APPLICATION OF QUANTUM CHEMICAL METHODS TO COMBUSTION, TAUTOMERISM, AND STACKING

by

BRYAN TOSHIRO SOTO

(Under the Direction of Steven Wheeler)

ABSTRACT

Computational chemistry is a cornerstone for understanding a wide variety of chemical phenomena and different computational methods are specifically tailored to effectively analyze chemical problems. By choosing the proper chemical methods, I have provided new insights to systems with applications to biofuels and pharmaceutical development. My first study employs high-accuracy *ab initio* methods to obtain energetics for the combustion pathways connecting *n*-butanol to the toxic byproducts formaldehyde and acetaldehyde. Results from "gold standard" coupled cluster methods and robust focal-point analysis indicate that the favored pathway is one long assumed to be unimportant.

Heterocycles abound in pharmaceutical compounds, and many heterocycles can exist as more than one tautomer. Identifying the most favorable among possible tautomers is vital for structure-based drug design, as different tautomers can engage in different non-covalent interactions in drug binding sites. At the same time, these same non-covalent interactions can impact the favored tautomer, in some cases leading to preferential formation of a tautomer that is disfavored for the isolated molecule. While

the role of hydrogen-bonding on tautomerism is well-appreciated, stacking interactions (roughly parallel interactions of planar π-systems)[1-4] have the potential to tune the preferred tautomeric state of drug-like heterocycles. To enhance our understanding of these effects, I first benchmarked density functional theory (DFT) methods against robust CCSD(T)/CBS computations, showing that M06-2X/def2-TZVP provides accurate tautomerization energies. Applying this level of theory to a large set of annular tautomeric systems (including both known heterocycles and those that should be synthetically viable but have not yet been synthesized), I classified and identified structural and energetic trends. Subsequently, I applied these findings to assess the impact of stacking interactions with the aromatic amino acids Phe, Tyr, and Trp on tautomeric equilibria, identifying a number of cases for which stacking interactions are predicted to change the preferred tautomer compared to that present in solution.

INDEX WORDS: computational chemistry, density functional theory, combustion chemistry, annular tautomers, noncovalent interactions

APPLICATION OF QUANTUM CHEMICAL METHODS TO COMBUSTION, TAUTOMERISM, AND STACKING

by

BRYAN TOSHIRO SOTO

B.S., Biochemistry and Mathematics, University of Puget Sound, 2014

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2019

APPLICATION OF QUANTUM CHEMICAL METHODS TO COMBUSTION, TAUTOMERISM, AND STACKING

by

BRYAN TOSHIRO SOTO

Major Professor:	Steven E. Wheeler
Committee:	Henry F. Schaefer
	Gary E. Douberly

Electronic Version Approved:

Ron Walcott
Interim Dean of the Graduate School
The University of Georgia
December 2019

DEDICATION

To Janet Soto Mukai, my mother. You spent your life providing me with opportunities, allowing me to be where I am today.

ACKNOWLEDGEMENTS

I am deeply grateful to all those who helped me along this path; by no means could I have accomplished the things I have without the people around me. I'd like to start by thanking my undergraduate advisor Dr. Steven Neshyba for 4 years of sound guidance. Two of my undergraduate organic chemistry professors also deserve recognition, as I wouldn't have considered attending graduate school without the advice of Dr. Eric Scharrer, nor would have I discovered computational chemistry if not for Dr. John Hanson.

Dr. Henry F. Schaefer was the driving force in bringing me to the University of Georgia, and I am thankful to have had him serve as my advisor for the first 2 years of my time here. The remainder of my time here was under the unparalleled guidance of Dr. Steven E. Wheeler, to whom I am eternally grateful. His unending patience and remarkable mentorship created an atmosphere to cultivate my success.
Additional thanks to Dr. Gary E. Douberly for sitting on my committee and providing additional scientific insight to my work.

I must thank all the excellent researchers in the Schaefer and Wheeler research groups who have been constant sources of support and assistance, both in academics and in life. I am especially thankful for the connections I made with Drew Harding, Andrew Launder, Sharath Chandra Mallojjala, Stephen Miller, Kevin Moore, and Walter Turner. I am also thankful to all the people from my non-academic pursuits – especially Athens Swing Night, The Ballroom Performance Group, and Georgia Ultimate – who helped

provide balance to my life. Special recognition goes to Natalie Cox and Michael Fulford, who were not only teachers in dance and performance, but irreplaceable sources of life advice.

No acknowledgements would be complete if I didn't thank my family. My mom Janet is my biggest supporter in every accomplishment or failure in life and she is my primary source of motivation to keep working hard. My sister Lisa is always there to help me when times are tough, and sometimes I think she's more excited about my accomplishments than I am. My father Robert has constantly provided me with kind encouragement during these stressful years. My brother-in-law Alex has always taken a fond interest in my pursuits. Thank you to my paternal grandfather Miguel, who I have not gotten to see often enough while living across the country. Thank you to my maternal grandmother Kiyo. I wish we could have celebrated this accomplishment together, but I know you are always with me. Finally, many thanks to my extended Soto and Yamada families for your curiosity about my work, despite it being difficult to understand.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS v

CHAPTER

1 INTRODUCTION AND LITERATURE REVIEW 1

Computational Methods 1

Butanol Combustion 4

Annular Tautomers 6

Tuning Tautomerization Energies Through Stacking Interactions 8

2 ABSTRACTION-INITIATED COMBUSTION PATHWAYS CONNECTING *n*-BUTANOL TO FORMALDEHYDE AND ACETALDEHYDE FORMATION 10

Abstract 11

Introduction 11

Theoretical Methods 16

Results and Discussion 17

Conclusions 23

3 TRENDS IN ANNULAR TAUTOMERIZATIONS OF HETEROAROMATIC RINGS IN THE GAS PHASE AND IN SOLUTION 25

Abstract 26

Introduction....26
Theoretical Methods....30
Results and Discussion....31
Conclusions....41
4 ANNULAR TAUTOMERISM MEDIATED BY STACKING INTERACTIONS....43
Abstract....44
Introduction....44
Theoretical Methods....46
Results and Discussion....49
Conclusions....55
5 CONCLUSIONS....58
REFERENCES....59
APPENDICES
A SUPPLEMENTARY INFORMATION RELATED TO ANNULAR TAUTOMERIZATIONS OF HETEROAROMATIC RINGS....78
A SUPPLEMENTARY INFORMATION RELATED TO ANNULAR TAUTOMERISM MEDIATED BY STACKING INTERACTIONS....81

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Computational chemistry has proved vital for tackling many chemical problems, including conformational analysis, mechanistic studies, quantifying the strength of non-covalent interactions, and structure-based design. By operating in tandem with experiment, computational chemistry can provide unique insights that are inaccessible or too time-consuming to obtain using conventional laboratory techniques. Over the last decade, computational methods have been refined to the point that they can now be used to tackle complex problems often with high levels of accuracy. However, computational cost scales with system size, limiting the depth to which larger systems may be studied. Coupling system size with appropriate computational methods is paramount to achieving quality results via a powerful chemical tool.

Computational Methods

There are a number of computational tools available that provide a wide range of accuracies. In terms of high-accuracy predictions for small molecules, coupled cluster theory reigns supreme. With its propensity for providing accurate energy computations, coupled cluster is the most widely used post-Hartree-Fock Method. In coupled cluster theory, the exact electronic wavefunction is represented by an exponential ansatz,

$$\psi^{CC} = e^{\hat{T}}\psi^{HF}$$

where ψ^{HF} is the reference Hartree-Fock (HF) determinant and the cluster operator $\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3$ ... is a sum of excitation operators, *e.g.*

$$\hat{T}_1\psi^{HF} = \sum_{ia}^{\substack{single \\ excitations}} t_i^a \psi_i^a$$

This cluster operator is truncated at a given level of excitations, giving rise to CCSD (coupled cluster with single and double excitations), CCSDT (coupled cluster with single, double, and triple excitations), *etc*. The effect of triple excitation is often treated perturbatively, leading to the so-called 'gold-standard' of quantum chemistry, CCSD(T). While coupled cluster theory provides highly accurate electronic energies, these methods scale steeply with the size of the system [*e.g.* CCSD(T) scales as N^7, where N is the number of basis functions].

Even higher accuracy can be obtained through the focal point approach.[5-11] This approach exploits the systematic convergence of Hartee-Fock and correlation energies computed using the Dunning correlation-consistent basis sets.[12] This allows energies using a given electron correlation method to be extrapolated to the complete basis set (CBS) limit. This extrapolation to the complete one-particle basis set limit is combined with systematic increases in the level of excitations in the coupled cluster expansion, resulting in a dual extrapolation of the one-particle and N-particle basis sets and a systematic march toward the *ab initio* limit. In practice, focal point combines CBS extrapolated HF, MP2 (second order Møller-Plesset perturbation theory), CCSD, and CCSD(T) energies, often combined with additive approximations for the inclusion of higher-level excitations in addition to corrections for zero-point vibrational energy and non-Born-Oppenheimer effects.

On the other end of the spectrum are density functional theory (DFT) methods. With DFT, we can tackle larger systems due to its reduced computational cost. This reduced computational cost arises because DFT is a mean field theory. DFT is based on the first Hohenberg-Kohn theorem, which states that the exact electronic energy is a functional of the exact electron density. While the form of this exact functional is unknown, progress over the last 50 years has led to increasingly accurate functionals including M06-2X[13] and wB97X-D.[14] The latter includes an empirical correction for dispersion interactions, which are not well-described by conventional DFT functionals. While DFT is often criticized for its inability to systematically approach the exact electronic energy (a feature of the *ab initio* methods discussed above), in practice DFT can often provide relative energies within a few tenths of a kcal mol^{-1} of coupled cluster computations. The benefit is that DFT can be readily applied to systems with 100s of atoms, or to 100s of systems with a few dozen atoms.

The applicability of coupled cluster methods has been extended to considerably larger systems in recent years through the development of DLPNO coupled cluster methods by Neese *et al*.[15-18] These DLPNO methods exploit the local nature of electron correlation effects, resulting in the capture of ~99.9% of the electron energy described by traditional coupled cluster methods, but with much more favorable scaling with system size. This allows for high-accuracy computations on systems with ~100 atoms, bridging the gap between coupled cluster and DFT methods. Unfortunately, analytic gradients are not available for DLPNO methods, so a common approach is to use DFT to optimize molecular geometries and then use DLPNO-CCSD(T) with a large basis set to compute accurate single point energies at these DFT geometries.

I have used these methods to tackle three challenging chemical problems: the combustion of *n*-butanol to form formaldehyde and acetaldehyde, annular tautomerism in drug-like heterocycles, and the impact of stacking interactions on annular tautomerism. The motivation for studying these systems is described briefly below.

Butanol Combustion

Combustion systems are particularly complex, as many reactions are occurring simultaneously. Many of these reactions involve radicals or other fleeting intermediates and it is very difficult to isolate individual reactions experimentally. On the other hand, pinpointing species present in the combustion of common or potential fuels allows for a more complete understanding of the formation of potentially harmful byproducts and opens the door for the development of improved combustion technology designed to minimize potentially harmful emissions. Traditionally, the species present in combustion systems have been elucidated using shock tubes and jet-stirred reactors or with pyrolysis techniques. While these experiments can often identify which molecules are present, they are generally unable to provide mechanistic details and often struggle with relative distributions of isomeric species. Computational quantum chemistry can help fill these gaps.

Concerns over limited fossil fuel resources have driven studies into potential long-term renewable fuel sources. Currently, reduction of fossil fuel consumption is accomplished through fuel additives, with ethanol as the most common gasoline additive. Bioethanol is created from sugarcane or corn biomass,[19-23] giving it some long-term

benefits over fossil fuels. However, ethanol is miscible with water and corrosive, which complicates its transportation.[24-25]

Recent advances in acetone-butanol-ethanol (ABE) fermentation have greatly increased the efficiency for production of butanol from corn biomass.[26-30] ABE fermentation produces the straight-chain alcohol *n*-butanol, which offers several advantages over ethanol. The longer carbon chain in *n*-butanol allows for better mixing with conventional fuels such as gasoline and eliminates the need for specialized transportation infrastructure, as butanol itself can be transported in existing gasoline pipelines. Additionally, butanol is 25% more energy dense than ethanol, which leads to reduced fuel consumption.[31]

The combustion of *n*-butanol has been explored both experimentally[32-35] and computationally.[33-34, 36] One of the primary concerns with *n*-butanol combustion is the formation of the carcinogenic inhalants formaldehyde and acetaldehyde. Experimental work has demonstrated significant mole fractions of these species, while theoretical kinetic models[33-34, 36] provide branching ratios and proposed intermediates. Consensus between experimental and computational data points to H-atom abstraction-initiated pathways in *n*-butanol combustion.

In Chapter 2, I use the gold-standard CCSD(T) method to examine H-abstraction with multiple initiating radicals and to propose transition state structures for *n*-butanol combustion pathways leading in formaldehyde and acetaldehyde. Using a robust focal point analysis[5-11] and adding in zero-point vibrational energy and diagonal Born-Oppenheimer corrections, I created a high-accuracy potential energy surface for these reactions, providing insight useful for refining existing kinetic models.

Annular Tautomerism

Structure-based drug design (SBDD) is instrumental for streamlining the pharmaceutical development process.[37] Identifying structural trends allows researchers to spend fewer resources on in-lab procedures, cutting costs in a multi-billion-dollar process. Quantitative structure activity relationships (QSAR) can provides additional insight through the use of multivariate parametric models to predict experimental outcomes. In the case of drug-design, QSAR models can be used to predict binding affinities, selectivities, and inhibitor concentrations, among other things. Coupling these two techniques can inform design processes and provide an understanding of interactions present in substrate binding.

Prototropic tautomerism is characterized by H-atom migration and a change in bond order between heavy atoms (see Figure 1.1). Structural changes resulting from tautomerism affect properties considered in both SBDD and QSAR. A common example for these differences is found with hydrogen bonding, as tautomerism will cause a H-bond donor to become an acceptor, or vice versa, potentially disrupting hydrogen bonding interactions vital for effective drug binding. In 2009, approximately 25% of reported pharmaceutical compounds contained at least one tautomerizable site,[38] underscoring the importance of both understanding and predicting tautomeric states in drug development.

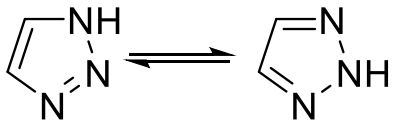


Figure 1.1 Annular tautomerization of 1,2,3-triazole.

With hundreds of unique ring systems identified in currently marketed pharmaceuticals,[39-45] heterocyclic rings also have important implications to both SBDD and QSAR. Compared to their carbocyclic counterparts, heterocyclic rings contain functional groups involved in many stabilizing non-covalent interactions useful in substrate binding. At the same time, most heterocyclic rings, particularly those containing nitrogen, can be present in more than one tautomeric state, with a greater number of tautomers available in more nitrogen-rich systems. Annular tautomerism refers to tautomeric transformations in which the moving proton is constrained to a ring system (*e.g.* the tautomerization of 1,2,3-triazole shown in Figure 1.1 is an example of annular tautomerism). Many ring systems commonly utilized in pharmaceuticals can exhibit multiple annular tautomers, and the ever-increasing list of synthetically tractable heterocycles opens the door for new annular tautomeric systems.

While tautomerization in general in drug-like molecules has been widely studied,[46-48] and has even been the subject of blind computational challenges,[49] annular tautomerism in particular has been much less well studied. The result is that we still lack simple predictive models of which among a group of annular tautomers will be most favorable. That is, given a pair of annular tautomeric structures, in most cases we are unable to predict which tautomer will be most prevalent in solution without expensive quantum chemical computations. To address this, Pitt *et al*.[50] developed ten annular tautomer classes and studied common heterocycles using *ab initio* methods in an effort to develop rules for predicting the preferred tautomeric state for each of these classes. Unfortunately, the set of common heterocycles considered by Pitt *et al*. provided insufficient data to cover all ten tautomer classes. The result is that we still lack simple

predictive tools for the preferred tautomer for many of these classes. Furthermore, even when such structural trends have been identified, these trends often change for different solvents, inciting questions about the robustness of Pitt's guidelines[50] for different solvent environments.

In Chapter 3, I first benchmark density functional theory methods against CCSD(T)/CBS computations to demonstrate the effectiveness of DFT for efficiently capturing relative tautomer energies. Following this benchmarking, I illustrate the importance of accounting solvent effects in annular tautomer systems. Finally, computations on 959 heterocyclic rings comprising 631 tautomer pairs from the VEHICLe database[51] are used to assess the annular tautomer classes of Pitt *et al.*,[50] resulting in a more clear picture of which tautomer classes exhibit strong preferences for one tautomeric state over another.

Tuning Tautomerization Energies through Stacking Interactions

Given the sensitivity of tautomeric equilibria to environmental factors, there is the possibility that the preferred tautomer among a given set of tautomeric states can be changed through non-covalent interactions.[38, 46, 52-55] The most obvious means of doing this is through hydrogen-bonding interactions. Placing a hydrogen-bond donating group near a nitrogen on a heterocycle with different possible tautomeric states will favor the tautomer with this nitrogen unprotonated. This situation is common in drug binding sites, leading to cases in which the dominant tautomer in solution differs from that in the bound state due to hydrogen bonding interactions with binding site residues or the protein amide backbone.

A much less well-studied opportunity to tune tautomeric equilibria is through stacking interactions.[1-4] In 2016, An *et al*.[56] predicted that several heterocycles commonly found in pharmaceuticals will be present in a different tautomeric state when stacked with 9-methyladenine compared to the isolated heterocycle. This arises because the disfavored tautomer is able to engage in stronger stacking interactions than the favored tautomer, with the difference in stacking interactions exceeding the tautomerization energy of the isolated heterocycle. Stacking interactions with the aromatic amino acid residues Phe, Tyr, and Trp are common in drug binding sites, yet we still do not understand the many factors that impact the strength of these interactions. This is despite tremendous advances in our understanding of stacking interactions in simple aromatic systems.[57-77]

Earlier this year, Bootsma *et al*.[78] developed a predictive model of the maximum possible stacking interaction between drug-like heterocycles and the aromatic amino acids Phe, Tyr, and Trp. This allows for the rapid prediction of stacking interactions for heterocycles, opening the door to identifying additional heterocycles that could exhibit different preferred tautomers when engaged in stacking interactions than in isolation

In Chapter 4, I use dispersion-corrected DFT methods and DLPNO-CCSD(T) to predict accurate stacking interactions for tautomeric heterocycles for which differences in stacking interaction energies are expected to outweigh the tautomerization energy. Such systems could exhibit a different tautomeric state when bound to a protein than in solution. In addition to verifying which of these systems are predicted to change tautomeric state upon binding, these new computational data has allowed us to assess the predictive model of Bootsma *et al*.[78]

CHAPTER 2

ABSTRACTION-INITIATED COMBUSTION PATHWAYS CONNECTING *n*-BUTANOL TO FORMALDEHYDE AND ACETALDEHYDE FORMATION[a]

[a] Soto, B. T., Turney, J. M. and H. F. Schaefer. *To be submitted to the J. Chem. Theory Comput.*

Abstract

The increased importance of biofuels has called for a deeper understanding of combustion properties of potential biofuels. Compounds such as *n*-butanol can be synthesized from biomass, mix well with conventional fuels, and have an energy density comparable to that of gasoline. We employed high-level computational techniques to explore the mechanisms for formation of formaldehyde and acetaldehyde from *n*-butanol. Reaction barriers for different hydrogen abstractions followed by β-scission of the resulting radicals were obtained at the CCSD(T)/CBS//CCSD(T)/ANO1 level of theory. These data show that, in contrast to previous assumptions, the most viable route to formaldehyde from *n*-butanol proceeds by the abstraction of the hydroxy H. New computational data also show that the relative ordering of the H-abstraction transition states is dependent on the nature of the abstracting radical. These and other results should prove useful in refining kinetic models of n-butanol combustion.

Introduction

Concerns sparked by finite amounts of fossil fuels and increased emissions of greenhouse gases have catalyzed research into alternative fuel sources. One solution to decreasing fossil fuel consumption is adding renewable compounds to conventional gasoline or diesel.[27-28, 34, 79] Small alcohols, namely ethanol and butanol, are being studied extensively, as they raise octane rating and boast many desired sustainability features. A biologically renewable source of ethanol, or bioethanol, may be produced from corn feedstock through fermentation by yeast,[22-23] while acetone-butanol-ethanol (ABE) fermentation can provide high yields of biobutanol from corn biomass.[26] Unlike ethanol,

butanol has four distinct isomers, the structures of which result in unique physical properties that influence fuels in different ways.[80] Modern biobutanol synthetic pathways result primarily in *n*-butanol, *i.e.*, straight chain 1-butanol. Recent advances in synthetic pathways of biobutanol such as ABE fermentation have improved efficiency to a point where it is practical for industrial use,[27-30] and BP and DuPont have started marketing fuels containing biobutanol in the U.K.[27]

The physical properties of *n*-butanol make it an attractive candidate for study as a feasible fuel additive, or even as a standalone fuel.[32, 34, 80] Due to its longer carbon chain, *n*-butanol provides approximately 25% more energy than ethanol and mixes better with conventional fuels,[28-29] resulting in decreased fuel consumption.[31] The unbranched chain also causes *n*-butanol to be the least polar isomer of butanol, allowing it to intersolubilize well with both diesel fuel and gasoline, thus increasing the commercial viability of biobutanol. Additionally, an energy density much closer to that of gasoline coupled with its physical similarities to gasoline allows for unmixed *n*-butanol to be used in conventional engines without significant engine modification. The low vapor pressure point (2.27 kPa compared to 13.8 kPa of ethanol and 31.01 kPa of gasoline) and high flash point (35 °C compared to 8 °C of ethanol and -45 to -38 °C of gasoline) also make *n*-butanol a safer fuel than ethanol or gasoline.[27]

Theoretical kinetic models for the combustion of *n*-butanol developed by Westbrook,[34] Black,[36] and Dagaut[33] have been compared to data from jet-stirred reactors (JSR) and shock tube experiments,[35] providing important insights into the combustion chemistry of *n*-butanol. At lower temperatures (≤ 800 K), Dagaut *et al*.[33] found that *n*-butanol showed no signs of reaction in a JSR. Consequently, they examined temperatures

between 800 K and 1250 K, at a pressure of 1 atm, with varying fuel concentrations. Based on JSR data, as well as shock tube experiments from Vranckx and coworkers,[81] H-atom abstraction has been determined to be the primary method by which *n*-butanol combusts.

The mechanism of Black *et al*.[36] predicts H-atom abstractions which can be followed by radical decomposition with particular importance to fission of the C_α-C_β bond. However, experimental species profiles measured by Karwat *et al*.[82] indicated that the Black mechanism did not accurately capture ethene formation and removal. Karwat *et al*.[83] used *ab initio* computations performed by Zhang *et al*.[84] to provide improved reaction rates for post-abstraction fission pathways in the Black mechanism. Subsequent predictions exhibit improved agreement with the low-pressure pyrolysis product species measurements made by Stranic *et al*.[35]

n-butanol contains five unique H-atom types,[80, 85] α, β, γ, δ, and OH, and the abstraction of these hydrogen atoms by various small radicals has been studied theoretically by a number of groups.[80, 84, 86-88] Sarathy et al.[80] used computations on ethanol and *n*-butane to estimate abstraction rates by ·H, but did not include energetic barriers. Abstraction by ·OH,[80, 86, 89] HO_2·,[80, 87, 90] and ·CH_3[88] have been studied by *ab-initio* computation. Both Moc and Simmie[86] and Zhou, Simmie, and Curran[89] computed barrier heights for abstraction by ·OH at the CCSD(T)/cc-pVQZ//MP2/6-31G(d,p) level of theory. Zhou, Simmie, and Curran[89] then used these computations to estimate rate constants, which indicated a lack of importance to abstraction from the alcoholic group. In a similar study, Zhou, Simmie, and Curran[90] computed energetic barriers for abstraction by HO_2· using CCSD(T)/cc-pVTZ//MP2/6-38G(d,p) methods and reported

rate constants which showed abstraction at the α site to be most common. Conversely, Katsikadakos *et al.*[88] used ROCBS-QB3 computations to examine reaction barriers for abstraction by $\cdot CH_3$, finding the alcoholic proton to have the second lowest TS barrier.

Following abstraction, β-scission of C–C bonds is a primary pathway by which the resultant radicals dissociate. The radicals resulting from C–H bond cleavage are commonly referred to as aC_4H_8OH, bC_4H_8OH, cC_4H_8OH, and dC_4H_8OH for radicals centered on the α, β, γ, and δ carbons, respectively.[34] Zhang *et al.*[91] conducted master equation analyses for the decomposition of the resultant hydroxybutyl and butyoxy radicals over temperatures ranging from 300 to 2500K and pressures from $1.3x10^{-3}$ to 100 atm based on reaction barriers computed at the RQCISD(T)/CBS//B3LYP/6-311++G(d,p) level of theory.

Despite its advantages, *n*-butanol combustion leads to formation of the toxic byproducts formaldehyde and acetaldehyde - known carcinogens with inhalation being the primary means by which they enter the human body. For *n*-butanol to become a safe fuel alternative it is imperative to gain an understanding of the processes by which these compounds are created. The primary combustion pathways leading from n-butanol to formaldehyde and acetaldehyde are shown in Figure 2.1, along with predicted branching ratios from Westbrook *et al.*[34] and Dagaut *et al.*[33] Formaldehyde is most directly produced from C_α–C_β bond scission after the formation of the *n*-butoxy radical. With the hydroxyl hydrogen being the most strongly bound hydrogen in *n*-butanol,[36, 86] formation of the hydroxy radical is slow, or considered unimportant.[89] Given the significant formation of formaldehyde observed experimentally,[32-34] there must be other viable pathways. An supported explanation is derived from the analysis set forth by Zhang et

al.,[91] which shows the pathway for dC_4H_8OH with the lowest energy barrier is isomerization to the butoxy radical. While they note the reverse reaction is more favorable than C_α–C_β bond scission, ultimately they conclude that a major pathway to formaldehyde passes through dC_4H_8OH.[91] One possibility is the β-scission of the C_α–C_β bond in cC_4H_8OH, which creates propene and the methylhydroxy radical. Upon H-atom abstraction, the methylhydroxy radical can form formaldehyde.[92] Further evidence supporting this mechanism is found in the species profiles measured by Stranic *et al.*,[35] who proposed that the formation of CH_3CHCH_2 and $\cdot CH_2OH$ is underpredicted by a factor of three to five in the Black mechanism.[36]

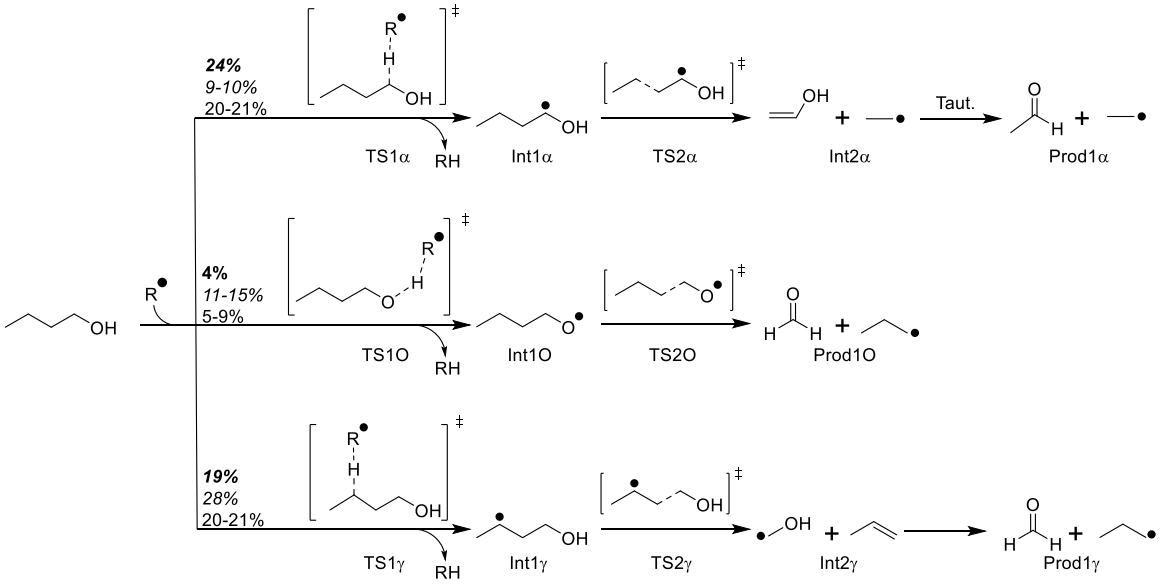


Figure 2.1. Proposed combustion mechanisms for abstraction-initiated pathways of *n*-butanol resulting in formaldehyde and acetaldehyde. Branching ratios from the Black model[36] (bold text), Dagaut model[33] (italicized text) and Westbrook model[34] (normal text) provided (pathways that do not lead to formaldehyde or acetaldehyde omitted).

Acetaldehyde is primarily produced by the β-scission of the C_2–C_3 bond after creation of aC_4H_8OH. It is important to note that vinyl alcohol is the direct product of the C_2–C_3 bond breakage, but it ultimately tautomerizes to acetaldehyde at elevated temperatures.[32] However, dC_4H_8OH may also undergo a H-atom migration to create aC_4H_8OH, thereby increasing the production of acetaldehyde from *n*-butanol.[32-34] While the likelihood of H-atom migration for dC_4H_8OH is highly dependent on the fuel/air ratio (φ) and the temperature, it is consistently above 85% in all of the aforementioned studies, with the exception of the one conducted by Zhang et al.,[84] where it is the pathway with the second lowest barrier.[91]

Herein, we employ high-level theoretical methods to elucidate the abstraction-initiated β-scission processes in *n*-butanol combustion (Figure 2.1) to provide a more robust understanding of the major pathways resulting in formaldehyde and acetaldehyde.

Theoretical Methods

All computations were performed using CFOUR,[93] MRCC,[94-97] and Q-Chem[98] quantum chemistry packages. Geometries were optimized using the coupled cluster singles, doubles, and perturbative triples level [CCSD(T)] of theory with the atomic natural orbital (ANO) basis set ANO1.[99] For all radical species computations, an unrestricted Hartree-Fock(UHF) reference was used to allow for greater computational efficiency and analytic calculations of second derivatives in frequency calculations.[100] For H-Atom abstraction by H· transition states, geometries were optimized with the ANO0 basis set, as the ANO1 basis set was too computationally demanding. Similar reasoning resulted in optimization of abstraction transition states by ·OH and ·CH_3 being

computed with the B3LYP functional and a 6-311G* basis set. Single point energies of these large transition states were computed with CCSD(T)/cc-pVTZ.

The focal point analysis (FPA) scheme developed by Allen and co-workers was utilized to achieve a high level of accuracy for relative energy determinations.[5-11] At the Hartree-Fock level of theory, a three-parameter extrapolation was used such that:

$$E_X^{\mathrm{HF}} = E_{\mathrm{CBS}}^{\mathrm{HF}} + ae^{-bX}$$

Where X ∈ [2,3,4,5] as it corresponds to the Dunning cc-pVXZ basis sets (X = D,T,Q,5).[101] Correlation energies were extrapolated via the two-parameter extrapolation:

$$E_X^{\mathrm{corr}} = E_{\mathrm{CBS}}^{\mathrm{corr}} + aX^{-3}$$

Based on cc-pVTZ and cc-pVQZ data. Zero-point vibrational energy corrections have been added based on harmonic vibrational frequencies computed at the same level of theory as the corresponding geometry optimizations. Diagonal Born-Oppenheimer corrections (DBOC) were appended to the final energies, evaluated at the Hartree-Fock level of theory. Correlation of the core electrons was corrected by employing a core basis set and computing a single-point energy both with and without the frozen core approximation.

$$\Delta E_{\mathrm{core}} = E_{\mathrm{CCSD(T)(AE)}}^{\mathrm{cc\text{-}pVTZ}} - E_{\mathrm{CCSD(T)(FC)}}^{\mathrm{cc\text{-}pVTZ}}$$

Results and Discussion

As is the case with molecules containing multiple rotatable bonds, *n*-butanol contains various conformers. Ohno et al.[102] studied the 14 possible conformers computationally and spectroscopically, identifying the TGt conformer as the lowest in

energy. The geometry of this conformer has been re-optimized at with CCSD(T)/ANO1 and will be used as a starting structure for all further computations (Figure 2.2).

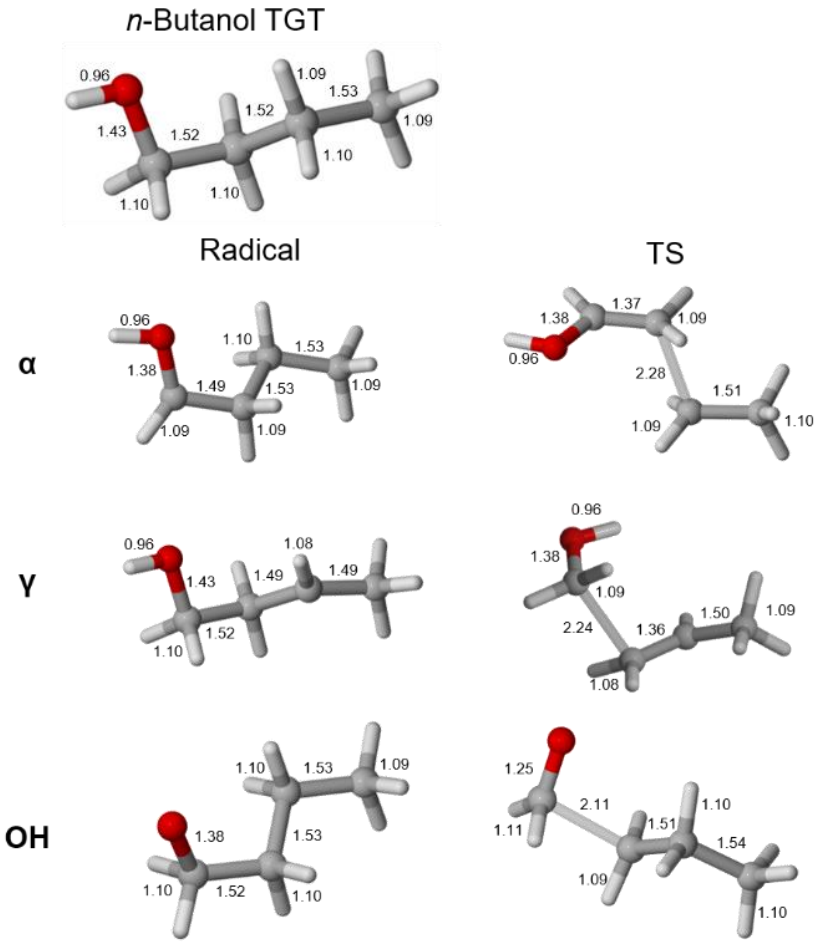


Figure 2.2. Resultant radicals of H-atom abstraction from *n*-butanol, along with their corresponding β-scission transition states computed at the CCSD(T)/ANO1 level of theory. All bond lengths are in Angstroms.

Reaction profiles for each pathway depicted in Figure 2.1 have been computed and are shown in Figure 2.3. In these reaction profiles, ·H is used at the abstracting radical to initiate each reaction sequence via cleavage of a C–H or O–H bond.

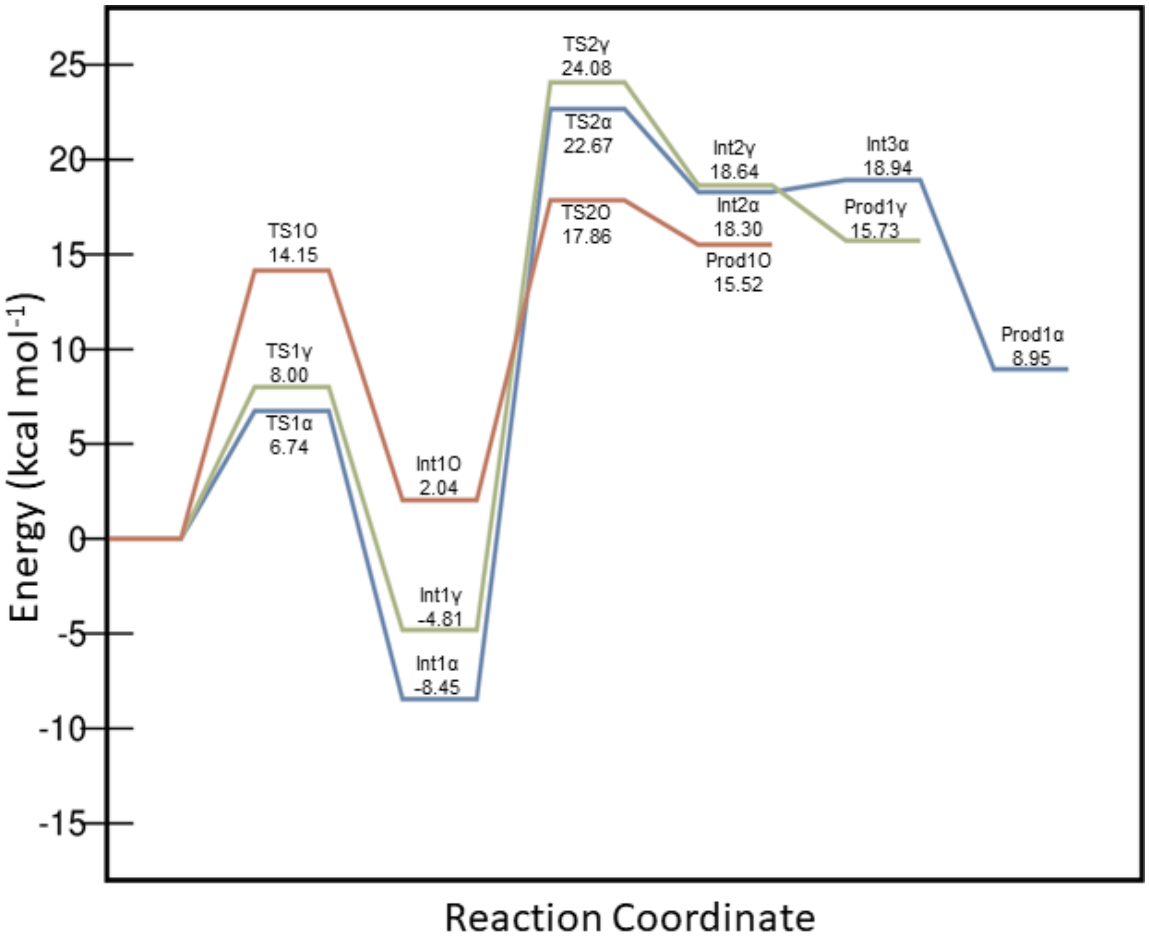


Figure 2.3. Reaction profile for formaldehyde and acetaldehyde formation using H as the abstracting radical. Red represents abstraction from the hydroxy site, green from the γ site, and blue from the α site.

Based on bond dissociation energies computed by Black et. al[36], one would expect the O-H bond to be the strongest, followed by the γ C-H bond, and finally the α C-H bond. When ·H is used as the abstracting radical, the abstraction barriers display the

expected trend. However, other radicals will be present under typical combustion conditions, all of which will be capable of abstracting a hydrogen atom from *n* -butanol. As such, we have also computed barriers using ·OH and ·CH_3.

Barriers for abstraction by ·OH and ·CH_3 have been computed with the CCSD(T)/cc-pVTZ//B3LYP/6-311G* level of theory. Comparison with CCSD(T)/CBS//CCSD(T)/ANO1 barriers for abstraction by ·H are used to demonstrate consistency of trends across multiple levels of theory. Differing by less than 1.6 kcal·mol^{-1}, non-extrapolated barriers exhibit an identical trend as extrapolated ones.

Table 2.1. Reaction barriers (kcal·mol^{-1}) for H-atom abstraction from various sites.

	CBS	CCSD(T)/cc-pVTZ				CCSD(T)/cc-pVQZ	
Abstracting Radical	H	H	CH_3	OH	OH[a]	OH[a]	OH[b]
TS1α	6.58	6.52	11.37	1.69	0.55	0.10	-1.13
TS1γ	7.86	8.04	12.62	0.03	0.72	0.47	-1.95
TS1O	14.15	12.54	12.54	3.12	2.84	3.19	3.19

[a] Ref [86]; single point energies at MP2/6-311G(d,p) geometries
[b] Ref [89]; single point energies at MP2/6-311G(d,p) geometries

Unexpectedly, the ordering of computed barrier heights for abstraction of the different the different hydrogens is strongly dependent on the abstracting radical (Table 1). When ·CH_3 is used as the abstracting radical, removing a hydrogen from the γ-carbon has the highest barrier, lying ~1 kcal mol^{-1} higher than for removal of the hydrogen on either the α-carbon or hydroxy group. Katsikadakos et al.[88] observed a similar trend. They examined the molecular orbitals of *n*-butane and compared them to those of *n* -butanol, noting that the HOMO and LUMO of *n*-butanol are more localized over the

methanol moiety, which contributes to the ease in removing the hydrogen from the hydroxy and α-carbon sites.

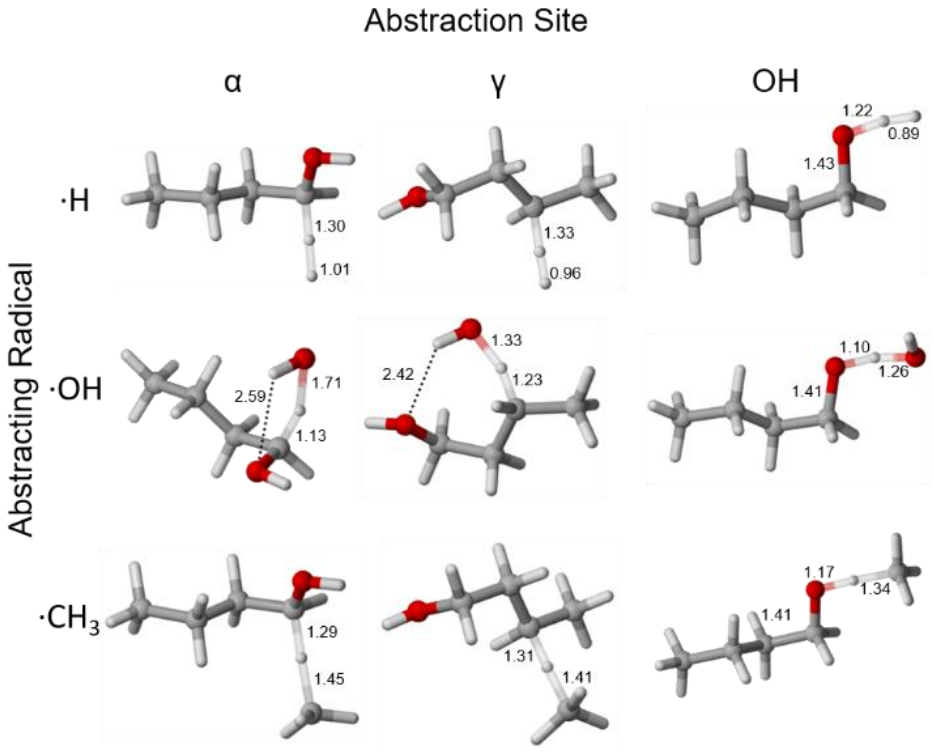


Figure 2.4. Transition state structures for H-atom abstraction by ·H, ·OH, and ·CH_3. Selected bond lengths in Angstroms.

Computed results for abstraction by ·OH from this study are not in agreement with those from Moc and Simmie,[86] who reported that the α site abstraction has a lower barrier than γ site abstraction. In a follow-up study, Zhou, Simmie, and Curran[89] identified abstraction barriers between *n*-butanol and lower-lying transition state structures ~1 kcal·mol^{-1} lower than the starting materials. Energetic barriers computed in the latter study demonstrate trends consistent with the present study with the barrier for γ-

site abstraction lower than that of α-site abstraction. The latter reaction channel computed[89] accounts for an additional stabilizing interaction arising from the ability of the OH moiety to engage in hydrogen bonding interactions with *n*-butanol. Such a stabilizing interaction is not available at the hydroxy site, but Zhou et al.[89] have demonstrated barrierless reaction pathways. Pathways computed in this study are nearly barrierless, but the B3LYP/6-311G* geometries are different than those computed with MP2/6-311G(d,p), which explain differences in single-point energies. Table 2.1 contains our lowest energy transition states, which include hydrogen bonding interactions. Additional transitions states can be found in the supporting information.

Potential energy surfaces for Zhang et al.[91] are set relative to 1-butene + OH so initial abstraction barriers cannot be considered, but the barriers to β-scission can be compared. In their study, Zhang et al.[91] used a unique extrapolation scheme to avoid computation of the time-consuming QCISD(T)/cc-pVQZ energy. The present study computes the energy of the barriers using the comparable CCSD(T)/cc-pVQZ level of theory without any extrapolation scheme. We find the energy barriers computed in this manner are consistently <0.5 kcal·mol^{-1} lower in energy than those computed by Zhang et al.[91] (Table 2.2). The present work has refined the barriers computed in previous work, while still exhibiting consistency amongst the trends displayed.

Table 2.2. Energy barriers for C—C β-Scission

	Energy (kcal mol^{-1})	
Radical	Present	Zhang
aC_4H_8OH	29.9	30.1
cC_4H8OH	27.1	28.6
C_4H_9O	14.3	14.6

Energy barriers for the subsequent β-scission are 29.86, 27.07, and 14.32 kcal·mol^{-1}, for the α, γ and butoxy radicals, respectively, so the initial abstraction step is not rate limiting regardless of the identity of the abstracting radical. For low temperature combustion systems (< 1100K), in which ·OH radicals are the primary reaction initiators, these initial hydrogen abstractions are expected to be extremely rapid. The high energetic barriers for β--scission of α and γ radicals indicate the creation of acetaldehyde or formaldehyde is unlikely to occur via these particular pathways. Instead, initial abstraction at the hydroxy site appears to be a viable pathway to formaldehyde. This is in disagreement with the Westbrook model,[34] which shows less than 9% of *n*-butanol following this pathway, and the Dagaut model,[33] which shows less than 15% of *n*-butanol following the pathway. The Westbrook kinetic model also exhibits a strong preference towards the creation of Int1γ over Int1α. Abstraction barrier heights we have computed do not support this, but rather support the Dagaut model, which shows nearly equal branching ratios.

Concluding Remarks

High accuracy barriers for reactions involved in acetaldehyde and formaldehyde creation have been computed. We have successfully used higher level computations than previously attempted. Comparison to existing kinetic models demonstrates that our computations may be used as a benchmark for further study to refine such models. The choice of abstracting radical for the abstraction-initiated pathway is found to directly affect the height of abstraction barriers. Despite claims that abstraction at the hydroxy

site is never important, our reaction profile indicates that the β-scission of the hydroxy radical is the most favorable pathways to formaldehyde.

CHAPTER 3

TRENDS IN ANNULAR TAUTOMERIZATIONS OF HETEROAROMATIC RINGS IN THE GAS PHASE AND SOLUTION[b]

[b] Soto, B.T., Doney, A. N., and S.E. Wheeler. To be submitted to *J. Chem. Inf. Model.*

Abstract

Identification of the most favorable annular tautomer for a given heterocycle is vital for effective structure-based drug design, as tautomeric state can profoundly impact binding affinities. A set of 959 heterocyclic rings comprising 631 tautomer pairs were examined with quantum mechanical computations. First, benchmark *ab initio* computations extrapolated to the complete basis set (CBS) limit were used to evaluate the reliability of annular tautomerization energies for a set of ten tautomers predicted with the M06-2X, B3LYP, and wB97X-D functionals paired with the def2-TZVP basis set. Tautomerization energies computed at the SMD-M06-2X/def2-TZVP//M06-2X/def2-TZVP level of theory are compared to a smaller set of available literature values. Results are used to demonstrate the importance of including solvent models when studying these systems. Using a scheme of ten different annular tautomerism classes from Pitt *et al.*,[50] accurate computed energies for 631 tautomer pairs reveal six classes of annular tautomerism that demonstrate a clear structural preference, providing a simple means of predicting a wide range of tautomerization states without resorting to expensive computations.

Introduction

Heterocyclic rings are prevalent in biologically active molecules and pharmaceuticals, with hundreds of unique ring systems identified in currently marketed drugs.[39-41] As many as 25% of pharmaceutical compounds can exhibit multiple tautomeric states[38] and the consideration of more nitrogen-rich heterocycles also increases the number of possible tautomers. When compared to their carbocyclic

counterparts, heterocyclic rings contain functional groups capable of stronger non-covalent stabilizing interactions useful in substrate binding. One such interaction, hydrogen bonding, is one of the most thoroughly studied and well-characterized non-covalent interactions in chemistry. Tautomerism alters hydrogen bonding donor and acceptor sites, driving interest to its study.[45]

Tautomerism comes in various forms but generally refers to the movement of a proton accompanied by a change in bond order between two heavy atoms - known as prototropic tautomerism. Restricting proton movement to annular nitrogen atoms is a specific form of tautomerism referred to as annular tautomerism. Preservation of aromaticity in annular tautomerism[103] increases its viability in pharmaceutical compounds. Balabin *et al*.[104] used *ab initio* methods to examine two of the most common tautomerizable heterocycles, tetrazole and triazole, determining high-accuracy relative energies between tautomers of these simple heterocycles in the gas phase.

Non-covalent interactions with binding site functional groups can strongly influence tautomeric equilibrium.[38, 46, 52-55] In particular, for systems with small tautomeric enthalpy differences, differences in non-covalent interactions in the bound state can lead to the preferential formation of tautomers that would be disfavored in solution. Cruz-Cabeza and Groom[105] examined all compounds in the Cambridge Structural Database (CSD), identifying tautomeric forms in approximately 10% of the structures. Within the tautomeric subset, Cruz-Cabeza and Groom[105] found that a mixture of tautomers could be present provided the relative energies were smaller than that of a strong hydrogen bond. Miletti and Vulpetti[43] note that binding site tautomer preference can differ from tautomer preference in water in annular tautomers or

tautomers with a low (<2 kcal mol^{-1}) relative free energies based on analysis of complexes in the Protein Data Bank (PDB).

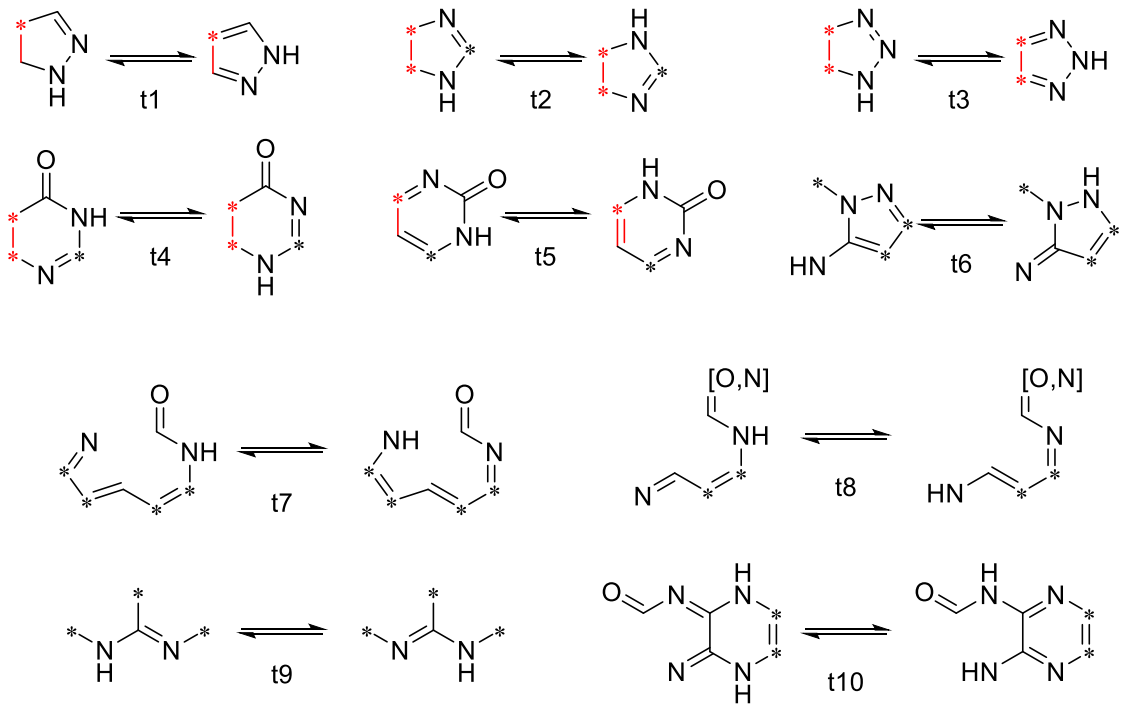


Figure 3.1. Ten annular tautomer classes characterized by unique proton movement as created by Pitt and coworkers.[50] Starred atoms can be carbon or nitrogen and red bonds are locations for ring fusion. Classes 1-5 can occur in monocycles or bicyclic heterocycles, while 6-10 are only possible in bicycles. For each class, the structure on the left will be referred to as tautomer X, while the right structure will be referred to as tautomer Y.

There have been a number of efforts to delineate the full space of chemically accessible heteroaromatic rings.[51, 106-107] Pitt *et al.*[51] generated a database of over 24,000 heteroaromatic rings (the VEHICLe database), along with a score of synthetic tractability

for each heterocycle. Of these possible heterocycles, only ~1700 have been synthesized. This database provides a prime opportunity to explore trends in tautomerization energies across diverse heterocycles, considering both those commonly encountered in pharmaceuticals[41] and those not yet synthesized. A follow-up study by Pitt *et al.*[50] examined 84 annular tautomers from the VEHICLe database comprising 34 different clusters (a tautomeric cluster contains all tautomeric forms of a molecule). Pitt *et al*.[50] categorized monocyclic and bicyclic structures commonly found in the CSD, the PDB, and the MSD into ten tautomer classes (see Figure 2.1), where each tautomer class is defined by a specific proton movement. Using the test set of 34 clusters, Pitt and coworkers computed tautomerization energies for 60 transformations falling into one of these classes and identified basic trends. Unfortunately, their set of heterocycles did not provide sufficient coverage of the ten tautomer classes, and Pitt *et al*. concluded that there was a need to expand the available data.[50] To our knowledge, this has not been done.

Below, we proceed by first providing rigorous computational benchmarks of tautomerization energies and enthalpies of selected heterocycles from the VEHICLe database. This benchmark is followed by the examination of the relative energies of 631 tautomer pairs involving 959 common and uncommon heteroaromatic rings, including include a direct comparison to the tautomers studied by Pitt *et al*. (Figure 2.2). Data for this much more expansive set of heterocycles provide ample examples of each of the ten tautomer classes, revealing six classes that exhibit a clear energetic preference for one tautomer over the other.

Theoretical Methods

A representative set of ten heterocyclic pairs was chosen for a benchmark analysis of three different DFT methods (Figure 2.3). Structures have been assigned identifiers based on those used by Pitt *et al.*,[50] and new identifiers have been created for those not included in the original Pitt set. Table A.1 in Appendix A lists the identifiers used in this work as well as identifiers used in the VEHICLe database.[50]

For each heterocycle, geometries were optimized with the B3LYP,[108] M06-2X,[13] and ωB97X-D[14] functionals paired with the def2-TZVP basis set.[109] To assess the performance of these three functionals, benchmark tautomerization energies were derived using the focal-point analysis scheme of Allen and coworkers.[5-11] For this analysis, the geometry of each heterocycle was computed at the MP2 level of theory with the Dunning aug-cc-pVTZ basis set.[12] The convergent nature of the Dunning cc-pVXZ (X = 2-5) basis sets was used to extrapolate single-point energies to the complete basis set limit.

This benchmark set was also used to demonstrate the importance of solvent effects on annular tautomerization energies. Using the MP2/aug-cc-pVTZ geometry, energies were computed with M06-2X/def2-TZVP in the gas phase, diethyl ether (an approximation for the environment inside a protein), and water. The SMD solvent model[110] provided the necessary solvent corrections.

A collection of 959 heterocycles comprising 423 tautomeric clusters was selected from the VEHICLe database.[50] Heterocycles were chosen based on a measure of synthetic feasibility (P_{good}) such that $P_{good} \geq 0.95$. The energies of each tautomer were computed in water with SMD-M06-2X/def2-TZVP after optimizing structures in the gas-phase at the M06-2X/def2-TZVP level of theory. Each tautomeric cluster was organized

in terms of increasing energy, and solution-phase tautomerization energies were computed. For clusters with more than two tautomers, tautomerization energies are reported relative to the lowest energy tautomer.

Results and Discussion

A. Tautomerization Energy Benchmark

To assess the performance of DFT methods in the prediction of tautomerization energies and enthalpies for heterocycles, we have computed benchmark values using the focal point approach of Allen and co-workers[9-11] for ten representative tautomeric pairs (see Scheme 3.1 and Table 3.1). The tautomerization energies of these tautomeric pairs range from 0.4 kcal mol^{-1} for **16b → 16a** to 9.8 kcal mol^{-1} for **35a→35b**. Focal point tables are available in Table A.2 in Appendix A. Overall, the data show reasonable convergence both with respect to the complete one-particle basis set limit and the inclusion of electron correlation.

Balabin[104] previously reported extrapolated tautomerization energies for **3b→3a**, **1b→1a**, and **2b→2a**. They utilized MP2/aug-cc-pVTZ for geometry optimization, adding corrections from MP*n* (*n* = 2-4), CCSD/CCSD(T) single-point energy computations with aug-cc-pVXZ (X = 2-5) basis sets and core corrections. Computed focal-point tautomerization energies from the current study exhibit good agreement with Balabin's work, with differences below 0.1 kcal mol^{-1}.

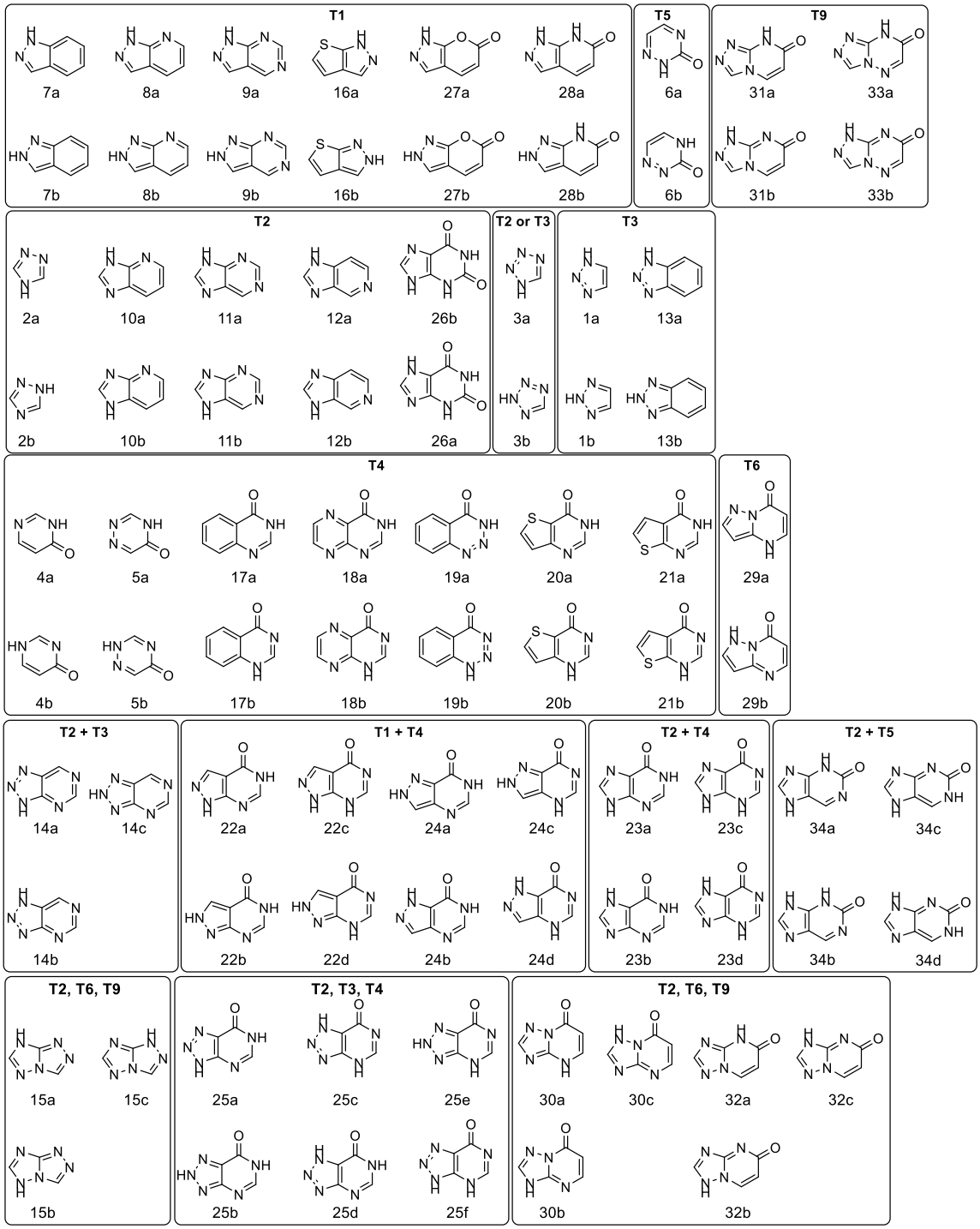
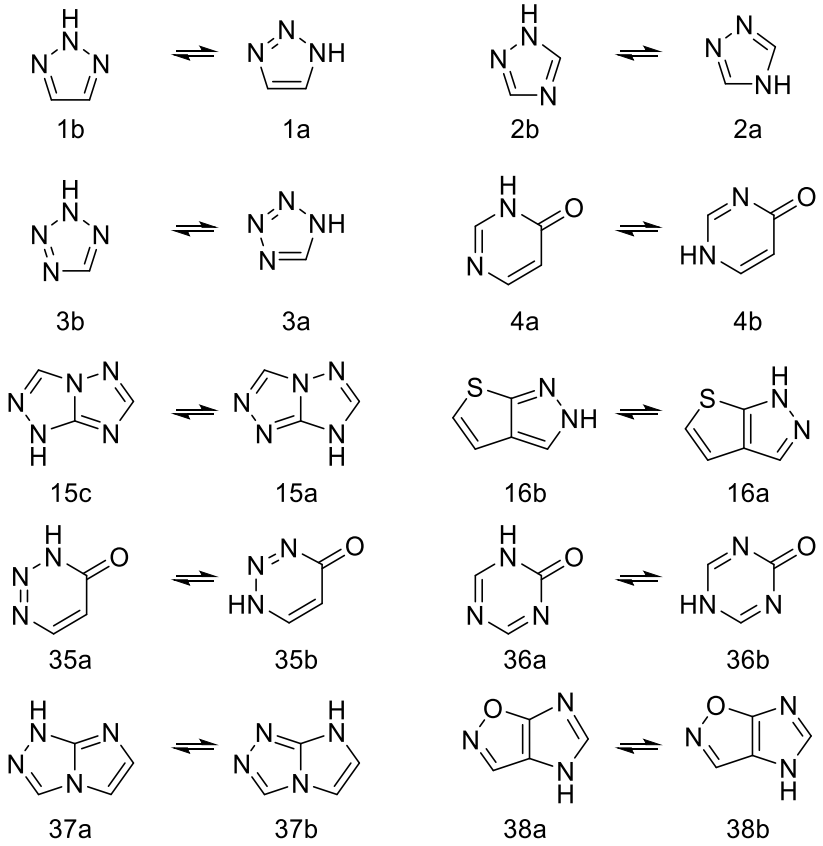


Figure 3.2. The set of heterocycles studied by Pitt and coworkers.[50] Each box contains a tautomeric cluster or clusters with transitions fitting into one, two, or three tautomerization classes.



Scheme 3.1. Tautomer pairs used for benchmark analysis of tautomerization energies. The B3LYP, M06-2X, and ωB97X-D functionals were compared to focal point extrapolated energies to assess how properly tautomerization energies of heterocyclic systems can be determined.

Tautomerization energies and enthalpies computed using ωB97X-D, B3LYP, and M06-2X (all paired with the def2-TZVP basis set) are also provided in Table 3.1. All three functionals provide reliable tautomerization energies and enthalpies relative to focal-point benchmark values. Notably, the M06-2X functional yields the smallest mean errors, justifying its exclusive use below.

Table 3.1. Focal point tautomerization enthalpies (ΔH) and energies (ΔE) along with the corresponding wB97X-D, B3LYP, and M06-2X values, as well as mean signed errors (MSE) and mean unsigned errors (MUE), relative to the focal point values, all in kcal mol^{-1}. Values from ref. [104] are also included.

	Focal Point		ωB97X-D		B3LYP		M06-2X		Focal Point[a]
Transition	ΔH	ΔE	ΔH	ΔE	ΔH	ΔE	ΔH	ΔE	ΔE
16b → 16a	0.4	0.4	0.2	0.2	0.5	0.5	0.1	0.1	–
3b → 3a	1.8	2.2	2.2	2.6	2.1	2.4	2.0	2.3	2.1
38a → 38b	2.5	2.6	2.5	2.8	2.4	2.6	2.5	2.7	–
1b → 1a	3.7	4.1	4.1	4.5	3.8	4.3	3.9	4.3	4.0
37a → 37b	3.8	4.1	4.4	4.6	4.3	4.6	4.3	4.6	–
2b → 2a	5.8	6.3	6.3	6.7	5.9	6.4	6.0	6.5	6.3
15c → 15a	8.4	8.8	8.9	9.4	8.5	9.0	8.7	9.1	–
36a → 36b	8.5	8.9	9.3	9.5	9.0	9.3	8.9	9.1	–
4a → 4b	9.0	9.1	9.4	9.5	9.2	9.4	9.2	9.4	–
35a → 35b	9.6	9.8	9.7	9.9	9.3	9.6	9.5	9.7	–
MSE			0.3	0.3	0.2	0.2	0.2	0.1	
MUE			0.4	0.4	0.3	0.2	0.2	0.2	

[a] From Ref [104]; CCSD(T)/aug-cc-pVXZ single points used for extrapolation

B. Tautomerization Enthalpies

Based on the above benchmark, we computed tautomerization energies at the SMD-M06-2X/def2-TZVP level of theory for 423 tautomeric clusters from the VEHICLe database. Tautomerization energies computed in this manner allow further analysis of data found in the literature and new insights previously into structural trends in tautomeric equilibria.

Annular tautomer systems are often found to be in equilibrium in solution.[38] The importance of computing solution-phase tautomerization energies is assessed by computing the tautomerization energies of the benchmark set in the gas phase, in diethylether, and in water (see Table 3.2). Overall, there is a significant reduction in the tautomerization enthalpies upon solvation, with changes as large as 7.1 kcal mol^{-1}

between predictions for the gas phase versus water. Often, the result is that tautomers that are thermodynamically inaccessible in the gas phase are expected to be present in non-negligible quantities in solution. This is exemplified in cluster **4**. In the gas phase, the 9.5 kcal mol^{-1} difference in energy between **4b** and **4a** should result in zero population of **4b**. In aqueous solution, however, this energy difference is only 1.9 kcal mol^{-1}, leading to the presence of both tautomers in water.

In some cases, the identity of the preferred tautomer is different in solution than in the gas phase. For example, in the gas phase **3a** lies 2.3 kcal mol^{-1} higher in energy than **3b**. In water, this energy difference is predicted to reverse, with **3a** lying 1.2 kcal mol–1 lower than **3b**. In diethyl ether, these two tautomers are predicted to be nearly isoenergetic, suggesting that both tautomers could be present in appreciable quantities in a protein-like dielectric environment.

Transitions **16a → 16b** and **38a → 38b** demonstrate an unexpected increase in the tautomerization enthalpy upon solvation. The difference in behavior of these two systems can be attributed to the inclusion of a chalcogen heteroatom in the aromatic ring systems. Upon further study, 77 systems containing group 16 heteroatoms in the ring have been examined, but only 25 of such systems exhibit the same trend. Despite the increase in relative tautomer energies from gas phase to water, clusters **36** and **38** maintain sufficiently low tautomerization energies (<3.6 kcal mol^{-1}) that both tautomeric states should be present in solution.

Table 3.2. Tautomerization energies in kcal mol^{-1} computed with M06-2X/def2-TZVP//MP2/aug-cc-pVTZ in gas phase, diethylether (DEE), and water.

	ΔE		
Transition	Gas	DEE	Water
1b → 1a	4.4	1.9	1.2
2b → 2a	6.6	3.5	2.8
3b → 3a	2.3	-0.4	-1.2
4a → 4b	9.5	5.1	1.9
15c → 15a	9.3	5.4	4.6
16b → 16a	0.0	0.5	0.9
35a → 35b	9.7	5.6	2.7
36a → 36b	9.4	5.3	2.5
37a → 37b	4.7	1.3	0.2
38a → 38b	2.7	3.1	3.5

C. Comparison of Tautomerization Energies

Relative tautomer energies of most clusters from the Pitt set are given in Table 3. Tautomerization energies computed in this work have been compared against the literature,[50] exhibiting a mean absolute difference (MAD) of 1.3 kcal mol^{-1} between energies computed with MP2 and those computed with M06-2X. While convenient, a low absolute error fails to reveal the more important conclusion—which tautomer is lower in energy. In 12 cases, M06-2X predicted a different ordering of tautomeric structures than MP2, casting some doubt on the veracity of Pitt's conclusions.

Table 3.3. Relative tautomer energies for the tautomer clusters in Figure 3.2.[50] All energies are relative to the lowest energy tautomer of the cluster. Absolute differences (AD) between SMD-M06-2x/def2-TZVP and PCM-MP2/6-311++G(d,p) energies and the mean absolute difference (MAD) are shown (outlier clusters **14** and **25** are omitted; See Table 3.4). All energies are in kcal mol^{-1}.

Transition	E_{MP2}[a]	$E_{M06\text{-}2X}$	AD	Transition	E_{MP2}[a]	$E_{M06\text{-}2X}$	AD
1b → 1a	1.0	1.3	0.4	22a → 22d	6.9	6.9	0.0
2b → 2a	1.8	3.0	1.2	22a → 22c	7.9	8.9	1.0
3a → 3b	3.2	0.8	2.4	23a → 23b	0.5	-0.1	0.6
4a → 4b	2.0	2.3	0.3	23a → 23d	5.0	5.1	0.1
5b → 5a	2.2	2.8	0.6	23a → 23c	8.3	9.4	1.1
6a → 6b	9.4	7.8	1.6	24b → 24a	0.8	0.5	0.3
7a → 7b	6.3	3.9	2.4	24b → 24d	4.1	4.4	0.3
8a → 8b	6.3	3.6	2.7	24b → 24c	5.7	6.1	0.4
9a → 9b	6.2	3.3	2.9	26a → 26b	2.0	2.7	0.6
10a → 10b	0.3	-0.1	0.3	27b → 27a	3.3	3.3	0.1
11a → 11b	0.8	0.3	0.5	28b → 28a	1.1	1.7	0.6
12b → 12a	0.9	-0.6	1.5	29a → 29b	7.9	6.0	1.9
13a → 13b	4.8	2.0	2.8	30a → 30b	3.4	3.5	0.1
15c → 15a	3.9	4.9	1.0	30a → 30c	13.1	9.9	3.3
15c → 15b	12.5	13.2	0.7	31a → 31b	2.1	1.4	0.7
16b → 16a	0.2	1.1	0.9	32a → 32c	5.5	5.6	0.1
17a → 17b	2.9	2.8	0.1	32a → 32b	16.7	17.9	1.1
18a → 18b	3.2	3.3	0.0	33b → 33a	1.6	1.7	0.1
19a → 19b	7.1	6.3	0.8	34d → 34c	2.6	2.7	0.1
20a → 20b	3.6	3.6	0.0	34d → 34a	2.8	2.4	0.4
21a → 21b	6.2	7.6	1.4	34d → 34b	4.5	4.8	0.3
22a → 22b	2.1	0.8	1.4			MAE	1.2

[a]Ref [50]; Relative tautomer energies computed with PCM-MP2/6-311++G(d,p)

Tautomers clusters **14** and **25** exhibit eight of the 12 instances of differences in tautomer ordering between MP2 and M06-2X. In these clusters, the data show large discrepancies between computed tautomerization energies with a radically different ordering for the lowest energy tautomer. In efforts to elucidate these differences, we attempted to reproduce the results of Pitt *et al.*[50] (Table 3.5), but were unable to do so.

Instead, recomputed tautomerizations are used as a correction to the Pitt set, reducing the number of inconsistently ordered tautomers between the two levels of theory from 12 to ten. Although there is no drastic improvement in tautomer ordering, the MAD between the MP2 and M06-2X data for the recomputed values reduces from 3.5 to 1.5 for literature values compared to the corrected values. The corrected MAD of 1.5 is in far better agreement with the MAD 1.2 from the test set with outliers removed (Table 3.3).

Table 3.4. Tautomerization energies computed in this work using the same PCM-MP2/6-311++G(d,p) as Pitt *et al*.[50] and the previously mentioned SMD-M06-2X/def2-TZVP//M06-2X/def2-TZVP. Mean absolute errors are computed and all energies are in kcal mol^{-1}.

Transition	Pitt[a]	$MP2_{adj}$	M06-2X	AD_{Pitt}	AD_{adj}
14a → 14c	3.5	-0.9	2.0	1.5	2.9
14a → 14b	6.4	-0.4	1.1	5.3	1.6
25a → 25b	3.6	-1.9	-0.2	3.8	1.7
25a → 25c	4.0	6.8	7.1	3.1	0.3
25a → 25d	7.9	-0.4	1.3	6.7	1.7
25a → 25f	8.0	11.1	9.2	1.1	2.0
25a → 25e	9.3	6.9	6.5	2.8	0.4
			MAD	3.5	1.5

[a]Ref [50]; Relative tautomer energies computed with PCM-MP2/6-311++G(d,p)

D. Extension of Tautomer Classes

To derive a more complete characterizations of the ten tautomer classes devised by Pitt *et al*. (Figure 3.1), additional examples in each class were examined for tautomerization energy and structural preference. For each tautomer class, tautomeric preference towards tautomer X (the left structure in Figure 3.1) is examined. The

tautomer energies relative to tautomer X and frequencies of tautomer X being lower in energy were computed and are listed in (Figure 3.3 and Table 3.5).

Table 3.5. Overall count of tautomerizations falling into each class from the expanded set of 631 transitions. Frequency of structural preference to tautomer X is given, demonstrating overall structural preferences within certain classes.

Class	Cases	% Taut X
T1	89	47
T2	113	19
T3	82	24
T4	138	99
T5	66	44
T6	47	100
T7	16	100
T8	26	96
T9	49	41
T10	5	0

Out of ten tautomer classes, four demonstrate a distinct structural preference, two demonstrate a likely preference, and four remain inconclusive. Previously unreported were structural preferences for classes T5, T6, T7, and T8, but the data gathered in this study have resolved some of these structures. Classes T6, T7, and T8 strongly favor tautomer X (with near 100%) and generally feature larger relative tautomer energies. As demonstrated in the Pitt Set, class T4 continues to exhibit a preference towards its amino form in all but one of 138 cases, where the relative tautomer energy is less than 1 kcal mol^{-1}. Pitt *et al.*[50] recognized that classes T2 and T9 are symmetric about an axis, and distinguishing tautomer X is arbitrary. To address this issue, we applied the condition that the tautomer X will be defined as the N-H closest to a higher priority group. Despite the introduction of this condition, there is no indication of a structural preference for class

T9. However, with class T2, only 19% of tautomer pairs favor tautomer X, indicating an opposing preference for tautomer Y. Similarly, tautomer class T3 shows a 76% preference towards tautomer Y. In most clusters, two T3 transitions are found due to asymmetry but effects on relative tautomer energies from substituents attached to the fused ring have not yet been studied. Class T10 is the only class characterized by movement of 2 hydrogen atoms and is preferential to tautomer Y in all 5 cases. This can be attributed to changes in aromaticity, as a Clar sextet ring can be drawn in tautomer Y but is lost in tautomer X. Tautomer classes T1, T5, and T9 give no clear indication of a structural preference.

While it lacks a structural preference, class T1 contains the smallest spread of relative tautomer energies (< 5 kcal mol^{-1}) whereas T5 and T9 exhibit a wider range. Higher relative tautomer energies are more difficult to overcome with differences in other interactions. Therefore, transitions falling into classes T1 and T3 are most likely to experience a change in major tautomer depending on the environment, while T2, T4, and T9 are less likely to undergo such a change. The high relative tautomer energies from classes T6, T7, and T8 indicate a resistance to changing tautomeric states based on environment. For further study, tautomer class T4 provides the most promising results since it is the most prevalent tautomerization class among synthetically tractable heterocycles, has a clear structural preference, and has a reasonable number of cases with low relative tautomer energies.

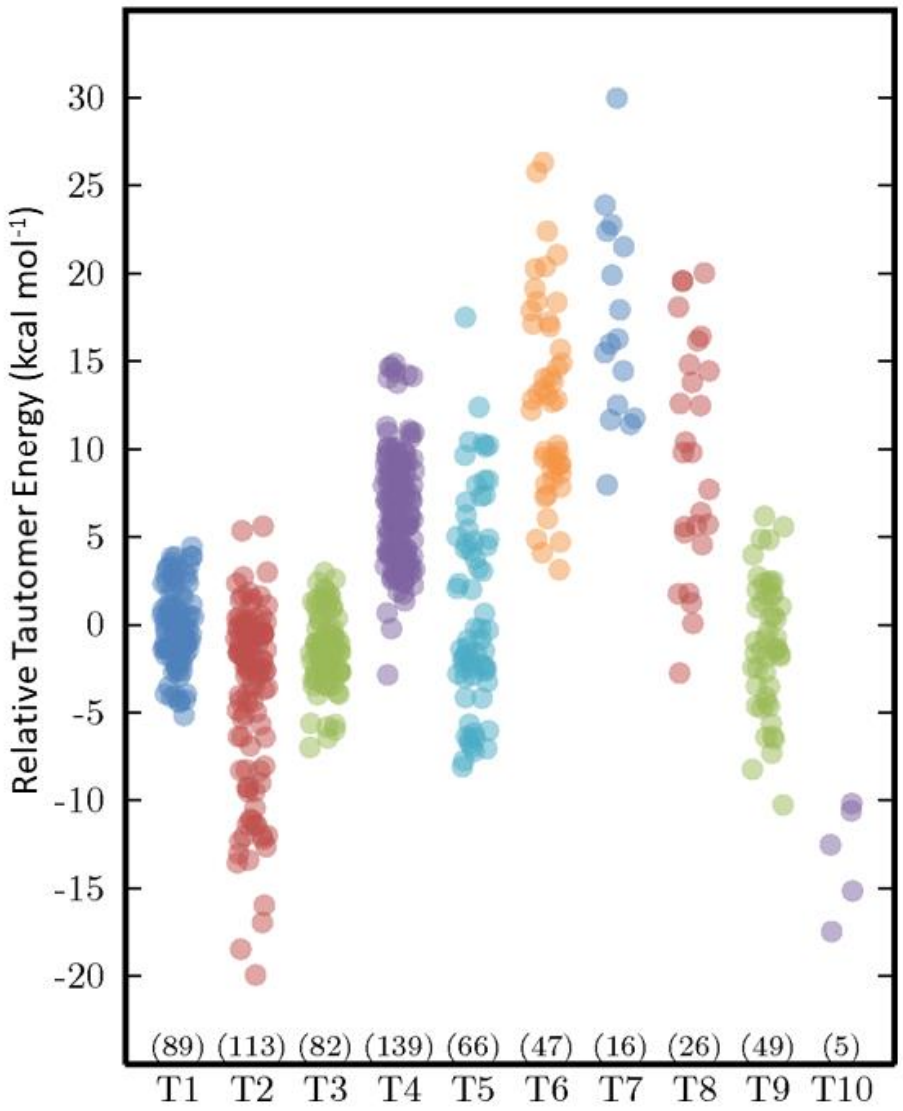


Figure 3.3. Strip plot showing relative tautomer energies for transitions in each tautomer class computed with SMD-M06-2X/def2-TZVP//M06-2X/def2-TZVP. Numbers in parenthesis give the number of transitions in each tautomer class. Positive relative energies indicate preference towards tautomer X.

Conclusions

Annular tautomerism is an important form of tautomerism that has not been fully characterized. The present study has expanded upon previous research[50] to gain insight

on various annular tautomerisms in common ring systems. The data indicate that six of the ten tautomer classes delineated by Pitt *et al*.[50] contain clear structural preferences. We also examined the effects of solvent on relative tautomer energies, identifying several cases in which the preferred tautomer can change depending on the dielectric environment. Noting that tautomeric states are highly dependent on the environment, we identify specific proton movements which are more prone to changes in the tautomeric equilibrium.

CHAPTER 4

ANNULAR TAUTOMERISM MEDIATED BY STACKING INTERACTIONS[c]

[c] Soto, B.T. and S. E. Wheeler. *To be submitted to J. Am. Chem. Soc.*

Abstract

Many heterocycles that commonly occur in pharmaceuticals can exhibit multiple tautomeric states. The preference for one tautomeric form over others can be perturbed by non-covalent interactions that occur in drug binding sites, including stacking interactions with aromatic amino acid residues. This can lead to cases in which the preferred tautomer in the bound state will differ from that in solution. We considered14 tautomer pairs of drug-like heterocycles for which stacking energies with amino acid side chains of Phe, Trp, and Tyr are expected to change the tautomeric equilibrium. High-level *ab initio* computations confirm that 13 of these tautomer pairs should exhibit a different preferred tautomeric state in the bound state than in solution.

Introduction

The ability to predict the preferred tautomeric state of potential pharmaceutical compounds in both solution and in drug binding sites is vital for structure-based drug design (SBDD). In cases in which the tautomer present in the bound state differs from that of the drug in solution, properly predicting drug binding affinities will require an accounting of the energetic cost of adopting the higher-energy conformer.

Tautomerism occurs in ~25% of pharmaceuticals[38] and can alter molecular properties including the ability to engage in different non-covalent interactions in drug binding sites. For example, in the case of hydrogen or halogen bonding interactions tautomerization can switch donor and acceptor character. Furthermore, Wheeler *et al*.[56, 111-113] demonstrated the local nature of substituent and heteroatom effects in stacking

interactions, suggesting that tautomerism can also change the strength of stacking interactions energies.

At the same time, the preferred tautomeric state can be impacted by non-covalent interactions. The most commonly observed means of tuning tautomeric equilibria is through hydrogen bonding interactions. For instance, analyses of the PDB[43] and CSD[105] have revealed many cases in which the preferred tautomer in the solid state differs from that in solution. This has been attributed to the preferential stabilization of the 'disfavored' tautomer through non-covalent interactions (primarily hydrogen-bonding interactions). However, stacking interactions can also impact tautomeric equilibria. For example, An *et al*.[56] identified several small heterocycles commonly found in pharmaceuticals for which differences in stacking energies with 9-methyladenine exceeded the tautomerization energies, indicating that stacking interactions alone can switch the preferred tautomer for some heterocycles. Additionally, Seifert et al.[114] demonstrated how stacking drives dimerization of the 1-napthol dimer over hydrogen bonding. There have been many advances in our understanding of the effects of substituents and heteroatoms on stacking interactions;[111-113, 115-131] however, there has been no study to our knowledge of the impact of stacking on tautomerization energies.

In Chapter 3, we predicted tautomerization energies for 631 annular tautomeric pairs and identified quantitative and structural trends within the ten classes of Pitt *et al*.[50] We found clear structural biases in six of these classes and that for four of these classes the tautomerization energies are systematically very low. Such cases will be ripe for qualitative changes in tautomeric equilibria due to differential stacking interactions. By combining the tautomerization energies from Chapter 3 with recently predicted stacking

interaction energies for drug-like heterocycles with the aromatic amino acid side chains Phe, Tyr, and Trp,[78] we identify 14 tautomeric pairs of heterocycles for which stacking interactions with binding site residues have the potential to change the nature of the preferred tautomer. Rigorous computational studies of stacked dimers of these drug-like heterocycles with model Phe, Tyr, and Trp side chains confirm that many are indeed expected to exhibit a different preferred tautomer when engaged in stacking interactions than when in isolation.

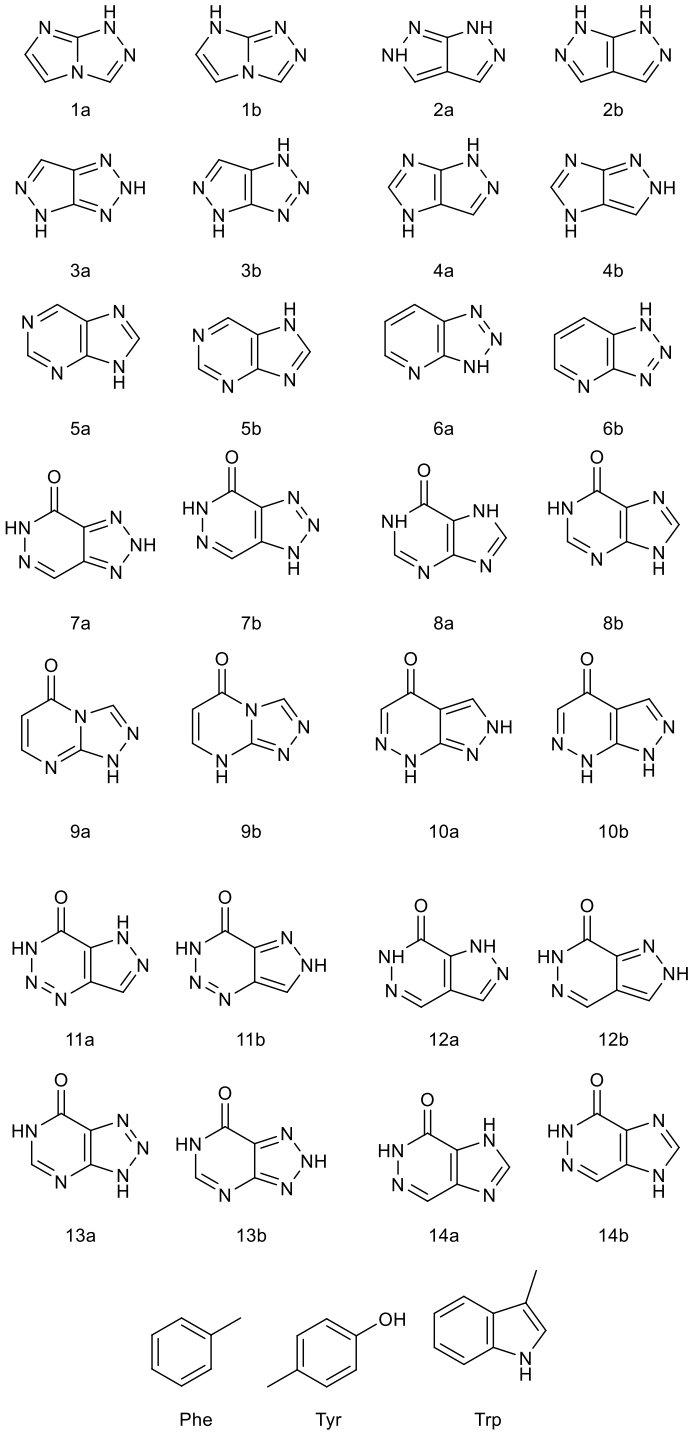
Theoretical Methods

We systematically search for stacked dimers of 14 tautomeric pairs of heterocycles (see Scheme 4.1) with the aromatic amino acid side chains Phe, Tyr, and Trp starting from 12 initial stacked geometries (see Bootsma *et al*.[132] for details) at the wB97X-D/def2-TZVP level of theory.[14, 109] To ensure all dimers were stacked, the heavy atoms were constrained to lie in parallel planes during geometry optimizations. Unique local minima were identified based on an RMSD cutoff of 0.4 Å. Single point energies of all unique stacked minima were evaluated at the DLPNO-CCSD(T)/cc-pVQZ level of theory[12, 15-18] to determine the lowest energy stacked dimer for each heterocycle with the three amino acid side chains.

Relative tautomer energies (E_{rel}, in water) were taken from Chapter 3, since in biochemical and pharmaceutical systems heterocycles generally exist in an aqueous medium before entering the binding pocket. To simulate stacking interactions within the dielectric environment of a typical protein, we used diethylether as a solvent. Since solvent corrections are not implemented for DLPNO-CCSD(T) computations, solvent

effects on stacked systems were computed at the SMD-wB97X-D/def2-TZVP level of theory. All DFT computations were completed using the Gaussian 09 quantum chemical package.[133] Computed gas-phase and solution-phase stacking interaction energies for these global minimum structures are listed in Tables B.1-B.3 in Appendix B.

Predicted stacking energy differences from the model of Bootsma *et al*.[78] are referenced as ΔE_{stack}^{pred}. DLPNO-CCSD(T)/cc-pVQZ stacking interactions were computed as the difference in energy between the (constrained) optimized stacked dimer and the separated optimized monomers, and are referred to as ΔE_{stack}^{gas} and ΔE_{stack}^{soln} for gas-phase and solution-phase (diethyether), respectively. Relative tautomer energies (in water) from Chapter 3 are designated as E_{rel} and the influence of stacking on the relative tautomer energy ($E_{rel} + \Delta E_{stack}^{z}$) will be referred to as ΔE_{r+s}^{z} where z is either pred or soln for predicted and solvent-corrected energies.



Scheme 4.1. 14 tautomer pairs for which Bootsma *et al*. predict the higher energy tautomer is will engage in stacking interaction larger than the relative tautomer energy. Also shown are the models of the amino acid side chains Phe, Tyr, and Trp.

Results and Discussion

Recently, Bootsma *et al.*[78] developed a multivariate linear model (equation 4.1) to predict stacking interactions between drug-like heterocycles and Phe, Trp, and Tyr. Equation 4.1 relies on simple heterocycle descriptors derived from the electrostatic potential (ESP) in the vicinity of the heterocycle as well as the heavy atom count. Consequently, stacking interactions between heterocycles and amino acid side chains can be predicted rapidly, without performing geometry optimizations of stacked dimers.

$$\mathrm{E}_{\mathrm{stack}}^{\mathrm{pred}} = N_{HA}^{AA}(-0.036ESP_{mean}^{Het} - 0.013ESP_{range}^{Het} - 0.095N_{HA}^{Het}) - 1.36 \qquad (4.1)$$

We used equation 4.1 to predict stacking interactions of Phe, Tyr, and Trp with the 631 annular tautomer pairs whose relative energies (E_{rel}) were evaluated in Chapter 3 above. By combining the E_{rel} values from Chapter 3 with the $\mathrm{E}_{\mathrm{stack}}^{\mathrm{pred}}$ values from equation 4.1, we identified 14 tautomer pairs for which the difference in $\mathrm{E}_{\mathrm{stack}}^{\mathrm{pred}}$ values between the tautomers ($\Delta\mathrm{E}_{\mathrm{stack}}^{\mathrm{pred}}$) exceeds the tautomerization energy (see Scheme 4.1):

$$\mathrm{E}_{\mathrm{rel}} + \mathrm{E}_{\mathrm{stack}}^{\mathrm{pred}} < 0 \qquad (4.2)$$

In other words, for these 14 tautomer pairs equation 4.1 predicts that differences in stacking interactions can overcome the difference in tautomer energies, changing the identity of the preferred tautomer. As expected, these are all cases in which the difference in energy between the major and minor tautomers in solution is small (< 0.5 kcal mol^{-1}).

To both verify the predictions from equation 4.1 and to test whether these 14 tautomeric equilibria will really be reversed through stacking interactions, we explicitly computed the maximum stacking interaction between these 28 heterocycles and Phe, Tyr,

and Trp. First, we performed an extensive search of local stacked energy minima with toluene, *p*-methylphenol, and 3-methylindole as models of the side chains of Phe, Tyr, and Trp, respectively to identify the global minimum stacked structures. We then evaluated accurate gas-phase stacking interaction energies (E_{stack}^{gas}) for all unique stacked dimers using DLPNO-CCSD(T). To account for the dielectric environment of a typical protein, we then used wB97X-D/def2TZVP to compute solvent corrections for diethyether, whose dielectric constant is similar to that in many proteins. The resulting (E_{stack}^{soln}) values should provide a reasonable estimate of the stacking expected in a protein binding site.

Impact of Stacking on Tautomeric Equilibria

Table B.1 lists relative tautomer energies (E_{rel}) and the relative stacking interaction energies in a protein-like environment (ΔE_{stack}^{soln}). The sum of these two values is the computed tautomerization energy in the stacked state, $E_{rel}^{stacked}$ (Figure 4.1). Negative $E_{rel}^{stacked}$ indicate cases in which the minor tautomer in aqueous solution is expected to be the major tautomer when stacked with the given amino acid side chain. In total, we identify 16 cases in which stacking interactions are able to reverse the preferred tautomer present in solution: five with Phe, two with Tyr, and nine with Trp.

Looking at the data more closely, the most negative ΔE_{stack}^{soln} values for Phe, Trp, and Tyr are -0.7, -1.5, and -0.7 for Phe, Tyr, and Trp, respectively. In two of three cases (Phe and Tyr) this maximum effect occurs for 5a → 5b and stacking with Trp yields the second-most negative ΔE_{stack}^{soln} of -0.9. In this case, the tautomerization energy in water is 0.3 kcal mol^{-1} favoring 5a. When stacked with Phe, Tyr, and Trp, the tautomerization

energy changes to –0.3, –0.6, and –0.3 (*i.e.* favoring **5b**). Taking this as the maximum impact of differential stacking interactions in tautomeric equilibria, stacking interactions have the potential to shift tautomeric equilibria for cases in which the one tautomer is preferred over the other by nearly 1 kcal mol^{-1}. If we use 1 kcal mol^{-1} as a cutoff, instead of 0 kcal mol^{-1} as in equation 4.2, then there are 58 examples from Chapter 3 for which stacking interactions could potentially change the identity of the tautomeric equilibrium. These could be examined in future work.

Of course, for many of the cases the impact of stacking interactions shifts the tautomeric equilibrium every further toward the major tautomer in solution. For example, for 8a → 8b, the tautomerization energy favors 8a by only 0.1 kcal mol^{-1}. Upon stacking with Phe, Tyr, and Trp, this energy increases to 0.6, 0.7, and 0.7, respectively, so will increase the proportion of 8a significantly. This could have important implications for the design of drugs containing this heterocyclic framework.

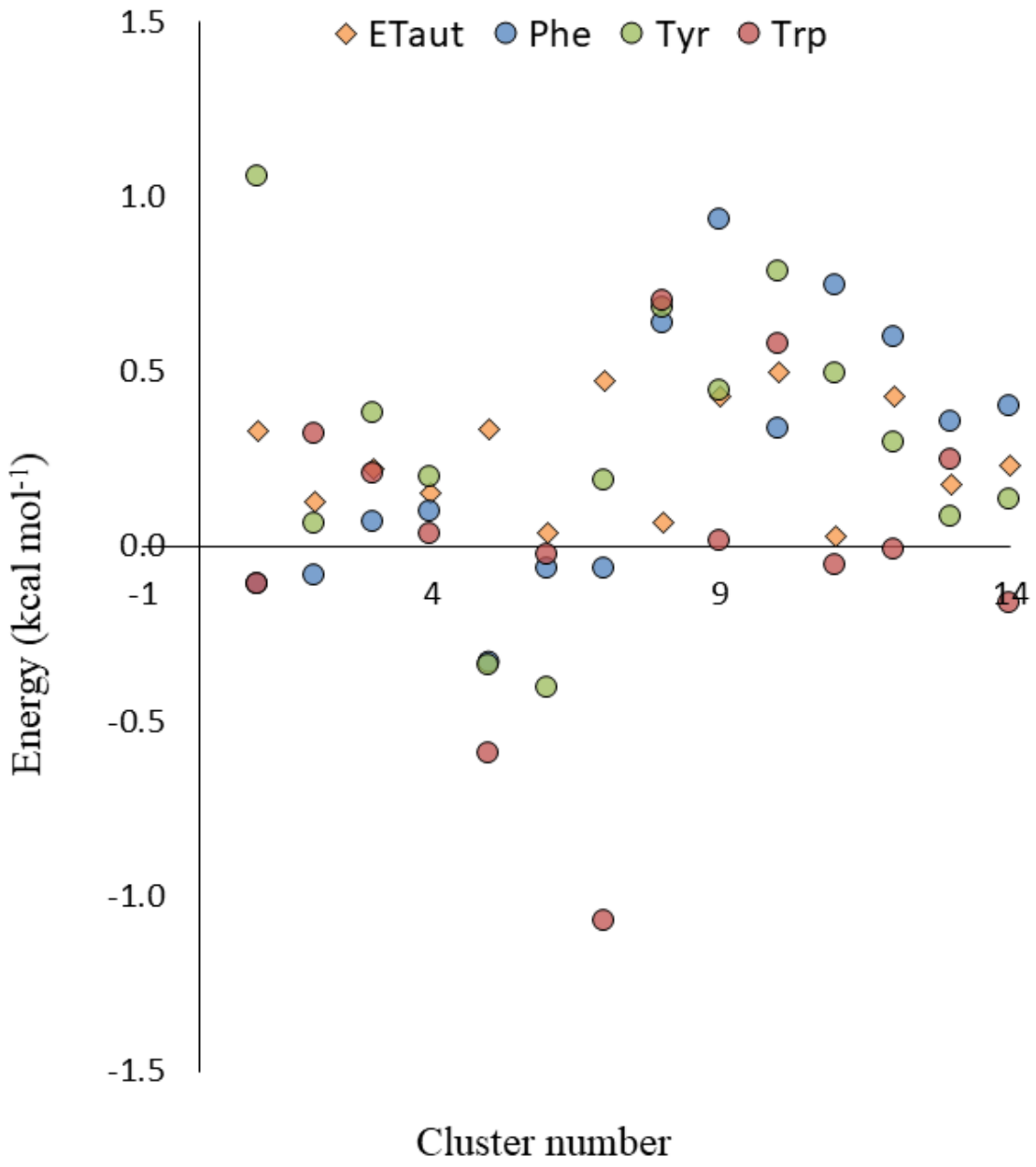


Figure 4.1 Influence ofSMD-ωB97X-D/def2-TXVP solvent-corrected DLPNO-CCSD(T)/cc-pVQZ stacking energies (ΔE_{stack}^{soln}) on tautomerization energies computed with SMD-M06-2X/def2-TZVP//M06-2X/def2-TZVP in Chapter 3 for clusters in scheme 4.1. Negative values indicate systems where the stacking is found to change the tautomeric equilibrium.

It's also instructive to consider these data in the context of the tautomer classes from Pitt et al.[50] (Figure 3.1). The 14 tautomeric pairs in Scheme 4.1 below to classes T1(**2**,**4**,**10**,**11**,**12**), T2(**5**,**6**,**8**,**14**), T3(**3**,**7**,**13**), and T9(**1**,**9**) transitions. These four classes were shown in Chapter 3 to exhibit systematically small tautomerization energies, so provide a clear means of identifying heterocycles in which stacking (and other non-covalent interactions) have the potential to reverse tautomeric equilibria.

It is worth noting that this work does not provide the complete or even realistic picture of the impact of drug binding on tautomeric equilibrium, as other non-covalent interactions as well as the heterogeneous electrostatic environment will impact these results. Regardless, these data provide proof of principle that stacking interactions alone have the potential to qualitatively change tautomeric equilibria of drug-like heterocycles. The use of the tautomer classes from Pitt *et al*.[50] in tandem with the predictive model from Bootsma *et al*.[78] (equation 4.1) provide a potential process for identifying such systems.

Analysis of Predicted Vs. Computed stacking Interaction Strengths

The reliable gas-phase stacking data computed above also provide an opportunity to assess the accuracy of the predictive model of Bootsma and co-workers.[78] Figure 4.2 shows stacking interactions energies computed with DLPNO-CCSD(T)/cc-pVQZ energies plotted against those predicted using equation 4.1. Generally, stacking energies for Phe are well-predicted. Conversely, Tryptophan stacking energies are overpredicted

while Tyrosine stacking energies are underpredicted. These deviations are potentially due to the planar constraints applied in the present work, compared to the more lenient requirement that the planes of the two arenes lie within 30° used in the work of Bootsma *et al*.[78]

In later work on DNA base stacking, Harding *et al*.[134] parametrized a model using simple ESP descriptors and a planar geometric constraint (Equation 4.3).

$$\mathrm{E}_{\mathrm{stack}}^{\mathrm{pred}} = 0.003 N_{HA}^{AA} N_{HA}^{Het} (ESP_{range}^{AA} + ESP_{range}^{Het}) - 4.181 \quad (4.3)$$

With this model, new stacking predictions are made (figure 4.3) and the prediction accuracy is compared to the Bootsma model.

Overall, the Bootsma model demonstrates good correlation with each individual amino acid with R^2 values of 0.85, 0.80, and 0.86 for Phe, Tyr, and Trp respectively. Conversely, the Harding model demonstrates poor correlation with each amino acid, showing R^2 values of 0.55, 0.44, and 0.74 for Phe, Tyr, and Trp, respectively. However, the Harding model performs well across all three amino acids, with a total $R^2 = 0.80$ while the Bootsma model is fragmented such that overall correlation is clearly very poor. Future work could examine a model created as a hybrid of the Bootsma and Harding models.

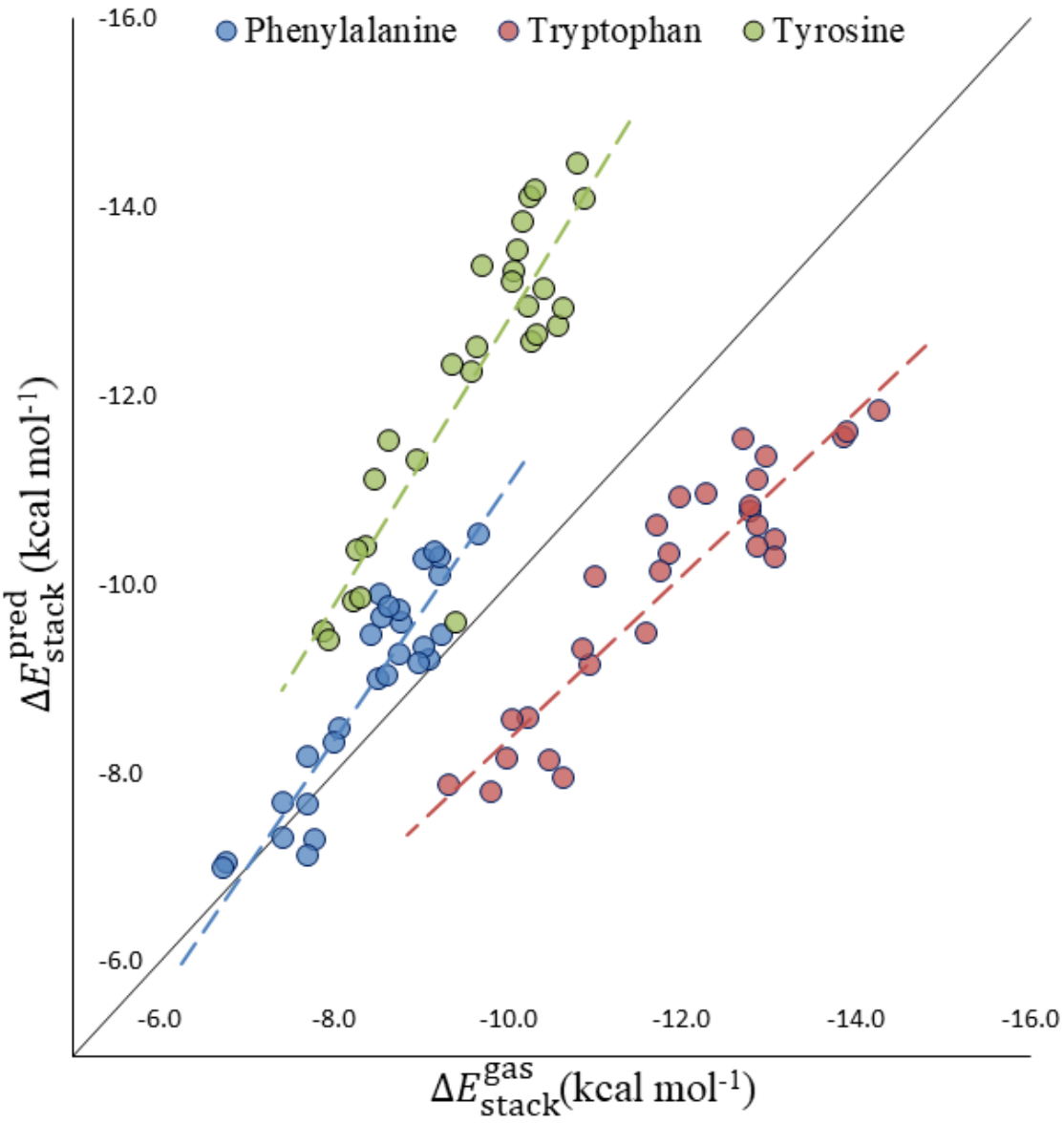


Figure 4.2. Interaction energies in kcal mol^{-1} for stacking with model side chains of Phe, Trp, and Tyr predicted by the model of Bootsma *et al.*[78] and computed in the gas phase at the DLPNO-CCSD(T)//cc-pVQZ level of theory. Data points above the trendline indicate overprediction of stacking interactions by the model, while points below the line indicate underprediction.

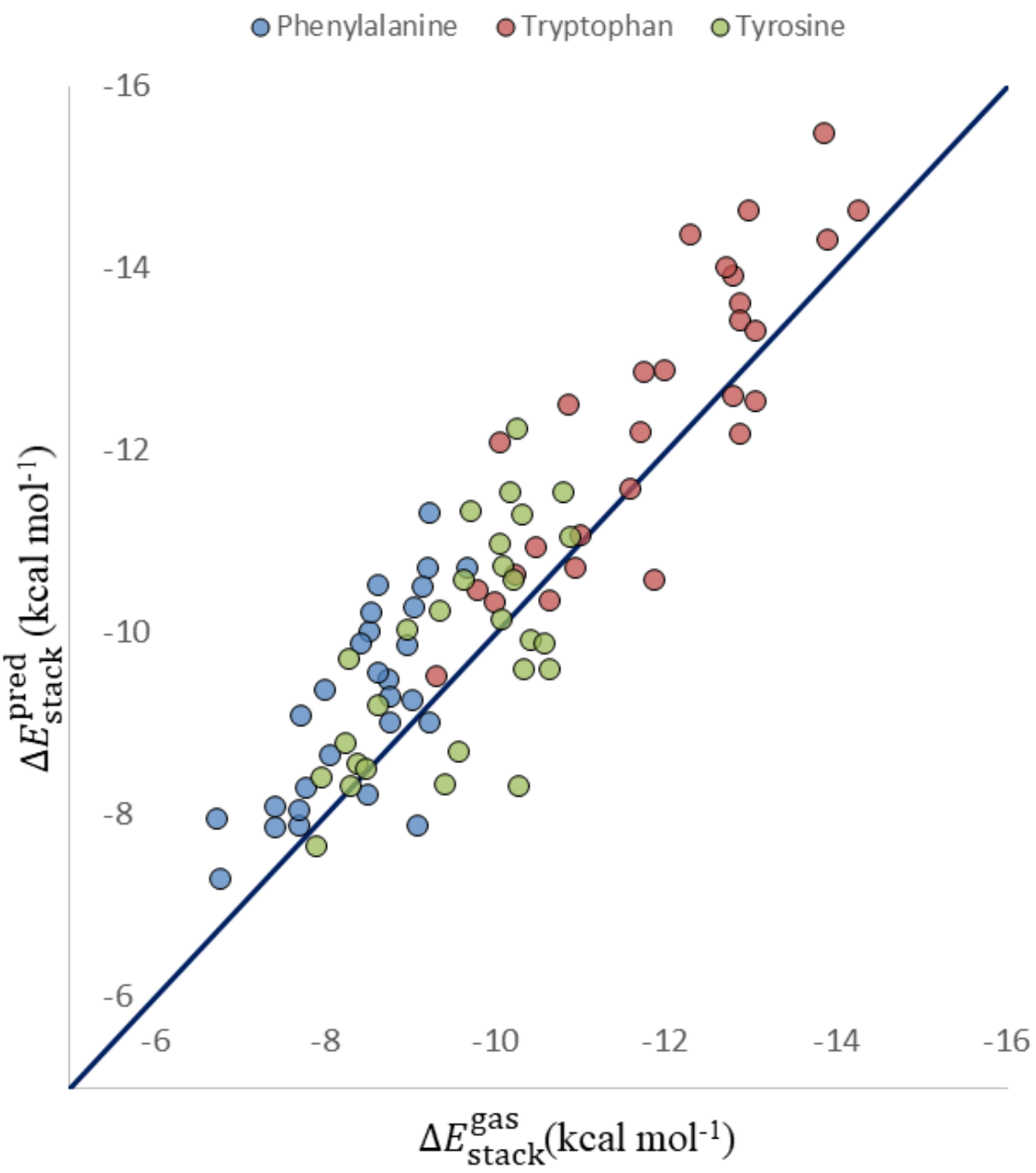


Figure 4.3. Interaction energies in kcal mol^{-1} for stacking with model side chains of Phe, Trp, and Tyr predicted by the model of Harding *et al.*[134] and computed in the gas phase at the DLPNO-CCSD(T)//cc-pVQZ level of theory. Data points above the trendline indicate overprediction of stacking interactions by the model, while points below the line indicate underprediction.

Conclusions

Using the predictive model of stacking interactions of Bootsma *et al.*[78] combined with tautomerization energies from Chapter 3, we have identified candidates where stacking interactions with aromatic amino acid residues can potentially shift tautomeric equilibria towards the higher energy tautomer. Accurate *ab initio* computations of stacked dimers confirm many cases in which stacking interactions alone can lead to the minor tautomer present in solution to become the major tautomer in a bound state. These cases occur most often for certain classes of annular tautomers examined in Chapter 3, providing a simple means of identifying tautomeric heterocycles for which there could be a switch of tautomeric states upon binding to a protein. This has important implications for structure-based drug design, as changes in tautomeric state must be accounted for both in terms of the energetic cost of forming the higher-energy tautomer and any resulting changes in non-covalent interactions with binding site functional groups.

CHAPTER 5

CONCLUSIONS

Each of these projects utilized computational methods to address a range of chemical problems. From combustion chemistry providing insights to the future of biofuels to studies on tautomeric systems in drug-like molecules, computational methods drive continued advancement in understanding chemical systems. Using high-accuracy coupled-cluster methods and focal-point analysis, I have reiterated the importance of considering all steps in a chemical pathway, as the pathway with the slowest initiation ultimately ends up as the lowest energy pathway. By benchmarking multiple density functional methods against CBS energies, I determined which method would perform best for computations on annular tautomer systems. I then demonstrated the importance including solvent effects in computations on tautomeric systems. Furthermore, I expanded upon existing data to identify structural and energetic trends across ten classes of annular tautomerization transitions. Finally, I examined the impact of stacking interactions with the aromatic amino acids Phe, Tyr, and Trp on tautomeric equilibria, identifying a number of cases in which stacking interactions are predicted to change the identity of the preferred tautomer upon binding. Overall, each of these projects has provided a unique application of computational methods yielding useful results for improving chemical data available for solving complex problems.

REFERENCES

1. Hunter, C. A.; Sanders, J. K. M., The Nature of π-π Interactions. *J. Am. Chem. Soc.* **1990,** *112*, 5525-5534.

2. Grimme, S., Do Special Noncovalent π-π Stacking Interactions Really Exist? *Angew. Chem. Int. Ed.* **2008,** *47*, 3430-3434.

3. Martinez, C. R.; Iverson, B. L., Rethinking the Term "π-Stacking". *Chem. Sci.* **2012,** *3*, 2191-2201.

4. Bloom, J. W.; Wheeler, S. E., Taking the aromaticity out of aromatic interactions. *Angew Chem Int Ed Engl* **2011,** *50* (34), 7847-9.

5. Allinger, N. L.; Fermann, J. T.; Allen, W. D.; Schaefer, H. F., The torsional conformations of butane: Definitive energetics from ab initio methods. *The Journal of Chemical Physics* **1997,** *106*, 5143-5150.

6. Csaszar, A. G.; Allen, W. D.; Schaefer, I., Henry F., In pursuit of the ab initio limit for conformational energy prototypes. *The Journal of chemical physics* **1998,** *108*, 9751-9764.

7. East, A. L. L.; Allen, W. D., The heat of formation of NCO. *J. Chem. Phys.* **1993,** *99*, 4638-4650.

8. Gonzales, J. M.; Pak, C.; Cox, R. S.; Allen, W. D.; Schaefer III, H. F.; Császár, A. G.; Tarczay, G., Definitive ab initio studies of model SN2 reactions CH(3)X+F- (X=F, Cl, CN, OH, SH, NH(2), PH(2)). *Chemistry (Weinheim an der Bergstrasse, Germany)* **2003,** *9*, 2173-92.

9. Allen, W. D.; East, A. L. L.; Császár, A. G., In *Structures and Conformations of Non-Rigid Molecules*, Laane, J.; Dakkouri, M.; van der Veken, B.; Oberhammer, H., Eds. Kluwer: Dordrecht, 1993; pp 343-373.

10. Császár, A. G.; Allen, W. D.; Schaefer, H. F., In Pursuit of the ab Initio Limit for Conformational Energy Prototypes *J. Chem. Phys.* **1998,** *108*, 9751-9764.

11. East, A. L. L.; Allen, W. D., The Heat of Formation of NCO. *J. Chem. Phys.* **1993,** *99*, 4638-4650.

12. Dunning, T. H., Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989,** *90* (2), 1007-1023.

13. Zhao, Y.; Truhlar, D. G., The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008,** *120* (1-3), 215-241.

14. Chai, J. D.; Head-Gordon, M., Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008,** *10* (44), 6615-20.

15. Pinski, P.; Riplinger, C.; Valeev, E. F.; Neese, F., Sparse maps-A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals. *J. Chem. Phys.* **2015,** *143* (3), 034108.

16. Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F., Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *J. Chem. Phys.* **2013,** *139* (13), 134101.

17. Riplinger, C.; Neese, F., An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013,** *138* (3), 034106.

18. Neese, F.; Hansen, A.; Liakos, D. G., Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. *J. Chem. Phys.* **2009,** *131* (6), 064103.

19. Balat, M.; Balat, H., Recent trends in global production and utilization of bio-ethanol fuel. *Applied Energy* **2009,** *86* (11), 2273-2282.

20. Mojović, L.; Nikolić, S.; Rakin, M.; Vukasinović, M., Production of bioethanol from corn meal hydrolyzates. *Fuel* **2006,** *85* (12), 1750-1755.

21. Sarkar, N.; Ghosh, S. K.; Bannerjee, S.; Aikat, K., Bioethanol production from agricultural wastes: An overview. *Renewable Energy* **2012,** *37* (1), 19-27.

22. Lam, F. H.; Ghaderi, A.; Fink, G. R.; Stephanopoulos, G., Biofuels. Engineering alcohol tolerance in yeast. *Science (New York, N.Y.)* **2014,** *346*, 71-75.

23. Bothast, R. J.; Schlicher, M. A., Biotechnological processes for conversion of corn into ethanol. *Applied Microbiology and Biotechnology* **2005,** *67*, 19-25.

24. Deyab, M. A., Adsorption and inhibition effect of Ascorbyl palmitate on corrosion of carbon steel in ethanol blended gasoline containing water as a contaminant. *Corrosion Science* **2014,** *80*, 359-365.

25. Niven, R. K., Ethanol in gasoline: environmental impacts and sustainability review article. *Renewable and Sustainable Energy Reviews* **2005,** *9* (6), 535-555.

26. Jones, D. T.; Woods, D. R., Acetone -Butanol Fermentation revisited. *Microbiol rev* **1986,** *50*, 484-524.

27. Jin, C.; Yao, M.; Liu, H.; Lee, C. F. F.; Ji, J., Progress in the production and application of n-butanol as a biofuel. *Renewable and Sustainable Energy Reviews* **2011,** *15*, 4080-4106.

28. Dürre, P., Biobutanol: An attractive biofuel. In *Biotechnology Journal*, 2007; Vol. 2, pp 1525-1534.

29. Wu, M.; Wang, M.; Liu, J.; Huo, H., Assessment of Potential Life-Cycle Energy and Greenhouse Gas Emission Effects. *Biotechnology Progress* **2008**, 1204-1214.

30. Väisänen, S.; Havukainen, J.; Uusitalo, V.; Havukainen, M.; Soukka, R.; Luoranen, M., Carbon footprint of biobutanol by ABE fermentation from corn and sugarcane. *Renewable Energy* **2016,** *89*, 401-410.

31. Wallner, T.; Miers, S. A.; McConnell, S., A Comparison of Ethanol and Butanol as Oxygenates Using a Direct-Injection, Spark-Ignition Engine. *Journal of Engineering for Gas Turbines and Power* **2009,** *131*, 32802.

32. Harper, M. R.; Van Geem, K. M.; Pyl, S. P.; Marin, G. B.; Green, W. H., Comprehensive reaction mechanism for n-butanol pyrolysis and combustion. *Combustion and Flame* **2011,** *158*, 16-41.

33. Sarathy, S. M.; Thomson, M. J.; Togbé, C.; Dagaut, P.; Halter, F.; Mounaim-Rousselle, C., An experimental and kinetic modeling study of n-butanol combustion. *Combustion and Flame* **2009,** *156*, 852-864.

34. Veloo, P. S.; Westbrook, C. K.; Wang, Y. L.; Egolfopoulos, F. N.; Westbrook, C. K., A comparative experimental and computational study of methanol, ethanol, and n-butanol flames. *Combustion and Flame* **2010,** *157*, 1989-2004.

35. Stranic, L.; Pyun, S. H.; Davidson, D. F.; Hanson, R. K., Multispecies Measurements in 1-Butanol Pyrolysis Behind Reflected Shock Waves. *Combustion and Flame* **2012,** *159*, 3242-3250.

36. Black, G.; Curran, H. J.; Pichon, S.; Simmie, J. M.; Zhukov, V., Bio-butanol: Combustion properties and detailed chemical kinetic model. *Combustion and Flame* **2010,** *157*, 363-373.

37. Anderson, A. C., The Process of Structure-Based Drug Design. *Chemistry & Biology* **2003,** *10* (9), 787-797.

38. Martin, Y. C., Let's not forget tautomers. *Journal of Computer-Aided Molecular Design* **2009,** *23* (10), 693.

39. Vitaku, E.; Smith, D. T.; Njardarson, J. T., Analysis of the Structural Diversity, Substitution Patterns, and Frequency of Nitrogen Heterocycles among U.S. FDA Approved Pharmaceuticals. *J. Med. Chem.* **2014,** *57* (24), 10257-10274.

40. Broughton, H. B.; Watson, I. A., Selection of heterocycles for drug design. *J. Mol. Graph. Model.* **2004,** *23* (1), 51-58.

41. Taylor, R. D.; MacCoss, M.; Lawson, A. D. G., Rings in Drugs. *Journal of Medicinal Chemistry* **2014,** *57* (14), 5845-5859.

42. Martin, Y. C., Experimental and pKa prediction aspects of tautomerism of drug-like molecules. *Drug Discovery Today: Technologies* **2018,** *27*, 59-64.

43. Milletti, F.; Vulpetti, A., Tautomer Preference in PDB Complexes and its Impact on Structure-Based Drug Discovery. *Journal of Chemical Information and Modeling* **2010,** *50* (6), 1062-1074.

44. Ivanova, D.; Deneva, V.; Nedeltcheva, D.; Kamounah, F. S.; Gergov, G.; Hansen, P. E.; Kawauchi, S.; Antonov, L., Tautomeric transformations of piroxicam in solution: a combined experimental and theoretical study. *RSC Advances* **2015,** *5* (40), 31852-31860.

45. Pospisil, P.; Ballmer, P.; Scapozza, L.; Folkers, G., Tautomerism in Computer-Aided Drug Design. *Journal of Receptors and Signal Transduction* **2003,** *23* (4), 361-371.

46. Katritzky, A. R.; Hall, C. D.; El-Gendy, B. E.-D. M.; Draghici, B., Tautomerism in drug discovery. *Journal of Computer-Aided Molecular Design* **2010,** *24* (6), 475-484.

47. Kasetti, Y.; Bharatam, P. V., Tautomerism in drugs with benzimidazole carbamate moiety: an electronic structure analysis. *Theoretical Chemistry Accounts* **2012,** *131* (3), 1160.

48. Wojnarowska, Z.; Wlodarczyk, P.; Kaminski, K.; Grzybowska, K.; Hawelek, L.; Paluch, M., On the kinetics of tautomerism in drugs: New application of broadband dielectric spectroscopy. *The Journal of Chemical Physics* **2010,** *133* (9), 094507.

49. Skillman, A. G.; Geballe, M. T.; Nicholls, A., SAMPL2 challenge: prediction of solvation energies and tautomer ratios. *J Comput Aided Mol Des* **2010,** *24* (4), 257-8.

50. Cruz-Cabeza, A. J.; Schreyer, A.; Pitt, W. R., Annular tautomerism: experimental observations and quantum mechanics calculations. *Journal of Computer-Aided Molecular Design* **2010,** *24* (6), 575-586.

51. Pitt, W. R.; Parry, D. M.; Perry, B. G.; Groom, C. R., Heteroaromatic Rings of the Future. *Journal of Medicinal Chemistry* **2009,** *52* (9), 2952-2963.

52. Person, W. B.; Szczepaniak, K.; Szczesniak, M.; Kwiatkowski, J. S.; Hernandez, L.; Czerminski, R., Tautomerism of nucleic acid bases and the effect of molecular interactions on tautomeric equilibria. *Journal of Molecular Structure* **1989,** *194*, 239-258.

53. Cornago, P.; Cabildo, P.; Claramunt, R. M.; Bouissane, L.; Pinilla, E.; Torres, M. R.; Elguero, J., The annular tautomerism of the curcuminoid NH-pyrazoles. *New Journal of Chemistry* **2009,** *33* (1), 125-135.

54. Alkorta, I.; Goya, P.; Elguero, J.; P. Singh, S., *A simple approach to the tautomerism of aromatic heterocycles*. 2007; Vol. 30, p 139-159.

55. Elguero, J.; Katritzky, A. R.; Denisko, O. V., Prototropic Tautomerism of Heterocycles: Heteroaromatic Tautomerism—General Overview and Methodology. In *Advances in Heterocyclic Chemistry*, Katritzky, A. R., Ed. Academic Press: 2000; Vol. 76, pp 1-84.

56. An, Y.; Doney, A. C.; Andrade, R. B.; Wheeler, S. E., Stacking Interactions between 9-Methyladenine and Heterocycles Commonly Found in Pharmaceuticals. *Journal of Chemical Information and Modeling* **2016,** *56* (5), 906-914.

57. Sinnokrot, M. O.; Sherrill, C. D., Substituent Effects in π-π Interactions: Sandwich and T-Shaped Configurations. *J. Am. Chem. Soc.* **2004,** *126*, 7690-7697.

58. Cockroft, S. L.; Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J., Electrostatic control of aromatic stacking interactions. *J. Am. Chem. Soc.* **2005,** *127* (24), 8594-5.

59. Cockroft, S. L.; Perkins, J.; Zonta, C.; Adams, H.; Spey, S. E.; Low, C. M.; Vinter, J. G.; Lawson, K. R.; Urch, C. J.; Hunter, C. A., Substituent effects on aromatic stacking interactions. *Org. Biomol. Chem.* **2007,** *5* (7), 1062-80.

60. Wheeler, S. E.; Houk, K. N., Substituent effects in the benzene dimer are due to direct interactions of the substituents with the unsubstituted benzene. *J. Am. Chem. Soc.* **2008,** *130* (33), 10854-5.

61. Ringer, A. L.; Sherrill, C. D., Substituent Effects in Sandwich Configurations of Multiply Substituted Benzene Dimers Are Not Solely Governed By Electrostatic Control. *J. Am. Chem. Soc.* **2009,** *131*, 4574-4575.

62. Cockroft, S. L.; Hunter, C. A., Desolvation and substituent effects in edge-to-face aromatic interactions. *Chem Commun (Camb)* **2009,** (26), 3961-3.

63. Wheeler, S. E., Local nature of substituent effects in stacking interactions. *J. Am. Chem. Soc.* **2011,** *133* (26), 10262-74.

64. Watt, M.; Hardebeck, L. K. E.; Kirkpatrick, C. C.; Lewis, M., Face-to-Face Arene-Arene Bidning Energies: Dominated by Dispersion but Predicted by Electrostatic and Dispersion/Polarizability Substituent Constants. *J. Am. Chem. Soc.* **2011,** *133*, 3854-3862.

65. Muchowska, K. B.; Adam, C.; Mati, I. K.; Cockroft, S. L., Electrostatic modulation of aromatic rings via explicit solvation of substituents. *J. Am. Chem. Soc.* **2013,** *135* (27), 9976-9.

66. Wheeler, S. E.; Bloom, J. W., Toward a more complete understanding of noncovalent interactions involving aromatic rings. *J. Phys. Chem. A* **2014,** *118* (32), 6133-47.

67. Hwang, J.; Li, P.; Carroll, W. R.; Smith, M. D.; Pellechia, P. J.; Shimizu, K. D., Additivity of substituent effects in aromatic stacking interactions. *J. Am. Chem. Soc.* **2014,** *136* (40), 14060-7.

68. Mignon, P.; Loverix, S.; De Proft, F.; Geerlings, P., Influence of stacking on hydrogen bonding: Quantum chemical study on pyridine-benzene model complexes. *J. Phys. Chem. A* **2004,** *108* (28), 6038-6044.

69. Piacenza, M.; Grimme, S., Van der Waals interactions in aromatic systems: structure and energetics of dimers and trimers of pyridine. *ChemPhysChem* **2005,** *6* (8), 1554-8.

70. Swart, M.; van der Wijst, T.; Fonseca Guerra, C.; Bickelhaupt, F. M., Pi-pi stacking tackled with density functional theory. *J. Mol. Model.* **2007,** *13* (12), 1245-57.

71. Itahara, T.; Imaizumi, K., Role of nitrogen atom in aromatic stacking. *J. Phys. Chem. B* **2007,** *111* (8), 2025-32.

72. Hohenstein, E. G.; Sherrill, C. D., Effects of heteroatoms on aromatic pi-pi interactions: benzene-pyridine and pyridine dimer. *J. Phys. Chem. A* **2009,** *113* (5), 878-86.

73. Geng, Y.; Takatani, T.; Hohenstein, E. G.; Sherrill, C. D., Accurately characterizing the pi-pi interaction energies of indole-benzene complexes. *J. Phys. Chem. A* **2010,** *114* (10), 3576-82.

74. Karthikeyan, S.; Nagase, S., Origins of the stability of imidazole-imidazole, benzene-imidazole, and benzene-indole dimers: CCSD(T)/CBS and SAPT calculations. *J. Phys. Chem. A* **2012,** *116* (7), 1694-700.

75. Li, P.; Zhao, C.; Smith, M. D.; Shimizu, K. D., Comprehensive experimental study of N-heterocyclic pi-stacking interactions of neutral and cationic pyridines. *J. Org. Chem.* **2013,** *78* (11), 5303-13.

76. Jacobs, M.; Greff Da Silveira, L.; Prampolini, G.; Livotto, P. R.; Cacelli, I., Interaction Energy Landscapes of Aromatic Heterocycles through a Reliable yet Affordable Computational Approach. *J. Chem. Theory Comput.* **2018,** *14* (2), 543-556.

77. Gung, B. W.; Wekesa, F.; Barnes, C. L., Stacking interactions between nitrogen-containing six-membered heterocyclic aromatic rings and substituted benzene: studies in solution and in the solid state. *J. Org. Chem.* **2008,** *73* (5), 1803-8.

78. Bootsma, A. N.; Doney, A. C.; Wheeler, S. E., Predicting the Strength of Stacking Interactions between Heterocycles and Aromatic Amino Acid Side Chains. *Journal of the American Chemical Society* **2019,** *141* (28), 11027-11035.

79. Kohse-Höinghaus, K.; Oßwald, P.; Cool, T. A.; Kasper, T.; Hansen, N.; Qi, F.; Westbrook, C. K.; Westmoreland, P. R., Biofuel combustion chemistry: From ethanol to biodiesel. In *Angewandte Chemie - International Edition*, 2010; Vol. 49, pp 3572-3597.

80. Sarathy, S. M.; Vranckx, S.; Yasunaga, K.; Mehl, M.; Oßwald, P.; Metcalfe, W. K.; Westbrook, C. K.; Pitz, W. J.; Kohse-höinghaus, K.; Fernandes, R. X.; Curran, H. J., A comprehensive chemical kinetic combustion model for the four butanol isomers. *Combustion and Flame* **2012,** *159*, 2028-2055.

81. Vranckx, S.; Heufer, K. A.; Lee, C.; Olivier, H.; Schill, L.; Kopp, W. A.; Leonhard, K.; Taatjes, C. A.; Fernandes, R. X., Role of peroxy chemistry in the high-pressure ignition of n -butanol – Experiments and detailed kinetic modelling. *Combustion and Flame* **2011,** *158*, 1444-1455.

82. Karwat, D. M. A.; Wagnon, S. W.; Teini, P. D.; Wooldridge, M. S., On th Chemical Kinetics of *n*-Butanol: Ignition and Specation Studies. *The Journal of Physical Chemistry A* **2011,** *115*, 4909-4921.

83. Karwat, D. M. A.; Wooldridge, M. S.; Klippenstein, S. J.; Davis, M. J., Effects of New Ab Initio Rate Coefficients on Predictions of Species Formed during *n*-Butanol Ignition and Pyrolysis. *The Journal of Physical Chemistry* **2015,** (119), 543-551.

84. Zhang, P.; Klippenstein, S. J.; Law, C. K., Ab Initio Kinetics for the Decomposition of Hydroxybutyl andButoxy Radicals of *n*-Butanol. *The Journal of Physical Chemistry A* **2013,** *117*, 1890-1906.

85. Moc, J.; Simmie, J. M.; Curran, H. J., The elimination of water from a conformationally complex alcohol: A computational study of the gas phase dehydration of n-butanol. *Journal of Molecular Structure* **2009,** *928*, 149-157.

86. Moc, J.; Simmie, J. M., Hydrogen abstraction from n-butanol by the hydroxyl radical: High level Ab initio study of the relative significance of various abstraction channels and the role of weakly bound intermediates. *Journal of Physical Chemistry A* **2010,** *114*, 5558-5564.

87. Black, G.; Simmie, J. M., Barrier heights for H-atom abstraction by HO2 from n -butanol-A simple yet exacting test for model chemistries? *Journal of Computational Chemistry* **2010,** *31*, 1236-1248.

88. Katsikadakos, D.; Hardalupas, Y.; Taylor, A. K. M. P.; Hunt, P. A., Hydrogen abstraction from n-butanol by the methyl radical: high level ab initio study of abstraction pathways and the importance of low energy rotational conformers. *Physical chemistry chemical physics : PCCP* **2012,** *14*, 9615-29.

89. Zhou, C.-W.; Simmie, J. M.; Curran, H. J., Rate constants for hydrogen-abstraction by O˙H from n-butanol. *Combustion and Flame* **2011,** *158*, 726-731.

90. Zhou, C.-W.; Simmie, J. M.; Curran, H. J., Rate Constants for Hydrogen Abstraction by HO2 from n-Butanol. *International journal of chemical Kinetics* **2012,** *44*, 155-164.

91. Zhang, P.; Klippenstein, S. J.; Law, C. K., Ab Initio Kinetics for the Decomposition of Hydroxybutyl and Butoxy Radicals of *n*-Butanol. *The Journal of Physical Chemistry A* **2013,** *117*, 1890-1906.

92. Barnard, J. A., THE PYROLYSIS OF n-BUTANOL. *Transactions of the Faraday Society* **1957,** *53*, 1423-1430.

93. Harding, M. E.; Metzroth, T.; Gauss, J.; Auer, A. a., Parallel Calculation of CCSD and CCSD(T) Analytic First and Second Derivatives. *Journal of Chemical Theory and Computation* **2008,** *4*, 64-74.

94. Kállay, M.; Gauss, J., Approximate treatment of higher excitations in coupled-cluster theory. *The Journal of chemical physics* **2005,** *123*, 214105.

95. Kállay, M.; Y; Gauss, J., Approximate treatment of higher excitations in coupled-cluster theory. II. Extension to general single-determinant reference functions and improved approaches for the canonical Hartree-Fock case. *The Journal of Chemical Physics* **2008,** *129*, 144101.

96. M. Kállay, Z. R., J. Csontos, P. Nagy, G. Samu, D. Mester, J. Csóka, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, and B. Hégely *MRCC, a quantum chemical program suite written by*, 2013.

97. Z. Rolik, L. S., I. Ladjánszki, B. Ladóczki, and M. Kállay, *J. Chem. Phys.* **2013,** *139*, 094105.

98. Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuś, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio, R. A.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J. W.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A.; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T.

R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M., Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Molecular Physics* **2015,** *113*, 184-215.

99. Almlöf, J.; Taylor, P. R., General contraction of Gaussian basis sets. I. Atomic natural orbitals for first- and second-row atoms. *J. Chem. Phys.* **1987,** *86*, 4070-4077.

100. Szalay, P. G.; Gauss, J.; Stanton, J. F., Analytic UHF-CCSD(T) Second Derivatives: Implementation and Application to the Calculation of the Vibration-Rotation Interaction Constants of NCO and NCS. *Theor. Chem. Acc.* **1998,** *100*, 5-11.

101. Dunning Jr, T. H., Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989,** *90*, 1007.

102. Ohno, K.; Yoshida, H.; Watanabe, H.; Fujita, T.; Matsuura, H., Conformational Study of 1-Butanol by the Combined Use of Vibrational Spectroscopy and ab Initio Molecular Orbital Calculations. *The Journal of Physical Chemistry* **1994,** *98*, 6924-6930.

103. Katritzky, A. R.; Lagowski, J. M., Prototropic Tautomerism of Heteroaromatic Compounds: I. General Discussion and Methods of Study**The last two chapters in this volume are of a series of four, the remaining two to appear in the next volume. Cross references to these articles include, for easy identification, the roman numeral given in the title of chapters. In *Advances in Heterocyclic Chemistry*, Katritzky, A. R., Ed. Academic Press: 1963; Vol. 1, pp 311-338.

104. Balabin, R. M., Tautomeric equilibrium and hydrogen shifts in tetrazole and triazoles: Focal-point analysis and ab initio limit. *The Journal of Chemical Physics* **2009,** *131* (15), 154307.

105. Cruz-Cabeza, A. J.; Groom, C. R., Identification, classification and relative stability of tautomers in the cambridge structural database. *CrystEngComm* **2011,** *13* (1), 93-98.

106. Fink, T.; Reymond, J.-L., Virtual Exploration of the Chemical Universe up to 11 Atoms of C, N, O, F: Assembly of 26.4 Million Structures (110.9 Million Stereoisomers) and Analysis for New Ring Systems, Stereochemistry, Physicochemical Properties, Compound Classes, and Drug Discovery. *Journal of Chemical Information and Modeling* **2007,** *47* (2), 342-353.

107. Tobias, F.; Heinz, B.; Jean-Louis, R., Virtual Exploration of the Small-Molecule Chemical Universe below 160 Daltons. *Angewandte Chemie International Edition* **2005,** *44* (10), 1504-1508.

108. Becke, A. D., Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993,** *98* (7), 5648-5652.

109. Weigend, F.; Ahlrichs, R., Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005,** *7*, 3297-3305.

110. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G., Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* **2009,** *113* (18), 6378-96.

111. Wheeler, S. E.; Houk, K. N., Substituent Effects in the Benzene Dimer are Due to Direct Interactions of the Substituents with the Unsubstituted Benzene. *Journal of the American Chemical Society* **2008,** *130* (33), 10854-10855.

112. Wheeler, S. E., Local Nature of Substituent Effects in Stacking Interactions. *Journal of the American Chemical Society* **2011,** *133* (26), 10262-10274.

113. Wheeler, S. E.; Bloom, J. W. G., Toward a More Complete Understanding of Noncovalent Interactions Involving Aromatic Rings. *The Journal of Physical Chemistry A* **2014,** *118* (32), 6133-6147.

114. Seifert, N. A.; Hazrah, A. S.; Jäger, W., The 1-Naphthol Dimer and Its Surprising Preference for π–π Stacking over Hydrogen Bonding. *The Journal of Physical Chemistry Letters* **2019,** *10* (11), 2836-2841.

115. Sinnokrot, M. O.; Sherrill, C. D., Substituent Effects in π−π Interactions: Sandwich and T-Shaped Configurations. *Journal of the American Chemical Society* **2004,** *126* (24), 7690-7697.

116. Cockroft, S. L.; Perkins, J.; Zonta, C.; Adams, H.; Spey, S. E.; Low, C. M. R.; Vinter, J. G.; Lawson, K. R.; Urch, C. J.; Hunter, C. A., Substituent effects on aromatic stacking interactions. *Organic & Biomolecular Chemistry* **2007,** *5* (7), 1062-1080.

117. Cockroft, S. L.; Hunter, C. A., Desolvation and substituent effects in edge-to-face aromatic interactions. *Chemical Communications* **2009,** (26), 3961-3963.

118. Ringer, A. L.; Sherrill, C. D., Substituent Effects in Sandwich Configurations of Multiply Substituted Benzene Dimers Are Not Solely Governed By Electrostatic Control. *Journal of the American Chemical Society* **2009,** *131* (13), 4574-4575.

119. Watt, M.; Hardebeck, L. K. E.; Kirkpatrick, C. C.; Lewis, M., Face-to-Face Arene−Arene Binding Energies: Dominated by Dispersion but Predicted by Electrostatic and Dispersion/Polarizability Substituent Constants. *Journal of the American Chemical Society* **2011,** *133* (11), 3854-3862.

120. Muchowska, K. B.; Adam, C.; Mati, I. K.; Cockroft, S. L., Electrostatic Modulation of Aromatic Rings via Explicit Solvation of Substituents. *Journal of the American Chemical Society* **2013,** *135* (27), 9976-9979.

121. Hwang, J.; Li, P.; Carroll, W. R.; Smith, M. D.; Pellechia, P. J.; Shimizu, K. D., Additivity of Substituent Effects in Aromatic Stacking Interactions. *Journal of the American Chemical Society* **2014,** *136* (40), 14060-14067.

122. Mignon, P.; Loverix, S.; De Proft, F.; Geerlings, P., Influence of Stacking on Hydrogen Bonding: Quantum Chemical Study on Pyridine−Benzene Model Complexes. *The Journal of Physical Chemistry A* **2004,** *108* (28), 6038-6044.

123. Piacenza, M.; Grimme, S., Van der Waals Interactions in Aromatic Systems: Structure and Energetics of Dimers and Trimers of Pyridine. *ChemPhysChem* **2005,** *6* (8), 1554-1558.

124. Swart, M.; van der Wijst, T.; Fonseca Guerra, C.; Bickelhaupt, F. M., π-π stacking tackled with density functional theory. *Journal of Molecular Modeling* **2007,** *13* (12), 1245-1257.

125. Hohenstein, E. G.; Sherrill, C. D., Effects of Heteroatoms on Aromatic π−π Interactions: Benzene−Pyridine and Pyridine Dimer. *The Journal of Physical Chemistry A* **2009,** *113* (5), 878-886.

126. Geng, Y.; Takatani, T.; Hohenstein, E. G.; Sherrill, C. D., Accurately Characterizing the π−π Interaction Energies of Indole−Benzene Complexes. *The Journal of Physical Chemistry A* **2010,** *114* (10), 3576-3582.

127. Karthikeyan, S.; Nagase, S., Origins of the Stability of Imidazole–Imidazole, Benzene–Imidazole, and Benzene–Indole Dimers: CCSD(T)/CBS and SAPT Calculations. *The Journal of Physical Chemistry A* **2012,** *116* (7), 1694-1700.

128. Jacobs, M.; Greff Da Silveira, L.; Prampolini, G.; Livotto, P. R.; Cacelli, I., Interaction Energy Landscapes of Aromatic Heterocycles through a Reliable yet Affordable Computational Approach. *Journal of Chemical Theory and Computation* **2018,** *14* (2), 543-556.

129. Gung, B. W.; Wekesa, F.; Barnes, C. L., Stacking Interactions between Nitrogen-Containing Six-Membered Heterocyclic Aromatic Rings and Substituted Benzene: Studies in Solution and in the Solid State. *The Journal of Organic Chemistry* **2008,** *73* (5), 1803-1808.

130. Itahara, T.; Imaizumi, K., Role of Nitrogen Atom in Aromatic Stacking. *The Journal of Physical Chemistry B* **2007,** *111* (8), 2025-2032.

131. Li, P.; Zhao, C.; Smith, M. D.; Shimizu, K. D., Comprehensive Experimental Study of N-Heterocyclic π-Stacking Interactions of Neutral and Cationic Pyridines. *The Journal of Organic Chemistry* **2013,** *78* (11), 5303-5313.

132. Bootsma, A. N.; Wheeler, S. E., Tuning Stacking Interactions between Asp-Arg Salt Bridges and Heterocyclic Drug Fragments. *J. Chem. Inf. Model.* **2019,** *59* (1), 149-158.

133. Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H.; Izmaylov, A.; Bloino, J.; Zheng, G.; Sonnenberg, J.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., JA; Peralta, J.; Ogliaro, F.; Bearpark, M.; Heyd, J.; Brothers, E.; Kudin, K.; Staroverov, V.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J.; Iyengar, S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J.; Klene, M.; Knox, J.; Cross, J.; Bakken, V.; Adamo, C.; ramillo, J.; Gomperts, R.; Stratmann, R.; Yazyev, O.; Austin, A.; Cammi, R.; Pomelli, C.; Ochterski, J.; Martin, R.; Morokuma, K.; Zakrzewski, V.; Voth, G.; Salvador, P.; Dannenberg, J.; Dapprich, S.; Daniels, A.; Farkas, O.; Foresman, J.; Ortiz, J.; Cioslowski, J.; Fox, D. *Gaussian 09, Revision D.01,* Gaussian, Inc.: 2009.

134. Harding, D. P., Wheeler, S. E. , *In Preparation* **2019**.

APPENDIX A.

SUPPLEMENTARY INFORMATION RELATED TO ANNULAR TAUTOMERIZATIONS OF HETEROAROMATIC RINGS

Table A.1 Heterocycles from the Pitt Set and Corresponding VEHICLe Identifiers. Note that clusters **5** and **6** do not have associated VEHICLe identifiers, as they did not meet the synthetic viability criteria.

Pitt	VEHICLe	Pitt	VEHICLe	Pitt	VEHICLe	Pitt	VEHICLe
1a	S10	12b	S1866	22c	S1196	29b	S15601
1b	S11	13a	S2740	22d	S14413	30a	S5904
2a	S23	13b	S7955	23a	S2055	30b	S15623
2b	S24	14a	S2787	23b	S2054	30c	S15622
3a	S29	14b	S2786	23c	S2040	31a	S5714
3b	S30	14c	S9695	23d	S2039	31b	S14410
4a	S75	15a	S5686	24a	S15612	32a	S5864
4b	S71	15b	S5685	24b	S1205	32b	S14426
5a	N/A	15c	S5830	24c	S14412	32c	S14428
5b	N/A	16a	S828	24d	S1195	33a	S5740
6a	N/A	16b	S830	25a	S2875	33b	S20192
6b	N/A	17a	S7998	25b	S15630	34a	S1948
7a	S1071	17b	S7995	25c	S2864	34b	S1949
7b	S7952	18a	S15404	25d	S2874	34c	S1954
8a	S1075	18b	S14595	25e	S14442	34d	S1955
8b	S8043	19a	S8007	25f	S2865	35a	S92
9a	S1118	19b	S8019	26a	S2001	35b	S105
9b	S9690	20a	S937	26b	S2000	36a	S83
10a	S1867	20b	S932	27a	S1093	36b	S72
10b	S1868	21a	S938	27b	S1084	37a	S1789
11a	S1929	21b	S933	28a	S1089	37b	S5674
11b	S1928	22a	S1206	28b	S8626	38a	S1349
12a	S1865	22b	S15613	29a	S1193	38b	S1348

Table A.2. Focal point analysis for benchmark tautomerizations, in kcal mol^{-1}.

	HF	MP2	CCSD	CCSD(T)
16b → 16a				
cc-pVDZ	-0.60	0.70	0.40	0.65
cc-pVTZ	-0.69	0.55	0.21	0.47
cc-pVQZ	-0.70	0.51	0.17	0.43
cc-pV5Z	-0.71	0.48	[+0.16]	[+0.41]
CBS LIMIT	[-0.71]	[+0.46]	[+0.15]	**[+0.40]**
3b→3a				
cc-pVDZ	2.12	4.41	2.87	3.02
cc-pVTZ	1.40	3.89	2.28	2.45
cc-pVQZ	1.28	3.72	2.13	2.29
cc-pV5Z	1.23	3.61	[+2.08]	[+2.23]
CBS LIMIT	[+1.21]	[+3.52]	[+2.05]	**[+2.20]**
38a→38b				
cc-pVDZ	0.85	3.27	1.99	2.35
cc-pVTZ	1.12	3.50	2.18	2.57
cc-pVQZ	1.18	3.54	2.22	2.61
cc-pV5Z	1.21	3.55	[+2.24]	[+2.63]
CBS LIMIT	[+1.23]	[+3.54]	[+2.25]	**[+2.63]**
1b→1a				
cc-pVDZ	5.52	5.42	5.02	4.82
cc-pVTZ	4.81	4.98	4.56	4.35
cc-pVQZ	4.66	4.82	4.42	4.20
cc-pV5Z	4.59	4.71	[+4.35]	[+4.12]
CBS LIMIT	[+4.56]	[+4.64]	[+4.32]	**[+4.09]**
37a→37b				
cc-pVDZ	5.57	4.78	4.76	4.70
cc-pVTZ	5.15	4.27	4.39	4.28
cc-pVQZ	5.07	4.13	4.31	4.18
cc-pV5Z	5.03	4.03	[+4.27]	[+4.13]
CBS LIMIT	[+5.01]	[+3.95]	[+4.25]	**[+4.10]**
15c→15a				
cc-pVDZ	10.98	8.81	9.03	8.89
cc-pVTZ	10.75	8.55	8.99	8.75
cc-pVQZ	10.73	8.54	9.05	8.79
cc-pV5Z	10.74	8.54	[+9.08]	[+8.81]
CBS LIMIT	[+10.74]	[+8.53]	[+9.12]	**[+8.84]**

		36a→36b		
cc-pVDZ	11.48	9.21	9.04	9.18
cc-pVTZ	11.22	8.96	8.96	8.98
cc-pVQZ	11.07	8.84	8.89	8.90
cc-pV5Z	11.01	8.77	[+8.86]	[+8.87]
CBS LIMIT	[+10.98]	[+8.74]	[+8.86]	**[+8.86]**
		4a→4b		
cc-pVDZ	10.95	10.46	10.09	10.22
cc-pVTZ	10.21	9.71	9.42	9.51
cc-pVQZ	9.93	9.43	9.18	9.26
cc-pV5Z	9.83	9.30	[+9.09]	[+9.17]
CBS LIMIT	[+9.79]	[+9.24]	[+9.06]	**[+9.14]**
		35a→35b		
cc-pVDZ	11.21	11.35	10.72	10.81
cc-pVTZ	10.40	10.72	10.07	10.13
cc-pVQZ	10.12	10.46	9.83	9.89
cc-pV5Z	10.03	10.35	[+9.76]	[+9.82]
CBS LIMIT	[+9.99]	[+10.28]	[+9.74]	**[+9.80]**
		2b→2a		
cc-pVDZ	7.28	7.25	6.71	6.68
cc-pVTZ	6.88	6.82	6.39	6.32
cc-pVQZ	6.88	6.78	6.39	6.30
cc-pV5Z	6.89	6.73	[+6.40]	[+6.30]
CBS LIMIT	[+6.89]	[+6.68]	[+6.41]	**[+6.30]**

APPENDIX B.

SUPPLEMENTARY INFORMATION RELATED TO ANNULAR TAUTOMERISM MEDIATED BY STACKING INTERACTIONS

Table B.1. Comparison of relative tautomer energies (E_{rel}) computed in water with SMD-M06-2X/def2-TZVP solvent-corrected (in diethylether) DLPNO-CCSD(T)/cc-pVQZ stacking interaction energies (ΔE_{stack}^{soln}). The tautomerization energy in the stacked state ($E_{rel}^{stacked}$) is the sum of E_{rel} and ΔE_{stack}^{soln}.

				Phe		Tyr		Trp	
Transition			E_{rel}	ΔE_{stack}^{soln}	$E_{rel}^{stacked}$	ΔE_{stack}^{soln}	$E_{rel}^{stacked}$	ΔE_{stack}^{soln}	$E_{rel}^{stacked}$
1a	→	**1b**	0.3	-0.4	-0.1	0.7	1.1	-0.4	-0.1
2a	→	**2b**	0.1	-0.2	-0.1	-0.1	0.1	0.2	0.3
3a	→	**3b**	0.2	-0.2	0.1	0.2	0.4	0.0	0.2
4a	→	**4b**	0.2	-0.1	0.1	0.0	0.2	-0.1	0.0
5a	→	**5b**	0.3	-0.7	-0.3	-0.7	-0.3	-0.9	-0.6
6a	→	**6b**	0.0	-0.1	-0.1	-0.4	-0.4	-0.1	0.0
7a	→	**7b**	0.5	-0.5	-0.1	-0.3	0.2	-1.5	-1.1
8a	→	**8b**	0.1	0.6	0.6	0.6	0.7	0.6	0.7
9a	→	**9b**	0.4	0.5	0.9	0.0	0.4	-0.4	0.0
10a	→	**10b**	0.5	-0.2	0.3	0.3	0.8	0.1	0.6
11a	→	**11b**	0.0	0.7	0.7	0.5	0.5	-0.1	-0.1
12a	→	**12b**	0.4	0.2	0.6	-0.1	0.3	-0.4	0.0
13a	→	**13b**	0.2	0.2	0.4	-0.1	0.1	0.1	0.2
14a	→	**14b**	0.2	0.2	0.4	-0.1	0.1	-0.4	-0.2

Table B.2 Stacking interaction energies for each monomer with the model phenylalanine side chain. Predicted (E_{stack}^{pred}) and DLPNO-CCSD(T)/cc-pVQZ interaction energies (E_{stack}^{gas}) are compared graphically in Figure 4.1 and the gas-phase ωB97X-D/def2-TZVP ($E_{stack}^{\omega B97}$) and solution-phase (in diethylether) SMD- ωB97X-D/def2-TZVP (E_{stack}^{SMD}) can be used to compute the solvent corrections for (ΔE_{stack}^{soln}) given in Table 4.1.

Monomer	E_{stack}^{pred}	E_{stack}^{gas}	$E_{stack}^{\omega B97}$	E_{stack}^{SMD}
1a	-7.1	-7.4	-7.2	-2.6
1b	-7.9	-8.0	-8.0	-3.3
2a	-6.7	-6.7	-6.3	-4.0
2b	-6.9	-7.7	-7.4	-4.4
3a	-6.9	-6.8	-6.3	-4.5
3b	-7.5	-7.4	-7.1	-4.8
4a	-7.0	-7.8	-7.9	-4.3
4b	-7.2	-7.7	-7.8	-4.3
5a	-8.3	-8.1	-8.0	-4.4
5b	-8.8	-9.0	-8.7	-4.8
6a	-8.1	-7.7	-7.3	-5.3
6b	-8.7	-8.6	-8.2	-5.4
7a	-9.6	-8.8	-8.2	-6.1
7b	-10.3	-9.7	-9.2	-6.7
8a	-9.2	-9.0	-8.5	-4.8
8b	-9.4	-8.6	-8.2	-4.4
9a	-9.4	-9.2	-8.8	-4.5
9b	-10.1	-9.2	-8.8	-4.1
10a	-9.2	-8.4	-8.0	-5.3
10b	-9.8	-9.2	-9.3	-6.0
11a	-9.3	-9.1	-8.5	-6.6
11b	-9.7	-8.5	-8.0	-6.1
12a	-9.0	-8.5	-8.1	-6.0
12b	-9.4	-8.6	-8.3	-5.9
13a	-9.5	-8.8	-8.1	-5.3
13b	-10.1	-9.1	-8.4	-5.2
14a	-9.1	-8.8	-8.2	-5.7
14b	-9.9	-9.2	-9.0	-5.8

Table B.3 Stacking interaction energies for each monomer with the model tryptophan side chain. Predicted (E_{stack}^{pred}) and DLPNO-CCSD(T)/cc-pVQZ interaction energies (E_{stack}^{gas}) are compared graphically in Figure 4.1 and the gas-phase ωB97X-D/def2-TZVP ($E_{stack}^{\omega B97}$) and solution-phase (in diethylether) SMD- ωB97X-D/def2-TZVP (E_{stack}^{SMD}) can be used to compute the solvent corrections for (ΔE_{stack}^{soln}) given in Table 4.1.

Monomer	E_{stack}^{pred}	E_{stack}^{gas}	$E_{stack}^{\omega B97}$	E_{stack}^{SMD}
1a	-7.9	-10.0	-9.6	-4.0
1b	-8.9	-10.9	-10.6	-4.5
2a	-7.5	-9.8	-9.1	-5.7
2b	-7.7	-10.6	-10.1	-5.7
3a	-7.7	-9.3	-8.3	-5.7
3b	-8.3	-10.3	-9.7	-6.1
4a	-7.8	-10.5	-10.1	-5.1
4b	-8.1	-10.1	-9.9	-5.5
5a	-9.3	-11.6	-11.2	-5.7
5b	-9.9	-13.1	-12.7	-6.6
6a	-9.0	-11.0	-10.3	-6.9
6b	-9.8	-11.8	-11.4	-7.3
7a	-10.8	-12.0	-11.4	-8.1
7b	-11.6	-14.3	-13.6	-9.5
8a	-10.3	-13.1	-12.7	-6.8
8b	-10.5	-12.8	-12.5	-6.2
9a	-10.5	-11.7	-11.2	-6.1
9b	-11.4	-13.9	-13.7	-6.8
10a	-10.4	-12.9	-12.5	-8.0
10b	-11.0	-13.0	-12.8	-8.2
11a	-10.4	-11.9	-11.1	-8.4
11b	-10.9	-12.9	-12.1	-8.5
12a	-10.1	-11.0	-10.3	-7.1
12b	-10.6	-12.3	-11.9	-7.9
13a	-10.7	-12.8	-12.2	-7.3
13b	-11.3	-12.7	-11.9	-7.0
14a	-10.3	-12.9	-12.3	-7.6
14b	-11.1	-13.9	-13.5	-8.1

Table B.4 Stacking interaction energies for each monomer with the model tyrosine side chain. Predicted (E_{stack}^{pred}) and DLPNO-CCSD(T)/cc-pVQZ interaction energies (E_{stack}^{gas}) are compared graphically in Figure 4.1 and the gas-phase ωB97X-D/def2-TZVP ($E_{stack}^{\omega B97}$) and solution-phase (in diethylether) SMD- ωB97X-D/def2-TZVP (E_{stack}^{SMD}) can be used to compute the solvent corrections for (ΔE_{stack}^{soln}) given in Table 4.1.

Monomer	E_{stack}^{pred}	E_{stack}^{gas}	$E_{stack}^{\omega B97}$	E_{stack}^{SMD}
1a	-9.6	-8.3	-8.0	-3.6
1b	-10.8	-9.0	-9.0	-3.1
2a	-9.0	-8.0	-7.4	-4.8
2b	-9.3	-9.4	-9.2	-5.2
3a	-9.4	-7.9	-6.9	-4.7
3b	-10.1	-8.4	-8.1	-5.2
4a	-9.4	-8.2	-8.1	-4.3
4b	-9.8	-8.3	-8.2	-4.4
5a	-11.3	-8.6	-8.4	-4.7
5b	-12.1	-9.7	-9.3	-5.2
6a	-11.0	-8.5	-8.0	-5.6
6b	-12.0	-9.4	-8.9	-6.0
7a	-13.2	-10.1	-9.5	-7.0
7b	-14.2	-10.8	-10.3	-7.3
8a	-12.6	-10.6	-10.0	-5.5
8b	-12.9	-10.1	-9.7	-5.1
9a	-12.9	-10.6	-10.1	-4.7
9b	-13.9	-10.3	-10.2	-5.1
10a	-12.7	-10.2	-9.7	-6.2
10b	-13.5	-10.2	-9.9	-6.2
11a	-12.7	-10.3	-9.7	-7.1
11b	-13.3	-10.1	-9.4	-6.5
12a	-12.3	-9.6	-9.3	-6.6
12b	-13.0	-9.7	-9.2	-6.5
13a	-13.0	-10.4	-9.7	-6.3
13b	-13.9	-10.9	-10.3	-6.5
14a	-12.5	-10.3	-9.8	-6.1
14b	-13.6	-10.3	-9.9	-6.3