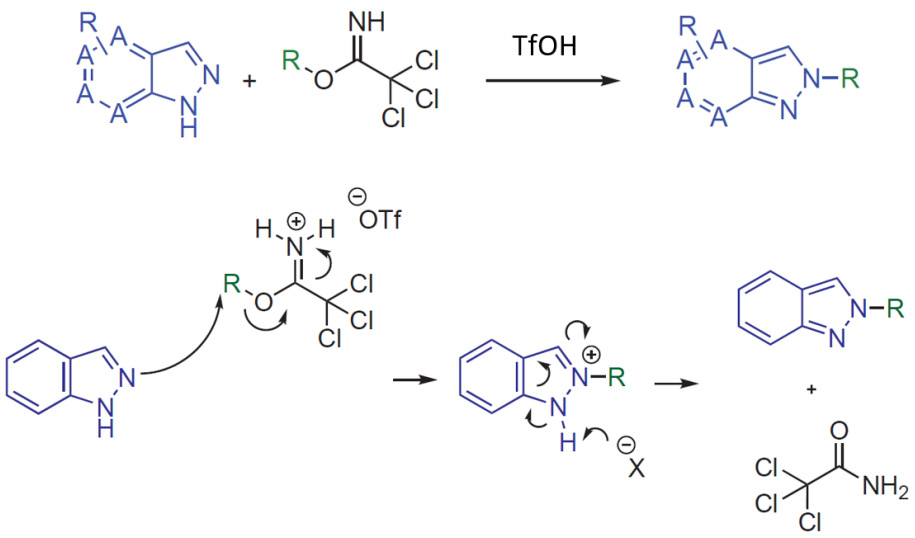
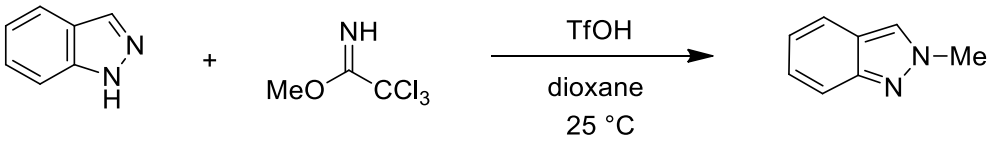


Figure 1. Proposed reaction mechanism for primary/secondary and tertiary alkyl 2,2,2-trichloroacetimidates (R is primary, secondary, and tertiary alkyl groups)[2]

QM analysis of the reaction mechanism



Initially, we selected two relatively simple structures in the article, indazole and methyl trichloroacetimidate as models. Using Spartan 20'(DFT B97X-D/6-31G*), we calculated the reaction energy profiles for N1 and N2 alkylation. We estimated that the activation energy for N1 alkylation is 12.76 kcal/mol (Figure 2), and for N2 alkylation, it is at 13.87 kcal/mol

(Figure 3). However, the lower activation energy for N1 alkylation by 1.11 kcal/mol contradicts the experimental observation of selective N2 alkylation. This discrepancy led us to explore the tautomeric forms of the indazole substrate, namely 1-H indazole and 2-H indazole, to understand their influence on the reaction selectivity.

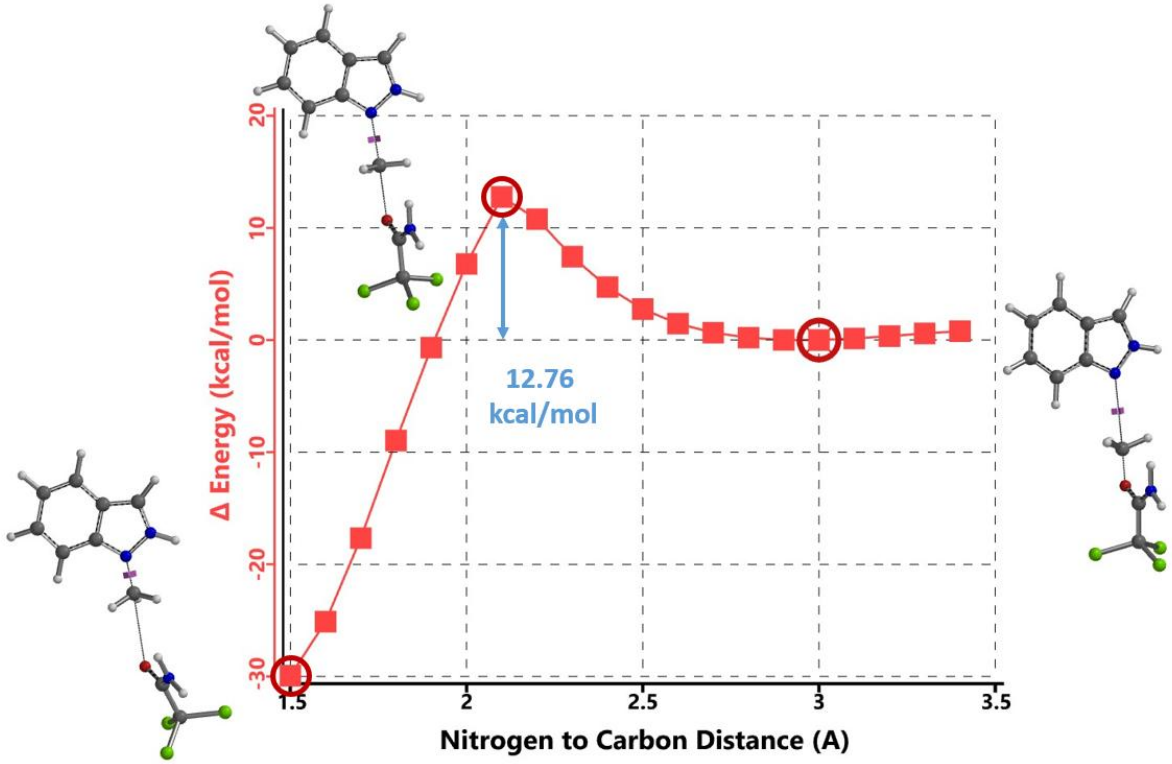


Figure 2. Reaction Energy Profile of N1 alkylation of Indazole

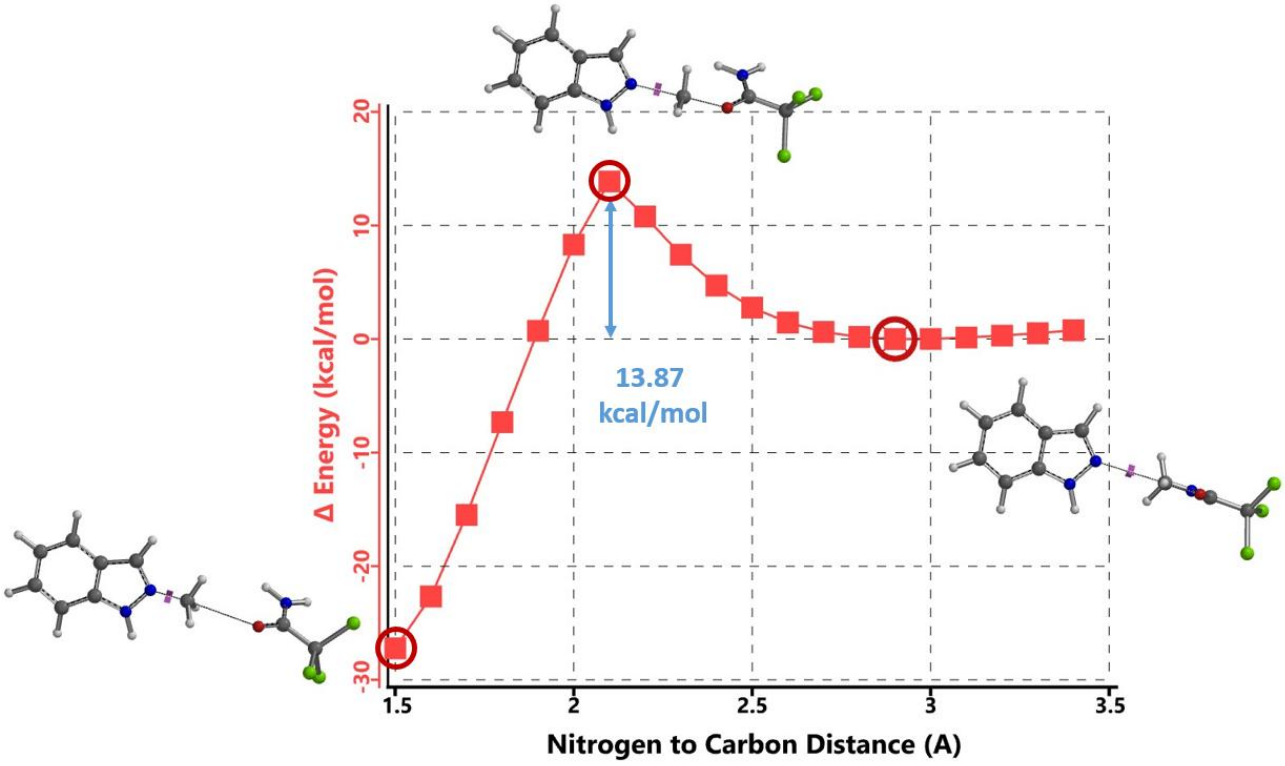


Figure 3. Reaction Energy Profile of N2 alkylation of Indazole

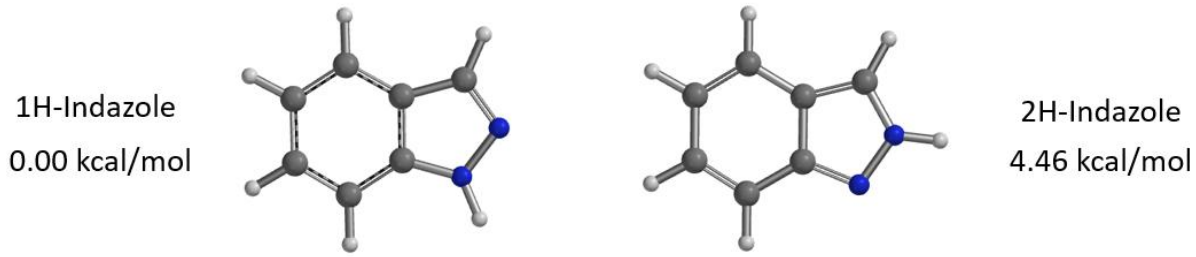


Figure 4. Relative energy of 1H- *vs* 2H-indazole tautomer

Calculations reveal that 1-H indazole is energetically more stable than its 2-H indazole tautomer by 4.46 kcal/mol (Figure 4). This energy difference must be included into the analysis of the two competing paths (Figure 5). For N1 alkylation, the substrate must convert from the lower-energy 1-H tautomer to the higher-energy 2-H form, which incurs an additional energy cost. Consequently, the total reaction energy barrier for N1 alkylation is 17.22 kcal/mol, which is 3.35 kcal/mol higher than that for N2 alkylation. This accounts for the high N2 selectivity observed (Figure 7)[3].

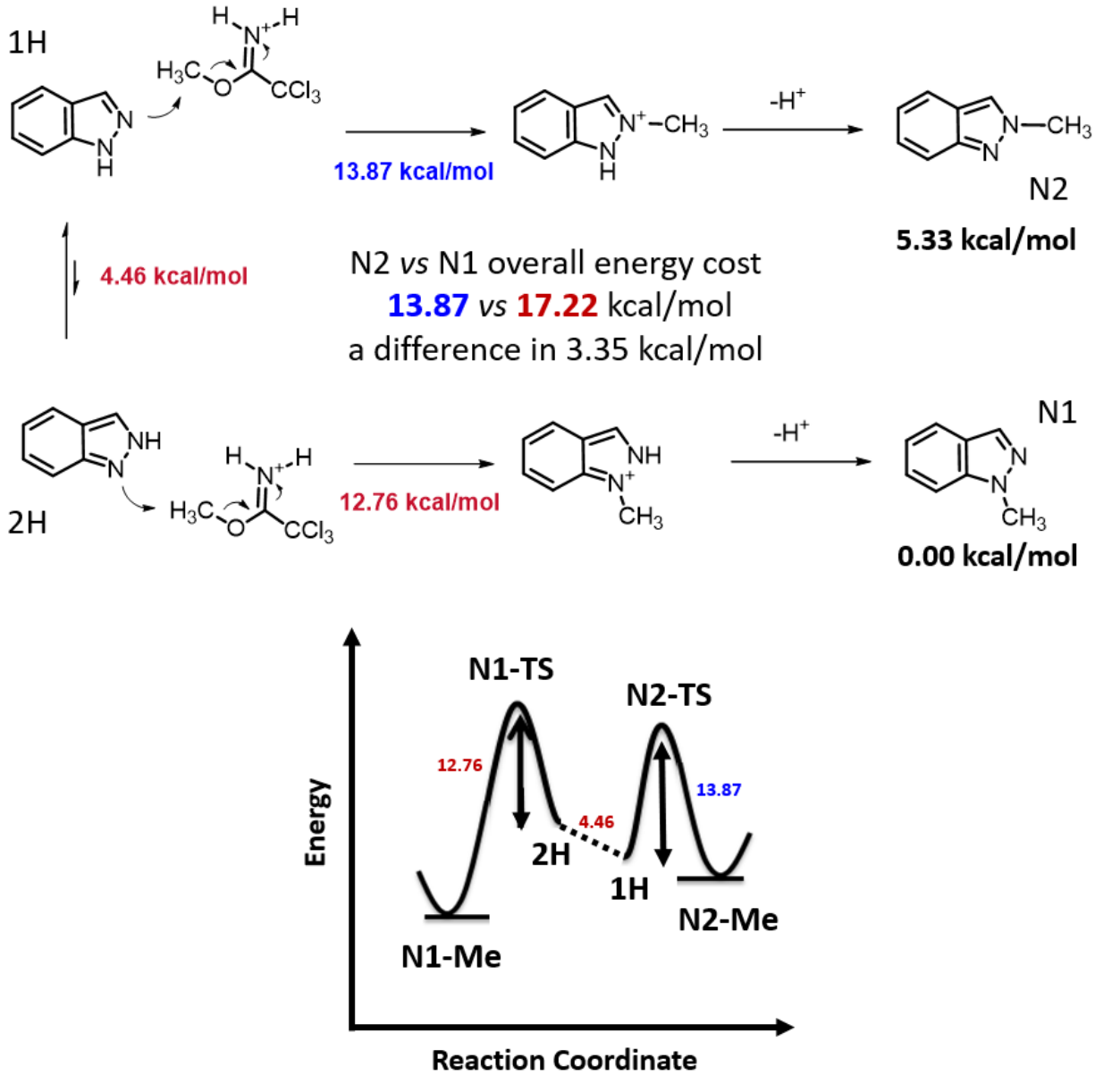


Figure 5. Comparison of energy changes in N1 vs N2 alkylation of Indazole

Transition state energy comparison

Based on the geometries at the highest energy points from the two reaction energy profiles (Figures 2 & 3), we calculated the relative energies of the transition states (TS) for further comparison. Each calculated TS exhibited only one imaginary frequency, corresponding to the vibration of the bonds being formed and broken (Figure 6)[4]. The energy difference between the N1 and N2 transition states is 3.42 kcal/mol, favoring N2. This is consistent with the earlier estimated energy difference that accounts for the observed N2 selectivity.

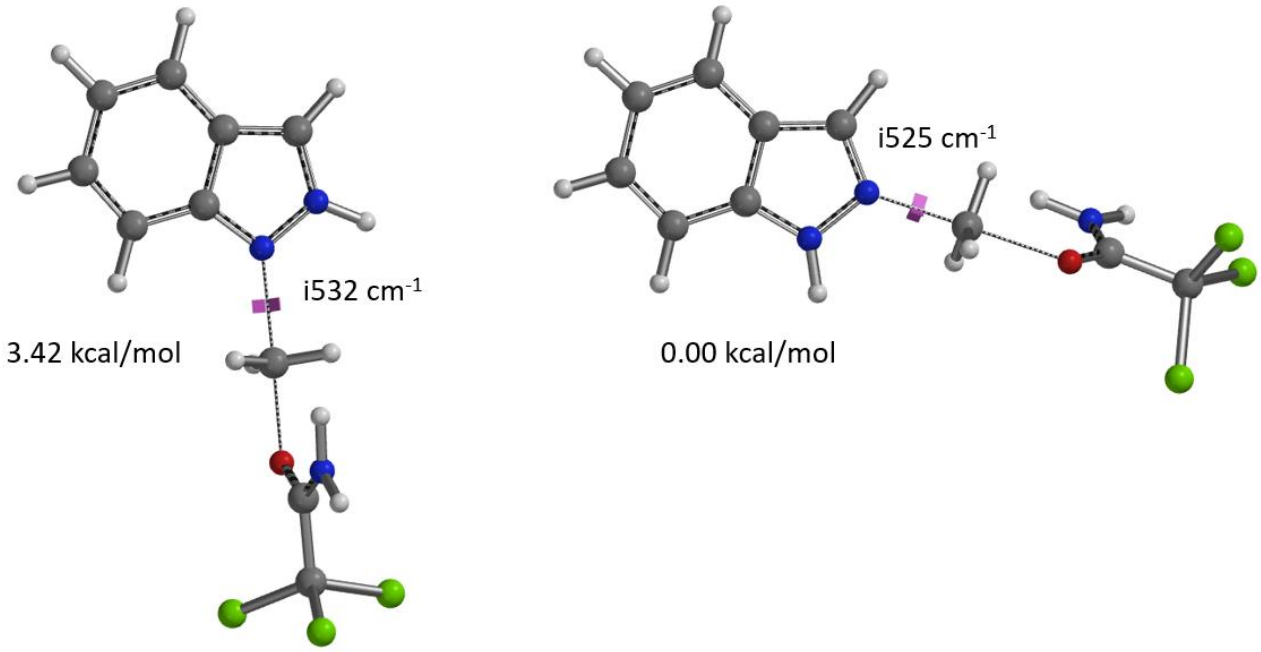


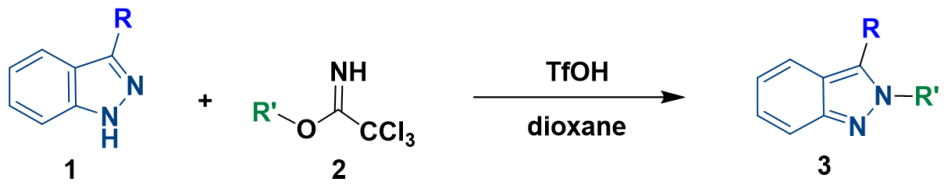
Figure 6. Relative energy comparison of indazole N1 and N2 alkylation transition states

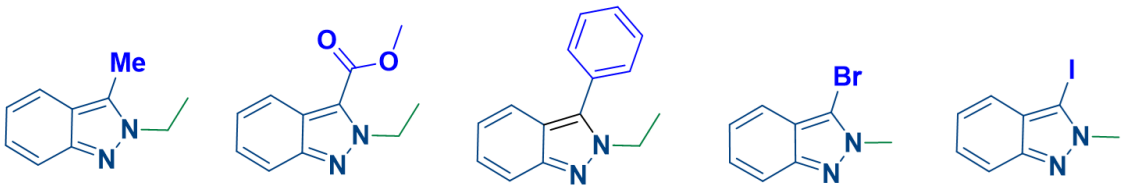
The difference translates to an N2 vs N1 product ratio of 322:1 for a reaction at 25 °C, aligning with experimental results. With the energy difference between the TS greater than 3 kcal/mol, there is no practical difference if the reaction was run under reflux (dioxane b.p. 101 °C). The calculated ratio of N2 vs N1 products is still as high as 100 to 1 (Figure 7). This demonstrates an intrinsic high selectivity for N2-alkylation.

T (°C)	Relative Energy Difference	(E1-E2)/RT	N2/N1 product ratio
25	3.35	5.66	286:1
25	3.42	5.78	322:1
101	3.42	4.60	100:1

Figure 7. Product ratio calculated for the reaction at 25 and 101 °C

3-Substituted indazoles





3A: 71% 3B: 85% 3C: 80% 3D: 37% 3E: 76%

The paper reported that alkylation of indazoles with large C3 substituents (compounds **3A** to **3C**) remains highly N2 selective. We were interested in alkylation of 3-bromo and 3-iodo-indazole (compounds **3D** & **3E**) and found them to exhibit high N2 selectivity as well.

Table 1 presents the calculated relative energies of the 1-H and 2-H tautomers with various C3 substitutions as shown in column 3. The 2-H tautomers are consistently higher in energy, approximately 4 kcal/mol compared to the corresponding 1-H tautomers. This energy difference is even more pronounced for the 3-bromo and 3-iodo indazoles, reaching around 6 kcal/mol.

Activation energy calculated (by adding the energy difference between the tautomers and energy barrier from reaction energy profile calculations) for N1 alkylation is usually higher than that for N2 alkylation (column 6). However, there are outliers (highlighted in red) that are not consistent with observations. We attributed this to the less precise estimates of relative energies between tautomers[5]. In contrast, calculated transition state energy differences **consistently correlate** with experimental observations, revealing that N2 alkylation is generally favored by >3 kcal/mol, (column 7). For 3-*tert*-butylindazole, the difference is smaller, presumably due to steric effect.

R-3	2H N1 REP	+ 1H to 2H tautomer	= N1 Overall	1H N2 REP	REP N1 vs N2	TS N1 vs N2	B value
H	12.76	4.46	17.22	13.87	**3.35**	**3.42**	2.2
Me	11.68	4.42	16.10	12.78	**3.32**	**3.14**	7.1
i-Pr	11.49	4.13	15.62	12.53	**3.09**	**3.55**	11.1
t-Bu	11.45	4.11	15.56	16.16	**-0.60**	**1.29**	15.6
Br	17.13	6.99	24.12	23.85	**0.27**	**4.93**	8.5
I	14.43	6.03	20.46	23.34	**-2.88**	**5.35**	9.9
Ph	12.67	4.39	17.06	16.94	**0.12**	**5.27**	**na**
CO_2Me	15.06	4.21	19.47	20.05	**-0.58**	**8.40**	**na**

Table 1. Calculated data for alkylation of indazole with different C3 substituents

For indazoles with simple C3 alkyl groups, the relative energy difference between the TS for N1 and N2 alkylation correlates with the ***B*** values - a parameter established to quantify steric effects[6]. As ***B*** value increases, which indicates greater steric hindrance from the substituent, the energy difference becomes smaller. However, for 3-bromo and 3-iodoindazole, their corresponding ***B*** values would be suggestive of smaller energy differences between the N2 and N1 TS than the calculated ones. This can be attributed to a non-covalent interaction between the bromo or iodo group and NH of the protonated imidate, which stabilizes the N2 alkylation TS, increasing the energy difference between the two TS. This could be illustrated with the electrostatic potential (ESP) surfaces of the reacting partners (left side of Figure 9), with notable electrostatic interaction between the iodine atom of 3-iodoindazole and the imine hydrogen of the protonated imidate, akin to a halogen bond (right side of Figure 9)[7].

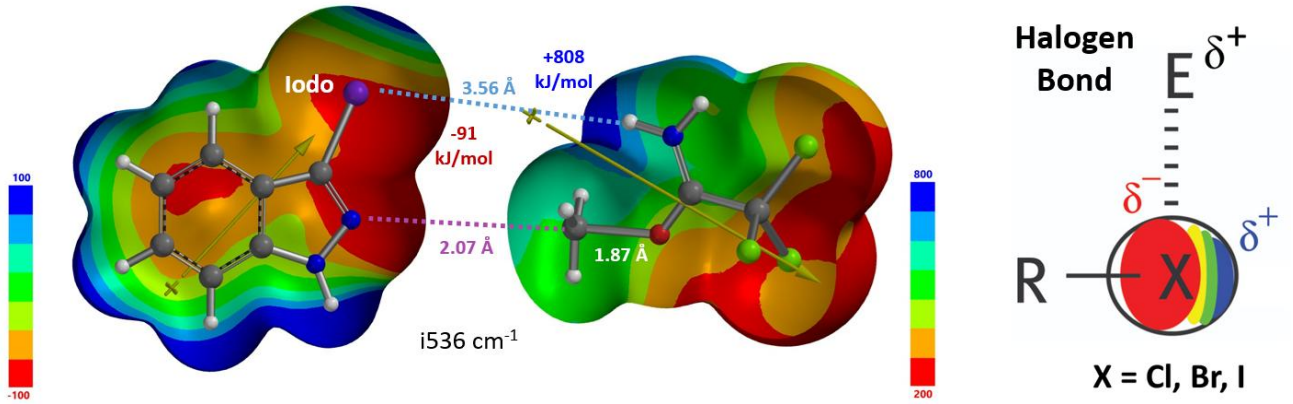


Figure 9. TS of 3-iodoindazole N 2 alkylation TS: ESP surfaces & dipole moments of indazole and imidate ester[8]

Although ***B*** values for phenyl and methoxycarbonyl groups were not reported[6], our calculations revealed the presence of non-covalent interactions (NCI) in the transition states (TS) of N2 alkylation of both the phenyl and ester substrates. Specifically:

1. **Phenyl indazole TS**: The electron density surface (EDS) of the TS (Figure 10, left) indicates an NCI between the C3 carbon of the phenyl group and the iminium H[9]. This interaction lowers the energy of the TS, facilitating N2 alkylation.
2. **Methyl indazole-3-carboxylate TS**: The EDS reveals a strong hydrogen bond between the carbonyl oxygen of and iminium H in the N2 TS (Figure 10, right), resulting in the highest energy difference between N1 and N2 TS of 8.40 kcal/mol in this study.

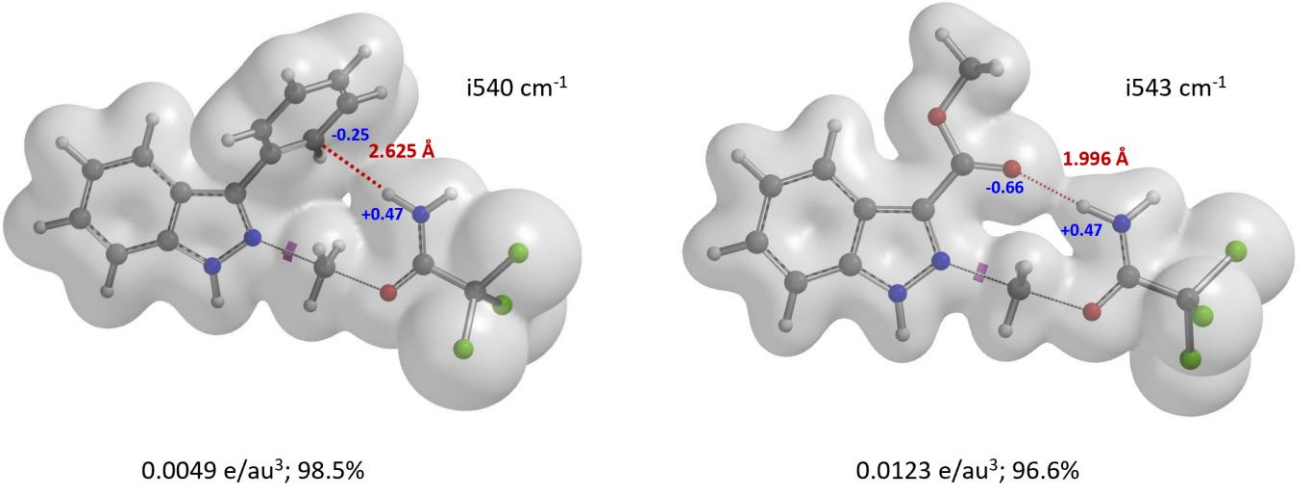
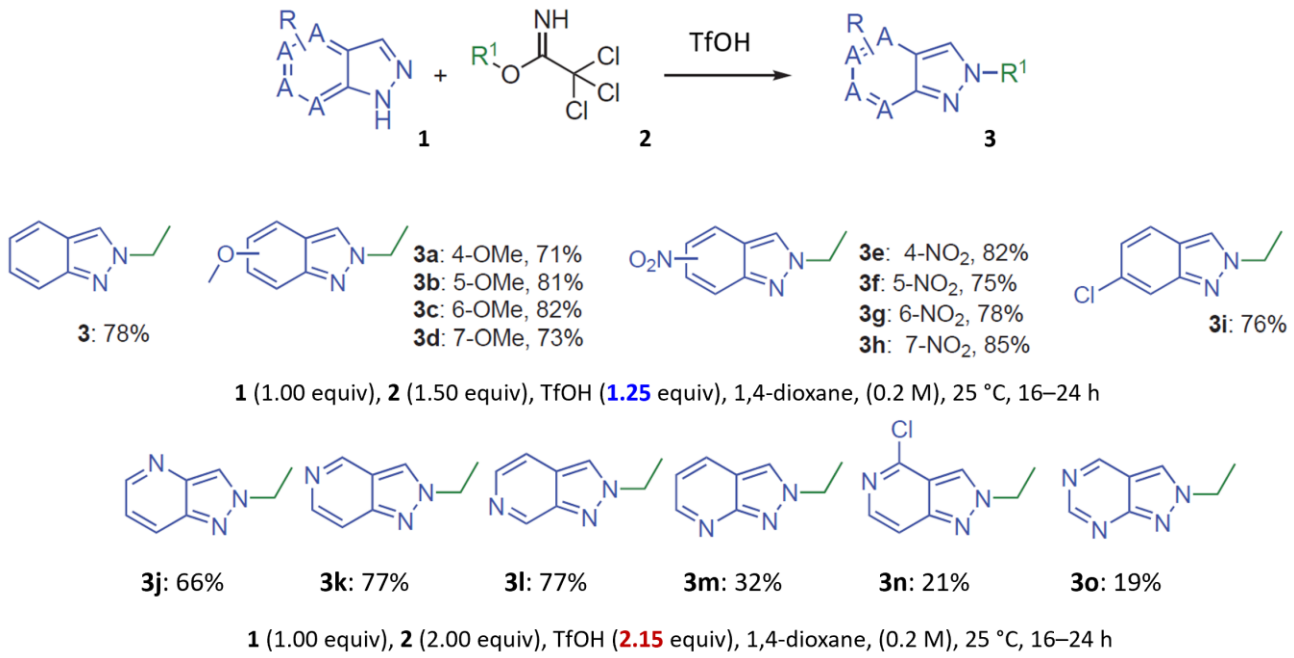


Figure 10. Electron density surface of the N2 alkylation TS: 3-phenylindazole & methyl indazole-3-carboxylate

Azaindazoles



For azaindazole substrates, the article mentioned that limited or no reaction was observed using **1.25 equivalents** of TfOH[1]. However, when TfOH was increased to **2.15 equivalents**, the N2 alkylation products were selectively formed. The authors proposed that the additional acid preferentially protonates the more basic nitrogen atoms in the pyridine or pyrimidine rings, thereby increasing the Lewis basicity of the N2 nitrogen[1]. Intuitively, protonation would render a nucleophile less reactive. We decided to apply the model developed above to look for more relevant explanations.

We'd like to understand why alkylation of azaindazoles with 1.25 equivalents of TfOH failed. There appear to be two ways to protonate the reacting species:

1. TfOH protonates the imidate and then reacts with the azaindazole. The calculated activation energy for the N2 methylation of azaindazole is 14.89 kcal/mol, only ~1 kcal/mol more than that of indazole. This could not account for the failed alkylation (Figure 11).

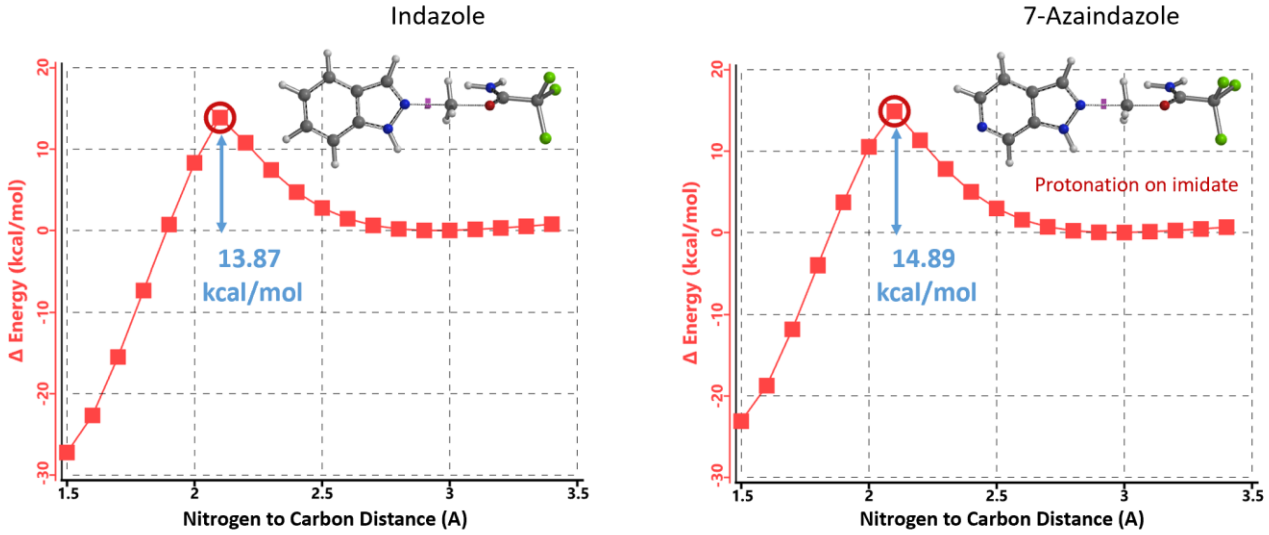


Figure 11. REP of N2 alkylation of protonated imidate with indazole vs 7- azaindazole

2. TfOH protonates the azaindazole and then reacts with the unactivated imidate.

Among the two possible sites for mono-protonation, equilibrium geometry calculation indicates that protonation on the **sp^2 nitrogen** is energetically more favorable than at the imine nitrogen, with a significant energy difference of 9.54 kcal/mol.

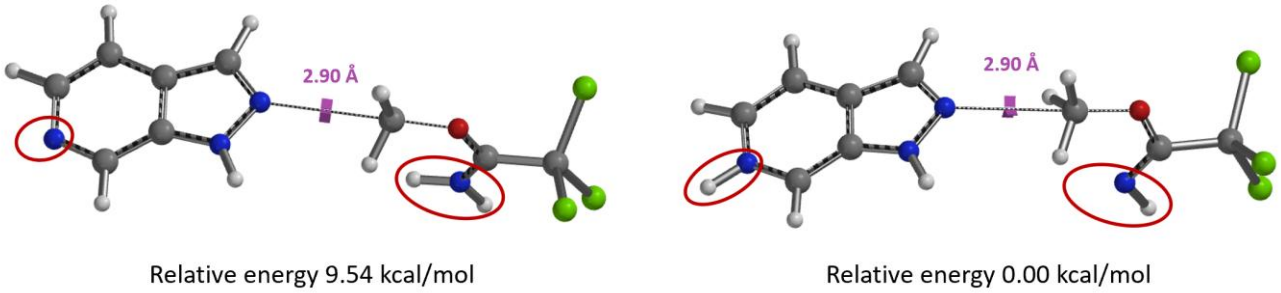


Figure 12. Relative energy of mono-protonation: on the imine N vs the sp2 N

The reaction energy profile for the alkylation of protonated azaindazole with unactivated imidate (Figure 13, right) indicates that the energy barrier of the reaction is 39.98 kcal/mol. This is 26.11 kcal/mol more than that with indazole, accounting for why the alkylation with azaindazole failed. When 1.25 equivalent of TFOH is used, it is the sp^2 pyridine nitrogen that gets protonated, leading to a decrease in the reactivity of azaindazole and the imidate is not activated for reaction.

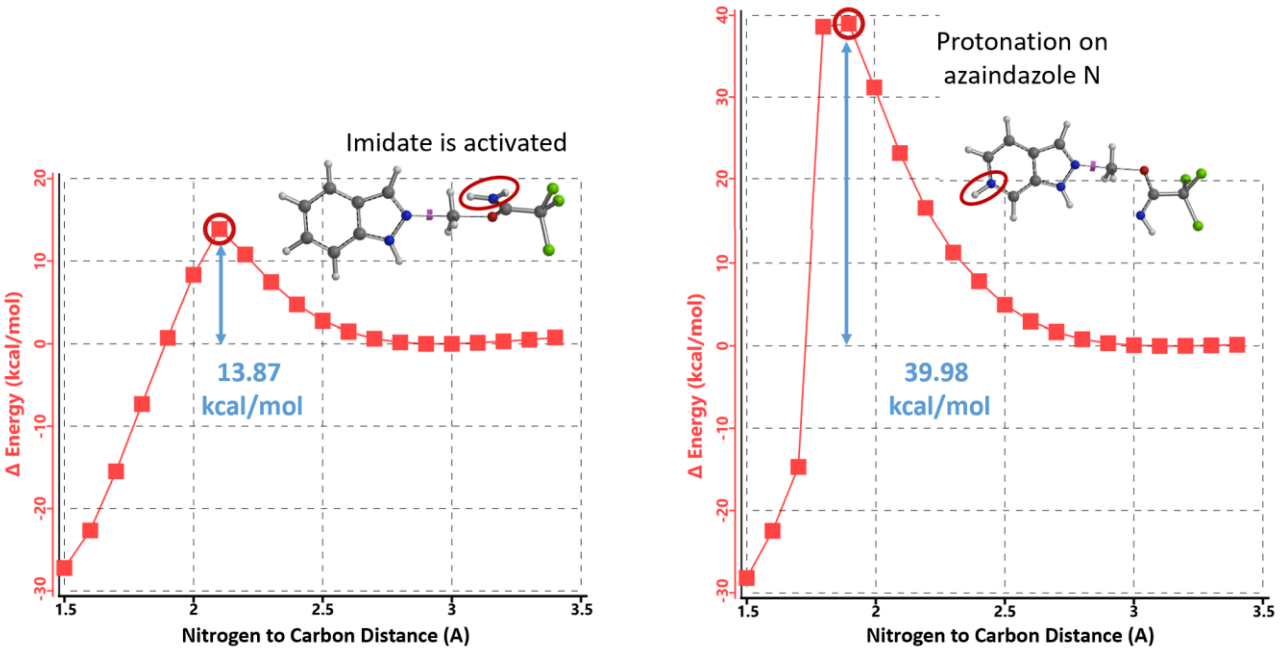


Figure 13. REP of N2 alkylation of protonated imidate with indazole vs unactivated imidate with protonated azaindazole

When the amount of TfOH is increased to 2.15 equivalents, it becomes possible for both the azaindazole and the imide ester to be protonated. The calculated energy barrier of 18.37 kcal/mol (Figure 14, right) suggests that the reaction is achievable. Despite the reduced reactivity of the protonated azaindazole, the activated imidate remains sufficiently reactive for alkylation to occur. The higher energy barrier is consistent with our observation that the reaction proceeds better when elevated from room temperature to 60 °C. Consequently, the original hypothesis posited by the authors—that the additional acid selectively protonates

the more basic pyridine/pyrimidine nitrogen, thereby enhancing the Lewis basicity of the N2 nitrogen[1]—does not hold up under these findings.

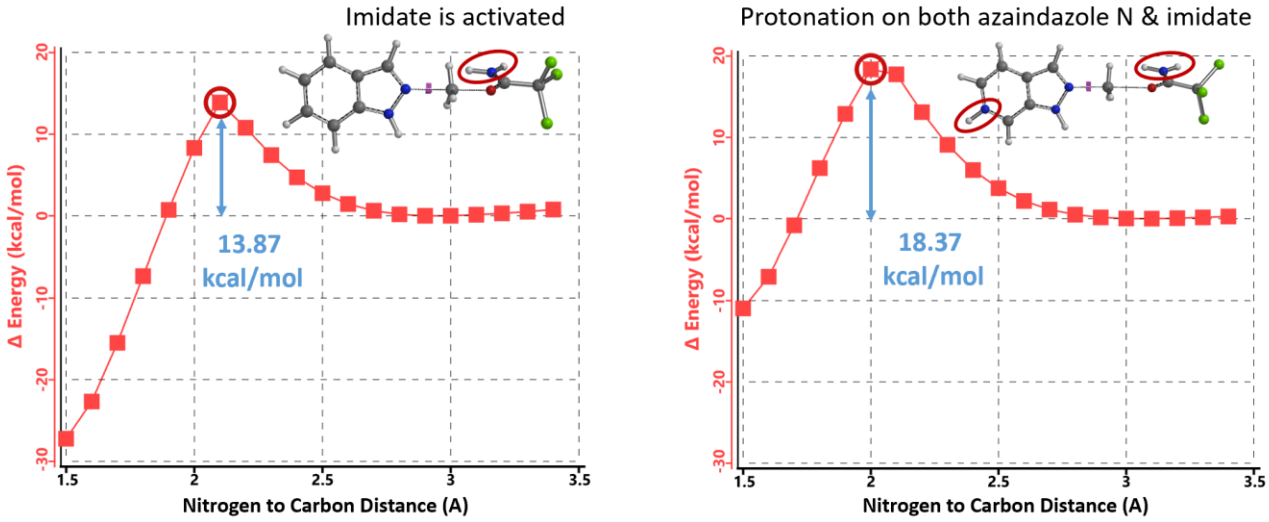
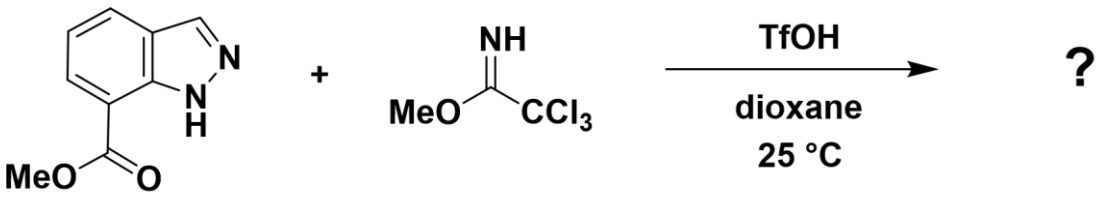


Figure 14. REP of alkylation of protonated azaindazole with protonated activated imidate

Summary

In this chapter, we used QM calculations to investigate why alkylation of indazole with imidate is highly N2 selective and to understand why the size of C3 substituent on the indazole ring has a negligible impact on this selectivity. When evaluating competing reactions involving tautomers, it is essential to include their relative energy difference into consideration. However, it should be noted that estimating the relative energies of tautomers may not be precise enough. In contrast, comparing the energies of their transition states offers a more dependable approach for correlation with experimental observations.

Building on What We Just Learned



For methylation of methyl indazole-7-carboxylate, based on the calculated results shown in Figures 15 and 16, will it be N1 or N2 selective?

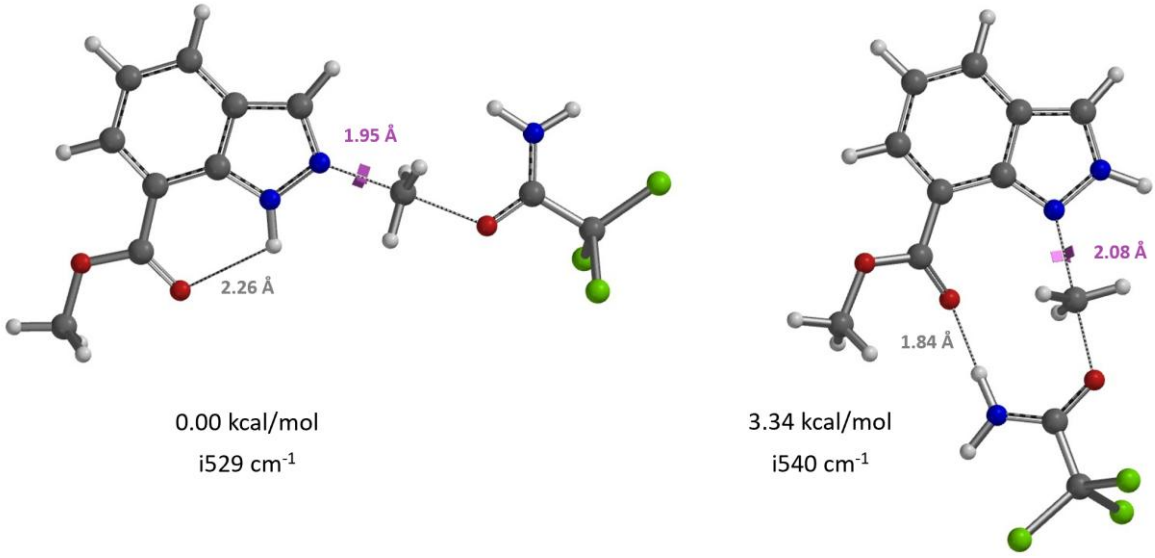


Figure 15. N1 and N2 alkylation transition states of methyl indazole-7-carboxylate

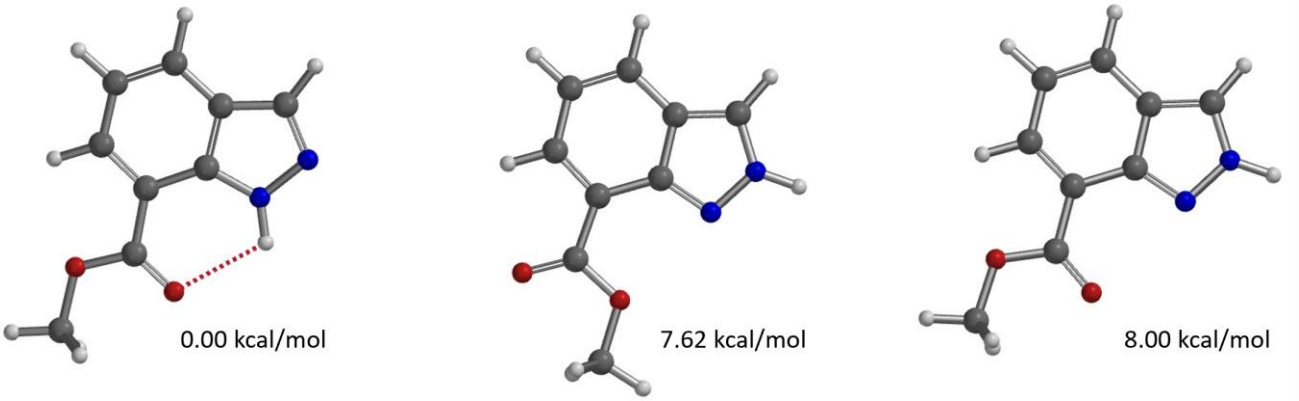


Figure 16. Relative energy of tautomers of methyl indazole-7-carboxylate

Return to Table of Contents

References:

[1] J. Clemens, E. L. Bell, A. T. Londregan, *Synthesis* **2022**, *54*, 3215.

[2] The Pfizer chemists believe that when R is a primary or secondary alkyl group, the reaction proceeds via an SN2 reaction mechanism. However, when R is a tertiary alkyl group, considering the steric hindrance of the tert-alkyl group, the reaction likely proceeds *via* a tert-alkyl cation.

[3] Chapter 9 on QM analysis of regioselectivity of alkylation of pyrazole.

[4] a) Spartan'20 Tutorial and User's Guide (2020). Irvine, CA, USA: Wavefunction, Inc. pp. 158, 442, 459 & 536.

b) Chapters 22 and 24 on the calculation of transition states and imaginary frequencies.

[5] Chapter 49 on Tautomers.

[6] R. Ruzziconi, S. Spizzichino, L. Lunazzi, A. Mazzanti, M. Schlosser, *Eur. J. Chem.* **2009**, *15*, 2645 "***B*** values as a sensitive measure of steric effects."

[6] P. J. Costa, *Phys. Sci. Rev.* **2017**, *2*, 136.

[8] Figure 9. Electrostatic potential surfaces are from transition state calculation, with a bond distance of 2.07 Å between indazole N2 and imidate CH_3O carbon. The indazole and imide ester were artificially separated further to show the ESP surfaces.

[9] Chapter 42 on visualizing non-covalent interactions with electron density maps.

Afterword

The **Remarkable Effectiveness** of Quantum Mechanics (QM) in analyzing organic reactions is exemplified with many examples in this book. QM calculations, even with the necessary approximations, lead to astoundingly accurate descriptions (predictions) of experimental observations, often surpassing our intuitive analyses of them. This leads to the question: Shall we learn organic chemistry with QM, abandon current empirical, intuitive approach that have confused generations of organic chemists? My personal journey of learning heterocyclic chemistry and QM for chemists was fraught with confusion. Yet, it was the QM in analyses of these reactions that rewarded me with precious insights, turning perplexing chemical puzzles into 'Eureka' moments. The most enlightening analyses often involve reactivities that challenge the principles we learned in school—those that form the basis of our intuitive reasoning. These instances highlight how little we know about the reactions we presumed to understand and the shortcomings of those empirical principles.

QM has been instrumental in bridging the gap from mere knowledge to deep understanding. An obvious question is how to turn QM from a retrospective tool into a prospective one. Learning fundamental organic reactions with QM. Day-to-day practices. We learn what to calculate, how to calculate, how to interpret the results, and how to optimize the calculation models, ensuring consistency with experimental conditions and findings. Successes are found upon curiosity and unwavering efforts to advance our understanding of reaction mechanisms.

Nature speaks the language of Quantum Mechanics - a language that is elegantly simple and immensely powerful. With high performance computers becoming increasingly accessible to laboratory scientists, it is a language, through curiosity and determination, we can and should endeavor to master.

John S. Wai

July 2024

About WuXi AppTec

WuXi AppTec is a trusted partner and contributor to the pharmaceutical and life sciences industries, providing R&D and manufacturing services that help advance healthcare innovation. With operations across Asia, Europe, and North America, we offer integrated, end-to-end services through our unique CRDMO (Contract Research, Development, and Manufacturing Organization) platform. We are privileged to work alongside nearly 6,000 partners across 30+ countries, supporting their efforts to bring breakthrough treatments to patients. Guided by our vision that every drug can be made and every disease can be treated, we are committed to advancing breakthroughs for patients—one collaboration at a time.

About Research Chemistry Services

Chemistry services are typical research services offered by WuXi AppTec.

The Research Chemistry Services (RCS) division has multiple R&D sites across the globe. As a competitive empowerment platform with advanced technologies, professional teams and the customer-first philosophy, RCS is committed to continuously enhancing research capabilities and efficiencies. We provide fast, flexible, reliable, and high-quality services and solutions to global customers in the pharmaceutical, biotechnological, and other chemistry-related industries. Our services span from the discovery of hits to preclinical candidates, empowering customers to accelerate small molecule drug research and innovation, reduce R&D costs, and benefit patients.

For more information, please visit: https://rcs.wuxiapptec.com

For business inquiries, please contact: chemistry_service@wuxiapptec.com

Index	Chapter

Hammond Postulate	8
Hydroaminomethylation	28
Hydrogen bond effect on rate of reactions	15, 24, 26, 33
Intrinsic Reaction Coordinate	28, 37
IR Calculation	5, 28
Ketalization	16
Leaving group	5, 12, 19, 20
Nitration	22
NMR Calculation	2, 6, 11, 17, 18, 30
Nucleophilic Aromatic Substitution	1, 7, 10, 18, 19, 27, 29, 35, 36
Nucleophilic Aromatic Substitution of polyhalogenated substrates	10, 29, 35, 45, 48
Oxidative Addition of polyhalogenated substrates	5
Pictet-Spengler reaction	25
pKa	3, 34
Product ratio estimation	9, 10, 13, 17
Pyrimidine	1, 7, 29, 48
QM Quiz	43, 47
Reaction Energy Profile	8, 9, 10, 13, 15, 16, 17, 19, 22, 23, 24, 26, 28, 29, 31, 32, 33, 35, 36, 37, 46, 48, 50
Reaction Rate Equation	8
Retrosynthetic analysis	5, 6
sp^2 S_N2 reaction	37
Tautomerism	49
Telesubstitution	7, 19
Tetrabutylammonium tribromide	6, 23
Thiol alkylation	33, 20
Thiol exchange in ADC Conjugation	12
Torsional scan	14, 40

Return to Table of Contents