Fall 1-1-2015

Computational Characterization of Carboxyphosphate

Traci M. Clymer

COMPUTATIONAL CHARACTERIZATION OF CARBOXYPHOSPHATE

A Dissertation

Submitted to the Bayer School of Natural and Environmental Sciences

Duquesne University

In partial fulfillment of the requirements for

the degree of Doctor of Philosophy

By

Traci M. Clymer

December 2015

COMPUTATIONAL CHARACTERIZATION OF CARBOXYPHOSPHATE

By

Traci M. Clymer

Approved November 16, 2015

Jeffrey D. Evanseck Professor of Chemistry and Biochemistry (Committee Chair)	Jeffry D. Madura Professor of Chemistry and Biochemistry (Committee Member)
Michael Cascio Associate Professor of Chemistry and Biochemistry (Committee Member)	Steven M. Firestine Associate Professor of Pharmaceutical Sciences Wayne State University (External Reader)
Phillip P. Reeder Dean and Professor, Bayer School of Natural and Environmental Sciences	Ralph A. Wheeler Chair, Department of Chemistry and Biochemistry Professor of Chemistry and Biochemistry

ABSTRACT

COMPUTATIONAL CHARACTERIZATION OF CARBOXYPHOSPHATE

By

Traci M. Clymer

December 2015

Dissertation supervised by Jeffrey D. Evanseck

Carboxyphosphate (CP) is an important intermediate involved in reactions catalyzed by acetyl-CoA carboxylase, pyruvate carboxylase, N5-Carboxyaminoimidazole ribonucleotide synthetase, propionyl-CoA carboxylase, urea amidolyase, and carbamoyl phosphate synthetase. Despite its important role, properties for CP have never been reported due to its short estimated half-life ($t_{1/2}$ < 70 ms). Thus, the high level *ab initio* methods, MP2 and CCSD(T), along with the DFT functionals: B3LYP, BB1K, M05-2X, M06-2X, and MPW1K were used to investigate the structure and energetics of CP in both vacuum and the PCM continuum solvation model of water. It was found that CP adopts a novel pseudo-cyclic structure featuring an intramolecular charge-assisted hydrogen bond (CAHB) that is reminiscent of chair cyclohexane. This structure is found to be the most stable in both vacuum and implicit solvation for both mono and dianionic

charge states. Additionally, the M06-2X/aug-cc-pVTZ level of theory was shown to give consistent agreement with ab initio methods for both geometric and energetic properties. The strengths of the CAHBs observed in mono- and dianionic CP were estimated to be within the range of -17.8 to -25.4 and -15.7 to -20.9 kcal/mol, respectively. This classifies them as short-strong but not low-barrier and makes them the dominant stabilizing feature for these conformations. pK_a values were computed to distinguish between different possible protonation states of CP. The predicted pK_a values were found to be be -3.43±0.81, 4.04±0.35, and 8.14±1.92 for the first, second and third acid dissociations of CP, respectively, indicating it is most likely to be present as a dianion or trianion in aqueous solution but more work is required to predict its charge state in the enzymatic pocket.

ACKNOWLEDGMENT

I would like to thank my advisor Jeffrey D. Evanseck for helping me develop as an independent scientist. I would like to thank Jeffrey D. Madura for his valuable insights and feedback both as a committee member and as part of the theory group. I would also like to thank Michael Cascio for his contributions as a committee member. His perspective as a biochemist has been invaluable to my work. Similarly I would like to thank Steven Firestine for his input.

I would like to thank Douglas J. Fox for his help and expertise with Gaussian. I would like to thank Scott Boesch for helping me get and keep my calculations up and running. Without his help none of this work would be possible. I thank Ralph Wheeler for his advice during Theory meetings as wells as the other members of the group.

I thank all the past and present members of the ERG. I'm glad to be able to pass along my knowledge about computational chemistry, and famous internet cats to a new group of Ph.D. candidates.

I thank the Department of Chemistry and Biochemistry at Duquesne University, the NIH, and the NSF for their financial assistance during my Doctoral work.

Last but not least, I would like to thank my boyfriend, my friends and family, for their continued emotional support throughout this entire process.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

Chapter 1

1 Introduction

Non-communicable or non-transmittable diseases are by far the leading cause of death in the world.[1] They are responsible for 63% of deaths, killing more than 36 million people every year.[2]

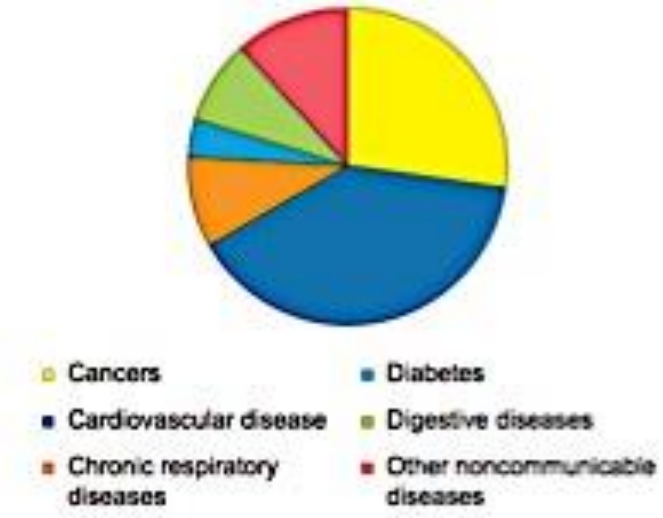


Figure 1.1. Proportion of global deaths by non-communicable diseases under age 70, by cause of death for 2008.[3]

There are four main categories of non-communicable diseases that contribute most substantially to this death toll: cardiovascular diseases, cancer, chronic respiratory diseases, and diabetes.[1] About 2.8 million people die every year from obesity with another 2.6 million and 7.5 million dying from related illnesses high cholesterol and high blood pressure, respectively.[4] In 2012, cancer was responsible for about 8.2 million deaths or 14.6% of all deaths globally and 14.1 million new cases were diagnosed.[5] As of 2014 an estimate 387 million people have been diagnosed with diabetes across the globe representing 8.3% of the total population.[6] Between 2012 and 2014 diabetes was estimated to be the cause of 1.5 to 4.9 million deaths per year.[7] The global cost of this disease was approximately $612 billion USD, and $245 billion was spent by the US in 2012.[8–10]

The ATP-grasp enzymes represent important targets for the development of new treatments for many of the diseases that are having the most substantial global impact on mortality and healthcare costs.[7,10–14] Acetyl-CoA carboxylase (biotin carboxylase), pyruvate carboxylase, N^5-Carboxyaminoimidazole ribonucleotide synthetase, propionyl-CoA carboxylase, urea amidolyase, and carbamoyl phosphate synthetase are believed to share the same general mechanism and catalyze reactions involved in important biological processes including: fatty acid biosynthesis, gluconeogenesis, purine biosynthesis, amino acid catabolism, urea hydrolysis, and arginine biosynthesis.[15] Unfortunately, the details of this mechanism remain poorly understood. Both acetyl-CoA carboxylase and pyruvate carboxylase have been suggested as targets for new diabetes treatments.[14] Additionally, there has been a great deal of interest in exploiting acetyl-CoA to develop treatments for obesity and cancer as well as microbial infections due to its critical role in fatty acid biosynthesis.[13,16] Thus, further elucidation of the mechanism shared by these enzymes has substantial implications in the treatment and understanding of many common and deadly diseases.

1.1 ATP-Grasp Enzymes

The ATP-grasp enzymes are a superfamily of 21 proteins characterized by a unique ATP-binding site called the ATP-grasp fold.[15,17] This fold consists of two α + β domains that "grasp" a molecule of ATP between them.[18,19] These proteins are present in several metabolic pathways including de novo purine biosynthesis, gluconeogenesis, and fatty acid biosynthesis.[20–23]

Table 1.1. ATP-grasp superfamily enzymes[15]

Enzyme	Biological process	Taxonomy	Nucleophilic substrate	Carboxylate substrate
Acetyl-CoA carboxylase (biotin carboxylase domain, BC) (ACC)	Fatty acid biosynthesis	Bacteria, archaea, eukaryota	Biotin-enzyme	HCO_3^-
D-Alanine-D-Alanine ligase (DDLigase)	Peptidoglycan biosynthesis	Bacteria	D-Alanine	D-Alanine
Glutathione synthetase (GSHase)	Glutathione biosynthesis	Bacteria, eukaryota	Glycine	c-Glutamyl-cysteine
Pyruvate Carboxylase (PC)	Gluconeogenesis	Bacteria, eukaryota	Biotin-enzyme	HCO_3^-
Glycinamide ribonucleotide synthetase (PurD)	Purine biosynthesis	Bacteria, eukaryota	Phosphoribosylamine	Glycine
Formylglycinamide ribonucleotide synthetase (PurT)	Purine biosynthesis	Bacteria	Glycinamide ribonucleotide	HCO_2^-
N^5-Carboxyaminoimidazole ribonucleotide synthetase (PurK)	Purine biosynthesis	Bacteria, yeast, fungi	Aminoimidazole ribonucleotide	HCO_3^-
Flavin 5-aminoimidazole-4-carboxamide (ribonucleotide synthetase (PurP)	Purine biosynthesis	Archaea	Aminoimidazole-4-carboxamide ribonucleotide	HCO_2^-
Ribosomal protein S6 modification protein (Rim K)	Ribosome biogenesis	Bacteria	Glutamate	Ribosomal protein S6
Propionyl-CoA carboxylase (PCC)	Amino acid catabolism	Bacteria, eukaryota	Biotin-enzyme	HCO_3^-
Urea amidolyase	Urea hydrolysis	Bacteria, eukaryota	Biotin-enzyme	HCO_3^-
Carbamoyl phosphate synthetase (CPS)	Arginine biosynthesis, pyrimidine biosynthesis	Bacteria, archaea, eukaryota	NH_3	HCO_3^- $NH_2CO_2^-$
Pyruvate phosphate dikinase (PPDK)	Gluconeogenesis, photosynthesis	Bacteria, eukaryota	Pyruvate	N/A*
Carnosine synthase	Dipeptide synthesis	Bacteria, archaea, eukaryota, viruses	L-Histidine	b-Alanine
Inositol 1,3,4-triphosphate 5/6-kinase (IP56 K)	Polyphosphate synthesis	Eukaryota	Inositol 1,3,4-triphosphate	N/A*
Synapsin 1a (SynC)	Neuronal function	Eukaryota	Unknown	Unknown
Tubulin-tyrosine ligase (TTL)	Microtubules assembly	Bacteria, eukaryota	Tyrosine	a-Tubulin
Succinyl-CoA synthetase b-chain (SCS)	Citric aced cycle	Bacteria, archaea, eukaryota	Coenzyme A	Succinate
Malate-CoA ligase b-chain	Growth on C-1 compounds	Bacteria, viruses	Coenzyme A	Malate succinate
ATP-citrate lyase	Lipid biosynthesis	Eukaryota	Coenzyme A	Citrate

Lysine biosynthesis enzyme (LysX)	Lysine biosynthesis pathway	Bacteria	a-Aminoadipate	Unknown

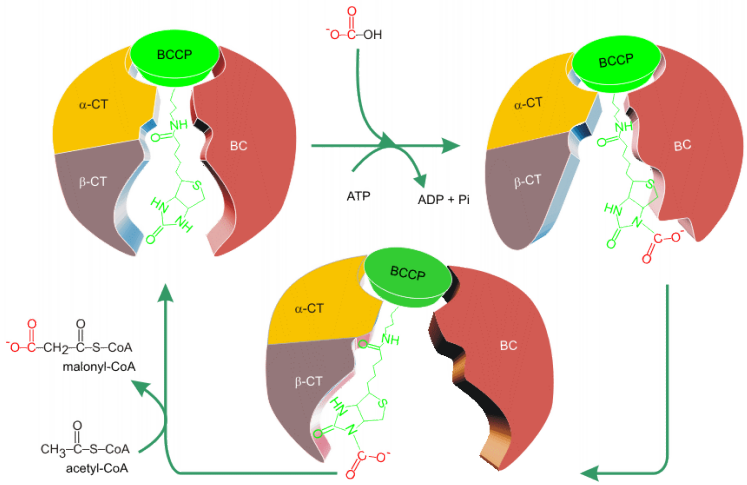
* Kinases are only involved in phosphoryl transfer and not carboxylation reactions

Members of this superfamily typically have three common conserved focal domains, commonly referred to as the A-C domains. The A domain is the N-terminal domain, the B domain is the central domain, and the C domain is the C-terminal domain.[18,19,24–26] While the ATP-grasp enzymes are highly conserved in the ATP-binding region (90% identical in the aligned regions of BC, GSHase, DDLigase, and SynC), they have an overall low sequence identity of only 10-20% in aligned regions between BC, DDLigase, GSHase, and SynC.[15,17]

With the exception of the kinases, the enzymes of the ATP-grasp superfamily all share the same general mechanism that consists of 2 basic steps.[27–30] In the first step, the carboxylic substrate reacts with ATP to generate a reactive acylphosphate intermediate. This step serves to generate an electrophile from the carboxylate substrate.[17,19,26,31–34] The second step is the transfer of the carboxylate group from the newly generated acylphosphate intermediate to the nucleophilic substrate.[35–40] Of the 19 enzymes that share this two-step, process, 6 of them all share the same carboxylate substrate: Acetyl-CoA carboxylase (biotin carboxylase), pyruvate carboxylase, N^5-Carboxyaminoimidazole ribonucleotide synthetase, propionyl-CoA carboxylase, urea amidolyase, and carbamoyl phosphate synthetase.[15,17,26,31]

1.1.1 Acetyl-CoA Carboxylase

Acetyl-CoA carboxylase (ACC) catalyzes the first committed step in fatty acid biosynthesis, the carboxylation of acetyl-CoA to form malonyl-CoA.[12] In most prokaryotes, ACC is a multi-subunit enzyme whereas in eukaryotes it is a multi-domain enzyme.[12,17,41]



Scheme 1.1. The first committed step of fatty acid biosynthesis. Carboxylation of acetyl-CoA to form malonyl-CoA[12]

There are three domains or subunits that make up ACC: biotin carboxylase (BC), biotin carboxyl carrier protein (BCCP), and carboxyl transferase (CT). BC catalyzes the first reaction of the overall reaction of ACC, which is the ATP-dependent carboxylation of biotin from bicarbonate. The biotin is bound to the BCCP via an amide bond to a lysine residue. Once carboxylated, the BCCP transfers the carboxybiotin to the carboxyl transferase functionality. Carboxyl transferase then catalyzes the transfer of the carboxyl group from biotin to acetyl-CoA producing malonyl-CoA.[17,30,41,42]

1.1.2 Pyruvate Carboxylase

Pyruvate carboxylase (PC) catalyzes the irreversible carboxylation of pyruvate to form oxaloacetate.[43] PC is utilized for the synthesis of phosphoenolpyruvate (PEP) in

gluconeogenesis.[27,31] The oxaloacetate produced by PC is decarboxylated and phosphorylated in a reaction catalyzed by phosphoenolpyruvate carboxykinase (PEPCK) to produce PEP.[29,44] In addition to its role in gluconeogenesis, PC also has an anaplerotic role in the citric acid cycle, where it provides the essential intermediate, oxaloacetate when concentrations are low.[23,45] PC is regulated by acetyl-CoA and aspartate.[23] Like ACC, it contains BC, BCCP, and CT domains. Additionally it also contains an allosteric effector domain.[23,29,46] The reaction process is essentially the same as ACC with the major difference being that in the second reaction the carboxyl is transferred from biotin to pyruvate rather than acetyl-CoA.

1.1.3 N^5-Carboxyaminoimidazole ribonucleotide synthetase

N^5-Carboxyaminoimidazole ribonucletide synthetase (N^5-CAIR synthetase) is an essential enzyme of the *de novo* purine biosynthesis in bacteria but is not found in humans.[11,47–52] In humans, *de novo* purine biosynthesis is a 10 step process beginning with phosphoribosyl pyrophosphate, while it is an 11-step process in bacteria, fungi, and yeast.[37,49,50,52] The key difference between these pathways is the conversion of 5-aminoimidazole ribonucleotide (AIR) to 4-carboxy-5-aminoimdazole ribonucleotide (CAIR) as shown in Figure 1.2.[37,49,52] For bacteria, fungi, and yeast, AIR is carboxylated to produce N^5-carboxy-5-aminoimidazole ribonucleotide (N^5-CAIR) by the enzyme N^5-CAIR synthetase. The carboxyl group is then shuffled from the amine to the imidazole ring by N^5-CAIR mutase. In humans, the imidazole is the initially site of carboxylation, generating CAIR directly from AIR and eliminating the need for multiple steps. This reaction is catalyzed by AIR carboxylase.

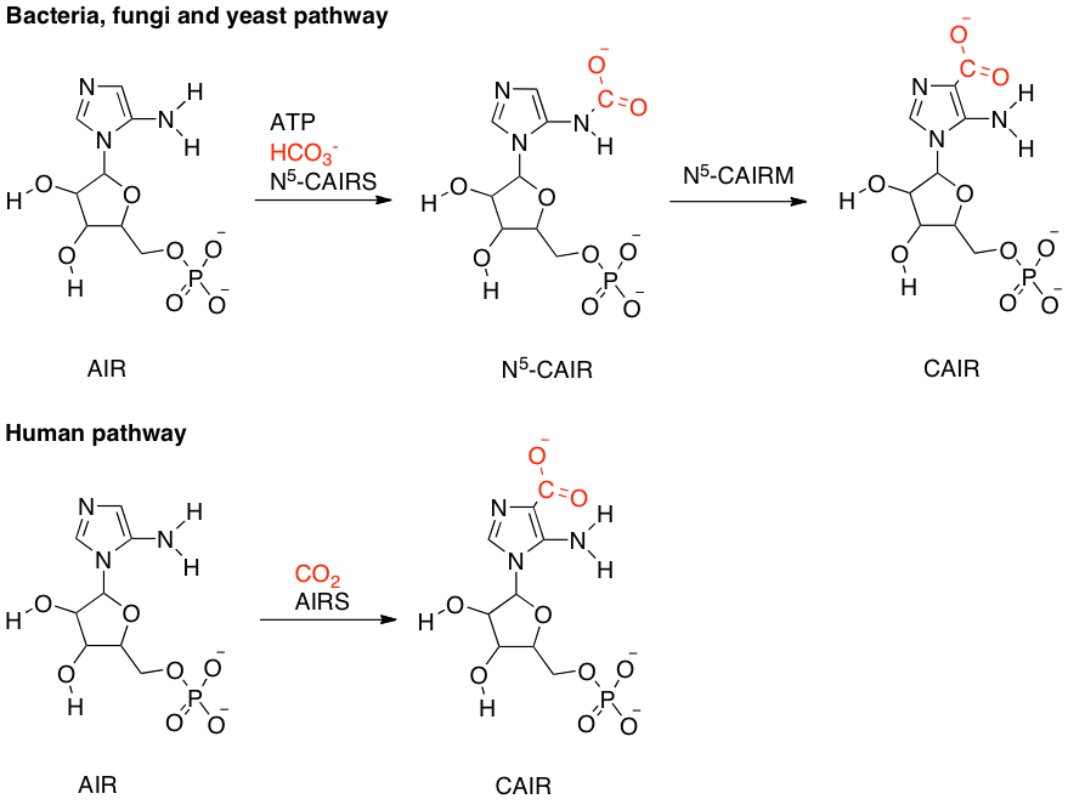


Figure 1.2. Contrasting reaction pathways for de novo purine biosynthesis in bacteria vs. humans

Although sequence analysis suggests an evolutionary link between AIR carboxylase and N^5-CAIR mutase, humans do not posses any enzyme homologous to N^5-CAIR synthetase.[50,52,53] Furthermore, it has been found that an N^5-CAIR synthetase deficit results in avirulant organisms, unable to proliferate in human serum or animal models used to predict disease progression.[54–56] Thus N^5-CAIR synthetase represents a novel target in the development of broad-spectrum antimicrobial agents to combat drug resistant infections.[57]

1.1.4 Propionyl-CoA carboxylase

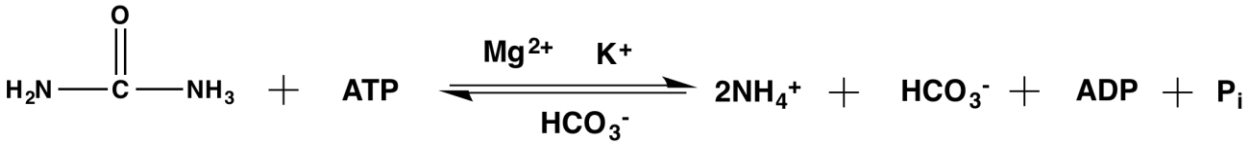
Propionyl-CoA carboxylase (PCC) catalyzes the carboxylation of propionyl-CoA to form (S)-methylmalonyl-CoA. Like ACC, and PC it is a biotin-dependent enzyme making it similar in

structure and function.[58,59] PCC is an alpha(6)-beta(6)-dodecamer.[60] The BC and BCCP domains responsible for carboxylation and transfer of biotin, respectively, are located in the alpha subunit of PCC, while the CT domain is in the beta subunit.[61,62]

Propionyl-CoA, the substrate that is carboxylated by PCC, is the final product of odd-chain fatty acid metabolism as well as a metabolite of most methyl-branched fatty acids.[63] Additionally it is the main metabolite of valine, and a metabolite of both isoleucine and methionine. It is also an important precursor to glucose. The (S)-methylmalonyl-CoA produced by PCC cannot be readily utilized by animals and is therefore converted first to (R)-methylmalonyl-CoA by a racemase then succinyl-CoA by methylmalonyl-CoA mutase.[59,60,64] The succinyl-CoA is then converted to oxaloacetate and malate in the Krebs cycle and the malate can then be utilized in gluconeogenesis.[59] The condition known as propionic academia results when PCC is not functional. As a result propionyl-CoA is converted to proprionic acid rather than (S)-methylmalonyl-CoA. This leads to a bulid-up of propionyl-CoA, proprionic acid, ketones, ammonia, and other toxic metabolites in the blood stream causing damage to many major organs.[63,65,66] This condition is characterized by vomiting, dehydration, acidosis, low muscle tone, seizures, and lethargy and quickly becomes life-threatening. Because this disease results from an inability to metabolize certain amino acids it is managed by avoiding consumption of the amino acids methionine, threonine, valine, and isoleucine and a low protein diet.[63,65] However the lack of these essential amino acids and limited protein consumption results in a new set of health concerns.

1.1.5 Urea amidolyase

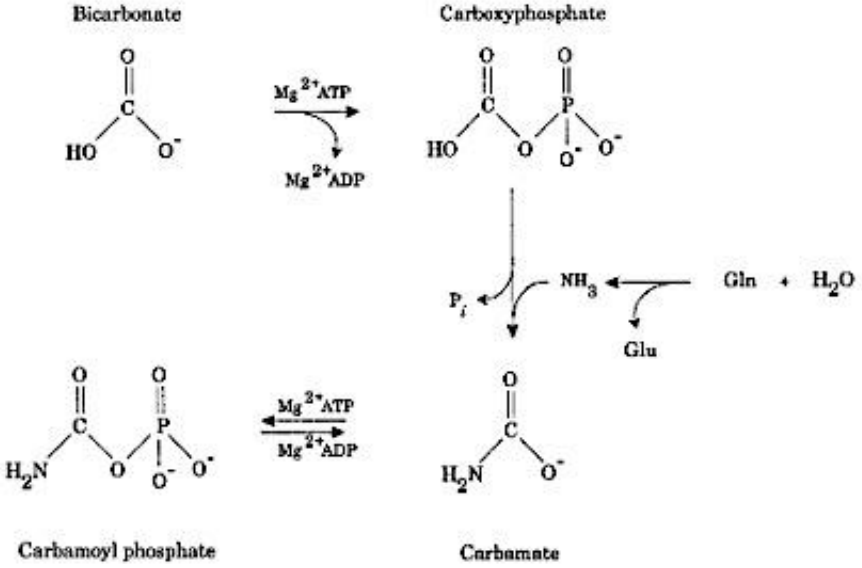
Most organisms generate urea as a result of degradation of nitrogen-containing molecules. For mammals the urea generated in these processes is simply excreted in urine. However plants, fungi, algae, and bacteria possess additional enzymes enabling them to convert urea into ammonia to be used as a source of nitrogen.[67–69] The enzyme urea amidolyase, and in some cases urease, is responsible for catalyzing this conversion of urea to ammonia.[67,70] Urea amidolyase has two activities, urea carboxylase (UC) and allophanate hydrolase (AH).[68,71–76] Along with its role in the utilization of urea, urea amidolyase plays an important role in many other processes including pyrimidine nucleic acid precursor degradation, and yeast-hyphal transition, an important mechanism in the pathogen *Candida albicans* used to escape host defense.[77–79]



Scheme 1.2. Reaction catalyzed by urea amidolyase. The decomposition of urea with formation of carbon dioxide and ammonia.[75]

Like ACC, PC, and PCC the UC activity is comprised of a BC, CT, and BCCP domain.[30,80] As expected, the BC and BCCP domains are highly conserved while the CT domain is distinct.[81,82] This is because the first step catalyzed by BC in which biotin is carboxylated by bicarbonate then transferred to the the CT domain by the BCCP is shared by each enzyme. However, in the second step catalyzed by CT the carboxyl group is transferred from biotin to varying substrates depending on the enzyme.[15,17]

1.1.6 Carbamoyl phosphate synthetase



Scheme 1.3. Reaction catalyzed by carbamoyl phosphate synthetase.[83]

Carbamoyl phosphate synthetase (CPS) catalyzes one of the most intriguing reactions ever described in biochemistry. That is the production of carbamoyl phosphate from bicarbonate, two molecules of ATP and either glutamine or ammonia.[83–85] In the first step, bicarbonate reacts with ATP to produce carboxyphosphate. The carboxyphosphate then reacts with ammonia or glutamine, releasing inorganic phosphate to form carbamic acid. The carbamic acid is then phosphorylated by a second molecule of ATP to produce carbamoyl phosphate.[38,83,86–88]

This is the first committed step in the biosynthesis of pyrimidine and arginine for prokaryotes and eukaryotes and in the urea cycle for most terrestrial vertebrates.

There are three known types of CPS. CPS-I is found in the liver, uses ammonia as a bubstrate, and only takes place in the presence of *N*-acetylglutamate.[86,89,90] CPS-II is located in

the cytosol of animals and has been found in *E. coli* and can utilize glutamine or ammonia as a substrate.[84] CPS-III is only present in the liver of spiny-dog fish and can utilize either glutamine or ammonia but requires the *N*-acetylglutamate cofactor.[91]

1.2 Carboxyphosphate

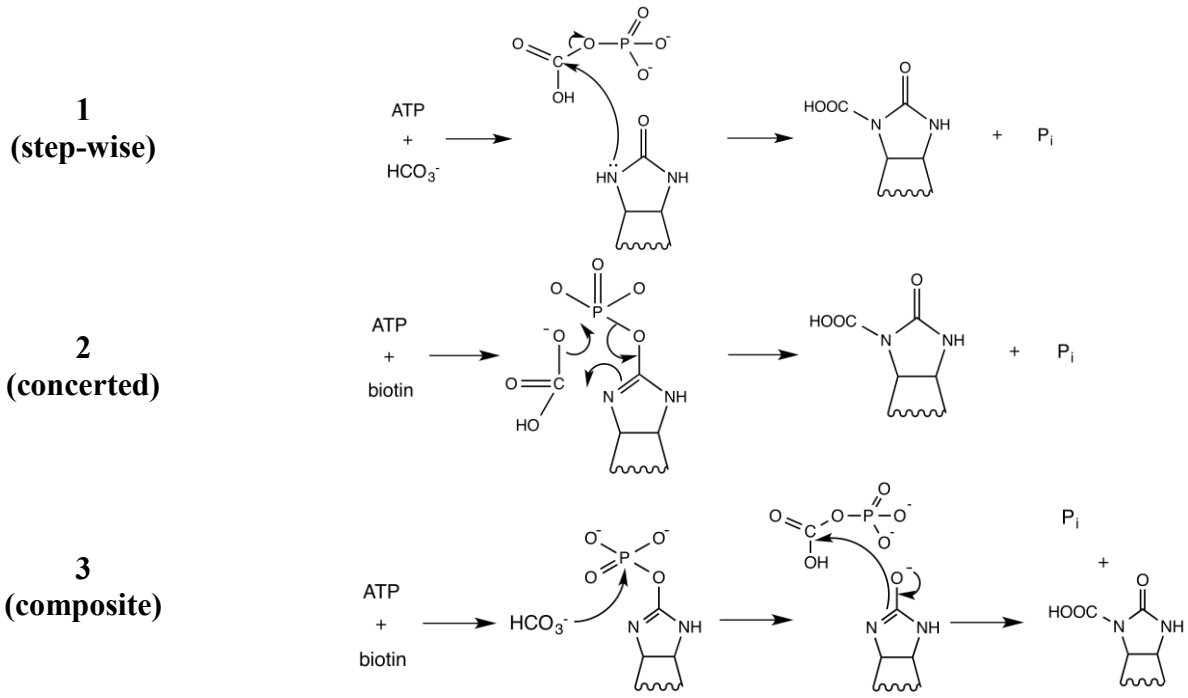
Carboxyphosphate (CP) is a common intermediate shared by all the reactions discussed above. It is generated via the SN2 reaction of bicarbonate with ATP.[15,17,92,93] Despite a great deal of interest in the mechanism by which this intermediate is generated and goes on to carboxylate a nucleophilic substrate it has never been directly detected. The majority of the work that has been done to understand this mechanism has been focused on the biotin-dependent enzymes: acetyl-CoA carboxylase, propionyl-CoA carboxylase, urea amidolyase, and pyruvate carboxylase.[30,58,94] These enzymes all share the same initial step in which biotin is carboxylated in an ATP-dependent reaction from bicarbonate. Due to the fact that bicarbonate is a poor electrophile and biotin a poor nucleophile, it was proposed that one of these substrates must be activated through phosphorylation by ATP.[41,93,95]

1.2.1 Mechanism proposed by Hansen & Knowles

Initially three possible mechanisms were proposed by Hansen and Knowles for the carboxylation of biotin from bicarbonate and ATP.[96]

In the first, bicarbonate is activated by ATP to form a carboxyphosphate intermediate that collapses to produce carboxybiotin and inorganic phosphate. In the second, biotin is activated by ATP to form a reactive, *O*-phosphobiotin intermediate that then undergoes an unprecedented six-

electron electrocyclic reaction with bicarbonate to form carboxybiotin while releasing inorganic phosphate.



Scheme 1.4. Three mechanisms proposed by Hansen & Knowles for the formation of carboxybiotin.[97]

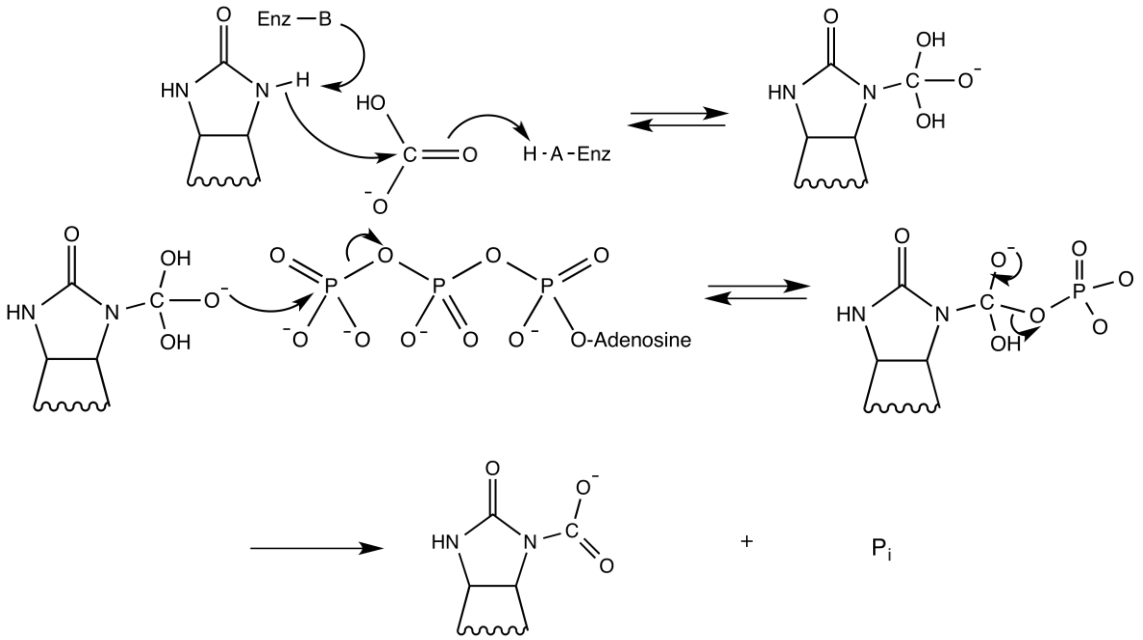
The third mechanism begins the same as the second with the formation of an *O*-phosphobiotin intermediate, which is then attacked by bicarbonate to form carboxyphosphate that reacts with the enolate form of biotin to form carboxybiotin and inorganic phosphate.

To differentiate between these three proposed mechanisms the oxygens of the terminal phosphate of ATP in the mechanism catalyzed by pyruvate carboxylase were isotopically labeled.[92] It was found that the conversion of ATP to ADP and P_i takes place with overall stereochemical inversion of the configuration at the phosphorus.[96] Both the step-wise and the concerted mechanisms are expected to occur with inversion of the phosphorus while the composite mechanism is not. To further differentiate between Mechanism 1 and 2, additional experimental evidence has been examined.

Kaziro *et al*. analyzed the reaction catalyzed by propionyl-CoA carboxylase by labeling bicarbonate with ^{18}O. They found that the labeled oxygen from bicarbonate is transferred to the P_i produced in this reaction as well the biotin-independent ATPase reaction catalyzed by biotin carboxylase strongly suggesting a direct reaction between bicarbonate and ATP that does not require biotin.[40,59] Additionally, biotin carboxylase and pyruvate carboxylase can produce ATP from the carboxyphosphate analogue carbamoyl phosphate and ADP in the reverse reaction and another carboxyphosphate analogue, phosphonoacetate is an inhibitor of pyruvate carboxylase.[44,98,99]

1.2.2 Mechanism proposed by Kluger and Taylor

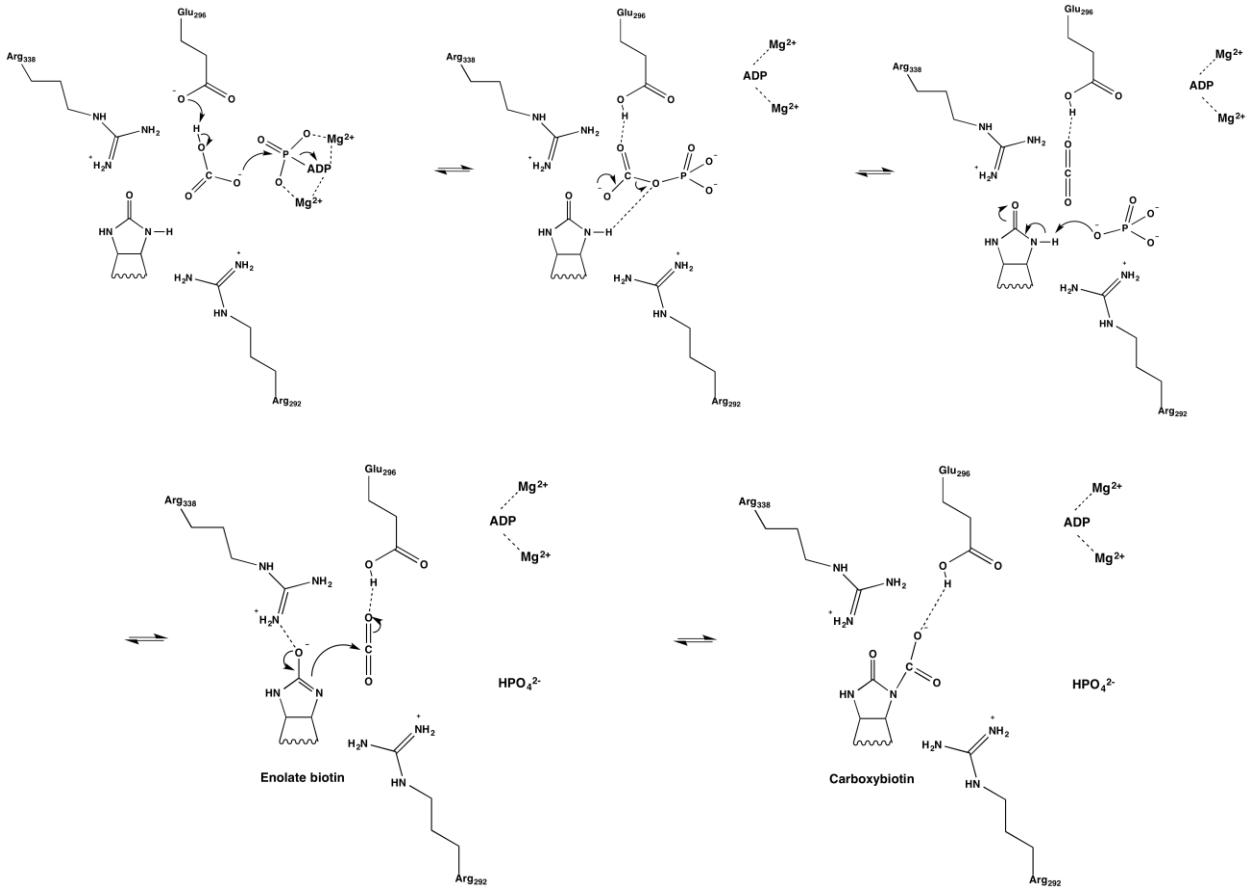
In 1991, Kluger and Taylor proposed another possible mechanism based on model studies in which the conjugate base of urea formed an adduct with an adjacent carboxylate and model studies of the intramolecular dephosphorylation of the carbonyl hydrate of methylacetoin diethyl phosphate.[100] In this mechanism, biotin reacts with bicarbonate to form an adduct. The adduct then reacts with the γ-phosphate of ATP producing a phosphocarboxy adduct of biotin. This adduct then undergoes an intramolecular reaction to expel P_i leaving the carboxybiotin product. While this mechanism is chemically feasible it is not consistent with evidence that has suggested a direct reaction between bicarbonate and ATP.[30,41,95]



Scheme 1.5. Mechanism for carboxylation of biotin proposed by Kluger & Taylor.[100]

1.2.3 Mechanism proposed by Chou et al.

Recently, Chou *et al.*[20] solved the crystal structure of biotin carboxylase in complex with magnesium, bicarbonate, and ATP and proposed a mechanism based on this structure. It has often been suggested that bicarbonate must be deprotonated by a catalytic base to initiate the reaction with ATP because carbonate is 10^6-fold more basic.[41,101] A great deal of research has been devoted to identifying this catalytic base but has been unsuccesful.[26,66,102] Based on the close proximity of glutamate 296 to bicarbonate, Chou *et al*. stated that it is the catalytic base responsible for the deprotonation of bicarbonate.[20]



Scheme 1.6. Mechanism for catalysis proposed by Chou *et al*.[20]

This mechanism has become the generally accepted mechanism for the reaction catalyzed by biotin carboxylase. However, it is questionable whether this mechanism is plausible for this enzyme. Bicarbonate has a pK_a value of 10.3[103] requiring a much stronger base than glutamate for deprotonation. Chou and co-workers suggest that the local environment alters the pK_a values significantly for bicarbonate and Glu296 to overcome this gap and facilitate deprotonation. To further support their mechanism they also performed mutagenic studies, but the kinetic data resulting from the E296A mutant showed only a 45-fold loss in activity measured from k_{cat}/K_m towards bicarbonate. In contrast, various mutations to Arg338 and Lys238 have been shown to render the enzyme essentially inactive, suggesting that they are much more likely to play critical

roles in the mechanism then Glu296.[20,26,66,102] Finally, there has been some contradictory evidence for the general suggestion that carbonate is more reactive towards ATP than bicarbonate.[101]

1.2.4 Herschlag and Jencks

In 1990, Herschlag and Jencks studied the reaction rates of carbonate and bicarbonate in phosphoryl transfer reactions with γ-picoline monoanion (PicP).[101] They found that the bicarbonate reaction was nine times faster than carbonate and is approximately 200 times faster when corrected for differences in basicity. It was also discovered that the proton from bicarbonate hydrogen bonds with PicP in the transition state. However, they concluded that it was not an added stabilization energy gained by this hydrogen bonding that accounted for the more reactive nature of bicarbonate (based on additional studies with formate and acetate) but the avoidance of the charge repulsion that would be present in a reaction between carbonate and PicP.[101]

1.2.5 Properties of CP

Sauers *et al*. made a series of predictions about the properties of carboxyphosphate via structure-activity studies on a series of aliphatic alcohols across a range of pK_a values.[104] In these studies, they used the pK_a values to predict the position of the equilibrium of Equation 1.1 and the rate of decomposition of alkyl monocarbonates (Equation 1.2).

$$HCO_3^- + ROH \rightleftharpoons ROCOO^- + H_2O \qquad (1.1)$$

$$ROCOO^- \rightleftharpoons RO^- + CO_2 \qquad (1.2)$$

The study of alkyl monocarbonate decomposition by Faurholt established that it occurs in a stepwise fashion.[105,106] Sauers *et al*. utilized the rate constants generated from this study along with pK_a values of the alcohols to develop Brønsted plots. From these plots along with the pK_a of the phosphate dianion, the expected rate and equilibrium constants for CP were generated. The ΔG° value for the hydrolysis of CP was estimated to be -3.6 kcal/mol through reversal of Equation 1.1. Based on this finding, Sauers *et al*. predicted that this value would become more negative in dilute solution or at lower pH, pushing the reaction toward hydrolysis through protonation of the phosphate dianion and bicarbonate. Additionally, they found the rate of decomposition of CP to be 10 sec^{-1} leading them to propose that CP is only present in very low concentrations in aqueous solution of bicarbonate and phosphate dianion. Thus, they concluded that CP has two roles in the active site: (1) a kinetic role, providing a supply of low-entropy, activated carbon dioxide from dehydration of bicarbonate; and (2) a thermodynamic role which forces the carboxylation reaction towards the formation of products.

Finally, Sauers *et al*. used their thermodynamic property calculations to estimate the half-life of CP to be *ca*. 70 miliseconds.[104] To build upon this work, Powers and Meister calculated the half-life of enzyme-bound CP in 67 % DMSO to be 2.5 minutes.[107] While, Rubio and Grisiola estimated the half-life of enzymatic CP in a mixture of water and acetone to be only 2.8 seconds based on studies with frog liver CPS.[108]

1.2.6 Synthesis and isolation attempts for CP

It is apparent based on the short estimated half-life, that CP is highly unstable in aqueous solution making isolation impractical, but attempts to isolate the enzyme-bound CP, which is expected to be more stable, have also proved unsuccessful.[37,88,92,100,104,109,110] However, the

trimethylated derivative of CP has been successfully trapped using diazomethane both in the both solution and the within the enzyme. Furthermore, all of the alkyl triesters listed in Figure 1.3. are found to be stable.

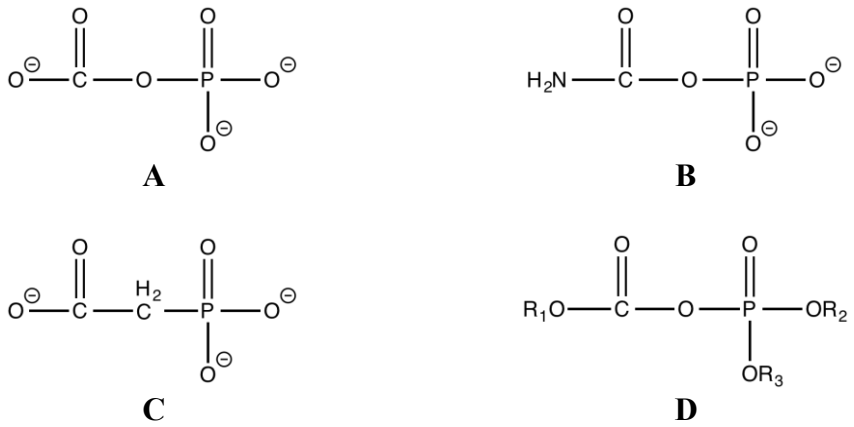


Figure 1.2. Carboxyphosphate and its structural analogs. (A) Carboxyphosphate, (B) carbamoylphosphate, (C) phosphonoacetate, and (D) ester of carboxyphosphate, R_1= *p*-nitrobenzyl, R_2=R_3=benzyl.

Many attempts have been made to synthesize CP as well, but have also proven unsuccessful. However, Griffith and Stiles were able to synthesize carbobenzoxy phosphate derivatives.[111] In these studies they discovered that metal ions doubled the rate of hydrolysis for the *p*-nitro analog. They observed that during hydrolysis by $H_2{}^{18}O$ there is the ^{18}O is not incorporated into the phosphate and proposed that hydrolysis of these derivatives must occur via C-O bond cleavage meaning that the water attacks the carbonyl oxygen.

In the structural analogue of CP, phosphonoacetate, the labile C-O bond is replaced by a C-C bond. This alteration results in inhibition of the phosphoenol pyruvate carboxylase enzyme, which is also believed to involve a carboxyphosphate intermediate but does not utilize ATP.[112] Carbamoylphosphate is a stable intermediate involved in the biosynthesis of arginine and pyrimidine nucleotides. Both BC and PC can use carbamoylphosphate in the reverse reaction to

convert ADP back to ATP.[98,112] It has been proposed by Allen and Jones that decomposition of carbamoylphosphate occurs via cleavage of the P-O bond in the monoanion and the C-O bond in the dianion.[113] Given the structural similarity between CP and carbamoylphosphate, it is reasonable to expect CP to follow a similar decomposition mechanism.

1.3 Objective of Dissertation

The general mechanism shared by several ATP-grasp enzymes involving the formation of a carboxyphosphate intermediate through reaction of bicarbonate and ATP has important implications for the development of new treatments for several of the most common and deadly diseases including cancer, diabetes, obesity and microbial infections. However, despite decades of experimental work, many details of these mechanisms remain unclear. While unsuccessful in fully elucidating many of the mechanistic details, the wealth of experimental data does support the involvement of a reactive CP intermediate. Most of the uncertainty surrounding this process is due to the instability of CP resulting in its failure to be synthesized, isolated, or analyzed directly. For this reason, this dissertation focuses on the use of computational methods to characterize unknown details of the CP intermediate.

Although the experimental evidence overwhelmingly supports the involvement of a CP intermediate in the mechanisms of these enzymes, very little is known about the structure, conformation, or even ionization state. Despite clear challenges to experimental characterization of CP, only one computational study has been reported to date.[114] Unfortunately, it lacked thorough analysis of the accuracy of the functional used for this system, investigation of different possible charge states and conformations. It also resulted in a calculated energy barrier of over 40 kcal/mol for the formation of carboxyphosphate. Thus, it is crucial to examine the details

absent from this study in an attempt to find a more energetically feasible pathway to carboxyphosphate formation. Specifically, it is crucial to determine not only the conformational stability of CP, but also identify the factors impacting differences in stability in order to deduce how ATP-grasp enzymes are able to generate and stabilize this reactive intermediate. The first task is to establish an inexpensive level of theory that can accurately model the structure and energetics of CP and be utilized in more sophisticated computations. Since there is no experimental data available to use as a reference, energetics and structures of CP computed with density functionals are evaluated relative to high-level quantum chemical calculations (MP2/aug-cc-pVQZ and CCSD(T)//MP2/aug-cc-pVQZ). Structural and energetic convergence of the mean average deviation (MAD) as a function of quantum chemical method and basis set was evaluated to determine appropriate level of theory.

Second, the stability of different possible arrangements and conformations for the monoanion, dianion and trianion forms of CP were examined. One of the most basic questions remaining in regards to CP is that of ionization state. Although it is most commonly depicted as a trianion, the stability of the different ionization states has yet to be seriously addressed. It is critical to establish from which charge state CP is most likely undergo decomposition, because different ionization forms may proceed through different mechanisms, as seen in monoanion and dianion acyl phosphates. Thorough analysis of different possible structures and conformations reveals that for each ionization state there are several different structures from which decarboxylation may take place. This structural analysis led to the discovery that CP is most stable when it forms an intramolecular charge-assisted hydrogen bond giving rise to a conformation resembling chair cyclohexane that has been dubbed pseudochair (PC). Both mono

and dianion charge states form this structure adding credence to the notion that the proton of bicarbonate must be retained during formation of CP.

Finally to further distinguish between the different ionization states of CP, the theoretical calculation of acid dissociation constants or pK_a values was undertaken. To establish an appropriate method and evaluate its accuracy, a series of test molecules of similar functionality were used to represent each of the three pK_a values of CP. Values within 1 pKa unit were achieved using the M06-2X/jul-cc-pVTZ/SMD and CCSD(T) corrections for aqueous solution. The results from this study indicate CP should exist majorly as a dianion at physiological pH. However, several basic residues are present in the active site of biotin carboxylase along with many of the other ATP-grasp enzymes where this reaction takes place. Future work should involve calculation of any shift in the pK_a values presented here to properly assess the impact the active site environment may have on the stability of different protonation states of CP.

The reactions catalyzed by acetyl-CoA carboxylase, pyruvate carboxylase, N5-CAIR synthetase, proprionyl-CoA carboxylase, urea amidoylase, and carbamoyl phosphatase are of critical importance to many diverse biological processes. A better understanding of the unknown mechanistic details shared by these enzymes will allow for substantial development in the treatment for several of the most common and deadly human diseases facing the world today.

Chapter 2

2 Electronic Structure Modeling

2.1 Quantum Mechanical Calculations

Quantum mechanics is the mathematical description of the principles that describe how and why electrons and nuclei combine to form atoms and those atoms chemical bond to form molecules.[115] Currently there are a wide variety of programs available for electronic structure calculation, the work described in this dissertation has been performed with Gaussian 09.[116] Theoretically, quantum mechanics is capable of calculating all observables associated with a molecular system. However in practice, exact solutions to these equations only exist for single electron systems (e.g. hydrogen).[117,118]

2.2 Schrödinger Equation

What differentiates quantum mechanics from classical mechanics is that electrons and nuclei are treated as waves rather than particles. Therefore, it is logical to begin the discussion of quantum mechanics with the De Broglie equation of a wavelength (λ) as defined below:

$$\lambda = \frac{h}{p} = \frac{h}{mv} \qquad (2.1)$$

where h, p, m, and v are Planck's constant, momentum, mass, and the speed of the particle, respectively.[119] Because Planck's constant is extremely small ($6.62606957(29)\times10^{-34}$ J•s), it follows that only very light particles (i.e. electrons) will have a measurable wavelength. Thus, we

may describe very small particles as waves, while those with larger mass should be treated as particles and governed by Newton's equations of motion.

Quantum mechanics allows us to calculate the energy and "location" of a wave of a given mass and charge. To understand how this works let's first consider the simplest equation of a wave moving in one direction (*x*).

$$\psi(x) = \sin\left[\frac{2\pi}{\lambda}x\right] \qquad (2.2)$$

where $\psi(x)$ is the wavefunction and λ is the wavelength. The energy of the wavefunction can be then obtained by taking the second derivative with respect to *x,* and using the De Broglie equation to substitute for λ.

$$\frac{\delta^2}{\delta x^2} = -\left[\frac{2\pi}{\lambda}\right]^2 \Psi = -\frac{p^2}{\hbar^2}\Psi = -\frac{2m}{\hbar^2}T\Psi \qquad (2.3)$$

where T is the kinetic energy ($\frac{1}{2}mv^2$) and ħ is h/2π. There is no potential energy because the wave is not interacting with anything so the total energy, E = T. Therefore the equation becomes:

$$\left[-\frac{h^2}{2m}\frac{d^2}{dx^2}\right]\Psi(x) = E\Psi(x) \qquad (2.4)$$

which is equivalent to:

$$-\frac{1}{2}\nabla^2\Psi(x) = E\Psi(x) \qquad (2.5)$$

Giving the Schrödinger equation for a single-electron system:

$$\hat{H}\Psi(x) = E\Psi(x) \qquad (2.6)$$

Where $\hat{H}$ is the Hamiltonian operator. In practice, the Schrödinger equation is applied in reverse. The Hamiltonian for the system of interest is first defined then a wavefunction that satisfies the Schrödinger equation is found. An isolated electron moving in the x direction is defined by the equation above and is solved by locating the wavefunction that satisfies it. This will be a function for which the second derivative equals the function multiplied by a constant (*E*) or a

sin(x) function. It is important to note that sin(2x), sin(3x), etc. are also solutions and that the Schrödinger equation has many solutions

$$\hat{H}\Psi_n(x) = E(n)\Psi_n(x) \qquad n = 1,2,\ldots \qquad (2.7)$$

each with different energies and wavefunctions

$$\Psi_n(x) = sin\left[\frac{2\pi}{\lambda}nx\right] \qquad (2.8)$$

where n is an integer and the quantum number. The energy increases with increasing quantum numbers and the energy associated with the lowest quantum number is the ground state energy.[115]

If we consider the case of a hydrogen atom, the Hamiltonian must now include a kinetic energy operator for both an electron and proton and a potential energy operator to describe the Coulombic interaction between the two particles. This interaction is -1/r in atomic unit where $r = |\boldsymbol{r} - \boldsymbol{R}|$ is the distance between the electron [at position $\boldsymbol{r}$ = (x, y, z) and nucleus (at position **R**).

$$\hat{H}\Psi_n(\boldsymbol{R},\boldsymbol{r}) = E(n)\Psi_n(\boldsymbol{R},\boldsymbol{r})$$

$$\left[-\frac{1}{2m_H}\nabla_{\boldsymbol{R}}^2 - \frac{1}{2}\nabla_{\boldsymbol{r}}^2 - \frac{1}{r}\right]\Psi(\boldsymbol{R},\boldsymbol{r}) = E(n)\Psi_n(\boldsymbol{R},\boldsymbol{r}) \qquad (2.9)$$

Unfortunately, the Schrödinger equation cannot be solved for more than one particle. To circumvent this problem the Born-Oppenheimer approximation can be employed.[120] The Born-Openheimer approximation states that because electrons are so much smaller than nuclei,[120] it can be assumed that the nuclei appears stationary and allows for the position of the nucleus and electrons to be treated separately with a separate Schrödinger equation for each as shown below:

$$\left[-\frac{1}{2m_H}\nabla_{\boldsymbol{R}}^2\right]\Psi(\boldsymbol{R}) = E_{nuc}(n_x, n_{y,} n_z)\Psi_{n_x n_y n_z}(\boldsymbol{R}) \qquad (2.10a)$$

$$\left[-\frac{1}{2}\nabla_{\boldsymbol{r}}^2 - \frac{1}{r}\right]\Psi(\boldsymbol{r}) = E_{elec}(n)\Psi_n(\boldsymbol{r}) \qquad (2.10b)$$

The total energy can then be calculated as the sum of the two equations ($E = E_{nuc} + E_{elec}$), where the Hamiltonian of equation 2.10a is that for a free particle but now in 3 dimensions instead of just one and *m* is the mass of a hydrogen atom, m_H. This equation can be solved separately for the x-, y-, and z-direction.[121]

The electronic Schrödinger equation can also be solved yielding:

$$\Psi(\boldsymbol{r}) = \Psi(r) = \frac{1}{\sqrt{\pi}} e^{-r} \qquad (2.11)$$

In kcal/mol:

$$E_{elec}(n) = -\frac{627.51}{2n^2} \qquad n = 1,2,\ldots \qquad (2.12)$$

However, for systems with two or more interacting electron's Schrödinger's equation cannot be solved exactly because electron-electron repulsions mean that the electrons must be coupled.[122,123] For this reason, further approximations are used to find solutions to the Schrödinger equation.

2.3 Variational Theorem

One useful tool for evaluating possible solutions to the Schrödinger equation is the variational theorem.[124] This theorem states that the value of an eigenvalue or energy ($\langle E \rangle$) predicted by a trial wavefunction will always be larger than the real energy (E_0) resulting from an exact solution to the Schrödinger equation.[125]

$$\langle E \rangle = \frac{\int \psi^* \hat{\mathrm{H}} \psi d\tau}{\int \psi^* \psi d\tau} \geq E_0 \qquad (2.13)$$

This is demonstrated in the equation (2.13) above where the expectation value or average energy, $\langle E \rangle$, of a system is calculated by taking the intergral over all space, $d\tau$, resulting in a value that is always greater than the true energy of the ground state wavefunction, E_0. Because in

most cases, the true wavefunction is not known a trial wavefunction is used. Since the energy calculated from this trial wavefunction will always be larger than the true wavefunction it follows that the lowest energy calculated corresponds to the most accurate wavefunction making it a better representation of the true wavefunction. In this way, the variational theorem can assess the accuracy of a given method for generation of a trial wavefunction.[126]

2.4 Hartree-Fock Theory

Hartree-Fock (HF) is an approximation method that allows for the determination of the wavefunction and energy for many-body systems, or those containing many electrons.[123,124,127] The HF method assumes that the exact wavefunction for an *N*-body system can be approximated by a single slater determinant of *N* spin-orbitals. Using the previously discussed variational method, a set of *N*-coupled equations can be derived for the *N*-spin orbitals. The solution to these equations gives the Hartree-Fock wavefunction and energy for the system. In other words, the wavefunction of a many-electron system is partitioned into single electron wavefunctions (or orbitals) that take the form of a Slater determinant. Expression of the wavefunction as a Slater determinant satisfies the antisymmetric requirement, in which the sign of the wavefunction must be inverted upon exchange of any two electrons.[117]

Equation 2.14 represents the wavefunction for a *N*-electron wavefunction as a Slater determinant of orbitals, where $\phi_2(1)$ indicates electron number "1" in the second spin orbital which is equivalent to the product of the spatial and spin functions. The coefficient $\frac{1}{\sqrt{N!}}$ is simply a normalization factor. It should be noted that an additionally important property of the Slater determinant is that if any column or row is equal to zero, the determinant will also equal zero.

$$\psi = \frac{1}{\sqrt{N!}}\begin{bmatrix} \phi_1(1) & \phi_2(1) & \cdots & \phi_N(1) \\ \phi_1(2) & \phi_2(2) & & \phi_N(2) \\ \vdots & & \ddots & \vdots \\ \phi_1(N) & \phi_2(N) & \cdots & \phi_N(N) \end{bmatrix} \qquad (2.14)$$

An important consequence expressing the wavefunction as a Slater determinant is that it does not account for electron correlation. As a result, an individual electron does not experience repulsion from each other electron in the system, but instead experiences this repulsion as an average field generated by all other electrons. This is the major weakness of Hartree-Fock theory.[117,123,124,127] Despite these shortcomings, Hartree-Fock represents a major breakthrough in quantum mechanics and electronic structure calculations by simplifying the unsolvable many-electron Schrödinger into multiple simpler single electron equations (Equation 2.15), that each correspond to single-electron wavefunction, consisting of an orbital, ϕ_i, and orbital energy, ε_i. With each orbital describing the behavior of an electron in the net field generated by all other electrons.

$$\hat{f}_i\phi_i = \varepsilon_i\phi_i \qquad (2.15)$$

Where $\hat{f}_i$ is the single-electron Hamiltonian or the Fock operator. This operator accounts for the one-electron core Hamiltonian, (H^{Core}) the Coulomb operator (J_i), and the exchange operator (K_i), as shown in Equation 2.16.

$$\hat{f}_i = H^{Core} + \sum_i^N [J_i - K_i] \qquad (2.16)$$

$$H^{Core} = -\frac{1}{2}\nabla_i^2 - \sum_A^M \frac{Z_A}{|\boldsymbol{r}_i - \boldsymbol{R}_A|} \qquad (2.17)$$

$$J_i\phi_i = \phi_i \int \phi_j^* \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}\phi_j d\tau_2 \qquad (2.18)$$

$$K_i\phi_i = \phi_j \int \phi_j^* \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}\phi_i d\tau_2 \qquad (2.19)$$

H^{Core} is the energy of an electron (i) and the Coulombic attraction between that electron and the nucleus (A). This term is a stabilizing contribution to the electronic energy as indicated by the negative sign in Equation 2.17. The Coulomb operator (Equation 2.18) defines the electron-electron repulsion energy caused by each of two electrons in the j_{th} orbital. The exchange operator accounts for the electron exchange energy caused by the antisymmetry of the total N-electron wavefunction and is an artifact of utilizing a Slater determinant.

When locating solutions to the Hartree-Fock equations, it is important to note that because the wavefunction is represented as multiple single-electron wavefunctions, each electron is assumed to move in an average field representative of the nuclei and electrons present rather than interact directly with individual particles. As a result, each solution to the single-electron eigenvalue equation will affect the solutions to the remaining single-electron eigenvalue equations. To solve these equations, the self-consistent field method is employed. In this method, a set of trial solutions, ϕ_i, to the HF equations are generated and used to calculate the Coulomb and exchange operators. The single-electron eigenvalue equations can then be solved yielding a new set of orbitals and associated energies. These solutions are then input into a second iteration of this process, which is repeated until the differences between the current solution and the last is below a given threshold referred to as convergence criteria. When the solutions reach convergence the system is self-consistent.

Although HF was a major breakthrough its utility was seriously limited by its failure to properly account for electron correlation. However HF remains useful for initial predictions of chemical systems and their geometric parameters and provides the foundation for higher-level methods like Møller-Plesset and coupled cluster.[122]

2.5 Møller-Plesset Perturbation Theory

Moller-Plesset perturbation theory (MP) is one of several methods that attempt to correct HF by adding electron correlation effects. MP adds electron correlation effects via special case of Schrödinger perturbation theory (RS-PT). The basis of perturbation theory is the division of the Hamiltonian into two parts: the Hamiltonian operator, and a small perturbation, V, that is the correlation potential.

$$\widehat{H} = \widehat{H_0} + \lambda V \qquad (2.20)$$

where λ is an arbitrary real parameter that varies from 0 to 1 and controls the size of the perturbation. Due to the assumption that V is small relative to H_0 the perturbed wavefunction and energy can be expressed as a power series in terms of the parameter λ.

$$\psi = \psi^{(0)} + \lambda\psi^{(1)} + \lambda^2\psi^{(2)} + \lambda^3\psi^{(3)} + \cdots \qquad (2.21)$$

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \lambda^3 E^{(3)} + \cdots \qquad (2.22)$$

The perturbed wavefunction and energy are then substituted back into the Schrödinger equation:

$$(H_o + \lambda V)\left(\psi^{(0)} + \lambda\psi^{(1)} + \cdots\right) = (E^{(0)} + \lambda E^{(1)} + \cdots)(\psi^{(0)} + \lambda\psi^{(1)} + \cdots) \qquad (2.23)$$

In MP2, the unperturbed Hamiltonian, $\widehat{H_0}$, is defined as the sum of single-electron Fock operators, which gives the zeroth-order energy, $E^{(0)}$, and consists only of the sum of orbital energies, ε_i. This means that a corrective perturbation is not present at the zeroth-order, or the first-order. Therefore higher-level perturbations must be used in order to obtain the desired corrections. The lowest order perturbation that accounts for electron correlation is second-order (MP2). The value of $E^{(2)}$, which is the first peturbative contribution to the HF energy will always

be negative and therefore lowers the HF energy. Higher-level third- and fourth- and fifth- order MP calculations are also possible; However, it has been shown that these calculations show little to no improvement over MP2 at a significantly higher cost and are therefore rarely used. MP2 calculations have been shown to recover 80-90% of electron correlation neglected by HF representing a marked improvement.[122] However, one major disadvantage of MP2 is that it is not variational, meaning the computed energies may be higher or lower than the exact energy. As previously mentioned, the correction applied by these perturbations lowers the HF energy, so in contrast to HF, MP2 can give energies lower than the true value due to overestimation of electron correlation. While MP2 can be problematic in some instances,[128] it has also been shown to model a variety of systems quite well and yields very accurate geometries.[121,122,124]

2.6 Coupled Cluster Theory

Coupled Cluster theory (CC) is similar to MP in that it corrects HF to include electron correlation, but at a higher-level treatment beyond MP4.[121,122] Essentially, CC takes basic HF method and constructs multi-electron wavefunctions utilizing the exponential cluster operator to account for electron correlation.[129] CC is based on the assumption that the wavefunction can be expressed as an exponential ansatz:

$$|\Psi\rangle = e^{T}|\Phi_0\rangle \tag{2.24}$$

where $|\Phi_0\rangle$ is the reference wavefunction, which is normally a Slater determinant constructed from Hartree-Fock orbitals although other methods can be used. *T* is the cluster operator, which gives a linear combination of excited determinants from the reference wave function. The cluster operator is written as:

$$T = T_1 + T_2 + T_3 + \cdots, \tag{2.25}$$

where T_1 is the operator of all single excitations, T_2 is the operator of all double excitations and so on. The notation for coupled cluster methods typically begins with the letters "CC" for coupled cluster followed by S for single excitations (singles), D for double excitations (doubles), T for triple excitations (triples), and Q for quadruple excitations (quadruples). Therefore, CCSDT would indicate that the cluster operator takes the following form:

$$T = T_1 + T_2 + T_3 \qquad (2.26)$$

In this work, the CCSD(T) method is used. The parenthesis around the "T" indicates that triples are calculated using perturbation theory. Although CCSD provides a relatively inexpensive method that performs better than MP2, it is often not very accurate without some inclusion of triples.[129–131] This is true even near the equilibrium geometry, but especially when describing single-bond cleavage or diradical species. The inclusion of perturbative triples provides a much better description for closed-shell molecules near equilibrium geometry but breaks down when describing more complicated systems.[121] However, when used appropriately this level of theory has been shown to provide highly accurate results relative to experiment and is the most frequently employed coupled-cluster method.[129,131–133]

2.7 Density Functional Theory

Density functional theory (DFT) is one of the most widely used and versatile quantum mechanical modeling methods. Its popularity is due the fact that it is a good compromise of accuracy and computational expense.[123,134–145] For example, many systems are too large to be modeled with MP2 or CCSD(T) methods. Like these higher-level, post-SCF methods, DFT accounts for electron correlation. However, it lacks a description of dispersive forces.[134,146,147] In DFT, the electronic structure properties are determined from functionals, or functions of

functions, of electron density.[148] Although originally developed in the 1960s, DFT did not become widely used in quantum mechanical calculations until the 1990s after exchange and correlation terms were improved.[134] DFT methods have been quite successful at producing results that are in good agreement with experimental results for a variety of different systems.[133,149,150] However, DFT is still problematic in its description of intermolecular forces (particularly van der Waals), charge transfer excitations, transition states, global potential energy surfaces,[151] dopant interactions, and certain strongly correlated systems.[133,147,152–154]

2.7.1 Thomas-Fermi, Hohenberg-Kohn, and Kohn-Sham

The conceptual origins of density functional theory are based on the work of Llewellyn Thomas and Enrico Fermi.[155,156] Their model was based on the idea that the electron distribution of a system could be represented as a series of small volume elements in which the electron distribution is uniform but is variable between different elements. Although this work was crucial first step in the development of DFT, this approximation was found to be a poor approximation and this method has limited utility due to its faulty representation of kinetic energy, exchange energy, and complete lack of electron correlation.[118,122,134]

The next major breakthrough came in the form of two theorems proposed by Pierre Hohenberg and Walter Kohn.[148] The first states that if a system of two electrons with potentials $v_1(\vec{r})$ and $v_2(\vec{r})$, respectively have the same ground state density, $n(\vec{r})$ then:

$$v_1(\vec{r}) - v_2(\vec{r}) = \text{constant} \qquad (2.27)$$

As a consequence, the potential, and therefore all properties of the system, are uniquely defined by the ground state density. The second theorem states that for any positive integer, N, and potential, $v(\vec{r})$, a functional of the density exists, $F[n]$ such that

$$E_{(v,N)}[n] = F[n] + \int v(\vec{r})n(\vec{r})d^3r \qquad (2.28)$$

The Schrödinger equation allows us to calculate energy for a molecular system, provided we have defined the Hamiltonian ($\widehat{H}$). It follows then that the energy may be defined as a function of the Hamiltonian, $E[\widehat{H}]$, i.e. the energy is a function of the Hamiltonian function. The Hamiltonian is in turn a function of the number of electrons, *N*, and the positions, $\{\boldsymbol{R}_A\}$, and charges of the nuclei, $\{Z_A\}$. Therefore the energy is a function of these three variables $E(N, \{\boldsymbol{R}_A\}, \{Z_A\})$. These three variables are extracted from the electron density, $\rho(\boldsymbol{r})$ and the energy is a functional of the density, $E[\rho(\boldsymbol{r})]$. The positions of the nuclei are obtained from peaks in the density and the charge comes from the slopes of the corresponding peaks.

$$\left[\frac{\delta\bar{\rho}(\boldsymbol{R}_A)}{\delta \boldsymbol{R}_A}\right]_{R_A=0} = -2Z_A\bar{\rho}(0) \qquad (2.29)$$

Where $\bar{\rho}$ is the spherical averaged density. The number of electrons, *N,* can then be found by integrating the density. There are three key properties of the energy density functional, $E[\rho(\boldsymbol{r})]$. First, it is unique. Therefore different densities will always result in different energies.

$$E[\rho_1] \neq E[\rho_2] \; when \; \rho_1 \neq \rho_2 \qquad (2.30)$$

An important consequence of this property is that the exact density will equal the exact energy.

$$E_{exact} = E[\rho_{exact}] \qquad (2.31)$$

The energy density functional obeys the variational principle:

$$E[\rho] \geq E[\rho_{exact}] \qquad (2.32)$$

and can be calculated from a 1-electron functional as follows:

$$E[\rho] = \int d\boldsymbol{r}\, \rho(\boldsymbol{r})v(\boldsymbol{r}) \;\; where \; v(\boldsymbol{r}) = \frac{\delta E[\rho(\boldsymbol{r})]}{\delta \rho(\boldsymbol{r})} \qquad (2.33)$$

These properties make it a simple way to calculate the exact energy. Because $v(\boldsymbol{r})$ is a single-electron function, it can be expressed in terms of orbitals.

$$E[\rho] = \sum_{i=1}^{N} \langle \phi_i(r) | v(r) | \phi_i(r) \rangle \tag{2.34}$$

The orbitals corresponding to the lowest energy can now be found utilizing the SCF procedure described previously.

$$\hat{\hat{F}}^{KS} \phi_i^{KS} = \varepsilon_i^{KS} \phi_i^{KS} \tag{2.35}$$

This is the Kohn-Sham (KS) SCF procedure.[157] In principle, it should have the same computational cost as the HF-SCF procedure, but the KS orbitals will add up to the exact density and the converged energy will be the exact energy. Unfortunately, we do not know the form of the $v(\boldsymbol{r})$, so we must guess. To minimize the uncertainty involved, the equations are typically rewritten so that the function that is being guessed only accounts for a small portion of the total energy of a system. Thus the equation is rewritten in terms of kinetic energy, electron-nuclear attraction energy, and electron-electron repulsion energy.

$$E[\rho] = T[\rho] + V_{ne}[\rho] + V_{ee}[\rho] \tag{2.36}$$

Therefore, the Kohn-Sham operator is as follows:

$$F^{KS} = -\frac{1}{2}\nabla^2 - \sum_A Z_A R_{1A}^{-1} + \frac{1}{2}\int dr_2 \, \rho(r_2) \frac{1}{r_{12}} + v^{XC}(r)$$

$$where \; v^{XC}(r) = \frac{\partial E_{XC}[\rho(r)]}{\partial \rho(r)} \tag{2.37}$$

In this equation, all of the terms are known except for E_{XC}, which is the exchange-correlation energy.[148,157] E_{XC} is typically further divided into exchange, E_X, and correlation energies, E_C. There is no systematic way to improve the approximation of these quantities, but many different methods have been developed.

Some of the most common approaches currently in use to approximate the value of E_{XC} in DFT are the local spin density approximation (LSDA),[157–159] the generalized gradient approximation (GGA),[160–165] meta-GGA (M-GGA),[137,166] hybrid-GGA (H-GGA),[137,150,160,164,167] and hybrid meta-GGA (HM-GGA).[150,168,169] The LSDA method is based on the assumption that $E_{XC}\{\rho]$ at any point in space depends only on the spin density in that particular spatial area. The GGA method improves upon LSDA by utilizing both the spin density and the gradient of the density. GGA is further improved in M-GG to include the spin-dependent electronic kinetic energy density, H-GGA includes a percentage of HF exchange, and HM-GGA incorporates approximations from both M-GGA and H-GGA.

2.7.2 B3LYP

B3LYP is the most widely used functional in the field of chemistry.[147,153,170,171] Its popularity is due in large part to its correction of some the deficiencies inherent in its predecessors at a low computational cost. Specifically, it has been shown to give predictions of chemical phenomena in better agreement with experiment including molecular structure, atomization energies, ionization potentials, proton affinities and total energies.[118,136,172] B3LYP utilizes the H-GGA method to calculate the exchange-correlation energy and is defined as

$$E_{B3LYP}^{XC} = (1-a)E_{LSDA}^{X} + aE_{HF}^{X} + b\Delta E_{B88}^{X} + cE_{LYP}^{C} + (1-c)E_{VWN3}^{C} \qquad (2.38)$$

The first term is the exchange energy determined by the local spin density approximation as proposed by Slater.[157–159] The second term is the exact Hartree-Fock exchange energy. The third term is the gradient corrected exchange energy proposed by Becke.[136] The fourth term represents the correlation energy, including the gradient term, determined by Lee, Yang, and Parr.[173] The final term is the standard correlation energy of the VWN3 functional developed by Vosko, Wilk

and Nusair.[174] The coefficients a, b, and c were derived from a least squares fit to 56 atomization energies, 42 ionization potentials, 8 proton affinities, and 10 first row, experimentally determined total atomic energies.[136,175] The optimized values for these coefficients are $a = 0.20$, $b = 0.72$, $c = 0.81$. (p164)

B3LYP remains the most widely used functional despite many advances in the development of new DFT methods, but it is not without deficiencies.[172] Some of the major limitations include (1) inaccurate calculation of bond lengths and energies that involve transition metals,[146,168] (2) underestimation of barrier heights,[146,176,177] and (3) poor estimation of interactions involving medium range correlation energies, including van der Waals attraction, aromatic-aromatic stacking and alkene isomerization energies.[134,170,171]

2.7.3 MPW1K

MPW1K stands for the modified Perdew-Wang 1-parameter functional kinetics. It was developed by Truhlar and coworkers in an attempt to improve the accuracy of computed transition state energies.[167] It is an H-GGA method, but it only contains one parameter whereas B3LYP employs three. It originates with a version of the Perdew-Wang functional,[178] modified by Adamo and Barone[135] and the Perdew-Wang correlation functional.[179] The single parameter used is the amount of HF exchange energy, which has been optimized to fit the experimental data for 20 forward barrier heights, 20 reverse barrier heights, and 20 energies of reaction.[167] The exchange-correlation energy is calculated according to Equation 2.39, where the amount of exchange energy, *X*, has been optimized to 42.8.

$$E_{MPW1K}^{XC} = \left(1 - \frac{X}{100}\right)\left(E_{LSDA}^{X} + \Delta E_{MPW}^{X}\right) + \frac{X}{100} E_{HF}^{X} + E_{LSDA}^{C} + \Delta E_{PW}^{C} \qquad (2.39)$$

The first and second terms in parenthesis are the exchange energy decided by the local spin density approximation and the modified version of the Perdew-Wang gradient corrected exchange energy established by Adamo and Barone. The third term is the exact HF exchange energy. The last two terms represent the correlation energy from the local spin density approximation and the Perdew-Wang corrected correlation energy, respectively.

MPW1K has been shown to calculate kinetic results that are in good agreement with experiment. However, it does a poor job of computing atomization energies.[133,147,153,167,180–189]

2.7.4 BB1K

BB1K stands for Becke88-Becke95 1-parameter model for kinetics and was developed by Truhlar and coworkers to address the shortcomings of MPW1K.[154] BB1K also contains one parameter that has been optimized against a database of three forward barrier heights, three reverse barrier heights and 3 energies of reaction from the BH6 database. The resulting optimized value for the parameter, X, is 28.0.[189] Like B3LYP and MPW1K it is also an H-GGA method. The exchange-correlation energy for BB1K is defined as

$$E_{BB1K}^{XC} = \left(1 - \frac{X}{100}\right)(E_{LSDA}^{X} + E_{B88}^{X}) + \frac{X}{100}E_{HF}^{X} + E_{LSDA}^{C} + \Delta E_{B95}^{C} \qquad (2.40)$$

The first term is the exchange energy, the third is the exact Hartree-Fock exchange energy, and the fourth is the correlation energy. Each of these terms is derived from the local density approximation. The second term is the gradient correction for the exchange functional determined by Becke's 1988 gradient corrected exchange functional (B).[160] (s29) The last term is the correlation energy of Becke's 1995 τ-dependent gradient-corrected correlation functional (B95).[137]

BB1K does indeed improve upon the weaknesses of MPW1K and has been shown to perform 40% better for calculation of atomization energies. Additionally it has been shown to outperform several other DFT and hybrid DFT methods.[150]

2.7.5 M05-2X and M06-2X

M05-2X and M06-2X are part of Truhlar's new generation of HM-GGA density functional methods developed in large part to address the deficiencies of B3LYP.[190] This new generation of functionals, also known as the M-functionals, includes the M05 and M06 functionals. The "2X" indicates that double the amount of Hartree-Fock exchange is included in the M05-2X and M06-2X functionals relative to the M05 and M06 functionals, respectively. The amount of HF exchange, X, in M05 and M06 is 28 and 27, respectively. Therefore the HF exchange in M05-2X and M06-2X is 56 and 54, respectively. The M06 suite of functionals are improved from the M05 suite by using a linear combination of both the M05 and VSXC functionals developed by Van Voohris and Scuseria.[166,191] The training and accuracy assessment data sets used for parameterization were also expanded for the M06 suite. As a result of these improvements, M06-2X has been shown to accurately describe long-range interactions where M05-2X fails. Additionally, analysis of 496 data points across 32 databases revealed that M06-2X, M06, and M05-2X all yield results in excellent agreement with experiment for main-group kinetics, hydrogen bonding, π-π stacking, interaction energy of nucleobases, and alkyl bond dissociation energies.[169]

2.8 Basis sets

Basis sets are linear combinations of functions or basis functions that provide a mathematical description of the orbitals in a system that in turn provide an approximation of the total electronic wavefunction. Although Slater-type functions provide a better description of atomic orbitals, it is very difficult to calculate orbital overlap. However it has been shown that Slater orbitals can be approximated using linear combinations of Gaussian type orbital functions (GTOs), making computations significantly less demanding.[192,193]

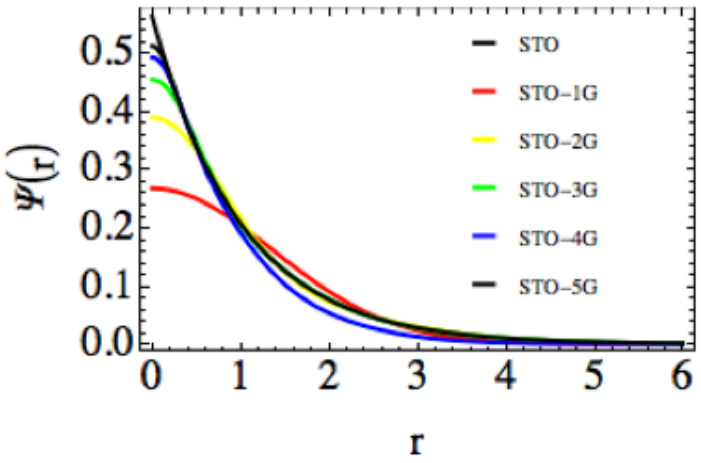


Figure2.1. Comparison of a 1s Slater-type orbital (STO) with five different linear combinations of Gaussian primitives truncated at specific orders (STO-1G, STO-2G, STO-3G, STO-4G, STO-5G) to approximate an STO.

The linear combination of Gaussian functions used to approximate each Slater orbital are referred to as contracted functions and the components of these functions are called primitives.

The main drawback of using GTOs instead of STOs is that GTOs decay as e^{-r^2} while STOs decay as e^{-r}. In Figure 2.1 a 1*s* STO is compared to several Gaussian primitives, truncated to one, two, three, four, and five GTOs defined as STO-1G, STO-2G, STO-3G, STO-

4G and STO-5G, respectively.[194] No matter how many GTOs are employed they will never correctly model the cusp at the origin shown in Figure 2.1 for the STO. As a result GTOs underestimate long-range overlap atoms as well as the electron density in close proximity to the nucleus.

There are many different basis sets that are made up of Gaussian functions. One of the most basic is the STO-*n*G basis set, which is considered a minimal basis set. In the STO-*n*G basis set *n* is an integer representing the number of Gaussian primitive functions that contribute to each basis function. This type of basis set is referred to as single-zeta, because each atomic orbital, core and valence, are computed from a single basis function. Using a single basis function to represent each atomic orbital is problematic because it can place strict limits on the movement of electrons in response to the molecular environment. These constraints are relieved as the number of basis functions used to represent each atom increases and the description of molecular orbitals becomes more accurate.[122,126]

Split-valence basis sets address some of the shortcomings of minimal basis sets by using multiple basis functions to represent each atomic orbital. They were designed based on the premise that chemical bonding is largely composed of valence orbital interactions.[194–199] Thus in these basis sets the core is represented by a single basis function while each valence orbital is represented by multiple basis functions. Following the notation used in minimal basis sets these basis sets are referred to as double, triple, or quadruple-zeta (and so on) basis sets where zeta corresponds to the number of basis functions used for each valence atomic orbital. Representing each valence orbital with multiple basis functions allows the electron density to adjust its spatial extent in response to its molecular environment.

One of the most popular basis sets are the Pople split-valence basis sets. The notation for this basis sets is typically X-YZg, where X is the number of primitives used for each core atomic orbital basis function, Y and Z indicate that the valence orbitals are represented by two different sets of basis functions, the first is made up of a linear combination of Y primitives, and the second is a linear combination of Z functions. In this example there are only two numbers after the hyphen indicating it is a double-zeta basis set. Triple and quadruple-zeta split valence basis sets would be denoted as X-YZWg and X-YZWVg, respectively.[197] Polarization or diffuse functions may also be added to these basis sets. Polarization functions add higher levels of angular momentum, allowing molecular orbitals to be more asymmetric. In the Pople basis sets these functions are denoted by an asterisk (*), or two asterisks (**) to indicate polarization functions have also been added to light atoms (H and He). Diffuse functions are denoted by a plus sign (+) or two plus signs (++) to indicate diffuse functions are also applied to light atoms. These functions aid in the description of the "tail" portion of atomic orbitals, or the area distant from the nuclei. These functions are important in the accurate description of anions or systems containing lone pair electrons.[200]

Another popular type of basis set is Dunning's correlation consistent basis set. The advantage of these basis sets over the Pople type basis sets is that they are designed to converge systematically to the complete basis set (CBS) limit with the aid of empirical extrapolation methods. The notation for these basis sets is aug-cc-pV*n*Z (*n* = D, T, Q, 5, 6,…).[177,201–203] The "aug" prefix indicates the presence of diffuse functions. Several different implementations now exist which remove select diffuse functions. For example, replacing the "aug" prefix with "jul" removes diffuse functions from the H and He atoms, "jun" also removes the highest angular momentum diffuse function from all other atoms, and "may" removes the two highest angular

momentum functions, to name a few. The "cc-p" in this notation stands for correlation-consistent polarized and the *V* signifies that they are valence-only. The *n* designates the number of polarization functions. The polarization functions used for each atom type for *n* = D, T, Q, 5, and 6 are summarized in Table 2.1.

Table 2.1. Polarization functions used in Dunning's correlation consistent polarized basis set.[177,201–203]

Atoms	cc-pVDZ	cc-pVTZ	cc-pVQZ	cc-pV5Z	cc-pV6Z
H	2s,1p	3s,2p,1d	4s,3p,2d,1f	5s,4p,3d,2f,1g	6s,5p,4d,3f,2g,1h
He	2s,1p	3s,2p,1d	4s,3p,2d,1f	5s,4p,3d,2f,1g	N/A
Li-Be	3s,2p,1d	4s,3p,2d,1f	5s,4p,3d,2f,1g	6s,5p,4d,3f,2g,1h	N/A
B-Ne	3s,2p,1d	4s,3p,2d,1f	5s,4p,3d,2f,1g	6s,5p,4d,3f,2g,1h	7s,6p,5d,4f,3g,2h,1i
Na-Ar	4s,3p,1d	5s,4p,2d,1f	6s,5p,3d,2f,1g	7s,6p,4d,3f,2g,1h	N/A
Ca	5s,4p,2d	6s,5p,3d,1f	7s,6p,4d,2f,1g	8s,7p,5d,3f,2g,1h	N/A
Sc-Zn	6s,5p,3d,1f	7s,6p,4d,2f,1g	8s,7p,5d,3f,2g,1h	9s,8p,6d,4f,3g,2h,1i	N/A
Ga-Kr	5s,4p,2d	6s,5p,3d,1f	7s,6p,4d,2f,1g	8s,7p,5d,3f,2g,1h	N/A

2.8.1 Basis Set Superposition Error

In quantum chemistry, basis sets are used to approximate the wavefunction. An infinite number of basis sets would be required to exactly represent the wavefunction accurately, therefore there can be some error associated with using finite or "incomplete" basis sets. This error is known as basis set superposition error (BSSE).[204] As atoms or interacting molecules move closer together their basis functions may overlap. These atoms or molecules can increase their basis set by borrowing functions from these neighboring atoms in order to lower the overall energy albeit artificially. When the short-range energies resulting from these mixed basis sets are compared to long range energies from unmixed sets there will be a mismatch resulting in an

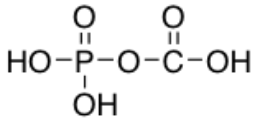
error. Of course as the CBS limit is approached this type of error is minimized which is exactly why Dunning's cc-pVnZ basis set was designed to converge to the CBS limit allowing one to systematically minimize this type of error. Truhlar's calendar basis sets also serve to address this type of error. Overuse of diffuse functions, particularly on small atoms like H and He, can lead to BSSE in smaller basis sets where larger atoms are not well described. For this reason, the "jul" basis set in which diffuse functions from H and He atoms is generally considered preferable to "aug" where full augmentation with diffuse functions is applied.[205]

Chapter 3

3 Determination of Appropriate Level of Theory

3.1 Introduction

Acetyl-CoA carboxylase, pyruvate carboxylase, N^5-Carboxyaminoimidazole ribonucleotide synthetase, propionyl-CoA carboxylase, urea amidolyase, and carbamoyl phosphate synthetase all exploit the abundant, but unreactive substrate, bicarbonate, phosphorylating it to generate a reactive carboxyphosphate intermediate, CP.[15] Unfortunately, the reactive nature of this intermediate has made direct detection and isolation impossible to date.[29,88,92,93,206] Failures to synthesize, isolate or characterize CP have been attributed to an estimated half-life of less than 70 ms in aqueous solution.[104]



carboxyphosphate (CP)

The experimental challenges of CP make computational studies one of the few available methods for interrogating this compound, yet to the best of our knowledge, no detailed computational studies of this intermediate have been reported. There are likely numerous reasons for this. First, the lack of experimental data presents a challenge for accurate benchmarking of computational data. Second, the pK_a values for CP are not known and thus, the exact charge of the molecule under physiological conditions is uncertain. Both the trianion and dianion of CP have been proposed to be involved in the mechanism of ATP-grasp enzymes;[20,41,92,101,102] However, chemical intuition suggests the dianion is most likely to be the significant intermediate utilized

by these enzymes. The presence of significant charge on the molecule presents substantial computational challenges.

Given that the protonation state of CP is fundamental to ATP-grasp enzymes and a detailed computational investigation has not been reported, it is both appropriate and timely to evaluate the three possible ionization states of CP and critically examine the structure and properties. In this study, we employ quantum mechanical calculations to interrogate the structures and energetics of all possible isomers of CP.

3.2 Computational Methods

Full geometry optimizations are carried out on different conformations of CP. For computations on all conformations of tri-, di-, and monoanionic CP, density functional theory (DFT),[148,157] and Moller-Plesset second-order perturbation theory (MP2)[207] were used to approximate the Schrödinger equation. Five different DFT methods have been employed using the Becke three parameter exchange functional[208] with the nonlocal correlation functional of Lee, Yang, and Parr[164,208] and local correlation of VWN functional[174] (B3LYP), Becke88-Becke95 1-parameter model kinetics (BB1K),[150] modified Perdew-Wang 1-parameter model for kinetics (MPW1K),[154] and Truhlar's hybrid meta-generalized gradient exchange correlation functionals (M05-2X and M06-2X).[168,169] Dunning's augmented correlation consistent polarized valence *n* zeta basis sets (aug-cc-pV*n*Z, *n* = D, T, Q, 5)[177,201–203,209] are combined with the above mentioned chemical methods to evaluate all ionization states and conformations in vacuum. These levels of theory have been thoroughly tested and benchmarked across a wide range of molecules and molecular properties.[16-19,21,22] Extrapolation to the complete basis set limit has been performed utilizing the extrapolation scheme proposed by Martin,[210] where the electronic

energy, *E(n)*, is a function of inverse powers of the angular momentum, *n* (*n* = D, T, Q and 5), plus the electronic energy in the complete basis set limit (E_∞). Due to the computer resources necessary for CCSD(T) geometry optimizations, CCSD(T) single point calculations with aug-cc-pVDZ and aug-cc-pVTZ basis sets were performed on the MP2/aug-cc-pV5Z optimized geometries. Frequency calculations were carried out to confirm all stationary points as minima on the potential energy surface, except when using the aug-cc-pV5Z basis set due to computer resource limitations.

The PCM continuum solvation model,[211,212] which has been described in detail elsewhere[26-30] was used to estimate the effect of solvent on the stability of each structure. Energy minimization has been carried out using the dielectric constant of water. Briefly in the PCM model, the solvent is treated as an infinite continuum of a specified dielectric. The solute or molecule is placed within a cavity defined as a series of interlocking van der Waals spheres centered on the atoms. The interaction between the solute and solvent, or free energy of solvation, is then calculated as a potential term added on to the molecular Hamiltonian. Specific to the PCM model, this potential term includes electrostatic, dispersive, and cavitation terms. Because self-consistency between the solute charge distribution and solvent reaction field is required to evaluate the amended Hamiltonian, this method is termed an SCRF or self-consistent reaction field method.

3.3 Results and Discussion

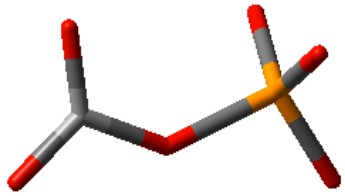


Figure 3.1. Trianionic conformation of carboxyphosphate at MP2/aug-cc-pVQZ.

The first step in our computational investigation was the systematic identification of all possible structures and conformations for the tri, di, and monoanionic states of carboxyphosphate. The different possible structures are largely defined by different arrangements of the protons. For this reason, only a single conformation exists as a stationary point for the trianion.

3.3.1 Dianionic carboxyphosphate

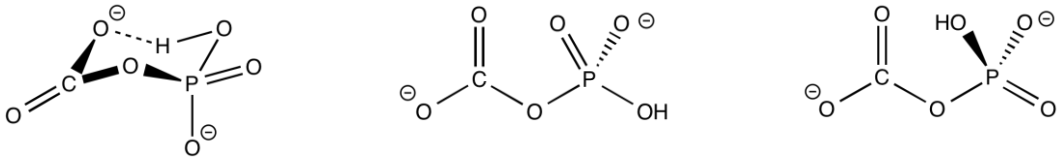


Figure 3.2. Three low energy structures of dianionic carboxyphosphate

For dianionic carboxyphosphate, six distinct structures have been located as energy minima. There are three lower energy conformers in which the phosphate side is protonated. The

lowest energy structure forms a charge assisted intramolecular hydrogen bond (CAHB) causing it to cyclize mimicking the structure of chair cyclohexane. For this reason, this conformation is referred to as *PC* for pseudochair. The other two structures are nearly identical energetically and are referred to as *PA* for anti- and *PG* for *gauche* based on the P-OH orientation with respect to the bridging oxygen O-C bond.

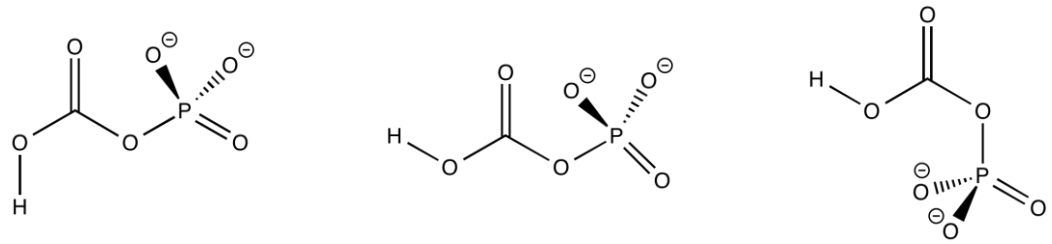
Table 3.1. Relative energies (ΔE^{elec} (kcal/mol) of *PC*, *PA*, and *PG* conformations of dianionic CP in vacuum and implicit water.

Level of theory	PC	PA^{vac}	PA^{wat}	PG^{vac}	PG^{wat}
MP2/aug-cc-pVDZ	0.0	10.9	5.0	11.5	6.1
MP2/aug-cc-pVTZ	0.0	12.2	5.9	12.1	6.5
MP2/aug-cc-pVQZ	0.0	12.4	5.9	12.2	6.5
CCSD(T)/aug-cc-pVTZ	0.0*	11.8*	$5.1^{\pm}$	11.7*	$5.8^{\pm}$

SP energy evaluation on *MP2/aug-cc-pV5Z $^{\pm}$ MP2/aug-cc-pVQZ

Three conformations in which the carboxylic acid side is protonated have also been located although they are significantly higher in energy due to the unstable concentration of charge on the phosphate group.

These conformations have been defined according to IUPAC convention as *EZ*, *ZZ*, and *ZE,* such that in *EZ* the *E*-conformation is defined for the carboxylic acid proton and the *Z*-conformation is defined for the phosphate group with respect to the ester. It should be noted that the *EE* conformation was also investigated but consistently underwent intramolecular hydrogen transfer to convert to the *PC* conformation at all levels of theory tested.



EZ **ZE** **ZZ**

Figure 3.3. Three high energy structures of dianionic carboxyphosphate

Table 3.2. Relative energies (ΔE^{elec} (kcal/mol) of *EZ*, *ZE*, and *ZZ* conformations of dianionic CP in vacuum and implicit water**.**

Level of theory	EZ^{vac}	EZ^{wat}	ZE^{vac}	ZE^{wat}	ZZ^{vac}	ZZ^{wat}
MP2/aug-cc-pVDZ	N/A	12.8	N/A	12.6	N/A	13.9
MP2/aug-cc-pVTZ	27.6	14.0	28.6	13.8	29.7	15.3
MP2/aug-cc-pVQZ	27.8	14.0	28.9	13.8	30.0	15.3
CCSD(T)/aug-cc-pVTZ	28.8*	$13.7^{\pm}$	29.5*	$13.5^{\pm}$	31.1*	$15.0^{\pm}$

SP energy evaluation on *MP2/aug-cc-pV5Z $^{\pm}$ MP2/aug-cc-pVQZ

3.3.2 Pseudochair carboxyphosphate

As mentioned previously, the PC conformation is named for its structural similarity to the chair conformation of cyclohexane. However, due to the heterocyclic nature of the ring and the tendency towards a more linear geometry about the hydrogen bond there are significant deviations. Most notably, non-uniform bind lengths, and the expansion of the ∠OHO and ∠COP angles.

Figure 3.4. Pseudo-chair conformation of dianionic carboxyphosphate calculated at MP2/aug-cc-pVTZ in vacuum. Bond lengths are shown on the left and angles on the right

It appears that the ring structure of PC adjusts to balance the forces between an optimal hydrogen bond and ring strain relief. The hydrogen bond is compressed by 23 degrees from normal the linear arrangement of the water dimer. The dihedral angle about the hydrogen bond (∠COHO = 4.6°) is relatively flat to maintain the strength of the hydrogen bond as shown in Figure 3.5.

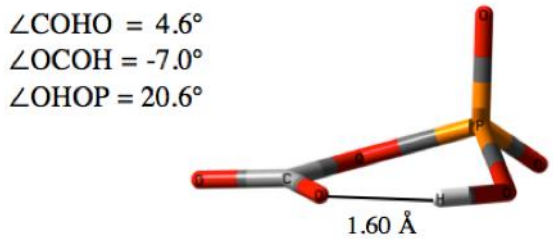


Figure 3.5. Dihedral angles for dianionic pseudo-chair carboxyphosphate calculated at MP2/aug-cc-pVTZ in vacuum.

3.3.3 Monoanionic carboxyphosphate

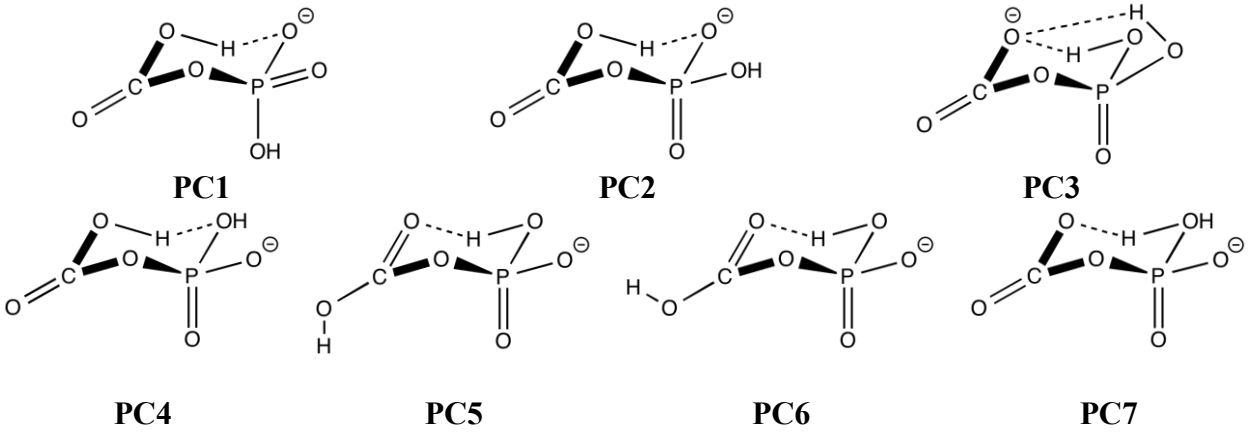


Figure 3.6. Pseudo-chair conformations of monoanionic carboxyphosphate in order of increasing energy (left to right/top to bottom)

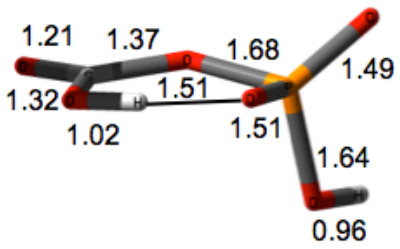
For monanion, 13 structures have been located as stationary points. More than half of these structures cyclize to form an intramolecular hydrogen bond as seen in the dianion. Therfore, these structures are referred to as PC followed by a number indicating order of stability, where PC1 is the lowest energy and PC7 the highest. As expected, the lowest energy conformation for monoanion CP is one in which both the carboxylic acid and phosphate groups

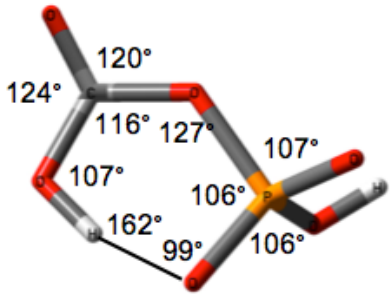
are protonated. The preference for the hydrogen bond donor (HBD) to be on the carboxyl side instead of the phosphate as seen in the dianion is due to the additional ionizable group present on the phosphate. As will be detailed in Chapter 4, the charge-assisted hydrogen bond (CAHB) is the largest contributing factor to the stability of pseudchair PC. Subsequently, the most stable conformation is the conformation in which the CAHB is strongest. When the carboxylic acid group is protonated but the phosphate acts as the HBD (*PC5*) then the hydrogen bond is no longer charge-assisted, resulting in an 8.5 kcal/mol rise in energy. The general trend across all of the pseudochair conformations confirms that as the intramolecular hydrogen bond is weakened, they become less stable.

Table 3.3. Relative energies (ΔE^{elec} (kcal/mol) of different conformations of monoanionic 1 in vacuum

Level of theory	PC_1	PC_2	PC_3	PC_4	PC_5	PC_6	PC_7	O_1
MP2/aug-cc-pVDZ	0.0	0.68	0.93	7.0	8.5	9.2	11.7	12.9
MP2/aug-cc-pVTZ	0.0	0.89	1.3	8.4	9.9	10.8	13.3	14.1
MP2/aug-cc-pVQZ	0.0	0.92	1.4	8.5	10.0	10.8	13.4	14.2
CCSD(T)/aug-cc-pVTZ*	0.0	1.0	1.7	8.6	9.6	10.5	12.8	13.6
SP energy evaluation on *MP2/aug-cc-pVQZ								

The geometry of the CAHB in PC1 and PC2 is nearly identical. However, the torsion angles are quite different.





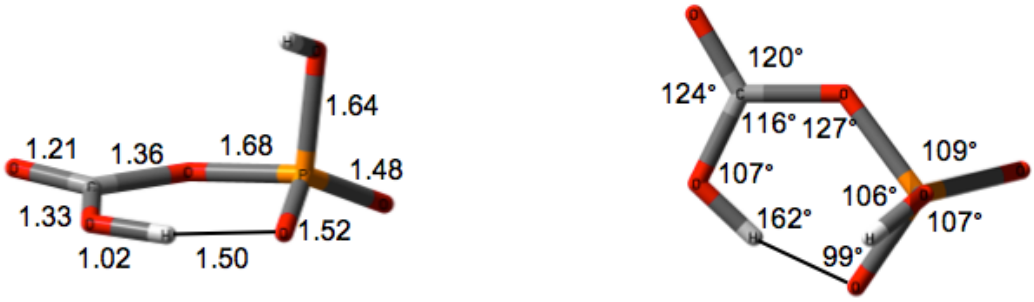


Figure 3.7. Geometric comparison of *PC1* (top) and *PC2* (bottom). Bond lengths (Å) are shown on the left and angles on the right.

It is clear that the CAHB makes the most significant contribution to the overall stability of these structures with the differing position of the OH group on the phosphate only resulting in about a 1 kcal/mol difference in energy.

Figure 3.8. Torsion angles of PC1 (left) and PC2 (right) at MP2/aug-cc-pVTZ

By contrast, although quite similar in stability, *PC3* is structurally quite different from *PC1* and *PC2*.

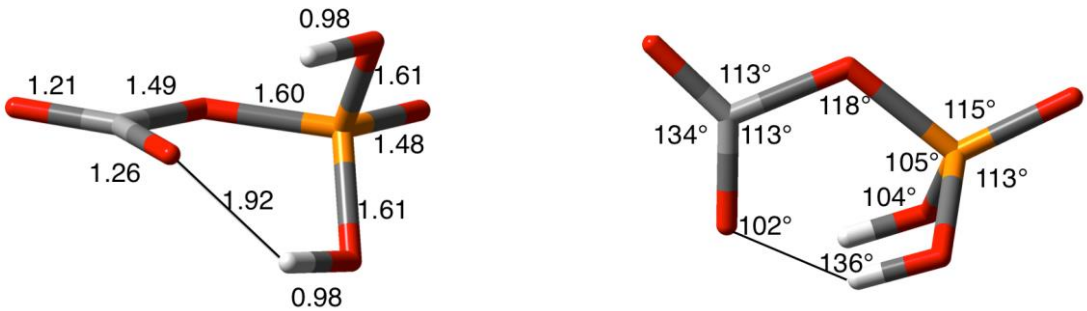


Figure 3.9. *PC_3* conformation of monoanionic carboxyphosphate calculated at MP2/aug-cc-pVTZ in vacuum. Bond lengths are shown on the left and angles on the right.

In this structure, both hydrogen atoms are located on the phosphate group forming two geometrically equivalent CAHBs with an oxygen atom of the carboxylic acid acting as the HBA. It appears that this structure is stabilized to a similar degree as *PC_1* and *PC2* with two weak, but equal CAHBs rather than a single strong CAHB.

The six higher energy conformations that do not contain an intramolecular hydrogen bond and are referred to as "open" and are designated O1 to O6 again in energetic order of lowest to highest.

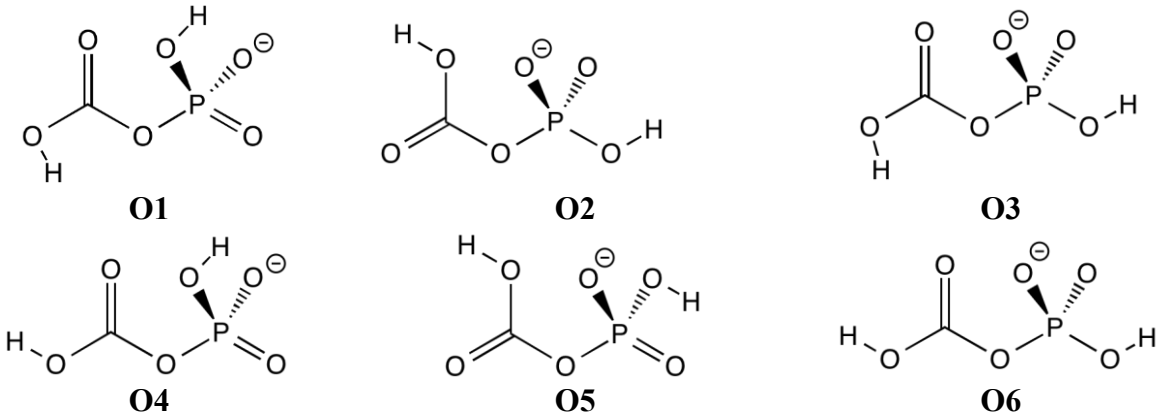


Figure 3.10. Open conformations of carboxyphosphate

3.3.4 Energetic analysis

The performance of five different density functionals combined with Dunning's augmented correlation consistent basis sets (n=D, T, Q, (5)) was evaluated relative to post-SCF MP2 and CCSD(T) calculations.

3.3.4.1 Dianion

A total of twenty-eight different levels of theory were evaluated for dianionic CP by combining the five DFT functionals: B3LYP, MPW1K, BB1K, M05-2X, and M06-2X with Dunning's aug-cc-pV*n*Z basis sets (*n* = D, T, Q, 5). The performance of these levels of theory across a broad range of molecular properties has been thoroughly tested and documented in the literature.[133,135,142,150,154,166,168,170,180,185,190,213] The performance of these five DFT functionals was compared to Moller-Plesset and coupled-cluster computations. The concentration of charge on phosphate in several conformations presents technical challenges for different approximation methods prompting the evaluation of different functionals.

As the number of basis functions approach the CBS limit computed values, such as energy, should converge to a reference value. An experimental value is typically used as this reference. In cases where there is no experimental reference available, a trusted level of theory may be used. The short half-life of CP has made it impossible to collect experimental data for this type of benchmarking. Therefore an appropriate level of theory has been determined by evaluating the convergence of computed quantities as a function of increasing basis functions relative to the MP2 CBS limit and CCSD(T)/aug-cc-pVTZ//MP2/aug-cc-pV5Z references.

Table 3.4. Relative energies (kcal/mol) of different conformations of dianionic carboxyphosphate with respect to PC conformation in vacuum.

Level of theory	**PC**	**PA**	**PG**	**EZ**	**ZE**	**ZZ**
B3LYP/aug-cc-pVDZ	0.0	11.1	11.7	N/A	26.1	27.2
B3LYP/aug-cc-pVTZ	0.0	11.6	11.6	N/A	N/A	N/A
B3LYP/aug-cc-pVQZ	0.0	11.6	11.5	N/A	N/A	N/A
B3LYP/aug-cc-pV5Z	0.0	11.5	11.5	N/A	N/A	N/A

MPW1K/aug-cc-pVDZ	0.0	12.3	12.9	28.8	29.2	31.0
MPW1K/aug-cc-pVTZ	0.0	12.9	12.8	30.3	31.0	32.7
MPW1K/aug-cc-pVQZ	0.0	12.9	12.7	30.3	31.1	32.7
MPW1K/aug-cc-pV5Z	0.0	12.9	12.7	30.4	31.1	32.8
BB1K/aug-cc-pVDZ	0.0	11.9	12.3	27.9	28.3	30.2
BB1K/aug-cc-pVTZ	0.0	12.4	12.2	29.4	30.0	31.8
BB1K/aug-cc-pVQZ	0.0	12.5	12.3	29.4	30.1	31.8
BB1K/aug-cc-pV5Z	0.0	12.5	12.3	29.5	30.2	31.9
M052X/aug-cc-PVDZ	0.0	11.9	12.3	28.6	28.7	30.6
M052X/aug-cc-PVTZ	0.0	12.7	12.5	30.1	30.6	32.3
M052X/aug-cc-PVQZ	0.0	12.8	12.6	29.9	30.5	32.2
M052X/aug-cc-PV5Z	0.0	12.8	12.6	30.0	30.5	32.2
M062X/aug-cc-pVDZ	0.0	11.6	12.0	28.3	28.4	30.3
M062X/aug-cc-pVTZ	0.0	12.2	12.0	29.6	30.3	31.9
M062X/aug-cc-pVQZ	0.0	12.3	12.1	29.6	30.3	31.9
M062X/aug-cc-pV5Z	0.0	12.2	12.0	29.6	30.3	31.9
MP2/aug-cc-pVDZ	0.0	10.9	11.5	N/A	N/A	N/A
MP2/aug-cc-pVTZ	0.0	12.2	12.1	27.6	28.6	29.7
MP2/aug-cc-pVQZ	0.0	12.4	12.2	27.8	28.9	30.0
MP2/aug-cc-pV5Z	0.0	12.4	12.2	27.9	29.0	30.1
CCSD(T) /aug-cc-pVDZ[*]	0.0	10.5	11.2	25.5	25.9	25.5
CCSD(T) /aug-cc-pVTZ[*]	0.0	11.8	11.7	28.8	29.5	31.1

*Single point energy evaluation of MP2/aug-cc-pV5Z

It should be noted from Table 3.4 that the conformations in which charge is localized to the phosphate could not be located as energy minima with the MP2/aug-cc-pVDZ level of theory. This can be attributed to the failure of the aug-cc-pVDZ basis set in the proper description of charge polarization effects of the phosphorous atom.

When extra *f*-functions for the phosphorous are added the EZ, ZZ, and ZE are able to be located as stationary points, confirming this as the source of error. The larger basis sets (n = T, Q, 5) include the necessary *f*-functions and therefore do not suffer from this issue.

The MP2 method is slower to converge than the DFT functionals, because the DFT methods are highly parameterized. The empirical nature of DFT means that the results are less dependent on the size of the basis set. As a result, all of the DFT functionals, excluding B3LYP

were able to locate the EZ, ZE, and ZZ conformations as energy minima wih the aug-cc-pVSZ basis set. B3LYP fails to characterize the EZ, ZE, and ZZ conformations, where charge is concentrated on the phosphate due to its inability to accurately model medium range correlation energies and systems with a high concentration of localized charge. As will be discussed in the following section, this inaccuracy is further evidenced by the extended length of the P-O bond in these conformations which is elongated to *ca*. 2.0 Å, much longer than a typical P-O bond of *ca*. 1.60 Å.[214] The other functionals investigated were developed to address these deficiencies. Therefore, they do not exhibit these problems in locating any of the conformations as stationary points.

The overall energetic accuracy for each DFT functional as a function of increasing basis functions was evaluated by calculating the mean absolute deviation (MAD) for each level of theory. MP2 and the five density functionals: B3LYP, BB1K, MPW1K, M05-2X, and M06-2X were combined with increasingly large basis sets ranging from 200 to 1050 basis functions, where each data point corresponds to the MAD for the energetics across the six conformations vs. the aug-cc-pV*n*Z (*n* = D, T, Q, 5) basis sets as shown in Figure 3.11. Extrapolation to the CBS limit was performed using the following Schwartz-type scheme recommended by Martin.[210]

$$E(l) = A + \frac{B}{\left(l+\frac{1}{2}\right)^4} + \frac{C}{\left(l+\frac{1}{2}\right)^6} \qquad (3.1)$$

In this scheme, *l* is the angular momentum. The CBS limit values are given next to the final data points on the right side of the graph. The CCSD(T) values are reference points from single point calculations and are not a function of increasing basis sets. Thus they are represented as straight lines with the MAD for each single point calculation to the right.

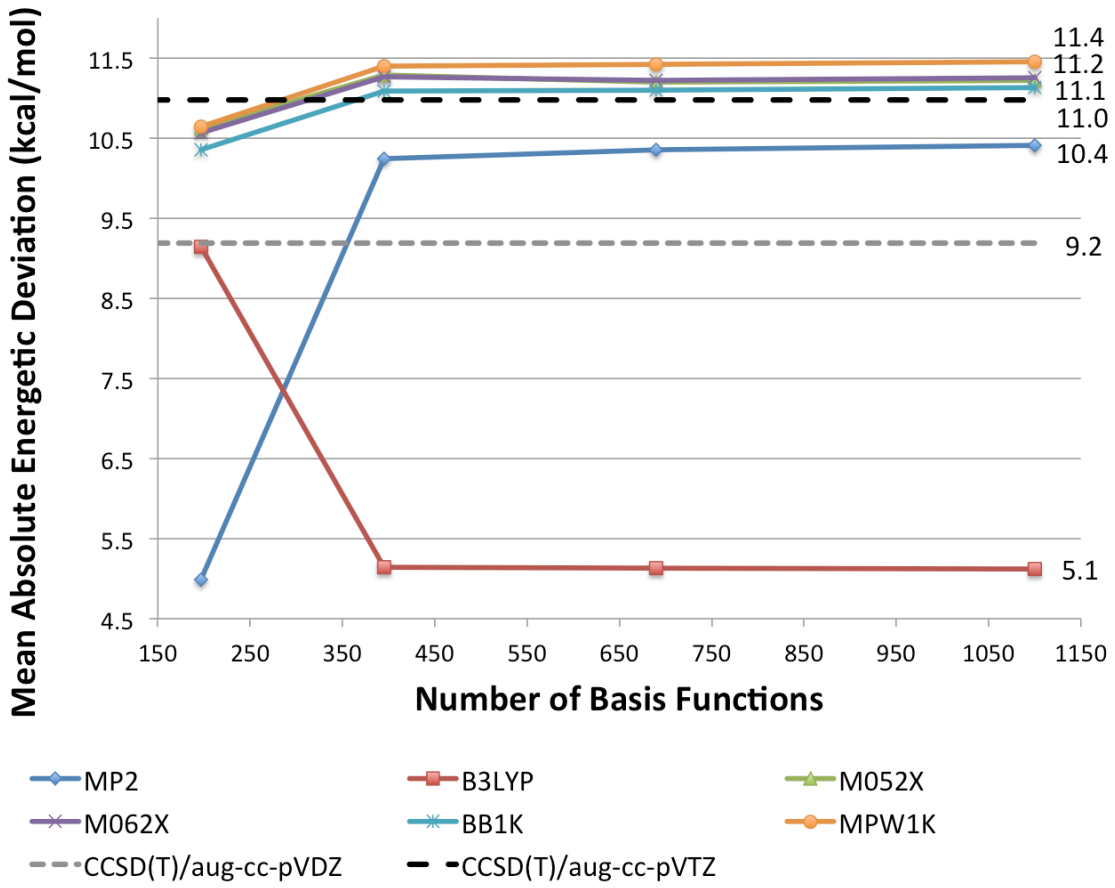


Figure 3.11. Basis set convergence of the mean average energy deviation for 6 dianionic structures in vacuum. Each data point is representative of Dunning's aug-cc-pVnZ basis sets where n = D, T, Q, and 5, from right to left. CCSD(T)/aug-cc-pVDZ and CCSD(T)/aug-cc-pVTZ are used as references representing single point energy calculations and are not a function of increasing basis functions.

The MAD at the MP2 CBS limit across the six dianionic structures is 10.4 kcal/mol, just below the calculated value of 11.0 for the CCSD(T)/aug-pVTZ//MP2/aug-cc-pV5Z single point calculation. It should be noted that the value for CCSD(T)/aug-cc-pVDZ is over a kcal/mol below both the MP2 CBS limit and CCSD(T)/aug-cc-pVTZ values. CCSD(T) suffers from the same issue as MP2 when it comes to basis set size making it slow to converge. Therefore, it is critical to use a basis set no smaller than the aug-cc-pVTZ with these approximation methods.

Although CCSD(T)/CBS calculations have been shown to provide accurate interaction energies for small noncovalent complexes the large computational cost makes them unrealistic, particularly for the more advanced calculations required to investigate the mechanism of ATP-grasp enzymes. Therefore, it is of interest to identify a DFT method that gives similar performance to these more advanced methods. The DFT calculations in Figure 3.9, with the exception of B3LYP, are all in very good agreement with CCSD(T)/aug-cc-pVTZ computations. Specifically, the error from each CBS limit compared to CCSD(T)/aug-cc-pVTZ is 0.4, 0.2, 0.2, 0.1 kcal/mol for the MPW1K, M05-2X, M06-2X, and BB1K functionals, respectively. As previously discussed, the B3LYP is problematic for some conformations and gives rise to a much larger error of 5.9 kcal/mol.

The performance of each functional was also evaluated using the PCM solvent model to represent water. It is important to evaluate these functionals in different environments that may be important in accurately evaluating the mechanism of ATP-grasp enzymes. While the enzyme pocket is largely hydrophobic, there are several charged or polar residues that have been implicated in the mechanism. Implicit water may give a rough representations of the stabilization provided by these residues. The same six structures found in vacuum were identified as stationary points when the implicit solvation was employed. The relative energies of these structures are given in Table 3.5. The overall energetic accuracy of each level of theory was evaluated in implicit water in the same manner as for vacuum. Calculations with Dunning's aug-cc-pV5Z basis set was excluded due to high computational demand and data from vacuum computations which indicated that little to no accuracy is gained by increasing from aug-cc-pVQZ to aug-cc-pV5Z.

Table 3.5. Relative energies (kcal/mol) of different conformations of dianionic carboxyphosphate with respect to PC conformation in PCM water.

Level of theory	**PC**	**PA**	**PG**	**EZ**	**ZE**	**ZZ**
B3LYP/aug-cc-pVDZ	0.0	5.0	6.1	12.8	12.6	13.9
B3LYP/aug-cc-pVTZ	0.0	5.9	6.5	14.0	13.8	15.3
B3LYP/aug-cc-pVQZ	0.0	5.9	6.5	14.0	13.8	15.3
MPW1K/aug-cc-pVDZ	0.0	5.0	6.0	12.7	12.6	13.6
MPW1K/aug-cc-pVTZ	0.0	5.2	5.8	13.2	13.2	14.4
MPW1K/aug-cc-pVQZ	0.0	5.2	5.7	13.2	13.1	14.4
BB1K/aug-cc-pVDZ	0.0	5.1	6.3	14.2	14.0	15.1
BB1K/aug-cc-pVTZ	0.0	5.6	6.2	14.7	14.6	15.8
BB1K/aug-cc-pVQZ	0.0	5.7	6.4	14.5	14.4	15.7
M052X/aug-cc-PVDZ	0.0	5.0	6.0	14.0	13.8	15.0
M052X/aug-cc-PVTZ	0.0	5.2	5.8	14.4	14.3	15.6
M052X/aug-cc-PVQZ	0.0	5.4	5.9	14.2	14.1	15.5
M062X/aug-cc-pVDZ	0.0	5.4	6.4	14.1	13.9	14.9
M062X/aug-cc-pVTZ	0.0	5.6	6.2	14.3	14.2	15.5
M062X/aug-cc-pVQZ	0.0	5.7	6.2	14.3	14.2	15.5
MP2/aug-cc-pVDZ	0.0	5.6	6.7	14.5	14.3	15.4
MP2/aug-cc-pVTZ	0.0	5.9	6.5	14.8	14.7	16.0
MP2/aug-cc-pVQZ	0.0	5.9	6.5	14.8	14.6	16.0
*CCSD(T) /aug-cc-pVDZ**	0.0	4.3	5.5	12.6	12.4	13.7
*CCSD(T) /aug-cc-pVTZ**	0.0	5.1	5.8	13.7	13.5	15.0

*Single point energy evaluation of MP2/aug-cc-pVQZ

The MADs for energetics of the dianion follow the same general trends as seen in the vacuum computations with a few exceptions. Overall, the MADs are about 2 kcal/mol lower in implicit solvation. Second, the issues that some of the B3LYP and MP2 computations suffered from are resolved with implicit solvation. This is expected since the solvent used is water, which is capable of stabilizing and mitigating the localized charge that can be problematic for these functionals.

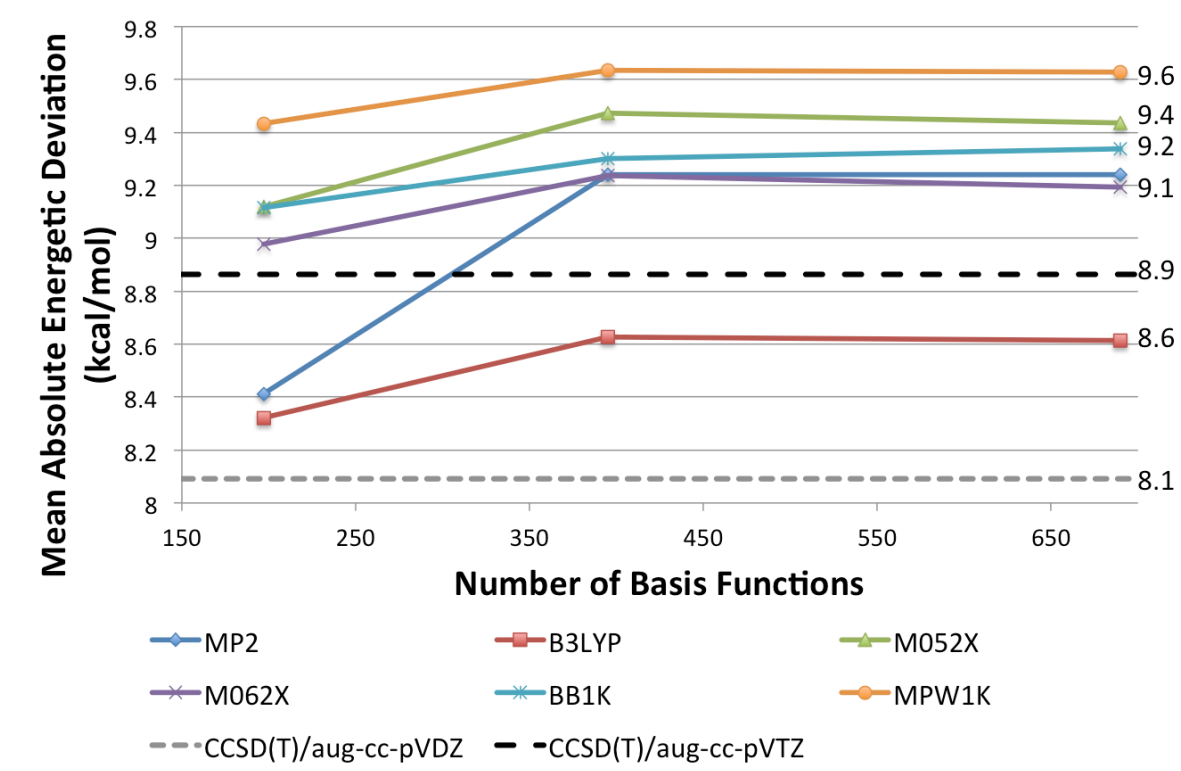


Figure 3.12. Basis set convergence of the mean average energy deviation for 6 dianionic structures in PCM water. Each data point is representative of Dunning's aug-cc-pV*n*Z basis sets where *n* = D, T, and Q, from right to left. CBS limit values are given to the right of each curve. CCSD(T) calculations are single point references and not a function of increasing basis functions.

Once again, all functionals give results in very good agreement with CCSD(T)/aug-cc-pVTZ single point calculations. Specifically, the error from each CBS limit compared to CCSD(T)/aug-cc-pVTZ is 0.7, 0.5, 0.3, 0.2 and 0.3 kcal/mol for the MPW1K, M05-2X, BB1K, M06-2X and B3LYP functionals, respectively.

3.3.4.2 Monoanion

MAD results for monanionic are generally consistent with dianionic computations. The most significant, but expected, change is the performance of B3LYP. For the dianionic structures where charge is concentrated on the phosphate, B3LYP fails to find stationary points and is in poor agreement with CCSD(T)/aug-cc-pVTZ calculations with a MAD value at the CBS limit nearly 6 kcal/mol lower. In contrast, B3LYP results are only 0.2 kcal/mol higher than the CCSD(T)/aug-cc-pVTZ value for the monoanion.

Table 3.6. Relative energies (kcal/mol) of different pseudo-chair structures of monoanionic carboxyphosphate with respect to PC1 in vacuum

Level of theory	**PC_1**	**PC_2**	**PC_3**	**PC_4**	**PC_5**	**PC_6**	**PC_7**
B3LYP/aug-cc-pVDZ	0.0	0.82	2.36	7.78	9.20	9.93	12.26
B3LYP/aug-cc-pVTZ	0.0	0.96	2.36	8.42	9.85	10.72	13.00
B3LYP/aug-cc-pVQZ	0.00	0.99	2.45	8.42	9.87	10.75	13.02
BB1K/aug-cc-pVDZ	0.00	0.85	0.70	8.08	9.76	10.41	12.57
BB1K/aug-cc-pVTZ	0.00	1.01	1.12	8.76	10.33	11.16	13.33
BB1K/aug-cc-pVQZ	0.00	1.03	1.21	8.76	10.34	11.20	13.42
M052X/aug-cc-PVDZ	0.00	0.71	1.06	8.49	9.94	10.56	12.77
M052X/aug-cc-PVTZ	0.00	0.94	1.20	9.45	10.87	11.69	13.96
M052X/aug-cc-PVQZ	0.00	0.96	1.45	9.40	10.89	11.71	14.03
M062X/aug-cc-pVDZ	0.00	0.72	0.79	7.94	9.38	10.06	12.14
M062X/aug-cc-pVTZ	0.00	0.92	0.86	8.74	10.09	10.98	13.11
M062X/aug-cc-pVQZ	0.00	0.93	1.00	8.70	10.12	11.02	13.20
MPW1K/aug-cc-pVDZ	0.00	0.87	1.90	8.73	10.52	11.17	13.60
MPW1K/aug-cc-pVTZ	0.00	1.03	2.43	9.49	11.24	12.05	14.40
MPW1K/aug-cc-pVQZ	0.00	1.06	2.56	9.46	11.20	12.04	14.43
MP2/aug-cc-pVDZ	0.0	0.68	0.93	6.96	8.49	9.15	11.69
MP2/aug-cc-pVTZ	0.00	0.89	1.30	8.39	9.94	10.77	13.28
MP2/aug-cc-pVQZ	0.00	0.92	1.44	8.50	9.97	10.83	13.35
CCSD(T)/aug-cc-pVDZ*	0.00	0.77	1.10	5.52	8.09	8.75	11.14
CCSD(T)/aug-cc-pVTZ*	0.00	1.01	1.71	8.59	9.63	10.49	12.78

*Single point energy calculations performed on MP2/aug-cc-pVQZ optimized geometries

The difference between the aug-cc-pVDZ and aug-cc-pVTZ basis sets is also decreased in the monanion calculations compared to the dianion. However, this is most likely due to the increase

in sample size from 6 to 13 different structures for monoanion CP leading to an reduction in overall deviation.

Table 3.7. Relative energies (kcal/mol) of different "open" structures of monoanionic carboxyphosphate with respect to PC1 in vacuum

Level of theory	**O_1**	**O_2**	**O_3**	**O_4**	**O_5**	**O_6**
B3LYP/aug-cc-pVDZ	13.47	12.83	13.35	13.95	13.71	13.30
B3LYP/aug-cc-pVTZ	13.79	13.72	14.02	14.42	14.36	14.08
B3LYP/aug-cc-pVQZ	13.75	13.76	14.06	14.41	14.37	14.11
BB1K/aug-cc-pVDZ	14.45	13.62	14.61	14.79	14.25	14.43
BB1K/aug-cc-pVTZ	14.71	14.49	15.20	15.26	14.92	15.18
BB1K/aug-cc-pVQZ	14.72	14.56	15.24	15.31	15.00	15.23
M052X/aug-cc-PVDZ	14.65	13.75	14.75	14.94	14.45	14.52
M052X/aug-cc-PVTZ	15.43	15.09	15.88	15.96	15.59	15.81
M052X/aug-cc-PVQZ	15.40	15.15	15.91	15.96	15.63	15.83
M062X/aug-cc-pVDZ	14.00	13.35	14.18	14.34	13.94	14.04
M062X/aug-cc-pVTZ	14.51	14.40	14.98	15.11	14.85	15.01
M062X/aug-cc-pVQZ	14.52	14.50	15.04	15.15	14.93	15.07
MPW1K/aug-cc-pVDZ	15.11	14.22	15.15	15.44	15.04	14.92
MPW1K/aug-cc-pVTZ	15.42	15.13	15.85	15.95	15.70	15.75
MPW1K/aug-cc-pVQZ	15.36	15.15	15.84	15.93	15.72	15.75
MP2/aug-cc-pVDZ	12.86	12.67	12.82	13.23	13.42	12.72
MP2/aug-cc-pVTZ	14.12	14.42	14.49	14.66	14.86	14.48
MP2/aug-cc-pVQZ	14.17	14.49	14.59	14.74	14.94	14.59
CCSD(T)/aug-cc-pVDZ*	12.31	11.99	12.32	12.69	12.76	12.21
CCSD(T)/aug-cc-pVTZ*	13.61	13.78	14.07	14.19	14.20	14.07

*Single point energy calculations performed on MP2/aug-cc-pVQZ optimized geometries

The relative energies relative to PC_1 for each level of theory tested are given in Table 3.6 for all pseudochair structures and in Table 3.7 for all "open" structures. Dianionic results indicated that little to no additional accuracy is achieved with the aug-cc-pV5Z basis set versus the aug-cc-pVQZ basis set. Thus, the significant additional computational cost required to

complete calculations with this basis set were not found to be justifiable for evaluation of monanionic CP.

Energetic MADs for the monanion are less than half those found for the dianion. This is expected, since there are twice as many conformations with much smaller deviations in energy between them. Once again the B3LYP and MP2 functionals do not exhibit the any of the issues seen with the charge localized dianionic conformations. The addition of a proton to these conformations eliminates the localization of charge that occurs with the dianion.

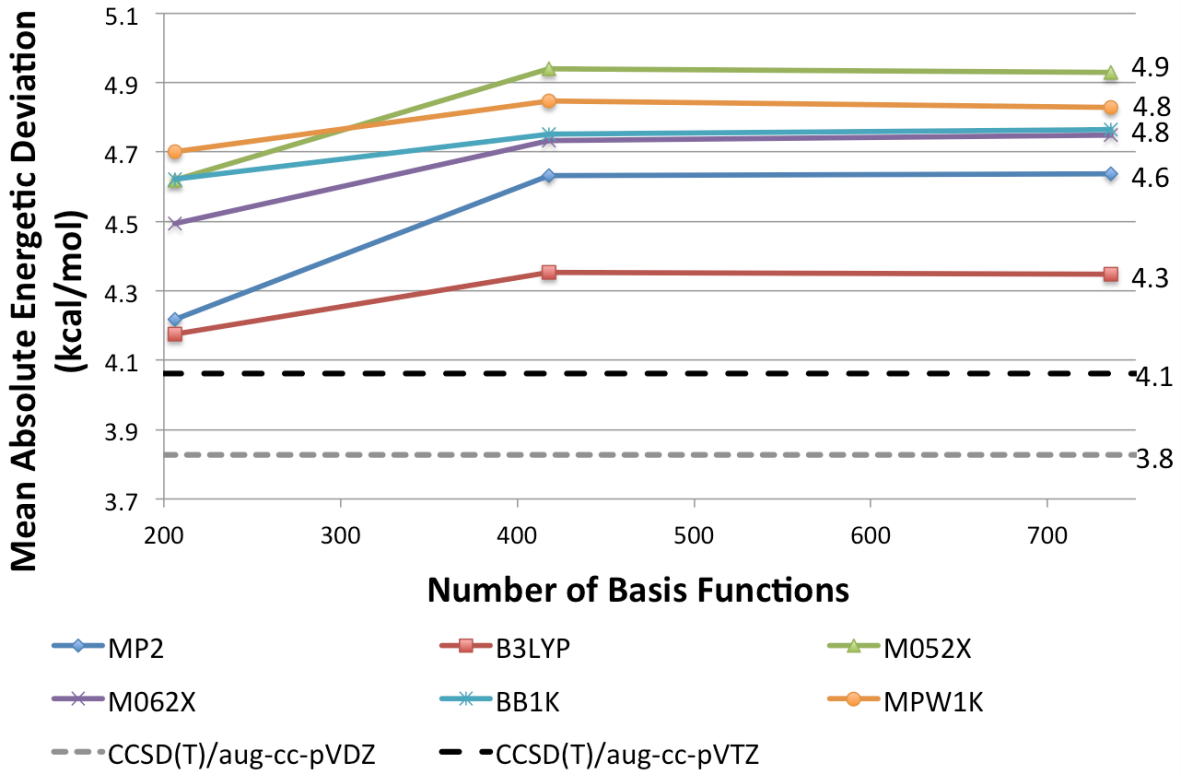


Figure 3.13. Basis set convergence of the mean average energy deviation for 13 monoanionic structures in vacuum. Each data point is representative of Dunning's aug-cc-pV*n*Z basis sets where *n* = D, T, and Q, from right to left. CBS values are given to the right of each curve. CCSD(T) calculations are single point references and not a function of increasing basis functions.

As in all the computations seen thus far, all functionals are in good agreement with the CCSD(T)/aug-cc-pVTZ reference point. The functionals exhibit errors of 0.8, 0.7, 0.7, 0.7, and

0.2 kcal/mol for M05-2X, MPW1K, BB1K, M06-2X, and B3LYP respectively. The MP2/CBS limit is 0.5 kcal/mol above the CCSD(T)/aug-cc-pVTZ reference point giving results in the same range as the DFT functionals.

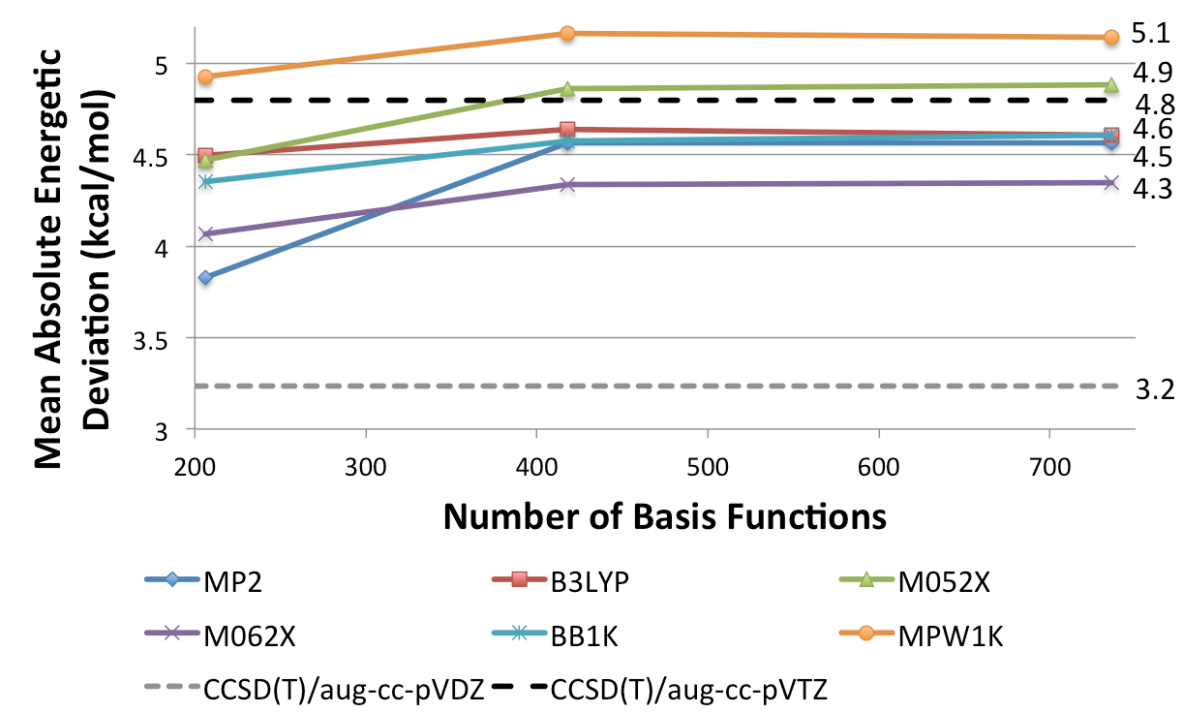


Figure 3.14. Basis set convergence of the mean average energy deviation for 13 monanionic structures in PCM water. Each data point is representative of Dunning's aug-cc-pV*n*Z basis sets where *n* = D, T, and Q, from right to left. CCSD(T) calculations are single point references and not a function of increasing basis functions.

Computations for the monanion with the PCM implicit solvent representation of water yield results consistent with those seen in vacuum and for the dianion. These calculations exhibit the smallest deviations seen so far relative to the CCSD(T)/aug-cc-pVTZ reference point. However, there is a larger difference between the CCSD(T)/aug-cc-pVDZ and CCSD(T)/aug-cc-pVTZ of 1.6 kcal/mol. As previously mentioned, it is very important to use sufficiently large basis sets with both the couple cluster and MP2 methods. It is clear from all of our calculations that the aug-cc-pVDZ basis set is not sufficiently converged and will not produce accurate results

particularly when paired with these higher-level methods that do not contain empirical parameters.

3.3.5 Geometric analysis

Mean average deviation was calculated for each functional as a function of increasing basis functions for each geometric value relative to MP2/aug-cc-pV5Z and MP2/aug-cc-pVQZ computations for di and monoanion, respectively. Consistent with the energetic analysis Dunning's aug-cc-pVTZ of the CBS limit (*ca*. 99%). Table 3.8 gives a summary of the geometric MAD results for dianion CP for each level of theory.

Table 3.8. Mean absolute deviation (MAD) across all bond lengths (Å) relative to MP2/aug-cc-pV5Z computations for dianionic carboxyphosphate in vacuum.

	Basis set aug-cc-pV*n*Z			
Functional	**D**	**T**	**Q**	**5**
MP2	0.030	0.010	0.003	0.000
B3LYP	0.257	0.715	0.714	0.713
M052X	0.018	0.007	0.008	0.009
M062X	0.016	0.006	0.009	0.010
BB1K	0.016	0.014	0.018	0.020
MPW1K	0.016	0.015	0.019	0.021

For bond lengths, B3LYP had the largest deviation from MP2/aug-cc-pV5Z computations with *ca*. 0.3, 0.7, 0.7, and 0.7 Å for aug-cc-pV*n*Z, where *n* = D, T, Q, and 5, respectively. The other functionals gave similar results with deviations *ca*. 0.02 Å for the smallest basis set (aug-cc-pVDZ), but as the basis set is increased to aug-cc-pVTZ the deviation for M05-2X and M06-2X is reduced to 0.007 and 0.006 Å, respectively. The deviation for the

BB1K and MPW1K remains more than double that of M05-2X and M06-2X even increasing as the CBS limit is approached.

Table 3.9. Mean absolute deviation (MAD) across all bond angles (°) relative to MP2/aug-cc-pV5Z computations for dianionic carboxyphosphate in vacuum.

	Basis set aug-cc-pV*n*Z			
Functional	**D**	**T**	**Q**	**5**
MP2	0.58	0.24	0.10	0.00
B3LYP	0.46	0.48	0.55	0.59
M052X	0.53	0.47	0.50	0.49
M062X	0.51	0.43	0.48	0.49
BB1K	0.66	0.60	0.63	0.60
MPW1K	0.51	0.63	0.68	0.70

All functionals give MADs within 1° of MP2/aug-cc-pV5Z results. In contrast to the results for bond lengths, B3LYP gives some of the lowest MADs relative to MP2/aug-cc-pV5Z, but increases rather than converging with increasing basis functions. As previously discussed, B3LYP gives abnormally large P-O bond lengths which is the reason it exhibits a large amount of error for bond lengths but not for other geometric values. The M05-2X and M06-2X functionals outperform the BB1K and MPW1K giving smaller MADs for angles just as they did for bond lengths. The smallest deviation is found for the M06-2X/aug-cc-pVTZ level of theory at 0.43°. In general, the MADs appear to converge with the aug-cc-PVTZ basis set for all levels of theory.

As expected, analysis of dihedral angles results in the largest MADs for each of the functionals tested. B3LYP once again is in good agreement with MP2/aug-cc-pV5Z calculations with the smallest basis set (aug-cc-pVDZ) at 5.33°, but its MAD rises to nearly double this value as the CBS limit is approached. The results for each functional are generally in line with those

seen for bond lengths and angles with the exception of MPW1K, which gives much better agreement with MP2/aug-cc-pV5Z results for dihedrals than any other geometric criteria evaluated thus far. Consistent with analysis of bond lengths and angles the M06-2X/aug-cc-pVTZ level of theory gives one of the smallest MADs at 5.39°.

Table 3.10. Mean absolute deviation (MAD) across all dihedral angles (°) relative to MP2/aug-cc-pV5Z computations for dianionic carboxyphosphate in vacuum.

	Basis set aug-cc-pV*n*Z			
Functional	**D**	**T**	**Q**	**5**
MP2	5.33	1.54	0.95	0.00
B3LYP	5.91	7.14	7.99	10.32
M052X	7.82	6.21	5.91	5.50
M062X	6.33	5.39	5.44	4.51
BB1K	9.45	7.25	6.75	4.31
MPW1K	5.23	5.90	5.85	6.63

MADs for geometric criteria were also evaluated for the monoanion and the results are given in Tables 3.11, 3.12, and 3.13.

Table 3.11. Mean absolute deviation (MAD) across all bond lengths (Å) relative to MP2/aug-cc-pVQZ computations for monoanionic carboxyphosphate in vacuum

	Basis set aug-cc-pV*n*Z		
Functional	**D**	**T**	**Q**
MP2	0.027	0.006	0.000
B3LYP	0.028	0.010	0.006
M052X	0.015	0.006	0.008
M062X	0.018	0.007	0.010
BB1K	0.028	0.015	0.025
MPW1K	0.013	0.014	0.017

All MADs analyzed are in fairly good agreement with MP2/aug-cc-pV5Z references for bond length at 0.015, 0.014, 0.010, 0.007, and 0.006 respectively for BB1K, MPW1K, B3LYP, M06-2X, and M05-2X, respectively with the aug-cc-pVTZ basis set. In contrast to dianion results, the MAD for B3LYP is shown to decrease as it approaches the CBS limit. The other functionals yield the best results when the aug-cc-pVTZ basis set is used. Similar to the dianion results, the M06-2X and M05-2X functionals produce MADs that are about half those given by the BB1K and MPW1K functionals for the aug-cc-pVTZ basis set.

Table 3.12. Mean absolute deviation (MAD) across all angles (°) relative to MP2/aug-cc-pVQZ computations for monoanionic carboxyphosphate in vacuum

	Basis set aug-cc-pVnZ		
Functional	**D**	**T**	**Q**
MP2	0.72	0.14	0.00
B3LYP	0.71	0.53	0.58
M052X	0.54	0.58	0.57
M062X	0.55	0.56	0.56
BB1K	0.50	0.52	0.57
MPW1K	0.47	0.55	0.62

All levels of theory give similar MADs for bond angles of *ca*. 0.5° relative to MP2/aug-cc-pVQZ. In contrast to previous results, BB1K and MPW1K give a slightly lower MAD values than M06-2X and M05-2X.

Table 3.13. Mean absolute deviation (MAD) across all dihedral angles (°) relative to MP2/aug-cc-pVQZ computations for monoanionic carboxyphosphate in vacuum

	Basis set aug-cc-pVnZ		
Functional	**D**	**T**	**Q**
MP2	1.43	0.27	0.00
B3LYP	1.55	1.80	1.78
M052X	1.04	1.07	1.15
M062X	1.40	1.24	0.99
BB1K	1.04	1.36	1.36
MPW1K	0.80	1.98	1.48

The MAD values for dihedral angles are much lower overall for the monoanion than those observed for the dianion. All MADs are found to deviate less than 2° from MP2/aug-cc-pVQZ results. The M06-2X and M05-2X functionals give the values in closest agreement with the MP2/aug-cc-pVQZ reference for the aug-cc-pVTZ basis set at 1.24° and 1.07°, respectively.

3.3.6 Jul basis sets

Over the past several decades, it has been recommended that diffuse functions should be used sparingly to minimize basis set superposition error and shorten calculation times.[204,205,210] Recently a series of "calendar" basis sets have been developed which incrementally reduce diffuse functions from Dunning's aug-cc-pV*n*Z basis sets.[205] In this implementation "aug" represents the standard basis set including diffuse functions on all atoms. The prefix "jul" removes diffuse functions from hydrogen and helium atoms. The use of diffuse functions on these small atoms is a common source of BSSE in smaller basis sets where larger atoms may lack be lacking in basis functions. For these reasons, it was deemed worthwhile to compare the effect of the "aug" vs "jul" methods for the smaller basis sets used in this study (aug-cc-pV*n*Z; *n*

= D, T). The results of this analysis confirm that the "jul" basis sets give results identical or superior to those achieved with full augmentation and require less computation time, making jul-cc-pVTZ the most logical choice to produce results consistent with CCSD(T)/aug-cc-pVTZ and MP2/CBS references.

Table 3.14. Comparison of aug- and jul- basis sets for mean absolute energetic deviation (MAD) in kcal/mol.

Functional:	M062X				CCSD(T)
Basis Set:	jul-cc-pVDZ	aug-cc-pVDZ	jul-cc-pVTZ	aug-cc-pVTZ	aug-cc-pVTZ
Basis Functions	193	197	386	395	395
MAD	10.6	10.6	11.3	11.3	11.0

3.4 Conclusions

Systematic analysis of different possible structures and conformations for mono-, di-, and trianionic CP was conducted using *ab initio* and density functional theory (DFT). This analysis revealed 13 distinct structures as minima for the monoanion, six for the dianion and a single conformation for the trianion. For both monoanion and dianion, novel structures referred to as pseudochair CP characterized by a stabilizing intramolecular hydrogen bond are located. These pseudochair conformations are significantly more stable than "open" conformations at *ca*. 12 and 14 kcal/mol for the di- and monoanion, respectively. These conformations are associated with strong, charge-assisted intramolecular hydrogen bonds and feature structural and energetic characteristics consistent with other low barrier and short-strong hydrogen bonds. For the EZ, ZE, and ZZ conformations of dianionic CP both negative charges are localized on the phosphate group. Both the MP2 and B3LYP functionals struggle to properly describe these structures. MP2

is slow to converge and cannot locate these conformations as energy minima with the aug-cc-pVDZ basis set unless additional *f*-functions are added. B3LYP struggles with these systems independent of basis set, and fails to locate these structures as minima at all levels of theory.

Mean average deviations (MADs) for both energy and geometric properties were used to evaluate the accuracy of each functional as a function of increasing basis functions. For all chemical methods tested, energetic and geometric MADs converge with the aug-cc-pVTZ basis set. The M06-2X and BB1K consistently deliver energies in close agreement with CCSD(T)/aug-cc-pVTZ reference values. However, M05-2X and M06-2X gave better agreement with MP2/aug-cc-pV*n*Z (*n* = Q, 5) calculations for geometric parameters than MPW1K and BB1K in most cases. With the exception of the issues previously discussed regarding B3LYP, none of the functionals tested exhibited significant deviation from MP2/CBS limit and CCSD(T)/aug-cc-pVTZ reference points. Overall, the M06-2X/aug-cc-pVTZ was observed to give the most consistent agreement with high-level reference calculations across both energetic and geometric parameters. It was also observed that the jul-cc-pVTZ basis set gave identical or superior results to aug-cc-pVTZ. Therefore, M06-2X/jul-cc-pVTZ is recommended as the most appropriate method to use for more complex calculations where CCSD(T) and MP2 methods are not computationally feasible.

This study has also established several different stable conformations across multiple charge states that carboxyphosphate may adopt through the course of its lifecycle, including the finding that it can gain a great deal of stabilization from the formation of a CAHB to yield a pseudochair structure. Establishing an appropriate level of theory that can produce accurate results for more advanced calculations is a critical first step in the study of the mechanism for the six ATP-grasp enzymes that have been discussed previously. Additionally, the identification of

stable structures provides a starting point to investigate possible mechanisms by exploring pathways through which these minima are connected. This work serves to lay a strong foundation for future computational investigation of the mechanism shared by ATP-grasp enzymes.

Chapter 4

4 Intramolecular Charge-Assisted Hydrogen Bond Strength in Pseudo-Chair Carboxyphosphate

Structural analysis of carboxyphosphate has revealed that both mono- and dianionic carboxyphosphate adopt a novel pseudo-chair conformation featuring an intramolecular charge-assisted hydrogen bond (CAHB). These conformations are consistently found to be significantly more stable in vacuum than those lacking this structural feature suggesting it is the largest contributor to the stabilization of these conformations. For this reason, it is of interest to examine the strength of the CAHB for mono- and dianionic carboxyphosphate.

4.1 Methods for Estimating CAHB Strength

The determination of hydrogen bond strength for intermolecular hydrogen bonds is straightforward, where donor and acceptor can be infinitely separated to provide a convenient reference. In contrast, determining intramolecular hydrogen bond (IMHB) strength can be an arduous task.[213,215–221] The added difficulty lies in the fact that the donor and acceptor cannot be easily separated without altering the general structure of the molecule and energetic features unrelated to the IMHB. Therefore choosing a reference structure becomes a significant challenge.

Although controversial, the most popular estimation scheme for IMHB strength is referred to as the open-closed method.[222] In this method, the IMHB strength is estimated as the difference in energy between the geometry optimized open and closed conformations, where the closed structure contains the hydrogen bond and it is disengaged in the open. The drawback of

this method is that a full geometry optimization of the open conformation has the potential to introduce or remove energetic contributions important in defining an accurate reference state.[223–228]

Therefore, there has been some effort into utilizing partially constrained optimizations as an alternative to full optimizations for the open reference system.[227,229] Another approach to address the difficulties with the standard open-close method is to assess and estimate the energetic terms crucial to the IMHB configuration and reinsert them into the open-closed energy. The process of correcting the open and closed energy differences by isolating and inserting energetic contributions is referred to here as the *additive* scheme.

Jablonski proposed that a partial optimization scheme chosen based on the system of interest would improve the open reference structure and the predicted accuracy of the IMHB.[230] The open reference structure is generated by breaching the hydrogen bond. This is done by altering the position of either the hydrogen bond donor, or acceptor only, without introducing any other changes to the closed system. An open reference generated by this method will be higher in energy and not a stationary state. Partial geometry optimizations on selected degrees of freedom are employed to lower the energy without significantly altering the open reference structure. Since the energy is lowered with increasing relaxation, this procedure is referred to as the *subtractive* scheme. Systematic partial relaxation of the initial conformation with the disengaged hydrogen bond produces a series of partially optimized open reference geometries and energies. This is done by removing geometric constraints one at a time. One example of these partial relaxations is to hold all dihedral angles except for that involved in disrupting the hydrogen bond are held fixed to prevent the introduction of new attractive or repulsive forces to the open reference form. This strategy is expected to improve estimation of the IMHB.[230]

However, the choice of which geometric features to hold fixed to produce the most accurate reference structure can be just as arbitrary as selecting a fully optimized reference structure.

Several other procedures have been developed and implemented outside of the open-closed method[222] to estimate IMHB strength. In particular, isodesmic reactions have been used to a large extent.[231] More recently, rotation barriers and the relationship between rotamers have been exploited.[217,227] Additionally, it has been shown that there is a direct correlation between ^{1}H NMR shifts and hydrogen bond strength.[219,221] NBO analysis has been used to investigate the orbital interactions involved in IMHB.[220,232] Finally, topological analysis with QTAIM has been used to evaluate changes in electron density as IMHBs are formed and disrupted.[176,220,222,233]

4.2 Computational Methods

All electronic structure calculations were performed with the Gaussian 09[116] program utilizing Truhlar's Minnesota M06-2X functional[169] and Dunning's augmented correlation consistent polarized valence triple-zeta basis set (aug-cc-pVTZ).[202] As previously discussed, this level of theory has been shown to capture the majority of structural and energetic convergence giving most consistent agreement with reference MP2 and CCSD(T) calculations. Frequency calculations were carried out to confirm all stationary points as minima on the potential energy surface. Atom types with epsilon and r_{min} values from the CHARMM force field[234,235] were used to compute the r^{-12} Pauli repulsion. All energy differences are reported as closed-open, $\Delta E_{C-O} = E_{closed} - E_{open}$, to give negative values indicating the stabilization energy provided to the closed structure by the IMHB.

The subtractive scheme follows Jablonski's notation and consists of a series of partial optimizations to locate an accurate open reference structure.[230] All model systems were modified

from the geometry optimized pseudochair structure by changing a single degree of freedom to maintain the original geometric features of the structure while disengaging the hydrogen bond as defined in Figure 4.1.

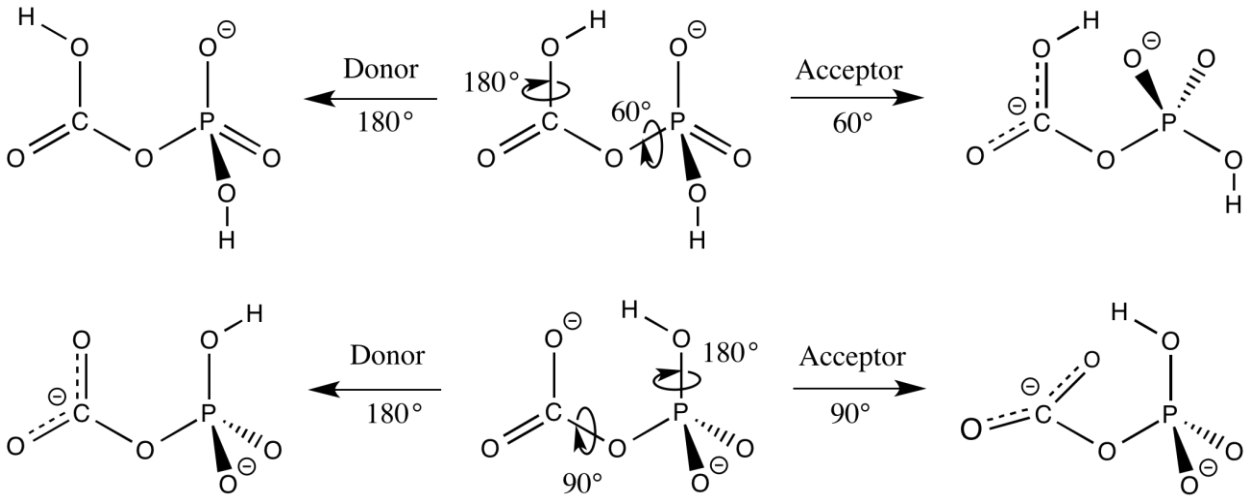


Figure 4.1. Possible conformational modification of pseudochair carboxyphosphate (center) to break the CAHB in the monoanion (top) and dianion (bottom).

To generate open conformations, the dihedral of the bond donor is rotated an idealized 180° to disrupt the hydrogen bond, while all other coordinates are held fixed. Alternatively, for hydrogen bond acceptors, the carboxylic acid is rotated 90° (dianion) or the phosphoryl group by 60° (monoanion). It is of note that this method of generating open structures with a single conformational change centered on either donor or acceptor leads to a different structure than the previously found lowest energy open conformation. However, given that these are two distinct approaches it is not necessary that the two reference structures be identical. The structures produced from this approach which have not been geometry optimized are referred to as SP. Those structures in which only all coordinates of the SP structure are held fixed except the bond lengths are referred to as B to indicate that bonds have been energy minimized. Structures in

which both bond lengths and angles are allowed to optimize are referred to as AB. Finally the structures that have been fully optimized are labeled as OPT. However, it should be noted that the fully optimized lowest energy open conformation is used as a common reference throughout this study.

4.3 Open-Closed Method

As previously stated, the open-closed method is one of the most widely used methods for estimating CAHB strength, where the energy of the hydrogen-bonded (closed) form is referenced to that of an open structure where the CAHB has been disengaged.[215,222,230,236–240] (35-39) To review, this method usually consists of rotation about one or more bonds of a system CAHB containing system to disrupt the hydrogen bond and produce an open reference structure. While there are several variations, the most primitive approach involves direct energetic comparison of a both fully geometry optimized open and closed reference structures. Structural relaxation of the open reference form may either add or remove or remove important interactions inherent to the closed structure containing CAHB leading to an inaccurate estimate if the CAHB strength. This is an issue that has been widely recognized and criticized.[223,225,226,241–243]

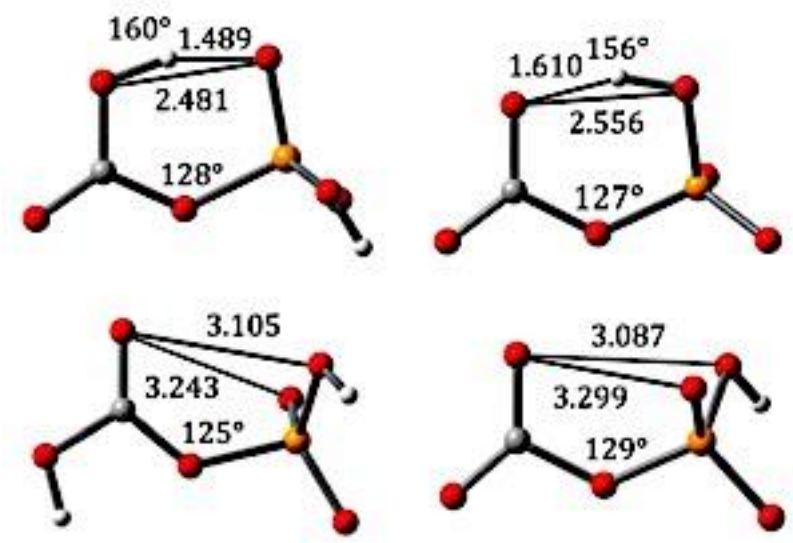


Figure 4.2. M06-2X/aug-cc-pVTZ open (bottom) and closed (top) monanionic (left) and dianionic (right).

To begin, we performed the basic open-closed method. The open reference system was chosen to be the geometry optimized lowest energy conformation of carboxyphosphate that does not contain an IMHB. The M06-2X/aug-cc-pVTZ geometry optimized structures utilized are shown in Figure 4.2. Stabilization of the pseudochair (closed) is defined to be the energy difference between closed and open conformations in both monanionic and dianionic states. CAHB strength as defined by the standard open-closed method is -14.5 kcal/mol for the monoanion and -12.0 kcal/mol for the dianion.

The open structures generated by the methods described in Figure 4.1. are slightly different from those shown in Figure 4.2. This variation occurs due to the differences in how each technique locates the open system. Thus, a more practical comparison is undertaken to represent the most prominent approaches described in the literature. This variation in the lowest

energy geometry optimized open structure confirms that essential geometric elements are often altered to varying degrees in the standard implementation of the open-closed method.

4.4 Additive Scheme

Structural comparison of the geometry minimized pseudochair (closed) conformations relative to the open structures reveals that beyond the CAHB, Pauli repulsion interactions and ring strain are the most likely contributors to the energetic differences between the two forms shown in Figure 4.2. To estimate the effect each of these contributions has on the overall stability of the pseudochair conformations, a series of model systems were employed.

4.4.1 Ring Strain

To estimate the energetic contribution of differential ring strain ($\Delta E^{rs}_{C-O} = E^{rs}_{closed} - E^{rs}_{open}$) between the closed and open forms of carboxyphosphate, dihydrogen phosphate is used as a model system. Replacement of the carboxylic acid with hydrogen produces dihydrogen phosphate and the angle containing the bridging oxygen of carboxyphosphate ∠COP becomes ∠HOP and is used to model ring strain.

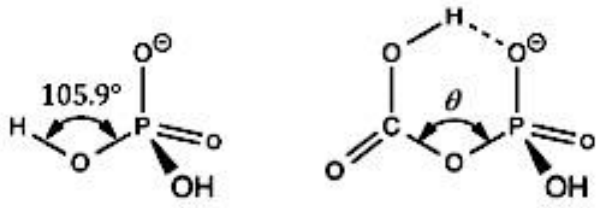


Figure 4.3. Angle representing ring strain in the dihydrogen phosphate model carboxyphosphate.

The geometry optimized ∠HOP angle of dihydrogen phosphate is 105.9° (M06-2X/aug-cc-pVTZ). This represents the strain free reference as shown in Figure 4.3.

The ∠HOP angle is then fixed to that of the bridging oxygen in the open and closed conformations of both mono- and dianionic carboxyphopshate to mimic the ring strain for each structure. The rest of the structure was then allowed to optimize. The relative energies of these partially constrained models are given in Table 4.1.

Table 4.1. Bridging oxygen angles and contributions of ring strain (kcal/mol) in the closed and open states for mono and dianionic carboxyphoshate modeled by H_2PO_4

	Open	Closed	ΔE^{rs}_{C-O}
Monanion	125.3°	127.7°	-0.8 kcal/mol
Dianion	129.3°	127.3°	0.7 kcal/mol

Although the ca. 20° expansion across the bridging oxygen relative to the standard chair form of cyclohexane is computed to give rise to *ca*. 4.0 kcal/mol, closer inspection reveals that this expanded angle is present in both open and closed forms. Thus, the computed energy difference between the open and closed forms is found to be less than 1 kcal/mol due to the small difference in angles between the open closed systems. Consequently, we found that ring strain plays only a minor role in the accurate prediction of the CAHB from the open-closed method.

4.4.2 Pauli Repulsion

Comparison of open and closed conformations optimized with M06-2X/aug-cc-pVTZ reveals that the distance between the oxygen atoms *(r_{OO})* and between the hydrogen bond donor and acceptor (r_{HO}) are substantially shortened as evident in Figure 4.4.

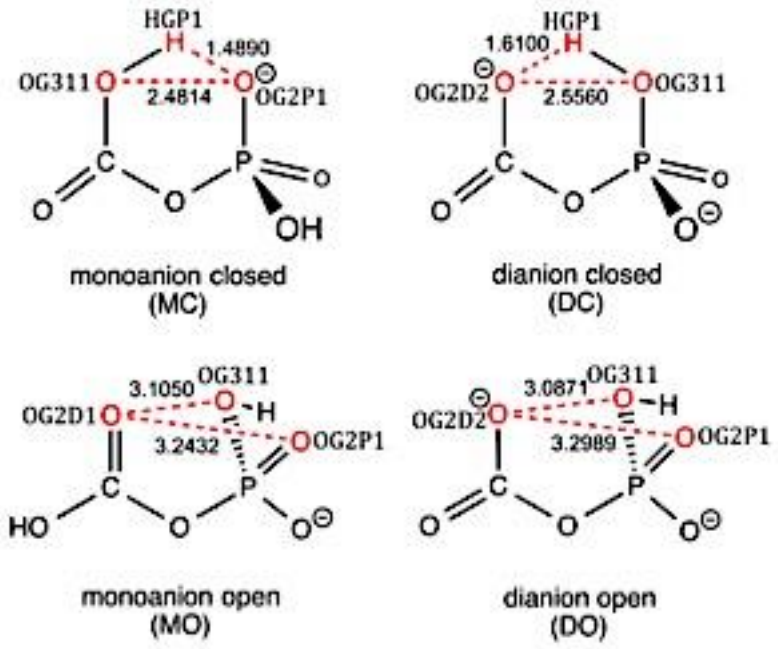


Figure 4.4. Specific nonbond interactions (red) computed to determine the Pauli repulsion in carboxyphosphate. CHARMM atom types are included.

In the geometry optimized closed conformations of both charge states, the r_{OO} distances of ca. 2.5 Å is less than the van der Waals sum of 3.04 Å $\left(r_O^{vdW} = 1.52 \text{ Å}\right)$. Additionally, the hydrogen bond r_{OH} distances are 1.49 and 1.61 Å for mono- and dianion, respectively which are shorter than the idealized van der Waals value of 1.09 Å, as given in Table 4.2.

In the open conformations, the hydrogen bond is absent and the distances between atoms expand to relieve strain caused by unfavorable intramolecular nonbonding interactions. This is evident by the r_{OO} distances that are found to be greater than the van der Waals sum in contrast to the closed conformations.

Table 4.2. r_{OO} and r_{HO} distances (Å) in the open and closed forms of the monoanion and the dianion and the difference Δr_{OO}

	open		closed		
	r_{OO}^{A}	r_{OO}^{B}	r_{OO}	r_{OH}	Δr_{OO}
monoanion	3.105	3.243	2.481	1.489	0.624
dianion	3.087	3.299	2.556	1.610	0.531

The difference in Pauli repulsion between closed and open conformations is estimated as the sum of the repulsive r^{-12} part of the Lennard-Jones approximation,[235] as given in eq 4.1.

$$E_{Pauli} = \epsilon_{ij}\left[\frac{r_{ij}^{min}}{r_{ij}}\right]^{12} \qquad (4.1)$$

Lennard-Jones atomic parameters ϵ and r^{min} are combined with appropriate mixing rules for a pair of atoms i and j. In the CHARMM[234] force field these parameters are defined as follows and listed in table 4.3.

$$\epsilon_{ij} = (\epsilon_i \epsilon_j)^{\frac{1}{2}} \text{ (geometric mean)} \quad \text{and} \quad r_{ij}^{min} = \frac{r_i^{min} + r_j^{min}}{2} \text{ (arithmetic mean)}$$

Table 4.3. CHARMM force field Lennard-Jones Parameters (ϵ, r^{min}).[234]

atom type	description	epsilon	$\frac{r_{min}}{2}$
OG311	hydroxyl oxygen	-0.1921	1.7650
OG2P1	=O in phosphate	-0.1200	1.7000
OG2D1	carbonyl oxygen	-0.1200	1.7000
OG2D2	carboxylate oxygen	-0.1200	1.7000
HGP1	polar hydrogen	-0.0460	0.2245

The CHARMM atom types for the parameters used to calculate Pauli repulsion according to Equation 4.1 are given in Table 4.3 and shown in Figure 4.4. The calculated Pauli repulsion for

closed and open forms of the dianion are 6.48 and 0.78 kcal/mol, respectively. Thus, the difference in Pauli repulsion energy for the dianion is 5.70 kcal/mol. Likewise, the computed Pauli repulsions for the closed and open forms of the monoanion are 9.96 and 0.78 kcal/mol, respectively. Therefore, the difference in Pauli energies for the monoanion is 9.18 kcal/mol.

4.4.3 Intramolecular Charge-Assisted Hydrogen Bond

Analysis of the energetic contributions of differential ring strain and Pauli repulsions of the open and closed forms allow for improvement of the standard open-closed method, as described in Equation 4.2.

$$\Delta E_{C-O} = \Delta E_{CAHB} + \Delta E_{C-O}^{Pauli} + \Delta E_{C-O}^{rs}$$

$$\Delta E_{CAHB} = E_{C-O} - \Delta E_{C-O}^{Pauli} - \Delta E_{C-O}^{rs} \qquad (4.2)$$

The CAHB strengths estimated from Equation 4.2 are listed below along with a summary of the values used in the calculation.

Table 4.4. Components of the CAHB energy (kcal/mol) of dianionic and monoanionic carboxyphosphate.

	ΔE_{C-O}^{rs}	ΔE_{C-O}^{Pauli}	ΔE_{C-O}	ΔE_{CAHB}
monoanion	0.8	9.2	-14.5	-22.9
dianion	0.7	5.7	-12.0	-18.4

This additive correction of the open-closed method for Pauli repulsion estimates the CAHB strength as -22.9 and -18.4 kcal/mol for the mono- and dianion, respectively. The addition of corrective terms accounting for Pauli repulsion significantly increases the predicted CAHB strength compared to the traditional implementation of the open-closed method.

4.5 Subtractive Scheme

A major criticism of the traditional open-closed method is that full geometry optimization of the open form may lead to an inaccurate reference structure.[223–226,228,230,231,242] To this point, Jablonski has shown that the IMHB strength calculated via the open-closed method is highly dependent on the degree of optimization.[229] As previously discussed, it is recommended that the open structure be generated by modifiying the closed geometry optimized structure with minimal structural changes required to break the IMHB to yield a more accurate estimate of CAHB strength. This is typically accomplished by rotation of the dihedral angle to reposition the hydrogen bond donor out of range of the acceptor (Figure 4.1). A single point energy evaluation (SP) is calculated for this structure, giving an upper bound to the IMHB estimation. At the other extreme, full relaxation of the SP structure leads to a geometry optimized structure (OPT). Ideally, there are two basic pathways to break the IMHB, consisting of modification of either the hydrogen bond donor or acceptor. Interestingly each pathway yields a different reference system as well as IMHB values.

Table 4.5. Computed energies (kcal/mol) for the subtractive method

	monoanion donor	monoanion acceptor	dianion donor	dianion acceptor
SP	-25.4	-18.2[a]	-20.9	-16.8[a]
B	-21.2	-16.2[a]	-18.8	-14.8[a]
AB	-17.8	-14.8[a]	-15.7	-13.5[a]
-TS	-19.5	-9.0[a]	-12.1	13.0[a]
OPT	-14.5	0.0	-12.0	0.0

[a]Inappropriate reference state that is not open.

Partial optimizations analogous to those described by Jablonski were performed for SP, B, AB, and OPT. First, the initial open structures were generated by disengaging the CAHB present in the closed strucutres via an 180° rotation of the hydrogen bond donor ($\angle HOCO$ for the monoanion and $\angle HOPO$ for the dianion). In a second method, initial open structures were generated through rotation of the hydrogen bond acceptor (60° for the $\angle OPOC$ of the monoanion and 90° for the $\angle HOPO$ of the dianion). However, it was found that only modification of the donor yields an appropriate open state. Alteration of the acceptor gives rise to a structure that connects closed states rather than generating an appropriate open reference structure. Thus, the results generated from this method are designated as inappropriate in the following table summarizing the results of these computations. Based on these findings, only the donor-modified structures are utilized in CAHB estimation. As reported in Table 4.5, the SP calculations were found to give the highest estimations the OPT gave the lowest. Removal of bond and angle constraints (B and AB) gave intermediate values relative to the SP and OPT calculations and provide the lower bounds for CAHB estimation.

This subtractive method estimates the CAHB strength to be between -17.8 and -25.4 kcal/mol for the monanion, and between -15.7 and -20.9 kcal/mol for the dianion. Theses ranges computed with the subtractive scheme are in good agreement with the values achieved through additive corrections to the open-closed method.

4.4.4 Transition structures

The transition structures for the process of disrupting the CAHB were computed and verified using M06-2X/aug-cc-pVTZ, and are shown in Table 5 under the label TS. The transition structure represents the highest point on the potential energy surface along the reaction

coordinate moving from the closed to open conformations. As such, the activation energy defined from the pseudochair structure is representative of the energy required to break the CAHB. The calculated activation energy required to disrupt the CAHB through donor rotation was found to be 19.5 kcal/mol and 9.0 kcal/mol via rotation of the acceptor for the monoanion. As previously discussed, rotation through the acceptor has been found to yield inappropriate results and the transition structure generated with this method simply leads to the transfer of the CAHB between two equivalent phosphate oxygens. Thus, the predicted CAHB strength for the monoanion is 19.5 kcal/mol. The computed activation energy for disengaging the dianion CAHB through the donor is 12.1 kcal/mol and 13.0 kcal/mol through the acceptor. As with the monoanion the transition structure connecting the pseudochair and open structure is only achieved through alteration of the donor. Thus, the predicted CAHB strength for the dianion is 12.1 kcal/mol.

4.4.5 CAHB Model Comparison

Comparison was made between the generated transition structures, AB minimized structures and geometry optimized closed structures to evaluate what key geometric features change in the CAHB models (Figure 4.5). It was found that the proton involved in the CAHB for the geometry optimized closed structures (monoanion and dianion) is positioned nearly directly between the donor and acceptor (Figure 4.2). This serves to alleviate lone-pair congestion between donor and acceptor oxygen atoms.

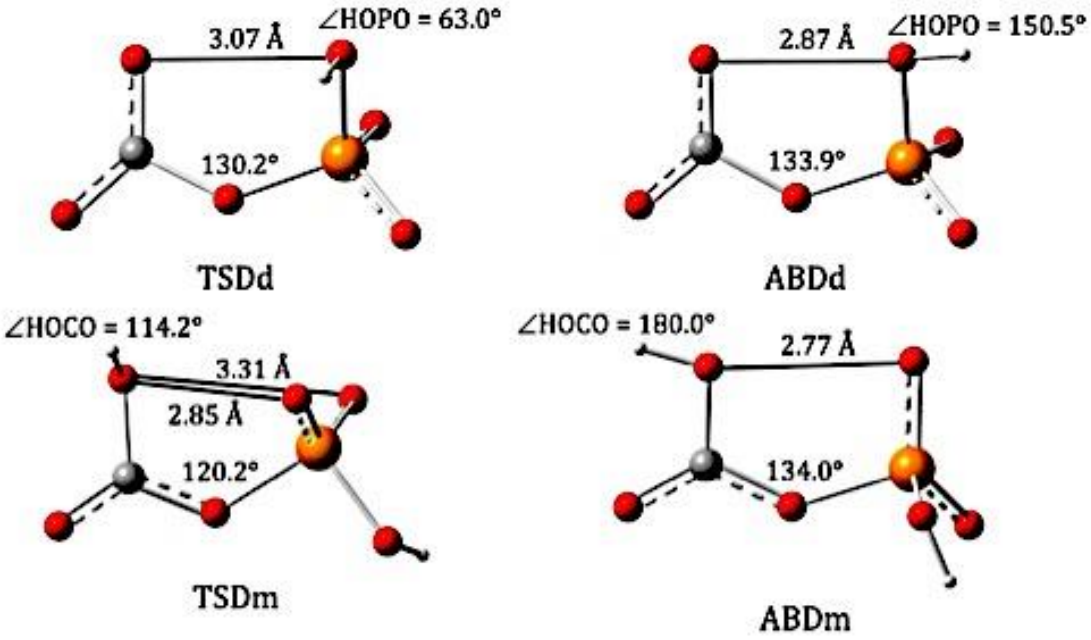


Figure 4.5. Dianionic transition structure computed using M06-2X/aug-cc-pVTZ for the breaking of the CAHB donor (TSDd) and the partially minimized dianion donor (ABDd) reference structure. Monoanionic transition structure for the breaking of the CAHB donor (TSDm) and the partially minimized monoanion donor (ABDm) reference structure.

Additionally, the closed structures exhibit two nonbond interactions that could lead to strong Pauli repulsion, between the hydrogen bond donor and acceptor oxygen atoms as well as the hydrogen atom involved in the hydrogen atom and the hydrogen bond acceptor oxygen. The dianion models (TSDd and ABDd) maintain the eclipsing hydrogen bond donor and acceptor seen in the closed structure. However, the main geometric difference in the dianion models is in the position of the hydrogen. TSDd directs the lone pairs away the lone pairs away from the acid group ($\angle HOPO = 63°$), whereas ABDd directs the lone pairs toward the acid ($\angle HOPO = 151°$). Thus, the dianion ABDd model should overestimate the contribution of repulsion to the CAHB estimation. The additional repulsion is exhibited by the lengthening of the hydrogen bond donor and acceptor oxygen atoms from 2.48 Å in the closed dianion to 3.07 Å (TSDd) and 2.87

Å (ABDd). The monoanion models (TSDm and ABDm) display different orientations between the hydrogen bond donor and acceptor. The phosphate group is rotated by *ca*. 60° in TSDm relative to the closed structure, lengthening the distance between donor and acceptor from 2.48 to 2.85 Å and yielding a second weak interaction with a second oxygen atom 3.31 Å away. The hydrogen is rotated to $\angle HOPO = 114°$, reorienting its lone pairs away from the phosphate group. However, ABDm maintains the eclipsing hydrogen bond donor acceptor interaction, since it is constrained to that of the closed structure. The distance between the hydrogen bond donor and acceptor expands to 2.77 Å, and the hydrogen is rotated to a $\angle HOPO = 180°$ increasing the congestion of lone pairs between the donor an acceptor. As a result, the contribution of repulsion is expected to be overestimated in this model just as it was in the dianion. Nevertheless, the error caused by this overestimation is somewhat mitigated when the energy difference is determined in reference to the OPT structures as the open reference structures.

In a final comparison of the activation energies with the subtractive and additive corrections to the open-closed method (Table 4.6), an agreement to the estimated CAHB strength is observed, but is greater than that found from the standard open-closed method.

Table 4.6. Comparison of three highlighted methods for CAHB estimation used in this study (kcal/mol)

	additive	subtractive	-TS	open-closed
monoanion	-24.5	-17.8 to -25.4	-19.5	-14.5
dianion	-18.4	-15.7 to -20.9	-12.1	-12.0

As expected, the CAHB strength predicted from the computed transition structures are on the low end of the additive and subtractive schemes for the monoanion and below the lower bounds of the additive and subtractive methods for the dianion. The dihedral angle constraints, or

release of other degrees of freedom, in the partial optimization procedure for the subtractive method are the likely source of error causing an overestimation, as highlighted in the discussion above. Similarly the additive corrections to the open-closed method may contain error due to the parameterization of nonbond terms, where the parameters were developed for the intermolecular interactions of phosphate esters. Despite the apparent overestimation of CAHB strength relative to those predicted by transition state structures, the importance of Pauli repulsion interactions is highlighted for the first time by the additive corrections to the open-closed method.

Based on the above estimation schemes the CAHB can be classified as strong (greater than 12 kcal/mol)[244–246] The class of hydrogen bonds known as short-strong hydrogen bonds (SSHBs) is often associated with those that are also low barrier (LBHBs).[245,247–249] Although hydrogen bonds are often both short-strong and low barrier, these classifications are based on separate criteria.[249,250] Classification of SSHBs is made based on the distance between heteroatoms (less than 2.5 Å for an OHO hydrogen bond). In contrast, classification as an LBHB is based on equivalent pK_a values of the two donor atoms, resulting in a low barrier to proton transfer between the two.

All three independent estimation schemes predict a CAHB strength greater than 12 kcal/mol consistent with SSHB energetic criteria.[244–246] The heteroatom distance is 2.48 Å in the monoanion and 2.56 Å in the dianion. Keeping previously discussed potential sources of error in mind, carboxyphosphate meets both the geometric and energetic criteria to be classified as a SSHB. While the pK_a values of donor and acceptor are similar, dependent upon charge state, they are far enough apart that the proton does not readily transfer and the CAHB cannot be classified as low barrier. In summary, the bond is considered to be short-strong but not low barrier for both monoanion and dianion pseudochair carboxyphosphate.

4.5 Conclusions

The CAHB strength for the pseudochair conformation of monanionic and dianionic carboxyphosphate has been estimated by independent additive and subtractive corrections to the standard open-closed method. The additive correction scheme shows differential Pauli repulsion contributions between pseudochair (closed) and open conformations of carboxyphosphate that are critical in accurate estimation of CAHB strength. This study has produced CAHB estimates of -24.5 kcal/mol for the monoanion and -18.4 kcal/mol for the dianion. Results from the subtractive technique reinforce those from our additive procedure, where the predicted CAHB strength ranges from -17.8 to -25.4 kcal/mol for the monoanion and from-15.7 to 20.9 kcal/mol for the dianion. Ultimately, we find the CAHB in carboxyphosphate meets the criteria for short-strong hydrogen bonds. However, carboxyphosphate has a unique energy profile that does not result in the symmetric symmetric double-well behavior characteristic of low-barrier hydrogen bonds. The CAHB in carboxyphosphate is a short-strong but cannot be classified as low-barrier, because the proton is not equally shared between donor and acceptor atoms.

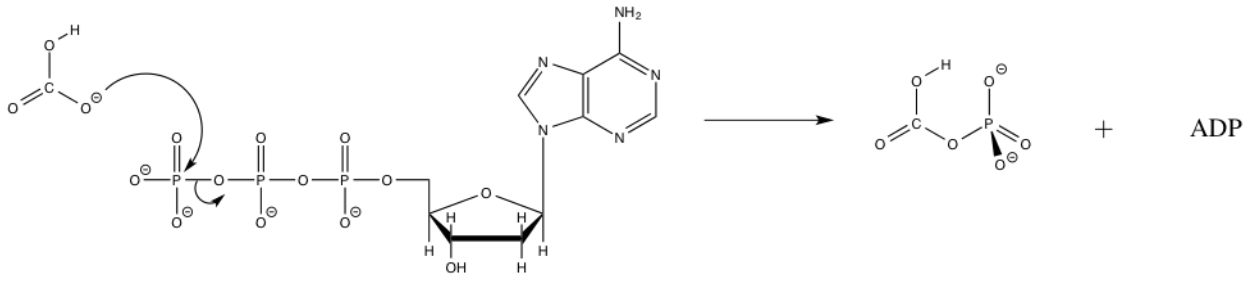
Chapter 5

5 Prediction of pK_a values for carboxyphosphate and its mechanistic implications for ATP-dependent carboxylase enzymes

To date, all attempts to synthesize, isolate, and detect carboxyphosphate directly have proved unsuccessful leaving a great deal of uncertainty as to the protonation state it may exist in under physiological conditions.[88,92,105,109,206,251] It has been a commonly held notion that a general base deprotonates bicarbonate to generating the more reactive carbonate and initiating reaction with ATP to generate carboxyphosphate as a trianion.[20,41,102] However, Herschlag and Jencks have shown that the reaction bicarbonate with a model for ATP, Υ-picoline monoanion (PicP), is 6x faster than reaction with carbonate, despite its greater basicity.[101] Furthermore, when corrected for differences in basicity, it is estimated that bicarbonate is *ca*. 200 times more reactive than carbonate towards PicP. The preference for the traditionally less reactive substrate, bicarbonate, can be explained by the charge repulsion that would result from reaction with the deprotonated form, carbonate. The highly unfavorable electrostatic interactions that would arise between carbonate and the phosphoryl oxygen atoms of ATP are mitigated by the presence of a hydrogen atom in bicarbonate.

A simple examination of pK_a values for bicarbonate and ATP can also provide insight into the likely ionization state of carboxyphosphate. The pKa value for bicarbonate is 10.3. ATP has three highly acidic protons and the remaining proton of the γ-phosphate has a value of 6.6. Thus at physiological pH, bicarbonate would retain its proton and ATP would exist majorly as the fully deprotonated form, with a small concentration retaining a proton on the γ-phosphate. This

simple analysis suggest that carboxyphosphate ought to be formed principally as a dianion as shown in Scheme 5.1.



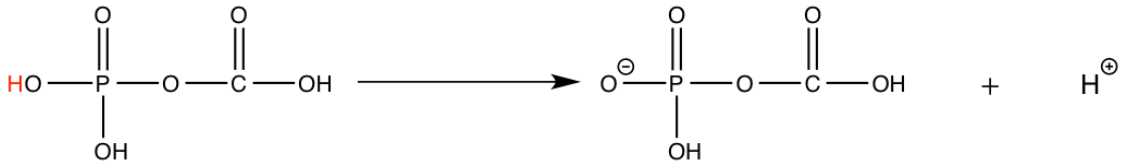
. **Scheme 5.1.** S_N2 reaction of bicarbonate with ATP to form carboxyphosphate as a dianion.

Despite a wealth of data and speculation spanning several decades, the molecular mechanism surrounding CP that is shared by several ATP-grasp enzymes remains poorly understood.[26,29,30,41,64,66,92,93,102,252] Although, the formation of carboxyphosphate has been widely adopted by the scientific community, there is a great deal of contradictory scientific evidence.[20,30,87] Carboxyphosphate is often represented as a trianion, but it is expected that it can also exist and may be more stable as a monoanion or dianion. To establish the most probable charge state for CP, theoretical pK_a values have been calculated.

5.1 Test sets

Unfortunately, there is no universal method for pK_a calculation that gives accurate results for all systems. Therefore to evaluate the accuracy of different methods for calculating the pK_a of carboxyphosphate, a series of test sets, molecules with published pK_a values representative of the functional groups present in carboxyphosphate, were used. The first pK_a value for

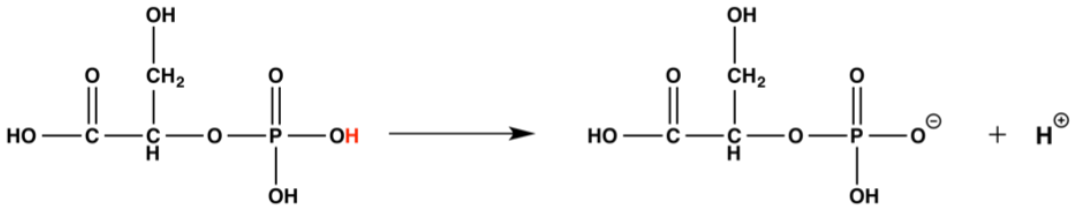
carboxyphosphate represents the loss of a proton from the phosphate to go from a neutral structure to the monoanion.



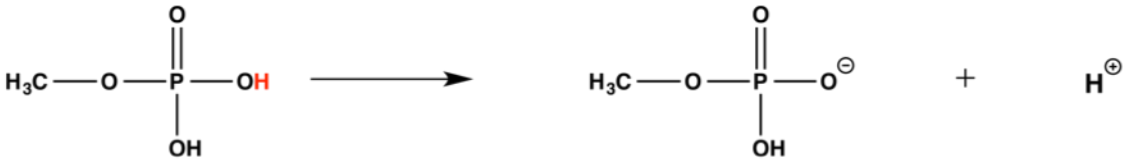
neutral **monoanion**

Scheme 5.2. Acid dissociation reaction of neutral carboxyphosphate to give the monoanion form.

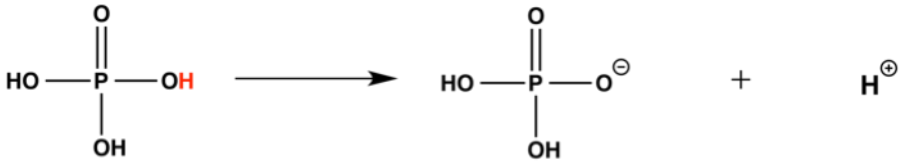
The test sets chosen to model this pK_a are 2-phosphoglyceric acid (2-PG), methyl phosphonic acid, and phosphoric acid, as shown in Figure 5.1.



2-phosphoglyceric acid (2-PG) **pK_{a1} = 1.42**



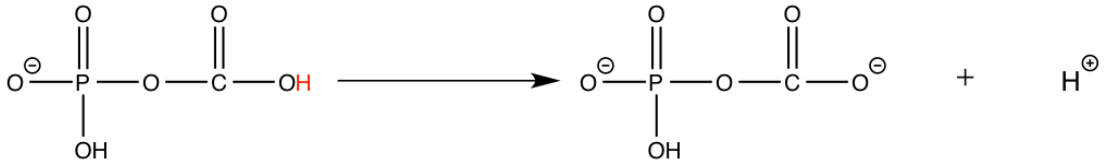
methyl phosphonic **pK_{a1} = 1.54**



phosphoric acid **pK_{a1} = 2.12**

Figure 5.1. Test sets used to model the first acid dissociation constant of carboxyphosphate.

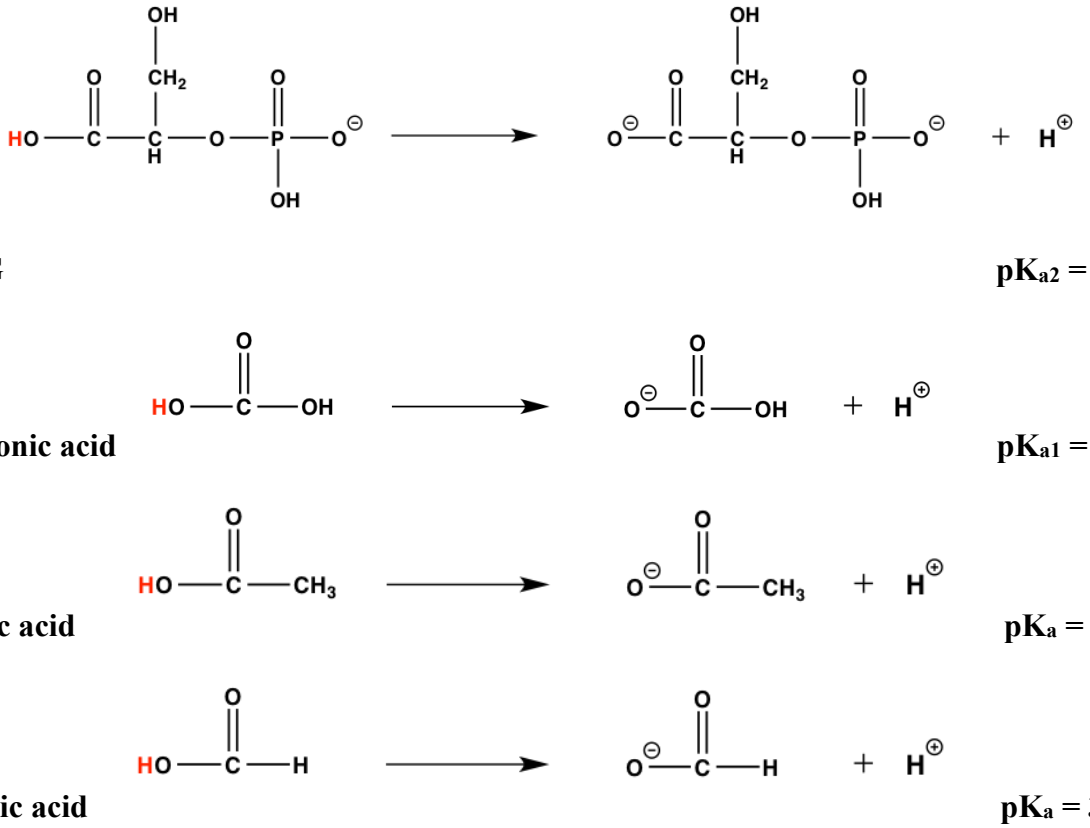
The second acid dissociation constant for carboxyphosphate describes the loss of a proton from the carboxylic acid of the monoanion to yield the dianion, as shown in Scheme 5.3.



monoanion **dianion**

Scheme 5.3. Acid dissociation reaction of carboxyphosphate from monoanion to dianion.

Acid dissociation reactions of for the loss of the second proton of 2-PG, carbonic acid, acetic acid, and formic acid are used to model the second pK_a of carboxyphosphate, as shown in Figure 5.2.



2-PG **pK_{a2} = 3.55**

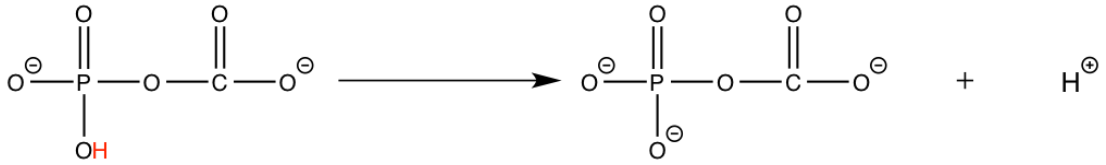
carbonic acid **pK_{a1} = 3.60**

acetic acid **pK_a = 4.76**

formic acid **pK_a = 3.77**

Figure 5.2. Test sets used for the second pK_a of carboxyphosphate.

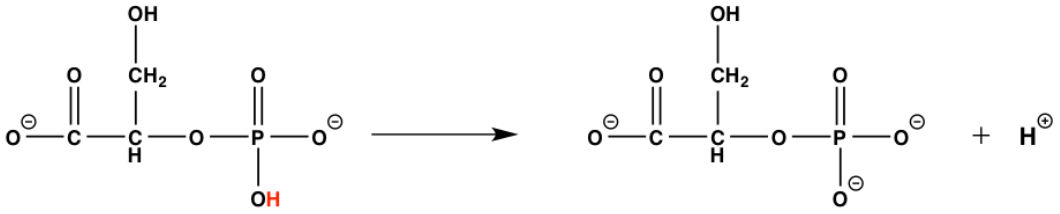
The final pK_a value for carboxyphosphate represents the loss of the second proton from the phosphate to go from dianion to trianion, as shown in Scheme 5.3.



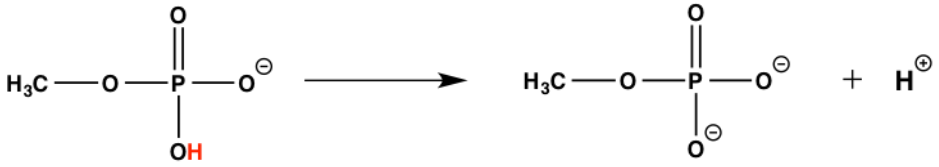
dianion **trianion**

Scheme 5.4. Acid dissociation reaction of dianion carboxyphosphate to trianion.

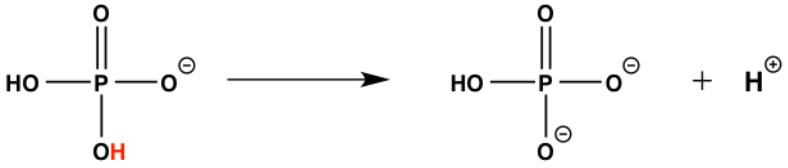
Since the final pK_a describes the loss of the last proton from the phosphate, it employs the molecules generated from the initial deprotonation.



2-PG **pK_{a3} = 7.10**



methyl phosphonic acid **pK_{a2} = 6.31**



phosphoric acid **pK_{a2} = 7.21**

Figure 5.3. Test sets used to model the third pK_a of carboxyphosphate.

5.2 Computing Acid Dissociation Free Energies from Solvation Free Energies

For the acid dissociation reaction shown below in Equation 5.1,

$$AH_{(aq)} \overset{K_a}{\rightleftharpoons} A^-_{(aq)} + H^+_{(aq)} \tag{5.1}$$

the equilibrium constant is defined according to the activities of the species involved:

$$K_a = \frac{\gamma_{A^-}[A^-]\gamma_{H^+}[H^+]}{\gamma_{AH}[HA]} \tag{5.2}$$

where the activity is the product of concentration and the activity coefficient, γ. For ideal conditions, the activities are replaced with concentrations to give

$$K_a \approx \frac{[H^+][A^-]}{[AH]} \tag{5.3}$$

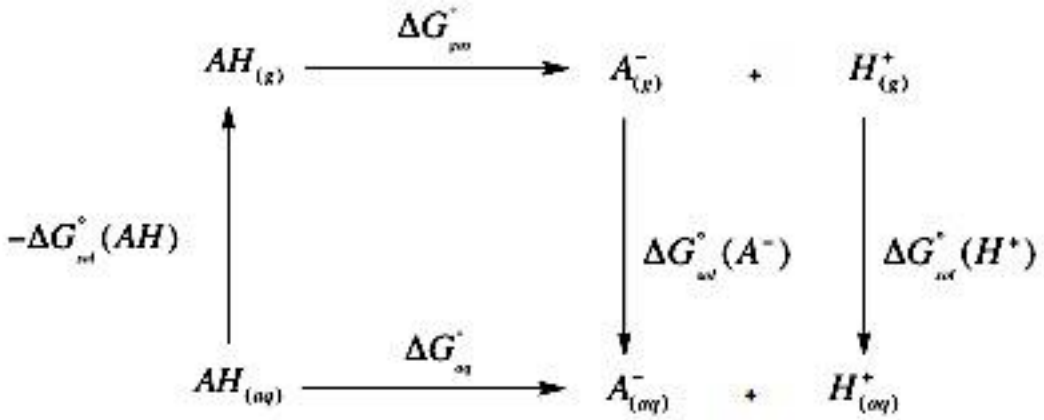
The equilibrium constant, or more specifically in this case, the acid dissociation constant is related to the Gibbs free energy as follows:

$$\Delta G^\circ = -RT(lnK_a) = -2.303RT(log_{10}K_a) \tag{5.4}$$

where R is the universal gas constant and T is the temperature.

Therefore, the standard-state free energy change, ΔG^*_{aq} is related to the aqueous pK_a as follows in Equation 5.5:

$$pK_a = \frac{\Delta G^*_{aq}}{RTln(10)} \tag{5.5}$$



Scheme 5.5. Thermodynamic cycle for acid dissociation reaction in eq 5.1.

Using the thermodynamic cycle shown in Scheme 5.5, ΔG^{*}_{aq} can be rewritten in terms of the aqueous solvation free energies of the acid AH and its conjugate base A^- as follows:

$$\Delta G^{\circ}_{(aq)}(AH) = \Delta G^{\circ}_{gas}(AH) + \Delta G^{\circ}_{sol}(A^-) - \Delta G^{\circ}_{sol}(AH) + \Delta G^{\circ}_{sol}(H^+) \qquad (5.3)$$

where $\Delta G^{\circ}_{sol}(AH)$ and $\Delta G^{\circ}_{sol}(A^-)$ are the standard-state aqueous solvation free energies of AH and A^-, respectively, $\Delta G^{\circ}_{sol}(H^+)$ is the aqueous solvation free energy of H^+, $\Delta G^{\circ}_{gas}(A)$ is the gas-phase acidity of AH, and is defined as follows:

$$\Delta G^{\circ}_{(aq)} = \Delta G^{\circ}_{gas}(A^-) + G^{\circ}_{gas}(H^+) - G^{\circ}_{gas}(AH) \qquad (5.4)$$

Experimental values are used for $\Delta G^{\circ}_{sol}(H^+)$ and $G^{\circ}_{gas}(H^+)$. Therefore, pK_a can be calculated according to Equation 5.5.

$$pK_a = \frac{G^{\circ}_{gas}(A^-) - G^{\circ}_{gas}(AH) + \Delta G^{\circ}_{sol}(A^-) - \Delta G^{\circ}_{sol}(AH) - 270.28567\ kcal/mol}{1.36449\ kcal/mol} \qquad (5.5)$$

5.3 *Computational Methods*

Calculation of free energies was performed using Gaussian 09[116] with M06-2X/jul-cc-pVTZ[169,202] level of theory. High-level corrections were applied to gas phase quantities consisting of the single point energy calculated CCSD(T)/jul-cc-pVTZ//MP2/jul-cc-pVQZ[130,253] with free energy corrections calculated at M06-2X/jul-cc-pVTZ and scaled by 0.94 to reproduce free energy corrections from CCSD(T)/jul-cc-pVTZ. Aqueous solvation free energies were calculated using the SMD implicit solvation method developed by Truhlar and co-workers.[254]

5.4 *Results and Discussion*

2-Phosphoglyceric acid (2-PG), methyl phosphonic acid, and phosphoric acid were selected as test sets to model the first pK_a of CP. These test sets were chosen to represent the loss of the first proton of CP from the phosphate to yield a monoanion. The results from performing all calculations with M06-2X/jul-cc-pVTZ level of theory and those achieved by using CCSD(T)/jul-cc-pVTZ single point energy calculations with free energy corrections calculated at M06-2X/jul-cc-pVTZ and scaled by 0.94 to represent values consistent with CCSD(T)/jul-cc-pVTZ are shown in Table 5.1.

The M06-2X/jul-cc-pVTZ method gave pK_a values that were in good agreement with the reference values producing absolute errors of 1.40, 1.54, and 2.12 for 2-PG, methyl phosphonic acid, and phosphoric acid, respectively. Giving a mean absolute error (MAE) across all 3 systems in this test set of 1.08. However, this error can be reduced further by incorporating

higher level CCSD(T) computations for the gas phase values. When CCSD(T) single point energy evaluations are combined with scaled DFT frequencies the MAE is reduced to 0.81.

Table 5.1. Test set results for 1st pK_a of CP.

Test set	Experimental pK_a	DFT Calculated pK_a	Error	CCSD(T) Scaled pK_a	Error
2-phosphogyceric acid (2-PG)	1.42	0.02	1.40	0.40	1.02
methyl phosphonic acid	1.54	1.07	0.47	1.85	0.31
phosphoric acid	2.12	0.79	1.37	1.06	1.10
			MAE = 1.08		MAE = 0.81

As expected the larger system, 2-PG produces a larger error than methyl phosphonic acid. In general, because all computational methods are approximations there will be some inherent error and that error will increase with system size. The calculation of the pK_a for deprotonation of phosphoric acid to dihydrogen phosphate is also nearly three times that of methyl phosphonic acid. This is likely due to the additional stabilization of the anionic oxygen atoms in methyl hydrogen phosphate that can be provided by the hydrogen atoms of the methyl group as shown in Figure 5.5.

Figure 5.4. Comparison of geometry optimized (M06-2X/jul-cc-pVTZ) methyl hydrogen phosphate (left) vs. dihydrogen phosphate.

The major weakness of implicit solvation methods used in these calculations is that they cannot properly account for strong solute-solvent interactions. The hydrogen atoms of the methyl group in methyl hydrogen phosphate stabilize the charge present on the anionic oxygens.

Table 5.2. Test set results for 2nd pK_a of CP

Acid	Experimental pK_a	DFT Calculated pK_a	Error	CCSD(T) Scaled pK_a	Error
2-PG (2nd deprotonation)	3.55	1.89	1.66	2.94	0.61
carbonic acid	3.60	2.43	1.17	3.69	0.09
acetic acid	4.76	6.16	1.40	7.44	2.68
formic acid	3.77	4.41	0.64	5.87	2.10
		MAE = 1.22/1.42*		MAE = 1.23/0.35*	

*Values calculated from 2-PG and carbonic acid only

Results for the test sets selected to represent the second pK_a of CP, loss of the second proton from the carboxylic acid to give a dianion were mixed. The method of using scaled M06-2X/jul-cc-pVTZ free energy corrections with CCSD(T)/jul-cc-pVTZ single point calculations for gas phase values produced pK_a values in very close agreement with experiment for the second deprotonation of 2-PG and carbonic acid. The pK_a values calculated with this method for the second deprotonation of 2-PG and carbonic acid only deviated from experiment by 0.61 and 0.09, respectively. The values calculated using the M06-2X/jul-cc-pVTZ method for acetic acid and formic acid yielded errors consistent with those previously seen for this method at 1.40 and 0.64 respectively. However, the error for these systems rises dramatically (>2) when higher-level calculations (CCSD(T)/jul-cc-pVTZ) are used for the gas phase. Comparing each of these systems, there is a noticeable difference in functionality for these two molecules. For example, in carbonic acid the HPO_3^- group of CP is replaced by H, but in acetic acid and formic acid, HPO_4^-

is replaced with CH_3 and H, respectively. The elimination of the bridging oxygen appears to have a large impact on calculation of pK_a and indicates that these are inappropriate references for this system. While the errors for these systems are in the typical range for other test molecules when only the M06-2X/jul-cc-pVTZ method is used, it is likely just a convenient cancellation of errors that leads to this value and the fact that they no longer follow the trend seen in the other test sets when CCSD(T)/jul-cc-pVTZ computations are included indicates that they are not appropriate references for this system.

The results for the last test set, which represents the loss of a second proton from dianionic CP to produce a trianion are given in Table 5.3. As expected, the use of implicit solvation is not sufficient to represent the strong solute solvent interactions that occur in the highly charged anions produced for this last pK_a value.

Table 5.3. Test set results for 3rd pKa of CP.

Acid	Experimental pK_a	DFT Calculated pKa	Error	CCSD(T) Scaled pK_a	Error
2-PG (3rd deprotonation)	7.10	14.44	7.34	14.29	7.19
methyl hydrogen phosphate	6.31	11.51	5.20	11.88	5.57
dihydrogen phosphate	7.21	12.35	5.14	12.40	5.19
			MAE = 5.89		MAE = 5.98

5.4.1 Cluster continuum method

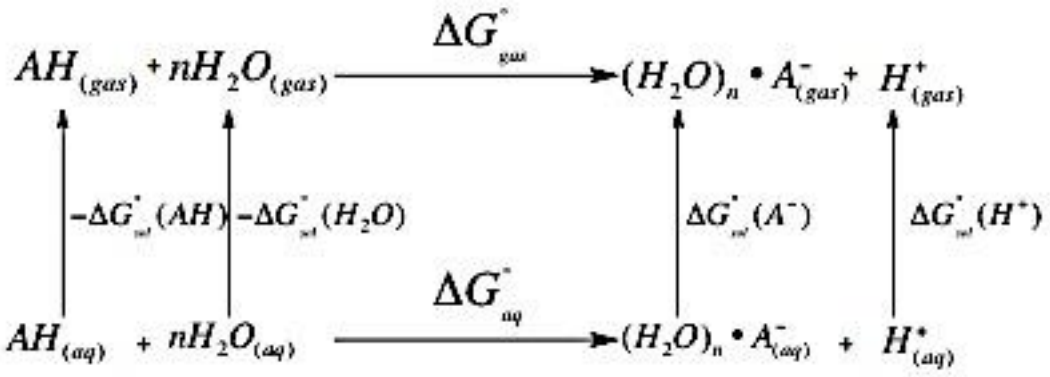
To improve pK_a predictions for the last value of CP, explicit waters must be introduced to account for strong solute-solvent interactions not properly described by the SMD implicit solvation method.[254] It is of interest to calculate accurate pK_a values at the lowest possible computational cost. Therefore, the accuracy of the calculated pK_a values as a function of increasing the number of explicit water molecules was systematically investigated for the following acid dissociation reaction:

$$H_2PO_4^- \rightleftharpoons HPO_4^{2-} + H^+ \qquad pK_a = 7.21 \qquad (5.5)$$

For explicit solvation, the cluster continuum model is used and the acid dissociation reaction given is altered from Equation 5.1 to incorporate explicit water molecules as follows:

$$AH_{(aq)} + nH_2O_{(aq)} \rightleftharpoons (H_2O)_n \cdot A^-_{(aq)} + H^+_{(aq)} \qquad (5.6)$$

which gives rise to the thermodynamic cycle shown in Scheme 5.6



Scheme 5.6. Thermodynamic cycle for the cluster continuum model.

From the thermodynamic cycle given in Scheme 5.6. pK_a is now calculated according to Equation 5.7.

$$pK_a = \frac{G^{\circ}_{gas}(A^-) - G^{\circ}_{gas}(AH) + \Delta G^{\circ}_{BE}[(H_2O)_n \cdot A^-] + \Delta G^{\circ}_{sol}[(H_2O)_n \cdot A^-] - \Delta G^{\circ}_{sol}(AH) - n\Delta G^{\circ}_{sol}(H_2O) - 270.3}{1.36449}$$

(5.7)

where $\Delta G^{\circ}_{BE}[(H_2O)_n \cdot A^-] = G^{\circ}_{gas}[(H_2O)_n \cdot A^-] - G^{\circ}_{gas}(H_2O)_n - G^{\circ}_{gas}(A^-)$. This term represents the specific interaction energy associated with the binding of the cluster (water) molecules to the anion.

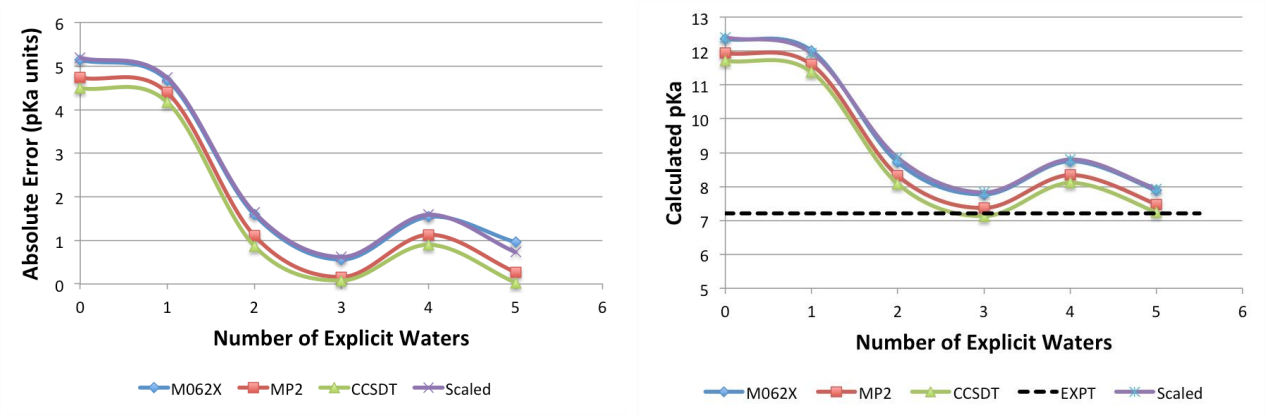


Figure 5.5. Plots of absolute error (left) and computed pK_a values (right) as a function of increasing explicit water molecules for the dissociation of dihydrogen phosphate to hydrogen phosphate. The curve for M06-2X indicates that all values were calculated with the M06-2X level of theory. The curve for MP2 represents the effect of calculation of gas phase values with MP2/jul-cc-pVQZ. The curve marked CCSD(T) gives the effect utilizing CCSD(T)/jul-cc-pVTZ for gas phase calculations. The curve marked scaled indicates that single point CCSD(T)/jul-cc-pVTZ single point energies were combined with M06-2X/jul-cc-pVTZ free energy corrections scaled by a factor of 0.94.

Figure 5.6 shows the convergence of calculated pK_a values to the experimental value of 7.21 as a function of increasing cluster molecules for the dissociation of dihydrogen phosphate to hydrogen phosphate. Along with testing the effects of increasing cluster molecules, different variations of the inclusion of higher-level computations for gas phase free energies were tested. Specifically, the use of MP2/jul-cc-pVQZ, CCSD(T)/jul-cc-pVTZ and CCSD(T)/jul-cc-pVTZ single point energies combined with free energy corrections scaled by 0.94 and calculated at M06-2X/jul-cc-pVTZ. It is clear that complete frequency analysis at the CCSD(T)/jul-cc-pVTZ level of theory gives results in closest agreement with the experimental value, followed closely

by MP2/jul-cc-pVQZ. However, these calculations require a significant amount of additional computational time for only a modest increase in accuracy.

The effect of adding of a single, explicit water is minimal, reducing the error from approximately 5 pK_a units to only about 4.5 pK_a units across all methods. However, addition of a second molecule, allows the formation of a cluster, the interaction between the two molecules reduces the energy thereby significantly reducing the error to only *ca*. 1.5 pK_a units.

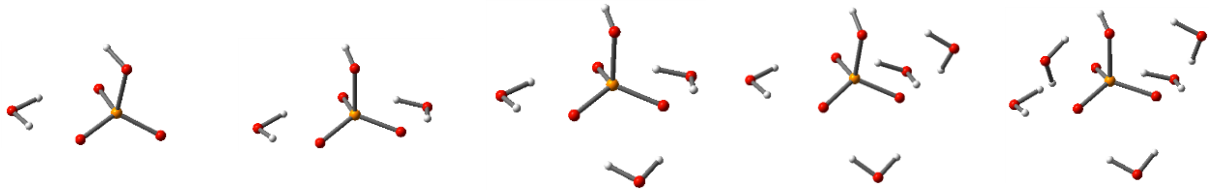


Figure 5.6. Hydrogen phosphate with one, two, three, four, and five waters bound (right to left).

The addition of three water molecules gives the best results. This can be easily rationalized by considering the structure of hydrogen phosphate. Due to resonance there will be three anionic oxygen atoms. Therefore three explicit water molecules is the minimum required to provide stabilization for each anionic oxygen atom.

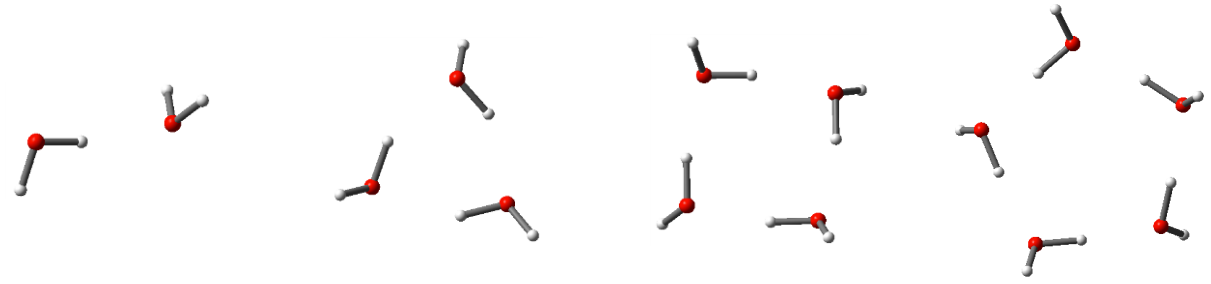


Figure 5.7. Water clusters used to calculate $\Delta G^{\circ}_{BE}[(H_2O)_n \cdot A^-]$ where n = 2, 3, 4, and 5 (left to right).

When four explicit waters are used, the error rises because a cluster of four water molecules is known to be less stable, than those formed with 2, 3, or 5.[255]

Table 5.4. Calculated pK_a values for the deprotonation of dihydrogen phosphate to hydrogen phosphate as a function of increasing explicit water molecules.

Number of explicit waters	Experimental pK_a	DFT Calculated pK_a	Error	CCSD(T) Scaled pK_a	Error
0	7.21	12.35	5.14	12.40	5.19
1		11.89	4.68	11.94	4.73
2		8.80	1.59	8.85	1.64
3		7.77	0.56	7.83	0.62
4		8.75	1.54	8.80	1.59
5		6.25	0.96	7.94	0.73

The observation that including one explicit water molecule per anionic oxygen is adequate for description of strong-solute interactions was followed for the other two molecules in this test set, the third deprotonation of 2-PG and methyl phosphate. However for the third deprotonation of 2-PG, it was unclear whether four or five water molecules should be included since one of the five anionic oxygen atoms present forms an intramolecular hydrogen bond, which may reduce its need for additional stabilization from explicit water molecules. The results for adding zero, four, and five explicit water molecules are listed in Table 5.5.

Table 5.5. Calculated pK_a values for the 3rd deprotonation of 2-PG for 0, 4, and 5 explicit waters.

Number of explicit waters	Experimental pK_a	DFT Calculated pK_a	Error	CCSD(T) Scaled pK_a	Error
0	7.10	14.44	7.34	14.29	7.19
4		10.02	3.65	9.87	3.60
5		9.55	3.48	9.26	3.43

It is clear that including five cluster molecules gives a slightly more accurate result than four with errors of 3.60 and 3.43, respectively. However, the error involved in these calculations is still much higher than those seen in hydrogen phosphate. This could be due to a number of reasons. First and foremost, 2-PG is a much larger system containing more than double the number of heavy atoms relative to hydrogen phosphate. The error could be the result of small errors compounding in a larger system.[122,141,204] However, it is more likely that the larger error in this calculation is an indication that there is a more stable arrangement of 2-PG bound to 4 and 5 water molecules. Because the molecule is much larger, there are more possible arrangements for both 2-PG and 2-PG bound to the explicit water molecules. Ideally, a systematic analysis of different possible arrangements and conformations would identify the most stable arrangement and reduce the errors seen for this system, but is beyond the scope of this study.

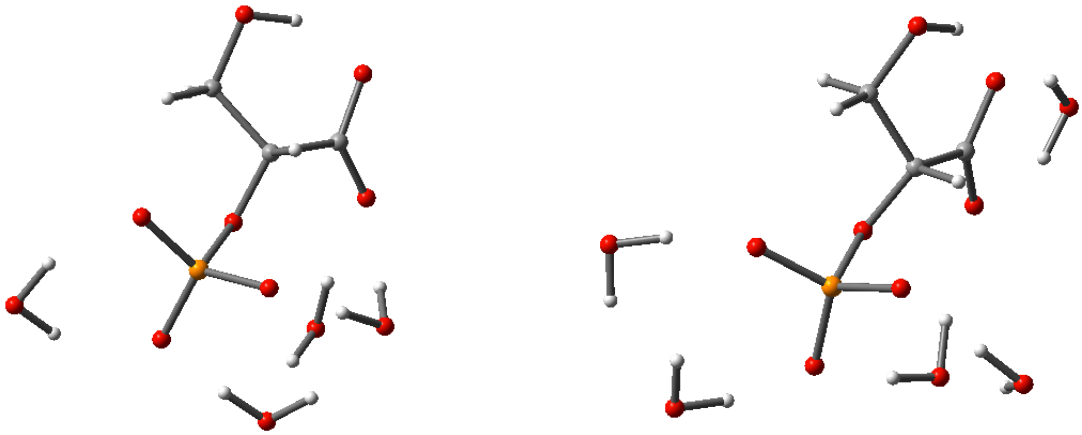


Figure 5.8. Geometric comparison of 2-PG (3rd deprotonation) bound to four (left) and five (right) explicit water molecules computed at M06-2X/jul-cc-pVTZ.

Table 5.6. Test set results for 3rd pKa of CP utilizing the cluster continuum method where 1 cluster molecule is added for each anionic oxygen present.

Acid	Experimental pK_a	DFT Calculated pK_a	Error	CCSD(T) Scaled pK_a	Error
2-PG (3rd deprotonation)	7.10	9.55	3.48	9.40	3.43
methyl hydrogen phosphate	6.31	7.64	1.33	8.01	1.70
dihydrogen phosphate acid	7.21	7.77	0.56	7.83	0.62
		MAE = 1.64			MAE = 1.92

The results for the cluster continuum method are summarized in Table 5.6. Although the error has been reduced to by more than half what it was when only implicit solvation was used, it is noticeably higher for these test sets than any of the others. In contrast to the other 2 systems, computations for dihydrogen phosphate are in very close agreement with experiment with errors of only 0.56 and 0.62 pKa units for M06-2X/jul-cc-pVTZ and CCSD(T)/jul-cc-pVTZ single point energies with scaled M06-2X/jul-cc-pVTZ free energy corrections, respectively. This is logical because dihydrogen phosphate was more thoroughly investigated than the other two molecules in this test set. More work may be required to optimize the calculation of pK_a for these other test sets. An energetic error of 1.36 kcal/mol corresponds to an error of 1 pK_a unit therefore small energetic errors can quickly contribute to large errors in pK_a estimates. As mentioned previously regarding 2-PG, these errors are probably due to the non-ideal placement of explicit water molecules and could be minimized with a thorough evaluation of the stability of different possible arrangements. It should also be noted that for all calculations using only the DFT method yields better results. Based on previous results and knowledge of the expected accuracy

of each of these methods, this is likely due to a convenient cancellation of errors that occurs when the cluster continuum method is used.

5.4.2 Carboxyphosphate

Results from test sets used to model each pKa of CP have indicated that using the SMD/M06-2X/jul-cc-pVTZ method with CCSD(T)/jul-cc-pVTZ single point energies and scaled M06-2X free energy corrections for gas phase quantities consistently produces results within *ca*. 1 pK_a unit of experimental values. For the third pK_a, explicit water molecules must be included via the cluster continuum method to achieve similar accuracy. Accordingly, pK_a values have been calculated for CP using these methods and are listed in Table 5.7.

Table 5.7. Calculated pK_a values for the 1st, 2nd, and 3rd acid dissociation constants of carboxyphosphate.

pK_a	DFT Calculated pK_a	pK_a range	CCSD(T) Scaled pK_a	pK_a range
1st	-4.95±1.08	-6.03 to -3.87	-3.43±0.81	-4.24 to -2.62
2nd	3.29±1.42	1.87 to 4.71	4.04±0.35	3.69 to 4.39
3rd	7.85±1.64	6.21 to 9.49	8.14±1.92	6.22 to 10.06

Following the observation that one explicit water per anionic oxygen is sufficient to account for strong solute-solvent effects, five explicit water molecules were bound to trianionic carboxyphosphate for the calculation of the specific interaction energy associated with the binding of the cluster (water) molecules to the anion and free energy of solvation.

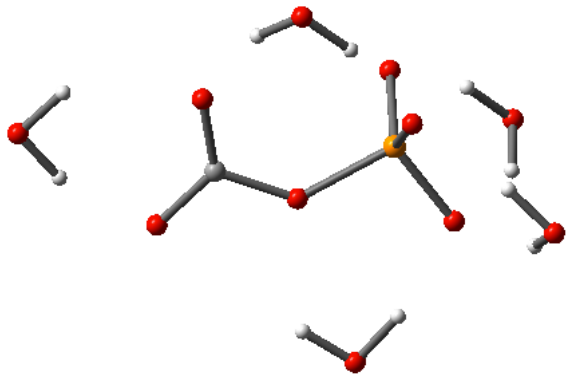


Figure 5.9. Trianionic carboxyphosphate with 5 explicit water molecules bound, geometry optimized at M06-2X/jul-cc-pVTZ.

5.5 Conclusions

Prediction of pK_a values for test sets representative of CP reveal that the first two values can be calculated within 1 pK_a unit using only implicit solvation. It is also clear that the M06-2X/jul-cc-pVTZ method paired with CCSD(T)/jul-cc-pVTZ scaled and corrected energies for gas phase computations gives results in good agreement with experiment at modest computational cost. For more basic systems with pK_a values near 7 that involve highly charged anions, a cluster continuum approach must be used. Accurate pK_a values have been achieved with this method by including one explicit water molecule for each anionic oxygen atom present.

The predicted pK_a values for CP indicate that it most likely be present as a dianion but may also be stable as a trianion in aqueous solution. Unfortunately, the calculated value for the last pK_a of CP has the largest uncertainty with a computed range between *ca*. 6-10. However, a more detailed investigation of the most stable binding modes of explicit water molecules would reduce this uncertainty. Additionally, it is of interest to investigate how these values may shift

when CP is located in the active site in order to distinguish between different proposed mechanisms.

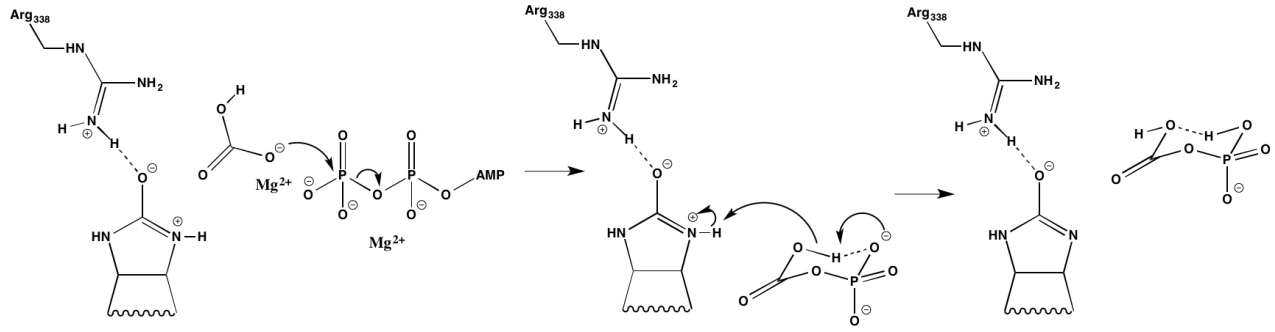
Chapter 6

6 Conclusions and Future Directions

This dissertation has identified and characterized many important structural features of carboxyphosphate in vacuum and aqueous solution. The finding that M062X/jul-cc-pVTZ consistently produces values in close agreement with CCSD(T) and MP2 computations provides the necessary foundation for more complex calculations involving carboxyphosphate and its role in the mechanism of ATP-grasp enzymes. The identification of a unique pseudochair conformation that features a short, strong CAHB is believed to be an important mechanistic feature of the six ATP-grasp enzymes previously discussed. We propose a possible mechanism for the reactions catalyzed by biotin carboxylase based on our findings.

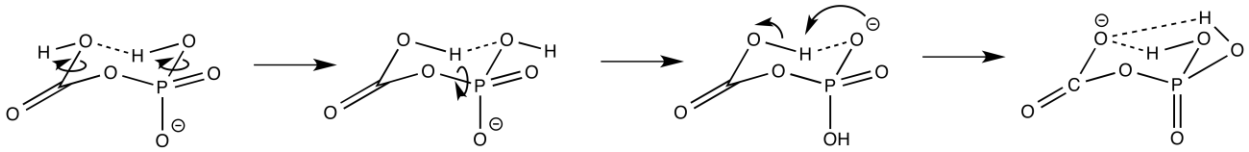
6.1 Proposed Mechanism for Biotin Carboxylase

Structural analysis for dianion CP reveals that CP is *ca*. 20 kcal/mol more stable when the phosphate side is protonated rather than the carboxy. However, the formation of CP is most likely to occur through reaction of bicarbonate and ATP, which would result in a CP intermediate where the carboxy side is protonated. Structural analysis indicates that the proton is not stable and will instantaneously transfer if oriented in position to form a CAHB to generate the low energy PC structure where the HBD is the phosphate. The shuttling of the proton across a CAHB in this manner is referred to as charge-assisted proton shuttling (CAPS).



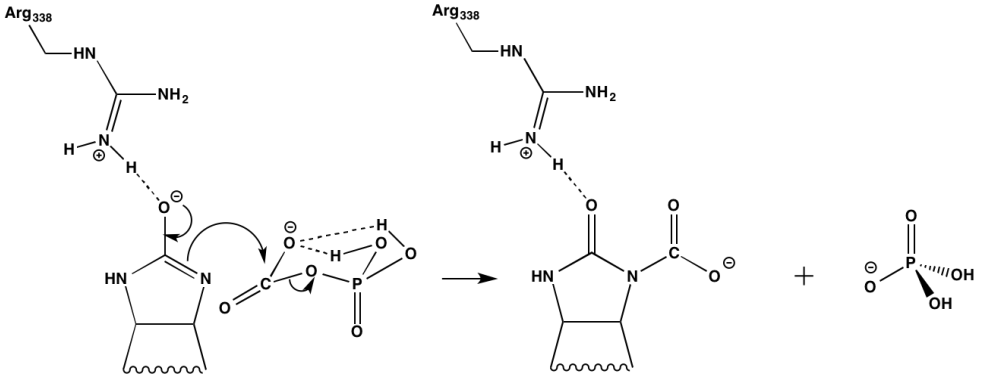
Scheme 6.1. Formation of carboxyphosphate as a dianion followed by deprotonation of biotin to give monoanionaic carboxyphosphate the enolate form of biotin.

This could allow for the deprotonation of biotin by the carboxy group of CP, a previously unconsidered avenue for one of the longstanding mechanistic problems of the biotin-dependent ATP-grasp enzymes. Although, it has been previously suggested that the phosphate group is the catalytic base responsible for deprotonating biotin, bicarbonate is positioned closest to biotin. One of the obstacles to identifying a base to deprotonate biotin is the estimated pK_a value for biotin that has been estimated to be 17.4. However, the interaction of the carbonyl with arginine 338 reduces this value considerably, making deprotonation by the carboxy group of CP feasible. This step would generate the *PC7* conformation of CP whose energy could be lowered through rotation of the HBD and HBA to give *PC4*. Further rotation of the phosphate group to reform a CAHB leads to either the *PC1* or *PC2* conformation.



Scheme 6.2. Rotation of hydrogen bond donor and acceptor of PC_7 to PC_4 followed by further rotation to give PC_1 which then becomes PC_3 through the CAPS mechanism.

CAPS can now take place for a second time leading to *PC₃*. The concentration of both protons on the phosphate makes it an ideal leaving group and initiates the collapse of **1** into carbonate and inorganic phosphate. The *N1* of biotin can then attack the carbonate during or after this collapse to yield carboxybiotin as shown in Scheme 6.3.



Scheme 6.3. Collapse of carboxyphosphate into carboxybiotin and inorganic phosphate.

6.2 Future Work

The discovery of several stable states of carboxyphosphate has allowed us to propose a possible mechanism by which they are connected in order to produce carboxybiotin and inorganic phosphate from bicarbonate and ATP. The identification of an appropriate level of theory allows future researchers to explore the energetics of the mechanism proposed in this work versus that proposed by Chou *et al*. Energetic comparison of these different possible

pathways will yield a better understanding of the how CP is involved in the mechanisms of ATP-grasp enzymes which is critical in the development of new treatments for many serious human diseases.

Additionally, we have made a great deal of progress in predicting the charge state of CP. However, several issues need to be addressed. First, a more detailed assessment of the ideal binding modes for explicit water molecules needs to be developed for the cluster continuum method in order to reduce the uncertainty in the prediction of pK_a values near 7. There are many well-established methods available to sample different possible configurations that range from standard statistical methods to quantum or molecular dynamics.

The other major goal of pK_a prediction for CP that has yet to be addressed is how these values will shift from those calculated for aqueous solution in the enzymatic pocket. While our estimates indicate that CP is most likely to exist as a dianion in aqueous solution, pK_a values have been shown to shift by several units, depending on their interaction with key residues in the enzymatic pocket.[256,257] This is an issue of great interest in the field of biochemistry. Therefore, many possible methods have been used to estimate this shift to varying degrees of success.[87,258,259] While a thorough investigation to assess the accuracy of different possible methods should be performed, it seems most logical to explore a method that builds upon the calculations used for aqueous solution.

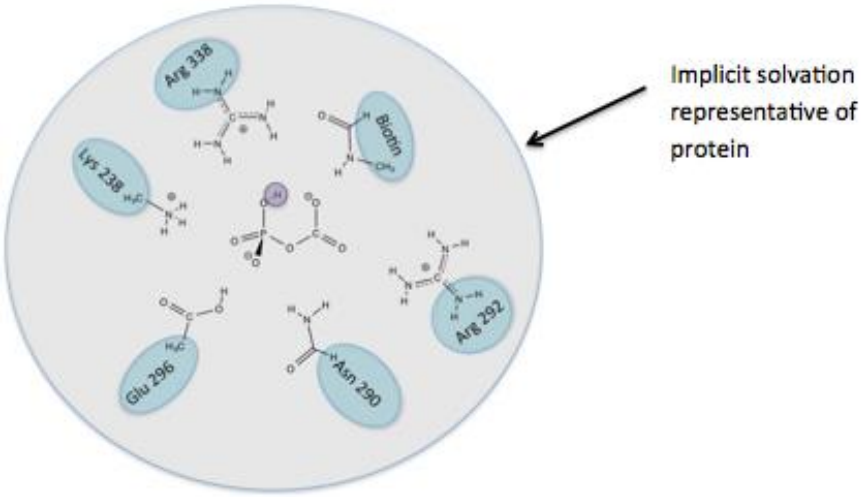


Figure 6.1. Representation of the enzymatic environment modeled with explicit representations of the functional groups for selected amino acids and the SMD solvation method using a dielectric constant representative of the enzyme.

Previous work suggests that only the residues that directly interact with a ligand need to be represented explicitly in order to estimate accurate pK_a values.[256,259,260] The rest of the protein can therefore be represented as a dielectric constant. This is the same concept that we have utilized in the cluster continuum approach so it follows that a modified version of this approach can produce accurate pK_a shifts for the enzymatic environment. Recently, NMR chemical shift perturbations have been used to directly determine the optimal protein dielectric constant, ε_p, and found a range of 2-5 with an optimum of 3.[257] This range is consistent with the ε_p measured from protein powder giving confidence that it is an appropriate value to use in further calculations. The residues Lys238, Asn290, Arg292, Glu296, and Arg338 of the biotin carboxylase component of ACC are all in position to interact with carboxyphosphate and have been identified as playing key roles in this mechanism from mutagenic studies.[20,26,58,66,102] Thus, a first step in estimating the pK_a shifts should be to utilize a cluster continuum type method where the functional groups of the previously named residues are treated at the quantum level and the bulk

protein environment is represented as a dielectric constant of 3. The accuracy of this method can then be compared to other possible methods for validation.

Chapter 7

7 References

(1) World Health Organization. Noncommunicable diseases http://www.who.int/topics/noncommunicable_diseases/en/ (accessed Jun 20, 2015).

(2) World Health Organization. The top 10 causes of death http://www.who.int/mediacentre/factsheets/fs310/en/index2.html (accessed Jun 26, 2015).

(3) United Nations Department of Public Information. Non-commmunicable diseases deemed development challenge of "epidemic proportions" in political declaration adopted during landmark general assembly summit http://www.un.org/press/en/2011/ga11138.doc.htm (accessed Jul 11, 2015).

(4) Haslam, D. W.; James, W. P. T. Obesity. *Lancet* **2005**, *366* (9492), 1197–1209.

(5) Stewart, B. W.; Wild, C. P. *World Cancer Report*; IARC Nonserial Publications: Lyon, France, 2014.

(6) Shi, Y.; Hu, F. B. The Global Implications of Diabetes and Cancer. *Lancet* **2014**, *383* (9933), 1947–1948.

(7) World Health Organization. The 10 leading causes of death in the world, 2000 and 2012 http://www.who.int/features/factfiles/noncommunicable_diseases/en/.

(8) Vos, T.; Flaxman, A. D.; Naghavi, M.; Lozano, R.; Michaud, C.; Ezzati, M.; Shibuya, K.; Salomon, J. A.; Abdalla, S.; Aboyans, V.; et al. Years Lived with Disability (YLDs) for 1160 Sequelae of 289 Diseases and Injuries 1990-2010: A Systematic Analysis for the Global Burden of Disease Study 2010. *Lancet* **2012**, *380* (9859), 2163–2196.

(9) International Diabetes Federation. IDF Diabetes Atlas: key findings 2014

http://www.idf.org/diabetesatlas/update-2014.

(10) Economic Costs of Diabetes in the U.S. in 2012. *Diabetes Care* **2013**, *36* (4), 1033–1046.

(11) Škedelj, V.; Tomašić, T.; Mašič, L. P.; Zega, A. ATP-Binding Site of Bacterial Enzymes as a Target for Antibacterial Drug Design. *J. Med. Chem.* **2011**, *54* (4), 915–929.

(12) Tong, L. Acetyl-Coenzyme A Carboxylase: Crucial Metabolic Enzyme and Attractive Target for Drug Discovery. *Cell. Mol. Life Sci.* **2005**, *62* (16), 1784–1803.

(13) Tong, L.; Harwood, H. J. Acetyl-Coenzyme A Carboxylases: Versatile Targets for Drug Discovery. *J. Cell. Biochem.* **2006**, *99* (6), 1476–1488.

(14) Supuran, C. T.; Fiore, A. Di; Simone, G. De. Carbonic Anhydrase Inhibitors as Emerging Drugs for the Treatment of Obesity. *Expert Opin. Emerg. Drugs* **2008**, *13* (2), 383–392.

(15) Fawaz, M. V.; Topper, M. E.; Firestine, S. M. The ATP-Grasp Enzymes. *Bioorg. Chem.* **2011**, *39* (5-6), 185–191.

(16) Wakil, S. J.; Abu-Elheiga, L. A. Fatty Acid Metabolism: Target for Metabolic Syndrome. *J. Lipid Res.* **2009**, *50* (Suppl.), S138–S143.

(17) Galperin, M. Y.; Koonin, E. V. A Diverse Superfamily of Enzymes with ATP-Dependent Carboxylate-Amine/thiol Ligase Activity. *Protein Sci.* **1997**, *6* (12), 2639–2643.

(18) Thoden, J. B.; Firestine, S.; Nixon, a.; Benkovic, S. J.; Holden, H. M. Molecular Structure of Escherichia Coli purT-Encoded Glycinamide Ribonucleotide Transformylase. *Biochemistry* **2000**, *39* (30), 8791–8802.

(19) Wang, W.; Kappock, T. J.; Stubbe, J.; Ealick, S. E. X-Ray Crystal Structure of Glycinamide Ribonucleotide Synthetase from Escherichia Coli. *Biochemistry* **1998**, *37* (45), 15647–15662.

(20) Chou, C.-Y.; Yu, L. P. C.; Tong, L. Crystal Structure of Biotin Carboxylase in Complex

with Substrates and Implications for Its Catalytic Mechanism. *J. Biol. Chem.* **2009**, *284* (17), 11690–11697.

(21) Li, H.; Fast, W.; Benkovic, S. J. Structural and Functional Modularity of Proteins in the de Novo Purine Biosynthetic Pathway. *Protein Sci.* **2009**, *18* (5), 881–892.

(22) Murzin, A. G. Structural Classification of Proteins : New Superfamilies. *Curr. Opin. Struct. Biol.* **1996**, *6* (3), 386–394.

(23) Jitrapakdee, S.; St Maurice, M.; Rayment, I.; Cleland, W. W.; Wallace, J. C.; Attwood, P. V. Structure, Mechanism and Regulation of Pyruvate Carboxylase. *Biochem. J.* **2008**, *413* (3), 369–387.

(24) Thoden, J. B.; Holden, H. M.; Paritala, H.; Firestine, S. M. Structural and Functional Studies of *Aspergillus Clavatus* N^5 -Carboxyaminoimidazole Ribonucleotide Synthetase,. *Biochemistry* **2010**, *49* (4), 752–760.

(25) Wolodko, W. T.; Fraser, M. E.; James, M. N.; Bridger, W. A. The Crystal Structure of Succinyl-CoA Synthetase from Escherichia Coli at 2.5-A Resolution. *J. Biol. Chem.* **1994**, *269* (14), 10883–10890.

(26) Sloane, V. Site-Directed Mutagenesis of ATP Binding Residues of Biotin Carboxylase. Insight into the Mechanism of Catalysis. *J. Biol. Chem.* **2001**, *276* (27), 24991–24996.

(27) Lasso, G.; Yu, L. P. C.; Gil, D.; Xiang, S.; Tong, L.; Valle, M. Cro-EM Analysis Reveals New Insights into Mechanism of Action of Pyruvate Carboxylase. *Structure* **2010**, *18* (10), 1300–1310.

(28) Erck, C.; Peris, L.; Andrieux, A.; Meissirel, C.; Gruber, A. D.; Vernet, M.; Schweitzer, A.; Saoudi, Y.; Pointu, H.; Bosc, C.; et al. A Vital Role of Tubulin-Tyrosine-Ligase for Neuronal Organization. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (22), 7853–7858.

(29) Attwood, P. V. The Structure and the Mechanism of Action of Pyruvate Carboxylase. *Int. J. Biochem. Cell Biol.* **1995**, *27* (3), 231–249.

(30) Attwood, P. V; Wallace, J. C. Chemical and Catalytic Mechanisms of Carboxyl Transfer Reactions in Biotin-Dependent Enzymes. *Acc. Chem. Res.* **2002**, *35*, 113–120.

(31) Murzin, A. G. Structural Classification of Proteins: New Superfamilies. *Curr. Opin. Struct. Biol.* **1996**, *6* (3), 386–394.

(32) Zhang, Y.; Morar, M.; Ealick, S. E. Structural Biology of the Purine Biosynthetic Pathway. *Cell. Mol. Life Sci.* **2008**, *65* (23), 3699–3724.

(33) Eroglu, B.; Powers-Lee, S. G. Mutational Analysis of ATP-Grasp Residues in the Two ATP Sites of Saccharomyces Cerevisiae Carbamoyl Phosphate Synthetase. *Arch. Biochem. Biophys.* **2002**, *407* (1), 1–9.

(34) Li, H.; Xu, H.; Graham, D. E.; White, R. H. Glutathione Synthetase Homologs Encode Alpha-L-Glutamate Ligases for Methanogenic Coenzyme F420 and Tetrahydrosarcinapterin Biosyntheses. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (17), 9785–9790.

(35) Healy, V. L.; Mullins, L. S.; Li, X.; Hall, S. E.; Raushel, F. M.; Walsh, C. T. D-Ala–d-X Ligases: Evaluation of D-Alanyl Phosphate Intermediate by MIX, PIX and Rapid Quench Studies. *Chem. Biol.* **2000**, *7* (7), 505–514.

(36) Marolewski, A. E.; Mattia, K. M.; Warren, M. S.; Benkovic, S. J. Formyl Phosphate: A Proposed Intermediate in the Reaction Catalyzed by Escherichia Coli PurT GAR Transformylase. *Biochemistry* **1997**, *36* (22), 6709–6716.

(37) Mueller, E. J.; Meyer, E.; Rudolph, J.; Davisson, V. J.; J, J. S. M-Carboxyaminoimidazole Ribonucleotide: Evidence for a New Intermediate and Two New Enzymatic Activities in

the de Novo Purine Biosynthetic Pathway of Escherichia Coli. *Biochemistry* **1994**, *33*, 2269–2278.

(38) Mullins, L. S.; Lusty, C. J.; Raushel, F. M. Alterations in the Energetics of the Carbamoyl Phosphate Synthetase Reaction by Site-Directed Modification of the Essential Sulfhydryl Group. *J. Biol. Chem.* **1991**, *266* (13), 8236–8240.

(39) Mullins, L. S.; Zawadzke, L. E.; Walsh, C. T.; Raushel, F. M. Kinetic Evidence for the Formation of D-Alanyl Phosphate in the Mechanism of D-Alanyl-D-Alanine Ligase. *J. Biol. Chem.* **1990**, *265* (16), 8993–8998.

(40) Ogita, T.; Knowles, J. R. On the Intermediacy of Carboxyphosphate in Biotin-Dependent Carboxylations. *Biochemistry* **1988**, *27* (21), 8028–8033.

(41) Knowles, J. R. The Mechanism of Biotin-Dependent Enzymes. *Annu. Rev. Biochem.* **1989**, *58* (1), 195–221.

(42) Tipton, P. a; Cleland, W. W. Carbon-13 and Deuterium Isotope Effects on the Catalytic Reactions of Biotin Carboxylase. *Biochemistry* **1988**, *27* (12), 4325–4331.

(43) Utter, M. F.; Keech, D. B. Formation of Oxaloacetate from Pyruvate and Carbon Dioxide. *J Biol Chem* **1960**, *235*, PC17–PC18.

(44) Attwood, P. V.; Graneri, B. D. L. A. Pyruvate Carboxylase Catalysis of Phosphate Transfer between Carbamoyl Phosphate and ADP. *Biochem. J.* **1991**, *273*, 443–448.

(45) Jitrapakdee, S.; Vidal-Puig, A.; Wallace, J. C. Anaplerotic Roles of Pyruvatecarboxylase in Mammalian Tissues. *Cell. Mol. Life Sci.* **2006**, *63* (7-8), 843–854.

(46) Cohen, N. D.; Beegen, H.; Utter, M. F.; Wrigley, N. G. A Re-Examination of Electron Microscopic Appearance of Pyruvate Carboxylase from Chicken Liver. *J Biol Chem* **1979**, *254* (5), 1740–1747.

(47) Meyer, E.; Kappock, T. J.; Osuji, C.; Stubbe, J. Evidence for the Direct Transfer of the Carboxylate of N5-Carboxyaminoimidazole Ribonucleotide (N5-CAIR) to Generate 4-Carboxy-5-Aminoimidazole Ribonucleotide Catalyzed by Escherichia Coli PurE, an N5-CAIR Mutase. *Biochemistry* **1999**, *38* (10), 3012–3018.

(48) Meyer, E.; Leonard, N. J.; Bhat, B.; Stubbe, J.; Smith, J. M. Purification and Characterization of the purE, purK, and purC Gene Products: Identification of a Previously Unrecognized Energy Requirement in the Purine Biosynthetic Pathway. *Biochemistry* **1992**, *31*, 5022–5032.

(49) Firestine, S. M.; Poon, S.-W.; Mueller, E. J.; Stubbe, J.; Davisson, V. J. Reactions Catalyzed by 5-Aminoimidazole Ribonucleotide Carboxylases from Escherichia Coli and Gallus Gallus: A Case for Divergent Catalytic Mechanisms?. *Biochemistry* **1994**, *33* (39), 11927–11934.

(50) Firestine, S. M.; Davisson, V. J. Carboxylases in de Novo Purine Biosynthesis. Characterization of the Gallus Gallus Bifunctional Enzyme. *Biochemistry* **1994**, *33* (39), 11917–11926.

(51) Constantine, C. Z.; Starks, C. M.; Mill, C. P.; Ransome, A. E.; Karpowicz, S. J.; Francois, J. A.; Goodman, R. A.; Kappock, T. J. Biochemical and Structural Studies of N 5 - Carboxyaminoimidazole Ribonucleotide Mutase from the Acidophilic Bacterium Acetobacter Aceti. *Biochemistry* **2006**, *45* (27), 8193–8208.

(52) Firestine, S. M.; Misialek, S.; Toffaletti, D. L.; Klem, T. J.; Perfect, J. R.; Davisson, V. J. Biochemical Role of the Cryptococcus Neoformans ADE2 Protein in Fungal de Novo Purine Biosynthesis. *Arch. Biochem. Biophys.* **1998**, *351* (1), 123–134.

(53) Watanabe, W.; Sampei, G.; Aiba, A.; Mizobuchi, K. Identification and Sequence Analysis

of Escherichia Coli purE and purK Genes Encoding 5'-Phosphoribosyl-5-Amino-4-Imidazole Carboxylase for de Novo Purine Biosynthesis. *J. Bacteriol.* **1989**, *171* (1), 198–204.

(54) Kirsch, D. R.; Whitney, R. R. Pathogenicity of Candida Albicans Auxotrophic Mutants in Experimental Infections. *Infect. Immun.* **1991**, *59* (9), 3297–3300.

(55) Perfect, J. R.; Toffaletti, D. L.; Rude, T. H. The Gene Encoding Phosphoribosylaminoimidazole Carboxylase (ADE2) Is Essential for Growth of Cryptococcus Neoformans in Cerebrospinal Fluid. *Infect. Immun.* **1993**, *61* (10), 4446–4451.

(56) Donovan, M.; Schumuke, J. J.; Fonzi, W. A.; Bonar, S. L.; Gheesling-Mullis, K.; Jacob, G. S.; Davisson, V. J.; Dotson, S. B. Virulence of a Phosphoribosylaminoimidazole Carboxylase-Deficient Candida Albicans Strain in an Immunosuppressed Murine Model of Systemic Candidiasis. *Infect. Immun.* **2001**, *69* (4), 2542–2548.

(57) Firestine, S. M.; Paritala, H.; Mcdonnell, J. E.; Thoden, J. B.; Holden, H. M. Identification of Inhibitors of N5-Carboxyaminoimidazole Ribonucleotide Synthetase by High-Throughput Screening. *Bioorg. Med. Chem.* **2009**, *17* (9), 3317–3323.

(58) Chou, C.-Y.; Tong, L. Structural and Biochemical Studies on the Regulation of Biotin Carboxylase by Substrate Inhibition and Dimerization. *J. Biol. Chem.* **2011**, *286* (27), 24417–24425.

(59) Kaziro, Y.; Hass, L. F.; Boyer, P. D.; Ochoa, S. Mechanism of the Propionyl Carboxylase Reaction. II. Isotopic Exchange and Tracer Experiments. *J. Biol. Chem.* **1962**, *237*, 1460–1468.

(60) Kalousek, F.; Darigo, M. D.; Rosenberg, L. E. Isolation and Characterization of

Propionyl-CoA Carboxylase from Normal Human Liver. Evidence for a Protomeric Tetramer of Nonidentical Subunits. *J Biol Chem* **1980**, *255* (1), 60–65.

(61) Diacovich, L.; Mitchell, D. L.; Pham, H.; Gago, G.; Melgar, M. M.; Khosla, C.; Gramajo, H.; Tsai, S. C. Crystal Structure of Theβ-Subunit of Acyl-CoA Carboxylase: Structure-Based Engineering of Substrate Specificity. *Biochemistry* **2004**, *43* (44), 14027–14036.

(62) Huang, C. S.; Sadre-Bazzaz, K.; Shen, Y.; Deng, B.; Zhou, Z. H.; Tong, L. Crystal Structure of the α6β6 Holoenzyme of Propionyl-Coenzyme A Carboxylase. *Nature* **2010**, *466* (7309), 1001–1005.

(63) Deodato, F.; Boenzi, S.; Santorelli, F. M.; Dionisi-Vici, C. Methylmalonic and Propionic Aciduria. *Am J Med Genet C Semin Med Genet* **2006**, *142* (2), 104–112.

(64) Wolf, B.; Kalousek, F.; Rosenberg, L. E. Essential Arginine Residues in the Active Sites of Propionyl CoA Carboxylase and Beta-Methylcrotonyl CoA Carboxylase. *Enzyme* **1979**, *24* (5), 302–306.

(65) Ravn, K.; Chloupkova, M.; Christensen, E.; Brandt, N. J.; Simonsen, H.; Kraus, J. P.; Nielsen, I. M.; Skovby, F.; Schwartz, M. High Incidence of Propionic Acidemia in Greenland Is Due to a Prevalent Mutation, 1540insCCC, in the Gene for the β-Subunit of Propionyl CoA Carboxylase. *Am. J. Hum. Genet.* **2000**, *67* (1), 203–206.

(66) Sloane, V.; Waldrop, G. L. Kinetic Characterization of Mutations Found in Propionic Acidemia and Methylcrotonylglycinuria: Evidence for Cooperativity in Biotin Carboxylase. *J Biol Chem* **2004**, *279* (16), 15772–15778.

(67) Mobley, H. L.; Hausinger, R. P. Microbial Ureases: Significance, Regulation, and Molecular Characterization. *Microbiol. Rev.* **1989**, *53*, 85–108.

(68) Mobley, H. L.; Island, M. D.; Hausinger, R. P. Molecular Biology of Microbial Ureases.

Microbiol. Rev. **1995**, *59*, 451–480.

(69) Sirko, A.; Brodzik, R. Plant Ureases: Roles and Regulation. *Acta Biochim. Pol.* **2000**, *47*, 1189–1195.

(70) Singer, M. A. Do Mammals, Birds, Reptiles, and Fish Have Similar Nitrogen Conserving Systems? *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **2003**, *134*, 543–558.

(71) Kanamori, T.; Kanou, N.; Atomi, H.; Inmanaka, T. Enzymatic Characterization of a Prokaryotic Urea Carboxylase. *J. Bacteriol.* **2004**, *186*, 2532–2539.

(72) Kanamori, T.; Kanou, N.; Kusakabe, S.; Atomi, H.; Imanaka, T. Allophanate Hydrolase of Oleomonas Sagaranensis Involved in an ATP-Dependent Degradation Pathway Specific to Urea. *FEMS Microbiol. Lett.* **2005**, *245*, 61–65.

(73) Navarathna, D. H.; Harris, S. D.; Roberts, D. D.; Nickerson, K. W. Evolutionary Aspects of Urea Utilization by Fungi. *FEMS Yeast Res.* **2010**, *10*, 209–213.

(74) Strope, P. K.; Nickerson, K. W.; Harris, S. D.; Moriyama, E. N. Molecular Evolution of Urea Amidolyase and Urea Carboxylase in Fungi. *BMC Evol. Biol.* **2011**, *11*, 80.

(75) Roon, R. J.; Levenberg, B. Urea Amidolyase. I. Properties of the Enzyme from Candida Utilis. *J. Biol. Chem.* **1972**, *247* (4107-4113), 4107–4113.

(76) Whitney, P. A.; Cooper, T. G. Urea Carboxylase and Allophanate Hydrolase: Two Components of Adenosine Triphosphate:urea Amidolyase in Saccharomyces Cerevisiae. *J Biol Chem* **1972**, *247*, 1349–1353.

(77) Andersen, G.; Björnberg, O.; Polakova, S.; Pynyaha, Y.; Rasmussen, A.; Møller, K.; Hofer, A.; Moritz, T.; Sandrini, M. P.; Merico, A. M.; et al. A Second Pathway to Degrade Pyrimidine Nucleic Acid Precursors in Eukaryotes. *J. Mol. Biol.* **2008**, *380*, 656–666.

(78) Ghosh, S.; Navarathna, D. H.; Roberts, D. D.; Cooper, J. T.; Atkin, A. L.; Petro, T. M.;

Nickerson, K. W. Arginine-Induced Germ Tube Formation in Candida Albicans Is Essential for Escape from Murine Macrophage Line RAW 264.7. *Infect. Immun.* **2009**, *77*, 1596–1605.

(79) Vylkova, S.; Carman, A. J.; Danhof, H. A.; Collette, J. R.; Zhou, H.; Lorenz, M. C. The Fungal Pathogen Candida Albicans Autoinduces Hyphal Morphogenesis by Raising Extracellular pH. *MBio* **2011**, *2*, e00055–00011.

(80) Jitrapakdee, S.; Wallace, J. C. The Biotin Enzyme Family: Conserved Structural Motifs and Domain Rearrangements. *Curr. Protein Pept. Sci.* **2003**, *4*, 217–229.

(81) St Maurice, M.; Reinhardt, L.; Surinya, K. H.; Attwood, P. V.; Wallace, J. C.; Cleland, W. W.; Rayment, I. Domain Architecture of Pyruvate Carboxylase, a Biotin-Dependent Multifunctional Enzyme. *Science* **2007**, *317*, 1076–1079.

(82) Xiang, S.; Tong, L. Crystal Structures of Human and Staphylococcus Aureus Pyruvate Carboxylase and Molecular Insights into the Carboxyltransfer Reaction. *Nat. Struct. Mol. Biol.* **2008**, *15*, 295–302.

(83) Holden, H. M.; Thoden, J. B.; Raushel, F. M. Carbamoyl Phosphate Synthetase: An Amazing Biochemical Odyssey from Substrate to Product. *Cell. Mol. Life Sci.* **1999**, *56* (5-6), 507–522.

(84) Meek, T. D.; Karsten, W. E.; Debrosse, C. W. Carbamoyl-Phosphate Synthetase II of the Mammalian CAD Protein: Kinetic Mechanism and Elucidation of Reaction Intermediates by Positional Isotope Exchange. *Biochemistry* **1987**, *26* (9), 2584–2593.

(85) Fahien, Leonard, A.; Cohen, P. P. A Kinetic Study of Carbamyl Phosphate Synthetase. *J Biol Chem* **1964**, *239* (6), 1925–1934.

(86) Elliott, K. R.; Tipton, K. F. Kinetic Studies of Bovine Liver Carbamoyl Phosphate

Synthetase. *Biochem. J.* **1974**, *141* (3), 807–816.

(87) Herschlag, D.; Jencks, W. P. The Effects of Mg2+, Hydrogen Bonding, and Steric Factors on Rate and Equilibrium Constants for Phosphoryl Transfer between Carboxylate Ions and Pyridines. *J. Am. Chem. Soc.* **1990**, *112* (1), 1942–1950.

(88) Powers, S. G.; Meister, a. Identification of Enzyme-Bound Activated CO2 as Carbonic-Phosphoric Anhydride: Isolation of the Corresponding Trimethyl Derivative from the Active Site of Glutamine-Dependent Carbamyl Phosphate Synthetase. *Proc. Natl. Acad. Sci. U. S. A.* **1976**, *73* (9), 3020–3024.

(89) Jones, M. E.; Lipmann, F. Chemical and Enzymic Synthesis of Carbamoyl Phosphate. *Proc. Natl. Acad. Sci. U. S. A.* **1960**, *46*, 1194–1205.

(90) Metzenberg, R. L.; Hall, L. M.; Marshall, M.; Cohen, P. P. Biosynthesis of Carbamyl Phosphate. *J Biol Chem* **1957**, *229* (2), 1019–1025.

(91) Anderson, P. M. Purification and Properties of the Glutamine- and N-Acetyl-L-Glutamate-Dependent Carbamoyl Phosphate Synthetase from Liver of Squalus Acanthias. *J. Biol. Chem.* **1981**, *256* (23), 12228–12238.

(92) Ogita, T.; Knowles, J. R. On the Intermediacy of Carboxyphosphate in Biotin-Dependent Carboxylations. *Biochemistry* **1988**, *27* (21), 8028–8033.

(93) Hansen, D. E.; Knowles, J. R. The Stereochemical Course at Phosphorus of the Reaction Catalyzed by Phosphoenolpyruvate Carboxylase. *J Biol Chem* **1982**, *257* (24), 14795–14798.

(94) Knowles, J. R. Biotin-Dependent Enzymes. *Annu. Rev. Biochem.* **1989**, *58* (1), 195–221.

(95) Tipton, P. A.; Cleland, W. W. Catalytic Mechanism of Biotin Carboxylase: Steady-State Kinetic Investigations. *Biochemistry* **1988**, *27* (12), 4317–4325.

(96) Hansen, D. E.; Knowles, J. R. N-Carboxybiotin Formation by Pyruvate Carboxylase: The Stereochemical Consequence at Phosphorus. *J. Am. Chem. Soc.* **1985**, *107* (26), 8304–8305.

(97) Hansen, D. E.; Knowles, J. R. N-Carboxybiotin Formation by Pyruvate Carboxylase: The Stereochemical Consequence at Phosphorous. *J. Am. Chem. Soc.* **1985**, *107* (5), 8304–8305.

(98) Polakis, S. E.; Guchhait, R. B.; Lane, M. D. On the Possible Involvement of a Carbonyl Phosphate Intermediate in the Adenosine Triphosphate-Dependent Carboxylation of Biotin. **1972**, *247*, 1335–1337.

(99) Ashman, L. K.; Keech, D. B. Sheep Kidney Pyruvate Carboxylase: Studies on the Coupling of Adenosine Triphosphate-Dependent Carboxylation of Biotin. *J Biol Chem* **1975**, *250* (1), 14–21.

(100) Kluger, R.; Taylor, S. D. Mechanisms of Carbonyl Participation in Phosphate Ester Hydrolysis and Their Relationship to Mechanisms for the Carboxylation of Biotin. *J. Am. Chem. Soc.* **1991**, *113* (3), 996–1001.

(101) Herschlag, D.; Jencks, W. The Effects of Mg2+, Hydrogen Bonding, and Steric Factors on Rate and Equilibrium Constants for Phosphoryl Transfter between Carboxylate Ions and Pyridines. *J. Am. Chem. Soc.* **1990**, *2* (1), 1942–1950.

(102) Levert, K. L.; Lloyd, R. B.; Waldrop, G. L. Do Cysteine 230 and Lysine 238 of Biotin Carboxylase Play a Role in the Activation of Biotin?†. *Biochemistry* **2000**, *39* (14), 4122–4128.

(103) Greenwood, N. N.; Earnshaw, A. *Chemistry of Elements*, 2nd ed.; Butterworth-Heinemann: Oxford, UK, 1997.

(104) Sauers, C. K.; Jencks, W. P.; Groh, S. Alcohol-Bicarbonate-Water System. Structure-Reactivity Studies on the Equilibriums for Formation of Alkyl Monocarbonates and on the Rates of Their Decomposition in Aqueous Alkali. *J. Am. Chem. Soc.* **1975**, *97* (19), 5546–5553.

(105) Faurholt, C. Studies on Monoalkyl Carbonates. II. The Formation of Monoalkyl Carbonic Acids or Their Salts on Dissolving Carbon Dioxide in Aqueous Solutions of Alcohols of Different Degrees of Acidity. *Z. Phys.* **1927**, *126*, 85–104.

(106) Faurholt, C. Studies on Monoalkyl Carbonates. I. The Formation of Monoalkyl Carbonates from Sodium Bicarbonate in Aqueous Solutions of Alcohols. The Equilibrium between Alcohol, Monoalkyl Carbonate, Carbonate and Carbon Dioxide in Aqueous Solution. *Z. Phys.* **1927**, *126*, 72–84.

(107) Meister, A.; Powers, S. G. Glutamine-Dependent Carbamyl Phosphate Synthetase: Catalysis and Regulation. *Adv. Enzyme Regul.* **1978**, *16*, 289–315.

(108) Rubio, V.; Grisolia, S. Mechanism of Mitochondrial Carbamoyl-Phosphate Synthetase: Synthesis and Properties of Active CO2, Precursor of Carbamoyl Phosphate. *Biochemistry* **1977**, *16* (2), 321–329.

(109) Wimmer, M. J.; Rose, I. A.; Powers, S. G.; Meister, A. Evidence That Carboxyphosphate Is a Kinetically Compentent Intermediate in the Carbamyl Phosphate Synthetase Reaction. *J. Biol. Chem.* **1978**, *254* (6), 1854–1859.

(110) Metzenberg, R. L.; Hall, L. M.; Marshall, M.; Cohen, P. P. Studies on the Biosynthesis of Carbamyl Phosphate. *J Biol Chem* **1957**, *229*, 1019–1025.

(111) Griffith, D. L.; Stiles, M. Phosphoric-Carbonic Anhydrides. The Hydrolysis of P-Nitrocarbobenzoxy Phosphate. *J. Am. Chem. Soc.* **1965**, *87* (16), 3710–3715.

(112) Ashman, L. K.; Keech, D. B. Sheep Kidney Pyruvate Carboxylase: Studies on the Coupling of Adenosine Triphosphate Hydrolysis and CO2 Fixation. *J. Biol. Chem.* **1975**, *250* (1), 14–21.

(113) Allen, C. M.; Jones, M. E. Decomposition of Carbamylphosphate in Aqueous Solutions. *Biochemistry* **1964**, *3* (303), 1238–1247.

(114) Ito, Y.; Kondo, H.; Shiota, Y.; Yoshizawa, K. Theoretical Analysis of the Reaction Mechanism of Biotin Carboxylase. *J. Chem. Theory Comput.* **2008**, *4* (2), 366–374.

(115) Ratner, M. A.; Schatz, G. C. *Introduction to Quantum Mechanics in Chemistry*; Prentice Hall: Upper Saddle River, NJ, 2001.

(116) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. Gaussian 09. Wallingford, CT, 2009.

(117) McQuarrie, D. A.; Simon, J. D. *Physical Chemistry: A Molecular Approach*, 1st ed.; University Science Books: Sausalito, CA, 1997.

(118) Trindle, C.; Shillady, D. *Electronic Structure and Modeling, Connections Between Theory and Software*, 1st ed.; CRC Press: Boca Raton, FL, 2008.

(119) de Broglie, L. The Reinterpretation of Wave Mechanics. *Found. Phys.* **1970**, *1* (1), 5–15.

(120) Born, M.; Oppenheimer, R. Quantum Theory of the Molecules. *Ann. der Phys. (Berlin, Ger)* **1927**, *84*, 457–484.

(121) Jensen, J. H. *Molecular Modeling Basics*; CRC Press: Taylor & Francis Group: Boca Raton, FL, 2010.

(122) Foresman, J. B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods*, 2nd ed.; Gaussian Inc: Pittsburgh, PA, 1996.

(123) Jensen, F. *Introduction to Computational Chemistry*, 2nd ed.; Wiley: Hoboken, NJ, 2008.

(124) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications: Mineloa, NY, 1996.

(125) Eckart, C. The Theory and Calculation of Screening Constants. *Phys. Rev.* **1930**, *36*, 878–892.

(126) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 2nd ed.; John Wiley & Sons Ltd.: West Sussex, UK, 2004.

(127) Leach, R. A. *Molecular Modeling Principles and Applications*, 2nd ed.; Pearson Education: Harlow, UK, 2001.

(128) Leininger, M. L.; Allen, W. D.; Schaefer, H. F.; Sherrill, C. D. Is Moller–Plesset Perturbation Theory a Convergent Ab Initio Method? *J. Chem. Phys.* **2000**, *112* (21), 9213.

(129) Monkhorst, H. J. Chemical Physics without the Born-Oppenheimer Approximation: The Molecular Coupled-Cluster Method. *Phys. Rev. A* **1987**, *36* (4), 1544–1561.

(130) Purvis, G. D.; Bartlett, R. J. A Full Coupled-cluster Singles and Doubles Model: The Inclusion of Disconnected Triples. *J. Chem. Phys.* **1982**, *76* (4), 1910–1918.

(131) Hobza, P. Calculations on Noncovalent Interactions and Databases of Benchmark Interaction Energies. *Acc. Chem. Res.* **2012**, *45* (4), 663–672.

(132) Poblador-Bahamonde, A. I.; Raynaud, C.; Eisenstein, O. Structures of d4 MH3X: A Computational Study of the Influence of the Metal and the Ligands. *Inorg. Chem.* **2012**, *51* (10), 5705–5715.

(133) Parthiban, S.; de Oliveira, G.; Martin, J. M. L. Benchmark Ab Initio Energy Profiles for the Gas-Phase SN2 Reactions Y− + CH3X → CH3Y + X− (X,Y = F,Cl,Br). Validation of

Hybrid DFT Methods. *J. Phys. Chem. A* **2001**, *105* (5), 895–904.

(134) Koch, W.; Houlhausen, M. C. *A Chemist's Guide to Density Functional Theory*; Wiley-VCH: Weinheim, Germany, 2001.

(135) Adamo, C.; Barone, V. Exchange Functionals with Improved Long-Range Behavior and Adiabatic Connection Methods without Adjustable Parameters: The mPW and mPW1PW Models. *J Chem Phys* **1998**, *108* (2), 664.

(136) Becke, A. D. Density Functional Thermochemistry . III . The Role of Exact Exchange. *J. Chem. Phys* **1993**, *98*, 5648–5652.

(137) Becke, A. D. Density Functional Theormochemistry. IV. A New Dynamical Correlation Functional and Implications for Exact Exchange Mixing. *J. Chem. Phys.* **1996**, *104* (January), 1040.

(138) Becke, A. D. Density-Functional Thermochemistry. V. Systematic Optimization of Exchange-Correlation Functionals. *J Chem Phys* **1997**, *107*, 8554–8560.

(139) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868.

(140) Boese, A. D.; Handy, N. C. New Exchange-Correlation Density Functionals: The Role of the Kinetic-Energy Density. *J. Chem. Phys.* **2002**, *116* (22), 9559.

(141) Boese, A. Daniel, Martin, Jan M L; Handy, N. C. The Role of the Basis Set: Assessing Density Functional Theory. *J. Chem. Phys.* **2003**, *119* (6), 3005–3014.

(142) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. Hybrid Exchange-Correlation Functional Determined from Thermochemical Data and Ab Initio Potentials. *J. Chem. Phys.* **2001**, *115* (20), 9233–9242.

(143) Handy, N. C.; Cohen, A. J. Left-Right Correlation Energy. *Mol. Phys.* **2001**, *99* (5), 403–

412.

(144) Hoe, W.-M.; Cohen, A. J.; Handy, N. C. Assessment of a New Local Exchange Functional OPTX. *Chem. Phys. Lett.* **2001**, *341* (3-4), 319–328.

(145) Perdew, J. P.; Kurth, S.; Zupan, A.; Blaha, P. Accurate Density Functional with Correct Formal Properties: A Step beyond the Generalized Gradient Approximation. *Phys. Rev. Lett.* **1999**, *82* (12), 2544–2547.

(146) Lynch, B. J.; Truhlar, D. G. How Well Can Hybrid Density Functional Methods Predict Transition State Geometries and Barrier Heights? *J. Phys. Chem. A* **2001**, *105* (13), 2936–2941.

(147) Ren, Y.; Wolk, J. L.; Hoz, S. The Performance of Density Function Theory in Describing Gas-Phase SN2 Reactions at Saturated Nitrogen. *Int. J. Mass Spectrom.* **2002**, *221* (1), 59–65.

(148) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136* (3B), B864–B871.

(149) Anjos, I. C.; Vasconcellos, M. L. a. a.; Rocha, G. B. A DFT and Natural Resonance Theory Investigation of the Electronic Structure of Mesoionic Compounds. *Theor. Chem. Acc.* **2012**, *131* (12), 1294.

(150) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. Development and Assessment of a New Hybrid Density Functional Model for Thermochemical Kinetics. *J. Phys. Chem. A* **2004**, *108* (14), 2715–2719.

(151) Gilbert, T. M. Computational Studies of Complexation of Nitrous Oxide by Borane-Phosphine Frustrated Lewis Pairs. *Dalton Trans.* **2012**, *41* (30), 9046–9055.

(152) Simón, L.; Goodman, J. M. How Reliable Are DFT Transition Structures? Comparison of

GGA, Hybrid-Meta-GGA and Meta-GGA Functionals. *Org. Biomol. Chem.* **2011**, *9* (3), 689–700.

(153) Claes, L.; François, J.-P.; Deleuze, M. S. Theoretical Study of the Internal Elimination Reactions of Xanthate Precursors. *J. Comput. Chem.* **2003**, *24* (16), 2023–2031.

(154) Zhao, Y.; Truhlar, D. G. Hybrid Meta Density Functional Theory Methods for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions: The MPW1B95 and MPWB1K Models and Comparative Assessments for Hydrogen Bonding and van Der Waals Interactions. *J. Phys. Chem. A* **2004**, *108* (33), 6908–6918.

(155) Fermi, E. Un Metodo Statistico per La Determinazione Di Alcune Proprieta Dell'atomo. *Rend. Accad. Naz. Lincei* **1927**, *6*, 602–607.

(156) Thomas, L. H. The Calculation of Atomic Fields. *Proc. Camb. Philol. Soc.* **1927**, *23* (5), 542.

(157) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev. A* **1965**, *140* (4), 1133–1138.

(158) Painter, G. S. Density Functional Description of Molecular Bonding within the Local Spin Density Approximation. *J. Phys. Chem.* **1986**, *90* (22), 5530–5535.

(159) Slater, J. C. A Simplification of Hartree-Fock Method. *Phys. Rev.* **1951**, *81* (3), 385.

(160) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38* (6), 3098–3100.

(161) Perdew, J. P.; Yue, W. Accurate and Simple Density Functional for the Electronic Exchange Energy: Generalized Gradient Approximation. *Phys. Rev. B* **1986**, *33* (12), 8800.

(162) Becke, A. D. Density Functional Calculations of Molecular Bond Energies. *J. Chem.*

Phys. **1986**, *84* (8), 4524.

(163) Perdew, J.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868.

(164) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Physical Review B*. **1988**, 37, 785–789.

(165) Becke, A. D. Correlation Energy of an Inhomogeneous Electron Gas: A Coordinate-Space Model. *J. Chem. Phys.* **1988**, *88* (2), 1053.

(166) Van Voorhis, T.; Scuseria, G. E. A Novel Form for the Exchange-Correlation Energy Functional. *J. Chem. Phys.* **1998**, *109* (2), 400–410.

(167) Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar, D. G. Adiabatic Connection for Kinetics. *J. Phys. Chem. A* **2000**, *104* (21), 4811–4815.

(168) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. Design of Density Functionals by Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2006**, *2* (2), 364–382.

(169) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Function. *Theor. Chem. Acc.* **2008**, *120* (1-3), 215–241.

(170) Zhao, Y.; Truhlar, D. G. Density Functionals with Broad Applicability in Chemistry. *Acc. Chem. Res.* **2008**, *41* (2), 157–167.

(171) Plumley, J. a.; Evanseck, J. D. Hybrid Meta-Generalized Gradient Functional Modeling of Boron-Nitroqen Coordinate Covalent Bonds. *J. Chem. Theory Comput.* **2008**, *4* (8), 1249–

1253.

(172) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. General Performance of Density Functionals General Performance of Density Functionals. **2007**, *111* (August), 10439–10452.

(173) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785–789.

(174) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58* (8), 1200–1211.

(175) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. Gaussian-2 Theory for Molecular Energies of First- and Second-Row Compounds. *J Chem Phys* **1991**, *94* (11), 7221.

(176) Schiøtt, B.; Iversen, B. B.; Hellerup Madsen, G. K.; Bruice, T. C. Characterization of the Short Strong Hydrogen Bond in Benzoylacetone by Ab Initio Calculations and Accurate Diffraction Experiments. Implications for the Electronic Nature of Low-Barrier Hydrogen Bonds in Enzymatic Reactions. *J. Am. Chem. Soc.* **1998**, *120* (46), 12117–12124.

(177) Peterson, K. a; Woon, D. E.; Dunning, T. H. Benchmark Calculations with Correlated Molecular Wave Functions. IV. The Classical Barrier Height of the H+H2→H2+H Reaction. *J. Chem. Phys.* **1994**, *100* (10), 7410–7415.

(178) Perdew, J. P.; Burke, K.; Wang, Y. Generalized Gradient Approximation for the Exchange-Correlation Hole of a Many-Electron System. *Phys. Rev. B* **1996**, *54* (23), 16533.

(179) Perdew, J. P.; Wang, Y. Accurate and Simple Analytic Representation of the Electron-Gas Correlation Energy. *Phys. Rev. B* **1992**, *45* (23), 13244.

(180) Lynch, B. J.; Truhlar, D. G. How Well Can Hybrid Density Functional Methods Predict Transition State Geometries and Barrier Heights? *J. Phys. Chem. A* **2001**, *105* (13), 2936–2941.

(181) Rybtchinski, B.; Oevers, S.; Montag, M.; Vigalok, A.; Rozenberg, H.; Martin, J. M. L.; Milstein, D. Comparison of Steric and Electronic Requirements for C−C and C−H Bond Activation. Chelating vs Nonchelating Case. *J. Am. Chem. Soc.* **2001**, *123* (37), 9064–9077.

(182) Lynch, B. J.; Truhlar, D. G. What Are the Best Affordable Multi-Coefficient Strategies for Calculating Transition State Geometries and Barrier Heights? *J. Phys. Chem. A* **2002**, *106* (5), 842–846.

(183) Claes, L.; François, J.-P.; Deleuze, M. S. From Sulfoxide Precursors to Model Oligomers of Conducting Polymers. *J. Am. Chem. Soc.* **2002**, *124* (25), 7563–7572.

(184) Dibble, T. S. Isomerization of OH-Isoprene Adducts and Hydroxyalkoxy Isoprene Radicals. *J. Phys. Chem. A* **2002**, *106* (28), 6643–6650.

(185) Ren, Y.; Wolk, J. L.; Hoz, S. Hybrid DFT Study on the Gas-Phase SN2 Reactions at Neutral Oxygen. *Int. J. Mass Spectrom.* **2003**, *225* (2), 167–176.

(186) Cohen, R.; Rybtchinski, B.; Gandelman, M.; Rozenberg, H.; Martin, J. M. L.; Milstein, D. Metallacarbenes from Diazoalkanes: An Experimental and Computational Study of the Reaction Mechanism. *J. Am. Chem. Soc.* **2003**, *125* (21), 6532–6546.

(187) Claes, L.; François, J.-P.; Deleuze, M. S. Theoretical Study of the Conversion of Sulfonyl Precursors into Chains of Poly(p-Phenylene Vinylene). *J. Am. Chem. Soc.* **2003**, *125* (23), 7129–7138.

(188) Iron, M. A.; Martin, J. M. L.; van der Boom, M. E. Cycloaddition Reactions of

Metalloaromatic Complexes of Iridium and Rhodium: A Mechanistic DFT Investigation. *J. Am. Chem. Soc.* **2003**, *125* (38), 11702–11709.

(189) Lynch, B. J.; Truhlar, D. G. Robust and Affordable Multicoefficient Methods for Thermochemistry and Thermochemical Kinetics: The MCCM/3 Suite and SAC/3. *J. Phys. Chem. A* **2003**, *107* (19), 3898–3906.

(190) Zhao, Y.; Truhlar, D. G. Applications and Validations of the Minnesota Density Functionals. *Chem. Phys. Lett.* **2011**, *502* (1-3), 1–13.

(191) Van Voorhis, T.; Scuseria, G. E. Exchange Energy Functionals Based on the Density Matrix Expansion of the Hartree-Fock Exchange Term. *Mol. Phys.* **1997**, *92* (3), 601–608.

(192) Boys, S. F. Electronic Wave Functions. II. A Calculation for the Ground State of the Beryllium Atom. *Proc. R. Soc. London, Ser. A* **1950**, *201*, 125.

(193) Boys, S. F. A General Method of Calculation for the Stationary States of Any Molecular System. *Proc. R. Soc. London, Ser. A* **1050**, *200*, 542.

(194) Hehre, W. J.; Stewart, R. F.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals. *J Chem Phys* **1969**, *51* (6), 2657.

(195) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J Chem Phys* **1972**, *56* (5), 2257.

(196) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XX. A Basis Set for Correlated Wave Functions. *J Chem Phys* **1980**, *72* (1), 650.

(197) Binkley, J. S.; Pople, J. A.; Hehre, W. J. Self-Consistent Molecular Orbital Methods. 21.

Small Split-Valence Basis Sets for First-Row Elements. *J. Am. Chem. Soc.* **1980**, *102* (3), 939.

(198) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods IX. Extended Gaussiain-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J Chem Phys* **1971**, *54* (2), 724.

(199) Hariharan, P. C.; Pople, J. A. Influence of Polarization Functions on MO Hydrogenation Energies. *Theor. Chima. Acta* **1973**, *28* (3), 213.

(200) Davidson, E. R.; Feller, D. Basis Set Selection for Molecular Calculations. *Chem Rev* **1986**, *86* (4), 681.

(201) Wilson, A. K.; van Mourik, T.; Dunning Jr, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. VI. Sextuple Zeta Correlation Consistent Basis Sets for Boron through Neon. *J. Mol. Struct.* **1996**, *388*, 339–349.

(202) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90* (2), 1007.

(203) Woon, D. E.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum through Argon. *J. Chem. Phys.* **1993**, *98* (2), 1358–1371.

(204) Kestner, N. R.; Combariza, J. E. Basis Set Superposition Errors: Theory and Practice. *Rev. Comput. Chem.* **1999**, *13*, 99.

(205) Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. Perspectives on Basis Sets Beautiful: Seasonal Plantings of Diffuse Basis Functions. *J. Chem. Theory Comput.* **2011**, *7* (10), 3027–3034.

(206) Anderson, P. M.; Meister, a. Evidence for an Activated Form of Carbon Dioxide in the

Reaction Catalyzed by Escherichia Coli Carbamyl Phosphate Synthetase•. *Biochemistry* **1965**, *4* (12), 2803–2809.

(207) Head-Gordon, M.; Pople, J. a.; Frisch, M. J. MP2 Energy Evaluation by Direct Methods. *Chem. Phys. Lett.* **1988**, *153* (6), 503–506.

(208) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Results Obtained with the Correlation Energy Density Functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **1989**, *157* (3), 200–206.

(209) Kendall, R. a; Dunning, T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. **1992**, *96* (9), 6796–6806.

(210) Martin, J. M. L. Ab Initio Total Atomization Energies of Small Molecules — towards the Basis Set Limit. *Chem. Phys. Lett.* **1996**, *259* (5-6), 669–678.

(211) Improta, R.; Scalmani, G.; Frisch, M. J.; Barone, V. Toward Effective and Reliable Fluorescence Energies in Solution by a New State Specific Polarizable Continuum Model Time Dependent Density Functional Theory Approach. *J. Chem. Phys.* **2007**, *127* (7), 74504.

(212) Improta, R.; Barone, V.; Scalmani, G.; Frisch, M. J. A State-Specific Polarizable Continuum Model Time Dependent Density Functional Theory Method for Excited State Calculations in Solution. *J. Chem. Phys.* **2006**, *125* (5), 54103.

(213) Hehre, W. J.; Ditchfield, R.; Radom, L.; Pople, J. A. Molecular Orbital Theory of the Electronic Structure of Organic Compounds. V. Molecular Theory of Bond Separation. *J. Am. Chem. Soc.* **1970**, *92* (16), 4796–4801.

(214) Gamoke, B.; Neff, D.; Simons, J. Nature of PO Bonds in Phosphates. *J. Phys. Chem. A* **2009**, *113* (19), 5677–5684.

(215) Schuster, P. LCAO-MO-Beschreibung intramolekularer Wasserstoffbrücken. *Monatshefte für Chemie / Chem. Mon.* **1969**, *100* (6), 2084–2095.

(216) Pople, J.; Radom, L.; Hehre, W. Molecular Orbital Theory of the Electronic Structure of Organic Compounds. VII. Systematic Study of Energies, Conformations, and Bond Interactions. *J. Am. Chem. Soc.* **1971**, *13* (3), 289–300.

(217) Nowroozi, A.; Raissi, H.; Farzad, F. The Presentation of an Approach for Estimating the Intramolecular Hydrogen Bond Strength in Conformational Study of β-Aminoacrolein. *J. Mol. Struct.* **2005**, *730* (1-3), 161–169.

(218) Buemi, G.; Zuccarello, F.; Venuvanalingam, P.; Ramalingam, M.; Salai Cheettu Ammal, S. Abinitio Study of Formazan and 3-Nitroformazan. *J. Chem. Soc. Faraday Trans.* **1998**, *94* (22), 3313–3319.

(219) Scheiner, S. *Hydrogen Bonding: A Theoretical Perspective*; Truhlar, D. G., Ed.; Oxford University Press: New York, 1997.

(220) Jesus, A. J. L.; Redinha, J. S. Charge-Assisted Intramolecular Hydrogen Bonds in Disubstituted Cyclohexane Derivatives. *J. Phys. Chem. A* **2011**, *115* (48), 14069–14077.

(221) Kumar, G. A.; McAllister, M. A. Theoretical Investigation of the Relationship between Proton NMR Chemical Shift and Hydrogen Bond Strength. *J. Org. Chem.* **1998**, *63* (20), 6968–6972.

(222) Grabowski, S. J. Hydrogen Bonding Strength - Measures Based on Geometric and Topological Parameters. *J. Phys. Org. Chem.* **2004**, *17*, 18–31.

(223) Craw, J. S.; Bacskay, G. B. Quantum-Chemical Studies of Hydrogen Bonding Involving Thioxoketones, Thienols, Thioformaldehyde and Hydrogen Sulfide with Specific Reference to the Strength of Intramolecular Hydrogen Bonds. *J. Chem. Soc. Faraday*

Trans. **1992**, *88* (16), 2315–2321.

(224) Scheiner, S.; Kar, T.; Cuma, M. Excited State Intramolecular Proton Transfer in Anionic Analogs of Malonaldehyde. *J. Phys. Chem. A* **1997**, *101* (33), 5901–5909.

(225) Chung, G.; Kwon, O.; Kwon, Y. Theoretical Study on 1,2-Dihydroxybenzene and 2-Hydroxythiophenol: Intramolecular Hydrogen Bonding. *J. Phys. Chem. A* **1997**, *101* (49), 9415–9420.

(226) Cuma, M.; Scheiner, S.; Kar, T. Effect of Adjoining Aromatic Ring upon Excited State Proton Transfer, O-Hydroxybenzaldehyde. *J. Mol. Struct.* **1999**, *467*, 37–49.

(227) Buemi, G.; Zuccarello, F. Is the Intramolecular Hydrogen Bond Energy Valuable from Internal Rotation Barriers? *J. Mol. Struct.* **2002**, *581*, 71–85.

(228) Jabłoński, M.; Kaczmarek, A.; Sadlej, A. J. Estimates of the Energy of Intramolecular Hydrogen Bonds. *J. Phys. Chem. A* **2006**, *110* (37), 10890–10898.

(229) Jablonski, M. Full vs. Constrain Geometry Optimization in the Open-Closed Method in Estimating the Energy of Intramolecular Charge-Inverted Hydrogen Bonds. *Chem. Phys.* **2010**, *376*, 76–83.

(230) Jabłoński, M. Full vs. Constrain Geometry Optimization in the Open–closed Method in Estimating the Energy of Intramolecular Charge-Inverted Hydrogen Bonds. *Chem. Phys.* **2010**, *376* (1-3), 76–83.

(231) Varnali, T.; Hargittai, I. Geometrical Consequences of Resonance-Assisted Hydrogen Bonding in 2-Nitrovinyl Alcohol and Indication of a Slight Attractive O…H Interaction in 2-Nitroethanol. An Ab Initio Molecular Orbital Investigation. *J. Mol. Struct.* **1996**, *388*, 315–319.

(232) Glendening, E. D.; Weinhold, F. Natural Resonance Theory: II. Natural Bond Order and

Valency. *J. Comput. Chem.* **1998**, *19* (6), 610–627.

(233) Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, 1990.

(234) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31* (4), 671–690.

(235) MacKerell, A. D. J.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102* (18), 3586–3616.

(236) Schuster, P. Energy Surfaces for Hydrogen Bonded Systems. In *The Hydrogen Bond: Recent Developments in Theory and Experiments*; Schuster, P., Zundel, G., Sandorfy, C., Eds.; North-Holland Publishing Co: Amsterdam, 1976; pp 25–163.

(237) Hajiabadi, H.; Nowroozi, A.; Hasani, M.; Mohammadzadeh Jahani, P.; Raissi, H. A Comparative Study of Open-Close and Related Rotamers Methods to Evaluate the Intramolecular Hydrogen Bond Energies in 3-Imino-Propen-1-Ol and Its Derivatives. *Int. J. Quantum Chem.* **2012**, *112* (5), 1384–1391.

(238) Nowroozi, A.; Hajiabadi, H. How to Estimate the Intramolecular Hydrogen-Bond Energy of Complex RAHB Systems? A Theoretical Study. *Struct. Chem.* **2014**, *25* (1), 215–220.

(239) Nowroozi, A.; Hajiabadi, H.; Akbari, F. OH···O and OH···S Intramolecular Interactions in Simple Resonance-Assisted Hydrogen Bond Systems: A Comparative Study of Various Models. *Struct. Chem.* **2014**, *25* (1), 251–258.

(240) Nowroozi, A.; Raissi, H.; Hajiabadi, H.; Jahani, P. M. Reinvestigation of Intramolecular Hydrogen Bond in Malonaldehyde Derivatives: An Ab Initio, AIM and NBO Study. *Int. J. Quantum Chem.* **2011**, *111* (12), 3040–3047.

(241) Scheiner, S.; Kar, T.; Čuma, M. Excited State Intramolecular Proton Transfer in Anionic Analogues of Malonaldehyde. *J. Phys. Chem. A* **1997**, *101* (33), 5901–5909.

(242) Buemi, G.; Zuccarello, F. Is the Intramolecular Hydrogen Bond Energy Valuable from Internal Rotation Barriers? *J. Mol. Struct.* **2002**, *581* (1-3), 71–85.

(243) Jablonski, M.; Kaczmarek, A.; Sadlej, A. J. Estimates of the Energy of Intramolecular Hydrogen Bonds. *J. Phys. Chem. A* **2006**, *110* (37), 10890–10898.

(244) Emsley, J. Very Strong Hydrogen Bonding. *Chem. Soc. Rev.* **1980**, *9*, 91–124.

(245) Perrin, C. L.; Nielson, J. B. Strong Hydrogen Bonds in Chemistry and Biology. *Annu. Rev. Phys. Chem.* **1997**, *48*, 511–544.

(246) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: New York, 1997.

(247) Kreevoy, M. M. Structures and Isotopic Fractionation Factors of Complexes AHA-1. *J. Am. Chem. Soc.* **1977**, *99* (15), 5207–5209.

(248) Cleland, W. W. Low-Barrier Hydrogen Bonds and Low Fractionation Factor Bases in Enzymic Reactions. *Biochemistry* **1992**, *31* (2), 317–319.

(249) Perrin, C. L. Are Short, Low-Barrier Hydrogen Bonds Unusually Strong? *Acc. Chem. Res.* **2010**, *43* (12), 1550–1557.

(250) Remer, L. C.; Jensen, J. H. Toward a General Theory of Hydrogen Bonding: The Short, Strong Hydrogen Bond [HOH···OH]-. *J. Phys. Chem. A* **2000**, *104* (40), 9266–9275.

(251) Hall, L. M.; Metzenberg, R. L.; Cohen, P. P. Isolation and Characterization of Naturally

Occuring Cofactor of Carbamyl Phosphate Biosynthesis. *J Biol Chem* **1958**, *230* (2), 1013–1021.

(252) Kluger, R.; Davis, P. P.; Adawadkar, P. D. Mechanism of Urea Participation in Phosphonate Ester Hydrolysis. Mechanistic and Stereochemical Criteria for Enzymic Formation and Reaction of Phosphorylated Biotin. *J. Am. Chem. Soc.* **1979**, *101* (20), 5995–6000.

(253) Moller, C.; Plesset, M. S. Note on the Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.

(254) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

(255) Bryantsev, V. S.; Diallo, M. S.; Goddard III, W. A. Calculation of Solvation Free Energies of Charged Solutes Using Mixed Cluster/Continuum Models. *J. Phys. Chem. B* **2008**, *112* (32), 9709–9719.

(256) Lippert, B. Ligand-pKa Shifts through Metals: Potential Relevance to Ribozyme. *Chem. Biodivers.* **2008**, *5* (8), 1455–1474.

(257) Kukic, P.; Farrell, D.; McIntosh, L. P.; García-Moreno E., B.; Jensen, K. S.; Toleikis, Z.; Teilum, K.; Nielsen, J. E. Protein Dielectric Constants Determined from NMR Chemical Shift Perturbations. *J. Am. Chem. Soc.* **2013**, *135* (45), 16968–16976.

(258) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Adding Explicit Solvent Molecules to Continuum Solvent Calculations for the Calculation of Aqueous Acid Dissociation Constants. *J. Phys. Chem. A* **2006**, *110* (7), 2493–2499.

(259) Kuhn, B.; Kollman, P. A.; Stahl, M. Prediction of pKa Shifts in Proteins Using a Combination of Molecular Mechanical and Continuum Solvent Calculations. *J. Comput. Chem.* **2004**, *25* (15), 1865–1872.

(260) Ullmann, G. M.; Noodleman, L.; Case, D. a. Density Functional Calculation of P K(a) Values and Redox Potentials in the Bovine Rieske Iron-Sulfur Protein. *J. Biol. Inorg. Chem.* **2002**, *7* (6), 632–639.