

COMPUTATIONAL STUDY OF THE REACTIONS OF  
PHOSPHINYLIDENES, FLUOROQUINOLONES AND  
HYDRAZONES - MOLECULES OF COMMERCIAL AND  
PHARMACEUTICAL INTEREST

by

MARK JOSEPH BRIDLE

Bachelor of Science, 2004  
Loughborough University  
Loughborough, Leicestershire, UK  
LE11 3TU

Submitted to the Graduate Faculty of the  
College of Science and Engineering  
Texas Christian University  
in partial fulfillment of the requirements  
for the degree of

Doctor of Philosophy

August, 2017



## Acknowledgement

The pressure is perhaps felt the greatest while writing this part of my dissertation to acknowledge all those that have helped me and supported me while I have been pursuing a doctoral degree. I would first like to express my deepest thanks and gratitude to my advisor, Dr. Benjamin Janesko, who has shown great patience and offered invaluable guidance during my time at TCU. Working with me and being flexible with my time balancing as a part-time student, while maintaining full-time employment. I am certain that I was a headache at times but I am glad he stuck it out with me.

I am also indebted to my committee members, Drs. Jean-Luc Montchamp, Kayla Green, and Youngha Ryu, for their support and advice during my time in their classes and while carrying out my research. I particularly enjoyed the discussions surrounding my research proposal.

I would also like to convey my thanks and appreciation to those that I have done collaborative research with. Drs. Henry Fisher and Jean-Luc Montchamp for their work regarding phosphinylidene tautomerization, and Vishal Sharma and Dr. Erik Simanek for their work on the hydrazone hydrolysis project.

I want to acknowledge Drs. Bimal Pudasaini, Richard Sniatynsky, John Determan, Andrew Mahler and Emil Proynov, as well as Aimee Garcia, Stephanie Jones and Arshad Mehmood, members of the Janesko Group that I have been able to get to know and work with

over the past six years. Working nights and keeping unsociable hours is made easier by those you spend it with.

I want to thank my family for their support for being available for early morning international phone calls. We may live in different countries but starting the day off before dawn is made a little easier when you can use the drive to catch up. I also want to acknowledge my children, Lilly, Caden, Caitlin, Evelyn and Harvey. They have made sacrifices and put in their time over the last few years and I look forward to being able to spending more time with them and being more involved in all their sports.

My biggest thanks and appreciation goes to my wife, Sarah. Coming to the end of this journey, I don't think we had any idea what we were getting ourselves into. We have been through a lot, much of it not related to graduate school. We bought a house, welcomed two new additions to the family, and heard the clinical diagnosis of Holoprosencephaly that forever changed our lives. We could have given up and been divorced multiple times over. I am so thankful that she stuck by me and supported me through every pit of despair and every triumph. I couldn't have gotten to this point on my own and I can't wait to spend evenings with her and give her back the time, thanks and attention that she most definitely deserves. Thank you, from the bottom of my heart. This is as much your doing as mine. I love you.

Finally, I want to thank the TCU Department of Chemistry and Biochemistry for their financial support.

## Author's Declaration

This is an original work, except where references have been made. No part of this work has been previously submitted as part of a requirement for an academic degree.

Part of this dissertation is based on the following:

- 1) Janesko, B.G., Fisher, H.C., Bridle, M.J., Montchamp, J-L. *P(=O)H to P-OH Tautomerism: A Theoretical and Experimental Study*. Published in *J. Org. Chem.* **2015**, 80(20), 10025-10032
- 2) Bridle, M.J., Fisher, H.C., Montchamp, J-L., Janesko, B.G. *Catalysis of Phosphinylidene Tautomerization by Small Organic Molecules: A Theoretical and Experimental Study*. In preparation for submission to *J. Org. Chem.*
- 3) Sharma, V.R., Bridle, M.J., Janesko, B.G., Simanek, E.E. *Hydrolysis Rates of Triazinyl Hydrazones with Different Groups at N2 at Various pHs*. In preparation for submission to *Tetrahedron*
- 4) Bridle, M.J., Janesko, B.G. *Computational Study of Fluoroquinolone Binding to Mg(H<sub>2</sub>O)<sub>n</sub><sup>2+</sup> and its Applicability to Future Drug Design*. Accepted to *Int. J. Quantum Chem.* June **2017**

# Contents

<b>Acknowledgement .....</b>	<b>ii</b>
<b>Author's Declaration .....</b>	<b>iv</b>
<b>Contents .....</b>	<b>v</b>
<b>List of Figures.....</b>	<b>ix</b>
<b>List of Schemes.....</b>	<b>xv</b>
<b>List of Tables .....</b>	<b>xvii</b>
<b>List of Equations .....</b>	<b>xx</b>
<b>List of Symbols and Abbreviations .....</b>	<b>xxiii</b>
<b>Chapter 1: Introduction .....</b>	<b>1</b>
1.1    General Applications of Computational Chemistry .....	1
1.2    Reducing the Cost of Pharmaceutical Development.....	3
1.3    Pharmaceutical Application of Screening and Computational Chemistry.....	5
1.4    Goal .....	9
<b>Chapter 2: Phosphinylidene Tautomerization .....</b>	<b>11</b>
2.1    Phosphorus Chemistry.....	11
2.2    Commercial Applications of Organophosphorus Compounds.....	13
2.3    Phosphinylidene Chemistry.....	14
2.4    Why the Interest in Phosphinylidenes for Phosphorus-Carbon Bonds? .....	17
2.5    Tautomerization Rate .....	23

2.6	Effect of Substituents on Phosphinylidene Tautomerization Rate .....	23
2.7	Computational Modelling of Phosphinylidene Tautomerization .....	28
2.8	Concluding Remarks .....	36
2.9	Computational Section .....	37
<b>Chapter 3: Catalysis of Phosphinylidene Tautomerization .....</b>		<b>38</b>
3.1	Industrial Catalysis .....	38
3.2	Catalysis of Phosphinylidene Tautomerization.....	40
3.3	Catalyzed Phosphinylidene Tautomerization Rate .....	43
3.4	Calculating the Catalyzed Tautomerization Barriers .....	48
3.5	Concluding Remarks .....	53
3.6	Computational Section .....	54
<b>Chapter 4: Fluoroquinolone Binding to Mg(H<sub>2</sub>O)<sub>n</sub><sup>2+</sup> and its Applicability to Future Drug Design.....</b>		<b>56</b>
4.1	Fluoroquinolone Chemistry and the Significance of Mg <sup>2+</sup> Binding .....	56
4.2	Fluoroquinolone Geometry Benchmarks .....	59
4.3	Mg <sup>2+</sup> Complex Geometry Benchmarks .....	66
4.4	Fluoroquinolone-Mg <sup>2+</sup> Binding Affinity Benchmarks.....	69
4.5	Dependence of Fluoroquinolone-Mg <sup>2+</sup> Binding on Protonation State.....	74
4.6	Effect of Carboxylic Acid Modification on Mg <sup>2+</sup> Binding .....	77
4.7	Concluding Remarks .....	79

4.8 Computational Section .....	81
<b>Chapter 5: Effect of N<sub>2</sub> Substituent on Hydrolysis of Hydrazones.....</b>	<b>87</b>
5.1 Hydrazine Chemistry.....	87
5.3 Hydrolysis of Hydrazones.....	89
5.4 Effect of N <sub>2</sub> Substituents on Hydrazone Stability to Hydrolysis.....	95
5.5 Computational Study of N <sub>1</sub> Proton Affinities.....	100
5.6 Concluding Remarks .....	111
5.7 Computational Section .....	112
<b>APPENDIX A .....</b>	<b>113</b>
Electronic Structure Theory .....	113
Hartree-Fock Theory.....	118
Semi-Empirical Methods .....	119
Density Functional Theory (DFT) .....	120
Basis Sets .....	126
<b>APPENDIX B .....</b>	<b>132</b>
From Electronic Structure to Thermochemistry .....	132
Calculation of Thermochemical Data .....	133
<b>APPENDIX C .....</b>	<b>138</b>
Computational Data .....	138
<b>References .....</b>	<b>139</b>

## **VITA**

### **ABSTRACT**

## List of Figures

Figure 1.1: Fischer-Tropsch process.....	2
Figure 1.2: Costs and time associated with bringing a new drug product to market.....	3
Figure 1.3: The “quick win, fast fail” drug development paradigm. ....	5
Figure 1.4: Generation of SMILES input. ....	7
Figure 2.1: A) Kinase inhibitor. B) Structure of farnesyl diphosphate (FPP). C) FPP analogs that are known inhibitors of SQS. D) Chemical structures of bisphosphonates of clinical relevance for the treatment of osteoporosis. ....	12
Figure 2.2: <i>H</i> -phosphinate detergent.....	13
Figure 2.3: Melamine polyphosphate (MPP) flame retardant. ....	13
Figure 2.4: Hydraulic cement retarding polymer.....	14
Figure 2.5: AIBN decomposition to form 2-cyanoprop-2-yl radicals. ....	15
Figure 2.6: Structures of (a) <i>H</i> -phosphine oxides, (b) <i>H</i> -phosphonates and (c) <i>H</i> -phosphinates, where R=H.....	19
Figure 2.7: Calculated transition state geometries for H <sub>3</sub> PO tautomerization. From left to right, the figure shows uncatalyzed tautomerization, tautomerization in continuum water, and tautomerization catalyzed by one and two explicit H <sub>2</sub> O molecules. ....	22
Figure 2.8: B3LYP/6-311++G(3df,3pd) calculated geometries and potential energy surfaces for methyl <i>H</i> -phosphinate tautomerization, both uncatalyzed and catalyzed by a single explicit water molecule, in a water continuum solvent.....	22
Figure 2.9: Phosphinylidenes investigated in experimental tautomerization study.....	24
Figure 2.10: Decay of phosphinylidene compounds (1M in CH <sub>3</sub> CN) during deuteration with D <sub>2</sub> O (half-lives > 3 hours).....	25

Figure 2.11: Decay of phosphinylidene compounds (1M in CH <sub>3</sub> CN) during deuteration with D <sub>2</sub> O (half-lives ~1 hour or less).....	25
Figure 2.12: Structures of phosphinylidenes used in computational tautomerization study. .	30
Figure 2.13: B3LYP/6-311++G(3df,3pd) optimized geometries of P(V), TS and P(III) for EtO,EtO, Modeled as MeO,MeO.....	31
Figure 2.14: B3LYP/6-311++G(3df,3pd) optimized geometries of P(V), TS and P(III) for OH,Ph.....	31
Figure 2.15: B3LYP/6-311++G(3df,3pd) optimized geometries of P(V), TS and P(III) for CF <sub>3</sub> ,CF <sub>3</sub> .....	31
Figure 2.16: Calculated barriers for catalysis by a single explicit water molecule vs. experimental half-lives.....	32
Figure 2.17: Plot of ln t <sub>1/2</sub> vs. -ln (k). .....	34
Figure 3.1: General overview of lowering of activation energy by addition of a catalyst for A) exothermic or B) endothermic reaction. ....	39
Figure 3.2: Predicted energy surfaces for methyl methylphosphite tautomerization; uncatalyzed, catalyzed by a single water molecule, and catalyzed by representative hard acid-soft base molecule thioformic acid.....	41
Figure 3.3: Tautomerization catalysts considered experimentally and computationally.....	42
Figure 3.4: B3LYP/6-311++G(3df,3pd) optimized geometries of P(V) and P(III) for OMe,Me. ....	44
Figure 3.5: B3LYP/6-311++G(3df,3pd) optimized geometries of P(V) and P(III) for OMe,OMe. ....	44

Figure 3.6: Decay of <i>tert</i> -butyl octylphosphinate (1M in CH <sub>3</sub> CN) during deuteration with D <sub>2</sub> O catalyzed by various organic compounds, with decay of di <i>tert</i> -butyl phosphite shown with dashed lines. No additive was used in the control experiment. ....	45
Figure 3.7: Calculated catalyzed barriers vs. ln t <sub>1/2</sub> (experimental). ....	49
Figure 3.8: Calculated catalyzed barriers vs. ln t <sub>1/2</sub> (experimental) for <i>tert</i> -butyl octylphosphinate with a smaller subset of catalysts (Water, VL, SC, AA, BA, FA and L-Pro). ....	49
Figure 3.9: B3LYP/6-311+G(3df,3pd) optimized geometries of representative organo-catalyzed OMe,Me P(V) to P(III) tautomerization transition states. ....	51
Figure 3.10: Interaction of phosphinylidene with thiol. ....	53
Figure 3.11: B3LYP/6-311+G(3df,3pd) optimized geometry of thioformic acid catalyzed OMe,Me P(V) to P(III) tautomerization transition state. ....	53
Figure 4.1: A) Structure of nalidixic acid and B) essential fluoroquinolone structure. ....	56
Figure 4.2: Diagram of proposed water-metal ion bridge that mediates critical interactions between quinolones and topoisomerase IV. ....	57
Figure 4.3: Binding of ciprofloxacin zwitterion to hydrated magnesium ion, geometry optimized using B3LYP method with 6-311+G(2d,2p) basis set (atoms are color coded, with C gray, H white, Mg yellow, O red, N blue and F teal). ....	58
Figure 4.4: Fluoroquinolones used in benchmark set. ....	60
Figure 4.5: Quinolones used in benchmark set. ....	61
Figure 4.6: Mean absolute errors in bond lengths for B3LYP, semiempirical and DFT approximations using the 6-311+G(2d,2p) basis set. ....	65

Figure 4.7: Mean absolute errors in bond angles for B3LYP, semiempirical and DFT approximations using the 6-311+G(2d,2p) basis set. ....	66
Figure 4.8: Magnesium complexes used in benchmark set .....	67
Figure 4.9: Fluoroquinolones used for magnesium-binding benchmarking.....	69
Figure 4.10: $\Delta G_{\text{expt}}$ (kcal mol <sup>-1</sup> ) vs. $\Delta G_{\text{calc}}$ (kcal mol <sup>-1</sup> ) for experimental binding affinities reported by Sissi <i>et al.</i> .....	71
Figure 4.11: Fluoroquinolones with modifications to the carboxylic acid .....	73
Figure 4.12: B3LYP/6-311+G(2d,2p) optimized geometry of ciprofloxacin zwitterion binding to Mg(H <sub>2</sub> O) <sub>N</sub> <sup>2+</sup> . ....	75
Figure 4.13: B3LYP/6-311+G(2d,2p) optimized geometry of ciprofloxacin anion binding to Mg(H <sub>2</sub> O) <sub>N</sub> <sup>2+</sup> . ....	75
Figure 4.14: B3LYP/6-311+G(2d,2p) optimized geometry of ciprofloxacin cation binding to Mg(H <sub>2</sub> O) <sub>N</sub> <sup>2+</sup> . ....	76
Figure 4.15: Average fluoroquinolone-Mg <sup>2+</sup> binding for fluoroquinolones in Figure 4.9 modeled at various protonation states.....	77
Figure 4.16: Ciprofloxacin derivatives (1) 7-chloro-1-cyclopropyl-6- uoro-1,4-dihydro-4-oxo-quinoline-3-methylcarbamate and (2) 1-cyclopropyl-7-(4-ethyl-1-piperazinyl)-6- uoro-1,4-dihydro-5-oxopyrazolo)[4,3-c]quino-line.....	81
Figure 4.17: Modelling of ciprofloxacin zwitterion (a) with the addition of one (b), three (c) or five (d) explicit water molecules (atoms are color coded, with C gray, H white, N blue, O red and F teal). .....	84
Figure 5.1: B3LYP/6-31+G(d,p) optimized geometry of hydrazine (atoms are color coded with N blue and H white). ....	87

Figure 5.2: Examples of 7-hydrazino-1-ethyl-6,8-difluoroquinolone-3-carboxylate derivatives.	88
.....	.....
Figure 5.3: pH dependence of triazine and acetyl hydrazone stability to hydrolysis. ....	89
Figure 5.4: Example of triazine protonation. ....	90
Figure 5.5: Structure of hydrazones studied and the designation of nitrogen N1 and N2.....	90
Figure 5.6: Experimental triazinyl hydrazines with 1) CH <sub>3</sub> , 2) H and 3) phenyl substituents at N2, where R = 2-ethoxyethanol.....	91
Figure 5.7: Aldehyde and ketones (a-d) used for hydrolysis studies. ....	93
Figure 5.8: Experimental hydrolysis of triazinylhydrazones as measured by HPLC. ....	96
Figure 5.9: Relative experimental traizinylhydrazine stability to hydrolysis at pH 4.0 (t <sub>1/2</sub> shown in minutes). ....	97
Figure 5.10: Triazinyl and acetyl hydrazine models (1-6) used for computational protonation studies. ....	101
Figure 5.11: B3LYP/6-31+G(d,p) optimized geometries and relative energies of keto-enol conformations and pyrazole ring for reaction of triazinylhydrazine 2 with diketone b.....	101
Figure 5.12: B3LYP/6-31+G(d,p) optimized geometries and relative energies of keto-enol conformations and pyrazole ring for reaction of triazinylhydrazine 2 with diketone c .....	102
Figure 5.13: Structures of hydrazones investigated computationally and the N1 proton affinity (kcal/mol), with the proton affinity for the protonated triazine in brackets.....	104
Figure 5.14: Nitrogen in plane lone pair conjugation with aromatic ring. ....	105
Figure 5.15: B3LYP/6-31+G(d,p) optimized geometries of the triazinyl hydrazones formed by reaction with benzaldehyde.....	105

Figure 5.16: HOMO of the protonated phenyl complex for acylhydrazine reacted with benzaldehyde.....	106
Figure 5.17: Computed N1 proton affinities for hydrazines with aldehyde or ketone. ....	108
Figure 5.18: Computed N1 proton affinities for protonated traizines with aldehyde or ketone. .....	108
Figure 5.19: Experimental half-lives for hydrolysis of hydrazones at pH 7.0.....	109
Figure 5.20: Experimental half-lives for hydrolysis of hydrazones at pH 4.0.....	110
Figure A.1: Potential energy surface for H <sub>2</sub> .....	115

## List of Schemes

Scheme 1.1: General mechanism of Suzuki-Miyaura coupling.....	2
Scheme 2.1: Central features of the human isoprenoid biosynthetic pathway. ....	11
Scheme 2.2: Preparation of Fosinopril by radical initiation. ....	15
Scheme 2.3: One pot process for phosphonic acid synthesis (2) from hypophosphorous acid reaction with an alkene via a phosphinic acid intermediate (1).....	15
Scheme 2.4: Hydrophosphinylation of terminal alkynes. ....	16
Scheme 2.5: Possible cross-coupling and transfer hydrogenation pathways.....	16
Scheme 2.6: Generation of phosphinylidene P(III) reactive feedstocks for cross-coupling reactions. ....	17
Scheme 2.7: Industrial process for the manufacture of organophosphorus compounds.....	18
Scheme 2.8: Alternative organophosphorus compound manufacturing process using phosphinylidenes (shown in green box) as feedstock, while bypassing $\text{PCl}_3$ pathway (shown in red box). ....	18
Scheme 2.9: Prototropic tautomerization of phosphinylidenes. ....	19
Scheme 2.10: Deuteration of phosphinylidene compounds using an excess of $\text{D}_2\text{O}$ . ....	23
Scheme 3.1: General catalytic reaction for phosphinylidene tautomerization. ....	38
Scheme 3.2: Experimental probes of phosphinylidene tautomerizaton kinetics: Deuteration of tert-butyl octylphosphinate (Scheme 1a) or di-tert-butyl phosphite (Scheme 1b), in the presence of excess $\text{D}_2\text{O}$ and 5-10 mol% catalyst.....	43
Scheme 4.1: Binding model simulating single charge cation protonation state. ....	72
Scheme 4.2: Binding model simulating neutral charged zwitterion .....	72
Scheme 4.3: Binding model simulating single charge anion protonation state .....	72

Scheme 5.1: Proposed mechanism of acid catalyzed hydrolysis of hydrazones (R = acetyl or triazinyl).....	90
Scheme 5.2: Synthesis of Triazinylhydrazine 1 <sup>a</sup> .....	92
Scheme 5.3: Synthesis of Triazinylhydrazine 2 <sup>a</sup> .....	92
Scheme 5.4: Synthesis of Triazinylhydrazine 3 <sup>a</sup> .....	93
Scheme 5.5: Formation of hydrazone. ....	93
Scheme 5.6: Formation of a pyrazole ring via a second dehydration when N2 is hydrogen and reacted with a diketone. ....	94
Scheme 5.7: Keto enol tautomerism as observed for acetylacetone .....	98

## List of Tables

Table 1.1: Input data comparison for IDEA and GASTROPLUS.....	8
Table 2.1: B3LYP/6-311++G(3df,3pd) total energy barrier ( $\Delta E^\ddagger$ kcal/mol) and Gibbs free energy barrier ( $\Delta G^\ddagger$ kcal/mol) for H <sub>3</sub> PO P(V) to P(III) tautomerization at 298.15K. ....	20
Table 2.2: B3LYP/6-311++G(3df,3pd) Tautomerization energy barrier ( $\Delta E^\ddagger$ kcal/mol) and TS P-H and P=O bond lengths (Å) for the phosphinylidenes in Figure 2.9 and two test cases... .	27
Table 2.3: Computed P(V) to P(III) tautomerization energy ( $\Delta E$ kcal/mol) from B3LYP calculations on phosphine oxide with various basis sets. ....	28
Table 2.4: B3LYP/6-311++G(3df,3pd) Tautomerization energy ( $\Delta E$ kcal/mol) and P(V) P-H and P=O bond lengths (Å) for the phosphinylidenes in Figure 2.11 and two test cases. ....	35
Table 3.1: Kinetics data for decay of <i>tert</i> -butyl octylphosphinate after addition of D <sub>2</sub> O using various catalysts and the corresponding computed tautomerization barriers. ....	46
Table 3.2: Kinetics data for decay of di <i>tert</i> -butyl phosphite after addition of D <sub>2</sub> O using various catalysts and the corresponding computed tautomerization barriers. ....	46
Table 4.1: Statistical errors in Hartree-Fock bond lengths for the computed geometries of the fluoroquinolones in Figure 4.4, evaluated with various basis sets (n is the number of basis functions used to describe ciprofloxacin). ....	61
Table 4.2: Statistical errors in Hartree-Fock bond angles for the computed geometries of the fluoroquinolones in Figure 4.4, evaluated with various basis sets (n is the number of basis functions used to describe ciprofloxacin). ....	62
Table 4.3: Statistical errors in B3LYP bond lengths for the computed geometries of the fluoroquinolones in Figure 4.4, evaluated with various basis sets (n is the number of basis functions used to describe ciprofloxacin). ....	62

Table 4.4: Statistical errors in B3LYP bond angles for the computed geometries of the fluoroquinolones in Figure 4.4, evaluated with various basis sets (n is the number of basis functions used to describe ciprofloxacin) .....	63
Table 4.5: Statistical errors in bond lengths for the computed geometries of the fluoroquinolones in Figure 4.4 and Figure 4.5, evaluated with different methods.....	64
Table 4.6: Statistical errors in bond angles for the computed geometries of the fluoroquinolones in Figure 4.4 and Figure 4.5, evaluated with different methods.....	64
Table 4.7: Statistical errors in B3LYP bond lengths for the magnesium-containing bonds in Figure 4.8. ....	68
Table 4.8: Statistical errors in B3LYP bond angles for the magnesium-containing bonds in Figure 4.8. ....	68
Table 4.9: Experimental and computed binding free energy $\Delta G$ (kcal mol <sup>-1</sup> ) for fluoroquinolone-Mg <sup>2+</sup> complexation for the fluoroquinolones in Figure 4.9.....	70
Table 4.10: Comparison of Predicted $\Delta G_{\text{calc}}$ (kcal mol <sup>-1</sup> ) fluoroquinolone-Mg <sup>2+</sup> binding for different protonation states for the fluoroquinolones in Figure 4.9. ....	74
Table 4.11: Comparison of predicted $\Delta G_{\text{calc}}$ (kcal mol <sup>-1</sup> ) fluoroquinolone-Mg <sup>2+</sup> binding for neutral molecules (theoretical zwitterion in brackets). ....	77
Table 4.12: Computational methods used for this study.....	82
Table 4.13: Modelling the ciprofloxacin zwitterion by addition of up to five explicit water molecules. ....	85
Table 5.1: First order reaction constants, half-lives and DFT-computed N1 proton affinities $\Delta E_{\text{prot}}$ relative to hydrazone 4d. ....	95
Table 5.2: Effect of solvent on keto-enol equilibrium constant for acetylacetone .....	99

Table 5.3: Computed N1 proton affinities $\Delta E_{\text{prot}}$ relative to 4d .....	103
Table 5.4: Computed N1 partial charges. ....	106
Table A.1: Computational methods used during this study and their level of theory .....	125
Table A.2: Correspondence of orbital type with sum of exponents .....	127
Table A.3: Orbitals required to describe the first 6 atoms using minimal basis sets .....	128

## List of Equations

Equation 2.1: Enthalpy.....	20
Equation 2.2: Ideal gas law.....	20
Equation 2.3: Gibbs energy.....	21
Equation 2.4: Arrhenius equation.....	33
Equation 2.5: Natural logarithm of the Arrhenius equation. ....	33
Equation 2.6: Eyring equation. ....	33
Equation 2.7: Relationship of half-life to rate constant. ....	34
Equation A.1: Coulomb potential .....	113
Equation A.2: Newton's 2 <sup>nd</sup> Law.....	113
Equation A.3: Schrödinger equation.....	114
Equation A.4: Sum of kinetic (T) and potential energy (V) .....	114
Equation A.5: Kinetic energy .....	114
Equation A.6: Laplacian operator .....	114
Equation A.7: Potential energy .....	114
Equation A.8: Born-Oppenheimer approximation.....	115
Equation A.9: Energy as an expectation value of the Hamiltonian operator.....	116
Equation A.10: Integration over all coordinates .....	116
Equation A.11: Energy.....	116
Equation A.12: Antisymmetric wave function to describe the ground state .....	117
Equation A.13: Interchange of electron coordinates.....	117
Equation A.14: Two-electron determinant .....	117
Equation A.15: Slater determinant (SD).....	117

Equation A.16: Fock operator.....	118
Equation A.17: Electron density .....	120
Equation A.18: Number of electrons .....	120
Equation A.19: Ground state energy.....	120
Equation A.20: Kinetic energy approximation .....	121
Equation A.21: Coulombic attraction .....	121
Equation A.22: Electron-electron repulsion energy.....	121
Equation A.23: Local density approximation for exchange energy .....	122
Equation A.24: PBE exchange.....	123
Equation A.25: Exchange enhance factor.....	123
Equation A.26: MO6L general functional .....	123
Equation A.27: B3LYP exchange-correlation functional.....	124
Equation A.28: General DFT method with dispersion correction .....	125
Equation A.29: Dispersion correction.....	125
Equation A.30: Slater-type orbitals.....	126
Equation A.31: Gaussian-type orbitals .....	127
Equation A.32: STO-3G basis set.....	128
Equation A.33: Calculating the binding energy of complexation of two species, A and B .	130
Equation B.1: Hooke's law .....	132
Equation B.2: Newton's second law of motion .....	132
Equation B.3: Simple harmonic motion.....	133
Equation B.4: Second derivative of potential energy .....	133
Equation B.5: Potential energy of state n.....	133

Equation B.6: Enthalpy .....	133
Equation B.7: Gibbs energy .....	134
Equation B.8: Total internal energy .....	134
Equation B.9: Energy from translational motion .....	134
Equation B.10: Energy from molecular rotation.....	135
Equation B.11: Energy from molecular vibration.....	135
Equation B.12: Electronic energies.....	136
Equation B.13: Calculation of total internal energy .....	136
Equation B.14: Ideal gas law .....	136
Equation B.15: Calculation of Enthalpy .....	136
Equation B.16: Calculation of total entropy .....	137
Equation B.17: Calclation of Gibbs energy .....	137

## **List of Symbols and Abbreviations**

ACAT	Advanced compartmental absorption and transit
ADME	Absorption, distribution, metabolism, and excretion
Alk	Alkyl
API	Active pharmaceutical ingredient
B3LYP	Becke 3 parameter Lee Yang Parr method
BDM	Biorelevant dissolution media
Bn	Benzyl
CACO-2	Human colon carcinoma cell line
CBS	Complete basis set
DFT	Density functional theory
DMAPP	Dimethylallyl diphosphate
DOPO	6 <i>H</i> -Dibenzo[ <i>c,e</i> ]oxaphosphinine 6-oxide
DP	Pharmaceutical drug product
DS	Pharmaceutical drug substance
FDA	Food and Drug Administration
FDPS	Farnesyl pyrophosphate synthase
FPP	Farnesyl pyrophosphate
FT	Fischer-Tropsch synthesis
FTase	Farnesyltransferase
FQ	Fluoroquinolone
GGA	Generalized gradient approximation
GTO	Gaussian-type orbital

HMG-CoA	3-Hydroxy-3-methylglutaryl-coenzyme A
HMGCR	HMG-CoA reductase
HTS	High throughput screening
IDEA	<i>In vitro</i> determination for the estimation of ADME
IDI	IPP isomerase
IPP	Isopentyl diphosphate
IVIVC	<i>In vitro – in vivo</i> correlation
LCAO	Linear combination of atomic orbitals
LDA	Local density approximation
LSDA	Local spin density approximation
MAE	Mean absolute error
Me	Methyl
ME	Mean error
MVD	Diphosphomevalonate decarboxylase
MVK	Mevalonate kinase
NME	New molecular entity
NMR	Nuclear magnetic resonance
Ph	Phenyl
PMVK	Phosphomevalonate kinase
PoC	Proof of concept studies
PSA	Polar molecular surface area
SAR	Structure-activity relationship
SD	Slater determinant

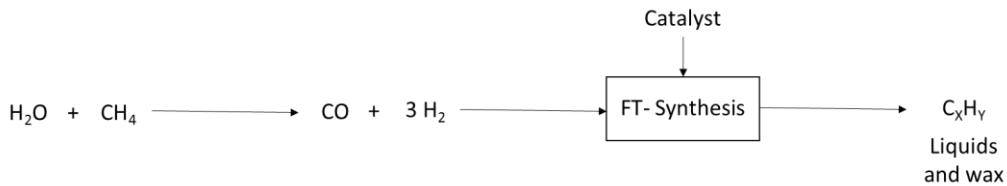
SMILES	Simplified molecular-input line-entry system
SQS	Farnesyl-diphosphate farnesyltransferase 1
STO	Slater-type orbital
SYBYL	Molecular modeling line notation
TLC	Thin layer chromatography
XC	Exchange-correlation
XRD	X-ray diffraction

# Chapter 1: Introduction

## 1.1 General Applications of Computational Chemistry

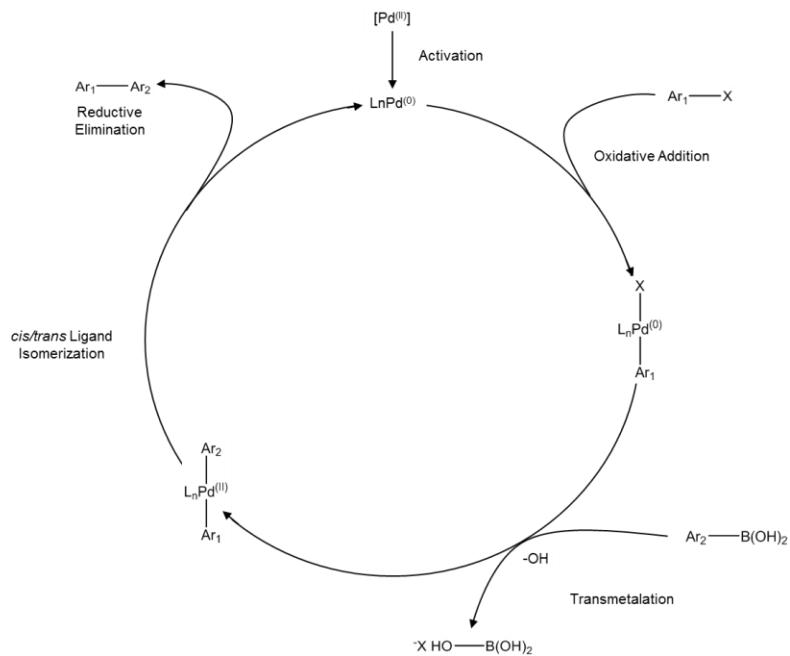
Computational chemistry uses computer simulation to aid in answering chemical questions. The ability to answer reaction and mechanistic problems theoretically is obviously extremely appealing. Not having to conduct experiments can potentially save time and money. However, computational approaches cannot be used on their own, as they must be validated against experimental data. Once validated, the model can be adapted to non-experimental systems as a screen to direct future research and development, with a number of industrially relevant applications using computational chemistry to better understand processes.

Much work has been done on the application of computational chemistry to physical chemistry problems. Surface catalysis is one such area of study. We may know that the addition of a catalyst to a particular reaction can increase the rate of a reaction by lowering the activation energy but we do not know the molecular mechanism by which this is achieved. Computational chemistry has been employed to look at a number of reactions that use catalysis to probe the mechanism and interaction with the surface. Looking at work carried out by the Janesko group, we see a number of recent publications regarding catalysis. Determan *et al.* investigated gold catalysis<sup>1</sup> by considering the dissociation of H<sub>2</sub> on gold clusters, surfaces and nanowires. They then went on to look at H<sub>2</sub> dissociation on group 10 and group 11 metal nanoribbons<sup>2</sup> to assess their catalytic potential, as determined computationally, against experimentally demonstrated catalytic utility. More recently Mahler and Janesko<sup>3</sup> worked to improve modeling of the initial carbon monoxide dissociation on ruthenium, of the Fischer-Tropsch (FT) process (Figure 1.1).



**Figure 1.1:** Fischer-Tropsch process.

Data from the computational studies may help in optimizing synthesis by providing better understanding of the carbon monoxide dissociation on FT catalysts. Computational methods are also employed to understand a particular reaction or reaction step. Pudasaini and Janesko<sup>4,5</sup> performed a number of computational studies into Suzuki-Miyaura coupling,<sup>6</sup> which is shown in Scheme 1.1. The study concentrated on the specific mechanism of the coupling of secondary alkyls and the ligand effects on  $\beta$ -hydride elimination.

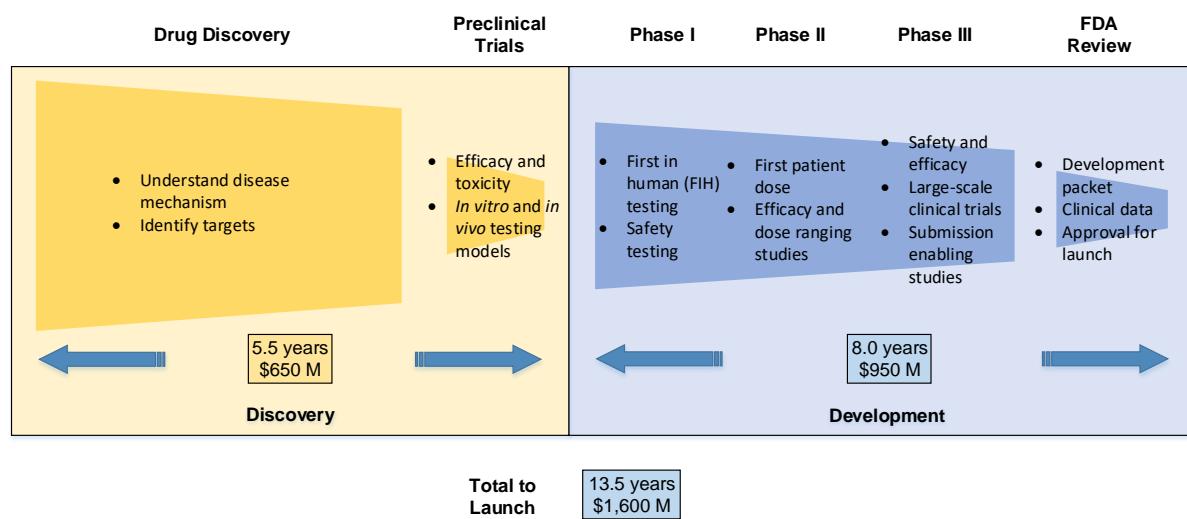


**Scheme 1.1:** General mechanism of Suzuki-Miyaura coupling.

Due to the computational time and cost associated with modeling large, complex systems and the requirement to validate models against experimental data, it is usual to focus on a single part of a reaction or mechanism to improve understanding and insight, rather than on an entire reaction scheme. The following section details how computational chemistry has been applied to pharmaceutical development.

## 1.2 Reducing the Cost of Pharmaceutical Development

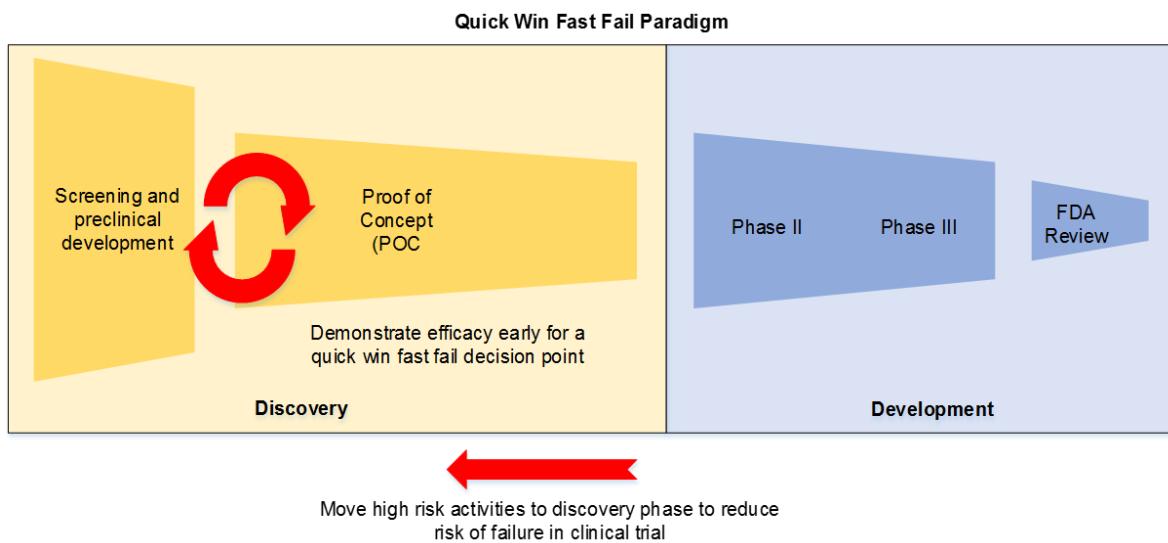
The development of a new pharmaceutical drug product (DP) for commercialization is a long and expensive process. The average time from discovery to market is between 7.5 and 15 years<sup>7–9</sup> depending on the target disease. Those drug substances (DS) developed for critical applications, such as cancer, tend to have a more rapid approval than those with a lower risk/reward ratio. Reviews of the drug development process have yielded a broad range of costs and times associated with bringing a new drug to market.<sup>10–12</sup> An overview of the compiled data is shown in Figure 1.2



**Figure 1.2:** Costs and time associated with bringing a new drug product to market.

Data used to estimate development costs are all based on self-reported numbers from companies, so there are issues associated with the data collection and accuracy. However, a generally accepted figure for the capitalized development in 2000 was around \$800 million, with costs potentially being as high as \$1,700 million in 2010.<sup>7-12</sup> As can be seen in Figure 1.2, the costs are divided between discovery and development. The majority of the costs are associated with the development phase, where clinical trials account for most of the costs in Phase I through Phase III.

The discovery and pre-clinical trials represent the points at which a new chemical entity is most likely to fail. What the figure highlights is the large risk associated with advancing an unsuitable candidate to Phase I or Phase II trials. From Figure 1.2, the total capitalized costs for these two phases is up to \$600 million. A failure at this stage represents a significant loss on investment for a pharmaceutical company. One approach to reduce the exposure to risk of failure in the development stage is to move high risk activities into the discovery stage and generate earlier go/no-go decision points in small scale proof of concept (PoC), first human dose (FHD) studies. This may be achieved by increasing the number of drugs in the candidate screening (CS) stage. The strategy is described by Owens *et al.*<sup>13</sup> and summarized in Figure 1.3. By screening more molecules, the probability of identifying a potential drug is increased and by conducting small-scale PoC studies creates an early timepoint for making a go/no-go decision regarding development. This is known as the “quick win, fast fail” approach, aimed at reducing the risk of a more costly fail later in the development stages, when more significant investment has already been made.



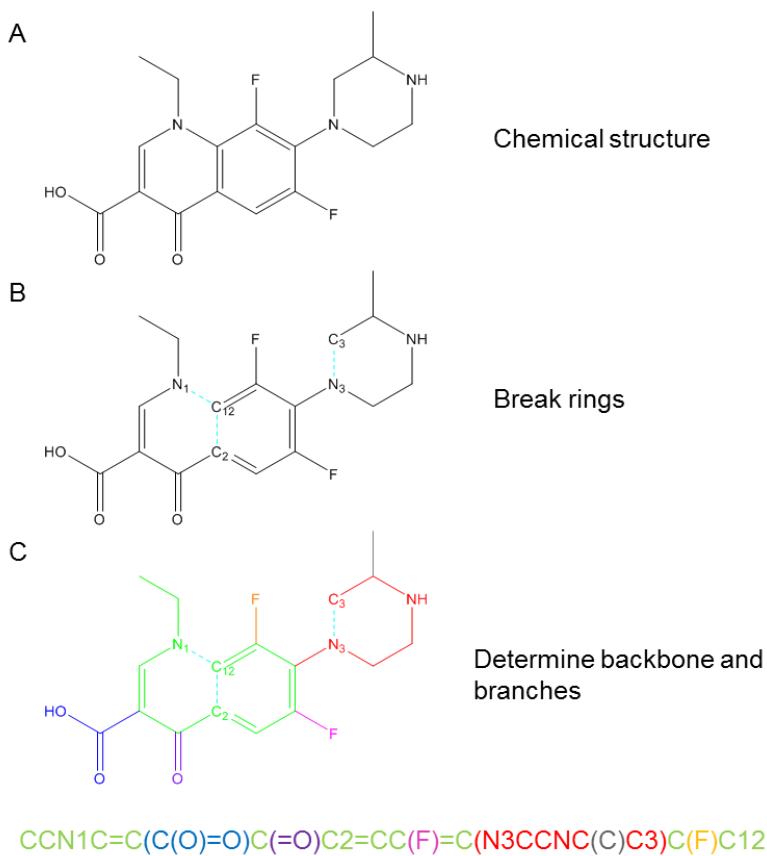
**Figure 1.3:** The “quick win, fast fail” drug development paradigm.

Paul *et al.*<sup>8</sup> put some estimated cost savings for successful implementation of the quick win, fast fail paradigm. They projected a 25% reduction in development costs by reducing the failure rate in Phase II from 66% to 50%. For a \$1,700 million development, this represents a saving of \$425 million.

### 1.3 Pharmaceutical Application of Screening and Computational Chemistry

Pharmaceutical development is concerned with two major outcomes. The first is identification of a new drug substance (DS), which is the active pharmaceutical ingredient (API). Secondly, a drug product (DP), which is the final dosage form, will be developed for evaluation in clinical trials. In order to evaluate the increased number of molecules entering the funnel in the quick win paradigm, some screening must be employed as not all the potential candidates can be exhaustively tested. This approach runs the risk of rejecting a potential DS but comes with a reduced cost. The most common screening methods employed by

pharmaceutical companies are relatively simple and are based on easily measured parameters. They also tend to focus on the applicability of the candidate to parenteral or oral dosages, screening against ADME (adsorption, distribution, metabolism, and excretion) properties, as these represent the major parameters for these routes of drug administration. One pharmacokinetic high throughput screen used during pharmaceutical development is Lipinski's rule of 5.<sup>14</sup> Based on an original 2000 drug candidate database<sup>15</sup> of physico-chemical properties, the method screens candidates based on an estimate of their solubility and permeability, using easily calculated drug properties. Poor absorption or permeation is predicted to be more likely when the candidate molecule has more than 5 H-bond donors, 10 H-bond acceptors, molecular weight greater than 500 and lipophilicity (Log P or Clog P) greater than 5. While the molecular weight, number of H-bond donors and H-bond acceptors may be established from the molecular structure, the Log P must be determined experimentally. For screening purposes, where exact values are not as important, lipophilicity may be calculated computationally, commonly by the group contribution method (Clog P).<sup>16</sup> This method calculates a value for LogP based on non-overlapping molecular fragments, using experimental data and assuming that properties of these fragments are the same in different molecules. These fragment contributions are determined experimentally in pure components and mixtures. Two other common screens may also be employed at this stage. Clark<sup>17</sup> developed a screen based on polar molecular surface area (PSA), after demonstrating correlation between PSA and intestinal absorption. The method converts the SMILES representation<sup>18-20</sup> (as shown in Figure 1.4), to a 3-D structure via the CONCORD program<sup>21</sup> and then energy-minimized in a molecular mechanics model using the SYBYL® line notation.<sup>22</sup>



**Figure 1.4:** Generation of SMILES input.

The van der Waals molecular surface area of the SYBYL® conformation is then calculated in MOLVOL,<sup>23</sup> prior to the addition contributions of the polar atoms (N, O, and H attached to N or O) calculated in Fortran, to generate the final PSA. The screening method can be entirely automated by implementation of a script for HTS. Candidates with a PSA  $\geq 140 \text{ \AA}^2$  are considered to likely exhibit poor intestinal absorption. The Lewis screening method<sup>24</sup> focuses on drug metabolism by P450s, based on structure-activity relationships.<sup>25,26</sup>

There have been more screens developed around intestinal absorption as part of pharmacokinetic (PK) *in vitro* screening. IDEA<sup>27</sup> and the ACAT<sup>28</sup> model in GASTROPLUS are two such screening tools.<sup>29,30</sup> Both use a combination of *in vitro* and *in silico* inputs for

molecular properties such as pK<sub>a</sub>, solubility, lipophilicity to estimate intestinal absorption, as shown in Table 1.1.

<b>Input</b>	<b>IDEA</b>	<b>GASTROPLUS</b>
Chemical structure	SMILES structure (or ChemDraw)	SMILES structure (or ISIS)
Dose and formulation	Dose (mg)	Initial dose (mg) Subsequent doses (mg) Dosing interval (hr) Dose volume (mL) Drug particle density (g/mL) Effective particle radius ( $\mu\text{m}$ ) Dosage form
Solubility	Solubility at pH values 1.5 to 7.5	Solubility at pH values 1.5 to 7.5 (or solubility at one known pH and a table of pK <sub>a</sub> values)
Permeability	CACO-2 or rabbit intestinal tissue	Permeability measure transformed based on correlation to human permeability or in silico estimate
pK <sub>a</sub>	N/A	Table of pK <sub>a</sub> values
Lipophilicity	N/A	LogD at known pH or LogP

**Table 1.1:** Input data comparison for IDEA and GASTROPLUS.

GASTROPLUS is a little more sophisticated in that it is based on a compartmental model, with nine compartments corresponding to different parts of the digestive tract. In developing screening models, it is important to show that the model and *in vitro* testing results have some correlation with *in vivo* data. Lobenberg *et al.*<sup>31</sup> demonstrated that glibenclamide formulations (a class of antidiabetic drugs), demonstrated good correlation between *in vitro* dissolution and *in vivo* absorption (IVIVC), as long as a biorelevant dissolution media (BDM) was used, such as fasting state simulated intestinal fluid (FaSSIF). These screening methods are able to be used for HTS as they focus on the ADME properties rather than on whether any of the

candidates have any pharmacological targets. The other item of note in these screening methods is that the optimized geometry of the molecule is not required.

In order to screen for pharmacological activity, structure-activity relationship (SAR) screening must be employed. Screening based on these principles can take longer as they are screened against specific targets and known active drug pharmacophores. Quantitative SAR (QSAR)<sup>32,33</sup> can be used to screen potential drugs using the pharmacophores of known active molecules. The approach is based on the assumption that molecules with similar structural shape will exhibit similar efficacy against the same biological target. Developing this concept leads to the evaluation of molecular docking of a molecule with a target receptor.<sup>34</sup> The molecular interactions are evaluated and binding estimated as a screen for molecules likely to have biological efficacy. This screening tool requires more detailed input. Firstly, the structure of the receptor must be determined. The 3-D molecular structure of the drug molecule must then be determined, either by X-ray or NMR. The molecular docking program<sup>35</sup> is then designed to evaluate drug-protein binding free energies by moving the drug around on a molecular model of the protein. The downside to these more advanced screening methods is that they are target specific and require a greater level of input, so are generally used further along in the discovery phase.

#### **1.4 Goal**

This work aims at identifying commercial application and aspects of the pharmaceutical development process to which computational approaches can be applied synergistically with experimental work to increase efficiency. To attain the goal, four projects are detailed to show how the application of computational methods can be of benefit. These areas are commercial synthesis, which also has potential for the development of new

pharmaceutical DS (Chapter 2 and Chapter 3), drug substance interaction with target (Chapter 4) and modifications to an existing drug substance (Chapter 5).

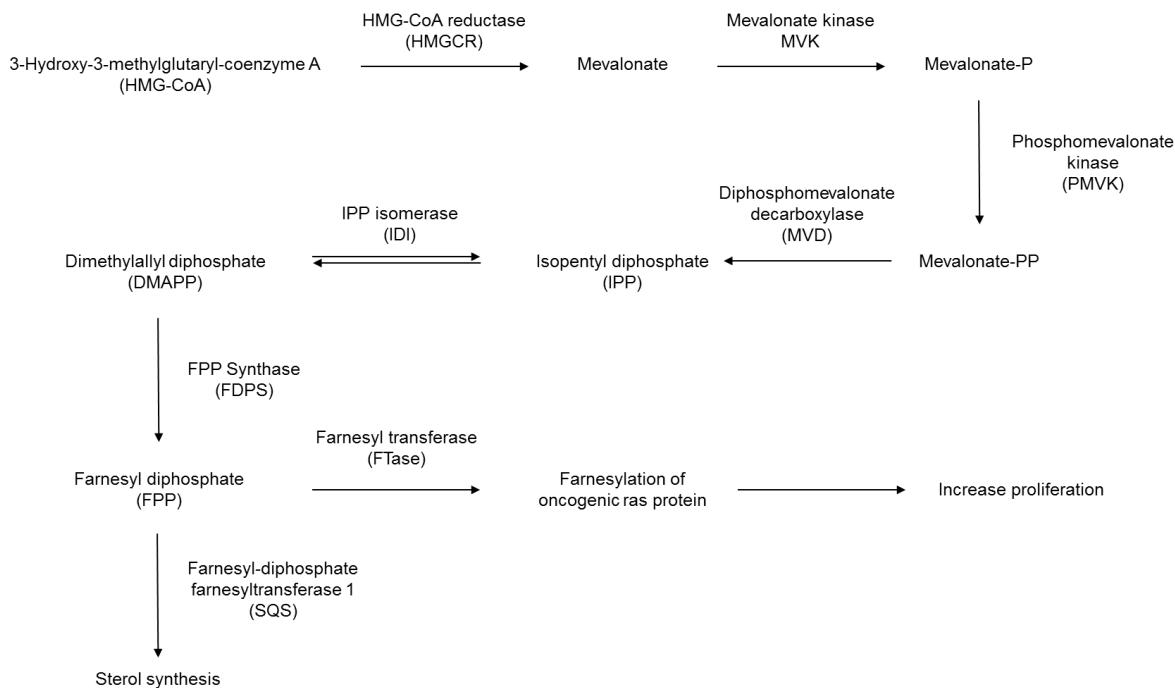
To achieve the desired outcomes, Density Functional Theory (DFT) calculations on model reactants, catalysts and ligands will be used to build insights into those reactions and interactions. While a number of screening and modeling approaches have been discussed, this work can aid in a number of areas: such as identifying a viable alternative to current manufacturing processes used to create new P-C bonded compounds, as well as new catalysts to improve reaction efficiency; fluoroquinolone binding to Mg<sup>2+</sup> as a screening method for new antibacterial entities with the same mode of action; stability of hydrazones formed after reaction with ketones or diketones, which may be of utility in the formation of prodrugs from existing DS.

A brief account of the computational theories involved are included in Appendix A and Appendix B. All computed geometries can be found in Appendix C. All calculations use Gaussian 09 software.<sup>36</sup>

## Chapter 2: Phosphinylidene Tautomerization

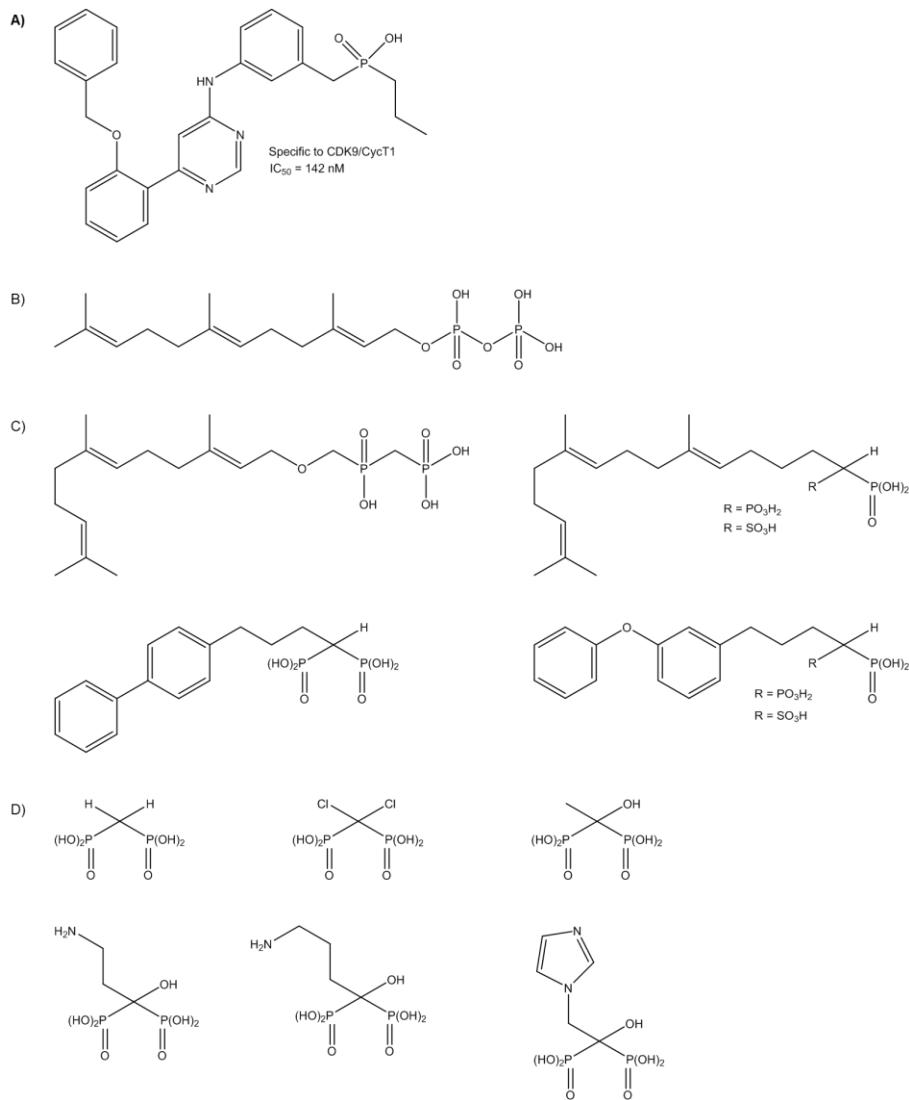
### 2.1 Phosphorus Chemistry

Organophosphorus compounds are organic compounds that contain phosphorus. They have broad utility in a number of applications, including a role in organic synthesis as components or intermediates, in agriculture as pesticides<sup>37–39</sup>, as flame retardants and ligands in organometallic catalysis. They also feature in pharmaceuticals, with phosphorus-containing compounds being employed as kinase inhibitors,<sup>40</sup> squalene synthesis inhibitors, anti-cancer agents,<sup>41</sup> and for the treatment of osteoporosis,<sup>42</sup> as shown in Figure 2.1. A common element of the phosphorus-containing drugs in Figure 2.1, is their action on the isoprenoid biosynthetic pathway,<sup>41</sup> as shown in Scheme 2.1.



**Scheme 2.1:** Central features of the human isoprenoid biosynthetic pathway.

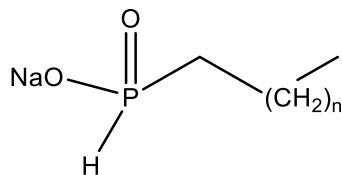
The pathway is important for a number of diseases. Farnesyl-diphosphate farnesyltransferase 1 (SQS) inhibitors can be used for the treatment of high cholesterol, by blocking the synthesis of sterols. They may also exhibit binding to and inhibit farnesyl transferase (FTase), which leads to a proliferation of ras protein observed in cancer. Bisphosphonates bind to farnesyl diphosphate synthase (FDPS) and exhibit high bone affinity.



**Figure 2.1:** A) Kinase inhibitor. B) Structure of farnesyl diphosphate (FPP). C) FPP analogs that are known inhibitors of SQS. D) Chemical structures of bisphosphonates of clinical relevance for the treatment of osteoporosis.

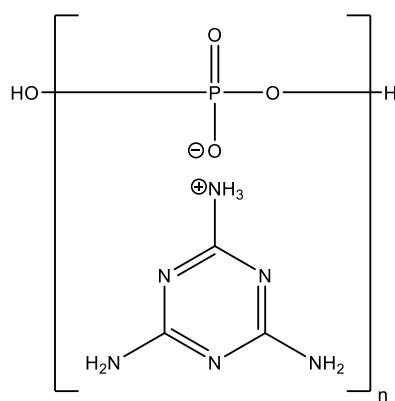
## 2.2 Commercial Applications of Organophosphorus Compounds

As mentioned, organophosphorus compounds have many commercial uses, outside of pharmaceutical applications. *H*-Phosphinate, with structures following that in Figure 2.2 that have an alkyl chain of 8 or more carbon atoms, have been shown to have surfactant and detergent properties.<sup>43</sup>



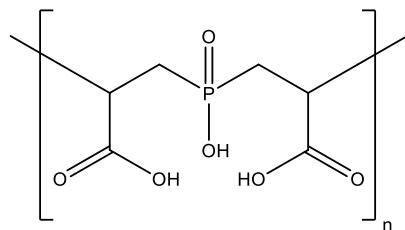
**Figure 2.2:** *H*-phosphinate detergent.

Phosphorus-containing compounds have also been used as flame retardants.<sup>44</sup> One such example is melamine polyphosphate (MPP), as shown in Figure 2.3. MPP acts as a retardant by undergoing endothermic decomposition. In this fashion, they act as a heat sink. Nitrogen gases such as ammonia are also released during decomposition, diluting oxygen and flammable gases. A char of phosphoric acid is also formed, acting as an insulating layer.



**Figure 2.3:** Melamine polyphosphate (MPP) flame retardant.

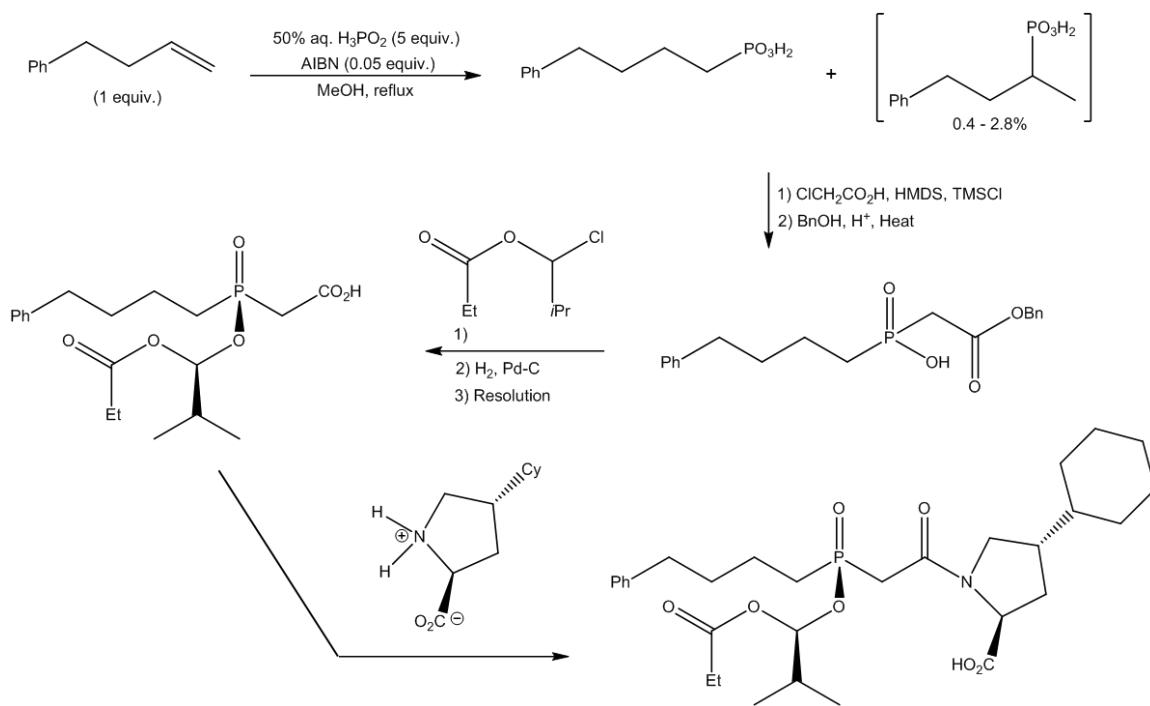
Organophosphorus compounds have also been used in the oil and gas industry. Phosphorus-containing polymers, such as the one shown in Figure 2.4, have been used as retarding additives, to extend the setting time for cement at the high temperatures (200-450°F) encountered when sealing casings in bore wells following drilling.<sup>45</sup>



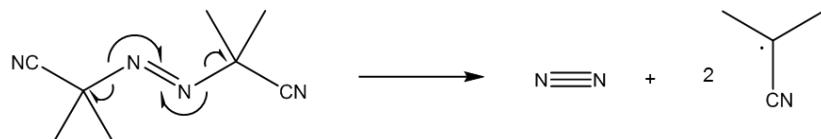
**Figure 2.4:** Hydraulic cement retarding polymer.

### 2.3 Phosphinylidene Chemistry

There have been a number of studies looking into the use of P(O)H as reactants for creating new phosphorus-carbon bonds via numerous pathways, including transition-metal catalysis and free-radical processes. Coudray and Montchamp<sup>46</sup> reviewed the landscape surrounding the addition of phosphinylidene compounds to unsaturated hydrocarbons. An example of addition to a terminal alkene is shown in Scheme 2.2. Here, the pharmaceutical drug Fosinopril, an angiotensin converting enzyme (ACE) inhibitor used for the treatment of hypertension, is made under radical conditions.<sup>47</sup> A radical initiator, azobisisobutyronitrile (AIBN) is used, with radical formation coming from its decomposition, as shown in Figure 2.5.

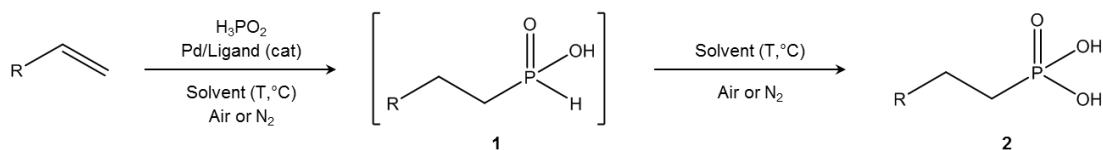


**Scheme 2.2:** Preparation of Fosinopril by radical initiation.



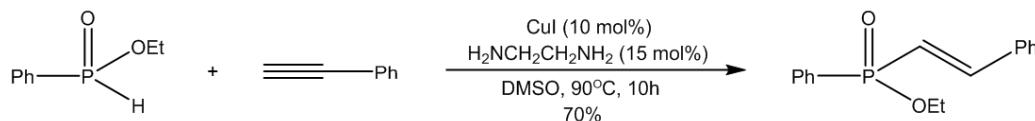
**Figure 2.5:** AIBN decomposition to form 2-cyanoprop-2-yl radicals.

Palladium catalysts can also be employed for the phosphinylidene addition to alkenes to form phosphonic acids, as described by Bravo-Altamirano and Montchamp.<sup>48</sup> An overview of the one pot process is shown in Scheme 2.3.



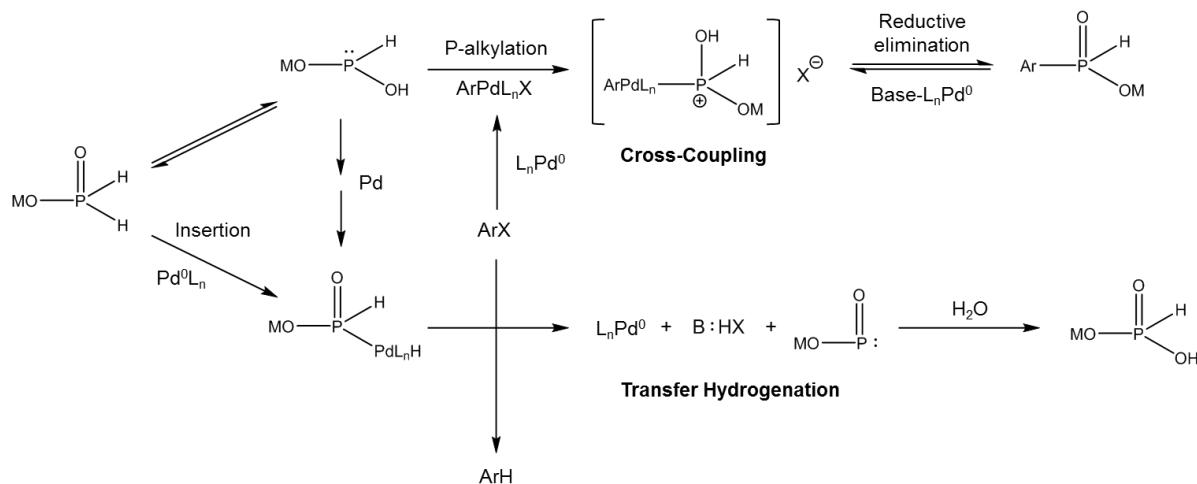
**Scheme 2.3:** One pot process for phosphonic acid synthesis (2) from hypophosphorous acid reaction with an alkene via a phosphinic acid intermediate (1).

New P-C bonds can also be formed through addition of phosphinylidenes to alkynes, with the use of transition metal catalysts, such as the reaction shown in Scheme 2.4.



**Scheme 2.4:** Hydrophosphinylation of terminal alkynes.

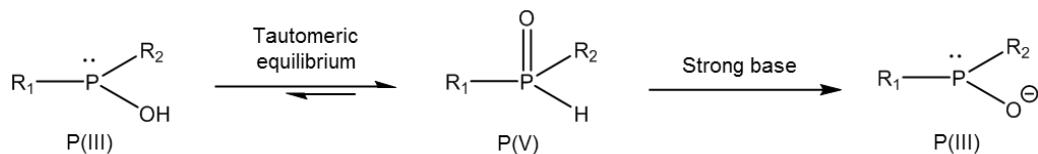
The majority of phosphinylidene reaction literature, as highlighted in the examples above, focuses on radical and transition metal catalyzed pathways. While looking at palladium-catalyzed cross-coupling of anilinium hypophosphite, Montchamp and Dumond<sup>49</sup> proposed a competing set of reactions for cross-coupling and reductive transfer hydrogenation, as shown in Scheme 2.5.



**Scheme 2.5:** Possible cross-coupling and transfer hydrogenation pathways.

Of particular interest is the initial step of the cross-coupling reaction. It involves the tautomerization of P(V) phosphinylidene to a reactive P(III) intermediate, prior to cross-coupling. This starts the journey towards using phosphinylidene chemistry to form reactive

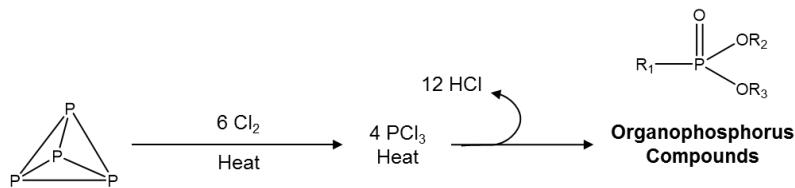
P(III) intermediates, without the requirement for radical reactions, metal catalysts or harsh reaction conditions. This work was expanded on by Berger *et al.*<sup>50</sup> when it was proposed that P(III) phosphinylidene tautomers could be used as feedstocks in cross-coupling reactions for the formation of new P-C bonds. Two paths for the formation of P(III) tautomers were presented in Scheme 2.6. One via tautomeric equilibrium, the other through reaction with a strong base. The fact that phosphinylidenes can undergo tautomerization to a reactive intermediate forms the basis for the subsequent study.



**Scheme 2.6:** Generation of phosphinylidene P(III) reactive feedstocks for cross-coupling reactions.

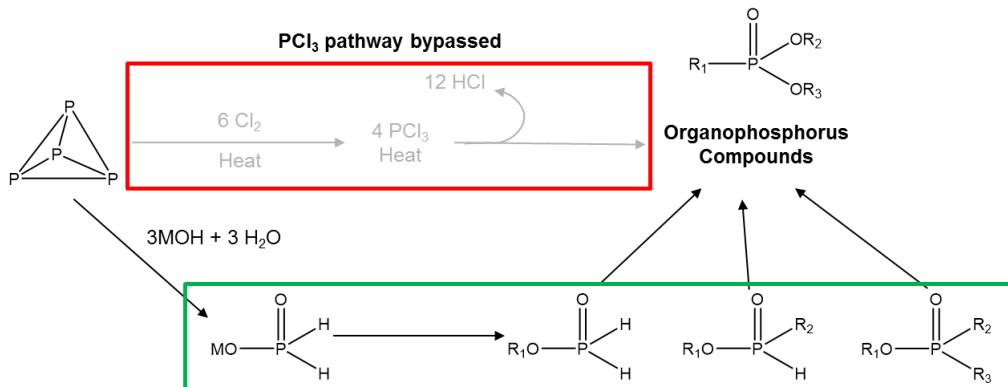
## 2.4 Why the Interest in Phosphinylidenes for Phosphorus-Carbon Bonds?

The varied use of organophosphorus compounds makes the formation of P-C bonds an area of great interest and of relevance for commercial processes and to the development of new pharmaceutical drug substances. To date, the industrial process employed in the manufacture of P-C bond containing compounds, relies on the use of phosphorus trichloride ( $\text{PCl}_3$ ).<sup>51,52</sup> The process, shown in Scheme 2.7, requires large amounts of energy and generates hydrochloric acid (HCl) as a by-product. While no products from the industrial process contain P-Cl bonds, the reaction is hazardous, with its generation of toxic and corrosive HCl.



**Scheme 2.7:** Industrial process for the manufacture of organophosphorus compounds.

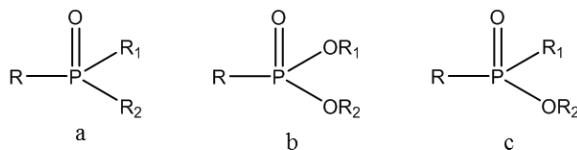
As such, a safer and more efficient approach would be more desirable. For a process not requiring  $\text{PCl}_3$ , phosphinylidene compounds, with the general formula  $\text{R}^1\text{R}^2\text{P}(\text{O})\text{H}$ , are proposed as a starting feedstock. They are produced commercially through alkaline hydrolysis of white phosphorus ( $\text{P}_4$ ), presenting an already available alternative feedstock to  $\text{PCl}_3$ , as shown in Scheme 2.8.



**Scheme 2.8:** Alternative organophosphorus compound manufacturing process using phosphinylidenes (shown in green box) as feedstock, while bypassing  $\text{PCl}_3$  pathway (shown in red box).

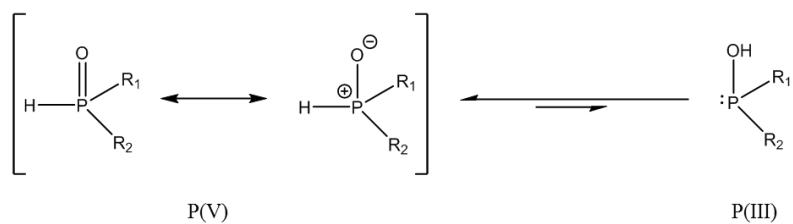
It has been previously demonstrated that phosphinylidenes undergo useful reactions but the interest in them as a starting feedstock for P-C bond formation is the fact they can undergo prototropic tautomerization. The work discussed primarily focuses on secondary

phosphine oxides ( $\text{R}^1(\text{R}^2)\text{P}(\text{O})\text{H}$ ), *H*-phosphonates ( $\text{R}^1\text{O}(\text{R}^2\text{O})\text{P}(\text{O})\text{H}$ ) and *H*-phosphinates ( $\text{R}^1(\text{OR}^2)\text{P}(\text{O})\text{H}$ ), the structures of which are shown in Figure 2.6.



**Figure 2.6:** Structures of (a) *H*-phosphine oxides, (b) *H*-phosphonates and (c) *H*-phosphinates, where R=H.

These two classes of organophosphorus compound are of interest in the current study as the presence of the phosphinylidene moiety [P(O)H] allows for the prototropic tautomerization from a stable P(V) to a reactive P(III) structure, as shown in Scheme 2.9, that presents a possible starting feedstock for industrial P-C bond formation, without the requirement for the use of  $\text{PCl}_3$ .<sup>53</sup> Looking at the tautomers shown in Scheme 2.9, we would expect to see the stabilization of the P(V) resonance structures by the inclusion of electron-donating substituents at R<sub>1</sub> and R<sub>2</sub>. Likewise, we anticipate that electron-withdrawing substituents would destabilize the P(V) tautomer, with the P(III) form being more favored. If viable as a starting feedstock, this would represent a more environmentally friendly and possibly more efficient process than the current method and could mark the first step in a commercially viable synthesis for the production of P-C bond containing compounds.



**Scheme 2.9:** Prototropic tautomerization of phosphinylidenes.

Unfortunately, the tautomerization energy barrier is approximately 60 kcal/mol when no catalyst is employed. Water can act as a simple catalyst, as shown in Table 2.1, which is consistent with previous work.<sup>54</sup> Uncatalyzed tautomerization of H<sub>3</sub>PO takes place via a strained 3-membered ring transition state (TS), with a high Gibbs free energy barrier of 59 kcal/mol. Catalysis, as modeled by one explicit water molecule, proceeds via a less strained 5-membered ring transition state, reducing the barrier to approximately 37 kcal/mol, as shown in Figure 2.7 and Figure 2.8.

Catalyst	$\Delta E^\ddagger$	$\Delta G^\ddagger$
Uncatalyzed	62.3	59.1
Continuum water	65.3	62.0
One explicit H <sub>2</sub> O molecule	28.2	36.8
Two explicit H <sub>2</sub> O molecules	9.3	28.7

**Table 2.1:** B3LYP/6-311++G(3df,3pd) total energy barrier ( $\Delta E^\ddagger$  kcal/mol) and Gibbs free energy barrier ( $\Delta G^\ddagger$  kcal/mol) for H<sub>3</sub>PO P(V) to P(III) tautomerization at 298.15K.

$\Delta E^\ddagger$  relates to the enthalpy, as shown in Equation 2.1, where  $U$  is the internal energy. From the ideal gas law, shown in Equation 2.2, we obtain an expression for  $pV$ .

$$H = U + pV$$

**Equation 2.1:** Enthalpy.

$$pV = nRT = nk_B T$$

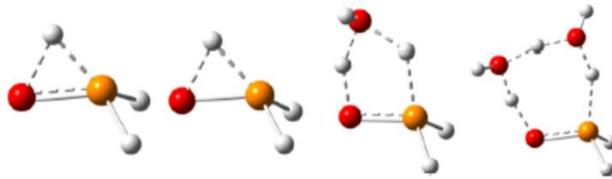
**Equation 2.2:** Ideal gas law.

Since the calculations are run at 0K, there is no entropy term included, hence  $\Delta E^\ddagger = \Delta H = \Delta U$ .  $\Delta G^\ddagger$  shows the Gibbs free energy change for tautomerization. The calculation is shown in Equation 2.3. In this case, a gas phase entropy term is included.

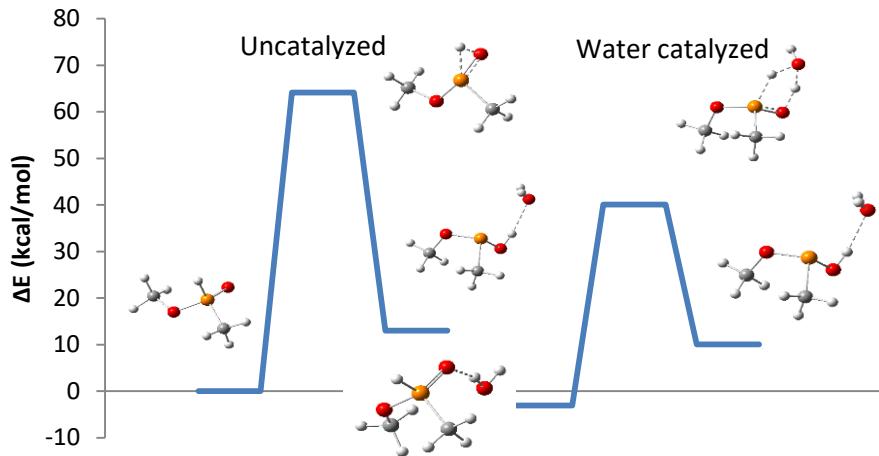
$$G = H - TS_{Tot}$$

**Equation 2.3:** Gibbs energy.

Adding a second explicit water molecule reduced the strain in the TS further by producing a 7-membered ring, yielding a Gibbs free energy barrier of 29 kcal/mol. Modelling with an explicit water molecule is required to stabilize the TS. This stabilization effect of water on the TS is not captured by using a continuum solvent model, where the Gibbs free energy barrier did not approach the 22 kcal/mol lowering achieved by catalysis with a single explicit water molecule but actually resulted in an increase in barrier energy by 3 kcal/mol to a calculated barrier energy of 62 kcal/mol. As with any time order in a system is increased, there is an entropic penalty incurred. We see this entropic penalty by ordering a second water molecule in the TS Gibbs free energy, with only an 8 kcal/mol decrease in  $\Delta G^\ddagger$ , compared to catalysis with only a single water molecule. While  $\Delta G^\ddagger$  is reported for the tautomerization barrier energies, it must be noted that the entropic penalty for the addition of explicit water molecules is an overestimation due to the use of gas phase approximations that do not take into account solvent effects.



**Figure 2.7:** Calculated transition state geometries for  $\text{H}_3\text{PO}$  tautomerization. From left to right, the figure shows uncatalyzed tautomerization, tautomerization in continuum water, and tautomerization catalyzed by one and two explicit  $\text{H}_2\text{O}$  molecules.

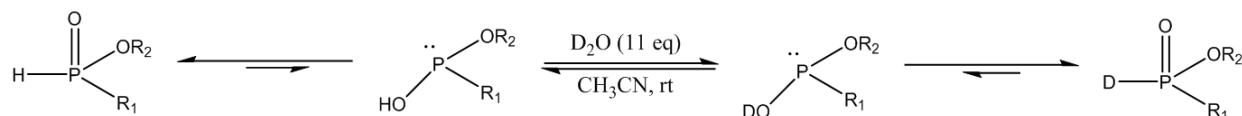


**Figure 2.8:** B3LYP/6-311++G(3df,3pd) calculated geometries and potential energy surfaces for methyl *H*-phosphinate tautomerization, both uncatalyzed and catalyzed by a single explicit water molecule, in a water continuum solvent.

While there may be a debate regarding the value of including gas phase entropies and that the inclusion of two explicit water molecules may yield  $\Delta E^\ddagger$  in line with experiment, all subsequent tautomerization computational models consider only a single explicit water molecule as a catalyst and report the Gibbs energy of tautomerization.

## 2.5 Tautomerization Rate

To measure the rate of tautomerization, we require a way to follow the progress of the reaction. We are able to follow the reaction progress indirectly by measuring the deuteration of phosphinylidene in D<sub>2</sub>O, as shown in Scheme 2.10.

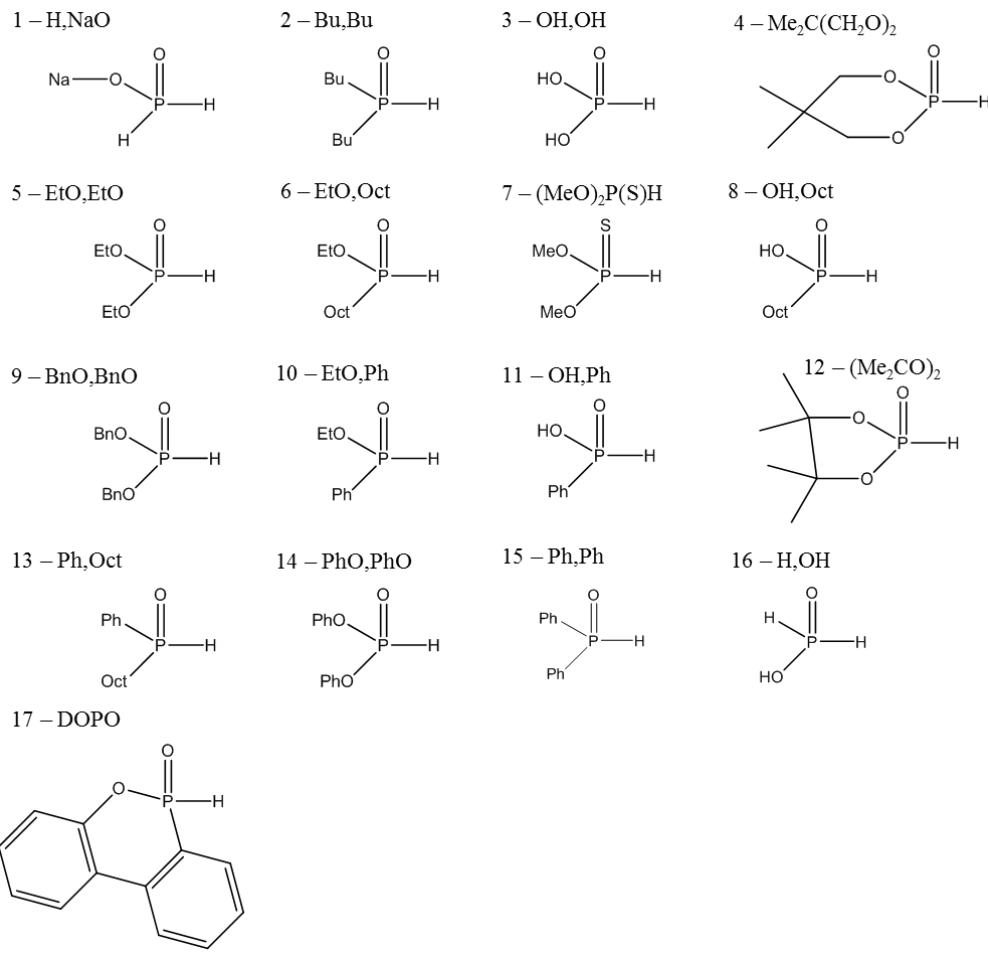


**Scheme 2.10:** Deuteration of phosphinylidene compounds using an excess of D<sub>2</sub>O.

This is due to the fact that deuteration is directly correlated with tautomerization, as the deuteration takes place via the P-O-H of the P(III) tautomer, rather than via the P-H of the P(V) tautomer, as this has a high energy barrier. It has been shown previously that phosphinylidene P(III) tautomers will exchange the P-O-H proton for deuterium extremely rapidly. The experimental portion of the study was carried out by Henry Fisher of the Montchamp Group. The use of excess D<sub>2</sub>O employed to maintain reaction conditions where tautomerization and not deuteration is the rate-limiting step. The <sup>31</sup>P NMR spectra were recorded on a 400 MHz spectrometer. NMR kinetics were recorded at room temperature. The NMR yields are determined by integration of all the resonances in the <sup>31</sup>P spectra. Phosphinylidene compounds were diluted to 1 M using freshly distilled CH<sub>3</sub>CN. NMR kinetic experiments were completed in duplicate for each phosphinylidene sample.

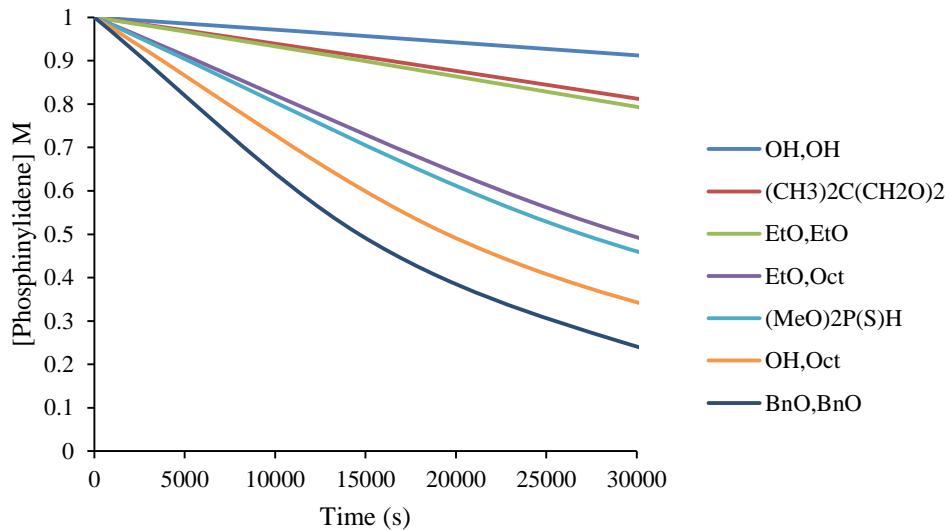
## 2.6 Effect of Substituents on Phosphinylidene Tautomerization Rate

To observe the effect of phosphinylidene R<sup>1</sup> and R<sup>2</sup> substituents on the rate of tautomerization, a combined experimental and computational study was carried out on the molecules in Figure 2.9.

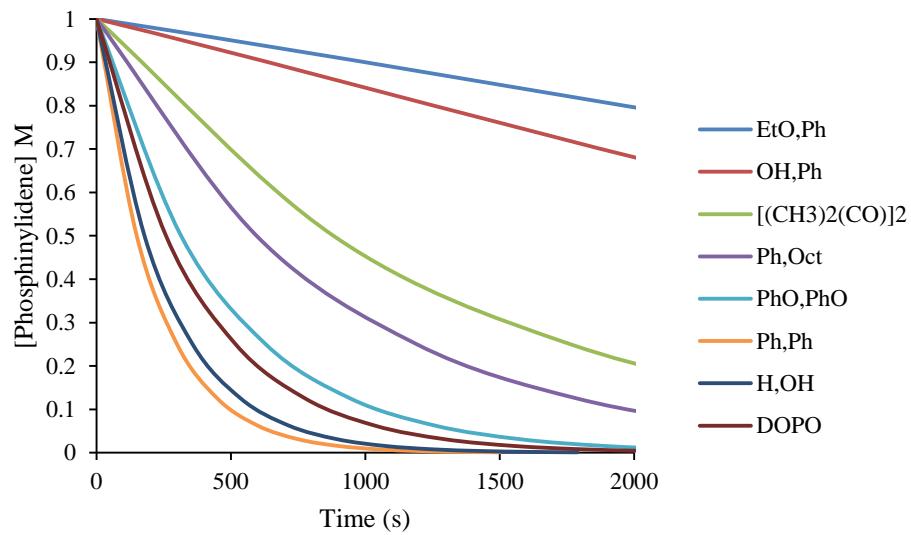


**Figure 2.9:** Phosphinylidenes investigated in experimental tautomerization study.

Decay data from  $^{31}\text{P}$  NMR spectra obtained by Henry Fisher are summarized in Figure 2.10, Figure 2.11 and Table 2.2.



**Figure 2.10:** Decay of phosphinylidene compounds (1M in CH<sub>3</sub>CN) during deuteration with D<sub>2</sub>O (half-lives > 3 hours).



**Figure 2.11:** Decay of phosphinylidene compounds (1M in CH<sub>3</sub>CN) during deuteration with D<sub>2</sub>O (half-lives ~1 hour or less).

Molecule 1 (H<sub>3</sub>NaO; entry 1) and molecule 2 (Bu<sub>2</sub>Bu; entry 2) did not undergo deuteration. Half-lives observed in experiment range from 49 hours (OH,OH; entry 3) to 150 seconds (Ph,Ph; entry 15). Phosphinic acid (H<sub>3</sub>OH; entry 16) and DOPO (entry 17) have similarly fast tautomerization rates, with half-lives of 179 and 259 seconds respectively. The difference in tautomerization rates of phosphorous acid (OH,OH) and phosphinic acid (H<sub>3</sub>OH) somewhat follows the difference in pK<sub>a</sub> (1.3 versus 1.2 respectively) of the acidic hydrogen. This makes the phosphorus atom in phosphinic acid (H<sub>3</sub>PO<sub>2</sub>) slightly more positively charged and increases the tautomerization rate via breaking of the P(V) P-H bond. There is a general trend where *H*-phosphinates (R<sup>1</sup>O)(R<sup>2</sup>)P(O)H undergo faster tautomerization than the corresponding *H*-phosphonate (R<sup>1</sup>O)(R<sup>2</sup>O)P(O)H.

Entr y	<b>R<sup>1</sup></b>	<b>R<sup>2</sup></b>	<b>Half-Life (s)</b>	<b>ΔE<sup>‡</sup> (kcal/mol)</b>	<b>R(P-H)</b>	<b>R(P=O)</b>
1	H	NaO	Not Observed	-	-	-
2	Bu	Bu	Not Observed	31.3	1.596	1.572
3	OH	OH	177366 (49 h)	-	-	-
4	(CH <sub>3</sub> ) <sub>2</sub> C(CH <sub>2</sub> O) <sub>2</sub>		84695 (24 h)	36.8	1.712	1.537
5	EtO	EtO	77025 (21 h)	35.4	1.725	1.529
6	EtO	Oct	29446 (8.2 h)	35.1	1.646	1.559
7 <sup>a</sup>	(MeO) <sub>2</sub> P(S)H		26992 (7.5 h)	33.5	1.842	2.046
8	OH	Oct	19525 (5.4 h)	34.9	1.668	1.554
9	BnO	BnO	14648 (4.1 h)	-	-	-
10	EtO	Ph	5200 (1.4 h)	33.5	1.640	1.560
11	OH	Ph	3324 (55 min)	33.4	1.676	1.553
12	[(CH <sub>3</sub> ) <sub>2</sub> (CO)] <sub>2</sub>		882 ( 15.6 min)	33.1	1.717	1.537
13	Ph	Oct	596 (9.9 min)	30.8	1.615	1.570
14	PhO	PhO	315 (5.2 min)	34.4	1.749	1.538
15	Ph	Ph	150	30.8	1.625	1.569
16	H	OH	179	32.7	1.695	1.547
17 <sup>b</sup>	DOPO		259	31.6	1.682	1.548
	H	H	-	28.2	1.668	1.557
	CF <sub>3</sub>	CF <sub>3</sub>	-	24.2	1.745	1.527

<sup>a</sup>Dimethyl *H*-thiophosphonate. <sup>b</sup>DOPO = 6*H*-Dibenzo[*c,e*]oxaphosphinine 6-oxide

**Table 2.2:** B3LYP/6-311++G(3df,3pd) Tautomerization energy barrier ( $\Delta E^{\ddagger}$  kcal/mol) and TS P-H and P=O bond lengths (Å) for the phosphinylidenes in Figure 2.9 and two test cases.

This can be seen when we compare *H*-phosphonates (OH,OH; entry 3) and (EtO,EtO; entry 4), with half-lives of 49 hours and 21 hours respectively, to their counterpart *H*-phosphinates (EtO,Oct; entry 6), (OH,Oct; entry 8), (EtO,Ph; entry 10) and (OH,Ph; entry 11) with half-lives of 8.2 hours, 5.4 hours, 1.4 hours and 55 minutes respectively. This represents between a 3 and 50 fold increase in tautomerization rate between *H*-phosphonates and corresponding *H*-phosphinates. When we look at the mechanism that phosphinylidene tautomerization proceeds through, shown in Scheme 2.9, the electron donating/accepting character of the R<sup>1</sup> and R<sup>2</sup> substituents affect the stability of the P(V) resonance structure proceeding to the P(III) TS,

with electron accepting substituents destabilizing the  $P^+-O^-$  in the resonance structure. This explains the tautomerization rate increase observed for *H*-phosphinates versus *H*-phosphonates. Also displayed is increased reactivity for aryl- compared to alkyl-substituents. Comparing (Bu,Bu, entry 2), to (Ph,Oct; entry 13) and (Ph,Ph; entry 15), we see a decrease in half-life from no deuteration, to 9.9 minutes and 150 seconds respectively.

## 2.7 Computational Modelling of Phosphinylidene Tautomerization

Evaluation of basis sets was done by comparing to high accuracy calculations of  $H_3PO$  tautomerization, performed by Wesolowski and coworkers.<sup>54</sup> While  $H_3PO$  has not been isolated experimentally, the reference is a viable test, with calculations approaching the complete basis set limit (CBS). In evaluating coupled cluster methods, up to full single, double and triple excitations (CCSDT) and large basis sets (cc-pV6Z), with angular momentum up to  $i$  on phosphorus and oxygen and  $h$  on hydrogen, the authors concluded that very large basis sets would be required to accurately reproduce relative energies when modeling the transition state. Even CCSD(T) method with cc-pVQZ basis set did not return exact energy.

Basis Set	No. Basis Functions	$\Delta E$
6-31G(d)	40	-3.29
6-311++G(2d,2p)	96	-2.59
6-311++G(3df,3pd)	155	-0.35
cc-pV(Q+d)Z	255	0.31
cc-pV6Z <sup>54</sup>	557	1.22

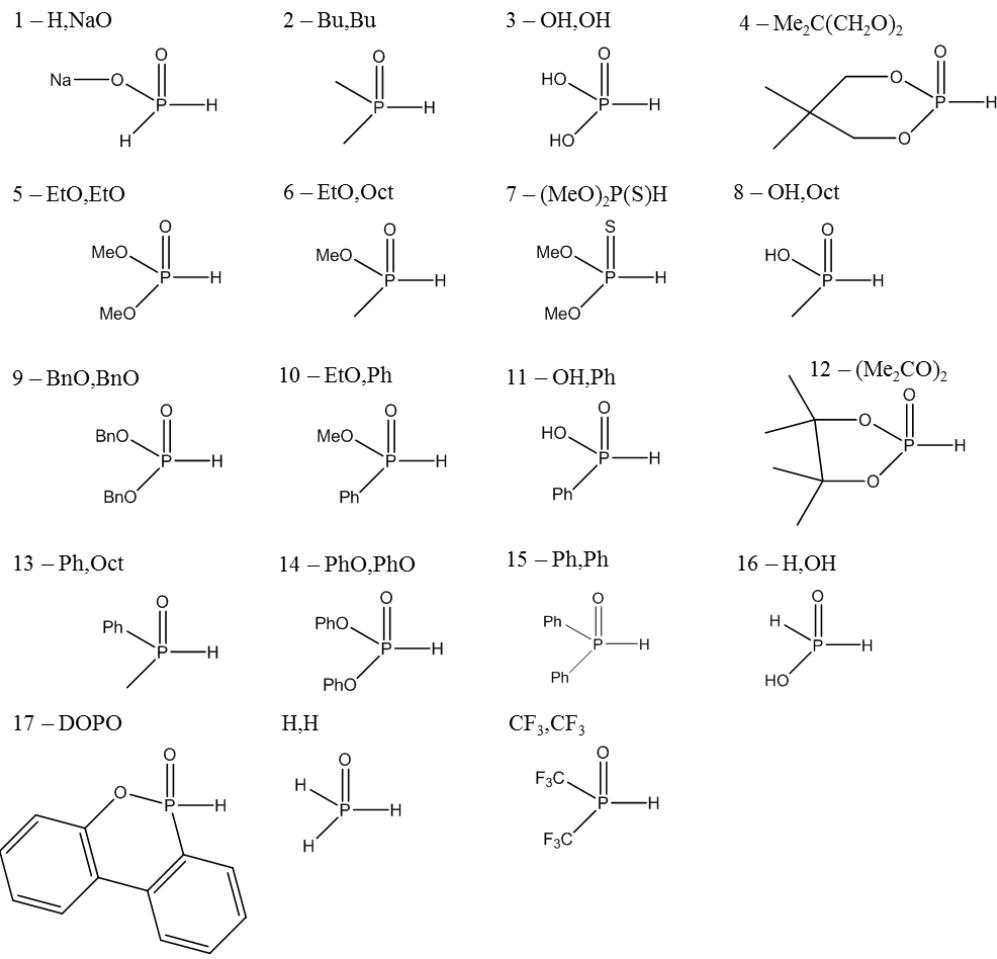
**Table 2.3:** Computed P(V) to P(III) tautomerization energy ( $\Delta E$  kcal/mol) from B3LYP

calculations on phosphine oxide with various basis sets.

In Table 2.3, we show that the commonly used B3LYP method with the smaller 6-311++G(3df,3pd) basis set returned energies within 1.6 kcal/mol of the reference. While the

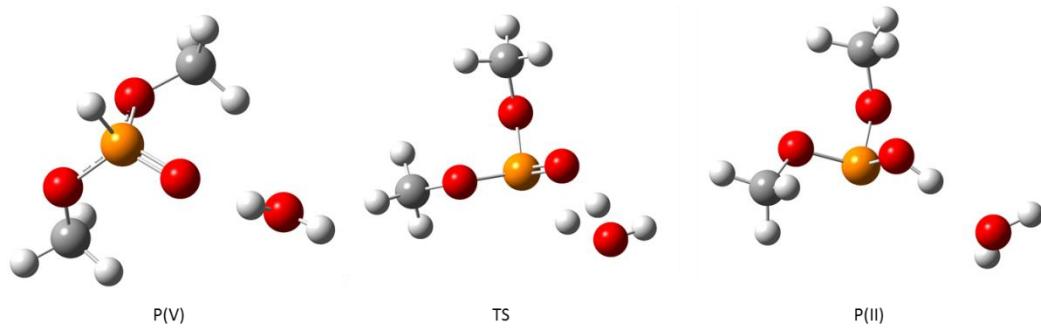
addition of *f*-polarization to the basis set for phosphorus may not be relevant for modeling the phosphine oxide reactant or phosphinous acid product, as noted by Wesolowski and coworkers, basis sets with higher angular momentum polarization were required for modeling the transition state. This approach represents a trade-off between accuracy and computational time. The identified method and basis set return results and energies that fit general trends and reactivity orders when compared to experiment.

The phosphinylidenes modeled computationally were simplified by replacing alkyl groups with methyl as shown in Figure 2.12. Two additional phosphinylidenes were considered computationally, with the inclusion of H<sub>2</sub>H and CF<sub>3</sub>CF<sub>3</sub> to model the effect of electron withdrawing effect on the P(V) to P(III) tautomerization.

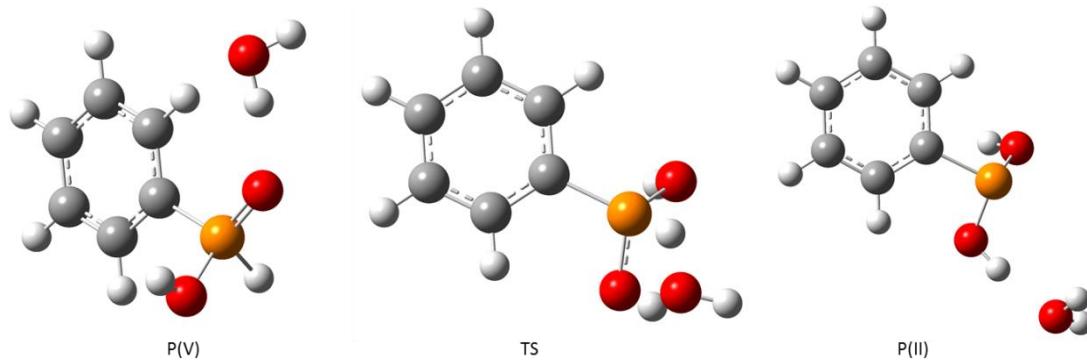


**Figure 2.12:** Structures of phosphinylidenes used in computational tautomerization study.

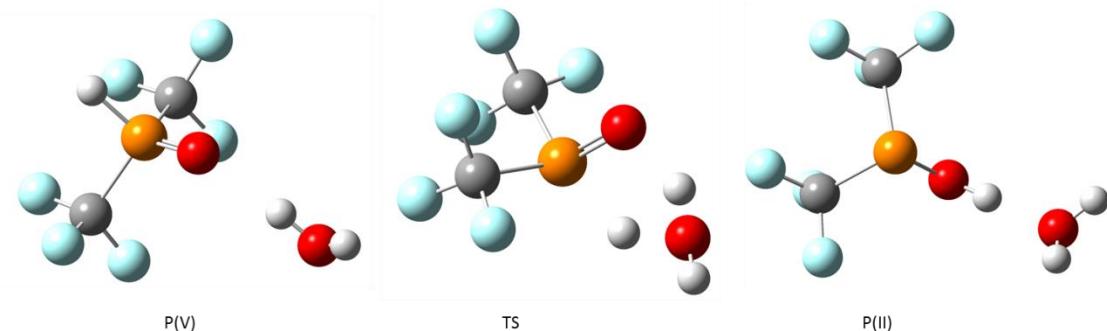
Representative computational reaction schemes for three P(V) to P(III) tautomerizations are shown in Figure 2.13, Figure 2.14 and Figure 2.15. Atoms are color coded, with C gray, H white, P orange, O red and F teal.



**Figure 2.13:** B3LYP/6-311++G(3df,3pd) optimized geometries of P(V), TS and P(III) for EtO<sub>2</sub>,EtO, Modeled as MeO,MeO.

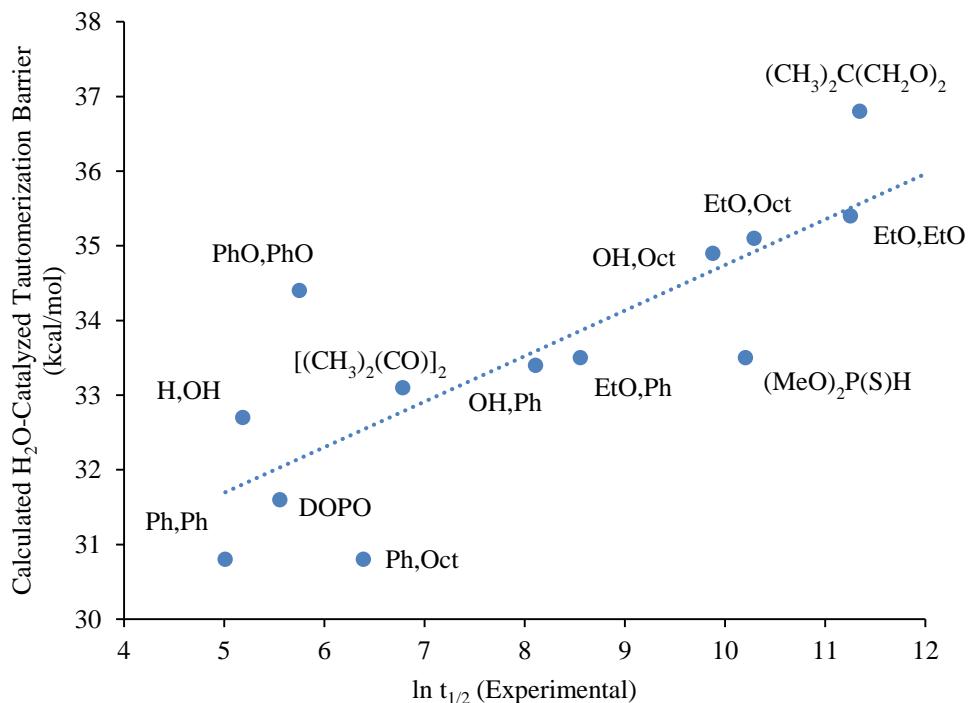


**Figure 2.14:** B3LYP/6-311++G(3df,3pd) optimized geometries of P(V), TS and P(III) for OH,Ph.



**Figure 2.15:** B3LYP/6-311++G(3df,3pd) optimized geometries of P(V), TS and P(III) for CF<sub>3</sub>,CF<sub>3</sub>.

We can see the correlation between the experimental half-life and the calculated tautomerization barrier, when we plot the barrier energy against the log of the experimental half-life, as shown in Figure 2.16. Again, with the experimental half-lives plotted on a logarithmic scale, returns a good correlation between the computed and experimental data, with  $r^2=0.65$ . This correlation suggests this is a valid approach for modeling phosphinylidene tautomerization and predicting general trends in reactivity. Looking at the data, we see two calculations that appear to be outliers. Those for PhO,PhO and Ph,Oct. In the case of PhO,PhO, the computational model does not fully account for resonance in the PhO substituent, with a larger electron-donating effect from oxygen than seen experimentally. In the case of Ph,Oct, the computational model overestimates the resonance stabilization of the lone pair with a phenyl substituent, as compared to experiment.



**Figure 2.16:** Calculated barriers for catalysis by a single explicit water molecule vs. experimental half-lives.

This relationship is expected from the relationship between barrier and half-life. The Arrhenius equation (Equation 2.4) gives the relationship between the rate constant ( $k$ ) and the activation energy ( $E_a$ ) in J/mol, where  $A$  is the pre-exponential factor,  $R$  is the universal gas constant and  $T$  is the absolute temperature in Kelvin.

$$k = A e^{-\frac{E_a}{RT}}$$

**Equation 2.4:** Arrhenius equation.

Taking the natural logarithm of the Arrhenius equation, gives the expression shown in Equation 2.5.

$$\ln(k) = \ln(A) - \frac{E_a}{R} \frac{1}{T}$$

**Equation 2.5:** Natural logarithm of the Arrhenius equation.

The Arrhenius equation is most widely used in calculating the activation energy of a reaction based on reaction rates at varying temperatures. Plotting  $\ln(k)$  against  $\frac{1}{T}$  yields a straight line, with gradient  $\frac{-E_a}{R}$  and y intercept  $\ln(A)$ . We calculated the Gibbs free energy of the transition rather than activation energy. The Eyring equation (Equation 2.6) shows the relationship between  $k$  and  $\Delta G^\ddagger$ , where  $k_B$  is the Boltzmann constant and  $h$  is Planck's constant.

$$k = \frac{k_B T}{h} e^{-\frac{\Delta G^\ddagger}{RT}}$$

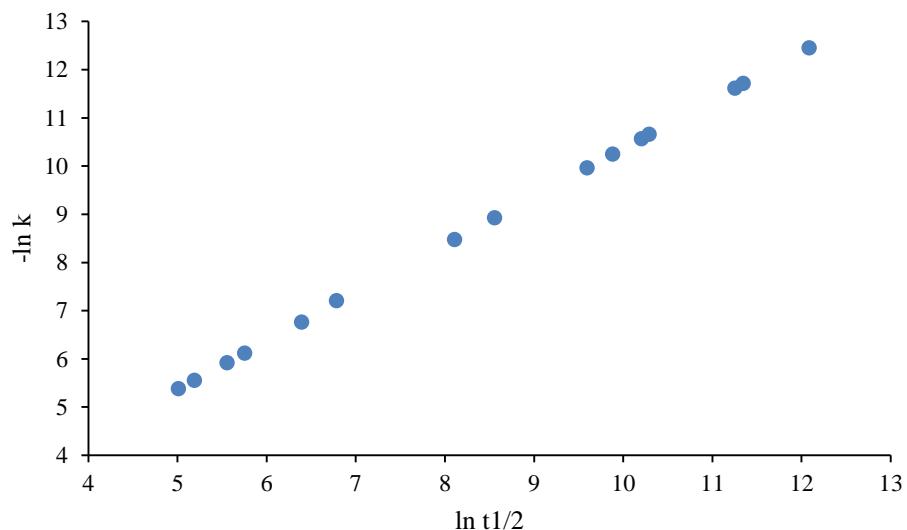
**Equation 2.6:** Eyring equation.

We can see again that taking the natural logarithm of the Eyring equation will again yield a relationship where  $-\ln(k)$  is proportional to  $\Delta G^\ddagger$ . The rate constant is related to the tautomerization half-life by the relationship shown in Equation 2.7.

$$t_{1/2} = \frac{\ln(2)}{k}$$

**Equation 2.7:** Relationship of half-life to rate constant.

From this relationship we can plot the natural logarithms from Table 2.2 as shown in Figure 2.17, since  $\ln t_{1/2}$  is proportional to  $-\ln(k)$ .



**Figure 2.17:** Plot of  $\ln t_{1/2}$  vs.  $-\ln(k)$ .

Having shown that  $-\ln(k)$  is proportional to  $\Delta G^\ddagger$  we can plot the barrier energy in place of  $-\ln(k)$  and expect a straight line when plotted against the natural logarithm of the half-life.

<b>Entry</b>	<b>R<sup>1</sup></b>	<b>R<sup>2</sup></b>	<b>ΔE (kcal/mol)</b>	<b>R(P-H)</b>	<b>R(P=O)</b>
1	H	NaO	24.7	1.439	1.499
2	Bu	Bu	7.8	1.417	1.481
3	OH	OH	10.0	1.393	1.465
4	EtO	EtO	6.4	1.393	1.463
5	(CH <sub>3</sub> ) <sub>2</sub> C(CH <sub>2</sub> O) <sub>2</sub>		10.2	1.396	1.466
6	EtO	Oct	10.2	1.413	1.472
7	(MeO) <sub>2</sub> P(S)H		9.8	1.399	1.923
8	OH	Oct	9.8	1.409	1.471
9	BnO	BnO	10.8	1.395	1.468
10	EtO	Ph	9.9	1.404	1.472
11	OH	Ph	9.2	1.404	1.471
12	[(CH <sub>3</sub> ) <sub>2</sub> (CO)] <sub>2</sub>		7.5	1.397	1.459
13	Ph	Oct	7.2	1.416	1.481
14	PhO	PhO	9.2	1.395	1.462
15	Ph	Ph	6.8	1.415	1.482
16	H	OH	7.0	1.404	1.468
17	DOPO		5.4	1.403	1.467
	H	H	-0.3	1.412	1.477
	CF <sub>3</sub>	CF <sub>3</sub>	-5.7	1.409	1.465

**Table 2.4:** B3LYP/6-311++G(3df,3pd) Tautomerization energy ( $\Delta E$  kcal/mol) and P(V) P-H and P=O bond lengths ( $\text{\AA}$ ) for the phosphinylidenes in Figure 2.9 and two test cases.

Table 2.4 shows the stability of the P(V) form, with positive values indicating the P(V) form is more stable than the P(III). The results are generally consistent with the experimental observations and calculated tautomerization energy barriers. The stabilities are in line with the electron donating/accepting principle, with those that are donating having a more stable P(V) form, compared to those that are more electron accepting. For this explanation, the computational prediction agrees with the hypothesis, where H<sub>2</sub>P(O)H does not have electron-donating or withdrawing substituents, while (CF<sub>3</sub>)<sub>2</sub>P(O)H, with two highly electron withdrawing groups strongly favors the P(III) form. Weakly donating and electron withdrawing substituents lead to the lowest calculated  $\Delta E$  between P(V) and P(III). It is likely

that these substituents stabilize the P(V) resonance structure shown in Scheme 2.9, leading to softer P-H bonds. This in turn makes breaking of the P-H bond more likely, leading to a more stable TS intermediate leading to the reactive P(III) form.

## 2.8 Concluding Remarks

The study demonstrated good correlation between experimental and computational results, as well as supporting previously documented trends in phosphinylidene  $R^1R^2P(O)H$  reactivity,<sup>51,55-59</sup> that generally correlate with the rate of prototropic tautomerism. Since tautomerization could not be measured directly, the initial rate of deuteration was determined. This approach works since deuteration proceeds via the P-H of the P(III) tautomer and is directly correlated with P(V) to P(III) tautomerization. By having D<sub>2</sub>O in excess, ensured that tautomerization and not deuteration was the rate limiting reaction. By comparing the initial rates of deuteration, a quantitative measure of relative phosphinylidene reactivity was able to be established. The ease with which phosphinylidene underwent water-catalyzed tautomerization was determined to be H<sub>3</sub>PO<sub>2</sub> > (Ph)<sub>2</sub>P(O)H > (PhO)<sub>2</sub>P(O)H > (Ph)(OAlk)P(O)H > (Alk)(OAlk)P(O)H ≈ (AlkO)<sub>2</sub>P(O)H > (Alk)<sub>2</sub>P(O)H. Tautomerization energy barriers, derived from computed single explicit water molecule catalyzed tautomerization transitions, calculated using the B3LYP method with the 6-311++G(3df,3pd) basis set captured most experimental trends. Based on the optimized reactant geometries, computational calculations suggest that R<sup>1</sup> and R<sup>2</sup> substituents also play a role, with weakly donating, and electron withdrawing groups that yield softer P(V) P-H bonds generally increasing tautomerization rates. Although exact tautomerization barrier energies were not returned, due to high energy differences associated with gas phase calculations, the study highlighted the promise of computation using a common DFT method and a mid-sized basis

set for predictive studies of phosphinylidene reactivity, increasing utility by reducing computational cost. Absolute accuracy of tautomerization energy barrier was sacrificed, while still returning usable trends that match experiment from an easily implemented computational approach.

## 2.9 Computational Section

All calculations use the Gaussian 09 electronic structure program,<sup>36</sup> generalized Kohn–Sham density functional theory<sup>60,61</sup> with the B3LYP hybrid exchange-correlation functional,<sup>62,63</sup> and Pople-type<sup>64</sup> or cc-pV(Q+d)Z<sup>65–67</sup> basis sets. Calculations carried out in a continuum solvent made use of the SMD model.<sup>68</sup> To decrease the number of conformations due to bond rotation and to minimize computational costs, methyl groups were used in place of alkyl chains. All reported geometries were fully optimized. Following geometry optimization, it was confirmed that the P(V) reactant and P(III) product had no negative eigenvalue in the vibrational Hessian for the local minima. Likewise, it was confirmed that there was one negative eigenvalue for the vibrational Hessian of the transition state. To confirm that the true transition state had been achieved, test calculations were performed on select TS optimized geometries, whereby the transition state geometry was displaced along the Hessian eigenvector with a negative eigenvalue. Doing this essentially makes the transition state more reactant like or more product like. Reoptimizing these geometries will converge to recover either the product or initial reactant geometry. H<sub>2</sub>O-catalyzed tautomerization barriers are evaluated as the difference between the energy of the optimized transition-state and the energy of free H<sub>2</sub>O and the free reactant P(V) phosphinylidene.

## Chapter 3: Catalysis of Phosphinylidene Tautomerization

### 3.1 Industrial Catalysis

The takeaway from our work on phosphinylidenes as an alternative feedstock for use in making new P-C bond containing compounds, is that it is plausible, although not yet practical. This is due to the high transition energy barrier. We saw that the uncatalyzed tautomerization for  $\text{H}_2\text{P}(\text{O})\text{H}$  was 59 kcal/mol. While water can act as a catalyst, the transition energy barrier was only reduced 22 kcal/mol to 37 kcal/mol. When we look at the effect of substituents on the water catalyzed transition barrier energy, we see an energy range of 30 to 37 kcal/mol. Although this is close to half the uncatalyzed barrier, it is not a sufficient reduction for a viable industrial approach. For the phosphinylidene approach to gain acceptance, it needs to show a number of advantages over the current manufacturing process. These may include: cost; yield; efficiency; reduction of side reactions; decreased production of toxic waste products. Therefore, we need to look at catalysis as another way to decrease the transition barrier energy to phosphinylidene tautomerization.

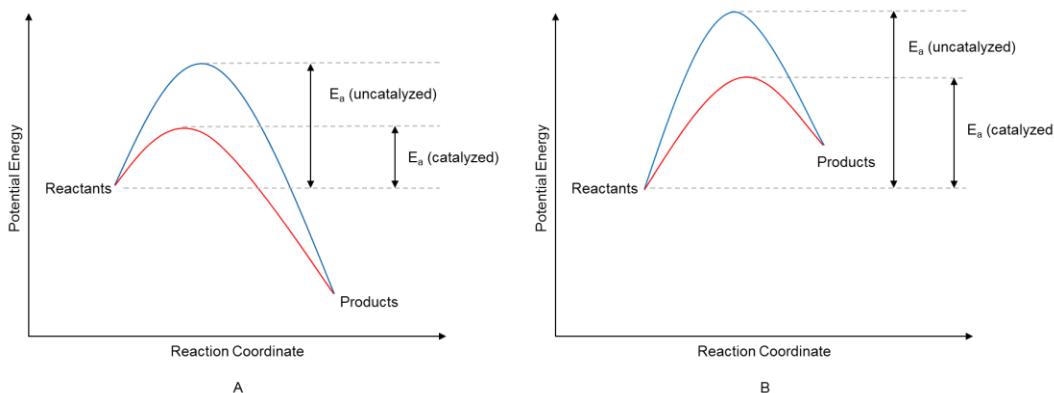
Catalysts, from chemistry first principles, increase the rate of a chemical reaction but are not consumed in the reaction but are released to act repeatedly. Reaction rate increases due to a decrease in activation energy. This does not affect the extent of reaction or the chemical equilibrium. For phosphinylidene tautomerization, a general catalytic reaction is shown in Scheme 3.1.



**Scheme 3.1:** General catalytic reaction for phosphinylidene tautomerization.

By decreasing the barrier energy associated with phosphinylidene tautomerization, we increase not only the forward reaction rate ( $P(V)$  to  $P(III)$ ) but also the back reaction ( $P(III)$  to  $P(V)$ ).

The generalized effect of catalysts on activation energy is shown in Figure 3.1.



**Figure 3.1:** General overview of lowering of activation energy by addition of a catalyst for  
A) exothermic or B) endothermic reaction.

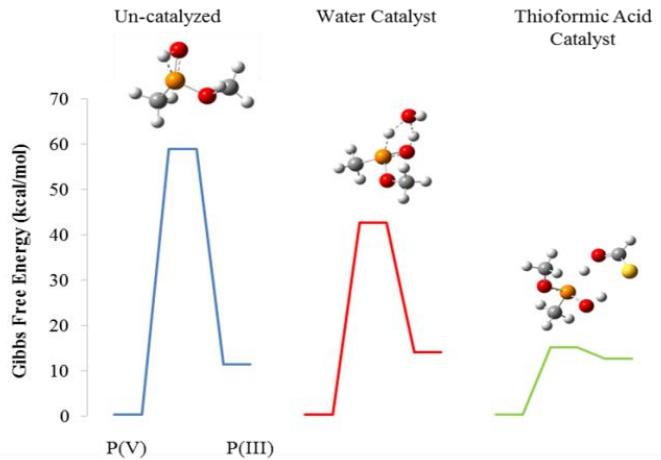
In order to understand other factors that can affect chemical equilibrium, we consider Le Chatelier's principle, where a system at equilibrium, subjected to change, will readjust to attempt to counteract the effect of the change, resulting in a new equilibrium. A number of factors can affect the equilibrium. Increasing the concentration of a reactant can push the equilibrium towards the product to counteract the change. The same effect can be achieved by removing product, or in the case of a reactive  $P(III)$  phosphinylidene tautomer, rapidly reacting the  $P(III)$  tautomer before it is able to undergo tautomerization back to the stable  $P(V)$  form. This may include ensuring that the second reactant is in excess so it is not rate limiting in the reaction with  $P(III)$ . Temperature can also move the equilibrium position. If the reaction is exothermic, increasing the reaction temperature would increase the heat in the system and cause the equilibrium to move towards the reactants to compensate. Phosphinylidene tautomerization requires an input of energy to overcome the tautomerization barrier, so

increasing the reaction temperature could shift the equilibrium towards the P(III) product. Of course, for an industrial process, there may be a number of reactions taking place, which is especially true in “one-pot” reactions. There is often no perfect answer to shifting equilibrium due to competing reactions. Reaction conditions are usually at a compromise where favorable equilibrium conditions for one reaction do not have too great a negative impact on a backwards equilibrium of a subsequent reaction.

### 3.2 Catalysis of Phosphinylidene Tautomerization

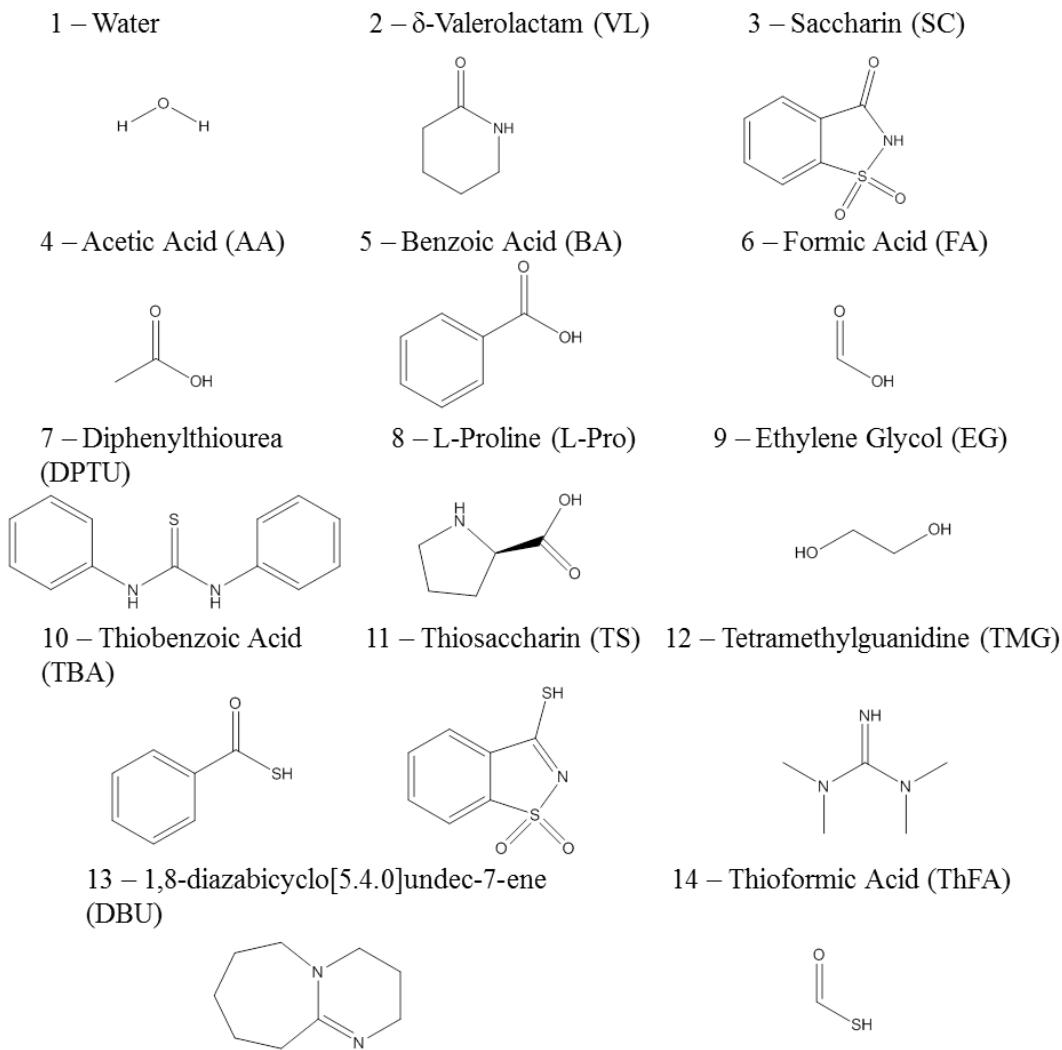
In chapter 2, we examined the potential for utilizing phosphinylidene tautomerization as the basis for the creation of P-C bonds, without the need for  $\text{PCl}_3$ . Although of potential utility, the uncatalyzed tautomerization activation energy is extremely high. Water was shown to act as a weak catalyst, that in conjunction with modifications to the phosphinylidene feedstock, showed promise for future investigation. In this chapter, we study another approach to improve P-C bond yields. That is to employ a catalyst in order to accelerate P(V) to P(III) tautomerization. This tautomerization appears to require catalysis.<sup>53</sup>

There have been several computational studies of how  $R_1$  and  $R_2$  affect the thermodynamics of phosphinylidene tautomerization.<sup>69–72</sup> However, relatively few works have considered tautomerization kinetics,<sup>54,71,73</sup> and we are unaware of a systematic study of tautomerization catalysts. Calculations in Figure 3.2 introduce the potential power of tautomerization catalysts.



**Figure 3.2:** Predicted energy surfaces for methyl methylphosphite tautomerization; uncatalyzed, catalyzed by a single water molecule, and catalyzed by representative hard acid-soft base molecule thioformic acid.

While tautomerization of the isolated gas-phase molecule is predicted to have a prohibitively high barrier, even a single water molecule significantly accelerates tautomerization by simultaneously donating a proton to oxygen and accepting a proton from phosphorus. Thioformic acid is predicted to further accelerate tautomerization, by donating a proton from the soft sulfur and accepting a proton at the hard oxygen.

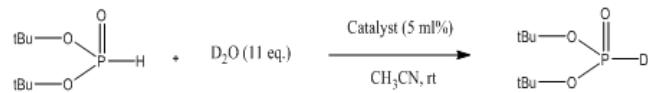
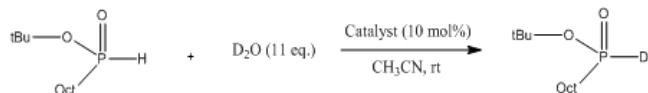


**Figure 3.3:** Tautomerization catalysts considered experimentally and computationally.

Here we present a systematic experimental and computational study of how fourteen small organic molecules (shown in Figure 3.3) catalyze phosphinylidene tautomerization. Our results suggest that Brønsted acid/base molecules,<sup>73–75</sup> particularly hard acid – soft base systems, show promise as practical phosphinylidene tautomerization catalysts. These results provide a potential new paradigm for designing more broadly applicable synthetic methods based on phosphinylidene feedstocks.

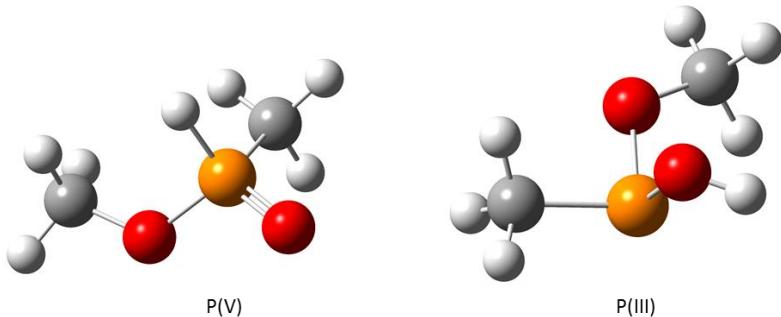
### 3.3 Catalyzed Phosphinylidene Tautomerization Rate

We experimentally measure the effects of various organic catalysts on phosphinylidene tautomerization, using the model reaction of alkyl H-phosphinate deuteration in the presence of excess D<sub>2</sub>O (Scheme 3.2). We consider the organocatalysts shown in Figure 3.3. Thioformic acid is not stable at the experimental conditions used in this study, so is only considered computationally. As discussed in Chapter 2, we assume that deuteration of the phosphorus proceeds through tautomerization. The experimental reaction rate is monitored through phosphorus <sup>31</sup>P NMR. Experiments consider the model phosphinylidene molecule *tert*-butyl octylphosphinate, as its un-catalyzed tautomerization half-life of ~8 hours reduced the amount of instrument time required. The *tert*-butyl ester was selected as a protecting group to protect against de-alkylation typically observed in the presence of base/nucleophile.

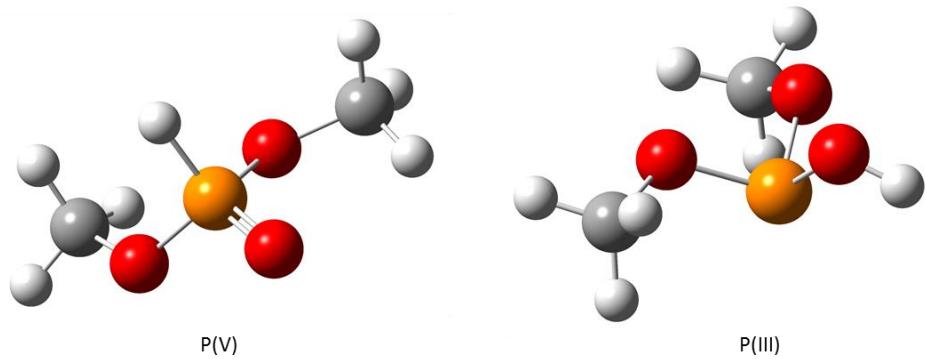


**Scheme 3.2:** Experimental probes of phosphinylidene tautomerizaton kinetics: Deuteration of *tert*-butyl octylphosphinate (Scheme 1a) or di-*tert*-butyl phosphite (Scheme 1b), in the presence of excess D<sub>2</sub>O and 5-10 mol% catalyst.

The *tert*-butyl octylphosphinate and di-*tert*-butyl phosphite are simplified for computational modeling, by the substitution of alkyl groups for methyl. The Models are shown in Figure 3.4 and Figure 3.5. Atoms are color coded, with C gray, H white, P orange and O red.

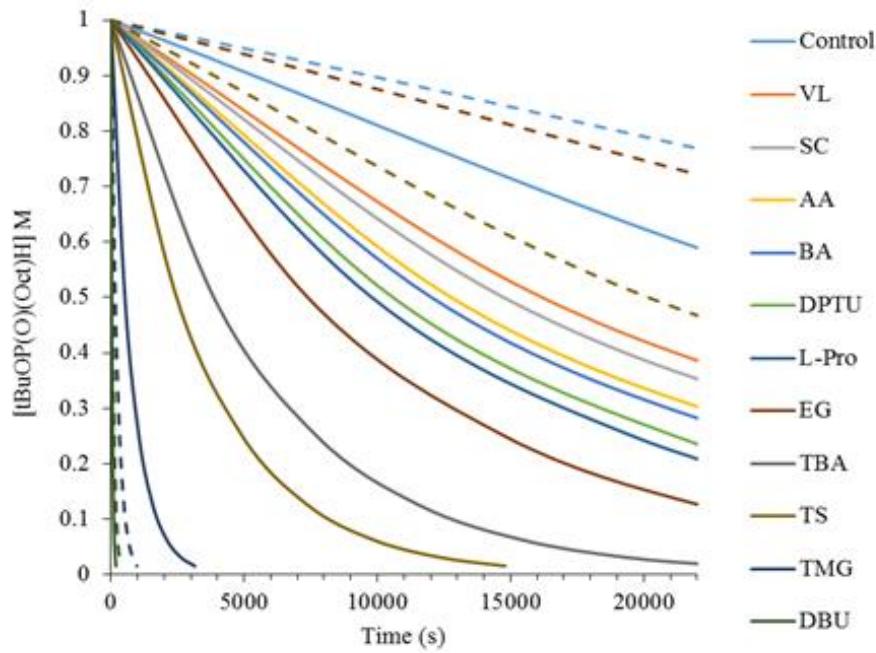


**Figure 3.4:** B3LYP/6-311++G(3df,3pd) optimized geometries of P(V) and P(III) for OMe,Me.



**Figure 3.5:** B3LYP/6-311++G(3df,3pd) optimized geometries of P(V) and P(III) for OMe,OMe.

We compute predicted tautomerization energy barriers of the tautomerization catalysts in Figure 3.3, using standard density functional theory (DFT) calculations. Data collected from  $^{31}\text{P}$  NMR kinetics experiments are presented in Figure 3.6 and summarized in Table 3.1 and Table 3.2 for reactions with *tert*-butyl octylphosphinate and di *tert*-butyl phosphite respectively.



**Figure 3.6:** Decay of *tert*-butyl octylphosphinate (1M in  $\text{CH}_3\text{CN}$ ) during deuteration with  $\text{D}_2\text{O}$  catalyzed by various organic compounds, with decay of di *tert*-butyl phosphite shown with dashed lines. No additive was used in the control experiment.

Entry	Catalyst <sup>a</sup>	Half-Life (s)	ln (Half-Life)	$\Delta G^\ddagger$ (kcal/mol)
1	Control	27365 (7.6h)	10.2	44.2
2	VL	16241 (4.5h)	9.7	31.1
3	SC	14817 (4.1h)	9.6	30.9
4	AA	13163 (3.7h)	9.5	21.7
5	BA	11229 (3.3h)	9.4	21.4
6	FA	10800 (3.0h)	9.3	20.7
7	DPTU	10583 (2.9h)	9.3	-
8	L-Pro	10378 (2.9h)	9.2	21.5
9	EG	7561 (2.1h)	8.9	-
10	TBA	4073 (1.1h)	8.3	24.6
11	TS	2803	7.8	20.4
12	TMG	121	6.3	29.9
13	DBU	41	3.5	22.5
14	ThFA	-	-	23.2

**Table 3.1:** Kinetics data for decay of *tert*-butyl octylphosphinate after addition of D<sub>2</sub>O using various catalysts and the corresponding computed tautomerization barriers.

Entry	Catalyst <sup>a</sup>	Half-Life (s)	ln (Half-Life)	$\Delta G^\ddagger$ (kcal/mol)
1	Control	50484 (14h)	10.8	47.2
2	EG	49405 (13.7h)	10.6	-
3	TS	19438 (5.4h)	9.9	26.7
4	TMG	162	5.1	24.2
5	DBU	50	4.2	15.3

**Table 3.2:** Kinetics data for decay of di *tert*-butyl phosphite after addition of D<sub>2</sub>O using various catalysts and the corresponding computed tautomerization barriers.

The two best organocatalysts were experimentally determined to be the strong bases 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU; entry 13) and tetramethylguanidine (TMG; entry 12) with deuteration rates approximately 700 and 200 times faster than the water-catalyzed control, respectively. However, strong bases such as DBU and TMG could be potentially problematic for real experiments due to increased side reactions. We find that the hard acid –

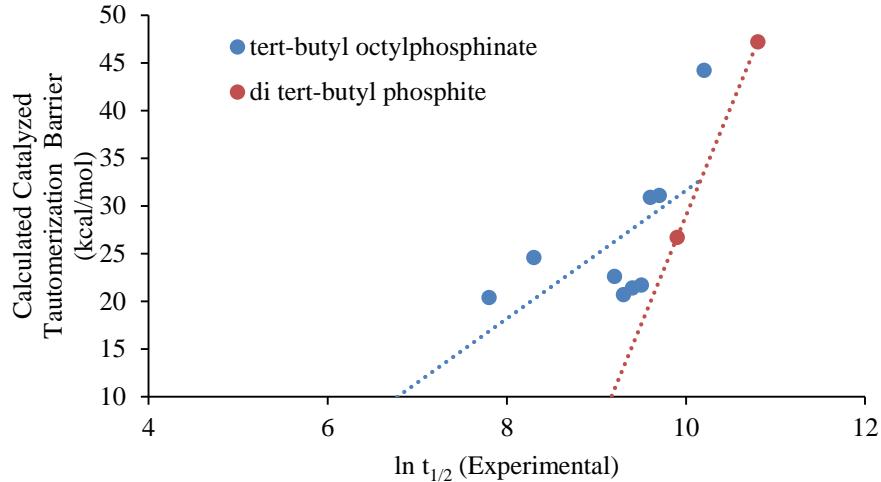
soft base molecule thiosaccharin (TS; entry 11) had the third-highest tautomerization rate with tautomerization 10 times faster than the water-catalyzed control, providing the best experimental choice of a mild tautomerization reagent. Saccharin (SC; entry 3) was found to be much less reactive in comparison, with a calculated rate only 2 times faster than the control. A similar trend was experimentally observed for thiobenzoic acid (TBA; entry 10) and benzoic acid (BA; entry 5), clearly demonstrating the importance of hard acid – soft base effects in an experimentally effective catalyst. Ethylene glycol (EG; entry 9) yielded a rate 3.6 times faster than the control. This is consistent with previous work on palladium–catalyzed cross coupling and hydrophosphinylation reactions,<sup>76,77</sup> where an increase in reactivity of *H*-phosphinates was noted when using ethylene glycol. Formic acid (FA; entry 6), diphenylthiourea (DPTU; entry 7) and L-proline (L-pro; entry 8) all performed similarly, with tautomerization rates approximately 2.5 times faster than the control. Acetic acid (AA; entry 4) had a rate 2 times faster than with water catalyst. Cyclic amides δ-valerolactam (VL; entry 2) and saccharin (SC; entry 3), both produced similar results, with calculated rates 1.7 and 1.8 times faster than the control respectively.

To determine if the catalytic trends are the same with a different phosphinylidene, we measured the deuteration of di-*tert*-butyl phosphite (Scheme 1b). Table 3.2 shows that di-*tert*-butylphosphite has a slower uncatalyzed deuteration rate, consistent with the stabilization of the P(V) reactant by the additional alkyl ether, as previously discussed in Chapter 2. Bases DBU and TMG are again the best tested organocatalysts. TS (entry 11) was less efficient as a catalyst, with only a 2.6 times increase in rate. However, we again note the potential of TS as a mild reagent for phosphinates with ester groups that react with strong bases. In contrast to

the tert-butyl octylphosphinate, EG (entry 9) did not significantly accelerate deuteration of di-*tert*-butyl phosphinate.

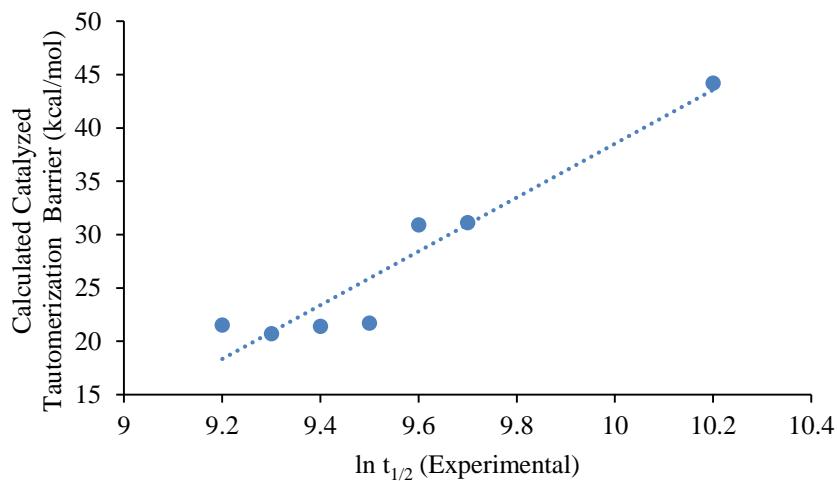
### 3.4 Calculating the Catalyzed Tautomerization Barriers

Table 3.1 and Table 3.2 also report the tautomerization barriers computed with standard DFT calculations. The computed tautomerization barriers show a strong correlation with the experimental reaction rates. These results clearly indicate the potential of simulations in designing new orgaocatalysts. The calculations clearly capture the general experimental trends for catalysis of the tautomerization of *tert*-butyl octylphosphinate and di-*tert*-butyl phosphite. The two outliers are the predicted high tautomerization barrier for the strong bases TMG and DBU. The computational model also returned slightly higher than expected barriers for TBA and ThFA. These disagreements between experiment and simulations may arise from multiple mechanisms involved in the tautomerization catalysis due to the catalysts' dual acid/base functionality. Solubility effects and the formation of the initial complex to P(V) may also play a role. The latter effect is discussed in greater detail below. Interestingly, the experimental trend for reactivity with strong bases was replicated reasonably well for *di* *tert*-butyl phosphite but not for *tert*-butyl phosphinate. Outside of those results, overall computational trends were captured by the computational model. As discussed in Chapter 2, *di* *tert*-butyl phosphite has a slower deuteration rate than *tert*-butyl phosphinate and this effect was again backed up computationally. Experimental and computational data are shown in Figure 3.7. DPTU and EG data are not shown as these did not optimize to a stable transition state. Calculated data for TMG and DBU are not shown in the graph below. There is agreement between the general trends of the calculated transition state energies and the experimental half-lives, for both *tert*-butyl octylphosphinate and di-*tert*-butyl phosphite, as shown in Figure 3.7.



**Figure 3.7:** Calculated catalyzed barriers vs.  $\ln t_{1/2}$  (experimental).

The calculated barrier energies capture the general experimental trends and, as expected, predict a slower tautomerization rate for the di tert-butyl phosphite compared to the tert-butyl octylphosphinate. Looking at a smaller subset of catalysts (Water, VL, SC, AA, BA, FA and L-Pro) with tert-butyl octylphosphinate, we see very good agreement with the experimental data, as shown in Figure 3.8.

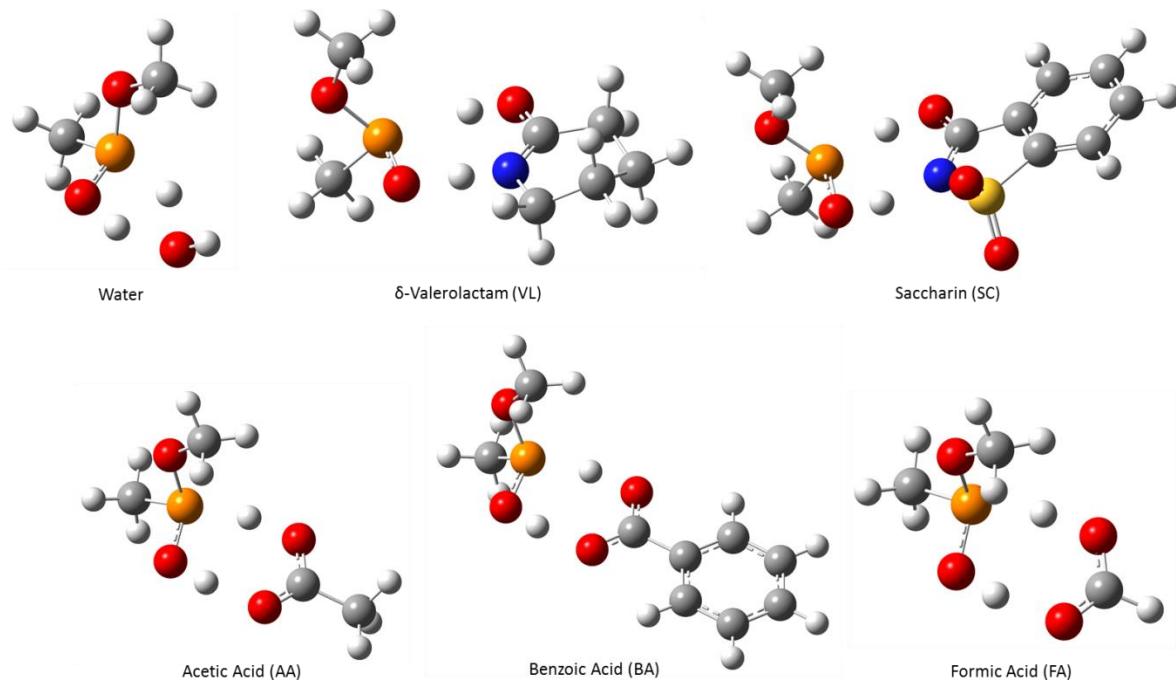


**Figure 3.8:** Calculated catalyzed barriers vs.  $\ln t_{1/2}$  (experimental) for tert-butyl octylphosphinate with a smaller subset of catalysts (Water, VL, SC, AA, BA, FA and L-Pro).

This correlation further validates the approach and the computational model's ability to reproduce experimental trends and suggests that geometry and binding refinement of the other catalysts may further improve the agreement with experiment.

L-proline (Entry 8, L-Pro), with an experimental half-life of 2.9 hours and a calculated barrier of 21.5 kcal/mol may be expected to have a lower barrier energy based on experimental rate data. This is likely due to the fact l-proline has a  $pK_a \sim 2$  and exists as a zwitterion. Modeling of the deprotonated carboxylic acid is not trivial, as will be discussed in Chapter 4. The anionic portion is modeled with three explicit waters to minimize the energy. This is more computationally demanding and is starting to move beyond the simple approach we used for the rest of the study. Future work could look at the interaction of phosphinylidenes with zwitterionic species and further refine the model with explicit solvent.

The data in Table 3.1 and Table 3.2 clearly suggest a general principle, that hard acid – soft base molecules provide improved tautomerization kinetics, as theorized in Chapter 2. The computed tautomerization barrier in the presence of TS is significantly lower than SC. Unexpectedly, this was not observed for ThFA and TBA over FA and BA. As mentioned, refinement of the binding should be expected to recover these lower tautomerization barriers given the agreement of the other computational data. This trend has led to the aforementioned experimental identification of thiosaccharin as a relatively mild, yet effective, experimental tautomerization catalyst. Representative catalyzed transition states are shown for select organic molecules in Figure 3.9. Atoms are color coded, with C gray, H white, P orange, O red and N blue.

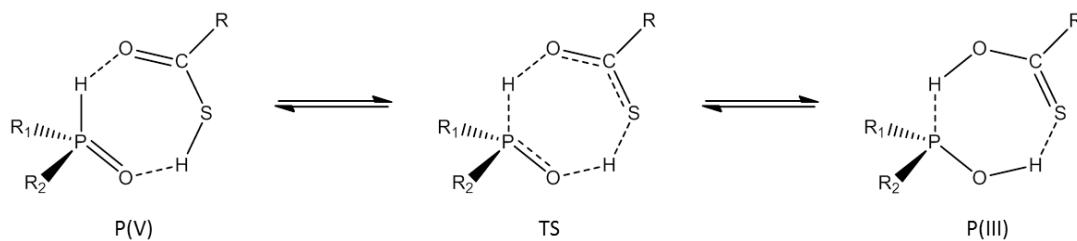


**Figure 3.9:** B3LYP/6-311+G(3df,3pd) optimized geometries of representative organocatalyzed OMe,Me P(V) to P(III) tautomerization transition states.

In order to better understand what makes a catalyst efficient, we focus on simulations of the critical initial step: catalyst complexation with the P(V) reactant. As previously discussed in Chapter 1 and described in Scheme 2.6, it is known that strong bases form the reactive P(III) tautomer. The strong bases DBU and TMG appear to simply stabilize the direct transfer of a proton from phosphorus to oxygen. This may be why the barrier energies were not reproduced by calculation. The transition state of the DBU catalyzed tautomerization appears to go through a strained 3-membered ring that is thermodynamically unfavorable, as discussed in Chapter 2. In this case, further modeling will be carried out to investigate the deprotonation of the phosphinylidene to produce the P(III) tautomer. The other catalysts transfer protons to the P=O bond and allow the hard or soft base in the molecule to bind to the proton transferring from the phosphorus to the oxygen. TS and other thiolated

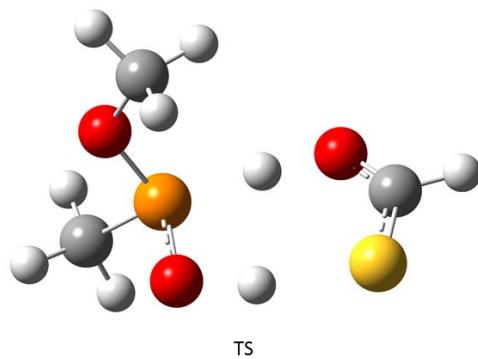
molecules, such as TBA, are expected to form a stable complex with the phosphinate. There are two possible tautomer structures for these catalysts, with the most likely having the acidic hydrogen on the sulfur, with the lone pair on the nitrogen and oxygen for TS and TBA respectively, with these bases stabilizing the transferring P-H hydrogen. This fact has been well documented.<sup>78</sup> Ethylene glycol appears to work as a decent experimental catalyst but calculations did not return a stable transition state.

Another point to make regarding the catalysis of phosphinylidene tautomerization is that for catalysts investigated in both a thiolated and an unthiolated form, we see a much faster tautomerization rate, and subsequent lower energy barrier, for the thiolated form. In Table 3.1, we see saccharin (Entry 3, SC) with an experimental half-life of 4.1 hours and calculated barrier energy of 30.9 kcal/mol has a much slower tautomerization rate than the corresponding thiosaccharin (Entry 11, TS) with an experimental half-life of 41 minutes and calculated barrier energy of 20.4 kcal/mol. The trend is repeated experimentally with benzoic acid (Entry 5, BA) with an experimental half-life of 3.3 hours and the corresponding thiobenzoic acid (Entry 10, TBA) with an experimental half-life of 1.1 hours. This is to be expected based on literature. The corresponding thiolated acid is more acidic than the carboxylic acid. Nguyen and co-workers<sup>79</sup> found that the thiol C(O)SH was the more stable form compared to the tautomeric thione C(S)OH. The S-H has a lower pK<sub>a</sub> than corresponding carboxylic acids and can be up to 10 times more acidic. This would increase phosphinylidene tautomerization rate by increasing the rate of transfer of hydrogen from S-H to the P(V) P=O and then increasing the transfer of hydrogen from P(V) P-H to the thiol C=O, stabilizing the transition state, as shown in Figure 3.10.



**Figure 3.10:** Interaction of phosphinylidene with thiol.

For comparison, the optimized computational geometry of the thioformic acid catalyzed transition state is shown in Figure 3.11. Atoms are color coded, with C gray, H white, P orange, O red and S yellow.



**Figure 3.11:** B3LYP/6-311+G(3df,3pd) optimized geometry of thioformic acid catalyzed OMe,Me P(V) to P(III) tautomerization transition state.

### 3.5 Concluding Remarks

The study presented in this chapter highlights the synergistic computational and experimental studies used to understand the catalyzed P(V) to P(III) tautomerization. The computationally predicted energy barrier data mostly followed the same trend as the experimentally calculated half-lives determined by  $^{31}\text{P}$  NMR. We demonstrate how computational chemistry, in collaboration with an experimental approach, may be used to help understand and develop efficient tautomerization catalysts via calculating the transition

state energies of phosphinylidene-catalyst complexes. Strong bases such as DBU and tetramethylguanidine provide the fastest experimental rates of tautomerization, although they cannot be employed for general catalytic reactions for phosphorus compounds with base-labile groups. Computational studies identified neutral thiosaccharin as a promising mild tautomerization catalyst due to favorable acid/base interactions with the phosphorus tautomers.

Expanded experimental and computational investigations are needed to understand how ethylene glycol works as a catalyst. Further refinement and investigation is warranted to examine the reason the thiolated catalysts did not return the lower tautomerization energy barriers that would be expected based on experimental data and on hard/soft acid/base understanding and to further explore the interaction of the phosphinylidenes with the strong bases.

### 3.6 Computational Section

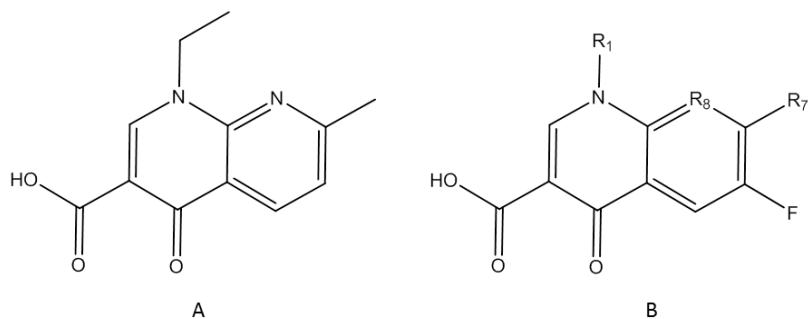
All calculations use the Gaussian 09 electronic structure program,<sup>36</sup> generalized Kohn–Sham density functional theory<sup>60,61</sup> with the B3LYP hybrid exchange-correlation functional,<sup>62,63</sup> and Pople-type<sup>64</sup> 6-311++G(3df,3pd) basis set. Calculations were carried out in a continuum solvent and made use of the SMD model.<sup>68</sup> To decrease the number of conformations due to bond rotation and to minimize computational costs, methyl groups were used in place of alkyl chains. There were no constraints applied to molecular symmetry and all reported geometries were fully optimized. Following geometry optimization, it was confirmed that the P(V) reactant and P(III) product had no negative eigenvalue in the vibrational Hessian for the local minima. Likewise, it was confirmed that there was one negative eigenvalue for the vibrational Hessian of the transition state. To confirm that the true transition state had been achieved, test calculations were performed on select TS optimized geometries, whereby the

transition state geometry was displaced along the Hessian eigenvector with a negative eigenvalue. Doing this essentially makes the transition state more reactant like or more product like. Reoptimizing these geometries will converge to recover either the product or initial reactant geometry. The catalyzed tautomerization barriers are evaluated as the difference between the energy of the optimized transition-state and the energy of free catalyst and the free reactant P(V) phosphinylidene.

## Chapter 4: Fluoroquinolone Binding to Mg(H<sub>2</sub>O)<sub>n</sub><sup>2+</sup> and its Applicability to Future Drug Design

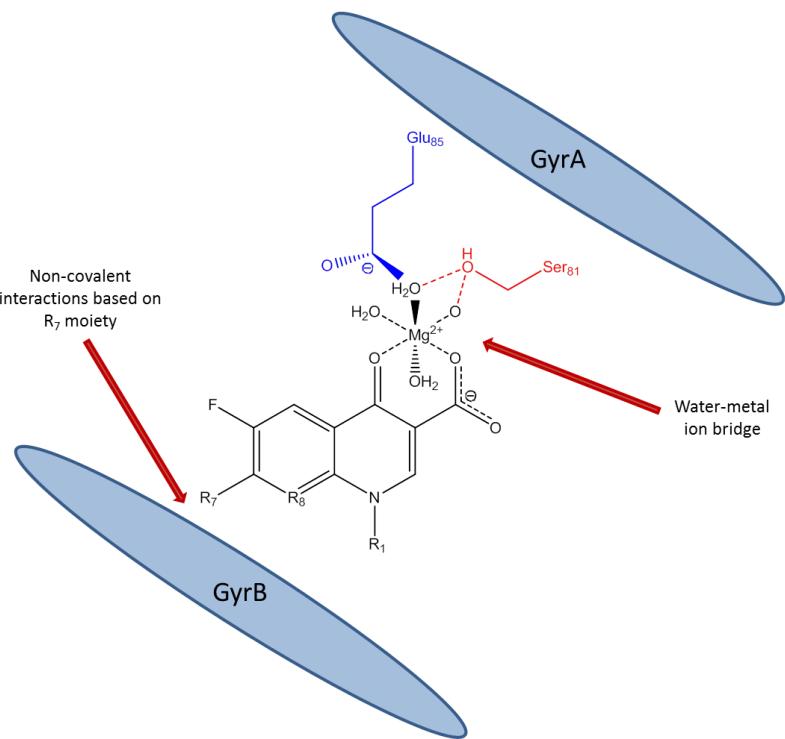
### 4.1 Fluoroquinolone Chemistry and the Significance of Mg<sup>2+</sup> Binding

Fluoroquinolones are a class of synthetic, broad spectrum, antibiotic drugs based on nalidixic acid and derived from the quinolone pharmacophore by fluorination at carbon 6, as shown in Figure 4.1.



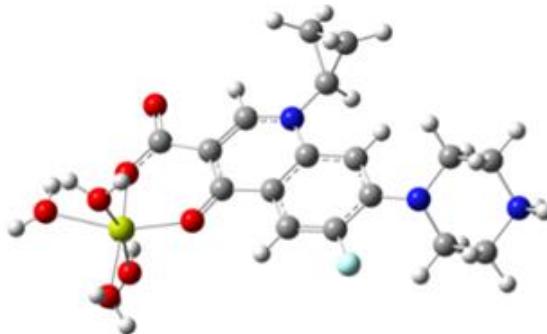
**Figure 4.1:** A) Structure of nalidixic acid and B) essential fluoroquinolone structure.

Quinolones have antibacterial activity but are associated with some adverse reactions.<sup>80,81</sup> Many fluoroquinolone drugs have been synthesized, showing varying effectiveness against gram-negative and gram-positive bacteria.<sup>82,83</sup> Quinolones target the ligase domain of bacterial type II topoisomerase (DNA gyrase or topoisomerase IV).<sup>84</sup> A drug-enzyme complex is formed, resulting in bacteriostasis.



**Figure 4.2:** Diagram of proposed water-metal ion bridge that mediates critical interactions between quinolones and topoisomerase IV.

The currently accepted model has binding mediated by magnesium ions Mg<sup>2+</sup> through a water-metal ion bridge<sup>85–87</sup> as shown in Figure 4.2. Fluoroquinolones can also bind Mg<sup>2+</sup> outside of the cell. This extracellular interaction is undesirable as it can decrease drug bioavailability, observed as a decrease in efficacy.<sup>88–90</sup> With the increase in resistant strains of bacteria,<sup>91–96</sup> there has been continued activity, focused on synthesizing new fluoroquinolones and derivatives. Since binding to magnesium has been shown to be required for antibiotic efficacy, this motivates investigation into the development of a screening tool for predicting binding of new fluoroquinolones and the systematic study into how fluoroquinolone derivatization affects Mg<sup>2+</sup> binding (Figure 4.3) and biological activity.



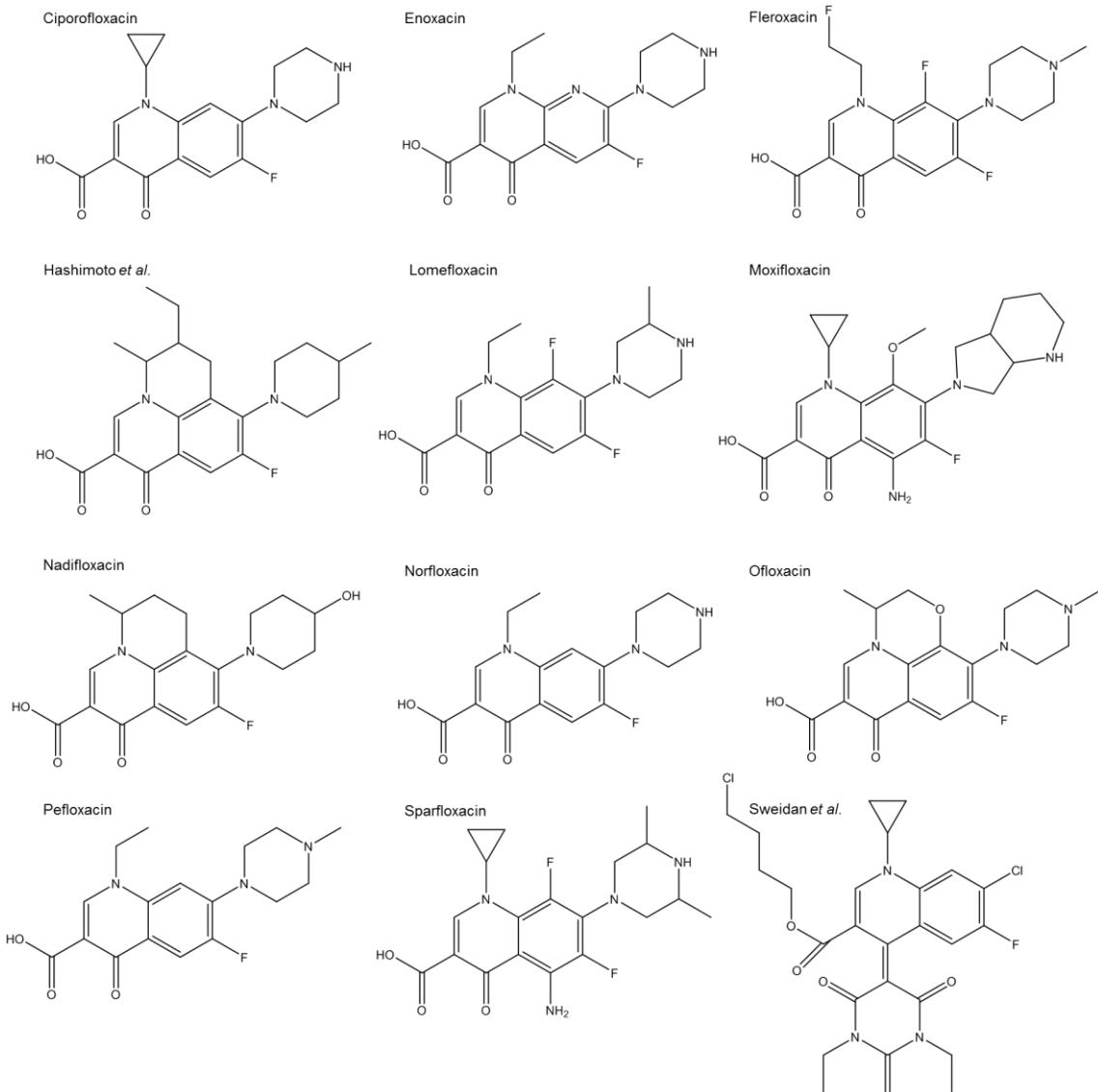
**Figure 4.3:** Binding of ciprofloxacin zwitterion to hydrated magnesium ion, geometry optimized using B3LYP method with 6-311+G(2d,2p) basis set (atoms are color coded, with C gray, H white, Mg yellow, O red, N blue and F teal).

The  $Mg^{2+}$  binding to fluoroquinolones has traditionally been studied by observing changes in antibiotic efficacy<sup>97</sup> in the presence of varying concentrations of magnesium. However, there are few systematic experimental or computational studies available in the open literature that consider the magnesium binding directly. There have been computational studies interested mainly with rationalizing fluoroquinolone phototoxicity, such as the study into the phototoxicity of lomefloxacin by Freccero and coworkers.<sup>98</sup> Some other examples include Vitorino and coworkers<sup>99</sup> had interest in modeling the hydration of neutral and zwitterionic norfloxacin, to identify the predominant species that may be expected in neutral solution. The study came to the conclusion that in fact both protonation states may coexist in solution. Toro-Labbe and coworkers<sup>100</sup> characterized the protonation site of nalidixic acid, while protonation and the photochemistry of norfloxacin was studied computationally by Musa and coworkers.<sup>101</sup> Meanwhile Xiao-Hong Li and coworkers<sup>102,103</sup> have published several studies on structure-activity relationships of a number of fluoroquinolones.

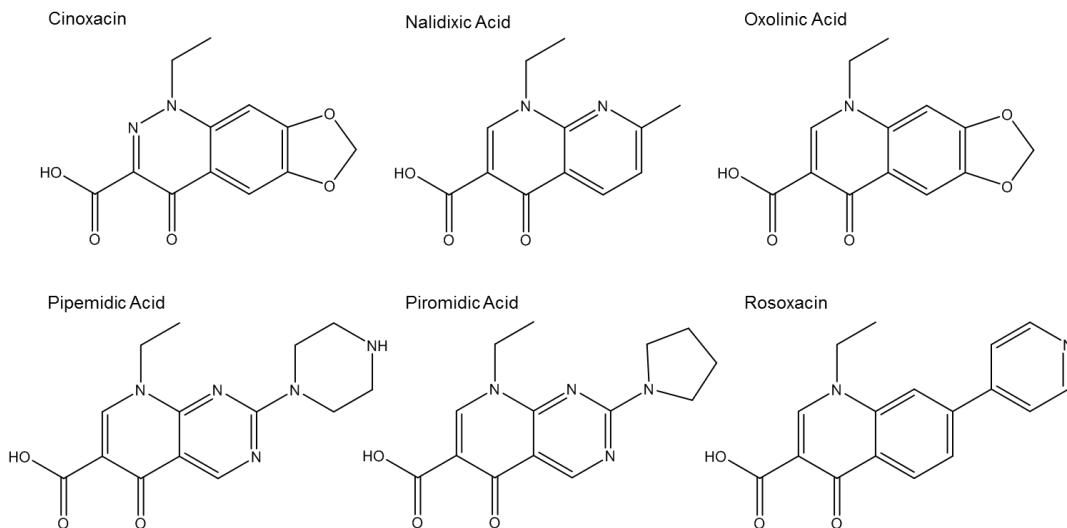
The study described here attempts to address the lack of systematic studies of fluoroquinolones' Mg<sup>2+</sup> binding. We benchmark several computational methods against experimental crystal structures of fluoroquinolones and magnesium complexes. We use the best identified method and basis set to compute Mg<sup>2+</sup> affinities, and compare these to available experimental data. We also explore the Mg<sup>2+</sup> affinities' dependence on fluoroquinolone protonation state. The results give insight into how fluoroquinolone derivatization affects pH-dependent Mg<sup>2+</sup> binding affinities. Our results should aid ongoing work in optimizing fluoroquinolone-magnesium and fluoroquinolone-target interactions.

## 4.2 Fluoroquinolone Geometry Benchmarks

We first benchmark semiempirical and DFT geometry approximations to 12 fluoroquinolone and 6 quinolone experimental XRD crystal structures<sup>104–121</sup> for the molecules shown in Figure 4.4 and Figure 4.5 respectively. Hydrogen atoms were added in GaussView 5.0, as the experimental hydrogen atom positions were not reported. Incidentally, the bonds to hydrogen were not considered when evaluating the errors in bond lengths and bond angles. The reported XRD precision is reported to be 0.001 to 0.01 Å, which equates to 0.1 to 1 pm. Therefore, we would expect errors of this order in bond length. Bond length errors below 0.1 pm would be below the XRD instrument precision and of an accuracy not able to be confirmed experimentally.



**Figure 4.4:** Fluoroquinolones used in benchmark set.



**Figure 4.5:** Quinolones used in benchmark set.

Our first benchmarking study evaluated geometry errors when carrying out Hartree-Fock and DFT calculations with various Pople-type basis sets. The mean error (ME), mean absolute error (MAE), maximum error (MAX) and minimum error (MIN) results for bond lengths and bond angles are summarized in Table 4.1, Table 4.2, Table 4.3, and Table 4.4.

<b>Basis Set</b>	<b>n</b>	<b>Job CPU Time (hours)</b>	<b>Bond lengths (pm)</b>			
			<b>ME</b>	<b>MAE</b>	<b>MAX</b>	<b>MIN</b>
cc-pVTZ	972	110:51:04	0.5	1.8	11.8	-11.5
6-311+G(2d,2p)	810	61:12:30	0.5	1.8	11.8	-11.5
6-311+G(d,p)	636	31:20:37	0.2	1.8	11.7	-11.7
6-31++G(d,p)	564	22:26:00	0.1	1.7	11.0	-11.7
6-31+G(d,p)	546	15:08:11	0.1	1.7	11.0	-11.7
cc-pVDZ	426	8:10:03	0.1	1.8	11.4	-11.7
6-31G(d)	396	1:51:10	0.1	1.8	11.2	-11.8
3-21G(d)	252	0:44:51	-0.3	1.6	9.6	-12.4
STO-3G	138	0:09:48	-2.3	3.2	11.0	-14.7

**Table 4.1:** Statistical errors in Hartree-Fock bond lengths for the computed geometries of the fluoroquinolones in Figure 4.4, evaluated with various basis sets (n is the number of basis functions used to describe ciprofloxacin).

Basis Set	n	Job CPU Time (hours)	Bond Angles (degrees)			
			ME	MAE	MAX	MIN
cc-pVTZ	972	110:51:04	-0.05	1.09	5.19	-7.73
6-311+G(2d,2p)	810	61:12:30	-0.04	1.08	5.13	-7.64
6-311+G(d,p)	636	31:20:37	-0.04	1.09	5.18	-7.69
6-31++G(d,p)	564	22:26:00	-0.03	1.08	5.28	-7.57
6-31+G(d,p)	546	15:08:11	-0.03	1.08	5.28	-7.56
cc-pVDZ	426	8:10:03	-0.04	1.11	5.26	-7.70
6-31G(d)	396	1:51:10	-0.03	1.09	5.36	-7.73
3-21G(d)	252	0:44:51	0.03	1.23	6.95	-8.32
STO-3G	138	0:09:48	0.04	1.33	6.41	-9.42

**Table 4.2:** Statistical errors in Hartree-Fock bond angles for the computed geometries of the fluoroquinolones in Figure 4.4, evaluated with various basis sets (n is the number of basis functions used to describe ciprofloxacin).

Basis Set	n	Job CPU Time (hours)	Bond lengths (pm)			
			ME	MAE	MAX	MIN
cc-pVTZ	972	48:52:51	-0.7	1.4	9.5	-11.7
6-311+G(2d,2p)	810	30:31:07	-0.8	1.4	9.4	-11.8
6-311+G(d,p)	636	26:19:36	-1.0	1.5	9.3	-12.0
6-31++G(d,p)	564	17:07:30	-1.2	1.6	8.6	-12.2
6-31+G(d,p)	546	12:51:11	-1.2	1.6	8.6	-12.2
cc-pVDZ	426	4:55:22	-1.2	1.7	8.9	-12.2
6-31G(d)	396	4:18:13	-1.2	1.6	8.8	-12.3
3-21G(d)	252	1:34:43	-1.2	1.6	8.8	-12.3
STO-3G	138	0:28:48	-4.7	4.7	4.4	-16.9

**Table 4.3:** Statistical errors in B3LYP bond lengths for the computed geometries of the fluoroquinolones in Figure 4.4, evaluated with various basis sets (n is the number of basis functions used to describe ciprofloxacin).

Basis Set	n	Job CPU Time (hours)	Bond Angles (degrees)			
			ME	MAE	MAX	MIN
cc-pVTZ	972	48:52:51	-0.03	0.90	5.69	-8.26
6-311+G(2d,2p)	810	30:31:07	-0.02	0.88	5.54	-8.19
6-311+G(d,p)	636	26:19:36	-0.01	0.89	5.67	-8.25
6-31++G(d,p)	564	17:07:30	-0.01	0.86	5.48	-8.16
6-31+G(d,p)	546	12:51:11	-0.01	0.87	5.47	-8.17
cc-pVDZ	426	4:55:22	-0.01	0.90	5.85	-8.28
6-31G(d)	396	4:18:13	-0.01	0.92	5.68	-8.29
3-21G(d)	252	1:34:43	0.06	1.21	7.23	-9.04
STO-3G	138	0:28:48	0.10	1.55	8.47	-10.9

**Table 4.4:** Statistical errors in B3LYP bond angles for the computed geometries of the fluoroquinolones in Figure 4.4, evaluated with various basis sets (n is the number of basis functions used to describe ciprofloxacin).

The results are largely consistent with previous studies. Merz and coworkers<sup>122</sup> study of small-molecule geometries found B3LYP/6-31++G\* bond length MAE ~0.01 Å, bond angle MAE ~1.3 degrees. Bauschlicher<sup>123</sup> found similar absolute errors for B3LYP with double and triple zeta basis sets ranging from 0.008 to 0.013 Å and ~0.6 degrees for bond lengths and bond angles respectively. That data set represents 53 molecules, with 71 bond lengths and 26 bond angles. Our B3LYP/6-31++G(d,p) bond length MAE is 0.017 Å, while our B3LYP/6-311+G(2d,2p) bond length MAE is 0.018 Å, perhaps reflecting the quality of experimental data or the lack of C-H bonds in our data set. HF/ 3-21G\* calculations benefit from error cancellation between Hartree-Fock's under binding and the over binding of minimal basis sets.<sup>124</sup> We conclude that B3LYP/6-311+G(2d,2p) calculations gives the best overall balance of cost and accuracy.

We next consider the effect that computational method has on the errors, with the results summarized in Table 4.5 and Table 4.6.

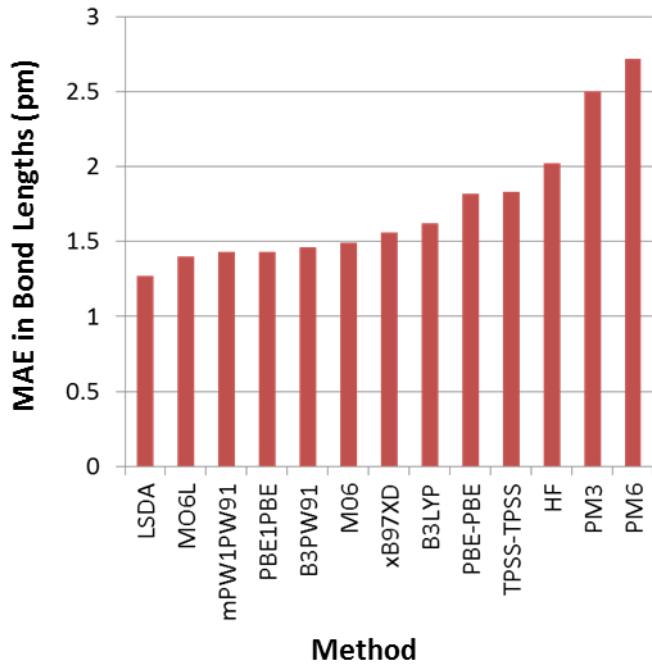
Method	Job CPU Time (hours)	Bond Lengths (pm)			
		ME	MAE	MAX	MIN
HF	61:12:30	0.5	1.8	11.8	-11.5
PM3	0:02:33	-1.4	2.4	9.0	-12.2
PM6	0:01:27	-1.8	2.6	9.1	-12.0
LSDA	14:35:02	0.3	1.2	9.0	-9.9
PBEPBE	118:56:19	-1.4	1.7	8.3	-12.0
B3PW91	57:16:33	-0.4	1.3	9.5	-11.3
PBE1PBE	58:43:20	-0.1	1.3	9.7	-11.1
B3LYP	30:31:07	-0.8	1.4	9.4	-11.8
mPW1PW91	59:17:27	-0.1	1.3	9.8	-11.1
M06	31:54:58	0.0	1.3	10.1	-10.6
TPSSTPSS	19:38:26	-1.4	1.7	8.4	-12.3
M06L	19:04:57	-0.3	1.3	9.5	-10.7
$\omega$ B97XD	85:25:08	-0.2	1.4	9.9	-11.4

**Table 4.5:** Statistical errors in bond lengths for the computed geometries of the fluoroquinolones in Figure 4.4 and Figure 4.5, evaluated with different methods.

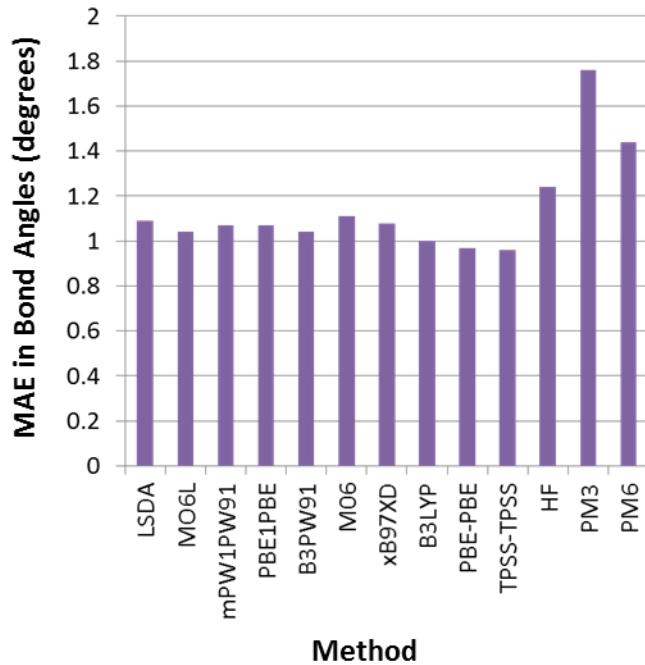
Method	Job CPU Time (hours)	Bond Angles (degrees)			
		ME	MAE	MAX	MIN
HF	61:12:30	-0.04	1.08	5.13	-7.64
PM3	0:02:33	0.04	1.80	13.01	-12.66
PM6	0:01:27	0.08	1.31	9.11	-11.14
LSDA	14:35:02	0.06	0.98	6.57	-8.11
PBEPBE	118:56:19	0.00	0.89	5.90	-8.41
B3PW91	57:16:33	-0.01	0.90	5.65	-8.13
PBE1PBE	58:43:20	0.01	0.92	5.66	-8.08
B3LYP	30:31:07	-0.02	0.88	5.54	-8.19
mPW1PW91	59:17:27	0.00	0.92	5.63	-8.07
M06	31:54:58	0.01	0.95	5.79	-8.16
TPSSTPSS	19:38:26	0.03	0.87	5.76	-8.39
M06L	19:04:57	0.01	0.91	6.11	-8.42
$\omega$ B97XD	85:25:08	0.01	0.92	5.60	-8.05

**Table 4.6:** Statistical errors in bond angles for the computed geometries of the fluoroquinolones in Figure 4.4 and Figure 4.5, evaluated with different methods.

This was evaluated by considering a number of electronic methods, using the best performing 6-311+G(2d,2p) basis set. An overview of the methods considered is shown in Table 4.12. Table 4.5 and Table 4.6 summarize the results. The PM3 and PM6 semi-empirical methods give the largest errors. Hartree-Fock theory exhibits its usual overestimate of bond lengths. The tested DFT methods give generally comparable accuracies, improving on Hartree-Fock and semi-empirical theories. The LSDA gives rather accurate geometries, though it is known to be problematic for other properties.<sup>125</sup> Merz and coworkers found that the LSDA in small basis sets has relatively small errors for bond length and larger errors in bond angle, consistent with Table 4.5 and Table 4.6. The MAE errors in bond length and bond angle for the different computational methods using the 6-311+G(2d,2p) basis set are shown visually in Figure 4.6 and Figure 4.7 respectively.



**Figure 4.6:** Mean absolute errors in bond lengths for B3LYP, semiempirical and DFT approximations using the 6-311+G(2d,2p) basis set.

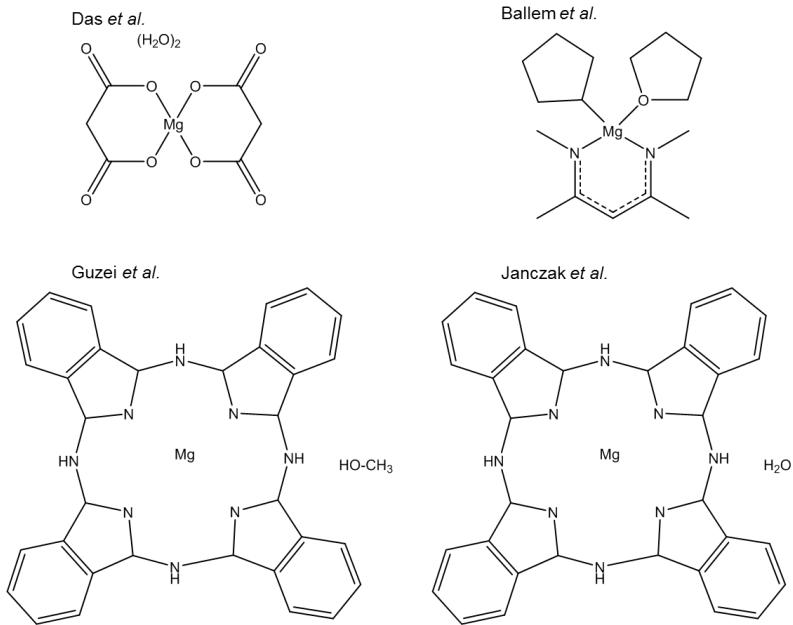


**Figure 4.7:** Mean absolute errors in bond angles for B3LYP, semiempirical and DFT approximations using the 6-311+G(2d,2p) basis set.

Given the comparable performance of the different DFT methods, and the extensive use of B3LYP in other areas, we elected to focus on B3LYP DFT for the remainder of this work.

### 4.3 Mg<sup>2+</sup> Complex Geometry Benchmarks

We then benchmark the best performing methods against the four experimental crystal structures<sup>126–129</sup> of Mg<sup>2+</sup> complexes shown in Figure 4.8. We benchmark to all bond lengths and bond angles involving a magnesium atom, for a total of 20 bonds and 41 angles. While this test set is relatively small, it suffices to confirm that the tested methods are not “pathological” for bonds to magnesium.



**Figure 4.8:** Magnesium complexes used in benchmark set

We next test B3LYP DFT in a variety of basis sets for the for the  $Mg^{2+}$  complexes in Figure 4.8. Table 4.7 and Table 4.8 summarize the results. While the STO-3G basis gives relatively large errors, the other basis sets all give comparable performance. This is consistent with the results of Cundari and coworkers, who found that basis sets are broadly more important for ligands than for transition metals.<sup>130</sup> Test calculations benchmarking against all bond lengths, not just those to magnesium, give error statistics similar to Table 4.7 and Table 4.8 (data not shown).

Basis Set	Bond Length (pm)			
	ME	MAE	MAX	MIN
6-311+G(2d,2p)	-2.0	4.3	7.4	-12.5
6-311+G(d,p)	-2.6	4.3	6.2	-12.6
6-31++G(d,p)	-2.6	4.3	6.4	-12.8
6-31+G(d)	-2.7	4.3	6.3	-13.0
cc-pVDZ	-1.9	3.8	6.7	-9.8
6-31++G	-1.5	2.9	6.2	-7.8
6-31+G	-1.5	2.9	6.2	-7.9
3-21G	1.5	1.9	9.1	-1.2
LANL2DZ	0.0	2.2	7.8	-4.2
STO-3G	7.6	7.6	14.6	3.5

**Table 4.7:** Statistical errors in B3LYP bond lengths for the magnesium-containing bonds in Figure 4.8.

Basis Set	Bond Angle (degrees)			
	ME	MAE	MAX	MIN
6-311+G(2d,2p)	-0.11	1.96	4.82	-5.05
6-311+G(d,p)	-0.11	2.01	4.28	-5.30
6-31++G(d,p)	-0.11	1.96	4.68	-5.35
6-31+G(d)	-0.11	2.00	5.30	-5.31
cc-pVDZ	-0.10	1.79	5.43	-5.42
6-31++G	-0.18	2.55	5.41	-7.21
6-31+G	-0.17	2.55	5.38	-7.17
3-21G	-0.14	2.09	5.28	-6.56
LANL2DZ	-0.21	2.77	6.59	-8.14
STO-3G	-0.32	4.14	10.16	-11.64

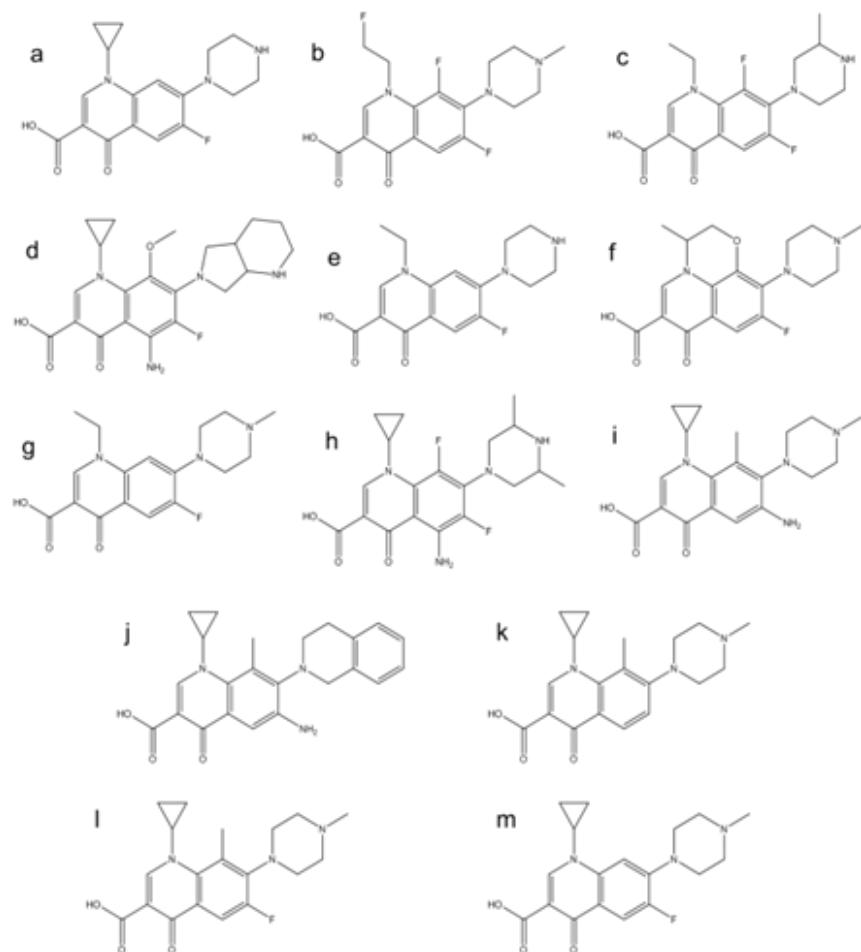
**Table 4.8:** Statistical errors in B3LYP bond angles for the magnesium-containing bonds in Figure 4.8.

While LANL2DZ performs slightly better than 6-311+G(2d,2p) when considering errors in bond length to magnesium, the 6-311+G(2d,2p) basis set performed very well in modeling fluoroquinolones. The use of a single basis set to model both fluoroquinolones and their magnesium complexes will increase the applicability of the screening approach. Therefore,

B3LYP method with the 6-311+G(2d,2p) basis set will be used for all computational calculations.

#### 4.4 Fluoroquinolone-Mg<sup>2+</sup> Binding Affinity Benchmarks

The best-performing methods' ability to reproduce experimental Mg<sup>2+</sup> binding affinities was tested against the fluoroquinolones in Figure 4.9.



**Figure 4.9:** Fluoroquinolones used for magnesium-binding benchmarking.

Comparison is made between the experimental  $\Delta G$ , as calculated from the measured  $K_{Mg}$  and converted by the relationship  $\Delta G_{(Exp)} = RT \ln(K_{Mg})$ , to  $\Delta G_{(Calc)}$  computed at neutral pH, evaluated as shown in Scheme 4.2. We assume that at neutral pH, all carboxyl groups are

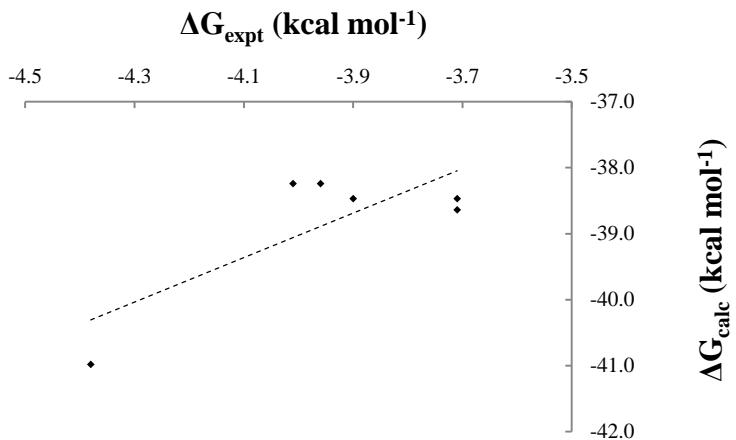
deprotonated and all amine groups are protonated, such that *e.g.* ciprofloxacin exists as a zwitterion. Molecule j was modeled as an anion, as described experimentally.<sup>97</sup>

Table 4.9 shows the calculated fluoroquinolone-Mg<sup>2+</sup> binding affinities and compares them to experimental values computed at neutral pH. We evaluate the binding affinities of a number of fluoroquinolones shown in Figure 4.9. The ability to reproduce experimental Mg<sup>2+</sup> binding affinities is assessed by comparing the binding energies against available experimental data.<sup>86,89,97</sup>

Molecule	$\Delta G_{\text{expt}}$	$\Delta G_{\text{calc}}$
Molecule a (Ciprofloxacin)	-3.90	-38.5
Molecule b (Fleroxacin)	-	-36.9
Molecule c (Lomefloxacin)	-	-37.1
Molecule d (Moxifloxacin)	-	-38.4
Molecule e (Norfloxacin)	-4.08	-37.8
Molecule f (Ofloxacin)	-	-38.4
Molecule g (Pefloxacin)	-4.53	-38.1
Molecule h (Sparfloxacin)	-4.10	-35.8
Molecule i	-4.01	-38.2
Molecule j	-4.38	-41.0
Molecule k	-3.71	-38.6
Molecule l	-3.96	-38.2
Molecule m	-3.71	-38.5

**Table 4.9:** Experimental and computed binding free energy  $\Delta G$  (kcal mol<sup>-1</sup>) for fluoroquinolone-Mg<sup>2+</sup> complexation for the fluoroquinolones in Figure 4.9.

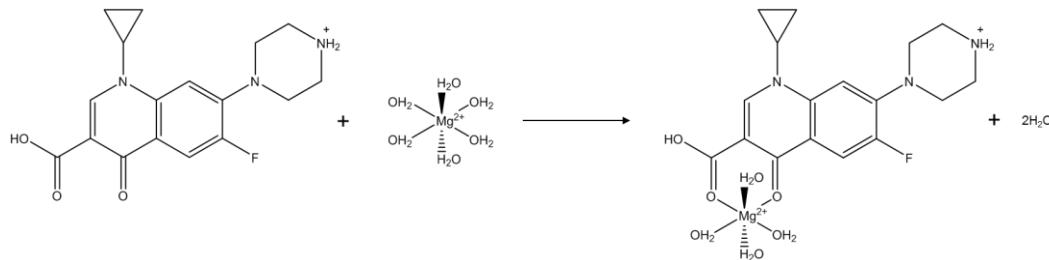
Binding energies for molecules a, i, j, k, l and m reported by Sissi *et al.*<sup>97</sup> reproduced general experimental trends, as shown in Figure 4.10.



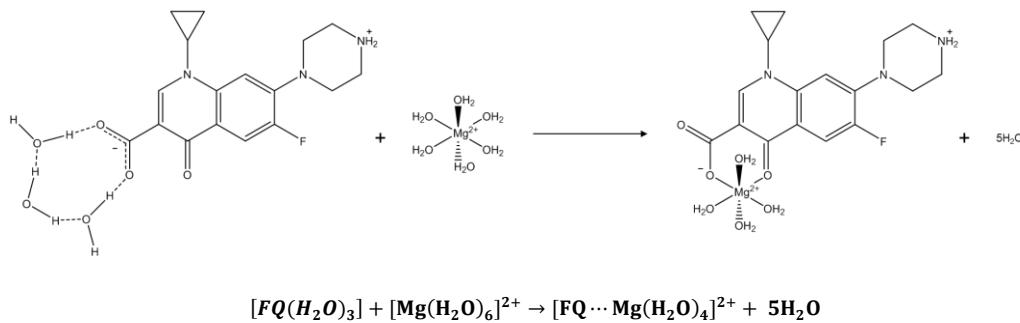
**Figure 4.10:**  $\Delta G_{\text{expt}}$  (kcal mol<sup>-1</sup>) vs.  $\Delta G_{\text{calc}}$  (kcal mol<sup>-1</sup>) for experimental binding affinities reported by Sissi *et al.*

Examining the magnesium binding affinity data in Table 4.9, we see that efficacious fluoroquinolones all have a calculated magnesium binding affinity of -35 to -38 kcal/mol when modeled in their zwitterionic form. When the molecule exists as an anion, as is the documented case for molecule j, a slightly more favorable binding affinity of -41 kcal/mol is predicted. This observation broadly agrees with the experimental magnesium binding affinities. There appears to be a clustering of magnesium binding affinities of -38.4 kcal/mol for the fluoroquinolone zwitterions. Based on this clustering, it appears our model overestimates binding by approximately 34 kcal/mol. The model returns high energies as all calculations are gas phase calculations. There are multiple charged species that are being dealt with, where the inclusion of explicit water molecules of solvation is not an accurate model of real world solvent stabilization effects. At this point, it is important to note that the benchmarking is not indicative of antibiotic efficacy but is intended to identify a target magnesium binding affinity, based on the benchmarking against experimental data. Sissi *et al.* also made note of this when comparing

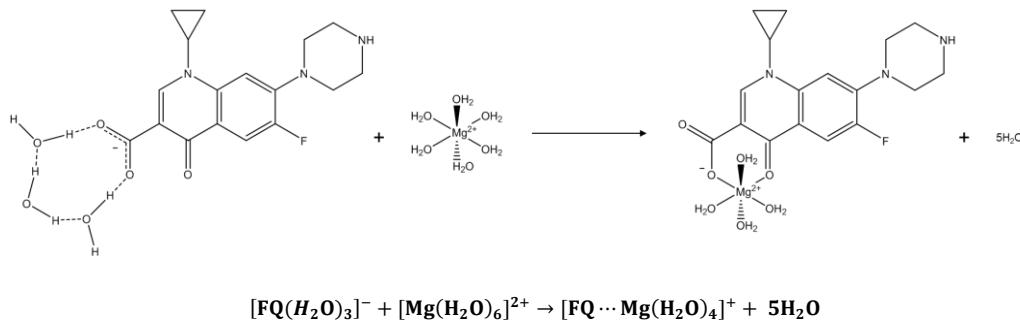
magnesium binding affinity to experimental IC<sub>50</sub> data. From this, we can use the target for screening purposes, projecting that a fluoroquinolone which exhibits binding in the target range may exhibit antibacterial efficacy, while a fluoroquinolone with a lower binding affinity is unlikely to exhibit antibiotic efficacy, based on the accepted binding model.



**Scheme 4.1:** Binding model simulating single charge cation protonation state.

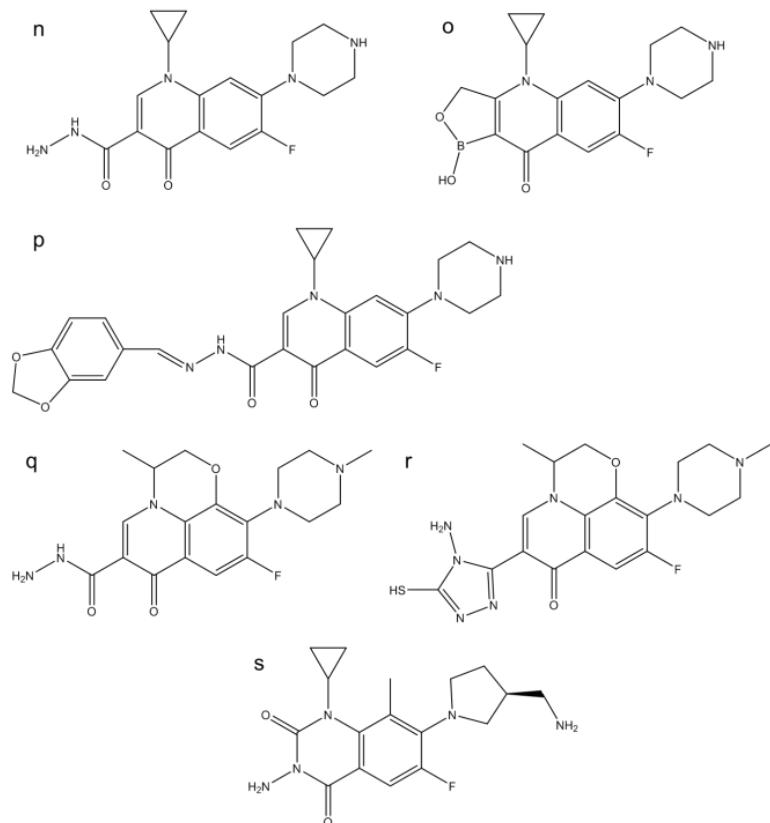


**Scheme 4.2:** Binding model simulating neutral charged zwitterion



**Scheme 4.3:** Binding model simulating single charge anion protonation state

We also explore the  $Mg^{2+}$  affinities' dependence on fluoroquinolone protonation state. The results give insight into how fluoroquinolone derivatization affects pH-dependent fluoroquinolone- $Mg^{2+}$  binding affinities. Our results should aid ongoing work in optimizing fluoroquinolone- $Mg^{2+}$  and fluoroquinolone-target interactions. We consider fluoroquinolone- $Mg^{2+}$  binding affinities for a broader range of fluoroquinolones, with modifications of the carboxylic acid,<sup>131–135</sup> as shown in Figure 4.11. All molecules are modeled as being neutral, without an ionizable group.



**Figure 4.11:** Fluoroquinolones with modifications to the carboxylic acid

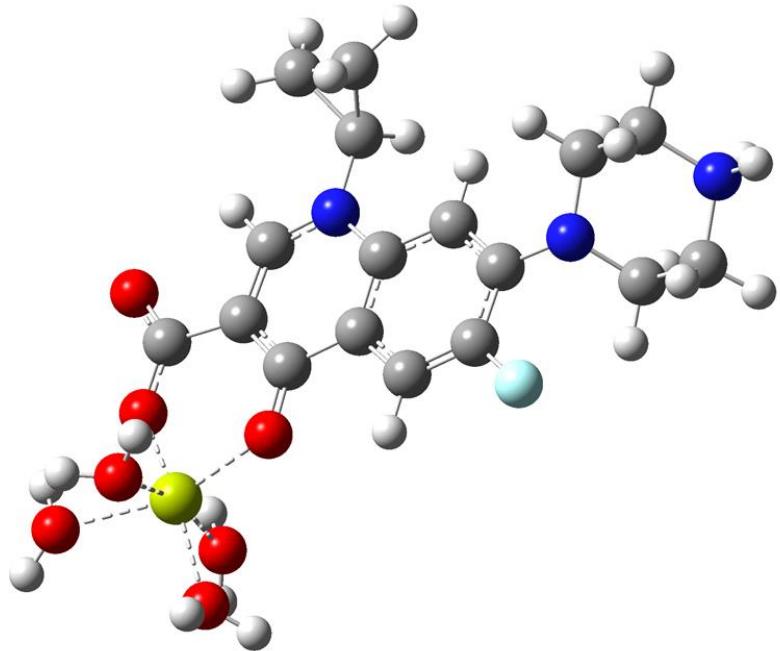
## 4.5 Dependence of Fluoroquinolone-Mg<sup>2+</sup> Binding on Protonation State

Next, we consider the dependence of binding on protonation state by taking into account the protonation state of the fluoroquinolone, with the molecule existing as a cation or an anion, as shown in Scheme 4.1 and Scheme 4.3. The results for fluoroquinolone-Mg<sup>2+</sup> binding in different protonation states, cation, zwitterion or anion, are shown in Table 4.10. Predicted binding affinities are reported. These are calculated by the addition of 34 kcal/mol to calculated affinities, based on deviation from experiment. Cationic forms are protonated at the amine. The anionic form is deprotonated at the carboxylic acid.

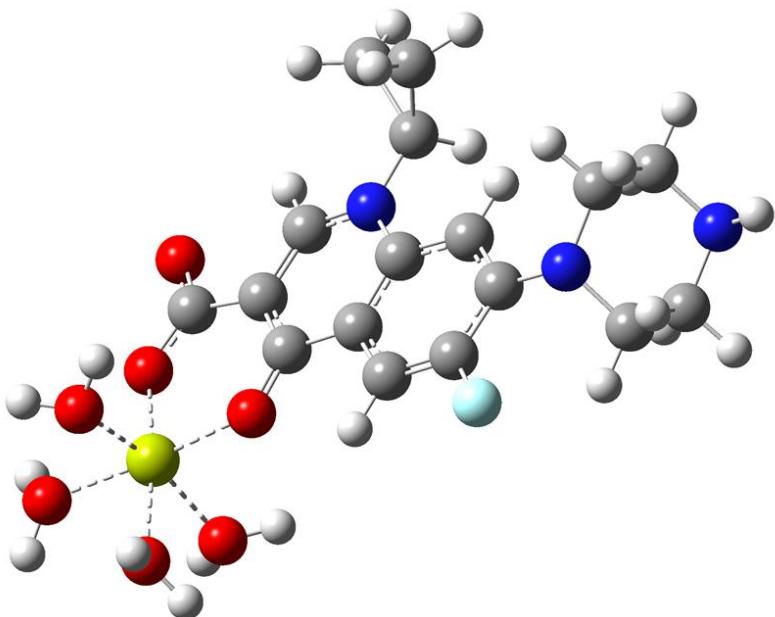
Molecule	Cation	Zwitterion	Anion
Molecule a	26.0	-4.5	-5.7
Molecule b	21.6	-2.9	-4.6
Molecule c	22.2	-3.1	-5.2
Molecule d	20.7	-4.4	-5.5
Molecule e	20.6	-3.8	-5.4
Molecule f	27.0	-4.4	-6.2
Molecule g	21.6	-4.1	-5.0
Molecule h	25.2	-1.8	-3.0
Molecule i	-	-4.2	-6.1
Molecule j	-	-	-7.0
Molecule k	-	-4.6	-6.0
Molecule l	-	-4.2	-4.5
Molecule m	-	-4.5	-5.9

**Table 4.10:** Comparison of predicted  $\Delta G_{\text{calc}}$  (kcal mol<sup>-1</sup>) fluoroquinolone-Mg<sup>2+</sup> binding for different protonation states for the fluoroquinolones in Figure 4.9.

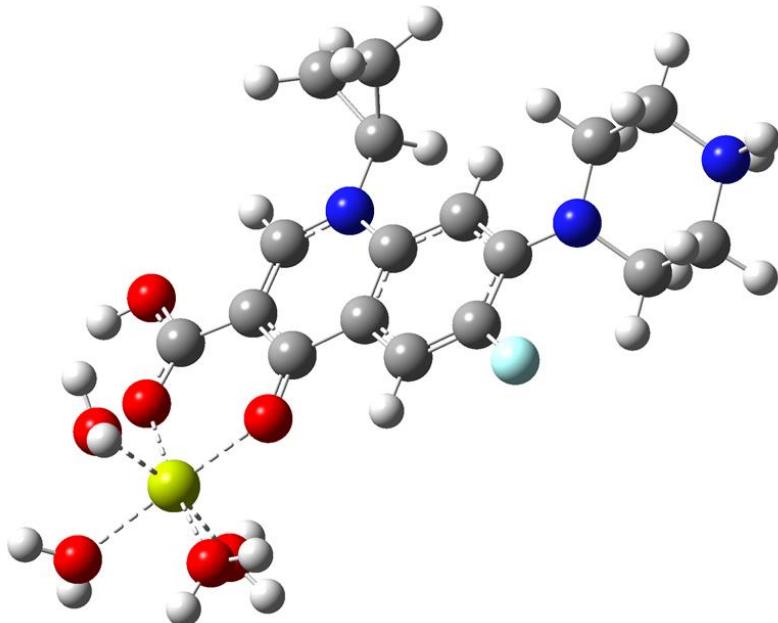
Representative structures for the binding schemes are shown in Figure 4.12, Figure 4.13, and Figure 4.14. Atoms are color coded, with C gray, H white, Mg yellow, O red, N blue and F teal.



**Figure 4.12:** B3LYP/6-311+G(2d,2p) optimized geometry of ciprofloxacin zwitterion binding to  $Mg(H_2O)_N^{2+}$ .

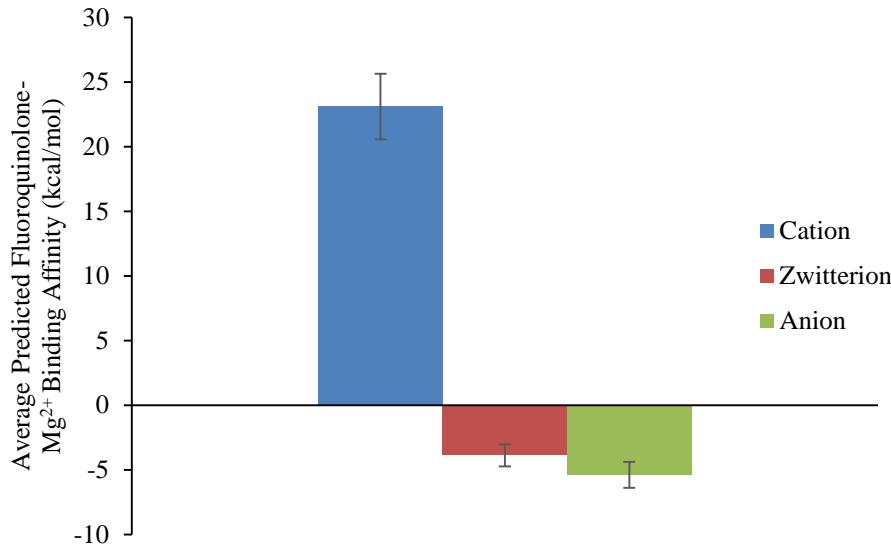


**Figure 4.13:** B3LYP/6-311+G(2d,2p) optimized geometry of ciprofloxacin anion binding to  $Mg(H_2O)_N^{2+}$ .



**Figure 4.14:** B3LYP/6-311+G(2d,2p) optimized geometry of ciprofloxacin cation binding to  $\text{Mg}(\text{H}_2\text{O})\text{N}^{2+}$ .

Anion binding is more favorable as compared to zwitterions by approximately 1 kcal/mol. This is consistent with what was observed experimentally and computationally for molecule j. Predicted binding energy for cations, as would be expected at low pH such as in the stomach, exhibited less favorable binding, by approximately 30 kcal/mol, as compared to predicted binding at neutral and high pH. These results are again consistent with published experimental data. Tulkens and coworkers,<sup>136</sup> Stubbings and coworkers,<sup>137</sup> Azoulay-Dupuis and coworkers<sup>138</sup> and Van Bambeke and coworkers<sup>139</sup> have all described the decrease in efficacy with lower pH. Experimentally, this is determined by an increase in MIC. The average binding affinities, based on fluoroquinolone protonation state, are shown in Figure 4.15.



**Figure 4.15:** Average fluoroquinolone-Mg<sup>2+</sup> binding for fluoroquinolones in Figure 4.9

modeled at various protonation states.

## 4.6 Effect of Carboxylic Acid Modification on Mg<sup>2+</sup> Binding

Recent work has been carried out on modifying fluoroquinolones. Some of these are antibacterial but others have been for alternative applications, with anti-tumor activity being a focus. The molecules in Figure 4.11 represent a selection of these molecules with modifications at the carboxylic acid.

Molecule	Predicted ΔG <sub>calc</sub> (kcal mol <sup>-1</sup> ) Fluoroquinolone-Mg <sup>2+</sup> Binding
n	16.3
o	25.1 (-9.7)
p	17.7
q	17.4
r	19.7
s	Does not form Mg <sup>2+</sup> complex

**Table 4.11:** Comparison of predicted ΔG<sub>calc</sub> (kcal mol<sup>-1</sup>) fluoroquinolone-Mg<sup>2+</sup> binding for neutral molecules (theoretical zwitterion in brackets).

Binding of the modified fluoroquinolones from Figure 4.11 are shown in Table 4.11. All the molecules are modeled in their neutral form, as they do not possess an ionizable group. In all cases, predicted magnesium binding was less than for zwitterionic species. Molecules n, p and q all have an amide in place of the carboxylic acid, with binding taking place via the amide bond and carbonyl. They all exhibit similar predicted binding energies (16 to 17 kcal/mol), suggesting that the type of bond may be more important in binding magnesium than the molecular substituents on the fluoroquinolone molecule. This was also the case for unmodified fluoroquinolones, regardless of the C7 moiety. Molecule r has an amino triazole. In this case, bonding takes place via the carbonyl and lone pair on the triazole nitrogen. The affinity is less favorable than for the molecules containing an amide by approximately 2 kcal/mol. Molecule s is discussed in an experimental study by Drlica and coworkers<sup>135</sup> as not exhibiting antibiotic activity. Based on the binding model, this was likely due to an inability to form a magnesium bridge, by replacing C=O with a primary amine.

Computationally, we show that this molecule does not bind to magnesium, as shown experimentally, supporting the mechanism where by fluoroquinolones must bind via a water-metal ion bridge in order to exhibit antibiotic efficacy. In this test group, molecule o is of particular interest. It has an oxaborole ring<sup>140,141</sup> in place of the carboxylic acid. The molecule did not demonstrate antibiotic activity compared to ciprofloxacin at neutral pH. The computational data is aligned with experimental observation, with a disfavorable predicted magnesium binding affinity of 25 kcal/mol. Examination of the properties of the molecule, the oxaborole ring is less acidic than the carboxylic acid of a fluoroquinolone such as ciprofloxacin. This means the molecule is neutral and the hydroxyl does not deprotonate to form a zwitterion. The authors suggest that manipulating the pKa may produce a more

efficacious molecule. This was investigated computationally by deprotonating the hydroxyl on the oxaborole ring and modeling the compound as a zwitterion. Our predicted magnesium binding affinity shows more favorable binding, than zwitterionic ciprofloxacin at neutral pH, by approximately 5 kcal/mol. This fact is likely due to the less electrophilic nature of boron, resulting in a more negative oxygen as compared to a deprotonated carboxylic acid. With modification of the hydroxyl, perhaps with a better leaving group, the oxaborole ring may represent a future direction for fluoroquinolone drug design.

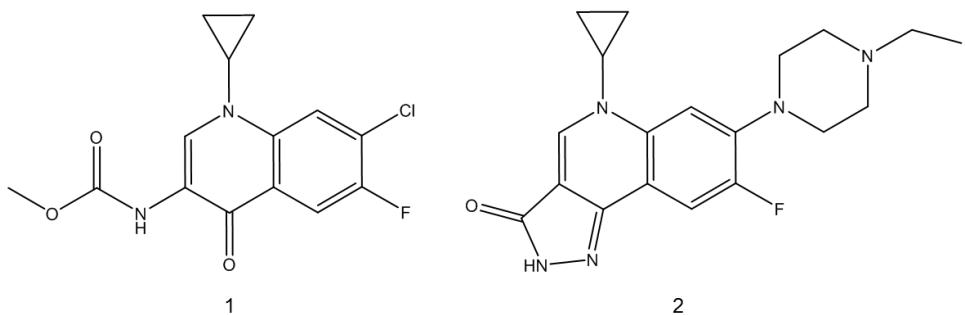
#### 4.7 Concluding Remarks

As previously discussed in Chapter 1, much screening of a development candidate involves ADME and simple application of the rule of 5.<sup>14,142</sup> While useful for estimating permeability, they are not a model for estimating efficacy. This is particularly true in the case of fluoroquinolones, where a binding via Mg<sup>2+</sup> has also been demonstrated to be required for efficacy. The aim of the study was to propose a simple pre-development screen of Mg<sup>2+</sup> binding affinity. This screening method could then be used to evaluate the impact of substituent modification on the likelihood of formation of a water-metal ion bridge, as required for efficacy.

Formation of the fluoroquinolone-Mg<sup>2+</sup> complex was shown to be more favorable for the anionic and zwitterionic forms, as compared to the protonated form, under the conditions modeled by Scheme 4.1, Scheme 4.2 and Scheme 4.3. The project aim was not necessarily to calculate exact magnesium binding affinities but to identify a target for magnesium binding, based on that calculated for efficacious molecules. This could then be used as a screening tool during drug development. Yielding exact binding affinities may not be of much use outside of purely academic study, as being able to accurately predict efficacy may not be possible for a

number of reasons: It has been shown there is not a direct correlation between magnesium binding affinity and antibiotic efficacy; efficacy (and gram-positive or gram-negative activity) are also dependent on other substituents and the C-7 moiety; efficacy has been shown to be substrate specific.

This work identified B3LYP method with the 6-311+G(2d,2p) triple-zeta basis set as being able to accurately optimize the geometries of both fluoroquinolone molecules and magnesium-fluoroquinolone complexes. Computational prediction of magnesium binding affinity, based on ionization state, in different biologically relevant conditions, showed agreement with published experimental data. This agreement suggests the screening approach developed in this study can be of use in designing future iterations of fluoroquinolones, ensuring that substituents or modifications do not adversely impact magnesium binding affinity. It was previously shown that binding to magnesium occurred through the ketone and carbonyl of the carboxylic acid. When the carboxylic acid was replaced, efficacy was not seen. This effect was seen computationally also, where a ciprofloxacin dione derivative did not bind to magnesium. Another study by Kojic-Prodic *et al.*<sup>143</sup> showed that *in vitro* activity of two ciprofloxacin derivatives, with modifications to the carboxylic acid (Figure 4.16), was lower than that of ciprofloxacin. Based on our studies and other literature, we can say that magnesium binding is required for fluoroquinolone antibacterial activity and that the screening approach we propose is valid.



**Figure 4.16:** Ciprofloxacin derivatives (1) 7-chloro-1-cyclopropyl-6-oxo-1,4-dihydro-4-oxo-quinoline-3-methylcarbamate and (2) 1-cyclopropyl-7-(4-ethyl-1-piperazinyl)-6-oxo-1,4-dihydro-5(3H)-oxopyrazolo[4,3-c]quino-line.

We also identify oxaborole rings as potential targets for future fluoroquinolone modifications. The identification of leaving groups and the tuning of pKa are potential areas for investigation in tuning antibiotic efficacy.

## 4.8 Computational Section

All calculations use the Gaussian 09 electronic structure package.<sup>144</sup> Tested methods (Table 4.12) include the PM3 and PM6 semi-empirical methods, Hartree-Fock theory, and generalized Kohn-Sham density functional theory.<sup>60,61,145,146</sup>

<b>Method</b>	<b>Description</b>
HF <sup>147–149</sup>	Exchange Only
PM3 <sup>150–152</sup>	Semi-Empirical
PM6 <sup>153</sup>	
SVWN5 <sup>60,61,154,155</sup>	Local Spin Density Approximation
PBEPBE <sup>156,157</sup>	Generalized Gradient Approximation
B3PW91 <sup>158–163</sup>	Global Hybrid-GGA
PBE1PBE <sup>156,157,164</sup>	
B3LYP <sup>158,165,166</sup>	
mPW1PW91 <sup>159–163,167</sup>	
M06 <sup>168</sup>	
TPSSTPSS <sup>169</sup>	Meta-GGA
M06L <sup>170</sup>	
$\omega$ B97XD <sup>171,172</sup>	Dispersion-Corrected Range-Separated Hybrid

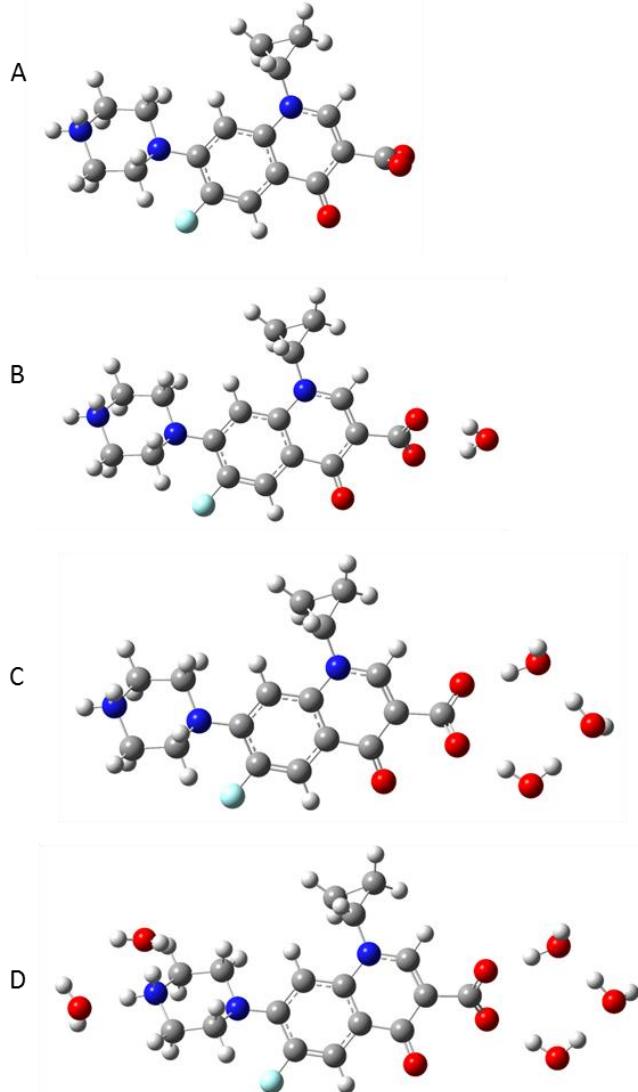
**Table 4.12:** Computational methods used for this study.

DFT calculations test several different approximations for the ground-state exchange-correlation (XC) density functional: the local spin-density approximation (SVWN5); the Perdew-Burke-Ernzerhof (PBEPBE) generalized gradient approximations (GGAs); the Tao-Perdew-Staroverov-Scuseria (TPSSTPSS) and Minnesota M06L meta-GGAs; the B3PW91, B3LYP, mPW1PW91, and M06 global hybrids, and the  $\omega$ B97X-D dispersion-corrected range-separated hybrid. Hartree-Fock and B3LYP calculations are performed in several basis sets: the STO-3G<sup>173,174</sup> minimal basis set, consisting of 3 primitive Gaussian orbitals fitted to a single Slater-type orbital; the 3-21G(d),<sup>175–180</sup> 6-31G(d),<sup>181–188</sup> 6-31+G(d,p), 6-31++G(d,p), 6-311+G(d,p),<sup>189,190</sup> 6-311+G(2d,2p) Pople<sup>175,176,181,182,191,192</sup> basis sets, cc-pVDZ double-zeta and cc-pVTZ triple-zeta Dunning<sup>65,193</sup> basis sets, and the Los Alamos LANL2DZ<sup>194–196</sup> basis set.

All benchmark calculations treat isolated molecules, and do not consider crystal packing effects. Isolated fluoroquinolones' geometries are optimized in the neutral state, with no solvent model. Fluoroquinolone-Mg<sup>2+</sup> complex geometries are optimized in a continuum model<sup>197–206</sup> for water solvent, as well as up to six explicit water molecules bound to Mg<sup>2+</sup>. All

calculations use stability analysis<sup>207</sup> to confirm that the self-consistent field calculations converge to a stable state. Benchmarks of geometries consider all distances between bonded atoms and all angles between pairs of bonds. Bonds to hydrogen are not included in the benchmarks, as these are often not well-resolved in crystal structures. "ME", "MAE", "Max", and "Min" denote mean, mean absolute, maximum and minimum errors. Magnesium binding affinities were calculated using B3LYP method with the 6-311+G(2d,2p) basis set in a continuum water solvent. Computed geometries and total energies of all species, at all tested levels of theory, are included in Appendix C. Hydrogen atoms were added in GaussView 5.0, as the experimental hydrogen atom positions were not reported.

Modelling the zwitterion or anionic forms of the fluoroquinolones presented some issues. The negatively charged carboxyl group gave a molecule of higher energy than anticipated. This fact lead to erroneous results that did not make sense when considering the expected magnesium binding, based on experimental data.



**Figure 4.17:** Modelling of ciprofloxacin zwitterion (a) with the addition of one (b), three (c) or five (d) explicit water molecules (atoms are color coded, with C gray, H white, N blue, O red and F teal).

Explicit water molecules were added to the fluoroquinolone to simplistically model solvent effects. Three explicit solvation models were considered, as shown in Figure 4.17. The first two models considered either one or three explicit water molecules bound at the deprotonated carboxyl group. The third model included two explicit waters bound to the protonated amine,

as well as three explicit waters bound at the carboxyl. The assessment of the models are shown in Table 4.13

Explicit Water	Gibbs Free Energy (Hartree)	O=C-O (Degrees)	O-H-O (Degrees)	N-H-O (Degrees)	H-O-H (Degrees)	R O-H (Å)
0	-1148.469739	127.1	-	-	-	-
1	-1224.928239	126.2	143.1	-	96.9	1.97
3	-1377.849959	126.1	174.0	-	105.6	1.78
5	-1530.777144	126.1	174.4	176.3	105.5	1.78
H <sub>2</sub> O	-76.465371	-	-	-	104.6	-
Ciprofloxacin	-1148.473087	122.4	-	-	-	-

**Table 4.13:** Modelling the ciprofloxacin zwitterion by addition of up to five explicit water molecules.

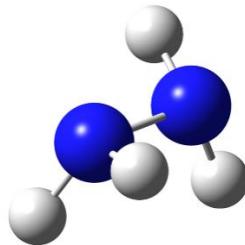
The first takeaway is that the zwitterion without any explicit waters is higher in energy than the neutral reference ciprofloxacin molecule. The O=C=O bond angle for the deprotonated carboxyl group is approximately 4 degrees more than for the neutral carboxylic acid O=C-O bond angle. The O=C-O bond angle is what we would expect based on the free zwitterion but the O-H-O hydrogen bond angle is strained at only 143 degrees. The H-O-H bond angle in the water molecule is also strained, recovering a bond angle of only 96.9 degrees compared to 104.6 degrees in an unbound water molecule. This is as a result of a strained 6-membered ring formed between the explicit water and deprotonated carboxyl. Addition of two more explicit water molecules, for a total of three bound via the carboxyl yields more respectable results compared to expectation. The O=C-O bond angle of 126.1 degrees returns that of an optimized ciprofloxacin zwitterion. The O-H-O hydrogen bond angle of 174 degrees is closer to the 180 degrees we may expect. The H-O-H bond angle in the water molecules averages at 105.6 degrees, returning the 104.6 degrees in the free water molecule returned at this level of theory.

Finally, adding two explicit water molecules bound to the protonated amine did not change the geometry at the carboxyl but returned N-H-O hydrogen bond angles of 176.3 degrees, close to an expected 180 degrees. Again the H-O-H bond angle in the water molecules returns that of the unbound water. Models using 3 or 5 explicit water molecules both returned more stable configurations than without explicit solvation and both had water and hydrogen bond angles closer approximating what we may expect in real life. With this in mind, all calculations of zwitterionic or anionic fluoroquinolones, use three explicit water molecules bound to the deprotonated carboxyl, as the model approximates a real system but is not too computationally demanding of cost or time. The model with 3 explicit water molecules bound at the carboxyl is also of greater interest due to the fact these waters are involved with magnesium binding and will be lost upon complexation with  $[Mg(H_2O)_6]^{2+}$ , having an effect of increasing entropy in the products. It is worth also noting, that in all cases, the hydrogen bond lengths calculated were between 1.78 to 1.97 Å. This is within the typical range for a hydrogen bond of 1.5 to 2.5 Å. The detailed computational benchmarks compiled in this study are included as adjunctive information.

# Chapter 5: Effect of N2 Substituent on Hydrolysis of Hydrazones

## 5.1 Hydrazine Chemistry

Hydrazine is a simple organic compound with the chemical formula  $\text{N}_2\text{H}_4$ , as shown in Figure 5.1. Hydrazine applications generally employ the hydrated form. Through 2000, approximately 120,000 tons of hydrazine hydrate were produced per year globally.<sup>208</sup>



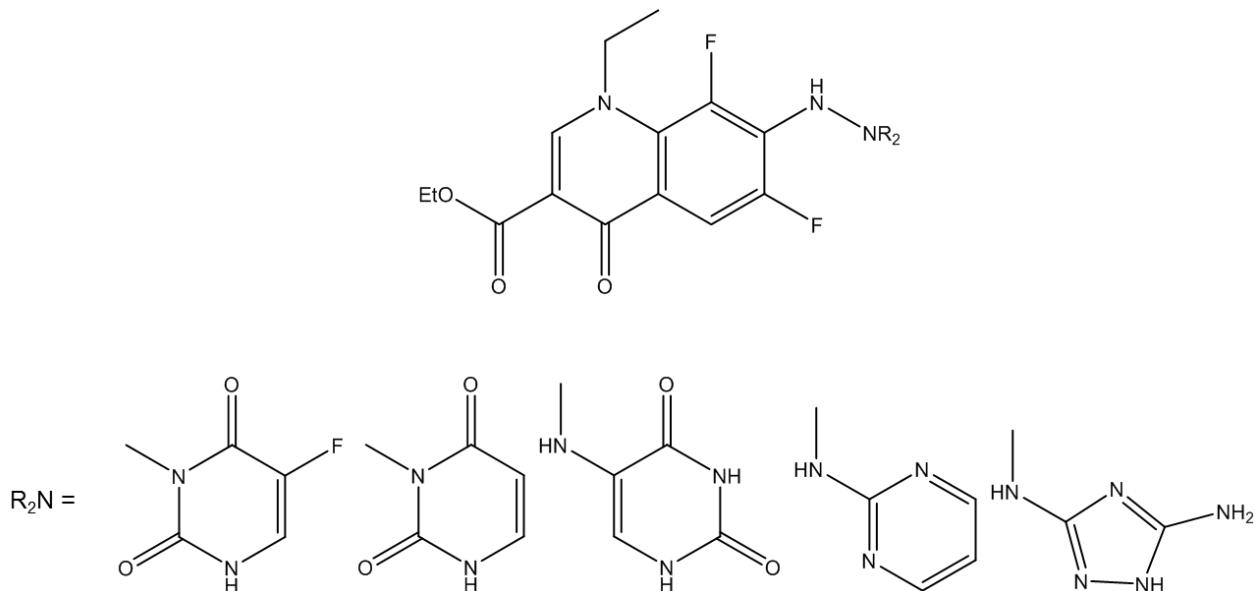
**Figure 5.1:** B3LYP/6-31+G(d,p) optimized geometry of hydrazine (atoms are color coded with N blue and H white).

As can be seen from Figure 5.1, hydrazine adopts a geometry where the angle between the lone pairs is approximately 60 degree. This is favored due to the gauche effect. Stabilization from the  $n \rightarrow \sigma^*$  interactions between antiperiplanar donor lone pairs and acceptor bonds. Hydrazine applications include foaming agents, rocket fuels, in the preparation of the gas precursors for air bags, pharmaceutical synthesis, and the Wolff-Kishner reduction.<sup>209,210</sup>

In the current study, we are interested in the hydrazone derivatives. The pH labile nature of hydrazones has made them useful in wide variety of applications. Hydrazones have been used in creating dynamic combinatorial libraries<sup>211–213</sup> and as reagents in organic synthesis<sup>214,215</sup>. Hydrazone chemistry has been used to get fragmented aldehydes and ketones

into colognes<sup>216</sup>. In addition, hydrazones have been used in molecular switches, metallo-assemblies, and sensors<sup>217</sup>. Furthermore, hydrazone based coupling methods have been used in medical biotechnology to attach drugs to target antibiotics.<sup>218–222</sup>

Of interest to the fluoroquinolone binding study discussed in Chapter 4, hydrazine derivatives of fluoroquinolones have been synthesized,<sup>223</sup> with a view to study their antibacterial and anti-cancer activities. A selection of the derivatives are shown in Figure 5.2.

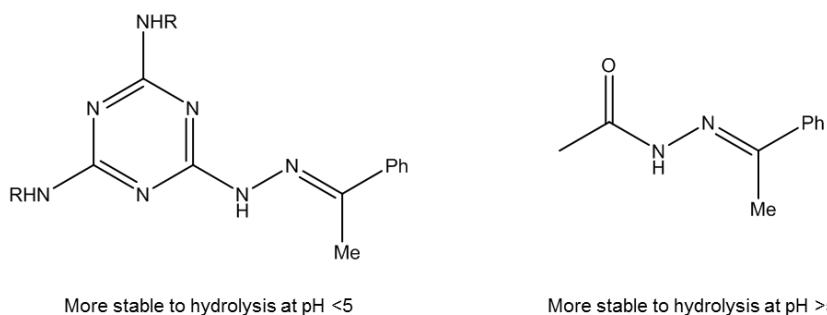


**Figure 5.2:** Examples of 7-hydrazino-1-ethyl-6,8-difluoroquinolone-3-carboxylate derivatives.

Hu *et al.*<sup>132</sup> have explored fluoroquinolone anti-tumor activities of Schiff-Mannich<sup>224,225</sup> bases derived from ofloxacin.

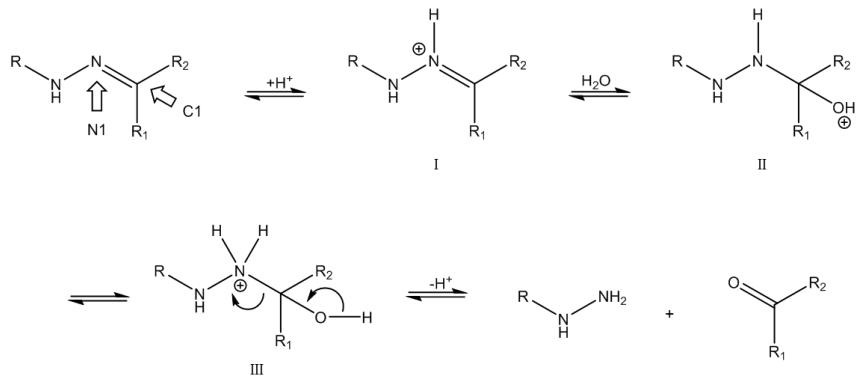
## 5.2 Hydrolysis of Hydrazones

Previously, the Simanek and Janesko groups in collaboration demonstrated a reversal in stability to hydrolysis of triazinyl hydrazones around the triazine pKa ( $\sim\text{pH } 5$ )<sup>226</sup> as shown in Figure 5.3.



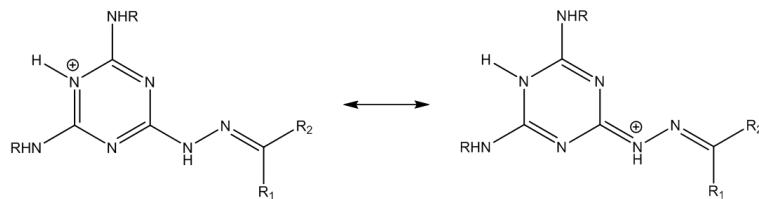
**Figure 5.3:** pH dependence of triazine and acetyl hydrazone stability to hydrolysis.

The varying theories regarding hydrazine stability to hydrolysis are based on thermodynamics,<sup>227</sup> resonance stabilization based on the reduction of electrophilicity of C1, and finally resistance to protonation of N1.<sup>228</sup> Results in the preceding study showed that resistance to protonation at N1 was the larger stability driver than electrophilicity of C1. Computational studies agreed with the experimentally determined stability, based on the hydrazone hydrolysis rate. The mechanism for hydrazone hydrolysis this study, and previous work<sup>226</sup> is based on, is shown in Scheme 5.1. Initiation step is protonation of nitrogen N1 to yield I. Addition of a molecule of water yields carbinolamine intermediate II. Protonation at N1 to give III leads to C–N bond cleavage to complete the hydrolysis. The mechanism is consistent with catalysis observed in acid.<sup>227,228</sup>



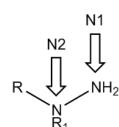
**Scheme 5.1:** Proposed mechanism of acid catalyzed hydrolysis of hydrazones ( $R = \text{acetyl or triazinyl}$ ).

The triazine, with a  $pK_a$  of  $\sim 5$ , can undergo protonation at low pH, as shown in Figure 5.4, with the formation of an imine resonance structure. For comparison, the  $pK_a$  of the amines on the triazine is  $\sim 10\text{-}11$ .



**Figure 5.4:** Example of triazine protonation.

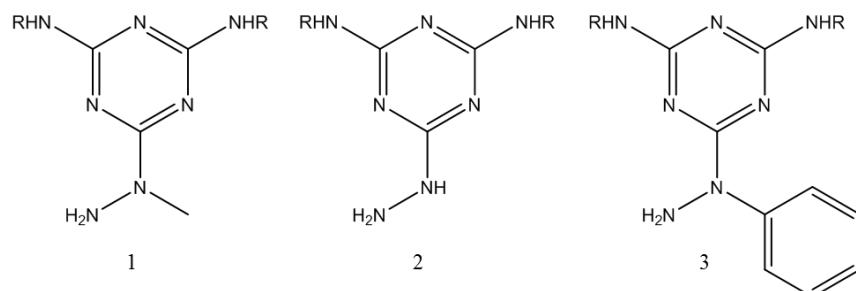
In this study, a collaboration with Vishal Sharma of the Simanek group, we compare the hydrolysis rates of a number of triazinylhydrazines, with different substituents at N2 (Figure 5.5).



$R = \text{Triazine, } \text{CH}_3\text{CO}$   
 $R1 = \text{CH}_3, \text{H, Ph, CF}_3, \text{CN}$

**Figure 5.5:** Structure of hydrazones studied and the designation of nitrogen N1 and N2.

Experimentally, we considered three hydrazines with different substituents at N2, as shown in Figure 5.6.



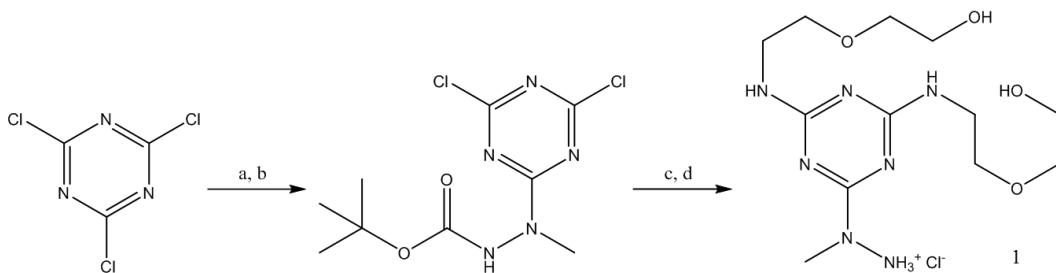
**Figure 5.6:** Experimental triazinyl hydrazines with 1) CH<sub>3</sub>, 2) H and 3) phenyl substituents at N2, where R = CH<sub>2</sub>CH<sub>2</sub>OCH<sub>2</sub>CH<sub>2</sub>OH.

Aminoethoxyethanol substituents R were used to enhance the solubility of the hydrazones in water. Installation of aminoethoxyethanol followed by substitution of monochlorotriazine intermediate by tert-butyl carbazole at elevated temperature gives the desired product.

Triazinylhydrazine 2 was previously reported by Ji *et al.*<sup>226</sup> Addition of a methyl group at N2 is expected to make N1 more basic and therefore more susceptible to hydrolysis. We also looked at a triazinylhydrazine with a phenyl substituent at N2. The phenyl group is electron withdrawing by induction, such that electronic effects alone are predicted to give these species the slowest hydrolysis rate.

N2-methylated and unmethylated triazinyl hydrazines were synthesized by adopting two different schemes each with four steps. In both schemes, aminoethoxyethanol was used as substituent to replace two out of three chlorine atoms of starting material triazine chloride. The last chlorine was replaced by tert-butyl carbazole. I carried out the synthesis of triazinylhydrazine 1 along with Vishal Sharma. Synthesis of triazinylhydrazines 2 and 3 were

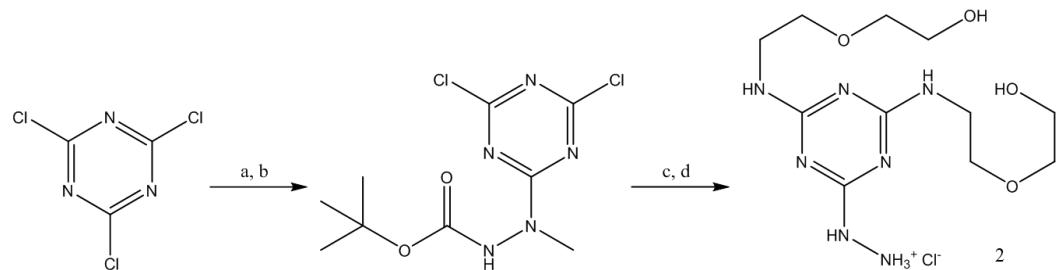
carried out solely by Vishal Sharma. Synthesis of the first triazinylhydrazine is shown in Scheme 5.2.



**Scheme 5.2:** Synthesis of Triazinylhydrazine 1.<sup>a</sup>

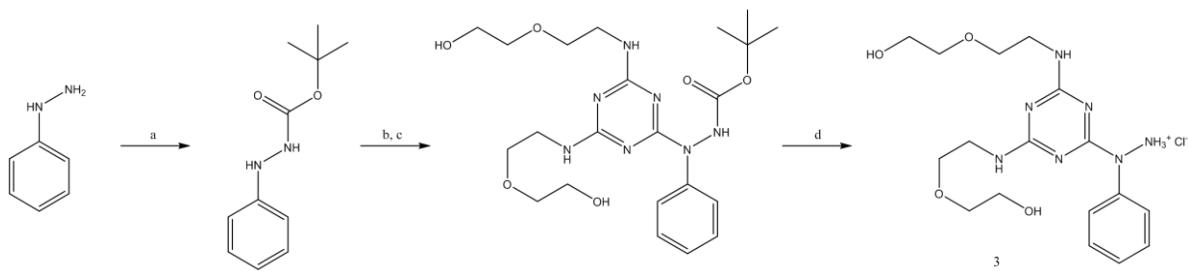
<sup>a</sup>Reagents and conditions: (a) Tert-butylcarbazate, THF, DIPEA, -20°C, 1 h. (b) Methyl iodide, DIPEA, rt, 12 h. (c) 2 equiv. Aminoethoxyethanol, CsCO<sub>3</sub>, dioxane, 95°C, microwave, 30 mins. (d) 4M HCl, CH<sub>3</sub>OH, rt, 12 h.

Similarly, unmethylated triazinyl hydrazine was made, however, without the methyl iodide step as shown in Scheme 5.3.



**Scheme 5.3:** Synthesis of Triazinylhydrazine 2.<sup>a</sup>

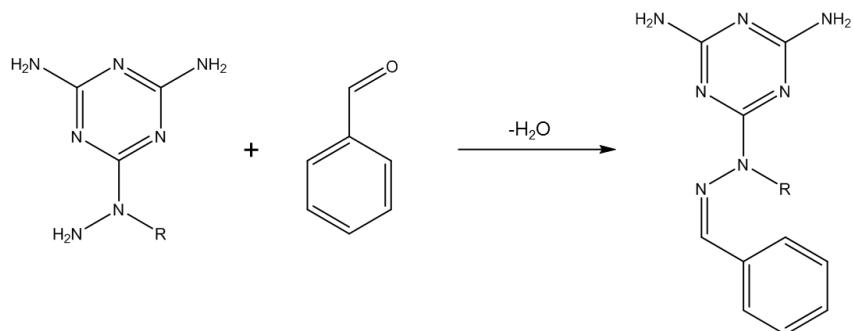
<sup>a</sup>Reagents and conditions: (a) Tert-butylcarbazate, THF, DIPEA, -20°C, 1 h. (b) 1 equiv. Aminoethoxyethanol, THF, DIPEA, rt, 12 h. (c) 1 equiv. Aminoethoxyethanol, dioxane, CsCO<sub>3</sub>, microwave, 95°C, 30 mins. (d) 4M HCl, CH<sub>3</sub>OH, rt, 12 h.



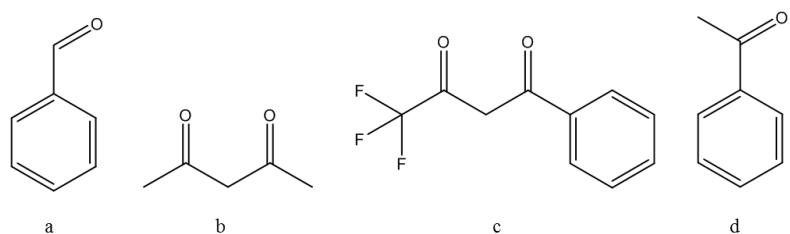
**Scheme 5.4:** Synthesis of Triazinylhydrazine 3.<sup>a</sup>

<sup>a</sup>Reagents and conditions: (a) Di-tert-butyl dicarbonate, THF, Triethylamine, refluxed, 12 h. (b) 1 equiv. Triazine chloride, THF, DIPEA, 0°C, 30 mins, rt, 2 h. (c) 2 equiv. Aminoethoxyethanol, dioxane, CsCO<sub>3</sub>, microwave, 95°C, 60 mins. (d) 4M HCl, DCM, rt, 12 h.

Upon reaction between a hydrazine and an aldehyde or ketone, a hydrazone is formed as shown in Scheme 5.5.



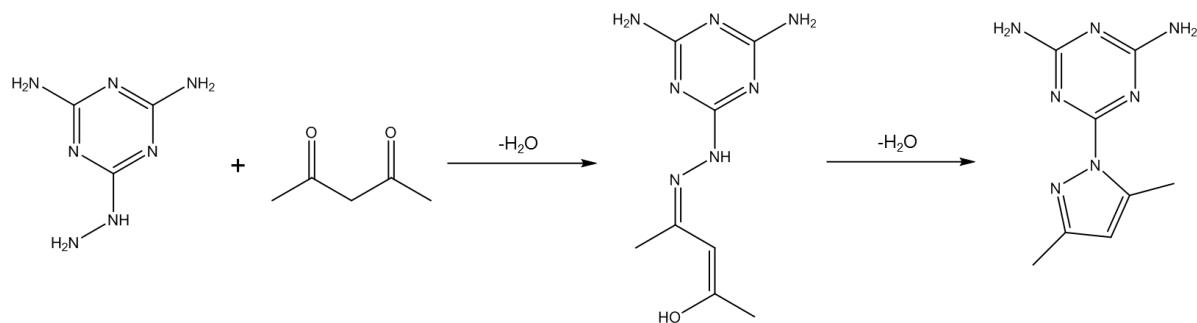
**Scheme 5.5:** Formation of hydrazone.



**Figure 5.7:** Aldehydes and ketones (a-d) used for hydrolysis studies.

We consider hydrazones 1a,1b...3d formed from hydrazines 1-3 (Figure 5.10) and aldehydes and ketones a-d (Figure 5.11). Structures of all tested hydrazones are shown in Figure 5.13.

Reacting hydrazine 1 with a diketone can form a pyrazole ring via a second dehydration, as shown in Scheme 5.6.



**Scheme 5.6:** Formation of a pyrazole ring via a second dehydration when N2 is hydrogen and reacted with a diketone.

All the hydrazones were characterized by using <sup>1</sup>H-NMR, <sup>13</sup>C-NMR, and mass spectroscopy. Vishal Sharma of the Simanek group evaluated the hydrolysis rates of hydrazones 1a-3d via HPLC. Prior to reaction, all hydrazones were dissolved in enough THF to make 30 to 50 mM stock solutions and diluted to 0.5 mM with respective buffer at 37 ± 2 °C. Buffer solutions at pH 4.00, 5.00, and 7.00 were made by mixing different amounts of 0.2M Na<sub>2</sub>PO<sub>4</sub> and 0.1M citric acid. Ten-fold formaldehyde was added to trap the released triazinyl hydrazine. A small sample of solutions were directly collected from the vials and ran in HPLC at regular intervals. The hydrolysis rates of all the hydrazones were monitored in HPLC using an H<sub>2</sub>O/AcCN (75/25) solvent system. Hydrazones 2b and 2c did not undergo hydrolysis, while hydrazones 1b, 3b, 1d, and 3d hydrolyzed too fast to be monitored by HPLC.

### 5.3 Effect of N2 Substituents on Hydrazone Stability to Hydrolysis

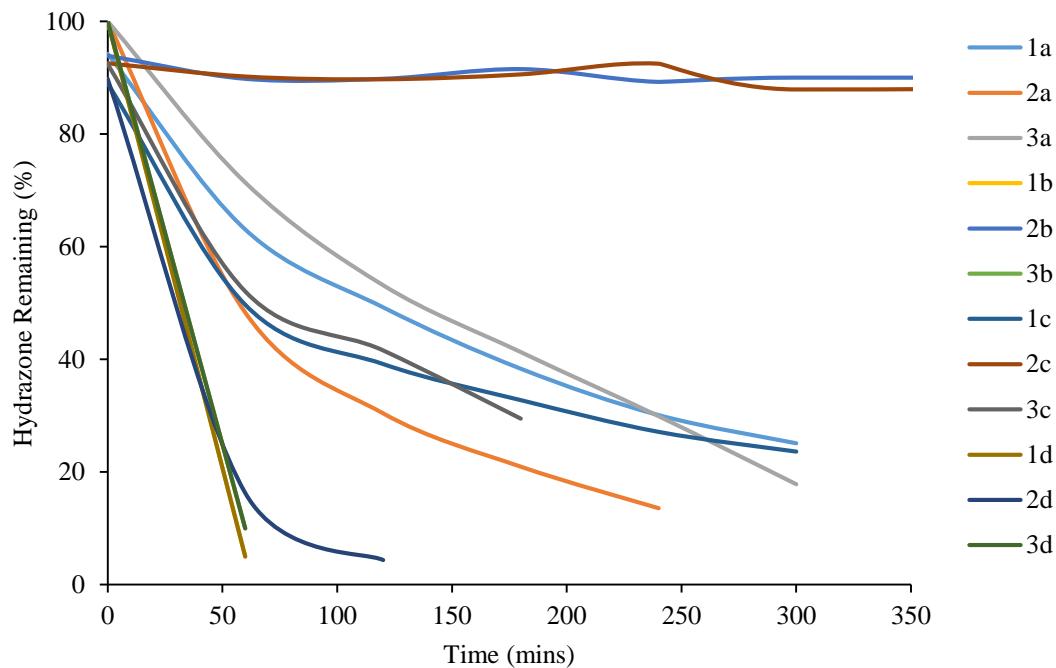
Results for experimental hydrolysis rates, as well as calculated protonation affinities are shown in Table 5.1. Progress of hydrazone hydrolysis is shown in Figure 5.8, while relative hydrazone stabilities are shown in Figure 5.9.

Compound	pH 4.0		pH 5.0		pH 7.0		$\Delta E_{\text{prot}}$ (kcal/mol) <sup>a</sup>
	$K$ ( $10^{-3}$ min $^{-1}$ )	$t_{1/2}$ (min)	$K$ ( $10^{-3}$ min $^{-1}$ )	$t_{1/2}$ (min)	$k$ ( $10^{-3}$ min $^{-1}$ )	$t_{1/2}$ (min)	
1a	5.3	131	3.1	224	0.2	3465	1.7 (-4.9)
2a	8.5	82	8.1	86	0.2	3500	-2.2 (-7.4)
3a	4.8	144	3.3	210	0.2	3500	1.3 (-5.4)
1b	>70	<10	>70	<10	>70	<10	12.3 (5.4)
2b	-	-	-	-	-	-	-22.6
3b	-	<20	-	<20	-	<20	10.0 (3.2)
1c	7.0	99	4.1	169	7.3	95	7.6 (-4.7)
2c	-	-	-	-	-	-	-30.6
3c	6.1	113	-	-	-	-	6.6 (-7.6)
1d	-	<10	-	<10	-	<60	7.6 (-1.4)
2d	24.9	28	14.9	47	0.5	1386	2.1 (-4.5)
3d	-	<10	-	<20	-	<90	6.7 (-1.9)

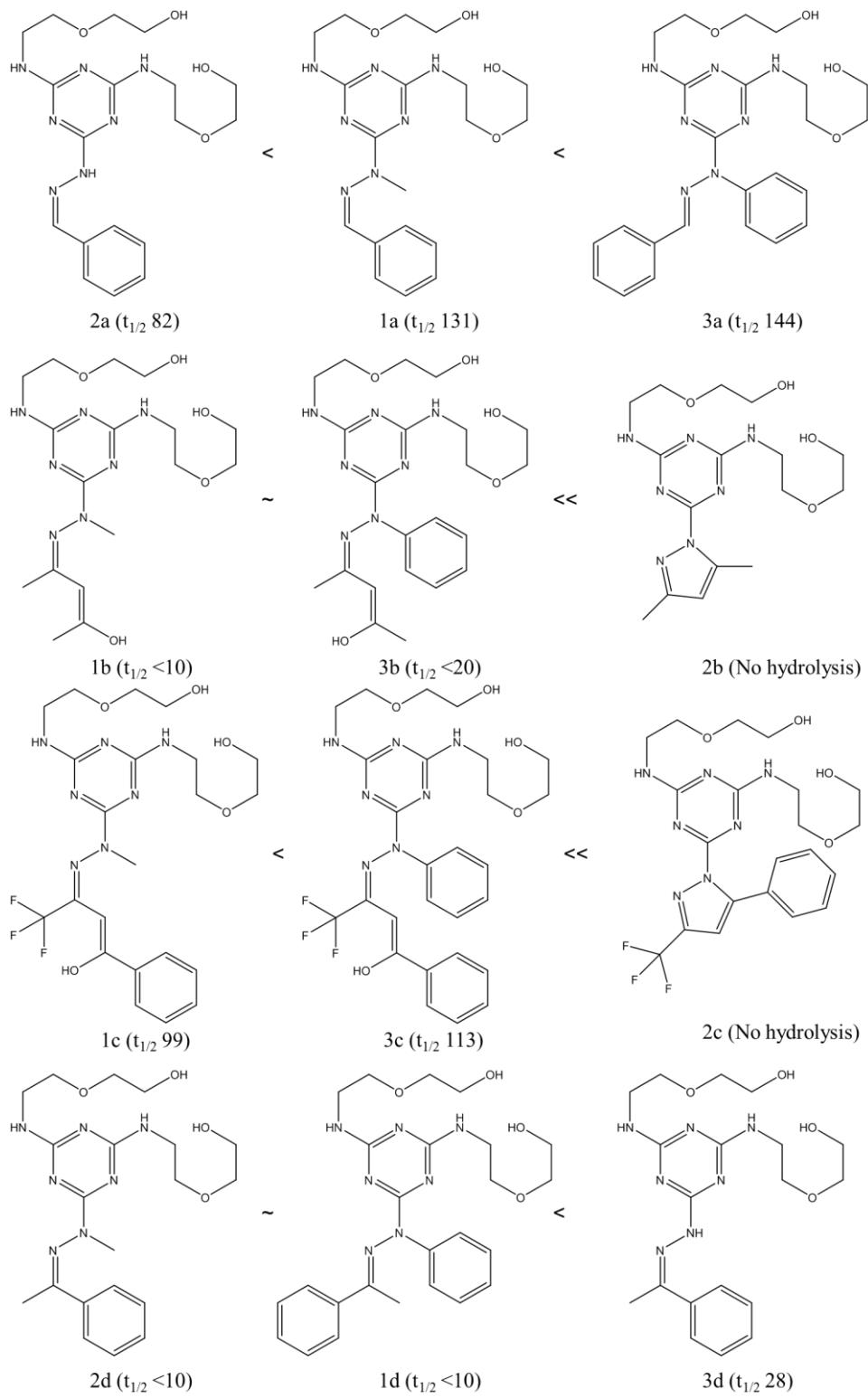
<sup>a</sup>Changes in  $\Delta E_{\text{prot}}$  upon traizine protonation, as shown in Figure 5.4, are given in parentheses.

**Table 5.1:** First order reaction constants, half-lives and DFT-computed N1 proton affinities

$\Delta E_{\text{prot}}$  relative to hydrazone 4d.



**Figure 5.8:** Experimental hydrolysis of triazinylhydrazones as measured by HPLC.

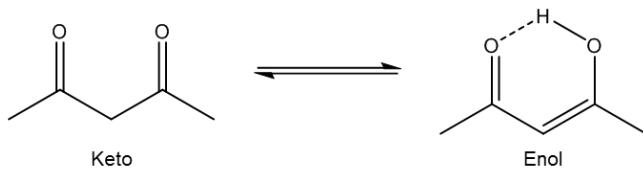


**Figure 5.9:** Relative experimental triazinylhydrazine stability to hydrolysis at pH 4.0 ( $t_{1/2}$  shown in minutes).

Experimentally, N2-methyl hydrazones hydrolyzed the fastest, followed by the hydrolysis of N2-phenyl hydrazones and then N2-hydrogen hydrazones. Methyl is electron donating, making N2 more basic. This has the effect of increasing the proton affinity of N1 and increasing the hydrolysis rate of hydrazones with N2-methyl. Trends in hydrolysis rates with phenyl are discussed below

Hydrazone 2b formed a stable pyrazole ring and did not hydrolyze. We would expect 2c to form a pyrazole ring also, however NMR analysis identified an open ring. Despite this, 2c did not hydrolyze. Possible explanations are discussed below.

The diketones are able to undergo keto-enol tautomerization, as shown in Scheme 5.7. In the hydrazone, the enol hydrogen atom capable of H-bonding with N1 of the hydrazone. There have been many studies looking at keto-enol tautomerization and solvent effects.<sup>229–232</sup>



**Scheme 5.7:** Keto enol tautomerism as observed for acetylacetone

Keto-enol tautomerization equilibrium is greatly affected by solvent effects. The enol form predominates in low polarity solvents, where there are no solvent interactions competing with the formation of the hydrogen bond. In high polarity solvents, the keto form dominates. This is due to competition with the H-bond, preventing the stabilization of the enol by intramolecular H-bonding. Experimental data<sup>229–231</sup> from spectroscopic studies on keto-enol tautomerism are shown in Table 5.2.

Solvent	K <sub>T</sub>
Gas Phase	11.7
Cyclohexane	7.3
Tetrahydrofuran	7.2
Deuterochloroform	6.1
Dichloromethane	4.2
Ethanol	2.7
Methanol	2.3
Dimethylsulfoxide	1.7
Acetonitrile	1.2
Water	0.2

**Table 5.2:** Effect of solvent on keto-enol equilibrium constant for acetylacetone<sup>229-231</sup>

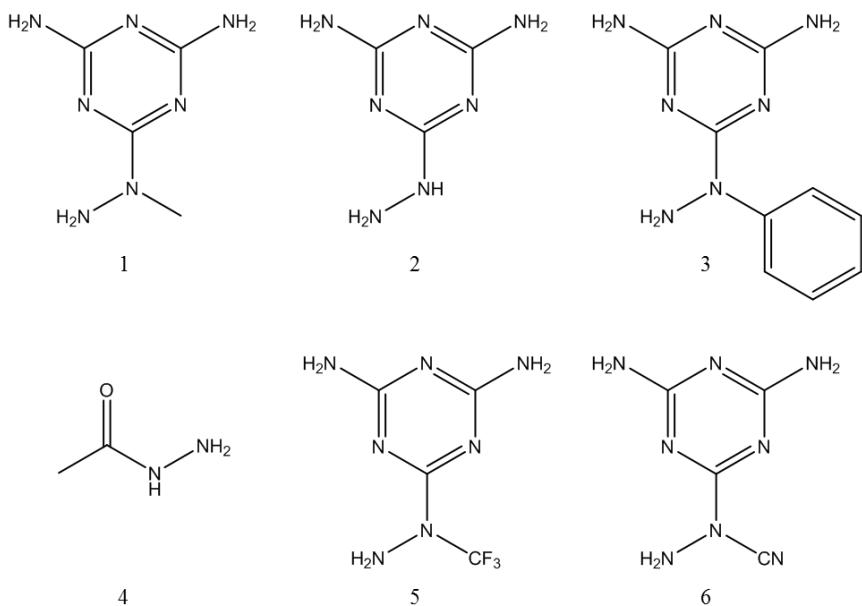
NMR analysis of hydrazones formed with a diketone showed the presence of a hydrogen bond to N1. Although in water this may not be the predominant species, we see from Table 5.2 that in deuterated chloroform, the enol is favored. Hydrolysis is observed in an HPLC mobile phase consisting of water/acetonitrile (75/25), so there will be a higher proportion of enol than for water alone. Also, the hydrazone is initially diluted in THF, where again the enol form is favored. This means that at time of addition, the enol form was instantaneously the predominant species in the hydrolysis media. The H-bond did not seem to be a big factor resisting hydrolysis as 1b, 3b, 1c, and 3c hydrolyzed. Experimentally, hydrazones 1b, 3b, 1d, and 3d hydrolyzed too fast to be monitored by HPLC, so the relative hydrolysis rate was estimated using mass spectroscopy.

Table 5.1 shows the half-lives ( $t_{1/2}$ ) in minutes and rate constants ( $k$ ) in  $\text{min}^{-1}$  at different pH conditions. It showed that compounds 1c hydrolyzed the fastest and compound 2b did not hydrolyze at all. Even though the Hydrolysis of 1c was fairly quick, it yielded multiple side products observed in HPLC which made integration of the peaks challenging. From table 1, we see the rate of hydrolysis decreases in the following order 1c>1a>2a>>1b~2c~2b.

Hydrazones 2c and 2b did not hydrolyze. The experimental half-life of 2a was determined to be 82 and 86 minutes at pH 4.0 and 5.0 respectively. The computational N1 proton binding affinity in Table 5.1 suggests that 2a should be the most stable hydrazone in this series. While the trends are the same, there is a discrepancy in the experimental data, with the half-life of 2a reported previously<sup>226</sup> as 390 and 770 minutes at pH 4.0 and 5.2 respectively. Relative orders of experimental stability to hydrolysis are shown in Figure 5.9. Table 5.1 also includes density functional theory (DFT) calculations of proton affinities ( $\Delta E_{\text{prot}}$ ) of N1 of the hydrazone. These values are assigned relative to 4d. Positive values correspond to stable proton binding. Acetyl hydrazine (4) was included as a reference for previous work carried out by Ji *et al.*<sup>226</sup>

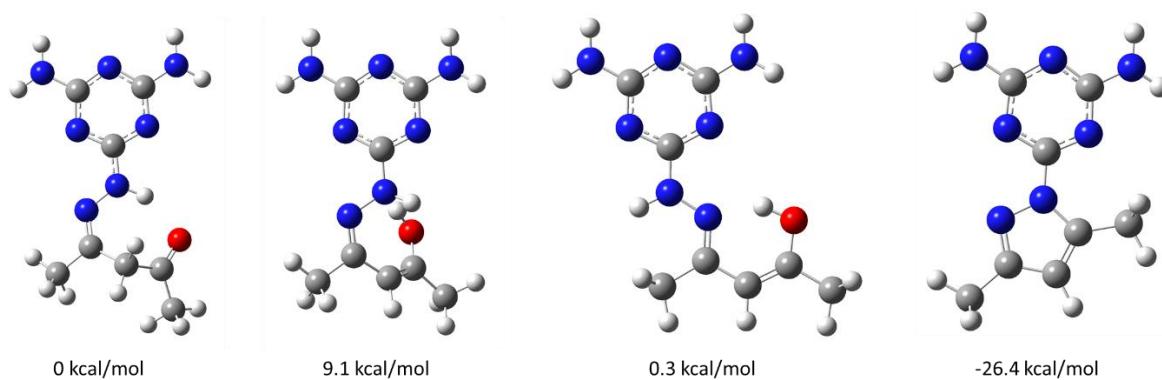
## 5.4 Computational Study of N1 Proton Affinities

The N1 proton affinity was investigated computationally for hydrazones formed from the six model hydrazines shown in Figure 5.10 and the aldehyde and ketones shown in Figure 5.7. Hydrazines 1-3 were also studied experimentally. Acetylhydrazine (4) was included for comparison to previous work<sup>226</sup> to confirm a stability reversal for triazinyl hydrazones at pH 5 compared to acyl hydrazones, which were found to be more stable to hydrolysis above pH 5. Hydrazines 5 and 6, with CF<sub>3</sub> and cyano substituents at N2 respectively, were included to show the effect of strongly electron withdrawing substituents at N2 on the N1 proton affinity more clearly.

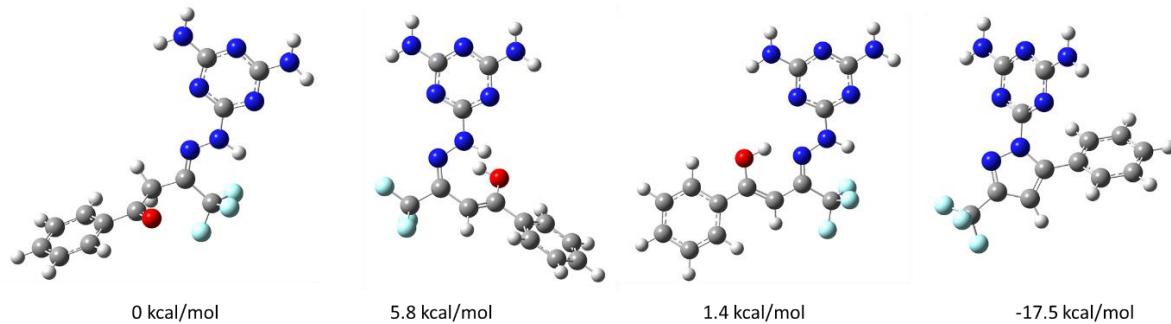


**Figure 5.10:** Triazinyl and acetyl hydrazine models (1-6) used for computational protonation studies.

Reaction with the diketones b and c can form a number of products, based on keto-enol tautomerization, as shown in Scheme 5.7. Relative energies of the triazinylhydrazine 2 products from reaction with diketone b and c are shown in Figure 5.11 and Figure 5.12 respectively. Atoms are color coded, with C gray, H white, N blue and O red.



**Figure 5.11:** B3LYP/6-31+G(d,p) optimized geometries and relative energies of keto-enol conformations and pyrazole ring for reaction of triazinylhydrazine 2 with diketone b



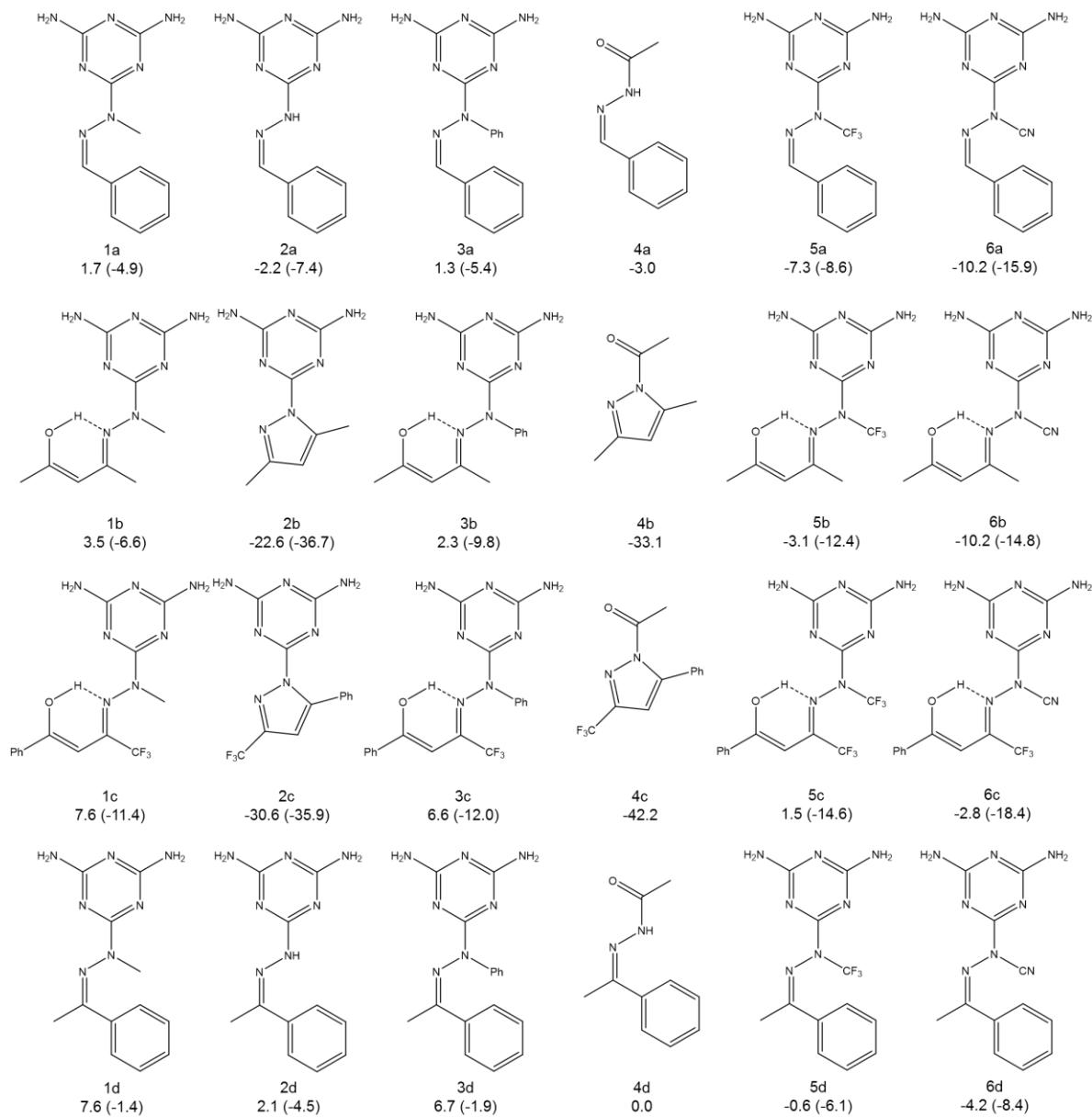
**Figure 5.12:** B3LYP/6-31+G(d,p) optimized geometries and relative energies of keto-enol conformations and pyrazole ring for reaction of triazinylhydrazine 2 with diketone c

Table 5.3 summarizes the calculated N1 proton affinities of the unprotonated and protonated hydrazone, as would be anticipated at pH 4. The relative proton affinities are shown graphically in Figure 5.17 and Figure 5.18.

Compound	$\Delta E_{\text{prot}}$ (kcal/mol)	$\Delta E_{\text{prot}}$ (kcal/mol) (Protonated triazine)
1a	1.7	-4.9
2a	-2.2	-7.4
3a	1.3	-5.4
4a	-3.0	-
5a	-7.3	-8.6
6a	-10.2	-15.9
1b	3.5	-6.6
2b	-22.6	-36.7
3b	2.3	-9.8
4b	-33.1	-
5b	-3.1	-12.4
6b	-10.2	-14.8
1c	7.6	-11.4
2c	-30.6	-35.9
3c	6.6	-12.0
4c	-42.2	-
5c	1.5	-14.6
6c	-2.8	-18.4
1d	7.6	-1.4
2d	2.1	-4.5
3d	6.7	-1.9
4d	0.0	-
5d	-0.6	-6.1
6d	-4.2	-8.4

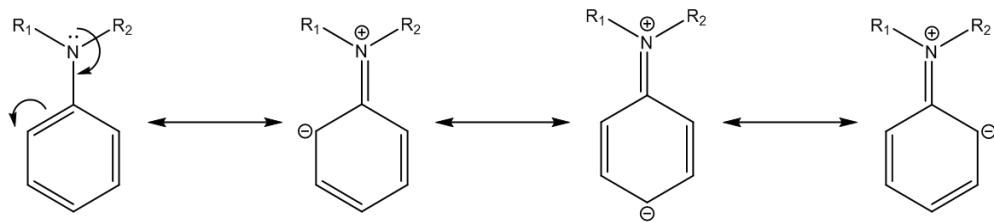
**Table 5.3:** Computed N1 proton affinities  $\Delta E_{\text{prot}}$  relative to 4d

Structures of the hydrazones investigated computationally are shown in Figure 5.13.



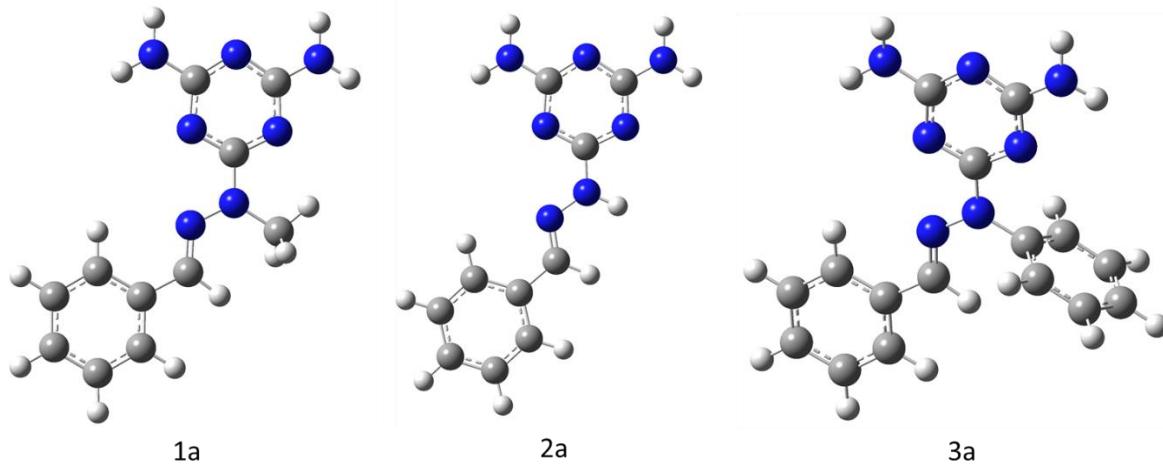
**Figure 5.13:** Structures of hydrazones investigated computationally and the N1 proton affinity (kcal/mol), with the protonated form in parentheses.

Phenyl is electron withdrawing by induction, which would tend to make nitrogen N2 less basic. We may also expect phenyl to be electron withdrawing via resonance, if the nitrogen lone pair is in plane with the aromatic ring, as shown in Figure 5.14.



**Figure 5.14:** Nitrogen in plane lone pair conjugation with aromatic ring.

If this were the case, we should see increased stability to hydrolysis for phenyl substituent at N2 compared to methyl or hydrogen. However, the phenyl lies out of the plane and therefore is not conjugated with the N-N bond, as shown in Figure 5.15.



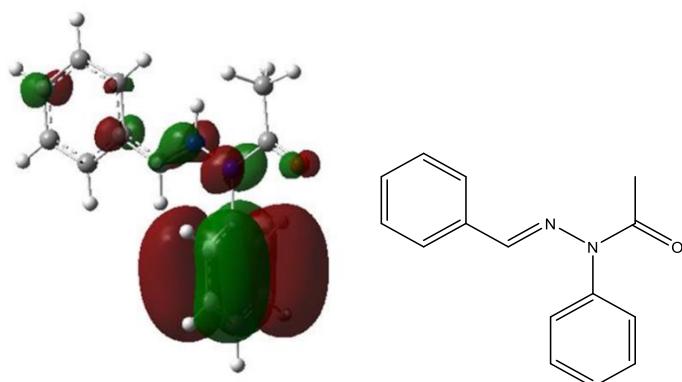
**Figure 5.15:** B3LYP/6-31+G(d,p) optimized geometries of the triazinyl hydrazones formed by reaction with benzaldehyde.

The phenyl also lies out of the plane for the protonated molecule. Hence, conjugation with the N-N bond cannot explain the results. The computed N1 partial charges for the neutral molecule and following triazine protonation are shown in Table 5.4.

Molecule	N2 Substituent	N1 Charge			
1	CH <sub>3</sub>	-0.69			
2	H	-0.76			
3	Ph	-0.53			
		Neutral Hydrazone	Protonated N1	Protonated Triazine	Protonated Triazine with Protonated N1
1a	CH <sub>3</sub>	0.08	0.50	0.13	0.52
2a	H	0.51	0.83	0.55	0.87
3a	Ph	0.62	1.01	0.66	1.03

**Table 5.4:** Computed N1 partial charges.

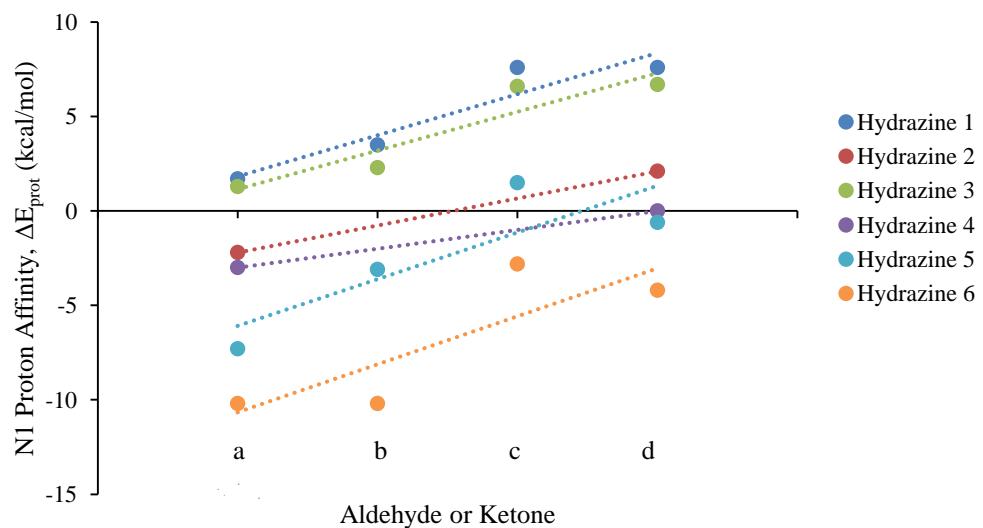
We can see from the computed N1 partial charges that methyl substituent at N2 is electron donating to N1 and that phenyl is electron withdrawing by induction as expected. This does not explain the increased hydrolysis rate or N1 proton affinity for hydrazones with N2 phenyl substituent. Looking at a simplified acylhydrazone, with a phenyl substituent at N2, we may be able to identify an area for more research to explain the proton affinity of N1. The HOMO of the protonated phenyl complex, may allow some hyperconjugation between the phenyl pi system and the C=N-N moiety, as shown in Figure 5.16.



**Figure 5.16:** HOMO of the protonated phenyl complex for acylhydrazine reacted with benzaldehyde along with chemical structure.

This may offer a way to delocalize the positive charge introduced upon protonation. The stabilization is not resonance but may be described as hyperconjugation. At this point there does not seem to be a perfect explanation for the influence of N2 phenyl substituent but further investigating this observation may be a focus of future work. As a result, we see a general experimental and computational stability of hydrazones formed from hydrazines 2>3>1. The N2 basicity effect is consistent with the accepted mechanism of hydrolysis of hydrazones as it starts with the protonation of the N1, as shown in Scheme 5.1.

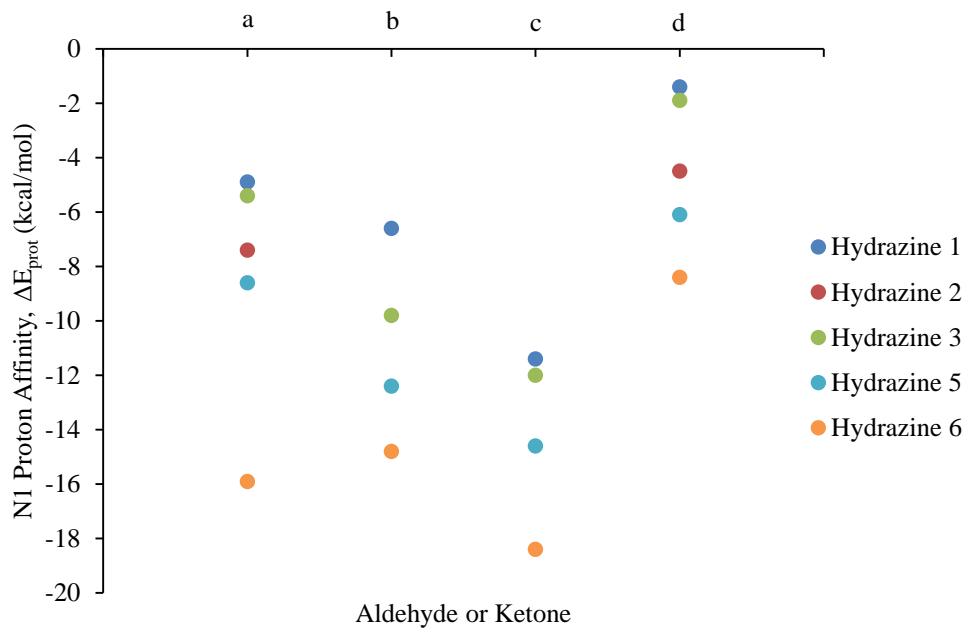
Computed binding affinities agree with experimental observations. The addition of  $\text{CF}_3$  and CN substituents at N2 further accentuate the effect of electron withdrawing substituents. The most stable geometries for 2b, 2c, 4b and 4c are pyrazole rings, formed via a second dehydration when the hydrazine reacts with a diketone. The unfavorable proton affinities for these hydrazones is based on protonation of N1 and the addition of water to break the pyrazole ring to form a reactive open chain molecule that can undergo hydrolysis via the mechanism shown in Scheme 5.1. The aldehyde and ketone reactants are shown in Figure 5.7.



Note. 2b, 2c, 4b and 4c are omitted for clarity of trends as their stability is based on formation of a pyrazole ring

**Figure 5.17:** Computed N1 proton affinities for hydrazines with aldehyde or ketone.

We can clearly see the stability trend in the hydrazines, with  $6>5>4>2>3>1$  in Figure 5.17, with benzaldehyde (a) giving the most stable hydrazone in each case. There is also an increase in stability at pH <5 consistent with previous work, as shown in Figure 5.18.

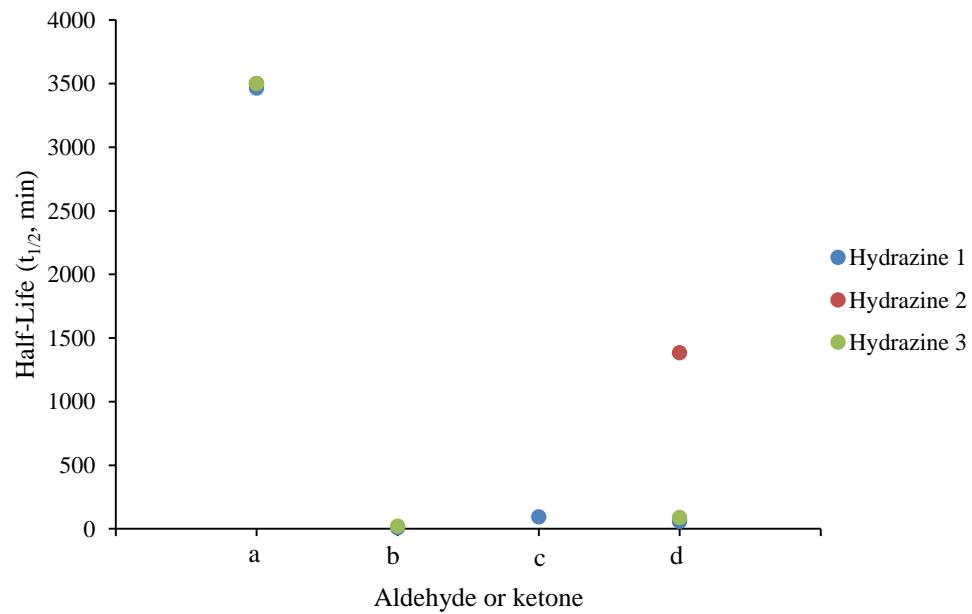


Note. 2b and 2c are omitted for clarity of trends as their stability is based on formation of a pyrazole ring

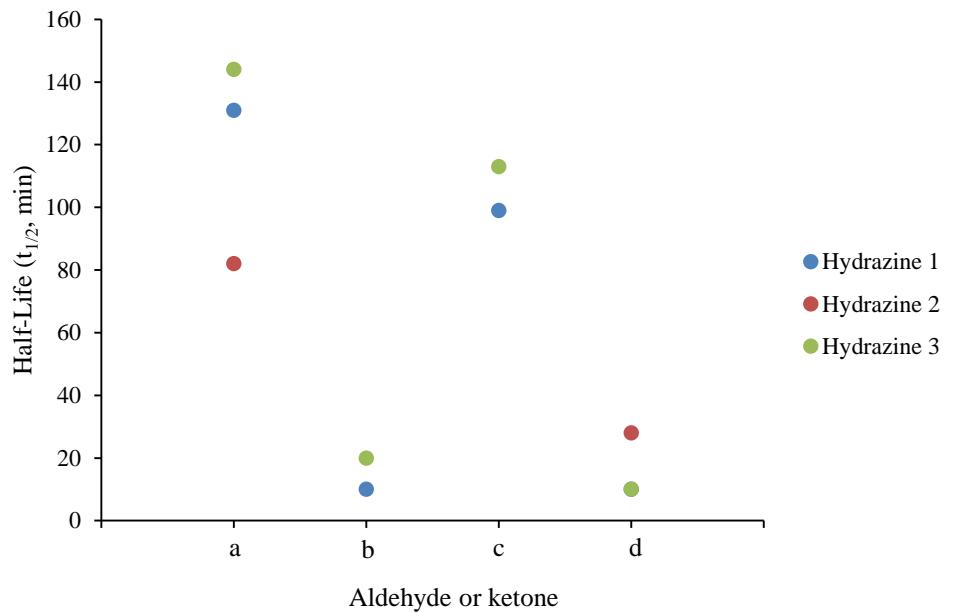
**Figure 5.18:** Computed N1 proton affinities for protonated triazines with aldehyde or ketone.

The relative stability trend of the hydrazines remains unchanged, with the omission of acetylhydrazine (4) that does not have a triazine for protonation, but the diketone with an electron withdrawing  $\text{CF}_3$  (c) now forms the most stable hydrazone. Table 5.3 shows that para protonation of the triazine reduces the proton affinity of N1 at pH 4 is reduced approximately 2-10 kcal/mol. This effect is greater in 5c and 6c where there are electron withdrawing groups

$\text{CF}_3$  and  $\text{CN}$  at N2 respectively and an electron withdrawing  $\text{CF}_3$  at C1. When compared to experimental data, as shown in Figure 5.19 and Figure 5.20, we see that the calculated data predicts the general experimental trends, although, it is worth noting that the experimental data set is smaller and the fast rates of hydrolysis in some cases lead to increased error due to over estimation of the hydrolysis rate by from mass spectroscopy.



**Figure 5.19:** Experimental half-lives for hydrolysis of hydrazones at pH 7.0.



**Figure 5.20:** Experimental half-lives for hydrolysis of hydrazones at pH 4.0.

In Figure 5.19, we can see the same general stability trend as for the calculations based on N1 proton affinity, again with benzaldehyde (a) forming the most stable hydrazone, when the triazine is not protonated at pH 7.0. At pH 4.0 (Figure 5.20), where the triazine can be protonated, experimentally we again see the same general trend in stability for the hydrazones formed from hydrazines, with the relative stability 2>3>1, although with a smaller data set and fast rates of hydrolysis, it is harder to discern a true trend as easily compared to the computational data. However, in general agreement with the calculated data, hydrazones formed with benzaldehyde (a) or the diketone with an electron withdrawing CF<sub>3</sub> (c) are the most stable to hydrolysis at pH 4.0. It should also be highlighted that, as previously noted, the experimentally determined half-life of 2d was lower than previously reported by Ji *et al.*<sup>226</sup> With hydrolysis rates determined previously, the stability trend predicted computationally is reproduced experimentally.

## 5.5 Concluding Remarks

The computational approach is validated through its agreement with experimental hydrolysis rates. As noted previously, it was found that triazinylhydrazones showed increased stability to hydrolysis at pH < 5 compared to a corresponding acetylhydrazone, which showed greater stability to hydrolysis at pH > 5. This study showed the same trends.

As far as substituents at N2, methyl increased the hydrolysis rate by making N1 more basic and susceptible to protonation, as shown by the calculated proton affinities, compared to hydrogen at N2 as a reference. Phenyl substituent at N2 decreased the hydrolysis rate compared to methyl due to the electron withdrawing nature of phenyl and the ability for conjugation of the nitrogen lone pair with the aromatic ring. CF<sub>3</sub> and CN substituents at N1 accentuated the effect, with CF<sub>3</sub><CN in terms of relative stability.

Para protonation of the triazine, as expected to occur at pH 4, reduced the N1 proton affinity relative to acetylhydrazone (4d) in all cases. This was also seen experimentally. The N2 substituent effects showed the same relative stability to hydrolysis at pH 4, with methyl stabilizing the charge at N2 by electron donation. These observations further support the conclusion in the previous study<sup>226</sup> that the Kalia/Raines mechanism for hydrolysis, resistance to protonation at N1, is the model that best explains hydrazone stability to hydrolysis.

At pH > 5, benzaldehyde (a), with its moderate electron donating, formed the most stable hydrazones, while diketone (c) is predicted to form the most stable hydrazones at pH 4.0. These same general trends were seen experimentally also. To further validate the computational model, a larger experimental sample set should be considered to cover more electron donating and withdrawing group effects at substituent on N2. Further study of the effect of aldehyde or ketone on hydrazone hydrolysis rate would also be warranted. The

stabilization of the protonated N1 by N2 phenyl substituents and the possible hyperconjugation may become a focus for future work on this topic.

## 5.6 Computational Section

All calculations use the Gaussian 09 electronic structure package.<sup>36</sup> Density functional theory (DFT) calculations<sup>62,63,68,158,165,182</sup> were used to calculate proton affinities ( $\Delta E_{\text{prot}}$ ) of the hydrazone N1 nitrogen. Calculations used the B3LYP exchange–correlation functional,<sup>62,63,158,165</sup> the 6-31+G(d,p) basis set,<sup>182</sup> and the SMD continuum model for aqueous solvent<sup>68</sup> as implemented in the Gaussian 09 suite of programs. To simplify the computational model aminoethoxyethanol groups on the triazinylhydrazine were replaced with primary amines. This reduced the number of rotational conformers that needed to be considered, as well as removing steric interactions with N2 and aldehyde or ketone substituents. This approach reduced calculation time and computational cost and meant that by having a rigid triazinylhydrazine, all changes in protonation energy at N1 were due to changes to the N2 substituent for a family of molecules. The N1 proton affinity values are assigned relative to molecule 4d, with positive values corresponding to stable proton binding. Computed geometries and total energies of all species are included in Appendix C.

## APPENDIX A

### Electronic Structure Theory

Quantum mechanics focuses on the theory of nature at small scales and low energy levels of atoms and subatomic particles. Unlike classical physics, energy in quantum mechanics is restricted to discrete packets, or energy levels. Quantum chemistry deals with the application of quantum mechanics in physical models and experiments or chemical systems.

To describe molecules, quantum chemistry does not just consider an arrangement of atoms but more generally a collection of charged particles, positive nuclei and negative electrons. The interaction between charged particles is described by the coulomb potential (Equation A.1).

$$V_{ij} = V(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} = \frac{q_i q_j}{r_{ij}}$$

**Equation A.1:** Coulomb potential

Classical mechanics is described by Newton's 2<sup>nd</sup> Law (Equation A.2: ).

$$F = ma$$

$$-\frac{dV}{dr} = m \frac{d^2r}{dt^2}$$

**Equation A.2:** Newton's 2<sup>nd</sup> Law

Where  $F$  is a force and  $a$  the acceleration.  $r$  is the position vector and  $m$  the particle mass.

Quantum chemistry deals with wave particle duality and is based on the Schrödinger equation (Equation A.3).

$$i\hbar \frac{\partial}{\partial t} \Psi(r, t) = \hat{H} \Psi(r, t) = \left[ \frac{-\hbar^2}{2\mu} \nabla^2 + V(r, t) \right] \Psi(r, t)$$

**Equation A.3:** Schrödinger equation

Where  $\hat{H}$  is the Hamiltonian operator, which characterizes the total energy of any given wave function, and  $\Psi$  the wave function of the system,  $\mu$  is the reduced mass of a particle,  $V$  is its potential energy, and  $\nabla^2$  the Laplacian of electron density.

The Hamiltonian ( $\hat{H}$ ) for a system with N-particles is dependent upon the sum of kinetic and potential energy (Equation A.4).

$$\hat{H} = \hat{T} + \hat{V}$$

**Equation A.4:** Sum of kinetic (T) and potential energy (V)

$$\hat{T} = \sum_{i=1}^N \hat{T}_i = - \sum_{j=1}^N \frac{\hbar^2}{2m_i} \nabla_i^2 = - \sum_{i=1}^N \frac{\hbar^2}{2m_i} \left( \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \right)$$

**Equation A.5:** Kinetic energy

$$\nabla_i^2 = \left( \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \right)$$

**Equation A.6:** Laplacian operator

$$\hat{V} = \sum_{i=1}^N \sum_{j>1}^N V_{ij} = \sum_{i=1}^N \sum_{j>1}^N \frac{q_i q_j}{r_{ij}}$$

**Equation A.7:** Potential energy

When used in the Schrödinger equation, the dynamics of all electrons and nuclei in a molecule or atom are taken into account.

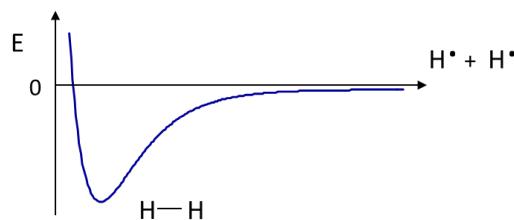
### Born-Oppenheimer Approximation

Nuclei have greater mass and smaller velocities as compared to electrons. Due to this fact, the Schrödinger equation is dealt with in two parts, 1) electronic wave function for a fixed nuclear geometry and 2) the nuclear wave function, where electronic energy replaces potential energy. The Born-Oppenheimer approximation is the assumption that the motion of atomic nuclei and electrons can be separated, with the wave function of a molecule being dealt with in terms of its electronic and nuclear (vibrational, rotational) components (Equation A.8). The approximation means that the electronic wave function depends only on the positions of the nuclei.

$$\Psi_{total} = \varphi_{electronic} \times \varphi_{nuclear}$$

**Equation A.8:** Born-Oppenheimer approximation

The physical implication is that the nuclei move on a potential energy surface (PES), which are solutions to the electronic Schrödinger equation (Figure A.1).



**Figure A.1:** Potential energy surface for  $\text{H}_2$

## Self-Consistent Field (SCF) Theory

We can only get an exact solution to the Schrödinger equation for one-electron systems. The solution to this problem is to generate approximate solutions. If an approximate wave function is used, the energy must be greater than or equal to the exact energy. This is achieved by generating a trial function that has a number of adjustable parameters and then minimizing the energy as a function of those parameters. Energy is expressed as an expectation value of the Hamiltonian operator (Equation A.9).

$$E = \frac{\int \Psi^* \hat{H}_e \Psi dr}{\int \Psi^* \Psi dr}$$

**Equation A.9:** Energy as an expectation value of the Hamiltonian operator

With bracket notation,  $\int \Psi^* \hat{H}_e \Psi dr = \langle \Psi | \hat{H}_e | \Psi \rangle$  and  $\int \Psi^* \Psi dr = \langle \Psi | \Psi \rangle$  we get a combined integration (Equation A.10).

$$E = \frac{\langle \Psi | \hat{H}_e | \Psi \rangle}{\langle \Psi | \Psi \rangle}$$

**Equation A.10:** Integration over all coordinates

If the wave functions are orthogonal and normalized (orthonormal), where  $\langle \Psi_i | \Psi_j \rangle = \delta_{ij}$  for  $\delta_{ij} = 1$  and  $\delta_{ij} = 0$ , we return (Equation A.11).

$$E = \langle \Psi | \hat{H}_e | \Psi \rangle$$

**Equation A.11:** Energy

Taking into account the fact that electrons are fermions, with  $S = 1/2$ , the total electronic wave function must be antisymmetric with respect to interchange of any two electron coordinates.

This is based on the Pauli exclusion principle, where no two electrons may have the same set of quantum numbers. For a two electron system, eg. He or H<sub>2</sub>, we get Equation A.12. Where for He  $\phi_1 = \phi_2 = 1s$  and for H<sub>2</sub>  $\phi_1 = \phi_2 = \phi_{bonding\ orbital}$ .

$$\Phi(1,2) = \phi_1\alpha(1)\phi_2\beta(2) - \phi_1\alpha(2)\phi_2\beta(1)$$

**Equation A.12:** Antisymmetric wave function to describe the ground state

Each electron occupies a spin orbital, which is a product of both spatial and spin functions. As spin functions are orthonormal, where  $\langle \alpha | \alpha \rangle = \langle \beta | \beta \rangle = 1$ ;  $\langle \alpha | \beta \rangle = \langle \beta | \alpha \rangle = 0$ , interchange of electron coordinates looks like Equation A.13.

$$\Phi(2,1) = \phi_1\alpha(2)\phi_2\beta(1) - \phi_1\alpha(1)\phi_2\beta(2)$$

$$\Phi(2,1) = -\Phi(1,2)$$

**Equation A.13:** Interchange of electron coordinates

More generally, antisymmetric electron wave functions are represented as determinants (Equation A.14 and Equation A.15).

$$\Phi(1,2) = \begin{vmatrix} \phi_1\alpha(1) & \phi_1\beta(1) \\ \phi_1\alpha(2) & \phi_1\beta(2) \end{vmatrix} = \phi_1\alpha(1)\phi_2\beta(2) - \phi_1\alpha(2)\phi_2\beta(1)$$

**Equation A.14:** Two-electron determinant

$$\Phi_{SD} = \begin{bmatrix} \phi_1(1) & \cdots & \phi_N(1) \\ \vdots & \ddots & \vdots \\ \phi_1(N) & \cdots & \phi_N(N) \end{bmatrix}, \langle \phi_i | \phi_j \rangle = \delta_{ij}$$

**Equation A.15:** Slater determinant (SD)

In the Slater determinant matrix, columns are one-electron wave functions, molecular orbitals and the rows contain electron coordinates. The trial wave function will consist of a single SD.

## Hartree-Fock Theory

The Hartree-Fock (HF) method is an approximation for the determination of the wave function and energy of a quantum many-body system in a stationary state. Since there are no exact solutions to the Schrödinger equation for many electron systems, the HF method uses five approximations:

- 1) The Born-Oppenheimer approximation is assumed.
- 2) Relativistic effects are completely neglected.
- 3) The variational solution is assumed to be a linear combination of a finite number of basis functions, which are usually chosen to be orthogonal. The finite basis set is assumed to be approximately complete.
- 4) Each energy eigenfunction is assumed to be describable by a single SD.
- 5) SCF is implied and coulomb interaction is neglected.

All terms of the Hamiltonian are expressed as the sum of one-electron operators (Equation A.16).

$$\hat{F}[\{\phi_j\}](1) = \hat{H}^{core}(1) + \sum_{j=1}^{N/2} [2\hat{j}_j(1) - \hat{K}_j(1)]$$

**Equation A.16:** Fock operator

Where  $\hat{F}[\{\phi_j\}](1)$  is the one-electron Fock operator, generated by orbitals  $\phi_j$  and  $\hat{H}^{core}(1)$  is the one-electron core Hamiltonian.  $\hat{j}_j(1)$  is the Coulomb operator (electron repulsion due to each of the two electrons in the  $j$  orbital) and  $\hat{K}_j(1)$  is the exchange operator (electron exchange

energy due to the antisymmetry of the total n-electron wave function. For a one electron system, the exchange term cancels out the electronic coulomb repulsion term, meaning the HF theory is exact for hydrogenic atoms. Of the approximations made in HF theory, it is generally the neglect of electron correlation that leads to the biggest deviations from experimental results. Many post HF methods include electron correlation in the multi-electron wave function.

## Semi-Empirical Methods

Semi-empirical quantum chemistry methods are based on HF but make other approximations and obtain some parameters from empirical data. They can prove useful for dealing with large molecules where HF method would be too costly. The inclusion of some empirical parameters also adds a degree of electron correlation. Two-electron integrals may be approximated or left out entirely. To correct for the loss, parameterization is employed. The results are fitted by a set of parameters to generate results that agree with experimental or *ab initio* results. While semi-empirical calculations may be fast, the results depend upon the molecule of question being similar enough to database molecules used to parameterize the method initially. Calculations that use the semi-empirical method to greatest effect have generally been organic molecules, of moderate size, and only containing a few elements.

Parameterized model number 3<sup>150</sup> (PM3) uses two Gaussian functions for core repulsion. It was originally produced with parameters for H, C, N, O, F, Al, Si, P, S, Cl, Br and I. Parametric method 6<sup>153</sup> (PM6) was developed with parameterization of 70 elements. In expanding the number of elements, a larger reference database was required. However, due to a lack of experimental data for a number of elements, much of the reference data was generated from *ab initio* calculations.

## Density Functional Theory (DFT)

DFT is a modelling method used to investigate ground state electronic structure of many-body systems. The system properties are determined by functionals (functions of another function), which in this case is the spatially dependent electron density. DFT is based on two Hohenberg-Kohn theorems. The first states that ground state properties of a many-electron system are determined by an electron density ( $\rho$ ) as shown in Equation A.17.

$$\rho(\vec{r}) = N \int d^3r_2 \dots \int d^3r_N \Psi^*(\vec{r}, \vec{r}_2, \dots \vec{r}_N) \Psi(\vec{r}, \vec{r}_2, \dots \vec{r}_N)$$

**Equation A.17:** Electron density

The number of electrons (N) can be found by integrating the electron density over all space (Equation A.18).

$$N = \int \rho(r) dr$$

**Equation A.18:** Number of electrons

The second defines an energy functional for the system where the ground state electron density gives the ground state energy (Equation A.19).

$$E_0 = T[\rho_0] + V_{ext}[\rho_0] + J[\rho_0] + E_{xc}[\rho_0]$$

**Equation A.19:** Ground state energy

Where  $T[\rho_0]$  is the electronic kinetic energy,  $J[\rho_0]$  is the mean field electronic coulomb repulsion energy on an electron, and  $E_{xc}[\rho_0]$  is an exchange-correlation functional. Kinetic

energy can be approximated using the wave function for non-interacting electronic system (Equation A.20).

$$T_M[\rho_0] = -\frac{1}{2} \sum_i \langle \varphi_i | \nabla^2 | \varphi_i \rangle$$

**Equation A.20:** Kinetic energy approximation

Coulombic attraction between electrons is given by Equation A.21.

$$V_{ext}[\rho_0] = \int \hat{V}_{ext} \rho(r) dr$$

**Equation A.21:** Coulombic attraction

The electron-electron repulsion energy is given by Equation A.22.

$$J[\rho_0] = \frac{1}{2} \int \int \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} dr_1 dr_2$$

**Equation A.22:** Electron-electron repulsion energy

The exchange-correlation functional  $E_{xc}[\rho_0]$  contains the remaining description of the electron-electron interaction. Different approximations may be used for  $E_{xc}[\rho_0]$  but the exact form is not known, except for the case of the uniform electron gas (UEG). This inability to properly describe exchange and correlation energy means DFT method calculations do not converge. For this reason, benchmarking is carried out against experiments or high level correlated methods to gauge the accuracy of a DFT method.

## Local Spin Density Approximation (LSDA)

Local spin density approximations (LSDA) (or local density approximation (LDA)) are a class of approximations to the exchange-correlation (XC) energy functional in DFT. The approximations depend only on the value of the electron density at each point in space. The exchange energy is shown in Equation A.23.

$$E_x^{LDA} = -\frac{3}{4} \left(\frac{3}{\pi}\right)^{1/3} \int \rho(r)^{4/3} d^3r$$

**Equation A.23:** Local density approximation for exchange energy

The approach is useful for systems that have a uniform electron density, such as bulk transition metals or systems that approximate the UEG model. It tends to break down in cases where there are rapid variations in the electron density, such as isolated atoms or molecules. As LDA assumes a uniform electron density, the method tends to overestimate the exchange-correlation energy. When the electron density is replaced with spin densities  $\rho_\alpha$  and  $\rho_\beta$ , the electron density becomes  $\rho = \rho_\alpha + \rho_\beta$  and the approximation is referred to as the local spin density approximation (LSDA).

## Generalized Gradient Approximation (GGA)

Generalized gradient approximation (GGA), as the name implies, includes the gradient of the electron density  $\nabla\rho$  in the exchange functional. GGA is intended to address the overestimation of exchange-correlation energy by LDA by using electron density gradient to account for non-homogeneity of the actual electron density. The general form of PBE exchange is shown in Equation A.24.

$$E_x^{GGA}[\rho(r), \nabla\rho(r)] = \int d^3r \rho(r) E_x^{LDA}(\rho) F_x$$

**Equation A.24:** PBE exchange

$F_x$  is the exchange enhance factor, which is a function of the dimensionless density gradient ( $s$ ), as shown in Equation A.25.  $\mu$  and  $k$  are empirical parameters equaling 0.235 and 0.804 respectively.

$$F_x = 1 + k - \frac{k}{1 + \frac{\mu s^2}{k}}$$

$$s = \frac{|\nabla\rho|}{2k_F\rho}$$

**Equation A.25:** Exchange enhance factor

### meta-Generalized Gradient Approximation (meta-GGA)

Meta-GGA functionals include the exchange functional's dependence on non-interacting kinetic energy density ( $\tau$ ), where  $\tau \equiv \frac{1}{2}\sum_i^{occ} |\nabla\varphi_i|^2$ . TPSS and MO6L are both examples of meta-GGA methods that were used in Chapter 4. The general MO6L functional form is shown in Equation A.26.

$$E_x^{MO6L}[\rho(r), \nabla\rho(r), \tau] = \int d^3r [E_x^{PBE}f(\omega) + E_x^{LDA}h_x(s, z)]$$

**Equation A.26:** MO6L general functional

$f(\omega)$  is the spin kinetic energy density enhancement factor  $f(\omega) = \sum_{i=0} \alpha_i \omega_i$ .  $\omega$  is the function of spin kinetic energy density  $\tau_\sigma$ , where  $\omega_\sigma = \frac{(\frac{\tau_\sigma^{LSDA}}{\tau_{\sigma-1}})}{(\frac{\tau_\sigma^{LSDA}}{\tau_{\sigma+1}})}$ . The spin kinetic energy density parameters are defined as,  $\tau_\sigma^{LSDA} \equiv \frac{3}{10} (6\pi^2)^{2/3} \rho^{5/3}$  and  $\tau_\sigma \equiv \frac{1}{2} \sum_i^{occ} |\nabla \varphi_{i\sigma}|^2$ .

### Global Hybrid Generalized Gradient Approximation (Hybrid-GGA)

Hybrid functionals combine exact exchange energy with LSDA, GGA or meta-GGA energy in different proportions. The inclusion of exact exchange makes these hybrid functionals a preferred method for quantum chemical calculations. The Becke 3 parameter Lee Yang Parr (B3LYP) method is used extensively throughout this work. It is a hybrid encompassing fractions of HF, LDA and GGA exchange, as shown in Equation A.27.

$$E_{xc}^{B3LYP} = E_x^{LDA} + \alpha_0(E_x^{HF} - E_x^{LDA}) + \alpha_x(E_x^{GGA} - E_x^{LDA}) + E_c^{LDA} + \alpha_c(E_c^{GGA} - E_c^{LDA})$$

**Equation A.27:** B3LYP exchange-correlation functional

Where  $\alpha_0 = 0.20$ ,  $\alpha_x = 0.72$  and  $\alpha_c = 0.81$ .  $E_x^{GGA}$  and  $E_c^{GGA}$  are generalized gradient approximations from the Becke 88 exchange functional and the correlation functional of Lee, Yang and Parr.  $E_c^{LDA}$  is the VWN local density approximation to the correlation functional. The three parameters were taken from Becke's original parameters in the B3PW91 functional.

### Dispersion-Corrected Hybrids

Density functional theory does not adequately describe intermolecular forces. These forces include van der Waals forces (dispersion). For systems where dispersion forces play a major role, incomplete treatment of dispersion can affect the accuracy of DFT calculations. Systems dominated by dispersion include interacting noble gas atoms and competing interactions in

biomolecules. The development of new DFT methods<sup>172</sup> designed to overcome this problem by including terms for dispersion are underway, with the general approach shown in Equation A.28.

$$E_{DFT-D} = E_{KS-DFT} + E_{disp}$$

**Equation A.28:** General DFT method with dispersion correction

The total energy is computed as the sum of a KS-DFT portion, using the  $\omega$ B97X functional, along with an empirical dispersion correction, given in Equation A.29.

$$E_{disp} = - \sum_{i=1}^{N_{at}-1} \sum_{j=i+1}^{N_{at}} \frac{C_6^{ij}}{R_{ij}^6} f_{damp}(R_{ij})$$

**Equation A.29:** Dispersion correction

## Methods Investigated

For reference, the methods used during this study are shown in Table A.1.

Method	Description
HF	Exchange Only
PM3	Semi-Empirical
PM6	
SVWN5	Local Spin Density Approximation
PBEPBE	Generalized Gradient Approximation
B3PW91	Global Hybrid-GGA
PBE1PBE	
B3LYP	
mPW1PW91	
M06	
TPSSTPSS	Meta-GGA
M06L	
$\omega$ B97XD	Dispersion-Corrected Range-Separated Hybrid

**Table A.1:** Computational methods used during this study and their level of theory

## Basis Sets

A basis set is a set of functions used to represent the electronic wave function. The wave function is described by a finite set of basis sets and using a larger basis set can improve the computational model but at increased computational cost. Since the molecular orbitals (MOs) are not known for most molecular calculations, the unknown MOs are expressed as a set of functions in a basis set. A linear expansion of Gaussian functions in the MO are commonly used as they are easy to handle computationally. Typically two types of atomic orbitals are used: Slater-type orbitals or Gaussian-type orbitals. There are two criteria for selection:

- 1) Physically meaningful
- 2) Computation of the integrals should be tractable

If the finite basis set is expanded to an infinite (complete) set of functions, calculations using such a set are said to approach the complete basis set (CBS) limit.

## Slater Type Orbital (STO)

Slater-type orbitals (STOs) are used as atomic orbitals in the linear combination of atomic orbitals molecular orbital method. STO basis functions have hydrogen-like features (Equation A.30).

$$x(\zeta, n, l, m, r, \theta, \phi) = N r^{n-1} e^{-\zeta r} Y_{lm}(\theta, \phi)$$

**Equation A.30:** Slater-type orbitals

N is the normalization factor and  $\zeta$  is the exponent factor that describes how fast the wave function decays. n, l and m represent principle, angular momentum and magnetic quantum

numbers respectively. Spherical coordinates are represented by  $r$ ,  $\theta$  and  $\phi$ . The orbital shape is determined by  $Y_{lm}(\theta, \phi)$ . This represents the angular dependenceportion as in orbitals of a hydrogenic atom.

### Gaussian Type Orbital (GTO)

Gaussian-type orbitals (GTOs) are functions used as atomic orbitals in the LCAO method for representation of electron orbitals in molecules. They are not as computationally expensive as STOs, so are commonly used. The general GTO function is shown in Equation A.31.

$$g(\alpha, a, b, c, x, y, z\lambda) = N \text{Exp}[-\alpha\lambda^2 r^2] x^a y^b z^c$$

**Equation A.31:** Gaussian-type orbitals

$N$  is the normalization constant,  $\alpha$  and  $\lambda$  are exponent coefficient and scale factor respectively. The sum of the exponents ( $K = a + b + c$ ) of Cartesian coordinates ( $x, y, z$ ) determines the type of basis function. Orbital types are shown in Table A.2.

<b>K</b>	<b>g Orbital type</b>
0	s-type
1	p-type
2	d-type
3	f-type

**Table A.2:** Correspondence of orbital type with sum of exponents

### Minimal Basis Sets

STO-nG basis sets are minimal basis sets. There are  $n$  primitive Gaussian orbitals fitted to a single STO. A minimal basis set only uses sufficient orbitals to contain all the electrons in a neutral atom. The orbital breakdown for the first 6 atoms are shown in Table A.3.

Atom	Orbitals described in minimal basis set
H	1s <sup>1</sup>
He	1s <sup>2</sup>
Li	1s <sup>2</sup> 2s <sup>1</sup>
Ba	1s <sup>2</sup> 2s <sup>1</sup>
B	1s <sup>2</sup> 2s <sup>2</sup> 2p <sup>1</sup>
C	1s <sup>2</sup> 2s <sup>2</sup> 2p <sup>2</sup>

**Table A.3:** Orbitals required to describe the first 6 atoms using minimal basis sets

Core and valence orbitals are represented the same number of primitive Gaussian functions  $\phi_i$ .

STO notation denotes the number of primitive functions. STO-3G basis sets are linear combination of 3 primitive Gaussian functions, as shown in Equation A.32.

$$\Psi_{STO-3G} = c_1 \phi_1 + c_2 \phi_2 + c_3 \phi_3$$

$$\phi_1 = \left(\frac{2\alpha_1}{\pi}\right)^{3/4} e^{-\alpha_1 r^2}$$

$$\phi_2 = \left(\frac{2\alpha_2}{\pi}\right)^{3/4} e^{-\alpha_2 r^2}$$

$$\phi_3 = \left(\frac{2\alpha_3}{\pi}\right)^{3/4} e^{-\alpha_3 r^2}$$

**Equation A.32:** STO-3G basis set

### Pople Basis Sets

These basis sets were developed by Sir John Anthony Pople, who, along with Walter Kohn, was awarded the Nobel Prize in chemistry in 1998. The basis sets are generally known as “split-valence” basis sets, due to the fact they have separate functions for core and valence orbitals. The basis set notation n-XYZG, conveys information about the core and orbital functions. n represents the number of GTOs comprising each core atomic orbital basis function. X, Y and Z represent the linear combination of GTOs used to make up the valence orbital basis functions. For example:

3-21G denotes a split-valence double zeta basis set

$n = 3$  GTOs describing core orbital,  $X = 2$  GTOs describing the first function of the double zeta and  $Y = 1$  GTOs describing the second function of the double zeta, equating to a total of 3 GTOs for valence orbitals.

6-311G denotes a slit-valence triple zeta basis set

$n = 6$  GTOs describing core orbital,  $X = 3$  GTOs describing the first function of the double zeta,  $Y = 1$  GTOs describing the second function of the double zeta and  $Z = 1$  GTOs describing the third function of the triple zeta, equating to a total of 5 GTOs for valence orbitals.

Further modifications can be made to the basis set, increasing the size to come closer to the exact electronic energy and wave function. One approach is to add polarization functions. A polarization function is any higher angular momentum orbital used in a basis function not normally occupied in the isolated atom. A hydrogen atom only occupies the s-type orbital. Adding p- and d-type orbitals to a hydrogen atom would be polarization functions. Likewise, for first row elements like carbon, only the s- and p-type orbitals are occupied. Adding d- and f-type orbitals would be polarization functions.

For example:

6-311G (d,p) denotes 1 d function and 1 p function to Hydrogen atoms

6-311G (3df,3pd) denotes 3 sets of d functions and 1 set of f functions on heavy atoms and 3

sets of p functions and 1 set of d functions on hydrogens

Diffuse functions can also be added to the basis set, meaning the electron is held far away from the nucleus. These functions are necessary for anions and very electronegative atoms, such as fluorine, with a lot of electron density. Diffuse functions in pople basis sets are indicated by + or ++.

For example:

6-311+G(d,p) denotes 1 set of sp-type diffuse basis functions is added to non-hydrogen atoms (4 diffuse basis functions per atom)

6-311++G (d,p) denotes that 1 set of sp-type diffuse functions is added to each non-hydrogen atom and 1 s-type diffuse function is added to hydrogen atoms.

### Correlation Consistent Basis Sets

The correlation-consistent basis sets developed by Thom Dunning are often favored in post-HF calculations as they are designed to converge to the complete basis set limit. The basis sets are typically denoted as cc-pVNZ (where N = D, T, Q, 5, 6 ....), with D, T and Q representing double, triple and quadruple respectively. cc-p represents “correlation-consistent polarized” and V indicates they are valence-only basis sets.

### Basis Set Superposition Error (BSSE)

Calculations using finite basis sets are susceptible to basis set superposition error (BSSE). This error can be described when looking at the complexation of two species, as shown in Equation A.33.

$$\Delta E_{bind} = E_{AB} - (E_A + E_B)$$

**Equation A.33:** Calculating the binding energy of complexation of two species, A and B

As the atoms of interacting molecules approach and interact, as in the complex AB, the basis functions of A and B overlap, meaning complex AB is described by more basis sets than A or B, which relatively stabilizes AB. This stabilization lowers the complexation energy, leading

to an overestimation of the binding energy,  $\Delta E_{bind}$ . BSSE can be corrected by using the additional basis sets generated by the overlap, to calculate the energy of the other isolated molecule. This approach is known as counterpoise correction of BSSE.<sup>233,234</sup> Calculations can also make use of larger basis sets to reduce BSSE.

## APPENDIX B

### From Electronic Structure to Thermochemistry

While a molecule has constant translational and rotational motion, atoms are in periodic motion, known as molecular vibration. The frequency of the periodic motion is the vibrational frequency. A molecule with  $N$  atoms has  $3N - 6$  normal modes of vibration, while a linear molecule has  $3N - 5$  modes. Molecular vibrations are treated by Newtonian mechanics, under the assumption that each vibration is a spring obeying Hooke's law (Equation B.1).

$$F = kX$$

**Equation B.1:** Hooke's law

Where  $F$  is an applied force,  $X$  is the displacement from the rest position and  $k$  is the force constant. For a spring,  $F$  is the restoring force. Since the restoring force is opposite to the direction of displacement, Hooke's law for a spring is written,  $F = -kX$ . Applying Newton's

**$d^2X/dt^2$**

**Equation B.2),** we recover reduced mass,  $\mu$ , multiplied by the acceleration.

$$F = ma$$

$$F = \mu \frac{d^2X}{dt^2}$$

**Equation B.2:** Newton's second law of motion

Combining Equation B.1 and Equation B.2, we recover a differential equation, from which the simple harmonic motion can be recovered (Equation B.3).

$$0 = kX + \mu \frac{d^2X}{dt^2}$$

$$X(t) = A \cos(2\pi\nu t)$$

**Equation B.3:** Simple harmonic motion

$A$  is maximum amplitude of vibration coordinate  $X$ , where  $\nu = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}}$  and  $\frac{1}{\mu} = \frac{1}{m_A} + \frac{1}{m_B}$ . The force constant is then equal to the second derivative of the potential energy (Equation B.4).

$$k = \frac{d^2E}{dX^2}$$

**Equation B.4:** Second derivative of potential energy

The energy states for each coordinate can be found by solving the Schrödinger equation (Equation B.5).

$$E_n = \left(n + \frac{1}{2}\right) \nu = h \left(n + \frac{1}{2}\right) \frac{1}{2\pi} \sqrt{\frac{k}{\mu}} = hc\nu \left(n + \frac{1}{2}\right)$$

**Equation B.5:** Potential energy of state n

## Calculation of Thermochemical Data

Statistical models are used to compute data such as Gibbs energies, enthalpy and entropy. From starting thermochemistry, we know that the Enthalpy ( $H$ ) is given by Equation B.6, and that the Gibbs energy ( $G$ ) is given by Equation B.7.

$$H = U + pV$$

**Equation B.6:** Enthalpy

$$G = H - TS_{Tot}$$

**Equation B.7:** Gibbs energy

Assuming degrees of freedom are independent, total internal energy,  $U$ , can be divided into energy contributions from molecular translation ( $U_{Tran}$ ), rotation ( $U_{Rot}$ ), vibration ( $U_{Vib}$ ) and electronic energy ( $E_{Elec}$ ), as shown in Equation B.8.  $E_{BO}$  is the electronic ground state using the Born-Oppenheimer approximation.

$$U = U_{Tran} + U_{Rot} + U_{Vib} + E_{Elec} + E_{BO}$$

**Equation B.8:** Total internal energy

Contribution of translational motion to total energy is shown in Equation B.9.

$$\begin{aligned} U_{Tran} &= \frac{(-N)}{q_{Tran}} \left( \frac{\partial q_{Tran}}{\partial \beta} \right)_V \\ q_{Tran} &= \left( \frac{2\pi\mu}{\beta} \right)^{2/3} \frac{V}{h^3} \\ U_{Tran} &= \frac{(-N)}{\left( \frac{2\pi\mu}{\beta} \right)^{2/3} \frac{V}{h^3}} \left( \frac{\partial \left( \left( \frac{2\pi\mu}{\beta} \right)^{2/3} \frac{V}{h^3} \right)}{\partial \beta} \right)_V = \frac{(-N)}{\left( \frac{1}{\beta} \right)^{2/3}} \left( \frac{\partial \left( \frac{1}{\beta} \right)^{3/2}}{\partial \beta} \right) = \frac{3N}{2\beta} \\ U_{Tran} &= \frac{3}{2} N k_B T \end{aligned}$$

**Equation B.9:** Energy from translational motion

Where  $\beta = 1/k_B T$ , where  $k_B$  is the Boltzmann constant and  $T$  is the temperature.

The contribution of molecular rotation to total energy is shown in Equation B.10.

$$U_{Rot} = \frac{(-N)}{q_{Rot}} \times \left( \frac{\partial(q_{Rot})}{\partial\beta} \right)_V = \frac{(-N)}{\left(\frac{1}{\beta}\right)^{2/3}} \left( \frac{\partial \left(\frac{1}{\beta}\right)^{3/2}}{\partial\beta} \right) = \frac{3N}{2\beta}$$

$$q_{Rot} = \frac{\sqrt{\pi}}{\sigma} \left( \frac{1}{\beta h c B_x} \right)^{1/2} \left( \frac{1}{\beta h c B_y} \right)^{1/2} \left( \frac{1}{\beta h c B_z} \right)^{1/2}$$

$$U_{Rot} = \frac{3}{2} N k_B T$$

**Equation B.10:** Energy from molecular rotation

Contribution of molecular vibration to total energy is shown in Equation B.11, where  $l = 1, 2, \dots$

From  $3N - 6$ .

$$\begin{aligned} U_{Vib} &= -N \left( -\frac{\partial \ln q_{Vib}}{\partial \beta} \right)_V \\ q_{Vib} &= \prod_l \frac{\text{Exp} \left[ -\frac{\beta h c v_l}{2} \right]}{1 - \text{Exp}[-\beta h c v_l]} \\ \ln q_{Vib} &= \sum_l \left( -\frac{\beta h c v_l}{2} \right) - \ln(1 - \text{Exp}[-\beta h c v_l]) \\ U_{Vib} &= N \sum_l \left[ \frac{h c v_l}{2} + \frac{h c v_l \text{Exp}[-\beta h c v_l]}{1 - \text{Exp}[-\beta h c v_l]} \right] \\ U_{Vib} &= N \sum_l \left[ \frac{1}{2} + \frac{1}{\text{Exp}[-\beta h c v_l] - 1} \right] h c v_l \end{aligned}$$

**Equation B.11:** Energy from molecular vibration

Contribution of molecular vibration to total energy is shown in Equation B.12 Electronic energy contributions to the total internal energy are negligible and so are not included.

$$U_{Elec} = -\frac{N}{q_{Elec}} \left( \frac{\partial q_{Elec}}{\partial \beta} \right)_V = 0$$

$$q_{Elec} = \sum_n g_n \text{Exp}[-\beta hcE_n] \approx g_0$$

**Equation B.12:** Electronic energies

The total internal energy calculation is shown in Equation B.13.

$$U = \frac{3}{2} N k_B T + \frac{3}{2} N k_B T + N \sum_l \left[ \frac{1}{2} + \frac{1}{\text{Exp}[-\beta hc\nu_l] - 1} \right] hc\nu_l + N E_{BO}$$

**Equation B.13:** Calculation of total internal energy

From the ideal gas law, Equation B.14, we can complete the enthalpy calculation.

$$pV = NRT = Nk_B T$$

**Equation B.14:** Ideal gas law

The total expression for enthalpy, is shown in Equation B.15.

$$H = U + pV = U + Nk_B T$$

$$H = \frac{3}{2} N k_B T + \frac{3}{2} N k_B T + N \sum_l \left[ \frac{1}{2} + \frac{1}{\text{Exp}[-\beta hc\nu_l] - 1} \right] hc\nu_l + N E_{BO} + N k_B T$$

**Equation B.15:** Calculation of Enthalpy

The total entropy ( $S_{Tot}$ ) can be calculated in the same way as the internal energy, from the entropy contributions from each of the molecular degrees of freedom, as shown in Equation B.16.

$$S_{Tot} = S_{Trans} + S_{Rot} + S_{Vib} + S_{Elec} = \left( \frac{\partial}{\partial T} (Nk_B T \ln q_{Tot}) \right)$$

$$S_{Trans} = k_B \left( \ln q_{Trans} + \frac{5}{2} \right)$$

$$S_{Rot} = k_B \left( \ln q_{Rot} + \frac{3}{2} \right)$$

$$S_{Vib} = k_B \sum_l \left( \frac{hc\nu_l}{Exp[-\beta hc\nu_l] - 1} \right) - \ln(1 - Exp[-\beta hc\nu_l])$$

$$S_{Elec} = k_B (\ln q_{Elec} + 0)$$

$$\begin{aligned} S_{Tot} &= k_B \left( \ln q_{Trans} + \frac{5}{2} \right) + k_B \left( \ln q_{Rot} + \frac{3}{2} \right) \\ &\quad + k_B \sum_l \left( \frac{hc\nu_l}{Exp[-\beta hc\nu_l] - 1} \right) - \ln(1 - Exp[-\beta hc\nu_l]) + k_B (\ln q_{Elec} + 0) \end{aligned}$$

**Equation B.16:** Calculation of total entropy

Combining all the expressions we now have a complete Gibbs energy calculation, as shown in Equation B.17.

$$\begin{aligned} G &= 4Nk_B T + N \sum_l \left[ \frac{1}{2} + \frac{1}{Exp[-\beta hc\nu_l] - 1} \right] hc\nu_l + NE_{BO} \\ &\quad - \left( k_B \left( \ln q_{Trans} + \frac{5}{2} \right) + k_B \left( \ln q_{Rot} + \frac{3}{2} \right) \right. \\ &\quad \left. + k_B \sum_l \left( \frac{hc\nu_l}{Exp[-\beta hc\nu_l] - 1} \right) - \ln(1 - Exp[-\beta hc\nu_l]) \right. \\ &\quad \left. + k_B (\ln q_{Elec} + 0) \right) T \end{aligned}$$

**Equation B.17:** Calculation of Gibbs energy

## **APPENDIX C**

### **Computational Data**

All calculated computational geometries are contained in the accompanying file:

*Appendix C – Computational Geometries.txt*

## References

1. Determan, J. J., Moncho, S., Brothers, E. N. & Janesko, B. G. *J. Phys. Chem. C* **118**, 15693–15704 (2014).
2. Determan, J. J., Moncho, S., Brothers, E. N. & Janesko, B. G. *Int. J. Quantum Chem.* **115**, 1718–1725 (2015).
3. Mahler, A. & Janesko, B. G. *Int. J. Quantum Chem.* **116**, 1451–1458 (2016).
4. Pudasaini, B. & Janesko, B. G. *Organometallics* **30**, 4564–4571 (2011).
5. Pudasaini, B. & Janesko, B. G. *Organometallics* **31**, 4610–4618 (2012).
6. Lennox, A. J. J. & Lloyd-Jones, G. C. *Chem Soc Rev* **43**, 412–443 (2014).
7. Sakolish, C. M., Esch, M. B., Hickman, J. J., Shuler, M. L. & Mahler, G. J. *EBioMedicine* **5**, 30–39 (2016).
8. Paul, S. M. *et al.* *Nat. Rev. Drug Discov.* (2010). doi:10.1038/nrd3078
9. Abrantes-Metz, R. M., Adams, C. P. & Metz, A. *Pharmaceutical development phases: a duration analysis*. (Bureau of Economics, Federal Trade Commission, 2004).
10. Adams, C. P. & Brantner, V. V. *Health Aff. (Millwood)* **25**, 420–428 (2006).
11. DiMasi, J. A., Hansen, R. W. & Grabowski, H. G. *J. Health Econ.* **22**, 151–185 (2003).
12. Morgan, S., Grootendorst, P., Lexchin, J., Cunningham, C. & Greyson, D. *Health Policy* **100**, 4–17 (2011).
13. Owens, P. K. *et al.* *Nat. Rev. Drug Discov.* **14**, 17 (2015).
14. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. *Adv. Drug Deliv. Rev.* **23**, 3–25 (1997).
15. Leeson, P. *Nature* **481**, 455–456 (2012).
16. Ghose, A. K., Viswanadhan, V. N. & Wendoloski, J. J. *J. Phys. Chem. A* **102**, 3762–3772 (1998).
17. Clark, D. E. *J. Pharm. Sci.* **88**, 807–814 (1999).
18. Weininger, D. *J. Chem. Inf. Comput. Sci.* **28**, 31–36 (1988).

19. Weininger, D., Weininger, A. & Weininger, J. L. *J. Chem. Inf. Comput. Sci.* **29**, 97–101 (1989).
20. Weininger, D. *J. Chem. Inf. Comput. Sci.* **30**, 237–243 (1990).
21. Balducci, R., McGarity, C. M. & Rusinko, A. *CONCORD v4. 02.* (Tripos, Inc.).
22. Morris, G. *SYBYL 6.9.* (Tripos Associates).
23. Dodd, L. R. & Theodorou, D. N. *Mol. Phys.* **72**, 1313–1345 (1991).
24. Lewis, D. F. *Biochem. Pharmacol.* **60**, 293–306 (2000).
25. Smith, D. A., Ackland, M. J. & Jones, B. C. *Drug Discov. Today* **2**, 406–414 (1997).
26. Smith, D. A., Ackland, M. J. & Jones, B. C. *Drug Discov. Today* **2**, 479–486 (1997).
27. Grass, G. M. *Adv. Drug Deliv. Rev.* **23**, 199–219 (1997).
28. Lawrence, X. Y., Crison, J. R. & Amidon, G. L. *Int. J. Pharm.* **140**, 111–118 (1996).
29. Neil, P. & Thierry, L. *Eur. J. Pharm. Sci.* **17**, 51–61 (2002).
30. Naylor, T. A. *et al.* *J. Appl. Ther. Res.* **6**, 15 (2006).
31. Löbenberg, R., Krämer, J., Shah, V. P., Amidon, G. L. & Dressman, J. B. *Pharm. Res.* **17**, 439–444 (2000).
32. Orsini, M. J. *et al.* *J. Med. Chem.* **50**, 462–471 (2007).
33. Srikanth, K., Anil Kumar, C., Ghosh, B. & Jha, T. *Bioorg. Med. Chem.* **10**, 2119–2131 (2002).
34. Kukol, A. *Molecular modeling of proteins.* (Humana Press, 2008).
35. Trott, O. & Olson, A. J. *J. Comput. Chem.* **31**, 455–461 (2010).
36. Frisch, M. J. *et al.* *Gaussian 09, Revision D.01.* (Gaussian, Inc., 2009).
37. Festa, C. & Schmidt, K. J. *The chemistry of organophosphorus pesticides.* (Springer Science & Business Media, 2012).
38. Dhansay, M. A., Linder, P. W., Torington, R. G. & Modro, T. A. *J. Phys. Org. Chem.* **3**, 248–254 (1990).
39. Wong, P. K. & Chang, L. *Environ. Pollut.* **55**, 179–189 (1988).
40. Nemeth, G. *et al.* *J. Med. Chem.* **57**, 3939–3965 (2014).

41. Wiemer, A. J., Hohl, R. J. & Wiemer, D. F. *Anti-Cancer Agents Med. Chem. Former. Curr. Med. Chem.-Anti-Cancer Agents* **9**, 526–542 (2009).
42. Eriksen, E. F., Diez-Perez, A. & Boonen, S. *Bone* **58**, 126–135 (2014).
43. Herranz, C. *J. Am. Oil Chem. Soc.* **64**, 1038–1039 (1987).
44. Rakotomalala, M., Wagner, S. & Doring, M. *Materials* **3**, 4300–4327 (2010).
45. Brothers, L. E. High temperature set retarded cement compositions and methods. U.S. Patent 5,484,478 (1996).
46. Coudray, L. & Montchamp, J.-L. *Eur. J. Org. Chem.* **2008**, 3601–3613 (2008).
47. Anderson, N. G. *et al. Org. Process Res. Dev.* **1**, 315–319 (1997).
48. Bravo-Altamirano, K. & Montchamp, J.-L. *Tetrahedron Lett.* **48**, 5755–5759 (2007).
49. Montchamp, J.-L. & Dumond, Y. R. *J. Am. Chem. Soc.* **123**, 510–511 (2001).
50. Berger, O., Petit, C., Deal, E. L. & Montchamp, J.-L. *Adv. Synth. Catal.* **355**, 1361–1373 (2013).
51. Montchamp, J.-L. *Acc. Chem. Res.* **47**, 77–87 (2013).
52. Borger, J., Ehlers, A. W., Slootweg, J. C. & Lammertsma, K. *Chem. - Eur. J.* (2017).
53. Janesko, B. G., Fisher, H. C., Bridle, M. J. & Montchamp, J.-L. *J. Org. Chem.* **80**, 10025–10032 (2015).
54. Wesolowski, S. S. *et al. J. Chem. Phys.* **116**, 112–122 (2002).
55. Kraihanzel, C. S. & Bartish, C. M. *J. Am. Chem. Soc.* **94**, 3572–3575 (1972).
56. Yakhvarov, D. *et al. Angew. Chem. Int. Ed.* **50**, 5370–5373 (2011).
57. Manca, G., Caporali, M., Ienco, A., Peruzzini, M. & Mealli, C. *J. Organomet. Chem.* **760**, 177–185 (2014).
58. Belabassi, Y., Antczak, M. I., Tellez, J. & Montchamp, J.-L. *Tetrahedron* **64**, 9181–9190 (2008).
59. Montchamp, J.-L. *Phosphorus Sulfur Silicon Relat. Elem.* **188**, 66–75 (2013).
60. Hohenberg, P. & Kohn, W. *Phys. Rev.* **136**, 864–871 (1964).
61. Kohn, W. & Sham, L. J. *Phys Rev A* **140**, 1133–1138 (1965).
62. Becke, A. D. *J. Chem. Phys.* **98**, 5648 (1993).

63. Stephens, P. J., Devlin, F. J., Chabalowski, Cf. & Frisch, M. J. *J. Phys. Chem.* **98**, 11623–11627 (1994).
64. Hehre, W. J., Ditchfield, R. & Pople, J. A. *J. Chem. Phys.* **56**, 2257 (1972).
65. Dunning, T. H. *J. Chem. Phys.* **90**, 1007 (1989).
66. Dunning, T. H. *J. Phys. Chem. A* **104**, 9062–9080 (2000).
67. Dunning, T. H., Peterson, K. A. & Wilson, A. K. *J. Chem. Phys.* **114**, 9244 (2001).
68. Marenich, A. V., Cramer, C. J. & Truhlar, D. G. *J. Phys. Chem. B* **113**, 6378–6396 (2009).
69. Chesnut, D. B. *Heteroat. Chem.* **11**, 73–80 (2000).
70. Babin, Y. V., Prisyazhnyuk, A. V. & Ustyynyuk, Y. A. *Russ. J. Phys. Chem. A* **82**, 94–100 (2008).
71. Prisyazhnyuk, A. V. & Babin, Y. V. *J. Struct. Chem.* **46**, 164–167 (2005).
72. Babin, Y. V. & Ustyynyuk, Y. A. *Russ. J. Phys. Chem. A* **81**, 1810–1819 (2007).
73. Cho, B., Tan, C.-H. & Wong, M. W. *Org. Biomol. Chem.* **9**, 4550 (2011).
74. Akiyama, T., Morita, H., Itoh, J. & Fuchibe, K. *Org. Lett.* **7**, 2583–2585 (2005).
75. Han, L.-B., Hua, R. & Tanaka, M. *Angew Chem Int Ed* **37**, 94–96 (1998).
76. Deal, E. L., Petit, C. & Montchamp, J.-L. *Org. Lett.* **13**, 3270–3273 (2011).
77. Petit, C., Fécourt, F. & Montchamp, J.-L. *Adv. Synth. Catal.* **353**, 1883–1888 (2011).
78. Cremlyn, R. J. *An introduction to organosulfur chemistry*. (John Wiley & Sons LTD, 1996).
79. Delaere, D., Raspoet, G. & Nguyen, M. T. *J. Phys. Chem. A* **103**, 171–177 (1999).
80. Sarro, A. D. & Sarro, G. D. *Curr. Med. Chem.* **8**, 371–384 (2001).
81. Van Bambeke, F., Michot, J.-M., Van Eldere, J. & Tulkens, P. M. *Clin. Microbiol. Infect.* **11**, 256–280 (2005).
82. Zhanel, G. G. *et al. Drugs* **62**, 13–59 (2002).
83. Oliphant, C. M. & Green, G. M. *Am. Fam. Physician* **65**, 455–464 (2002).
84. Cheng, G., Hao, H., Dai, M., Liu, Z. & Yuan, Z. *Eur. J. Med. Chem.* **66**, 555–562 (2013).
85. Pommier, Y., Leo, E., Zhang, H. & Marchand, C. *Chem. Biol.* **17**, 421–433 (2010).

86. Palù, G., Valisena, S., Ciarrocchi, G., Gatto, B. & Palumbo, M. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 9671–9675 (1992).
87. Aldred, K. J., McPherson, S. A., Turnbough, C. L., Kerns, R. J. & Osheroff, N. *Nucleic Acids Res.* **41**, 4628–4639 (2013).
88. Turel, I., Živec, P., Pevec, A., Tempelaar, S. & Psomas, G. *Eur. J. Inorg. Chem.* **23**, 3718–3727 (2008).
89. Lecomte, S., Baron, M. H., Chenon, M. T., Coupry, C. & Moreau, N. J. *Antimicrob. Agents Chemother.* **38**, 2810–2816 (1994).
90. Campbell, N. R., Kara, M., Hasinoff, B. B., Haddara, W. M. & McKay, D. W. *Br. J. Clin. Pharmacol.* **33**, 115–116 (1992).
91. Gao, L.-Z. *et al.* *Chin. Chem. Lett.* **26**, 149–151 (2015).
92. Li, T. *et al.* *Acta Pharm. Sin.* **50**, 569–573 (2015).
93. Ravi Kumar, A. *et al.* *J. Heterocycl. Chem.* **52**, 235–242 (2015).
94. Chen, P.-T., Lin, W.-P., Lee, A.-R. & Hu, M.-K. *Molecules* **18**, 7557–7569 (2013).
95. Rajulu, G. G. *et al.* *Med. Chem. Res.* **22**, 3843–3856 (2012).
96. McPherson III, J. C. *et al.* *Eur. J. Med. Chem.* **47**, 615–618 (2012).
97. Sissi, C. *et al.* *Bioorg. Med. Chem.* **6**, 1555–1561 (1998).
98. Freccero, M. *et al.* *Chem. – Eur. J.* **14**, 653–663 (2008).
99. Vitorino, G. P., Barrera, G. D., Mazzieri, M. R., Binning, R. C. & Bacelo, D. E. *Chem. Phys. Lett.* **432**, 538–544 (2006).
100. Pavez, P., Toro-Labbé, A. & Rojas, M. V. E. *Photochem. Photobiol.* **82**, 254–261 (2006).
101. Musa, K. A. K. & Eriksson, L. A. *J. Phys. Chem. A* **113**, 10803–10810 (2009).
102. Li, X., Zhang, X., Cheng, X., Yang, X. & Zhu, Z. *Chin. J. Chem. Phys.* **19**, 143–148 (2006).
103. Li, X., Zhang, R., Cheng, X. & Yang, X. *Chin. J. Chem. Phys.* **20**, 167–172 (2007).
104. Rosales, M. J., Toscano, R. A., Barba-Behrens, N. & García, J. *Acta Crystallogr. C* **41**, 1825–1826 (1985).

105. Turel, I., Leban, I., Zupancic, M., Bukovec, P. & Gruber, K. *Acta Crystallogr. Sect. C* **52**, 2443–2445 (1996).
106. Parvez, M., Arayne, M. S., Sultana, N. & Siddiqi, A. Z. *Acta Crystallogr. C* **56**, 910–912 (2000).
107. Zheng, X.-J., Zhao, H.-C. & Jin, L.-P. *Acta Crystallogr. C* **58**, o596–o597 (2002).
108. Hashimoto, K., Fujita, N., Tanaka, T. & Kido, M. *Acta Crystallogr. C* **51**, 519–521 (1995).
109. Jasinski, J. P., Butcher, R. J., Siddegowda, M. S., Yathirajan, H. S. & Hakim Al-arique, Q. N. M. *Acta Crystallogr. Sect. E Struct. Rep. Online* **67**, o483–o484 (2011).
110. Qian, J.-J., Gu, J.-M., Shen, J., Hu, X.-R. & Wu, S.-X. *Acta Crystallogr. Sect. E Struct. Rep. Online* **67**, o2773–o2774 (2011).
111. Kido, M. & Hashimoto, K. *Chem. Pharm. Bull. (Tokyo)* **42**, 872–876 (1994).
112. Huber, C. P., Sake Gowda, D. S. & Ravindra Acharya, K. *Acta Crystallogr. B* **36**, 497–499 (1980).
113. Florence, A. J., Kennedy, A. R., Shankland, N., Wright, E. & Al-Rubayi, A. *Acta Crystallogr. C* **56**, 1372–1373 (2000).
114. Drevenšek, P., Ulrih, N. P., Majerle, A. & Turel, I. *J. Inorg. Biochem.* **100**, 1705–1713 (2006).
115. Cygler, M. & Huber, C. P. *Acta Crystallogr. C* **41**, 1052–1055 (1985).
116. Parvez, M., Arayne, S., Sultana, N. & Siddiqi, A. Z. *Acta Crystallogr. C* **60**, o281–o283 (2004).
117. Fonseca, I., Martínez-Carrera, S. & García-Blanco, S. *Acta Crystallogr. C* **42**, 1618–1621 (1986).
118. Song, H., Shin, H.-S., Park, K.-I. & Cho, S.-I. *Acta Crystallogr. C* **54**, 1915–1917 (1998).
119. Sivalakshmidevi, A., Vyas, K. & Om Reddy, G. *Acta Crystallogr. C* **56**, e115–e116 (2000).
120. Sweidan, K., Haddad, S. F., AlDamen, M. A. & Al-Sheikhb, A. *Acta Crystallogr. Sect. E Struct. Rep. Online* **69**, o1191–o1191 (2013).
121. Marsh, R. E. Space group P1: an update. *Acta Crystallogr. B* **61**, 359 (2005).
122. Riley, K. E., Op't Holt, B. T. & Merz, K. M. *J. Chem. Theory Comput.* **3**, 407–433 (2007).
123. Bauschlicher Jr., C. W. *Chem. Phys. Lett.* 40–44 (1995).
124. Maroulis, G. & Simos, T. *Trends and Perspectives in Modern Computational Science*. (2006).
125. Patton, D. C. & Pederson, M. R. *Phys. Rev. A* **56**, R2495 (1997).

126. Janczak, J. & Idemori, Y. M. *Acta Crystallogr. C* **58**, m549–m550 (2002).
127. Ballem, K. H. D., Smith, K. M. & Patrick, B. O. *Acta Crystallogr. Sect. E Struct. Rep. Online* **60**, m408–m409 (2004).
128. Guzei, I. A., McGaff, R. W. & Kieler, H. M. *Acta Crystallogr. C* **61**, m472–m475 (2005).
129. Das, A. *et al.* *J. Phys. Chem. B* **114**, 4998–5009 (2010).
130. Cundari, T. R., Kurtz, H. A. & Zhou, T. *J. Phys. Chem. A* **102**, 2962–2966 (1998).
131. Akhtar, R. *et al.* *Synth. Commun.* **46**, 1849–1879 (2016).
132. Hu, G. *et al.* *Acta Pharm. Sin. B* **2**, 312–317 (2012).
133. Shi, Z. *et al.* *Acta Pharmacol. Sin.* **33**, 271–278 (2012).
134. Fu, Y. *et al.* *Int. J. Pharmacol.* **9**, 416–429 (2013).
135. Malik, M. *et al.* *Antimicrob. Agents Chemother.* **55**, 2335–2343 (2011).
136. Lemaire, S., Van Bambeke, F. & Tulkens, P. M. *Int. J. Antimicrob. Agents* **38**, 52–59 (2011).
137. Stubbings, W. *et al.* *Antimicrob. Agents Chemother.* **55**, 4394–4397 (2011).
138. Azoulay-Dupuis, E. *et al.* *Antimicrob. Agents Chemother.* **36**, 2698–2703 (1992).
139. Lemaire, S., Tulkens, P. M. & Van Bambeke, F. *Antimicrob. Agents Chemother.* **55**, 649–658 (2011).
140. Hernandez, V. S. *et al.* Boron-Containing Small Molecules. U.S. Patent 8,703,742 (2014).
141. Li, X. *et al.* *Bioorg. Med. Chem. Lett.* **23**, 963–966 (2013).
142. Lipinski, C. A. *J. Pharmacol. Toxicol. Methods* **44**, 235–249 (2000).
143. Tomisic, Z. B., Kujundzic, N., Krajacic, M. B., Visnjevac, A. & Kojic-Prodic, B. *J. Mol. Struct.* **611**, 73–81 (2002).
144. Frisch, M. J. *et al.* *Gaussian Development Version, Revision I.02.* (Gaussian, Inc., 2014).
145. Parr, R. G. & Yang, W. *Annu. Rev. Phys. Chem.* **46**, 701–728 (1995).
146. Salahub, D. R. & Zerner, M. C. Quantum Chemistry Throughout the Periodic Table. in *The Challenge of d and f Electrons* (eds. Salahub, D. R. & Zerner, M. C.) **394**, 1–16 (American Chemical Society, 1989).

147. Roothaan, C. C. *J. Rev. Mod. Phys.* **23**, 69 (1951).
148. Pople, J. A. & Nesbet, R. K. *J. Chem. Phys.* **22**, 571 (1954).
149. McWeeny, R. *J. Chem. Phys.* **49**, 4852 (1968).
150. Stewart, J. J. *J. Comput. Chem.* **10**, 209–220 (1989).
151. Stewart, J. J. *J. Comput. Chem.* **10**, 221–264 (1989).
152. Anders, E., Koch, R. & Freunsch, P. *J. Comput. Chem.* **14**, 1301–1312 (1993).
153. Stewart, J. J. *P. J. Mol. Model.* **13**, 1173–1213 (2007).
154. Slater, J. C. *The self-consistent field for molecules and solids.* **4**, (McGraw-Hill New York, 1974).
155. Vosko, S. H., Wilk, L. & Nusair, M. *Can. J. Phys.* **58**, 1200–1211 (1980).
156. Perdew, J. P., Burke, K. & Ernzerhof, M. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
157. Perdew, J. P., Burke, K. & Ernzerhof, M. *Phys Rev Lett* **78**, 1396 (1997).
158. Becke, A. D. *Phys Rev A* **38**, 3098–3100 (1988).
159. Perdew, J. P. *Electronic structure of solids.* (Akademie Verlag, 1991).
160. Perdew, J. P. *et al.* *Phys. Rev. B* **46**, 6671–6687 (1992).
161. Perdew, J. P. *et al.* *Phys. Rev. B* **48**, 4978 (1993).
162. Perdew, J. P., Burke, K. & Wang, Y. *Phys. Rev. B* **54**, 16533–16539 (1996).
163. Burke, K., Perdew, J. P. & Wang, Y. *Electronic Density Functional Theory: Recent Progress and New Directions.* (Plenum, 1998).
164. Adamo, C. & Barone, V. *J. Chem. Phys.* **110**, 6158–6169 (1999).
165. Lee, C., Yang, W. & Parr, R. G. *Phys. Rev. B* **37**, 785–789 (1988).
166. Miehlich, B., Savin, A., Stoll, H. & Preuss, H. *Chem Phys Lett* **157**, 200–206 (1989).
167. Adamo, C. & Barone, V. *J. Chem. Phys.* **108**, 664–675 (1998).
168. Zhao, Y. & Truhlar, D. G. *Theor. Chem. Acc.* **120**, 215–241 (2008).
169. Tao, J., Perdew, J. P., Staroverov, V. N. & Scuseria, G. E. *Phys. Rev. Lett.* **91**, (2003).
170. Zhao, Y. & Truhlar, D. G. *J. Chem. Phys.* **125**, 194101 (2006).
171. Grimme, S. *J. Comput. Chem.* **27**, 1787–1799 (2006).

172. Chai, J.-D. & Head-Gordon, M. *Phys. Chem. Chem. Phys.* **10**, 6615 (2008).
173. Hehre, W. J. *J. Chem. Phys.* **51**, 2657 (1969).
174. Collins, J. B. *J. Chem. Phys.* **64**, 5142 (1976).
175. Binkley, J. S., Pople, J. A. & Hehre, W. J. *J. Am. Chem. Soc.* **102**, 939–947 (1980).
176. Gordon, M. S., Binkley, J. S., Pople, J. A., Pietro, W. J. & Hehre, W. J. *J. Am. Chem. Soc.* **104**, 2797–2803 (1982).
177. Pietro, W. J. *et al. J. Am. Chem. Soc.* **104**, 5039–5048 (1982).
178. Dobbs, K. D. & Hehre, W. J. *J. Comput. Chem.* **7**, 359–378 (1986).
179. Dobbs, K. D. & Hehre, W. J. *J. Comput. Chem.* **8**, 861–879 (1987).
180. Dobbs, K. D. & Hehre, W. J. *J. Comput. Chem.* **8**, 880–893 (1987).
181. Ditchfield, R. *J. Chem. Phys.* **54**, 724 (1971).
182. Hehre, W. J. *J. Chem. Phys.* **56**, 2257 (1972).
183. Hariharan, P. C. & Pople, J. A. *Theor. Chim Acta* **28**, 213–222 (1973).
184. Franc, M. M. *J. Chem. Phys.* **77**, 3654 (1982).
185. Binning, R. C. & Curtiss, L. A. *J. Comput. Chem.* **11**, 1206–1216 (1990).
186. Blaudeau, J.-P., McGrath, M. P., Curtiss, L. A. & Radom, L. *J. Chem. Phys.* **107**, 5016 (1997).
187. Rassolov, V. A., Pople, J. A., Ratner, M. A. & Windus, T. L. *J. Chem. Phys.* **109**, 1223 (1998).
188. Rassolov, V. A., Ratner, M. A., Pople, J. A., Redfern, P. C. & Curtiss, L. A. *J. Comput. Chem.* **22**, 976–984 (2001).
189. McLean, A. D. & Chandler, G. S. *J. Chem. Phys.* **72**, 5639 (1980).
190. Krishnan, R., Binkley, J. S., Seeger, R. & Pople, J. A. *J. Chem. Phys.* **72**, 650 (1980).
191. Frisch, M. J., Pople, J. A. & Binkley, J. S. *SJ. Chem. Phys.* **80**, 3265 (1984).
192. Clark, T., Chandrasekhar, J., Spitznagel, G. W. & Schleyer, P. V. R. *J. Comput. Chem.* **4**, 294–301 (1983).
193. Kendall, R. A., Dunning, T. H. & Harrison, R. J. *J. Chem. Phys.* **96**, 6796 (1992).
194. Hay, P. J. & Wadt, W. R. *J. Chem. Phys.* **82**, 270 (1985).

195. Wadt, W. R. & Hay, P. J. *J. Chem. Phys.* **82**, 284 (1985).
196. Hay, P. J. & Wadt, W. R. *J. Chem. Phys.* **82**, 299 (1985).
197. Miertuš, S., Scrocco, E. & Tomasi, J. *Chem. Phys.* **55**, 117 (1981).
198. Pascual-ahuir, J.-L., Silla, E. & Tunon, I. *J. Comput. Chem.* **15**, 1127–1138 (1994).
199. Cossi, M., Barone, V., Cammi, R. & Tomasi, J. *Chem. Phys. Lett.* 327–335 (1996).
200. Barone, V., Cossi, M. & Tomasi, J. *J. Chem. Phys.* **107**, 3210 (1997).
201. Mennucci, B., Cances, E. & Tomasi, J. *J. Phys. Chem. B* **101**, 10506–10517 (1997).
202. Barone, V. & Cossi, M. *J. Phys. Chem. A* **102**, 1995–2001 (1998).
203. Cossi, M., Barone, V., Mennucci, B. & Tomasi, J. *Chem. Phys. Lett.* 253–260 (1998).
204. Barone, V., Cossi, M. & Tomasi, J. *J. Comput. Chem.* **19**, 404–417 (1998).
205. Cammi, R., Mennucci, B. & Tomasi, J. *J. Phys. Chem. A* **103**, 9100–9108 (1999).
206. Cossi, M., Barone, V. & Robb, M. A. *J. Chem. Phys.* **111**, 5295 (1999).
207. Seeger, R. & Pople, J. A. *J. Chem. Phys.* **66**, 3045 (1977).
208. Schirrmann, J.-P. & Bourdauducq, P. *Ullmann's Encyclopedia of Industrial Chemistry*. (Wiley-VCH, 2002).
209. Wolff, L. *Eur. J. Org. Chem.* **394**, 86–108 (1912).
210. Kishner, N. *J. Russ. Phys. Chem. Soc.* **43**, 582–595 (1911).
211. Corbett, P. T. *et al. Chem. Rev.* **106**, 3652–3711 (2006).
212. Li, J., Nowak, P. & Otto, S. *J. Am. Chem. Soc.* **135**, 9222–9239 (2013).
213. Ramström, O. & Lehn, J.-M. *Nat. Rev. Drug Discov.* **1**, 26–36 (2002).
214. Dirksen, A., Dirksen, S., Hackeng, T. M. & Dawson, P. E. *J. Am. Chem. Soc.* **128**, 15602–15603 (2006).
215. Lazny, R. & Nodzewska, A. *Chem. Rev.* **110**, 1386–1434 (2010).
216. Levrand, B., Fieber, W., Lehn, J.-M. & Herrmann, A. *Helv. Chim. Acta* **90**, 2281–2314 (2007).
217. Su, X. & Aprahamian, I. *Chem. Soc. Rev.* **43**, 1963 (2014).
218. Rollas, S. & Küçükgüzel, S. G. *Molecules* **12**, 1910–1939 (2007).

219. Hamann, P. R. *et al.* *Bioconjug. Chem.* **13**, 47–58 (2002).
220. Hamann, P. R. *et al.* *Bioconjug. Chem.* **13**, 40–46 (2002).
221. Zhang, X. *et al.* *J. Controlled Release* **174**, 209–216 (2014).
222. Krüger, H. R. *et al.* *J. Controlled Release* **194**, 189–196 (2014).
223. Leyva, S. & Leyva, E. *Tetrahedron* **63**, 2093–2097 (2007).
224. Liu, Y. & Yang, Z. *Eur. J. Med. Chem.* **44**, 5080–5089 (2009).
225. Shaw, A. Y. *et al.* *Eur. J. Med. Chem.* **45**, 2860–2867 (2010).
226. Ji, K., Lee, C., Janesko, B. G. & Simanek, E. E. *Mol. Pharm.* **12**, 2924–2927 (2015).
227. Wiberg, K. B. & Glaser, R. *J. Am. Chem. Soc.* **114**, 841–850 (1992).
228. Kalia, J. & Raines, R. T. *Angew. Chem.* **120**, 7633–7636 (2008).
229. Murthy, A. S. N., Balasubramanian, A., Rao, C. N. R. & Kasturi, T. R. *Can. J. Chem.* **40**, 2267–2271 (1962).
230. Mills, S. G. & Beak, P. *J. Org. Chem.* **50**, 1216–1224 (1985).
231. Cook, A. G. & Feltman, P. M. *J Chem Educ* **84**, 1827 (2007).
232. Emsley, J. & Freeman, N. J. *J. Mol. Struct.* **161**, 193–204 (1987).
233. Boys, S. F. & Bernardi, F. *Mol. Phys.* **19**, 553–566 (1970).
234. Sherrill, C. D. *Sch. Chem. Biochem. Ga. Inst. Technol.* (2010).

# VITA

MARK JOSEPH BRIDLE

## PERSONAL

Born 13<sup>th</sup> January, 1983  
Gloucester, UK

## EDUCATION

**B.Sc.** Chemistry with Forensic Analysis (2004)  
Loughborough University, Loughborough, UK  
Advisor: Dr. Paul Kelly  
Research advisor: Prof. J.N. Miller

**Ph.D.** Chemistry, Texas Christian University (2017)  
Fort Worth, TX  
Research advisor: Dr. Benjamin G. Janesko

## PROFESSIONAL EXPERIENCE

**Quality Assurance Chemist**, Alcon Laboratories, Inc., Fort Worth, TX  
August 2007 – July 2008

**Associate Scientist**, Alcon Research, Ltd., Fort Worth, TX  
July 2008 – February 2011

**Scientist I**, Alcon Research, Ltd., Fort Worth, TX  
March 2011 – February 2013

**CMC Team Lead – Otic**, Alcon Research, Ltd., Fort Worth, TX

**Scientist II**, Alcon Research, Ltd., Fort Worth, TX  
March 2013 – February 2016

**CMC Team Lead – Glaucoma**, Alcon Research, Ltd., Fort Worth, TX

**Senior Scientist I**, Alcon Research, Ltd. Fort Worth, TX  
March 2016 – December 2016

**Senior Scientist**, Novartis Pharmaceutical Corp. Fort Worth, TX  
January 2017 - Present

## PUBLICATIONS

1. Janesko, B.G., Fisher, H.C., Bridle, M.J., Montchamp, J-L. *J. Org. Chem.* **2015**. 80(20), 10025-10032. DOI: 10.1021/acs.molpharmaceut.5b00205
2. Bridle, M.J, Janesko, B.G. *Int. J. Quantum Chem.* Accepted June 2017

## PATENTS

1. Boudreaux, B.G., Bridle, M.J., Huynh, B.H., Ghosh, M., Chowhan, M.A., Alani, L., Kabra, B.K. *Finafloxacin Suspension Compositions*. U.S. Patent No. 9,504,691. **2016**.

## **ABSTRACT**

# **COMPUTATIONAL STUDY OF THE REACTIONS OF PHOSPHINYLIDENES, FLUOROQUINOLONES AND HYDRAZONES - MOLECULES OF COMMERCIAL AND PHARMACEUTICAL INTEREST**

by

Mark Joseph Bridle, Ph.D., 2017  
Department of Chemistry and Biochemistry  
Texas Christian University

Dissertation Advisor: Benjamin G. Janesko, Associate Professor

The formation of P-C bonds is of interest for many commercial, as well as pharmaceutical applications. Current production relies on the use of  $\text{PCl}_3$  in an energy intensive process that produces large amounts of environmentally toxic HCl as a by-product. The Montchamp group has long had an interest in an alternative process for the green synthesis of P-C bond-containing compounds. In a collaborative experimental and computational study, it was demonstrated that using phosphinylidenes are viable starting feedstock for P-C bond formation, taking advantage of their P(V) to reactive P(III) tautomerization. Computational and experimental results demonstrate the stabilization of the P(III) tautomer by electron withdrawing groups. Furthermore, with a high tautomerization energy, catalysis is necessary for viable commercial application. Computational screening of a variety of simple organic molecules showed good agreement with experimental data and highlighted their potential for future development.

Fluoroquinolones require the formation of a water-metal ion bridge between the molecule and topoisomerase IV for antibiotic efficacy. A computational study of the pH dependent binding to  $Mg(H_2O)_N^{2+}$  showed good agreement with experimental *in vitro* binding affinities and efficacy. The approach produced an average magnesium binding affinity for efficacious fluoroquinolones at neutral and basic pH, demonstrating viability as a screening method for future drugs within the class. The study also identified systematic modulation of  $pK_a$ 's as a possible future direction for drug design.

Hydrazones are of commercial and pharmaceutical interest due to their pH lability. Previously, it was demonstrated that triazinylhydrazones exhibited greater stability to hydrolysis at pH <5 compared to acetylhydrazones. In a collaborative follow-up, the effect of substituents at N2 were investigated. It was shown that electron donating groups increase the proton affinity of N1 and hence the rate of hydrolysis, while electron withdrawing groups decreased the proton affinity and hydrolysis rate. The same stabilization at pH <5 was observed, with computational and experimental results in good agreement. Outside of this trend lay N2 phenyl substituents. While electron withdrawing by induction they appeared to exhibit hyperconjugation stabilization of the protonated N1. This observation may become the focus of future work.